DISCLAIMER

Argonne NATIONAL LABORATORY

ANL-25/27

CRiSPPy: An advanced hydropower scheduling tool for the Colorado River Storage Project

Energy Systems and Infrastructure Assessment

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Lemont, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

DOCUMENT AVAILABILITY

Online Access: U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free at OSTI.GOV (http://www.osti.gov/), a service of the US Dept. of Energy's Office of Scientific and Technical Information.

Reports not in digital format may be purchased by the public from the National Technical Information Service (NTIS):

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312
www.ntis.gov
Phone: (800) 553-NTIS (6847) or (703) 605-6000
Fax: (703) 605-6900
Email: **orders@ntis.gov**

Reports not in digital format are available to DOE and DOE contractors from the Office of Scientific and Technical Information (OSTI):

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
www.osti.gov
Phone: (865) 576-8401
Fax: (865) 576-5728
Email: **reports@osti.gov**

ANL-25/27

CRiSPPy: An advanced hydropower scheduling tool for the Colorado River Storage Project

prepared by
Quentin Ploussard[1], Matija Pavičević[1], Ao Yu[1], Jerry Wilhite[2], Walter Henriquez[2]
[1]Energy Systems and Infrastructure Assessment, Argonne National Laboratory
[2]Western Area Power Administration

June 2025

Foreword

The Western Area Power Administration (WAPA) plays a vital role in delivering reliable and cost-effective hydroelectric power to millions of customers across the western United States. The Colorado River Storage Project (CRSP) carries out WAPA's mission in Arizona, Utah, Colorado, New Mexico, Nevada, Wyoming and Texas. Achieving this mission requires effective management of the Colorado River system, and depends on the use of advanced analytical tools and modeling methodologies.

For many years, CRSP has relied on the Generation and Transmission Maximization Superlite (GTMax SL) model for its mid-term and long-term hydroscheduling needs. However, the evolving energy market, power system operations, environmental rules, and hydrology conditions, coupled with advancements in computational capabilities, have necessitated the development of a more modern and robust solution.

This report introduces the Colorado River Storage Project Python-based (CRiSPPy) model, a new, advanced hydropower scheduling tool developed to address CRSP ever-evolving challenges. CRiSPPy represents a significant leap forward in our ability to model and optimize the operation of the Colorado River system. It incorporates state-of-the-art optimization algorithms, enhanced data management capabilities, and an advanced graphical user interface, providing WAPA CRSP personnel with unprecedented insights and decision-making support.

This document details the development, capabilities, and implementation of CRiSPPy. It is intended to serve as a comprehensive resource for WAPA staff, stakeholders, and anyone interested in the future of hydropower scheduling in the Colorado River Basin. We are confident that CRiSPPy will enhance WAPA's mission while adapting to the challenges of a dynamic and increasingly complex environment.

The version of CRiSPPy described in this report is the version 2.3. New versions of CRiSPPy will be developed as the tool keeps evolving to address CRSP challenges.

Contents

List of Figures

List of Tables

List of Acronyms and Abbreviations

1D one-dimensional. 13, 19, 20, 41

24MS 24-month study. 6, 37, 42, 63, 65

2D two-dimensional. 15, 52, 53

AF acre-feet. 8

ANL Argonne National Laboratory. 1, 2, 4, 17

BM Blue Mesa. 6, 10, 11, 23, 24, 26

CAISO California Independent System Operator. 27

CBRFC Colorado Basin River Forecast Center. 6

CCA Cooperative Conservation Alternative. 65

cfs cubic feet per second. 8, 9, 39, 63

CPWL continuous piecewise linear. 15, 17, 19, 20, 41, 52–55

CRiSPPy Colorado River Storage Project Python-based model. 2–5, 13–15, 17–20, 24, 25, 27, 29, 32, 34, 37, 39–41, 43–50, 56–59, 61, 63–66

CRMMS Colorado River Mid-term Modeling System. 6, 37, 65

CRSP Colorado River Storage Project. 1, 2, 4, 6, 8, 12, 14, 15, 20, 23–26, 37, 42, 43

CRSS Colorado River Simulation System. 6, 37

CSV comma-separated values. 43, 62

CY Crystal. 6, 11, 13, 17, 19, 23, 24

DC difference-of-convex. 15, 52–55

DSA Deliverable Sales Amount. 6, 8, 43

EIS Environmental Impact Statement. 1, 9, 10, 65

EMMO Energy Management and Marketing Office. 1, 6, 37, 39, 50, 63, 64

FES Firm Electric Service. 8, 27, 37, 41

FG Flaming Gorge. 6, 9–11, 17–19, 23, 25, 32, 39, 55

ft feet. 11, 12, 14, 20

GC Glen Canyon. 6, 8, 11, 19, 20, 23, 24, 26, 32, 34, 53, 56, 63–66

GLOP Google Linear Optimization Package. 5, 60

GTMax Generation and Transmission Maximization. 1

GTMax SL GTMax Superlite. 1, 2, 17, 46

GUI graphical user interface. 5, 37, 39, 44, 50–52, 54, 56

HFE high flow experiment. 8, 34–37, 40, 43, 59, 60, 63–66

hr hour. 8, 9

I/O Input/Output. 46

LP linear programming. 17, 20, 46–48, 59, 60, 66

LTEMP long-term experimental and management plan. 8, 34, 63

MB megabyte. 46

MILP mixed integer linear programming. 15, 19, 30

ML machine learning. 2–4, 13–15, 17, 18, 20, 25, 39, 46, 52–55

MP Morrow Point. 6, 11, 13, 23, 24, 26

MW megawatts. 8, 9, 11, 12, 20, 62

NEPA National Environmental Policy Act. 1

PCF water-to-power conversion factor. 14, 15, 17, 25, 37, 52–54

RAM Random Access Memory. 46, 49

ROD record of decision. 8, 34, 63

SCIP Solving Constraint Integer Programs. 5, 60

SLCA/IP Salt Lake City Area Integrated Projects. 1, 2, 6, 8, 43

SMB Smallmouth bass. 40, 63, 64

SSARR Streamflow Synthesis and Reservoir Regulation. 17

TAF thousands of acre-feet. 8, 63

TTD travel time distribution. 17, 18, 25

USBR U.S. Bureau of Reclamation. 1, 4, 6, 9, 13, 14, 17, 20, 37, 40, 52, 58, 63–65

USGS U.S. Geological Survey. 13, 18, 19, 53

WAPA Western Area Power Administration. 1, 2, 4, 6, 8, 9, 23, 39, 42, 65

WECC Western Electricity Coordinating Council. 27

1 Introduction

The U.S. Department of Energy's Western Area Power Administration (WAPA) markets electricity produced at hydroelectric facilities operated by the U.S. Bureau of Reclamation (USBR). WAPA plays a vital role in delivering reliable and cost-effective hydroelectric power to millions of customers across the western United States. The Colorado River Storage Project (CRSP) region, one of the five WAPA regions, markets power from the CRSP, its participating projects (Dolores and Seedskadee) and the Collbran and Rio Grande projects. These resources comprise eleven powerplants located in Arizona, Colorado, New Mexico, Utah and Wyoming and are marketed together as the Salt Lake City Area Integrated Projects (SLCA/IP).

USBR, responsible for managing the water resources and dam operations in the Colorado River Basin, determines the monthly volumes of water that must be released from the CRSP reservoirs. The USBR communicates these monthly volume targets to WAPA's Energy Management and Marketing Office (EMMO), and WAPA schedules hourly water release and power generation at each reservoir by accounting for hourly customer load, monthly release targets, flow rate limits and other operational environmental constraints. To achieves its mission, WAPA requires sophisticated tools and methodologies that helps them meet power system, reservoir, and plant operational requirements while addressing customer load requests.

In 2001, Argonne National Laboratory (ANL) developed the Generation and Transmission Maximization (GTMax) Model [1], a versatile power market dispatch model powered by the LINGO optimization solver [2]. GTMax was developed to help utilities and other stakeholders maximize the value of their power system assets while considering various constraints and opportunities in electricity markets. A few years later, WAPA tasked Argonne to use the GTMax model as the main simulation tool to dispatch the power from SLCA/IP hydropower plants and conduct several financial analysis of experimental releases [3, 4]. To accurately represent the SLCA/IP system's unique characteristics, the GTMax model underwent several iterative updates and was supported by multiple ancillary tools and databases.

While sophisticated, the GTMax model offered numerous features that were unnecessary for the SLCA/IP system and lacked specific functionalities that were needed to address the system unique requirements. As a result, in 2018, the GTMax Superlite (GTMax SL) model was developed to serve as a new, lighter, simulation tool tailored for WAPA CRSP [5,6]. The GTMax SL model is an Excel-based simulation model composed of a collection of supporting spreadsheets. As for GTMax, the GTMax SL optimization model is powered by the LINGO solver. For performance purposes, the model represents each month by a representative week (168 hours) with day-type weighing. The GTMax SL model underwent several updates to address the evolving needs of WAPA CRSP and to incorporate additional constraints. Specifically, the latest GTMax SL version, the "GTMax SL Experiments" model, accounts for all hours in a given month, allows the user to fix hourly releases in specific periods of time, and offers the option to exceed some operational constraints subject to user-defined penalty costs [7]. These modeling enhancements, however, come at the cost of significantly slower run times and increased debugging difficulty. Depending on the modeling parameters, a model run time with GTMax SL ranges from a few seconds to a few minutes.

In June 2023, USBR announced its intent to prepare an Environmental Impact Statement (EIS) as part of the required Post-2026 National Environmental Policy Act (NEPA) process [8]. The Post-2026 process is a multi-year process to determine long-term operations for Lake Powell and Lake Mead after the expiration of existing operating agreements in 2026. USBR

Figure 1.1: CRiSPPy logo[1]

proposed to analyze five alternatives. To accurately capture the impact of each alternative, the analysis scope covers 34 years of monthly operations (from 2027 to 2060) across 1,200 distinct hydrology conditions. WAPA tasked ANL with simulating the hourly operations of each of these alternatives. Each alternative represents 489,600 model runs, for a total of 2,448,000 model runs, resulting in over 5 billion hourly output values. It is estimated it would take the GTMax SL model thousands of hours of computing effort to solve and manage the output results of all these model runs. As a result, the Colorado River Storage Project Python-based model (CRiSPPy) was developed to address this challenge.

The CRiSPPy model is a Python-based hydropower scheduling model developed to address the ever-evolving challenges of WAPA CRSP. It uses advanced machine learning (ML) and optimization techniques to accurately represent the hydropower operations of the SLCA/IP system and significantly reduce solution times. Contrary to general power market dispatch models, it is specifically designed to model the unique features and ever-evolving operational rules of the SLCA/IP system. The tool allows WAPA staff to automatically learn and update the system physical rules, simulate hourly operations, visualize input and output data in a concise way, generate hourly and aggregated reports, and calculate financial and economic costs.

The official logo of the CRiSPPy software is shown in Figure 1.1. The logo features a French croissant in front of a dam, symbolically referencing both the software's name and the cultural background of its primary developer. The croissant, known for its crispy outer texture, is stylized in the shape of a "C" to allude to the software name.

This report provides a detailed description of the CRiSPPy software (version 2.3) and the hydropower system it is designed to model. Section 2 offers an overview of the software's capabilities and implementation. Section 3 introduces the CRSP system, including its operational guidelines and constraints. Section 4 presents the mathematical models that power the CRiSPPy

[1]Picture AI-generated with Google's ImageFX

software, including the core hydropower scheduling model and the ML models used to automatically update system physical rules. Section 5 explains how data are managed both within the software and through external sources. Section 6 highlights key improvements of CRiSPPy over the legacy platform. Section 7 describes the various user interfaces that make up the software. Finally, section 8 presents four use cases and their corresponding results.

2 Model overview

The CRiSPPy model is a highly-specialized hydropower scheduling tool designed to address the unique challenges faced by WAPA CRSP. As such, the tool is mainly intended for WAPA staff in their day-to-day operations and for ANL staff to support WAPA with periodic analyses. However, many of the mathematical models, algorithms, and computation techniques described in this report can, and are welcomed to be, used or adapted for other hydropower systems.

Below is a summary of the model capabilities and its technical implementation.

2.1 Model capabilities

The CRiSPPy model provides a comprehensive environment for hydropower scheduling and analysis. It facilitates the creation and management of case studies, allowing users to define the scope of their analysis by specifying the modeled plants, time period, and the number of price, load, and hydrology scenarios. The software streamlines data input by offering automated import from various offline and online sources, such as the USBR hydrologic database [9]. A key feature is its ability to automatically preprocess raw input data into parameters directly usable by the optimization solver, for instance, by calculating plant-specific water-to-power conversion factors based on hydrological conditions.

Users have extensive control over defining and customizing the operational rules of individual hydropower plants and the overall system, including parameters like release rate limits and spinning reserve requirements. To further refine these rules, the software assists users in computing, visualizing, and selecting physical plant characteristics (*e.g.*, reservoir storage-elevation relationships) using historical data and integrated ML algorithms. It also allows for the definition of specific rule changes that can be applied conditionally based on plant, time, load, price, and hydrology scenarios.

The core functionality of the software lies in its ability to compute optimal hourly water release schedules driven by customer loads and energy prices while taking into account monthly release targets and physical and operational rules. Users can also manually fix specific hourly power and non-power releases for certain plants. The model operates at an hourly level, simulating plant power and non-power release schedules, power generation, reservoir storage volume and forebay elevation, as well as the system's day-ahead and real-time energy transactions. For non-regular release events, the software can generate hourly release profiles based on user-defined event features.

To ensure robust and feasible solutions, the software includes an automatic infeasibility detection mechanism that can automatically apply adjustments to operational rules. Finally, the software offers comprehensive reporting capabilities, generating detailed hourly operation results and aggregated reports at various temporal scales (daily, weekly, monthly, yearly) using user-specified statistical measures. These results can be visualized through customizable graphs with different plot types and axis configurations. Furthermore, the software enables users to calculate the financial or economic impact of different operational rules or experimental releases by facilitating comparisons across multiple case studies.

2.2 Technical implementation

The CRiSPPy model was implemented in the Python programming language using several standard and third-party libraries, including: NumPy, Pandas, Matplotlib, OpenPyXL, xlwings, python-calamine, Shapely, SciPy, PyInstaller, and OR-Tools. The graphical user interface (GUI) and data visualization tools were developed with Tkinter and Matplotlib. Optimization models were formulated using Google OR-Tools MathOpt [10] and solved with the Google Linear Optimization Package (GLOP) solver and the Solving Constraint Integer Programs (SCIP) solver [11]. The model's Python script was compiled as a standalone executable file using PyInstaller. An illustration of the model implementation is depicted in Figure 2.1.

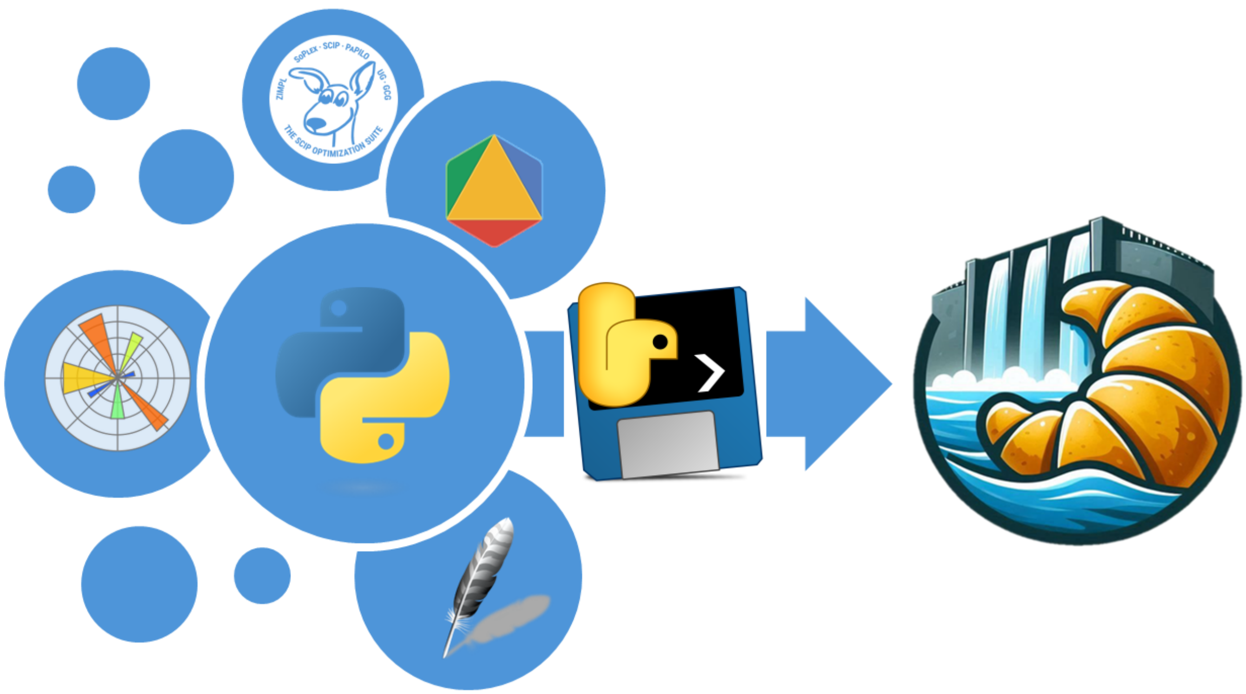

Figure 2.1: Implementation of the CRiSPPy software

3 The Colorado River Storage Project

The CRSP is a large-scale water management project authorized by the U.S. Congress in 1956. Its primary purpose is to comprehensively develop the water resources of the Upper Colorado River Basin (Colorado, New Mexico, Utah, and Wyoming). CRSP is composed of several storage units that include: the Glen Canyon (GC) dam, the Aspinall Unit composed of the three cascading reservoirs Blue Mesa (BM), Morrow Point (MP), and Crystal (CY), the Flaming Gorge (FG) dam, and other smaller reservoirs.

A geographical representation of the CRSP is depicted in Figure 3.1.

3.1 Monthly release targets

USBR sets the monthly water releases in the Upper and Lower Colorado River Basin to be consistent with various operating rules and guidelines, acts, international water treaties, consumption use requirements, State agreements, and the "Law of the River" [13]. In addition to power production, monthly release volumes are set considering other uses of the reservoirs, such as for flood control, river regulation, consumptive uses, water quality control, recreation, and fish and wildlife enhancement, and to address other environmental factors [14]. Release decisions are made based on simulation results from the Colorado River Simulation System (CRSS) or Colorado River Mid-term Modeling System (CRMMS) models [15]. Both models are implemented in the river modeling software RiverWare [16] and simulate release decisions using various forecasting methods, including runoff projections provided by the Colorado Basin River Forecast Center (CBRFC). Because future hydrologic conditions in the Colorado River Basin are not known with certainty and because events do not unfold as previously projected, USBR periodically adjusts its annual operating plan. Its release decisions are adjusted on a monthly basis to reflect projections made by a rolling 24-month study (24MS) that is updated monthly [17].

3.2 Scheduling guidelines

The hourly scheduling of SLCA/IP hydropower plant operations is performed by WAPA's EMMO located in Montrose, Colorado. Schedulers make decisions based on a set of scheduling priorities and guidelines, including a directive to comply with several reservoir-specific environmental operating criteria described in the following sections.

As operational constraints were imposed on SLCA/IP resources, scheduling guidelines and goals shifted from objectives driven primarily by market prices to objectives driven by customer loads. Within the boundaries of these operating constraints, SLCA/IP power resources are used to serve firm load. Because of operational limitations, WAPA staff may need either to purchase or sell varying amounts of energy on an hourly basis in the day-ahead and/or real-time market. In the day-ahead market, such transactions are scheduled in 16-hour, on-peak blocks, and 8-hour, off-peak blocks. In the past, the volume of these variable market purchases and sales used to be relatively small. However, evolving energy market and dry conditions led to more significant transactions. As a result, in 2022, WAPA implemented the Deliverable Sales Amount (DSA) [18]. The DSA is the amount of hydropower capacity and energy, determined quarterly, based on the forecasted SLCA/IP generation. Under DSA rules, each customer receives its proportionate share of the marketable hydropower capacity and energy.

Figure 3.1: The Colorado River Storage Project [12]

The load-following objective and DSA rules create a strong link between WAPA's contractual obligations and SLCA/IP hydropower plant operations, requiring dispatch among the hydropower plants to be closely coordinated. This interdependency exists because loads and hydropower resources are balanced whenever feasible. WAPA is able to affect the shape of its Firm Electric Service (FES) customer load requests indirectly through specifications in its contract amendments. In turn, these customer loads affect both SLCA/IP powerplant operations and hourly reservoir releases. The load-following dispatch directive minimizes scheduling problems and helps WAPA avoid noncompliant water releases.

An unwritten scheduling practice followed by WAPA involves maintaining a number of ramping cycles that is consistent with customer load patterns but not greater than two. This practice applies to all CRSP plants and is mainly followed to minimize turbine wear and tear.

3.3 Glen Canyon

The GC Powerplant, located at the base of the GC Dam on the Colorado River in Arizona, is a major hydroelectric facility with eight generators and a total capacity of 1,320 megawatts (MW). The 710-foot high dam creates the Lake Powell, the second largest man-made reservoirs in the U.S. with a capacity of more than 25 million acre-feet (AF).

The 2016 long-term experimental and management plan (LTEMP) record of decision (ROD) [19] describes the operating constraints at GC. These operating criteria are intended to temper the rate of change in hourly and daily water releases and were put into practice by WAPA from October 2017. Under this ROD, the release rate at GC is not allowed to exceed 25,000 cubic feet per second (cfs). Under exceptional circumstances, such as very wet hydrological conditions or high flow experiment (HFE), this flow rate limit may be exceeded. In addition, the 2016 ROD criteria require water release rates to be 8,000 cfs or greater between the hours of 7:00 a.m. and 7:00 p.m. and at least 5,000 cfs at night.

The criteria also limit how quickly the release rate can increase and decrease in consecutive hours. The maximum hourly increase (*i.e.*, up-ramp rate) is 4,000 cfs/hour (hr), and the maximum hourly decrease (*i.e.*, down-ramp rate) is 2,500 cfs/hr. The 2016 ROD operating criteria also restrict the range of release rates during rolling 24-hour periods. This change constraint depends on the monthly volume of water releases. Daily fluctuation (in cfs/24-hr) is limited to 10 times the monthly release specified in thousands of acre-feet (TAF) from June to August. As an example, if the target release volume in August is 700 TAF, the daily fluctuation limit in that month is 7,000 cfs/24-hr. From September to May, daily fluctuation is limited to 9 times the monthly release volume. The daily range can never exceed 8,000 cfs/24-hr. A summary of the 2016 ROD operating criteria, together with previous operating rules, is provided in Table 3.1.

In addition to the 2016 ROD operating rules, WAPA follows other scheduling practices. These practices fall within ROD operational boundaries but are not ROD requirements. One practice involves releasing the same volume of water during weekdays. It also requires weekend and holiday release volumes to be no larger than weekday release volumes but not smaller than 85 percent of weekday release volumes.

Table 3.1: Evolution of operating constraints at GC

Operational Constraint	Historical Flows (before 1991)	1996 ROD Flows (from 1997 to 2017)	2016 ROD Flows (after 2017)
Minimum flows (cfs)	*3,000 during the summer* *1,000 during the rest of the year*	*8,000 from 7:00 a.m. to 7:00 p.m.* *5,000 at night*	*8,000 from 7:00 a.m. to 7:00 p.m.* *5,000 at night*
Maximum non-experimental flows (cfs)[a]	*31,500*	*25,000*	*25,000*
Daily fluctuations (cfs/24-hr)	*28,500 during the summer* *30,500 during the rest of the year*	*5,000, 6,000, or 8,000 depending on release volume*	*Equal to 10 times the monthly water release (in TAF) during June–August,* *and equal to 9 times the monthly water release the rest of the year, but never exceeding 8,000*
Ramp rate (cfs/hr)	*Unrestricted*	*4,000 up* *1,500 down*	*4,000 up* *2,500 down*

[a] Except during very wet conditions

3.4 Flaming Gorge

The FG dam is a 502-foot high concrete arch dam on the Green River in Utah, creating the FG Reservoir for water storage, flood control, and recreation. Located at the dam's base, the FG Powerplant houses three turbines with a total capacity of 152 MW, generating hydroelectric power while also managing downstream flows for environmental needs.

The hourly water release from the FG Reservoir is required to be always maintained above 800 cfs. This rule resulted from an operating agreement between USBR and the Utah Division of Wildlife Resources described in the 2006 FG EIS [20] and designed to enhance the use of the river for fishing, fish spawning, and boating. Water ramp-rate restrictions at FG have been implemented per agreements between USBR and WAPA for system operations [20]. The up-ramp rate is currently limited to 800 cfs/hr and the down-ramp rate is limited to 1,000 cfs/hr.

In addition, current flow restrictions [20] impose FG to be operated in such a way that flow fluctuations at the Jensen gage are within a specified range. The Jensen gage is located on the Green River, about 29 miles downstream of the Yampa confluence and about 94 miles downstream of FG, as depicted in Figure 3.2. The intent of these river flow restrictions is to reduce the negative impacts of FG operations on endangered fishes. Under these flow restrictions, hourly releases at FG are patterned to produce no more than a 0.1-meter (about a 4-inch) stage

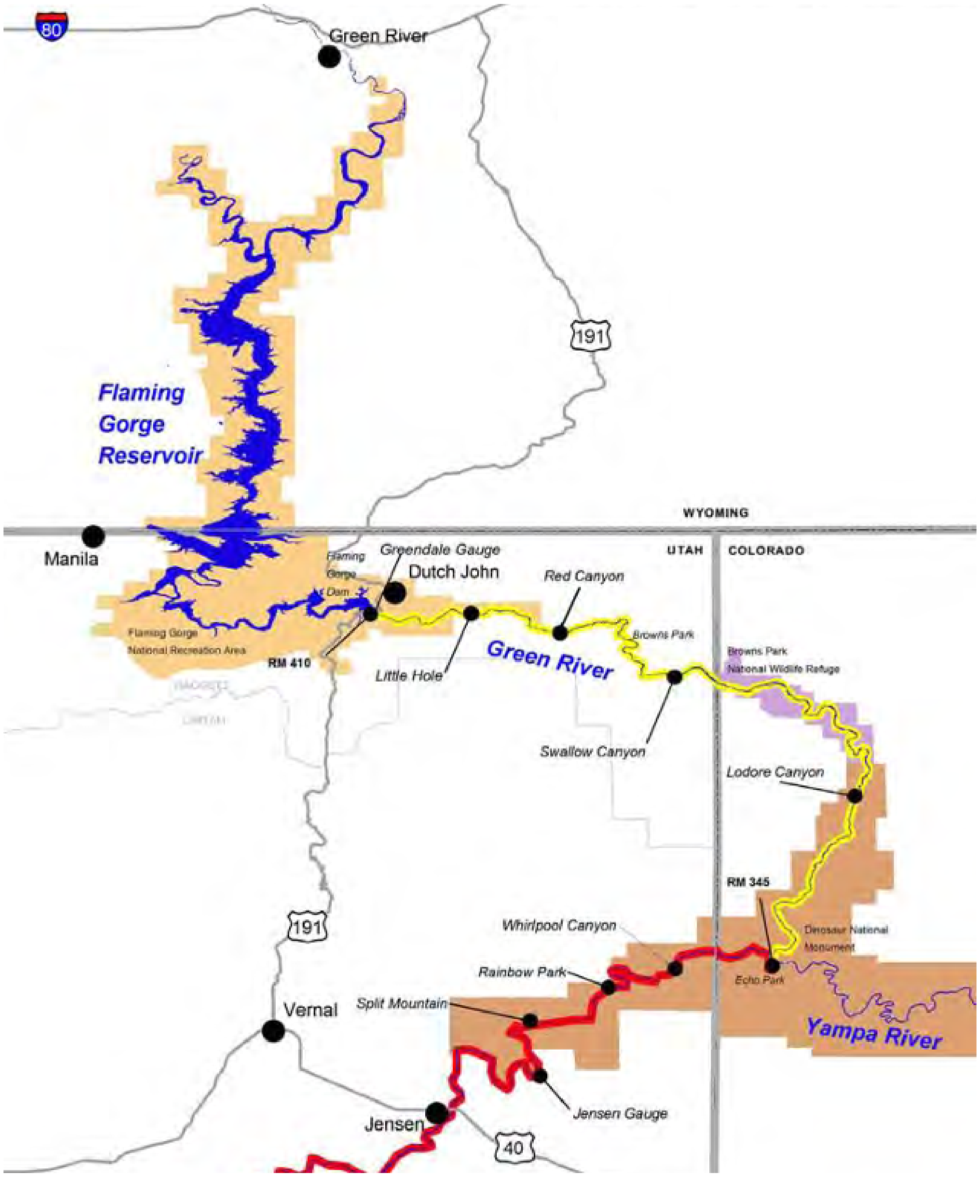

Figure 3.2: The Flaming Gorge reservoir, Yampa River, and Jensen gage [20]

change at Jensen within a 24-hour period except during emergency operations. Although Yampa River flows are accounted for when scheduling FG releases, the dam releases are not required to compensate for large and unpredictable changes in Yampa River flows per the EIS. In other words, from an operating standpoint, Yampa River flows are considered constant during a monthly period . As a result, the FG hourly release pattern is expected to be the same every day of the month.

