Arnold Schwarzenegger
Governor

EVALUATION OF ALTERNATIVE MODELS AND METHODS FOR PREDICTION OF HYDROPOWER RESOURCES IN CALIFORNIA AND THE PACIFIC NORTHWEST

PIER FINAL PROJECT REPORT

Prepared For:
California Energy Commission
Public Interest Energy Research Program

Prepared By:
Department of Civil and Environmental Engineering
University of Washington

March 2008
CEC-500-2007-104

Prepared By:
University of Washington
Department of Civil and Environmental Engineering
Dennis P. Lettenmaier, Alan F. Hamlet, Tazebe Beyene, and Se-Yeun Lee
University of Washington, Box 352700
Seattle, WA 98195

Commission Contract No. 500-02-004-02
Commission Work Authorization MR-043

Prepared For:
Public Interest Energy Research (PIER) Program
California Energy Commission

Gina Barkalow
Project Manager

Beth Chambers
Contract Manager

Kelly Birkinshaw
Program Area Lead
Energy-Related Environmental Research

Ken Koyama
Acting Office Manager
Energy System Research

Martha Krebs
Deputy Director
ENERGY RESEARCH & DEVELOPMENT DIVISION

Melissa Jones
Executive Director

Acknowledgments

The authors wish to acknowledge John Fazio at the Northwest Power and Conservation Council for his assistance in configuring and running the GENESYS model and Ryan Wilbur, Modeling Support Branch, Bay-Delta Office, Department of Water Resources, for his assistance in obtaining archived hydropower simulations from the CALSIM II model for the Sacramento San Joaquin system.

Please cite this report as follows:

Lettenmaier, Dennis P., Alan F. Hamlet, Tazebe Beyene, and Se-Yeun Lee. 2008. *Evaluation of Alternative Models and Methods for Prediction of Hydropower Resources in California and the Pacific Northwest.* California Energy Commission, PIER Energy-Related Environmental Research Program. CEC-500-2007-104.

Preface

The Public Interest Energy Research (PIER) Program supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The PIER Program, managed by the California Energy Commission (Energy Commission), conducts public interest research, development, and demonstration (RD&D) projects to benefit California.

The PIER Program strives to conduct the most promising public interest energy research by partnering with RD&D entities, including individuals, businesses, utilities, and public or private research institutions.

PIER funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency
- Energy Innovations Small Grants
- Energy-Related Environmental Research
- Energy Systems Integration
- Environmentally Preferred Advanced Generation
- Industrial/Agricultural/Water End-Use Energy Efficiency
- Renewable Energy Technologies
- Transportation

Evaluation of Alternative Models and Methods for Prediction of Hydropower Resources in California and the Pacific Northwest is the final report for the Evaluation of Alternative Models and Methods for Prediction of Hydropower Resources in California and the Pacific Northwest project (contract number 500-02-004, work authorization number MR-043) conducted by the Department of Civil and Environmental Engineering, University of Washington.

For more information about the PIER Program, please visit the Energy Commission's website at www.energy.ca.gov/pier or contact the Energy Commission at 916-654-5164.

Table of Contents

List of Figures

List of Tables

Abstract

Water resources used for hydropower can vary considerably. Hydropower simulation models can provide information to help hydropower and water managers plan future decisions. In this study, researchers applied three types of hydropower simulation models to the Sacramento San Joaquin Basin in California and the Columbia River Basin in the Pacific Northwest. Three models were evaluated: (1) multiple linear regression models, (2) medium complexity reservoir simulation models (CVMod in California and ColSim in the Pacific Northwest), and (3) operational reservoir simulation models (CALSIM II in California and GENESYS in the Pacific Northwest). The models were evaluated for six dams in California and eight dams in the Pacific Northwest, in both retrospective and forecast mode, and the evaluations incorporated the computational and human resources requirements of running each model. Although limited by the availability of observed data, regression models, where they can be implemented, are the least-cost alternative and outperformed both types of reservoir simulation models at most sites examined. In California, the operational CALSIM II model was superior to the medium complexity CVMod model at four of six sites evaluated. In the Pacific Northwest, the results of the operational GENESYS model and medium complexity ColSim model were more similar, with neither model superior to the other overall. In retrospective mode, regression models also consistently outperformed the medium complexity reservoir models for lead times less than six months.

Keywords: Hydropower, prediction, forecasts, regression models, simulation models, model performance, role of model complexity, cost of implementation, climate change

Executive Summary

Introduction

Hydropower is a clean, renewable energy source that is ideally suited to supplying peak energy needs and is usually much cheaper than replacement sources such as natural gas turbines or other fuel-based technologies. Hydropower resources in both the Pacific Northwest and California vary with hydrologic conditions within and between years, and may also vary over longer time horizons (decades to centuries) as a result of long-term climate variability and change.

The ability to forecast these variations in hydropower resource availability has considerable value for various energy-related applications, including retrospective analyses, short- and long-term planning, and energy marketing decisions, among others. The California Energy Commission (Energy Commission) would like to use simulation models or alternative methods to analyze and forecast hydropower resources in California and the Pacific Northwest.

Purpose

An important first step in developing analysis and forecast tools is to evaluate the performance of alternative predictive models and methods, and the resources necessary to implement, use, and maintain these tools.

Project Objectives

This study sought to evaluate the ability of three different kinds of hydropower simulation models to reproduce observed hydropower production at large hydropower projects in the Columbia River Basin in the Pacific Northwest and the Sacramento/San Joaquin basins in California. The study examined three different types of models (in order of increasing complexity and cost):

1. Multiple linear regression models.
2. Medium complexity reservoir simulation models (California-CVMod, Pacific Northwest-ColSim).
3. Operational reservoir simulation models (California-CALSIM II, PNW-GENESYS).

The models were evaluated for their potential use in the following contexts:

1. Retrospective studies of hydropower production (that is, their ability to reproduce observed past hydropower production).
2. Retrospective hydropower forecasts with lead times ranging from one to nine months.
3. Climate change studies.

The report graphically details the results of these comparisons.

Project Outcomes

Researchers examined a total of six sites on the Sacramento/San Joaquin river system and eight sites on the Columbia River system. Results were broadly similar in character at all sites, but they varied among the different sites.

- In general, simulation models were better able to reproduce historic observations for the Pacific Northwest sites than for the California sites.
- In retrospective tests, linear regression models performed the best at reproducing observations and offered the least expensive implementation costs.
- In California the operational CALSIM II model provided marginally better performance than the medium complexity CVMod model.
- In the Pacific Northwest, the operational GENESYS model and medium complexity ColSim model showed different performance at different dams, but neither was clearly superior to the other.
- In the context of seasonal forecasts, regression models showed superior performance at three to six months forecast lead time (depending on region and location), but the more complex simulation models were comparable or superior to them at the longer nine month lead times in California and at six month lead times in the Pacific Northwest.
- For climate change experiments and long-term retrospective simulations, the medium complexity simulation models were found to be the most appropriate choice because they can simulate more different types of projects and are not dependent on observed hydropower or reservoir elevation data (only streamflow data are needed). In addition, medium complexity reservoir models can be used for retrospective studies over a much longer time frame.

Key Findings

- Although limited in application by the availability of observed data, regression equations were found to provide both the lowest implementation cost and superior performance (in comparison to both medium complexity and operational reservoir models) at most projects evaluated. Exceptions are discussed in the report.
- In the Sacramento/San Joaquin river system, CALSIM II performed somewhat better than CVMod at four out of six projects, but CALSIM II is much more costly to implement (both computationally and in terms of staff resources), suggesting that future evaluations of these two models should consider the tradeoffs between cost and marginal performance increases. Furthermore, some straightforward modifications to CVMod appear likely to resolve its performance issues. The comparison between GENESYS and ColSim was somewhat inconclusive because GENESYS could not be run for scenarios that required data later than 1978—a critical limitation for planning studies, which should be resolved. Comparison of ColSim and GENESYS over the pre-1978 period showed that the performance of the two models varied at different projects; however, neither model was clearly superior overall.

- In retrospective forecast mode (with storage updating), regression models were typically superior to CVMod and ColSim at one- and three-month lead times (Pacific Northwest) and one to six months (California), but for lead times of six months or longer (ColSim), and nine months (CVMod), the performance of the medium complexity models was comparable or superior to the regression models. Forecast results were better overall in the Pacific Northwest than in California. The operational simulation models were not evaluated in forecast mode, but based on the retrospective comparisons, they would probably achieve performance comparable to that of the medium complexity models.
- In the context of long-term retrospective studies and climate change studies, regression models (as formulated here) cannot be used because reservoir elevation data is not available. Regression models using streamflow as a single explanatory variable are feasible for climate change work, but this approach raises concerns about how consistent the performance of regression models could be when using them to model an altered climate, and those concerns are difficult to address. Both operational and medium complexity reservoir simulation models are more appropriate choices for such studies. CALSIM II and GENESYS provide more operational detail (and in the case of CALSIM II, superior performance at most sites). However, CVMod and ColSim offer advantages because they automatically update reservoir rule curves in response to changing flow patterns and are easier to modify. Thus, a choice between these models should consider the tradeoffs between implementation cost, model flexibility, retrospective performance, and the need for operational detail in the analysis of water resources impacts.

Recommendations

The study has shown that regression approaches work very well, but that implementation is limited because of available data. One hybrid approach that would be worth exploring is improvement of the medium complexity models (for example, by incorporating more sophisticated regression approaches within the models) to exploit the strengths of the two approaches in an integrated simulation tool.

Benefits to California

Available hydropower resources in California and the Pacific Northwest affect energy use in California. Decision support systems used to forecast hydropower production are believed to be valuable for planning and market analysis, and the California Energy Commission has expressed interest in constructing such tools. This study evaluates simulation approaches that would be used in creating such forecasting and decision support systems.

1.0 Introduction

1.1. Project Overview

Hydropower is a clean, renewable energy source, is ideally suited to supplying peak energy needs, and is usually much cheaper than replacement sources for peaking energy such as natural gas turbines or other fuel-based technologies. Hydropower resources in both the Pacific Northwest (PNW) and California vary with hydrologic conditions within and between years, and may also vary over longer time horizons (decades to centuries) as a result of long-term changes in climate variability and global climate change. The ability to forecast these variations in hydropower resource availability has considerable value for various kinds of energy-related applications, including retrospective analyses, short and long-term planning, and energy marketing decisions, among others. The California Energy Commission (Energy Commission) has expressed interest in using simulation models or related methods to analyze and forecast hydropower resources in California and the PNW. An important first step in developing such tools is the evaluation of the performance of alternative predictive models and methods and the resources necessary to implement, use, and maintain these tools.

In this study, the performance of three types of hydropower simulation models was examined in detail, with respect to their ability to reproduce observed hydropower production at a number of large hydropower projects in the Columbia River Basin in the PNW and the Sacramento/San Joaquin basins in California. The models were evaluated in the following contexts:

1. Retrospective studies of hydropower production (that is, the ability to reproduce past hydropower generation).
2. Forecasts of hydropower generation with lead times ranging from one to nine months.
3. Climate change studies.

1.2. Primary Research Objectives

The primary objectives of this study were as follows:

- Compare and contrast the three modeling approaches and the human and computational resources needed to run and maintain them.
- Compare and contrast the performance of different modeling approaches with respect to their ability to reproduce observations in the context of retrospective studies and seasonal forecasts.
- Highlight the sources of uncertainty and skill in each simulation approach.
- Make specific recommendations regarding the most appropriate use of each model for various purposes.

2.0 Approaches and Methods

This section describes sources of observed data, an overview of the models used for comparison, and the methods used in comparing the model simulations to observations.

2.1. Sources of Observed Data

Hydropower production, reservoir elevation, reservoir storage, and naturalized or modified streamflow data needed for the study were obtained from the sources indicated in Table 1.

Table 1. Sources of observed hydropower, reservoir elevation, and streamflow data

Hydropower Production Data	**Source**
PNW	U.S. Army Corps of Engineers[1]
California	California Energy Commission
Reservoir Elevation and Storage Data	
PNW	U.S. Army Corps of Engineers[1]
California	California Dept of Water Resources[2]
Streamflow Data	
PNW	2000 Level Modified Streamflow, Bonneville Power Administration
California	CALSIM II reservoir model[3]

[1] www.nwd-wc.usace.army.mil/perl/dataquery.pl

[2] http://cdec.water.ca.gov/queryTools.html

[3] Draper et al. 2004; Van Rheenen et al. 2004

2.2. Projects and Time Periods Included in the Evaluation

Observed hydropower production, reservoir elevation, and modified streamflow data were available for eight projects in the PNW for water years[1] 1976–1999. For California, hydropower and naturalized streamflow data were available for water years 1980–1994. Observed reservoir elevation data and hydropower production data were available for the overlapping period from Water Year (WY) 1988–2005. Table 2 lists the projects included in the study, maps of which are included as Figure 1

[1] Water years begin in October and end in September. Water year 1976, for example, begins October 1975 and ends September 1976.

Table 2. Hydropower projects included in the model intercomparison. Storage dams are marked with an (S), run of river dams are marked with an (R).[2] Annual average hydropower production at each dam (in gigawatt-hours) is shown in square brackets.

PNW Hydropower Projects		California Hydropower Projects	
Albeni Falls (S)	[223]	Shasta (S)	[1730]
Libby (S)	[2040]	Folsom (S)	[580]
Hungry Horse (S)	[929]	Trinity (S)	[423]
Dworshak (S)	[1799]	New Melones (S)	[390]
Grand Coulee (S)	[20900]	Keswick (R)**	[420]
Lower Granite (R)	[2796]*	Nimbus (R)**	[58]
Lower Monumental (R)	[2642]*		
John Day (R)	[10600]*		
McNary (R)	[6500]		
The Dalles (R)	[7910]		
Bonneville (R)	[5110]		

* These dams were evaluated but not shown in the report results.

** Keswick is simulated as a reregulation dam below Shasta (including interbasin transfer from Trinity), and Nimbus is simulated as a reregulation dam below Folsom.

2.3. Overview of Simulation Models

Three types of simulation models were examined in this study. In this section, some of the key aspects of each of these model types are described, as well as the computer and human resources needed to construct, run, and maintain them.

2.3.1. Multiple Linear Regression Models

Regression models hypothesize a linear relationship between one or more physical variables that can be measured or forecasted in real time (e.g., water surface elevation in a reservoir, and modified (PNW) or naturalized (California) reservoir inflows) and hydropower production. For example:

(Equation 1) Hydropower = A* res_elev + B * inflow + C

Where *res_elev* is the starting reservoir elevation, and *inflow* is the total naturalized reservoir inflow over the time step. A, B, and C are fixed parameters to be determined using observations. Optimal estimates of the parameters are obtained by minimizing the squared errors of the model using well-established statistical techniques.

[2] A storage dam impounds a significant volume of water in comparison with the annual inflow and is usually operated in such a manner that the storage changes seasonally. A run-of-river dam has little storage in comparison to its inflow, and usually the storage changes only on very short time scales (e.g., a few hours); at monthly time scales the inflows and outflows are essentially equal.

Multiple linear regression models for four different forecast lead times[3] (1, 3, 6, and 9 months) were created for the model intercomparison studies. In each case the explanatory variables were the starting reservoir elevation at the beginning of the period (which was the ending daily value for the previous month), and the sum of the inflows over the forecast period. For example, a one-month-lead-time regression model forecast for hydropower production during October 1976 would use the reservoir elevation on September 30, 1976, and the total volume of natural (or modified) streamflow for October 1976 at the appropriate location. A separate model was developed for each starting month.

Regression models using flow as a single explanatory variable were also developed initially. For monthly simulations these very simple tools performed as well or nearly as well as those that used multiple linear regression (flow and reservoir elevation) in many cases; however, the goodness of fit of multiple linear regression models was found to be superior overall and was therefore used as the basis for the evaluations shown in subsequent sections. For climate change experiments or long-term retrospective studies, regression based only on flow would be preferable, because future reservoir elevation would not be available.

Models constructed for longer lead times were similar except that the streamflow was aggregated over the longer period, and the model was trained to predict the aggregate hydropower production over the longer time period. For the retrospective comparisons, the one-month lead time forecast was used. Note that this configuration assumes "perfect" (i.e., observed) information about both initial reservoir elevation and future streamflow.

Regression Model User Requirements

The resources needed to construct, run, and maintain these kinds of models are modest, and these approaches undoubtedly represent the least cost alternative among the three model types. Because the models are not physically based, observed data for both hydropower production and all explanatory variables chosen are required to train the model, and all explanatory variables are needed to run them. Regression models are a more problematic choice for retrospective studies before about 1975 (PNW) or 1980 (California), or for climate change studies, because reservoir elevation data consistent with current projects and reservoir operating policies (an important explanatory variable in some cases) were not available prior to those dates.

2.3.2. Medium Complexity Reservoir Operations Models

Medium complexity models are intended to represent, in a realistic and physically based manner, key aspects of the systems they represent, while at the same time ignoring many fine-scale operational details. This approach is appropriate when flexibility and transparency in the modeling system are at a premium and operational details are less important or unknown (i.e., in the context of climate change experiments). The two medium complexity reservoir operations models examined in this study were developed for climate change and seasonal-to-interannual forecasting research at the University of Washington. They are described briefly below.

[3] The "lead time" of a forecast refers to how far the forecast looks ahead in time. As defined here, a hydropower forecast with a three month lead time predicts future hydropower production for the three months immediately following the forecast date.

ColSim Model

ColSim (short for **Col**umbia **Sim**ulation) simulates the operation of 33 major storage and run-of-river dams in the Columbia River Hydropower system at a monthly time step[4] (The largest dams are shown in Figure 1). The key system elements simulated by the model are flood control operations, hydropower production, and use of storage for instream flow augmentation (Hamlet and Lettenmaier 1999). ColSim is driven by monthly time step modified flows at a number of system checkpoints, meaning that net diversions for irrigation are accounted for in the driving data rather than in the simulation process. The exception is in the Snake River basin, where diversions from two aggregate storage reservoirs (equal in size to the total storage upstream) are explicitly included in the simulations.

ColSim constructs flood control rule curves (a set of storage guidelines defining the amount of flood space required in each reservoir) using a set of lookup tables that define the relationship between inflow and flood evacuation rule curves.[5] It uses perfect inflow forecasts as the selection criterion. This approach simplifies operational modeling approaches, which use a similar set of lookup tables, but construct the rule curves based on inflow *forecasts*, which contain more realistic errors in seasonal storage evacuation requirements (see the discussion of the GENESYS model below).

ColSim represents minimum flow requirements from individual dams, and simulates releases from storage to support system wide instream flow targets for fish. In recent years, streamflow targets in the Columbia River basin have changed markedly in response to the Endangered Species Act's listings of salmon and other endangered fish (National Marine Fisheries Service 1995, 2000). However, as used here, the ColSim model is configured to simulate the streamflow targets that were typical in the 1970s–1990s, to best reproduce the conditions present at the time the hydropower observations were taken.

