

PNNL-39230

# Flow Forecasting on New England's Great River Hydro (CRADA 549)

April 2026

1Cameron Bracken  
2Youngjun Son  
3Mucun Sun (INL)  
4Vincent Tidwell

## **NOTICE**

This report was produced by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830 with the Department of Energy. During the period of commercialization or such other time period specified by the Department of Energy, the Government is granted for itself and others acting on its behalf a nonexclusive, paid-up, irrevocable worldwide license in this data to reproduce, prepare derivative works, and perform publicly and display publicly, by or on behalf of the Government. Subsequent to that period, the Government is granted for itself and others acting on its behalf a nonexclusive, paid-up, irrevocable worldwide license in this data to reproduce, prepare derivative works, distribute copies to the public, perform publicly and display publicly, and to permit others to do so. The specific term of the license can be identified by inquiry made to the Contractor or DOE. NEITHER THE UNITED STATES NOR THE UNITED STATES DEPARTMENT OF ENERGY, NOR ANY OF THEIR EMPLOYEES, MAKES ANY WARRANTY, EXPRESS OR IMPLIED, OR ASSUMES ANY LEGAL LIABILITY OR RESPONSIBILITY FOR THE ACCURACY, COMPLETENESS, OR USEFULNESS OF ANY DATA, APPARATUS, PRODUCT, OR PROCESS DISCLOSED, OR REPRESENTS THAT ITS USE WOULD NOT INFRINGE PRIVATELY OWNED RIGHTS.

**Printed in the United States of America**

**Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831-0062**

**[www.osti.gov](http://www.osti.gov)**

**ph: (865) 576-8401**

**fox: (865) 576-5728**

**email: [reports@osti.gov](mailto:reports@osti.gov)**

**Available to the public from the National Technical Information Service  
5301 Shawnee Rd., Alexandria, VA 22312 ph: (800) 553-NTIS (6847)**

**or (703) 605-6000**

**email: [info@ntis.gov](mailto:info@ntis.gov)**

**Online ordering: <http://www.ntis.gov>**

# Flow Forecasting on New England's Great River Hydro (CRADA 549)

April 2026

1Cameron Bracken  
2Youngjun Son  
3Mucun Sun (INL)  
4Vincent Tidwell

Prepared for the U.S. Department of Energy under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory  
Richland, Washington 99354

## CRADA Final Report Certification Memo

**Memo Completion Date: 8/29/2025**

**CRADA Number:** PNNL CRADA 549, INL CRADA 22CRA22

**CRADA Title:** Flow Forecasting on New England's Great River Hydro

**Responsible Technical Contact at DOE Lab(PNNL):** Vincent Tidwell  
(vincent.tidwell@pnnl.gov)

**Provide a list of publications, conference papers, or other public releases of results, developed under this CRADA:**

Conference Presentations:

Bracken, C., Tidwell, V. and Ragonese, J., 2024. Hydropower Scheduling Oriented Inflow Forecast Evaluation for Great River Hydro, HEPEX, Tuscaloosa, AL, , 2025.

Tidwell, V., Bracken, C., Son, Y., Canaday, J. and Ragonese, J., 2025. Hydropower Scheduling Oriented Flow Forecast Evaluation for Great River Hydro. 2025 CEATI HOPIG Meeting Streamflow Forecast Interest Group, Palm Springs, CA, March 27, 2025.

**Provide a detailed list of all subject inventions, to include patent applications, copyrights, and trademarks:**

No subject inventions were generated under this CRADA.

## Executive Summary of CRADA Work

Changing regulation, increasing demands and growing operational uncertainties are requiring more precise operations of our nation's reservoir and hydropower systems. This desire for increased precision has driven interest in improved streamflow forecasts and operational hydropower scheduling tools. While numerous commercial and in-house options are available, system operators are faced with the challenge of evaluating whether a more sophisticated forecasting and/or scheduling tool will actually improve their operations and justify the ongoing financial investment. Here an enhanced evaluation process is demonstrated as a case study aimed at assisting Great River Hydro with such decisions.

Great River Hydro (GRH) is New England's largest conventional hydropower generator, operating 13 generating stations and 3 non-generating storage reservoirs along the Upper Connecticut River in Vermont, New Hampshire, and Massachusetts. Three of the GRH managed reservoirs have recently been part of a FERC relicensing process. While the new license imposes tighter run-of-river constraints on reservoir operations, it also provides new opportunities in the form of flex hours where deviated operations are allowed to capture favorable market conditions or emergency operations. To maximize value from these new operations, GRH sought technical assistance to evaluate potential improvements in inflow forecasting and scheduling accuracy, particularly during short- to medium- duration periods (1-10 days). At issue was whether improved forecasting and scheduling accuracy would lead to enhanced efficiency in the way GRH utilizes water, improving their ability to schedule higher releases during the best priced hours throughout the system and enhancing revenues as a result.

Pacific Northwest National Laboratory (PNNL) in cooperation with Idaho National Laboratory (INL) provided support to GRH under DOE's Water Power Technologies Office's Notice of Opportunity for Technical Assistance: Improving Hydropower's Value through Informed Decision Making. This effort was organized into three tasks. Tasks 1 and 2 were the responsibility of PNNL while INL was responsible for Task 3.

The first task was to provide an objective evaluation of the accuracy of inflow forecasts available through commercial vendors. The evaluation involved the head-to-head comparison of two new vendor forecasts and the currently utilized forecast from the Northeast River Forecast Center all against the baseline of measured inflows at three relicensed GRH dams. Evaluation of the short-term forecast products utilized real time data collected over a roughly two-year period of time. Unique to the evaluation was the employment of multiple performance metrics that measured different aspects of goodness of fit (e.g., Nash-Sutcliffe Efficiency, King-Gupta Efficiency, Percent Bias, normalized Root Mean Square Error). Working directly with GRH operators, tailored metrics were also designed and implemented that address unique aspects of their operations. These tailored metrics focused on accuracy of forecasted reservoir storage recovery rates related to the utilization of flex hours. Evaluation results were published to an interactive visualization tool that allowed comparison across forecasts as well as across the different metrics. By using multiple traditional and tailored metrics as well as making the data easily accessible and comparable, hydropower operators like GRH are provided with the information they need to make informed decisions on forecast products with confidence. Note that this approach can easily be replicated on other river systems.

The second phase of the project involved the evaluation of a new scheduling tool to identify windows of opportunity to change operations in ways that increase profit through participation in the day ahead market. Compared was the current GRH Excel-based scheduler against a new higher-resolution operations model, each designed to optimize value across the system or at individual locations based upon available storage, inflow, water utilization and generation. The evaluation involved operating the lower three GRH dams according to the schedule developed from one of the two models. Schedulers were rotated on a weekly basis. Schedule efficacy was evaluated on the basis of system compliance, hydropower generation and revenue. An unforeseen issue caused the conclusion of the project prior to the collection of sufficient data (collected approximately 2 weeks' worth of data) for full system evaluation.

The third phase of the project involved identification of opportunities to increase market revenue through improved use of forecasts and scheduling models. Building on the forecast evaluation of Task 1 and the scheduling model comparisons of Task 2, this task focused on translating those technical improvements into financial outcomes. Specifically, the team worked with Great River Hydro operators to explore how enhanced forecasts and scheduling tools could be used to target discretionary Flex hours toward higher-value market periods while maintaining compliance with new licensing requirements. The evaluation considered a range of operational scenarios, from conservative schedules that prioritize strict compliance to more opportunistic strategies that adjust dispatch in response to market conditions. These scenarios were assessed with respect to both near-term (day-ahead and real-time) operations as well as potential future markets where multi-day (7–10 day) scheduling may be required. Although an unforeseen issue limited the ability to fully test these opportunities under extended real-time operations, the analysis demonstrated a clear framework for connecting forecast and scheduling accuracy to revenue gains. This framework provides hydropower operators with a structured approach to assess the economic benefits of investing in advanced forecasting and scheduling tools.

This project makes use of the laboratory's core competencies that are not available in the hydropower industry. The project reflects a common need in the hydropower industry to evaluate flow forecast system adequacy for day-to-day management needs. Enhancing and better integrating flow forecasts into hydropower operations will further position the grid for reliable delivery of electricity which is aligned with federal government goals. Importantly, improved accuracy in scheduling will result in increased hydropower generation, better flexibility to contribute needed grid services, and ultimately improve utility profits. In turn, the public benefits from reliable and affordable electricity service and improved coordination among other watershed resource sectors.

# Summary of Research Results

## 1.0 Introduction

### 1.1 Background

Great River Hydro (GRH) is by far New England’s largest conventional hydropower generator. As a private, independent power producer, GRH owns and operates 589 megawatts of nominal hydropower capacity in ISO New England. GRH’s fleet consists of 13 generating stations and 3 non-generating storage reservoirs under 5 separate FERC licenses. See Figure 1 for a map of Great River Hydro’s facilities.

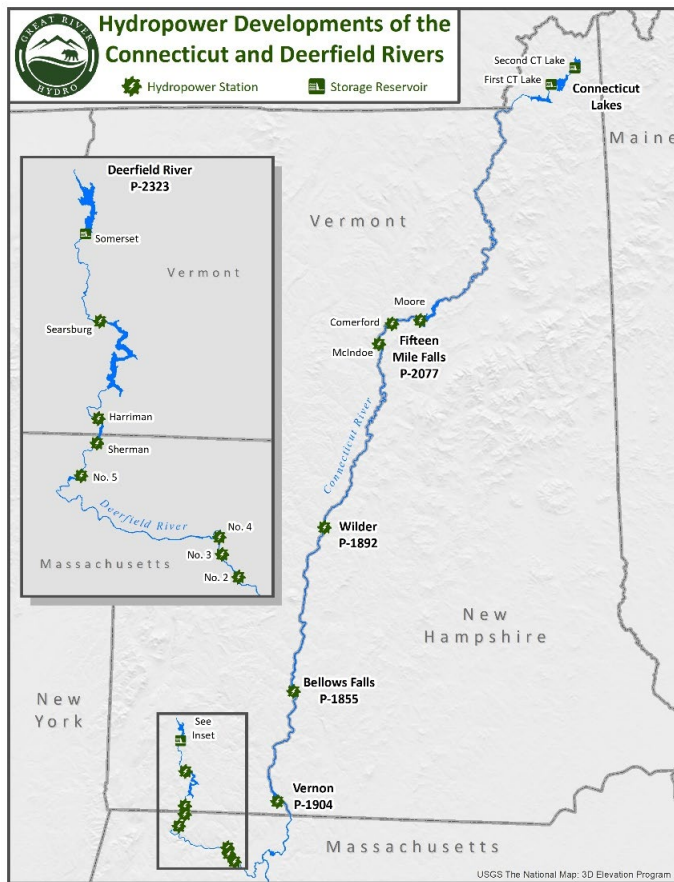


Figure 1 Map of Great River Hydro’s Facilities

Six of GRH’s generating stations and 2 storage-only reservoirs (with a third storage-only reservoir under operational control via a use agreement with the State of New Hampshire) are located on the Connecticut River in northern New England. GRH’s Connecticut River system reaches from the headwaters in the very northern tip of New Hampshire and continues south along the common boundary with Vermont nearly to the Massachusetts state line, with a total gross drainage area of 6,266 square miles.

GRH’s three most downstream Connecticut River plants, the Wilder Station, Bellows Falls Station, and the Vernon Station, are each individually licensed (Table 1). All three licenses were scheduled to expire on April 30, 2021. As such, these three lower Connecticut River hydro projects were recently the subject of relicensing with FERC, a process that began in 2012 and culminated with the issuance of new 40-year operating licenses in 2022.

Table 1. GRH hydropower stations under recent relicensing.

Station	Capacity (MW)	FERC License No.	Target Elevation (ft)
Wilder	41	P-1892	384.5
Bellows Falls	49	P-1855	291.1
Vernon	37	P-1904	219.63

## 1.2 Problem

Great River Hydro’s operations are manually scheduled based on National Weather Service Northeast River Forecast Center (NRFC) forecasts, regional weather forecasts, market forecasts, and operations experience spanning many decades. The tools, data sources, and existing models that Great River Hydro utilizes to manage its Connecticut River facilities are quite good at handling longer-duration (monthly and seasonal) water management. Additionally, available flow forecasts (i.e. Ensemble1 forecast of the NRFC), are particularly useful during high-flow events (during those events, elevation predictions are solid but flow predictions are not as accurate). The medium range (7-10 days) and short range (1-7 days) flow forecasting is much less accurate, especially under “normal” conditions, for this large and diverse watershed, particularly for several of the largest contributing tributaries. Furthermore, NRFC’s flow forecast locations do not directly forecast all tributaries and headwater inflows, rather account for them in a Connecticut River mainstem forecast as specific gage locations for NRFC’s flood forecast objectives. Current operating regimes of Great River Hydro’s assorted facilities on the Connecticut River are flexible enough to manage the variations in actual versus predicted flows; however, new operational regimes for the three lower Connecticut River projects substantially tighten the operational ranges that those facilities must operate within. GRH has been investigating third-party inflow forecasting services that would provide grid-scale inflows but has yet to hire or utilize their services independently.

