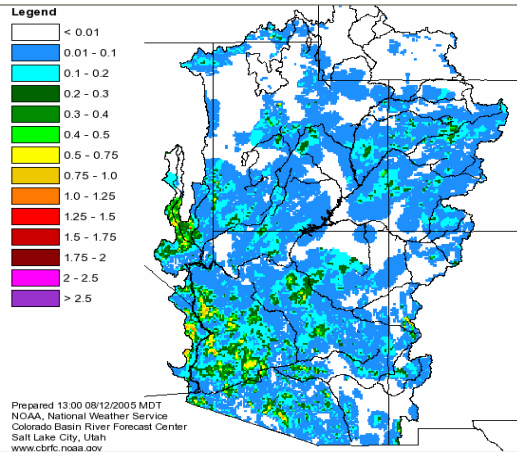
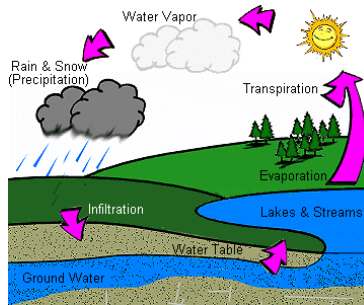




Office of Hydrologic Development Hydrology Laboratory Strategic Science Plan





Executive Summary

This Strategic Science Plan (Plan) establishes the directions for research in hydrology at the Hydrology Laboratory of the Office of Hydrologic Development. It first establishes a cross-reference between the Plan and NOAA, NWS, OHD policy documents, and to the National Research Council reviews of relevant NWS programs, especially those recommending the development of a strategic science plan for hydrologic research at the NWS. Furthermore, the Bulletin of the American Meteorological Society (BAMS) article by Welles *et al.* (2007) highlights the need for hydrologic forecast verification, and how this verification should be a driving force for new model development. A main section of the document is dedicated to the verification topic. This document is intended to be a “living document” to be updated on an annual basis.

The Plan follows closely the outline of the Integrated Water Science Team report. In each section the Plan offers several subsections: where we are, in which it details what the state of the science at OHD is; where we want to be, where the Plan sets the new directions or reaffirms existing directions; what are the challenges to get there, and a road map to reach those goals. The highlights of the Plan are:

The Plan directs the research on hydrologic modeling seeking:

- A more expeditious and cost-effective approach by reducing the effort required in model calibration.
- Improvements in forecast lead-time and accuracy.
- Provide for the rapid adjustment of model parameters to account for changes in the watershed, both rapid as the result from forest fires or levee breaches, and slow, as the result of watershed reforestation.
- A comprehensive ensemble approach to hydrologic forecasting.
- Data assimilation for lumped and distributed models.
- In future versions of this plan:
 - The impact of climate variability change on hydrologic forecasting.
 - Hydraulic modeling.
 - Water quality modeling.
 - Social Science research.

To this end, it places an emphasis on research of models with parameters that can be derived from physical watershed characteristics. Purely physically based models may be unattainable or unpractical, and, therefore, models resulting from a combination of physically and conceptually approached processes may be required. The plan acknowledges that the path in that direction is not clearly defined yet, as evidenced by the DMIP-1 and DMIP-2 results, and that OHD is proceeding in that direction, first, with the development of the *a priori* parameter es-

timates for the Sacramento model, and second, with the development of the Heat-Transfer modification to the Sacramento model.

For snow science, the research will be directed towards energy-budget models. This section will be addressed by a separate Snow Science Plan to be developed by the Snow Science Steering Team, once the present plan is approved.

The Hydrometeorological forcings section emphasizes the development of improved precipitation estimation techniques through the synthesis of radar, rain gauge, satellite, and numerical weather prediction model output, particularly in those areas where ground-based sensors are inadequate to detect spatial variability in precipitation. Better estimation and forecasting of precipitation are most likely to be achieved by statistical merging of remote-sensor observations and forecasts from high-resolution numerical prediction models. Enhancements to the satellite-based precipitation products will include use of TRMM precipitation data in preparation for information to be supplied by the Global Precipitation Mission satellites not yet deployed.

Because of a growing need for services in water resources, including low-flow forecasts for water supply customers, the plan directs research into coupled surface-groundwater models that will eventually replace the groundwater component of the existing models, and will be part of the new generation of models.

The Plan confirms the directions set forth in the respective planning documents of both the ensemble research and the verification system. In NWS operations, forecasters modify model state variables, forcings or even parameters to adjust model output to match observed streamflow. These adjustments, known as “modifications” or simply “mods,” are essentially a manual form of data assimilation. Moving into the realm of operational use of ensemble forecasting with distributed hydrologic models, “mods” become impractical, given the impossibility to what subcomponents (cells) of a watershed distributed model to adjust. Furthermore, doing “wholesale” (i.e. “lumped”) adjustments to cell contents or forcings defeats some of the advantages of distributed modeling. Therefore, it is imperative either to adapt existing or to develop new techniques to support automatic error correction in distributed models. Chapter 6 describes ensemble modeling. Chapter 7 covers data assimilation, and Chapter 8 delineates the verification plan.

Two key components of the implementation of this strategic plan are the Community Hydrologic Prediction System (CHPS), currently under development in collaboration with Deltares (http://www.deltares.nl/xmlpages/page/deltares_en), and the Hydrology testbed. The testbed will serve as an environment for the development of advanced science and software engineering techniques, moving the hydrologic forecasting applications developed since the 1970s for mainframe

computers, into distributed processing systems as the computational demands so require.

This first version of the plan does not cover the following topics: hydraulic models; water quality models; effects of climate change and variability on hydrologic forecasting; irrigation, and social sciences. These will be in the next year's update.

Reference

Welles E., S. Sorooshian, G. Carter, and B. Olsen. (2007). "Hydrologic Verification: A Call for Action and Collaboration." *BAMS*, 88(4), April, 503-511.

Acknowledgements

This report is a collaborative effort of the Strategic Science Plan Team, which consisted of:

Gary Carter, Team Mentor
Pedro Restrepo, Team Leader
John Kochendorfer, Coordinating Lead Author, and Appendices A and B main author
Mike Smith, Seann Reed, Victor Koren, Chapter 3 Contributing Authors
Dave Kitzmiller, David Riley, Denis Miller Chapter 4 Contributing Authors
D.J. Seo, Chapters 6 and 7 Contributing Author
Julie Demargne, Chapters 8 Contributing Author

Additional contributions and suggestions were also made by Geoff Bonnin, Ken Mitchell, John Schaake, Jay Day, Tim Love, Billy Olsen, Kevin Werner, Don Laurine, Mike Ek, Tim Schneider, and scientists from NSSL.

Contents

Executive Summary	i
Acknowledgements	v
Contents	vii
Figures	xi
Abbreviations and Acronyms	xiii
1. Introduction	1
1.1. The Role of OHD within the NOAA Mission Goals	1
1.1.1 NOAA Mission	1
1.2. OHD Strategic Science Goals	2
1.3. Organization of the Plan	3
1.4. References	6
2. A High-Level View	7
2.1. The Current Hydrologic Forecasting Process	7
2.2. The Future Hydrologic Forecasting Process	8
3. Watershed Models	14
3.1. Surface Properties	14
3.1.1 Where We Are	14
3.1.2 What Our Partners Are Doing	14
3.1.3 Where We Want to Be	14
3.1.4 Challenges to Getting There	14
3.1.5 A Road Map for Getting There	15
3.2. Infiltration and Surface Runoff	15
3.2.1 Where We Are	15
3.2.2 What Our Partners Are Doing	16
3.2.3 Where We Want to Be	16
3.2.4 Challenges to getting to where we want to be	17
3.2.5 A Road Map for Getting There	17
3.3. Soil Moisture and Temperature	18
3.3.1 Where We Are	18
3.3.2 What Our Partners Are Doing	20
3.3.3 Where We Want to Be	21
3.3.4 Challenges to Getting There	22
3.3.5 A Road Map for Getting There	22
3.4. Groundwater Storage and Base Flow	23
3.4.1 Where We Are	23
3.4.2 What Our Partners Are Doing	23
3.4.3 Where We Want to Be	23
3.4.4 Challenges to Getting There	24

3.4.5	A Road Map for Getting There	24
3.5.	Snow Accumulation, Sublimation and Melt	24
3.5.1	Where We Are	25
3.5.2	What Our Partners Are Doing	25
3.5.3	Where We Want to Be	26
3.5.4	Challenges to Getting There	27
3.5.5	A Road Map for Getting There	27
3.6.	Evapotranspiration	28
3.6.1	Where We Are	28
3.6.2	What Our Partners Are Doing	29
3.6.3	Where We Want to Be	29
3.6.4	Challenges to getting there	30
3.6.5	A Road Map for Getting There	30
3.7.	Hydraulics (Future)	30
3.8.	Water Quality (Future)	30
3.9.	References	31
4.	Forcings	37
4.1.	Observed Precipitation	37
4.1.1	Where We Are	37
4.1.2	What Our Partners Are Doing	38
4.1.3	Where we want to be	39
4.1.4	Challenges to Getting There	40
4.1.5	A Road Map for Getting There	40
4.2.	Forecasted Precipitation	42
4.2.1	Where We Are	42
4.2.2	What Our Partners Are Doing	42
4.2.3	Where We Want to Be	43
4.2.4	Challenges to getting there	44
4.2.5	A Road Map for Getting There	44
4.3.	Observed Air Temperature and Humidity	45
4.3.1	Where We Are	46
4.3.2	What Our Partners Are Doing	46
4.3.3	Where We Want to Be	47
4.3.4	Challenges to Getting There	47
4.3.5	A Road Map for Getting There	48
4.4.	Forecasted Air Temperature and Humidity	48
4.4.1	What Our Partners Are Doing	48
4.5.	Winds	49
4.5.1	Where We Are	49
4.5.2	What Our Partners Are Doing	50
4.5.3	Where We Want to Be	50
4.5.4	Challenges to Getting There	51
4.5.5	A Road Map for Getting There	51
4.6.	Shortwave/Longwave Radiation and Skin Temperature	51

4.6.1	Where We Are	51
4.6.2	What Our Partners Are Doing	52
4.6.3	Where We Want to Be	52
4.6.4	Challenges to Getting There	52
4.6.5	A Road Map for Getting There	53
4.7.	References	53
5.	Anthropogenic and Natural Perturbations to the Hydrologic Cycle	56
5.1.	Climate Change and Variability (Future)	56
5.2.	Irrigation (Future)	56
5.3.	Reservoir-based River Regulation	56
5.3.1	Where We Are	56
5.3.2	What Our Partners Are Doing	57
5.3.3	Where We Want to Be	58
5.3.4	Challenges to Getting There	58
5.3.5	A Road Map for Getting There	58
6.	Ensemble Forecasting	59
6.1.	Where We Are	61
6.2.	What our partners are doing	62
6.3.	Where We Want to Be	64
6.4.	Challenges to getting there	66
6.5.	A Road Map for Getting There	67
6.5.1	The eXperimental Ensemble Forecast System (XEFS)	67
6.5.2	Hydrology Test Bed	68
6.5.3	Collaborations with NCEP	69
6.5.4	Climate Prediction Project for the Americas (CPPA) Core Project	70
6.5.5	HEPEX	70
6.5.6	DMIP	71
6.6.	References	71
7.	Data Assimilation	75
7.1.	Where We Are	75
7.2.	What Our Partners Are Doing	75
7.3.	Where We Want to Be	76
7.4.	Challenges to Getting There	77
7.5.	A Road Map for Getting There	78
7.5.1	Hydrology and Water Resources Data Assimilation Projects (AHPS, CPPA, Water Resources, Hurricane Supplemental)	79
7.5.2	Hydrology Test Bed	79
7.5.3	Collaborative Projects	79
7.6.	References	80
8.	Verification	82
8.1.	Where We Are	85

8.2.	What Our Partners Are Doing	85
8.3.	Where We Want to Be	86
8.3.1	Forecast Services Verification	86
8.3.2	Forecast Verification	87
8.4.	Challenges to getting there	90
8.5.	A Roadmap for Getting There	91
8.5.1	Forecast Services Verification	91
8.5.2	Forecast Verification	92
8.6.	References	93
9.	Social Science Research (Future)	95
	Appendix A- Review of higher-level NOAA guidance and NRC recommendations	97
A-1	The NOAA Hydrology Program Core Goals	97
A-2	NOAA Strategic Plan for FY 2006-2011	98
A-3	NOAA Draft Five-Year Research Plan for FY 2007-2011	98
A-4	NOAA Twenty-Year Research Vision	100
A-5	The NOAA Annual Guidance Memorandum for FY 2010-2014	101
A-6	NWS Strategic Plan for FY 2005-2010	102
A-7	NWS 2004 Science and Technology Infusion Plan	104
A-8	National Research Council Reports	105
A-9	References	108
	Appendix B- : Hydrologic Modeling Literature Review Paper	111
B-1	Introduction	111
B-2	Defining Distributed and Physically Based along a Continuum	112
B-3	The State of Distributed Datasets	113
B-4	Arguments for Moving towards more Physically Based Approaches	115
B-5	Limitations of Operational Use of Complex, Highly Distributed Models	116
B-6	New and Revisited Paradigms for Physically Based Modeling	118
B-7	Examples of Distributed Models in Current Use	123
B-8	Results from Model Intercomparison Projects	126
B-8.1	The Distributed Model Intercomparison Project	126
B-8.2	The Model Parameter Estimation Experiment and A Priori Parameter Estimation	130
B-8.3	The North American Land Data Assimilation System	134
B-8.4	Other Model Intercomparison Studies	138
B-9	Summary and Conclusions	139
B-10	References	141

Figures

Figure 1-1.	Major science elements addressed by the Strategic Science Plan. (Adapted from Figure 2 of the Integrated Water Science Plan)	5
Figure 2-1.	Conceptual figure demonstrating how advances in predictability science transition to improved operational predictions.	12
Figure 3-1	Simplified Structure of the SAC-HT	19
Figure 3-2	Example 4km, Hourly Computed Volumetric Soil Moisture Grid	19
Figure 6-1	Uncertainty in hydrologic forecasts as a function of forecast horizon	59
Figure 6-2	Spatial Scales for Hydrologic Modeling	60
Figure 6-3	Uncertainties in Hydrologic Forecast	61
Figure 6-4	Current Experimental Operations of the Short Term Ensembles and Variational Data Assimilation (VAR)	62
Figure 6-5	Vision for Ensemble and Data Assimilation in Hydrologic Forecast Operations	65
Figure 6-6	Multi-model Ensemble Framework	66
Figure 6-7	XEFS Architecture	68
Figure 6-8	Hydrology Test Bed	69
Figure 6-9	Climate-to-Water Research-to-Operations Pathways	70
Figure 7-1	Data Assimilation Strategy	78
Figure 8-1	Setup of the verification system to evaluate the processing steps of the forecasting system (final team report, 2006)	83

Abbreviations and Acronyms

Abbreviation or Acronym	Meaning
2DVAR	Two-dimensional VAR
4DVAR	Four-dimensional VAR
ADSTAT	Advective-Statistical System
AFFDEF	A Spatially Distributed Grid Based Rainfall–Runoff model
AHPS	Advance Hydrologic Prediction Service
API	Antecedent Precipitation Index
AR	Auto Regressive
ARIMA	Auto Regressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
AWIPS	Advanced Weather Interactive Processing System
BAMS	Bulletin of the American Meteorological Society
CADSWES	Center for Advanced Decision Support for Water and Environmental Systems
CASA	Collaborative Adaptive Sensing of the Atmosphere
CASC2D	CASCade 2 Dimensional
CATFLOW	A physically based, distributed model for small catchments
CBRFC	Colorado Basin River Forecast Center
CHPS	Community Hydrologic Prediction System
CFS	Climate Forecast System
CIFLOW	Coastal and Inland Flooding Observation and Warning
CIRES	Cooperative Institute for Research in Environmental Sciences
CNRFC	California-Nevada River Forecast Center
CONUS	Conterminous United States
CO-OPS	Center for Operational Oceanographic Products and Services
CPC	Climate Prediction Center
CPPA	Climate Prediction Project for the Americas
CREW	Cooperative Community Catchment model based on REW
CRN	Climate Reference Network
CWMS	Corps Water Management System
DA	Data Assimilation
DEM	Digital Elevation Model
DHM	Distributed Hydrology Model
DHSVM	Distributed Hydrology-Soil-Vegetation Model
DMIP	Distributed Model Intercomparison Project
DPC	Dominant Processes Concept
DSS	Differential Split Sample
EC	Environment Canada
EFFS	European Flood Forecast System
EMC	Environmental Modeling Center
EnKF	Ensemble Kalman Filter

Abbreviations and Acronyms

ENSO	El Niño Southern Oscillation
EPP	Ensemble Post-Processor
EPP2	Ensemble Pre-Processor II
ESMA	Explicit Soil Moisture Accounting
ESP	Ensemble Streamflow Prediction
ESRL	Earth System Research Laboratory
ET	Evapotranspiration
EVS	Ensemble Verification System
FFG	Flash Flood Guidance
FH69	Freeze and Harlan's 1969 blueprint
GCM	Global Climate Model
GEFS	Global Ensemble Forecast System
GFS	Global Forecast System
GFZ	GeoForschungsZentrum
GIS	Geographic Information System
GLDAS	Global Land Data Assimilation System
GLUE	Generalized Likelihood Uncertainty Estimation
GOES	Geostationary Operational Environmental Satellite
GPM	Global Precipitation Mission
GR4J	Génie Rural, 4 paramètres
GR5H	A modified version of the GR4J model
GRACE	Gravity Recovery and Climate Experiment
GSFC	Goddard Space Flight Center
GSFLOW	Groundwater/Surface-Water Flow
GSSHA	Gridded Surface/Subsurface Hydrologic Analysis
GSWP	Global Soil Wetness Project
GVF	Green Vegetation Fraction
HAS	Hydrometeorological Analysis and Support
HEC-HMS	US Army Corps of Engineers' Hydrologic Engineering Center Hydro- logic Modeling System
HEH	Hydrologic Ensemble Hindcaster
HEPEX	Hydrologic Ensemble Prediction EXperiment
HL	Hydrology Laboratory
HL-RMS	Hydrology Lab-Research Modeling System
HMT	Hydrometeorological Testbed
HOSIP	Hydrologic Operations and Service Improvement Process
HPC	Hydrometeorological Prediction Center
HRAP	Hydrologic Rainfall Analysis Project
HRC	Hydrologic Research Center
HRCDM	Hydrologic Research Center Distributed Model
HRU	Hydrologic Response Unit
HSD	Hydrologic Services Division
HSEB	HL's Hydrologic Software Engineering Branch
HSPF	Hydrological Simulation Program - FORTRAN
HYDROTEL	A Spatially Distributed Rainfall–Runoff model specifically designed to take advantage of GIS and remote sensed data
ICP	Interactive Calibration Program

Abbreviations and Acronyms

IDMA	Interactive Double Mass Analysis
IHAC	A modified version of IHACRES
IHACRES	Identification of unit Hydrographs and Component flows from Rain-falls, Evaporation and Streamflow data
IHFS-DB	Integrated Hydrologic Forecast System database
Inform	Integrated Forecast And Reservoir Management
ISBA	Interactions between Soil–Biosphere–Atmosphere
IVP	Interactive Verification Program
IWSP	Integrated Water Science Plan
JWGV	WMO Joint Working Group on Verification
LIS	Land Information System
LISFLOOD	A Two-dimensional hydrodynamic model
LSM	Land Surface Model
LSPS	Land Surface Parameterization Scheme
LULC	Land Use-Land Cover
MADIS	Meteorological Assimilation Data Ingest System
MAP	Mean Areal Precipitation
MDL	Meteorological Development Laboratory
MODFLOW	U. S. Geological Survey Modular Ground-Water Model
MODIS	Moderate Resolution Imaging Spectroradiometer
MODs	Run-time Modifications
MODSPA	A Spatially Distributed Rainfall–Runoff model\
MOPEX	Model Parameter Estimation Experiment
MORDOR	A rainfall runoff model
MOS	Model Output Statistics
MPE	Multisensor Precipitation Estimator
MPN	Multisensor Precipitation Nowcaster
NAEFS	North American Ensemble Forecast System
NAM	North American Mesoscale model
NASA	National Atmospheric and Space Administration
NATSGO	National Soil Geographic database
NCAR	National Center for Atmospheric Research
NCDC	National Climate Data Center
NCEP	National Centers for Environmental Prediction
NCRFC	North-Central River Forecast Center
NDFD	National Digital Forecast Database
NDVI	Normalized Difference Vegetation Index
NESDIS	National Environmental Satellite Data and Information Service
NEXRAD	Next Generation Radar
NLCD	National Land Cover Database
NLDAS	North American Land Data Assimilation experiment
NMQ	National Mosaic and Multi-Sensor QPE
NOAA	National Oceanic and Atmospheric Administration
NOAA CREST	Cooperative Remote Sensing Science and Technology Center
NOHRSC	National Operational Hydrologic Remote Sensing Center
NOMADS	NOAA National Operational Model Archive and Distribution System

Abbreviations and Acronyms

NPOESS	National Polar-Orbiting Operational Environmental Satellite System
NRC	National Research Council
NRCS	Natural Resource Conservation Service
NS	Nash-Sutcliffe efficiency
NSSL	National Severe Storms Laboratory
NWP	Numerical Weather Prediction
NWRFC	Northwest River Forecast Center
NWS	National Weather Service
NWSRFS	National Weather Service River Forecasting System
OCWWS	Office of Climate, Water and Weather Services
OHD	Office of Hydrologic Development
PB	Proxy basin
PDO	Pacific Decadal Oscillation
PET	Potential Evapotranspiration
PILPS	Project for Intercomparison of Landsurface Parameterization Schemes
PQPF	Probabilistic QPF
PRISM	Precipitation-elevation Regressions on Independent Slopes Model
PSD	Physical Sciences Division
PUB	Prediction in Ungauged Basins
Q2	Next Generation QPE
QPE	Quantitative Precipitation Estimate
QPF	Quantitative Precipitation Forecast
QTE	Quantitative Temperature Estimate
QTF	Quantitative Temperature Forecast
R&D	Research and Development
RAX	RFC Archive machines
RES-J	Joint Reservoir Regulation Operation Program
ResSim	USACE Reservoir System Simulation Program
RES-SNGL	NWS Single-reservoir model Program
REV	Representative Elementary Volume
REW	Representative Elementary Watersheds
RFC	River Forecast Center
ROC	NWS Radar Operations Center
RTi	Riverside Technology, Inc.
RTMA	Real-time Mesoscale Analysis
RTO	Research-to-Operations
RUC	Rapid Update Cycle
SAC-HT	Sacramento Heat-Transfer model
SACSMA	Sacramento Soil Moisture Accounting model
SAR	Synthetic Aperture Radar
SCAN	Soil Climate Analysis Network
SCAN	System for Convection Analysis and Nowcasting
SHE	Systeme Hydrologique Europeen
SMAP	Short and Medium-term Priority Environmental Action Programme
SNODAS	Snow Data Assimilation System

Abbreviations and Acronyms

SNOTEL	Snowpack Telemetry
SNTHERM	SNow THERmal Model
SREF	Short-Range Ensemble Forecasts
SS	Split Sample
SSARR	Streamflow Simulation and Reservoir Regulation System
SSARRESV	SSARR Reservoir Regulation Operation
SSG	Strategic Science Goals
SSHP	Site-Specific Hydrologic Prediction system
SSST	Snow Science Steering Team
SSURGO	Soil Survey Geographic database
STAR	Center for Satellite Applications and Research
STATSGO	State Soil Geographic database
SVAT	Soil-Vegetation-Atmosphere Transfer
SWAP	Soil Water–Atmosphere–Plants
SWAT	Soil Water Assessment Tool
SWB	Simple Water Balance
SWE	Snow Water Equivalent
SWPC	Space Weather Prediction Center
TACD	Tracer-Aided Catchment Model
TDWR	Terminal Doppler Weather Radar
THORPEX	The Observing System Research and Predictability Experiment
TIN	Triangular Irregular Networks
TOPKAPI	Topographic Kinematic Approximation and Integration
TOPLATS	TOPMODEL-based Land-Atmosphere Transfer Scheme
TOPMO	An 8 parameter modified version of TOPMODEL
TOPNET	A Networked version of TOPMODEL
tRIBS	TIN-based Real-time Integrated Basin Simulator
TRMM	Tropical Radar Rainfall Measurement Mission
TVA	Tennessee Valley Authority
UHG	Unit Hydrograph
USACE	Unites States Army Corps of Engineers
USBR	Unites States Bureau of Reclamation
USDA	United States Department of Agriculture
USGS	United States Geological Survey
VAR	Variational Assimilation
VIC	Variable Infiltration Capacity
WAPA	Western Area Power Administration
WaterNet	Water Cycle Solutions Network
WATFLOOD	Waterloo Flood Forecast Model
WFO	Weather Forecast Office
WGRFC	West-Gulf River Forecast Center
WMO	World Meteorological Organization
WR	NWS Western Region
WSR 88D	Weather Surveillance Radar-1988 Doppler
XEFS	Experimental Ensemble Forecast System

1. Introduction

This National Weather Service (NWS) Office of Hydrologic Development (OHD) Strategic Science Plan identifies opportunities to meet research and operational goals set forth in National Oceanic and Atmospheric Administration (NOAA) and NWS guidance.

The overarching goal of this Strategic Science Plan is to provide the direction of research for OHD as well as for OHD-funded collaborative research for a period of 5 to 10 years, and, thus, to maintain the NWS as the world leader in hydrologic forecasting by using state-of-the-art science and technology. This plan is not intended to be a static road map that should, after that period, be abandoned in favor of a new plan. Rather, it is meant to be a dynamic plan to be updated on an annual basis, according to results of research, availability of data resources, user requirements, etc.

1.1 The Role of OHD within the NOAA Mission Goals

1.1.1 NOAA Mission

To understand and predict changes in Earth's environment and conserve and manage coastal and marine resources to meet our Nation's economic, social, and environmental needs.

The NOAA Strategic Plan for FY 2006-2011 organizes the scientific activities of the agency around the following four mission-directed goals:

- Protect, Restore, and Manage the Use of Coastal and Ocean Resources through an Ecosystem Approach to Management
- Understand Climate Variability and Change to Enhance Society's Ability to Plan and Respond
- Serve Society's Needs for Weather and Water Information
- Support the Nation's Commerce with Information for Safe, Efficient, and Environmentally Sound Transportation

While OHD activities are directed primarily at fulfilling NOAA's Weather and Water Mission Goal through their support of streamflow forecasting at the NWS River Forecast Centers (RFCs) and the Weather Forecast Offices (WFOs), they also contribute to the fulfillment of the other three mission goals. For example, stream water quality and quantity forecasts are critical to the objective of the Ecosystem Goal to forecast algal blooms and hypoxia in coastal waters. OHD research is also increasingly intertwined with the Climate Goal as, under the Advance Hydrologic Prediction Service (AHPS), the RFCs have begun to produce probabilistic long-term streamflow forecasts based on sub-seasonal to seasonal

climate forecasts from the NOAA Climate Prediction Center (CPC). Finally, OHD's flash-flood and streamflow forecasting research and product development help to meet the Commerce and Transportation Goal through the impact they have on the nation's ability to anticipate the conditions of both land and riverine transportation routes.

1.2 OHD Strategic Science Goals

The following Strategic Science Goals (SSG) summarize the requirements identified in the rest of the document, as well as higher-level NOAA and NWS guidance and two recent NRC reports that discuss OHD research and product development. *Appendix A* describes in detail the connections between these goals and those documents.

1. Leverage outside research, especially that from U. S. Government and academic partners;
2. Use the latest information technology, especially with regard to the manipulation and display of high-resolution spatial data;
3. Improve the quality of physical inputs and forcings;
4. Use new datasets, especially from remote-sensing, after the utility and quality of the datasets are established;
5. Evaluate and implement more complete and sophisticated data assimilation;
6. Reduce and quantify the uncertainty in forecasts at all timescales;
7. Verify both deterministic and probabilistic forecasts;
8. Better account for the impact of reservoirs and other forms of stream regulation (e.g., withdrawals, return flows and groundwater pumping) on streamflow forecasts;
9. Increase the lead times and accuracy of warnings and forecasts, especially for flash floods;
10. Produce climate timescale hydrologic forecasts;
11. Evaluate and implement new, higher resolution distributed models;
12. Support the integration and coupling of hydrologic models with weather and climate models;
13. Enlarge the suite of forecast elements, including soil moisture, low flows and water quality parameters;
14. Support coastal and marine ecosystem forecasting;
15. Better understand and respond to user needs;
16. Provide more explicit and targeted decision support;
17. Consider the role of climate change in hydrologic forecasting at all time scales.

The order of the above SSG is not intended to reflect priority, but rather provide a certain amount of logical grouping. For example SSG 1 is crosscutting in that all the other SSG will meet with greater success the more thorough are OHD collaborations. SSG 2-4 all involve the use of new, more accurate and more highly resolved data. SSG 5-7 deals with the rapidly increasing production of probabilistic forecasts. SSG 8-10 are focused on improving both the accuracy and lead time of forecasts at all time scales. SSG 11 and 12 relate to the movement towards more highly resolved and integrated models. SSG 13 and 14 emphasize the need for new forecast products (although a discussion of water quality and ecosystem forecasts are left to a future version of the plan). SSG 15 and 16 reflect the growing understanding of the need to engage users to a greater extent if NWS forecasts are to provide their greatest potential societal benefit. The climate change issue addressed by SSG 17 is perhaps the most challenging of all as the many changes to the Earth system that global warming is thought to be causing are only beginning to be fully understood. Consequently, climate change is addressed in the present version of the plan in only a cursory manner but will become an increasingly larger part of future versions as climate change science continues to mature.

1.3 Organization of the Plan

In order to put the rest of the plan into a single, coherent scheme, it is important to express up-front how the operational hydrologic forecasting process is currently done, and what the plan proposes for it in the future. Accordingly, *Chapter 2* presents a high-level view of both the current and future hydrologic forecasting systems in the NWS.

The Integrated Water Science Plan (IWSP; NWS, 2004) has an excellent description of the combination of processes that form part of the hydrologic cycle, and we chose to loosely follow the structure of that document. Some of those processes occur in the atmosphere, some on land, and some in the oceans. Forecasting river flows require that those processes be properly observed, forecasted, or otherwise estimated, including their inherent errors and uncertainties. *Chapters 3* through *5* of this plan will essentially follow *Chapter 4* and Figure 2 of the IWSP. Specifically, *Chapter 3* is devoted to the main storages and fluxes in hydrologic models; *Chapter 4* to the atmospheric forcings of surface hydrologic processes; and *Chapter 5* to anthropogenic and natural large-scale disturbances to and non-stationarities in hydrologic processes. Given that a major focus (although not the only one) of this Strategic Science Plan will be the research, evaluation and transition to operations of a suite of new distributed and physically based hydrologic models, a comprehensive literature review of such models was undertaken, a draft of which is included in the *Appendix B*. Some of the major findings and recommendations from that paper are summarized in *Section 2.2*, with the complete paper attached in the *Appendix B*. Part of the strategy involved in the research and implementation of physically based models is covered

by the Science Plan of Phase 2 of the Distributed Model Intercomparison Project (DMIP) (Smith *et al.*, 2006). **Section 2.2** also addresses the continuing need for lumped and conceptual models.

Section 4 of the IWSP also discusses streamflow. Although the hydraulic modeling of stream channels is an essential component of any flood forecasting system, streamflow processes and associated issues of water quality will be addressed in future releases of this plan. Those future releases will include a section on hydraulics and routing; the need for higher dimensional models to address problems of wind effects on shallow bodies of water and levee breaches; inundation mapping; coastal zone linkages; ice formation and breakup; dambreak scenarios; water diversion, irrigation, returns and aquifer pumping. The exception in this version of the plan is the impact of reservoir regulation on streamflow, which will be addressed in **Section 5.3**.

A number of the Strategic Science Goals identified in **Section 1.2** involve the overarching issues of the quantification and reduction of uncertainty in observations and forecasts. Accordingly, **Chapters 6, 7, and A-8** respectively address the related topics of ensemble modeling, data assimilation, and verification.

Finally, a future version of the plan will cover the directions for Social Science research in **Chapter 9**. OHD has already sponsored work in the Social Sciences, namely determining the economic value of water resources forecasting, and research on the improvement of web pages' probabilistic information.

Figure 1-1 of this plan (which is a revised version of Figure 2 in the IWSP) illustrates the overall structure of the Plan. Each Chapter is divided into the major processes and variables involved. For each process or variable, a subsection discusses the science and technology for observations, forecasting (i.e., modeling) or both. That discussion will come under five headings: where we are, what our partners are doing, where we want to be over the life of the plan (i.e., 5-10 years), what the challenges are, and a road map for getting there.

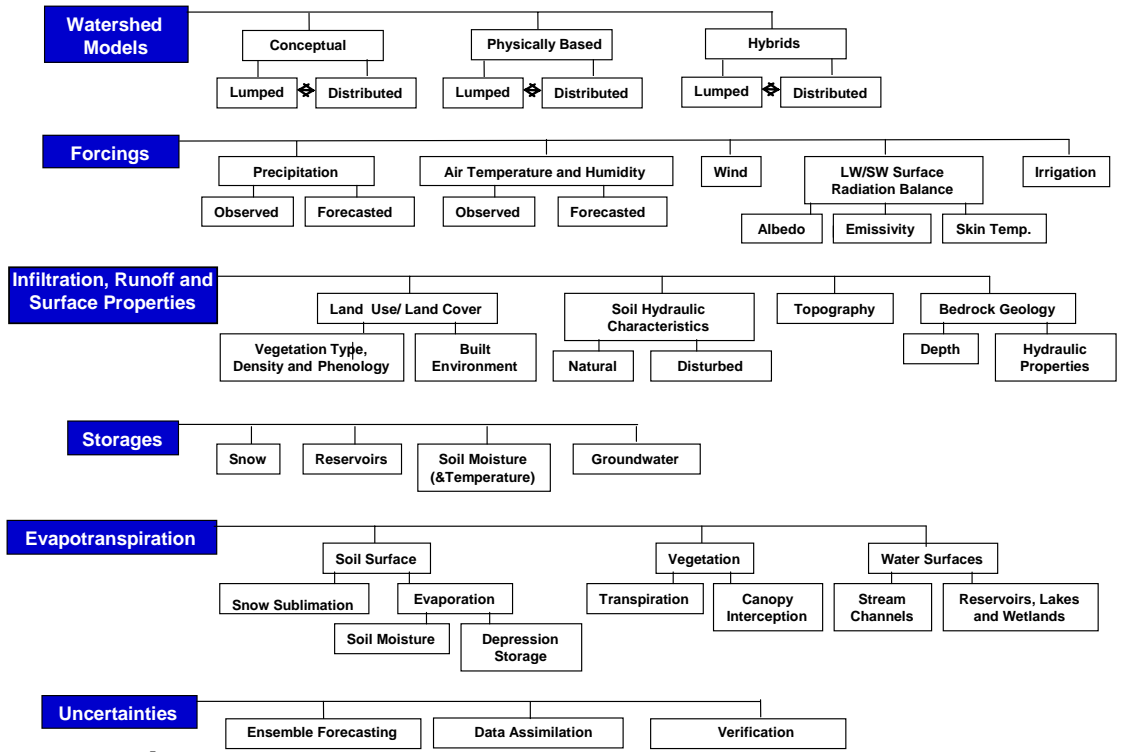


Figure 1-1. Major science elements addressed by the Strategic Science Plan. (Adapted from Figure 2 of the Integrated Water Science Plan)

1.4 References

NWS, 2004. The Integrated Water Science Plan, Report of the NWS IWSP team.

Smith, M., V. Koren, S. Reed, Z. Zhang, D. J. Seo, F. Moreda and Z. Cui, 2006.
The Distributed Model Intercomparison Project: Phase 2 Science Plan.
NWS Office of Hydrologic Development.

2. A High-Level View

In order to put the rest of the plan into a single, coherent scheme, it is important to express up-front how the operational streamflow forecasting process is currently done, and what the plan proposes in the future.

2.1 The Current Hydrologic Forecasting Process

Hydrologic forecasting in the National Weather Service is performed by a complex system of data management tools that combine observations and forecasts of a number of atmospheric processes in order to provide input to a suite of mathematical models of watershed processes, and to a suite of time series and data manipulation techniques that comprise the operations available in the National Weather Service River Forecasting System (NWSRFS).

Observations used in the forecasting process include precipitation, temperature, snow water equivalent, freezing level and snow cover from a variety of sensors, the details of which are described in the corresponding sections of the Plan. As far as potential evapotranspiration is concerned, current forecasts at some of the RFCs use long-term climatological means as opposed to actual observations. Quality assurance on the rain gauge observations is performed at each RFC by the Hydrometeorological Analysis and Support (HAS) forecaster.

At the heart of the land surface models are two lumped models: the Sacramento Soil Moisture Accounting (SAC-SMA) model, and the conceptual snow model known as SNOW 17. There are other models developed and made operational for the land surface process, but those two conceptual models are, by far, the most used by the River Forecasting Centers.

Despite the high-quality of the forecasts produced with lumped conceptual models, a major shortcoming is the need for long time series of good-quality observations required for calibration, and the high-level of training required to perform a good calibration. We should point out that calibration is a major step in the use of any model, hydrologic or hydraulic, lumped or distributed. Furthermore, many RFCs regard model calibration as a very important step in the development of good hydrologic forecasters (Smith *et al.*, 2003). Although the difficulties posed by the model calibration requirement have been somewhat alleviated by the development of prior parameter values (Koren *et al.*, 2003) and tools such as the Interactive Calibration Program (ICP) and the Interactive Double Mass Analysis (IDMA) tool, there is still the need to fine-tune those parameters either by manual or automatic calibration. Calibration on historical data is also problematic for non-stationary condition such as induced by climate and land-use/land-cover changes (see *Chapter 5*). Furthermore, apart from subdividing watersheds into elevation zones, the lumped models in the NWSRFS do not account for the dis-

tribution of forcings, surface properties, and runoff processes within the watershed.

A major step forward in accounting for subwatershed heterogeneities has been achieved with the development of a gridded distributed model that uses SAC-SMA for rainfall-runoff calculations, and kinematic overland flow and channel routing (Koren et al, 2004). These gridded modeling techniques are part of the Distributed Hydrologic Model (DHM), an operational component of the NWSRFS. Gridded implementations of SNOW-17 are also being evaluated. As demonstrated by results from the first phase of results of Phase 1 of the Distributed Model Intercomparison Project (DMIP1; Reed *et al.*, 2004), the DHM is capable of providing more accurate streamflow simulations than the NWS lumped SAC-SMA under certain circumstances, as well as forecasts at interior points. In addition, recent work with research versions of the DHM, modified versions of the SAC-SMA model and *in-situ* observations of soil moisture suggests that such observations may be helpful in calibrating the model and making it useful for the forecasting of soil moisture (Koren *et al.*, 2007) and data assimilation (Seo *et al.* 2007). Similarly, the North American Land Data Assimilation experiment (NLDAS, Mitchell *et al.* 2004), SnowMIP 1, 2, and Franz (2006) showed that the SNOW-17 model performed well in evaluations with energy budget and other snow models. Concurrently with this first version of the plan, a team is now actively working on defining the high-level requirements for an enhanced version of the NWS distributed hydrologic model.

(http://www.weather.gov/oh/rfcdev/projects/rfcDMOCT_chart.html)

In addition to the deterministic forecast process described above, under the Advanced Hydrologic Prediction Service (AHPS), the River Forecast Centers are implementing probabilistic forecasts based on an ensemble approach (Ensemble Streamflow Prediction or ESP). This approach uses historical observations of precipitation and temperature to produce medium term and seasonal forecasts of water supply, conditional to current conditions on the watershed. Furthermore, OHD is actively investigating short-term ensemble forecast techniques. More information about the current state of ensemble forecasts can be found in *Chapter 6*.

2.2 The Future Hydrologic Forecasting Process

The future of hydrologic forecasting at the NWS will include:

- Enhanced use of remotely sensed information on a wide range of atmospheric and land-surface characteristics, from both active and passive satellite-based sensors;
- Higher-resolution models;

- Explicit consideration of the uncertainty in the forcings and forecasts (An ensemble approach is currently being pursued and will be fully implemented for short-, medium- and long-term forecasting);
- Multi-model ensembles to address the problem of uncertainty in the forecasts arising from structural errors in the models (These ensembles may be formed by combinations of lumped or distributed, conceptual or physically based models);
- Explicit consideration of the errors introduced by sub-optimal parameter values and initial conditions;
- Data assimilation of *in-situ* and remote-sensed state variables; and
- Verification of single-value (deterministic) and ensemble (probabilistic) forecasts.

These characteristics of the future hydrologic forecasting process are reflected in the SSG identified in **Section 1.2** and can be thought of as synthesis of the longer list of SSG.

Many distributed, physically based models have been developed with the intent of improving the accuracy of hydrologic forecasts, in general, and minimizing the need for model calibration, specifically. Furthermore, it often has been the expectation that those models should be able to adapt their parameters to physical changes to a watershed, such as those resulting from large-scale disturbances to soil and vegetation such as occur during forest fires, without having to resort to recalibration.

By physically based, we mean models for which the parameters can be directly estimated, and the processes closely mimic those actually occurring in the watershed. Since not all properties of the watershed can be directly observable, it would be necessary for some of those parameters to be estimated by calibration. Similarly, some of the processes may be more efficiently modeled by a conceptual or an empirical approach. It follows then that models that blend purely physical parameters and processes with conceptual processes and calibrated parameters may be most appropriate. One example of models in this class is the Sacramento-Heat transfer model (SAC-HT) with *a priori* estimates for some of the original Sacramento model parameters, and a physically based model for the heat-transfer portion. By distributed models we mean any model that does not consider a watershed as a lumped system. This includes models in which a watershed is divided into regular or irregular grids, subwatersheds, hydrologic response units, Representative Elementary Watersheds (REW), etc. The various gradations and definitions used in the literature to talk about distributed and physically based models are discussed in greater detail in the **Appendix B**.

A number of studies, including the results of the DMIP-1 and initial results from the DMIP-2 indicate that physically based models have largely fallen short of

their goals as operational tools for a number of reasons. The limitations of models that are in particular highly distributed and physically complex are discussed in *Appendix B* and references therein and include:

- The models are typically based on small-scale hydrologic theory and thereby fail to account for larger-scale processes such as preferential flow paths;
- The data necessary to estimate parameter values are not available at high enough resolution, certainty, or both;
- The data necessary to drive the models are not available at high enough resolution, certainty or both; and
- Despite the rapid increase in computer power and decrease in hardware costs, the computational demands are still a barrier, particularly for performing data assimilation and ensemble modeling in real-time.

The operations and research communities are steadily making progress towards resolving the above limitations. Nevertheless, one thing is clear: the most highly resolved and most physically complex models are not necessarily the most appropriate for operational hydrologic forecasting for the foreseeable future. At the same time the anticipated benefits of models that are more highly resolved and more physically based than the lumped conceptual models that have been the mainstay of hydrological forecasting for the last several decades should be investigated further. Some progress in this direction has already been achieved. Recent research (e.g. DMIP) shows promising results for approaches that combine physically based and conceptual model components such as the DHM and the REW approach (see *Appendix B*). Continued research and development along these lines offers the potential for:

- More accurate forecasts in ungauged and poorly gauged basins;
- More accurate forecasts after changes in land use and land cover, such as forest fires and other large-scale disturbances to soil and vegetation;
- More accurate forecasts under non-stationary climate conditions;
- Modeling of interior states and fluxes, which are critical for forecasts of water quality, soil moisture, land slides, groundwater levels, low flows, etc.; and
- The ability to merge hydrologic forecasting models with those for weather and climate forecasting.

The above-anticipated benefits to more highly resolved and physically based models and the advances in data availability and modeling methodologies that are likely to continue to lead to their realization, are discussed in *Appendix B* and references therein. How these advances will be utilized and furthered by OHD is a major thrust, although not the only one, of the Plan. It is recognized that distributed, physically based modeling is not an end itself, but rather must be evaluated in recognition of operational requirements and capacities, and in the im-

provements achieved and cost effectiveness. Accordingly, OHD will focus on models that:

- Make use of the prior estimation of parameter values from existing distributed datasets;
- Have parameter sets that can be initially calibrated and then adjusted to account for changes in watersheds and stream channels in a computationally efficient, physically meaningful and robust manner;
- Are amenable to the assimilation and forecasting of both streamflow and internal watershed states (e.g., soil moisture and groundwater levels);
- Are amenable to ensemble modeling (including multi-model ensembles) and other forms of uncertainty propagation;
- Are appropriate for the resolution and certainty of both the observed and forecasted atmospheric forcings;
- Are appropriate for hydrologic forecasting across a range of space and time scales;
- Are amenable to real-time forecasting at NWS field offices given realistic levels of computer resources, personnel and training, and
- Work within the Community Hydrologic Prediction System.