3.5 Aspinall Unit

The Aspinall Unit, located on the Gunnison River in western Colorado, comprises three main dams and reservoirs, forming a cascading system. The BM reservoir is the uppermost and largest of the three, created by the BM Dam. It serves as the primary water storage for the unit and has two power generation units with a capacity of 86 MW. BM Reservoir is also a significant recreational resource and the largest lake in Colorado. Its initial operation focused on maximizing water storage and hydropower production. The MP reservoir is located downstream of BM, created by the MP Dam. This dam houses the most productive power plant within the Aspinall Unit, with two generation units totaling 173 MW. MP is considered the hydropower "workhorse" of the unit and can serve as a black start power source for the western grid. The final downstream reservoir is formed by the CY Dam. This is the last dam of the Aspinall Unit to have been completed and has the smallest power generation capacity within the unit, with one 32-MW unit. CY plays a crucial role in re-regulating the flow of the Gunnison River downstream, ensuring sufficient releases to meet the needs of water users while also aiming to avoid harming endangered fish habitat in the lower Gunnison River.

As opposed to GC and FG, there are relatively few rules that govern the water releases of the Aspinall Unit. There is no restriction regarding minimum and maximum hourly water releases for any of the three dams. BM and MP have no ramp-up and ramp-down water releases rate restrictions. However, water releases from CY are required to be flat to stabilize the flow of water through Gunnison National Park.

Table 3.2: Forebay elevation constraints at the Aspinall Unit

Operational Constraint	BM	MP	CY
Minimum elevation (ft)	*7,393*	*7,099.8*	*6,747 from March to June* *6,733 the rest of the year*
Maximum elevation (ft)	*7,519.4*	*7,160*	*6,756*
Trigger elevation (ft)	-	*7,144*	*6,748 from March to June* *6,733 the rest of the year*[a]
Maximum drawdown above trigger elevation (ft/24hr)	*Unrestricted*	*Unrestricted*	*4 from March to June* *10 the rest of the year*
Maximum drawdown below trigger elevation (ft/24hr)	*Unrestricted*	*3*	*0.5 from March to June* *5 the rest of the year*

[a] In cases of dry conditions when CY forebay elevation falls below the minimum limit

The main operating guidelines followed by the Aspinall Unit are related to forebay elevation constraints to insure dam safety [21]. The reservoir elevation at BM must always be between 7,393 feet (ft) and 7,519.4 ft. The reservoir elevation at MP must be between 7,099.8 ft and 7,160 ft. If the reservoir elevation is lower than 7,144 ft, the reservoir cannot be drawn down more than 3 ft per day. The reservoir elevation at CY must be between 6,747 ft and 6,756 ft during the wet season (from March to June), and between 6,733 ft and 6,756 ft during the dry season (from July to February). In addition, during the wet season, the reservoir drawdown is not to exceed 4 ft in a 24-hour period. Below a reservoir elevation of 6,748 ft, the reservoir may not be dropped more than 0.5 ft during any 24-hour period. During the dry season, there is a 10-ft per 24-hour period fluctuation limit. If the reservoir elevation is below 6,733 ft, there is a 5-ft per 24-hour period draw down limit. A summary of the Aspinall operating criteria is provided in Table 3.2.

3.6 Other reservoirs and plants

Apart from the three main projects mentioned above, CRSP is composed of several other, smaller, projects. Four of these projects have hydroelectric units:

- the Seedskadee project, with the Fontenelle powerplant (10 MW),
- the Collbran project, with the Lower Molina (4.86 MW) and Upper Molina (8.64 MW) powerplants,
- the Rio Grande project, with the Elephant Butte powerplant (28 MW),
- and the Dolores project, with the McPhee (1.28 MW) and Towaoc (11.5 MW) powerplants.

These reservoirs generate less power than the three main projects and are subject to less operating rules. Therefore, their power generation is not modeled by the hydropower scheduling model but is instead simulated in the preprocessing step.

4 Mathematical models

This section presents the mathematical models that underpin the CRiSPPy software. These models fall into three main categories: (1) ML models used to learn and represent the physical rules governing the reservoirs, power plants, and river system; (2) the hydropower scheduling model, which serves as the core optimization engine of CRiSPPy; and (3) ancillary optimization models that support the preprocessing and post-processing phases of the main model.

4.1 Machine learning models

CRiSPPy uses an empirical and automated approach to model the physical rules that govern the reservoirs, plants, and river system. Instead of being manually entered by the user, the physical relationships are automatically learned and updated from observed data. To achieve this, the CRiSPPy model uses various ML techniques tailored to each rule. The model also directly interfaces with online USBR and U.S. Geological Survey (USGS) databases to retrieve the most recent hydrology and power system data used to train the ML algorithms.

4.1.1 Relationship between reservoir storage volume and forebay elevation

The volume of water in a reservoir obeys a water budget, *i.e.*, the change in storage volume ΔS is equal to the difference between water inflow I and outflow Q.

$$\Delta S = I - Q$$

In CRiSPPy, the storage volume of a reservoir can be modeled at the hourly level using this simple equation. For some reservoirs with operating constraints on elevation changes (*e.g.* MP, CY), it is also important to model the hourly forebay elevation level. This can be achieved by using the relationship between the water storage volume and forebay elevation of a reservoir.

The storage volume of a reservoir is entirely determined by its geometry. As a result, on reasonable time scales with limited sedimentation (*i.e.* a few years), the relationship between the storage volume S of a reservoir and its forebay elevation $E_{forebay}$ is typically deterministic. That is, the forebay elevation can be accurately deduced from the storage volume.

$$E_{forebay} = f(S) \tag{4.1.1.1}$$

Leveraging this fact, CRiSPPy offers a ML method to identify this one-dimensional (1D) relationship using polynomial regression. The function estimating $E_{forebay}$ from S is expressed by:

$$f(x) = \sum_{j=0}^{J} a_j x^j \tag{4.1.1.2}$$

The optimal polynomial function f^* is calculated algebraically using the ordinary least squares method [22]. This calculation is typically instantaneous from the user viewpoint (a few milliseconds).

$$f^* = \text{argmin} \sum_{i \in \{1,\ldots,N\}} (y_i - f(x_i))^2 \tag{4.1.1.3}$$

The degree of the polynomial must be specified by the user. In practice, a degree 2 or 3 proves to be sufficient to accurately capture the relationship between S and $E_{forebay}$, with an

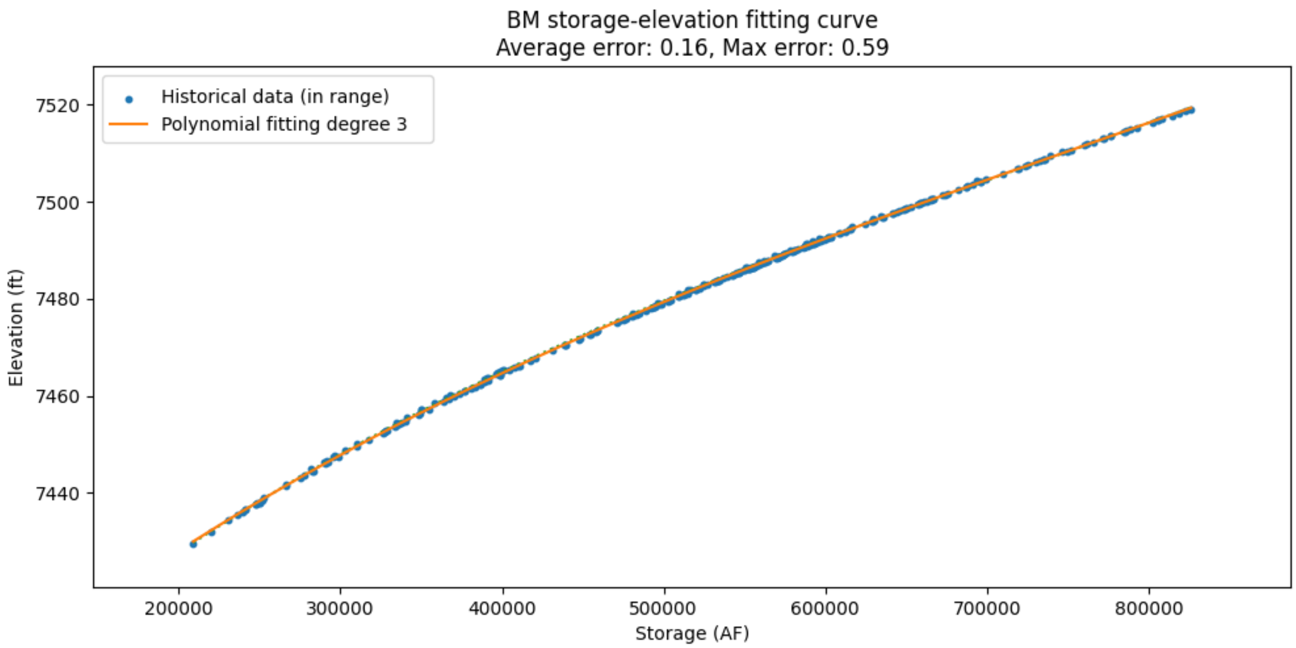

Figure 4.1: Polynomial regression of the storage and elevation data

average approximation error lower than 0.2 ft for all CRSP reservoirs. Similarly, a polynomial regression can also be applied to identify the inverse function, *i.e.* the function estimating S from $E_{forebay}$.

The ML module in CRiSPPy automatically updates this physical relationship by training on data retrieved from the USBR hydrologic database for user-specified time periods [9]. An example of polynomial fitting for the storage-to-elevation relationship is shown in Figure 4.1.

4.1.2 Relationship between release rate, forebay elevation, and water-to-power conversion factors

Hydropower harnesses the potential energy of water stored at a higher elevation. As water flows downhill, its potential energy is converted into kinetic energy. The power P generated by a hydropower plant follows a well-known formula that depends on the density of water ρ, the turbine flow rate Q_{power}, the hydraulic head H, and the turbine efficiency η.

$$P = \rho \cdot g \cdot \eta \cdot H \cdot Q_{power}$$

The water-to-power conversion factor (PCF) describe the amount of energy that can be harnessed for each unit of water flowing through the turbine, and can be expressed as:

$$PCF = \frac{P}{Q_{power}} = \rho \cdot g \cdot \eta \cdot H$$

The hydraulic head represents the difference between the forebay elevation $E_{forebay}$ and the tailrace elevation $E_{tailrace}$. The tailrace elevation generally increases with the total release rate Q_{total} (*i.e.* the water released through and outside the turbine). As the water flows through the penstock or headrace, it experiences friction against the pipe walls and other losses. Because of this, P and PCF are generally calculated using the net head, which is the effective head available at the turbine after accounting for these losses:

$$H = E_{forebay} - E_{tailrace}(Q_{total}) - H_{loss}(Q_{power})$$

The turbine efficiency is not static and varies with the turbine release rate Q_{power} and the hydraulic head H.

$$\eta = \eta(H, Q_{power})$$

From an empirical standpoint, the PCF can be reasonably estimated based on the of the total release rate Q_{total} and the forebay elevation $E_{forebay}$. Historical data from CRSP reservoirs for hydrology and power generation support this approximation.

$$PCF \approx f(Q_{total}, E_{forebay}) \tag{4.1.2.1}$$

To identify the function f that estimates the PCF based on Q_{total} and $E_{forebay}$, the CRiSPPy model offers two ML methods: the first ML method uses a two-dimensional (2D) polynomial regression, and the second ML method uses a 2D continuous piecewise linear (CPWL) fitting.

Polynomial regression is extended to two dimensions using the following expression:

$$f(x, y) = \sum_{j=0}^{J} \sum_{k=0}^{K} a_{j,k} x^j y^k \tag{4.1.2.2}$$

As in 4.1.1, the optimal polynomial function f^* is calculated algebraically using the ordinary least squares method [22].

$$f^* = \operatorname{argmin} \sum_{i \in \{1,\dots,N\}} (z_i - f(x_i, y_i))^2 \tag{4.1.2.3}$$

An example of 2D polynomial fitting for the PCF is shown in Figure 4.2.

A CPWL function is a continuous function composed of several linear pieces. An interesting property of CPWL functions is that they are universal approximators [23], that is, they can uniformly approximate any continuous function to an arbitrary degree of accuracy, given a sufficient number of linear pieces. As a result, they represent a pertinent group of fitting functions. A CPWL function can be formulated as a difference-of-convex (DC) function [24]:

$$f(x, y) = \max_{j \in \{1,\dots,P^+\}} \left(a_j^+ x + b_j^+ y + c_j^+\right) - \max_{k \in \{1,\dots,P^-\}} \left(a_k^- x + b_k^- y + c_k^-\right) \tag{4.1.2.4}$$

In CRiSPPy, this fact is leveraged to identify the best CPWL approximation of the PCF function. The optimal CPWL approximation can be identified by solving a mixed integer linear programming (MILP) problem [25]. The MILP problem identifies the linear coefficients that minimizes the maximum error between the PCF values and their CPWL approximation.

$$f^* = \operatorname{argmin} \max_{i \in \{1,\dots,N\}} |z_i - f(x_i, y_i)| \tag{4.1.2.5}$$

An example of 2D CPWL fitting for the PCF is shown in Figure 4.3.

Although easier to compute, the polynomial fitting method generally yields lower quality approximations compared to the CPWL fitting method due to its nonlocal behavior and poor

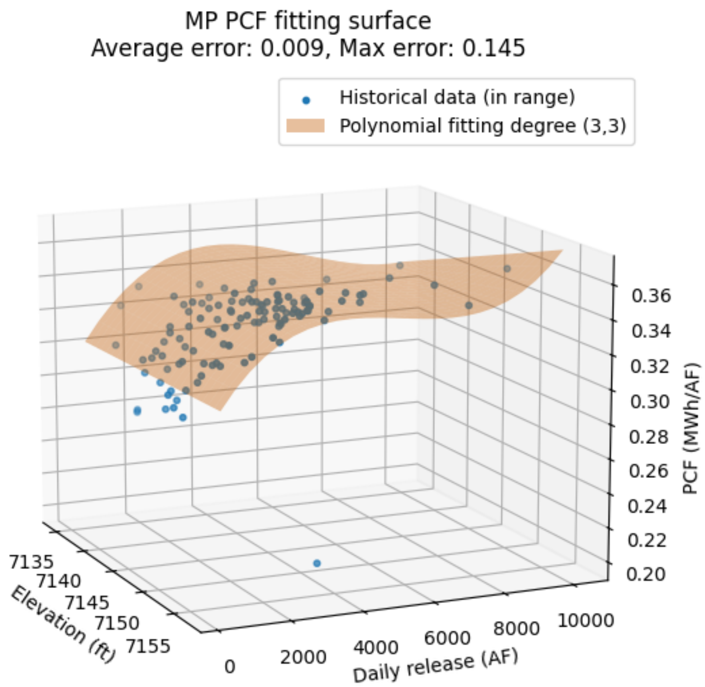

Figure 4.2: 2D polynomial regression of the PCF data

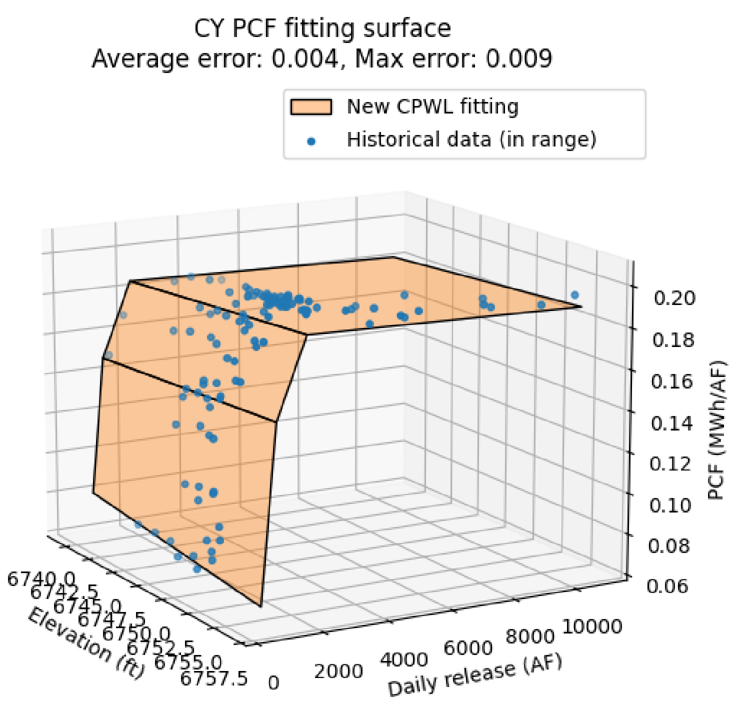

Figure 4.3: 2D CPWL regression of the PCF data

extrapolation [26]. This is notably the case for a few hydropower plants (*e.g.* CY), whose data set do not exhibit a polynomial underlying relationship and have sparse data points at the edges of the observed range. As a result, the CPWL fitting method is generally preferred when estimating the PCF based on the total release rate and the forebay elevation.

As for the previous relationship, the ML module in CRiSPPy automatically updates this physical relationship by training on data retrieved from the USBR hydrologic database for user-specified time periods [9].

4.1.3 Water travel time distribution

The water travel time distribution (TTD) describes the range of times it takes for water to travel from a specific upstream river source to a downstream river destination. Instead of a single travel time, it acknowledges that water particles originating from the same source can take different pathways and experience varying velocities, leading to a distribution of arrival times at the destination. At hourly timescales and for relatively short distances (a few river miles), the travel time is negligible and can be considered instantaneous. However, for longer distances, the travel time becomes significant and a TTD might be required to model the downstream impact of upstream flows.

Operational rules at the FG reservoir require CRiSPPy to model the flow and stage at the Jensen gage, located 94 river miles downstream of the FG reservoir (see section 3.4). In the previous GTMax SL model, the physics-based Streamflow Synthesis and Reservoir Regulation (SSARR) model [27] was used to estimate the impact of FG releases at the Jensen gage. However, hydrology conditions have evolved since the SSARR model was first applied to the region. Moreover, as a physics-based model, the SSARR model parameters need to be manually updated by conducting detailed studies about the new river shape and soil conditions, which is not always feasible or practical. As a result, a linear programming (LP)-based ML algorithm was developed by ANL to automatically identify and update the TTDs involved at the Jensen gage using recent historical flow data [28]. Owing to the assumption of constant Yampa flows during a monthly period (see section 3.4), only the TTD corresponding to the reach between FG and the Jensen gage needs to be calculated. The LP problem is formulated as follow:

$$\min \left(\sum_{h \in H} e(h) + w_C \cdot C^{max} \right) \tag{4.1.3.1}$$

$$-e(h) \leq \widehat{Q}_{down}(h) - Q_{down}(h) \leq e(h), \quad h \in H \tag{4.1.3.2}$$

$$\widehat{Q}_{down}(h) = \sum_{h' \in W} TTD(h') \cdot Q(h - h') + I_{down}, \quad h \in H \tag{4.1.3.3}$$

$$-C^{max} \leq TTD(h+1) - 2TTD(h) + TTD(h-1) \leq C^{max}, \quad h \in W \tag{4.1.3.4}$$

The problem is formulated at an hourly time resolution h. H represents the time scope of the observed releases, whereas W represents the convolution window (typically 1 or 2 days). The LP model co-minimizes the downstream flow estimation error $e(h)$ and the maximum curvature C^{max} of the TTD. The weight w_C is chosen in such a way that the resulting TTD profile is relatively smooth while having a negligible impact on the average estimation error. The estimation error is the difference between the observed and estimated downstream flow $Q_{down}(h)$ and $\widehat{Q}_{down}(h)$, respectively. $\widehat{Q}_{down}(h)$ is estimated as the sum of two components: a constant

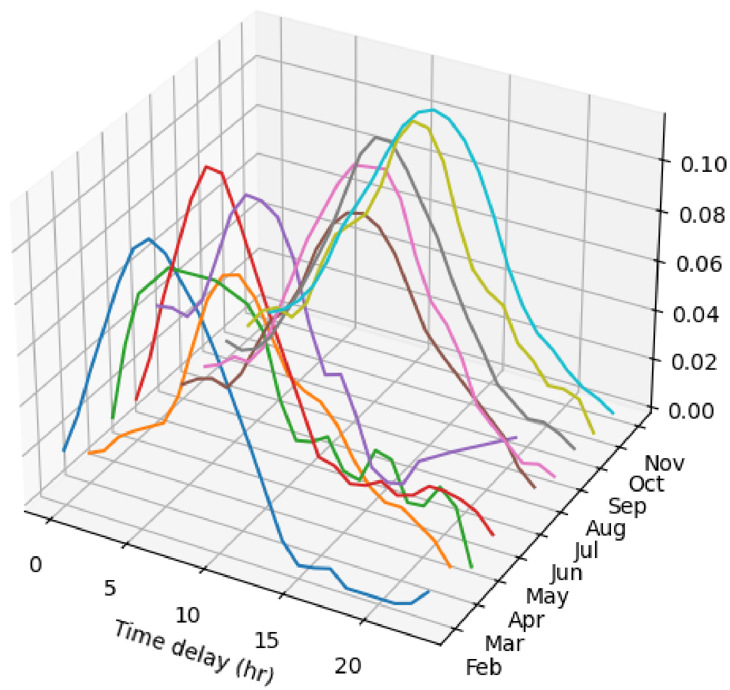

Figure 4.4: TTDs between FG and the Jensen gage across different months of the year

tributary flow I_{down} (Yampa flow), and the convolution of an upstream dam release (FG) and its corresponding TTD.

This ML model assumes a linear relationship between Yampa flows, FG releases, and the flows at the Jensen gage. In reality, the relationship between these flows is more complex and include nonlinearities. However, under similar hydrology conditions (*e.g.*, the same month of the year), the model produces reasonable downstream flow estimates with small error. In CRiSPPy, optimal TTD coefficients are calculated for each month of the year using historical data of the corresponding month.

The TTD coefficients can be automatically updated in CRiSPPy by training the ML model using data downloaded from the USGS water database at user-specified time periods [29]. The calculated TTDs between the FG dam and the Jensen gage are shown in Figure 4.4. Insufficient data were available to calculate TTDs in December and January, as the river at Jensen was frozen for most of that period.

4.1.4 Relationship between river flow and river stage

For a stable river channel at a specific cross-section, the relationship between the river stage and the flow rate is typically deterministic. As the water level rises (higher stage), the cross-sectional area of the flow increases, and the water velocity also increases, leading to a higher flow rate (discharge).

As seen in the previous section, operational rules at the FG reservoir require CRiSPPy to model the flow and stage at the Jensen gage. The stage at the Jensen gage can be estimated based on Jensen flow rate using an empirical stage-discharge curve. As in section 4.1.1, a polynomial regression can be used to identify the empirical stage-discharge curve. Another alternative

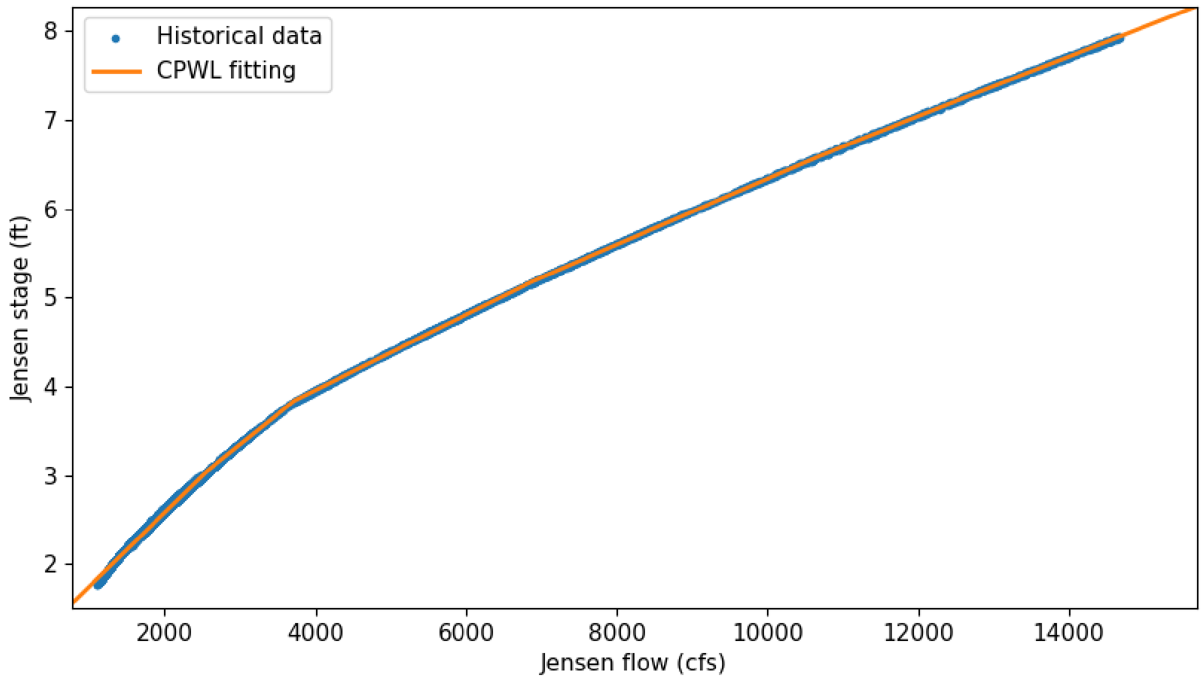

Figure 4.5: CPWL regression of the Jensen flow and stage data

consists in using a 1D CPWL approximation. The 1D CPWL approximation is particularly relevant when modeling the the stage-discharge curve more accurately in a MILP model. In CRiSPPy, the stage-discharge relationship at the Jensen gage is linearized for performance purposes. However, a spinoff software of CRiSPPy, called "MiniFG", was developed to more accurately identify the optimal release schedule at FG only. In "MiniFG", the 1D CPWL approximation of the stage-discharge relationship is used in the MILP formulation of the model.