---

[4] The "time step" of a model refers to the increment of time used in the model to calculate the behavior of a system or process as a function of time. In a monthly time step model, inputs (e.g., streamflow) and outputs (e.g., hydropower production) for the model are aggregated over each month in the simulation.

[5] Storage reservoirs are operated for flood control according to "rule curves" that specify the amount of space that must be present in the reservoir at the end of each month. In the Columbia basin these evacuation requirements are a function of expected flow in each year. That is, in dry years less storage is required than in wet years.

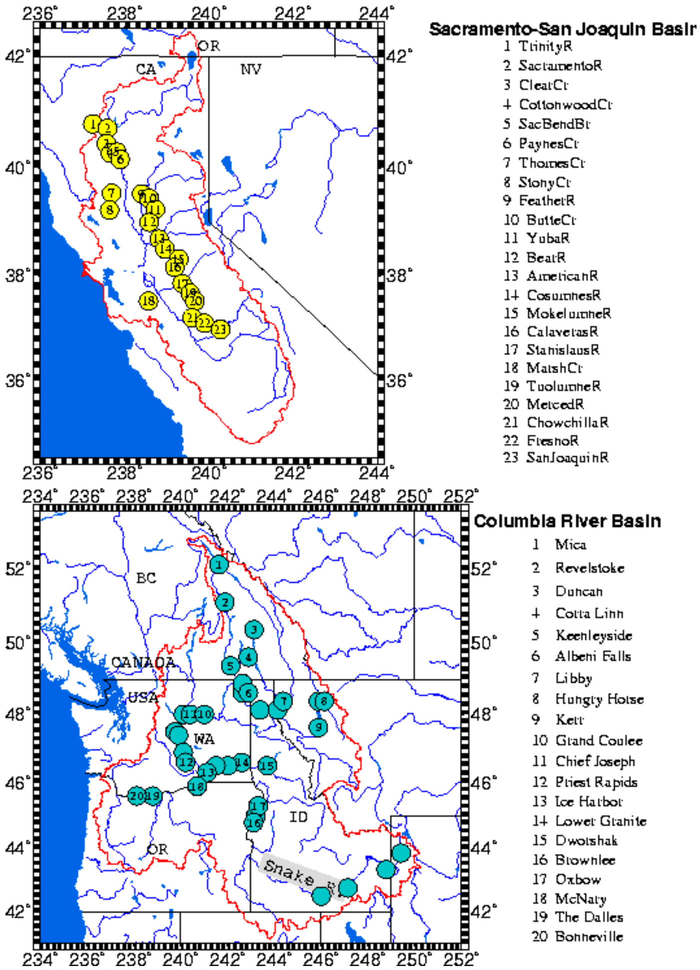

Figure 1. Map of projects and river checkpoints included in the CVMod (upper panel) and ColSim (lower panel) "medium-complexity" reservoir simulation models

As used in this study, ColSim uses a fixed "firm energy" (the maximum energy that can be produced in the worst drought in a period of record (1929–1999) and "non-firm energy" (total energy minus firm energy) targets in each year of the simulation. This approach simplifies operational approaches that include specific information about annual variations in energy demand, and their seasonal timing, due to each year's individual weather. ColSim also has the option (not used here) of including variable non-firm energy targets based on a time series of total demand forecasts (Voisin et al. 2006).

Hydropower production is simulated in the model by first estimating the net head at the project (which is the estimated reservoir elevation minus a typical tailwater elevation[6]) and then

[6] The net "head" for an operating hydropower turbine is a function of the surface elevation in the water impounded behind the dam, and the elevation of the water below the dam, which is called the "tailwater."

estimating the power production associated with the release from storage specified by the model for a fixed overall plant efficiency (usually assumed to be 80%).

ColSim User Requirements

ColSim is coded in STELLA (version 8.0), which is a proprietary modeling software package (www.iseesystems.com/softwares/Education/StellaSoftware.aspx) and which runs only on PC and Macintosh platforms (no Unix version is available). The code for ColSim itself is freely available, but the user must purchase the software for the STELLA environment to use all of the model's features (an inexpensive "run time" version with reduced functionality is also available). ColSim requires some training and experience to run and maintain successfully, but a high level of programming skill is not required. Most users of ColSim have been M.S. or Ph.D. level students in a university setting who have some basic familiarity with the Columbia basin's water resources system and moderate to advanced programming skills. Although the code is relatively easy to read and modify (one of the most desirable features of the STELLA environment), successful modification of the code requires knowledge of the reservoir system itself and a thorough understanding of the model's layout and computational structure.

CVMod

CVMod (short for **C**entral **V**alley **Mod**el) simulates the operation of 13 storage reservoirs in the Sacramento/San Joaquin River system in California (a subset of the 23 locations shown in Figure 1). Trinity dam (in the North Coast region of California) is also included in the model because this dam provides water to the SSJ. Energy production at two reregulation dams[7] is also simulated. As for ColSim, CVMod was developed for use in assessment of climate change impacts on California water resources, and is described in more detail, along with the California climate change application, by Van Rheenen et al. (2004). The main simplifying assumption in the model is that reservoir operations are constrained primarily by water supply requirements and instream flow targets at dams and other system checkpoints. The model bases simulated water supply targets on simplified versions of more detailed algorithms developed for the CALSIM II model (Draper et al. 2004). The approach simulates water allocation policies by relating "perfect" forecasts (i.e., observations) of reservoir inflows to estimated water allocations for irrigation for each water year. This is carried out in the model using a simple categorization scheme that classifies each water year into one of five levels (ranging from "critically dry" to "wet"), each associated with specific water allocation targets. The model then determines appropriate releases from storage reservoirs to supply these allocations in each month of the simulation. Reservoir operations decisions associated with energy production are ignored in this simplified modeling framework, and the model calculates hydropower production as a post-processing step, relating energy production to the natural logarithm of reservoir storage and simulated reservoir release via a regression approach. That is:

[7] A "reregulation dam" is a small run-of-river dam that is built below a larger storage dam, and whose purpose is to "re-regulate" large intermittent releases from the storage dam (e.g., for peak power needs) in such a way that a more continuous and steady flow is produced from the combined projects.

(Equation 2) $$HP = (A * [\ln(storage)] + B) * release$$

Where *HP* is hydropower production, *storage* is simulated reservoir storage volume at the beginning of the time step, and *release* is total release from the reservoir during the time step. (The first term on the right side in the parentheses represents the estimated overall plant efficiency as a function of storage).

Because the reservoir releases are completely independent of the estimated hydropower they produce in the model simulations, the model output can be post processed using other approaches, if desired.

Instream flow targets are not necessarily stationary through time in the SSJ (as in the case of Columbia basin discussed above), but the model simulates instream flow augmentation as if the dams were operated in the past according to current reservoir operating policies. This is consistent with the approach used in CALSIM II.

CVMod User Requirements

CVMod is coded in STELLA (version 8.0) and has similar user requirements to ColSim as discussed above.

2.3.3. Operational Models

In designing this study, two operational models that were felt to be reasonable choices for use by the Energy Commission for making hydropower simulations were selected for evaluation. These were the GENESYS model for the Columbia River basin and the CALSIM II model for the Sacramento/San Joaquin basins. Some important characteristics and features of these models in the context of this study are outlined below. The interested reader is directed to the references given below for a more complete description of these tools.

GENESYS

GENESYS is a Monte Carlo simulation tool for the Columbia River Hydropower system that is designed primarily for load studies or historic simulations of reservoir operations, regulated flow, and hydropower production (www.nwcouncil.org/GENESYS/tour.htm). The model uses a 14-period water year with split months in April and August. GENESYS uses the same modified flow data set as ColSim but includes many smaller hydropower projects that are not represented in ColSim. The model represents both conventional and hydropower energy resources and simulates the use of the combined system in meeting estimated electrical loads that are associated with a particular planning horizon and set of historic water years. The core of the model is the same as the operational reservoir model (called HYDSIM) used by the Bonneville Power Administration for medium and long-range planning.

A complete description of the differences between GENESYS and ColSim is beyond the scope of this brief overview, however, the primary differences between the two models that are relevant to this study are related to the specification of flood rule curves, energy targets, and spill requirements. Flood rule curves in GENESYS include estimates of streamflow forecasting errors that partly determine the changing trajectory of reservoir storage from January–April at most

dams. In the context of simulation of retrospective hydropower generation, these effects can influence the results in a significant manner. In the context of forecasting, when these kinds of errors cannot be specified in advance with great certainty, the effects are probably less important. Unlike ColSim, which uses fixed energy targets every year, GENESYS uses energy targets that are related to the specific weather sequences in each water year. Thus, water availability and energy demand are linked in a more realistic and consistent manner. GENESYS also includes more detail about operations in the lower Columbia River that require spill from several dams during the salmon outmigration season. ColSim currently ignores this aspect of the system's operation.

The GENESYS configuration that was available for this study includes simulation of dam releases to support fish flow targets specified by the 2000 Biological Opinion (NMFS 2000). ColSim also includes the option to simulate these recent operations, but because the simulation period corresponding to the hydropower observations was before these operations were put into effect, the ColSim configuration used here intentionally does not include these more recent aspects of system operation as discussed above.

An important limitation of the current GENESYS application encountered during the course of this study is that the period of record for the model-driving data was limited to water years 1929–1978. The limitation does not relate to available streamflow data, which are currently available through 1999 (as for ColSim), but rather to observed flood control rule curves which were unavailable for the most recent period from 1979–1999 in the current GENESYS implementation.[8] Hydropower observations with all current dams in place, which begin in water year 1976, therefore, include only three years of overlap with the GENESYS runs, from which it is difficult to draw meaningful conclusions. To overcome this limitation, ColSim runs were compared to GENESYS runs for the period 1970–1978 to gain a sense of the differences between the two models. GENESYS is only included in the retrospective comparison portion of the study and is not used for the forecasting experiments.

GENESYS User Requirements

GENESYS is easily run on a PC and includes user-friendly output summary graphics and data export features. Changes to the code or changes to the simulated reservoir operations are more difficult to make in comparison with models like ColSim and CVMod, and the source code is currently not available on the model's website.

CALSIM II

The CALSIM II (Draper et al. 2004; http://modeling.water.ca.gov/hydro/model) implementation for the Sacramento/San Joaquin River basins (SSJ) simulates the operation of the entire SSJ and San Francisco Bay Delta water system within a multi-objective management framework. Some important components of the model include: simulations of surface water operations at storage and run-of-river dams for water supply, instream flow, flood control, and hydropower

---

[8] At the time of this writing the required rule curves from 1979–1999 are under development to allow the GENESYS model simulations to be extended to 1999.

production, conjunctive use of groundwater, transfers of water between different portions of the system, and system operations designed to maintain water quality in the Bay Delta area. The model also contains optimization routines to allocate and transfer water to various uses and users within the system for irrigation and urban water supply. As for GENESYS, CALSIM II results are only included in the retrospective comparison portion of this study.

CALSIM II User Requirements

CALSIM II is coded in FORTRAN 90 and can be compiled and run on modern PCs or other systems. The code for the model is open source, however, some restrictions in obtaining the implementation for the Sacramento/San Joaquin may apply, and a FORTRAN 90 compiler and proprietary optimization software (XA Professional Linear Programming System; www.sunsetsoft.com/outgoing/CalsimDownload.html) must be obtained to run the model. CALSIM II is the most complex model examined in this study and requires extensive knowledge of the reservoir system and of the model itself to successfully exercise the tool. These requirements translate to higher implementation and maintenance costs when using the tool in a decision support system. Because of limited resources for this study and the model's complexity, it was decided, in designing the study, to use only archived output from a previous model run (water year 1980–1994).

2.4. Overview of the Model Intercomparison Procedures

Two model intercomparison procedures were used. The first evaluated the regression models, medium complexity simulation models, and operational models in terms of their ability to reproduce the observed time series of hydropower production at some specific projects for which appropriate observations were available. In the case of the regression models, a time series of observed end-of-month reservoir elevation and monthly naturalized streamflow served as driving data, and the models simulated the time series of hydropower production. For the other models, the primary input time series were the observed monthly naturalized streamflows at a number of system nodes. In the case of the operational models, some reservoir rule curves (e.g., flood rule curves), or other kinds of information related to operational details in the simulations were also derived at certain times from a time series of observations.

As discussed above, a direct comparison with hydropower observations was not possible for more than a few years, due to limited driving data for GENESYS. Instead, in a separate time series analysis, hydropower simulations from GENESYS and ColSim from 1970–1978 were compared to gain a sense of the differences between the two models.

The second model intercomparison used only the regression equations and medium complexity reservoir models and compared the simulation models in "forecast mode" by comparing retrospective forecasts ("hindcasts") to the observations. In these experiments, it was assumed that reservoir elevations were known at the beginning of the forecast period, and future streamflows were perfectly known for a range of forecast lead times (one to nine months). The ability of the two models to reproduce the observations for each retrospective forecast period was evaluated. To produce the hindcasts, the medium complexity reservoir simulation models were initialized with observed reservoir storage data at the beginning of the forecast period,

and a time series of naturalized streamflow for a particular lead time (one to nine months) drove the model. The models simulated a trajectory of reservoir storage through the forecast period. The multiple linear regression models used the same initial reservoir elevation and an aggregate streamflow volume for the forecast period as the two explanatory variables. In both cases, the aggregate hydropower production over the forecast period from each model was compared to the same quantity from observations.

To reduce the number of plots to a reasonable level, only four forecast starting dates (of 12 possible) were examined, as shown in Table 3.

Table 3. Forecast starting dates (when storage and reservoir elevation are updated to observed end-of-month values) and corresponding forecast periods used to summarize the forecast results

Start (update) month	1-month lead	3-month lead	6-month lead	9-month lead
Sep	O	OND	ONDJFM	ONDJFMAMJ
Dec	J	JFM	JFMAMJ	JFMAMJJAS
Mar	A	AMJ	AMJJAS	AMJJASOND
Jun	J	JAS	JASOND	JASONDJFM

2.4.1. Metrics of Performance

The performance of the various models is displayed graphically using monthly time series plots, rank plots (data ranked from highest to lowest and then plotted in order), and scatter plots (observed versus modeled data). All graphical plots are produced in a single page, one per project. Summary performance statistics such as the root mean square error (RMSE, the square root of the average squared model error), bias in the mean (the difference between the simulated and observed mean), the Nash Sutcliffe efficiency index (E—see Equation 3 below) and the familiar linear correlation coefficient (R) are calculated for each project and presented in tables. The Nash Sutcliffe efficiency index (Nash and Sutcliffe 1970), which may be unfamiliar to some readers, is a measure of the mean squared model errors in comparison with the variance of the observations. That is:

(Equation 3) E = 1 - (mean of squared model errors)/(variance of observations)

Perfect skill using this metric is a value of E=1.0 (i.e., model errors = 0), and "no skill" is a value of E=0.0 (average model errors equivalent to an assumption of an average value every year). Negative values mean that the forecast is inferior overall to assuming average conditions. RMSE and E are sensitive to both random and systematic errors (e.g., bias), whereas the correlation coefficient (R) is not very sensitive to bias. In some cases where the model bias is large, the goodness of fit measured by R may be relatively high, but both RMSE and E may suggest poor goodness of fit. In most cases, however, the metrics suggest a similar performance ranking of each approach and are consistent with the more qualitative graphical measures shown in the time series, rank, and scatter plots.

To keep the number of figures in this report within manageable limits, only scatter plots for six representative projects in each basin are used as a graphical illustration of retrospective evaluation performance, and only three representative projects in each basin are used to illustrate the retrospective hindcast evaluations.

3.0 Results and Discussion

Although the different simulation methods for all available projects were compared (Table 2) during the course of the study, to keep the number of figures to a manageable level, the results for only six representative projects in each region are shown in this report. Tabulated results are shown for six projects in California, and eight projects in the PNW.

3.1. Historical Intercomparison

In this section, the three modeling approaches were evaluated in the context of retrospective hydropower simulations.

3.1.1. California

In Figures 2 through 7 the three modeling approaches at six dams in California were evaluated, using (a) time series plots, (b) rank plots (with data ranked from highest to lowest and then plotted in order from left to right), and (c) scatter plots. Goodness of fit statistics are tabulated in Table 4.

Table 4. Retrospective goodness of fit statistics for three modeling approaches in California by project

		Regression		
DAM	**RMSE (MW-hr)**	**BIAS (MW-hr)**	**E[a]**	**R**
TRINITY	21348	1290	0.268	0.580
SHASTA	48631	9635	0.578	0.776
NEW MELONES	11313	610	0.777	0.882
KESWICK	8915	-1355	0.729	0.882
NIMBUS	7665	2683	0.680	0.852
FOLSOM	1363	398	0.637	0.821
		CVmod		
DAM	**RMSE (MW-hr)**	**BIAS (MW-hr)**	**E[a]**	**R**
TRINITY	30656	-3200	-0.427	0.326
SHASTA	85362	-8486	0.126	0.497
NEW MELONES	33383	16331	0.072	0.698
KESWICK	31564	-1754	0.094	0.630
NIMBUS	16088	2014	-0.218	0.353
FOLSOM	2586	-1490	-0.049	0.574
		CALSIM-II		
DAM	**RMSE (MW-hr)**	**BIAS (MW-hr)**	**E[a]**	**R**
TRINITY	30931	-10266	-0.453	0.407
SHASTA	64834	-3459	0.496	0.780
NEW MELONES	26192	-12320	0.429	0.768
KESWICK	31008	6796	0.126	0.651
NIMBUS	11649	-1778	0.361	0.785
FOLSOM	2718	-688	-0.158	0.580

[a]Nash-Sutcliffe efficiency factor
MW-hr = megawatt-hour

The results demonstrate that, where they can be applied, regression approaches generally produce superior historic hydropower simulations when compared to medium complexity and operational reservoir simulation models. The operational CALSIM II model was superior to CVMod for most of the sites examined. Quantitative measures of performance broadly corroborate the graphical comparisons, and these conclusions are believed to be robust for the sites examined. Strengths and weaknesses of each modeling approach are summarized in sections below.