GRH currently utilizes an in-house developed tool called the Megawatt Scheduler (MS) to manage each of its river systems. The MS is a Microsoft Excel based model that is compiled daily. Following seasonal river and reservoir management “rule curves”, operations and marketing staff input flow forecast data (using a combination NRFC data, USGS gauge data, and historical flow data), hourly day-ahead pricing data (using a combination of pricing forecast data from Enverus2 cross-referenced against past market data from recent similar days), and unit availability and efficiencies. The result is a Megawatt Run Schedule for the entire river system, which is then bid into the ISO New England day ahead market. Over the course of each day, GRH operations staff use the MS to track and confirm the actual hourly megawatt generation and hourly inflow to keep each site compliant with the respective FERC Licenses.

While the new licensing of the lower three reservoirs on the Connecticut River reduced flexibility in operations (move toward an inflow-equal-outflow regime), some flexibility was returned in the form of discretionary generation hours, which are referred to as Flex hours. GRH

is allowed a set number of Flex hours on a per month basis to target releases to favorable market conditions. These flex hours (i.e., a single hour in which reservoir releases are allowed to exceed inflow conditions) include strict operational requirements for up-ramping, down-ramping and impoundment refills. To aid GRH in relicensing negotiations, Hatch (a consulting engineering firm, see hatch.com) developed an operations and scheduling model to evaluate the effects of an inflow-equals-outflow (IEO) operation and IEO plus limited discretionary Flex generation (IEO/Flex). This tool is not currently implemented by GRH. Of interest is whether potential improvements to scheduling accuracy (i.e., through use of the Hatch model) would enhance the efficiency with which GRH utilizes available water, improving their ability to hit the best priced hours throughout their system and enhancing revenues as a result.

GRH currently schedules and markets much of its production through the markets established by the New England System Operator (ISO-NE). There are potential changes in market designs which would look to schedule or commit generation further into the future than present day-ahead; perhaps 7-10 days in advance. Thus, there are opportunities to increase revenues by evaluating the bidding strategies in the real-time, day-ahead and 7-10 day in advance markets.

### **1.3 Project Formulation**

Toward these challenges Great River Hydro sought technical assistance to evaluate potential improvements in inflow forecasting and scheduling accuracy in light of new operational procedures implemented as a result of relicensing of the three hydropower generation resources on the lower Connecticut River. Support was obtained through DOE/EERE/WPTO Notice of Opportunity for Technical Assistance (NOTA): Improving Hydropower's Value through Informed Decision Making: Topic Area 2, Assessing the Value of Inflow Forecasting Tools and Practices. Pacific Northwest National Laboratory and Idaho National Laboratory teamed with GRH to provide the requested technical support.

### **1.4 Objectives**

Based on the needs of GRH, three technical objectives were identified:

1. Increase the accuracy of medium (3-10 day) and short (0-48 hour) range flow forecast during a range of flow conditions that fall within normal (non-spill) operation.
2. Evaluate the risk of non-compliance operation and loss of opportunity to maximize the value of available discretionary hours when operating under new relicensing operation for three Lower Connecticut projects.
3. Explore alternate scheduling paradigms to maximize GRH's ability to manage new operation in compliance with operational constraints and opportunistically utilize discretionary flexible hours.

### **1.5 Project Tasks**

To accomplish these objectives work was organized in three separate but interacting tasks. A general outline of each task is given below.

Task 1: GRH system operations under flow forecast uncertainties and new regulation, enhance forecast value to GRH operations: PNNL was responsible for this task. This task evaluated flow forecast products from different providers based on their value to the GRH system and provided recommendations to improve the GRH flow forecast system. Standard flow forecast verification metrics were reviewed and evaluated in light of their efficacy under different flow regimes. Additionally, new metrics were developed with GRH that were tailored to address operational questions unique to GRH. While all flow conditions were addressed, inflow forecasts within the normal (non-spill) operating range of the hydro stations were given priority. Additionally, a review of the GRH flow forecast system was performed to identify system improvements that might enhance the accuracy of the forecasts, particularly the identification of finer resolution inflow forecast (e.g., tributary inflows) locations.

Task 2: GRH system operations under flow forecast uncertainties and new regulation, evaluation of the GRH system operability under flow forecast uncertainty and new operations: PNNL was responsible for this task. This task evaluated two alternative tools for operating and scheduling hydropower resources within the GRH system. Comparison was drawn between GRH's current Megawatt Scheduler and a high-resolution operations model developed by Hatch. The Hatch model has the advantage that it has been set up to evaluate flexibility under the new license rules.

Evaluations were performed under real-time operational conditions. Day-ahead schedules were developed using flow forecasts from Task 1. The hydropower system was then operated according to the developed schedule. Schedulers were alternated on a weekly basis (i.e., Hatch model was used to schedule for a week and then operations were scheduled with the Megawatt Scheduler the next week). The subsequent evaluation focused on the scheduler's ability to accommodate the new compliance operations when under forecast uncertainties. Of particular interest was whether the scheduler would enhance the value of the forecasts by either or both increasing GRH revenues and flexibility to address windows of market opportunities. GRH was responsible for scheduling with their Megawatt Scheduler while Hatch was solely responsible for the operations of their scheduler.

Task 3: Identification of opportunities to increase market revenue: INL was responsible for this task. The focus of this task was to identify opportunities to modify dispatch patterns to increase revenue throughout the entire GRH river system. This work built on the outcomes of the previous tasks by utilizing resultant flow forecasts and scheduling model outputs. Work progressed by first identifying if any opportunities exist and, if so, requirements for realizing them. Based on the outcomes of that, additional steps were taken based on prioritization and resources including activities such as performing additional analysis on specific scenarios identified as being of interest to GRH.

## **1.6 Potential Value to Industry**

Hydropower is recognized as a key energy source needed to reduce the price of electricity and support broad grid services. Enhancing and better integrating flow forecasts into hydropower operations will further position the grid for reliable and affordable delivery of electricity which is aligned with federal government goals.

This project uses the laboratory's core competencies that are not available in the hydropower industry. The project reflects a common need in the hydropower industry to develop flow forecast systems adequate for the day-to-day management needs, in addition to extreme high flow conditions. One anticipated outcome is to support the hydropower industry in better navigating new environmental operations through the use of project-specific flow forecasts.

Great River Hydro will benefit from hydropower-oriented flow forecast expertise to evaluate a range of flow forecast providers and options to enhance their in-house processing tool. Great River Hydro will also benefit from evaluating new environmental operations and their impact under flow forecast uncertainties.

## **1.7 Report Outline**

The remainder of the report is organized according to three chapters that parallel the primary objectives and tasks of this project. Each chapter is prepared as self-contained paper addressing related background, methods, results and recommendations.

## 2.0 Task 1: Evaluation of the GRH System Operability under Flow Forecast Uncertainty and New Operations

### 2.1 Introduction

This task focused on the evaluation of two commercially available forecast products, one being a traditional rainfall runoff model (forecast A) and the other an AI based forecast (forecast B). In addition, we compared the in-house GRH forecast which is based on NRFC forecasts plus in-house routing and operator expertise. Both forecasts extend to 10 days in the future, while the GRH forecast extends only 1-2 days in the future since it is used specifically to cover their bid into the day ahead market.

While probabilistic and ensemble forecasts and evaluation metrics are gaining in popularity, deterministic forecasts such as those issued by the National Weather Service River Forecast Centers are still widely used operationally and thus served as a benchmark of this evaluation. In addition, we focus on metrics which are feasible to compute in a real-time setting as opposed to a hindcast based evaluation. While they may be preferable statistically, hindcast evaluations require that forecasters/operators emulate their own behavior during hypothetical past conditions where other factors such as market prices may not be known. In the remainder of this section, we describe the components of the real-time forecast evaluation framework we used to evaluate the commercial and in-house forecasts.

### 2.2 Methods

The real-time forecast evaluation process can be described in the following steps:

1. **Collect real-time data** – We worked with GRH to develop a real-time data pipeline to access observed and forecast data. This included streamflow forecasts (natural or regulated), calculated inflows, scheduled and actual outflows, and other metadata for computing evaluation-relevant metrics such as forecast issue times and past forecasts.
2. **Compute evaluation metrics** – At the start of an evaluation, it was only possible to compute a limited number of generic evaluation metrics based on small sample sizes but as the available data grew it was possible to compute more reliable metrics as well as metrics tailored to GRH's system.
3. **Visualization** – We developed a convenient interactive web report with built in documentation that could be updated in near real time to provide information on the quality of the forecasts. Visuals were concise, clear, and required a limited amount of explanation. In addition, the report was a single html file which could be sent over email or hosted online.
4. **Solicit feedback** – We collected feedback on an ongoing basis from GRH operators which was critical to the success of a real-time forecast evaluation. This process was used to develop tailored metrics that were specific to GRH's system and operations.
5. **Iterate** – The steps above were repeated at a regular interval throughout the entire forecast evaluation period.

The iterative nature of this evaluation process with direct operator feedback was key to its success. As more data became available so did our understanding of the forecast performance, which helped us to develop additional tailored metrics. The following sections detail some additional aspects of the forecast evaluation.

### **2.2.1 Benchmark Forecasts**

It was beneficial to include certain benchmark forecasts in the evaluation to provide us with context for how well a specific forecast was performing. Two commonly used benchmarks are the “perfect” forecasts, which uses observed or calculated inflows in place of a forecast to get a sense of the best possible forecast, and “persistence” forecasts which use the last observed value carried out for all future timesteps to get a sense of the arguably worst possible forecast. We note that other simple climatologically driven forecasts could be used in place of the persistence forecast. Other benchmarks could include publicly available forecast products or any other forecast that is currently being used at the reservoir. GRH is already using the NRFC forecast for their in-house forecast, which is another reason we went with persistence as a benchmark.

### **2.2.2 Smoothing calculated inflows**

Most available reservoir inflow data is not directly measured, instead it is computed based on measured changes in forebay elevation and a known elevation-volume relationship. These calculations are often noisy due to short term changes in forebay elevations from wind, turbine and spillway release changes, and other factors such as high frequency waves propagating from upstream. Inflow computed in this way can often contain negative values and other non-physical fluctuations. In this situation we recommend the inflow data be smoothed prior to comparing with any forecasted inflow. We recommend a triangular smoother as a simple option which produces reasonable results; however, there are many more options available. Figure 2 shows an example of hourly computed inflow data which has been smoothed using a triangular smoother.