As for section 4.1.2, the optimal CPWL approximation of this relationship can be identified by solving a MILP problem. Owing to recent tightening algorithms, the optimal 1D CPWL approximation can be identified very rapidly [30]. As for section 4.1.3, the stage-discharge relationship can be identified using data downloaded from the USGS water database [29]. The optimal CPWL fitting for the Jensen gage's flow-to-stage relationship is shown in Figure 4.5.

4.1.5 Relationship between forebay elevation and power plant capacity

As seen in section 4.1.2, the capacity of an hydropower plant is typically proportional to its hydraulic head. Assuming little variations in the tailrace elevation, the relationship between forebay elevation and power capacity can be considered deterministic.

Contrary to the previous physical rules, historical data about the relationship between forebay elevation and power plant capacity is significantly more limited. This is mainly due to operating rules described in section 3 that prevent power plants from operating at full capacity and revealing the plant true power capacity. For example, the fact that CY must operate at constant flow rate typically prevents the powerplant from reaching peak power capacity. The daily change rule and recent low monthly release targets have a similar impact on GC operations.

As a result, in CRiSPPy, the relationship between forebay elevation and power plant capacity

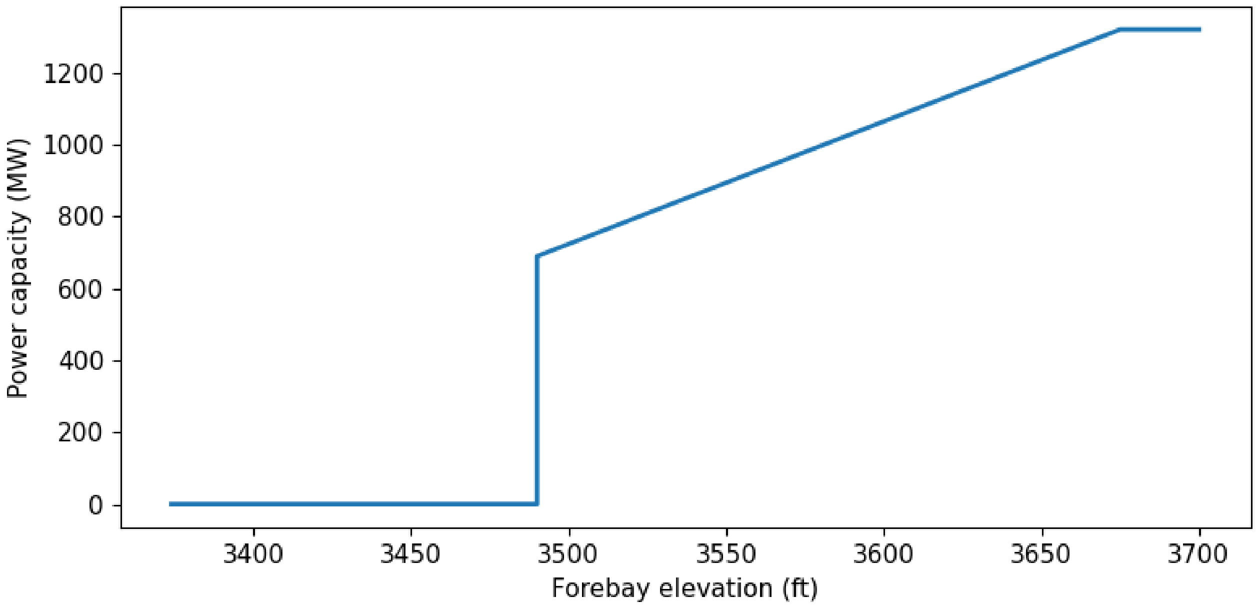

Figure 4.6: Relationship between forebay elevation and power capacity

is the only physical rule that is not updated by a ML algorithm but is instead manually entered by the user. This physical rule is modeled by a 1D CPWL function and can generally be found in USBR official reports [31]. For example, the relationship between forebay elevation and power capacity at GC is depicted in Figure 4.6. Above the minimum power pool elevation of 3,490 ft, the relationship between GC forebay elevation and power capacity is typically linear, and capped by the combined turbine limit of 1,320 MW. However, below this level, water can no longer reach the turbine intakes, causing the power capacity to drop to zero.

4.2 Hydropower scheduling model

The hydropower scheduling model is the core optimization model of CRiSPPy. It is formulated as an LP model for performance purposes. It computes the optimal hourly water release schedules of the CRSP reservoirs in a given month. These optimal hourly releases are driven by customer loads and energy prices while accounting for monthly release targets and physical and operational rules. This section describes the detailed mathematical equations that govern the hydropower scheduling model. The list of sets, parameters, and variables is summarized in Table 4.1.

Table 4.1: List of Sets, Parameters, and Variables of the Hydropower scheduling model

Symbol	**Unit**	**Description**
		Sets
$p \in \Pi$	-	Set of plants/reservoirs
$U(p) \subset \Pi$	-	Set of upper reservoirs of reservoir p
$h \in H$	-	Set of hours

Continued on next page

Symbol	Unit	Description
$H_{up} \subset H$	-	Set of up-ramping hours
$H_{DA} \subset H$	-	Set of hours in which day-ahead purchases and sales can change
$H_{SR} \subset H$	-	Set of hours in which spinning reserve purchases can change
$h \in W$	-	Set of hours of the convolution window
$d \in D$	-	Set of days
	Parameters	
$w(h)$	hr	Hourly weight representing the frequency of hour h in the time representation
$Q_{target}(p)$	AF	Monthly release volume of plant p
$Q^{min}(p,h)$	AF/hr	Minimum release rate of plant p in hour h
$Q^{max}(p,h)$	AF/hr	Maximum release rate of plant p in hour h
$R^{min}(p,h)$	AF/hr^2	Minimum release ramp rate of plant p in hour h
$R^{max}(p,h)$	AF/hr^2	Maximum release ramp rate of plant p in hour h
$R_{up}(p)$	AF/hr^2	Maximum up-ramp rate of plant p
$R_{down}(p)$	AF/hr^2	Maximum down-ramp rate of plant p
$DC(p)$	AF/24-hr^2	Maximum change in release rate at plant p in a 24-hr period
$DF(p,d)$	-	Minimum daily release volume factor of plant p in day d
$I(p,h)$	AF/hr	Unregulated water inflow in reservoir p in hour h
$S_{init}(p)$	AF	Initial reservoir storage volume of reservoir p
$S^{min}(p)$	AF	Minimum reservoir storage volume of reservoir p
$S^{max}(p)$	AF	Maximum reservoir storage volume of reservoir p
$DD(p)$	ft/24-hr	Maximum elevation drawdown of reservoir p in a 24-hr period
$StoE(p)$	ft/AF	Change in reservoir elevation per unit increase in storage volume of reservoir p
$TTD(p,h)$	-	Water travel time distribution of releases from reservoir p to the downstream river location
$I_{down}(p,h)$	AF/hr	Unregulated inflow from tributary source to the river location downstream of reservoir p in hour h
$QtoE(p)$	ft/AF/hr	Change in stage per unit increase in flow rate at the river location downstream of reservoir p
$\Delta E^{max}_{down}(p)$	ft	Maximum change in stage at the river location downstream of reservoir p
$PCF(p)$	MWh/AF	Water-to-power conversion factor of plant p
$P^{max}(p)$	MW	Power capacity of plant p
RD_{system}	MW	Regulation down required by the system
RU_{system}	MW	Regulation up required by the system
SR_{system}	MW	Spinning reserve required by the system
$t(p)$	-	Ratio of power generation from plant p not converted in transmission losses

Continued on next page

Symbol	Unit	Description
$L(h)$	MW	Total system load in hour h
$\pi(h)$	\$/MWh	Energy price in hour h
$SRC(p)$	\$/MWh	Penalty cost for providing spinning reserve with from plant p
$RUC(p)$	\$/MWh	Penalty cost for providing regulation up with from plant p
$RDC(p)$	\$/MWh	Penalty cost for providing regulation down with from plant p
$DASC$	\$/MWh	Penalty cost for selling energy on the day-ahead market
$DAPC$	\$/MWh	Penalty cost for purchasing energy on the day-ahead market
$RTSC$	\$/MWh	Penalty cost for selling energy on the real-time market
$RTPC$	\$/MWh	Penalty cost for purchasing energy on the real-time market
$RUPC$	\$/MWh	Penalty cost for purchasing regulation up
$RDPC$	\$/MWh	Penalty cost for purchasing regulation down
$SRPC$	\$/MWh	Penalty cost for purchasing spinning reserve
		Continuous Variables
$Q(p,h)$	AF/hr	Total release rate of plant p in hour h
$Q_{power}(p,h)$	AF/hr	Power (turbine) release rate of plant p in hour h
$Q_{nonpower}(p,h)$	AF/hr	Non-power release rate of plant p in hour h
$Q_{ref}(p)$	AF/hr	Reference release rate of plant p for the daily change constraint
$Q_{day}(p)$	AF	Reference daily release volume of plant p for the daily release volume constraint
$\Delta S(p,h)$	AF/hr	Change in storage volume of reservoir p in hour h
$\Delta E(p,h)$	ft/hr	Change in forebay elevation of reservoir p in hour h
$Q_{down}(p,h)$	AF/hr	Flow rate of the river point downstream of reservoir p in hour h
$Q_{down}^{min}(p)$	AF/hr	Lowest flow rate of the river point downstream of reservoir p
$Q_{down}^{max}(p)$	AF/hr	Highest flow rate of the river point downstream of reservoir p
$P(p,h)$	MW	Power generated by plant p in hour h
$RD(p,h)$	MW	Regulation down provided by plant p in hour h
$RU(p,h)$	MW	Regulation up provided by plant p in hour h
$SR(p,h)$	MW	Spinning reserve provided by plant p in hour h
$RDP(h)$	MW	Regulation down purchased in hour h
$RUP(h)$	MW	Regulation up purchased in hour h
$SRP(h)$	MW	Spinning reserve purchased in hour h
$DAP(h)$	MW	Amount of day-ahead purchases in hour h
$RTP(h)$	MW	Amount of real-time purchases in hour h
$DAS(h)$	MW	Amount of day-ahead sales in hour h

Continued on next page

Symbol	Unit	Description
$RTS(h)$	MW	Amount of real-time sales in hour h

4.2.1 Time representation

Owing to the reservoirs' monthly release volume constraints, CRSP hourly operations are typically independent between months. Consequently, a case study can be broken down into several model runs, each consisting of a single month of operations, that can be independently solved.

The user has two time representation options to model hourly operations:

- Model all hours in a month (*e.g.* 744 hours for January, 720 hours for April)
- Model the month using a representative week (Sunday to Saturday) composed of 168 hours ($h \in \{0, \dots, 167\}$)

The latter time representation is a model simplification intended to speed solution times for large case studies. In this time representation, hourly weights $w(h)$ are used to scale the results to represent the entire month. More specifically, $w(h)$ represents the frequency of the weekday of index $\lfloor h/24 \rfloor$. For example, $w(0)$ is equal to the number of Sundays and holidays in the month, whereas $w(24)$ is equal to the number of non-holiday Mondays in the month.

Due to their similarities in terms of price and customer load patterns, the following federal holidays are modeled identically to Sundays: New Year's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving Day, and Christmas Day.

4.2.2 Water release constraints

The hydropower scheduling model represents both power (*i.e.* turbine) and non-power (*i.e.* bypass, spillway) releases. Total water releases are the sum of power and non-power releases.

$$Q(p,h) = Q_{power}(p,h) + Q_{nonpower}(p,h) \quad \forall p, \forall h \tag{4.2.2.1}$$

WAPA is required to meet monthly volume targets that are specified for each reservoir, resulting in the weighted sum of hourly releases in a given model run to be fixed.

$$\sum_{h \in H} w(h) \cdot Q(p,h) = Q_{target}(p) \quad \forall p \tag{4.2.2.2}$$

When modeling all hours in a month, $w(h) = 1$. Otherwise, when modeling a representative week, $w(h)$ represents the frequency of the corresponding weekday in the month.

Certain CRSP reservoirs operate under specific flow rate constraints. For instance, GC is subject to both maximum and minimum flow limits (see section 3.3), while FG is only subject to a minimum flow rate (see section 3.4).

$$Q^{min}(p,h) \leq Q(p,h) \leq Q^{max}(p,h) \tag{4.2.2.3}$$

For other reservoirs (BM, MP, CY): $Q^{min}(p,h) = 0$ and $Q^{max}(p,h) = \infty$.

Certain CRSP reservoirs are subject to specific ramp constraints. This is the case of GC and FG.

$$R^{min}(p,h) \leq Q(p,h) - Q(p,h-1) \leq R^{max}(p,h) \tag{4.2.2.4}$$

In addition, all reservoirs in the CRSP system are subject to cycling constraints. More specifically, reservoir releases are required to ramp up no more than two times per day. In CRiSPPy, this cycling constraint is enforced by calculating in advance an optimal set of up-ramping hours H_{up}. The 24 hours of the days are partitioned into up-ramping hours H_{up} and down-ramping hours H_{down}. The partitioning is calculated in such a way that a 24-hour cycling period does not contain more than two up-ramping periods. More details about the ramping hour partitioning algorithm is provided in section 4.3.2. The following ramping limit values are used to simultaneously enforce ramping and cycling constraints.

$$R^{max}(p,h) = \begin{cases} R_{up}(p) & \text{if } h \in H_{up} \\ 0 & \text{if } h \notin H_{up} \end{cases}, \quad R^{min}(p,h) = \begin{cases} -R_{down}(p) & \text{if } h \notin H_{up} \\ 0 & \text{if } h \in H_{up} \end{cases}$$

For reservoirs with no ramping limits, *i.e.* BM and MP:

$$R_{up}(p) = R_{down}(p) = \infty$$

For CY, which is required to operate at a constant flow rate:

$$R_{up}(p) = R_{down}(p) = 0$$

GC water releases are subject to a maximum daily change $DC(p)$. Due to its ramping and cycling constraints, GC release is expected to reach the same minimum and maximum flow rate each day of the month. Owing to this, we can enforce the maximum 24-hour flow change using the equation below, where $Q_{ref}(p)$ is a model variable representing a reference release rate whose value is identified by the solver.

$$Q_{ref}(p) \leq Q(p,h) \leq Q_{ref}(p) + DC(p) \tag{4.2.2.5}$$

GC water releases are subject to daily volume restrictions. Specifically, the reservoir must maintain a daily release volume between $DF(p,d) \cdot Q_{day}(p)$ and $Q_{day}(p)$. $DF(p,d)$ is typically equal to 100% during the weekdays and 85% during the weekend. This constraint is expressed by the equation below.

$$DF(p,d) \cdot Q_{day}(p) \leq \sum_{24d \leq h \leq 24(d+1)} Q(p,h) \leq Q_{day}(p) \tag{4.2.2.6}$$

As for $Q_{ref}(p)$, the value of $Q_{day}(p)$ is not known in advance but is identified by the solver.

4.2.3 Storage and elevation constraints

The storage volume of a reservoir is governed by a water budget equation. Specifically, the change in storage volume $\Delta S(p,h)$ at a given reservoir p is equal to the difference between inflows and outflows. A reservoir's outflow is essentially its total release rate $Q(p,h)$. Its inflow consists of unregulated inflows and, where applicable, water releases from upstream reservoirs $U(p)$.

$$\Delta S(p,h) = I(p,h) - Q(p,h) + \sum_{p' \in U(p)} Q(p',h) \tag{4.2.3.1}$$

Some reservoirs are subject to specific forebay elevation limits. Owing to the deterministic relationship between reservoir storage and forebay elevation (section 4.1.1), these limits can be equivalently expressed as storage volume limits.

$$S^{min}(p) \leq S_{init}(p) + \sum_{h'<=h} \Delta S(p, h') \leq S^{max}(p) \tag{4.2.3.2}$$

The explicit representation of hourly changes in elevation is necessary to formulate the maximum drawdown constraint below.

$$\sum_{0 \leq h'' < h'} \Delta E(p, h + h'') \geq -DD(p), \quad 0 \leq h < N_{hours}, \quad 1 \leq h'' < DD_{hours}(p) \tag{4.2.3.3}$$

Albeit deterministic, the relationship between storage and elevation is non-linear. However, it can be reasonably approximated as linear over a monthly time frame.

$$\Delta E(p, h) = StoE(p) \cdot \Delta S(p, h) \tag{4.2.3.4}$$

4.2.4 Downstream flow constraints

The flow rate $Q_{down}(p, h)$ at a river point located downstream of a reservoir p can be reasonably estimated as the sum of two flow components: a tributary component $I_{down}(p, h)$, and the convolution of the reservoir release rate $Q(p, h)$ and its corresponding TTD on the river reach $TTD(p, h)$. The TTD can be thought as the probabilistic time it takes for a unit of water released from the reservoir to reach the downstream location. In CRiSPPy, the empirical TTD is identified from historical data using ML techniques, as seen in section 4.1.3.

$$Q_{down}(p, h) = \sum_{h' \in W} Q(p, h - h') \cdot TTD(p, h') + I_{down}(p, h) \tag{4.2.4.1}$$

In reservoirs like FG, modeling the downstream flow rate is necessary to constrain the maximum change in downstream flow height $\Delta E^{max}_{down}(p)$. The relationship between the flow height of a river and its flow is typically deterministic. Although non-linear, this relationship can be linearized for relatively small variations in flow.

$$Q^{min}_{down}(p) \leq Q_{down}(p, h) \leq Q^{max}_{down}(p) \tag{4.2.4.2}$$

$$QtoE(p) \cdot (Q^{max}_{down}(p) - Q^{min}_{down}(p)) \leq \Delta E^{max}_{down}(p) \tag{4.2.4.3}$$

4.2.5 Power system constraints

At monthly timescales, the power generation $P(p, h)$ of CRSP hydropower plants is typically proportional to their power release rate $Q_{power}(p, h)$. The PCF is calculated during the preprocessing step based on the hydrology conditions and the PCF function identified via the ML model described in section 4.1.2.

$$P(p, h) = PCF(p) \cdot Q_{power}(p, h) \tag{4.2.5.1}$$

The power generation is primarily limited by the power capacity $P^{max}(p)$ of the plant, which depends on the elevation and unit availability and is calculated during the preprocessing step.

Scheduled or forced hourly outages can be modeled by using the hourly power capacity parameter $P^{max}(p,h)$. In addition, ancillary services such as up-regulation $RU(p,h)$, down-regulation $RD(p,h)$, and spinning reserve $SR(p,h)$ further limit the operational range of the power plants.

$$RD(p,h) \leq P(p,h) \leq P^{max}(p) - RU(p,h) - SR(p,h) \tag{4.2.5.2}$$

The ancillary service requirements of the system can either be met by the CRSP plants themselves or via external purchases.

$$\sum_p RU(p,h) + RUP(h) \geq RU_{system} \tag{4.2.5.3}$$

$$\sum_p RD(p,h) + RDP(h) \geq RD_{system} \tag{4.2.5.4}$$

$$\sum_p SR(p,h) + SRP(h) \geq SR_{system} \tag{4.2.5.5}$$

The system obeys a power balance equation. Because CRSP generation resources may not perfectly align with customer loads, purchases and sales must necessary be made to balance the power system. At any time, the sum of plant power generation and energy day-ahead and real-time purchases $(DAP(h), RTP(h))$ must meet the customer loads $L(h)$ and energy sales $(DAS(h), RTS(h))$. Transmission losses of each plant are accounted for using a transmission factor $t(p)$.

$$\sum_p t(p) \cdot P(p,h) + DAP(h) + RTP(h) = DAS(h) + RTS(h) + L(h) \tag{4.2.5.6}$$

Finally, day-ahead energy purchases and sales, in addition to spinning reserve purchases, are typically made in "blocks", that is, their values only change in few hours of the day.

$$DAP(h) - DAP(h-1) = 0 \quad \forall h \notin H_{DA} \tag{4.2.5.7}$$

$$DAS(h) - DAS(h-1) = 0 \quad \forall h \notin H_{DA} \tag{4.2.5.8}$$

$$SRP(h) - SRP(h-1) = 0 \quad \forall h \notin H_{SR} \tag{4.2.5.9}$$

4.2.6 Objective function

The hydropower scheduling model aims to meet customer loads with CRSP resources to the extent possible. This is achieved by minimizing purchases and sales made to balance the power system. Purchases and sales are driven by the market price $\pi(h)$ while being minimized using penalty costs ($DASC$, $DAPC$, $RTSC$, $RTPC$). Similarly, external purchases of ancillary services are minimized using penalty costs ($RUPC$, $RDPC$, $SRPC$), prioritizing the use of CRSP power plants to deliver these services. Some CRSP power plants have higher priority in delivering ancillary services. Typically, the three power plants expected to deliver such services are, in decreasing order of priority: GC, BM, and MP. This priority order is reflected in the plant-based penalty costs $RUC(p)$, $RDC(p)$, $SRC(p)$.

$$
\begin{aligned}
\max \sum_h \sum_p w(h) \cdot \pi(h) \cdot P(p,h) \\
+ \sum_h \sum_p w(h) \cdot (-SRC(p) \cdot SR(p,h) - RUC(p) \cdot RU(p,h) - RDC(p) \cdot RD(p,h)) \\
+ \sum_h w(h) \cdot (DAS(h) \cdot (\pi(h) - DASC) + RTS(h) \cdot (\pi(h) - RTSC)) \\
+ \sum_h w(h) \cdot (DAP(h) \cdot (-\pi(h) - DAPC) + RTP(h) \cdot (-\pi(h) - RTPC)) \\
+ \sum_h w(h) \cdot (-RUP(h) \cdot RUPC - RDP(h) \cdot RDPC - SRP(h) \cdot SRPC)
\end{aligned}
\tag{4.2.6.1}
$$

4.2.7 Constraint priority ranking

The model formulation described above is feasible for most hydrology conditions. However, in cases of extreme (wet or dry) hydrology conditions, the hydropower scheduling model described above might be infeasible, *i.e.*, no solution exists. In addition, CRiSPPy allows the user to fix the hourly power and non-power release profiles in user-defined time periods, which may conflicts with some operating rules. To address this, the CRiSPPy model may relax some constraints based on their predefined priority level. While physical rules cannot be lifted, some operational rules have higher priority than others. For example, monthly release volumes and ramping constraints have higher priority than flow rate limits which, in turn, have higher priority than cycling constraints. As a result, infeasible model runs can be made feasible by relaxing the constraints with lowest priority. The priority level associated with each constraint is described in table 4.2.

4.3 Other optimization models

In addition to the core hydropower scheduling model, CRiSPPy uses several other optimization models in the preprocessing and post-processing steps.

4.3.1 Hourly price generation model

The hydropower scheduling model uses hourly market price profiles to identify optimal energy purchases and sales. In some cases, such as forward (*i.e.*, future-looking) analyses, available input price data are limited. Instead of detailed hourly price profiles, the only available input data are monthly average values during "on-peak" and "off-peak" periods. The "on-peak" and "off-peak" periods refer to periods of time that were traditionally linked to "on-peak" and "off-peak" price values. For the purposes of this section, the set of "on-peak" hours H_{on} are defined as the hours from 8 to 23 on Mondays through Saturdays (excluding federal holidays), whereas the set of "off-peak" hours H_{off} refer to the rest of the hours.

The CRiSPPy model implements an hourly price generation model that converts input monthly "on-peak" and "off-peak" values into a reasonable hourly price profiles. A reference hourly profile $L(h)$ (*e.g.*, FES customer load, Western Electricity Coordinating Council (WECC) load, or California Independent System Operator (CAISO) price) is used to shape the output

Table 4.2: Priority ranking of physical, user-defined, and operating constraints

Constraint Type	Equations	Priority level
Physical constraints	4.2.2.1, 4.2.3.1, 4.2.3.2[a], 4.2.3.4, 4.2.4.1, 4.2.5.1, 4.2.5.2, 4.2.5.6	*Strictly enforced*
User-defined hourly power and non-power releases	-	*Strictly enforced*
Monthly release targets	4.2.2.2	*Highest*[b]
Ramping constraints	4.2.2.4	*Very high*
Daily flow change	4.2.2.5	*Very high*
Maximum drawdown	4.2.3.3	*Very high*
Downstream stage limits	4.2.4.2, 4.2.4.3	*Very high*
Ancillary service requirements	4.2.5.3, 4.2.5.4, 4.2.5.5	*Very high*
Block purchase requirements	4.2.5.7, 4.2.5.8, 4.2.5.9	*Very high*
Flow rate limits	4.2.2.3	*High*
Forebay elevation limits	4.2.3.2[c]	*High*
Cycling constraints	4.2.2.4	*Medium*
Daily release constraints	4.2.2.6	*Medium*

[a] When storage limits refer to dead and crest pool
[b] Only lifted if they conflict with user-defined hourly releases
[c] When storage limits refer to regulatory limits

hourly price profile $\pi(h)$. More specifically, it is assumed that there is a deterministic monotonic piecewise linear function f such that $\pi(h) = f(L(h))$. The goal of the hourly price generation model is to identify the optimal price profile $\pi(h)$ such that the price range and the curvature of f are co-minimized while matching the average "on-peak" and "off-peak" price targets. To achieve this, the function f is divided into K linear pieces f_k whose domain is defined as $D_k = [\widetilde{L}(k), \widetilde{L}(k+1)]$, where:

$$\widetilde{L}(k) = \min_h L(h) + \frac{k}{K}(\max_h L(h) - \min_h L(h)), \quad 0 \leq k \leq K$$

The optimal hourly price profile is then identified by solving the following LP model.

$$\min \left(w_C \cdot C^{max} + w_\pi \cdot (\pi^{max} - \pi^{min}) \right) \tag{4.3.1.1}$$

$$\sum_{h \in H_{on}} \pi(h) = |H_{on}| \cdot \pi_{on} \tag{4.3.1.2}$$

$$\sum_{h \in H_{off}} \pi(h) = |H_{off}| \cdot \pi_{off} \tag{4.3.1.3}$$

$$\pi^{min} \leq \pi(h) \leq \pi^{max}, \quad \forall h \tag{4.3.1.4}$$

$$\pi(h) = a(k) \cdot L(h) + b(k), \quad \forall (h,k) : L(h) \in D_k \tag{4.3.1.5}$$

$$a(k+1) \cdot \widetilde{L}(k+1) + b(k+1) = a(k) \cdot \widetilde{L}(k+1) + b(k), \quad 0 \leq k < K-1 \tag{4.3.1.6}$$

$$-C^{max} \leq a(k+1) - a(k) \leq C^{max}, \quad 0 \leq k < K-1 \tag{4.3.1.7}$$

$$a^{min} \leq a(k) \leq a^{max}, \quad 0 \leq k < K \tag{4.3.1.8}$$

The user can adjust the weights w_C and w_π to prioritize minimizing either the curvature of the function f or the price range.

Note the model is only feasible if the average "on-peak" and "on-peak" values of the reference profile $L(h)$ are oriented in the same direction as the target "on-peak" and "off-peak" price values. If this is not the case, an additional step is required to adjust the reference profile. For example, if the load profile is used as the reference profile, and the order of the average "on-peak" and "off-peak" values is opposite to the order of the target "on-peak" and "off-peak" prices, an adjusted "net load" profile can be used, in which additional solar generation is accounted for. In CRiSPPy, this additional step is performed automatically. An illustration of the hourly price generation model result is depicted in Figure 4.7.

In future versions of the CRiSPPy model, a new version of the hourly price generation model, based on deep learning techniques, will be developed.

4.3.2 Ramping time partitioning model

The CRiSPPy model implements a ramping time partitioning model designed to identify the optimal time for reservoir releases to ramp up and down during a typical day. This partitioning model, executed during the preprocessing step (before the hydropwer scheduling model) is necessary to impose a predefined number of ramping cycles while keeping the hydropower scheduling model linear.