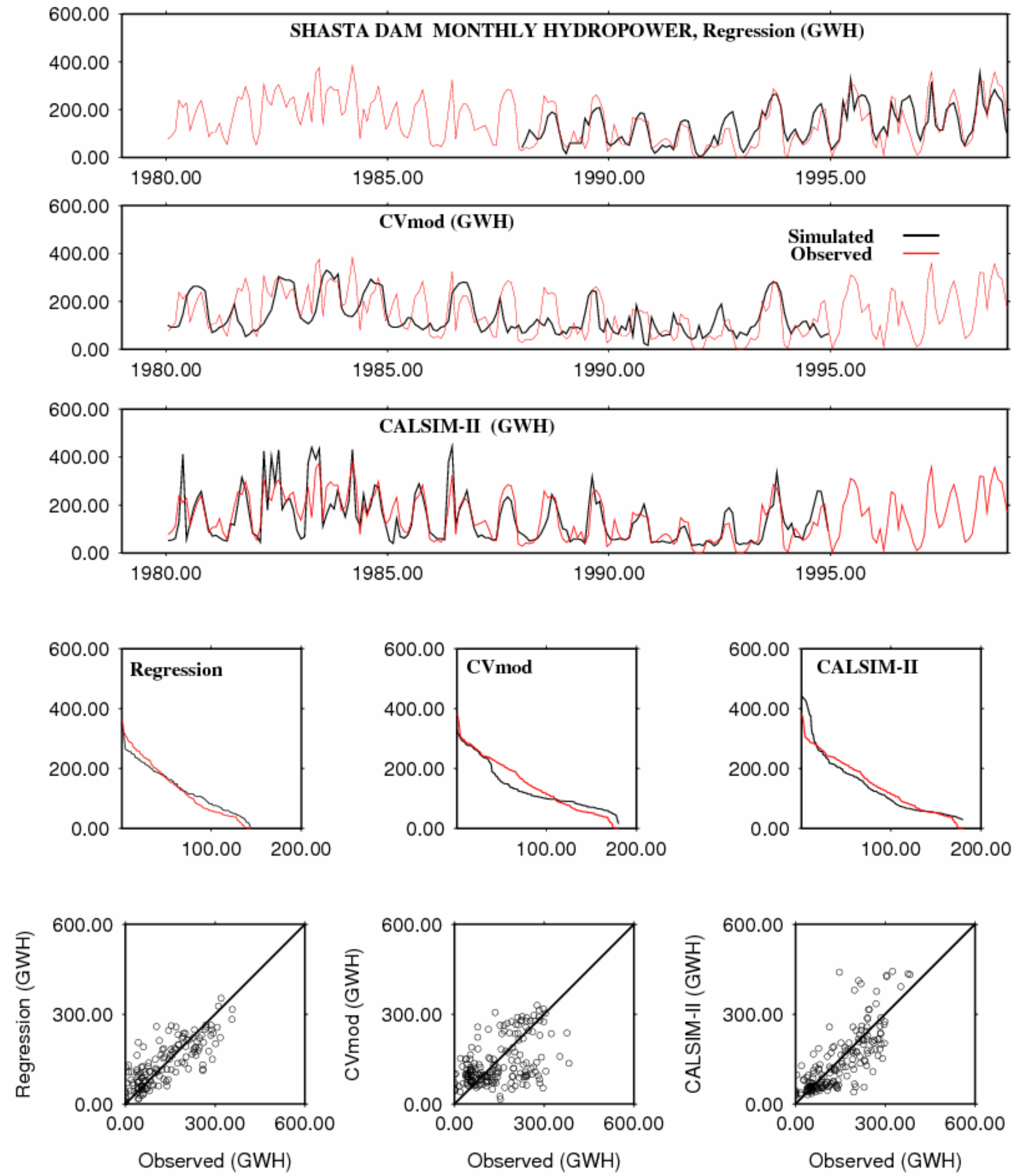

Figure 2. Graphical comparison of three hydropower modeling approaches for Shasta Dam in the Sacramento San Joaquin basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

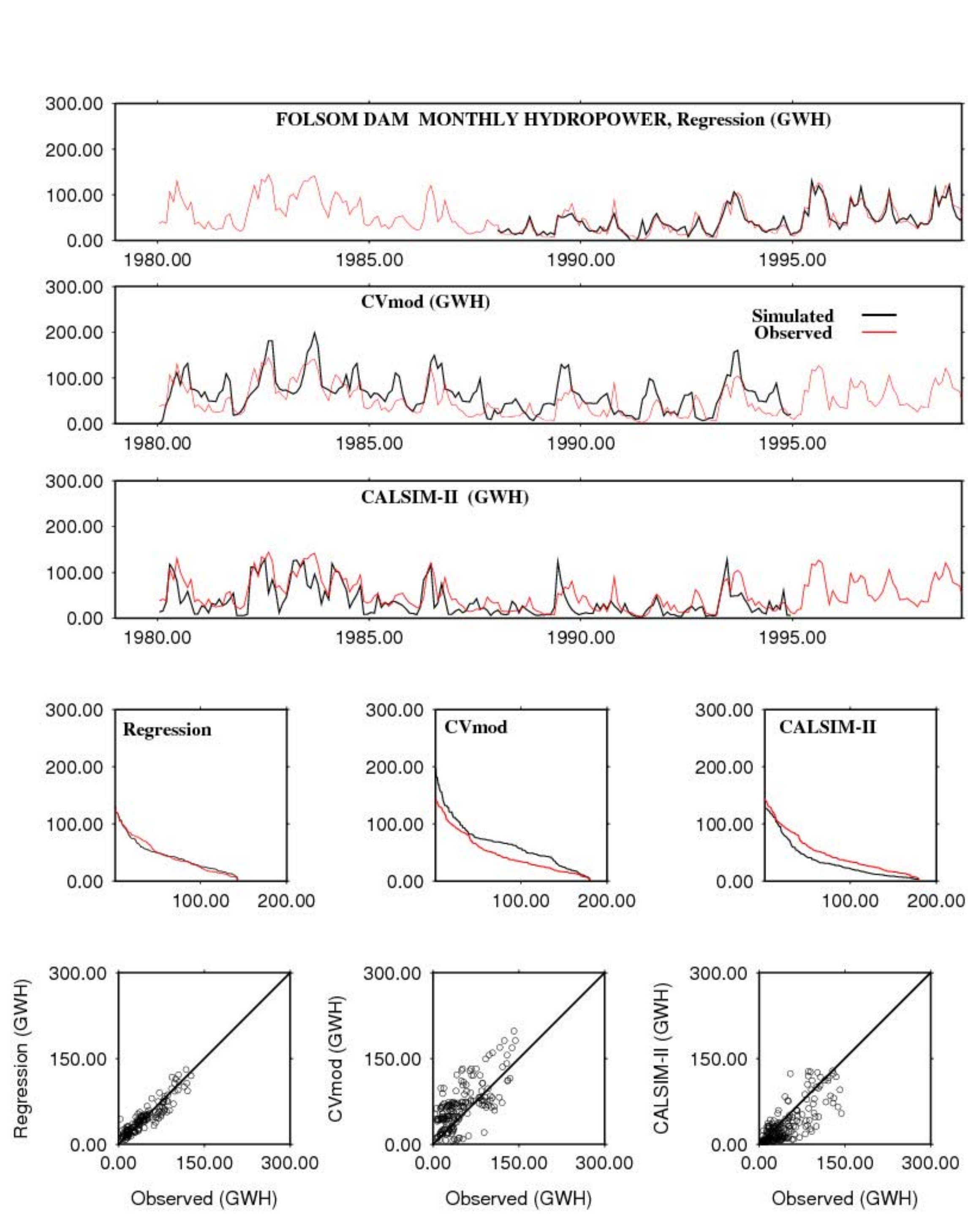

Figure 3. Graphical comparison of three hydropower modeling approaches for Folsom Dam in the Sacramento San Joaquin basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

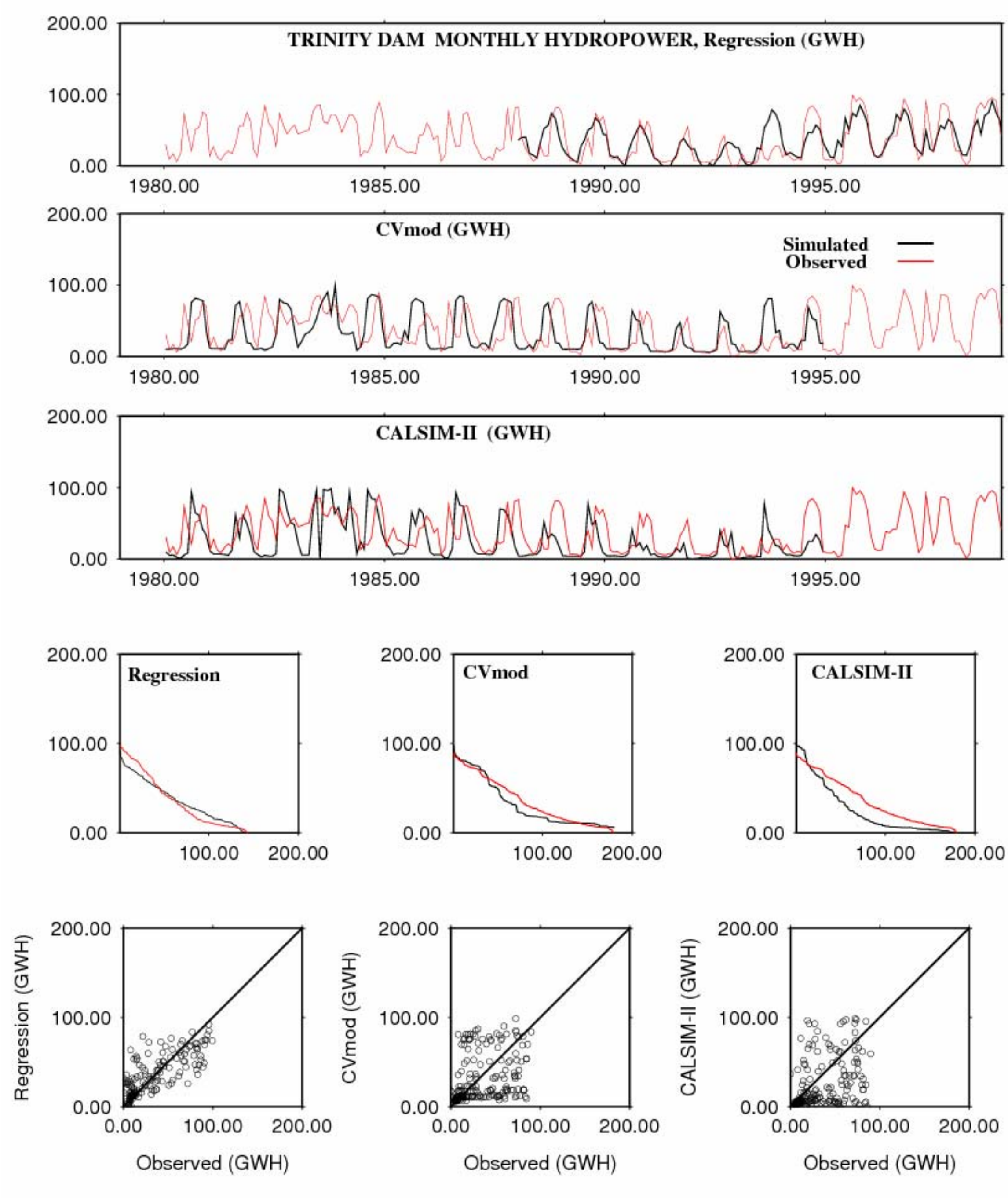

Figure 4. Graphical comparison of three hydropower modeling approaches for Trinity Dam in California's North Coast region. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

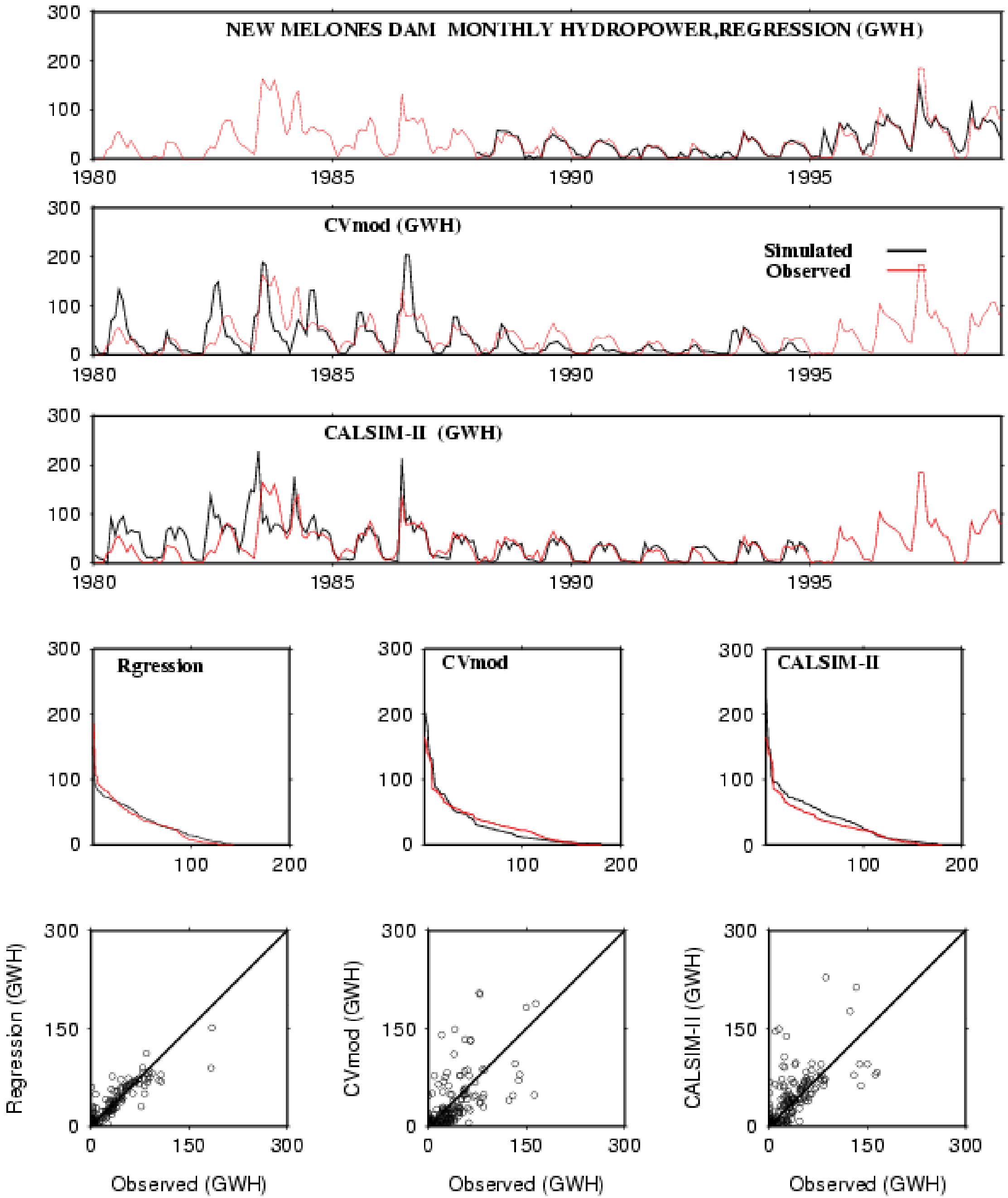

Figure 5. Graphical comparison of three hydropower modeling approaches for New Melones Dam in the Sacramento San Joaquin basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

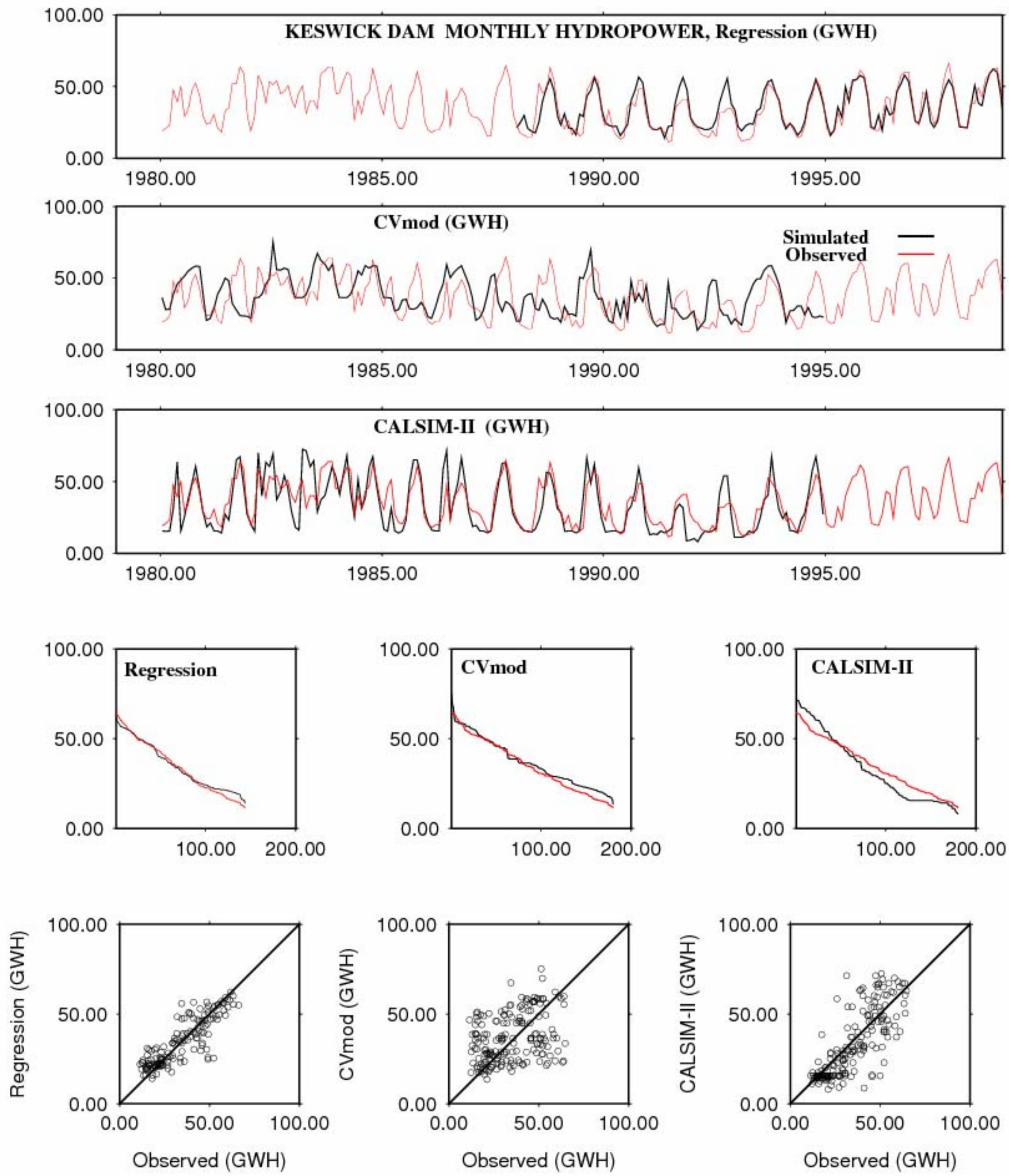

Figure 6. Graphical comparison of three hydropower modeling approaches for Keswick Dam in the Sacramento San Joaquin basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