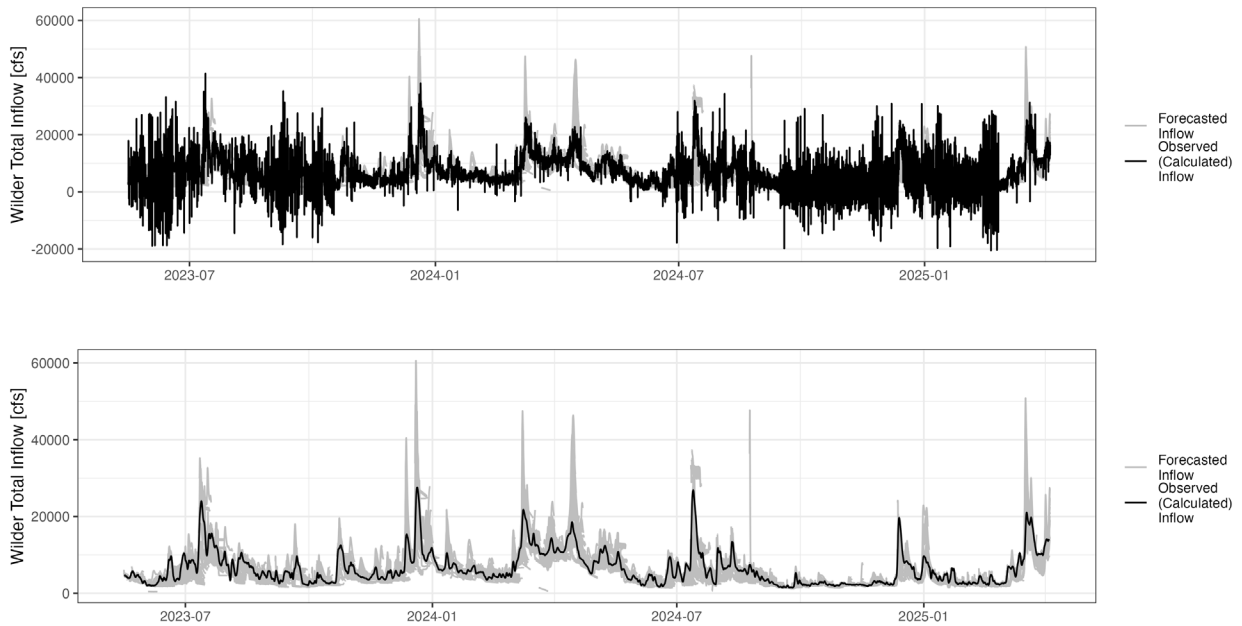


Figure 2: Wilder calculated (observed) inflow unsmoothed (top) and smoothed with a 24 hour triangular smoother (bottom) which was used in this evaluation. Operational forecasts are included for context.

**2.2.3 Traditional inflow evaluation metrics**

Traditional deterministic forecast evaluation metrics that are particularly relevant to inflow forecasting include measures of average behavior, correlation, and bias. Additional metrics include higher order moments of the forecast distributions, categorical forecast performance, and conditional statistics. This section is not intended to be a comprehensive review of deterministic forecast evaluation metrics. For operational applications, simple metrics were computed for a variety of lead times as well as seasonally to fully assess forecast performance across a range of conditions.

Average hydrologic forecast behavior is typically measured with an error metric such as root mean squared error (RMSE), Nashe-Sutcliffe Efficiency (NSE), or King-Gupta Efficiency (KGE). We recommend KGE because it is not skewed by large values like the NSE and RMSE, which is the result of their use of a squared error term.

The NSE is given by

$$NSE = 1 - \frac{\sum(o_i - f_i)^2}{\sum(o_i - \mu_o)^2} \tag{1}$$

Where  $o_i$  is the observation and  $f_i$  is the forecast at time  $i$  and  $\mu_o$  is the mean of the observations and the summation indicates a summation for all timesteps  $i$ . NSE can range from negative infinity to 1 with values greater than 0 indicating a forecast that is better than the mean of the observations.

The KGE is given by

$$\text{KGE} = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_f}{\sigma_o} - 1\right)^2 + \left(\frac{\mu_f}{\mu_o} - 1\right)^2} \quad [2]$$

where  $r$  is the linear correlation coefficient (given below),  $\sigma_f$  is the standard deviation of the forecast,  $\sigma_o$  is the standard deviation of the observations,  $\mu_f$  is the mean of the forecast, and  $\mu_o$  is the mean of the observations. Values of KGE that are greater than -0.41 are better than the mean of the observations, but larger values are preferred up to the maximum value of 1 for a perfect forecast. We note that NSE and KGE have different properties and so should not be used interchangeably

The bias of a single forecast point is the difference between observed and forecasted values but this may be expressed as a percentage for a set of forecasts by averaging across all bias values

$$\% \text{PBIAS} = 100[\%] \times \frac{\sum(f_i - o_i)}{\sum(o_i)} \quad [3]$$

where %PBIAS measures the tendency of a forecast to be greater than (positive values) or less than (negative values) than observations.

Correlation measures the strength of a linear relationship between observations ( $o$ ) and forecasts ( $f$ ) is computed using the Pearson correlation coefficient

$$r = \frac{\sum(o_i - \mu_o)(f_i - \mu_f)}{\sqrt{\sum(o_i - \mu_o)^2 \sum(f_i - \mu_f)^2}} \quad [4]$$

where  $r$  can range from -1 to 1 but for forecast evaluation, negative values would indicate poor forecast performance.

#### 2.2.4 Interactive visualization report

The results of the evaluation were summarized in an interactive web report. Figure 3 shows example screenshots from the web report. The report was contained in a single html file which can be hosted online or easily sent via email. The full html file containing the final evaluation report is included within the project files for reference.

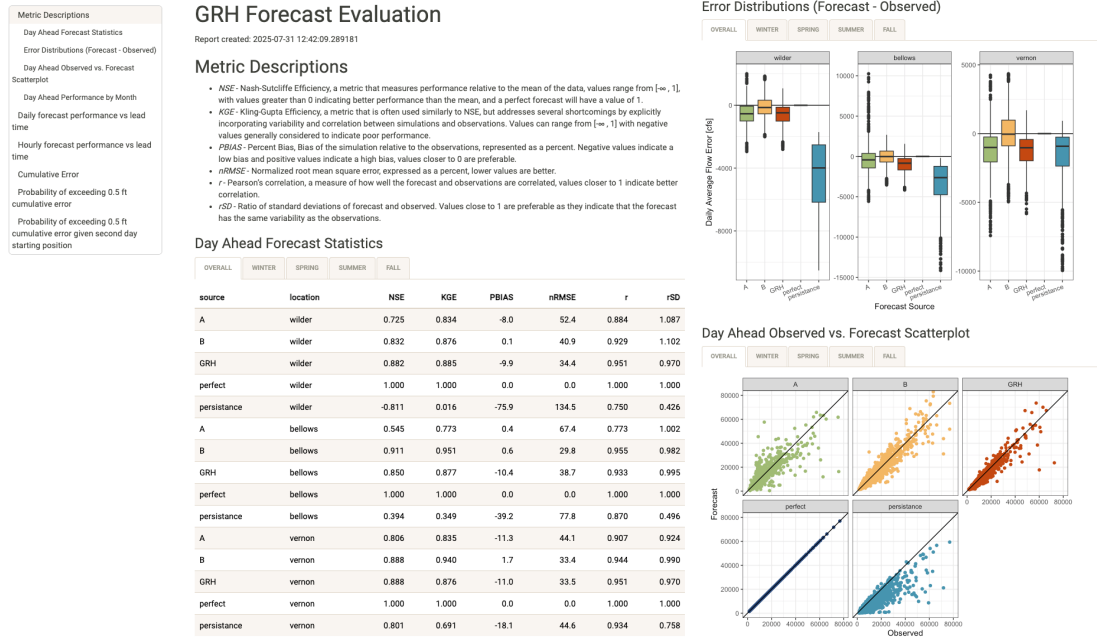


Figure 3: Screenshots from the interactive visualization report.

## 2.3 Results

We ran the real-time evaluation for almost 2 years from July 2023 to May 2025 with regular feedback from the operators. We evaluated two commercial forecast products which are anonymized and labeled as A and B. In addition, we evaluated the GRH in-house forecast which is a combination of NRFC forecasts, routing, and operator judgment. During the initial few months of the evaluation, we focused on generic metrics of forecast performance. As the quantity of data increased, we transitioned to seasonally computed metrics and tailored metrics developed in collaboration with the system operators.

While both commercial inflow forecasts were available hourly for the next 7-10 days, the volumetric forecast for the next day is the most important for GRH. This largely sets what they can bid into the day ahead market. Table 2 shows the overall performance metrics for the forecast evaluation for first day forecast. Overall, we found that for this time period, GRH's in-house forecast was roughly similar to forecast B but better than forecast A overall. This overall tabular view is a good way to get quick comparisons between forecasts but can hide certain forecast behaviors, especially seasonal.

Figure 4 shows the performance of all the available forecasts and benchmarks out to a 10-day lead time. The GRH in-house forecast is only used to develop bids for the day ahead market, so it does not extend the full 10-day period. At Wilder and Vernon all three forecasts perform roughly similar with 1-3 days of lead time with the A and B forecasts diverging after that period. At Vernon, the furthest downstream point, the three forecasts are distinct in terms of quality with the B showing the best performance at most lead times followed by the GRH in-house forecast and then forecast A. Notably all forecasts exhibited some amount of bias at the downstream forecast locations which may be due to errors propagating from upstream.

Table 2: Overall forecast evaluation metrics for the real time evaluation conducted from July 2023 to May 2025 for the two commercial forecasts (A, B), the GRH in-house forecast (GRH), and the benchmarks perfect and persistence forecasts. The table shows metrics for the next day ahead forecast. See section 2.2.1 for details on the metrics.

Forecast	Location	NSE	KGE	PBIAS	r
A	wilder	0.72	0.83	-8.00	0.88
B	wilder	0.83	0.88	0.10	0.93
GRH	wilder	0.88	0.89	-9.90	0.95
perfect	wilder	1.00	1.00	0.00	1.00
persistence	wilder	-0.81	0.02	-75.90	0.75
A	bellows	0.55	0.77	0.40	0.77
B	bellows	0.91	0.95	0.60	0.95
GRH	bellows	0.85	0.88	-10.40	0.93
perfect	bellows	1.00	1.00	0.00	1.00
persistence	bellows	0.39	0.35	-39.20	0.87
A	vernon	0.81	0.83	-11.30	0.91
B	vernon	0.89	0.94	1.70	0.94
GRH	vernon	0.89	0.88	-11.00	0.95
perfect	vernon	1.00	1.00	0.00	1.00
persistence	vernon	0.80	0.69	-18.10	0.93

Figure 4 shows metrics by month for the entire evaluation period which reveals some trends in forecast quality in certain seasons. Namely forecasts tend to perform worse in the winter, which is the wet season and worse in the summer which is the dry season. Forecasts tended to perform well during the snow melt season (April-July). Although forecast B tends to perform better overall, there are cases in which the A forecast is better, highlighting the difficulty of selecting a single deterministic forecast.

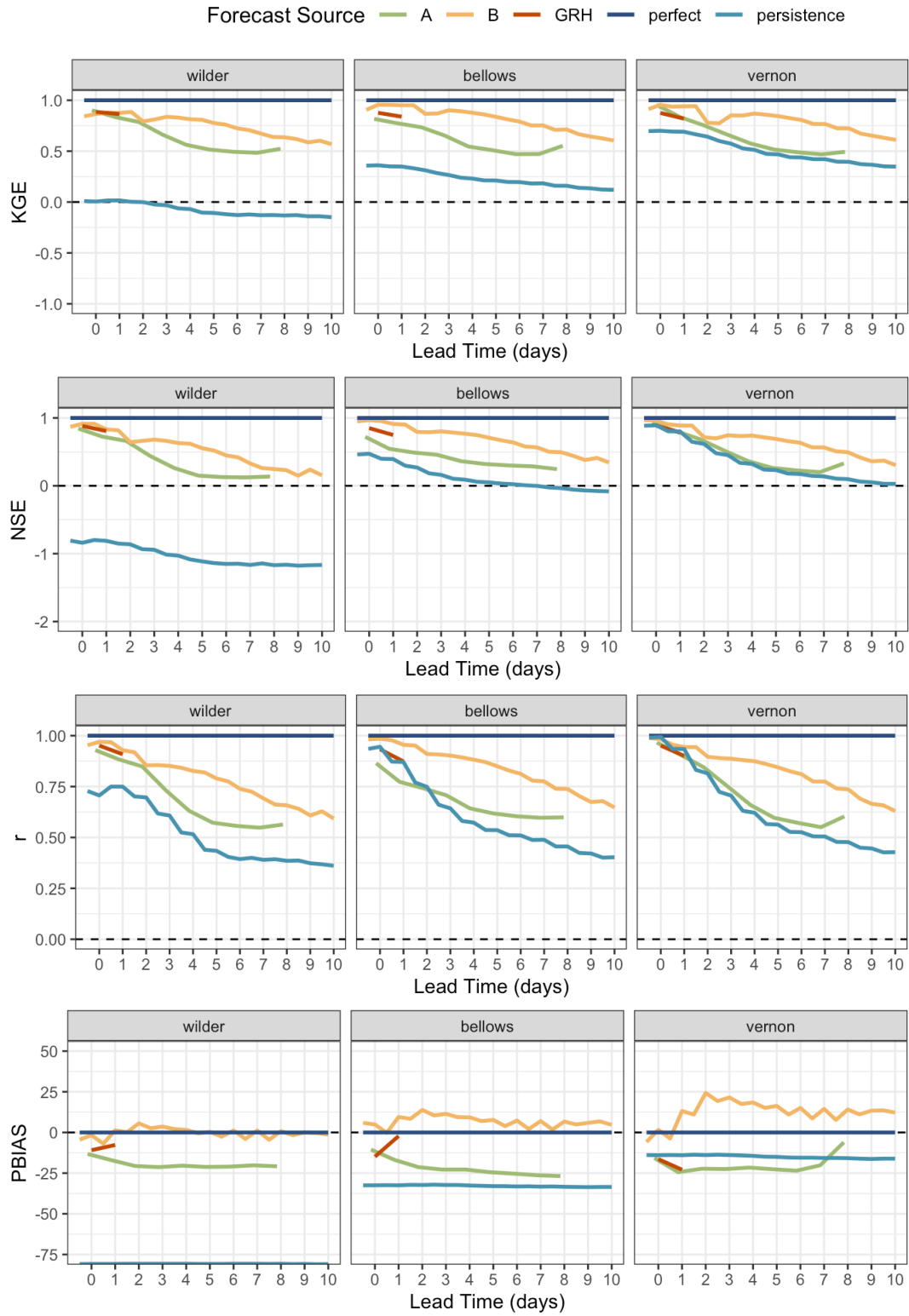


Figure 4 Summary of daily forecast performance with 4 generic metrics, see section 2.2.3 for descriptions of the metrics.