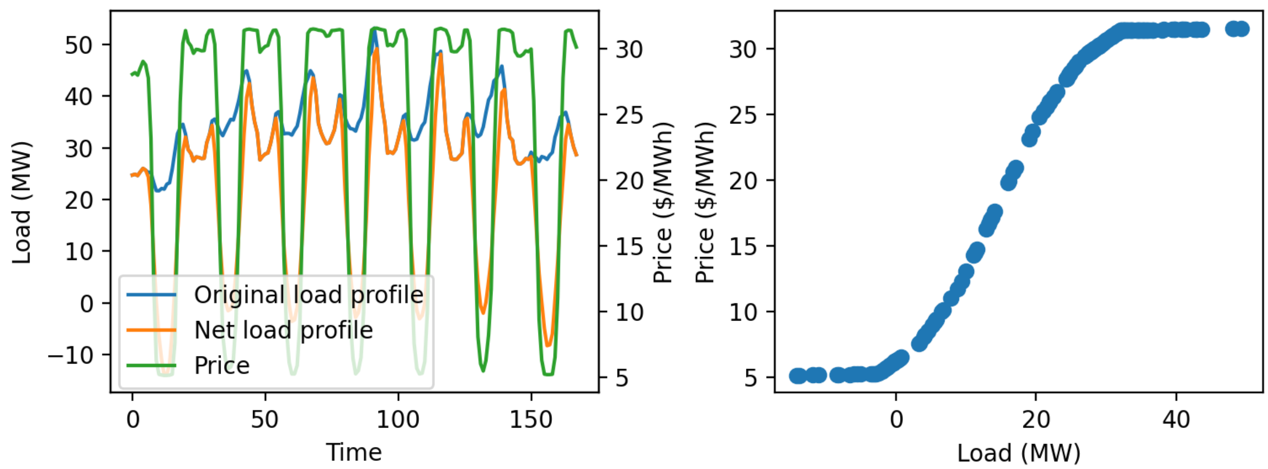

Figure 4.7: Hourly price generation model

Let $D(h)$ be the hourly driver behind water release ramping pattern, *e.g.* the price or load profile. The driver ramp $R(h)$ is defined as the discrete derivative of the driver $D(h)$, where the modulo operation is used to account for the cyclical nature of the driver.

$$R(h) = D(h) - D((h-1) \bmod N_{hours}), \quad 0 \le h < N_{hours}$$

The set of hours H of a water release profile can be partitioned into two subsets H_{up} and $\overline{H_{up}}$, where H_{up} is the subset of hours in which the water release of a plant is ramping up and $\overline{H_{up}}$ where it is ramping down. The number of cycles of a water release profile can be defined as the number of intervals in H_{up} composed of consecutive hours (with cyclical ordering).

The goal of the ramping time partitioning model is to identify the optimal partitioning $\left(H_{up}, \overline{H_{up}}\right)$ that is consistent with the value of the driver ramp $R(h)$ while maintaining a number of cycles that is lower or equal than a target number N_{cycles}. This can be achieved by minimizing the sum of hourly deviations $d(h)$ between the driver ramp and the release ramp using the following MILP formulation.

$$\min \sum_{0 \le h < N_{hours}} d(h) \tag{4.3.2.1}$$

$$-M_{down} \cdot (1 - T_{up}(h)) - d(h) \le R(h) \le M_{up} \cdot T_{up}(h) + d(h), \quad 0 \le h < N_{hours} \tag{4.3.2.2}$$

$$T_{down\_to\_up}(h) \ge T_{up}(h) - T_{up}((h-1) \bmod N_{hours}), \quad 0 \le h < N_{hours} \tag{4.3.2.3}$$

$$\sum_{0 \le h < N_{hours}} T_{down\_to\_up}(h) \le N_{cycles} \tag{4.3.2.4}$$

$$d(h) \ge 0, \quad 0 \le h < N_{hours} \tag{4.3.2.5}$$

$$T_{up}(h) \in \{0, 1\}, \quad 0 \le h < N_{hours} \tag{4.3.2.6}$$

$T_{up}(h)$ is a binary variable indicating whether the hour h is a ramping hour. $T_{down\_to\_up}(h)$ indicates whether the release switches from the down to up direction in hour h. M_{up} and M_{down} are big-M parameters. In a given hour, when the driver ramp and the release ramp are in the same direction ($T_{up}(h) = 1$ and $R(h) \ge 0$, or $T_{up}(h) = 0$ and $R(h) \le 0$), we have $d(h) = 0$.

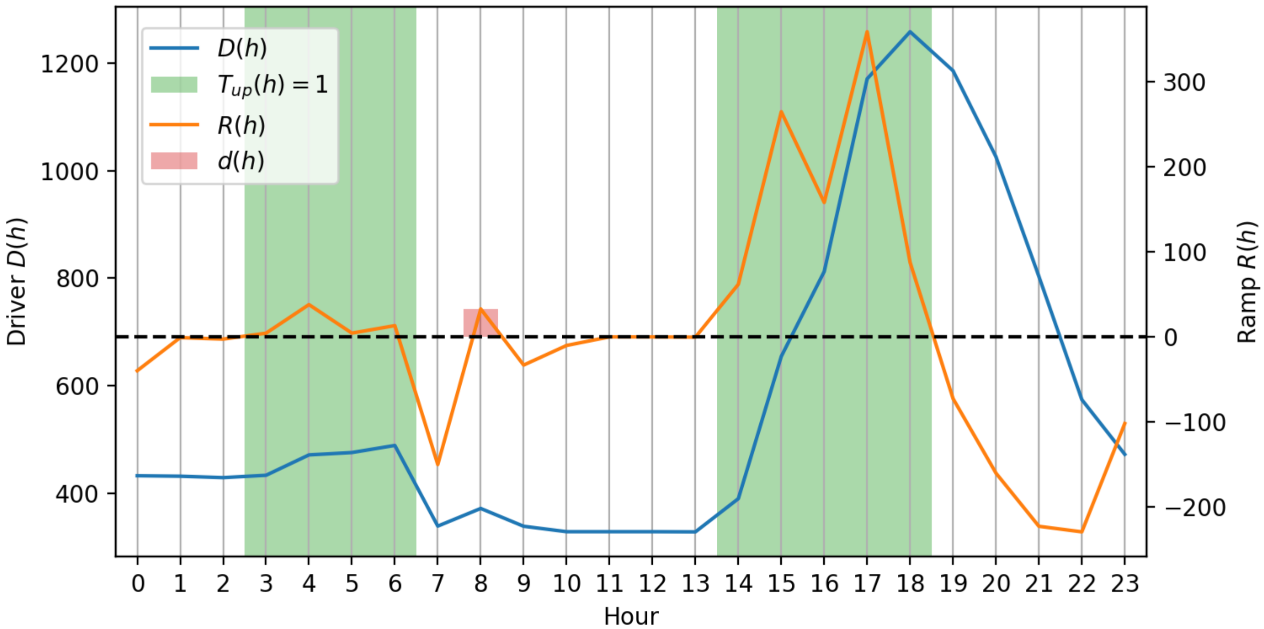

Figure 4.8: Ramping time partitioning model

Otherwise, $d(h)$ is equal to the driver ramp intensity $|R(h)|$. The maximum number of times that the release ramp can switch from the "down" to "up" direction, *i.e.*, reaches a local minimum, is equal to the target number N_{cycles}.

Upon solving the MILP problem, the subset of up-ramping hours H_{up} is given by:

$$H_{up} = \{h : T_{up}(h) = 1\}$$

An illustration of the ramping time partitioning model result is depicted in Figure 4.8.

4.3.3 Infeasibility detection model

As mentioned in section 4.2.7, under extreme hydrology conditions, the hydropower scheduling model might be infeasible. The most common source of infeasibility is a conflict between the monthly volume target and the flow rate limit. This source of infeasibility can be detected in advance, *i.e.* in the preprocessing step, by running a simple LP model. This LP model identifies the smallest and largest feasible monthly release volumes subject to the flow rate limits, ramping constraints, and cycle constraints.

$$\min | \max \sum_h Q(h) \tag{4.3.3.1}$$

$$Q^{min}(h) \leq Q(h) \leq Q^{max}(h) \tag{4.3.3.2}$$

$$R^{min}(h) \leq Q(h) - Q(h-1) \leq R^{max}(h) \tag{4.3.3.3}$$

We call $Q^{*,min}(h)$ and $Q^{*,max}(h)$ the optimal solution of the minimization and maximization problem, respectfully. If $Q_{target} < \sum_h Q^{*,min}(h)$, the problem is infeasible because the monthly volume target conflicts with the minimum release rate constraint. Conversely, if $Q_{target} >$

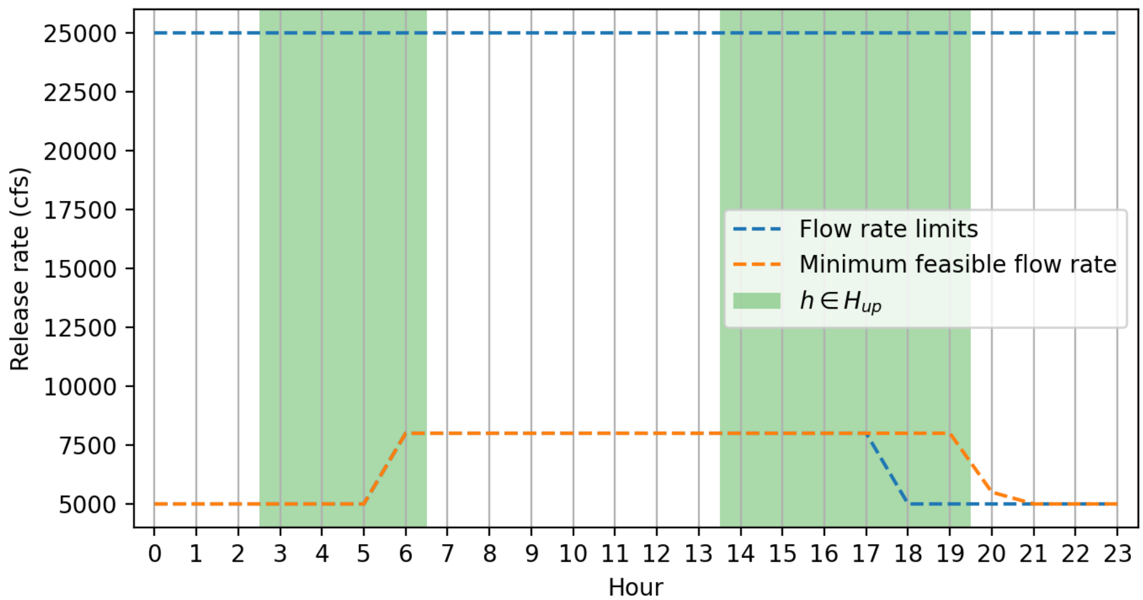

Figure 4.9: Real flow limits at GC

$\sum_h Q^{*,max}(h)$, the problem is infeasible because the monthly volume target conflicts with the maximum release rate constraint.

In future versions of the CRiSPPy model, other sources of infeasibility will be detected, such as infeasibility due to elevation constraints and fixed hourly releases.

An illustration of the real flow rate limits at GC resulting from ramping and cycling constraints is depicted in Figure 4.9.

4.3.4 Sub-minimum flow shaping model

When an infeasibility is detected, the constraints of the hydropower scheduling problem must be adapted to make the problem feasible. The infeasibility is removed by relaxing the constraints with the lowest priority, as described in table 4.2. Monthly release volumes and ramping constraints have higher priority than flow rate limits which, in turn, have higher priority than cycling constraints. As a result, model runs whose flow rate limits conflict with monthly release volumes can be made feasible by adapting these rate limits.

Maximum flow rate limits currently only apply to GC and consist of a single, constant, value Q^{max}. If the target monthly release volume Q_{target} is too large with respect to the value of the maximum flow rate Q^{max}, the maximum flow rate is lifted and a constant flow rate equal to Q_{target}/N_{hours} is imposed, where N_{hours} is the number of hours in the month.

Minimum flow rates, on the other hand, apply to GC and FG, and can vary hourly. For GC, they are typically composed of two limits: a night limit (5,000 cfs), and a larger day limit (8,000 cfs). If the infeasibility detection model identifies a conflict between the target monthly release volume and the minimum flow rate, an adjusted hourly release profile is calculated that complies with the monthly release target. The hourly release profile is built by breaching a list of conflicting constraints according to their priority order (see Table 4.2) until the target monthly release volume is met. The list of constraints, ordered from lowest to highest priority is: the ramping cycle constraint, the day limit constraint, imposed during day hours, and the night limit constraint, imposed during all hours of the day. This ordered breaching logic can be formulated as a single LP model.

$$\min \left(w_{cycle} \cdot Z_{cycle}^{max} + w_{day} \cdot Z_{day}^{max} + w_{night} \cdot Z_{night}^{max} \right) \tag{4.3.4.1}$$

$$Z_{cycle}^{max} \geq Z_{cycle}(h), \quad \forall h \in H \tag{4.3.4.2}$$

$$Z_{day}^{max} \geq Z_{day}(h), \quad \forall h \in H \tag{4.3.4.3}$$

$$Z_{night}^{max} \geq Z_{night}(h), \quad \forall h \in H \tag{4.3.4.4}$$

$$\sum_{h} Q(h) = Q_{target} \tag{4.3.4.5}$$

$$-R_{down} \leq Q(h) - Q(h-1) \leq R_{up}, \quad \forall h \in H \tag{4.3.4.6}$$

$$R^{min}(h) - Z_{cycle}(h) \leq Q(h) - Q(h-1) \leq R^{max}(h) + Z_{cycle}(h), \quad \forall h \in H \tag{4.3.4.7}$$

$$Q(h) + Z_{night}(h) \geq Q_{night}, \quad \forall h \in H \tag{4.3.4.8}$$

$$Q(h) + Z_{day}(h) \geq Q_{day}, \quad \forall h \in H_{day} \tag{4.3.4.9}$$

The value of the weights w_{cycle}, w_{day}, and w_{night} can be arbitrary set in a way that is consistent with the constraint priority order. An illustration of the sub-minimum flow shaping model results for various target average release rates is depicted in Figure 4.10.

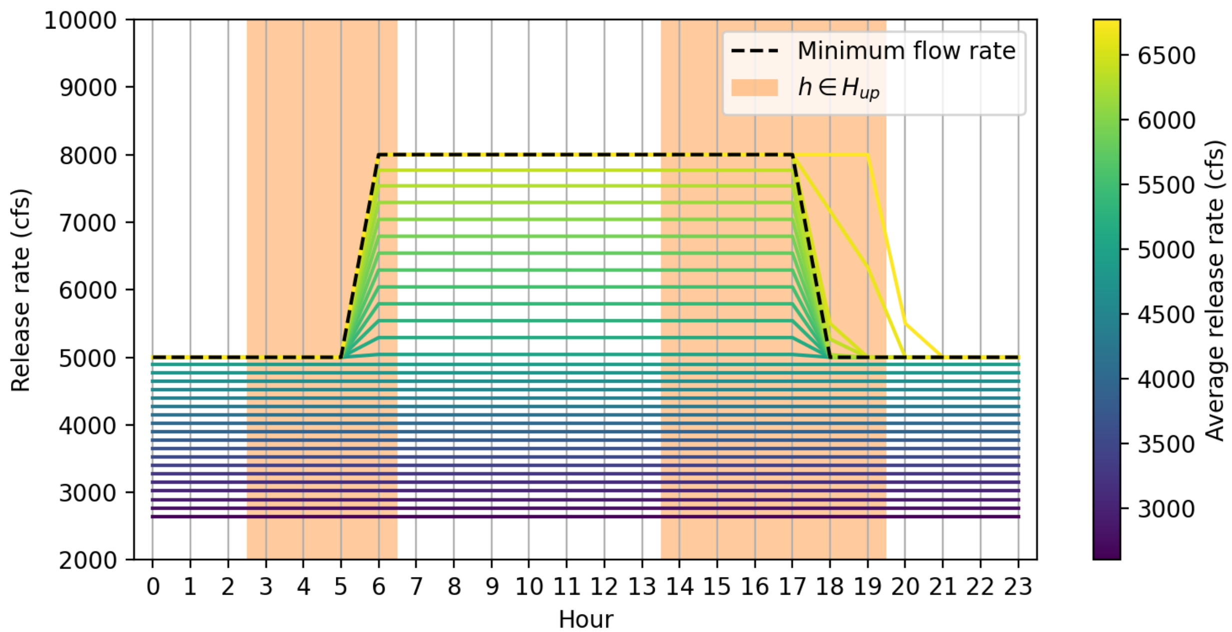

Figure 4.10: Sub-minimum flow shaping model

4.3.5 High flow shaping model

HFE are periodically scheduled at the GC dam and involve releasing a large volume of water over a relatively short period [32]. These controlled floods are designed to mimic natural flood events that occurred in the Colorado River before the GC dam was built. The primary goals of HFEs are to redistribute sand and sediment, restore and maintain habitat, and study the effects of high flows. The LTEMP ROD [19] establishes a framework to decide when and how to implement these HFEs.

A HFE typically consists of three main phases: a ramp-up, constant flow, and ramp-down phase. The ramp-up and ramp-down phases are generally following regular GC ramp limits described in Table 3.1. In some forward (*i.e.*, future-looking) analyses, HFEs are only described by a small list of features instead of a detailed release profile. These features include: the duration of the HFE $HFE_{duration}$, the magnitude of the HFE at peak release HFE_{peak}, and the total volume of water released during the HFE period HFE_{volume}.

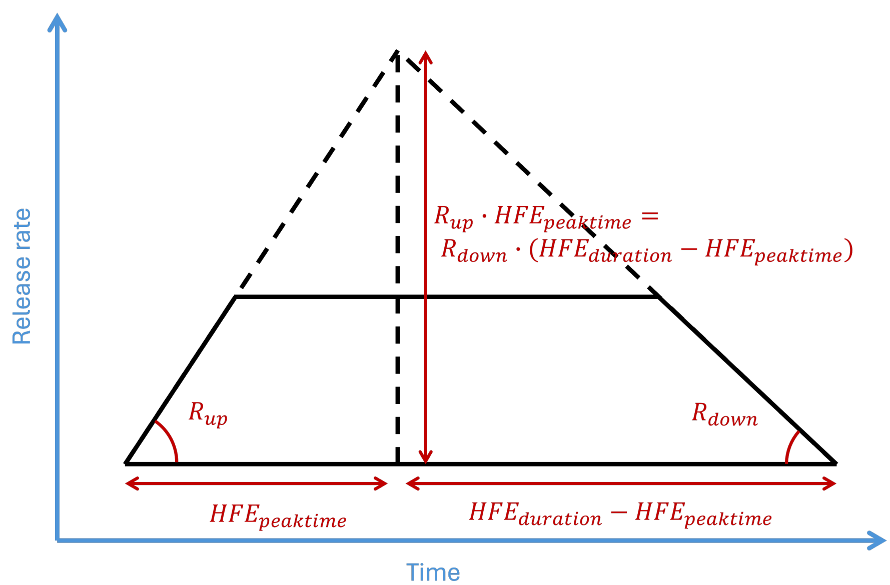

Figure 4.11: Calculation of the HFE peak time

CRiSPPy implements a high flow shaping model that allows the user to build a HFE hourly profile based on this list of features. This shaping model is executed after the hydropower scheduling model and uses water releases that occur in the hours just before and after the HFE period. To formulate this model, it is necessary to estimate the HFE peak time, *i.e.* the time when the HFE profile is expected to reach its peak release value. We assume that the water release rate before and after the HFE period is the same, that the HFE release rate ramps up and down at a constant rate R_{up} and R_{down}, and that it remains constant between the two ramping phases. Then, the HFE hourly profile is expected to have a trapezoidal shape, as illustrated in Figure 4.11. The two ramping lines of this trapezoid intersect at a point directly above the constant flow phase of the HFE release profile. The time-coordinate of that point, called $HFE_{peaktime}$, is therefore always contained in the peaking period of the HFE. Using simple

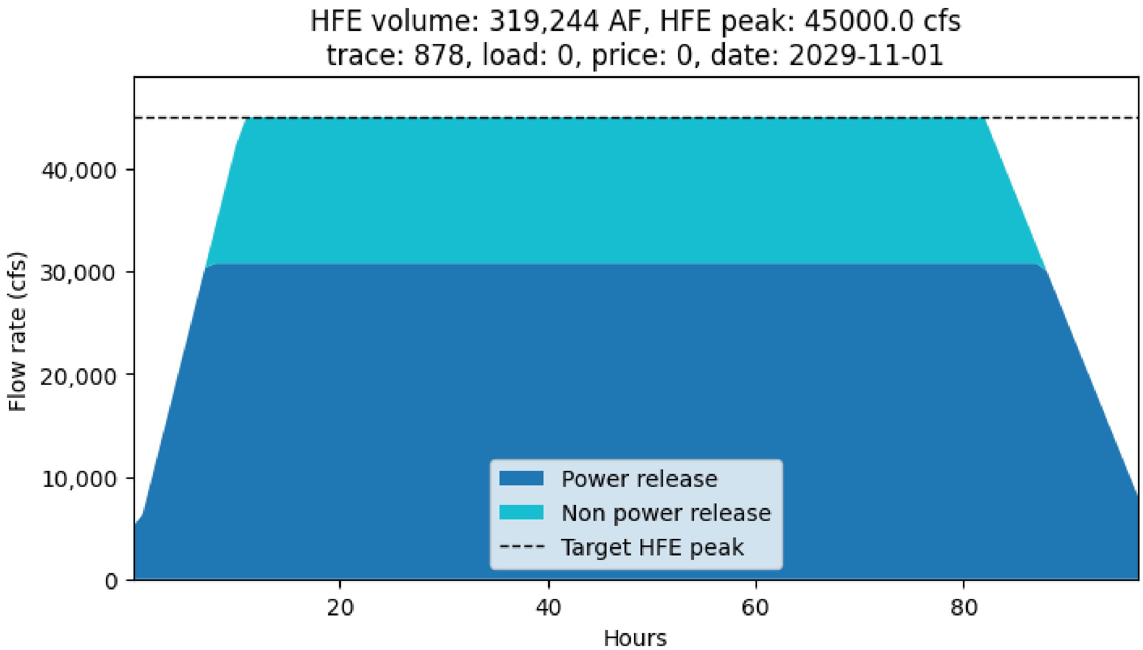

Figure 4.12: High flow shaping model

geometry (see Figure 4.11), we show that:

$$HFE_{peaktime} = \left\lfloor \frac{R_{down} \cdot HFE_{duration}}{R_{up} + R_{down}} \right\rfloor$$

The HFE hourly release profile can be built by solving the following LP problem.

$$\max \sum_{0 \le h < HFE_{duration}} (w_{power}(h) \cdot Q_{power}(h) + w_{nonpower}(h) \cdot Q_{nonpower}(h)) \tag{4.3.5.1}$$

$$Q(h) = Q_{power}(h) + Q_{nonpower}(h), \quad 0 \le h < HFE_{duration} \tag{4.3.5.2}$$

$$R^{min}(h) \le Q(h) - Q(h-1) \le R^{max}(h), \quad 1 \le h < HFE_{duration} \tag{4.3.5.3}$$

$$0 \le Q(0) - Q_{before} \le R_{up} \tag{4.3.5.4}$$

$$-R_{down} \le Q_{after} - Q(HFE_{duration} - 1) \le 0 \tag{4.3.5.5}$$

$$Q(h) \le HFE_{peak}, \quad 0 \le h < HFE_{duration} \tag{4.3.5.6}$$

$$\sum_{0 \le h < HFE_{duration}} Q(h) = HFE_{volume} \tag{4.3.5.7}$$

$$PCF \cdot Q_{power}(h) \le P^{max}, \quad 0 \le h < HFE_{duration} \tag{4.3.5.8}$$

Where $R^{min}(h)$ and $R^{max}(h)$ are such that:

$$R^{max}(h) = \begin{cases} R_{up} & \text{if } h < HFE_{peaktime} \\ 0 & \text{if } h \ge HFE_{peaktime} \end{cases}, \quad R^{min}(h) = \begin{cases} 0 & \text{if } h < HFE_{peaktime} \\ -R_{down} & \text{if } h \ge HFE_{peaktime} \end{cases}$$

Q_{before} and Q_{after} are the release rate just before and after the HFE period. $w_{power}(h)$ and $w_{nonpower}(h)$ are arbitrary weights such that $w_{power}(h)$ and $w_{nonpower}(h)$ strictly increase before $HFE_{peaktime}$ and strictly decrease after. In addition,

$$w_{power}(h) > w_{nonpower}(h') > 0, \quad \forall h, h'$$

This formulation ensures that power releases have priority over non-power releases. It also ensures that the HFE reaches and remains at its peak target value for a maximum number of hours. An illustration of high flow shaping model results for a typical HFE is depicted in Figure 4.12.

5 Data management and processing

5.1 Input data

A CRiSPPy case study typically requires the following input data: hydrology data, load data, energy price data, power system rules, reservoir and plant rules. In addition, input data can include special rules and user-defined profiles but these data types are optional. Large inputs such as hydrology and load data are typically imported into CRiSPPy via import modules and can be collected from various data sources (Excel files, online databases). Conversely, smaller inputs such as power system and reservoir rules are directly entered into GUI entry fields. The data flow diagram of the data import and preprocessing phase is depicted in Figure 5.1.

5.1.1 Hydrology data

Hydrology data typically consist of monthly release volumes and end-of-month forebay elevations at each reservoir. For case studies that include HFEs, they can also include essential HFE features such as release volumes, durations, and peak flow rates. Hydrology data are the core input of the CRiSPPy model. They not only define monthly release targets for each reservoir, but also indirectly provide critical information such as: storage volumes, power plant capacities, PCFs, unregulated inflows, and maximum daily flow changes.

Hydrology data can be imported from various sources:

- the online USBR hydrologic database [9], either for historical or forecasted data,
- output from the USBR CRSS model,
- output from the USBR CRMMS model, including the 24MS [17].

5.1.2 Load data

Load data consist of project use load and hourly FES customer electricity demand. Customer loads are the main driver of the CRSP reservoirs release schedule. To simulate this, the CRiSPPy model aims to minimize purchases and sales made to compensate the difference between CRSP resources and customer loads, as seen in section 4.2.6.

Hourly FES customer demand can be imported from:

- the “Average Load” file, which represents an estimate of typical future loads for the 12 months of the year,
- EMMO “Load and Resource” files, which are daily reports of hourly power system operations.

5.1.3 Energy price data

Energy price data can be imported at a hourly or monthly time resolution. If imported at a monthly resolution, the price timeseries is upsampled to a hourly resolution using the hourly price generation model described in section 4.3.1. Energy price data represent the price at which electricity is exchanged on the electricity market and apply to the purchases and sales described in section 4.2.