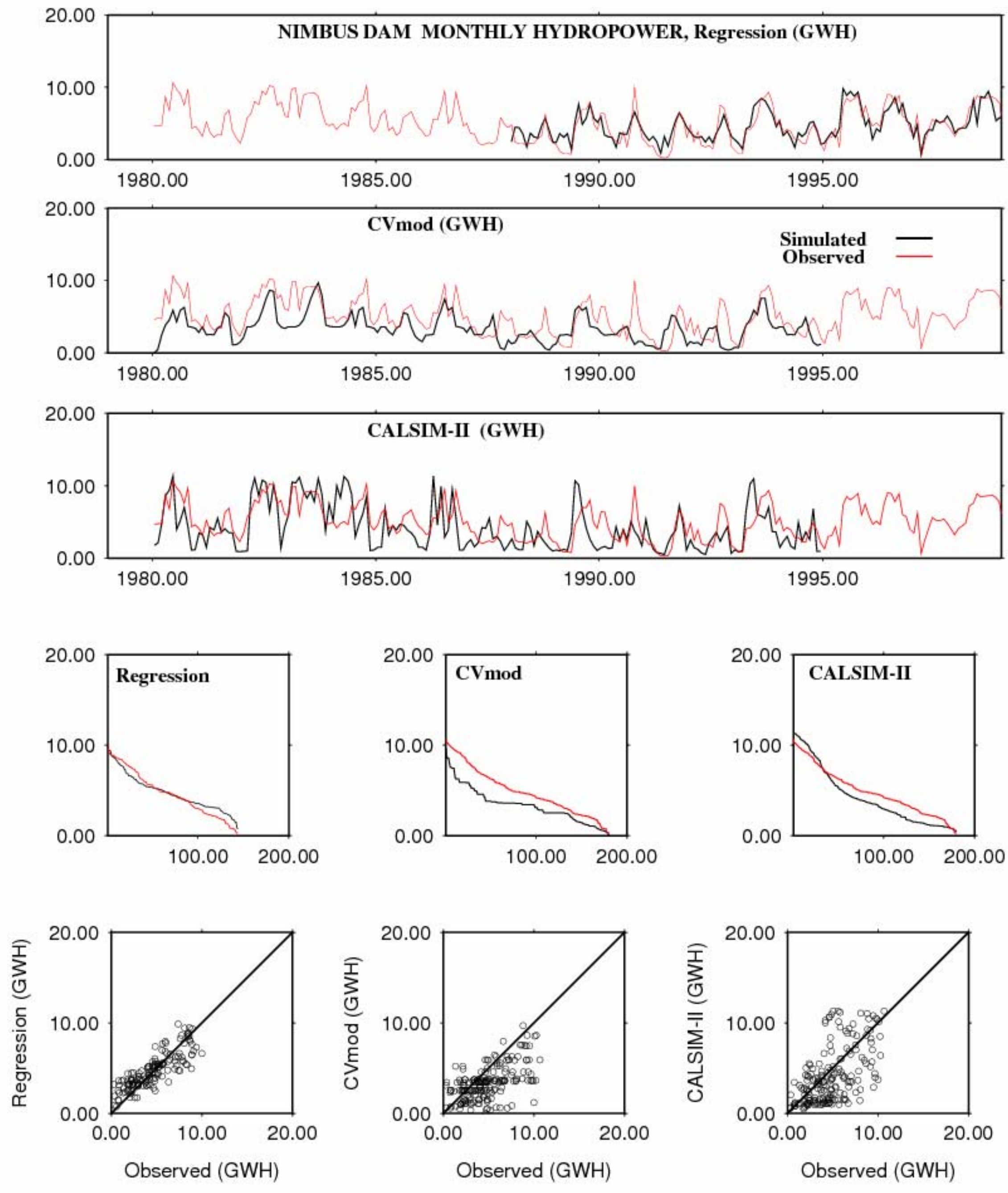

Figure 7. Graphical comparison of three hydropower modeling approaches for Nimbus Dam in the Sacramento San Joaquin basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

3.1.2. PNW

Figures 8 through 13 evaluate the three modeling approaches at six dams in the PNW using (a) time series plots, (b) rank plots (data are ranked and then plotted from highest to lowest from left to right), and (c) scatter plots. Due to the short overlap with observations, only time series plots of the simulations are shown for GENESYS. Goodness of fit statistics for regression and ColSim are tabulated in Table 5.

Table 5. Retrospective goodness of fit statistics for two modeling approaches in the PNW by project

		Regression		
DAM	**RMSE (MW-hr)**	**BIAS (MW-hr)**	**E[a]**	**R**
GRAND COULEE	275145	-150	0.607	0.779
DALLES	104823	470	0.622	0.788
MCNARY	72070	737	0.632	0.794
BONNEVILLE	61484	960	0.467	0.683
LIBBY	59602	580	0.620	0.787
DWORSHAK	69496	896	0.420	0.647
HUNGRY HORSE	48462	-246	0.181	0.425
ALBENI FALLS	3055	-666	0.776	0.881
		ColSim		
DAM	**RMSE (MW-hr)**	**BIAS (MW-hr)**	**E[a]**	**R**
GRAND COULEE	579551	-8505	-0.743	0.557
DALLES	234916	37000	-0.899	0.643
MCNARY	135928	23623	-0.310	0.698
BONNEVILLE	151892	45337	-2.252	0.459
LIBBY	88810	-13764	0.157	0.595
DWORSHAK	130028	2930	-1.029	0.168
HUNGRY HORSE	70593	-1533	-0.738	0.373
ALBENI FALLS	7468	-3627	-0.337	0.445

[a]Nash-Sutcliffe efficiency factor

As for California, the results demonstrate that, where they can be applied, regression approaches generally produce superior historic hydropower simulations in comparison to both the medium complexity ColSim reservoir simulation model and the operational GENESYS model. Quantitative measures of performance in Table 5 corroborate the graphical comparisons for the regression and ColSim results, and these conclusions are believed to be robust for the sites examined for these two approaches. Limited sample size for GENESYS simulations made robust comparison with observations problematic (see discussion of alternate approach below), and scatter plots were not prepared for this reason. Strengths and weaknesses of each of these approaches are summarized in sections below.

Available simulations from the GENESYS model are frequently different from ColSim during the overlapping period 1976–1978, although it is not clear from the small sample available which model is superior overall. To gain a better sense of the differences between the two models, GENESYS was compared to ColSim for a longer retrospective period (See Section 3.3 below).

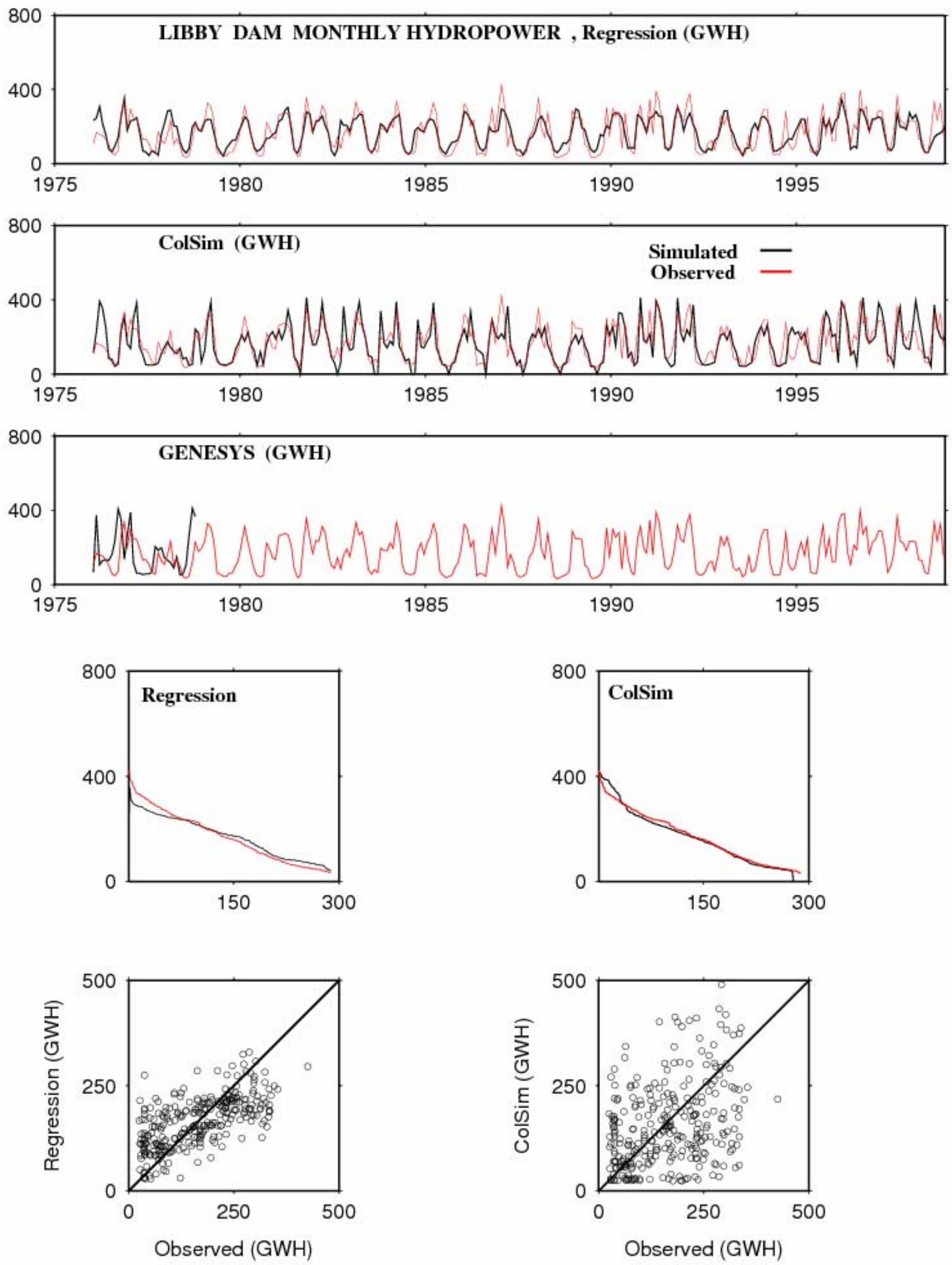

Figure 8. Graphical comparison of three hydropower modeling approaches for Libby Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

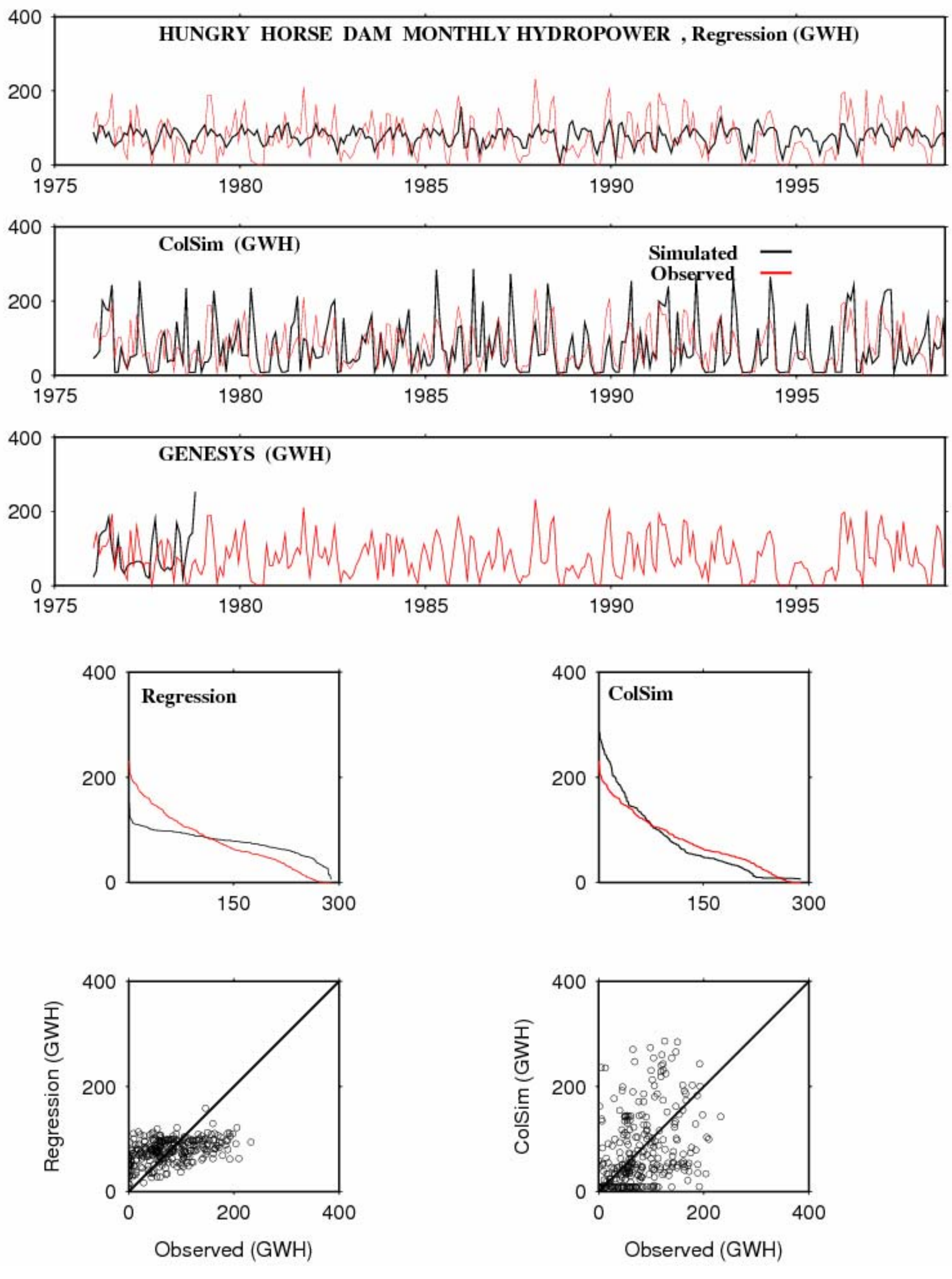

Figure 9. Graphical comparison of three hydropower modeling approaches for Hungry Horse Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

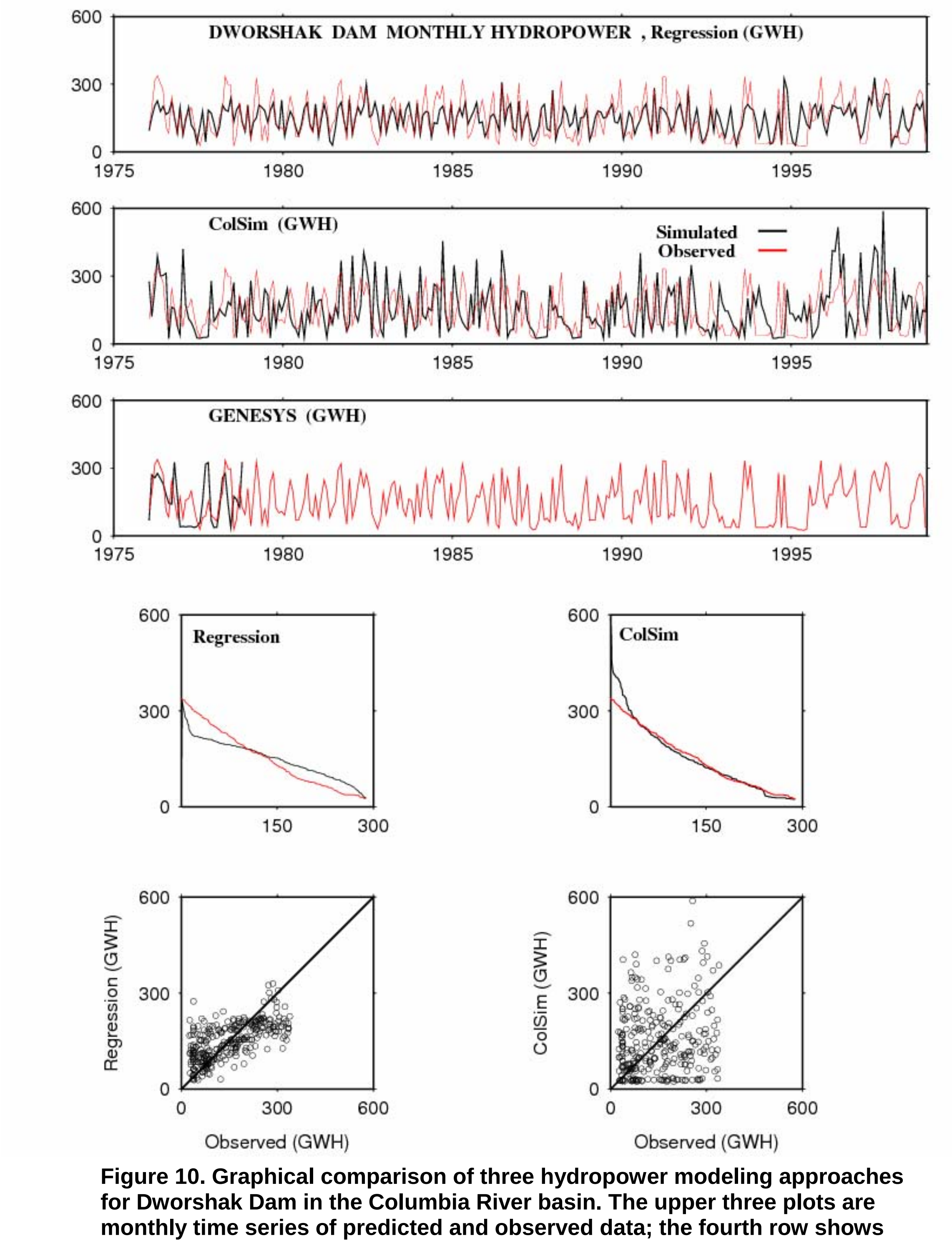

Figure 10. Graphical comparison of three hydropower modeling approaches for Dworshak Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed data; the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