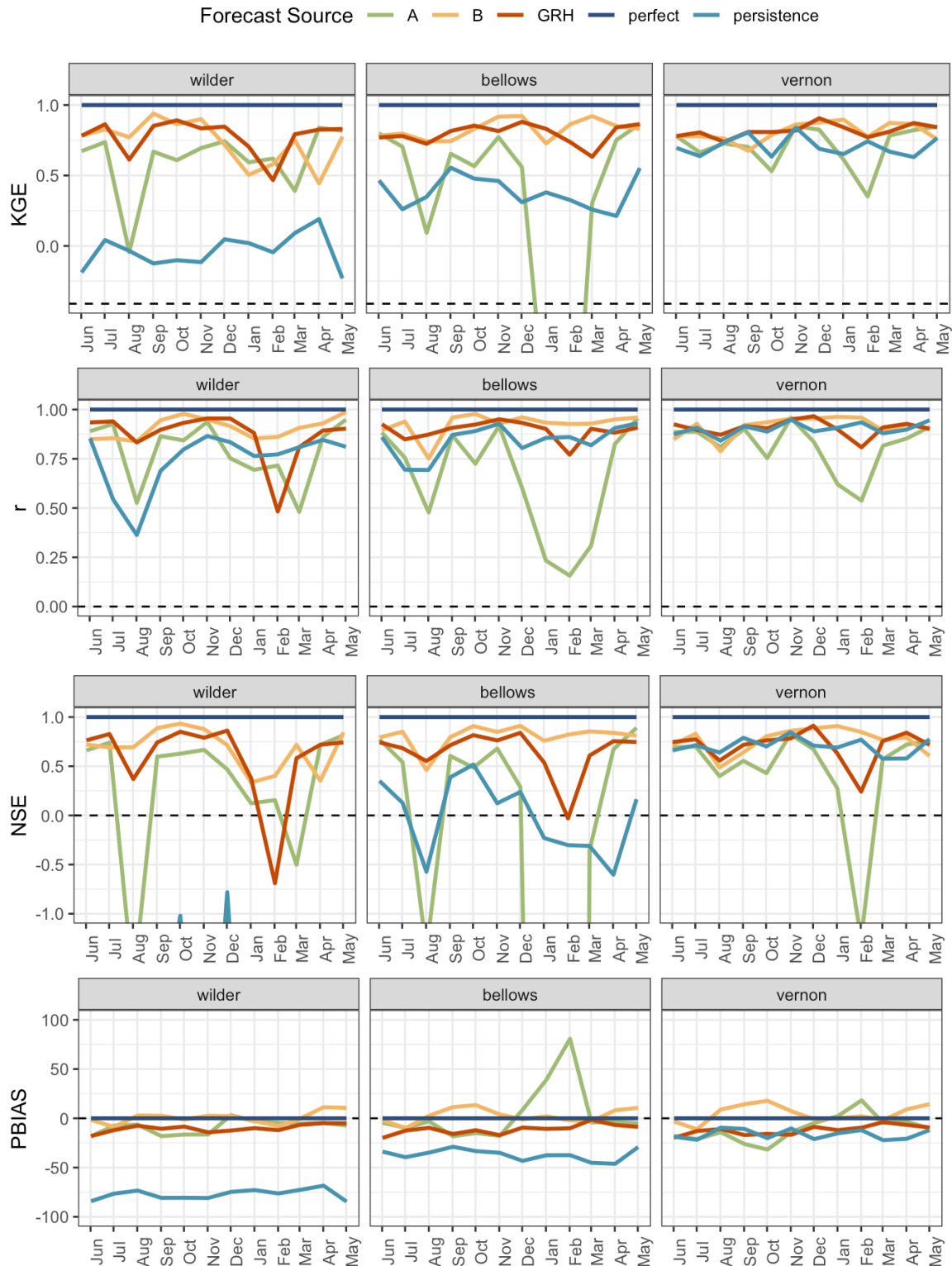


Figure 5: Summary of forecast performance by month, see section 2.2.3 for descriptions of the metrics.

### 2.3.1 Results – tailored metrics

In consultation with GRH, PNNL developed several tailored metrics that were relevant to GRH's operations. The first focused on day ahead forecast performance as shown in the previous section. This was motivated by the day ahead market that GRH bids into, and as such they focus their attention heavily on the 1-2 day ahead period where the forecasts matter the most. Operationally, they adjust their flows after this point, so no forecast is used after the 2 day mark.

During the FERC relicensing for the three main dams on the lower Connecticut river a strict limit on forebay fluctuations was imposed. GRH was given a +/- 0.5 foot operating range after which penalties are incurred. Given GRH's reliance on the inflow forecast, a metric was developed that captures the chance of the forebay limit being exceeded if the inflow forecast is used directly. Figure 6 shows where the forebay would be after a certain amount of time if the forecast was used directly set a pass inflow condition (i.e., inflow=outflow). The error bars extend to the 2.5<sup>th</sup> and 97.5<sup>th</sup> percentile of all forecasts. We can see that rarely would a forecast cause the forebay to exceed the 0.5 foot limit during the first day of operations, after which the operators can adjust.

Figure 7 shows the corresponding probability of exceeding the 0.5 foot limit which is less than 10% for all forecasts for all lead times in the 1-3 day range. GRH was also interested in seeing a more realistic situation involving a non-optimal forebay starting position on the second day, what was the probability of exceeding the 0.5 foot range, which is shown in Figure 8. We can see that the probabilities are much higher, up to 25% in some cases for the A forecast. This information provides a detailed view of the forecast performance that GRH can use to assess forecast quality.

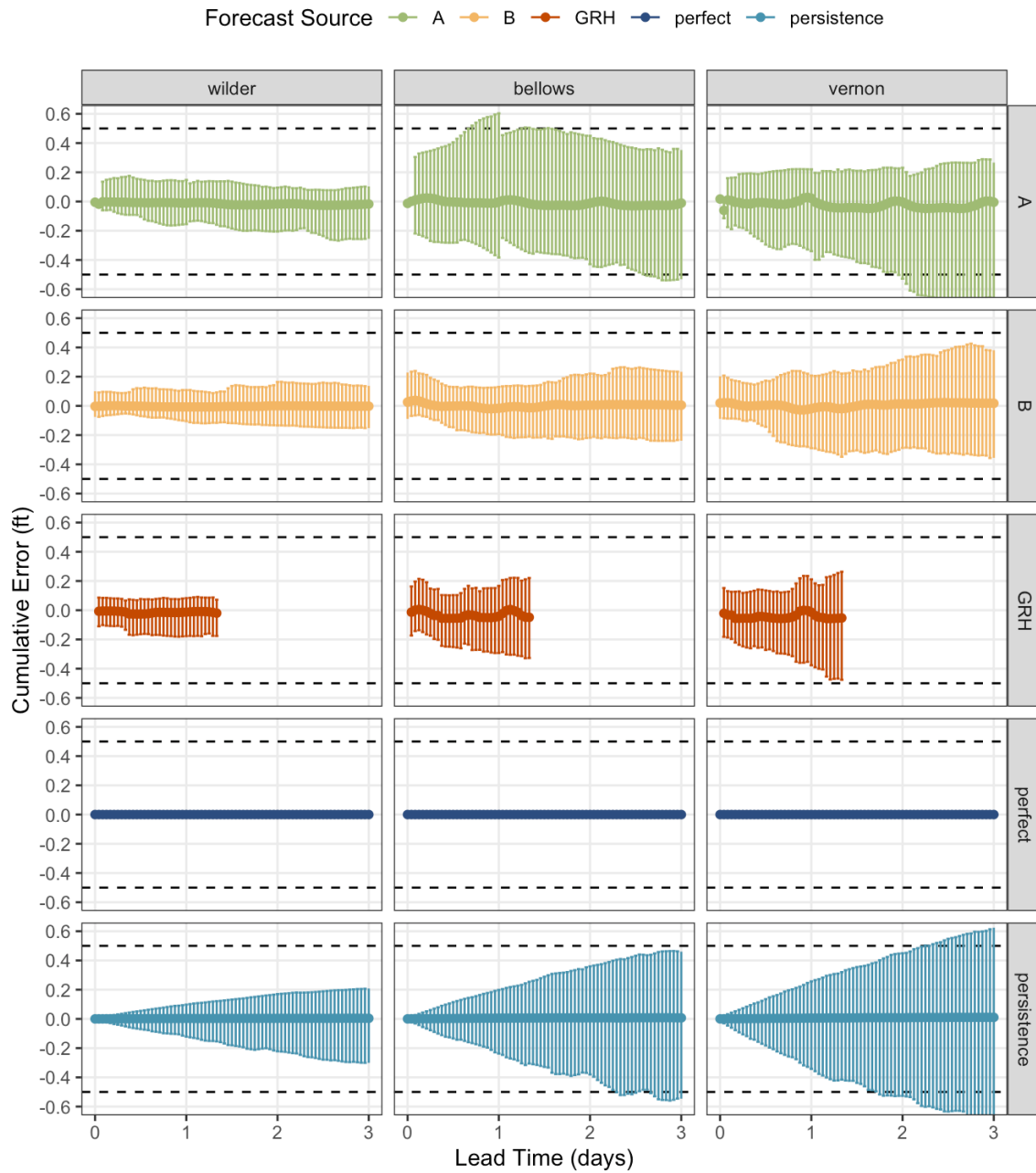


Figure 6 Cumulative chance of exceeding a 0.5 foot operating range if the operations were set to pass inflow (inflow=outflow) where inflow was determined by the various inflow forecasts.



Figure 7 Probability of exceeding a 0.5 foot operating range given a neutral forebay starting position on day 0.

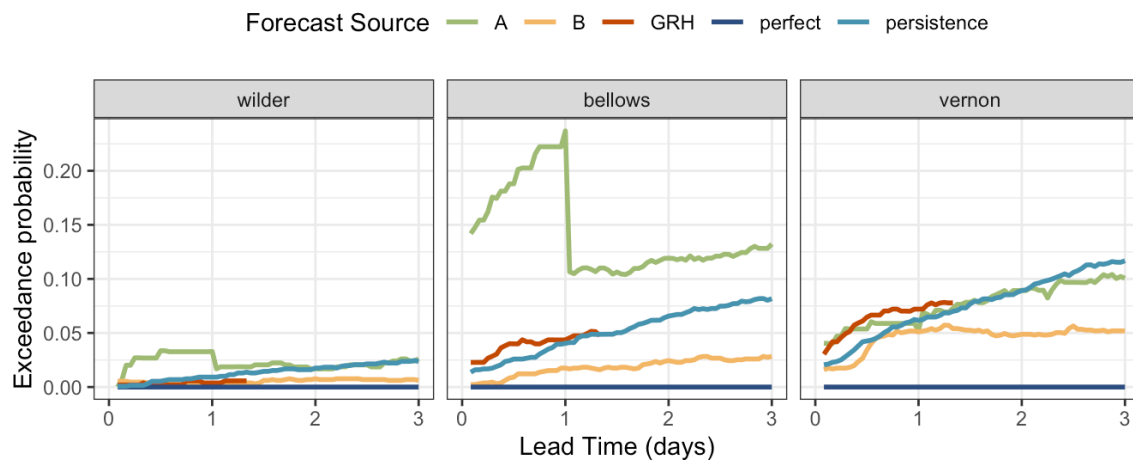


Figure 8 Probability of exceeding a 0.5 foot operating range on the second day, given the forebay starting position from the first day.

## 2.4 GRH forecast evaluation discussion

In this task we conducted a real time reservoir inflow forecast evaluation for the three main dams on the lower GRH system and three tributaries. We conducted general forecast verification and validation but also included other tailored metrics of forecast quality that were specific to GRH. The evaluation framework we developed involved an iterative approach which includes direct operator feedback in the process. This evaluation was able to provide GRH hydropower operators with information necessary for evaluating how much a particular forecast product will improve their operations and ultimately inform ongoing investment decisions (i.e. purchase external forecast products or improve existing internal forecasts). We consider this to be a successful application of a forecast evaluation framework that can be applied to other hydropower producers across the country.

In our real time evaluation of the GRH inflow forecasts there was one commercial forecast that stood out as having better performance overall. However, it was not always the best in every month and every lead time. This underscores the difficulty of picking a single best deterministic forecast. Moving to a probabilistic forecast may alleviate this somewhat but operators still need to bid fixed quantities into energy markets so overall improvements in deterministic forecast skill are valuable. Ultimately GRH must weigh the cost of a commercial forecast against improving their internal forecast. Importantly, this task did not address how much improvements in inflow forecasts will translate into improvements in hydropower operations and increased revenue. The next step in the evaluation framework (Section 3.0) is to couple this evaluation with hydropower scheduling models to quantify increases in revenue from improved inflow forecasts.

Real time forecast evaluations such as the one conducted for GRH are inherently limited by the amount of time they are conducted. Ideally a forecast model's performance would be assessed over full range of hydrologic conditions spanning multiple decades. While this was not practical to conduct for GRH in this project, we recognize the limitation. The decision of future forecast evaluations to pursue a real-time or hindcast evaluation should be made by the system operators depending on their specific system needs and resource availability.

The iterative nature of this evaluation was perhaps its most valuable component. We were able to develop tailored metrics that were relevant to GRH's operations. Having direct operator involvement in the process was critical to its success as PNNL was able to hone and develop the evaluation throughout the process. We hope this evaluation can serve as a template for further applications and neutral evaluation of the myriads of inflow forecasts available to hydropower system operators.

## 3.0 Task 2: Evaluation of the GRH System Operability under Flow Forecast Uncertainty and New Operations

### 3.1 Introduction

This task evaluated two hydropower scheduling tools, one currently implemented within GRH facilities and the other newly developed by Hatch specifically for GRH, to examine their performance and associated uncertainties in adapting to operational changes with limited flexibility under the new licensing rules.

### 3.2 Methods

#### 3.2.1 Scheduling Models

Two hydropower scheduling models were evaluated for the operations of three GRH facilities located along the lower Connecticut River:

- **GRH's Existing Megawatt Scheduler:** GRH currently manages operations across its river system by utilizing an in-house developed, spreadsheet-based scheduling tool, known as Megawatt Scheduler (MS). The MS tool provides daily hydropower operation schedules derived from empirical operation data, such as seasonal rule curves and historical records, along with a combination of inflow forecasts (e.g., data from NOAA Northeast River Forecast Center and USGS) and electricity market predictions (e.g., Enverus2 forecasts), under operational constraints, including unit availability and efficiency. Furthermore, throughout the day, GRH operations staff monitor and update actual hourly inflow and power generation to ensure compliance with FERC licensing requirements.
- **Hatch's New Higher-Resolution Scheduler:** Hatch has developed a new scheduling model designed to account for the impacts of two distinct operational scenarios: (a) Inflow-Equals-Outflow (IEO) operation, and (b) IEO combined with limited discretionary flexible generation (IEO/Flex). Although the Hatch's new scheduling model has not yet been implemented by GRH for operational use, it has the capability of incorporating the limited flexibility for operations, as well as utilizing inflow forecasts (Task 1) at multiple GRH facility locations.

The basis of the evaluation is system data collected by GRH. Specifically, system records, including calculated inflow, pool elevation, power generation, accessible through the PI-Tag data management system for the GRH facilities of interest: Wilder, Bellows, and Vernon.

#### 3.2.2 Evaluation Metrics

Evaluation of hydropower scheduling models were conducted from three primary perspectives: system compliance, power generation, and revenue optimization. During the initial phase of experiments, multiple evaluation metrics were proposed to identify gaps between hydropower schedules and actual operations, with the aim of further developing and refining the metrics as additional data were obtained through a series of experiments.

## System Compliance

Maintaining pool elevations within a specified range is essential for reservoir operations under the new licensing rules. To assess system compliance, hourly deviations in pool elevations from target water storage levels (Target Elevations include-Wilder: 384.5 ft, Bellows: 291.1 ft, and Vernon: 219.63 ft) were calculated, and corresponding daily statistics were derived according to the following equation:

$$\text{Pool Elevation Deviation [ft]} = \text{PI-Tag Record Elevation [ft]} - \text{Target Elevation [ft]} \quad [5]$$

Consequently, positive deviations indicate water storage exceeding the target levels, while negative deviations represent water storage below the target levels. The daily statistical measures include parameters such as the average, minimum, and maximum deviations of reservoir water levels from the target levels, as well as the total number of hours during which the elevation falls outside the favorable range (i.e.,  $\pm 0.5$  ft of the target level for all three reservoirs).

The discrepancy between actual and scheduled inflow or generation can result in deviations from target water storage levels. To examine this relationship, the pool elevation deviations and the actual differences from scheduler values were normalized using their own reference values (e.g., 0.5 ft for pool elevations and scheduler inflow or generation), and daily ratios. A negative sign was applied to the normalized ratio for generation, as additional generation reduces water storage in the reservoirs. Consequently, the normalized ratios can indicate the robustness of compliance with pool elevations in relation to extra inflow or generation over their scheduled values.

$$\begin{aligned} & \text{Normalized Pool Elevation Deviation per Extra Inflow [\%/ \%]} \\ & = \frac{\text{Pool Elevation Deviation/Predefined Threshold (0.5 ft) [\%]}}{[\text{PI-Tag Actual Inflow} - \text{Scheduler Inflow}]/\text{Scheduler Inflow} [\%]} \end{aligned} \quad [6]$$

$$\begin{aligned} & \text{Normalized Pool Elevation Deviation per Extra Generation [ft/MW]} \\ & = - \frac{\text{Pool Elevation Deviation/Predefined Threshold (0.5 ft) [\%]}}{[\text{PI-Tag Record Generation} - \text{Scheduler Generation}]/\text{Scheduler Generation} [\%]} \end{aligned} \quad [7]$$

## Power Generation

Power generation is a fundamental outcome of hydropower facilities, in which potential and kinetic energy from water are converted into electrical energy, contributing directly to revenues. This process is closely linked to facility operations and is influenced by the hydrologic conditions of river systems as well as the mechanical conditions of generating units. The ratio of generation to inflow was calculated, focusing primarily on water utilization. In the calculation, effective inflow (Wilder: < 11,000 cfs, Bellows: < 11,000 cfs, and Vernon: < 14,000 cfs), representing the share of inflow that can be harnessed for generation, was considered (i.e., excludes water lost through spills).

$$\text{Generation per Effective Inflow [MWh/cfs]} = \frac{\text{Daily Total Generation [MWh]}}{\text{Daily Total Effective Inflow [cfs]}} \quad [8]$$

where

$$\text{Daily Total Generation [MWh]} = \sum_{hour=1}^{24} \sum_{unit=1}^n \text{Pi-Tag Record Generation for Each Unit} \quad [9]$$

$$\text{Daily Total Effective Inflow [cfs]} = \sum_{hour=1}^{24} \min[\text{Inflow, Harnessable Maximum Inflow}] \quad [10]$$

According to the basic equation for hydropower ( $P$ ), the efficiency ( $\eta$ ) term can be calculated as follows:

$$P = \rho g H Q \eta \rightarrow \eta = \frac{P}{\rho g H Q} \quad [11]$$

where  $\rho$ ,  $g$ , and  $H$  denote water density, gravitational acceleration, and hydraulic head, respectively. Although the hydropower equation is typically applied at the turbine level, it was adapted to evaluate overall generation efficiency at the facility level. The facility generation efficiency was calculated hourly and then averaged daily.

$$\text{Facility Generation Efficiency } (\eta) \text{ [\%]} = \frac{\text{Generation [MWh]}}{\text{Pi-Tag Record Head [ft]} \times \text{Effective Inflow [cfs]}} \times 11,810 \left[ \frac{\text{ft} \cdot \text{cfs}}{\text{MWh}} \right] \quad [12]$$

### Revenue Optimization

Revenue was calculated by multiplying power generation by the electricity market price. Due to the day-ahead bidding process, daily total revenue consists of two portions: the expected revenue, based on the generation scheduled with ISO New England subject to day-ahead market prices, and the extra revenue, which accounts for surpluses or deficits in actual generation and real-time market prices.

$$\text{Daily Total Revenue [\$]} = \text{Daily Expected Revenue} + \text{Daily Extra Revenue (Surplus or Deficit)} \quad [13]$$

where

$$\text{Daily Expected Revenue [\$]} = \sum_{hour=1}^{24} [\text{Submitted Generation [MW]} \times \text{Day-Ahead Price [\$ / MW]}] \quad [14]$$

$$\begin{aligned} \text{Daily Extra Revenue [\$]} &= \sum_{hour=1}^{24} [(\text{Pi-Tag Generation} - \text{Submitted Generation})[\text{MW}]] \\ &\quad \times \text{Real-Time Price [\$ / MW]} \end{aligned} \quad [15]$$

Similar to the ratio of generation to inflow, the ratio of revenue to inflow was calculated to evaluate the revenue potential under varying inflow conditions. Also, effective inflow was used for the ratio calculations.

$$\text{Revenue per Effective Inflow [\$ / cfs]} = \frac{\text{Daily Total Revenue [\$]}}{\text{Daily Total Effective Inflow [cfs]}} \quad [16]$$

## Other Metrics

The metrics for system compliance, power generation, and revenue optimization were calculated individually for each facility, even though the facilities are interconnected through the Connecticut River system. To identify the propagation of uncertainty across facilities, discrepancies in pool elevation, inflow, and power generation between the actual and scheduler data were normalized using scheduler values (0.5 ft predefined threshold for pool elevation) and then plotted together.

$$\frac{\Delta h}{h} = \frac{[\text{PI-Tag Record Elevation} - \text{Target Elevation}]}{\text{Predefined Threshold (0.5 ft)}} \quad [17]$$

$$\frac{\Delta Q}{Q} = \frac{[\text{PI-Tag Actual Inflow} - \text{Scheduler Inflow}]}{\text{Scheduler Inflow}} \quad [18]$$

$$\frac{\Delta P}{P} = \frac{[\text{PI-Tag Record Generation} - \text{Scheduler Generation}]}{\text{Scheduler Generation}} \quad [19]$$

### 3.2.3 Experimental Scheme

Experiments alternating between the GRH MS and Hatch schedulers were designed to collect sufficient datasets that incorporate diverse inflow, generation, and market conditions, which would aid in the development and refinement of robust evaluation metrics. The initial setup of maintaining pool elevations as close as possible to the target levels was considered to constrain the experimental conditions. Additionally, any operational changes, such as flex hours and spills, were to be noted in the actual records to distinguish such cases.

## 3.3 Preliminary Results

### 3.3.1 Benchmark

The initial experimental dataset provided by GRH contained hourly records spanning 8 days distributed across two weeks, covering February 26-28, March 4-7, and March 11, 2025. Due to initial challenges in experimental operations, the collection of the experimental dataset did not strictly align with the designed setup, such as maintaining pool elevations at near-constant target levels. Moreover, high-flow conditions exceeding the effective operational range were periodically experienced (March 8-10) and then persisted into late June, restricting further experimental trials. As a result, evaluations were preliminarily focused on the initial dataset that was only available with limited duration.

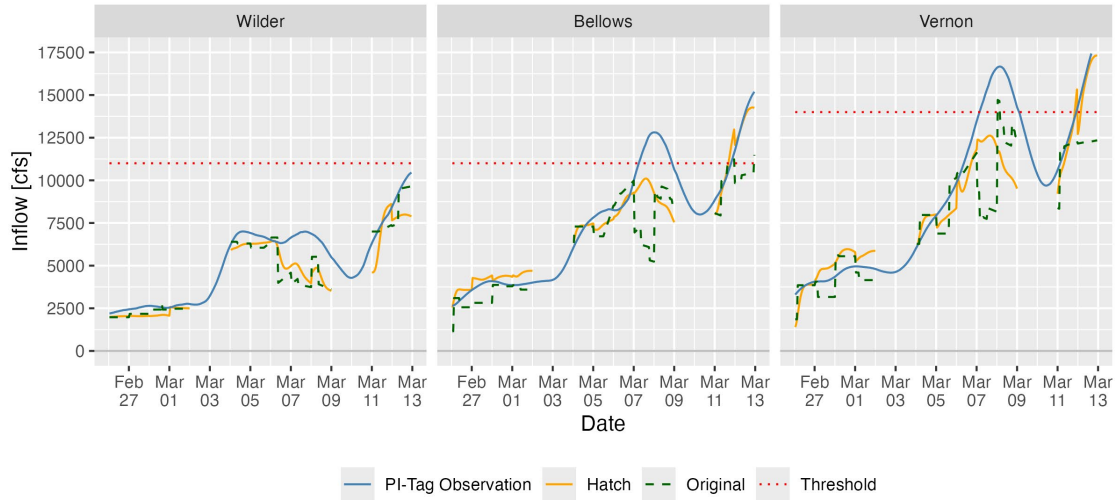
Analyses relied in part on inflow data from the selected forecast tool (see section 2 above) and projected generation data from the Hatch scheduler. Additionally, the dataset contained actual generation records, which are identical to those stored in the PI-Tag system. Estimated power generation from the Hatch scheduler was submitted to ISO New England for bidding in the day-ahead electricity market after GRH operators reviewed and processed the data (e.g., rounding adjustments). Consequently, the dataset includes four categories:

- **Hatch** scheduler inflow: inflow data that serves as the basis for calculating hydropower generation for scheduling, derived along with forecast inflow (see Section 2)

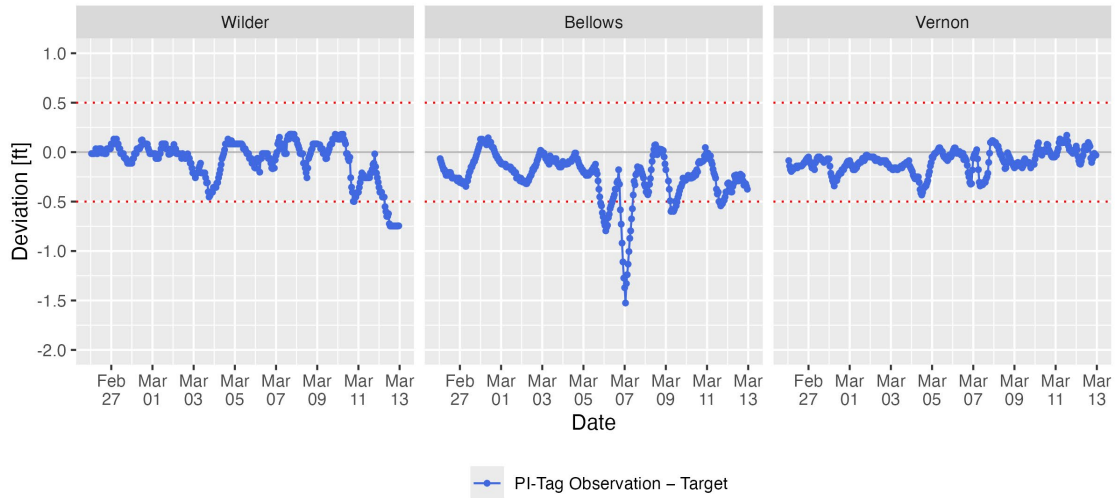
- **Hatch** scheduler generation: generation data calculated by the Hatch scheduler
- **Submitted** generation: generation data submitted to ISO New England for bidding in the day-ahead electricity market, based on the Hatch scheduler during the initial experimental periods, which can be slightly different from Hatch scheduler generation due to rounding or operational discretion
- **Original** generation: generation data that GRH MS would have scheduled for operation if no experiments were conducted
- **Actual** generation: actual generation data (identical to the PI-Tag system records)

Figure 9 shows fundamental data at three GRH facilities, including hourly inflow, pool elevation deviations, and power generation, compared with the PI-Tag system records. During the initial experimental periods, the inflows fluctuated significantly over a series of days, later exceeding effective ranges at the Bellows and Vernon facilities. While both GRH forecast (marked original in the figures) and Hatch scheduler inflows aligned closely with each other as well as actual values recorded by the PI-Tag system for most days, exceptions were observed between March 7 and 8 when inflows approached the threshold values for effective generation capacity. For the same period, discrepancies were also noticeable in terms of power generation schedules. Moreover, substantial drops in pool elevation were observed at the Bellows facility. Although specific operational notes were not discussed, it may be beneficial to further examine the differences in scheduler outputs (e.g., reduction in power generation on March 7 and 8, according to the MS scheduler), especially when approaching high-flow conditions. Notably, such inflow conditions were considered outside the effective operational range for evaluating the scheduling models. Nevertheless, the metrics were calculated for the later periods as a reference, since the valid data was limited only for a few days in the earlier period.

(a) Hourly Inflow



(b) Hourly Pool Elevation Deviations



(c) Hourly Generation (MAX Capacity: Wilder 41 MW, Bellows 49MW, Vernon 37MW)

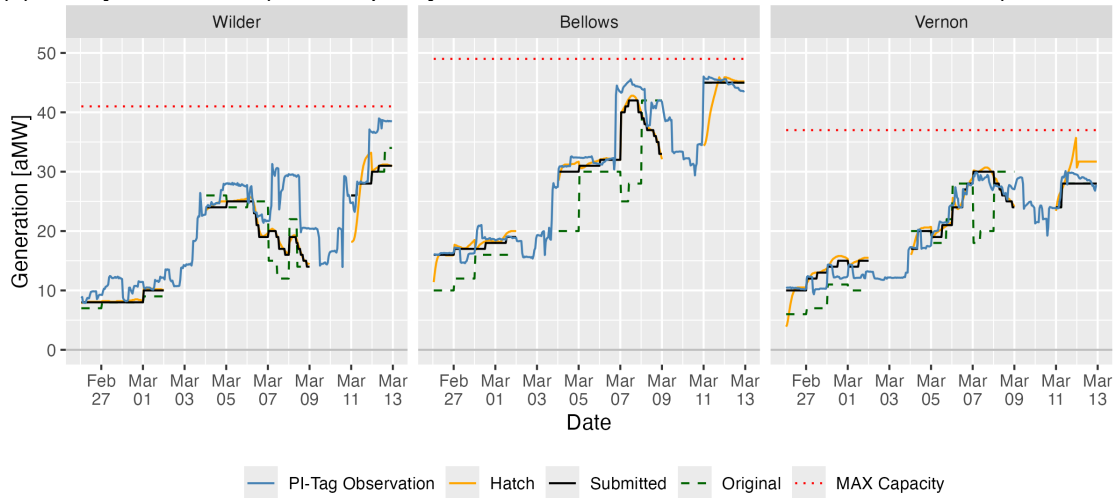


Figure 9 Hourly data at Wilder, Bellows, and Vernon facilities during the initial experimental periods: (a) inflow, (b) pool elevation deviations, (c) power generation.

### 3.3.2 Preliminary Results and Recent Data Shared with GRH

As shown in Figure 10, the hourly inflows presented in the previous subsection were aggregated daily to align with the scheduler cycles. Additionally, the effective inflows were considered for the daily aggregation, as inflows exceeding the thresholds (Wilder: 11,000 cfs; Bellows: 11,000 cfs; and Vernon: 14,000 cfs) do not generate additional power beyond the maximum capacity. Therefore, most of the evaluation metrics related to power generation and revenue optimization were based on these effective inflows for their calculations. The gray-shaded portions in the figure below represent the dates recorded by the scheduler experiment, while the PI-Tag system provided the records for the remaining dates.

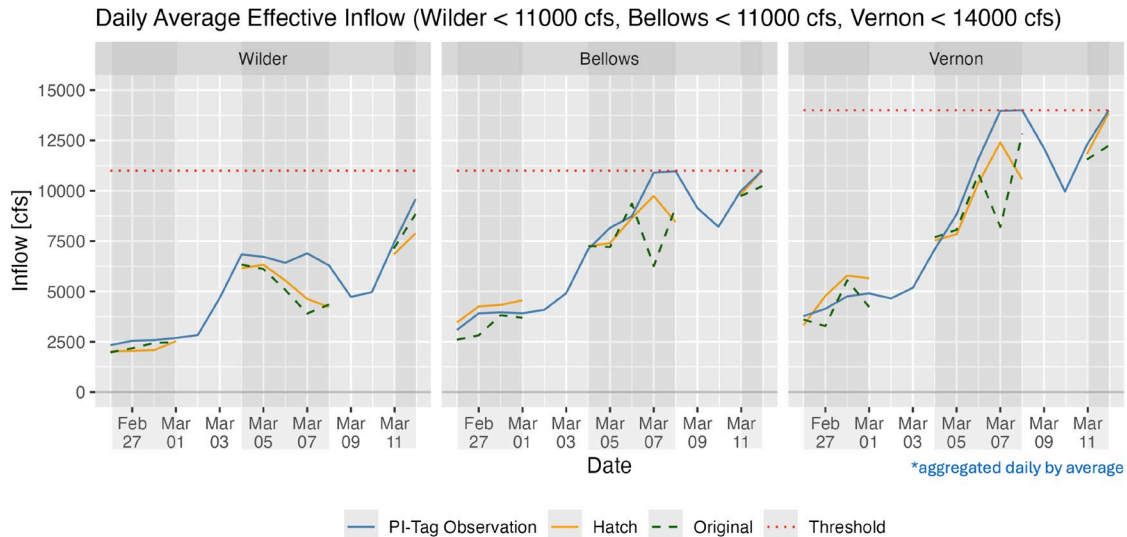


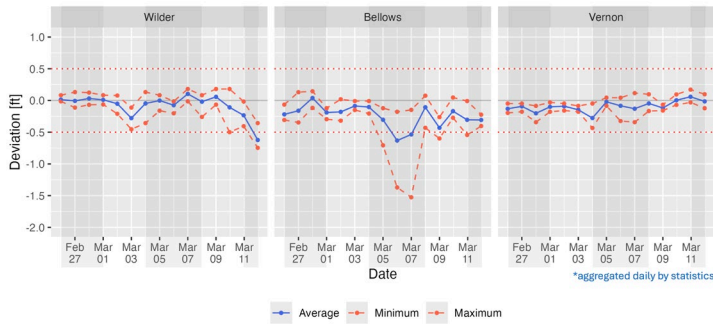
Figure 10 Daily average of effective inflow, generating hydropower at Wilder, Bellows, and Vernon facilities. The darker gray shades represent the periods for the initial scheduler experiment.

### System Compliance

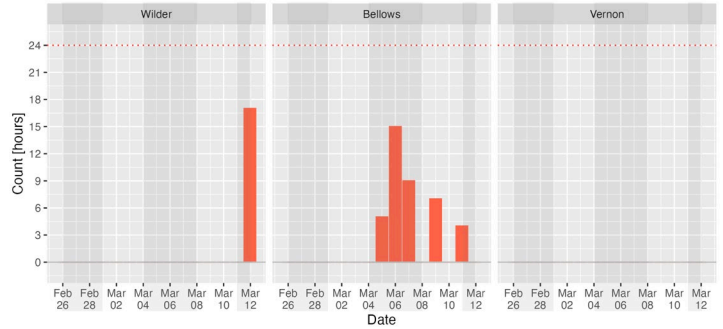
In Figure 11, the daily statistics indicate that the average pool elevation deviations remained consistently close to zero, with maximum and minimum variations ranging from approximately 0.1 to 0.3 ft, except on March 12 for the Wilder facility and March 5 to 11 for the Bellows facility. A comparison across the facilities shows that pool elevations at the Vernon facility only remained within the target ranges, while both Bellows and Vernon facilities experienced inflow conditions surpassing the normal ranges for power generation. Conversely, although inflow conditions were still normal at the Wilder facility, its pool elevations dropped below the target range for several hours on March 12.

The normalized ratios of pool elevation deviations to the differences between actual and scheduled values are plotted for inflow and generation, respectively. The negative ratio values imply that the pool elevations remained below target despite higher actual inflow or lower actual generation than their scheduled values, which occurred around March 06. However, additional experimental data within normal inflow ranges are required to draw robust conclusions or to develop more sophisticated metrics for assessing system compliance.

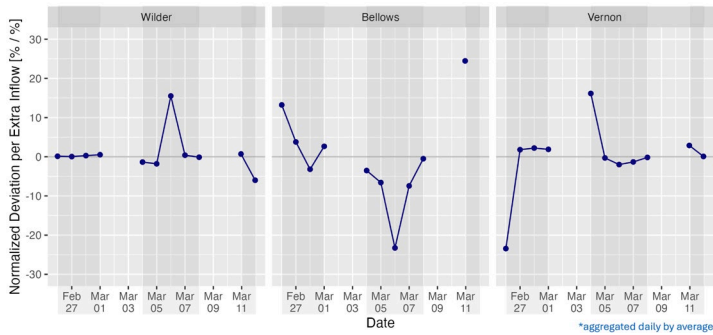
(a) Pool Elevation Deviation: Daily Statistics



(b) Pool Elevation Deviation: # of Exceedance Hours



(c) Daily Ratio: Pool Elevation Deviation per Extra Inflow



(d) Daily Ratio: Pool Elevation Deviation per Extra Generation

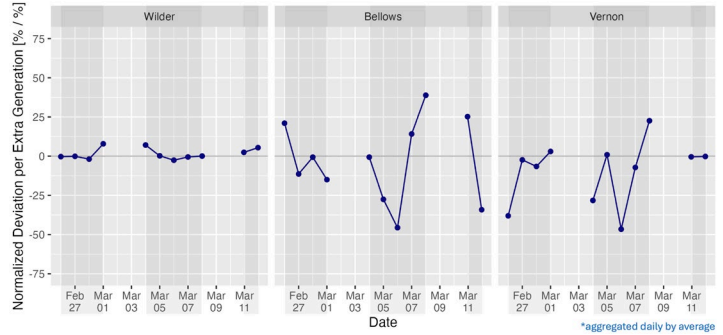


Figure 11 Evaluation of system compliance with relevant records and potential metrics: (a) daily statistics of pool elevation deviation, (b) number of exceedance hours, (c) daily ratio of pool elevation deviation to extra inflow, and (d) daily ratio of pool elevation deviation to extra generation.

## Power Generation

As shown in Figure 12, actual power generation largely followed the Hatch’s scheduler during the initial experimental period, except on March 7 to 8, when relatively higher generation was observed at the Wilder facility. Overall, power generation is closely correlated with effective inflow. However, due to mechanical constraints, the maximum generation capacity was approximately 90% of the nameplate capacity, even when inflow conditions exceeded the threshold. The daily ratios of generation to effective inflow (currently calculated using the Hatch scheduler inflow) indicate water utilization rates ranging from 1 to 6 kWh per unit cfs, with higher rates observed at the Wilder facility (upstream) between 4 and 6 kWh/cfs, compared to 2 to 3 kWh/cfs at the Vernon facility (downstream). These variations are partially attributed to differences in hydraulic head across the facilities (i.e., Bellows > Wilder > Vernon). Facility generation efficiencies (calculated using PI-Tag inflow data) show similar levels across all facilities, with slightly higher efficiency observed at Wilder compared to the other facilities. Those metrics can be similarly calculated for alternating experiments using the MS scheduler to enable comparison with the Hatch scheduler.

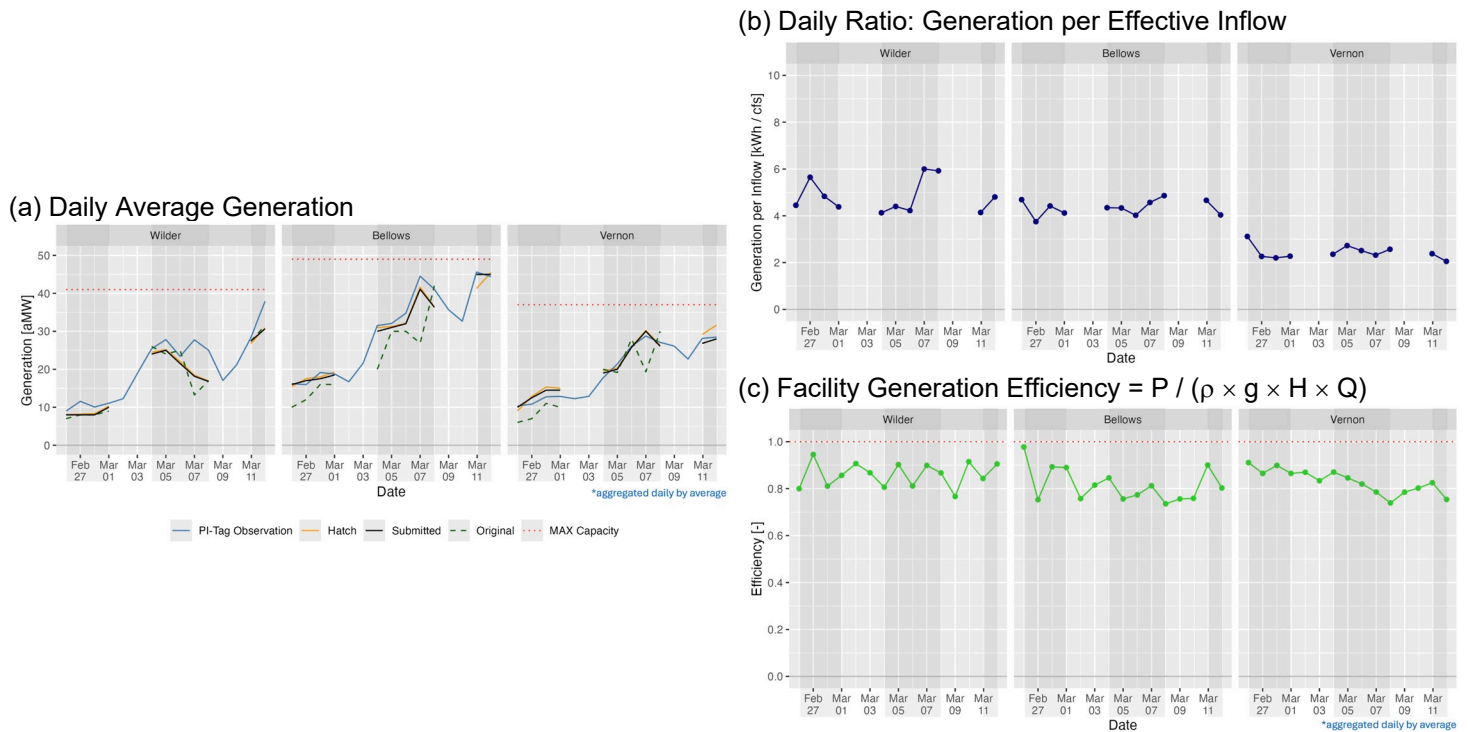
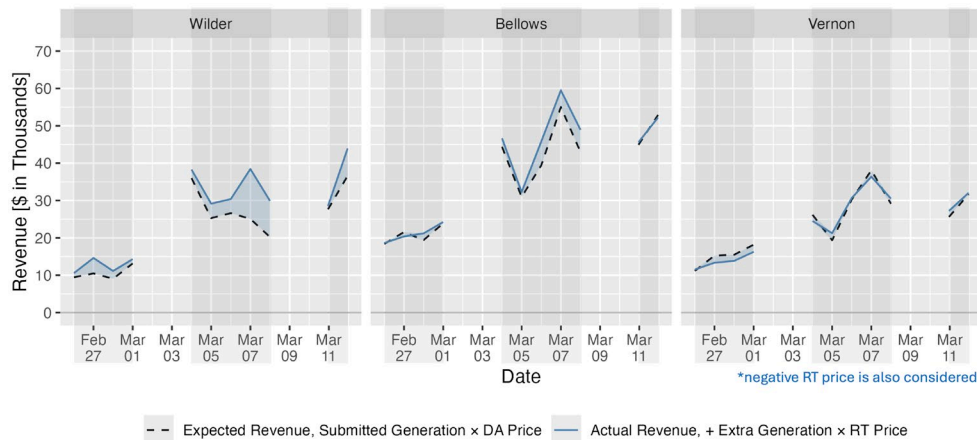


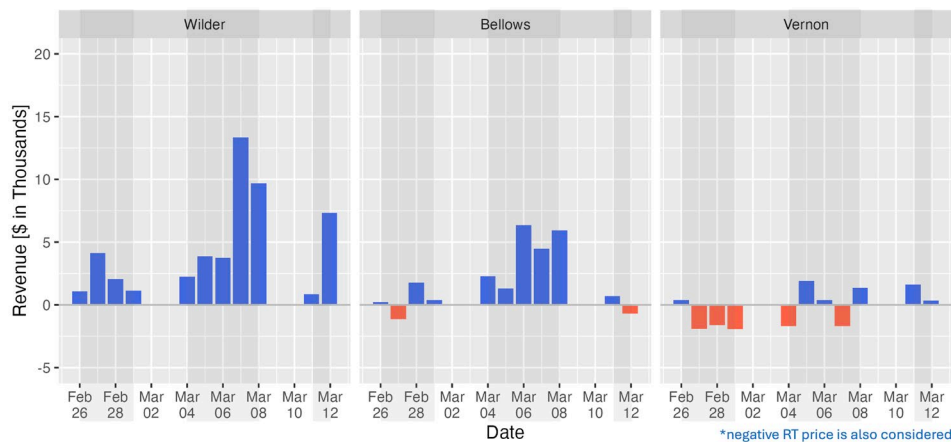
Figure 12 Evaluation of power generation with relevant records and potential metrics: (a) daily average generation, (b) daily ratio of generation to effective inflow, and (c) daily average facility generation efficiency.

## Revenue Optimization

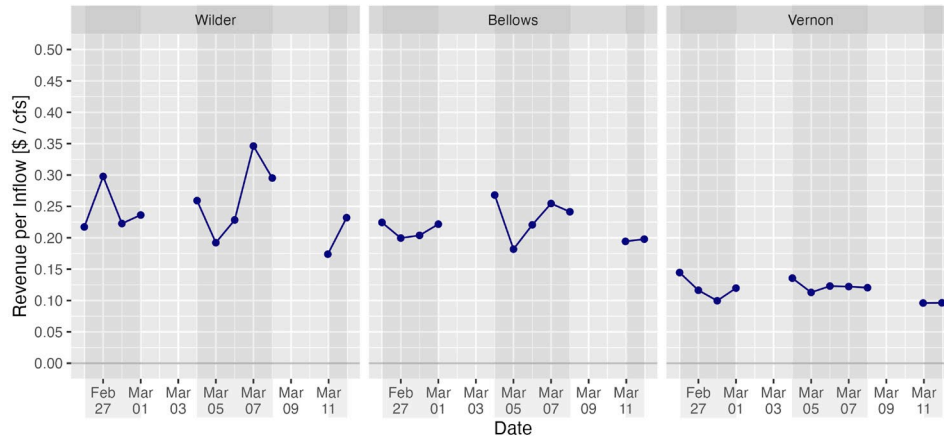
In the first panel of Figure 13 below, the blue shading represents extra revenue earned or lost beyond the expected revenue based on day-ahead bidding. At the Wilder and Bellows facilities, significant additional revenue was observed from March 5 to 8, as more power was generated than the submitted generation to ISO New England. The actual generation at the Vernon facility closely matched the submitted generation, resulting in a combination of surpluses and deficits in revenue. Their total revenue was often slightly lower than the expected revenue, as shown by the red histograms in the plot for daily extra revenue. These revenue decreases were due to unmet generation, which was multiplied by real-time electricity market prices. The daily ratios of revenue to effective inflow (calculated using forecasted inflow-see Section 2) showed revenue rates ranging from 0.1 to 0.35 \$ per unit cfs. Higher rates were observed at the Wilder facility (upstream), ranging between 0.2 and 0.35 kWh/cfs, while lower rates of 0.1 to 0.15 kWh/cfs were seen at the Vernon facility (downstream). These variations are consistent with a similar ratio for generation, which reflects water utilization.



**(a) Daily Total Revenue**



**(b) Daily Extra Revenue**



**(c) Daily Ratio: Revenue per Effective Inflow**

Figure 13 Evaluation of revenue optimization with relevant records and potential metrics: (a) daily total revenue, (b) daily extra revenue from generation surplus or deficit, and (c) daily ratio of revenue to effective flow.

## Other Metrics

In Figure 14, the discrepancies in pool elevation, inflow, and power generation between the actual and scheduler data were plotted with respect to facilities, along with extra generation, for cross-comparison. According to the plots for inflow and generation, the biases between the actual and scheduler values were similar for the Bellows and Vernon facilities, whereas the Wilder facility had more positive biases (i.e., scheduler underestimation). With extended experimental datasets, correlations across facilities can potentially be established to quantify bias propagation and guide improvements to scheduler models.

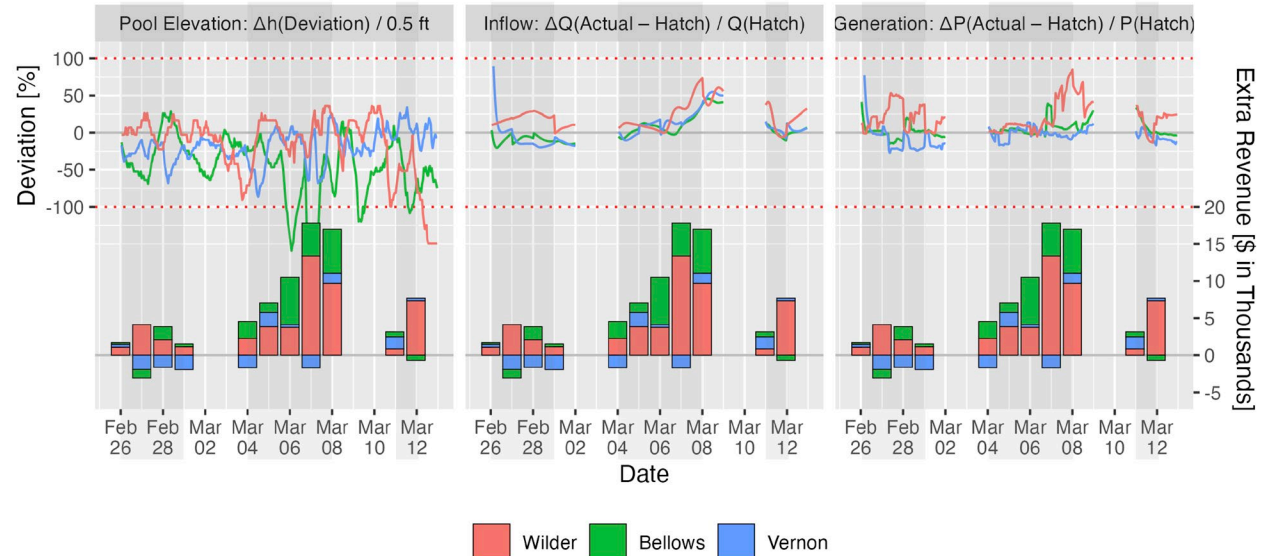


Figure 14 Deviation of pool elevation, inflow, and generation, along with daily extra revenue. The hourly records are normalized by scheduler values (0.5 ft predefined threshold for pool elevation).

## 3.4 Summary

Here we have demonstrated an approach for evaluating tools for operating and scheduling hydropower resources within the GRH system. Evaluations were drawn on the basis of three measures: compliance, generation and revenue. A wide range of metrics were developed to explore the efficacy of the new scheduling tool. Unfortunately, only a couple of weeks of data were able to be collected due to environmental and resource limitations. As such no definitive conclusions can be drawn concerning the value of the tool at this time.

## 4.0 Task 3: Identification of opportunities to increase market revenue

### 4.1 Introduction

Hydropower operators increasingly face the challenge of balancing compliance with regulatory requirements against the need to maximize economic value in competitive electricity markets. For GRH, this challenge has intensified following the relicensing of its lower Connecticut River projects, which impose stricter run-of-river operating rules but allow a limited number of discretionary Flex hours. These Flex hours create opportunities to deviate from strict inflow-equals-outflow operations, enabling operators to shift water to higher-value market periods. However, realizing this value requires careful planning and the ability to quantify trade-offs between compliance, operational feasibility, and revenue optimization. This task sought to quantify the incremental value that enhanced scheduling approaches can provide relative to GRH's current practices. Specifically, we compared a run-of-river baseline model—representing business-as-usual operations—with an enhanced reservoir-augmented scheduling model, which explicitly incorporates storage dynamics, reservoir constraints, and intertemporal optimization. The contrast between these two approaches provides insight into how additional operational flexibility can be harnessed to capture value in ISO-NE's day-ahead and real-time markets. The importance of this analysis extends beyond GRH. Many U.S. hydropower operators are entering relicensing processes that restrict traditional flexibility, while at the same time market structures increasingly reward resources that can shift generation across hours or days. Quantifying the benefits of reservoir-based scheduling under realistic flow, price, and regulatory conditions provides a replicable framework for hydropower operators to evaluate investment in forecasting and optimization tools.

### 4.2 Model development

Hydropower revenues in ISO-NE are ultimately realized through multi-horizon bidding: (i) emerging 7–10-day-ahead forward commitments (FWD), (ii) day-ahead (DA) hourly schedules, and (iii) real-time (RT) balancing around DA positions. For Great River Hydro (GRH), the economic question is not only how much flexibility exists (via storage, ramping, and Flex hours), but how to allocate that flexibility across market horizons under uncertain inflows and prices.

We build the run-of-river (RoR) and cascade reservoir formulations into a market-integrated, three-stage optimization that (1) plans tentative block bids 7–10 days ahead, (2) converts them into hourly DA quantity bids with Flex-hour choices and refill commitments, and (3) executes RT recourse to minimize imbalance costs when inflows or prices deviate. This structure preserves the water balance, storage limits, ramping, spill bounds, and generation constraints in reservoir models and adds market-specific bid variables, deviation settlements to produce price-aligned yet compliance-robust schedules. Together, the three horizons convert physical flexibility into bankable bidding strategies while respecting GRH's new licensing constraints and limited Flex hours.

### 4.3 Model Description

This section details the objective and constraints involved in scheduling cascading hydropower plants along a river.

### 4.3.1 NOMENCLATURE

---

Indices	
t	Index of time
i	Index of hydropower plant along a river
Variables	
$p_{i,t}^H$	Power generation of plant i at time t
$q_{i,t}$	Power generation flow of plant i at time t
$s_{i,t}$	Spill flow of plant i at time t
$r_{i,t}$	Storage of reservoir I at time t
Parameters	
$w_{i,t}$	Mountain stream natural incoming water into hydropower plant I at time t
$q_{i,t}^{FWD}$	7-10 day-ahead forecast inflow
$q_{i,t}^{DA}$	Day-ahead forecast inflow
$q_{i,t}^{RT}$	Real-time forecast inflow
$q_{i,t}^{obs}$	Observed inflow
$c_{i,t}^{FWD}$	7-10 day-ahead forecast price
$c_{i,t}^{DA}$	day-ahead forecast price
$c_{i,t}^{RT}$	Real-time forecast price
$c_{i,t}^{RT,obs}$	Observed real-time settlement price
$B_{i,k}^{FWD}$	emerging 7–10-day-ahead forward commitments (FWD) for plant I at time block k
$B_{i,t}^{DA}$	DA hourly bid for plant I at time t
$D_{i,t}^{RT}$	Real-time deviation $D_{i,t}^{RT} = p_{i,t}^H - B_{i,t}^{DA}$
$U_i$	Set of upstream hydropower plants for plant i
$p_f$	Minimal power generation of all the hydropower plants along the river
$\Delta p_{i,max}$	Maximal hourly up-ramping/down-ramping rates
$r_{i,min/max}$	Maximal/minimal storage of reservoir i
$q_{i,min/max}$	Maximal/minimal power generation flow for plant i
$s_{i,min/max}$	Maximal/minimal spill flow for plant i
$p_{i,min/max}$	Maximal/minimal power generation for plant i
$T_i^{exp/imp,max}$	Maximal export/import power for plant i

---

### 4.3.2 Problem Formulation

**Objective function:** The objective of this project is to maximize the total revenues generated by all the run-off-river hydropower plants along the river. Run-off-river hydropower plants have very small reservoir. This objective can be mathematically represented as:

$$\max \sum_{k,t \in k} \sum_i c_{i,t}^{FWD} B_{i,k}^{FWD} + \sum_t \sum_i c_{i,t}^{DA} B_{i,t}^{DA} + \sum_t \sum_i c_{i,t}^{RT} D_{i,t}^{RT} \quad [20]$$

### 4.3.3 Constraints

**Water balance constraint:** As shown in Figure 15, the storage of reservoir  $i$  at time  $t+1$  is updated based on the following: the storage of reservoir  $i$  at time  $t$ , mountain stream natural incoming water, power generation flow, and spill flow at time  $t$ , power generation flow and spill flow from upper hydropower plants at time  $t - \tau$ :

$$r_{i,t+1} = r_{i,t} + w_{i,t} - q_{i,t} - s_{i,t} + \sum_{j \in U_i} (q_{j,t-\tau} + s_{j,t-\tau}) \quad [21]$$

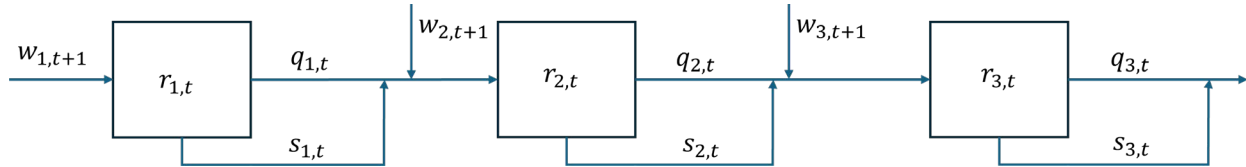


Figure 155 Storage update for reservoirs in cascade hydro power plants, considering inflows and outflows

**Storage bounds:** The storage of reservoir  $i$  must remain within the defined lower and upper bounds:

$$r_{i,min} \leq r_{i,t} \leq r_{i,max}$$

**Power generation flows for each plant:** Minimal and maximum power generation flows are defined by:

$$q_{i,min} \leq q_{i,t} \leq q_{i,max}$$

**Spill flows for each plant:** Minimal and maximum spill flows are defined by:

$$s_{i,min} \leq s_{i,t} \leq s_{i,max}$$

**Electricity generation for each plant:** Minimal and maximum electricity generation of plant  $i$  is defined by:

$$p_{i,min} \leq p_{i,t} \leq p_{i,max}$$

**Total electricity generation:** Total electricity generation of cascade hydropower plant is within a defined range:

$$p_f \leq \sum_{i=1} p_{i,t} \leq \sum_{i=1} p_{i,max}$$

**Maximum ramping capacity constraint:** The maximum hourly up-ramping and down-ramping rates are defined by:

$$|p_{i,t+1} - p_{i,t}| \leq \Delta p_{i,max}$$

The role of each market stage is described below:

**FWD (7–10 days):** Uses 7–10-day price and inflow forecasts to pre-position storage trajectories so water is available for forecasted high-value blocks, while staying within reservoir, ramping, spill, and eco-flow limits across the cascade.

DA (next day): Converts block intent into hourly bids with Flex-hour placement and explicit refill-by-deadline constraints, honoring turbine/spill limits, ramping, and upstream-to-downstream travel-time lags.

RT (intra-day): Executes feasible dispatch with real-time inflow and price forecasts, trimming deviations from DA via reservoir headroom (and optional battery), under the same hydropower physics used in planning

We use forecasts as first-class drivers. The formulation explicitly consumes 7–10-day inflow/price forecasts in FWD, day-ahead forecasts in DA, and real-time nowcasts in RT. These inputs determine storage targets, hourly bid levels, and flexible use. Potential performance indicators may use: (1) peak-capture ratio: share of scheduled energy in highest forecasted price hours. (2) real-time deviation magnitude: absolute deviation from DA, by hour/plant.

#### **4.4 Summary**

Here we have outlined a framework that provides hydropower operators with a structured approach to assessing the economic benefits of investing in advanced forecasting and scheduling tools. An unexpected issue arose that caused the project to end early thus limiting the ability to fully test these opportunities under extended real-time operations; nevertheless, the analysis demonstrated a clear framework for connecting forecast and scheduling accuracy to revenue gains.

# Pacific Northwest National Laboratory

902 Battelle Boulevard  
P.O. Box 999  
Richland, WA 99354  
1-888-375-PNNL (7665)

***[www.pnnl.gov](http://www.pnnl.gov)***