Energy price data can be imported from:

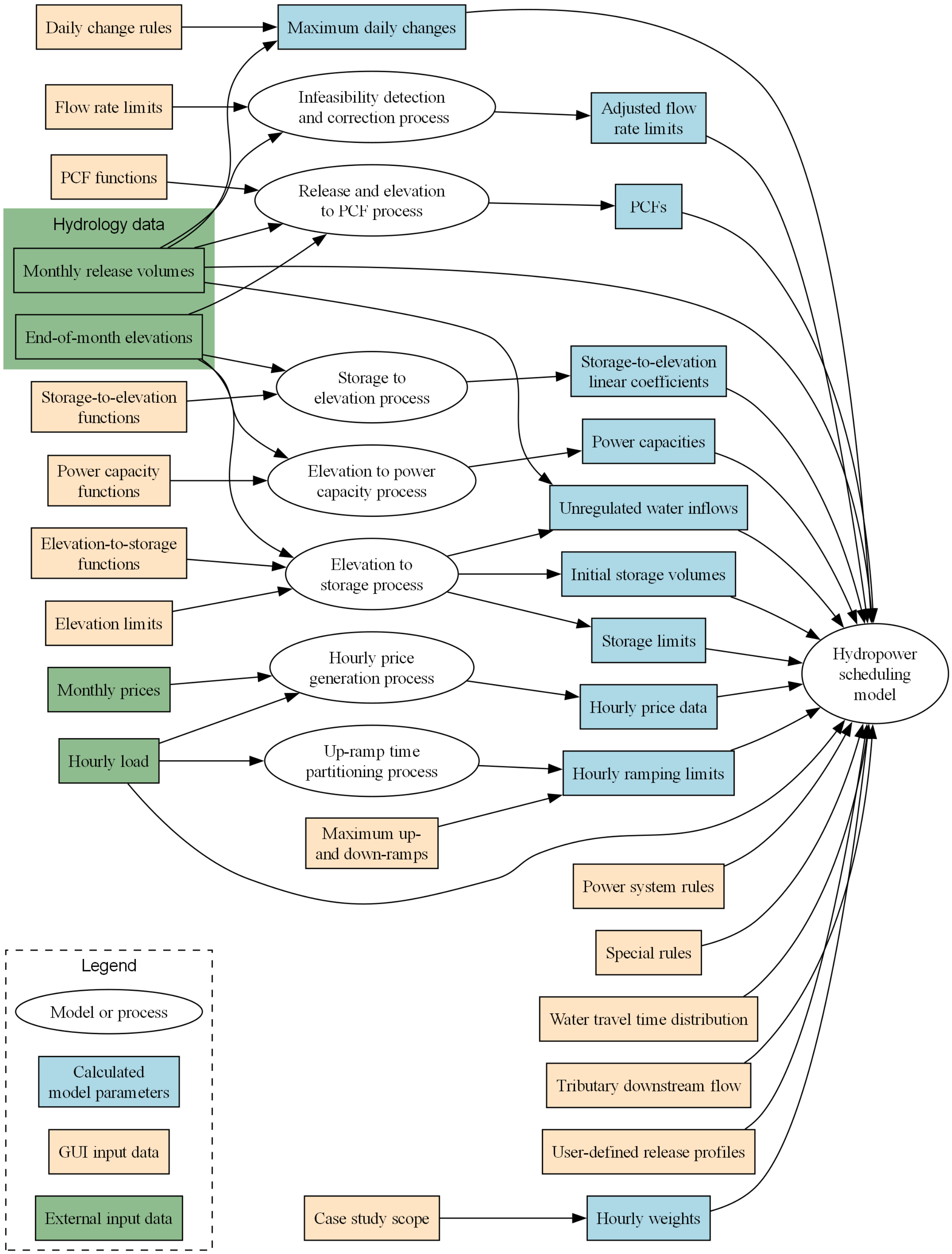

Figure 5.1: Data flow of the data import and preprocessing phase

- the “Forward curves” file, which represents an estimate of future monthly-average “on-peak” and “off-peak” market prices,

- EMMO “Daily Deals” report files, which record historical day-ahead and real-time energy transactions made by WAPA.

5.1.4 Power system rules

The power system rules are the power system physical rules, market operation guidelines, and ancillary service requirements described in section 4.2. They consist of a small set of parameters which is provided by the user via the GUI.

5.1.5 Reservoir and plant rules

Reservoir and plant rules are the physical and operating rules that govern each reservoir and plant. Reservoir operating constraints are described in details in section 3 and can be represented by a small set of parameters which is provided by the user via the GUI. Similarly, physical rules can be represented by a small set of parameters describing the physical relationships between reservoir/plant attributes (*e.g.*, storage-to-elevation curve). These parameters can be either manually updated by the user, or automatically calculated via ML modules accessible via the GUI (see section 4.1). The data flow diagram of the ML process involved in the reservoir physical rules is depicted in Figure 5.2.

5.1.6 Special rules

The CRiSPPy software integrates a “Special events” module which allows the user to define special rules, *i.e.*, few instances where a power system or reservoir rule may be different from the general case. For example, the user may want to impose a minimum flow rate of 1,000 cfs at FG for a single month, while retaining the general 800 cfs rule for the rest of the study period. Special rules are defined at specific plants, for specific time periods, and specific trace, price, and load scenarios.

In the current version of CRiSPPy, special rules can be applied to:

- daily release volume factors,

- flow rate limits,

- maximum up-ramp and down-ramp rates,

- up-ramping times (*i.e.*, cycling patterns),

- maximum daily flow changes,

- transmission losses.

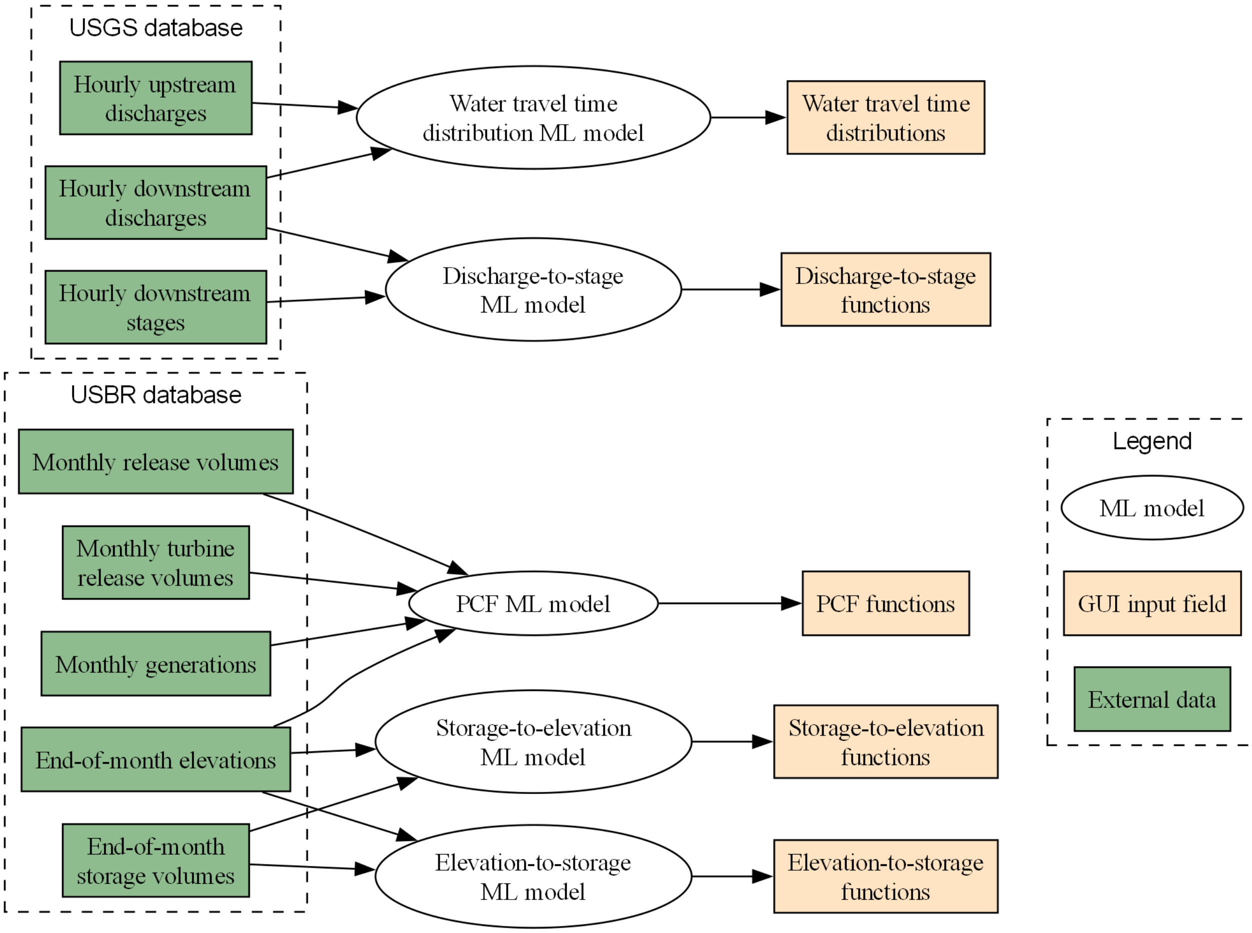

Figure 5.2: Data flow of the ML phase

5.1.7 User-defined release profiles

The CRiSPPy software allows the user to specify hourly power and non-power release profiles at specific time periods. This feature is typically used to model experimental release events such as HFEs or Smallmouth bass (SMB) bypass flows. The user can specify these hourly releases via a temporary tabular file generated by CRiSPPy. The user can choose to edit historical hourly releases automatically downloaded from the USBR hydrologic database or to manually enter hourly values. To ensure model feasibility, any operating constraint that conflicts with user-specified hourly releases is relaxed.

5.2 Data preprocessing

The main goal of the data preprocessing phase is to prepare the parameter values for the hydropower scheduling model described in section 4.2. First, the input data imported by the user and entered via the CRiSPPy interface are converted into model parameters. Second, infeasible model runs are detected in advance with the specific reservoirs/plants and constraints causing the infeasibility. Third, for these model runs, the conflicting constraints with lower priority are relaxed to resolve the infeasibility. The data flow diagram of the preprocessing phase is depicted in Figure 5.1.

5.2.1 Calculation of hydropower scheduling model parameters

The first step of the data preprocessing phase is to convert input data into model parameters. If the case study uses the "representative week" time representation, hourly weights $w(h)$ are calculated for each of the 168 hours of the week. If the case study includes experimental releases that are calculated in the post-processing step (see section 5.4.1), the time period of the experiment is removed from the model run and the monthly release volume target is adjusted accordingly.

The system total load $L(h)$ is calculated as the sum of FES customer load and project use load. If price data are provided with a monthly resolution, the hourly price generation model described in section 4.3.1 is used to generate the hourly price profile $\pi(h)$ based on the total load. Optimal up-ramping time for all reservoirs is calculated using the model described in section 4.3.2 and the load profile as hourly driver. Hourly release ramp limits are defined using the optimal up-ramping time and the up-ramp and down-ramp rate values. Maximum daily changes $DC(p)$ are calculated based on monthly release targets and operating rules described in Table 3.1.

Initial storage volumes $S_{init}(p)$ and storage volume limits $S^{min}(p)$, $S^{max}(p)$, are calculated based on forebay elevations, forebay elevation limits, and the elevation-to-storage functions described in section 4.1.1. Storage volumes are used to back-calculate unregulated water inflows $I(p, h)$ using water budget equations. Linear storage-to-elevation coefficients $StoE(p)$ are calculated based on storage-to-elevation functions and monthly hydrology conditions. Similarly, linear discharge-to-stage coefficients $QtoE(p)$ are calculated based on discharge-to-stage functions and monthly discharge values.

$PCF(p)$ values are calculated based on monthly hydrology conditions and the learned CPWL functions described in section 4.1.2. Plant power capacities $P^{max}(p)$ are calculated based on average monthly forebay elevations and the 1D CPWL functions described in section 4.1.5.

5.2.2 Infeasibility detection and constraint adjustments

The second step of the data preprocessing phase is to detect infeasible model runs before running the hydropower scheduling model. In the current version of CRiSPPy (version 2.3), the infeasibility detection model described in section 4.3.3 can detect flow rate limits that are in conflict with monthly release targets. In future versions, the infeasibility detection logic will be extended to other sources of infeasibility, such as infeasibility due to elevation constraints and fixed hourly releases.

The third step of the data preprocessing phase is to relax the conflicting constraints with lower priority to resolve the infeasibility. In the current version of CRiSPPy, the cycling constraint and flow rate limits are sequentially lifted to meet the monthly release targets. This is done by solving the model described in section 4.3.4. Specifically, the model identifies the optimal release schedule that meets both monthly release targets and ramp constraints. This optimal release schedule is then used as a fixed release pattern in the hydropower scheduling model.

5.3 Scheduling model output

The main outputs of the hydropower scheduling model are the hourly power and non-power releases $Q_{power}(p,h)$ and $Q_{nonpower}(p,h)$ of each reservoir p. In short-term future-looking studies, *e.g.*, 24MS, these output data guide WAPA in their water release scheduling decisions. Power generation results $P(p,h)$ inform WAPA about CRSP energy and capacity resources and are used to estimate WAPA's financial position.

Additional outputs relevant to WAPA staff typically include hourly forebay elevations, Jensen gage's hourly flow and stage, and day-ahead and real-time transactions. A few additional steps are required to calculate output in a useful form. If the case study uses the "representative week" time representation, model outputs are extended to all the hours of the study period. Hourly storage volumes $S(p,h)$ are calculated based on the cumulative change in storage volume $\Delta S(p,h)$. Hourly forebay elevations $E(p,h)$ are calculated by applying the storage-to-elevation functions to to $S(p,h)$. The data flow diagram of the output data and post-processing phase is depicted in Figure 5.3.

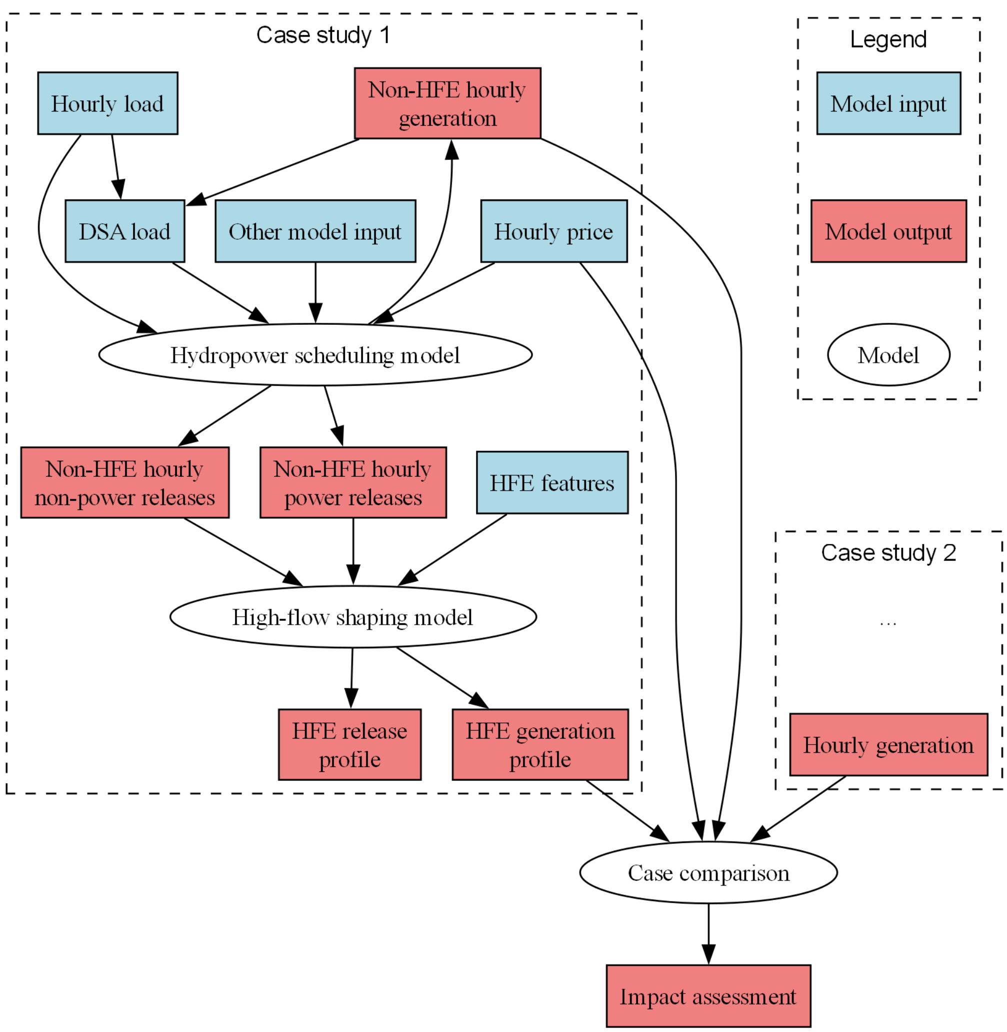

Figure 5.3: Data flow of the post-processing phase

5.4 Post processing step

Some post processing steps may be required depending on the case study. These post processing steps involve calculations that are not directly modeled into the hydropower scheduling model but depend on the hydropower scheduling model results. The data flow diagram of the post-processing phase is depicted in Figure 5.3.

5.4.1 Experimental releases

As seen in section 4.3.5, some case studies involve the modeling of future experimental releases, such as HFEs, that are calculated based on a small set of features. These experimental releases are typically calculated in the post-processing steps, since they are based on water release conditions occurring before and after the event and calculated by the hydropower scheduling model.

5.4.2 DSA load

As seen in section 3, customer loads are based on DSA, which is the amount of hydropower capacity and energy based on the forecasted SLCA/IP generation. In short-term future-looking studies, the customer load profile is not initially known. It is estimated by running the hydropower scheduling model a first time. The model is then run a second time, as a post-processing step, using the adjusted customer load profile.

As a first step, the hydropower scheduling model is solved using a typical future load. Then, modeled power generation results are aggregated into a single hourly generation profile. Accounting for transmission losses, this hourly profile informs about CRSP available energy and capacity resources. The DSA load profile is calculated by scaling the typical future load using modeled CRSP energy and capacity. Finally, the estimated DSA load profile is used as the customer load profile in a second run iteration of the hydropower scheduling model.

5.4.3 Case comparison

Several CRiSPPy analyses consist of analyzing the impact of implementing certain operating rules or experimental releases. Impacts are typically assessed by conducting a counterfactual analysis, *i.e.*, by comparing two scenarios: one scenario in which these rules or experiments are implemented, and another scenario where they are not. Impacts are typically quantified by calculating the net difference in economic or financial value between the two scenarios.

In CRiSPPy, this counterfactual analysis is performed in the post-processing step. Each scenario is computed by the user as an independent case study. Then, the user can access the counterfactual analysis module where they select the "Base case" scenario and the counterfactual scenarios. Financial or economic impacts are then calculated for each counterfactual case and can be aggregated on a daily, monthly, or yearly basis.

5.5 Data storage and management

Data are stored in various ways depending on the processing phase. Outside the CRiSPPy software, input data are typically collected from online database and offline spreadsheets, as seen in section 5.1, and model results can be exported as comma-separated values (CSV) or

Excel files. Small-size parameters, such as reservoir or power system rules, are typically entered via the software GUI.

Within the software's Python environment, the data are stored as a dictionary of Pandas MultiIindex Series [33]. A Python dictionary is a data structure that maps keys to values and is implemented as a hash table which allows for constant-time average lookup. In CRiSPPy, keys are strings (*e.g.*, "Load") and their corresponding values are MultiIndex Series (*e.g.*, the hourly load profile). Pandas MultiIindex Series are one-dimensional labeled arrays whose index is made up of multiple levels, allowing hierarchical indexing. This data structure is used for most of the preprocessing and post-processing phases and enables organizing and accessing data along multiple dimensions. This data structure was preferred over relational databases owing to its tight integration with Python ecosystem and its high flexibility, allowing for easy and rapid software updates without needing to normalize the data. Items of the dictionary include hydrology data (indexed by date, trace, plant, and datatype), hourly price (indexed by datetime and price scenario), etc. A summary of the index levels used for the MultiIindex Series is provided in Table 5.1. Note that not all levels are used for each type of data, as seen earlier for hydrology and price data.

Table 5.1: Index levels of MultiIndex Series in the CRiSPPy Python environment

Index level	Type	Description	Examples
Trace	*Integer*	*Index of the hydrology trace*	*0, 1, 2, ...*
Load	*Integer*	*Index of the load scenario*	*0, 1, 2, ...*
Price	*Integer*	*Index of the load scenario*	*0, 1, 2, ...*
Date[a]	*Date*	*Year and month*	*2025-04-01, 2025-05-01, ...*
Hour[a]	*Integer*	*Hour of the month, or hour of the representative week*	*0, 1, 2, ..., 167, ..., 744*
Datetime[a]	*Datetime*	*Date and time*	*2025-04-01 00:00:00, 2025-04-01 01:00:00, 2025-04-01 02:00:00, ...*
Plant	*String*	*Plant associated with the data*	*'GC', 'FG', 'BM', ...*
Datatype	*String*	*Physical quantity the data represents*	*'Power release (AF/hr)', 'Generation (MW)', ...*

[a] The date level (year and month) labels the time period for monthly data, while the datetime level is used to timestamp hourly data. In the "typical week" time representation, hourly data are indexed using the date level and an hour level ranging from 0 to 167.

The hydropower scheduling model parameters are organized using a reference-based data structure to improve memory efficiency and enable faster queries. This approach takes advantage of the fact that many model runs share identical parameter values across certain dimensions (*e.g.*, traces or time). Model parameters are referenced via a Pandas DataFrame called 'Run_parameter_ID'. This DataFrame is indexed by model run dimensions (*i.e.* trace, price, load, and date) and reservoir, and each column represents a specific model parameter (*e.g.*, monthly release volume, hourly weights, hourly load). Rather than storing the full parameter values, each cell contains an ID that refers to the row index of a separate lookup table (NumPy array) containing the unique profiles for that parameter. For example, if the cell in row 0 of the 'Hourly weight' column contains the value 4, this means that the hourly weight profile for the model run index 0 is stored at row index 4 in the 'Hourly weight' array. When running the hydropower scheduling model, each model run can be rapidly updated by querying the parameter IDs from the 'Run_parameter_ID' table and by querying the model parameter values from their corresponding lookup tables.

At any stage of the case study workflow—whether during data import, preprocessing, model execution, or post-processing—the user can save the case study to the storage disc. This is achieved using Python's built-in pickle module, which serializes the dictionary representing the case study into a byte stream and writes it to a file named 'database.pkl'. This file preserves the complete state of the case study, enabling seamless reloading into CRiSPPy at a later time. The 'database.pkl' file is stored within a folder named after the case study, as specified by the user. Upon launching the CRiSPPy software, the user can reopen an existing case study simply by selecting the corresponding folder.

6 Key improvements over legacy software

The CRiSPPy software demonstrates a performance improvement which, in some cases, is more than two orders of magnitude compared to the previous GTMax SL software. This significant improvement is due to several optimized process designs described below.

6.1 Streamline data import and physical rule updates

A key improvement of the CRiSPPy model lies in its substantially reduced reliance on manual data entry and its streamlined data import processes. As described in section 5.1, a significant part of input data can be automatically retrieved from online databases using built-in import modules, eliminating the need for user-driven file preparation. For input data that reside in offline spreadsheets, efficient extraction is performed using Rust-based routines interfaced through the python-calamine library [34], enabling fast and reliable access. As an example, in a specific case study, load data from 150 1-megabyte (MB) Excel files were imported in a few seconds. Additionally, physical rules governing the reservoirs, power plants, and river systems can be automatically updated using embedded ML modules. These models only require the user to specify the historical time range of the training data and a few learning parameters, thereby minimizing configuration effort.

6.2 In-memory data workflow

The GTMax SL software frequently reads from and writes to storage disc throughout its workflow. Conversely, unless explicitly saved by the user, CRiSPPy maintains the full case study in Random Access Memory (RAM) throughout the workflow—from data import and preprocessing to model execution and result analysis. This design is enabled by Python's rich in-memory data structures (*e.g.*, dictionaries, DataFrames, arrays).

Below are key advantages of keeping the case study in RAM:

- Performance gains: In-memory operations eliminate slow disk Input/Output (I/O) bottlenecks, enabling faster data access, preprocessing, and model evaluation—especially when handling large batches of runs.
- Real-time interaction: The user can interactively update parameters, re-run models, and explore results without waiting for repeated file access operations.
- Simplified data management: Python data structures (*e.g.*, dictionaries, NumPy arrays, Pandas DataFrames) can be passed directly between modules, avoiding serialization overhead until explicitly saved.
- Efficient parallelization: Multiprocessing benefits from fast shared data access (or copies) in memory, supporting the parallel LP warm-start method described later in section 6.4.

6.3 Single preprocessing batch and efficient access to model run parameters

Instead of generating model parameters one run at a time—as was done in the GTMax SL software—all parameters are now generated in the preprocessing phase (see section 5.2) in a single batch using NumPy vectorization. This approach leverages NumPy's underlying C-based operations to perform bulk array computations without explicit Python loops. As a result, the preprocessing phase is significantly more efficient, both in terms of execution speed and memory access patterns. By structuring the data as multi-dimensional arrays and applying operations across entire slices at once, the new method reduces overhead and enables scalable preparation of model inputs, even when handling thousands of model runs.

In addition, once computed during the preprocessing step, model parameters are stored in lookup tables implemented as NumPy arrays. These arrays are referenced by a central Pandas DataFrame indexed by model runs, as detailed in section 5.5. This data structure provides high memory efficiency and enables fast access to model parameters, thereby further reducing the software's overall computation time.

6.4 Parallelized LP warm-start

CRiSPPy employs a parallelized LP warm-start strategy to significantly reduce the solution time of large batches of model runs. More specifically, CRiSPPy leverages multiprocessing by distributing groups of model runs across available CPU cores. Within each core, runs are solved sequentially, with the solver applying LP warm-start techniques [35] to accelerate the solution of subsequent runs. This warm-start capability is enabled by the GLOP solver [10], which allows reusing solution information from previously solved runs. The combined use of multiprocessing and warm-starting substantially reduces overall computation time. This parallelized warm-start approach is particularly effective when runs assigned to the same core share similar parameter values and are pre-sorted accordingly, thereby increasing the likelihood that the solution from one run provides a useful starting point for the next. The parallelized LP warm-start strategy is depicted in Figure 6.1.

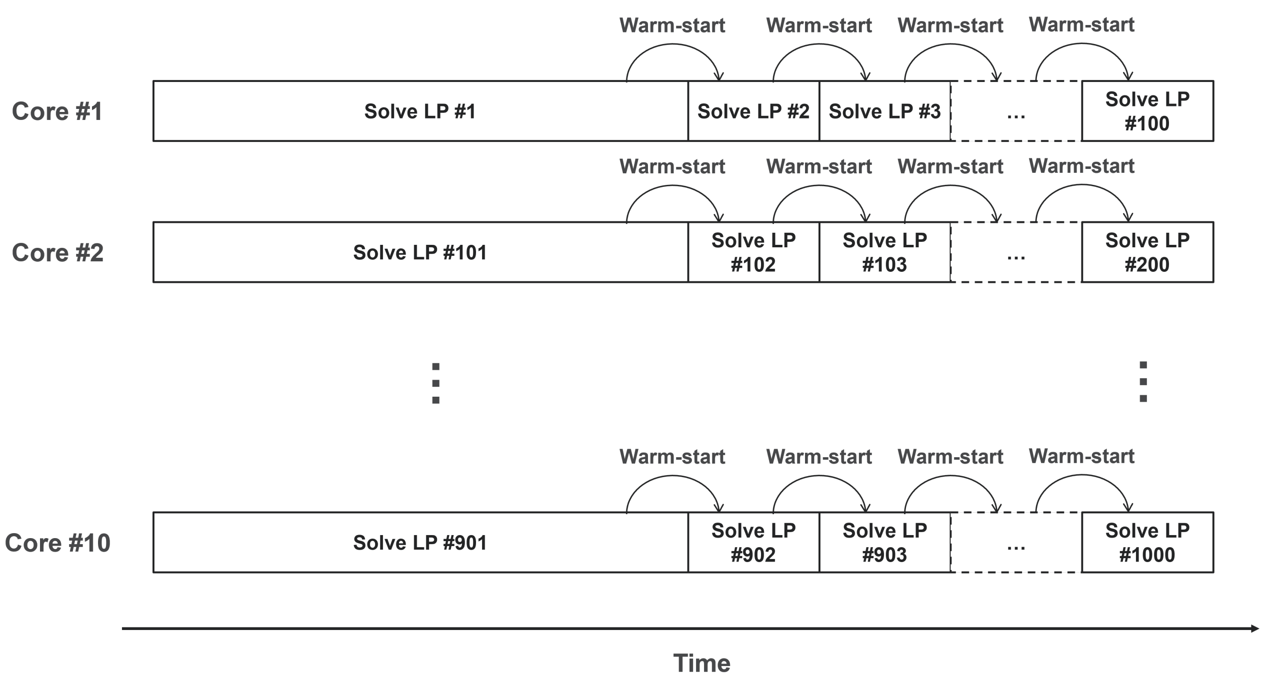

Figure 6.1: Parallelized LP warm-start

6.5 Infeasibility detection and correction

An additional key enhancement in the CRiSPPy model is its ability to proactively identify infeasible model runs (see section 5.2.2) and adjust the associated model parameters to restore feasibility. This feature offers two main advantages. First, early detection prevents infeasible model runs from interrupting the LP warm-start process (section 6.4), thereby preserving the efficiency of the parallelized warm-start strategy. Second, the corrected model runs tend to be less constrained and are generally solved more quickly than runs that were initially feasible. In some cases, the applied corrections yield a model run whose solution is fully determined by the adjusted constraints, making further evaluation by the hydropower scheduling model unnecessary. Such model runs can be excluded from the solution batch, thereby reducing overall computation time.