It is very likely that no single model will be capable of meeting all of the above requirements, and so this plan envisions a suite of models, including distributed and lumped models, that will be integral components of the hydrologic forecasting system for the foreseeable future.

In evaluating any new technology—hydrologic or otherwise—against a well-established one, it is critical to recognize that there is almost always an unavoidable period of maturation before the new technology reaches its full potential. This process in the context of paradigm shifts in hydrologic forecast systems is very well illustrated in **Figure 2-1**, which is modified from a figure in the National Research Council (NRC) Report of a Workshop on Predictability and Limits-to-Prediction in Hydrologic Systems (NRC, 2002). Quoting from the caption of the latter figure:

“The effectiveness of the predictions is measured with a skill score. Prediction systems are based on existing paradigms or scientific understanding. Initially the system has a slow rate of increase in skill. Errors in implementation, uneven completion of auxiliary systems, and gradual training of personnel in the prediction system are some of the reasons why the initial increase in prediction skill can be modest or even negative. As the prediction system matures it undergoes a period of rapid improvement in its effectiveness. As the prediction system and its supporting science paradigm mature, the system again experiences slower rates of skill increase with time. In this phase the prediction system has essentially

reached its highest potential for characterizing and predicting hydrologic phenomena.”

While one might argue about the exact shape of the trajectory of the increase in skill of a prediction system based on a new paradigm, comparing its forecast skill to a system based on a better-established paradigm will likely provide an unfairly pessimistic view of the ultimate potential of the new system. Therefore, transferring from the existing paradigm once it reaches or approaches maturity, to a new paradigm, may result in a decrease in skill. However, once the technology matures its forecast skill should overpass that of the existing one. Some examples of new paradigms in hydrologic forecasting that are showing promise of following this trajectory include: multi-model ensemble forecasts, multivariate and distributed parameter calibration, and assimilation of distributed and multi-variate data.

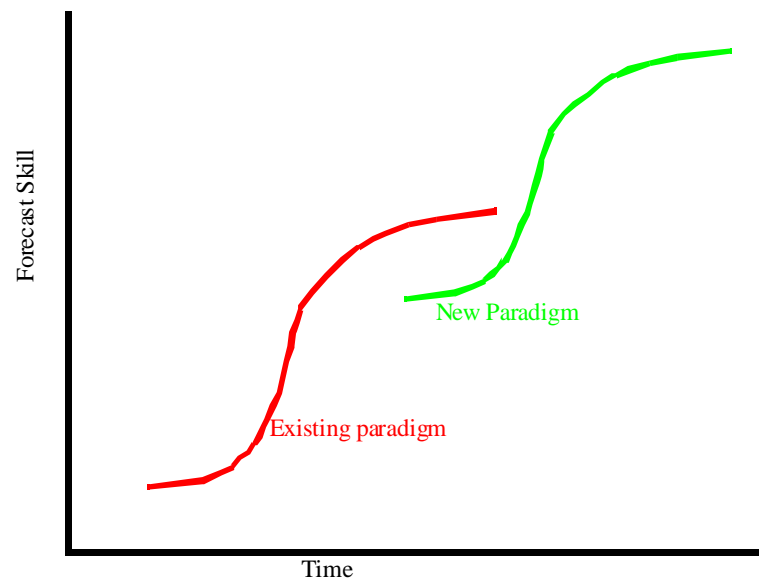


Figure 2-1.

Conceptual figure demonstrating how advances in predictability science transition to improved operational predictions. Adapted from NRC (2002).

2.3 References

Franz, K.J. (2006). “Characterization of the comparative skill of conceptual and physically-based snow models for streamflow prediction.” PhD dissertation, U. California at Irvine, CA.

Koren, V., S. Reed, Z. Zhang, D. Seo, F. Moreda, Kuzmin, V. (2003). “Use of spatially variable data in river flood prediction.” *AGU-EGS-EUG Assembly*, April 9-14, Nice, France.

- Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.-J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US national weather Service. *Journal of Hydrology*, Vol. 291, 297-318.
- Koren, V., Fu, F., and Zhang, J. (2007). "Results from a Simple Parameterization of the Snow-Vegetation Interaction in Energy-Budget and Temperature Index Models." *IUGG Meeting*, Perugia, Italy.
- Mitchell, K.E. and co-authors, 2004. The Multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *JGR*, Vol., 109, D07S90, doi:10.1029/2003JD003823
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and DMIP Participants. (2004). "Overall distributed model Intercomparison project results." *J. of Hydrol.*, Vol. 298, Nos. 1-4, 27-60.
- Seo, D.-J., L. Cajina, R. Corby, and T. Howieson. (2007). "Automatic State Updating for Operational Streamflow Forecasting via Variational Data Assimilation." Submitted to *J. of Hydrol.*
- Smith, M.B., Laurine, D.P., Koren, V.I., Reed, S.M., and Zhang, Z., 2003. Hydrologic Model Calibration in the National Weather Service. In: *Calibration of Watershed Models, Water Science and Application Volume 6*, AGU Press, 133-152.

3. Watershed Models

This Chapter covers watershed properties (surface properties), fluxes (infiltration, surface runoff, base flow, snow sublimation and melt and evapotranspiration), and natural storages (soil moisture, groundwater, snow accumulation). The chapter contains place holders for future sections on hydraulic and water quality modeling.

3.1 Surface Properties

The feasibility of operational use of the new suite of models envisioned by this Plan largely depends on the availability of highly resolved, accurate, and nationwide observations of land-surface properties. These properties include: albedo, land use/land cover (especially vegetation type, density and phenology, but also including features of the built environment), soil characteristics, topography, and bedrock geology. The current availability of such datasets is discussed in *Appendix B*.

3.1.1 Where We Are

Currently, we make very limited use of the available datasets of surface properties. For example, as noted in *Section 2.1*, the lumped models in the NWSRFS, use topographic data only for subdividing watersheds into elevation zones. However, the DHM does make greater use of datasets of surface properties, especially soil characteristics in its *a priori* parameterization scheme. The Noah model also makes use of many coarse-resolution datasets. However, all the current models in use at OHD fail to make use of the data at their highest available resolution, especially topography, soils and vegetation, which are the principal controls of surface runoff.

3.1.2 What Our Partners Are Doing

As discussed in *Appendix B*, the remote sensing community continues to develop new, more accurate and more highly resolved remotely sensed datasets. In addition, the NRCS has nearly completed digitizing the county soil surveys into the SSURGO database.

3.1.3 Where We Want to Be

In general, OHD should make optimal use of state-of-the-art datasets of surface properties, particularly those for soils, topography and land-use/land-cover.

3.1.4 Challenges to Getting There

As noted in *Section 2.2* and *Appendix B*, databases of surface properties are often not as accurate as their resolution implies. And so, as with all data involved in hydrologic forecasting, dealing with uncertainty is a key challenge. Much of that uncertainty arises in translating surface properties into model parameters, as for example when using pedotransfer functions to translate soil textural properties to soil hydraulic parameters, particular for non-standard conditions such as hydrophobic soils and highly disturbed soils, as for example resulting from fire, urban development, cultivation or overgrazing. A strategy for capturing the impact of land-use and land-cover changes on hydrologic processes will be included in *Section 5.2* of a future version of the Plan.

In some cases, the desired datasets are simply not available. The most significant are those dealing with the subsurface, such as deeper soil layers and bedrock depth and hydraulic properties.

3.1.5 A Road Map for Getting There

OHD should develop and evaluate models that make the most of the available data on physical properties of the land surface. As discussed in *Appendix B*, those models may not be the most highly resolved ones, but should at least represent subgrid heterogeneities in a physically realistic manner (see the discussion in *Appendix B* of the Representative Elementary Watershed as one promising approach). For those data which observations are poor, OHD should consider model-based datasets. An example is the geomorphology model of Dietrich *et al.* (1995), which makes watershed scale predictions of soil depth. It has been used with some success in the Distributed Hydrology-Soil-Vegetation Model (DHSVM; Whitaker *et al.*, 2003). OHD should also build on the work of Koren *et al.* (2003) and thoroughly examine the extensive literature on pedotransfer functions (e.g., Elsenbeer, 2001).

3.2 Infiltration and Surface Runoff

Modeling infiltration correctly is necessary principally for the estimation of infiltration-excess runoff. Infiltration-excess and saturation-excess runoff form the two major components of fast responding overland flow. Downslope infiltration of overland flow can also be a major process in flood dynamics. In general, runoff is water that enters a stream channel either directly from overland flow or after storage and release from the soil or deeper subsurface. Hydrologists commonly estimate runoff by computing the water balance on a control volume bounded on top by the land surface, on the bottom by either an impermeable boundary or a sink boundary, and laterally by the stream channel boundaries.

3.2.1 Where We Are

As noted in *Section 2.1*, the Sacramento Soil Moisture Accounting Model (SAC-SMA) has proven to be an effective conceptual approach to modeling infiltration and runoff at large scales for operational forecasting. Although traditionally applied in lumped mode to watersheds larger than 300 km², recent research has shown effective implementation of a gridded SAC-SMA using 16 km² and 4 km² grid cells. A key development that has made this feasible is a technique to estimate *a priori* parameter values from soil and land use data (Koren *et al.*, 2003). These *a priori* parameter values describe physically meaningful patterns of parameter variation within a basin and from basin-to-basin within a region and serve as a starting point for model calibration. Simple techniques to calibrate gridded SAC-SMA models have proven effective. In a limited set of basins, forecasters have operationally begun to look at distributed model forecasts alongside lumped model forecasts to aid in decision-making. Improvements at forecast points are seen under certain conditions. Distributed SAC-SMA implementations show much promise for improving flash flood forecasts and providing new products such as gridded soil moisture and temperature (*Section 3.3*).

3.2.2 What Our Partners Are Doing

Working with our partners, OHD has been evaluating alternative model components, model structures, and model parameter estimation schemes through the DMIP and collaborative research agreements.

The National Centers for Environmental Prediction (NCEP) Environmental Modeling Center (EMC) is a key partner in our evaluation of hydrologic models. Although their primary goal is to derive accurate boundary conditions for atmospheric models (e.g. latent and sensible heat fluxes), analyzing the runoff produced by their model will help them evaluate if the other components are working properly.

Section B-8 details some of the work our partners are doing in model intercomparison studies, which are expected to yield valuable information towards specific model improvements.

3.2.3 Where We Want to Be

Our goals with respect to infiltration and runoff prediction are as follows:

- Quantify uncertainty in forecasts due to errors in models and forcings (*Chapters 6 , 7 and 8*);
- Improve the accuracy and reduce the uncertainty in our runoff models by making better use of distributed observations and forecasts of surface properties (*Section 3.1*), snow (*Section 3.5*), evapotranspiration (*Section 3.6*) and related forcings (*Chapter 4*);
- Predict runoff in ungauged basins;

- More easily and accurately account for land use and land cover changes including rapid changes in land cover due to wild fires, land use changes due to deforestation and urbanization, and seasonal vegetation changes (*Section 5.2*).

3.2.4 Challenges to getting to where we want to be

Achieving the above goals are hampered by the fact that:

- Infiltration and runoff are generally not directly observable at small scales, but rather inferred at large scales from stream gauging;
- We must rely on remote sensing for the estimation of soil moisture and other surface properties that control infiltration and runoff dynamics;
- The parameters of conceptual models tuned for application at one spatial and temporal scale are often not applicable at other scales;
- The literature does not show clear evidence that, currently, streamflow forecasts at well-gaged locations can be improved with more highly distributed and physically based approaches over lumped, conceptual ones (see *Appendix B*); and
- Different geographic regions may require different solutions. The viability of new methods will vary by geographic region due to differences in hydroclimatology, terrain, soils, land use, geology, data availability, and data quality.

3.2.5 A Road Map for Getting There

The following are waypoints on the roadmap for achieving the goals identified in *Section 3.2.2*:

- Generate runoff estimates at the highest resolution where informative results can be supported by our data;
- Provide forecasters with parallel lumped and distributed modeling tools;
- Continue efforts to develop improved model parameterization and calibration schemes for lumped and distributed models and to validate these schemes;
- More fully utilize existing *in-situ* and remote sensing observations to improve models and parameterizations and develop scientifically-based requirements for the next generation of sensors;
- Initiate large area model runs, including on a national domain on a 3-4 km grid;
- Build extensive model validation databases over diverse regions;
- Develop a hindcaster that will allow us to assess how improvements in runoff models translate into forecast improvements; and
- Continue collaborative model evaluation and development.

3.3 Soil Moisture and Temperature

Accurate representation of soil moisture and temperature—including their distribution in three dimensions—in hydrologic forecast models is important for many reasons. As a whole, soil moisture can be argued to be the state variable that has the most control on vadose-zone water and energy fluxes. Soil moisture in-and-of-itself is a variable worth forecasting. Soil moisture and temperature are also important controls on biotic processes and thus are important to water quality, agricultural and ecosystem forecasts. Finally, heat transfer processes in the vadose zone are an especially important component of runoff generation mechanisms in regions where seasonal soil freezing/thawing occurs.

3.3.1 Where We Are

For many years the NWSRFS has had a conceptual modification to the Sacramento Soil Moisture Accounting Model (SAC-SMA) that simulates a frost index and makes SAC-SMA runoff adjustment depending on this index. As a conceptual model, this approach requires the calibration of seven parameters.

Capitalizing on successful collaboration between OHD and NCEP/EMC, the Hydrology Laboratory (HL) developed a physically based parameterization that addresses two problems: i) modification of a storage-type model such as SAC-SMA to be compatible with a theoretical heat transfer model, and ii) parameterization of frozen ground effects on runoff. Hereafter, we refer to this model as the SAC Heat Transfer model (SAC-HT; see *Figure 3-1*), details of which are available in Koren (2006) and Koren *et al.* (2006).

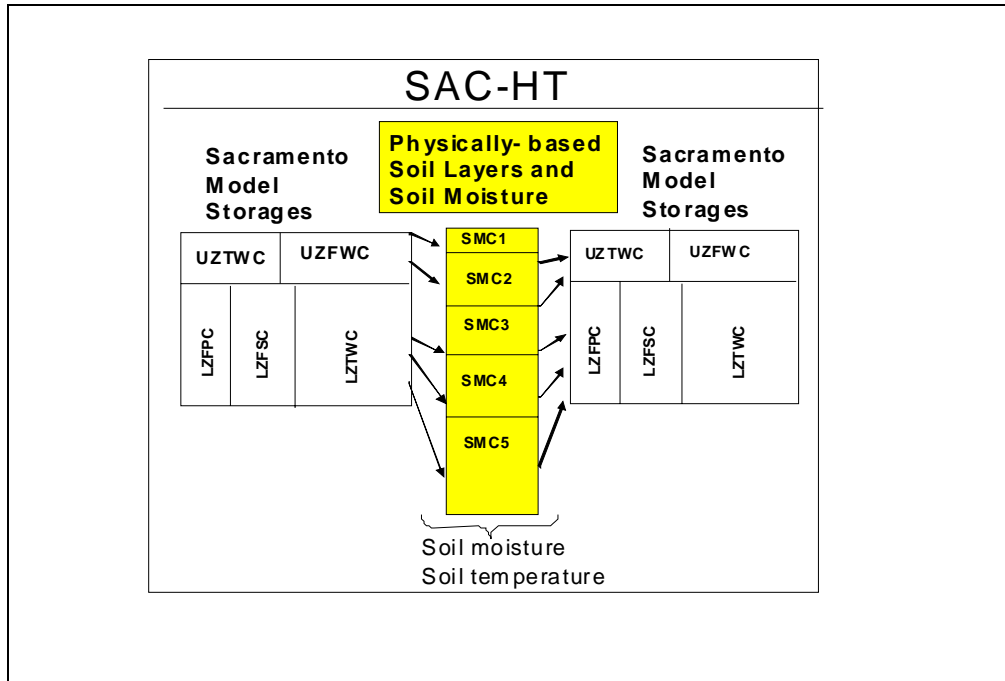


Figure 3-1 Simplified Structure of the SAC-HT

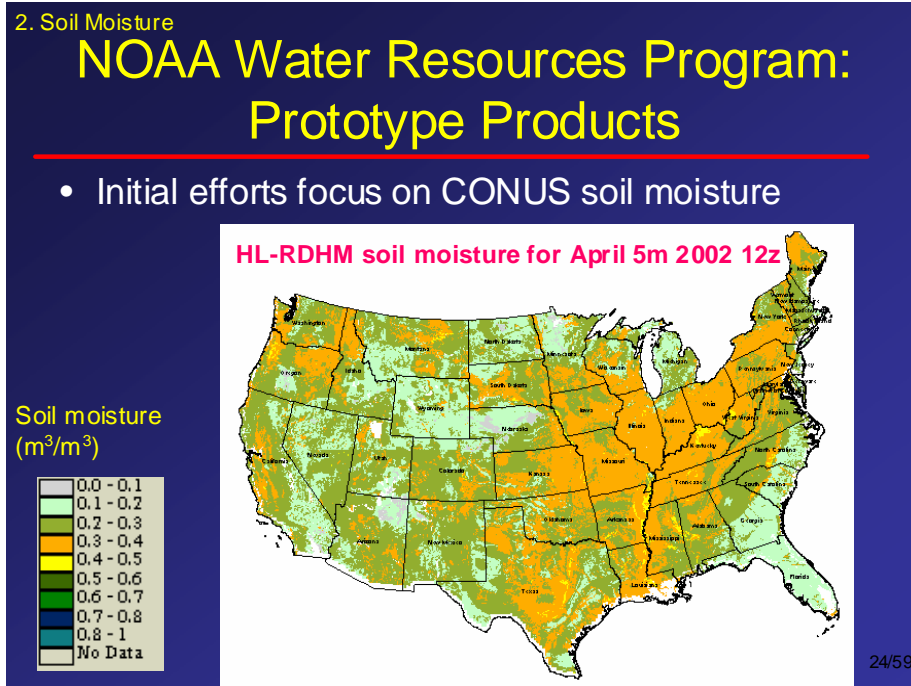


Figure 3-2 Example 4km, Hourly Computed Volumetric Soil Moisture Grid

3.3.2 What Our Partners Are Doing

NCEP/EMC produces realtime operational 3-hourly 4-layer simulations of soil moisture and soil temperature fields (including frozen soil moisture) over both 1) the North America continent at 12-km resolution and 2) the global domain at 0.5-degree resolution from the Noah LSM of the data assimilation components of its mesoscale (North American Mesoscale) NAM model and its global GFS model, respectively. Over the CONUS domain, the NAM soil moisture fields are driven by the Stage III RFC hourly precipitation analyses. Additionally, the Noah LSM component of the NCEP N. American Regional Reanalysis (NARR) produces N. American soil moisture fields in both realtime and in reanalysis mode back to 1979 at 32-km resolution. The 3rd NCEP Global Reanalysis, now in production mode and scheduled for completion in early 2009 (including an ongoing realtime extension), produces global 4-layer soil moisture and temperature fields at 0.5-deg resolution from its Noah LSM component, spanning from 1979 to present realtime.

Similarly, ESRL produces multi-layer soil moisture and temperature fields from the LSM component of its 13-km Rapid Update Cycle (RUC) mesoscale analysis system. The RUC executes not only in realtime developmental mode at ESRL, but also in realtime operational mode at NCEP in partnership with EMC. Recently, the National Operational Hydrologic Remote Sensing Center (NOHRSC) began displaying simulated soil moisture values generated by the realtime execution of ESRL RUC at 13-km resolution at hourly time steps.

The NLDAS suite (Mitchell et al. 2004) of EMC continues to execute in realtime, since October 1996, and produces realtime hourly simulations of multi-layer soil moisture and soil temperature with four different land models (Noah, SAC, VIC, and Mosaic) over the CONUS domain at 1/8-th degree resolution. These soil moisture states are also depicted as percentiles and anomalies with respect to the climatology of the 10-year NLDAS executions since 1996. The historical simulations of the 4-model NLDAS are presently being extended back almost 30-years to 1979, to provide a more robust depiction of soil moisture percentiles and anomalies from climatology. These NLDAS soil moisture products are being provided by CPC to the national drought monitor at <http://www.drought.gov>.

NCEP currently generates soil moisture estimates as part of the Noah model executions that support the production of numerical weather predictions. NCEP is participating in the DMIP 2 soil moisture experiments with the Noah model and the results will be forthcoming. The NOAA Drought Monitor is a joint project among several federal agencies and produced by the NCEP Climate Prediction Center (CPC). In this application, soil moisture is estimated by a one-layer hydrological model (Huang *et al.*, 1996, van den Dool *et al.*, 2003). The model

takes observed precipitation and temperature and calculates soil moisture, evaporation, and runoff. The potential evaporation is estimated from observed temperature.

NASA has a heritage of observing and modeling soil moisture. NASA scientists have considerable experience with satellite remote sensing of soil moisture and other hydrologic variables (e.g., de Gonçalves *et al.*, 2006). In addition, NASA was a major partner collaborating in land surface modeling over large areas with a view towards initializing numerical weather prediction models (NLDAS: Mitchell, K.E., *et al.*, 2004). A prominent focus of NASA continues to be the assimilation of hydrologic observations, including soil moisture, into land surface models.

Similarly, the Agricultural Research Service (ARS) has a broad legacy with observing and modeling soil moisture. For example, the ARS laboratory in Beltsville, Maryland is conducting an experiment entitled Soil Moisture Retrieval and Mapping Using Two-Dimensional Synthetic Aperture Radiometry (2d-Star). ARS scientists are also heavily involved in the evaluation of in-situ soil moisture sensors. A recent study compared the neutron soil moisture probe with several commercial soil moisture sensing systems, including four based on the electromagnetic properties of soil as influenced by its water content. (http://www.ars.usda.gov/research/publications/Publications.htm?seq_no_115=209943). Moreover, ARS is heavily involved with the use of satellite soil moisture observations with land surface models.

The Hydrometeorology testbed (HMT, <http://hmt.noaa.gov/>), lead by NOAA's Earth System Research Laboratory (ESRL), is evaluating improvements on QPE, QPF, snow, hydrologic applications and decision support tools for extreme precipitation in mountainous areas. OHD is actively participating on the HMT to quantify the benefits to streamflow forecasting obtained from that research.

3.3.3 Where We Want to Be

Reliable estimates of soil moisture and temperature are envisioned as fundamental products of the NOAA Water Resources Program, as for example shown in **Figure 3-2**. These will be produced at 'fine' scales that are useful for various applications, such as drought monitoring and forecasting, planting scheduling for optimal agricultural seed germination, mitigating plant disease transmission through the soil, crop management, trafficability planning for civilian and military purposes, construction planning, and others. Ultimately, there could be a convergence between hydrologic forecasting models and land surface models used for weather prediction to generate such products. Indeed, the IWSP foresees this eventuality, in which the land-surface component of the numerical weather prediction models is used for hydrologic forecasting as well, although in an uncoupled fashion.

Correspondingly, observed soil moisture and soil temperature values should be assimilated into gridded distributed models to update internal model states. Initial work in this area is already underway in OHD (see *Chapter 7*). In addition, observed values of these variables should be more routinely used for hydrologic model calibration and model development and testing.

Another opportunity that has not been explored in the NWS is to use 2-m air temperature as an input to a data assimilation algorithm to estimate soil moisture. During summer, air temperatures are strongly affected by soil moisture. This relationship is used by Meteo France in this way. However, notable difficulties have been encountered with this approach since errors in the simulated 2-m air temperature in the coupled background assimilating model are often more due to errors in solar insolation forcing, horizontal air advection, or the physics of boundary layer mixing than to errors in soil moisture.

3.3.4 Challenges to Getting There

- More in situ observations such as the Climate Reference Network (CRN) and the Soil Climate Analysis Network (SCAN) sites are needed to augment state and regional mesonets.
- There are numerous science issues surrounding point-to-grid scale transformations such as point scale soil moisture observations compared to remotely sensed surface soil moisture and computed, gridded soil moisture.
- There are numerous scale and process modeling difficulties between land surface models for numerical weather prediction and hydrologic models. The research challenge is to demonstrate why models with a more physical basis do not perform as is expected.
- Current-technology airborne and satellite sensors are unable to provide soil moisture observations at more than a few cm depth.
- The time horizons to plan and launch space-borne sensors and experiments such as Short and Medium-term Priority Environmental Action Programme (SMAP) are long.

3.3.5 A Road Map for Getting There

Waypoints along the roadmap include:

- Soil moisture and temperature normals need to be computed to serve as a basis for computing anomalies. Doing this in a climate-changing environment will be a challenge.
- Expansion of existing observational networks such as SCAN are needed.
- OHD will continue to monitor land-surface model developments within the numerical weather community, with testing of such models in experiments

such as DMIP 2. In particular, as part of DMIP-2, OHD will continue working intensively with the HMT to evaluate model results and compare them with field observations.

- OHD will participate with the sensor-development community in planning future space-borne efforts.

3.4 Groundwater Storage and Base Flow

Historically, the focus of NWS streamflow forecasts has been primarily on high flows for flood forecasting nation-wide and for water supply in the west. Therefore, detailed knowledge/modeling of groundwater has not been necessary, given its negligible contribution to the hydrograph during high flows. However, with the new emphasis on low flow forecasting for drought and water resources services, the groundwater contribution to streamflow becomes substantial. It follows then that improving NWS ability to forecast low flows depends on the quality of groundwater models we use.

3.4.1 Where We Are

The hydrologic forecast operations by the National Weather Service do not currently include explicit groundwater hydrology models. Only the base flow components of the two hydrologic models (Sacramento and Continuous API) provide some degree of information about groundwater conditions. .

3.4.2 What Our Partners Are Doing

The USGS is currently pursuing research on joint surface-ground water simulation models. Recently, Niswonger *et al.* (2006) developed an unsaturated zone flow package for MODFLOW. In this new package, aquifer recharge is handled by modeling flow through the unsaturated zone, as opposed to applying the recharge directly to the aquifer. It also partitions evaporation from the saturated and unsaturated zones, and accounts for land surface runoff to streams and lakes. Additional work along the same lines is being pursued by the USGS, coupling the Precipitation Runoff Modeling System (PRMS), a distributed rainfall-runoff model, with the U. S. Geological Survey Modular Ground-Water Model (MODFLOW) under the name, Groundwater/Surface-Water Flow (GSFLOW) Model. MODFLOW has also been coupled with Hydrological Simulation Program - FORTRAN (HSPF) at the University of South Florida (Said *et al.*, 2005). At the University of Texas/Austin, a groundwater component and Topmodel approach to runoff has been added to the Noah LSM, in partnership with NCEP/EMC.

3.4.3 Where We Want to Be

NWS should be in a position to provide reliable forecasts of groundwater contribution to stream flow, by using two- or three-dimensional groundwater models in conjunction with other surface and hydraulic routing models. Furthermore, NWS

should make use of the wealth of information on groundwater levels provided by observations wells (as of November 2007, the USGS obtains real-time data (5 – 60 minutes) at 1,035 sites, and daily data at 4,953 sites.) Use of these data would provide information for model calibration, verification, and data assimilation.

3.4.4 Challenges to Getting There

Obtaining estimates of groundwater pumping in rural areas will be a major challenge. Those records are not available in real-time, although the slow-responding times of groundwater systems make the availability of real-time information less critical.

Verification of the performance of the groundwater models requires observations of the water table elevation. Although new remote-sensing techniques, based on the NASA/ GeoForschungsZentrum (GFZ) Gravity Recovery and Climate Experiment (GRACE) satellite system are now being developed, those observations have a very high vertical resolution, (of the order of cm), but a horizontal resolution of the order of about 80 km, much too coarse for practical NWS forecasting applications. Furthermore, GRACE observations measure total change in water (including snow, soil moisture and groundwater). It is, therefore, a challenge to estimate how changes are distributed among the components.

The modeling of karstic and fractured aquifers is particularly challenging. Obtaining geologic information and calibration of groundwater models for such aquifer is the main difficulty.

Finally, Operational use of coupled groundwater and surface-water models will require additional training by NWS forecasters.

3.4.5 A Road Map for Getting There

In collaboration with the USGS and other agencies, OHD should research the issues of coupling surface and groundwater forecasting models, specifically how to consider flow in the unsaturated zone,; how to couple the one-dimensional soil moisture accounting models with two- and three-dimensional groundwater models; and how to deal with widely different response times.

Once the issues have been identified, OHD will produce prototypes to be field-tested in those RFCs that have watersheds in which low flows and water extraction for irrigation and water supply are important.

3.5 Snow Accumulation, Sublimation and Melt

In much of the U. S., particular the West, snowmelt is a major—and often the dominant—contributor to stream flow. Accurately modeling the water balance of

the snowpack at high spatial and temporal resolutions are critical to both flood and water supply forecasts.

3.5.1 Where We Are

As stated earlier, SNOW-17 is the predominant snow accumulation and melt model used for RFC forecasting. SNOW-17 is a well known and widely used conceptual model using temperature and precipitation as the sole forcings. Temperature is the driving forcing for the snowpack dynamics with the exception that during rain on snow events, assumptions are made so that energy budget computations are used. SNOW-17 is implemented over lumped basins or in elevation zones in the mountains. Some RFCs use NOHRSC's snow water equivalent (SWE) values to update their SNOW-17 states.

A variant of SNOW-17 exists (SNOW-43) in NWSRFS that uses Kalman-Filtering to account for the relative uncertainties of observed and simulated water-equivalent values in a procedure that optimally updates the model simulated states using areal estimates of snow water equivalent based on observations.

SNOW-17 has demonstrated notable performance compared to the Variable Infiltration Capacity (VIC), Mosaic, and Noah models in NLDAS (Sheffield *et al.*, 2003; Pan *et al.*, 2003), to the Noah model (Lei *et al.*, 2007) and in the SnowMIP 1,2 experiments (Koren *et al.*, 2007). These comparisons add to the original World Meteorological Organization (WMO) snow model intercomparison project (WMO Operational Hydrology Report No. 23, WMO - No. 646, 1986).

OHD has recently developed a physically based model of the effects of rain over frozen ground. This model is closely tied to and enhances the Sacramento model by mapping the conceptual soil moisture reservoirs to physical layers of the soil, and it is known as the Heat Transfer (HT) model. Because of this close connection with the SAC-SMA model, the HT model is covered in *Section 3.3*.

3.5.2 What Our Partners Are Doing

NOHRSC executes a full-energy-budget snow accumulation and melt model over CONUS at the hourly time scale and a 1km spatial resolution within its Snow Data Assimilation System (SNODAS). SNODAS uses a variant of the Snow Thermal Model (SNTHERM) model, developed by the U. S. Army Corps of Engineers (Jordan, 1991) from the work of (Anderson, 1976). One goal of the SNODAS modeling effort is to generate the best possible SWE estimates using all available data. These SWE estimates are then sent to RFCs for use in updating the SNOW-17 states. Forcings from the Rapid Update Cycle (RUC) model are downscaled to the 1-km grid scale.

NCEP executes the Noah energy budget snow model (Koren et al., 1999) that is a part of the Noah Land Surface Model (LSM; Mitchell *et al.*, 2002). The Noah Land Surface Model (LSM) is a component of the operational NAM numerical weather prediction model. The Noah snow model accounts for the effects of frozen ground, patchy snow cover, and temporal/spatial variability in snow properties such as density and thermal conductivity. It does not include conceptual-type parameters and no (or very little) calibration is needed. The output variables from the model include: snow depth, snow water equivalent, snow melt rate, liquid water content, bottom runoff, etc. As a community model, Noah has ability to utilize new science and data sources. The NLDAS study (Mitchell et al. 2004) and companion papers listed therein documented an early bias in the timing of early springtime snowpack depletion in the Noah LSM. This early bias has since been substantially reduced by modifying the Noah treatment of sublimation when snow cover is patchy, changing the Noah treatment of albedo over snow, and modifying the treatment of aerodynamic conductance in stable planetary boundary layer regimes. Additionally, at the University of Texas/Austin, a multi-layer snowpack treatment has been added to the Noah LSM to replace its traditional single bulk-layer treatment of snowpack physics.

3.5.3 Where We Want to Be

The Natural Disaster Survey Report for the Northeast floods of January 1996 provides several recommendations for the NWS to include the use of energy budget snow modeling for RFC and water resources forecasting (Office of Hydrology, 1998). The January 1996 floods were characterized by a large snowmelt contribution, resulting from above average snow cover and high melt rates produced by latent and sensible heat exchange. The conditions in the January 1996 event were beyond those for which the SNOW-17 model was calibrated. It is clear from this and other reports that the NWS needs to include the use of energy budget snow modeling for RFC and water resources forecasting. We envision the complementary use of the SNOW-17 model (and/or SNOW-43) at the RFCs with a more complete utilization of the capabilities of the SNODAS model run at NOHRSC. The SNOW-17 model may continue to be used as it has proven to perform well in the conditions for which it is calibrated until its performance is exceeded by that of the new energy-budget snow models. For highly unusual events, information from the SNODAS model could be used explicitly in runoff calculations or to guide run-time modifications to SNOW-17 to adjust for non-standard conditions. NOHRSC products such as SWE will continue to be used in everyday SNOW-17 updating. Anderson (2003, 2006a) provides many suggestions on these issues.

We envision the following two broad goals:

1. Development of a new suite of snow forecasting models, based on energy balance and physical principles. Until it can be demonstrated that those new

models outperform SNOW-17, this model may continue to be operated at the RFCs in parallel with the SNODAS (run at NOHRSC). SNOW-17 could be applied as a lumped model, to elevation zones, or in a gridded format as appropriate;

2. Exploitation of SNODAS to generate a broader array of data products for RFC use, most notably for storms with non-standard meteorologic conditions. This could range from use of SNODAS products directly or as guidance to RFC forecasters on making run-time modifications for SNOW-17 (Anderson, 2006b).

3.5.4 Challenges to Getting There

Mountainous area hydrology has been identified by some as containing some of the largest knowledge gaps. For distributed versions of SNOW-17, parameterization and calibration strategies need to be developed and tested. Initial work has been completed for deriving gridded estimates of the SNOW-17 melt factors. Franz (2006) and Lei *et al.*, (2007) outline some of the challenges regarding the full use of energy budget snow models for operational forecasting. One of the major challenges is the sensitivity of energy budget models to errors in forcing data. NOHRSC has established a detailed list of improvements needed for the SNODAS. Among these is the need to examine alternatives to the current Quantitative Precipitation Estimate (QPE) from the Rapid Update Cycle system. Data scarcity remains a large problem in the intermountain west and is the dominant problem in Alaska. A corollary question is “how much data are needed to improve RFC forecasts in the mountainous areas?” The DMIP-2 science plan (Smith *et al.*, 2006) outlines several of the dominant questions for modeling in mountainous areas.

3.5.5 A Road Map for Getting There

We identify the following waypoints on the roadmap for achieving the two broad goals identified in *Section 3.5.3*:

In the short term (1-2 years):

- Development of a detailed snow science plan by the Snow Science Steering Team, following the guidelines developed during the 2007 cold regions conference;
- Analysis of the SnowMIP 1, 2 results and development of next steps;
- Continuation the development of gridded SNOW-17 parameters;
- Development calibration strategies for gridded SNOW-17 as requested by RFCs;
- Evaluation of gridded SNOW-17 in mountainous terrain;

- Evaluation of conceptual and other snow models in mountainous areas via DMIP-2;
- Evaluation of SNODAS products for RFC use (Anderson, 2006b);
- Analysis of advanced data collection strategies (e.g., QPE from gap-filling radars, vertically pointing radars for rain/snow level detection) for mountainous areas via the HMT experiment in conjunction with DMIP-2;
- Complete the current joint NASA-OHD project on assimilation of snow cover derived from MODIS;
- Exploration the SNOW-43 National Weather Service River Forecast System (NWSRFS) operation to understand if the perceived limitations can be overcome; and
- Coordination via the Snow Science Steering Team (SSST) to address needed SNODAS improvements (Carroll, 2005).

In the long term (2-10 years):

- Development of new or adapted energy-budget models

3.6 Evapotranspiration

In all climates, evapotranspiration (ET) is often the dominant flux leaving a watershed during interstorm periods. In dry climates, it is almost always the dominant flux on monthly to interannual time scales. Therefore, accurate modeling of ET is critical to accurate modeling of the soil water balance, which in turn is critical to accurate representation of conditions antecedent to flooding events. Accurate observations and forecasts of ET and soil moisture are also important as the NWS moves into soil moisture, low flow, and water quality forecasting.

3.6.1 Where We Are

ET in the SAC-SMA is typically estimated as a storage-controlled fraction of potential evapotranspiration (PET), with PET taken as a function of climatological observations of pan evaporation. The associated PE adjustment factors, used to compensate for the effects of vegetation on ET, are usually treated as calibration parameters in the SAC-SMA. For operational purposes, there is an unfortunate loss of climate observations, leading to a climatology that may be too old, especially for a relatively rapidly changing climate.

Digital fields of monthly PE adjustment factors are available and these are based in part on mean monthly vegetation fraction fields used in the Noah model. Also, monthly digital PET fields based on annual and May-October free water surface evaporation fields from the NOAA Evaporation Atlas are available.

NWSRFS also includes a preprocessor to estimate PET. This program is called SYNTRAN and uses cloud cover as an input variable.

In collaboration with NASA's Marshall Space Flight Center, OHD is carrying out a research project that uses remotely sensed observations of cloud cover and radiation to estimate PET. (Restrepo *et al.*, 2007).

3.6.2 What Our Partners Are Doing

In addition to remote-sensing estimates of ET, the NCEP operational regional and global NWP models provide operational forecasts of ET and PET, presently at 3-hourly temporal resolution (or better), at 12-km resolution out to 84 hours in the regional NWP model, and at 60-km resolution out to 15 days in the global model. The regional and global model forecasts of ET and PET are also available as ensemble forecasts at lower spatial resolution. Additionally, the NCEP seasonal climate forecast system provides ET and PET forecasts out to 45 days at 12-hourly intervals (to be extended to a 365-day forecast range at 6-hourly intervals in the next generation seasonal forecast system). The ET forecasts are based on the Noah land model and the PET forecasts utilize a particular form of the Penman equation that yields excellent diurnal variability. Earlier OHD studies have shown that the NCEP model PET forecast values have a high bias on the order of 20-25% in the warm season compared to NOAA pan evaporation climatologies but OHD applications of NCEP PET forecasts can address this by applying a scaling coefficient.

NASA is currently sponsoring a number of collaborative research projects that are addressing the estimation of actual or potential evapotranspiration through the use of remote sensing observations. For example, Aggett (2007) is studying the use of remotely sensed estimates of actual evapotranspiration amounts for improving water management in the west. Hendricks *et al.* (2007) are looking at the effects of land use and crop coefficients in the estimation of evapotranspiration from space; and Houser (2007) is developing the Water Cycle Solutions Network (WaterNet), whose vision is to "To improve our collective ability to routinely interact with and harness the results of scientific research so as to address water assessment and management challenges".

3.6.3 Where We Want to Be

LSM such as Noah and similarly physically based coupled energy and water balance models should be the primary basis of ET (and soil moisture calculations; see *Section 3.3*) in the next generation of models envisioned under this Strategic Science Plan. OHD needs to strengthen its ties with NASA research to take advantage of the wealth of observations and science being developed by its Earth Sciences program.

3.6.4 Challenges to getting there

As all LSM intercomparison studies have shown, such models typically produce widely varying estimates of energy and water fluxes. Those studies have attributed differences in model results to a myriad of causes. In general, divergent results are traceable to the many ways all the different sources of the fluxes are represented. Specifically, total ET from a watershed over the course of a year is typically composed of: evaporation from the soil surface—both from soil moisture and depression storage; vegetal transpiration; evaporation from canopy interception of rain; snow sublimation from canopy interception and surface storage; evaporation from stream channels and open bodies of water; and evaporation from storages in the built environment. Typically, LSM represent only a few of these sources of ET. In addition, they are typically one-dimensional representations applied at large-scales, or account very crudely for subgrid heterogeneities.

Remote sensing estimates of ET are dependent on the remote sensing of the forcings discussed in *Chapter 4*, especially shortwave/longwave radiation and skin temperature, and thus suffer from all the challenges discussed in that chapter.

3.6.5 A Road Map for Getting There

In the immediate term, we will continue to work with NCEP in the development and implementation of Noah. We should also increase our collaboration with partners that have models that account for a greater number of sources of ET, as well as more realistically represent subgrid heterogeneities. An example of such a model is NASA's Topographically based Land Atmosphere Transfer Scheme model (TOPLATS). Continued participation in LSM model intercomparison projects such as NLDAS is critical for finding the best performing and most suitable models.

With regards to remote sensing of ET, this roadmap depends largely on those issues described in the Forcings Chapter. We also need to increase our collaboration with partners that are working on the problem, especially those at NASA, NESDIS, and NOAA Cooperative Remote Sensing Science and Technology Center (CREST).

3.7 Hydraulics (Future)

3.8 Water Quality (Future)

3.9 References

- Aggett, G. (2007). "Enhancing Water Management DSS with High Spatio-temporal Resolution Mapping of Actual ET". *NASA Water Management Program PI Meeting*.
- Anderson, E.A. (1976). "A Point Energy and Mass Balance Model of a Snow Cover". NOAA Technical Report, *NWS HYDRO-17 NWS 19*.
- Anderson, E.A. (2003). "Reflections on Future Directions in Snow Modeling for the Hydrology Program of the National Weather Service". *Internal OHD document*.
- Anderson, E.A. (2006a). "Further Comments on Future Directions in Snow Modeling for the NWS Hydrology Program." *Internal OHD document*.
- Anderson, E.A. (2006b). "SNODAS-based MODS for SNOW-17 – FINAL DRAFT Project Plan." *Internal OHD document*.
- Carroll, T. (2005). "Areas of Potential Collaboration between the NOHRSC and Hydrology Laboratory." *Internal NOHRSC document sent to OHD Jan 10, 2005*.
- de Goncalves, L. G. G., W. J. Shuttleworth, S. C. Chou, Y. Xue, P. R. Houser, D. L. Toll, J. Marengo, and M. Rodell: 2006. Impact of different initial soil moisture fields on Eta model weather forecasts for South America, *J. Geophys. Res.*, 111, D17102, doi:10.1029/2005JD006309
- Dietrich, W., R. Reiss, M. Hsu, D. Montgomery (1995). A process-based model for colluvial soil depth and shallow landsliding using digital elevation data. *Hydrological Processes*, 22(3-4), 383-400
- Elsenbeer, H. (2001). Preface to a special issue on pedotransfer functions in hydrology. *J. of Hydrol.*, 251(3-4), 121-122.
- Franz, K.J. (2006). "Characterization of the comparative skill of conceptual and physically-based snow models for streamflow prediction." PhD dissertation, U. California at Irvine, CA.
- Georgakakos, K.P., and Carpenter, T.M. (2004). "Potential value of operationally available and spatially distributed ensemble soil water estimates for agriculture." *J. of Hydrol.*, in review.

- Hendricks J, Stodt, R., Allen, R. Byrd, A., Ogden F. (2007). "Integrating Best Estimates of Regional Evapotranspiration into Hydrologic DSS". *NASA Water Management Program PI Meeting*.
- Houser, P., Belvedere, D., Pozzi, W., Imam, B., Schiffer, R., Schosser, C.A., Gupta, H., Welty, C., Vorosmary, C., Matthews, D., Lawford, R. (2007). "WaterNet: The NASA Water Cycle Solutions Network". *NASA Water Management Program PI Meeting*.
- Huang, J., van den Dool, H., and Georgakakos, K.P. (1996). "Analysis of Model-calculated soil moisture over the United States (1931-1993) and applications to long-rang temperature forecasts." *Journal of Climate*, Vol. 9, 14350-14362.
- IWSP, The Integrated Water Science Plan, Report of the NWS IWSP team, NWS, November 24, 2004.
- Johnson, D., Smith, M., Koren, V., and Finnerty, B. (1999). "Comparing mean areal precipitation estimates from NEXRAD and rain gauge networks." *Journal of Hydrologic Engineering*, Vol. 4, No. 2, 117-124.
- Jordan, R. (1991). "A one-dimensional temperature model for a snow cover: Technical documentation for SNTHERM. 89." *U. S. Army Corps of Engineers, Special Rep. 91-16*, 64 pp. [Available from *Cold Regions Research and Engineering Laboratory*, Hanover, NH 03755-1290.].
- Koren V., J. Schaake, K. Mitchell, Q.-Y. Duan, F. Chen, J. M. Baker. (1999). "A Parameterization of Snowpack and Frozen Ground Intended for NCEP Weather and Climate Models." *Journal of Geophysical Research*, Vol. 104, NO.D16, 19569- 19,585.
- Koren, V., S. Reed, Z. Zhang, D. Seo, F. Moreda, Kuzmin, V. (2003). "Use of spatially variable data in river flood prediction." *AGU-EGS-EUG Assembly*, April 9-14, Nice, France.
- Koren, V., 2006. Parameterization of frozen ground effects: sensitivity to soil properties. In: *Predictions in Ungauged Basins: Promise and Progress*, IAHS Publ. 303, 125-133.
- Koren, V., Moreda, F., Reed, S., Smith, M., Zhang, Z., 2006. Evaluation of a grid-based distributed hydrological model over a large area. In: *Predictions in Ungauged Basins: Promise and Progress*, IAHS Publ. 303, 47-56.

- Koren, V., Fu, F., and Zhang, J. (2007). "Results from a Simple Parameterization of the Snow-Vegetation Interaction in Energy-Budget and Temperature Index Models." *IUGG Meeting*, Perugia, Italy.
- Kulic, V. Y. (1978). "Water Infiltration into Soil." (In Russian), *Gidrometeoizdat*, Moscow.
- Lei, F., Koren, V., Smith, M., and Moreda, F. (2007). "A Sensitivity Study of an Energy Budget Snow Accumulation and Ablation Model." *87th Meeting of the AMS*. Paper J6.4A.
- Mitchell, K.E. and co-authors, 2004. The Multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *JGR*, Vol., 109, D07S90, doi:10.1029/2003JD003823
- National Research Council (NRC). (2000). "Grand Challenges in Environmental Sciences." Committee on Grand Challenges in Environmental Sciences, Oversight Commission for the Committee on Grand Challenges in Environmental Sciences, *National Academy Press*, Washington, D.C., 96pp.
- Niswonger, R.G., Prudic, D.E., and Regan, R.S. (2006). "Documentation of the Unsaturated-Zone Flow (UZFI) Package for modeling unsaturated flow between the land surface and the water table with MODFLOW-2005: U. S. Geological Survey Techniques and Methods." 6-A19, 62 p.
- NOAA. (2004). "New Priorities for the 21st Century-NOAA's Strategic Plan." Available from <http://www.spo.noaa.gov/pdfs/NOAA%20Strategic%20Plan.pdf>.
- NWS. (2004a). "National Weather Service Strategic Plan for 2005-2010." Available from http://www.weather.gov/sp/NWS_draft_strategic_plan_10-15-04.pdf.
- NWS. (2004b). "The NWS Integrated Water Science Plan (IWSP)." Report of the IWSP Team.
- Pan, M., Sheffield, J. and co-authors, 2003. Snow Process Modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of Model Simulated Snow Water Equivalent. *JGR*, Vol. 108, No. D22.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., and DMIP Participants. (2004). "Overall distributed model Intercomparison project results." *J. of Hydrol.*, Vol. 298, Nos. 1-4, 27-60.

- Restrepo, P., Limaye, A., Peters-Lidard, C., Laymon, C., Houser, P., Toll, D., Rodell, M., Cosgrove, B., Garcia, M., Kumar, S., Dong, J., Arsenault, K. de Goncalves, L. (2007). "NASA Earth Science Radiation and Snow Data Products and LIS Modeling for NOAA Flood Assessment/Forecast". *NASA Water Management Program PI Meeting*.
- Ryu, D., Crow, W.T., Zhan, X. 2007. Assimilation of coarse-scale satellite soil moisture observations into a fine-scale hydrological model [abstract]. Abs. 9. BARC Poster Day.
- Said, A., D. Stevens and G. Sehlke (2005). Estimating water budget in a regional aquifer using HSPF-MODFLOW integrated model. *J. Am. Water Res Assoc.*, 41, 1-55.
- Seo, D.-J., Koren, V., and Cajina, L. (2003). "Real Time Assimilation of Radar-based Precipitation and Streamflow Observations in a Distributed Hydrologic Model." *IAHS Pub. 282*, 138-142.
- Sheffield, J, Pan, M., and co-authors. (2003). "Snow Process Modeling in the North American Land Data Assimilation System (NLDAS): 1." Evaluation of Model Simulated Snow Cover Extent. *JGR*, Vol. 108, D22.
- Smith, M.B., Seo, D. -J., Koren, V. I., Reed, S., Zhang, Z., Duan, Q.-Y., Moreda, F., and Cong, S. (2004). "The distributed model intercomparison project (DMIP): motivation and experiment design." *J. of Hydrol.*, Vol. 298, Nos. 1-4, 4-26.
- Smith, M.B., Koren, V., Finnerty, B., Johnson, D. (1999). "Distributed Modeling: Phase 1 Results", *NOAA Technical Report NWS 44*, February, 250pp.
- Stellman, K.M., Fuelberg, H.E., Garza, R., and Mullusky, M. (200). "An examination of radar and rain gauge-derived mean areal precipitation over Georgia watersheds." *Journal of Hydrometeorology*, Vol. 16, 133-144.
- Tsintikidis, D., Georgakakos, K.P., Sperflage, J.A., Smith, D.E., and T.M. Carpenter. (2002). "Precipitation Uncertainty and Raingage Network Design within the Folsom Lake Watershed." *ASCE Journal of Hydrologic Engineering*, 7(2), 175-184.
- van den Dool, H., Huang, J., and Fan, J. (2003). "Performance and analysis of the constructed analogue method applied to U. S. soil moisture over 1981–2001." *JGR* Vol. 108, NO. D16, 8617, doi:10.1029/2002JD003114, 2003.
- Wang, D., Smith, M.B., Zhang, Z., Reed, S., and Koren, V.I. (2000). "Statistical comparison of mean areal precipitation estimates from WSR-88D, opera-

tional, and historical gage networks.” *15th Annual Conference on Hydrology, 80th Meeting of the AMS*, Long Beach, CA., J2.17.

Whitaker, A., Y. Alila, J. Becker and D. Toews (2003). Application of the Distributed Hydrology Soil Vegetation Model to Redfish Creek, British Columbia: model evaluation using internal catchment data. *Hydrological Processes*, 17, 199-224.

World Meteorological Organization. (1992). “Simulated real-time intercomparison of hydrological models.” *Operational Hydrology Report No. 38*, WMO-No. 779, Secretariat of the World Meteorological Organization, Geneva, Switzerland.

WMO Operational Hydrology Report No. 23, WMO - No. 646, 1986.

Young, C.B., Bradley, A.A., Krajewski, W.F., Kruger, A. (2000). “Evaluating NEXRAD multi-sensor precipitation estimates for operational hydrologic forecasting.” *Journal of Hydrometeorology*, Vol. 1, 241-254.

4. Forcings

The IWSP describes the role of forcings in hydrologic forecasting as follows:

“The choice of land surface model applied in a given hydrological prediction system determines the scope of the required land surface forcing fields. These surface forcing fields will include all or some of the following fundamental forcings: precipitation, incoming surface shortwave radiation, incoming surface longwave radiation, surface pressure, and near-surface air temperature, humidity and wind speed....

Each forcing field requires both a monitoring thrust and a prediction thrust. The prediction thrust must span the forecast ranges of nowcasting (hours or less), short-range (days), medium-range (weeks), and seasonal (months) and include ensemble/probabilistic forecast approaches. To support high-resolution geospatial WRPS, accurate determination of high-resolution precipitation forcing is critical. In the monitoring thrust for precipitation, a major emphasis must be the expansion of our observing capabilities with gauge, radar, and satellite observations and the identification and correction of the systematic biases in observations, especially in mountainous terrain and cold season snowfall regimes. In the prediction thrust, we must develop cutting-edge methods to take full advantage of the emerging ensemble forecast approach in NWP and seasonal climate prediction.”

We should emphasize that the role of observations is to estimate initial conditions for use in forecasting. This is basically true for both flash flood forecasting as well as major river forecasting. The initial conditions include soil water and snow water equivalent state variable as well as water flowing in streams throughout the drainage area.

4.1 Observed Precipitation

Precipitation is the primary driver for streamflow, affecting discharge through surface runoff, subsurface flow, groundwater recharge, and snowmelt (see **Chapter 3**). The timescale of the response of streamflow to observed precipitation varies greatly with the spatial scale of the storm and the watershed and the season of the year, ranging from minutes in small, flashy watersheds to weeks on the mainstems of large basins to months in the case of snowmelt runoff and groundwater contributions to base flow.

4.1.1 Where We Are

For prediction of larger streams that feature a lag greater than six hours, NWS hydrologic forecasters rely primarily on a combination of rain gauge and radar

estimates, supplemented by infrared satellite estimates where the other two sensors are lacking. The system to couple these precipitation estimates is known as the Multi-sensor Precipitation Estimator (MPE). MPE uses an optimal estimation algorithm to blend the estimates from rain gages and radar into a single gridded estimate. Two of the River Forecast Centers have developed their own algorithms. The Arkansas-Red River Forecast Center developed the P3 application, based on an algorithm originally developed by the Army Corps of Engineers. P3 takes advantage of the spatial information from radars and uses it to interpolate the precipitation field observed by rain gages. The Colorado Basin River Forecast Center developed Mountain Mapper, which incorporates a climatology produced by the Precipitation-elevation Regressions on Independent Slopes Model (PRISM; Daly *et al.*, 1994). Mountain Mapper is currently used at the Colorado Basin, California-Nevada, Northwest and Alaska-Pacific River Forecast Centers. In some areas where sufficient radar coverage is lacking, rain gauge reports are used exclusively, to create both gridded and basin-scale precipitation estimates. Advanced Weather Interactive Processing System (AWIPS) has interactive software facilities for quality control and, when necessary, modification of, input data and final precipitation output. For monitoring of flash floods, on basins with lag times less than six hours, radar and, to a limited extent, rain gauge reports are used, in combination with Flash Flood Guidance (FFG).

4.1.2 What Our Partners Are Doing

Key advances in weather radar are imminent, including dual-polarization and operational access of units other than Weather Surveillance Radar-1988 Doppler (WSR 88D), the current workhorse for NWS surface radar. Precipitation and target identification algorithm enhancements are necessary to effectively use the radars; work is ongoing at the National Severe Storms Lab (NSSL), The National Center for Atmospheric Research (NCAR), and the NWS Radar Operations Center (ROC). Development of dual-polarization-based precipitation estimation and quality control are ongoing at NSSL (Ryzhkov *et al.* 2005a-b) and NCAR; OHD is providing validation support.

Advances in satellite instrumentation and algorithm development for both operational and future platforms are carried out and supported by the National Aeronautic and Space Administration (NASA) and the National Environmental Satellite Data and Information Service (NESDIS), as well as by NSSL and many research institutions and universities. Planning has begun for a new satellite constellation, the Global Precipitation Mission, GPM.

New methods of integrating and quality-controlling conventional and Doppler radar reflectivity data and associated precipitation estimates are ongoing through the National Mosaic and Multi-Sensor QPE (NMQ) and the Next Generation QPE (Q2) development effort managed by NSSL with development support also coming from NESDIS (Zhang *et al.* 2005; Vasiloff *et al.* 2007). Experiments to

validate the relative value of several approaches to radar and multisensor interpolation are being shared among NSSL, NESDIS, (Center for Satellite Applications and Research (STAR) and National Climate Data Center (NCDC)) and OHD. Local experiments in integrating radar data other than WSR 88D are underway at several WFOs and RFCs.

Work in quantifying the uncertainty of radar estimates has been ongoing in several institutions, in particular the University of Iowa (Krajewski *et al.* 2005), and at McGill University (Lee *et al.*, 2006).

NCEP is using procedures developed by OHD to produce gridded hourly precipitation estimates on a national 1/8 degree grid for use as input to its Land Data Assimilation system and for assimilation into its regional forecast systems.

Opportunities for testing some of these evolving techniques and better utilizing existing ones, have been expanded through the operation of the HMT in California (HMT-West). A medium-size river basin has been heavily instrumented with *in-situ* equipment and radar units during three winter seasons, making observations available in an area with otherwise limited coverage. HMT is also providing high quality gridded QPE and QTE. These gridded fields are being used in the evaluation of hydrologic models as part of the DMIP2 mountain component. OHD will partner with HMT to evaluate new and evolving techniques under differing conditions, as testbeds are established in other representative regions in the near and distant future.

4.1.3 Where we want to be

We are working towards continuous, routine integration of all available sensor data and where needed, numerical prediction model estimates. At any one place the most statistical weight will be given to the most reliable sources available at that place. Data will be quality controlled automatically to the extent possible, with a final human intervention step whenever feasible. Data from newly-developed and validated sources (radar, satellite, surface sites) will be ingested and integrated as soon as logistically feasible. Simple characterizations of the statistical distributions of estimation error will be available to end users.

For radar input, several developments are crucial, including implementation of dual-polarization algorithms, introduction of reflectivity profile and range corrections, and automated selection of Z-R relationships (for single-polarization radars). For satellite input, we need implementation of algorithms for automatic real-time calibration of infrared temperature vs. rain-rate relationships based on collocated satellite and radar data. This approach shows some promise for improving satellite estimates in regions with appreciable radar coverage. OHD will assist with the development and validation of algorithms preparatory to deploy-

ment of the GPM. Anticipating its deployment, OHD will explore applications of the Tropical Radar Rainfall Measurement Mission (TRMM) observations.

Precipitation is a random field that is observed with sensors of different noise characteristics. Consequently, precipitation will be estimated using optimal estimation theory and an ensemble approach that will quantify the uncertainty of that estimate. This ensemble of precipitation analyses will work seamlessly with the ensembles of precipitation and temperature forecasts, including the results of numerical weather prediction models, including short-range precipitation and temperature ensemble forecasts, to improve accuracy and reduce uncertainty in precipitation and temperature analyses especially in mountainous areas. Because of the contribution of noisy precipitation observations to uncertainty in streamflow and water resources products forecasts, it is necessary to develop techniques compatible with OHD's ensemble products, which explicitly consider that source of uncertainty in model calibration and real-time operations

(see *Chapter 6*).

4.1.4 Challenges to Getting There

Reduction and quantification of uncertainty are the overarching challenges.

For mountainous areas, it is difficult to make more effective use of satellite and radar estimates in areas where radar coverage exists but is incomplete due to terrain beam blockage or beam overshooting. Current NWS algorithms don't permit logical extrapolation of radar rainfall estimates to nearby areas. The success of the Collaborative Adaptive Sensing of the Atmosphere (CASA) project in developing inexpensive radars will be a key to the widespread implementation of gap-filling radars in mountainous regions in a cost-effective manner, and to the improvement of precipitation estimation in those areas.

In general, effective quality control of rain gauge and radar data occupies a great amount of time, and automated algorithms to insure reliability and decrease human workload are needed.

4.1.5 A Road Map for Getting There

In mountainous areas, there should be more effective use of high-resolution numerical model simulations of precipitation and temperature. Passive microwave and to some extent space-based radar (TRMM and later GPM) estimates can be used at roughly multi-hour intervals to locally calibrate continuously available infrared satellite observations.

For flash flood monitoring, operations are evolving to integrate data from other radar systems (Terminal Doppler Weather Radar (TDWR), CASA radars). In

some areas, lightning observations, both cloud-to-ground and in-cloud, might be used to supplement radar estimates.

Automated surface observing networks are becoming more economical and widely available as time goes on. OHD staff has significant experience in precipitation estimation algorithm development, implementation, and maintenance. We envision a continued role in validating, developing, and implementing new multisensor algorithms proposed by our partners, particularly NMQ/Q2, gap-filling radars, and satellite-based algorithms.

Specific waypoints on the proposed road map are:

- Enhance transfer of research to operations from the NOAA labs to OHD and NWS field offices;
- Implementation of an High-resolution Precipitation Estimator package to serve flash flood monitoring operations;
- Development of a concept of operations for river forecast center use of centrally produced multisensor precipitation estimates such as Q2, particularly in geographic regions where current radar estimation techniques do not function well; this package should include gauge, radar, satellite, and NWP model output as precipitation estimators, and automated and manual quality control procedures;
- Assistance in the implementation and use of QPE from newly available radar systems, including TDWR, Air Route Surveillance Radar, and CASA units;
- Implementation of dual-polarization radar QPE algorithms and, after suitable operational validation, advice to field offices on the transition to exclusive use of these products;
- Evaluation of the HMTs observational data, starting with HMT-West, to refine techniques for extracting precipitation information from existing radar, satellite and rain gauge sensors; and evaluation and quantification of the benefits of additional sensors, such as wind profilers and gap-filling radars to streamflow forecasting;
- As applicable, operational implementation of Q2 processing for estimation on a national scale;
- Routine dissemination of uncertainty information with operational QPE products, such as parameters for error distributions;
- Investigation of updates to existing Open Radar Product Generator QPE capabilities, including range correction;
- Support for a community precipitation estimation and forecasting platform, as proposed for Q2 development;
- Routine reanalysis of precipitation using data (particularly rain gauge reports) of increasing latency, including daily updates to the Analysis of Re-

cord, with a time lag of 1-2 days, possibly involving staff resources at RFCs and WFOs for quality control;

- For hydrologic model calibration, a gridded precipitation dataset of at least 10 years' duration, using gauge, radar, and satellite data; possibly building on the Analysis of Record methodology; and
- Techniques compatible with OHD's ensemble products, which explicitly consider uncertainty of noisy precipitation observations in model calibration and real-time operations.
- Support for the development and dissemination of GPM products.

4.2 Forecasted Precipitation

Forecasts of precipitation are produced over a wide range of spatial and temporal scales. Uncertainty in these forecasts generally increases with lead time and decrease with larger spatial averaging.

4.2.1 Where We Are

Forecasts in the very short range (less than three hours) are based on automatic or subjective extrapolation of current radar and satellite features. Automated algorithms for 0-1 hour precipitation in 4-km gridded form are available in the AWIPS System for Convection Analysis and Nowcasting (SCAN). Probabilistic and categorical forecasts of point rainfall exceeding a fixed threshold up to 50 mm are available from an Advective-Statistical System (ADSTAT). However the latter forecasts are not specific to individual basins.

Forecasts for the near-term (beyond 3-12 hours) are based on a combination of radar-feature extrapolation and output from the Global Forecast System (GFS) and North American Mesoscale (NAM) models of NCEP. Manual modifications based on experience and physical logic are made to gridded precipitation fields by NCEP/HPC forecasters, and by HAS forecasters at RFCs. Output from the two models is subjectively weighted according to recent performance in the areas of interest. Longer-term forecasts are based on GFS and NAM output, again with the most significance attributed to that which has performed best in handling the weather system of interest.

Statistical interpretation (i.e., probabilistic forecasting) is based on operational Model Output Statistics (MOS) and/or ensemble output of the GFS, derived by applying random error fields to the most recent observed initial conditions.

4.2.2 What Our Partners Are Doing

Research and development in very short-range prediction is ongoing in institutions including NCAR, NSSL, and universities. There is much support for this

activity, since phenomena associated with heavy rainfall have a major impact on transportation, power generation, and public safety.

Research and development in numerical QPF is undertaken at NCEP/EMC and numerous academic and government institutions. Implementing organizations within NOAA include NCEP and ESRL. Expertise in statistical guidance based on NWP models resides in the Meteorological Development Laboratory (MDL).

In NCEP/EMC, ensemble prediction of QPF (and ensemble prediction of the other forcing fields discussed in this chapter) is being enhanced by applications of 1) ensemble short-range predictions (1-4 days) with an ensemble of mesoscale models in the Short-Range Ensemble Forecast (SREF) system, 2) ensemble medium-range prediction (1-2 weeks) from the Global Ensemble Forecast System (GEFS) based on the GFS, and 3) ensemble seasonal-range predictions (1-12 months) from the global coupled ocean-land-atmosphere Climate Forecast System (CFS). Moreover, the suite of medium-range ensemble forecasts is being expanded by the realtime acquisition of ensemble global forecasts from Environment Canada (EC). The combination of the GEFS and EC global ensemble forecast systems is denoted at NCEP as the North American Ensemble Forecast System (NAEFS). Efforts are underway to expand the NAEFS by realtime acquisition of ensemble global forecasts from other national and international NWP centers. Downscaling and bias correction techniques are being developed and applied by EMC and EMC partners for SREF, NAEFS, and CFS.

OHD presently collaborates with the Atmospheric Sciences Institute of the Czech Academy of Sciences in the refinement and testing of new QPF techniques.

A Collaborative agreement has started with the New Mexico Institute of Mining and Technology to study short-term precipitation forecasting and its impact on hydrologic modeling.

NOAA HMT is evaluating deterministic model ensembles for QPF and other state variables out to three days at basin-scale resolution. Techniques extending forecasts to five days and beyond are also being evaluated. Various and new QPF verification methods are employed to evaluate new and existing model performance

4.2.3 Where We Want to Be

As with antecedent precipitation estimates, forecasts will be integrated with operations, including flash flood monitoring, in which rather limited use of objective forecasts has been made. Forecast uncertainty will be adequately quantified for end users. Output from newly-updated Numerical Weather Prediction (NWP) models will routinely be downscaled and corrected for statistical biases prior to use using an adequately long archive of updated model forecasts to calibrate the

downscaling and bias removal algorithms and assure reliable processing of atmospheric forecasts.

For very short-term forecasts, blending of current remote-sensor data and the most recent operational forecasts will offer the greatest potential for improvements in flash flood forecasts. It is possible that for some time into the future, the very greatest benefit will be from a combination of remote-sensor and human input into algorithms that forecast convective initiation and decay. However, this human input is often not possible for short lead time and high resolution models.

For numerical modeling itself, advances in data assimilation, particularly radar assimilation, appear to offer the best hope for major improvements in flash flood forecast lead times, particularly if the numerical models can accurately simulate convective initiation and phenomena such as back-building convective cells. Retrospective ensemble forcings (both analyses and forecasts) will be used to produce retrospective ensemble streamflow simulations and forecasts (see *Chapter 6*).

4.2.4 Challenges to getting there

For all forecasts, reduction of uncertainty and quantification of uncertainty are overarching challenges. Two specific challenges for very short-term forecasts (out to approximately 9 hours), particularly for flash flooding, include: (1) implementation of documented methods for radar extrapolation using multiple radars and incorporation of gauge/radar bias (Multisensor Precipitation Nowcaster), and (2) implementation and utilization of recently-developed multisensor integrator systems that offer the potential to predict convection initiation (Autonowcaster); However, for short-term forecasts (9-18 hours), NWP assimilation systems for radar reflectivity data (e.g., Hu and Xue 2007) have some promise for predicting the evolution of mesoscale convective systems and storm formation. For longer-term forecasts, challenges include generation of unbiased deterministic and ensemble forecasts. In particular, ensemble forecasts of precipitation and temperature must feature the climatic degree and form of intercorrelation. Another challenge is to understand how to use a combination of dynamic downscaling and statistical processing of atmospheric forecasts for all forecast lead times from a few hours out to about a year.

4.2.5 A Road Map for Getting There

OHD expertise and resources dictate that our primary focus for the next several years will continue to be in short-term prediction based on extrapolative, advective, and statistical techniques. Our role will evolve to include development of statistical guidance (ensemble QPF and QTF) to serve river forecasting and ensemble river forecasting, in partnership with NCEP and MDL.

Specific waypoints on the proposed road map might include:

- Implementation of the High-resolution Precipitation Nowcaster;
- Refinement of the existing 0-3h advective-statistical Probabilistic QPF (PQPF) package (operated by MDL with development work done by OHD) to: produce higher-spatial resolution output; incorporate better physics through regionalization of the statistical equations that generate probabilities; and utilize better radar analyses than are currently used, in particular using data from the national 3-D reflectivity mosaic rather than the 10-km Radar Coded Message mosaic;
- Extension of extrapolative or advective techniques and blending with NWP model output to cover the 2-9 hour time domain, which is important to NCEP and RFC QPF operations, with work done primarily by OHD;
- To cover the 9-18 hour time window, incorporation of output from cloud-resolving models, particularly those which assimilate radar data, in collaboration with NSSL, ESRL's Physical Sciences Division (PSD), and NCEP;
- Refined Model Output Statistics forecasts, including error distribution information, developed through collaboration between MDL and the hydrometeorology and ensemble prediction groups in OHD; and
- External research on potential impacts of climate change on precipitation frequency and runoff processes.

4.3 Observed Air Temperature and Humidity

Temperature influences runoff by controlling precipitation type and snowmelt, and affects river discharge by controlling river ice formation and breakup. Over time-scales ranging from days to weeks, both temperature and humidity influence evapotranspiration. The most obvious influence of near-surface temperature is on warming of snowpack and subsequent snowmelt. In areas with rugged surface topography, near-surface temperature or freezing level height affect the extent of the surface area receiving rain, which can generate runoff quickly, or snow, which might persist in frozen form for a considerable period before melting. In many river basins record floods are caused by rain on snow events.

Temperature also controls the formation and breakup of river ice, which in turn has marked effects on discharge and flooding through the development of ice jams.

Temperature and humidity both influence surface evaporation from bare soil and surface water and, during the growing season, plant transpiration. In general, higher temperatures and lower relative humidity lead to increased rates of both evaporation and transpiration.

Finally, both temperature and humidity strongly affect precipitation development, particularly through convection, over periods of 0-12 hours.

4.3.1 Where We Are

Temperature and humidity are automatically measured at fixed reporting sites and reported to RFCs and WFOs on an hourly or sub-hourly basis. Human observers report maximum and minimum temperature on a daily basis. Sounding information including freezing level is reported by rawinsonde and automated sensors on some commercial aircraft. In precipitation situations, estimates of the height and depth of the melting layer are inferred from radar data and used to modify previous estimates. Most automated temperature and humidity data are automatically given some quality control and then assimilated into numerical prediction models.

After human quality control, observed surface temperature and freezing level data are used to estimate snow accumulation, snowpack temperature, and snowmelt. Temperature and humidity also drive models for potential evapotranspiration (PET). Physically or statistically based spatial downscaling is used to account for terrain and climatological effects in interpolating point values to grids or basins.

In some RFCs, 6-hour basin average temperature is derived from empirical relationship with daily maximum and minimum temperature. OHD provided the factors for this empirical relationship using more than 10-year's worth of data.

4.3.2 What Our Partners Are Doing

To get observations with high spatial and temporal resolution, more local mesonet observation network have been and are being added. For example, the Oklahoma Mesonet, which has 110 stations, can provide data up to every 5 minutes. Other local mesonets, such as Texas mesonet by Texas A&M University and Western Texas mesonet by Texas Tech. University, are being expanded to get data with higher spatial and temporal resolution. The University of Washington developed a procedure that depends on latitude and time of year (to account for timing of solar energy forcing) in estimating hourly temperatures from daily maximum and minimum temperatures.

A program for collecting and, as appropriate, disseminating data from multiple networks is maintained through the Meteorological Assimilation Data Ingest System (MADIS) of ESRL.

NCEP/EMC and its data assimilation partners such as ESRL continue to improve the techniques and temporal and spatial resolution of gridded national mesoscale analysis systems. The major such current thrust in EMC and ESRL is the devel-

opment and realtime execution of a national hourly mesoscale analysis system at 2-5 km resolution, known as the Real Time Mesoscale Analysis System (RTMA). The RTMA applies modern data assimilation algorithms, observational quality control, and concerted application of the mesonet observing networks such as those cited above, including MADIS. The RTMA can be a key source of the gridded analyses of observed temperature, humidity, and wind fields needed by NWS hydrological models, and at the high spatial resolution needed to resolve crucial orographic patterns in mountainous terrain. This is further emphasized in Section 4.5.5.

The RTMA is not presently available over the multi-decadal historical periods needed for hydrological model calibration. To fill this need for retrospective mesoscale analysis, NCEP/EMC has recently completed the nearly 30-year North American Regional Reanalysis, or NARR (Mesinger *et al.*, 2006), whose domain spans all of North and Central America at 3-hourly, 32-km resolution for the period 1979-present. NCEP/CPC maintains an ongoing daily realtime update of the NARR mesoscale analyses. Additionally, the multi-institution NLDAS project of the CPPA program has developed algorithms to downscale the historical and realtime NARR analyses of surface forcings to the 1/8-deg NLDAS grid (nominally 14-km) at hourly resolution, thereby producing a 1979-to-present realtime surface forcing suite to drive executions of the NLDAS multiple land models from 1979 to present. These NLDAS algorithms to downscale NARR 32-km surface forcings could easily be adapted to downscale NAAR forcing fields to the 3-4 km national HRAP grid of OHD. Moreover, given the multi-decade extent and frozen configuration of the NARR, bias correction algorithms could be developed and continually applied to the realtime NARR-based surface forcing, for purposes of driving hydrological models in historical and realtime mode.

4.3.3 Where We Want to Be

We would like to have access to or produce ourselves the following products: routine updates to freezing level data based on dual-polarization radar precipitation phase classification; temperature prediction systems such as Model Output Statistics for creation of gridded fields to support distributed modeling; an alternative temperature and humidity data source to fill gaps in mountainous areas far from any observation sites; and uncertainty information or ensemble forecast of air temperature, freezing height, precipitation type, and humidity.

4.3.4 Challenges to Getting There

The challenges are mainly the lack of: real-time updates to freezing level, particularly in mountainous regions; statistically unbiased gridded forecasts; ensemble temperature forecasts that are physically consistent with temperature/precipitation correlations; temperature observations and forecasts outside the territory of the U.S; temperature and humidity observations and forecasts with

higher spatial and temporal resolution; and an adequate objective-analysis algorithm or data-assimilation algorithm to convert point observation data to gridded data. Also needed are procedures to estimate uncertainty in precipitation type.

4.3.5 A Road Map for Getting There

The proposed waypoints on the road map are:

- To improve observing capabilities, routine updating of data acquisition to ingest mesonet surface temperature and humidity observations as they become available with automated quality control to incorporate these observations seamlessly.
- Utilization of evolving techniques for temperature interpolation in mountainous terrain, including high-resolution climatology and selective objective interpolation using elevation, land use, land cover, and aspect information, including the University of Washington's temporal interpolation algorithm.
- Use of assimilation techniques such as those of the Real-time Mesoscale Analysis (RTMA) system, which generates high-resolution temperature fields through interpolation based on physical constraints.
- In the long term (5+ years), creation of a dataset of surface temperature of sufficient duration (more than 10 years) on a national scale, for hydrologic model calibration, to support retrospective simulation and ensemble forecasting.

4.4 Forecasted Air Temperature and Humidity

Because the improvement of the tools to issue forecasts of air temperature and humidity is outside of the purview of the NWS Hydrology Program, we provide below only a brief summary of what are partners are doing and some suggestions for improved collaboration with those partners

4.4.1 What Our Partners Are Doing

Forecasting temperature and humidity from mesoscale numerical models (like NAM, WRF) is the major responsibility of NCEP. As computing capability and modeling techniques advance, the gridded temperature and humidity from the numerical model will better support distributed hydrological modeling.

The Meteorological Development Laboratory maintains and develops Model Output Statistics (MOS) systems for the production of statistically unbiased temperature and humidity information.

Suggestions for improving the utility of air temperature and humidity forecasts for hydrologic forecasting include:

- Continue ongoing work to support HL's Hydrologic Software Engineering Branch (HSEB) upgrades to operational systems, such as those for estimating 6-h average temperature from daily maximum/minimum values, or from the National Digital Forecast Database (NDFD) grids. Some RFCs are starting to issue routine forecasts with 1-hour time intervals. Therefore, there is a need to develop procedures to produce gridded 1-hour forecasts of temperature, that support lead times as low as 1-hour. These procedures must be able to support reforecasts beginning in 1979, and to process observations beginning in 1948.
- Gridded MOS temperature forecasts should be made available for basins flowing into U. S. territory from outside the U. S., especially Canada.
- OHD and NCEP will collaborate to generate forecast ensembles of temperature fields to support ensemble streamflow prediction. One aspect of this task is to insure proper statistical relationships between temperature and precipitation anomalies, which have significant correlations that vary in time and with respect to location. Another aspect is to have an adequate archive of reforecasts to support ESP hindcasts.

4.5 Winds

The chief influence of near-surface winds is through its effect on evaporation from bare surfaces, sublimation of snow and plant transpiration. In general, stronger winds increase the rates of both evaporation and transpiration. Outside the growing season, strong winds are a prime factor in soil desiccation.

Winds can influence discharge in large rivers directly through surface drag, or through tidal effects in adjoining lakes and estuaries. Winds exert control on the movement of ice pack in the larger water bodies into which rivers flow. Finally, winds are a prime controlling influence on wildfires, which in turn change surface characteristics through destruction of vegetation and the creation of burn scars.

Finally, winds control the movement of heat energy and water vapor. These processes are particularly important in the development of convective precipitation.

4.5.1 Where We Are

Most near-surface winds over land are derived from *in-situ* observations, collected mainly by automated observing systems. Over oceans, estimates of surface wind vectors can be derived from satellite-based scatterometer observations. Spatial interpolation of wind vectors over land is difficult, particularly in rugged terrain. Therefore, gridded wind-field estimates are generally derived from assimilation of point observations in a numerical weather prediction model such as the

RUC-2 (Benjamin *et al.* 2004a-b). These models yield wind fields that are in balance with the observed temperature and mass fields while reflecting topographic influences. This process is presently used to create the operational RTMA wind field. In addition there is also the Regional Reanalysis winds and all other surface forcing variables for the period 1979 - present.

Forecasts of winds are output by numerical prediction models including the NAM and GFS (Kalnay *et al.* 1990; Moorthi *et al.* 2001). These forecasts, though generally realistic, may contain statistical biases. Therefore the NWS also issues forecasts of winds at specific points, and gridded wind fields, derived from the basic model output through the Model Output Statistics (MOS) technique.

4.5.2 What Our Partners Are Doing

Real-time ingestion and quality control of near-surface observations over land and adjacent coastal waters are controlled by NCEP, other NOAA offices (for example ESRL through MADIS in the western U. S.) and many other government entities (Bureau of Land Management, U. S. Department of Agriculture, state environment agencies). These organizations maintain and expand mesonets, support communications infrastructure, and maintain and refine quality control procedures.

Wind forecast capabilities are maintained and developed by NCEP and MDL. Modeling of wind effects on estuaries is carried out in MDL, which maintains an operational system for tide departure forecasts, and other partners such as North Carolina State University. We collaborate with the latter organization through the Coastal and Inland Flooding Observation and Warning (CIFLOW) program.

NOAA's National Ocean Service through the Center for Operational Oceanographic Products and Services (CO-OPS) also has an operational system at a number of estuaries that forecasts the effect of winds on water levels and currents. The winds are downscaled from the latest runs of the North American Mesoscale model (NAM).

4.5.3 Where We Want to Be

We would like to have real-time access to accurate and unbiased estimates and forecasts of wind speed and direction, or vector winds, on a high-resolution grid mesh over all land and coastal areas. In the immediate future, forecasts need to be produced through 10 days. Eventually, we'll need wind forecasts out to one year. A particular need is the availability of spatially continuous ensemble wind fields for modeling of evapotranspiration processes and movement of water in estuaries and large lakes. We also need techniques for bias removal and downscaling wind ensemble forecasts to produce ensemble wind forcing members that are consistent with all other ensemble forcing members.

4.5.4 Challenges to Getting There

As with most weather elements, reduction and quantification of uncertainty in estimates and forecasts are primary challenges.

While the availability of automated surface observations is generally increasing (with the deployment of surface mesonets and the development and deployment of sounding systems operated on commercial aircraft), the quality of the observations is often compromised by undesirable siting, mechanical equipment failures, or communications failures. Therefore, automated quality control is a prime concern, as is a method for robust estimates of the error distribution. Description of error bounds is complicated by the fact that wind is an essentially a two-dimensional vector quantity.

4.5.5 A Road Map for Getting There

In general, we should coordinate with NCEP and MDL to insure that our hydrologic prediction systems can take advantage of developments to observing and prediction capabilities.

To improve observing capabilities, we should use routine updating of data acquisition to ingest mesonet surface wind observations as they become available. Automated quality control will be needed to incorporate these observations seamlessly. Assimilation techniques such as those of the RTMA system, which generates high-resolution wind fields through interpolation based on physical constraints, including topography, should be used.

An effort to expand gridded MOS temperature forecasts to include some basins outside the immediate conterminous U. S. is being investigated with MDL staff. It will be possible to make a similar expansion of humidity or dew point temperature grids, though a scarcity of ground-truth data for equation development outside the U. S. complicates this effort.

4.6 Shortwave/Longwave Radiation and Skin Temperature

Incoming shortwave radiation and outgoing longwave radiation are prime drivers of surface evaporation, plant transpiration, and snowmelt. Diagnosis of these radiative fluxes is crucial to estimating evapotranspiration and snowmelt processes, which processes are very difficult to measure directly and which must be estimated through empirical or physical models.

4.6.1 Where We Are

As recently as the early 1990's, shortwave radiative input to hydrologic prediction models came primarily from human estimates of sky cover, and longwave estimates from surface air temperature. These estimates were terminated with the introduction of ASOS. Although there are Solar and Infrared Radiation Observation Stations and a Baseline Surface Radiation Network, the radiation observations networks are sparse and difficult to use to estimate spatially-continuous radiative flux fields. Geostationary Satellite and/or multi-satellite estimations are expected to be the primarily sources of shortwave/longwave radiation data.

4.6.2 What Our Partners Are Doing

Advances in satellite algorithm development for both operational and future platforms are carried out and supported by NASA and NESDIS as well as the Space Weather Prediction Center (SWPC) and many research institutions and universities. Derivation of radiation and temperature from multi-satellite observations, data quality control, and development new satellite algorithms and new satellite technology are going on at NASA and NESDIS. At present some proxies for cloud-cover estimates do exist (for example the ASOS Geostationary Operational Environmental Satellite (GOES)-derived skycover product). Extensive research is ongoing to explicitly estimate surface radiative balance from geostationary satellites. OHD has supported some research into the impact of these estimates on hydrologic models (e.g., Jacobs, 2005). Operational NWP models, such as the NCEP GFS, simulate radiative transfer within their physics packages and the results are generally included as part of the forecast product suite. There have been some evaluations of these radiation simulations (Yang *et al.* 2006). NESDIS produces real-time estimates of hourly downwelling solar insolation that account for effects of cloud cover. Retrospective analyses are available since the early 1980's. These are being use at NCEP as an input to its NLDAS system.

4.6.3 Where We Want to Be

We would like to be able to use satellite shortwave/longwave radiation and skin temperature observations in combination with physically based models to estimate soil moisture, evapotranspiration, and the surface radiation balance. We would also like to be able to forecast the same quantities using NWP models.

4.6.4 Challenges to Getting There

Reduction and quantification of uncertainty are overarching challenges here as well.

For radiative flux estimates, cloud cover plays a crucial role in determining shortwave/longwave radiation and skin temperature. Because of very large cloud variability, insufficient sampling, lack of diurnal cycle coverage, accuracy across all cloud types will be very difficult and expensive to achieve. Improvement of

satellite algorithms and development of new satellite technology are the best choices for radiation observations. The greatly improved measures of cloud ice water path are required. Both spectral and broadband measurements of solar and thermal infrared radiation from space and ground sites are needed to improve and develop satellite algorithms. The data quality control is also a challenge. If satellite-based estimates of short-wave radiation are used, there is an additional challenge to get long wave estimates (net or downwelling) that are consistent with the shortwave estimates.

As new sources of forcing data become available it is essential that the climatological statistics of the new forcing be consistent with the climatological statistics of the forcing used for model calibration. Moreover, this needs to be done in a way that it is possible to make retrospective simulations and hindcasts so that we, and our users, understand the strengths and limitations of our products.

4.6.5 A Road Map for Getting There

While OHD does not work directly on radiation observing and simulation systems, it does have a role in testing the impact of such systems on hydrologic prediction, specifically on soil moisture and runoff estimation and energy-balance snowpack models. These impacts can be tested in-house or by working with external partners. One avenue is to test the utility of the ASOS skycover product, in gridded form, as a proxy for manual surface estimates of sky cover. The NWSRFS already has logic for ingesting such skycover input. The previous external collaboration in testing the impacts of satellite-based shortwave/longwave balance on NWSRFS simulations can be extended to distributed hydrologic models and snowpack models. For longer-term (5+ years) the development of a PET dataset of 10 years' duration on a national scale must be considered for hydrologic model calibration.

Investigation of the impact of NWP forecasts of radiation components is required, followed by operational use in PET forecasts.

4.7 References

- Benjamin S. G., Coauthors. (2004a). "An hourly assimilation/forecast cycle: The RUC". *Mon. Wea. Rev.*, 132, 495–518.
- Benjamin S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck. (2004b). "Mesoscale weather prediction with the RUC hybrid isentropic/terrain-following coordinate model". *Mon. Wea. Rev.*, 132, 473–494.
- Daly, C., R.P. Neilson, and D.L. Phillips, 1994: A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *J. Appl. Meteor.*, **33**, 140–158.

- Hu, M., and M. Xue. (2007). "Impact of configurations of rapid intermittent assimilation of WSR-88D radar data for the 8 May 2003 Oklahoma City tornadic thunderstorm case". *Mon. Wea. Rev.*, 135, 507–525.
- Jacobs, J. (2005). "Development of Enhance Techniques for Computing Potential Evapotranspiration". *Report to the Office of Hydrologic Development*, 77 pp.
- Kalnay E., M. Kanamitsu, and W. E. Baker. (1990). "Global numerical weather prediction at the National Meteorological Center". *Bull. Amer. Meteor. Soc.*, 71, 1410–1428.
- Krajewski, W.F., Ciach G.J., and Villarini, G. (2005). "Towards Probabilistic Quantitative Precipitation WSR-88D Algorithms: Data Analysis and Development of Ensemble Generator Model: Phase 4". *Report to OHD under NOAA Contract: DG133W-02-CN-0089*, 202pp.
- Lee G., Seed, A. and Zawadski, I. (2006). "Generating ensembles of radar estimates of precipitation and ensemble nowcasts". *AGU Spring Conference*, Baltimore, MD.
- Mesinger, F., G. DiMego, E. Kalnay, P. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, K. Mitchell, E. Rogers, M. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish and W. Shi (2006), North American Regional Reanalysis, *Bulletin of the American Meteorological Society*, 87, 343-360.
- Moorthi S., H.-L. Pan, and P. Caplan. (2001). "Changes to the 2001 NCEP operational MRF/AVN global analysis/forecast system". *Tech. Procedures Bull. 484*, Office of Meteorology, National Weather Service, 14 pp. [Available online at <http://205.156.54.206/om/tpb/484.htm>.].
- Ryzhkov, A.V., S.E. Giangrande, and T.J. Schuur. (2005a). "Rainfall estimation with a polarimetric prototype of WSR-88D". *J. Appl. Meteor.*, 44, 502–515.
- Ryzhkov, A.V., T.J. Schuur, D.W. Burgess, P.L. Heinselman, S.E. Giangrande, and D.S. Zrnich. (2005b). "The Joint Polarization Experiment: Polarimetric Rainfall Measurements and Hydrometeor Classification". *Bull. Amer. Meteor. Soc.*, 86, 809–824.
- Vasiloff, S., and Coauthors. (2007). "Improving QPE and very short-term QPF: An initiative for a community-wide integrated approach." Accepted for publication in *Bulletin of the American Meteorological Society*.

Yang, F., H.L. Pan, S.K. Krueger, S. Moorthi, and S.J. Lord. (2006). "Evaluation of the NCEP Global Forecast System at the ARM SGP site." *Mon. Wea. Rev.*, 134, 3668–3690.

Zhang, J., K. Howard, and J.J. Gourley. (2005). "Constructing three-dimensional multiple-radar reflectivity mosaics: examples of convective storms and stratiform rain echoes". *J. Atmos. Oceanic Technol.*, 22, 30–42.

5. Anthropogenic and Natural Perturbations to the Hydrologic Cycle

Perturbations to the hydrologic cycle involve natural and anthropogenic changes to the climate, the effect of large-scale irrigation, and the effects of reservoir regulation on river forecasts. Only the latter is cover in this version of the plan.

5.1 Climate Change and Variability (Future)

5.2 Irrigation (Future)

5.3 Reservoir-based River Regulation

Streamflow regulation refers to the man-made changes to natural flow regimes. Those changes may be the result of reservoir operations, water withdrawals, water returns, and pumping from aquifers. River regulation is a complex problem, because it involves legal, economic, and technical considerations, as opposed to the natural laws that govern the flow of water in unregulated rivers. NWS often receives short-term reservoir release projections for major reservoirs operated by its cooperating partners. However, there are many reservoirs which, for a number of reasons, NWS receives no information concerning projected releases. In addition, NWS must be able to anticipate reservoir releases beyond these short-term projections in order to provide longer-term probabilistic forecasts at downstream forecast points. A further complication is the administration of water rights. In western states, water rights follow the prior appropriation doctrine as opposed to the riparian system. Under the prior appropriation doctrine, water rights are based on seniority. In times of shortages, senior rights must be satisfied first, regardless of their location on the river. Furthermore, the problem is compounded by the fact that those water rights are a marketable commodity, which allows owners to buy and sell the right to withdraw water.

5.3.1 Where We Are

Although deterministic models can capture much of the streamflow variability due to regulation, deterministic approaches cannot account for uncertainty caused by external factors (e.g., power market, legal mandates), human factors (e.g., maintenance decisions, subjective operations), and small-scale complexity. During 2005 and 2006, Riverside Technology, inc. (RTi), under contract with OHD, addressed the problem of river regulation using a deterministic approach. Although the project identified some areas of the National Weather Service River Forecast System (NWSRFS) that could be improved, the net outcome was to confirm that a pure deterministic approach is not suitable for modeling river regulation. It follows, then, that deterministic models must be complemented with probabilistic approaches in a manner that is compatible with the ensemble approach that NWS is implementing.

The NWSRFS currently includes three deterministic models for modeling reservoir regulation. RES-SNGL was designed to model short-term reservoir releases for single reservoirs. The RES-SNGL model was primarily designed to capture the reservoir functionality needed to model United States Army Corps of Engineers (USACE) reservoirs in the Southeastern U. S. The model provides limited utility for long-term ensemble forecasting, and it is not able to model systems of reservoirs operated for flood control or the water supply/flood control reservoirs common in the West. The SSARRESV is based on the Streamflow Simulation and Reservoir Regulation System (SSARR) developed by the Northwest River Forecast Center (NWRFC) and the North Pacific Division of the USACE. This model relies on regulation options specified at run-time to forecast reservoir releases. It was specifically designed to model USACE reservoirs in the Northwest. The Joint Reservoir Regulation Operation (RES-J) model is an object-oriented network reservoir model that was designed to model a network of reservoirs and control points as a system. RES-J was also designed to support long-term ensemble forecasting. Users can combine a flexible set of methods to mimic historical operations.

An important project is also underway to integrate the HEC Reservoir System Simulation (ResSim) model with NWSRFS. This project will allow better coordination between CNRFC, USACE, the California Department of Water Resources, and local stakeholders by ensuring that all project participants have an identical representation of the river system at all times. The integration is being done using a “service-based” architecture to ensure compatibility with CHPS.

5.3.2 What Our Partners Are Doing

In the mid-1990s the United States Bureau of Reclamation (USBR), the Tennessee Valley Authority (TVA), and the Western Area Power Administration (WAPA) supported the development of the RiverWare river basin modeling system. RiverWare is a tool for scheduling, forecasting, and planning reservoir operations, and is currently supported by the University of Colorado’s Center for Advanced Decision Support for Water and Environmental Systems (CADSWES). RiverWare includes controllers for: solving a completely specified problem, rule-based simulation, linear goal-programming optimization, and multiple run management. USBR uses RiverWare as a long-term policy and planning model for the Colorado River, as well as a daily operations model for both the Upper and Lower Colorado regions. USBR also uses RiverWare on the Yakima, Rio Grande, and Truckee River Basins.

HEC has developed the Corps Water Management System (CWMS) to support real-time flood control operations by the USACE. ResSim is the model included in CWMS for modeling reservoir operations. HEC is currently in the process of

incorporating RiverWare into CWMS, because some USACE district offices use RiverWare for reservoir and river basin modeling.

5.3.3 Where We Want to Be

NWS requires a range of deterministic models and probabilistic modeling approaches that can be selected based on an assessment of the streamflow regulation to be modeled and consideration of an appropriate level of operational complexity, cost, and benefit in terms of forecast accuracy. The approaches for modeling streamflow regulation must be compatible with an ensemble-forecasting framework, and they must be robust across short-, mid-, and long-term time scales. Furthermore, forecasts of streamflow regulation must account for uncertainties due to water rights administration, human factors and known or unknown external factors.

5.3.4 Challenges to Getting There

Some of the challenges are:

- Consideration of the effect of water rights administration;
- Short-term operations can fluctuate dramatically based upon non-hydrologic variables;
- Actual regulation of reservoirs frequently does not follow the operating rules, because the operators often have latitude in operations;
- Basins include overlapping federal, state, and private projects with competing objectives;
- Individual effects of small regulator operations may be insignificant, but cumulatively the operations can completely alter the river's flow regime; and
- Private companies may be reluctant to share operational data with outside parties, including the NWS.

5.3.5 A Road Map for Getting There

As of late 2007, the problem of river regulation remains a major obstacle in the implementation of AHPS forecast points throughout the country. It was identified by the Hydrologists-in-Charge as their top priority. Furthermore, the AHPS Innovation goal team gave it its top priority. The first step on the road map is to award at least one collaborative research grant to an institution that clearly understands the problem, and with the capacity to formulate a procedure that could be implemented into operations. This procedure must be completely compatible with the short- and long-term ensemble work currently in active development and testing in the NWS (see *Chapter 6*). The second step will be the creation of a CHPS model that will incorporate the approach developed by the research project into the operational system.

6. Ensemble Forecasting

Hydrologic states are generally sparsely observed (e.g., soil moisture), hydrologic processes are highly nonlinear (e.g., surface runoff) and boundary conditions are highly variable in space and time (e.g., precipitation). Quantification of uncertainties associated with the major sources of error, understanding of how uncertainty propagates through the hydrology and water resources systems, and quantification of the integrative uncertainties associated with the products and services are necessary not only for risk-based decision making by the forecasters and users of the operational hydrology and water resources products but also for cost-effective improvement of forecast systems and processes.

The need for reliable and skillful ensemble and probabilistic hydrology and water resources forecasts has grown greatly in recent years as more users practice risk-based decision making. The range of spatio-temporal scale for which such probabilistic forecast information needs to be produced is very large. **Figure 6-1**, which depicts the overarching service goal for the Advanced Hydrologic Prediction Service (AHPS; McEnery *et al.*, 2005), shows the range of forecast lead-time for which reliable and skillful ensemble and probabilistic information must be produced to meet the needs of the multitude of customers and users.

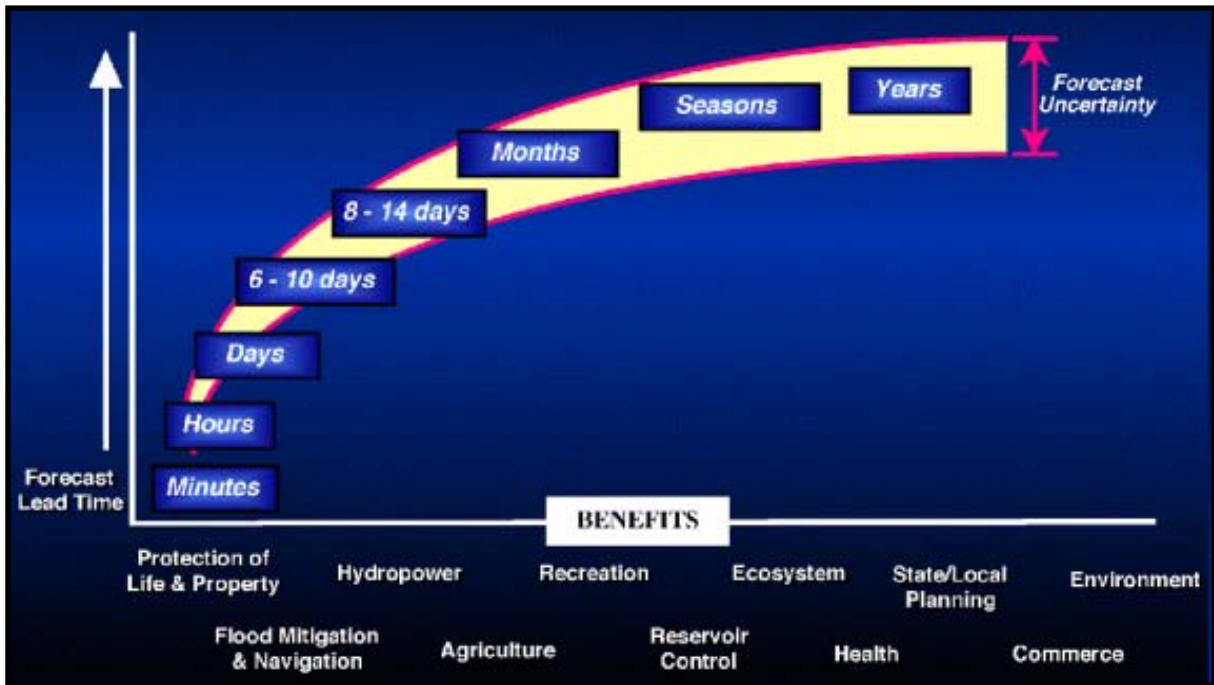


Figure 6-1

Uncertainty in hydrologic forecasts as a function of forecast horizon

The spatial scale at which the ensemble and probabilistic information needs to be produced ranges from local, regional to national, spanning several orders of magnitude (*Figure 6-2*). The reader is referred to McEnery *et al.* (2005) for specific examples of the multi-scale nature of hydrologic modeling that is necessary to meet the service needs.

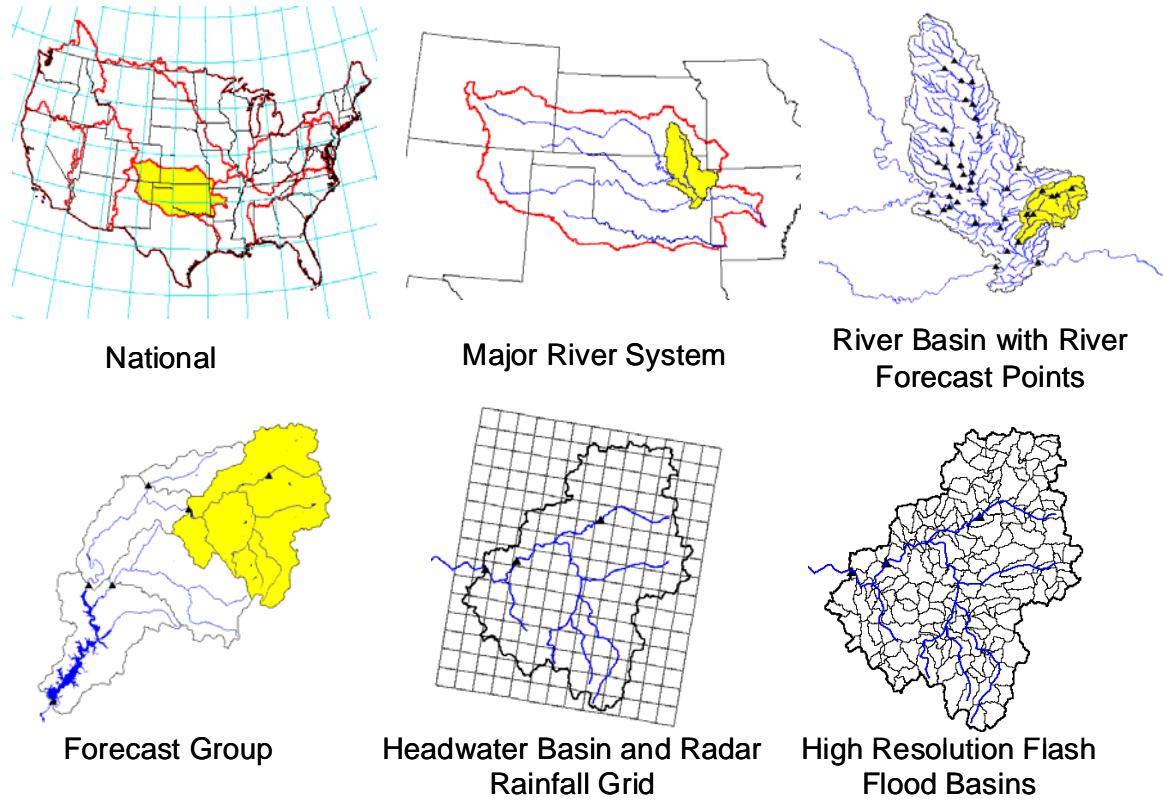


Figure 6-2

Spatial Scales for Hydrologic Modeling

Operational hydrologic ensemble forecasting has two overarching science goals. The first is to accurately quantify the integrative uncertainty associated with the principal forecast elements in hydrology and water resources products, such as streamflow and soil moisture. The second is to minimize the constitutive uncertainties cost-effectively. The left-hand side of *Figure 6-3* shows the major sources of error in hydrologic forecasting, and illustrates qualitatively how the uncertainty may increase as the forecast lead-time increases. The right-hand side of the figure identifies the components of the hydrologic ensemble forecast system (see below) that address reduction and quantification of the uncertainties.

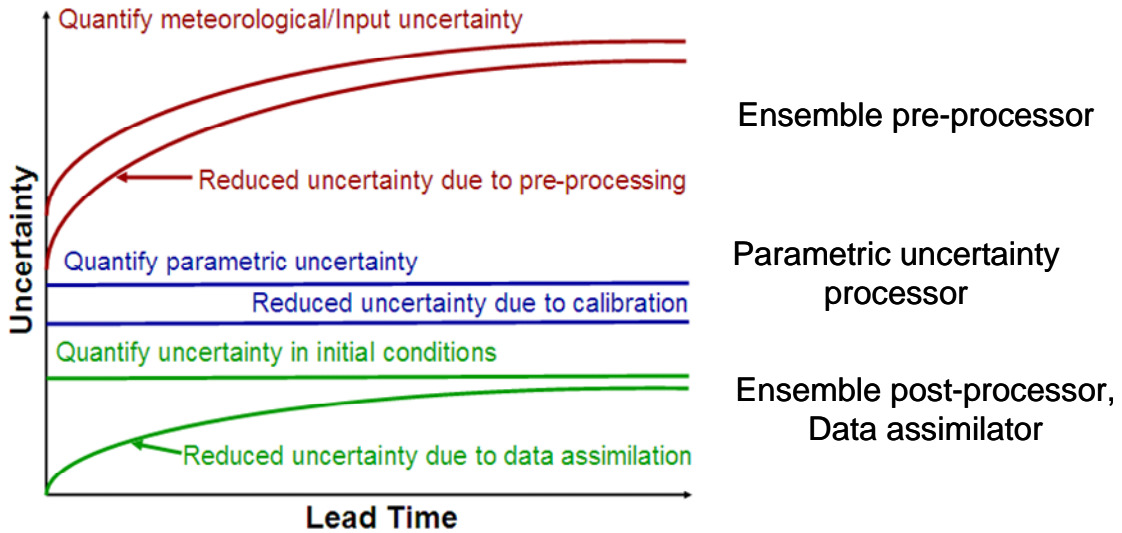


Figure 6-3

Uncertainties in Hydrologic Forecast

6.1 Where We Are

ESP has been in operational use by the NWS RFCs for over 20 years (Day, 1985). The ESP process was initially designed and implemented within the NWSRFS to serve as a long-range probabilistic forecasting tool. Although there are shortcomings, the technique and tools have served some of the RFCs and customers interested in long-range forecasts well. However, many customers are interested in short-term probabilistic forecasts for which ESP is not suitable.

With the implementation of AHPS, the NWS Hydrology Program has committed to meeting customer requirements for hydrologic forecasts and information, including uncertainty information, at all time scales (*Figure 6-1*). In mid-1990s, new techniques were developed to assimilate the monthly and seasonal outlook forecasts from NCEP/CPC (Perica, 1998).

Feedback from NWS customers and partners in recent years clearly indicates that the provision of reliable uncertainty estimates, particularly at shorter time scales, will dramatically increase the value of forecast services. Additionally, knowledge of forecast uncertainty will provide benefits to the forecast and warning decision processes within the NWS and cost-effective improvement of them.

In an effort to produce the uncertainty information for short-term forecasts, NWS initiated development of prototype capabilities for short-term ensemble forecasting in the late 1990s through early 2000s. They include the Ensemble Pre-Processor (EPP) for generation of ensembles of future precipitation and temperature from single-value quantitative precipitation and temperature forecasts (QPF, QTF; Clark *et al.*, 2004; Schaake *et al.*, 2007), the Ensemble Post-Processor for accounting of hydrologic uncertainties (Seo *et al.*, 2006), the Hydrologic Ensemble Hindcaster (HEH) to enable hindcasting and to support large-sample verification of streamflow ensembles (Demargne *et al.*, 2007), and the Ensemble Verifi-

ation System (EVS) for verification of precipitation, temperature and streamflow ensembles (Demargne *et al.*, 2007). Since the mid-2000's, NWS has expanded the capability of EPP to generate mid-range (from Day 1 through Day 14) precipitation and temperature ensembles from the mean of the ensemble forecasts from the NCEP's GFS (Schaake *et al.*, 2007). A number of RFCs have been operating these prototype tools experimentally (*Figure 6-4*).

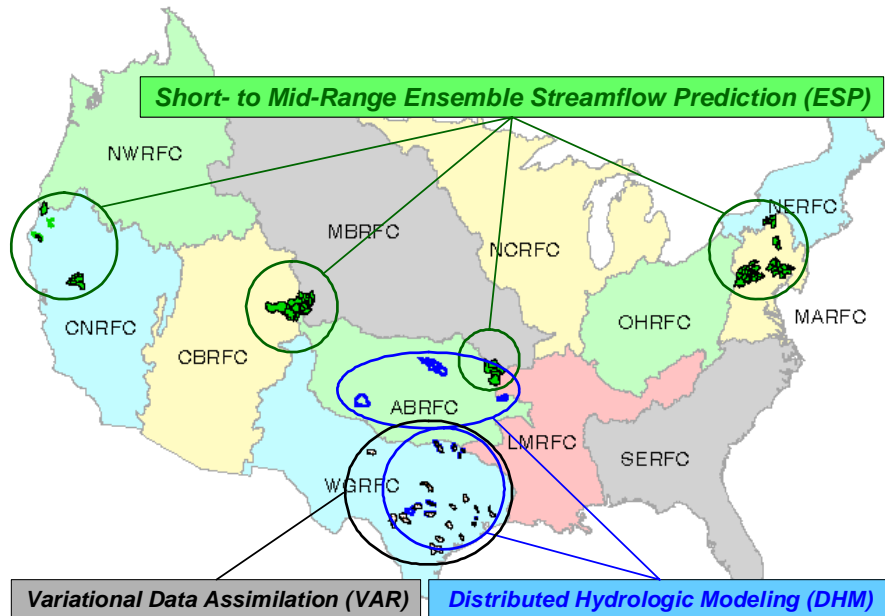


Figure 6-4

Current Experimental Operations of the Short Term Ensembles and Variational Data Assimilation (VAR)

6.2 What our partners are doing

- NCEP/EMC – NLDAS (Mitchell *et al.* 2004). The NLDAS community has been developing multi-model land surface modeling capability over CONUS. The models included are Mosaic, NOAH, SAC, and VIC. Of particular interest to OHD is large-scale evaluation of multi-model hydrologic ensembles driven by both analysis and ensemble forecasts of hydrometeorological forcings. OHD is actively collaborating with the EMC's land surface modeling group through the Climate Prediction Project for the Americas (CPPA) Core Project and with the NLDAS community for cost-effective research and development and transition of proven capabilities to the RFC operations.
- NCEP/EMC – Global Forecast System ensemble forecasting (Buizza *et al.* 2005). Reliable forcing ensembles can extend the lead-time of hydrologic forecasts in that longer-range hydrometeorological ensemble forecasts may

be input to the hydrologic models to produce hydrologic forecasts that reflect the lead-time dependence of uncertainty in the forcing forecasts that is necessary for decision-making. In addition to efforts to improve model physics and data assimilation, EMC is developing statistical techniques for post processing, including bias correction and downscaling, to produce reliable and skillful hydrometeorological ensemble forecasts. OHD is actively collaborating with the EMC's global ensemble forecast system group through the The Observing System Research and Predictability Experiment (THORPEX) program Hydro project to improve the quality of hydrometeorological ensembles.

- NCEP/Hydrometeorological Prediction Center (HPC) - Confidence Interval Estimation for Quantitative Precipitation Forecasts (QPF) Using Short-Range Ensemble Forecasts (SREF) (Im *et al.* 2006). It is recognized that capturing flow-dependent skill is one of the most important aspects of assimilating hydrometeorological ensemble forecasts into operational hydrologic forecasting. HPC has developed a statistical technique that estimates confidence intervals for SREF. While such estimates are necessarily tied to the space-time scale at which the regression is developed, they provide RFCs with an additional context information necessary for interpretation of the hydrometeorological ensemble/probabilistic forecasts.
- NCRFC - Use of HPC QPF confidence interval forecasts to produce a hydrologic ensemble of river forecasts (Halquist 2006). The HPC confidence interval estimates have been made available to a number of RFCs for experimental use. NCRFC has been using them to generate stratified hydrologic conditional ensemble forecasts. While such hydrologic forecasts do not lend themselves to straightforward probabilistic interpretations, the experience can help assess the value of the flow dependence information in SREF inferable through the confidence interval.
- CBRFC, WR, CIRES, University of Colorado, NCEP, Princeton University, University of Washington – ensemble techniques (Clark *et al.* 2004, Gangopadhyay *et al.* 2004, Werner *et al.* 2004, 2005). CBRFC and WG, in collaboration with CIRES and the University of Colorado, developed and experimentally implemented a technique for assimilating precipitation and temperature ensemble forecasts from the frozen version of the Global Forecast System (GFS). They demonstrated the value of medium-range hydrometeorological ensemble forecasts for hydrologic forecasting, which motivated development of the Ensemble Pre-Processor II (EPP2) GFS Subsystem. These capabilities are being integrated/implemented into EPP3, the ensemble pre-processing component for XEFS.
- NASA/Goddard Space Flight Center (GSFC) – Land Information System (LIS) (Kumar *et al.* 2006). The Land Information System is a high-performance land-surface modeling and data assimilation system based on GSFC's Land Data Assimilation Systems. Through the NASA-NWS/OHD project (see Chapter 7), OHD is actively collaborating with members of the

LIS Team to enhance and transition multi-model ensemble and data assimilation capabilities to operations

- University of Washington – A testbed for new seasonal forecasting approaches in the western U. S. (Wood and Lettenmaier, 2005), Princeton University - A seasonal hydrologic ensemble forecast system over the eastern U. S. (Luo and Wood, 2006). Referred to as the westwide and eastwide seasonal hydrologic forecast systems, respectively, these testbeds generate experimental, real-time hydrologic and streamflow forecasts using a macroscale hydrologic simulation model. The westwide system uses ESP, ESP conditioned on El Niño Southern Oscillation (ENSO) and Pacific Decadal Oscillation (PDO), and ensemble forecasts downscaled from the CPC's seasonal outlooks. The eastwide system uses multimodel ensemble forcings merged via the Bayesian model averaging, among others. These systems are being implemented at NCEP, and OHD is actively collaborating with NCEP, the University of Washington, and Princeton University for enhancement and transition of proven capabilities to operations.
- Hydrologic Research Center (HRC), Georgia Tech - Integrated Forecast And Reservoir Management (Inform) For Northern California (Georgakakos *et al.* 2006). The primary objective is to demonstrate the utility of present-day meteorological/climate and hydrologic forecasts for the Northern California River and reservoir system. The system contains real-time short-range forecast components, off-line longer-range forecast components, and off-line decision components that span forecast and decision time scales from hours to seasons. Forecast uncertainty is explicitly characterized and used for risk-based decision support.
- A major challenge (current and future) is the successful transmission of probabilistic information to the users. By successful transmission, we mean that the ultimate users of the information are capable of understanding and using that information. At the time this version of the plan is being written, Aptima, Inc. a small business specialized in cognitive psychology and human factors engineering, was awarded an OHD Collaborative Research grant to evaluate how probabilistic information is successfully transmitted to the users. Along a separate path, the NWS Western Region developed a web page specifically directed to the users of its long-range probabilistic seasonal water supply forecasts.

6.3 Where We Want to Be

The vision is to be able to produce reliable and skillful ensembles for a wide spectrum of hydrology and water resources services (see *Figure 6-1* and *Figure 6-2*) from minutes to years into the future and over a range of spatial scales where the service needs exist. The envisioned hydrologic ensemble forecast system must be able not only to capture the integrative uncertainty associated with the hydrology and water resources variables over this range of spatio-temporal

scale, but also to reduce the various uncertainties in the forecast process (*Figure 6-3*) through pre-processing, data assimilation, and post-processing. *Figure 6-5* depicts this vision.

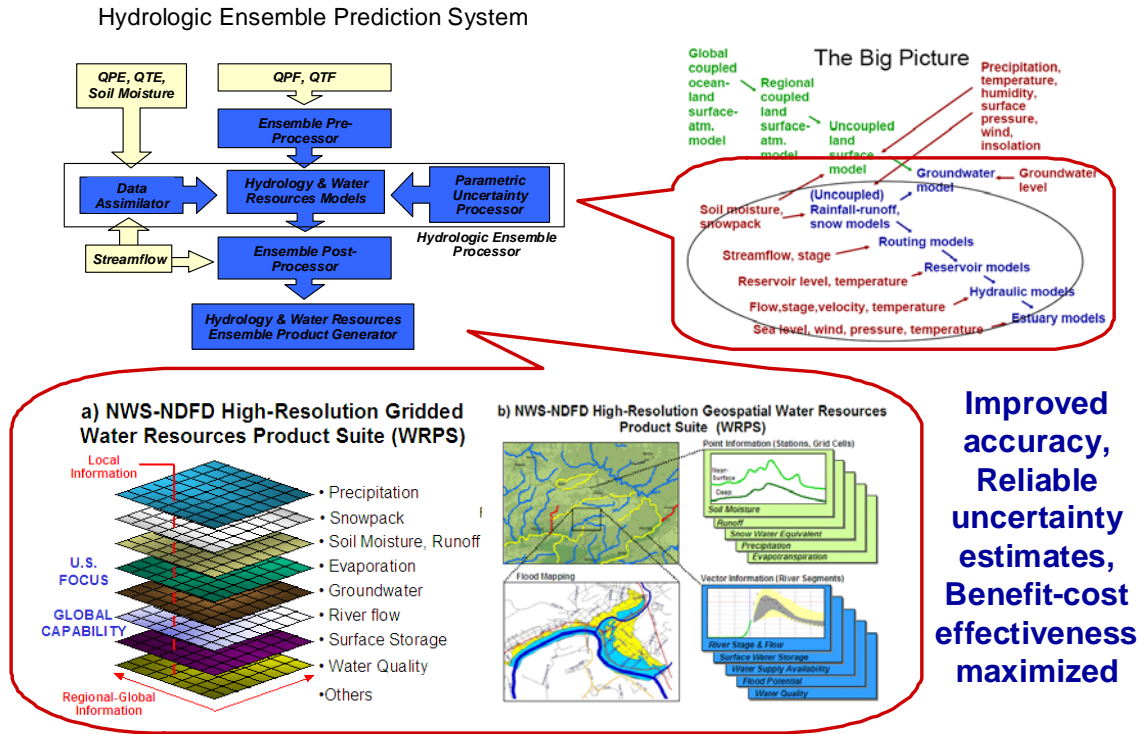


Figure 6-5

Vision for Ensemble and Data Assimilation in Hydrologic Forecast Operations

It is well known that structural errors in hydrology and water resources (in particular, rainfall-runoff) models are a major source of uncertainty. The multimodel ensemble approach (Georgakakos *et al.*, 2004) provides a framework in which the major sources of uncertainty (*Figure 6-3*) may be quantified and reduced while maintaining dynamic and statistical consistency of the processes modeled and products generated. Given the wide range of spatio-temporal scales over which ensemble and probabilistic information must be produced, the space-time scale at which the hydrology and water resources models can operate cost-effectively and the level science and technology can support may vary (e.g., at time steps of hourly, 6-hourly, etc. and at spatial scales of, e.g., HRAP, MAP, etc.). As such, the ensemble forecasting framework must allow operation of hydrology and water resources models at different space-time scales, and the science capabilities need to be developed to produce ensemble and probabilistic in-

formation from multiscale models that is statistically consistent across scale. **Figure 6-6** depicts this envisioned multi-model ensemble framework through which each of the major sources of uncertainty (**Figure 6-3**) may be accounted for, propagated and integrated.

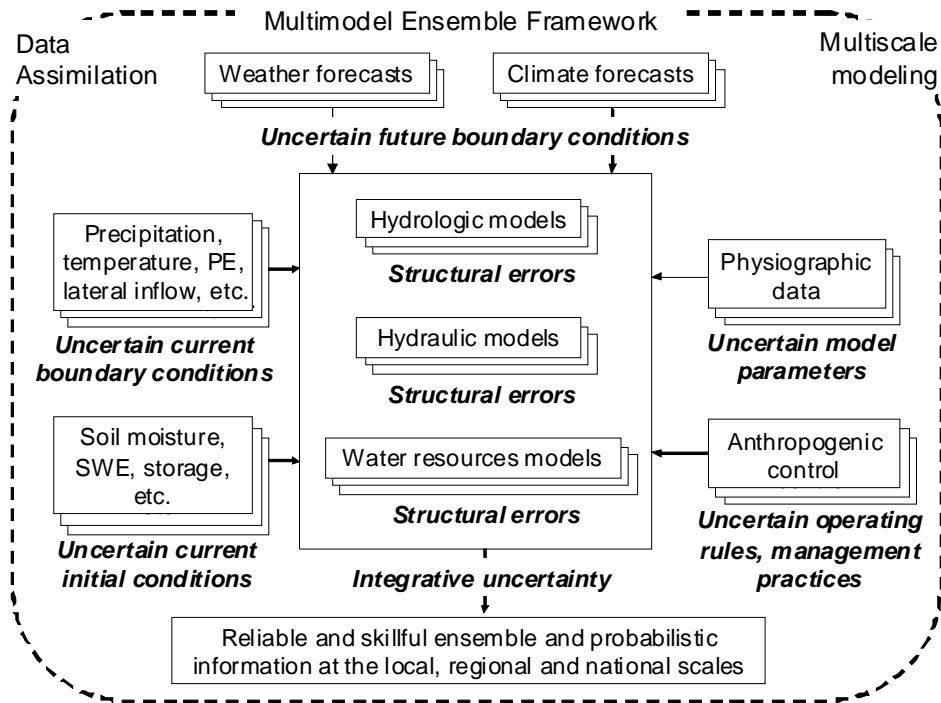


Figure 6-6 Multi-model Ensemble Framework

6.4 Challenges to getting there

For accurate and space-time-specific monitoring and prediction of water resources, hazards, and quality, comprehensive and integrated modeling of water flow, storage, and quality from hillslope to ocean is necessary. Such modeling should be comprehensive and multi-scaled to close the water budget from local to national scales, and integrated across all natural and man-made hydrologic, hydraulic, limnological, and estuarine processes and systems that impact availability, quality, supply, and demand of water. Modeling of such processes and systems should include all science elements of storage and flow as well as a number of other elements (see **Figure 1-1** above, and Figure 2 of IWSP, NWS 2004). The time scales associated with these processes and systems range from minutes to years and beyond. To integrate these diverse models with dynamical and statistical consistency over a wide range of scale, wide-ranging interdisciplinary science and systems expertise are required. Of particular challenge is to couple

the water resources models with the decision support systems under the ensemble paradigm for uncertainty-based prediction and decision-making. For these, closer and expanded partnerships and collaborations with the research and the user communities are essential.

The rationale for an uncoupled, rather than coupled, hydrologic/land surface model, as depicted in *Figure 6-6*, is based on the assessment that, within the strategic science planning horizon of this document, only the uncoupled framework is likely to provide the flexibility and modularity necessary to meet the NWS service goals. The advantages of an uncoupled hydrologic/land surface model include full utilization of the expanded ensemble forcing, broadening of the forcing sources, and easier correction of model biases, increase in model resolution, support to RFC operations, and development and implementation of multi-model ensemble methodologies.

6.5 A Road Map for Getting There

Modeling uncertainty is complex. It requires, in addition to the attendant science capabilities, substantial increase in computational and data storage and retrieval resources. Also, interpreting uncertainty takes training and experience, and communicating uncertainty effectively requires close interactions with the customers and users. Furthermore, while a suite of new ensemble forecasting capabilities are developed and tested, the existing single-value forecast system must operate and be improved to the extent that is cost-effective. As such, research and development and research-to-operations (RTO) transition of ensemble forecasting capabilities require careful planning that encompasses the end-to-end operational hydrologic forecasting process, collaboration and coordination with NCEP for atmospheric ensembles, and collaborations with the research and operational communities at large for cost-effective research and development, infusion and implementation of multi-model ensemble capabilities. This section describes the key activities that OHD is leading or engaged in toward meeting that goal.

6.5.1 The eXperimental Ensemble Forecast System (XEFS)

To hasten the pace through which an integrated system of short-, medium-, and long-range ensemble streamflow forecasting capability can be delivered to RFCs, Gary Carter, Director of OHD, formed and charged the Experimental Ensemble Forecast System (XEFS) Design and Gap Analysis Team in January 2007. *Figure 6-7* shows the basic design of the system and the 5 principal components (see NWS 2007 for details).

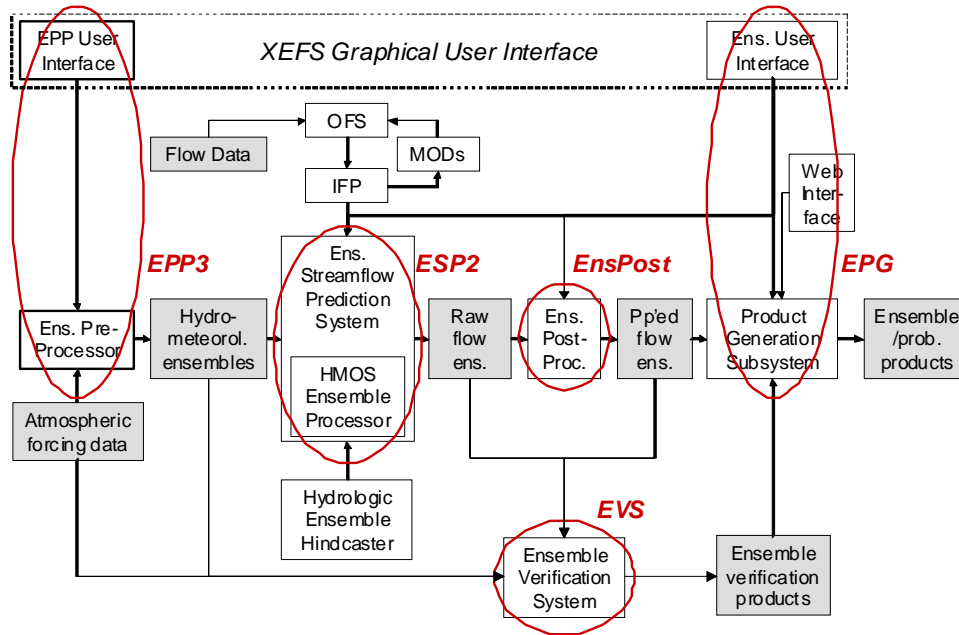


Figure 6-7

XEFS Architecture

The overriding science objective of XEFS is to produce reliable and skillful streamflow ensembles from 1 hour to 2 years into the future. As a prototype for the operational hydrologic ensemble forecast system, XEFS has a relative short development and implementation horizon of 2 to 3 years for its Phase 1 capabilities.

6.5.2 Hydrology Test Bed

The scientific and technological resources and infrastructure necessary for hydrologic ensemble forecasting is significantly more demanding and complex than single-value forecasting. To support efficient and cost-effective in-house and collaborative Research and Development (R&D) and RTO of ensemble forecasting capabilities, an integrated and unified end-to-end development platform and environment is necessary that serves the OHD, the RFCs, and the external collaborators. Such a capability supports not only development of prototype science algorithms but also prototyping of the operational forecast system envisioned in the CHPS paradigm (*Figure 6-8*).

Hydrology Test Bed

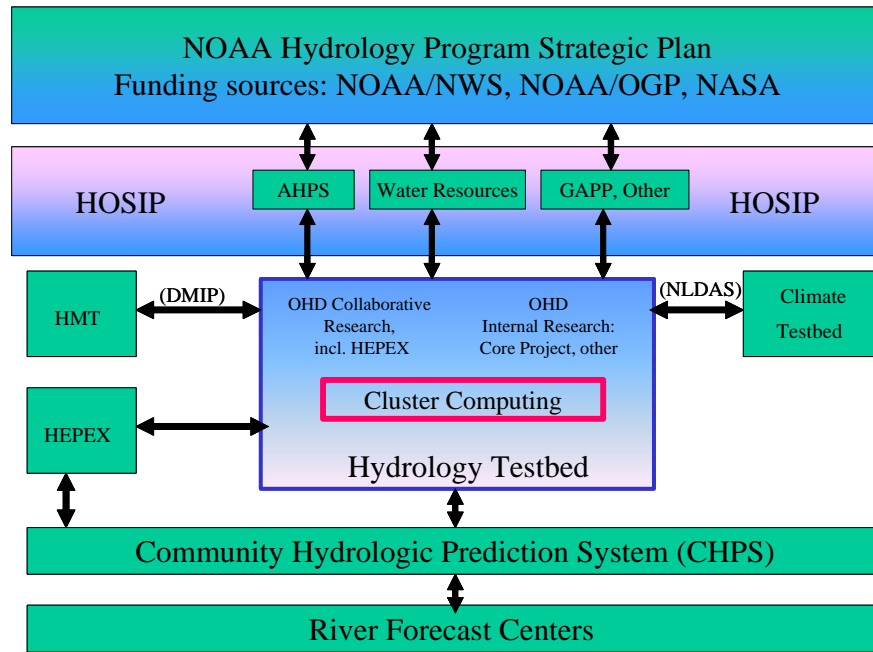


Figure 6-8

Hydrology Test Bed

6.5.3 Collaborations with NCEP

Reliable and skillful forcing ensembles are a necessary ingredient for reliable and skillful ensembles of hydrology and water resources variables such as streamflow and soil moisture. Also, to leverage advances in land surface modeling, including new and improved physics, use of new data sources (in particular, remotely-sensed) and multi-model ensembles toward the unified hydrologic/land-surface modeling paradigm of IWSP (NWS, 2004), close collaboration with NCEP is necessary, in particular with the Environmental Modeling Center (EMC), the Climate Prediction Center (CPC) and the Hydrometeorological Prediction Center (HPC).

Collaboration between NCEP/EMC and OHD, the THORPEX-Hydro Project seeks to develop capabilities to produce reliable and skillful hydrometeorological ensembles cost-effectively, to demonstrate the value of such ensemble forecasts for hydrology and water resources applications, and to expedite the delivery and operational use of the hydrometeorological ensemble products in hydrologic ensemble forecasting operations at the RFCs. The key science issues targeted include bias correction, downscaling, and hindcasting.

Through AHPS, HPC has developed the confidence interval product (Im *et al.*, 2006; Halquist, 2006). While it provides probabilistic guidance on their single-value products, further work is necessary to provide information usable by XEFS.

6.5.4 Climate Prediction Project for the Americas (CPPA) Core Project

To produce reliable and skillful ensembles of hydrology and water resources variables from an hour to about two years, seamless assimilation of climate forecasts and reduction and accounting of hydrologic uncertainties are critical. Through the CPPA Core Project, OHD collaborates with NCEP and leverages the climate research community to translate climate forecasts into water resources information that the NWS water customers can use for their decision-making. Figure 13 depicts this R&D and RTO transition framework.

Climate-to-Water RTO Pathways

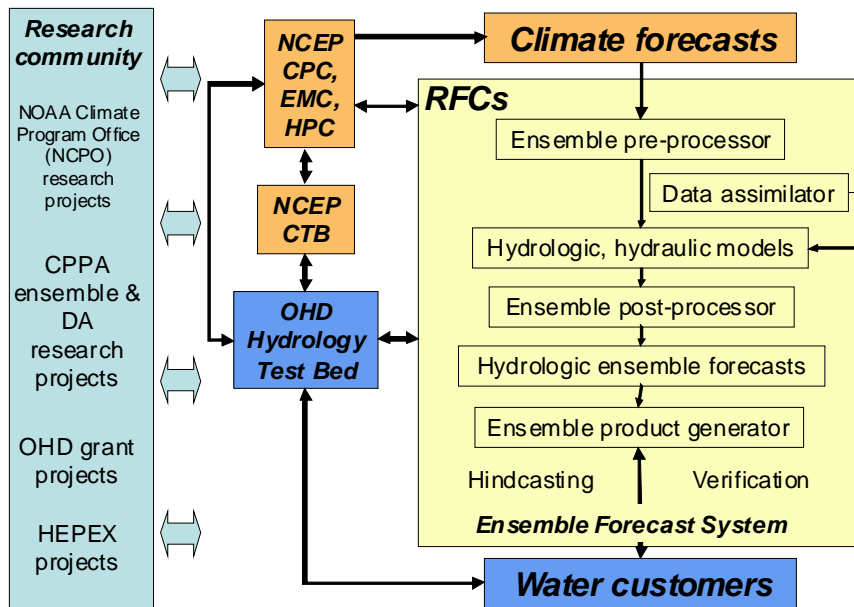


Figure 6-9 Climate-to-Water Research-to-Operations Pathways

6.5.5 HEPEX

The Hydrologic Ensemble Prediction EXperiment (HEPEX) is an international effort that brings together hydrological and meteorological communities from around the globe to build a research project focused on advancing probabilistic

hydrologic forecast techniques (Schaake *et al.* 2007). The HEPEx mission is to demonstrate how to produce reliable hydrological ensemble predictions that can be used with confidence by emergency management and water resources sectors to make decisions that have important consequences for economy, public health and safety (from <http://hydis8.eng.uci.edu/hepex/>). OHD will continue to leverage the global expertise on -and experience with- ensemble prediction through HEPEx and other collaborative efforts.

6.5.6 DMIP

Analysis from DMIP-1 showed that multi-model ensemble streamflow simulation improves skill and reduces bias, and may reduce the effort necessary for calibration necessary to attain the level of skill obtainable from a single model (Georgakakos *et al.*, 2004). OHD will continue to leverage DMIP to assess the value of multi-model ensembles under forecast and/or real-time updating scenarios, including those of soil moisture and runoff.

6.6 References

- Buizza, R., P.L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, Y. Zhu. (2005). “A comparison of the ECMWF, MSC and NCEP Global ensemble prediction systems.” *Mon. Wea. Rev.* 133, 1076-1097.
- Clark, M., S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby. (2004). “The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields.” *Journal of Hydrometeorology*, 243-262, 2004.
- Day, G. N. (1985). “Extended streamflow forecasting using NWSRF.” *J. Water Resour. Plann. Manage.*,111(2), 157–170.
- Demargne, J., L. Wu, D.-J. Seo, and J. Schaake. (2007). “Experimental hydrometeorological and hydrological ensemble forecasts and their verification in the U. S. National Weather Service in Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management.” *Proceedings of Symposium HS2004 at IUGG2007*, Perugia, IAHS Publ. 313, 2007, p.177-187.
- Gangopadhyay, S., M. Clark, K. Werner, D. Brandon, and B. Rajagopalan. (2004). “Effects of Spatial and Temporal Aggregation on the Accuracy of Statistically Downscaled Precipitation Estimates in the Upper Colorado River Basin.” *Journal of Hydrometeorology*, 5, 1192-1206.

- Georgakakos, K., D.-J. Seo, H. Gupta, J. Schaake, and M. B. Butts. (2004). "Towards the characterization of streamflow simulation uncertainty through multimodel ensembles." *J. Hydrol.* 298, 222-241, 2004.
- Georgakakos, K. P., N. E. Graham, T. M. Carpenter, E. Shamir, J. Wang, J. A. Sperflage, and S. V. Taylor, A. P. Georgakakos, H. Yao and M. Kistenmacher. (2006). "Integrated Forecast And Reservoir Management (Inform) For Northern California: System Development And Initial Demonstration." *HRC Technical Report No. 5*.
- Halquist, J. (2006). "Use of HPC QPF confidence interval forecasts to produce a hydrologic ensemble of river forecasts." National Weather Service, Chanhassen, Minnesota, AMS.
- Im, J.-S., K. Brill, and E. Danaher. (2006). "Confidence Interval Estimation for Quantitative Precipitation Forecasts (QPF) Using Short-Range Ensemble Forecasts (SREF)." *Weather and Forecasting*, 24-41.
- Luo, L., and E. F. Wood. (2006). "A seasonal hydrologic ensemble forecast system over the eastern U. S." *1st CPPA PI Workshop*, Tucson, AZ.
- McEnery, J., J. Ingram, Q. Duan, T. Adams, and L. Anderson. (2005). "NOAA'S Advanced Hydrologic Prediction Service, Building Pathways for Better Science in Water Forecasting." *BAMS*, March, p. 375-385.
- Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, D. P. Lettenmaier, C. H. Marshall, J. K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A. A. Bailey. (2004). "The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system." *J. Geophys. Res.*, 109, D07S90, doi:10.1029/2003JD003823.
- NWS, the Integrated Water Science Plan, 2004.
- NWS, the eXperimental Ensemble Forecast System design and gap analysis Report, 2007.
- Schaake, J., T. M. Hamill, and R. Buizza. (2007). "HEPEX, the Hydrological Ensemble Prediction Experiment." Accepted for publication in *Bull. Am. Meteor. Soc.*
- Schaake, J., J. Demargne, R. Hartman, M. Mullusky, E. Welles, L. Wu, H. Herr, X. Fan, and D. J. Seo. (2007). "Precipitation and temperature ensemble

forecasts from single-value forecasts.” *Hydrol. Earth Syst. Sci. Discuss.*, 4, 655-717.

Seo, D.-J., H. D. Herr, and J. C. Schaake. (2006). “A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction.” *Hydrol. Earth Syst. Sci. Discuss.*, 3, 1987-2035.

Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay. (2005). “Incorporating Medium-Range Numerical Weather Model Output into the Ensemble Streamflow Prediction System of the National Weather Service.” *Journal of Hydrometeorology*, 6, 101-114.

Werner, K., D. Brandon, M. Clark, and S. Gangopadhyay. (2004). “Climate Index Weighting Schemes for NWS ESP-Based Seasonal Volume Forecasts.” *Journal of Hydrometeorology*, 5, 1076-1090.

Wood, A., and D. Lettenmaier. (2005). “A testbed for new seasonal forecasting approaches in the western U. S.” Submitted to *Bull. Am. Meteor. Soc.*

7. Data Assimilation

Hydrologic forecasts are subject to uncertainties in the initial conditions (soil moisture, snowpack, channel storage, etc.), observed boundary conditions (observed precipitation, observed temperature, etc.), future boundary conditions (future precipitation, future temperature, etc.), and pedologic and physiographic boundary conditions (soil properties, basin geomorphology, channel geometry, vegetation, etc.). Cost-effective assimilation of all available informative data sources is essential to reducing these uncertainties, and hence to improving and increasing the skill and lead-time of hydrologic and water-resource forecasts.

7.1 Where We Are

Operational hydrologic data assimilation has been dominated by manual techniques (known as run-time modifications, or MODs) that are—while very effective in the hands of an experienced forecaster—generally labor-intensive and often subjective. As the space-time scale of modeling gets finer, and more new (and presumably informative) data sources become available, the sheer volume of data to be assimilated into the forecast process is increasing very rapidly.

Automatic hydrologic data assimilation, often referred to as state updating, is not new in hydrology (Kitanidis and Bras, 1980; Day, 1990). While much attention has been paid in the last quarter century to the topic, the operational reality in NWS is that automatic hydrologic data assimilation is yet to be recognized as an essential element in the forecast process and has not been adequately exploited as a complement to the forecaster MODs. Currently, some form of automatic data assimilation (DA) techniques is used at CNRFC (SS-SAC; Sperflage and Georgakakos, 1996) and WGRFC (VAR; Seo *et al.*, 2007) for updating of soil moisture states, and at NOHRSC (SNODAS; Carroll *et al.*, 2001) and NWRFC (NWSRFS Snow Updating System, 2003) for updating of snow states.

7.2 What Our Partners Are Doing

- NOHRSC – snow data assimilation (Carroll *et al.*, 2001). The National Operational Hydrologic Remote Sensing Center (NOHRSC) ingests daily ground-based, airborne, and satellite snow observations from all available electronic sources for the coterminous U. S. These data are used along with estimates of snowpack characteristics generated by a physically based snow model to generate the operational, daily NOAA National Snow Analyses (NSA) for the coterminous U. S.
- NCEP/EMC – NLDAS (Mitchell *et al.*, 2003). While currently NLDAS is a land-surface modeling system, ultimately it will employ data assimilation techniques to constrain model predictions with observations of LDAS storages (soil moisture, temperature, snow) and fluxes (evaporation, sensible heat flux, runoff). Many of the capabilities necessary for assimilating data

into the participating models have already been developed or under development through the Land Information System (LIS: see Chapter 8). The recent NASA public release of LIS Version 5.0 formally includes an Ensemble Kalman Filter (EnKF) capability. NCEP/EMC plans to explore the assimilation of satellite derived snow cover, SWE, soil moisture and vegetation density into the Noah LSM of NLDAS using the new EnKF capability of LIS. EMC has discarded its pursuit of variational land data assimilation methods via adjoint and tangent linear models, because such models were found to be very difficult to derive and update for largely physically-based LSMs such as the Noah LSM.

- NASA - The Global Land Data Assimilation System (Rodell *et al.* 2004), updating a land surface model with MODIS-derived snow cover (Rodell and Houser 2004), the Land Information System (Kumar *et al.* 2006). The goal of the Global Land Data Assimilation System (GLDAS) is to ingest satellite- and ground-based observational data products using advanced land surface modeling and data assimilation techniques to generate optimal fields of land surface states and fluxes (Rodell *et al.*, 2004). Data assimilation techniques for incorporating satellite based hydrological products, including snow cover and water equivalent, soil moisture, surface temperature, and leaf area index, are now being implemented.
- CIRES - Snow Data Assimilation via an Ensemble Kalman Filter (Slater and Clark, 2005), Assimilation of snow covered area information into hydrologic and land-surface models (Clark *et al.* 2006). The aim is to improve the model's (SNOW-17) estimate of SWE by merging the uncertainties associated with meteorological forcing data and SWE observations within the model. An ensemble square root Kalman filter is applied to perform assimilation on a 5-day cycle. Once the temporal persistence inherent in a snowpack is removed from both model and assimilated observations during the update cycle, a result is produced that is consistently superior to either the model or interpolated observations within the limits of available information.
- University of Arizona –Uncertainty in Hydrologic Modeling: Towards an Integrated Data Assimilation Framework (Liu and Gupta 2007). The key to properly addressing hydrologic uncertainty is to understand, quantify, and reduce uncertainty involved in hydrologic modeling in a cohesive, systematic manner. They propose and are developing an integrated hierarchical framework for pursuing hydrologic data assimilation in several progressive steps to maximally reduce uncertainty in hydrologic predictions.

7.3 Where We Want to Be

As the spectrum of hydrology and water resources products that are necessary to meet the customer needs increases and the breadth of models and data that are necessary to support generation of such products increases, it is increasingly clear that a shift is necessary toward a paradigm that fully capitalizes on advances in

computing power and availability of new data sources (remote sensing in particular). The new paradigm must reduce the burden of manual DA on the part of the forecasters, particularly in time-critical situations, while fully recognizing the need for and utilizing the value of forecaster control, intervention and override of automatic DA results in the forecast process.

A comprehensive data assimilation capability, that is a companion to the multi-scale hydrology and water resources modeling system (see **Figure 6-5** and **Figure 6-6**) and that supports ensemble prediction, is necessary to fully utilize all available *in-situ* and remotely sensed hydrometeorological and hydrologic data, and to exploit advances in hydrologic land surface models and NWP. Such a system would optimally blend multisensor data with model states to produce informative and accurate hydrology and water resources products that are dynamically and statistically consistent from local, regional, to national scales.

The types of observations that would be assimilated include: a) hydrologic observations, such as hydrologic states of the hydrology and water resources models (soil moisture and temperature, snowpack, reservoir/lake storage and others) and observations of output of the models (streamflow, discharge from reservoir and others); and b) hydrometeorological observations of the forcing variables (precipitation, air temperature, solar insolation and others). Data assimilation techniques may also be used to better-utilize *in-situ* and remotely sensed physiographic and phenological data, particularly for estimation of distributed parameters. Capability for routine objective assimilation of large amounts of different types of data is also essential to routine objective assessment of marginal value of new and improved observational capabilities.

7.4 Challenges to Getting There

Comprehensive data assimilation necessary to produce informative and accurate hydrology and water resources products that are dynamically and statistically consistent across all physical elements, hydrology and water resources models and space-time scales is a major challenge. More basically, getting the observations to support data assimilation on a timely basis is also another major challenge. Water-focused, integrated, and interdisciplinary data assimilation research and development among the NWS entities, and existing and new research partners is essential to meeting this challenge. For example, present space-based microwave estimates of soil moisture sense only the top several centimeters of the soil column, far short of the deeper depths needed for land-state initialization. The data assimilation system would blend sparse land observations with the background fields of the hydrologic model to produce the optimally estimated initial state.

The history of automatic state updating in hydrology offers valuable lessons in meeting these challenges. First, it is important to recognize that DA problems in

operational hydrology are often quite different in nature from those, e.g., in operational meteorology due to very large degrees of freedom in the system modeled and high nonlinearity of the processes involved. Close communications, and interactions and collaborations between the research and the operational communities are needed to better formulate the hydrologic DA problems, to leverage more effectively advances in DA methodology and to develop solution techniques that are viable in operational hydrologic forecasting.

7.5 A Road Map for Getting There

Operational hydrologic forecasting involves a large number of models and data sets, and physical processes that operate over a wide range of time scales. While the ultimate goal should be an integrated DA system that encompasses all models and data sets, it is likely that development of such a system is too large and complex to yield, at the current level of understanding, cost-effective solutions for operational hydrologic forecasting. A preferred strategy is to reduce the size of the DA problem by decomposing the very large problem into smaller ones (see *Figure 7-1*) and to phase the development and infusion of DA capabilities in such a way that they may support deterministic prediction but can be extended to ensemble prediction.

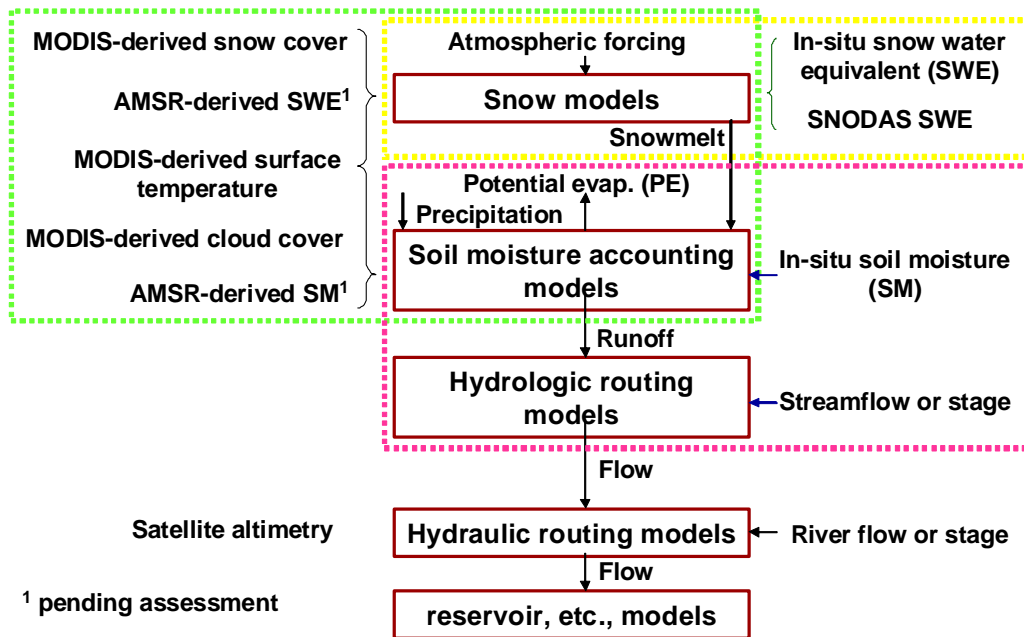


Figure 7-1 Data Assimilation Strategy

7.5.1 Hydrology and Water Resources Data Assimilation Projects (AHPS, CPPA, Water Resources, Hurricane Supplemental)

OHD has been carrying out R&D and RTO of hydrologic DA capabilities for lumped and distributed hydrologic models. The prototype capabilities developed thus far include a 2DVAR technique for assimilation of streamflow, precipitation and PE into lumped SAC and UHG models at 1-hour timestep (Seo *et al.* 2007), and a 4DVAR technique for assimilation of streamflow, *in-situ* soil moisture, gridded precipitation and PE into distributed SAC and kinematic-wave routing models at 1-hour timestep (Seo *et al.* 2003). The 2DVAR technique has been implemented at WGRFC for experimental operation (Seo *et al.* 2007) and in the Site-Specific Hydrologic Prediction (SSHP) System for operation in the Gulf States. Enhancements are needed to these prototypes for routine operational implementation and to include updating of routing models.

7.5.2 Hydrology Test Bed

The next big step in operational hydrologic DA is the development and implementation of ensemble DA capabilities. While the computational requirements for ensemble DA for lumped models are relative modest, those for distributed models are not. To develop and implement operation-worthy ensemble DA capabilities for distributed models, a significant increase in computational capability is necessary. The Hydrology Test Bed will provide a development platform for an integrated end-to-end ensemble forecast system in which ensemble DA is an integral part.

7.5.3 Collaborative Projects

NASA-NWS/OHD Project

Satellite data and derived products offer spatially continuous information that may potentially reduce the uncertainties in the initial and boundary conditions. As with all remotely-sensed observations, satellite data are measurements of radiometric intensity at different wavelengths that may be related to the hydrologic variables of interest with varying degree of uncertainty. This project assesses the value of NASA satellite data-derived products, including MODIS cloud cover and MODIS-derived snow cover (Dong and Peters-Lidard 2007), to the NWS operational models.

CPPA External Project

This project develops and evaluates methods for producing high-resolution ensemble atmospheric forcing data sets for distributed hydrologic and land-surface models, and evaluates the relative importance of uncertainties in model inputs for modeling streamflow in the western U. S., develops and evaluates methods for obtaining error estimates in SWE data, such that they can be used for assimilation

purposes and infuses new scientific advances/methods developed from this study in future versions of NWSRFS.

7.6 References

- Carroll, T., D. Cline, G. Fall, A. Nilsson, L. Li, and A. Rost. (2001). "NOHRSC operations and the simulation of snow cover properties for the coterminous U. S." *Proc. 69th Annual Western Snow Conference*, Sun Valley, ID, Western Snow Conference, 1–14.
- Clark, M. P., A. G. Slater, A. P. Barrett, L. E. Hay, G. J. McCabe, B. Rajagopalan, G. H. Leavesley. (2006). "Assimilation of snow covered area formation into hydrologic and land-surface models." *Advances in Water Resources* 29, 109–1221.
- Day, G. N. (1990). "A methodology for updating a conceptual snow model with snow measurements." *NOAA Tech. Rep. NWS 43*, 133 pp.
- Dong, J., and C. Peters-Lidard. (2007). "On the Relationship between Temperature and MODIS Snow Cover Retrieval Errors." Submitted to *Geophysical Research Letters*.
- Rodell, M., and P. R. Houser. (2004). "Updating a land surface model with MODIS-derived snow cover." *J. Hydrometeor.*, 5, 1064–1075.
- Kitanidis, P. K., and R. L. Bras. (1980). "Real-time forecasting with a conceptual hydrologic model. I. Analysis of Uncertainty." *Water Resour. Res.*, 16(6), 1025-1033.
- Kumar, S. V., C. D. Peters-Lidard, Y. Tian, P. R. Houser, J. Geiger, S. Olden, L. Lighty, J. L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E. F. Wood and J. Sheffield. (2006). "Land Information System - An Interoperable Framework for High Resolution Land Surface Modeling." *Environmental Modelling & Software*, Vol. 21, pages 1402-1415.
- Mitchell, K. E., D. Lohmann, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R. T. Pinker, J. D. Tarpley, D. P. Lettenmaier, C. H. Marshall, J. K. Entin, M. Pan, W. Shi, V. Koren, J. Meng, B. H. Ramsay, and A. A. Bailey. (2004). "The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system." *J. Geophys. Res.*, 109, D07S90, doi:10.1029/2003JD003823.

- Liu, Y, and H. V. Gupta. (2007). "Uncertainty in Hydrologic Modeling: Towards an Integrated Data Assimilation Framework." *Water Resour. Res.*, 43, W07401, doi:10.1029/2006WR005756.
- RTi, NWSRFS Snow Updating System Version 2.00.02 Snow Updating User Manual, 2003.
- Rodell, M., and Coauthors. (2004). "The Global Land Data Assimilation System." *Bull. Amer. Meteor. Soc.*, 85, 381–394.
- Seo, D. J., V. Koren, and N. Cajina. (2003). "Real-time variational assimilation of 18 hydrologic and hydrometeorological data into operational hydrologic forecasting." *J. Hydrometeorol.*, 4(3), 627-641.
- Seo, D.-J., V. Koren, and L. Cajina. (2003). "Real-time assimilation of radar-based precipitation data and streamflow observations into a distributed hydrologic model." *IAHS Publications Series (Red Books)* 282, 109-113.
- Seo, D.-J., L. Cajina, R. Corby, and T. Howieson. (2007). "Automatic State Updating for Operational Streamflow Forecasting via Variational Data Assimilation." Submitted to *J. of Hydrol.*
- Slater, A. G., and M. P. Clark. (2006). "Snow Data Assimilation via an Ensemble Kalman Filter." *Journal of Hydrometeorology*, 7 478-493.
- Slater A.G., and Clark, M.P. (2007). "Uncertainty and Complexity in a Snow Hydrology Models, Session HW2004, Towards Improved Evaluation of Hydrological Models: The Need to Understand and Characterize Uncertainties in the Modelling Process." *IUGG*, Perugia, Italy.
- Sperflslage, J. A., and K. P. Georgakakos. (1996). "Implementation and testing of the HFS operation as part of the National Weather Service River Forecast System (NWSRFS)." *HRC Tech. Rep. 1*, Hydrologic Research Center, San Diego, CA, 213 pp.

8. Verification

The inability to systematically and routinely verify hydrologic forecasts and communicate effectively verification information to all users is one of the long-standing deficiencies in the NWS Hydrologic Services Program. More generally, little verification of hydrologic forecasts has been conducted to date (Welles *et al.*, 2007). Operational forecasts produced by the RFCs and experimental forecast developed by the NWS have been verified by various limited verification applications. Generally verification information is not communicated to most user groups. However, verification helps us answer the following key questions:

- What forecast products are the NWS providing and where?
- How good are the forecasts?
- What are the sources of uncertainty in the forecasts?
- What are the strengths and weaknesses in the forecasts?
- How is new science improving the forecasts?
- What should be done to improve the forecasts?

The need for a comprehensive hydrologic forecast verification system for the NWS has been outlined by the National Research Council in 1996, which stated that verification of hydrologic forecasts was inadequate. In 2005, the Department of Commerce, Office of Inspector General, published an Inspection Report entitled “The Northeast River Forecast Center Is Well Managed, But Some Improvements Are Needed.” The report recommended that the NWS develop, document, and implement a timeline and action plan for completing a comprehensive river forecast verification system. Additionally, verification was emphasized in 2006 by the National Research Council (see *Appendix A*). To address the need for verification, the NWS created an advisory team, the Verification System Requirements team, to develop requirements for a comprehensive hydrologic verification system and propose a verification plan. The findings and recommendations of the team are summarized in the final report published in October 2006. (http://www.nws.noaa.gov/oh/rfcdev/docs/Final_Verification_Report.pdf).

Regarding the system requirements, the verification component in the river forecasting system should be a comprehensive national system to evaluate the quality of delivered forecast services and the quality of all the hydrologic forecasts and guidance products (inputs and outputs) which satisfy the needs of all users. The forecasts to be verified include deterministic hydrologic forecasts, ensemble forecasts, statistical water supply forecasts, and gridded forecasts, with a wide range of lead times from minutes (e.g. for flash flood purposes) to years (e.g. for water supply). The system would improve forecast services by analyzing the sources of uncertainty and skill across the entire river forecasting system. It

should also provide easy access to forecast verification data to improve our scientific and operational techniques and services.

The hydrologic verification system supports:

- Scientists/researchers and hydrologic program managers, by identifying needs to improve the forecasting system and clarifying the value of products and results from current and new science
- Hydrologic forecasters, by defining acceptable methods to generate forecasts and products and satisfying user demands
- Emergency and water resources managers, and the general public, by quantifying forecast performance and uncertainty for better decision making

The goals of the River Forecast Verification System are to:

- Quantify the quality of the river forecasts and the quality of the forecast services
- Monitor the forecast quality over time
- Monitor the quality at various steps during the forecasting process to pinpoint the different sources of uncertainty and skill
- Identify the best ways to improve the forecast quality

To analyze the different uncertainty sources, the verification system needs to assess the different components of the forecasting system: the model set-up component, the state updating component, the forecast computation, and the product review and issuance component.

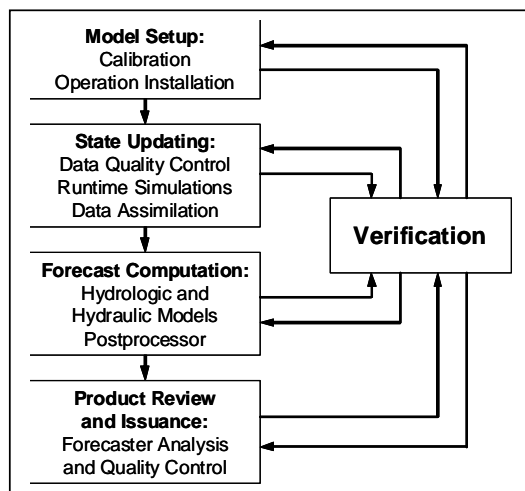


Figure 8-1

Setup of the verification system to evaluate the processing steps of the forecasting system (final team report, 2006)

There are two main components of the River Forecast Verification System:

- Forecast services verification (or logistical verification) component, to evaluate the quality of delivered forecast services in terms of the usability of the forecasts (number of forecasts locations, new type of forecasts, effort to issue forecast, forecast timeliness, etc.); and
- The forecast verification component, to quantify the quality of forecasts, which includes deterministic and probabilistic verification (for ensemble forecasts and water supply forecasts) on different space-time domains (for example for point/area forecasts or gridded forecasts). This component needs to include diagnostic and prognostic verification. Diagnostic verification evaluates the quality of past forecasts given certain conditions (time, variable value, event, methodology, etc.). Prognostic verification evaluates the quality of live forecasts in real-time (before the observation occurs) using analog forecasts from the past to evaluate potential performance of these live forecasts.

In order to achieve the goals stated above, the River Forecast Verification System must provide the following specific capabilities:

1. Data archiving for forecasts and associated observations, as well as attributes (relative to time, service, basin, events...);
2. Computing verification metrics to evaluate the different aspects of forecast quality (i.e., accuracy, bias, association, skill, reliability, resolution, discrimination, sharpness, uncertainty), which requires a functionality to aggregate and stratify the forecast samples according to the forecast values, observation values, or other attributes;
3. Displaying verification data and metrics with graphics, numerical results, and reports to examine the metrics, which should include a functionality to screen large volumes of verification results and find the meaningful information;
4. Disseminating verification data and metrics along with training material and documentation of verification results, to understand the quality and usefulness of the delivered forecasts;
5. Real-time access to verification metrics to understand uncertainties in recent forecasts and over the long-term;
6. Uncertainty analysis to identify the strengths and weaknesses of the forecasts, which requires the use of multiple forecast scenarios, including hind-casting experiments (to produce large sample of forecasts to be verified), and the analysis of both input and output; and
7. Performance measure tracking to evaluate the level of success of river forecasting.

8.1 Where We Are

The Hydrology Program has begun operational implementation of verification capabilities and will phase in the comprehensive river forecast verification system over the next five years.

For logistical verification, a new interactive application called Forecast Services Manager enables the RFCs to define point forecast services in the Integrated Hydrologic Forecast System database (IHFS-DB).

Regarding the archiving capability, the arrival of the RFC Archive machines (RAX) in 2004 has enabled all the RFCs to store in a common and standardized database the observations and deterministic forecasts for forcing inputs and hydrologic outputs for forecast points (capability 1). Information from the archive database is then extracted to run an operational deterministic verification capability, the Interactive Verification Program (IVP), which verifies deterministic forecasts of precipitation, temperature, streamflow, and stage at forecast points (capabilities 2 and 3).

For probabilistic forecasts, two capabilities have been developed within the XEFS project: the Ensemble Verification System (EVS) prototype, which verifies ensemble forecasts for forcing inputs and hydrologic outputs (capabilities 2 and 3) at forecast points; its supporting application, the Hydrologic Ensemble Hindcaster prototype, which retroactively applies existing and new methodologies to generate hydrologic ensemble hindcasts/reforecasts with large enough samples, for uncertainty tracking purposes (capability 6). These two prototypes are currently used to analyze the performance of precipitation, temperature, and streamflow ensemble forecasts generated with different methodologies. Demaragne *et al.* (2007) compared the performance of streamflow ensembles generated from climatological input ensembles (similarly to the operational ESP) and QPF-based EPP2 ensembles and showed the improvement of using QPF-based precipitation ensembles and the need to account and reduce hydrological uncertainty.

For water supply forecasts, the Western Region has developed a website (www.cbrfc.noaa.gov/westernwater) to archive and provide water supply forecasts for six RFCs within the western U. S., as well as information on forecast verification. This application has capabilities for data visualization and computation of error metrics, skill metrics, and categorical metrics. It also provides access to all forecast and verification data, as well as analysis of various climate change scenarios.

8.2 What Our Partners Are Doing

- AHPS Verification System developed by the University of Iowa at http://www.ihr.uiowa.edu/ahps_ver (capabilities 6, 2 and 3): web-based tools for online access, analysis, and comparison of retrospective long-term ensemble forecasts within the operational setting of the RFC; interactive exploration of verification results; instant access to forecasts and quality measures for verification of forecast sites. Additionally papers were published regarding distribution-oriented verification metrics, which were used to evaluate long-term streamflow ensemble forecasts (Bradley *et al.* 2004 and Hashino *et al.* 2007);
- Development of an ensemble verification application to evaluate NWS streamflow ensemble forecasts by the University of California-Irvine (Franz *et al.* 2003);
- Grid forecast verification at NCEP (for example for QPF grid verification with the National Precipitation Verification Unit, <http://www.hpc.ncep.noaa.gov/npvu/>), ESRL (Real Time Verification System, <http://rtvs.noaa.gov/>), and the NCAR Research Applications Laboratory (MODE object-based verification tool, spatial verification method inter-comparison project, and Model Evaluation Tools project, <http://www.rap.ucar.edu/research/verification/>); and
- WMO Joint Working Group on Verification (JWGV), an international group of scientists who supports the development and testing of new methods, offers web resources (discussion group via email, reference website http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.shtml), and organizes verification workshops (including tutorials).

8.3 Where We Want to Be

8.3.1 Forecast Services Verification

A comprehensive description of service efficiency and forecast usability includes the following logistical verification measures:

- Characterizing point forecasts by service type, frequency, and location;
- Characterizing areal forecasts by service type, frequency, and location;
- Identifying daily the number of issued forecasts by type and location;
- Quantifying the person effort required to set up a basin for forecasting, including data gathering, calibration, model setup and implementation efforts;
- Quantifying the person effort required to issue each type of forecast, including manual quality control of input data, forecaster run-time modifications and forecaster review and analysis; and
- Quantifying the timeliness of issued forecasts.

The logistical verification measures will improve the management of the hydrology science and service programs.

The logistical verification component of the River Forecast Verification System needs to include the six capabilities described earlier, from archiving to dissemination (capabilities 1-5), and performance measure tracking (capability 7).

8.3.2 Forecast Verification

The forecast verification component of the verification system needs to assess the quality of all the different forecasts, which could be either deterministic or probabilistic (including ensemble and statistical), and which could be relative to different space and time domains (including point forecasts and gridded forecasts).

Data Archiving: the verification system needs to include an archiving capability for standardized archive datasets in a common format. This capability will systematically archive all forecasts and observations used/generated by the forecasting system, attributes of the forecasts/observations, as well as information on the forecaster inputs (such as MODs), to study the impact of different processes (such as one or a combination of MODs) on the forecast quality. Since uncertainties from the rating curves need to be accounted for, it is useful to archive forecasts both for streamflow and stage, as well as the rating curves (which may change with time). Additionally, the system should also capture any modification of the forecasting process (e.g., model parameters, segment definition, station/area definition). A visualization tool with a data quality control capability needs to be developed to analyze the forecast and observations data, and detect and potentially eliminate incorrect input data.

Computing Verification Metrics: the verification system needs to use a variety of verification metrics to capture all the different aspects of forecast performance and meet the multiple needs of the users. The Verification System Requirements team has recommended various metrics for both deterministic and probabilistic forecasts, which are defined from seven categories: categorical, error, correlation, distribution, skill score, conditional, and statistical significance (see final report). For metrics based on threshold values, the system should include a capability to offer guidance to the user to select meaningful threshold values given the forecasts to be verified. Forecast verification should also be done on different space and time domains. For example, uncertainties for streamflow forecasts could be defined for specific time step and lead time (e.g., the 6-hour streamflow forecasts for lead days 1 to 5); uncertainties could also be defined by analyzing the streamflow time series to define different types of errors (timing error, shape error, peak value error, etc.) for specific events.

Additionally, the system should help the user to determine how to pool and stratify the forecasts to be verified. Verification requires a trade-off between a large

sample to compute reliable verification metrics and stationarity of the sample, which is assumed at some level during verification. A capability should offer guidance on which forecast samples to verify for results that would be robust (from large sample size) and meaningful (from quasi-homogeneous subsets).

Also, verification should allow the user to compare performance of single-value forecasts and probabilistic forecasts. Since a direct comparison is difficult, probabilistic forecasts are generally converted into single-value forecasts (using ensemble median for example), which leads to information loss. A better approach is to convert single-value forecasts into probabilistic forecasts by estimating the uncertainties in the single-value forecasts.

Uncertainty Analysis: the uncertainty sources are mainly the input data (observations, forecasts and outlooks, rating curves, reservoir outflows and releases, etc.), the hydrologic and hydraulic models (model parameters, model states, and model structure), and the forecaster analysis. Although each step in the forecast development process is assumed to improve the performance of the final forecast product, the individual contributions of the input data, the forecast models, and the forecaster need to be evaluated. The comparison of the forecast system performance with and without a specific process increases understanding of the relative impact of that process on the forecast quality. For example, when evaluating the impact of the data assimilation process, the scientist needs to evaluate the streamflow forecasts with and without the data assimilation process. Also, not only observed values could be used as reference to evaluate the forecasts. For example, when evaluating the impact of the input forecasts on streamflow forecasts, the reference forecasts could be the streamflow generated from the perfect inputs (i.e. simulated flows). By comparing streamflow forecasts with simulated flows as well as observed flows, it is possible to separate the uncertainty from the forcing inputs and the uncertainty from other sources.

Besides, the verification system needs to include the capability to hindcast/re-forecast all the forecast data and time series required for the uncertainty analysis work to apply the current state of the science retroactively. It should produce hydrologic hindcasts, forcing input hindcasts, and retrospective model states over a time period up to multiple years. It should also provide real-time access to the available hindcast archive and robust metadata to fully describe each hindcasting scenario. This hindcasting capability is crucial since forecast verification requires a sample size large enough to compute robust metrics. The hindcasts to be generated for a given forecasting scenario would reflect a single forecasting system, with no changes relative to the models. A capability for routine, systematic, and rigorous hindcasting is necessary to assess and validate any new forecasting methodology.

Regarding the forecaster inputs, it is extremely important to evaluate the impact of runtime modifications, MODs, on the forecast quality given the number of

MODs used by the forecaster and the effort put into this operation. This includes the definition of a raw model (or baseline model) to generate hydrologic forecasts by running only those parts that are automated, requiring no interaction with the forecaster during operational use. Besides, other reference forecasts (climatology, persistence, etc.) need to be used in the verification studies to evaluate the benefits of using forecasts produced by the forecasting system under evaluation vs. forecasts from other sources. Reference forecasts (to compute skill scores for example) are also useful to evaluate whether the forecasts perform better because the events are more easily predictable, or because of the “smarts” of the forecast system itself.

Diagnostic and Prognostic Verification: the verification system should include both diagnostic and prognostic verification. Diagnostic verification evaluates the quality of past forecasts given certain conditions (time, variable value, event, methodology, etc.). Prognostic verification evaluates the quality of live forecasts in real-time (before the observation occurs) using similar (analog) forecasts from the past to evaluate potential performance of live forecasts. It could lead to optimized forecasts by correcting some of the expected forecast errors as detected in the prognostic verification results. Also verification of analogs (past forecasts for similar hydrologic events) from different models could be used to evaluate what the forecast performance is for each individual model; then live forecasts could be improved by using/merging forecasts from the different models according to the past models’ performance. This prognostic capability needs to be built on top of the operational forecasting system. It will require a process to determine from the forecast archive analog forecasts, which could be based on multiple criteria.

Besides, since the verification process will produce a huge amount of data, the verification system should include a functionality to screen large volumes of verification results and find the meaningful information. Methods such as data mining or artificial intelligence could be developed to help the user analyze verification metrics results for various forecasting situations.

Communicating Results: the verification system needs to archive verification data and results and to effectively communicate the information to the various users. A graphical capability will display verification results for the metrics in both run-time mode and inter-comparison mode for the uncertainty analysis work and according to various characteristics (e.g., lead time, verification time window, spatial location, type of variable). The system should be flexible to accommodate various methods of dissemination of forecast/guidance products and verification results to different users. This information should be provided with various degrees of sophistication; experienced users could do their own verification analysis to answer their specific questions; more basic users without a high knowledge of statistics need to access verification information expressed in “common language”. Additionally the system should include comprehensive documentation (including verification case studies) about the interpretation and

meaning of the verification metrics and the methods used to develop and analyze the verification results.

8.4 Challenges to getting there

The forecast products and verification information should be provided to end users for best decision making, as well as scientists and forecasters for collaboration on forecast verification and forecast improvement. This could be done via a community verification system and a community archiving system. These community components would be developed within CHPS.

Archiving all data and verification information is challenging, given the number of forecast points and areas, forecast types, as well as the number of processes and methodologies that are available to produce input and output forecasts and that need to be evaluated. This is even more challenging for probabilistic forecasts than deterministic forecasts given the huge volume of data when storing probabilistic forecasts. Different options based on file systems or databases will need to be evaluated, using examples such as the NOAA National Operational Model Archive and Distribution System (NOMADS).

Meanwhile, most of verification studies use available archived data that are limited. This underlines the need to describe the validity of verification results (especially for rare events), which includes the estimation of confidence intervals for verification metrics, using analytical or numerical methods. Also accounting for observational uncertainty in verification is a current subject of research. For example in light precipitation amounts, the bias in observed data could affect the probability of precipitation among other variables and should be taken into account when computing verification metrics.

The forecast verification on different space and time domains will need to be expanded, to include for example gridded forecast verification, which will become more important with distributed modeling approaches. This will involve more advanced verification metrics, such as an intensity-scale verification approach, object-oriented methods, or event-oriented methods to analyze spatial objects. Scale issues (e.g., observation scale vs. forecast scale) need to be accounted for; this area requires further research.

Regarding user-oriented verification measures, new approaches need to be developed to address specific operational questions. Verification experts should work closely with those users who have specific needs to jointly develop techniques that address their verification problems.

Additional educational opportunities regarding statistics and forecast verification should be made available through short courses, workshops, and web-based material, for better understanding and use of forecast verification.

An additional challenge in forecast verification is the need to have retrospective atmospheric forecasts from the same forecast system that is used operationally. This is a major hydrologic forecast requirement that is not being met by the current NWS forecast system.

8.5 A Roadmap for Getting There

The verification plan proposed in 2006 identified the roadmap to develop the comprehensive verification system by 2011. This verification plan will be updated regularly by the Verification System Requirements and Planning team to define and prioritize the different activities of research, development, and implementation given on-going work and findings, as well as available funding and resources.

8.5.1 Forecast Services Verification

The current HOSIP project on Logistical Verification focuses on forecast services to measure what hydrologic services the NWS provides, where these services are provided, and how often (HOSIP documentation). One of the first goals is to standardize and automate the collection of these measures (capabilities 1 and 2).

With the new interactive application called Forecast Services Manager, four types of point forecast services could be defined and managed in the IHFS-DB:

- Data point service: all locations on a river/stream for which observed data are input to RFC or WFO hydrologic forecast procedures, or included in public hydrologic products;
- Deterministic Forecast Service: all forecast points for which a single-value forecast is produced;
- Ensemble Forecast Service: all forecast points for which ensemble forecasting is used to generate forecasts and associated uncertainty information; and
- Water Supply Forecast Service: all forecast points for which water supply forecasts are provided.

This information will be used for management of the services provided by the NWS hydrology program and will be incorporated into the broader verification measures effort managed by the Office of Climate, Water and Weather Services (OCWWS) Hydrologic Services Division (HSD).

In the future, the logistical measures will be expanded to include measures for areal forecast services, forecast timeliness, and forecaster efforts to set up basins and issue forecasts. These logistical measures will be developed to be meaningful

for various users (e.g. managers, general public), including the selection of performance tracking measures (capability 7). Future work will also be needed to develop capabilities for display and dissemination of logistical measures (capabilities 3-5).

8.5.2 Forecast Verification

There are currently 4 HOSIP projects to develop different capabilities of the River Forecast Verification System:

- The RFC Archive Database on the RAX machines to address the capability 1 for deterministic forecasts;
- The Hydrologic Deterministic Verification project to develop IVP, which is available in AWIPS, and address the capabilities 2 and 3 for deterministic forecasts;
- The Ensemble Verification and Validation project to develop EVS, which is part of XEFS, and address the capabilities 2 and 3 for ensemble forecasts; and
- The Ensemble Hindcaster project to develop the Hydrologic Ensemble Hindcaster, which is part of XEFS, to address the capability 6 for ensemble forecasts.

Within the XEFS project, the existing applications IVP and EVS will be combined and expanded to develop the National Baseline Verification System (NBVS) that could verify both single-valued and ensemble forecasts that are operational or experimental. Also a capability for prognostic verification will be developed to define analog forecasts (i.e. past forecasts for similar hydrologic events) from forecast archive, estimate the quality of live forecasts using specific prognostic metrics, and effectively communicate results to forecasters and customers. Besides, the hindcasting capability will be expanded to retroactively generate hindcasts/re-forecasts that are either deterministic or ensemble, based on various forecast scenarios (e.g., different preprocessing or post processing techniques); this would include the capability to retrospectively generate raw model forecasts without any MODs.

Also the Western Region in collaboration with RFCs and WFOs is enhancing the Western Water Supply Forecast website to address capabilities 1 to 4. The Western Region water supply team and OHD will work closely to maintain consistency between the different verification functionalities to compute metrics and display results for various types of forecasts. This will help the development of a unified verification system to verify all types of forecasts and meet all user needs.

Closer collaboration with the RFCs and WFOs will help:

- Develop a verification system consistent with the operational river forecasting system and the verification needs of forecasters; it includes the definition of a raw model forecast application at each office following some general guidance, and the use of raw model forecasts as a baseline to evaluate the impact of MODs on the forecast performance; and
- Develop standardized verification strategies to effectively communicate results to end users while ensuring verification needs are met. The WR Hydrology Verification team and the NWS Hydrologic Forecast Verification team, which include the RFC verification focal points at the 13 RFCs, have started to work on verification case studies to propose standardized metrics and graphics to present verification results to identified end users.

Closer collaboration with agencies such as NCEP, ESRL, and NCAR will help develop the gridded forecast verification component by leveraging existing applications.

Also collaborating with a Hydrology Testbed will help the NWS to systematically verify all the existing and newly developed forecasting processes, to determine the most cost-effective methodologies to improve the forecasts, and to determine ways to effectively communicate the forecasts and verification information to user communities, with more user-oriented verification measures.

8.6 References

- Bradley A. A., S. S. Schwartz, and T. Hashino. (2004). "Distributions-Oriented Verification of Ensemble Streamflow Predictions." *Journal of Hydrometeorology*, 5(3), 532-545.
- Demargne J., L. Wu, D.-J. Seo, & J. Schaake. (2007). "Experimental hydrometeorological and hydrological ensemble forecasts and their verification in the U. S. National Weather Service." Quantification and Reduction of Predictive Uncertainty for Sustainable Water Resources Management, *Proceedings of Symposium HS2004 at IUGG2007*, Perugia, IAHS Publ. 313, 177-187.
- Franz K. J., H. C. Hartmann, S. Sorooshian, and R. Bales. (2003). "Verification of National Weather Service Ensemble Streamflow Predictions for Water Supply Forecasting in the Colorado River Basin." *Journal of Hydrometeorology*, 4(6), 1105-1118.
- Hashino T., A. A. Bradley, and S. S. Schwartz. (2007). "Evaluation of bias-correction methods for ensemble streamflow volume forecasts." *Hydrology and Earth System Sciences*, 11, 939-950.

National Research Council, Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts. (2006). "Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts." *The National Academies Press*, 124p.

Welles E., S. Sorooshian, G. Carter, and B. Olsen. (2007). "Hydrologic Verification: A Call for Action and Collaboration." *BAMS*, 88(4), April, 503-511.

9. Social Science Research (Future)

Appendix A- Review of higher-level NOAA guidance and NRC recommendations

A-1 The NOAA Hydrology Program Core Goals

The NOAA Mission Goals are each supported by a number of NOAA-wide programs. Under the Weather and Water Mission Goal, the most relevant to the activities of OHD is the Hydrology Program (NOAA Hydrology Program, 2005). The Hydrology Program has recently formulated 21 Core Goals. Those Core Goals and the OHD Strategic Science Goals (SSG; see *Section 1.2*) to which they are tied are as follows:

1. Improve the quality of physical inputs and forcings, e.g., Quantitative Precipitation Estimation and Forecasting (QPE, QPF), temperature, snow, evapotranspiration, soil conditions, burn data, etc. (SSG 3, 4).
2. Improve river forecasts by improving hydrologic models (Note: “river forecasts” include water supply forecasts) (SSG 11).
3. Improve forecasts of fast response hydrologic events (SSG 9).
4. Improve forecasts based on the effect of dam failures.
5. Improve hydrologic forecasts impacted by reservoirs and regulation (SSG 8).
6. Improve the routing techniques used to connect forecast locations (includes coastal effects) (SSG 14).
7. Improve flood forecast inundation maps.
8. Quantify the uncertainty of our forecast information (SSG 6).
9. Generate and disseminate information to and for our users (SSG 7, 15 and 16).
10. Provide, then improve, gridded water resource data production capability (SSG 11 and 13).
11. Provide, then improve, water quality forecasting capability (SSG 13).
12. Disseminate hydrometeorological data to the field (e.g. HADS).
13. Software refresh – enhance the usability and/or internal workings of existing software.
14. Allow the hydrology community to participate more fully in research to operations (e.g. CHPS) (SSG 1).
15. Archive information required to support the Hydrology Program now and in the future.
16. Verify our forecast and uncertainty information (SSG 7).
17. Provide science and software training on Hydrology Program applications throughout the research to operations cycle.

18. Inform customers of our information and services, assess their satisfaction, and incorporate comments and feedback into Hydrology Program planning (SSG 15 and 16).
19. Improve the efficiency and effectiveness of Hydrology Program management, including an understanding of logistical measures.
20. Update and maintain the nation's precipitation frequency estimates.
21. Define and coordinate Hydrology Program requirements with other NOAA programs (SSG 1).

Core Goals 4, 6 and 7 involve hydraulic modeling. Related SSG will be included in a future version of the Plan.

A-2 NOAA Strategic Plan for FY 2006-2011

The NOAA Strategic Plan identifies three high-level outcomes for the Weather and Water Goal, which are encapsulated in a number of the SSG. All of the SSG are directed at improving hydrologic forecasts and their use and, therefore, will help lead to *“reduced loss of life, injury and damage to the economy.”* Most of the SSG are directed at *“better, quicker and more valuable...water information...,”* with SSG 15 and 16 specifically directed at the use of that information *“to support improved decisions.”* SSG 15 and 16 will also lead to *“increased customer satisfaction”*.

The achievement of the Weather and Water outcomes is to be evaluated in terms of seven performance objectives. The NOAA Strategic Plan also lists six strategies for achieving the weather and water outcomes and performance objectives. Most of the above performance objectives and strategies can be mapped to the SSG. For example, SSG 6 and 9 come out of a recognition of the need to *“increase lead time and accuracy for...water warnings and forecasts,”* to *“improve predictability of the onset, duration and impact of hazardous and severe...water events,”* and to *“reduce uncertainty associated with...water decision tools.”* Several of the other performance objectives and strategies for the Weather and Water Goal emphasize users and their decisions, as do SSG 15 and 16. Likewise, the emphasis of several of the performance objectives and strategies on partnerships corresponds to SSG 1, and their emphasis on new data, science, and technology corresponds to SSG 2, 4, 11, and 12.

A-3 NOAA Draft Five-Year Research Plan for FY 2007-2011

The NOAA Research Plan identifies six overarching research questions for NOAA's mission:

1. *What factors, human and otherwise, control ecosystem processes and impact our ability to manage marine ecosystems and forecast their future state?*

2. *What is the current state of biodiversity in the oceans, and what impacts will external forces have on this diversity and how we use our oceans and coasts?*
3. *What are the causes and consequences of climate change?*
4. *What improvements to observing systems, analysis approaches, and models will allow us to better analyze and predict the atmosphere, ocean, and hydrological land processes?*
5. *How are uncertainties in our analyses and predictions best estimated and communicated?*
6. *How can the accuracy and warning times for severe weather and other high-impact environmental events be increased significantly?*

All six questions are relevant to varying degrees to OHD's activities over the next 5-10 years. OHD will help answer the first two questions related to coastal and marine ecosystems through SSG 14. The climate change issue is addressed by SSG 17. Although it is not OHD's mission to research the causes of climate change, a changing climate is potentially significant for the hydrological forecasting (i.e., a "consequence") at seasonal and shorter time scales that it is OHD's mission to support, particularly when that forecasting relies on an assumption of a stationary climate system. As noted in Section 2.2, this is one of the reasons this plan directs the research to the development of models that require a minimum amount of calibration. The fourth research question is directly addressed by SSG 3-8, 11 and 12; the fifth question by SSG 5,6,15 and 16; and the sixth question by SSG 9.

Under the Weather and Water Goal, the Research Plan provides the following discussion of the development and application of research tools:

"NOAA research focuses on technological developments in the major components of prediction: observational science, quality control, analysis, and ingestion of the observational data (e.g., data assimilation), improved numerical modeling, and user products and other services. Beyond reducing errors, a new emphasis will be on the description of uncertainty at all stages in the forecast process. Observations drive improved understanding of important processes. NOAA will integrate multi-purpose observing systems, especially those involving radars, satellites, and profilers, and obtain better observations of environmental parameters. The new observations will be digested by advanced data assimilation methods, reducing the error in the ensuing forecasts. Numerical modeling, including ensemble techniques, will focus on reducing and representing all forecast uncertainty for use in existing and new forecasts and warnings. Altogether, these improvements will lead to enhancements in NOAA's flagship weather and water forecast products to better serve the needs of the user community."

The above brief paragraph discusses many of the essential aspects of SSG 3-6.

The Research Plan discusses four specific research areas under the Weather and Water Goal, the most relevant to OHD being the improvement of water resources forecasting capabilities, which is described in Section 8.4.2 of the Research Plan. Nearly all the activities described in that section will be supported by OHD. The mention in the first paragraph of the need for an “*expanded suite of water resource predictions*” is covered by SSG 13. The longer timescales that water-resource predictions involve is addressed by SSG 10. The need for “*increasing the lead time for flood warnings and flow predictions, and quantifying and reducing uncertainty*” directly maps to SSG 6 and 9. The second paragraph goes on to discuss the type of modeling and data analysis the improvement in water-resources forecasting capabilities will involve. It identifies the need to model reservoir operations and water balances (SSG 8). The improvement of physical inputs and forcings (SSG 3) and the use of new data sets (SSG 4) are also discussed, particularly with regard to remote sensing. Improved data assimilation and uncertainty analysis (SSG 5 and 6) are also mentioned. Finally, it is noted that the improved modeling will involve “*a new generation of distributed rainfall-runoff models*” (SSG 11) and “*the coupling of ocean, atmospheric and hydrologic models*” (SSG 12). The former will in particular be necessary to account for the effects of groundwater pumping and irrigation on streamflows.

A-4 NOAA Twenty-Year Research Vision

The Twenty-Year Research Vision discusses advances in the four key technology sectors on which NOAA relies to fulfill its mission. The most relevant to OHD’s development and use of new models and computer systems over the next 5-10 years is:

“Information Technology will continue to advance with computer processing speed doubling every 18 months. There will be better frameworks for constructing complex modeling systems, as well as better data management and analysis tools. This will allow NOAA to advance model-based analysis techniques (through data assimilation) that will exploit the data acquired from new sensors. NOAA will employ high resolution, holistic models that include information on land-based activities, estuaries, coasts, oceans, living marine resources, and the atmosphere. These holistic models will enable NOAA to describe, understand, and predict the interactions of all parts of the environment at increasingly finer resolution.”

The above advances in information technology and the new types of modeling they will engender are reflecting in the SSG 2 - 6, 11 and 12.

In a table of sample NOAA products and services in 2025, the Twenty-Year Research Vision lists three examples for water-resource and hydrologic forecasting:

- “Water resource and drought forecasts including nutrient runoff;
- Improved stream flow forecasting models that cover flow levels from droughts to floods, including interactions with groundwater, water resources applications, estuaries and coasts; and
- New soil moisture forecasting models for agricultural applications and mud-slide warnings.”

While all the SSG are directed at development of improved water-forecasting products and services, the three above examples will in particular be outcomes of the fulfillment of SSG 8-13. Although water quality is part of SSG 13, it will only be discussed in detail in a future version of the Plan. Likewise, under SSG 14, it is anticipated that OHD will provide support to its NWS partners whose mission it is to conduct ecological forecasting for the Great Lakes and the ocean coasts, a detailed discussion of which will be included in the same update to the Plan in which water quality forecasting is discussed.

A-5 The NOAA Annual Guidance Memorandum for FY 2010-2014

The Annual Guidance Memorandum “*identifies the most urgent and compelling NOAA-wide programmatic and managerial priorities for FY 2010-2014...*” The introduction to the Memorandum identifies and discusses several “*external pressures to change.*” The most relevant to the Plan are discussed below.

Heightened awareness and acceptance of the scientific basis of climate change: The Memorandum points to the high-level of confidence (>90%) the Fourth Assessment Report of the Intergovernmental Panel on Climate Change places on the fact that the climate is exhibiting the effects of an anthropogenic warming. It then identifies NOAA’s role in providing regional-scale climate information. SSG 10 and 17 of this plan relate directly to the provision of climate information.

Demand for a strategy for improved operational forecasts of high-impact events: Although the discussion of this driver of change focuses on hurricanes, flash floods (SSG 9) can also be considered a high-impact event. In addition, the improvements in flood forecasting engendered in particular by SSG 3-9 and 11 will lead to better forecasts during extreme precipitation events such as occur during a hurricane.

Regional collaboration: The Memorandum identifies several state-led initiatives that are requiring NOAA to improve capabilities at the regional-scale. The drive towards greater regional collaboration is tied to the user-related SSG 15 and 16. The National Integrated Drought Information System (NIDIS) is discussed under both this heading and the above “climate

change”heading. It is anticipated that NIDIS will be a major interagency partner for OHD.

The remainder of the Memorandum provides a discussion of specific priorities for FY 2010-2014. In addition to expanding on the themes of climate-scale information, high-impact weather and water events and regional decision support, those priorities include the “*management and integration of observational data*”(SSG 2-5), “*forecasts of ecosystem health and productivity*”(SSG 14), “*Earth system modeling*”(SSG 12) and “*strategic use of information technology*”(SSG 2).

A-6 NWS Strategic Plan for FY 2005-2010

NWS Mission

The National Weather Service provides weather, water, and climate forecasts and warnings for the United States, its territories, adjacent waters, and ocean areas for the protection of life and property and the enhancement of the national economy.

And

NWS data and products form a national information data base and infrastructure, which can be used by other government agencies, the private sector, and the global community.

The NWS Strategic Plan identifies a number of major “*forces for change*” that will shape the context for the NWS over the life of the plan. The most relevant to OHD is:

“Requirements for a broader range of environmental information services from NWS, and more broadly, from NOAA including:

- ***Expanded climate information*** – *in all meanings of the term, i.e. retrospective studies of past and current climate; seasonal and longer forecasts of climate variations; and improved long-range predictions of climate change.*
- ***Expanded water information*** – *initially as part of the Advanced Hydrologic Prediction Service initiative already underway, but ultimately expanded to include a wider range of environmental information such as soil moisture and water quality forecasts for fresh water, estuaries, and the coastal zone.*
- ***True “ecosystem” forecasts including biological, chemical, and physical conditions.***

- *Expanded digital services* – allow communication of forecast information with greater resolution in time and space and facilitate the integration of data in all service program areas.
- An overall push, affecting all NWS service programs, to provide more explicit and more useful measures of forecast certainty.”

The expansion of water information is covered by SSG 11; ecosystem forecasts by SSG 14; the expansion of digital services by SSG 2, 3, 11 and 12; and the provision of more explicit and useful measures of forecast certainty by SSG 6 and 16.

Under other subheadings of Forces for Change, the NWS Strategic Plan discusses “continued advance in numerical models”(SSG 11 and 12), “expanding sources of observational data”(SSG 3-4), and “continued integration of environmental sciences”(SSG 12). Under the remaining two headings in the same section, the Strategic Plan discusses “responding to society’s needs”(SSG 15) and “our commitment’s to work together”(SSG 1).

The remainder of the NWS Strategic Plan discusses the role of the NWS in the NOAA Mission Goals. Of relevance to SSG 10, the anticipated improvements in intraseasonal to interannual climate forecasts and their application to hydrologic forecasting are highlighted under the Climate Mission Goal. Under the Ecosystem Goal, the “...greater emphasis on contributions of our ...water...forecasts for ecosystem forecasting...”is highlighted (SSG 14). Under the Weather and Water Mission Goal, the Plan covers: the need “...to better communicate information to the public”(SSG 6, 7 and 16), the “...move into a new direction of forecasts, including...water quality prediction...” (SSG 13); the “...need to be at the limits of the skill which science, technology, and a highly-trained workforce can provide”(SSG 9); “...improving data assimilation to use effectively all the relevant data we and others collect”(SSG 4 and 5); “... improving collaboration with the research community through creative approaches like community modeling (e.g., establish an Earth System Model Framework)”(SSG 1, 11 and 12); “...evolving our services from a text based paradigm to one based on making NWS and NOAA information available quickly, efficiently, and in convenient and understandable forms (e.g., National Digital Forecast Database and GIS)”(SSG 2); “...including information on forecast uncertainty to enhance customer decision processes”(SSG 6, 7 and 16); and the dependence “...on partners in the private, academic, and public sectors to acquire data, conduct research...”(SSG 1).

For each of the NOAA Mission Goals, the NWS Strategic Plan identifies specific NWS activities and links them to the NOAA strategies listed in the NOAA Strategic Plan for achieving mission-specific outcomes. The Weather and water activities most relevant to OHD science include:

“radar; surface observations (Automated Surface Observation System (ASOS), cooperative observer program (COOP), Hydrometeorological Automated Data System (HADS)); snow survey; Integrated Flood Observing and Warning System (IFLOWS); fire and soil observations; statistical forecast models (e.g., MOS); collaborative/common modeling infrastructures (e.g., WRF, Distributed Modeling Intercomparison Project DMIP); hydrologic research (e.g., VAR); forecast applications research; visualization and verification; collaborative forecasting; Service Programs (...Hydrology (water level, water supply, snow)...); Advanced Hydrologic Prediction Service (AHPS); IT Infrastructure (Advanced Weather Interactive Processing System (AWIPS, High performance computers); National Digital Forecast Database (NDFD).”

The involvement of OHD in many of the specific programs and projects identified in the activities list are discussed in the main body of the Plan.

A-7 NWS 2004 Science and Technology Infusion Plan

The Science and Technology Infusion Plan (STIP) states, *“The long-term (2025) S&T vision of the NWS is to provide the Nation with forecasts, warnings, and other environmental data, products, and information with lead times, specificities, and accuracy meeting thresholds established by risk managers and by careful socio-economic research.”* Specific types of warnings and forecasts are identified for this vision. The ones most relevant to OHD research and product development are:

- *“**Flash Floods:** Warning lead time increases from an average of 43 minutes in 2000 for counties to as much as 1 hour for specific portions of counties...” (SSG 9)*
- *“**Water Resources:** River, lake, and estuary forecasts; as well as other high-resolution water resource and soil moisture information are provided to customers where and when needed”. (SSG 13 and 14)*
- *“**Water Quality:** Reliable surface and estuarine water quality forecasts are provided to support maintenance, enhancement, and restoration of water supplies for aquatic habitat and domestic, agriculture, and industrial use.”(SSG 13 and 14)*
- *“**Climate:** Reliable probabilistic forecasts of temperature and precipitation indicating weekly departures from normal are issued months in advance. This allows better management of resources including water, fisheries.”. (SSG 10 and 14)*
- *“**Environmental Impacts:** Predictions of weather, water and climate variability and change, at time and space scales relevant to ecosystem models, provide resource managers with forecasts of natural impacts on ecosystems*

and scenarios of ecosystem responses to management decisions.” (SSG 10, 11, 13-16)

The STIP is organized around four S&T-dependent strategies discussed in an earlier NOAA Strategic Plan. The four strategies are treated as steps to collect, produce, deliver, sustain, and improve weather, water, climate, and related environmental information:

- **Monitor and Observe** elements that define the Earth environment (space, atmosphere, land-surface, ocean, coastal, and inland water), archive these data, and make them available and accessible to users;
- **Assess and Predict** the current and future state (from minutes to months and years) of the Earth environment by transforming observational data into forecast and warning products and information through data assimilation and numerical prediction models...
- **Engage, Advise, and Inform** users of these observations, warnings, forecasts, and other information to promote appropriate responses to changing hazardous and routine environmental conditions; and
- **Understand and Describe** the Earth system, develop new and improved observational systems, forecast models, and technologies, and demonstrate advances...

While the improvement of primary observations is not part of OHD’s mission, OHD will be involved in the development of secondary data products under SSG 2-4. OHD’s principal activities fall mostly under the assessment and prediction step. Some of the objectives the STIP identifies for this step are to: “*advance data assimilation technique*”(SSG 5); “*improve and couple numerical modeling systems*”(SSG 12); “*improve probabilistic predictions systems*”(SSG 6 and 7); and “*improve gridded forecast preparation applications*”(SSG 2). Under the second step, the STIP also discusses a vision for “*integrated probabilistic environmental forecasts and information,*” which involves a common Earth-system model (SSG 12) and “*integrated environmental forecasts that span minutes to months and seasons*”(SSG 10) and “*new types of forecast products such as...water quality...and harmful algal blooms*”(SSG 13 and 14). OHD will be involved in the third step primarily through SSG 15 and 16, and in the fourth step through SSG 12.

A-8 National Research Council Reports

Beginning with its 1996 report entitled, **Assessment of Hydrologic and Hydrometeorological Operations and Services** (NRC, 1996), The National Research Council (NRC) has undertaken, at the request of NOAA, a number of evaluation studies of NOAA operations and research programs. With regard to research and development, the 1996 report recommends that the NWS develop a

formal, long-term plan for hydrologic science research, which is part of an ongoing dialogue between NWS headquarters and its field offices as to the most appropriate research and product development for hydrologic services. This Strategic Science Plan represents the first such effort exclusively for and by OHD. Below, two recent NRC reports that specifically reference OHD activities are reviewed and related to the SSG.

The first NRC report is a review of the NWS Advanced Hydrologic Prediction Service (AHPS; NRC, 2006a). Because OHD is an integral part of the AHPS program and the report was commissioned by OHD, many of the findings and recommendations in the NRC report are applicable to the Plan. The NRC review is organized around three elements of AHPS; its programmatic foundations, its scientific and technical aspects, and its users. The five main recommendations for the science and technology of AHPS are most relevant to this Plan and are:

1. *AHPS developers are encouraged to work closely with satellite precipitation groups to ensure that AHPS hydrologic requirements for precipitation are considered in other federal activities, such as the National Aeronautics and Space Administration's Global Precipitation Measurement mission.*
2. *The NWS should strengthen quantitative precipitation estimation (QPE) and quantitative precipitation forecasts (QPF) for hydrologic prediction through an end-to-end evaluation that assesses QPE/QPF quality and impacts on flood and streamflow products for basins of diverse size and topography.*
3. *The NWS should strengthen connections between DMIP Phase I/DMIP Phase II and AHPS goals.*
4. *The NWS should clarify the criteria and decision-making process for selecting the next generation of hydrologic model(s) for AHPS, using an advisory group that involves modeling experts from inside and outside of the NWS to ensure that the state-of-the-art modeling advances are incorporated objectively into NWSRFS.*
5. *The NWS should invest in the next generation of NWSRFS that includes a flexible framework that allows alternative models, methods, or features that can be tested, verified, and implemented expeditiously. A total redesign of the NWSRFS is needed for AHPS to fulfill its scientific and technical goals.*

NOAA prepared a formal response to all of the NRC recommendations (NOAA, 2006). In concurrence with the first science-and-technology (S&T) recommendation, the response states that OHD is looking at including the Global Precipitation Measurement Mission in its Multi-sensor Precipitation Estimator (MPE). To that end, OHD will assess the potential of the Tropical Rainfall Measurement Mission (TRMM) data to improve the estimation of precipitation forcings (see **Section 4.1**). The first recommendation and the activity identified in the NOAA response are reflected in SSG 3 and 4. OHD's partnership with NASA is discussed throughout the Plan. OHD's vision for QPE and QPF is discussed in **Sections 4.1**

and 4.2, respectively. The use of probabilistic QPE and QPF in ensemble modeling is discussed in **Chapter** 6. NOAA further concurs with the second science S&T recommendation. NOAA's response to this recommendation identifies OHD's sponsorship of a Hydrology Verification Requirements team, whose mission encompasses the entire hydrologic forecast process including an assessment of the impact of the quality of QPE and QPF on hydrologic forecasts. Verification is the focus of SSG 7 and **Chapter** 8. The NOAA response to the third S&T recommendation discusses how the recommended connections already exist. Phase 2 of the Distributed Modeling Intercomparison Project is underway and is an integral part of the fulfillment of SSG 11. SSG 11 will also address the fourth S&T recommendation. In NOAA's response to this recommendation, OHD discusses its plans to set up an advisory group. NOAA's response to the fifth S&T recommendation discusses how the Community Hydrologic Prediction Service (CHPS) will become the software platform for the next generation of the NWSRFS. The hydrologic modeling paradigm underpinning the transition to the next generation of the NWSRFS is discussed in **Chapter 2.2**. The philosophy and design of CHPS will especially facilitate SSG 1 and 11. The importance of partnerships highlighted by SSG 1 is also the focus of one of the NRC programmatic recommendations for AHPS. In NOAA's response to that recommendation, a number of OHD federal agency and academic partnerships are identified, many of which are discussed in the remainder of this Plan. Finally, the user-related NRC recommendations are relevant to SSG 15 and 16. The Plan has been written with RFC and WFO needs and capabilities in mind and includes feedback from the RFCs on earlier draft.

The second NRC report (NRC, 2006b) is also heavily focused on users, namely how they understand and incorporate probabilistic forecasts into their decision-making. As such it is invaluable guidance for SSG 6, 7, 15 and 16. In particular, an entire section of the chapter on "*estimating and validation uncertainty*" is devoted to OHD. Three recommendations are developed in that section:

1. *OHD should implement operational hydrology databases that span a large range of scales in space and time. The contribution of remotely sensed and onsite data and the associated error measures to the production of such databases should be delineated.*
2. *OHD should organize workshops with participation from all sectors of the Enterprise to design alternatives to the AHPS ensemble prediction system components and develop plans for intercomparisons through retrospective studies, demonstration with operational data, and validation, and for participation in testbed demonstration experiments.*
3. *OHD should develop methods for seamlessly blending short-term (weather) with longer-term (climate) ensemble predictions of meteorological forcing within the operational ensemble streamflow prediction system. This will re-*

quire NCEP model output downscaling and bias adjustment, and real-time data availability.

Although the content of the databases is not made clear in the first recommendation, based on the preceding text it appears that the recommendation mostly refers to the observed and forecasted drivers of the hydrologic models (SSG 3 and 4). “...A large range of scales in space and time” refers to making the data applicable to forecasts ranging from flash floods in small basins to seasonal flows in large basins (SSG 9 and 10). The reference to error measures is with regard to making the data amenable to assimilation with model states (SSG 5 and 6). With regard to the second recommendation, OHD has already begun the process of replacing Extended Streamflow Prediction (ESP) with the Experimental Ensemble Forecast System (XEFS). An XEFS team comprised of personnel from NWS headquarters and RFCs has completed a Design and Gap Analysis. The XEFS contains a verification system (SSG 7). Ensemble test beds are part of the international Hydrological Ensemble Prediction Experiment (HEPEX), of which OHD is a co-leader. Both the XEFS and HEPEX are further discussed in *Chapter 6*. With regard to the third recommendation, blending of forecast across time-scales is an active area of research in OHD and is further discussed in *Chapter 6*. NCEP as an OHD partner is discussed throughout the Plan

At the end of NRC report on probabilistic forecasting, a number of overarching recommendations are made for the forecasting “enterprise” as a whole. The recommendations most relevant to OHD science and product development focus on: effective communication of forecast uncertainty (SSG 15 and 16); collaboration with users and partners (SSG 1, 15 and 16); production of objective uncertainty information over a range of scales (SSG 6); and verification studies and measures that are easily available to and understood by users (SSG 7 and 15; *Chapter 7*).

A-9 References

- NOAA (National Oceanic and Atmospheric Administration), 2006. “Toward a New Advanced Hydrologic Prediction Service”(AHPS) Report (National Research Council-NRC) Recommendation and Responses from NOAA’s Hydrology Program, 8 pp.
- NRC (National Research Council), 1996. Toward a new national weather service: Assessment of hydrologic and hydrometeorological operations and services. National Weather Modernization Committee, Commission on Engineering and Technical Systems, National Research Council of the National Academies. National Academy Press, Washington, D.C., 62 pp.
- NRC (National Research Council), 2006.: Toward a new Advanced Hydrologic Prediction Service (AHPS). Committee to assess the NWS AHPS Initiative, Water Science and Technology Board, Division on Earth and Life Studies,

National Research Council of the National Academies. National Academy Press, Washington, D.C., 74 pp.

NRC (National Research Council), 2006b. Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts. Committee on Estimating and Communicating Uncertainty in Weather And Climate Forecasts, Board on Atmospheric Sciences and Climate, Division on Earth and Life Studies, National Research Council of the National Academies. National Academy Press, Washington, D.C., 124 pp.

Appendix B- : Hydrologic Modeling Literature Review Paper

DISTRIBUTED AND PHYSICALLY BASED RAINFALL-RUNOFF MODELING FOR HYDROLOGIC FORECASTING: QUO VADIS?

*John Kochendorfer, Pedro Restrepo
NOAA NWS Office of Hydrologic Development*

B-1 Introduction

Rainfall-Runoff modeling for the purposes of hydrologic forecasting has a long and venerable history. Beven (2000), Singh and Woolhiser (2002) and Todini (2007) trace the historical development of models, from use of the rational method as early the mid-nineteenth century, through development of methods based on the unit hydrograph in the middle of the last century, to the numerically solved—and more physically realistic and complex—schemes most widely employed today. Our purpose in this paper is to examine where the science is at in physically based rainfall-runoff modeling as applicable to real-time hydrologic forecasting, particularly as implemented in distributed environments. Our interest is not in providing a comprehensive review of the hundreds of models in the literature—many thorough reviews of models already exist (e.g., Singh and Frevert, 2006b). Rather we are interested in a more general characterization of the types and performances of models most widely in use (or proposed for use) in operations. We are not only interested in mature models but also those in development. In other words, we are casting our eyes towards the next generation of models.

As the computing-power limitations to implementation of complex rainfall-runoff models at larger scales and finer spatiotemporal resolution have decreased, and more and more uncertainty and sensitivity analyses have been performed, a lively debate has ensued in the literature as to the quality of the information that such models produce. Questions being asked include: How well are actual hydrologic processes being represented? How well do models perform outside the range of observed conditions on which they are calibrated? What is the physical significance of the “effective” parameter values that result from model calibration? The reader is referred to Beven and Feyen (2002), Uhlenbrook *et al.* (2003) and Smith *et al.* (2004a) for a more comprehensive list of some of the most salient questions. In this paper, we do not attempt to provide novel answers to such questions, but rather to summarize the answers provided by the most recent literature. We rely heavily on the many special issues of journals and several IAHS “redbooks” that have been devoted to various topics related to physically based and distributed modeling, as well as to hydrologic forecasting (Andréassian, *et al.*, 2006b; Beven and Feyen, 2002; Blöschl, 2003; Franks *et al.*, 2005; Ghazi, 2005; Mitchell, *et al.*, 2004; Montanari and Uhlenbrook, 2004; Schaake, *et al.*,

2006; Schertzer, *et al.*, 2007; Sivapalan, 2006; Sivapalan, *et al.*, 2003b; Smith, *et al.*, 2004a; Tachikawa, 2003; Zehe and Sivapalan, 2007). Many of these publications are the outcomes of international projects devoted to furthering the art and science of physically based, distributed modeling. These projects include the Distributed Model Intercomparison Project (DMIP), the Model Parameter Estimation Experiment (MOPEX), the North American Land Data Assimilation System (NLDAS), the European Flood Forecast System (EFFS), and Predictions in Ungauged Basins (PUB) initiative. We provide brief reviews of the outcomes of the projects, with more depth given to those involving model intercomparisons.

B-2 Defining Distributed and Physically Based along a Continuum

Rainfall-runoff models are often defined as being either conceptual or physically based. The former consist of relationships that are typically developed more for their parsimonious use of parameters than their physical meaning. Nonetheless, conceptual models almost always make use of the water balance equation, which, as a conservation of mass equation, is physically based.

The most widely used models track the water balance in the unsaturated zone, leading O’Connell (1991) to coin them explicit soil moisture accounting (ESMA) models. ESMA models often make use of other equations with a strong physical basis such as the Penman-Monteith equation (which is based on the conservation of energy) for evapotranspiration losses and the kinematic wave equation (which is based on the conservation of mass and momentum) for the routing of overland, subsurface and stream flows. ESMA models were some of the first numerical rainfall-runoff models developed for operational purposes and are still the most widely used (see Section B-7 for examples).

In the minds of many, the use of Darcy’s law for subsurface flows is what most defines a truly physically based model (Beven, 2002). Freeze and Harlan (1969) were the first to lay out what could be considered the full set of equations and boundary conditions necessary to implement a completely physically based model—hereafter referred to as the FH69 blueprint following Beven (2002). The Systeme Hydrologique Europeen (SHE) of Abbott *et al.* (1986a; 1986b) is widely considered to be the first implementation of the FH69 blueprint. In criticizing the FH69 blueprint, Beven (2002) suggests defining physically based using two criteria: consistency with hydrological theory and consistency with observations. He argues that, given both the limited hydrological theory and the limited observations that we have for the scale at which they are applied, many so-called conceptual models qualify as physically based under those criteria, and many models that apply detailed small-scale physical theory to large scales are more appropriately called conceptual models. In the end, one can argue that all rainfall-runoff models are conceptual in that they are based on the authors’ “concept” of how the given watershed functions at the various scales at which the relevant processes operate (e.g., Blöschl and Sivapalan, 1995).

Likewise, there is no clear definition of what a lumped model is and what a distributed model is. In the simplest sense, a distributed model is any model that uses more than one computational element to calculate the runoff at a given watershed outlet. Many distributed models are lumped models applied over a rectangular grid, a patchwork of Hydrologic Response Units (HRUs), or a system of subwatersheds. It can be argued that for a model to be fully distributed, flows between computational elements along hill slopes must be represented. Explicit representation of topography would be a related requirement. It could also be argued that models with explicit topographic controls on infiltration, evapotranspiration and flows between hillslope elements are also more physically based. Distributed models without such representations might be more appropriately referred to as “semi-distributed,” although we caution the reader that this term means different things to different people. Although we prefer the term “quasi-distributed,” we’ll continue to use the “semi-distributed” connotation in this document.

Semi-distributed can also mean that watershed heterogeneity is characterized by distribution functions (Beven, 2000), such as with the infiltration storage capacity in the variable infiltration capacity (VIC) model of Wood *et al.* (1992) and topographic index of Beven and Kirkby (1979) in the suite of models collectively referred to as TOPMODEL. Much of the work in distribution-function based hydrologic models has been driven by the need to incorporate subgrid heterogeneities in the land-surface components of climate and atmospheric models in a computational efficient manner. The VIC model is an example of a model developed for that purpose. The models of Entekhabi and Eagleson (1991) and Stieglitz *et al.* (1997) are two other examples. The latter model is based on TOPMODEL and the topographic index. Because the topographic index and the related soil-topographic index are measures of the propensity of a location to become saturated and thereby generate saturated overland flow, they can be thought of as indices of hydrologic similarity (Sivapalan, *et al.*, 1987). Similarity of hydrologic response is also the idea behind dividing a watershed into HRUs based on topography, soils and vegetation (Beven, 2000). Given the wide range of methods and scales of dividing a watershed and the many algorithms used to estimate water fluxes to and from those divisions, we do not draw a bright line between models that are distributed or lumped, or conceptual or physically based, but rather use relative language, such as “more physically based” and “highly distributed.”

B-3 The State of Distributed Datasets

The boom in distributed modeling in the last couple of decades owes its existence to advances in computer hardware and software. In the former category are inexpensive parallel-processing platforms, and in the latter category are Geographic Information Systems (GIS) and other database management systems for storing and processing large geo-referenced datasets. In terms of populating those data-

sets, the single most important scientific advancement has been our ability to remotely sense precipitation, particularly from ground-based radar (e.g., Young, *et al.*, 2000)), but as well as from satellites (e.g., Grimes, *et al.*, 1999). Distributed datasets of the other atmospheric drivers of rainfall-runoff models—particularly ones containing the variables that determine the rates of snowmelt and evapotranspiration—have concomitantly increased in number, as well as in spatiotemporal resolution and coverage (e.g., Cosgrove, *et al.*, 2003; Maurer, *et al.*, 2002). Many distributed datasets of historical hydrometeorological variables owe their existence to the increasingly sophisticated art of re-analysis, which assimilates suites of surface, upper air and remotely sensed observations into global and/or regional climate models (e.g., Mesinger, *et al.*, 2006; Sheffield, *et al.*, 2006). The same advances in weather and climate modeling that have improved re-analysis data have produced more accurate and highly resolved weather and climate forecasts (e.g., Olson, *et al.*, 1995).

Model formulation and parameterization has been aided by distributed digital datasets of land-surface characteristics. Chief among these are high-resolution Digital Elevation Models (DEM) (e.g., Farr, T.G. and al., 2007), which have been invaluable for determining flow paths and channel networks (e.g., Tarboton, 1997). Also of great value have been datasets of soil characteristics digitized from soil-survey maps. In the US, the Natural Resource Conservation Service (NRCS) of the US Department of Agriculture (USDA) currently provides three polygon-based soil geographic databases, which are in order of increasing scale of mapping and decreasing level of detail: the Soil Survey Geographic (SSURGO) database, the State Soil Geographic (STATSGO) database, and the National Soil Geographic (NATSGO) database (USDA/NRCS, 1994). Only the latter two databases currently contain complete coverage of the coterminous US, although SSURGO is nearly complete. A one-kilometer raster database was developed from STATSGO by Miller and White, (1998) for use in hydrology models and the soil-vegetation-atmosphere transfer (SVAT) schemes of climate models. It has since been used in a wide range of rainfall-runoff model applications (e.g., Donner, *et al.*, 2004; Duan, *et al.*, 2006; Koren, *et al.*, 2003; Maurer, *et al.*, 2002; Smith, *et al.*, 2004a; Westrick, *et al.*, 2002; Yu, *et al.*, 2002)). Finally, distributed datasets of land use and land cover (LULC) continue to evolve; the USGS in particular has been producing successive generations of the 30-meter National Land Cover Database (NLCD) (Homer, *et al.*, 2004) from Landsat imagery. Additionally, a new 1-km global LULC database is available and updated every several years from MODIS satellite data (Friedl *et al.*, 2002). This MODIS-based 1-km LULC and its associated retrieval algorithms are the prototype for the global 1-km LULC database to be updated every several years from the next-generation National Polar-Orbiting Operational Environmental Satellite System (NPOESS). Also, NESDIS now operationally produces and delivers a weekly global 0.144 deg (~16 km) database of NDVI (Normalized Difference Vegetation Index) and GVF (Green Vegetation Fraction), in addition to a historical time series of these two weekly products from 1981 (Jiang *et al.*, 2008).

While the focus of remote sensing has mainly been on vegetation type and phenology, datasets characterizing the built environment in terms of impervious area are also being created (e.g., Elvidge, *et al.*, 2004; Goetz and Jantz, 2006).

Remote sensing of land-surface storages of water has also been the focus of considerable research over the last couple of decades. Given its presence above the soil surface, snow cover has proven the most amenable to remote sensing. The NOAA NWS National Operational Hydrologic Remote Sensing Center makes use of daily ground, airplane and satellite based observations to produce a range of snow products, including coverage, depth and water equivalent (Carroll, *et al.*, 1999). Soil moisture and ground water have proven more problematic for remote sensing, and such data have yet to be used for operational hydrologic forecasting. Both active and passive microwave have been studied for the measurement of soil moisture, with active sensing by Synthetic Aperture Radar (SAR) receiving the most attention (Moran, *et al.*, 2004). However, the bandwidths on currently deployed SAR satellite systems are only able to penetrate the first few centimeters of soil and are subject to interference from vegetation biomass and strongly dependent on surface roughness. Under the NASA/GFZ Gravity Recovery and Climate Experiment, monthly variations of total land-surface storage of water (including ground water) have been mapped at a two-degree resolution from variations in the Earth's gravity field (Han, *et al.*, 2005). By itself such coarse resolution data is unlikely to be of much value to hydrologic forecasting models, but may be valuable for assimilation in combination with point observations from monitoring wells.

B-4 Arguments for Moving towards more Physically Based Approaches

From the preceding discussion it is clear that the datasets necessary to drive, update and parameterize distributed, physically based rainfall-runoff models are increasing in number and spatiotemporal resolution. It is essential then that the models evolve along with these datasets.

For the next generation of models to make the most of our continually improving ability to remotely sense the physical states of the land surface, those models should be able to realistically represent those physical states in order to take advantage of the opportunities for data assimilation (e.g., Mitchell, *et al.*, 2004). We might even look to the day when streamflow forecasting is done with the same coupled models of the atmosphere and land surface that are used for weather and climate forecasting. Finally, being able to predict the state of vegetation, soil moisture and ground water—in addition to streamflow—is critical to hydrologic forecasting for water-resource management (e.g., Visser, *et al.*, 2006). For example, assessing the impacts of groundwater withdrawals on streamflow is best accomplished with a finite difference model of the saturated zone (e.g., Barlow, *et al.*, 2003).

It also has long been argued that more accurate representation of the physical states and processes in a watershed are key to accurate prediction of runoff for conditions outside the range of those under which a model was calibrated (e.g., Beven, 1989; Gan and Burges, 1990; Grayson, *et al.*, 1992; Kirchner, 2006; Klemes, 1986a; Wagener, 2003). In terms of the atmospheric drivers of models, these conditions include climate changes and extreme weather events. In terms of watershed characteristics, we can include any changes in LULC. Of particular concern are changes in soil hydraulic properties and vegetation cover that occur, for example, as a result of wildfire (e.g., Robichaud and Elsenbeer, 2001) or timber harvest (e.g., Andréassian, 2004; Croke, *et al.*, 2004). There is also increasing concern for the hydrologic changes that occur in rapidly urbanizing watersheds (e.g., Dougherty, *et al.*, 2007; Smith, *et al.*, 2002). Predicting flows with realistic runoff mechanisms—“getting the right answers for the right reasons” (Kirchner, 2006)—is also essential for predicting the concentrations of the matter and energy that is transported by the water, i.e., sediment, stream temperature and other water quality parameters (e.g., Quinn, 2004; Scanlon, *et al.*, 2004). Finally, physically based rainfall-runoff models also presumably perform better at predicting flows in ungauged basins (Sivapalan, 2003a), where *a priori*, uncalibrated (e.g., Koren, *et al.*, 2003) or regionally calibrated (e.g., Merz, *et al.*, 2006; Vogel, 2005; Wagener, *et al.*, 2004), parameters must be used. Regional calibration requires not only physically meaningful model structures, but parsimonious ones as well (Vogel, 2005).

B-5 Limitations of Operational Use of Complex, Highly Distributed Models

Increases in affordable computing power has been perhaps the single most important driver of the development and application of highly resolved rainfall-runoff models. It is now possible to run three-dimensional models over large watersheds discretized in the horizontal over elements on the order of 1000 m² and in the vertical over tens of levels, and at time steps of tens of minutes (e.g., Ivanov, *et al.*, 2004b). However, to run such models in operational time frames—even deterministically for forecast horizons of a few days—requires parallel-processor, distributed-memory systems with dozens of nodes. Thus, the computing demands of the most complex and highly resolved models are not trivial and can become prohibitive for any significant amount of model calibration or uncertainty analysis.

Whether the benefits of operational use of highly distributed, physically complex rainfall-runoff models outweigh their computational costs is an open question. It is a particularly important one given the resolution and certainty of the datasets of the hydrometeorological drivers identified in Section B-3. In particular, the NEXRAD Stage III precipitation is produced at a nominal 4-km resolution. Even at the resolution and after bias correction with gauge data, the data are subject to large uncertainties (Grassotti, *et al.*, 2003; Seo and Breidenbach, 2002; Young, *et*

al., 2000). Uncertainties in forecasted precipitation are even greater and mushroom with the length of the forecast horizon. Therefore, even if it can be argued that highly distributed, physically complex rainfall-runoff models provide more accurate predictions of runoff given highly certain precipitation inputs at small scales, those data are unlikely to ever exist outside of intensely gauged research sites. Indeed, even some of the biggest proponents of the art accept that highly distributed, physically based models remain, and should remain, mostly in the research domain (e.g., Grayson, *et al.*, 1992; Loague, *et al.*, 2006).

With regard to *a priori* estimation of parameters in a highly distributed environment, the 30-m resolution of the USGS DEMs and the National Land Cover Database (NLCD), along with the scale of the mapping units in SSURGO, suggests that it should be possible for horizontal elements as small as 1000 m². However, whether that is the most physically relevant scale is another question. Singh and Woolhiser (2002) state,

“a working concept of physical heterogeneity remains still elusive...the methods of subdivision are governed more by data availability than by physical meaning.”

Furthermore, while we have relatively good datasets characterizing the land surface, the subsurface remains the great unknown as a result of the difficulty of making large-scale measurements. In many watersheds, it is the subsurface heterogeneities that most control the runoff response to rainfall. For example, recent research at the Panola experimental watershed in Georgia points to the fact that, on hill slopes in humid climates, bedrock topography is equally or more important in the generation of storm flow than is surface topography (Freer, *et al.*, 2002), and that, in combination with pipe flow, is what is necessary to explain a threshold behavior (Tromp-van Meerveld and McDonnell, 2006). Beven (2002) offers some of the most trenchant criticism of the ability of models based on the FH69 blueprint to capture such flow pathways and runoff behavior:

“There is one very important limitation of the FH69 blueprint that will ultimately result in it being abandoned. Particularly in its description of unsaturated subsurface flow, it is based on Darcian theory that may be accurate at small scales but is certainly not applicable at large scales due to the effects of the nonlinearity of the unsaturated Darcy flow equation, the heterogeneity of soil properties and preferential flow of different types.”

Because even the most highly resolved and physical complex rainfall-runoff models still do not capture important heterogeneities and must rely on “effective” and highly uncertain parameter values, many commentators have criticized such models as being over-parameterized and over-fitted (e.g., Beven, 1989; Kirchner, 2006; Klemes, 1986a; Michel, *et al.*, 2006; Young, 1983). If all that is available for model calibration is a limited rainfall-runoff record, and streamflow at a single gauged site is the only desired predictand, then parsimonious conceptual

models may indeed be the order of the day (Michel, *et al.*, 2006). With potentially thousands of free parameters, highly distributed models cannot be calibrated in any optimal sense—and should not be given the uncertainty in and length of most rainfall-runoff records. In a seminal study of the information content in the rainfall-runoff records at seven catchments up to 90 km² in area, Jakeman and Hornberger (1993) concluded that a two-component linear model (representing slow and fast responses) with four parameters was adequate to capture the hydrograph behavior of all the catchments. In summary, the chief advantage of more conceptual and less distributed approaches is parameter parsimony and the resulting greater ease of (and justification for) finding optimal parameter values and of doing thorough uncertainty analysis.

B-6 New and Revisited Paradigms for Physically Based Modeling

As an alternative to the FH69 blueprint—which he views as inductive and aggregative—Beven (2002) proposes a “deductive” and “disaggregative” approach. He characterizes this alternative approach as one of mapping of the “landscape space” onto a “model space,” which has been defined broadly so as to encompass a range of plausible functional responses and parameter values. It is essentially a fuzzy classification of model structures and parameter sets into a behavioral set (i.e., those that are able to reproduce observations to some level of acceptability) and a non-behavioral set (i.e., all other model structures and parameter sets.) Therefore, rather than a specification of a particular model structure, Beven’s alternative blueprint is a methodology for selecting models based on their concordance with hydrologic theory and observations. Defining a behavioral set of models—as opposed to a single optimal model structure and parameter set—is closely allied with the equifinality thesis in rainfall-runoff modeling, which Beven and colleagues discuss in great detail elsewhere (e.g., Beven, 2006a). Most simply, equifinality implies that, given the limited observations available in any given watershed for model selection and calibration, many model structures and parameter sets can provide equally good predictions of watershed response. The behavioral set then becomes the basis for uncertainty analysis in model predictions. The Generalized Likelihood Uncertainty Estimation (GLUE) methodology of Beven and Binley (1992) is a prominent and widely used one for doing such analyses.

While behavioral modeling and equifinality concepts have been widely applied to parameter sensitivity and uncertainty analysis for a given model structure, they are only beginning to gain a foothold in the intercomparison of model structures (e.g., Vache and McDonnell, 2006). In the framework of Beven (2002) the challenge lies in defining the model space so as to include alternatives to the FH69 blueprint. As noted above, his rejection of that blueprint is primarily based on its use of the Darcian model of subsurface flow, which only holds at the scale of what is often referred at the representative elemental volume (REV). Although never well defined according to Beven, the REV scale is clearly much smaller

than any feasible control volume in a watershed-scale rainfall-runoff model. Any acceptable model structure should then reflect hydrologic functioning at the scale of the control volume that it uses. A specific alternative examined by Beven is the representative elemental watershed (REW) concept of Regianni *et al.* (1999; 1998). He describes the REW “*as essentially the area draining a link in the channel network.*” The uniqueness of the REW approach is not in using a large sub-unit of a watershed as a control volume, but rather in the formulation of scale-independent conservation equations for mass, energy and momentum. The challenge in the approach is finding the REW-scale closure relationships necessary for solving the conservations equations, particularly for energy. In a more recent critique of the REW approach, Beven (2006b) characterizes the closure problem as “*the Holy Grail of scientific hydrology,*” while observing that “*the relationship between internal state variables of an REW element and the boundary fluxes will be nonlinear, hysteretic and scale-dependent and may depend on the extremes of the heterogeneities within the REW.*” The REW approach as implemented in the Cooperative Community Catchment model based on the Representative Elementary Watershed (CREW) model (Lee, *et al.*, 2007) is discussed further in Section B-7.

In its disaggregative nature and focus on consistency with observations, the approach of Beven (2002) is akin to the “downward” or “top-down” approach to model development. The downward approach is seeing a resurgence (e.g., Sivapalan, *et al.*, 2003b), having been initially proposed by Klemes (1983). The downward approach is a systematic, hierarchal and iterative one in which a model is made successively more complex in an effort to match modeled to observed variables. In this way, the model structure ends up no more complex than is necessary to forecast the predictand(s) of interest to a desired level of certainty. The downward approach can be contrasted with the “upward” or “bottom-up” approach, which makes *a priori* assumptions about which processes are important and how they should be represented, usually based on small-scale physical theory (e.g., the FH69 blueprint). The upwards approach is thus often characterized as reductionist and mechanistic (Sivapalan, *et al.*, 2003a). Bottom-up model development often starts with attempting to reproduce hydrologic processes and observations at well studied and characterized experimental hill slopes and small, upland watersheds (e.g., VanderKwaak and Loague, 2001; Western, *et al.*, 1999). The difficulty is then in generalizing the results to larger scales. Although clearly not as well grounded in hydrologic theory as the upward approach, the downward approach can still be argued to be physically based because it relies on physical reasoning coupled to observations in selecting model structures. However, others have emphasized purely data-driven approaches (e.g., Young, 2003). In cases where the only data used are rainfall and discharge at single locations, the end result is usually a very simple lumped formulation with only a few storages (e.g., Farmer, *et al.*, 2003). As demonstrated by Jakeman and Hornberger (1993), this is consequence of the limited information content of most rainfall-runoff records.

That simple models with few parameters are often adequate to capture the rain-fall-runoff behavior of watersheds is one of the observations that led Grayson and Blöschl (2000b) to propose the dominant processes concept (DPC). As the name implies, the DPC recognizes the fact that typically no more than a few processes dominate the hydrologic response in a given watershed. Therefore it is not necessary “to model everything” to predict that response. Using the DPC thus requires abandoning the hope of a Hobbesian bargain with the bottom-up approach in which a single model structure can be developed that is applicable in all environments. The DPC is proposed at the end of an edited volume on “Spatial Patterns in Catchment Hydrology” Grayson and Blöschl (2000a), and it is in the modeling and observation of those spatial patterns (along with streamflow at the watershed outlet) that Grayson and Blöschl (2000b) see the potential for identifying what the dominant processes are. Most of the case studies in that volume are from heavily instrumented and well-studied small experimental watersheds. They recognize that the challenge for model development is in generalizing the results from such watersheds, especially given that the dominance of a given process appears to be a function of scale, climate, season and other environmental factors. Nevertheless, Woods (2002) is also encouraged by the potential for the DPC to serve as the basis for a system of hydrological classification and model selection.

The spatial patterns studied in Grayson and Blöschl (2000a) include precipitation, snowpack, evapotranspiration, soil moisture, overland flow, groundwater levels, and recharge/discharge areas. Grayson and Blöschl (2000b) note that much of the spatial data is binary (e.g., saturated vs. unsaturated, snow-covered vs. bare ground) or otherwise of a qualitative nature. Seibert and McDonnell (2002) categorize such data as “soft” because they cannot be used in traditional model calibration and validation. Nonetheless, they see soft data as valuable in improving the transfer of knowledge between experimentalists and modelers, in general, and in model calibration using fuzzy measures of model-simulation and parameter-value acceptability, specifically. They demonstrate their proposed methodology with a three-box model applied to a small study catchment in New Zealand. The hard data used consists of time series of streamflow and groundwater levels, and the soft data consists of isotopically estimated new-water contribution to peak runoff, fraction of saturated part of the soil, and frequency of groundwater levels above a certain level. Although their soft data can be expressed quantitatively, they are considered soft because they are discontinuous in time, highly uncertain, or both.

Vache and McDonnell (2006) take the use of isotopes further to show how model structures and parameter values in a distributed environment can be accepted or rejected based on their simulation of mean residence time. Uhlenbrook *et al.* (2004) use the results from tracer studies and other field investigations in a mountainous meso-scale watershed in Germany to develop the distributed tracer-aided catchment model (TACD). They use detailed process understanding and

GIS data on topography, soils and geology to delineate the 40 km² basin into eight “hydrological functional units”, each with a different dominant process. The same conceptual model and associated parameter values are applied to each of the 50-m grid cells with the same dominant process. Uhlenbrook *et al.* (2007) discuss the delineation of watersheds based on dominant processes in terms of potential relevance to prediction in ungauged basins.

Multivariate calibration and validation strategies do not necessarily have to involve soft data, nor are they applicable only in top-down modeling approaches. For example, Reefsgard (1997) shows how groundwater levels at several locations, in combination with discharge data, can be used to calibrate and validate MIKE SHE (a further development of SHE), as applied to a 440 km² watershed over 500-m grid cells. The prediction of groundwater levels using various versions of TOPMODEL has been particularly heavily studied (e.g., Campling, *et al.*, 2002; Lamb, *et al.*, 1997; Seibert, *et al.*, 1997)), and used for calibration of PRMS of an ephemeral watershed in Cyprus (Mazi *et al.*, 2003). For a highly resolved, physically based model applied to a small experimental watershed, Ebel and Loague (2006) compare observed and modeled pressure-head at three tensiometer locations. Among five sets of parameter values that provided good simulations of discharge over a seven-day sprinkler experiment, only one matched the near-saturation conditions observed at the three tensiometer locations. The authors suggest that use of such distributed data for model parameterization and validation is a means to “see through the fog of equifinality” in physically based modeling.

From the preceding discussion, it is clear that both upward and downward approaches to model development and testing have their advantages and disadvantages. Sivapalan (2003b) sees the strength in the upward approach as a means to gain detailed process understanding at the hill-slope scale, while the downward approach may prove better at creating model structures for the watershed scale. He sees the need to reconcile the two approaches in order to develop model structures that are generalizable. He argues that this can be achieved by finding “common threads” that link the hill slope to the watershed scale and that can be easily scaled. The examples he gives of such potential linkages include: travel time distributions, storage versus discharge relationships, storage versus saturated area relationships, and distribution functions of soil, vegetation and terrain characteristics. Elsewhere, Sivapalan (2003a) characterizes his proposed synthesis of the upward and downward approaches as philosophically in line with the combination of “the reductionist or mechanistic (e.g., Newtonian) and the holistic or ecological (e.g., Darwinian) worldviews” in the earth sciences, as recently argued for by Harte (2002). Accordingly, Sivapalan (2003a) sees the need for a paradigm shift in hydrologic theory, which among other criteria involves searching for patterns and laws in multiscale heterogeneities, particularly with regard to the co-evolution of climate, soils, vegetation and topography. It is out of those pat-

terns and laws that he expects expect to find the REW-scale equations of mass, momentum and energy balance.

Regardless of the paradigm under which one works, it is essential that any subsequently developed model be falsifiable in a hypothesis-testing framework. It has widely been recognized that many typical testing protocols, such as streamflow split-sample testing, have limited falsification power (e.g., (e.g., Kirchner, 2006; Klemes, 1986b; Kuczera and Franks, 2002; Oreskes, *et al.*, 1994; Refsgaard and Henriksen, 2004). Klemes (1986b) proposes several ways to make more rigorous use of streamflow data, such as calibrating on a dry period and validating on a wet one. Citing evidence in the literature that most rainfall-runoff records can support models of at most six free parameters, Michel *et al.* (2006) question whether complex, distributed models can be falsified at all. They advocate for rigorous hypothesis testing of parsimonious lumped model structures in a downward approach as being the most promising means of advancing basin-scale modeling. Loague and VanderKwaak (2004) see the ability to test hypotheses about flow pathways—particularly with regard to the FH69 blueprint—as the strength of distributed, physically based models. Both sets of authors recognize the need for data beyond streamflow at the watershed outlet, for the upward approach to be viable. Refsgaard (2000) discusses the use of spatial data in model calibration and validation. He develops a protocol for model conceptualization, coding and testing, for which he uses Refsgaard (1997) as a case study. Also noted above is the use by Vache and McDonnell (2006) of an isotopic tracer to accept or reject model structures in a “soft data” framework. Kuczera and Franks (2002) talk in general about the need to “augment” streamflow data with observations of other hydrological variables in testing a model as a hypothesis. They warn of the dangers of “fortifying” models with unjustified complexity without more thorough efforts at model falsification.

Kuczera and Franks (2002) also discuss the challenges and limitations of error and uncertainty analysis in model hypothesis testing, highlighting in particular the GLUE methodology and its use of a subjective likelihood function. Pappenberger and Beven (Pappenberger and Beven, 2006) and Beven (Beven, 2006a), on the other hand, argue strongly for the use of GLUE and other forms of probabilistic uncertainty analysis in model acceptance/rejection and hypothesis testing. Despite their limitations, GLUE and other methods based on Bayes’ Rule, Monte Carlo simulation or both (e.g., Kavetski, *et al.*, 2006a; b; Thiemann, *et al.*, 2001; Vrugt, *et al.*, 2003; Wagener, *et al.*, 2003) have increasingly been applied to the calibration and testing of distributed models with soft and spatial data. Largely because of the computational requirements, such studies been limited to smaller watersheds (e.g., Christiaens and Feyen, 2002; Freer, *et al.*, 2004; Vache and McDonnell, 2006), but have also seen their use with larger watersheds and more complex models (e.g., McMichael, *et al.*, 2006).

B-7 Examples of Distributed Models in Current Use

As discussed in Section B-2, models exist across a spectrum. So it is easy to argue one model or another is conceptual as opposed to physically based, lumped as opposed to distributed or semi-distributed. Most operational models would fall closer to the conceptual and semi-distributed end of the spectrum. Many of the cutting edge research models are physically based and fully distributed. We are most interested in models that have been used or are being proposed for use in real-time hydrologic forecasting.

Many others have catalogued and described to varying extents the large number of rainfall-runoff models in use. Loague and VanderKwaak (2004) provide a table of 20 “selected physically based and quasi-physically based models” with an emphasis on research models being employed “for concept-development purposes.” Singh and colleagues have been particularly prolific compilers of information on watershed models. Singh (1995) edited a volume with individual chapters on 25 models, most written by the author(s) of the given model. Singh and Frevert (2002a) and Singh and Frevert (2002b) contain similar numbers of chapters on individual models and are divided between models of “large watershed” and “small watershed” hydrology, respectively. A fourth volume (Singh and Frevert, 2006a) contains chapters on 24 different models and modeling systems, many developed since the earlier volumes. The same authors have compiled and an on-line inventory (<http://hydrologicmodels.tamu.edu/models.htm>) containing descriptions and references for more than 80 models as of late 2007. Singh and Woolhiser (2002) provide another lengthy compendium of models. They list 71 models in tabular form, including descriptors such as: physically based, process-oriented, event-based, continuous simulation, distributed, semi-distributed and lumped. Below we examine a cross-section of the most popular distributed models that fall along the continuums of conceptual to physically based and semi-distributed to fully distributed and that have been designed for application to large watersheds over operational time scales.

As noted above and in Section B-2, the models most popular today for operational use are lumped ESMA models applied in a semi-distributed environment. Prominent examples include: the Hydrology Lab-Research Modeling System (HL-RMS) (Koren, *et al.*, 2004) used operationally in the NWS as the Distributed Hydrology Model (DHM), the US Army Corps of Engineers’ Hydrologic Engineering Center Hydrologic Modeling System (HEC-HMS)(Feldman, 2000), the Precipitation Runoff Modeling System (PRMS) (Leavesley, *et al.*, 1983), used extensively at the USGS and USBR, the USDA Soil Water Assessment Tool (SWAT) (Arnold and Fohrer, 2005; Arnold, *et al.*, 1998), the USEPA Hydrology Simulation Program in Fortran (HSPF) (Bicknell, *et al.*, 2000), the Waterloo Flood (WATFLOOD) Forecast Model (Kouwen, 2002) used operationally by Environment Canada, and LISFLOOD (De Roo, *et al.*, 2000), a prominent model in the European Flood Forecast System (Pappenberger, *et al.*, 2005). In

application to large watersheds, the typical scale of the lumped computational element is 1-10 km². Both rectangular grid cells (e.g., HL-RMS, WATFLOOD) and irregular Hydrologic Response Unit (HRU) (e.g., PRMS, SWAT) are used in the listed models. As discussed further in Section B-8.1, results from the Distributed Model Intercomparison Project (DMIP) indicate that the performance of semi-distributed ESMA models vary widely, with HL-RMS and other semi-distributed models based on the lumped SAC-SMA performing particularly well. Other model intercomparison studies discussed in Section B-8 also suggest that semi-distributed ESMA are able to outperform more physically based and highly distributed models. Such results are evidence that the limitations of the latter models discussed in Section B-5 are real and daunting, and thus a ripe area for research, especially given that the arguments for using such models discussed in Section B-4 are also real. An example of an active area of research that is moving semi-distributed ESMA models in a more physical and more highly resolved direction is the coupling of such models with finite difference groundwater models (e.g., Niswonger, *et al.*, 2006; Said, *et al.*, 2005)

At the other end of the spectrum of physically based, distributed models are those models that are often referred to as “fully distributed,” which means that the computational elements used are at a high enough resolution that the controls of topography on runoff can be explicitly accounted for. They also typically follow the FH69 blueprint in applying small-scale physical theory to the scale of the computational element. Although SHE is often cited as the first implementation of the FH69 and includes explicit topographic controls on runoff, it and MIKE SHE are typically applied over rectangular grid cells 100 or more meters on a side (e.g., Refsgaard, 1997). At that resolution, the physically based algorithms in the model start approaching conceptual ones because of the mismatch in scales between theory and model. In particular, overland flow in MIKE SHE is based on a numerical solution of the equations of the conservation of mass and momentum (i.e., the de Saint Venant equations) in two dimensions. Another model that uses the 2-d de Saint Venant equations is CASC2D (Julien, *et al.*, 1995), which is used operationally by the US Army. CASC2D is typically implemented at the resolution of the 30-m DEM that it uses. In its original version, CASC2D models runoff on from the infiltration-excess runoff. Redistribution of subsurface soil moisture and saturation-excess runoff via interaction with a water table are part of the Gridded Surface/Subsurface Hydrologic Analysis (GSSHA) model (Downer and Ogden, 2004), an offspring of CASC2D. The TIN-based Real-time Integrated Basin Simulator (tRIBS) (Ivanov, *et al.*, 2004a) is a model that combines the high resolution of CASC2D and GSSHA with the more comprehensive physical representation of the entire land-surface of MIKE SHE. As the name implies, tRIBS discretizes the land surface into a triangulated irregular network (TIN), which also allows for variable sizes of triangular computational elements based on the complexity of the topography. Although yet to be used in operations, tRIBS has been extensively investigated for that purpose at the basin scale (e.g., Ivanov, *et al.*, 2004b).

In terms of capturing high-resolution spatial heterogeneities of physical properties and processes, the distribution-function variety of semi-distributed models discussed in Section B-2 are intermediate between semi-distributed models of the ESMA variety and fully distributed models. The VIC model and TOPMODEL were cited as examples of this class of model. VIC, which has its origins in the lumped, conceptual models of Zhao (1977) and Todini (1988), does not explicitly account for topography and thus could be argued to be less physically based than TOPMODEL. The Topographic Kinematic Approximation and Integration (TOPKAPI) model (Ciarapica and Todini, 2002) is essentially a hybrid of VIC and TOPMODEL. Since its conception, the VIC model has grown greatly in sophistication both in its representation of subsurface flow and the exchange of water and energy between the vegetation canopy and atmospheric boundary layer (e.g., Liang, *et al.*, 1994; Liang, *et al.*, 2003). While developed as a land-surface scheme for Global Climate Models (GCMs), it probably sees greater use today “off-line” as a macroscale hydrologic model. In particular, it is being used operationally with seasonal climate forecasts (Wood, *et al.*, 2002). Since its inception, numerous versions of TOPMODEL have also been developed in an effort to increase its physical realism. These include a fully distributed version. TOPMODEL has also been the basis of many land-surface schemes for atmospheric models. The first—and most sophisticated in terms of representation of spatial variations in soil and vegetation—is the TOPMODEL-based Land-Atmosphere Transfer Scheme (TOPLATS; Famiglietti *et al.*, 1992). Although TOPLATS has been used extensively in the investigation of the assimilation of remotely sensed soil moisture data, it has yet to be fully coupled to an atmospheric model, owing largely to its computational intensity. This led Stieglitz *et al.* (1997) to propose a simpler version.

The REW modeling approach discussed in Section B-7 can in many ways be thought of as intermediate in complexity between the fully distributed and distribution-function approaches. As an example of its kinship to the distribution function approach, the CREW model (Lee *et al.*, 2007) bases the relationship between saturated surface area and average depth to groundwater on TOPMODEL assumptions. Other REW-scale equations are based on the assumption that point processes can be scaled up using spatial distributions of parameters. For example, the REW-average infiltration capacity is derived from the Green-Ampt equation under the assumption of lognormally distributed saturated hydraulic conductivity. Similarly the lognormal distribution of saturated hydraulic conductivity is combined with the local-scale exfiltration capacity model of Eagleson (1978a,b) to derive a closure relationship for the capacity for exfiltration by evapotranspiration. The kinship of the CREW model to fully distributed models can be seen in the closure relationships that were obtained from numerical simulation in a bottom-up manner. For example, CATFLOW (Zehe *et al.*, 2001), a fine-scale physically based distributed model, was used to derive and parameterize a power law form for seepage outflow as a function of the ratio of the degree of saturation

over the entire REW to the average matric potential over the entire REW. In a similar manner Lee *et al.* (2007) use the CATFLOW model to parameterize REW-scale power laws for the water retention and hydraulic conductivity curves.

B-8 Results from Model Intercomparison Projects

B-8.1 The Distributed Model Intercomparison Project

In 2002, the National Weather Service (NWS) of the National Oceanic and Atmospheric Administration (NOAA) sponsored phase I of the Distributed Model Intercomparison Project (DMIP) (Smith, *et al.*, 2004a). The principal goal of the project was to ascertain the advantages and limitations of using distributed hydrologic modeling over lumped modeling for operational flood forecasting. A detailed discussion of the DMIP science questions can be found in (Smith, *et al.*, 2004c). The central hypothesis tested was that distributed models will lead to more accurate outlet hydrograph simulations. Related questions concerned calibration, prediction at ungauged interior points, model complexity and the use of distributed estimates of rainfall from radar. Although the emphasis was on a comparison of lumped versus distributed modeling, there appears to have been some interest in the performance of conceptual models against more physically based models. Regarding the former goal, (Smith, *et al.*, 2004c) conclude based on their survey of the literature that the advantages of distributed modeling are not always clear cut, often depending on the size and physiography of the basin, the type and resolution of the rainfall data, the dominant runoff mechanisms, among many other factors. They also find many studies in which increasing model complexity did not necessarily improve results, especially when driven by coarsely resolved and uncertain radar data. Given the widely varying and often conflicting results in the literature, the DMIP was designed to provide a set of common test watersheds typical of the size for which the NWS River Forecast Centers provide flow forecasts, along with the data that is typically used to parameterize, calibrate and force a semi-distributed version of the Sacramento Soil Moisture Accounting (SAC-SMA) model developed at the NWS Office of Hydrologic Development (OHD) (Koren, *et al.*, 2004) and which is part of the OHD Hydrology Lab Research Modeling System (HL-RMS). Most significant amongst the forcing data were seven years worth of high-quality hourly NEXRAD Stage III precipitation estimates gridded at a nominal 4 km resolution.

In addition to OHD's use of the HL-RMS distributed model and the lumped version of the SAC-SMA, six (out of a total of 11) other participants from government, academia and the private sector submitted calibrated results from their respective models for at least two of the four primary parent basins (Reed, *et al.*, 2004). Worthy of note is that one of those models, the Hydrologic Research Center Distributed Model (HRCMD; (Carpenter and Georgakakos, 2004)) is also a semi-distributed version of the SAC-SMA. In contrast to the HL-RMS version, which was run over 4-km grid cells, the HRCMD used sub-basins between 59

and 85 km². Two additional models are of a similar vein to the HL-RMS and HRCDM in that they are essentially lumped ESMA models applied in a distributed environment. The two models are the Soil Water Assessment Tool (SWAT; (Di Luzio and Arnold, 2004) and WATFLOOD (Kouwen, *et al.*, 1993). In the DMIP, the former model is divided into HRUs 6-7 km² in size, and the latter is applied over one-km grid cells. One of the models is of the distribution-function variety of semi-distributed models. It is the TOPNET model, which is a version of TOPMODEL and is applied to sub-basins on the order of 90 km² in the DMIP (Bandaragoda, *et al.*, 2004). TOPNET is thus one of the more physically based of the DMIP models. The remaining two models are also strongly physically based. The first is the r.water.fea model of (Vieux, *et al.*, 2004), which uses a kinematic wave approximation to route overland flow between one-km grid cells. As an event-based model, r.water.fea, could not be evaluated as completely as the continuous simulation models. With overland flow and lateral movement of soil moisture in both the saturated and unsaturated zones calculated within triangular irregular networks (TINs), The TIN-based real-time integrated basin simulator (tRIBS) used by Ivanov *et al.*, (2004b) can be argued to be the most physically based of all the models used in the DMIP. Use of the TINs allows for efficient representation of high-resolution topography such that the number of computational elements can be reduced by a factor of about 20 from a 30-m raster DEM (Vivoni, *et al.*, 2004). Nonetheless, the major basins in DMIP, which are on the order of 1000 km², still require tens of thousands of nodes.

Using a wide range of statistics, Reed, *et al.*, (2004) summarize model performance. For the calibrated results, they note that the three models that consistently exhibit the best performance on all but the smallest interior basin make use of the SAC-SMA model. It should be noted that statistics for a six-year calibration period were combined with the statistics for the fourteen month verification period and therefore no data are provided to indicate performance in the verification period alone and the consequent likelihood of overfitting. For the combined calibration/verification period, the lumped SAC-SMA model averaged a Nash-Sutcliffe (NS) efficiency of 0.79 for the five parent basins. The HL-RMS and the HRCDM achieved average NS efficiencies of 0.77 and 0.76, respectively. This suggests that there was no benefit in applying the SAC-SMA model in a distributed environment. We suspect that much of the success of the lumped SAC-SMA models is largely due to the highly refined manual and automated calibration strategies that have been developed over the years at the NWS and the large number of free parameters (fifteen) in the model. With an average NS efficiency of 0.70, TOPNET also performed well. Calibrated results for tRIBS were provided for only two of the four main basins, for which an average NS efficiency of 0.55 was achieved. This is still considerably better than the results obtained with the remaining two ESMA models, SWAT and WATFLOOD, with the former averaging an NS efficiency of 0.26 and the latter an efficiency of 0.42. Similar relative performance of models was found based on a modified correlation coefficient. So based on the results presented in Reed, *et al.*, (2004), it is tempting, yet difficult

to draw a blanket conclusion that more physically based and more highly distributed models perform less well compared to calibrated lumped or semi-distributed conceptual models, especially given the lack of a separate verification period.

Model performance was also compared at three interior basins for which streamflow data was available but not allowed to be used in calibration. The calibrated parameter values are those derived using streamflow at the outlet of the parent basin. Although prediction at interior basins is presumably a strength of distributed modeling, the lumped SAC-SMA performed nearly as well or better than the two distributed versions of the SAC-SMA model at two of the interior basins, with an average NS efficiency of 0.62 for the lumped model and 0.68 for the two distributed models. For the third and smallest interior basin, Christie, the lumped SAC-SMA performed poorly with an NS efficiency of -0.26, a decrease from the uncalibrated results. Of the two distributed versions of the SAC-SMA model, only results for the HL-RMS were presented, and it too showed a negative NS efficiency for the Christie basin that was worse than for the uncalibrated case. Other models also showed mixed results for three interior basins, with the exception of TOPNET, which had an NS efficiency of 0.59 for Christie, and 0.47 and 0.52 for the other two interior basins.

Reed, *et al.*, (2004) also present event statistics in the form of errors in total flood volume, flood peak runoff, and time to peak. They note that three best performing models for the calibrated results are again the lumped SAC-SMA model and its two distributed versions. Only flood peak using the HL-RMS was better than with use of the lumped model for all basins—with the percent improvements ranging from 0.3 to 11.0. Across the three event statistics, the best improvement with the HL-RMS was in the Blue River parent basin with a 9.9% improvement in peak flow, a 3.3 hour improvement in time to peak, and a slight worsening in flood volume of 2.3%. (Reed, *et al.*, 2004) also note modest peak flow improvements for the Blue River with the HRCDM and one other model. However, given the elongated shape of and the wide range of soil textures in the Blue River, and its more flashy nature, one would expect much greater benefit from distributed modeling.

Using three indices of rainfall spatial variability and basin rainfall filtering, Smith, *et al.*, (2004b) investigate how the basin characteristics of the Blue River and two other DMIP basins translate precipitation variability into runoff variability. They indeed find that the elongated shape of the Blue results in a high degree of variability in the location of rainfall relative to the centroid of the basin, whereas the more rounded shapes of other two basins tend to keep storms more evenly distributed over the basin. Dampening of rainfall appears to be substantially lower for the Blue only for rainfall centered in the lower half of the basin. This suggests that that is the flashiest part of the basin, which is in agreement with both the shorter travel times and the fact that the lower half of the basin contains a large area of clay soils that are likely to produce infiltration-excess over-

land flow. (Smith, *et al.*, 2004b) plot the differences in the performance of lumped SAC-SMA model and the HL-RMS against their three indices for the 23 to 31 events for which they were calculated. Although there is a high degree of scatter, there is some indication of a tendency for greater improvements with the HL-RMS when rainfall is located in the lower half of the basin. If indeed such events tend to be flashy with large peaks, then they are particularly worthy of improved forecasting, and performance statistics based on all events in the basin tend to underestimate the value of distributed modeling of the basin.

In comparison to the HL-RMS results for the Blue River, improvements in event statistics are much more profound for TOPNET with the interior Christie basin. With a 21 % improvement in flood volume and a 67 % improvement in peak flow, the TOPNET model again shows itself to be the best performer for Christie. With 6 % and 30%, respective improvements, the tRIBS model also outperforms the HL-RMS in the basin. Reed, *et al.*, (2004) suggest that one of the reasons TOPNET may do well in Christie is its tendency to produce lower flood volume estimates than the other models. However, that tendency is not great enough to explain the entire improvement, as well as the relatively high NS efficiencies noted above for all three interior basins. More than likely it is an issue of scale, given that, at 65 km², the Christie Basin is less than a fourth the size of the next larger interior basin. Results presented in Bandaragoda, *et al.*, (2004) for sub-basins of the Blue River suggest that there may be a certain amount of scale independence in the distribution of the wetness (i.e., topographic) index, which is in keeping with the Representative Elementary Area concept of Wood *et al.* (1988). So it may be the preservation of flow pathway information in the distribution of the wetness index that explains the relatively good performance of TOPNET for the interior basins. A similar judgement for the high-resolution topography in tRIBS is difficult to make as its event statistics for Christie were only modestly better than those for the HL-RMS and no calibrated runs were submitted for the other two interior basins.

In addition to the need for a better understanding of the impact of scale, one of the more significant conclusions of (Reed, *et al.*, 2004) is the following:

“Among calibrated results, models that combine techniques of conceptual rainfall–runoff and physically based distributed routing consistently showed the best performance in all but the smallest basin. Gains from calibration indicate that determining reasonable a priori parameters directly from physical characteristics of a watershed is generally a more difficult problem than defining reasonable parameters for a conceptual lumped model through calibration.”

That statement suggests the benefits of not only distributed modeling, but more physically based distributed modeling can be realized with improvement methods

of *a priori* parameter estimation. That possibility is examined in the next section, mostly within the context of MOPEX.

B-8.2 The Model Parameter Estimation Experiment and *A priori* Parameter Estimation

Initiated in 1996 with funding from the NOAA Office of Global Programs, the Model Parameter Estimation Experiment (MOPEX) is an international project devoted to the enhancement of techniques for the *a priori* estimation of parameters in hydrologic models and the land surface parameterization schemes (LSPSs) of atmospheric models (Schaake, *et al.*, 2006). The MOPEX science strategy involves: data preparation, *a priori* parameter estimation methodology development and improvement, and demonstration of parameter transferability. The first step has resulted in a comprehensive database that contains historical hydrometeorological data and land surface characteristics data for many river basins in the United States (US) and in other countries. (Duan, *et al.*, 2006) report on results from the second and third of five workshops that have been held to date. Those workshops focused on the second step of the MOPEX strategy. (Duan, *et al.*, 2006) identify three questions to be addressed:

- (1) “How do we define the relations between model parameters and basin characteristics?”
- (2) How can model calibration be used to refine the *a priori* parameters?
- (3) How do we evaluate the uncertainty due to model structure, calibration data and model parameters?”

To address the second-step questions a model intercomparison experiment was designed for 12 basins located in the eastern United States and encompassing a range of climatic regimes. The required model runs were completed with four rainfall-runoff models (SWB, GR4J, SAC-SMA, and PRMS) and three LSPSs (ISBA, SWAP, and Noah). An eighth model (VIC) has been used both as a watershed model and as an LSPS in atmospheric models. SWB and GR4J are simple conceptual models with two storages and four and five free parameters, respectively (Andréassian, *et al.*, 2006a). The infiltration and runoff algorithms in SWB have been incorporated into Noah. Like SAC-SMA, PRMS and VIC are relatively sophisticated ESMAs that have considerable physical basis. All three have been applied in distributed environments; however they presumably were run in lumped mode for the MOPEX experiments. The three LSPSs are typical of the genre of models in that they are one-dimensional with multiple soil layers and sophisticated representations of the vegetation canopy. Their main function is to model the exchange of moisture and energy with the atmosphere at time steps of minutes to hours. Runoff is typically viewed as a means for validating evapotranspiration rather than as a main predictand. When coupled to atmospheric models, they are used to calculate land-surface fluxes over grid cells of the size used for the overlying atmospheric column, which can be any where from a few kilome-

ters to a few degrees of latitude and longitude. For the MOPEX experiments, they were presumably run as lumped models.

In order not to make the experiment a ranking of model quality, Duan, *et al.*, (2006) anonymously assign letters to the results for each model. Their main intent is to establish benchmarks for current *a priori* parameter procedures. NS efficiencies were calculated for daily streamflow over the period 1960-1998. As one might expect, a wide range of NS efficiencies are seen for the *a priori* parameterizations, with no one model performing well in all basins. Each model achieved an NS efficiency near or above 0.8—which is often considered a good fit between modeled and observed values—for at least one basin. Five models produced negative NS efficiencies for one to a few basins, indicating that modeled values are on average worse predictors of the observed values than is the mean. The models were then calibrated to 19-year periods in the record. The original intent was for all models to use the same periods for model calibration and validation. Because not all participants did so, a traditional split-sample comparison was not possible. Instead, Duan, *et al.*, (2006) present NS efficiencies for the combined calibration and validation periods. As one would expect, calibration resulted in large improvements, with many more model-basin pairs achieving NS efficiencies near 0.8 and none dropping below 0.2. The main conclusion of Duan, *et al.*, (2006) is “that existing *a priori* parameter estimation procedures are problematic and need improvement.” As acknowledged by Duan, *et al.*, (2006), finding answers to the three questions that they pose clearly requires more detailed analysis of results.

A fourth MOPEX workshop was held in Paris in July 2004 in which a new series of 40 French river basins were added to the MOPEX database (Chahinian, *et al.*, 2006b). Chahinian, *et al.*, (2006a) summarize the results from the 13 models that participated in the simulation experiments. Three of the models (AFFDEF, HYDROTEL and MODSPA) were run in distributed environments. The ten lumped models are SAC-SMA (used by two participants), GR4J, GR5H, HBV, IHAC, MORDOR, TOPMO, Noah, SWB, and VIC. Andréassian, *et al.*, (2006a) provide a catalog of information on the models. The models range in complexity from a two-storage/four-parameter model (GR4J) to more physically based model such as the Noah LSPS, which uses physically based equations of the coupled energy and water balance, and HYDROTEL, which uses the Richards equation to compute vertical movement through three tilted soil layers and the kinematic wave equation to route runoff to and through stream channels. Participants were allowed to run their models on sets of 3, 12 or all 40 catchments. HYDROTEL, SAC-SMA, AFFDEF and MODSPA were run on the three-catchment set (3C); VIC, SAC-SMA, Noah and SWB on 11 catchments in common (11C); and the remaining six models on the 40-catchment set (40C). Based on several performance statistics for the 3C validation period, Chahinian, *et al.*, (2006a) conclude that SAC-SMA does the best, followed by HYDROTEL and MORDOR. The AFFDEF results and the SAC-SMA results from the other participant using the

model are judged to be the worst. The difference in the two uses of the SAC-SMA is attributed to different calibration techniques, although those techniques are not specified. Chahinian, *et al.*, (2006a) state that they cannot ascertain whether the differences in the performance of the other models are due to differences in model structure or calibration strategies. They also see no clear difference between the performance of the lumped and distributed models. For the 11C validation runs, (Chahinian, *et al.*, 2006a) place MORDOR, GR5H and SAC-SMA in the top three and see Noah and VIC as performing significantly worse than the other models. For the 40C validation runs, MORDOR, GR5H and SAC-SMA again rank at the top and Noah at the bottom. They note that model performance cannot be related to the number of calibrated parameters, as the top three models have, respectively, 10, 5 and 13 free parameters.

Models were also run in an ungauged (i.e., *a priori* parameterization) mode. Eight models were run in the ungauged mode for 3C. SAC-SMA, SWB, and MODSPA were able to produce NS efficiencies greater than 0.64 for all three catchments. Four models were tested in the ungauged mode for 11C. Based on NS and bias statistics, SAC-SMA was the clear winner and Noah the clear loser, with SWB and VIC ranking closer to SAC-SMA than Noah.

Duan, *et al.*, (2006) do not identify the *a priori* parameterization and subsequent calibration techniques used by each of the models, while Chahinian, *et al.*, (2006a) provide only a list of parameterization techniques in a table. However, some of the techniques are documented elsewhere, particularly in a special issue of the J. of Hydrol devoted to the second and third MOPEX workshops (Schaake, *et al.*, 2006) and an IAHS Red Book (Andréassian, *et al.*, 2006b) devoted to the fourth workshop. In general, regionalization methods are applied to the lumped, parsimonious models. Merz, *et al.*, (2006) give an overview of two major categories of regionalization methodologies: those based on spatial proximity, and those based on catchment attributes. Both typically involve calibration of the model in numerous gauged catchments, with the former method involving spatial interpolation of parameter values from nearby catchments, and the latter method involving development of empirical relations between catchment attributes and parameter values. Merz, *et al.*, (2006) also summarize numerous regionalization studies using large catchment samples located mostly in Europe. A wide range in performance of the various methods is reported. The HBV model is the focus of many of the studies. The methodology of Hundecha and Bardossy, (2004), in which functional relationships are developed between model parameters and catchment attributes (presumably by non-linear regression), appears to perform the best for HBV, with NS efficiencies between 0.79 and 0.90 for 30 calibration catchments, and between 0.76 and 0.92 for an unspecified number of validation catchments. (Merz, *et al.*, 2006) conclude by identifying some of the problems with regionalization approaches, including: low correlations between model parameters and catchment attributes, the lack of catchment attributes related to sub-

surface processes, and the difficulty of spatially interpolating parameters in a manner that accounts for the hydrologic organization of the landscape.

Models that are more physically based tend to use parameter values that are fundamental physical constants or measurable properties of the system. Some can be estimated directly from distributed datasets, such as the topographic index from DEMs. Others require using empirical relationships developed independent of the model, such as leaf area index from the remotely sensed Normalized Difference Vegetation Index and soil hydraulic properties from soil texture. The latter relationships are typically referred to as pedotransfer functions. LSPSs often use look-up tables which relate vegetation and soil-texture classes to model parameters. Gusev and Nasonova, (2006) describe the use of the 12-basin US MOPEX database with pedotransfer functions and a vegetation look-up table to estimate parameter values for the SWAP model. The resulting sets of *a priori* parameter values perform poorly at predicting daily runoff, with NS efficiencies below 0.6 in 11 basins and negative in two. Manual calibration of saturated hydraulic conductivity improved model performance nearly as much as optimal calibration of six soil parameters. Gusev and Nasonova, (2006) attribute this to the high sensitivity of runoff to saturated hydraulic conductivity and the high degree of spatial variability of the parameter. They propose deriving a pedotransfer function for effective saturated hydraulic conductivity based on calibration of the parameter. Xie and Yuan, (2006) present the soil and vegetation look-up tables used to apply VIC to the set of 12 French MOPEX basins. For the seven model parameters not in the look-up tables they use calibrated values from a humid basin in China. As with above-noted SWAP results, the performance of the *a priori* parameter values varies widely between basins, with calibration providing the greatest relative improvement for the worst-performing basins.

More conceptual models can make use of pedotransfer functions if a relationship can be inferred between model parameters and soil hydraulic properties. For example, Koren, *et al.*, (2003) describe a methodology for *a priori* estimation of the parameters of the SAC-SMA based on soil properties as extracted from the database of Miller and White, (1998) noted in Section B-3.

Gan and Burges, (2006) compare results for the SAC-SMA parameterized for the MOPEX basins using a combination of automated optimization and manual adjustment versus using the *a priori* methodology of Koren, *et al.*, (2003). They find the two methods produce very different parameter values. For both calibration and validation periods, the calibrated parameter values produce both substantially higher NS efficiencies and lower biases. The authors go on to test the transferability of the calibrated parameter values between basins. They find that the resulting 144 simulations averaged similar biases to the 12 *a priori* simulations and only slightly worse NS efficiencies. They also find that scaling the calibrated values based on the *a priori* values does not significantly improve their transferability. Anderson, *et al.*, (2006) adapts the *a priori* methodology of Koren, *et al.*, (2003) for use with the SSURGO and NLCD databases. They apply

parameter values estimated from the two soils databases to six basins in the Ohio River Basin. Improved streamflow simulations are seen with the SSURGO-based parameters for two of the three basins for which there were significant differences in soil textures between the two databases.

Along with the HL-RMS, PRMS is a distributed modeling system with a comprehensive schema for *a priori* estimation of parameter values. (Leavesley, *et al.*, 2003) describe the distributed topographic, vegetation, LULC and soil datasets and the software tools that are used to estimate parameter values within each HRU.

B-8.3 The North American Land Data Assimilation System

As part of the North American Land Data Assimilation System (NLDAS), a study was undertaken to generate and validate land-surface states and fluxes over the entire conterminous US (CONUS) on a 1/8° grid using four land surface models (LSMs)(Mitchell, *et al.*, 2004). The models were executed at an hourly time step over the period from October 1996 to September 1999, with the first year used for model spin up. The principal differences of the study with past LSM intercomparison studies (e.g., PILPS, GSWP) are its continental-scale coverage and its use of an operational rainfall-runoff model (SAC-SMA) along with two traditional LSPSs (Noah and Mosaic) and the hybrid VIC model. The latter three models have representations of the vegetation canopy that are involved in the calculation of the coupled energy and water balances. The SAC-SMA model on the other hand calculates only the water balance and uses externally estimated values of potential evapotranspiration (PET). Each of the models was essentially applied in a lumped manner to each 1/8° grid cell. However, both Mosaic and VIC account for subgrid variations in vegetation and soils with “tiles.” VIC also makes use of elevation bands for snowmelt calculations.

The NLDAS study differs from DMIP and MOPEX in that internal states of watersheds—in particular snow cover and soil moisture—are examined in addition to streamflow. As well as having three models in common with MOPEX (SAC-SMA, Noah, VIC), the NLDAS study also used *a priori* parameterizations in the models. VIC, Noah and Mosaic used soil and vegetation look-up tables that have been developed from the literature and past applications of the models. Vegetation seasonality and density were based on remote sensing. VIC also relied on a CONUS-wide calibration of some soil and runoff parameters to large river basins that was performed in a previous study over the NLDAS domain (Maurer, *et al.*, 2002). Although that study uses the same daily precipitation data as the NLDAS study, it involved running VIC at a 3-hour time step with daily precipitation evenly distributed across the eight time steps in a day. In the NLDAS study the gauge-based, gridded daily precipitation estimates of (Maurer, *et al.*, 2002) were disaggregated to hourly values using the diurnal distributions in NEXRAD data. Two features of Noah are worth pointing out: (1) as noted in the previous section,

it incorporates the infiltration and runoff algorithms from a parsimonious, conceptual rainfall-runoff model (Schaake, *et al.*, 1996), and (2) it is the LSPS used in the atmospheric model which produced all but precipitation and solar radiation in the NLDAS retrospective forcing data (Cosgrove, *et al.*, 2003). SAC-SMA was used with the *a priori* parameterization scheme of Koren *et al.* (2000), which is an earlier version of that of (Koren, *et al.*, 2003). Estimates of PE used in SAC-SMA were taken from Noah.

Lohmann, *et al.*, (2004) report on the streamflow and water balance results from the NLDAS study. They also describe the common routing models applied to the runoff outputs from each of the four models. Modeled streamflows are compared to observed values for 1145 basins ranging in size from 23 to 10,000 km². Maps of biases in daily streamflow show similar large-scale spatial patterns for SAC-SMA and Mosaic, with underestimation of streamflow in most basins and biases less than -0.6 mostly concentrated in the northern half of the country. VIC produces overestimation biases greater than 0.2 over most of the coastal plains of the Southeast, the Midwest and the Great Plains. The largest concentration of underestimation biases of -0.2 or less is in the Northwest. Noah exhibits similar, although more unevenly distributed biases, with most positive values in the Southeast, the lower Midwest and the Great Plains, and negative values in the Northeast, Appalachian Mountains, upper Midwest and Northwest. The biases are reflected in a comparison of modeled and observed total runoff over the two-year evaluation period averaged by quadrant. In the NE and SE quadrants, SAC-SMA produced less than half the observed runoff, VIC about 50% too much, and Noah close to the observed. In the SE quadrant, SAC-SMA and Mosaic underestimate total runoff by about 20%, VIC overestimates it by about 30%, and Noah is again very close to observed. In the NW quadrant all models underestimate total runoff in the range of 25 to 50%, with VIC the closest to observed. Lohmann, *et al.*, (2004) attribute some of the latter bias to the underestimation of precipitation (particular snowfall) by the interpolation of gauge-measured precipitation over mountains. Pan, *et al.*, (2003) compare the NLDAS precipitation to that measured at 110 SNOTEL sites and find that at all the sites the measured is greater than that for the corresponding NLDAS grid cell. They report that the mean NLDAS values average less than half the Snowpack Telemetry (SNOTEL) values.

Maps of NS efficiencies paint a slightly different picture of model accuracy. All four models show similar distributions, with positive values concentrated in the East and coastal West and negative values in between. It appears that no model produces NS efficiencies greater than 0.5 for more than a third of the basins, with the ranking of decreasing number of basins as follows: SAC-SMA, VIC, Noah, Mosaic. A variety of results presented in Lohmann, *et al.*, (2004) suggest that the models tend to perform the worst in basins where the annual hydrograph is dominated by snowmelt. Results presented in Lohmann, *et al.*, (2004), Pan, *et al.*, (2003) and Sheffield, *et al.*, (2003) show that Noah, Mosaic and SAC-SMA all

predict peaks too early in the snow-dominated basins of the West, with Noah showing peaks months in advance of the observed. It is suggested that the elevation banding in VIC is what allows it to better predict the snowmelt peaks. However, it should also be recognized that only VIC was calibrated to a version of the NLDAS forcing data and that the early peaks in the other models are at least in part due to the lower snowpack that results from the negative winter bias over the mountainous West in the NLDAS precipitation data. The latter conclusion is supported by the much better model results in the Northeast and upper Midwest for snow cover (Sheffield, *et al.*, 2003) and peak spring streamflow (Lohmann, *et al.*, 2004). As a whole, VIC seems to do the best at modeling snow accumulation and melt, and Noah the worst. The temperature-index based snow module in SAC-SMA performs about as well as the energy balance methodology in Mosaic. The poor performance of Noah is attributed to high rates of sublimation and a strongly positive albedo feedback to air temperature from snowmelt. It is also of interest that only Noah simulates soil freezing.

Lohmann, *et al.*, (2004) also analyze the monthly water balance by quadrant. In that analysis they observe that: (1) SAC-SMA produces more surface than subsurface runoff, while the reverse is true for the other three models; and (2) there are large differences in the seasonality of changes in soil moisture storage and evapotranspiration (ET), with SAC-SMA showing the least seasonality and Mosaic the greatest. Regarding the latter observation, they note that SAC-SMA tends to have a high cold season ET, and that Mosaic allows for a high rate of diffusion from lower to upper zone storages. Modeled soil moisture storage is compared against large-scale observations by Schaake, *et al.*, (2004) and Robock, *et al.*, (2003). Schaake, *et al.*, (2004) note that soil depths in Noah and Mosaic were set to 2.0 m in all grid cells, while they vary in both SAC-SMA and VIC according to their *a priori* and calibrated parameter values. They report CONUS-average total water storage capacities of 435, 917, 879 and 618 mm for SAC-SMA, Noah, Mosaic and VIC respectively. Schaake, *et al.*, (2004) computed a semimonthly, statewide average of total soil moisture over a two-meter depth using data from 17 monitoring sites managed by the Illinois State Water Survey. This resulted in 48 “snapshots” of soil moisture over the two year validation period. The authors plot those values against modeled total water storage averaged over the grids encompassing the state of Illinois. They find good agreement between measured and modeled values for SAC-SMA and Noah. However, because of the varying soil depths in SAC-SMA and the fact that it only accounts for “active” storage, the SAC-SMA results had to be adjusted to make them commensurable with the measurements. Schaake, *et al.*, (2004) only note that water storage at wilting point was added to the modeled values. We observe that while the two models may do a good job simulating soil moisture over Illinois, results in Lohmann, *et al.*, (2004) show that Noah tends to considerably overestimate runoff and SAC-SMA considerably underestimate it over most of the state. Mosaic shows a tendency to underestimate soil moisture and runoff, while VIC considerably underestimates soil moisture and significantly overestimates runoff. Mosaic also shows

about a 50% too large seasonal range in soil moisture, in agreement with the seasonal water balance analysis of Lohmann, *et al.*, (2004). Because the VIC parameters for the northwest and southeast halves of Illinois were derived from calibration to runoff from two different large river basins, the VIC water storage capacities are very different for the two halves. VIC underestimates total water storage by about 25% in northwest half and by about 50% in the southeast half. In both halves, VIC approximates well the seasonal range. As with SAC-SMA, it is not clear how the variable soil depths in VIC were made commensurable with the fixed 2-m depth of the observations.

Robock, *et al.*, (2003) compare modeled estimates of daily soil moisture over Oklahoma against statewide averages of volumetric soil moisture in the top 40 cm as estimated from observations at 72 Oklahoma Mesonet stations. Noah is seen to overestimate soil moisture systematically by about 7%. Consistent with the Illinois comparison, Mosaic tends to underestimate soil moisture and overestimate its seasonal variation, with summer soil moisture estimates 30-50% below observed. SAC-SMA tracks the Mosaic results most closely, with somewhat greater interstorm/storm variability. Robock, *et al.*, (2003) note that SAC-SMA models total water storage in an upper zone which has no explicitly assumed depth. Although it is not clear what depth the SAC-SMA upper zone water storage is divided into to calculate volumetric soil moisture, the tendency for the resulting values to underestimate the observations suggests that the assumed depth is too large, while the tendency to overestimate the seasonal and storm/interstorm variability suggests that it is too small. VIC does about as well as Noah in matching observed soil moisture, with a tendency to underestimate it slightly in winter and overestimate slightly in summer. All models overestimate the intra-seasonal variability in observed soil moisture, with Noah and VIC doing so only slightly. Robock, *et al.*, (2003) suggest that errors in modeled soil moisture arise from errors in soil hydraulic parameters. They first identify where the soil texture classes at the Mesonet stations differ from the soil texture classes assigned to the corresponding grid cell in the NLDAS database. However, differences in the values of the soil hydraulic parameters assigned by each of the LSPSs for a given texture class seem to be more significant. In particular, the saturated hydraulic conductivities used by Mosaic are an order of magnitude lower than the ones used by VIC for many of the texture classes, including ones that are commonly found at the Mesonet stations. Both the low saturated hydraulic conductivities and high wilting points assigned by Mosaic would tend to cause it to overestimated soil moisture. However, as noted before, it tends to underestimate soil moisture. It appears then that the high rate of upward diffusion from the lower soil layer in Mosaic more than offsets the differences in soil hydraulic parameters. In summary, we observe that Noah seems to do the best job of simulating observed soil moisture and runoff in Illinois and Oklahoma at both the daily and seasonal time scales.

B-8.4 Other Model Intercomparison Studies

There are many smaller studies sprinkled throughout literature. Reviews of many of them can be found in Smith, *et al.*, (2004c) and Michaud and Sorooshian, (1994). We discuss two that are of interest for their use of MIKE SHE which is the modeling system used in the studies of Refsgaard, (1997) and Christiaens and Feyen, (2002) noted in Section B-6. Refsgaard and Knudsen, (1996) apply MIKE SHE, along with a lumped conceptual model and a semi-distributed model of intermediate complexity, to three river basins in Zimbabwe, 254, 1040 and 1090 km² in size. The models are driven by daily precipitation measured at 5 to 7 rain gauges located adjacent to or within the given basin. Refsgaard and Knudsen, (1996) use the hierarchical scheme for model testing proposed by Klemes, (1986b). Their application of that scheme involves traditional split-sample (SS) calibration and validation, along with a differential split sample (DSS) test in which calibration is performed on a period of above normal flows and validation on a period of below normal flows. They also applied several versions of a proxy-basin (PB) test in which model parameters are transferred from one basin to another with limited or no calibration to the second basin. In each test, three quantitative evaluation criteria were examined: NS efficiencies of modeled monthly flows, an error index for the daily flow duration curves, and the absolute value of the mean bias in daily flows. No model clearly outperformed the others for all basins and all evaluation criteria. For the SS and DSS tests, NAM ranks the highest for four out of the six of criteria/basin combinations. For the PB and combined PB-DSS tests, WATBAL ranks the highest for 10 out of the 12 criteria/basin combinations. Refsgaard and Knudsen, (1996) draw two main conclusions from their study: (1) Given only a few years of runoff measurements, a lumped model of the NAM type is the most suitable, and (2) For ungauged catchments, a distributed model is expected to perform best contingent on the availability of the necessary information on watershed characteristics. Based on their results, it seems the second point applies most to the semi-distributed WATBAL model.

One of the features of MIKE SHE that distinguishes it from the other two models used by Refsgaard and Knudsen, (1996) is its use of a three-dimensional model of the saturated zone. In a semi-arid climate such as that of Zimbabwe, that may not be an advantage for streamflow simulation. For a humid 465 km² basin in Belgium, Abu El-Nasr, *et al.*, (2005) compare the performance of MIKE SHE against SWAT (Arnold, *et al.*, 1998), a semi-distributed model with considerable physical basis. Both models were driven by daily precipitation data from seven rain gauges in and around the basin and calibrated to daily discharge at the outlet of the basin. Groundwater levels at 8 wells within the basin and discharge at an internal stream gauge were also used in the calibration of MIKE SHE. Based on several evaluation criteria applied to daily streamflow, Abu El-Nasr, *et al.*, (2005) conclude that “*both models are able to simulate the hydrology of the catchment in an acceptable way*” with MIKE SHE predicting “*slightly better the*

overall variation of the river flow.” However, their plot comparing modeled to observed discharge curves suggests that the performance of MIKE SHE is more than slightly better. In addition, many of the evaluation statistics for the validation period are substantially better for MIKE SHE than SWAT. For example, the NS efficiency is 0.55 for SWAT and 0.76 for MIKE SHE. Abu El-Nasr, *et al.*, (2005) acknowledge that the ability to perform more detailed modeling of the saturated zone with MIKE SHE may have contributed to its better performance. Although a figure in Abu El-Nasr, *et al.*, (2005) indicates that the groundwater data may have been split for calibration and validation, no groundwater simulation results are presented for MIKE SHE.

B-9 Summary and Conclusions

From the above discussion of model intercomparison studies, it is clear that, to date, conceptual models hold their own against more physically based models, particularly in small well-gauged watersheds in which runoff at the basin outlet is the only predictand of interest. This has clearly been demonstrated to stem from the limited information content in most rainfall-runoff records and the resultant need for only a limited number of degrees of freedom in a model—be it a purely blackbox model (e.g., statistically based models of the AR, ARMA, ARIMA class, neural networks, etc.) or one that is derived from some physical basis (e.g., a storage-based model developed in a top-down manner). But what are the most appropriate models in ungauged basins? Here again, the evidence tilts towards more parsimonious, lumped models. At the same time, the estimation of parameter values in ungauged basins appears to be best accomplished with well-designed, physically based and data-intensive regionalization schemes and thus requires models with a strong physical basis. In particular, models and regionalization schemes that are based on a handful of similarity measures appear to hold the greatest promise.

Despite continued success of lumped and conceptual models, we should not sound the death knell for fully distributed and physically complex models. They have clearly shown their worth in the research arena as means of understanding runoff process over a range of scales and developing simpler models in a bottom-up approach. But what is their value to operations? Although it has been nearly 30 years since Freeze and Harlan (1969) outlined their blueprint for such models, they have seen surprisingly little testing—and to our knowledge no routine use—in operational environments. This can to a certain extent be traced to the computational demands of such models, but that constraint is rapidly being relaxed. Just as the three most important factors to the value of real estate are “location, location and location,” the three most important factors to the value of a distributed model are “data, data and data.” The data necessary to calibrate, drive and update fully distributed and physically complex models has historically been sorely lacking outside of small, well-instrumented and -studied experimental watersheds. However, the availability of distributed datasets, particularly from remote-

sensing platforms, has proliferated along with—and not independent of—the number of models in recent years.

The ever-increasing availability of distributed datasets means that the time for operational, distributed modeling has arrived. Results from Phase 1 of the DMIP have clearly shown that distributed models are capable of accounting for the impact of the large-scale spatial distribution of precipitation and soils on the stream-flow hydrograph at the outlet of operational-scale basins. The results also show the potential for distributed models to capture the hydrograph at smaller interior basins when the model is calibrated to streamflow from larger parent basins. Initial results from Phase 2 of the DMIP, show that interior states of basins, in particular soil moisture and snowpack, can also be well represented in distributed and physically based models. In addition to being able to predict these internal states, their observation may provide additional data for calibrating and updating distributed, physically based models.

Although distributed, physically based modeling is clearly the direction that operational hydrologic forecasting is taking—and should be taking, much research is needed to determine just how distributed and just how physically based such models should be. Ideally, that determination should be based on the needs of forecast users, but in reality, it is often determined by the resolution and quality of data. In particular, the NWS currently produces NEXRAD precipitation data for its hydrologic forecasters at a nominal 4-km resolution. The 4-km resolution of the NEXRAD data is primarily the reason that the NWS DHM has been implemented at the same resolution. Statistically downscaling NEXRAD precipitation data to higher resolutions is an active area of research (e.g., Mascaro, *et al.*, 2006). However, given the large uncertainties in the NEXRAD data, and the fact that a watershed is a natural low-pass filter in which information at high spatial and temporal resolution is considerably filtered out, one might question the value of such exercises. At a minimum, a large (and potentially computational unmanageable) number of ensemble runs of the downscaled precipitation will be necessary to account for those uncertainties. In addition, the increasing accumulated errors that accompany the higher resolution of highly uncertainty precipitation data may outweigh the benefits of modeling runoff dynamics at a higher resolution (Koren, 2007).

As Beven (2002) and others have argued, we should probably sound the death knell for distributed, physically based models based on small-scale physically theory, primarily because we will never have the necessary deterministic measures of the physical properties of watersheds at the scale that the theory holds. The promising alternative blueprint appears to be what we have described as the distributional variety of semi-distributed models. TOPMODEL was the first of this generation of models. However, TOPMODEL still ignores small-scale heterogeneities such as preferential flow paths, among other very constricting simplifying assumptions. What is needed is a whole suite of distributions of spatial het-

erogeneities and similarity measures beyond the soils-topographic index. We believe the greatest progress and promise towards that end is being made under the framework of the representative elementary watershed (REW) discussed in Sections B-7 and B-8. Rather than pursuing a futile effort to reduce irreducible uncertainties, the REW framework has the potential to quantify them in a relatively parsimonious and physically meaningful way.

In summary, the four greatest needs for realizing the long sought operational benefits for distributed, physical based rainfall-runoff modeling for operational purposes are:

1. parsimonious models that capture hydrologic functioning at the HRU or REW scale;
2. physically based methods for *a priori* estimation of distributed parameter values;
3. ways of using soft data and observations of internal states (particularly snowpack, soil moisture and groundwater) to calibrate and fine-tune distributed parameter values, and to update and validate modeled values of distributed internal states; and
4. methods of uncertainty analysis that are compatible with (1)-(3).

B-10 References

- Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O’Connell and J.Rasmussen. (1986a) “An introduction to the European Hydrological System – Systeme Hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modeling system”. *J. Hydrol.* 17: 45-59.
- Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O’Connell and J.Rasmussen (1986b). “An. introduction to the European Hydrological System – Systeme Hydrologique Europeen, “SHE”, 2: Structure of a physically-based, distributed modelling system”. *J. Hydrol.* 17: 61-77.
- Abu El-Nasr, A., JG Arnold, J Feyen and J Berlamont (2005), Modelling the hydrology of a catchment using a distributed and a semi-distributed model, *Hydrological Processes*, 19, 573-587.
- Anderson, R. M., VI Koren and SM Reed. (2006), Using SSURGO data to improve Sacramento Model *a priori* parameter estimates, *J. of Hydrol.*, 320, 103-116.
- Andréassian, V. (2004), Waters and forests: from historical controversy to scientific debate, *J. of Hydrol.*, 291, 1-27.

- Andréassian, V., S. Bergström, N. Chahinian, Q. Duan, Y. M. Gusev M., I. Littlewood, T. Mathevet, C. Michel, A. Montanari, G. Moretti, R. Moussa, O. N. Nasonova, K. O'Connor, E. Paquet, C. Perrin, A. Rousseau, J. Schaake, T. Wagener and Z. Xie (2006a), Catalogue of the models used in MOPEX 2004/2005, in Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX, edited by V. Andréassian, *et al.*, pp. 41-107, IAHS Press, Wallingford.
- Andréassian, V., A Hall, N Chahinian and J Schaake (Eds.) (2006b), Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX, 347 pp., IAHS Press, Wallingford.
- Arnold, J. G. and N. Fohrer (2005), SWAT2000: current capabilities and research opportunities in applied watershed modelling, *Hydrological Processes*, 19, 563-572.
- Arnold, J. G., *et al.* (1998), Large area hydrologic modeling and assessment - Part 1: Model development, *J. Am. Water Res Assoc.*, 34, 73-89.
- Bandaragoda, C., DG Tarboton and R Woods. (2004), Application of TOPNET in the distributed model intercomparison project, *J. of Hydrol.*, 298, 178-201.
- Barlow, P. M., D.P. Ahlfeld and D.C. Dickerman (2003), Conjunctive-management models for sustained yield of stream-aquifer systems, *Journal of Water Resources Planning and Management-ASCE*, 129, 35-48.
- Beven, K. (1989), Changing Ideas in Hydrology - the Case of Physically-Based Models, *J. of Hydrol.*, 105, 157-172.
- Beven, K. (2002), Towards an alternative blueprint for a physically based digitally simulated hydrologic response modelling system, *Hydrological Processes*, 16, 189-206.
- Beven, K. (2006a), A manifesto for the equifinality thesis, *J. of Hydrol.*, 320, 18-36.
- Beven, K. (2006b), Searching for the Holy Grail of scientific hydrology: $Q(t) = H((S)\text{under-left-arrow}, (R)\text{under-left-arrow}, \Delta t)A$ as closure, *Hydrology and Earth System Sciences*, 10, 609-618.
- Beven, K. and A. Binley (1992), The Future of Distributed Models - Model Calibration and Uncertainty Prediction, *Hydrological Processes*, 6, 279-298.

- Beven, K. and J. Feyen (2002), The future of distributed modelling - Special issue, *Hydrological Processes*, 16, 169-172.
- Beven, K. and M. J. Kirkby (1979), A physically based, variable contributing area model of basin hydrology, *Hydrology Science Bulletin*, 24, 43-69.
- Beven, K. J. (2000), Rainfall-runoff modelling : the primer, xi, 360 p. pp., Wiley, Chichester ; New York.
- Bicknell, B., J. Imhoff, J. Kittle Jr and A. Donigian Jr (2000), Hydrological Simulation Program-Fortran (HSPF): User Manual for Release 12.
- Blöschl, G. (Ed.) (2003), Water resources systems, X, 366 S. pp., IAHS Press, Wallingford.
- Bloschl, G. and M. Sivapalan (1995), Scale Issues in Hydrological Modeling - a Review, *Hydrological Processes*, 9, 251-290.
- Campling, P., A. Gobin, K. Beven and J. Feyen (2002), Rainfall-runoff modelling of a humid tropical catchment: the TOPMODEL approach, *Hydrological Processes*, 16, 231-253.
- Carpenter, T. M. and K. P. Georgakakos (2004), Continuous streamflow simulation with the HRCDHM distributed hydrologic model, *J. of Hydrol.*, 298, 61-79.
- Carroll, S. S., T.R. Carroll and R.W. Poston (1999), Spatial modeling and prediction of snow-water equivalent using ground-based, airborne, and satellite snow data, *Journal of Geophysical Research-Atmospheres*, 104, 19623-19629.
- Chahinian, N., V. Andréassian, Q. Duan, V. Fortin, H. Gupta, T. Hogue, T. Mahtevet, A. Montanari, G. Moretti, R. Moussa, C. Perrin, J. Schaake, T. Wagener and Z. Xie. (2006a), Compilation of the MOPEX 2004 results, in "Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment – MOPEX," edited by V. Andréassian, *et al.*, pp. 313-346, IAHS Press, Wallingford.
- Chahinian, N., T. Mathevet, F. Habets and V. Andreassian (2006b), The MOPEX 2004 French database: main hydrological and morphological characteristics, in "Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX," edited by V. Andréassian, *et al.*, pp. 29-40, IAHS Press, Wallingford.

- Christiaens, K. and J. Feyen (2002), Constraining soil hydraulic parameter and output uncertainty of the distributed hydrological MIKE SHE model using the GLUE framework, *Hydrological Processes*, 16, 373-391.
- Ciarapica, L. and E. Todini (2002), TOPKAPI: a model for the representation of the rainfall-runoff process at different scales, *Hydrological Processes*, 16, 207-229.
- Cosgrove, B. A., D. Lohman, K. Mitchell, D. Lohman, P. Houser, E.F. Wood, J.C. Schaake, A. Robock, C. Marshall, J. Sheffield, Q. Duan, L. Luo, R. W. Higgins, R.T. Pinker, J.D. Tarpley and J. Meng (2003), Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project, *Journal of Geophysical Research-Atmospheres*, 108, -.
- Croke, B. F. W., W.S. Merritt and A.J. Jakeman (2004), A dynamic model for predicting hydrologic response to land cover changes in gauged and ungauged catchments, *J. of Hydrol.*, 291, 115-131.
- De Roo, A. P. J., C. G. Wesseling and W. P. A. Van Deursen. (2000), Physically based river basin modelling within a GIS: the LISFLOOD model, *Hydrological Processes*, 14, 1981-1992.
- Di Luzio, M. and J. G. Arnold (2004), Formulation of a hybrid calibration approach for a physically based distributed model with NEXRAD data input, *J. of Hydrol.*, 298, 136-154.
- Donner, S. D., C. J. Kucharik and J. A. Foley (2004), Impact of changing land use practices on nitrate export by the Mississippi River, *Global Biogeochemical Cycles*, 18, -.
- Dougherty, M., R. L. Dymond, T. J. Grizzard, Jr. A. N. Godrej, C. E. Zipper and J. Randolph (2007), Quantifying long-term hydrologic response in an urbanizing basin, *Journal of Hydrologic Engineering*, 12, 33-41.
- Downer, C. W. and F. L. Ogden (2004), GSSHA: Model to simulate diverse stream flow producing processes, *Journal of Hydrologic Engineering*, 9, 161-174.
- Duan, Q., J. C. Schaake, V. Andréassian, S. Franks, G. Goteti, H. V. Gupta, Y. M. Gusev, F. Habets, A. Hall, L. Hay, M. Huang, G. Leavesley, X. Liang, O. N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener and E. F. Wood (2006), Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, *J. of Hydrol.*, 320, 3-17.

- Ebel, B. A. and K. Loague (2006), Physics-based hydrologic-response simulation: Seeing through the fog of equifinality, *Hydrological Processes*, 20, 2887-2900.
- Elvidge, C. D., C. Milesi, J. B. Dietz, B. T. Tuttle, P. C. Sutton, R. Nemani, and J. E. Vogelmann (2004), US Constructed Area Approaches the Size of Ohio, *EOS, Trans., AGU*, 85, 233-233.
- Entekhabi, D. and P. S. Eagleson (1991), Climate and the Equilibrium State of Land Surface Hydrology Parameterizations, *Surveys in Geophysics*, 12, 205-220.
- Farmer, D., M Sivapalan and C Jothityangkoon (2003), Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: Downward approach to water balance analysis, *Water Resources Research*, 39, -.
- Farr, T.G., *et al.* (2007), The Shuttle Radar Topography Mission, *Reviews of Geophysics*, 45, RG2004.
- Feldman, A. D. (2000), Hydrologic Modeling System HEC-HMS: Technical Reference Manual, US Army Corps of Engineers, Hydrologic Engineering Center.
- Franks, S. W. M. Sivapalan, K. Takeuchi and Y Tachikawa (2005), Predictions in ungauged basins : international perspectives on the state of the art and pathways forward, xi, 348 p. pp., Oxfordshire, UK, IAHS, Wallingford.
- Freer, J., J. J. McDonnell, K. J. Beven, N. E. Peters, D. A. Burns, R. P. Hooper, B. Aulenbach and C. Kendall (2002), The role of bedrock topography on subsurface storm flow, *Water Resources Research*, 38, -.
- Freer, J. E., H. McMillan, J. J. McDonnell and K. J. Beven (2004), Constraining dynamic TOPMODEL responses for imprecise water table information using fuzzy rule based performance measures, *J. of Hydrol.*, 291, 254-277.
- Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, A. Cooper et al. (2002). Global land cover mapping from MODIS: algorithms and early results. *Remote Sensing of Environment*, 83(1-2), 287-302.
- Freeze, R. A. and R. L. Harlan (1969), Blueprint for a physically-based, digitally-simulated hydrologic response model, *J. of Hydrol.*, 9, 237-258.
- Gan, T. Y. and S. J. Burges (1990), An Assessment of a Conceptual Rainfall-Runoff Models Ability to Represent the Dynamics of Small Hypothetical

Catchments .2. Hydrologic Responses for Normal and Extreme Rainfall, *Water Resources Research*, 26, 1605-1619.

Gan, T. Y. and S. J. Burges (2006), Assessment of soil-based and calibrated parameters of the Sacramento model and parameter transferability, *J. of Hydrol.*, 320, 117-131.

Ghazi, A. (2005), Advances in flood forecasting - Foreword, *Hydrology and Earth System Sciences*, 9, 279-279.

Goetz, S. J. and P. Jantz (2006), Satellite Maps Show Chesapeake Bay Urban Development, *EOS, Trans., AGU*, 87, 149-152.

Grassotti, C., R. N. Hoffman, E. R. Vivoni and D. Entekhabi (2003), Multiple-timescale intercomparison of two radar products and rain gauge observations over the Arkansas-Red River basin, *Weather and Forecasting*, 18, 1207-1229.

Grayson, R. and G. Bloschl (Eds.) (2000a), Spatial patterns in catchment hydrology : observations and modelling, xii, 404 p. : ill. (some col.), maps (some col.) ; 426 cm. pp., Cambridge University Press.

Grayson, R. B. and G. Bloschl (2000b), summary of pattern comparisons and concluding remarks, in spatial patterns in catchment hydrology, edited by R. B. Grayson and G. Bloschl, pp. 355-367, Cambridge University Press, Cambridge.

Grayson, R. B., I. D. Moore and T. A. McMahon (1992), Physically Based Hydrologic Modeling .2. Is the Concept Realistic?, *Water Resources Research*, 28, 2659-2666.

Grimes, D. I. F., E. Pardo-Igúzquiza and R. Bonifacio. (1999), Optimal areal rainfall estimation using raingauges and satellite data, *J. of Hydrol.*, 222, 93-108.

Gusev, Y. M. and O. N. Nasonova (2006), Simulating runoff from MOPEX experimental river basins using the SWAP land surface model, in Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX, edited by V. Andréassian, *et al.*, pp. 188-195, IAHS Press, Wallingford.

Han, S. C., C. K. Shum, C. Jekeli and D. Alsdorf (2005), Improved estimation of terrestrial water storage changes from GRACE, *Geophys Res Lett*, 32, -.

- Harte, J. (2002), Toward a synthesis of the Newtonian and Darwinian world-views, *Physics Today*, 55, 29-34.
- Homer, C., C. Huang, L. Yang, B. Wylie and M. Coan (2004), Development of a 2001 National Land-Cover Database for the United States, *Photogrammetric Engineering and Remote Sensing*, 70, 829-840.
- Hundecha, Y. and A. Bardossy (2004), Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. of Hydrol.*, 292, 281-295.
- Ivanov, V. Y., E. R. Vivoni, R. L. Bras and D. Entekhabi (2004a), Catchment hydrologic response with a fully distributed triangulated irregular network model, *Water Resources Research*, 40, -.
- Ivanov, V. Y., E. R. Vivoni, R. L. Bras and D. Entekhabi. (2004b), Preserving high-resolution surface and rainfall data in operational-scale basin hydrology: a fully-distributed physically-based approach, *J. of Hydrol.*, 298, 80-111.
- Jakeman, A. J. and G. M. Hornberger (1993), How Much Complexity Is Warranted in a Rainfall-Runoff Model, *Water Resources Research*, 29, 2637-2649.
- Jiang, L., J. D. Tarpley, K. E. Mitchell, S. Zhou, F. N. Kogan, W. Guo (2008). Adjusting for long-term anomalous trends in NOAA's global vegetation index data sets. *IEEE Transactions on Geoscience and Remote Sensing*, 46(2).
- Julien, P. Y., B. Saghaian and F. L. Ogden (1995), Raster-Based Hydrologic Modeling of Spatially-Variied Surface Runoff, *Water Resources Bulletin*, 31, 523-536.
- Kavetski, D., G. Kuczera and S. W. Franks (2006a), Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory, *Water Resources Research*, 42, -.
- Kavetski, D., G. Kuczera and S. W. Franks (2006b), Bayesian analysis of input uncertainty in hydrological modeling: 2. Application, *Water Resources Research*, 42, -.
- Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resources Research*, 42, -.
- Klemes, V. (1983), Conceptualization and scale in hydrology, *J. of Hydrol.*, 65, 1-23.

- Klemes, V. (1986a), Dilettantism in Hydrology - Transition or Destiny, *Water Resources Research*, 22, S177-S188.
- Klemes, V. (1986b), Operational Testing of Hydrological Simulation-Models, *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 31, 13-24.
- Koren, V., S. Reed, M. Smith, Z. Zhang and D. J. Seo (2004), Hydrology Laboratory Research Modeling System (HL-RMS) of the US National Weather Service, *J. of Hydrol.*, 291, 297-318.
- Koren, V., M. Smith and Q. Y. Duan (2003), Use of *a priori* parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models, in "Calibration of Watershed Models Water Science and Applications," edited by Q. Duan, *et al.*, pp. 239-254, *AGU*.
- Kouwen, N. (2002), WATFLOOD/SPL9 Hydrological Model & Flood Forecasting System, in Mathematical models of large watershed hydrology, edited by V. Singh and D. K. Frevert, pp. 649-685, Water Resources Publications, LLC, Highlands, CO.
- Kouwen, N., E. D. Soulis, A. Pietroniro, J. Donald and R. A. Harrington (1993), Grouped Response Units for Distributed Hydrologic Modeling, *Journal of Water Resources Planning and Management-ASCE*, 119, 289-305.
- Kuczera, G. and S. Franks (2002), Testing hydrologic models: fortification or falsification?, in Mathematical Models of Large Watershed Hydrology, edited by V. P. Singh and D. K. Frevert, pp. 141-181, Water Resources Publications, LLC, Highlands Ranch, CO.
- Lamb, R., K. Beven and S. Myrabo (1997), Discharge and water table predictions using a generalized TOPMODEL formulation, *Hydrological Processes*, 11, 1145-1167.
- Leavesley, G. H., L. E. Hay, R. J. Viger and S. L. Markstrom (2003), Use of *a priori* parameter-estimation methods to constrain calibration of distributed-parameter models, in Calibration of Watershed Models Water Science and Applications, edited by Q. Duan, *et al.*, pp. 255-266, *AGU*.
- Leavesley, G. H., R. W Lichty, B. M. Troutman and L. G. Saindon (1983), Precipitation-Runoff Modeling System: User's Manual, 207 pp., USGS Water Resources Investigations Report 83-4238.
- Lee, H., E. Zehe and M. Sivapalan (2007), Predictions of rainfall-runoff response and soil moisture dynamics in a microscale catchment using the CREW model, *Hydrology and Earth System Sciences*, 11, 819-849.

- Liang, X., D. P. Lettenmaier, E. F. Wood and S. J. Burges (1994), A Simple Hydrologically Based Model of Land-Surface Water and Energy Fluxes for General-Circulation Models, *Journal of Geophysical Research-Atmospheres*, 99, 14415-14428.
- Liang, X., Z. Xie and M. Huang (2003), A new parameterization for surface and groundwater interactions and its impact on water budgets with the variable infiltration capacity (VIC) land surface model, *Journal of Geophysical Research-Atmospheres*, 108, -.
- Loague, K., C. S. Heppner, B. B. Mirus, B. A. Ebel, Q. Ran, A. E. Carr, S. H. BeVillie and J. E. VanderKwaak (2006), Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology, *Hydrological Processes*, 20, 1231-1237.
- Loague, K., and J. E. VanderKwaak (2004), Physics-based hydrologic response simulation: platinum bridge, 1958 Edsel, or useful tool, *Hydrological Processes*, 18, 2949-2956.
- Lohmann, D., K. E. Mitchell, P. R. Houser, E. F. Wood, J. C. Schaake, A. Robock, B. A. Cosgrove, J. Sheffield, Q. Duan, L. Luo, W. Higgins, R. T. Pinker and J. D. Tarpley (2004), Streamflow and water balance intercomparisons of four land surface models in the North American Land Data Assimilation System project, *Journal of Geophysical Research-Atmospheres*, 109
- Mascaro, G., G. Mascaro, E. R. Vivoni and R. Deidda (2006), Evaluation of Uncertainty in Nested Flood Forecasts by Coupling a Multifractal Precipitation Downscaling Model and a Fully-Distributed Hydrological Model, in AGU Fall Meeting, edited.
- Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *Journal of Climate*, 15, 3237-3251.
- Mazi, K, A. D. Koussis, P. J. Restrepo and D. Koutsoyiannis (2003), A groundwater-based, objective-heuristic parameter optimisation method for a precipitation-runoff model and its application to a semi-arid basin, *J. of Hydrol.*, 290, 243-258.
- McMichael, C. E., A. S. Hope and H. A. Loaiciga (2006), Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation, *J. of Hydrol.*, 317, 307-324.

- Merz, R., G. Blöschl and J. Parajka. (2006), Regionalization methods in rainfall-runoff modelling using large catchment samples, in *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX*, edited by V. Andréassian, *et al.*, pp. 117-125, IAHS Press, Wallingford.
- Mesinger, F., G. DiMego, E. Kalnay, P. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, K. Mitchell, E. Rogers, M. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish and W. Shi (2006), North American Regional Reanalysis, *Bulletin of the American Meteorological Society*, 87, 343-360.
- Michaud, J. and S. Sorooshian (1994), Comparison of Simple Versus Complex Distributed Runoff Models on a Midsized Semiarid Watershed, *Water Resources Research*, 30, 593-605.
- Michel, C., C. Perrin, V. Andréassian, L. Oudin and T. Mathevet (2006), Has basin-scale modelling advanced beyond empiricism?, in *Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX*, edited by V. Andréassian, *et al.*, pp. 108-116, IAHS Press, Wallingford.
- Miller, D. A. and R. A. White (1998), A Conterminous United States Multilayer Soil Characteristics Dataset for Regional Climate and Hydrology Modeling, edited, pp. 1-26.
- Mitchell, K. E., et al. (2004), The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research-Atmospheres*, 109, -.
- Montanari, A. and S. Uhlenbrook (2004), Catchment modelling: towards an improved representation of the *Hydrological Processes* in real-world model applications, *J. of Hydrol.*, 291, 159-159.
- Moran, M. S., C. D. Peters-Lidard, J. M. Watts and S. McElroy (2004), Estimating soil moisture at the watershed scale with satellite-based radar and land surface models, *Canadian Journal of Remote Sensing*, 30, 805-826.
- Niswonger, R. G., Niswonger, R.G., S.L. Markstrom, R.S. Regan, D.E. Prudic, G. Pohl and R.J. Viger. (2006), Modeling ground-water/surface-water interaction with GSFLOW—A new USGS model, in “MODFLOW and more,” Conference Proceedings 99–103. Golden, Colorado: International Ground Water Modeling Center.

- O'Connell, P. E. (1991), A historical perspective, in *Recent Advances in the Modelling of Hydrologic Systems*, edited by D. S. Bowles and P. E. O'Connell, pp. 3-30, Kluwer, Dordrecht.
- Olson, D. A., N. W. Junker, B. Korty (1995), Evaluation of 33 Years of Quantitative Precipitation Forecasting at the NMC, *Weather and Forecasting*, 10, 498-511.
- Oreskes, N., K. Shrader-Frechette, K. Belitz (1994), Verification, Validation, and Confirmation of Numerical-Models in the Earth-Sciences, *Science*, 263, 641-646.
- Pan, M., *et al.* (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent, *Journal of Geophysical Research-Atmospheres*, 108, -.
- Pappenberger, F. and K. J. Beven (2006), Ignorance is bliss: Or seven reasons not to use uncertainty analysis, *Water Resources Research*, 42, -.
- Pappenberger, F., K. J. Beven, N. M. Hunter, P. D. Bates, B. T. Gouweleeuw, J. Thielen and A. P. J. de Roo (2005), Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), *Hydrology and Earth System Sciences*, 9, 381-393.
- Quinn, P. (2004), Scale appropriate modelling: representing cause-and-effect relationships in nitrate pollution at the catchment scale for the purpose of catchment scale planning, *J. of Hydrol.*, 291, 197-217.
- Reed, S., *et al.* (2004), Overall distributed model intercomparison project results, *J. of Hydrol.*, 298, 27-60.
- Refsgaard, J. C. (1997), Parameterisation, calibration and validation of distributed hydrological models, *J. of Hydrol.*, 198, 69-97.
- Refsgaard, J. C. (2000), Towards a formal approach to calibration and validation of models using spatial data, in *spatial patterns in catchment hydrology*, edited by R. B. Grayson and G. Bloschl, pp. 329-354, Cambridge University Press, Cambridge.
- Refsgaard, J. C. and H. J. Henriksen (2004), Modelling guidelines - terminology and guiding principles, *Advances in Water Resources*, 27, 71-82.
- Refsgaard, J. C. and J. Knudsen (1996), Operational validation and intercomparison of different types of hydrological models, *Water Resources Research*, 32, 2189-2202.

- Reggiani, P., S. M. Hassanizadeh, M. Sivapalan and W. G. Gray (1999), A unifying framework for watershed thermo-dynamics: constitutive relationships, *Advances in Water Resources*, 23, 15-39.
- Reggiani, P., M. Sivapalan and S. M. Hassanizadeh (1998), A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics, *Advances in Water Resources*, 22, 367-398.
- Robichaud, P. R. and H. Elsenbeer (2001), Special issue: Wildfire and Surficial Processes - Preface, *Hydrological Processes*, 15, 2865-2866.
- Robock, A., *et al.* (2003), Evaluation of the North American Land Data Assimilation System over the southern Great Plains during the warm season, *Journal of Geophysical Research-Atmospheres*, 108, -.
- Said, A., D. K. Stevens and G. Sehlke (2005), Estimating water budget in a regional aquifer using HSPF-modflow integrated model, *J. Am. Water Res Assoc.*, 41, 55-66.
- Scanlon, T. M., G. Kiely and Q. Xie (2004), A nested catchment approach for defining the hydrological controls on non-point phosphorus transport, *J. of Hydrol.*, 291, 218-231.
- Schaake, J., Q. Duan, V. Andréassian, S. Franks, A. Halle and G. Leavesley (2006), The model parameter estimation experiment (MOPEX) - Preface, *J. of Hydrol.*, 320, 1-2.
- Schaake, J. C., *et al.* (2004), An intercomparison of soil moisture fields in the North American land data assimilation system (NLDAS), *Journal of Geophysical Research-Atmospheres*, 109, -.
- Schaake, J. C., V. I. Koren, Q. Y. Duan, K. Mitchell and F. Chen (1996), Simple water balance model for estimating runoff at different spatial and temporal scales, *Journal of Geophysical Research-Atmospheres*, 101, 7461-7475.
- Schertzer, D., P. Hubert, S. Koide and K. Takeuchi (Eds.) (2007), Prediction in Ungauged Basins: PUB Kick-Off 20-22 November 2002, International Association of Hydrological Sciences.
- Seibert, J., K. H. Bishop and L. Nyberg (1997), A test of TOPMODEL's ability to predict spatially distributed groundwater levels, *Hydrological Processes*, 11, 1131-1144.

- Seibert, J. and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resources Research*, 38, -.
- Seo, D. J. and J. P. Breidenbach (2002), Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements, *Journal of Hydrometeorology*, 3, 93-111.
- Sheffield, J., G. Goteti and E. F. Wood (2006), Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of Climate*, 19, 3088-3111.
- Sheffield, J., *et al.* (2003), Snow process modeling in the North American Land Data Assimilation System (NLDAS): 1. Evaluation of model-simulated snow cover extent, *Journal of Geophysical Research-Atmospheres*, 108, -.
- Singh, V. P. (Ed.) (1995), Computer models of watershed hydrology, Rev. ed., xiv, 1130 pp., Water Resources Publications, Highlands Ranch, Colo.
- Singh, V. P. and D. F. Frevert (Eds.) (2002a), Mathematical models of large watershed hydrology, xxii, 891 pp., Water Resources Publications, Highlands Ranch, Colo.
- Singh, V. P. and D. F. Frevert (Eds.) (2002b), Mathematical models of small watershed hydrology and applications, 972 pp., Water Resources Publications, Highlands Ranch, Colo.
- Singh, V. P. and D. K. Frevert (2006a), Watershed models, xxiv, 653 p. pp., Taylor & Francis, Boca Raton.
- Singh, V. P. and D. K. Frevert (Eds.) (2006b), Watershed models, xxiv, 653 pp., CRC/Taylor & Francis, Boca Raton, FL.
- Singh, V. P. and D. A. Woolhiser (2002), Mathematical modeling of watershed hydrology, *Journal of Hydrologic Engineering*, 7, 270-292.
- Sivapalan, M. (2003a), Prediction in ungauged basins: a grand challenge for theoretical hydrology, *Hydrological Processes*, 17, 3163-3170.
- Sivapalan, M. (2003b), Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection?, *Hydrological Processes*, 17, 1037-1041.
- Sivapalan, M. (2006), Predictions in ungauged basins : promise and progress, viii, 520 p. pp., International Association of Hydrological Sciences, Wallingford, UK.

- Sivapalan, M., K. Beven and E. F. Wood (1987), On Hydrologic Similarity .2. A Scaled Model of Storm Runoff Production, *Water Resources Research*, 23, 2266-2278.
- Sivapalan, M., G Bloeschl, L Zhang, R Vertessy (2003a), Downward approach to hydrological prediction, *Hydrological Processes*, 17, 2101-2111.
- Sivapalan, M., L. Zhang, R. Vertessy and G. Blöschl (2003b), Preface - Downward approach to hydrological prediction, *Hydrological Processes*, 17, 2099-2099.
- Smith, J. A., M.L. Baeck, J.E. Morrison, P. Sturdevant-Rees, D.F. Turner-Gillespie, and P.D. Bates (2002), The regional hydrology of extreme floods in an urbanizing drainage basin, *Journal of Hydrometeorology*, 3, 267-282.
- Smith, M. B., K. P. Georgakakos and X. Liang (2004a), The distributed model intercomparison project (DMIP), *J. of Hydrol.*, 298, 1-3.
- Smith, M. B., V. I. Koren, Z. Zhang, S. M. Reed, J.-J. Pan and F. Moreda (2004b), Runoff response to spatial variability in precipitation: an analysis of observed data, *J. of Hydrol.*, 298, 267-286.
- Smith, M. B., D. J. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda and S. Cong (2004c), The distributed model intercomparison project (DMIP): motivation and experiment design, *J. of Hydrol.*, 298, 4-26.
- 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.
- Tachikawa, Y. (2003), Weather radar information and distributed hydrological modelling : proceedings of an international symposium (Symposium HS03) held during IUGG 2003, the XXIII General Assembly of the International Union of Geodesy and Geophysics : at Sapporo, Japan, from 30 June to 11 July, 2003, ix, 323 p. pp., International Association of Hydrological Science, Wallingford, Oxfordshire, UK.
- Tarboton, D. G. (1997), A new method for the determination of flow directions and upslope areas in grid digital elevation models, *Water Resources Research*, 33, 309-319.
- Thiemann, M., M. Trosset, H. Gupta and S. Sorooshian (2001), Bayesian recursive parameter estimation for hydrologic models, *Water Resources Research*, 37, 2521-2535.

- Todini, E. (2007), Hydrological catchment modelling: past, present and future, *Hydrology and Earth System Sciences*, 11, 468-482.
- Tromp-van Meerveld, H. J. and J. J. McDonnell (2006), Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, *Water Resources Research*, 42, -.
- Uhlenbrook, S., J. Didszun, N. Tilch, J. McDonnell and Kevin Mcguire (2007), Breaking up is always difficult - landscape discretization as a process-transfer approach, in PUB Kick-Off Meeting held in Brasilia, 20-22 November 2002IAHS publication; no. 309., edited by D. Schertzer, *et al.*, pp. 102-109, International Association of Hydrological Sciences.
- Uhlenbrook, S., J. McDonnell and C. Leibundgut (2003), Preface: Runoff generation and implications for river basin modelling, *Hydrological Processes*, 17, 197-198.
- Uhlenbrook, S., S. Roser and N. Tilch (2004), Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model, *J. of Hydrol.*, 291, 278-296.
- USDA/NRCS (1994), State soil geographic (STATSGO) data base : data use information, ii, 35, A-36, B-13, C-49, D-34 p. pp., U.S. Dept. of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, [Washington, D.C.].
- Vache, K. B. and J. J. McDonnell (2006), A process-based rejectionist framework for evaluating catchment runoff model structure, *Water Resources Research*, 42, -.
- VanderKwaak, J. E. and K. Loague (2001), Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, *Water Resources Research*, 37, 999-1013.
- Vieux, B. E., Z. Cui and A. Gaur (2004), Evaluation of a physics-based distributed hydrologic model for flood forecasting, *J. of Hydrol.*, 298, 155-177.
- Visser, A., R. Stuurman and M. F. P. Bierkens (2006), Real-time forecasting of water table depth and soil moisture profiles, *Advances in Water Resources*, 29, 692-706.
- Vivoni, E. R., V. Y. Ivanov, R. L. Bras and D. Entekhabi (2004), Generation of triangulated irregular networks based on hydrological similarity, *Journal of Hydrologic Engineering*, 9, 288-302.

- Vogel, R. M. (2005), Regional Calibration of Watershed Models, Chapter 3, in Watershed Models,, edited by V. P. Singh and D. F. Frevert, pp. 47-71, CRC Press.
- Vrugt, J. A., H. V. Gupta, W. Bouten and S. Sorooshian (2003), A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research*, 39, -.
- Wagener, T. (2003), Evaluation of catchment models, *Hydrological Processes*, 17, 3375-3378.
- Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrological Processes*, 17, 455-476.
- Wagener, T., H. S. Wheater and H. V. Gupta (2004), Modelling Ungauged Catchments - Regional Procedures, in "Rainfall-Runoff Modelling in Gauged and Ungauged Catchments," edited, pp. 169-240, Imperial College Press, London.
- Western, A. W., R. B. Grayson and T. R. Green (1999), The Tarrawarra project: high resolution spatial measurement, modelling and analysis of soil moisture and hydrological response, *Hydrological Processes*, 13, 633-652.
- Westrick, K. J., P. Storck, C. F. Mass (2002), Description and Evaluation of a Hydrometeorological Forecast System for Mountainous Watersheds, edited, pp. 250-262.
- Wood, A. W., E. P. Maurer, A. Kumar and D. P. Lettenmaier (2002), Long-range experimental hydrologic forecasting for the eastern United States, *Journal of Geophysical Research-Atmospheres*, 107, -.
- Wood, E. F., D. P. Lettenmaier and V. Zartarian (1992), A Land-Surface Hydrology Parameterization with Subgrid Variability for General-Circulation Models, *Journal of Geophysical Research-Atmospheres*, 97, 2717-2728.
- Woods, R. (2002), Seeing catchments with new eyes. Spatial patterns in catchment hydrology: observations and modelling *Hydrological Processes*, 16, 1111-1113.
- Xie, Z. and F. Yuan (2006), A parameter estimation scheme of the land surface model VIC using the MOPEX database, in Large Sample Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter Experiment - MOPEX, edited by V. Andréassian, *et al.*, pp. 169-179, IAHS Press, Wallingford.

- Young, C. B., A. A. Bradley, W. F. Krajewski, A. Kruger and M. Mark L. Morrissey (2000), Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting, *Journal of Hydrometeorology*, 1, 241-254.
- Young, P. (2003), Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale, *Hydrological Processes*, 17, 2195-2217.
- Young, P. C. (1983), The validity and credibility of models for badly defined systems, in *Uncertainty and Forecasting of Water Quality*, edited by M. B. Beck and G. Van Straten, pp. 69-100, Springer-Verlag, Berlin.
- Yu, Z., E. J. Barron, B. Yarnal, M. N. Lakhtakia, R. A. White, D. Pollard and D. A. Miller (2002), Evaluation of basin-scale hydrologic response to a multi-storm simulation, *J. of Hydrol.*, 257, 212-225.
- Zehe, E. and M. Sivapalan (2007), Towards a new generation of hydrological process models for the meso-scale: an introduction, *Hydrology and Earth System Sciences*.