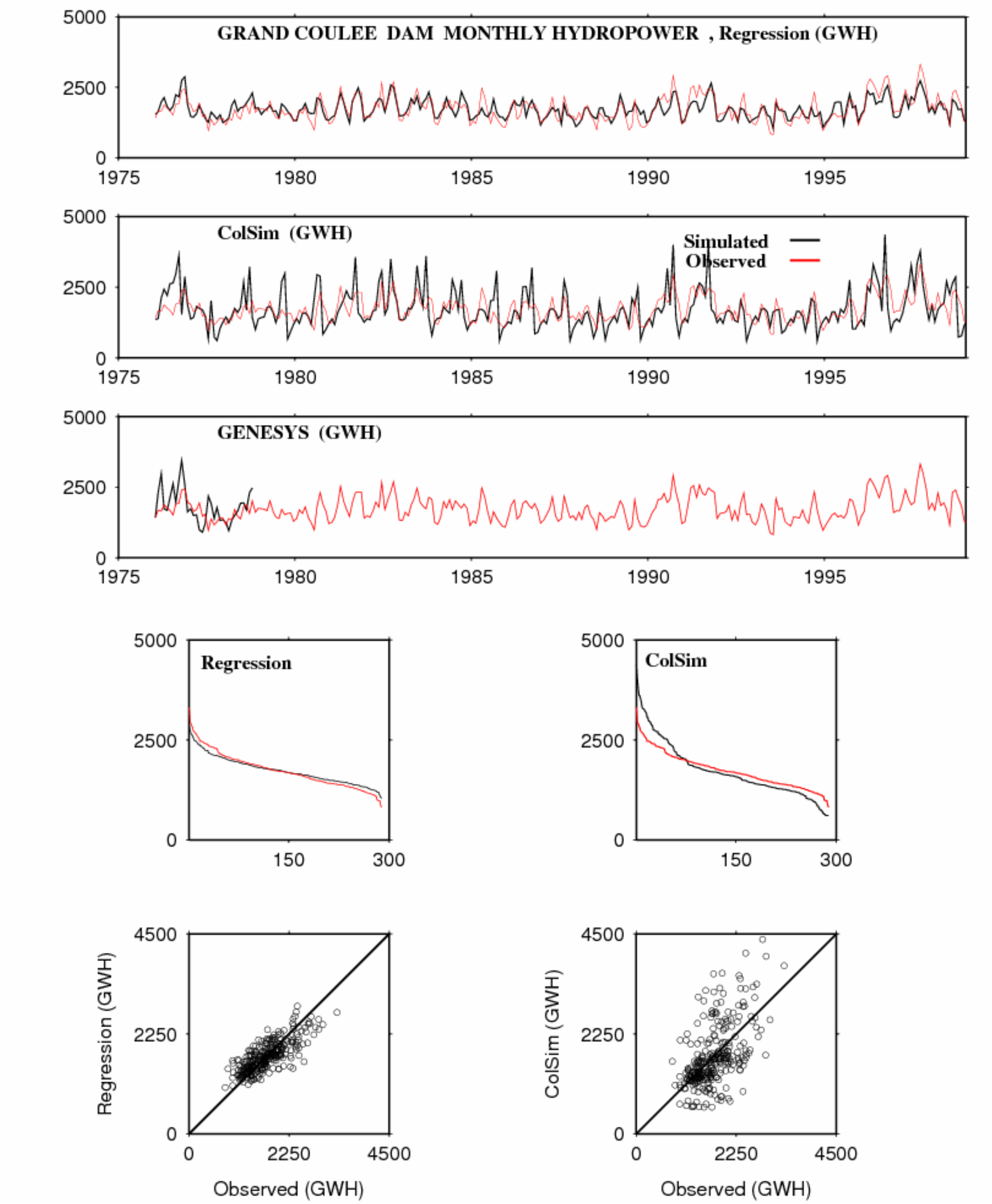

Figure 11. Graphical comparison of three hydropower modeling approaches for Grand Coulee Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

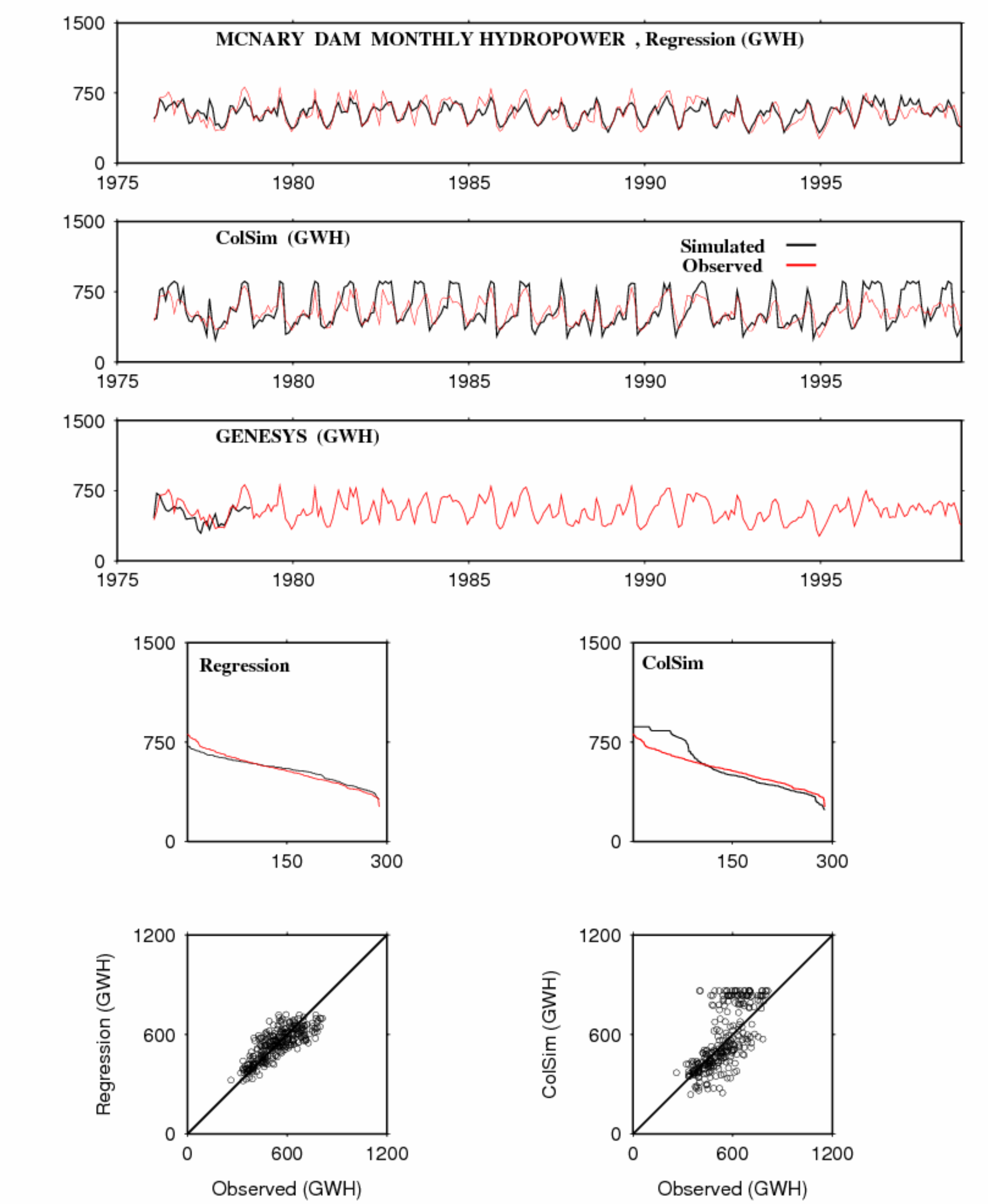

Figure 12. Graphical comparison of three hydropower modeling approaches for McNary Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed data; the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

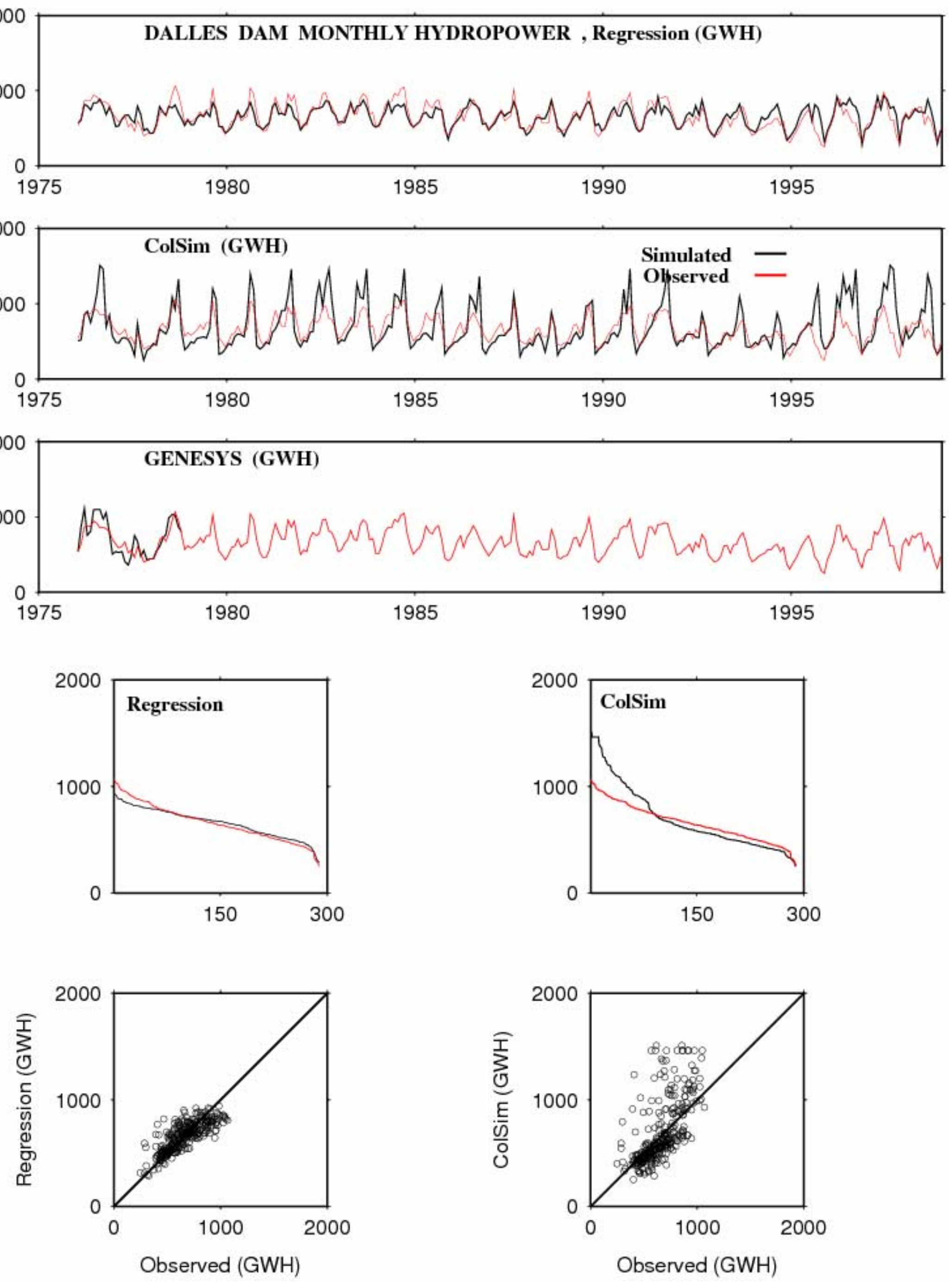

Figure 13. Graphical comparison of three hydropower modeling approaches for The Dalles Dam in the Columbia River basin. The upper three plots are monthly time series of predicted and observed hydropower production (units in gigawatt-hours); the fourth row shows ordered plots of ranked time series data; and the fifth row shows scatterplots of predicted vs. observed data.

3.2. Retrospective Forecast Intercomparison

In this section, the regression and medium complexity models were evaluated in the context of retrospective forecasts, with lead times ranging from one to nine months.

3.2.1. California

In Figures 14 through 19, graphical comparisons are presented of retrospective forecasts produced using regression and CVMod approaches for four forecast dates and four forecast lead times. Results for three representative dams are shown. Although results vary with location and forecast date, overall, the regression approaches are superior to CVMod forecasts for one, three, and six month lead times; whereas CVMod and regression approaches are more comparable for the nine month lead time. In the case of Folsom, regression is superior to CVMod simulations at all lead times. Both regression and CVMod performed relatively poorly for the one-month forecasts, in comparison with the PNW forecasts. Numerical goodness of fit statistics shown in the plots generally confirm the goodness of fit shown in the scatter plots, however the regression approaches are generally more successful at eliminating bias and the Nash Sutcliffe efficiency scores are therefore often higher for regression, even though qualitatively the forecasts from the two different methods may look comparable in the scatter plots and have similar correlation coefficients. (Note that the regression and CVMod simulations are for different, partly overlapping time periods.)

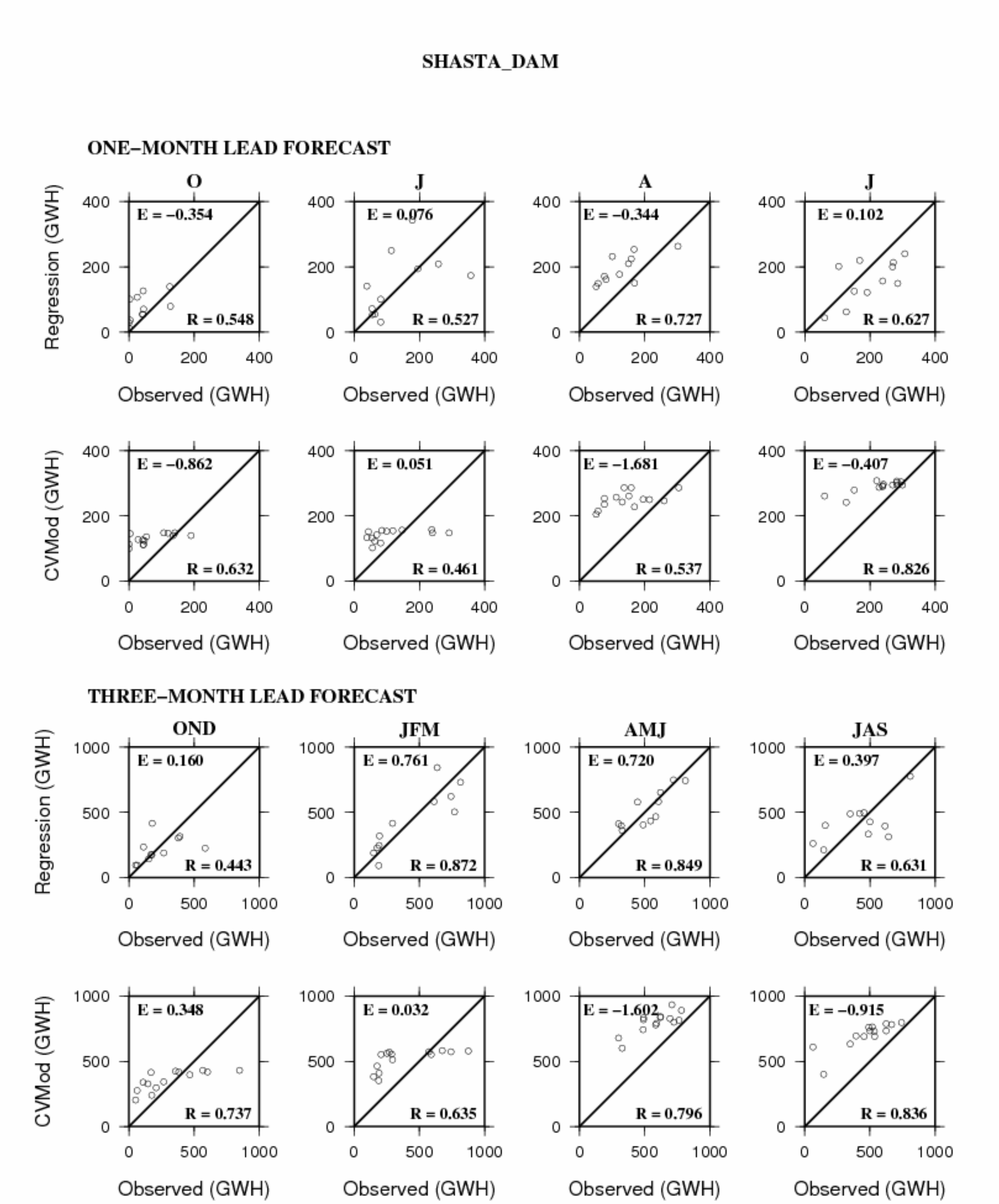

Figure 14. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for Shasta Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

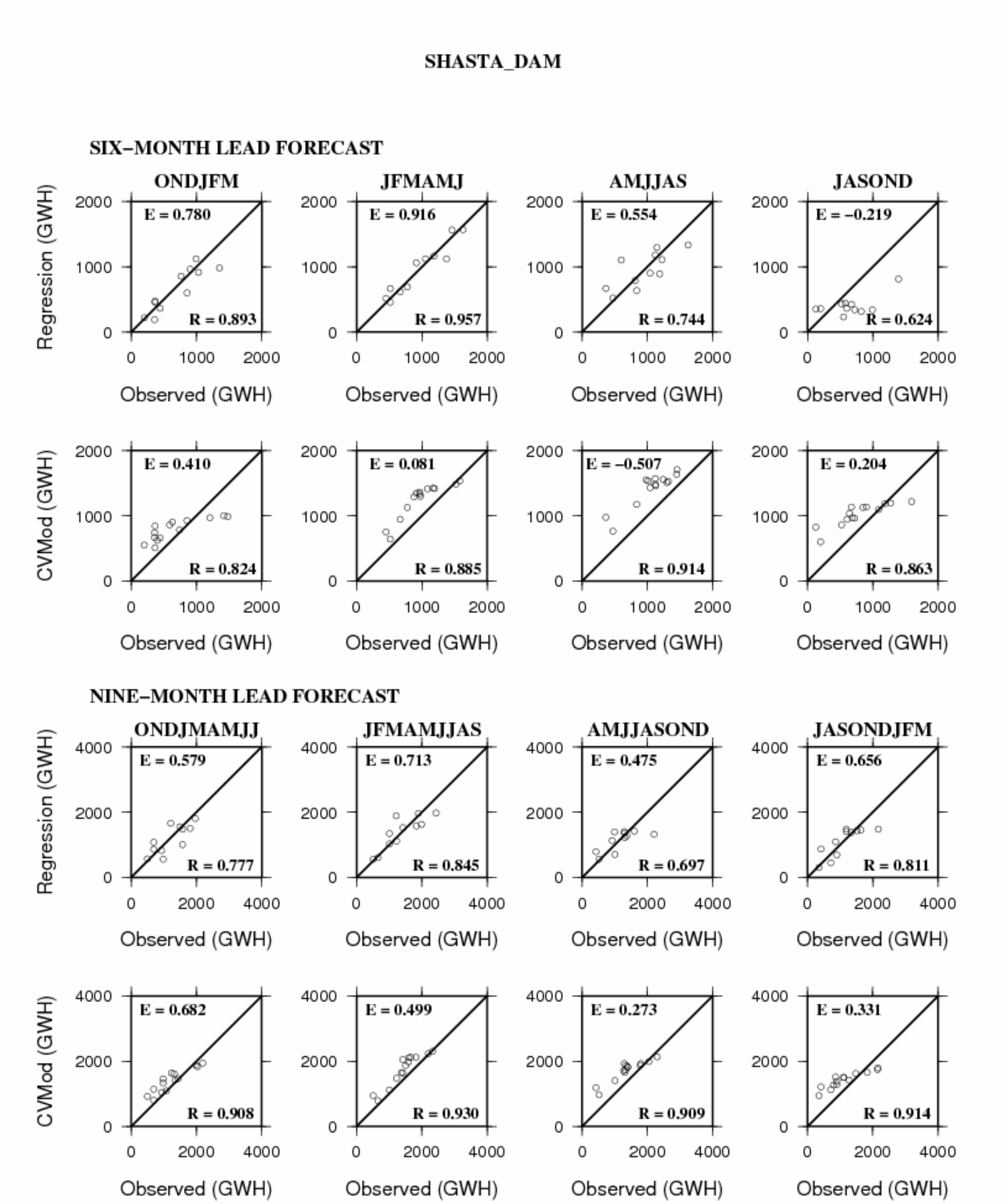

Figure 15. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for Shasta Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

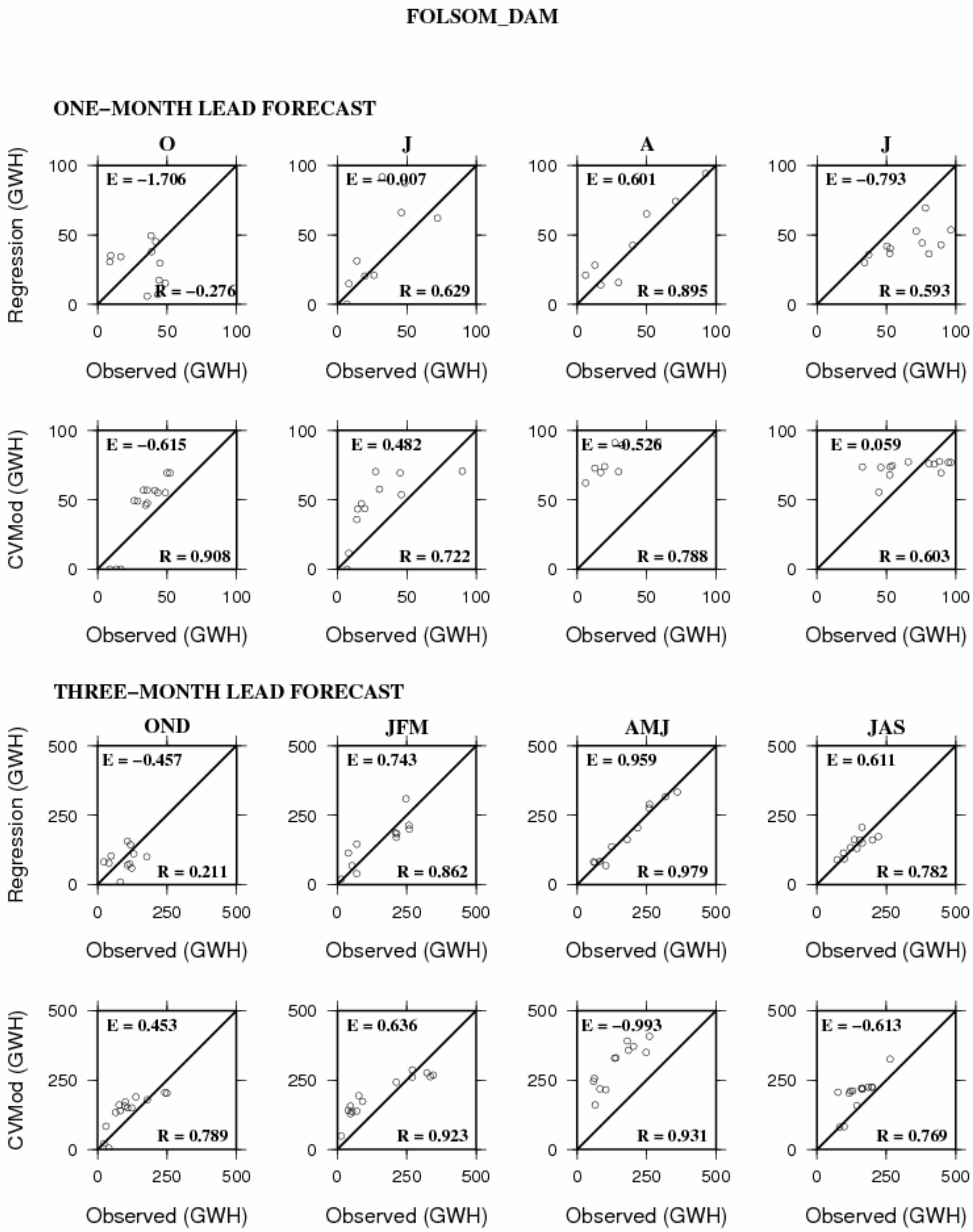

Figure 16. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for Folsom Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

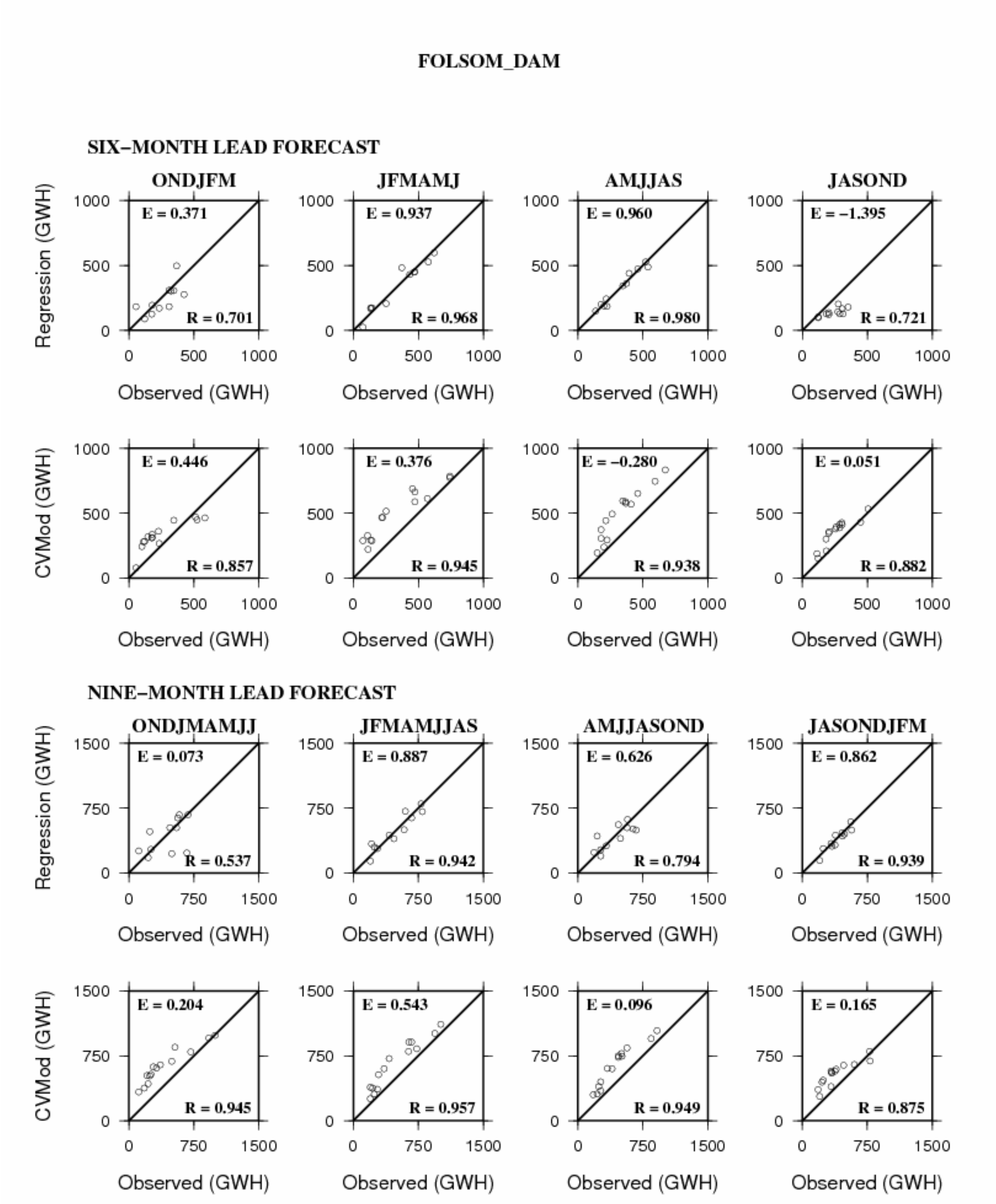

Figure 17. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for Folsom Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

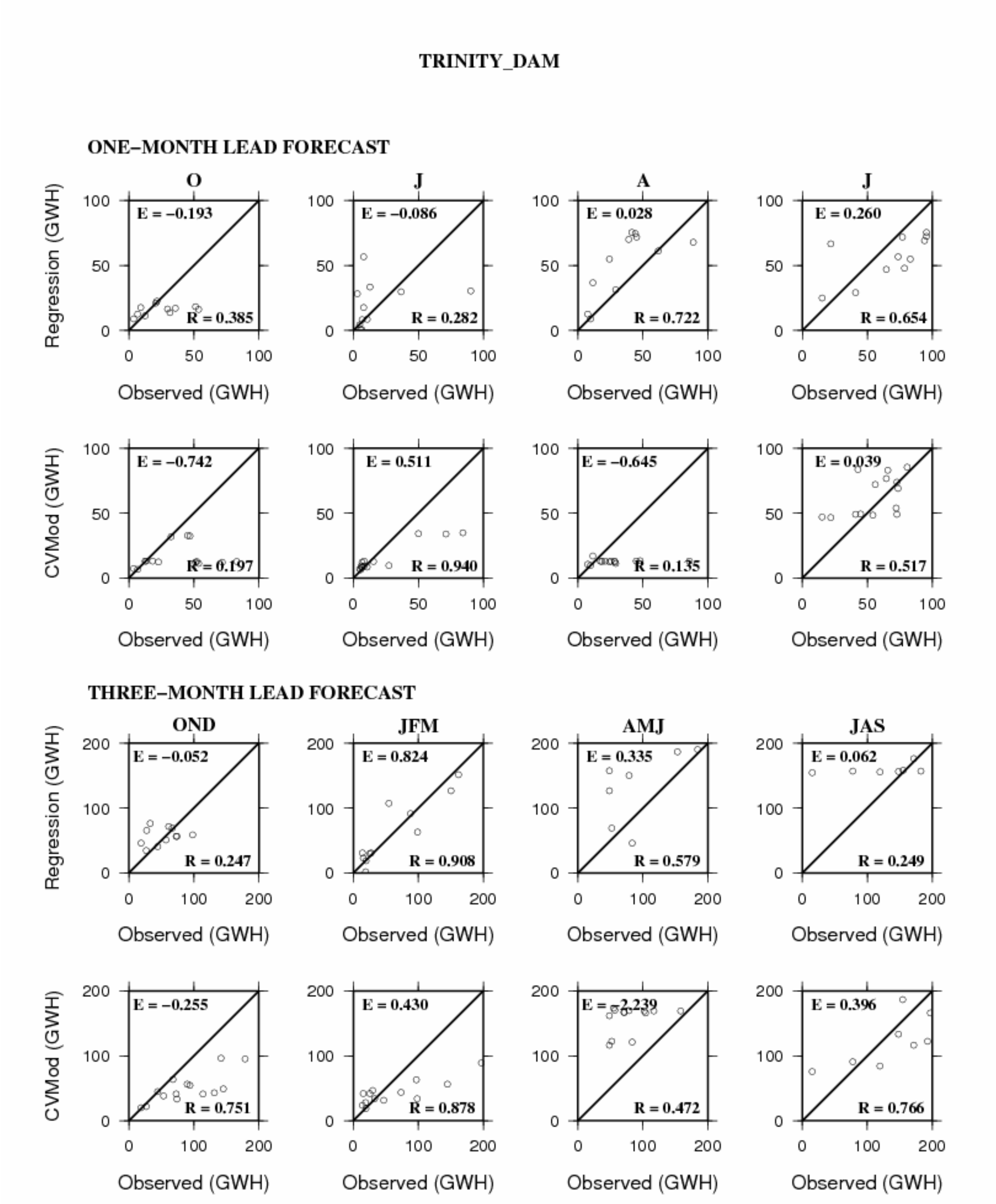

Figure 18. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for Trinity Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

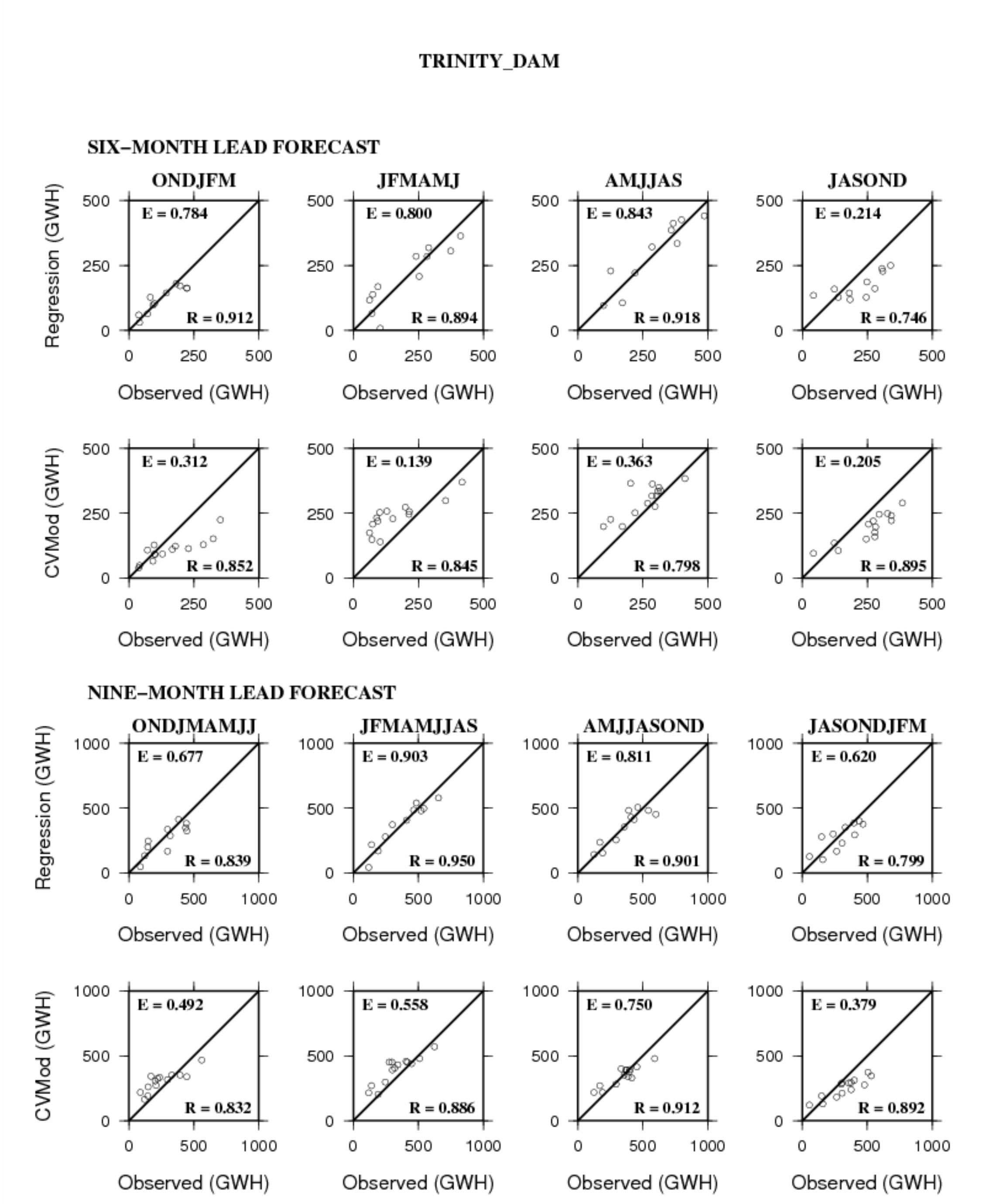

Figure 19. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for Trinity Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

3.2.2. PNW

In Figures 20 through 25, graphical comparisons are presented of retrospective forecasts produced using regression and ColSim approaches for four forecast dates and four forecast lead times. Results for three representative dams are shown. Although results vary with location and forecast date, overall, the regression approaches are frequently equal to or superior to ColSim forecasts for one and three month lead times, whereas ColSim and regression approaches are generally comparable for six and nine month lead times. Numerical goodness of fit statistics shown in the plots generally confirm the goodness of fit shown in the scatter plots, however the regression approaches are generally more successful at eliminating bias, and the Nash Sutcliffe efficiency scores are therefore often higher for regression, even though qualitatively the forecasts from the two different methods may look comparable in the scatter plots and have similar correlation coefficients.