7 User interface

This section details the various software interfaces the user can interact with covering the entire workflow, from case creation and data processing to report generation.

7.1 Creating a new case

Upon lauching the CRiSPPy software, the user can create a new case study by clicking on the "Case study" menu, followed by the "Create new case" option. This opens a new window that is depicted in Figure 7.1.

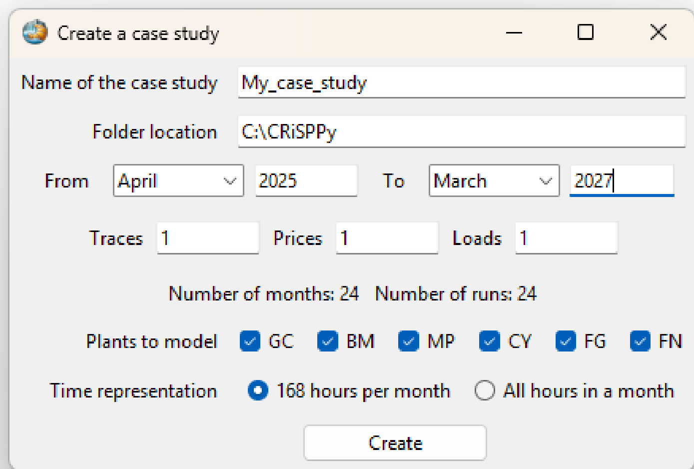

Figure 7.1: Creating a new case study

In this window, the user can specify a name for the case study and modify the folder location if desired. They specify the time period, the plants to be modeled, as well as the number of trace, price, and load scenarios. The user can choose between two time representations: 168 hours per month, or all hours in a month. As described in section 4.2.1, the former time representation is a model simplification intended to speed solution times for large case studies. Upon clicking on the "Create" button, a new folder named after the case study is generated at the specified location and contains a "database.pkl" file that stores the study scope information.

7.2 Case info

The first tab of the main CRiSPPy interface is the "Case info" tab. This tab provides general information about the case study: folder location, study scope (time period and number of trace, price and load scenarios), number of model runs, modeled plants, time representation, and number of model runs already computed (Figure 7.2). The case study location is also displayed at the top of the main window, together with the "Saved" status. This status informs the user whether the case study is currently stored into the "database.pkl" on the storage disc ("Saved") or only in the RAM ("Not saved"). Upon closing the software, information stored in the RAM is erased, and case study data can only be retrieved from the Pickle file stored on the storage disc. The user can save the case study at any time by clicking on the "Save" button. This action is also available via the "Case study" menu under the "Save current case" option.

The user can open an existing case study via the "Case study" menu under the "Open existing case" option. This action prompts the user to select a folder corresponding to the desired case study. Upon selection, the case study is automatically loaded in-memory into the CRiSPPy environment.

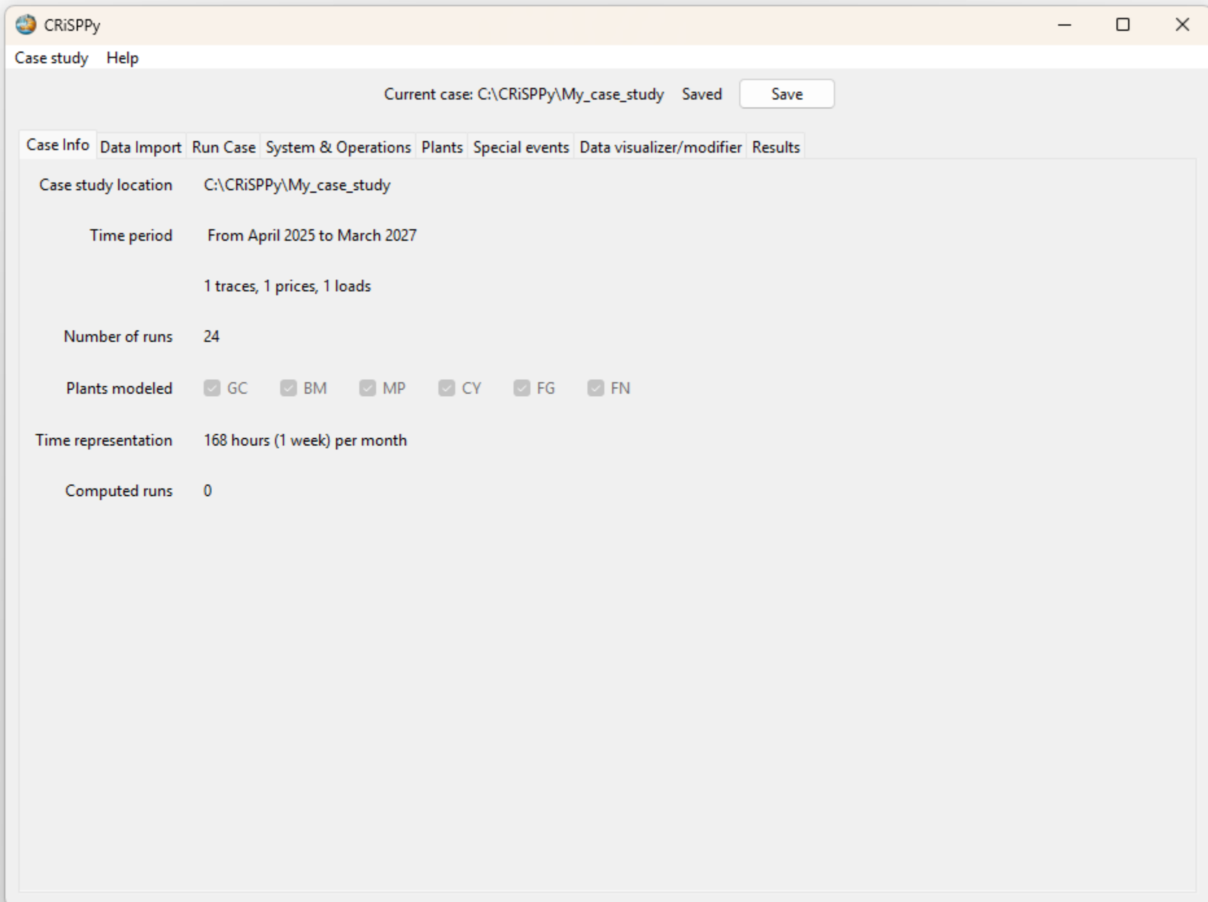

Figure 7.2: Case study information

7.3 Data import

The "Data import" tab is designed to import large external input data, as opposed to smaller-size data that can be directly entered into the GUI (Figure 7.3). The "Import folder" is the default folder location for offline data sources, such as the "Average Load" file and the EMMO "Load and Resource" files (see section 5.1.2). By default, the import period is identical to the study period, and imported data apply to all trace, price, and load scenarios, unless specified in the input data. In this module, the user can specify the data source ("Import option") of each type of input data. More details about the data sources is provided in section 5.1. The user can either import data one type at a time using the individual "Import" buttons, or perform a "Bulk Import" after selecting multiple data types.

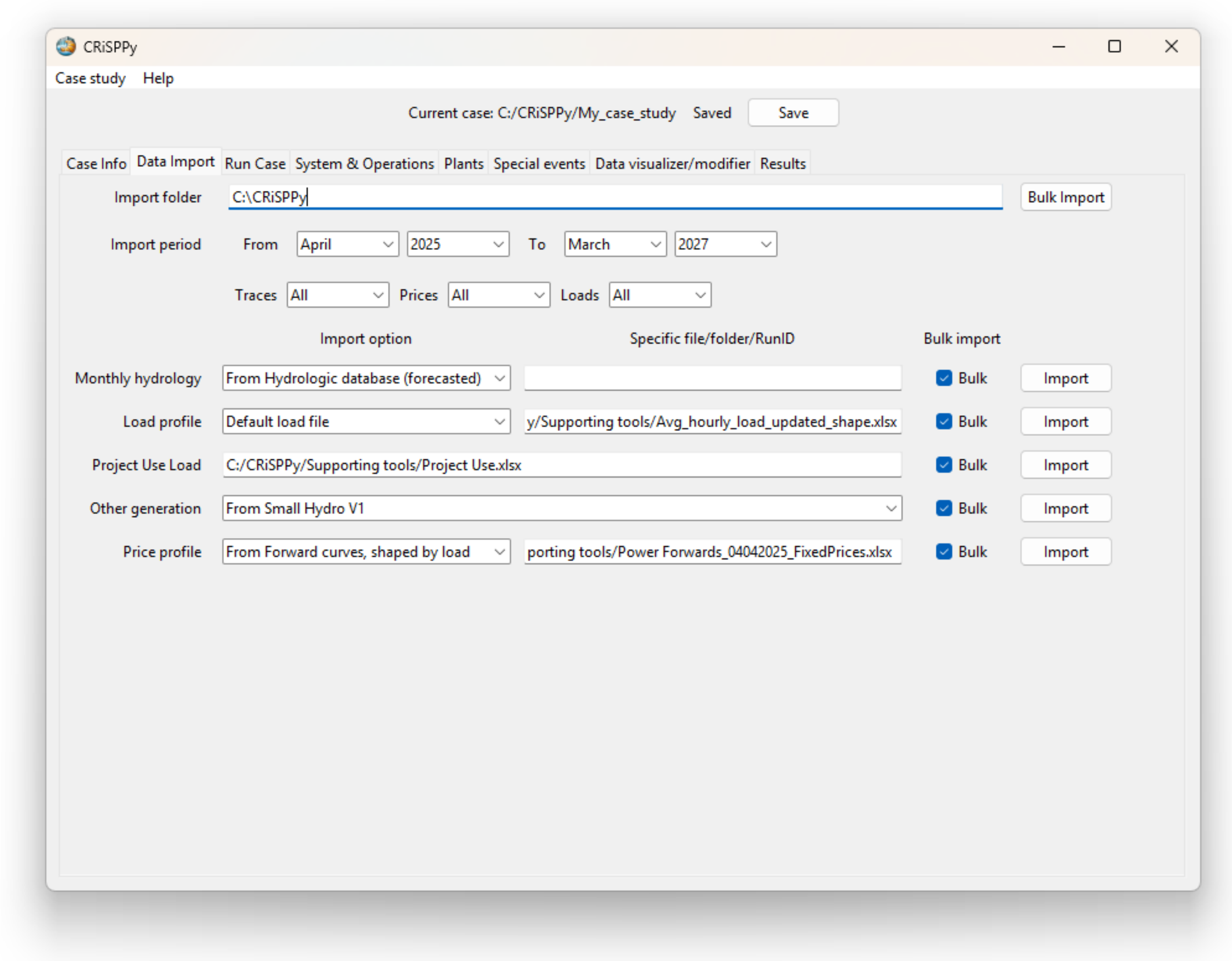

Figure 7.3: Data import

7.4 System & Operations

The "System & Operations" tab is the interface through which the user defines the physical rules, market operation guidelines, and ancillary service requirements governing the power system. Unlike the "Data Import" tab, these inputs consist of a small set of numerical values that can be entered directly via the GUI (Figure 7.4). The parameters representing these power system rules are described in more details in section 4.2.

7.5 Plants

The "Plants" tab is the interface through which the user defines the reservoir and hydropower plant rules. The "Plants" tab itself is divided into multiple sub-tabs, each representing a specific reservoir of the system. Similar to the "System & Operations" tab, reservoir and plant rules

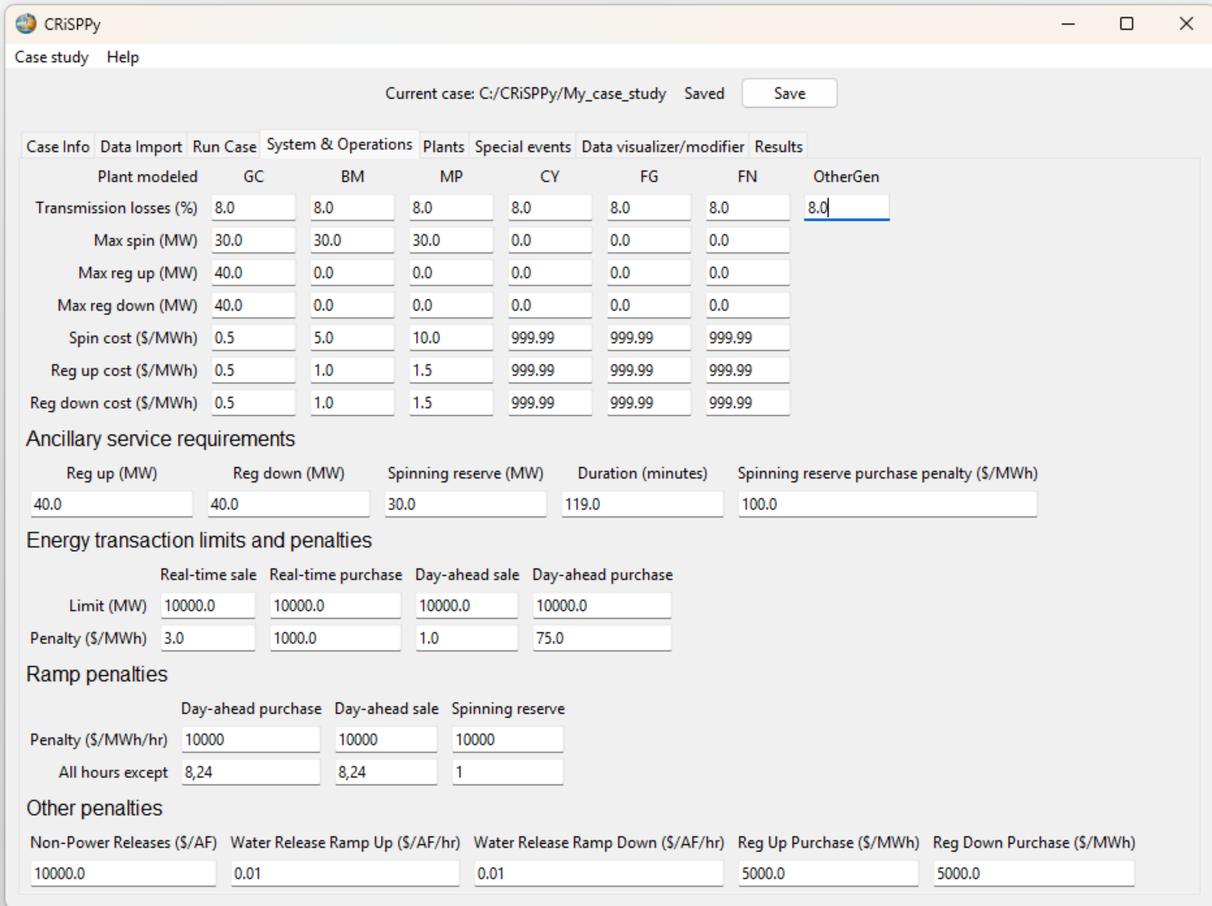

Figure 7.4: Power system rules

consist of a small set of numerical values that are directly entered via the GUI (Figure 7.5). As described in section 5.1.5, reservoir and plant rules can be categorized into operating rules and physical rules. Reservoir operating constraints can be entered in the top part of each sub-tab and are described in details in section 3. Physical rules are stored in the lower part of the sub-tabs and consist of the PCF 2D functions coefficients (for both polynomial and DC representation), the storage-to-elevation and elevation-to-storage polynomial coefficients, and the elevation-to-capacity table (CPWL representation). These coefficients can be updated either manually by the user or automatically using the ML models described in section 4.1. To use the ML models, the user can simply click one of the "Update" buttons within the corresponding sub-tab.

7.5.1 Updating the storage-to-elevation coefficients

Clicking on the "Update" button next to the storage-to-elevation coefficients opens a new window, shown in Figure 7.6. This window serves as the GUI for the ML module described in section 4.1.1. At the top of the interface, the user specifies the historical time period to be used for training the model. Upon clicking the "Import" button, monthly data are retrieved from the USBR database and displayed as a scatter plot within the window. At the bottom of the inter-

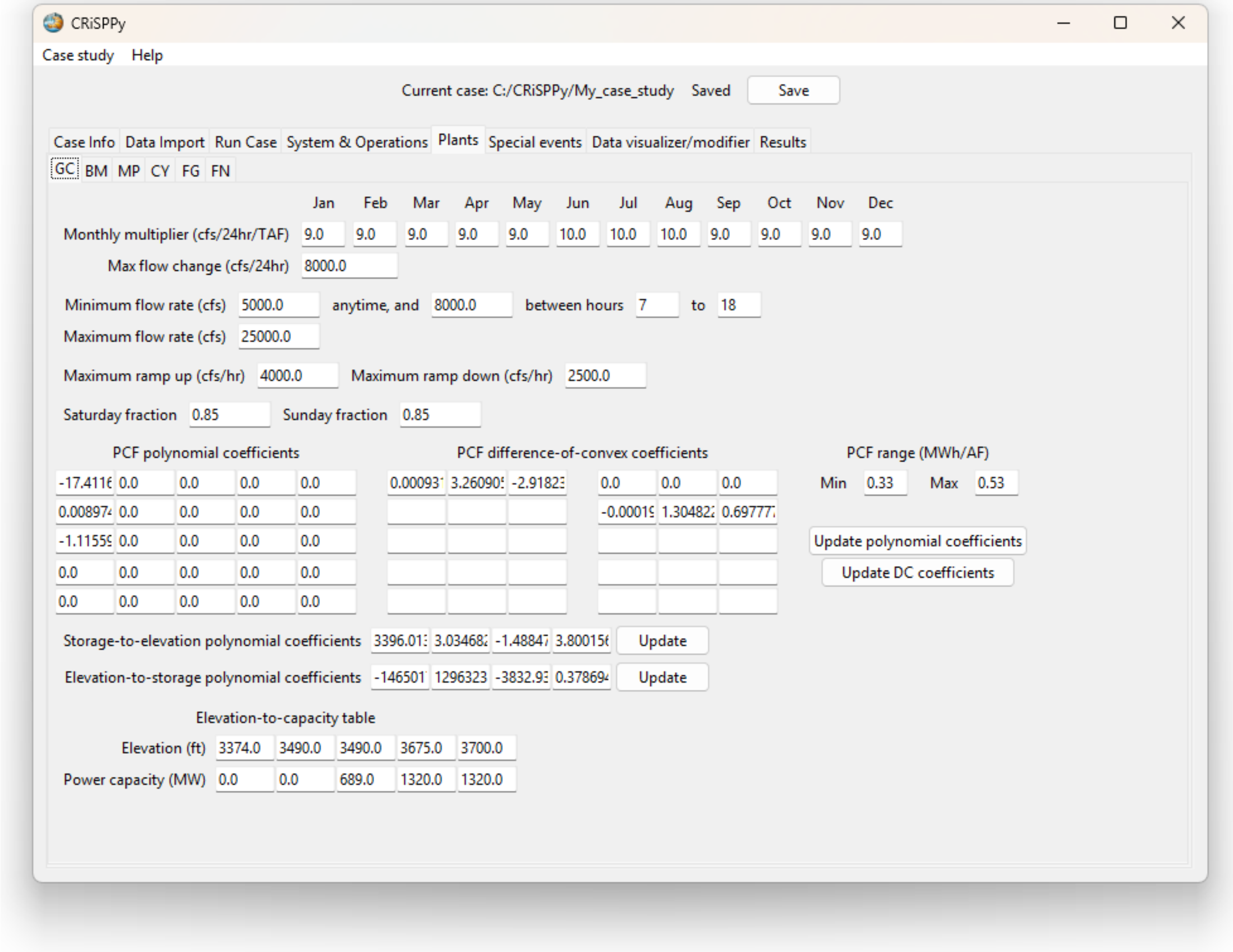

Figure 7.5: Reservoir and plant rules

face, the user specifies the degree of the polynomial used to represent the storage-to-elevation relationship. Below the scatter plot, a date-range slider is available to filter out undesired data points. For example, as shown in 7.6, the storage-to-elevation relationship at GC differs before and after June 2022—a shift attributed to storage capacity loss from sedimentation, as documented in a recent USGS report [36]. The fitted curve updates in real time as the user adjusts the date range and polynomial degree. The graph title simultaneously displays fitting quality metrics, *i.e.*, average and maximum error, updated dynamically with each change. Once satisfied with the fit, the user can click “Update Coefficient” to close the ML window and automatically apply the updated storage-to-elevation coefficients to the corresponding plant sub-tab.

The elevation-to-storage coefficients are updated in a similar manner, with the roles of the x- and y-axes being reversed.

7.5.2 Updating the PCF function coefficients

PCF functions can be modeled using either 2D polynomial functions or 2D CPWL functions. As discussed in section 4.1.2, the DC CPWL representation yields better fitting results compared to

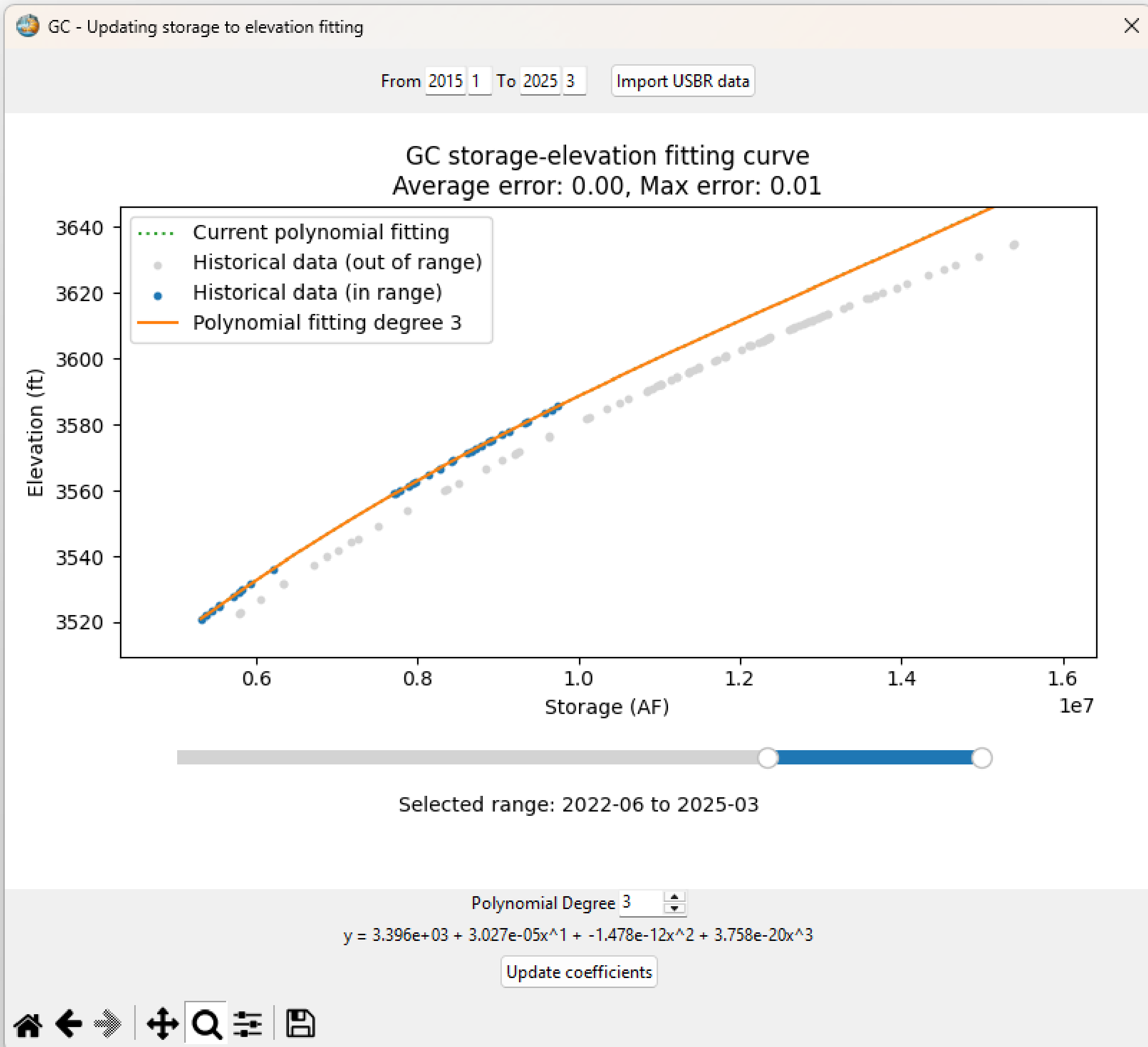

Figure 7.6: Storage-to-elevation ML module

the polynomial approach. Consequently, the CPWL representation is given precedence. Specifically, the polynomial approach is only used when the DC CPWL coefficients are not defined in the GUI. For clarity and simplicity, only the DC CPWL implementation of the ML module is described in this section.

Clicking on "Update DC coefficients" opens a new window, shown in Figure 7.7. This window serves as the GUI for the ML module described in section 4.1.2. Similar to the previous window, the user specifies the historical time period to be used for training the model. At the bottom of the interface, the user specifies the number of "+" and "-" planes used in the DC representation of the PCF function. Below the scatter plot, a date-range and PCF-range slider is available to filter out undesired data points. Contrary to the polynomial fitting, the fitted CPWL function is not updated in real time. Instead, the user must specify a computation time limit and the number of processing cores to be used before launching the ML algorithm by clicking the "Compute CPWL" button. Upon completion of the ML algorithm, if satisfied with the fit, the user can click "Update Coefficients" to close the ML window and automatically apply the updated DC CPWL coefficients to the corresponding plant sub-tab.

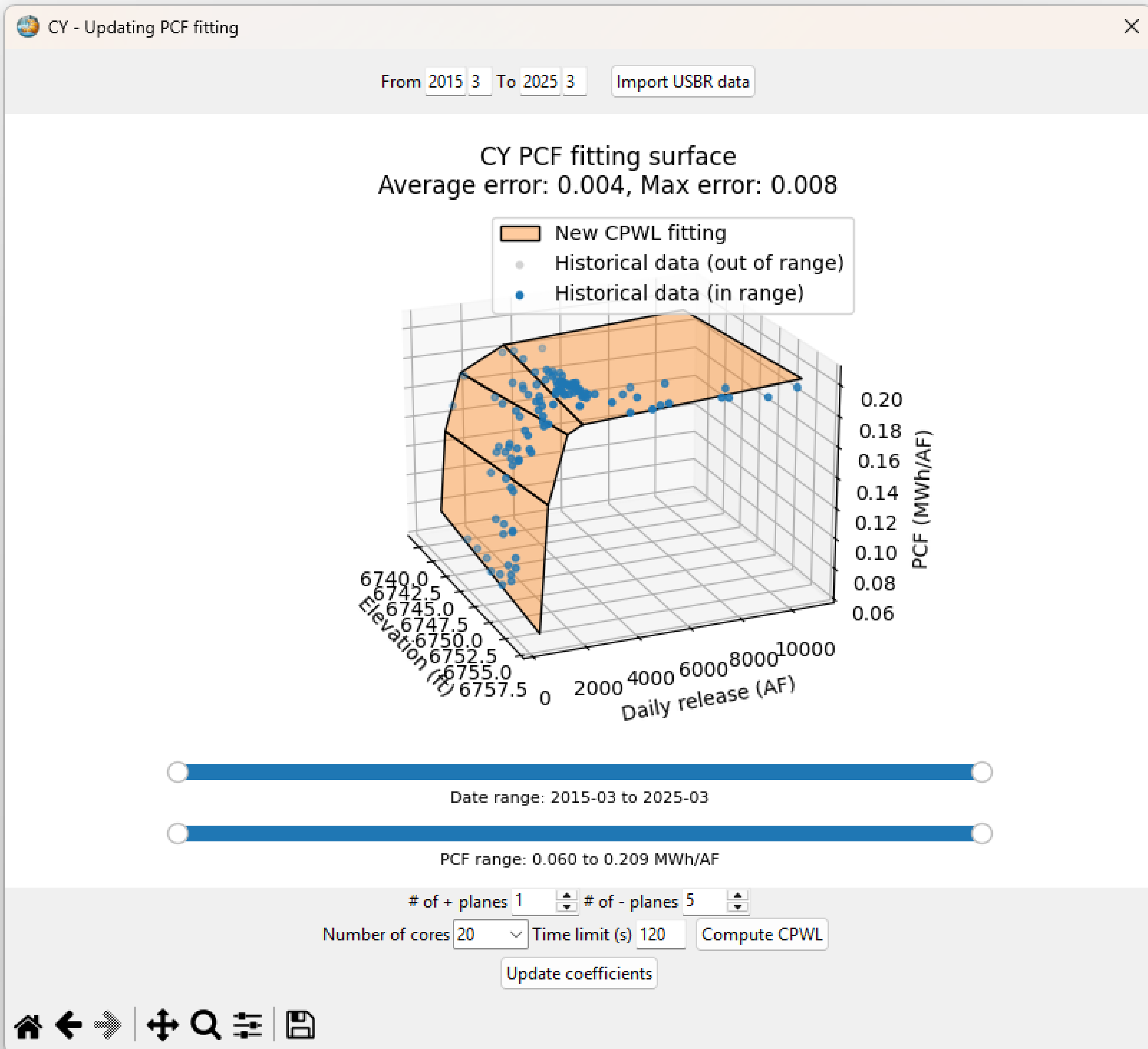

Figure 7.7: PCF function ML module

7.5.3 Updating the water travel time distribution

The ML model identifying the water travel time distribution between the FG dam and the Jensen gage, described in section 4.1.3, is implemented in the software codebase but is not yet accessible to end users. The GUI for this feature will be made available in a future version of CRiSPPy.

7.6 Special events

The "Special events" tab allows the user to define special rules that apply to a subset of model runs. These special rules are described in more details in section 5.1.6. In the example shown in Figure 7.8, the user defines a special rule for the minimum flow rate at GC that only applies during the period from June to July 2025. Upon clicking on the "Special value" field, a window appears where the user can define two flow levels, and the hours of the day these levels apply to.

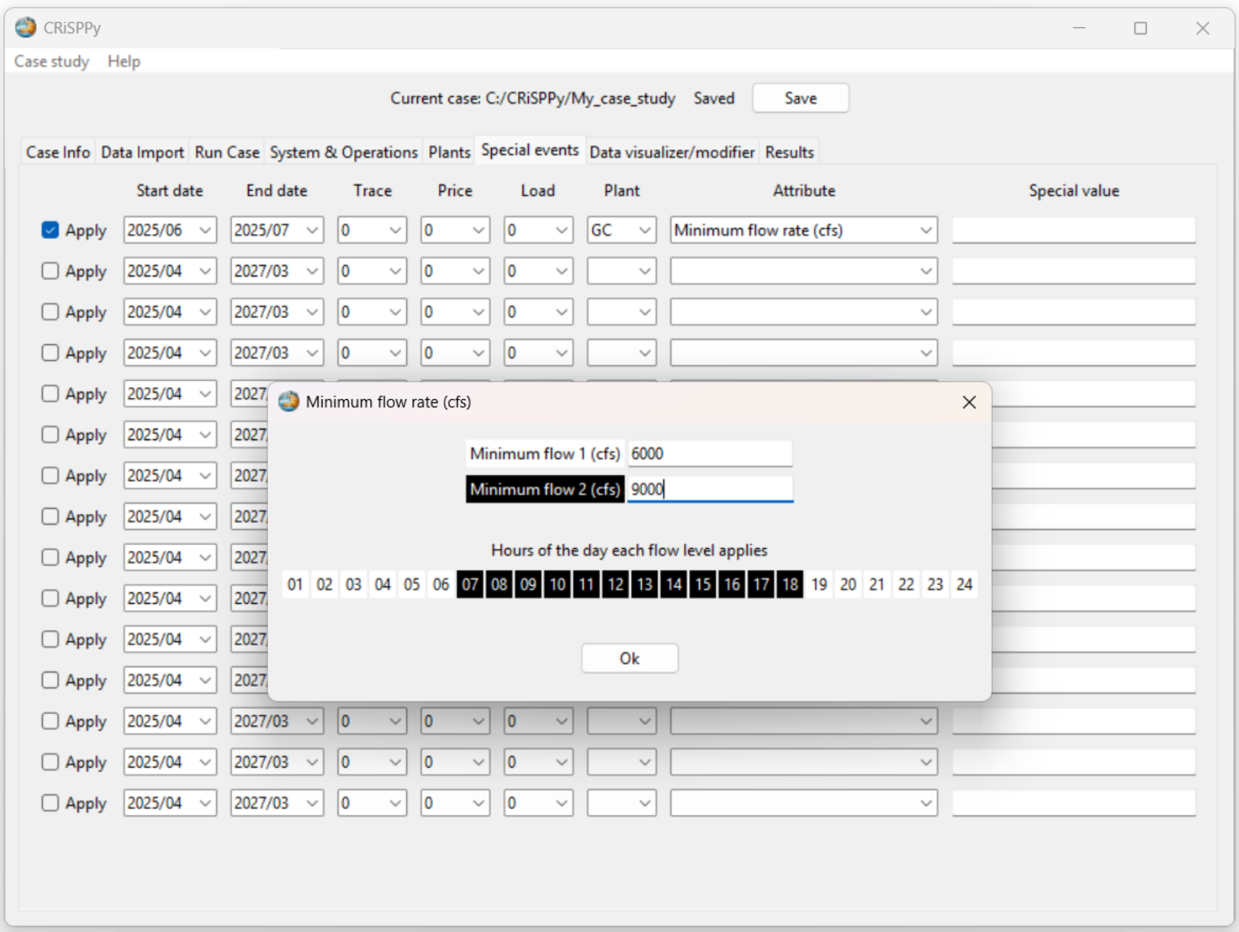

Figure 7.8: Special rules

7.7 Data visualizer/modifier

The "Data visualizer/modifier" is designed to provide the user with refined control over the hourly and monthly input data. This tab is divided into three sub-tabs: "Hourly data", "Monthly data", and "Results scanner". The "Hourly data" and "Monthly data" sub-tabs allow for both reading and writing data, whereas the "Results scanner" only allows the user to read the data and is designed to perform a "quick check" of the model results. The "Monthly data" is not implemented yet but will be available in a future version of CRiSPPy.

7.7.1 Hourly data

The "Hourly data" sub-tab allows the user to view and edit input hourly data,as well as visualize model output data (Figure 7.9). The primary purpose of this module is to enable the user to specify fixed hourly releases over selected time periods. The user begins by defining a time window and choosing between two data sources: "User Specified" or "Output Results". "User specified" data include load, price, but also user-defined power and non-power releases. "Output Results" refer to model-generated hourly outputs such as power generation, turbine releases, and non-power releases. The user can then select one or more data types, along with their associated plant (if applicable), and the corresponding trace, price, and load scenarios. To deactivate a data item, the user simply leaves the "Datatype" field blank.

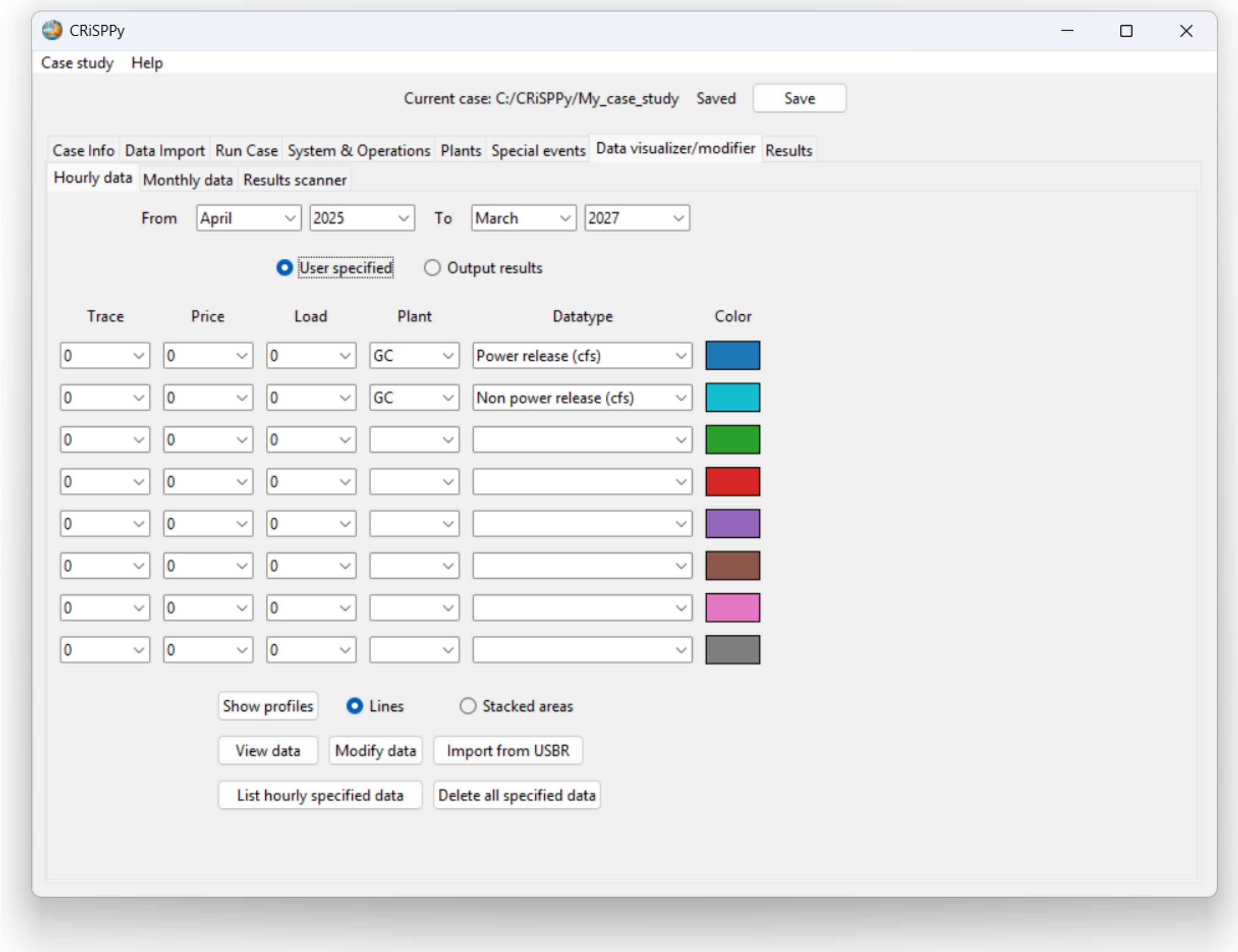

Figure 7.9: Viewing and editing hourly data

The selected data can be visualized using the "Show Profiles" button, with plots displayed either as lines or stacked area charts, depending on the user's selection. The color of each line or area is specified via the "Color" field. The "View data" button opens a read-only tabular file displaying the selected hourly data and is available for both input and output data. The "Modify data" button opens an editable tabular file that allows the user to modify and save input data; this option is only available for "User specified" inputs. Once the editable table is closed, any modifications made to the hourly data are automatically imported into the CRiSPPy case study. "Import from USBR" functions similarly to "Modify data" but additionally pre-populates selected turbine and non-power release profiles with hourly data retrieved from the USBR database.

The "List hourly specified data" button displays a summary table of all user-defined profiles, including their associated start and end dates. The "Delete all specified data" button removes all user-specified profiles from the case study.

7.7.2 Results scanner

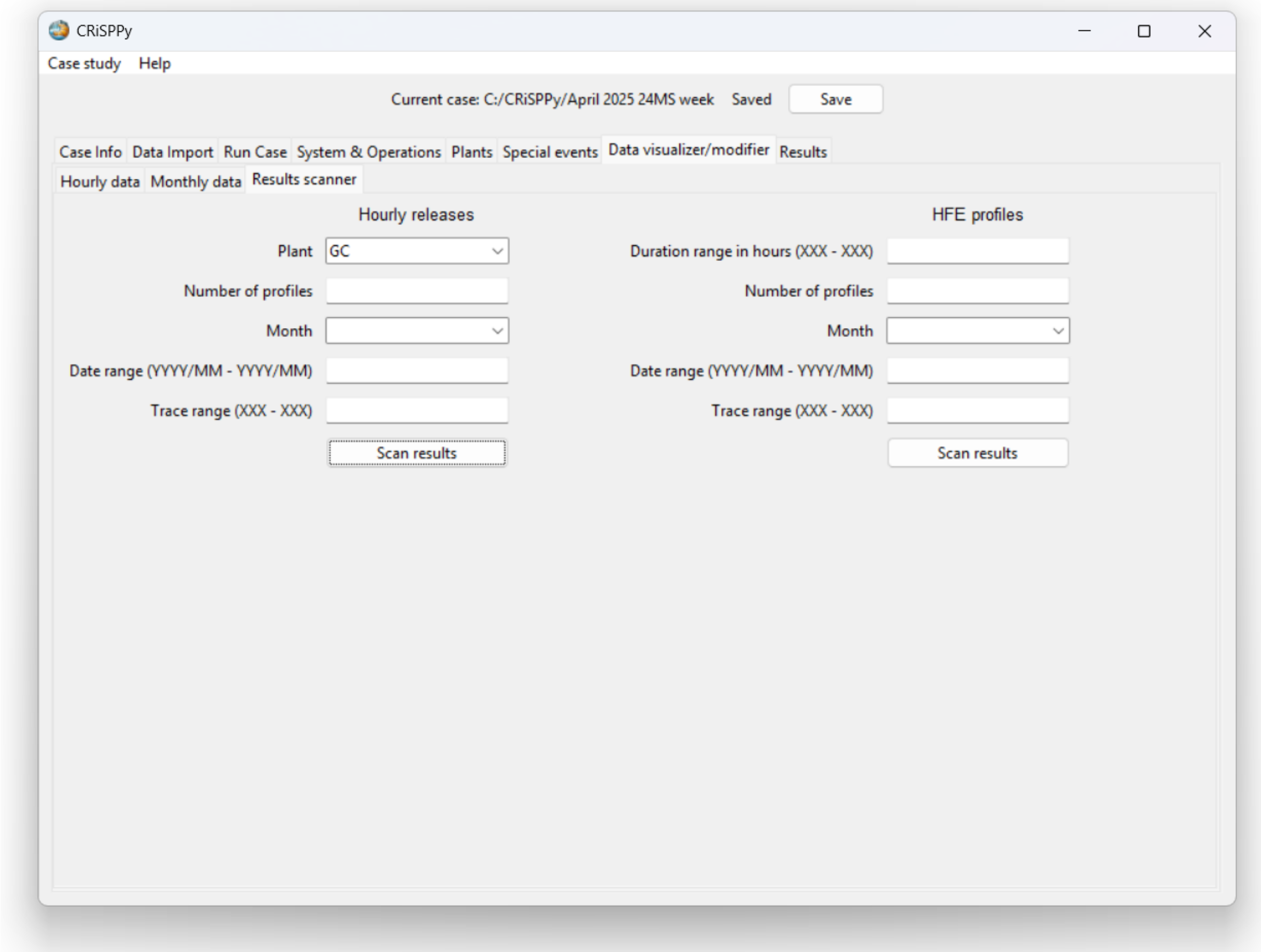

Figure 7.10: Results scanner

The "Results scanner" sub-tab allows users to quickly review the outputs of the hydropower scheduling and high-flow shaping models (Figure 7.10). For hydropower scheduling results, the user begins by selecting a reservoir and specifying the number of profiles to display. These profiles can be filtered by month (*e.g.*, "January"), date range, and trace range.

Upon clicking on "Scan results", a new window opens displaying the selected hourly release profiles, along with any applicable flow rate limits (Figure 7.11). A slider located beneath the graph enables the user to scan through the results, which are sorted in order of increasing weekly release volume.

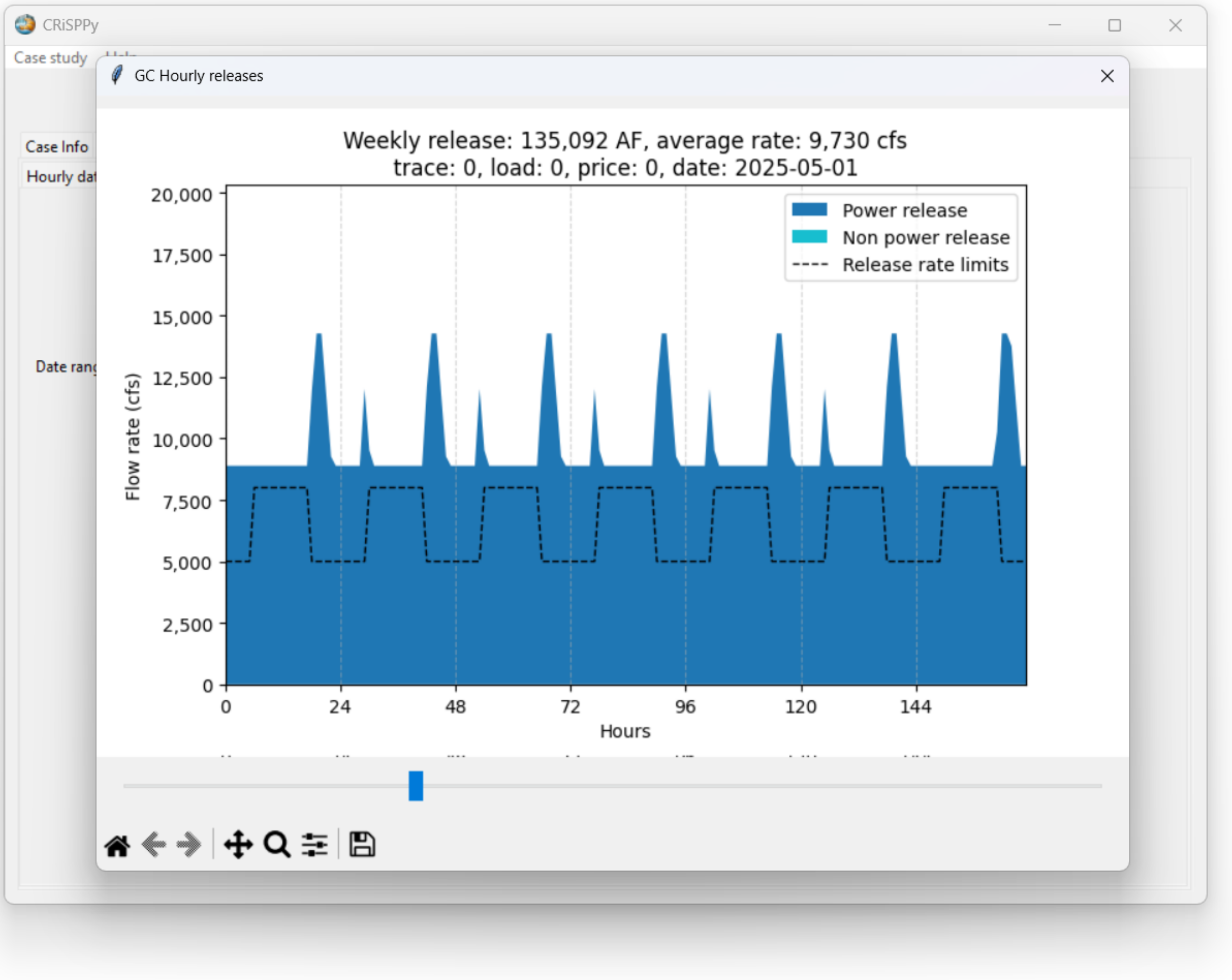

Figure 7.11: Scanning the hydropower scheduling model results

A similar window displays the high-flow release profiles, which can be filtered by HFE duration range.

7.8 Run case

The "Run case" tab is the interface through which the user can execute the data preprocessing step, the hydropower scheduling model, and the high-flow shaping model (Figure 7.12). The user begins by selecting the number of processing cores to allocate for the parallelized LP warm-start procedure (see section 6.4). An optimization solver must also be selected. By default, CRiSPPy uses the GLOP solver [10]; however, the user also has access to the SCIP [11], HiGHS [37], and Gurobi [38] solvers. Note that Gurobi is a commercial solver and requires a valid license.

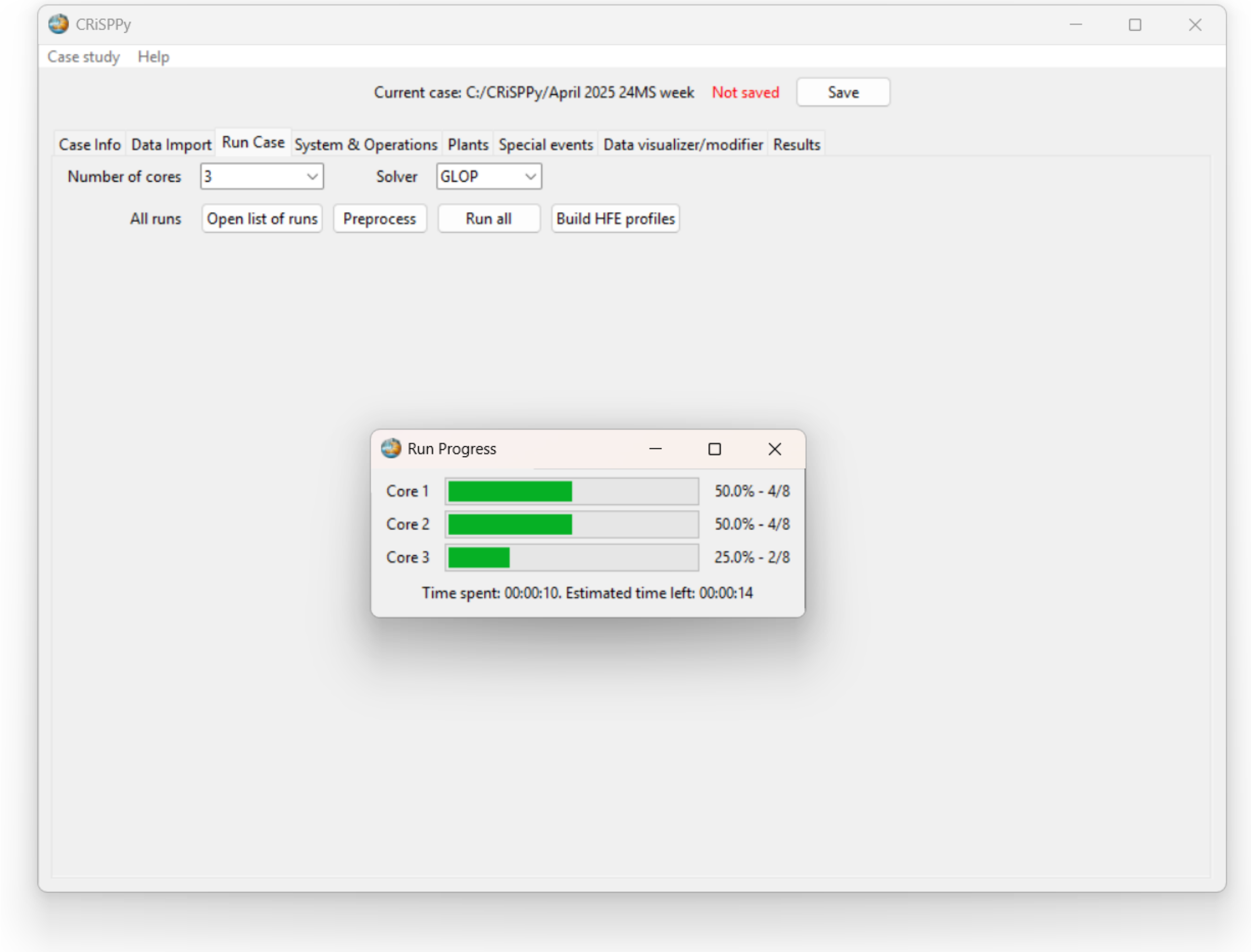

Figure 7.12: Running the model

The data preprocessing phase, described in section 5.2, is initiated by the "Preprocess" button and typically completes within a few seconds. If the user wishes to exclude specific model runs from the hydropower scheduling process, this can be done via the "Open list of runs" button. This option opens a tabular interface listing all model runs, where the user can set individual entries to "False" to exclude them from execution. Clicking "Run all" launches the parallelized LP warm-start execution of the hydropower scheduling model. During execution, a "Run progress" window appears, displaying real-time status updates for each processing core (Figure 7.12).

The "Build HFE profiles" button is part of the post processing step described in section 5.4.1 and is applicable only to case studies that include hydrology data with HFE features. This button initiates a parallelized LP warm-start execution of the high-flow shaping model, as detailed in section 4.3.5.

7.9 Results

The "Run case" tab is the final tab of the CRiSPPy interface and is intended to assist the user in retrieving both hourly model outputs and summary statistics (Figure 7.13). This tab is divided into two sub-tabs: "Hourly results" and "Report writer". The "Hourly results" sub-tab contains a single button that generates a tabular file in the case study folder with all hourly outputs (*e.g.*, turbine and non-power releases, generations, elevations). In contrast, the "Report writer" sub-tab is designed to support the generation of summary statistics and customized visualizations for user-defined data types.