LIBBY_DAM

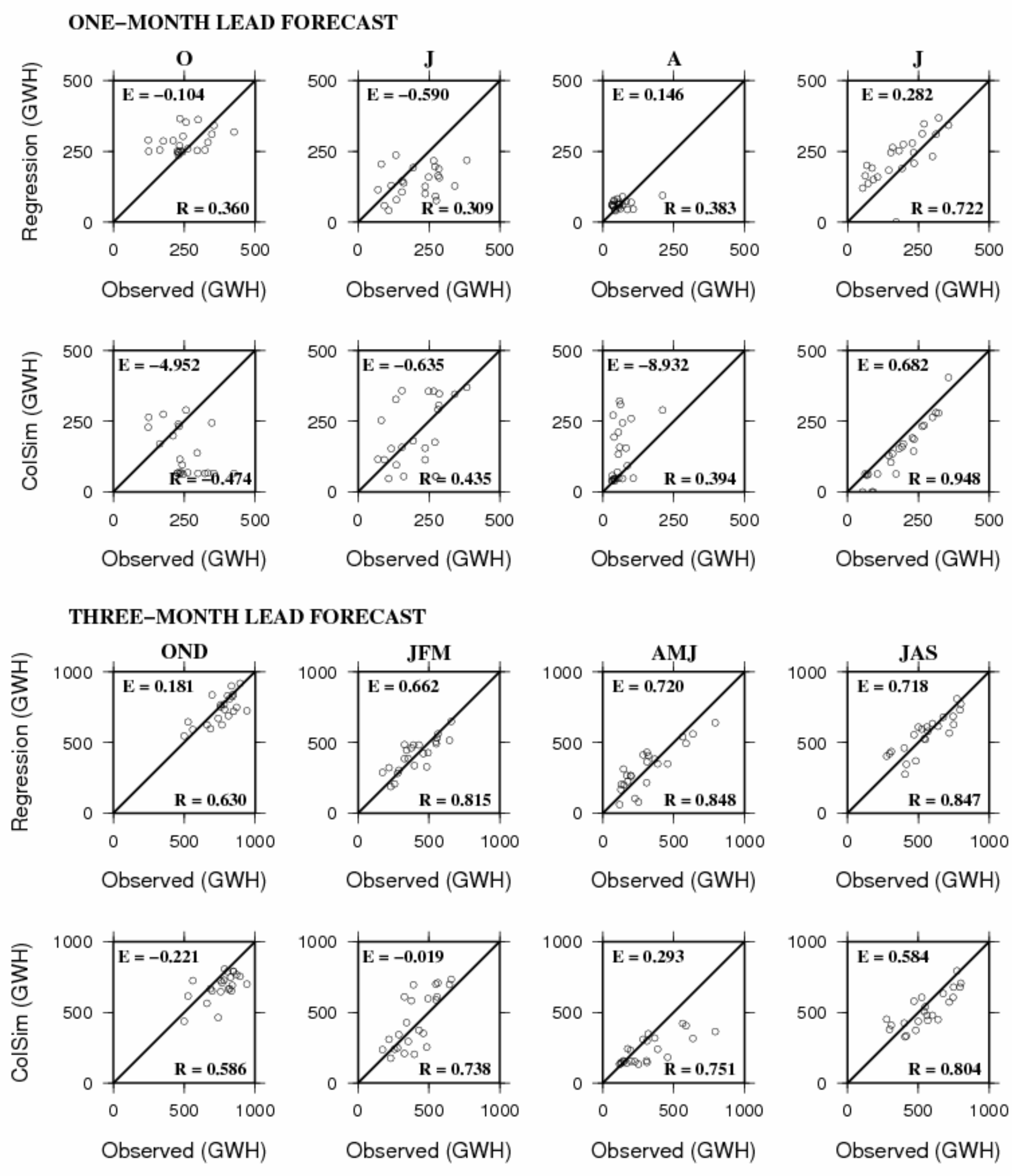

Figure 20. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for Libby Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

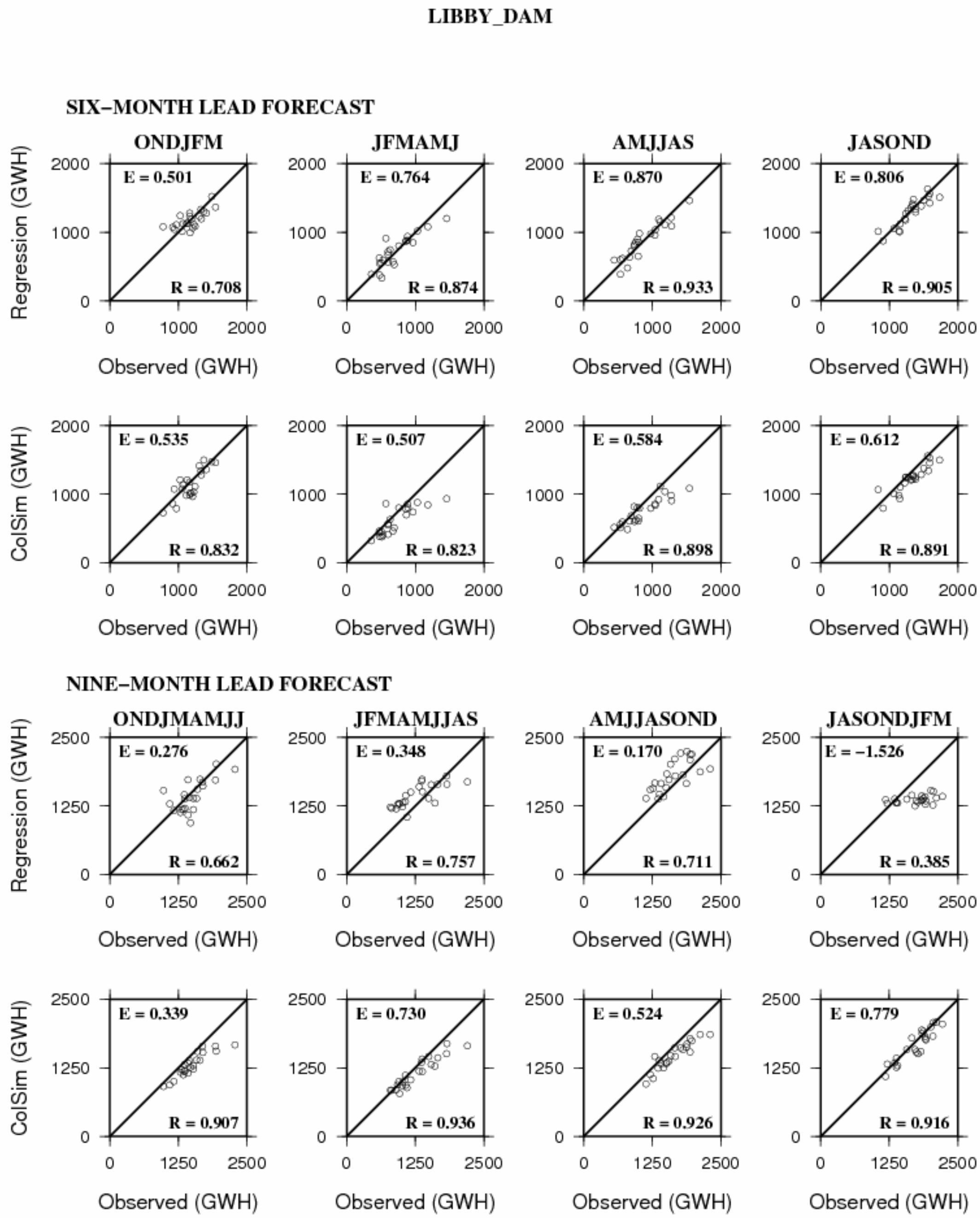

Figure 21. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for Libby Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

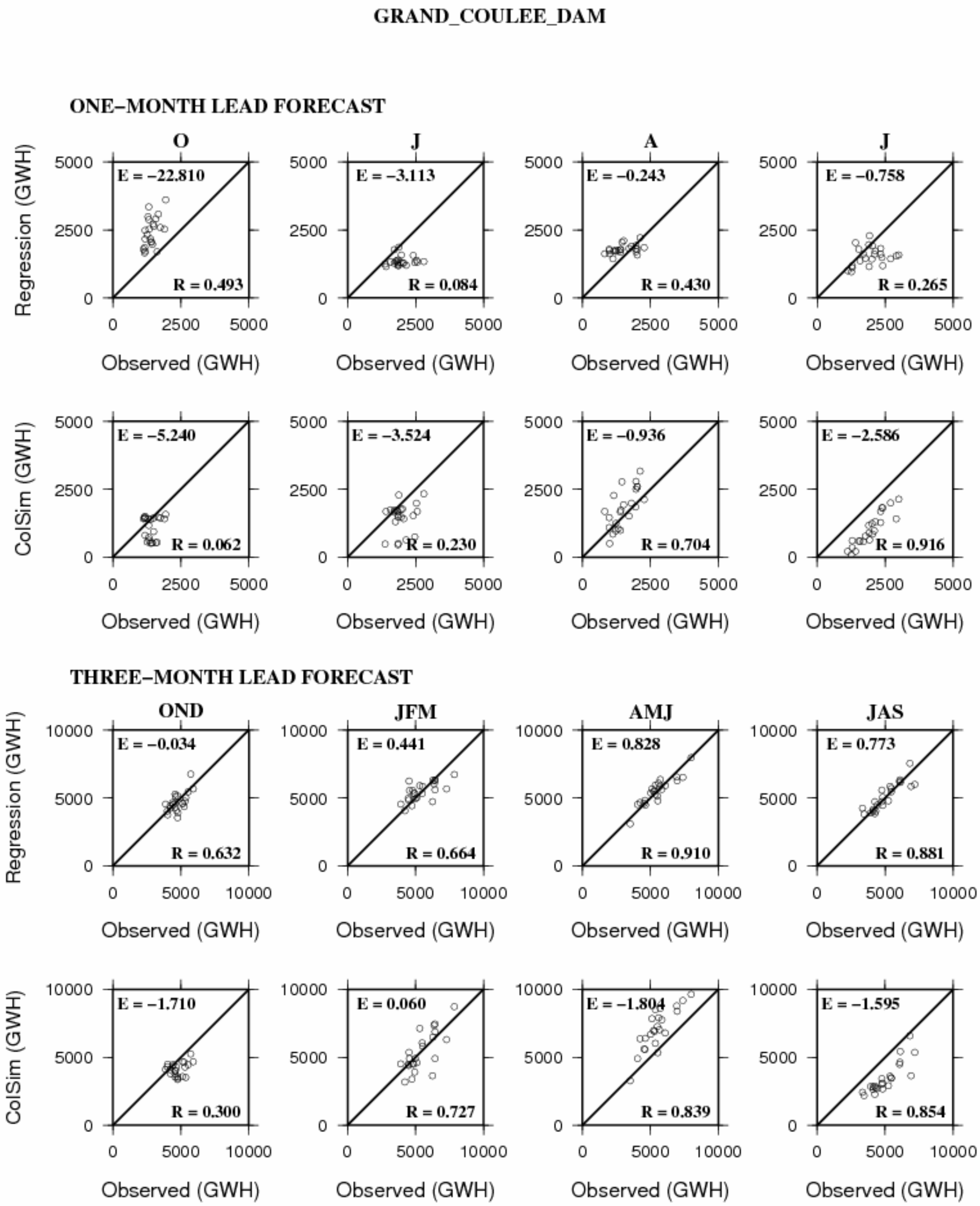

Figure 22. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for Grand Coulee Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

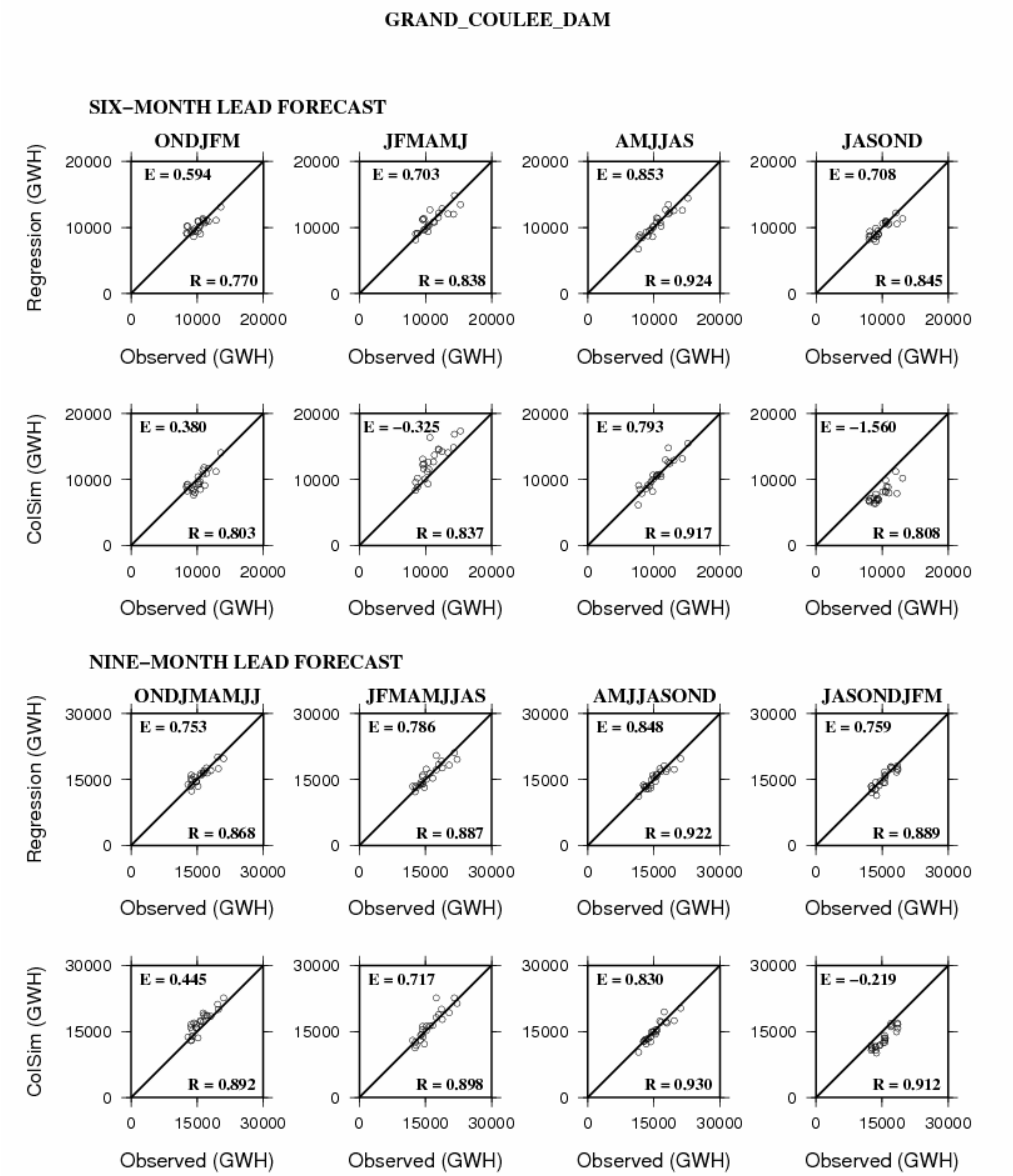

Figure 23. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for Grand Coulee Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

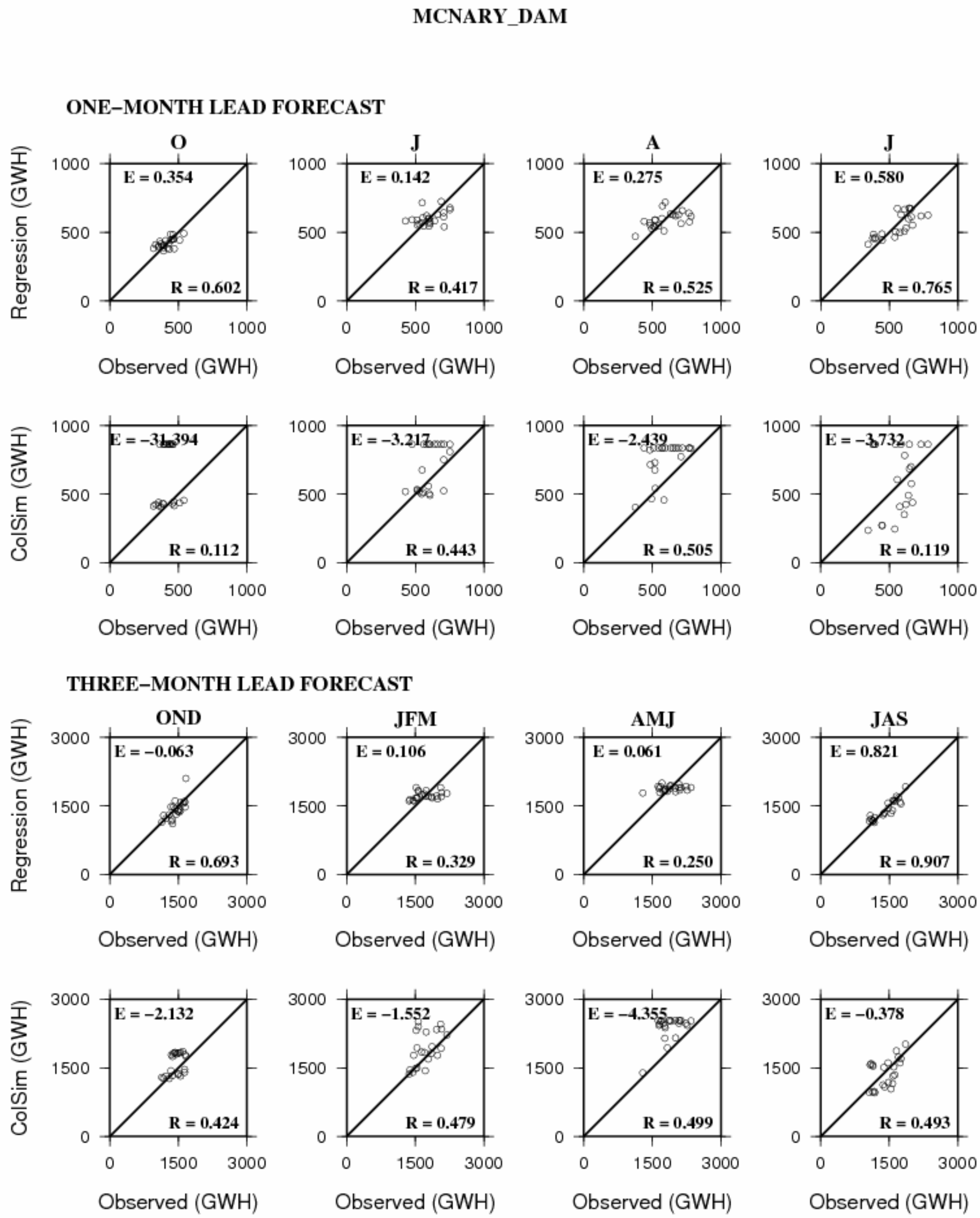

Figure 24. Scatter plots of one- and three-month lead time forecast performance (model vs. observations) for McNary Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

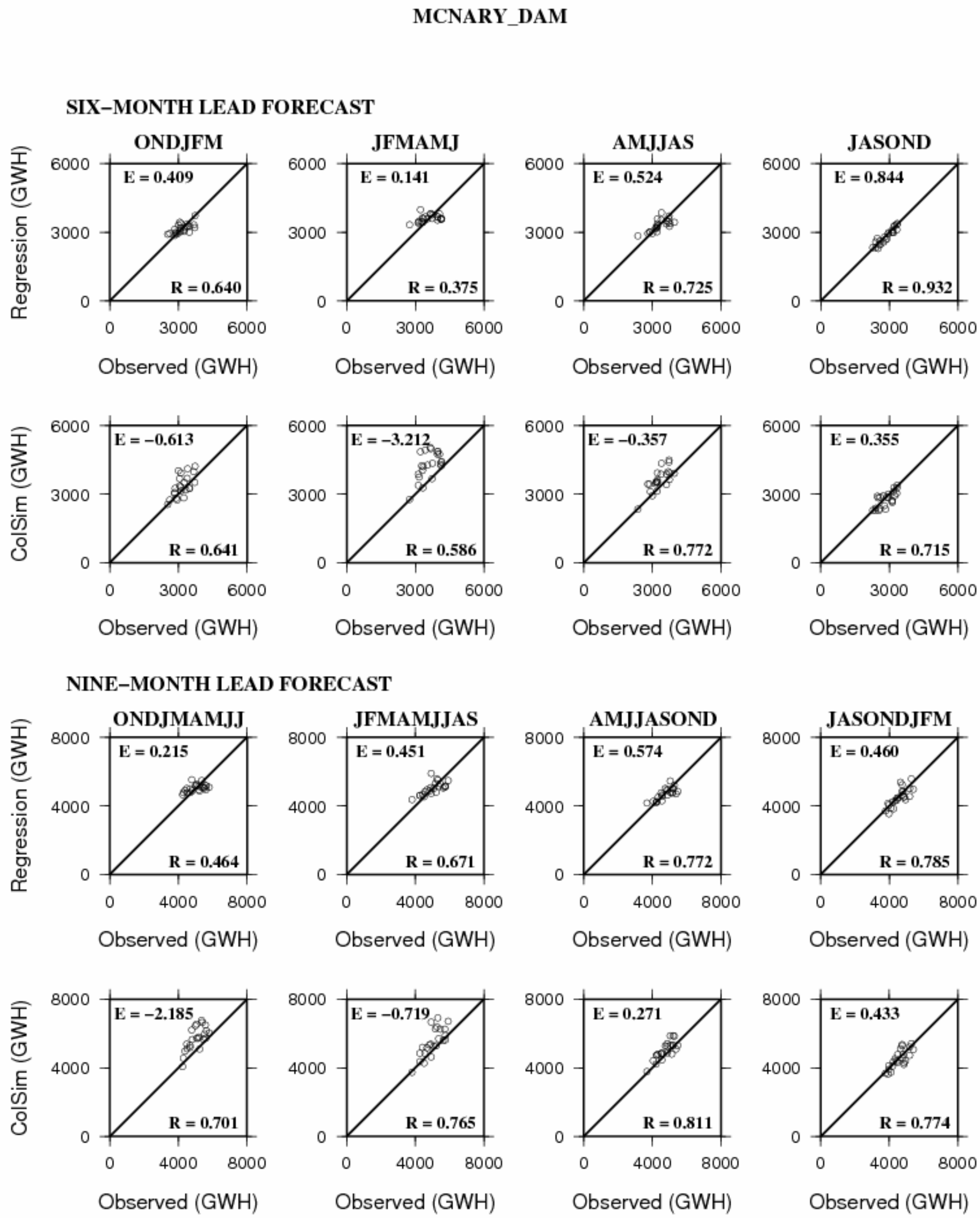

Figure 25. Scatter plots of six- and nine-month lead time forecast performance (model vs. observations) for McNary Dam for two modeling approaches and four forecast dates. E is the Nash Sutcliffe efficiency index; R is the correlation coefficient.

3.3. ColSim and GENESYS Intercomparison

As mentioned above, GENESYS can only be run from water year 1929 to 1978 in its current form. Since observations are only available from 1976 forward, there are only a few years of simulations that can be directly compared to observations. To gain a better sense of how GENESYS performs retrospectively, GENESYS and ColSim hydropower simulations were compared to observations for three representative projects: Libby, Grand Coulee, and The Dalles (Figure 26). Although all three time series only overlap for a few years, it is still possible to gain a reasonably clear sense of the GENESYS model's performance from these comparisons.

At Libby Dam, GENESYS has a systematic timing shift in comparison with ColSim, and comparison with post-1976 observations suggests the ColSim simulation is more accurate at this project. At Grand Coulee the performance of the two models is broadly comparable. Comparison with observations at The Dalles post 1976 suggests that ColSim systematically overpredicts hydropower production during spring peak flows, and GENESYS appears to simulate a more realistic seasonal pattern of hydropower production at this project. Comparisons at a number of other projects (not shown) were broadly similar in character and neither model was clearly superior to the other in these comparisons. It is also clear from the results presented in previous sections that regression models, where they can be applied, perform better than both GENESYS and ColSim in retrospective simulations.

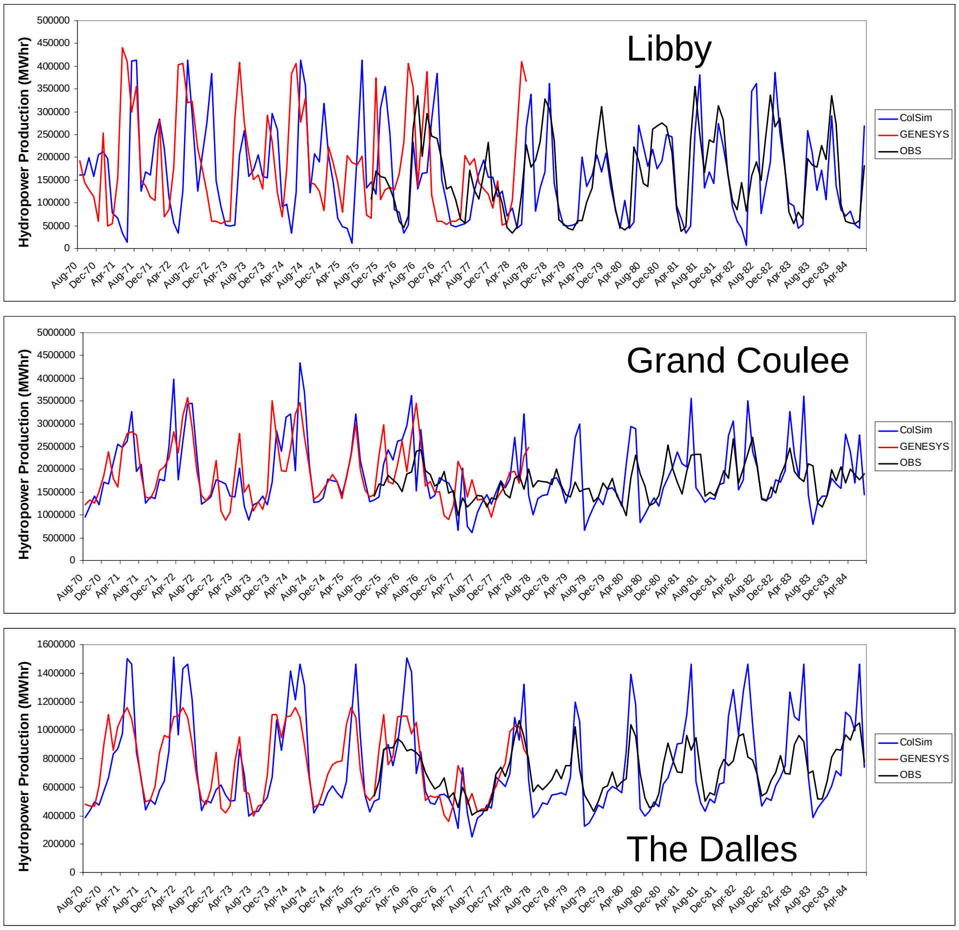

Figure 26. Comparison of ColSim and GENESYS hydropower simulations at three representative projects from 1970–1984.

3.4. Summary of the Strengths and Weaknesses of Different Approaches

This section discusses the strengths and weaknesses, as well as the sources of skill and uncertainty, in the three modeling approaches that were evaluated. Table 6 summarizes the findings.

Table 6. Overview of strengths and weaknesses of the three modeling approaches

Model Type	Strengths	Weaknesses
Regression	• Best overall performance • Lowest cost	• Limited project application due to available data • Application to long term retrospective studies and climate change experiments is problematic
Medium complexity reservoir model	• Simulation of projects limited only by availability of flow data • Flexible model structure and changes to operating rules are relatively easy • Intermediate costs of implementation • Comparable performance to operational model in PNW • Appropriate for long range retrospective studies and climate change assessments • Adapts some reservoir operations to changing flow patterns associated with climate change	• Performance inferior to regression in PNW and California • Higher implementation cost than regression • Lower performance than operational models in California • Less operational detail and simulated projects than operational models in PNW and California • Must purchase STELLA software
Operational reservoir model	• Simulation of projects limited only by availability of flow data • Include more comprehensive and realistic operational detail • Superior performance to medium complexity reservoir model in California • Integrated simulation of conventional energy resources in the PNW • Appropriate for long range retrospective studies and climate change assessments	• Performance inferior to regression in PNW and California • Highest cost of implementation • Decreased flexibility in comparison with medium complexity models • Source code is not available for the PNW • Must purchase FORTRAN 90 compiler and XA linear programming solver

3.4.1. Regression Models

Regression models present an attractive option where data for implementation are available, because they are easily developed and implemented, and, for recent retrospective studies, they outperform even the most sophisticated operational reservoir simulation models at most of the locations evaluated.

It is important to note that in the retrospective simulations, the regression models had a marked advantage over the other approaches because they had available as predictor variables "perfect" information about reservoir elevation and inflows at each time step. This was in contrast to reservoir simulation models which, in the retrospective evaluations, were expected to simulate the trajectory of reservoir storage without assimilation of observed reservoir contents. Thus, the regression models did not suffer from any cumulative errors in simulations of reservoir conditions.

In the forecasting evaluations, however, when all approaches assimilate observed initial reservoir contents, this characteristic of the regression models can become a potential disadvantage at the longer lead times, because the regression models do not incorporate information about how reservoir elevations change during the forecast period. This weakness was most pronounced at storage reservoirs, and less pronounced at run-of-river projects where the head did not vary greatly at monthly time scales.

Insufficient data was available for split sample evaluation of the regression results, and it is to be expected that actual forecast performance will be somewhat lower than what is shown in the tests.

The regression models were strongly limited in their applicability by available data resources. Only six sites in California and eleven sites in the PNW, for example, could make use of regression approaches, due to data limitations. Similarly, climate changes studies or retrospective studies prior to about 1975 (PNW) and 1980 (California) were problematic using regression approaches, because appropriate observed reservoir elevation data were not available.

Alternate approaches using only flow as a single explanatory variable (see section on methods) could provide appropriate retrospective simulations, however, and in some cases these approaches perform almost as well as the multiple linear regression approaches.

3.4.2. Medium Complexity Reservoir Simulation Models (ColSim, CVMod)

Medium complexity reservoir simulation models are an attractive alternative when the implications of changing conditions or reservoir operations need to be considered, because the models can easily be reconfigured to include these kinds of changes. The models also represent a wider range of projects and are not dependent on observed hydropower or reservoir elevation data (only streamflow data are needed). Thus, medium complexity reservoir models can be used for retrospective studies over a much longer time frame (e.g., 1915–2003) and can offer more complete coverage of projects within each system.

In retrospective studies without storage updating from observations, both ColSim and CVMod suffered markedly in comparison with regression approaches in part because of cumulative errors in simulated reservoir storage. This is particularly true of CVMod, where small systematic errors in irrigation demand estimates frequently resulted in large cumulative errors in the storage simulations at the end of long low-flow sequences (not shown). When storage updating was included in the forecasting experiments, however, the performance of the medium complexity models improved markedly, and for the longer forecast lead times (six and

nine months for ColSim and nine months for CVMod), ColSim and CVMod were generally equal to or superior to regression approaches.

The comparison between CVMod and CALSIM II simulations suggests that including additional operational details improves performance significantly in some cases (particularly at Folsom). However, the operational model was not superior at all sites, and it is less clear whether these marginal improvements at specific projects are worth the investment in additional resources needed to apply the more complex operational models in a decision support system. Such tradeoffs may also depend on the specific application details. Furthermore, it appears that with modest effort, some of the CVMod simulation errors could be markedly reduced by refining the regression equations used inside the model for hydropower calculations (i.e., put the superior regression equations developed in this study inside the model).

A comprehensive comparison between ColSim and GENESYS was not possible because only a few years of hydropower observations were available for the intercomparison period. However, based on the limited results, neither model was clearly superior to the other at all projects.

Simulations in the PNW using medium complexity tools were generally superior (in terms of reproducing observations) to those achieved in California.

3.4.3. Operational Models

Operational models present the most attractive alternative under the following circumstances:

- When changes to the modeled system are not expected (which would require expensive or potentially infeasible changes to the model code).
- When simulation of the greatest number of projects and/or specific operational details are most important to the decision processes being supported.
- When the greater resources needed to operate and maintain the more complex models are worth the marginal improvements in performance that these more sophisticated tools offer.

The comparison between GENESYS and ColSim (which showed substantially different time series results from 1970–1978 for some projects) suggests that including additional operational details may have some important implications in the PNW at some projects, although GENESYS did not produce better simulations overall. Similarly, a direct comparison between GENESYS and the regression approaches is not feasible at this point, but it is reasonably clear from the intercomparison with ColSim that regression approaches are generally superior in performance to GENESYS as well in the context of retrospective simulations.

In comparing CVMod and CALSIM II, the operational model was superior in several cases (most notably at Folsom), but for two of the six projects in the comparisons, CVMod was marginally superior to CALSIM II. For the other sites, the improvements in hydropower simulations over the medium complexity models were modest. CALSIM II was not superior to the least-cost regression approaches in the retrospective tests, but it simulated many more projects than were feasible using the regression approaches. In addition, for CVMod, long retrospective simulations and climate change investigations are feasible; whereas, this is not the case with regression approaches as formulated here.

Perhaps the most compelling case for using operational models is that they include specific management details that are not in the other models. In particular, GENESYS integrates the use of conventional energy sources with the operation of the Columbia's hydropower system. CALSIM II includes a wider range of management considerations associated with water quality and flow in the San Francisco Bay Delta. If these interactions between other system objectives and hydropower resources are of specific interest, then investment in these more sophisticated tools makes sense.

4.0 Conclusions and Recommendations

This section presents conclusions and recommendations related to the model performance of the models in the context of three types of potential uses outlined in the introduction:

- Retrospective studies.
- Forecasting.
- Climate change planning studies.

4.1. Application to Retrospective Studies

For retrospective hydropower studies over the last few decades, regression based approaches offer the least cost alternative and outperformed both CVMod in the Sacramento/San Joaquin (SSJ) and ColSim in the Pacific Northwest (PNW), as well as the operational CALSIM II implementation for the SSJ.

Regression approaches, however, are limited by available hydropower and reservoir elevation data. Using publicly available data resources, these tools could only be applied at six sites in the Sacramento/San Joaquin and eleven sites in the Pacific Northwest. Long-term retrospective studies using regression approaches before 1980 in the Sacramento San Joaquin and before 1975 in the Pacific Northwest were also problematic for the multiple linear regression approach evaluated above. However, as noted above, if regression using flow as a single explanatory variable provides comparable performance (true in many cases), then long retrospective studies are possible.

The CALSIM II operational model for the SSJ provided marginally better performance in comparison with the medium complexity CVMod model in the retrospective tests, most notably at Folsom. CVMod performed better at Trinity and Nimbus dams in the retrospective tests. Thus, the CALSIM II model offers marginal improvements in performance over CVMod for retrospective studies, albeit at a considerably higher cost. It is also worth noting that additional refinement and calibration of CVMod would probably improve performance without increasing the complexity of the model.

Although direct retrospective comparison with observations was not possible for GENESYS, comparison of hydropower simulations from GENESYS with ColSim simulations over an earlier period (1970–1978) showed substantial differences between the two models at various projects, but neither model was clearly superior at a majority of projects. Updating of GENESYS to allow direct comparison with observations is needed to more fully evaluate the model.

4.2. Application to Real-Time Forecasting

Forecast results were better overall in the PNW than in California. For forecast lead times of one to three months (PNW) and one to six months (California), regression models were typically equal to or superior to ColSim and CVMod (respectively) in reproducing observations in the retrospective hindcasts. If observed data can be obtained from other sources to permit

application to more projects, regression models would offer a simple, cost-effective approach to producing real time forecasts at these lead times.

Although implementation costs are higher for ColSim and CVMod, these tools offer more complete coverage of projects within each basin, and for lead times greater than three months (PNW) and nine months (California) (depending on location and forecast date), the performance is comparable to regression equations. Forecast results were also better overall in the PNW than in California, particularly at the one month lead time.

Operational models were not tested in forecast mode, but retrospective evaluations discussed above suggest that comparable results to those produced by CVMod and ColSim would be achieved in using the operational models.

4.3. Application to Climate Change Studies

For climate change studies, regression models (at least as implemented here) are problematic to apply because future reservoir elevation data is not available. Regression models using flow as a single explanatory variable are feasible for climate change work, but raise concerns about parameter stationary that are difficult to address. Both operational and medium complexity models have frequently been used in climate change investigations. However, because fidelity to current operational detail is probably not a major factor in assessing future hydropower resources with lead times of 50 years or more, flexible research models like CVMod and ColSim offer some practical advantages for climate change studies over operational models like CALSIM II and GENESYS. In particular, the ability of ColSim and CVMod to automatically create new operating parameters as a function of flow (based on current rules) and incorporate alternate operating strategies in response to changing seasonal flow volumes is a desirable feature in the context of climate change assessments. Thus, in further evaluating operational and medium complexity reservoir simulation models, tradeoffs between implementation costs, model flexibility, ability to reproduce observations, and the importance of operational detail would need to be considered.

5.0 References

Draper A. J., A. Munevar, S. K. Arora, E. Reyes, N. L. Parker, F. I. Chung, and L. E. Peterson. 2004. "CalSim: Generalized model for reservoir system analysis." *ASCE J. of Water Resources Planning and Management* 130: 480–489.

Hamlet, A. F., and D.P. Lettenmaier. 1999. "Effects of climate change on hydrology and water resources in the Columbia River Basin." *J. Am. Water Res. Assoc.* 35: 1597–1623.

Nash, J. E., and J. V. Sutcliffe. 1970. "River flow forecasting through conceptual models part I — A discussion of principles." J*ournal of Hydrology* 10 (3): 282–290.

National Marine Fisheries Service (NMFS). 1995. Biological Opinion: Reinitiation of Consultation on 1994–1998 Operation of the Federal Columbia River Power System and Juvenile Transportation Program in 1995 and Future Years. Endangered Species Act, Section 7 Consultation. National Marine Fisheries Service, Northwest Region. March 2, 1995.

National Marine Fisheries Service (NMFS), U.S. Army Corps of Engineers, Bonneville Power Administration and U.S. Bureau of Reclamation (NMFS et al.). 2000. Draft Biological Opinion. National Marine Fisheries Service, Northwest Region. July 27, 2000.

VanRheenen, N. T., A. W. Wood, R. N. Palmer, and D. P. Lettenmaier. 2004. "Potential implications of PCM climate change scenarios for Sacramento-San Joaquin River Basin hydrology and water resources." *Climatic Change* 62: 257–281.

Voisin, N., A. F. Hamlet, L. P. Graham, D. W. Pierce, T. P. Barnett, and D. P. Lettenmaier. 2006. "The role of climate forecasts in western U.S. power planning." *Journal of Applied Meteorology* 45 (5): 653–673.