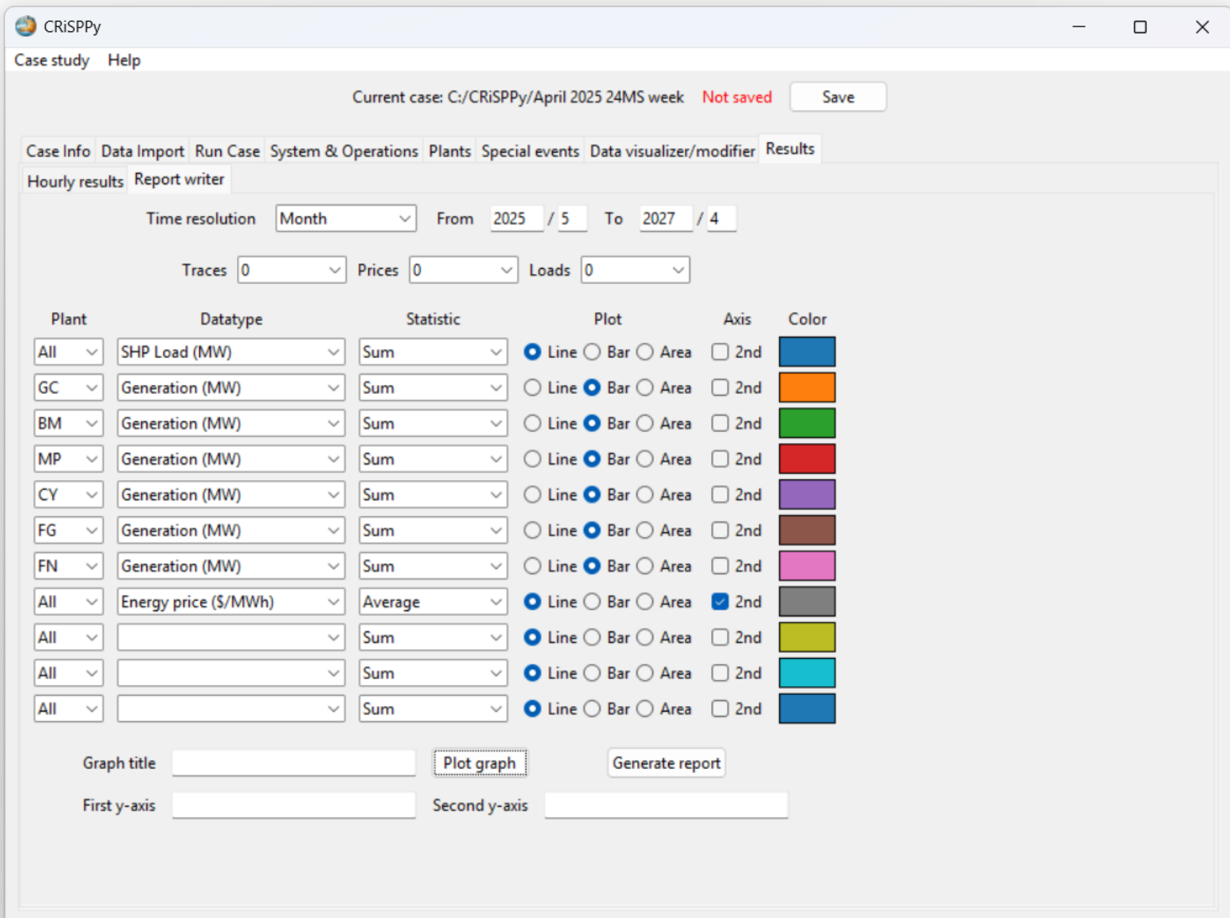

Figure 7.13: Model results

The user begins by selecting a time resolution—hourly, daily, weekly, monthly, or yearly—and specifying the time period of interest. In the current version of CRiSPPy, a specific trace, price scenario, and load scenario must be selected. Future versions will incorporate statistical tools to support the analysis of multiple scenarios simultaneously. Next, the user provides a list of data types along with their associated reservoir, if applicable. Selecting "All" as the reservoir aggregates the corresponding data type across all modeled plants (*e.g.*, total system generation). For each data type, the user specifies a statistical function to apply, including minimum, maximum, average, sum, or selected percentiles. When generating a chart, each data type can be visualized as a line, bar, or stacked area plot, with a custom color. Data types with different units (*e.g.*, \$/MWh vs. MW) may be plotted on a secondary y-axis. As in the "Data Visualizer/Modifier" tab, unused data type fields may be left blank. Before generating the plot, the user may define a title and labels for both the primary and secondary axes. The graph title is also used as the file name for the exported CSV report.

8 Use cases

This section introduces four practical use cases of the CRiSPPy software. The first case is an analysis of the HFE conducted at GC in April 2023 (2023 HFE). The second case is an analysis of the SMB bypass flows implemented at GC from July to November 2024 (2024 SMB). The third case is a 24MS conducted in April 2025 (Apr 2025 24MS). The fourth case is the study of one of the five alternatives proposed under the Post-2026 process (Post-2026 EIS). Table 8.1 provides a summary of the four use cases, including their study scope, time representation, number of model runs, and total model run time.

Table 8.1: Summary of use cases

Case study	Study scope	Time representation	Number of model runs	Total run time (s)
2023 HFE	*Apr 2023* *1 trace* *All plants*	*All hours in a month*	*1*	*11* *(1 core)*
2024 SMB	*Jul 2024 to Nov 2024* *1 trace* *All plants*	*All hours in a month*	*5*	*20* *(5 cores)*
Apr 2025 24MS	*Apr 2025 to Mar 2027* *1 trace* *All plants*	*Representative week*	*24*	*17* *(6 cores)*
Post-2026 EIS (CCA)	*Jan 2027 to Dec 2060* *1,200 trace* *Only GC*	*Representative week*	*489,600*	*3,429* *(20 cores)*

8.1 2023 HFE

As discussed in section 4.3.5, HFEs are controlled high-flow releases periodically scheduled at the GC dam to mimic pre-dam flood events in accordance with the LTEMP ROD framework [19,32]. One such release was conducted in spring 2023, beginning the morning of April 24 and ending the night of April 27, with a peak target flow of 39,500 cfs and an estimated 14 TAF of water bypassed.

The 2023 HFE event was analyzed in CRiSPPy by modeling all reservoirs throughout April 2023, modeling all hours in the month to cover both HFE and non-HFE periods. HFE hours were defined as all hours from April 24 to April 27. During these hours, turbine and non-power releases at GC were imported from USBR records and fixed as user-specified profiles using the "Data visualizer/modifier" module (see section 7.7.1). Historical monthly hydrology data were retrieved from the USBR database, while the load and price profiles were imported from the EMMO "Load and Resource" and "Daily Deals" reports, respectively (see sections 5.1.2 and 5.1.3). After preprocessing, CRiSPPy completed the model run in 11 seconds. Hourly

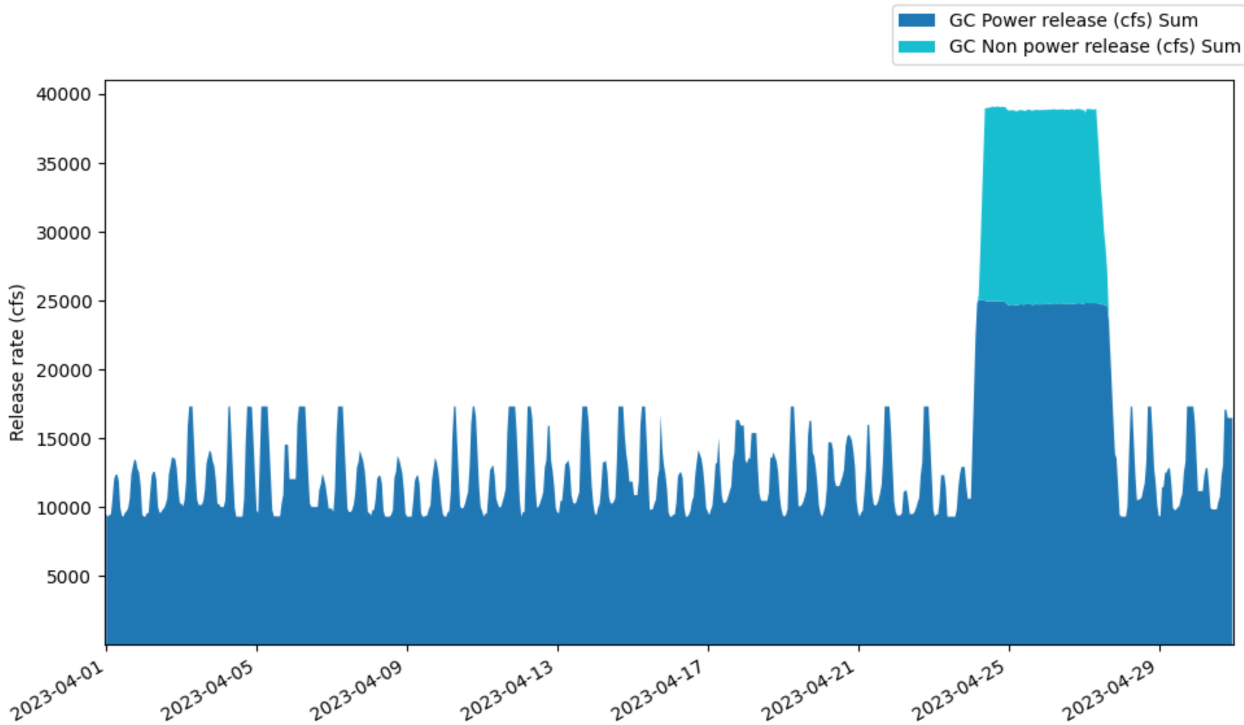

Figure 8.1: 2023 HFE

results for GC are shown in Figure 8.1, where releases during HFE hours reflect the fixed, user-specified data, and releases during non-HFE hours were determined by the optimization model.

8.2 2024 SMB

The SMB is an invasive species threatening native fish in the Colorado River below GC. Warmer surface water in Lake Powell, due to lower lake levels, increase the chance of these bass passing through the dam [39]. GC dam operators use experimental bypass releases of colder water to lower the river's temperature downstream. These "SMB bypass flows" aim to prevent the SMB from spawning, as they need warmer water to reproduce successfully, thus helping to mitigate the invasion. SMB bypass flows were implemented for the first time in 2024, beginning on July 9 and ending on November 18.

The 2023 SMB event was analyzed in CRiSPPy by modeling all reservoirs from July through November 2024, simulating all hours within that period. Throughout the study period, non-power releases (*i.e.*, bypass flows) at GC were imported from USBR records and fixed as user-specified profiles using the "Data visualizer/modifier" module (see section 7.7.1), while turbine releases were left unconstrained. Similar to the HFE 2023 analysis, historical monthly hydrology data were retrieved from the USBR database, while the load and price profiles were imported from the EMMO "Load and Resource" and "Daily Deals" reports, respectively (see sections 5.1.2 and 5.1.3). After preprocessing, CRiSPPy completed all model runs in 20 seconds using 5 processing cores. Daily release results for GC are presented in Figure 8.2, where non-power releases reflect the fixed, user-specified inputs, and power releases were optimized by the model.

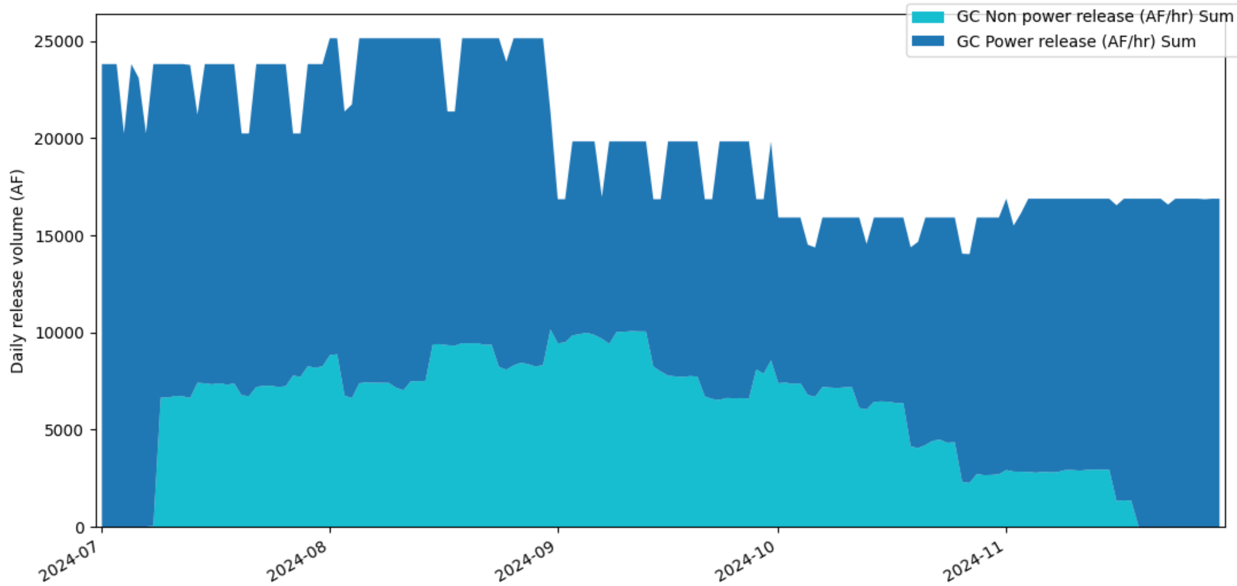

Figure 8.2: 2024 SMB

8.3 Apr 2025 24MS

As described in section 3, 24MS are conducted monthly by the USBR to determine reservoir release decisions 24 months into the future [17]. These monthly release decisions are then used by WAPA to forecast hourly release schedules at each reservoir.

The use case presented here is based on a 24MS with monthly release decisions specified for all plants from April 2025 through March 2027. Since all input data are forecasted, it is not necessary to simulate every hour of each month; instead, a typical week time representation is used. Monthly hydrology data were imported from CRMMS forecast outputs available in the USBR database. Load profiles were obtained from the "Average Load" file, while price profiles were generated using the hourly price generation model described in section 4.3.1, based on monthly price data from the "Forward Curves" files. CRiSPPy completed all model runs in 17 seconds using 6 processing cores. The resulting monthly outputs are shown in Figure 8.3.

8.4 Post-2026 EIS

As discussed in section 1, at the time of writing this report, USBR is preparing a Post-2026 EIS to evaluate long-term operational strategies for the GC dam and Lake Powell, following the expiration of the current guidelines in 2026 [8]. The ongoing analysis focuses on five alternatives over a 34-year planning horizon (2027–2060) and spans 1,200 hydrology traces.

The use case presented here is based on the Cooperative Conservation Alternative (CCA), one of the five alternatives under consideration. Due to its wide scope, the CCA alternative includes 489,600 model runs. As with the 24MS, a typical week time representation is used to balance resolution and computational efficiency. Monthly hydrology data were imported from CRMMS forecast output files provided by USBR. The data include 26,618 HFE events, which are distributed across the model runs and specified by release volume, duration, and peak flow rate, rather than detailed hourly profiles. Load profiles were sourced from the "Average Load"

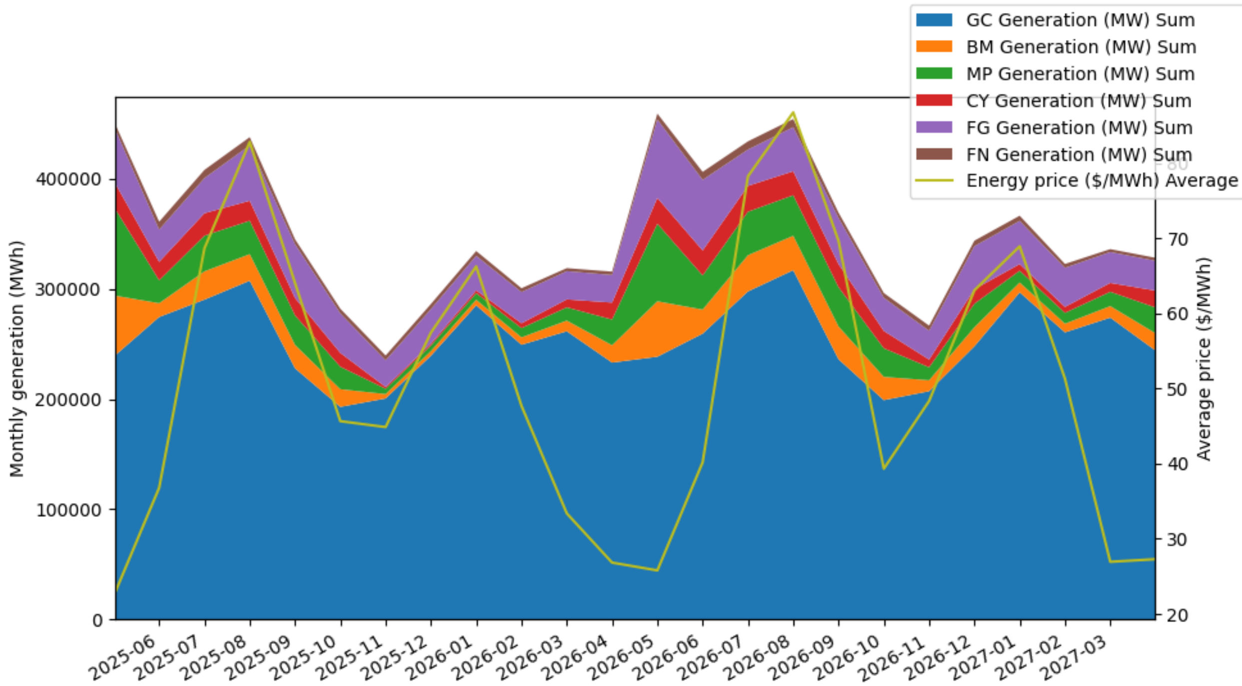

Figure 8.3: Apr 2025 24MS

file, and surrogate price profiles were generated based on the load profiles.

Unlike the previous use cases, a subset of runs in this analysis was found to be infeasible during preprocessing. Specifically, 7,543 runs (1.5%) were flagged for exceeding maximum flow constraints, and 26,548 runs (5.4%) violated minimum flow constraints relative to monthly release targets. These infeasibilities were automatically detected and corrected during the preprocessing phase by the infeasibility detection model and sub-minimum flow shaping model, respectively (see sections 4.3.3 and 4.3.4). Since only GC was modeled, and the infeasibility corrections fully determined the release schedule for affected runs, the 34,901 model runs were discarded from the run batch.

CRiSPPy completed the remaining model runs in 3,429 seconds (under one hour) using 20 processing cores. In the post-processing phase, the hourly release profiles for the 26,618 HFE events were computed in just a few seconds. Overall, the case study produced over one billion hourly output values, including power releases, non-power releases, and power generation. These performance levels were made possible by the parallelized LP warm-start strategy described in section 6.4.

References

[1] Thomas D. Veselka and Michael North. Generation and transmission maximization model, 04 2001. URL: https://www.osti.gov/biblio/1230613.

[2] Linus Scharage. *Optimization Modeling with LINGO*. Chicago, Ill. : LINDO Systems, 2006.

[3] T D Veselka, L A Poch, C S Palmer, S Loftin, B Osiek, Decision, Information Sciences, and Western Area Power Administration. Financial analysis of experimental releases conducted at glen canyon dam during water years 1997 through 2005. Technical report, Argonne National Lab. (ANL), Argonne, IL (United States), 04 2010. URL: https://www.osti.gov/biblio/977375, doi:10.2172/977375.

[4] L A Poch, T D Veselka, C S Palmer, S Loftin, and B and Osiek. Financial analysis of experimental releases conducted at glen canyon dam during water years 2006 through 2010. Technical report, Argonne National Lab. (ANL), Argonne, IL (United States), 08 2011. URL: https://www.osti.gov/biblio/1023155, doi:10.2172/1023155.

[5] Q. Ploussard and T. D. Veselka. Financial analysis of the 2018 glen canyon dam bug flow experiment. Technical report, Argonne National Laboratory (ANL), Argonne, IL (United States), 06 2019. URL: https://www.osti.gov/biblio/1526634, doi:10.2172/1526634.

[6] Q. Ploussard and T. D. Veselka. Financial analysis of experimental releases conducted at glen canyon dam during water year 2017. Technical report, Argonne National Laboratory (ANL), Argonne, IL (United States), 06 2019. URL: https://www.osti.gov/biblio/1526629, doi:10.2172/1526629.

[7] Q. Ploussard and T. D. Veselka. Financial analysis of the 2022 macroinvertebrate production flow experiment at glen canyon dam. Technical report, Argonne National Laboratory (ANL), Argonne, IL (United States), 02 2023. URL: https://www.osti.gov/biblio/1924855, doi:10.2172/1924855.

[8] Federal register. Notice of intent to prepare an environmental impact statement and notice to solicit comments and hold public scoping meetings on the development of post-2026 operational guidelines and strategies for lake powell and lake mead. URL: https://www.federalregister.gov/documents/2023/06/16/2023-12923/notice-of-intent-to-prepare-an-environmental-impact-statement-and-notice-to-solicit-co

[9] U.S. Bureau of Reclamation. Hydrologic database data query. URL: https://www.usbr.gov/lc/region/g4000/riverops/_HdbWebQuery.html.

[10] Laurent Perron and Vincent Furnon. OR-Tools, 2024. URL: https://developers.google.com/optimization/.

[11] Suresh Bolusani, Mathieu Besançon, Ksenia Bestuzheva, Antonia Chmiela, João Dionísio, Tim Donkiewicz, Jasper van Doornmalen, Leon Eifler, Mohammed Ghannam, Ambros Gleixner, Christoph Graczyk, Katrin Halbig, Ivo Hedtke, Alexander Hoen, Christopher Hojny, Rolf van der Hulst, Dominik Kamp, Thorsten Koch, Kevin Kofler, Jurgen Lentz, Julian Manns, Gioni Mexi, Erik Mühmer, Marc E. Pfetsch, Franziska Schlösser, Felipe

Serrano, Yuji Shinano, Mark Turner, Stefan Vigerske, Dieter Weninger, and Lixing Xu. The SCIP Optimization Suite 9.0. ZIB-Report 24-02-29, Zuse Institute Berlin, February 2024. URL: https://nbn-resolving.org/urn:nbn:de:0297-zib-95528.

[12] U.S. Bureau of Reclamation. Colorado river storage project. URL: https://www.usbr.gov/uc/rm/crsp/index.html.

[13] U.S. Department of the Interior. The law of the river, 2008. URL: http://www.usbr.gov/lc/region/g1000/lawofrvr.html.

[14] U.S. Department of the Interior. Annual operating plan for colorado river reservoirs 2014. Technical report, U.S. Department of the Interior, 2013.

[15] U.S. Bureau of Reclamation. General modeling information. URL: https://www.usbr.gov/lc/region/g4000/riverops/model-info.html.

[16] Edith A. Zagona, Terrance J. Fulp, Richard Shane, Timothy Magee, and H. Morgan Goranflo. Riverware: A generalized tool for complex reservoir system modeling. *JAWRA Journal of the American Water Resources Association*, 37(4):913–929, 2001. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.2001.tb05522.x, arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1752-1688.2001.tb05522.x, doi:10.1111/j.1752-1688.2001.tb05522.x.

[17] U.S. Bureau of Reclamation. 24-month study projections. URL: https://www.usbr.gov/lc/region/g4000/riverops/24ms-projections.html.

[18] Western Area Power Administration. Deliverable Sales Amount (DSA) and Western Replacement Firming (WRF): Q & A , 2021. URL: https://www.wapa.gov/wp-content/uploads/2023/04/wrf-questions-and-answers-updated-september-20-2021.pdf.

[19] U.S. Department of the Interior. Record of decision for the glen canyon dam long-term experimental and management plan final environmental impact statement. Technical report, U.S. Department of the Interior, 2016. URL: https://ltempeis.anl.gov/documents/docs/LTEMP_ROD.pdf.

[20] U.S. Department of the Interior. Operation of flaming gorge dam final environmental impact statement executive summary. Technical report, U.S. Department of the Interior, 2005. URL: https://www.uintahwater.org/sites/g/files/vyhlif1341/f/uploads/eis_executive_summary_september_2005.pdf.

[21] Western Area Power Administration. Methodology for Analyzing the Impacts of Aspinall EIS Alternatives on Power Economics. Technical report, Western Area Power Administration, 2008. URL: https://www.usbr.gov/uc/envdocs/eis/AspinallEIS/Vol2-Appdx-D.pdf.

[22] Raymond H Myers et al. *Classical and modern regression with applications*, volume 2. Duxbury press Belmont, CA, 1990.

[23] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. URL: https://www.sciencedirect.com/science/article/pii/0893608089900208, doi:10.1016/0893-6080(89)90020-8.

[24] Anita Kripfganz and R. Schulze. Piecewise affine functions as a difference of two convex functions. *Optimization*, 18(1):23–29, January 1987. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02331938708843210. doi:10.1080/02331938708843210.

[25] Kody Kazda and Xiang Li. Nonconvex multivariate piecewise-linear fitting using the difference-of-convex representation. *Computers & Chemical Engineering*, 150:107310, 2021. URL: https://www.sciencedirect.com/science/article/pii/S0098135421000880, doi:10.1016/j.compchemeng.2021.107310.

[26] Lonnie Magee. Nonlocal behavior in polynomial regressions. *The American Statistician*, 52(1):20–22, 1998. URL: https://www.tandfonline.com/doi/abs/10.1080/00031305.1998.10480531, arXiv:https://www.tandfonline.com/doi/pdf/10.1080/00031305.1998.10480531, doi:10.1080/00031305.1998.10480531.

[27] U.S. Army Corps of Engineers North Pacific Division. Program description and user manual for ssarr, streamflow synthesis and reservoir regulation: Program 724-k5-g001. Technical report, Army Engineer Division, North Pacific:, 1975.

[28] Jeremy Feinstein, Quentin Ploussard, Thomas Veselka, and Eugene Yan. Using data-driven prediction of downstream 1d river flow to overcome the challenges of hydrologic river modeling. *Water*, 15(21), 2023. URL: https://www.mdpi.com/2073-4441/15/21/3843, doi:10.3390/w15213843.

[29] U.S. Geological Survey. Green River Near Jensen, UT. URL: https://waterdata.usgs.gov/monitoring-location/09261000/.

[30] Quentin Ploussard. Piecewise linear approximation with minimum number of linear segments and minimum error: A fast approach to tighten and warm start the hierarchical mixed integer formulation. *European Journal of Operational Research*, 315(1):50–62, 2024. URL: https://www.sciencedirect.com/science/article/pii/S0377221723008585, doi:10.1016/j.ejor.2023.11.017.

[31] U.S. Bureau of Reclamation. Near-term colorado river operations final supplemental environmental impact statement. Technical report, U.S. Bureau of Reclamation, 2024. URL: https://www.usbr.gov/ColoradoRiverBasin/documents/NearTermColoradoRiverOperations/20240300-Near-termColoradoRiverOperations-FinalSEIS-508.pdf.

[32] Glen Canyon Dam Adaptive Management Program. High-Flow Experimental (HFE) Releases. URL: https://gcdamp.com/index.php/The_HFE_Page.

[33] Pandas. MultiIndex / advanced indexing, 2024. URL: https://pandas.pydata.org/docs/user_guide/advanced.html.

[34] Steve Klabnik and Carol Nichols. *The Rust Programming Language*. No Starch Press, USA, 2018.

[35] István Maros and Gautam Mitra. Strategies for creating advanced bases for large-scale linear programming problems. *INFORMS Journal on Computing*, 10(2):248–260, 1998. arXiv:https://doi.org/10.1287/ijoc.10.2.248, doi:10.1287/ijoc.10.2.248.

[36] Jonathan Casey Root and Daniel K. Jones. Elevation-area-capacity relationships of lake powell in 2018 and estimated loss of storage capacity since 1963. Technical report, U.S. Geological Survey, 2022. doi:10.3133/sir20225017.

[37] J. A. J. Hall Q. Huangfu. Parallelizing the dual revised simplex method. *Mathematical Programming Computation*, 10(1):119–142, 2018. arXiv:https://doi.org/10.1007/s12532-017-0130-5, doi:10.1007/s12532-017-0130-5.

[38] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2024. URL: https://www.gurobi.com.

[39] Glen Canyon Dam Adaptive Management Program. Smallmouth Bass Page. URL: https://gcdamp.com/index.php/Smallmouth_Bass_Page.

Energy Systems and Infrastructure Assessment
Argonne National Laboratory
9700 South Cass Avenue
Lemont, IL 60439

www.anl.gov

Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC