



Localization of water resources management technology to adapt to climate change in the Hong-Thai Binh River basin

Activity 4.1 Report on the determination of the scope of intervention of the selected technologies in transboundary water resource management and planning in the Hong-Thai Binh river basin

Request ID# AF-2021000099



October 2023

Prepared for UNFCCC Climate Technology Centre and Network, CTCN, as part of the Adaptation Fund Climate Innovation Accelerator



Localization of water resources management technology to adapt to climate change in Hong-Thai Binh River basin

Activity 4.1 Report on the determination of the scope of intervention of the selected technologies in transboundary water resource management and planning in the Hong-Thai Binh River basin

Request ID# AF-2021000099

Prepared for: UNFCCC Climate Technology Centre and Network, CTCN, as part of the Adaptation Fund Climate Innovation Accelerator Programme

Represented by: Ms Clara Landeiro, Regional Manager, Asia-Pacific

Quality Supervisor: Sílvia Leirião
Author: Bertrand Richaud, Elbys Meneses, Jannik Anthonj and Roar Jessen
Revision: Final 1.0
Classification: **Open**
File name: Act4_1_Report on determination of scope

Contents

Introduction to this Report	1
1 Background	2
2 Approach	4
3 Seasonal Forecasting System	7
3.1 Assessment of reservoir characteristics.....	7
3.2 Detecting timeseries of reservoir levels and storage from Earth Observation.....	10
3.3 Forecasting of unregulated river flows.....	13
3.4 Data retrieval and processing.....	17
3.5 Priestly-Taylor Potential Evaporation approach.....	18
3.6 Forecast bias-adjustment.....	19
4 Model Calibration and Validation	21
4.1 Calibration summary.....	21
4.2 Observation series and calibration/ validation periods.....	22
4.3 Calibration results.....	24
5 Results Presentation and Reporting	28
5.1 Results presentation.....	28
5.2 Operational report.....	28
6 Climate Change	30
6.1 Future scenarios.....	30
6.2 Calculation of Climate Change Factors.....	32
7 Next Steps	34
8 References	35

Figures

Figure 1.1	Outline of the scope of work and outputs of the technical assistance.....	3
Figure 2.1	Overview of the basin with the delineation of the catchments.....	4
Figure 2.2	Catchment delineation and key reservoirs of the Chinese part of the Hong-Thai Binh basin as described in the MIKE HYDRO Basin model.....	5
Figure 2.3	Principle in deriving reservoir characteristics from Earth Observation Data.....	5
Figure 3.1	Diagram of the final scope of intervention for the technology development.....	7
Figure 3.2	Illustration of the Calculation of the area and volume of a region between a surface and a reference plane.....	9
Figure 3.3	7-year timeseries of surface water extent from Sentinel-1 and Sentinel-2 fusion as described in Druce et al. (2021) for Malutang reservoir.....	12
Figure 3.4	Historical volume and level timeseries for Malutang reservoir estimated from the water extent observed by SAR-1 and calculated using the LAV lookup tables.....	12
Figure 3.5	Average Volume variation vs. reservoir inflows for the Yayangshan Reservoir.....	13
Figure 3.6	Conceptual diagram of DHI's Global Hydrological Model. The two green boxes describe the three water routing steps.....	14
Figure 3.7	Time coverage for the operational setup of the Hong-Thai Binh GHM setup.....	15
Figure 3.8	The resolution of the Upper Hong-Thai Binh Basin rainfall-runoff model, illustrated as corners of the single squared NAM model tiles, coloured according to the gauging station to which the tile contributes.....	16

Figure 3.9	ERA5 Precipitation product covering the study area.	18
Figure 3.10	Yearly average evaporation from the Priestley-Taylor evaporation dataset	19
Figure 3.11	GFS precipitation before and after bias correction for the period 2022-10-13 - 2022-12-05, compared to the reference dataset of ERA5 precipitation.....	20
Figure 4.1	Locations of the Vietnamese flow gauging stations (red Diamonds) used for calibration and the two rain gauges (yellow circles) that was not used in the model.	22
Figure 4.2	Comparison of the simulated and observed average monthly flow showing the average annual and seasonal model performance at the various gauging stations (in m ³ /s)....	25
Figure 4.3	Simulated hydrograph (blue curve) and observed flow hydrograph (orange line) of inflows to the new Lai Chau Reservoir.....	26
Figure 4.4	Simulated and observed cross boundary flow hydrographs at the Lao Cai gauging station in Vietnam.....	27
Figure 5.1	Report sent via email, listing the last job instances and simulations in a database.....	29
Figure 6.1	A summary of the Shared Socioeconomic Pathways (SSPs) that are used to form the latest IPCC AR6 climate scenarios (O'Neill et al., 2015).....	31
Figure 6.2	Future annual emissions of CO ₂ in the five new IPCC scenarios. Future emissions cause future additional warming, with total warming dominated by past and future CO ₂ emissions. Source: IPCC AR6 WG1 Summary for Policy Makers Box SPM.1.....	31
Figure 6.3	Summary of the five emissions scenarios that inform the latest IPCC AR6 report.	32

Tables

Table 3-1	Curve fitting parameters of archived GFS forecasts before and after bias-correction.	20
Table 4-1	List of NAM parameters adjusted during the calibration.....	23
Table 4-2	Statistics for validation of the calibration for the five Sub basins.....	24

Appendices

Appendix A	Test protocol – check list	1
-------------------	---	----------

Acronyms

AOI	Area of Interest
AR5	IPCC Fifth Assessment Report
AR6	IPCC Sixth Assessment Report
CDSAPI	Climate Data Store (CDS) Application Program Interface (API)
CEWAFO	Centre for Water Resources Planning and Investigation
CFS	Climate Forecast System
CMIP6	Coupled Model Intercomparison Project Phase 6
CTCN	Climate Technology Centre Network
DCC	Department of Climate Change
DEM	Digital Elevation Model
DFS2	2D Data File System
DTM	Digital Terrain Model
EO	Earth Observation
ERA5	fifth generation ECMWF reanalysis for the global climate and weather
ESA	European Space Agency
ESP	Extended Streamflow Prediction
FAO	Food and Agriculture Organization
GCM	Global Climate Models
GFS	Global Forecast System
GHM	Global Hydrological Model
GIS	Geographical Information System
IPCC	Intergovernmental Panel on Climate Change
LAV	Level-Area-Volume
MHB	MIKE HYDRO Basin
NASA	National Aeronautics and Space Administration
NAWAPI	National Center for Water Resources Planning and Investigation
NCEP	National Centers for Environmental Prediction
NDC	Nationally Determined Contribution
NDE	National Designated Entity
NetCDF	Network Common Data Form
NOAA	National Oceanic and Atmospheric Administration
NVWATER	North Division for Water Resources Planning and Investigation
PET	Potential Evapotranspiration
RCP	Representative Concentration Pathways



SAR Synthetic Aperture Rader
SRTM Shuttle Radar Topography Mission
SSP Shared Socioeconomic Pathways
TA Technical Assistance

Introduction to this Report

This report is the written deliverable of Activity 4.1 of the technical assistance to Viet Nam entitled '**Localization of water resources management technology to adapt to climate change in Hong-Thai Binh River basin**' (reference number AF-2021000099). The proponent and recipient of the Technical Assistance (TA) is the National Center for Water Resources Planning and Investigation (NAWAPI) via its North Division (NVWATER).

Activity 4.1 is the first activity of Output 4 which corresponds to the Customization of the selected technology also called the Forecasting System. The technology aims at improving transboundary flow calculations.

This report outlines the technical specifications of the technology by describing all the steps implemented in the modelling framework and operational system setup. Details on the data sources and assumptions are also given. Then results of the model calibration and validation presented in this report highlight the benefit of the Forecasting system.

The technology will be thoroughly tested during the following Activity 4.3. For this purpose, a draft template for test protocol is included in this report.

Key model outputs and reporting of operational system performance are presented.

Finally, a description of the recently published climate change scenarios is provided, as well as a methodology for how the impacts of climate change can be factored in using a modelling approach. The calculation of the climate change data required for such an approach is also explained in this report.

1 Background

The United Nations awarded funding for a technical assistance (TA) request to the Climate Technology Centre Network (CTCN) to select and customize technologies that can provide science-based information on transboundary water resources in the Hong-Thai Binh River basin in Viet Nam, to ensure more accurate and integrated planning. The proponent is the North Division of the National Center for Water Resources Planning and Investigation (NVWATER). The organization delivering the technical assistance is DHI. The TA is divided into several outputs as described in Figure 1.1 and its duration is from April 2022 to October 2023.

The main output is a seasonal forecasting system that can provide transboundary flow estimates coming from upstream catchments in China. The users of the technology to be developed are:

- Department of Water Resources Planning
- Department of Water Resources Monitoring
- Center for Water Resources Warning and Forecast (CEWAFO)

All these units are under the National Center for Water Resources Planning and Investigation (NAWAPI). The system will be installed at CEWAFO.

Output 2 led to the identification of the **climate change risks and challenges in the Hong-Thai Binh River basin**. During Activity 2.1, the consultant held a kick-off meeting and stakeholder consultations with part of the core team not based in Vietnam having visited Hanoi. It presented the opportunity for face-to-face meetings where the broader group of TA stakeholders were consulted. This was also an information and data collection effort to inform the remaining activities and a critical step in the identification of climate change risk and challenges to transboundary Water Resource Management (WRM) in the Hong-Thai Binh River basin.

Output 3 (Activities 3.1 and 3.2) served to identify **the most appropriate technology to provide science-based information was described**. The selection was fully based on the findings of the stakeholder consultation mission to Hanoi especially the final technical meeting held at the end of the visit with technical and high level NAWAPI decision-makers. The meeting included Mr. Tong Ngoc Thanh, General Director of NAWAPI, Mr. Nguyen Ngoc Ha, Deputy General Director of NAWAPI, and the key contact point and proponent for this TA, Mr. Nguyen Chi Nghia, Director General of NVWATER.

The main objective of Output 4 is the **Customization of the selected technology**. This report sets out the scope of intervention of the selected technology corresponding to the key deliverable under Activity 4.1. The selected technology has been implemented as described in this report during in Activity 4.2.

Section 2 of this report provides an overview of the approach to the system development. Section 3 describes the modelling framework and section 4 the key results of the model calibration and validation are outlined. Section 5 discusses result presentation and operational reporting and section 6 highlights the calculation of climate change factors to be used by NAWAPI in estimating climate change impacts and the recommended approach for building climate change scenarios.

Finally in section 7, this report summarizes the next steps and gives an overview on how Activity 4.3 will be carried out to verify and report the information provided by the developed technology. A proposed test protocol has been drafted and included in Appendix A of this report. This protocol is to be used by the system operators during the testing period from March to September 2023, and sent back to DHI in case there are issues to be reported.

Output 2: Identification of climate change risk and challenges to transboundary WRM

- Act 2.1: A kick-off meeting, stakeholder consultations and site visit
- Act 2.2: Review of challenges and climate change risk to transboundary water resource management
- Act 2.3: Investigation of case studies on the use of different models of data management and sharing for transboundary water resource management at the river basin level
- Act 2.4: Assessment of data needs

Output 3: Identification of the most appropriate technology to provide science-based information

- Act 3.1: Enumeration and review of the advanced technologies providing water resource data in a transboundary river basin context
- Act 3.2: Selection of the most appropriate advanced technologies to be used for providing transboundary water resource data

Output 4: Customization of the selected technology

- Act 4.1: Determination of the scope of intervention of the selected technologies in transboundary water resource management and planning in the Hong-Thai Binh river basin
- Act 4.2: Establishment of the selected technologies and baseline information
- Act 4.3: Report and verification of information on transboundary water resources

Output 5: Capacity enhancement for use of the selected technology

- Act 5.1: Development of the technical manual for use of the selected technologies
- Act 5.2: Capacity needs assessment and development of the capacity building program for use of the selected technologies
- Act 5.3: Training of government bodies and stakeholders for use of the selected technologies

Figure 1.1 Outline of the scope of work and outputs of the technical assistance.

2 Approach

The developed technology is a **Forecasting System** covering the basin outside Viet Nam based on software already in use at NAWAPI's Forecasting Center. This choice has been motivated to ensure a successful transfer of the technology to NAWAPI. Hence, the Forecasting Model is based on MIKE HYDRO Basin¹ modelling code and MIKE OPERATIONS² platform.

The aim of the Forecasting System model is to forecast the flow at the border between China and Viet Nam. The applied technology is a combination of the GHM (a new large-scale operationalisation of the traditional NAM hydrological model) and MIKE HYDRO Basin Water allocation model. The models are based on freely available data and both the NAM and the MIKE HYDRO Basin (MHB) models are being used extensively by NAWAPI in their daily operations.

To forecast the flows in the various transboundary rivers along the border with China, a flow routing model has been established of the upper parts of the Hong-Thai Binh Basin in China. The model relies on real-time reanalysed numerical weather prediction data as input to the hydrological model. Moreover, the model includes assessed characteristics and overall regulation rules of the reservoirs in China, and the flow routing in the main rivers from these reservoirs to the Vietnamese border.

The hydrological model, developed at a very fine scale (0.1° x 0.1° grid), provides runoff input to the routing model that is then used to compute the runoff to all the catchments defined in the water resources model. Figure 2.3 gives an overview of the different scales used in the modelling framework.

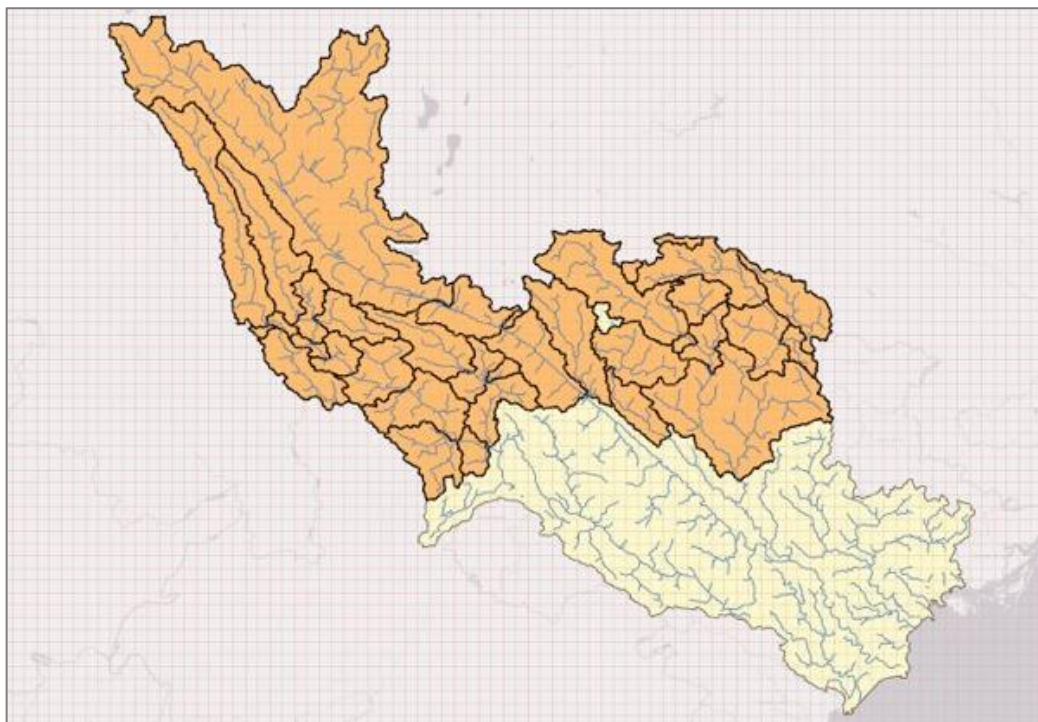


Figure 2.1 Overview of the basin with the delineation of the catchments.

The orange catchments are delineated with the location of the reservoirs and are integrated into the MIKE HYDRO Basin model. The red grid is the finer resolution used for computing runoff by the hydrological model. A routing model is used to convert gridded runoff to the catchment runoff.

¹ https://manuals.mikepoweredbydhi.help/latest/MIKE_HYDRO.htm

² <https://www.mikepoweredbydhi.com/products/mike-operations>

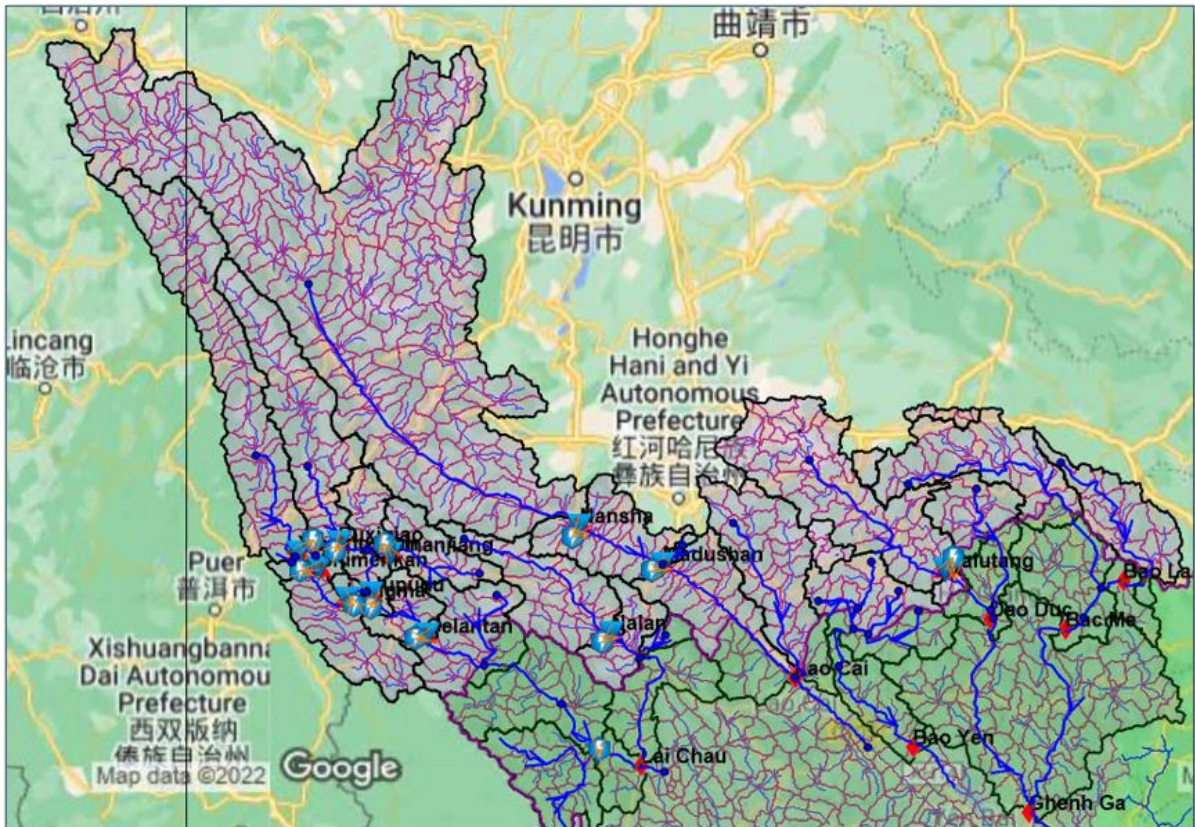


Figure 2.2 Catchment delineation and key reservoirs of the Chinese part of the Hong-Thai Binh basin as described in the MIKE HYDRO Basin model.

One important task of the configuration of the technology has been the identification of the reservoir characteristics. The reservoir storage characteristics of the selected Chinese reservoirs have been defined by its Level-Area-Volume relations derived from Earth Observation (EO) data products from reputable and widely used providers. Furthermore, EO data sources were also applied to assess historical time series of reservoir storage, water level and surface area as illustrated in Figure 2.3

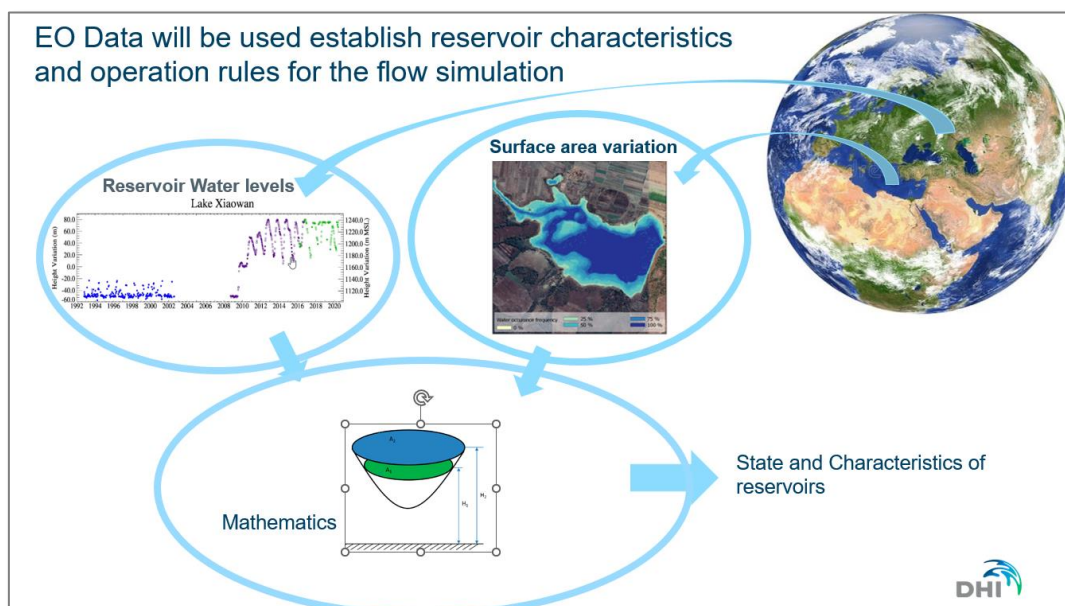


Figure 2.3 Principle in deriving reservoir characteristics from Earth Observation Data.

The calibration and validation of the technology is based on several metrics as presented in section 4.

The technology described in the report has been established and provides a baseline model setup for the Hong-Thai Binh River basin (Activity 4.2).

The developed technology has been demonstrated to the stakeholders during 4 online sessions organized in collaboration with NVWATER, NAWAPI and CEWAFO. During these sessions the technologies (tools and models) were presented through a live demonstration and showcased. These sessions allowed sufficient time for questions and answers and have been valuable in describing the technology.

Finally, the established technology will be tested for a period of 6 months following the implementation during **Activity 4.3 Report and verification of information on transboundary water resources in the Hong-Thai Binh River basin**. Activity 4.3 aims at finetuning/improving the operationalization of the selected technology and ensuring that any related issues or risks detected (including access to data) are addressed, in consultation with the operators. To monitor the system and model status, a test protocol has been developed – included in Appendix A. There, system operation actions are described, which should be checked on their robustness. The test protocol template will serve as a check list as well as for reporting of Activity 4.3.

3 Seasonal Forecasting System

This chapter describes the modelling and operational framework of the seasonal forecasting system that has been developed based on the technology selection made during the technical assistance. Figure 3.1 gives an overview of the steps for the implementation of the system. Each step is then described in a specific section.

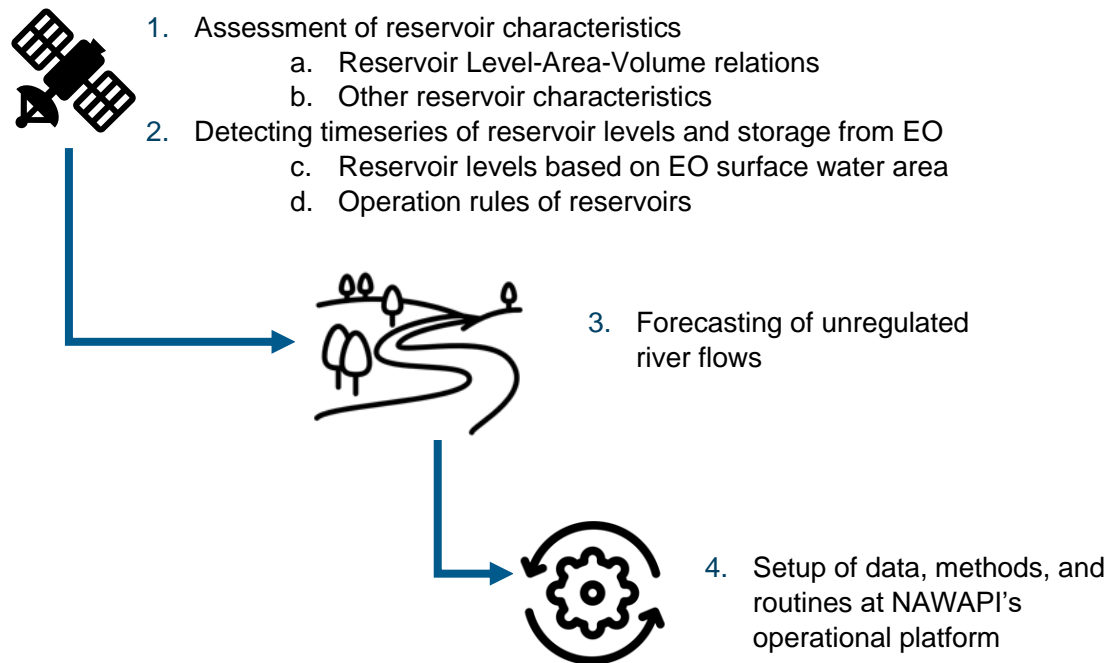


Figure 3.1 Diagram of the final scope of intervention for the technology development.

Steps 1 and 2: allow the assessment of **reservoir characteristics** by detecting the characteristics and states of many predefined reservoirs (including level-area-volume relations and timeseries of reservoir levels and storage). These steps are described in the following sections:

- Section 3.1 Assessment of reservoir characteristics
- Section 3.2 Detecting timeseries of reservoir levels and storage

Step 3: allows forecasting the **unregulated river flows** by implementing the Global Hydrological Model as well as reservoir operation rules

- Section 3.3 Forecasting of unregulated river flows

Step 4: Setup of data, methods, and routines at **NAWAPI's operational platform**

- Section 3.4 Data retrieval and processing
- Section 3.5 Priestly-Taylor Potential Evaporation
- Section 3.6 Forecast bias-adjustment

3.1 Assessment of reservoir characteristics

Surface water reservoirs are established on river systems to regulate their waters, either by raising the water levels (for power production or navigation interests) or to store water from one point in time to another. Knowledge of the characteristics and operation rules of larger reservoirs

are therefore relevant for the prediction of the downstream river flows, which not only depend on hydrology but also on the reservoir regulation.

The 11 largest reservoirs in the upper parts of the basin, found in the Global database of large Dams (Grand³) version 1.3 published by the Socioeconomic Data Center (SEDAC, NASA), have been included in the model. Additionally, the recently built Vietnamese Lai Chau Reservoir (on the Da River) has been included in the model because of its location upstream of the Lai Chau gauging station and its influence on the gauged flows at said station.

For the Lai Chau Reservoir the characteristics and timeseries of historical operations are available and its characteristics and operation rules were copied into the model from NAWAPI's MIKE HYDRO Basin model⁴ of the Vietnamese parts of the HongThai Binh Basin.

Contrary to the extensive knowledge we have available on the Lai Chau Reservoir, very little information was available on the 11 Chinese reservoirs upstream of the border. The Grand database gives only overall characteristics of the various reservoirs, and this information is not always exact. The Grand data have therefore been considered in the model establishment in combination with other information. Due to this lack of accurate data, our methodology has focussed on prioritizing the information deemed most reliable and fill in the knowledge gaps using other sources.

Reservoir Level-Area-Volume relations

A key input when simulating surface water reservoirs are the Level-Area-Volume (LAV) relations or curves, which describe how the surface area and storage volume vary with the reservoir water level. This information was not available, and therefore the curves were estimated through analyses of a Digital Elevation Model (DEM). A DEM holds information on terrain elevations in a dense grid (in this case every 30 by 30 metres). The method applied was to fill depressions in the DEM, and in that way estimate the storage and area for different elevations (also interpreted as water levels).

Our methodology is based on the use of a DEM obtained earlier than the construction of the dams and filling of the reservoirs. The applied DEM source from the Shuttle Radar Topography Mission⁵ (STRM) is from the year 2000 (product name STRM GL1 Global 30 m), which is earlier than the construction date of all dams in the study area, hence it represents the elevations of the terrain now constituting the hidden bottoms of the reservoirs.

At each grid point in the DEM, the distance between a certain hypothetical (water) level and the DEM level (now the reservoir bottom) represents the depth of the reservoir at that location of a given water level. Counting all the points with depths larger than zero and integrating the depths for these points (accounting for the distance between them) reveals the reservoir surface and the reservoir volume, respectively, at that elevation as illustrated in Figure 3.2. Repeating this exercise for multiple elevations from the lowest point of the dam up to its crest results in a full LAV look-up table for the reservoir – a table that is used as input to the MHB model.

³ <https://sedac.ciesin.columbia.edu/data/collection/grand-v1>

⁴ Reference is made to NAWAPI's existing model for the Vietnamese part of the basin, on which they base their water resources management and planning activities, including preparation of annual water resources plans.

⁵ Farr, T. G., and M. Kobrick, 2000, Shuttle Radar Topography Mission produces a wealth of data. Eos Trans. AGU, 81:583-583

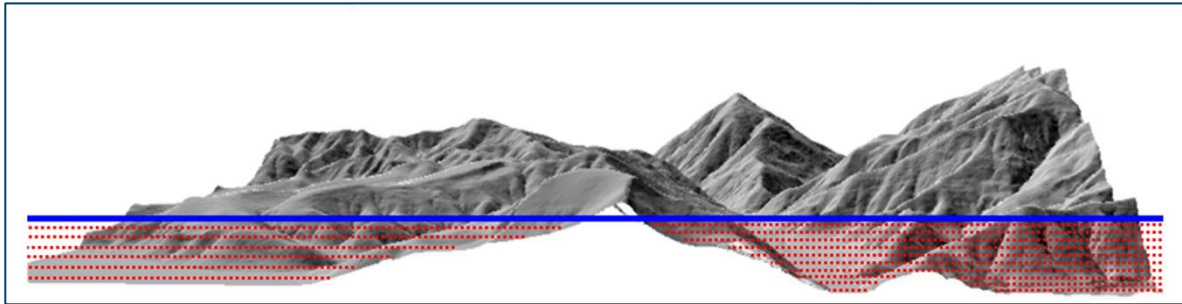


Figure 3.2 Illustration of the Calculation of the area and volume of a region between a surface and a reference plane⁶

The applied DEM has a spatial resolution of 30 m by 30 m resulting in a very large number of bottom level points within each of the reservoirs and we find that this is the most plausible methodology for the establishment of the LAV relations. We do, however, acknowledge the uncertainty, particularly in the absolute levels relative to the sea level. It is therefore important to realise that the LAV curves may have an offset when compared to more accurate absolute levels and keeping this in mind when postprocessing levels from other sources, such as EO data.

Using this described methodology we have developed LAV Curves detailing the surface area and stored volume at 5-meter increments for all 11 dams and included this information in the MHB model

Other reservoir characteristics

The various reservoir characteristics have been assessed as follows:

1. The bottom level of the dam has been assessed as the DEM elevation of the lowest point in the terrain just upstream of the dam⁷
2. The height of each dam has been taken from the Grand database. The height of a dam is an important detail in structural engineering and is therefore regarded plausible.
3. The dam crest level has been calculated as the bottom level of the dam plus its height.
4. The freeboard, the level difference from the full reservoir level (the flood control level) to the crest, has been *assumed* to be 6 m for all dams. This assumption complies well with the freeboard from the Lai Chau reservoir (5 m).
5. Assuming that the full reservoir level corresponds to the flood control level (assessed above), the *dead storage* of each dam has been assessed by subtracting the capacity of the dam (listed in Grand data base) from the LAV volume at flood control level. In other words, we assume that the Grand storage capacity represents the live volume of the reservoir, and it is correct. As this capacity is lower than the full storage at flood control level the remaining storage is assumed to be dead storage.
6. Alternatively, Earth Observation (EO) techniques can be used to assess the characteristics and filling states of remote reservoirs and this methodology has also been tested on the reservoirs in NAWAPI's upper Hong Thai Binh model. However, the very long and narrow shapes of the Chinese reservoirs have shown to introduce problems in the direct detection of water levels in five of the 11 reservoirs. The established LAV curves (based on the SRTM

⁶ Source : <https://pro.arcgis.com/en/pro-app/latest/tool-reference/3d-analyst/surface-volume.htm>

⁷ This concept was selected as the damsites are often located at narrow cross sections of the valleys where influence from the sides may create incorrect thresholds in the "valley floors"

DEM) are therefore found to be the most reliable methodology at this stage and will be used in combination with the Earth observations of water surface area to generate data for near real time data assimilation of the MHB model before each forecast.

3.2 Detecting timeseries of reservoir levels and storage from Earth Observation

Time series of historical water level are important to learn how the reservoirs are operated. Since such timeseries are not available, EO data and postprocessing techniques have been applied to obtain such historical timeseries of reservoir storage. The generated time series has in turn been used to guide the estimated dam operations rules in the MHB model to be applied in the forecast simulations.

Sentinel-1 products by ESA are important sources of data for monitoring surface water as it often provides imagery when is not available through optical sensors due to cloudy regions or light constrains. This makes the C-Band Synthetic Aperture Radar (SAR) of the Sentinel 1 mission attractive for operational purposes for water resource applications.

An operational workflow has been established for gathering the observed water surface areas for all 11 reservoirs. The workflow is in fact semi-real time, since satellite data is available every 5-12 days, only. The implemented workflow methodology is adapted from Druce et al. (2021) and uses a multivariate logistic regression model to estimate surface water probability from the combined usages of Sentinel-1 and Sentinel-2 imagery. In contrast to the original approach by Druce et al., the current implementation favours the SAR data to provide robust monitoring surface water extent dynamics irrespective of cloud cover. The resulting water probabilities are determined using image thresholding into a binary water/not-water classification, and the surface water area time series are produced.

Satellite altimetry observations from Sentinel-3 may be used to estimate water surface elevation (WSE) with a 27-day repeat cycle. The altimetry data will contain time stamped along-track WSEs, which can be reduced to a summary statistic; mean or median to obtain a WSE estimate with a better precision for a given crossing over an inland water body.

Altimetry measurements are point scale measurements and can only be done in vertically downwards (nadir) direction meaning that only points where the satellite ground track intersects with rivers or water bodies can be measured. The minimum size of a target is approximately 50-100 meters but also depends on surrounding topography and the orientation of the water body.

Due to these limitations only the water levels for 5 out the 11 reservoirs could be estimated from satellite altimetry.

If both altimetry and water extent observations are available, reservoir water volumes and water levels can be determined at the time of observations and LAV relations established for the operational parts of a given reservoir. A methodology for this is described in Textbox A, below.

Textbox A

Methodology for assessment of reservoir volumes from observed altimetry and water area data

The change in the reservoir water volume can be obtained by combing the surface water extent with the water surface elevation.

The relationship between water volume and water level is known if the bathymetry is available:

$$V = \sum_{i=1}^n (H - H_i) \times A \quad \text{Equation 1}$$

with V = total storage, H = water body surface height, H_i the bottom elevation of each volume, and A is the sectional area of volume. This can be used to establish the stage-storage curve. Thus, in principle the instantaneous storage can be calculated using altimetry data.

However, in most cases bathymetry data is not available. In such cases it may be more relevant to look at changes in water volumes and which can be estimated combining reservoir level with corresponding water extent. Sometimes, the estimation of storage change based on a constant extent may provide a reasonable approximation. However, with the high temporal resolution from the Sentinels it is possible to match updated water extent with the altimetry observations to estimate lake storage changes dynamically:

$$S = \frac{1}{3} (H_2 - H_1) (A_1 + A_2 + \sqrt{A_1 \times A_2}) \quad \text{Equation 2}$$

where, S is the storage change; H_2 , H_1 and A_2 , A_1 are lake levels and areal extents at different dates, respectively.

Reservoir levels based on EO surface water area

Since only surface area observations can be obtained for all 11 reservoirs from Sentinel-1 and Sentinel-2, the established LAV relations are applied to obtain corresponding time series of reservoir water levels and volumes from observed surface areas, a methodology that is also intended for operational data assimilation in the real time system.

An example of an EO generated time series of water surface area of the Malutang reservoir is displayed in Figure 3.3. The established LAV curve is applied to relate the EO timeseries of surface area to water levels and volumes. As the observed areas are not necessarily consistent with those in the LAV curve, the EO based area needs adjustment. The EO area representing the full reservoir is assessed from the timeseries (here around 9 km²) by relating this EO area to the area in the LAV curve corresponding to the full reservoir (the Flood Control level) an adjustment factor is obtained.

Assuming linearity all EO areas are hereafter adjusted with the same factor and a new adjusted surface area timeseries is obtained (Figure 3.3) that can be converted to corresponding water levels and reservoir volumes using LAV curve as a lookup table (Figure 3.4).

From the adjusted areas, the water levels and reservoir volumes are calculated from the LAV curve as input for the real time operational model. In Figure 3.4 below, the 7-year historical level and volume timeseries are displayed for the Malutang reservoir.

From Figure 3.4 also a representative lowest water level can be detected at a level which has been introduced in MHB, as minimum operational level, for simulation of the reservoir operations during the forecasts.

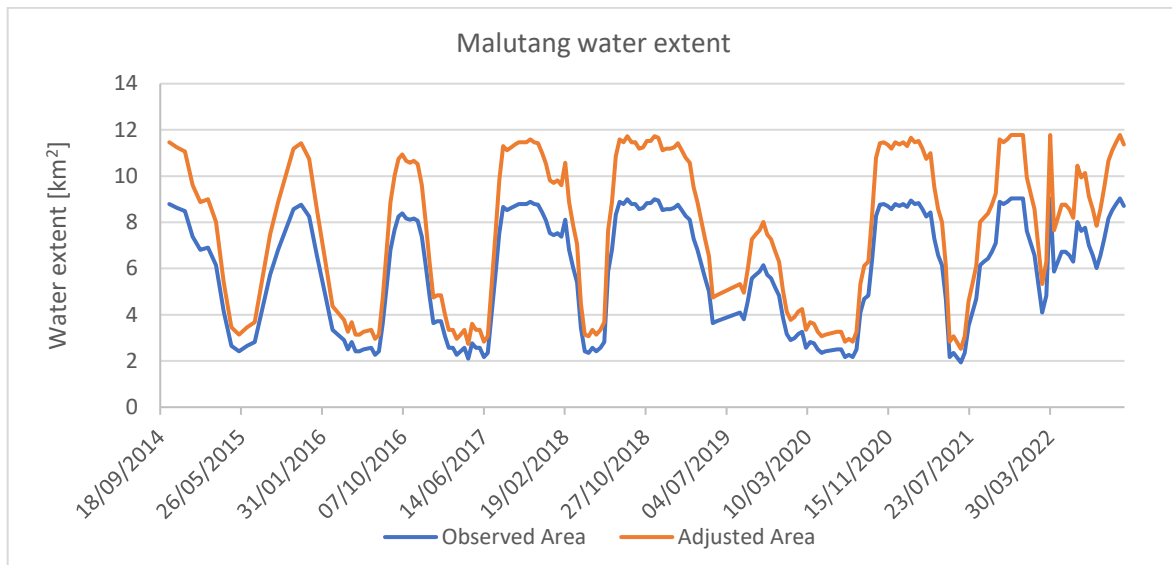


Figure 3.3 7-year timeseries of surface water extent from Sentinel-1 and Sentinel-2 fusion as described in Druce et al. (2021) for Malutang reservoir.

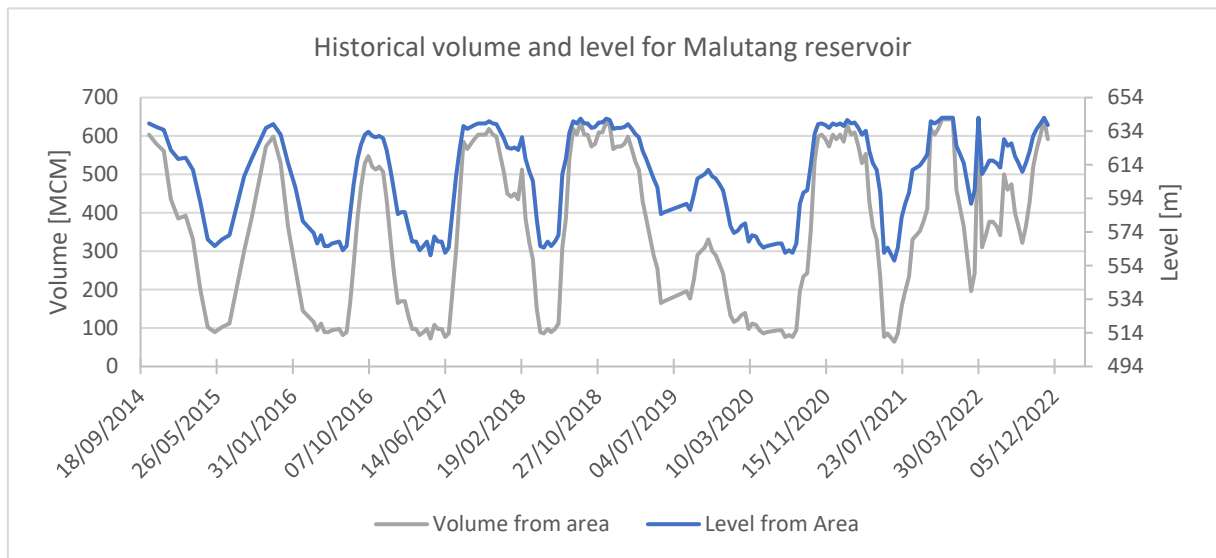


Figure 3.4 Historical volume and level timeseries for Malutang reservoir estimated from the water extent observed by SAR-1 and calculated using the LAV lookup tables.

Operation rules of the reservoirs

The operation rules in MHB are based on rule curves of reservoir water levels. A possible rule curve may be revealed from the average monthly reservoir water levels which can be processed from the generated water level time series. Although a lot can be learned from studying the detected water level time series it is also obvious from Figure 3.4 that operation rules may change from year to year and not only from season to season. At other reservoirs such as the Yayangshan, the upper reservoir on the Da River cascade, volume seems to be kept low during the flood season, presumably for flood protection purposes (Figure 3.5).

The reservoir operations in MHB are framed by the following characteristic water levels:

- The flood control level: defines the max operational level (only to be exceeded during flood operations); this level has been assumed to be 6 m below the crest level of the dam.

- The minimum operation level: defines the reservoir level at which the extraction from the reservoir will stop; this level has been determined from the water level timeseries derived from the earth observations as described above.
- The dead storage level: under this level no water can be extracted.; in this case it has been defined as the level corresponding to the bottom of the active storage (the capacity volume in the GranD database). It was found to be lower than the detected minimum operation level (which it should be). So, in fact it has no influence on the results, but it is required by the model.

In accordance with the information from the GranD database (Lehner et al, 2011) and its technical references, all dams are assumed to be operated exclusively for power production for which water is extracted at a constant rate when the reservoir level is between minimum operation level and flood control level. The rate has been adjusted through calibration to stretch the supply over the full dry season. When the water level is above flood control level the reservoir inflow is turbinated up to the full capacity of the turbines while any remaining inflow is spilled.

It should be noted that for determination of downstream flows the separation between turbinated flow and spill will not matter.

The extraction from the reservoirs for hydropower has been adjusted during the calibration together with the hydrological parameters to obtain a good fit to the observed flows at the gauging stations.

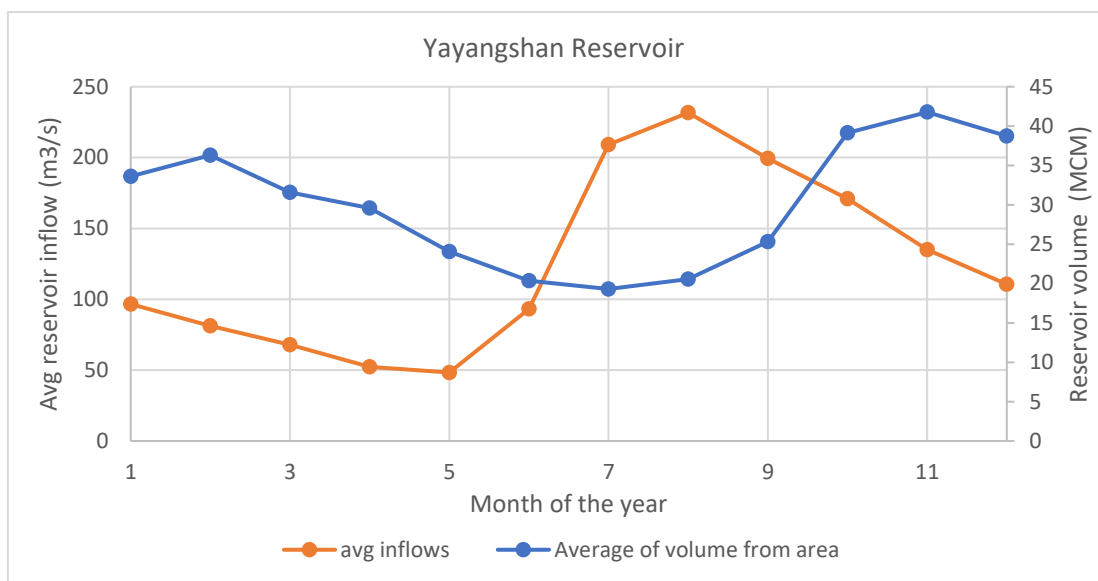


Figure 3.5 Average volume variation vs. reservoir inflows for the Yayangshan reservoir.

3.3 Forecasting of unregulated river flows

Historical transboundary inflows to Vietnam are known from daily (or more frequent) observations at Vietnamese gauging stations located at the international rivers close to the border. Assuming that the long-term historical observations will also be representative for future flows (a stationary climate) they have traditionally been used as probabilistic forecast of the future transborder flows the so called 'Extended Streamflow Prediction' (ESP) approach. However, under a changing climate or in situations where the upstream developments may impact the transboundary flow pattern, such approach is not necessarily valid.

As the Hong-Thai Binh basin is subjected to both these types of changes, hydrological rainfall modelling is necessary for predicting the climate impacts, and modelling of upstream reservoir operations is necessary to assess the impact of the most significant upstream developments on transboundary flows.

Furthermore, the ESP approach does, even in a non-changing climate, only result in an ensemble of *possible* future flows. It will not qualify if one member of the historical years will provide a better representation of the coming year than other historical years. By using hydrological rainfall runoff modelling driven by near real time information and seasonal forecast of meteorological forecasts, the forecasting system has a built-in qualification as it benefits from the best available meteorological prediction at the time the forecast is made.

Considering the lack of hydro-meteorological data from upstream ground stations, we apply DHI's Global Hydrological Model⁸ for the runoff simulations. The model is a distributed version of the NAM hydrological model that is already in use by NAWAPI's planning and forecasting departments. This model is tailored to operate based on globally available meteorological data from satellites and numerical weather models. Transfer functions have been developed converting information on land use, terrain, and soils to local model parameters.

While the EO routines specified in Sections 3.1 and 3.2 will provide the necessary reservoir characteristics for the reservoir simulation, the hydrological model is necessary for provision of the inflow to each reservoir which is necessary for assessing the reservoir regulation rules. The model will also provide the necessary inflows to the river between the upstream reservoir(s) and the border.

The Global Hydrological model (GHM)

The model is distributed on cells of approximately 0.1° by 0.1°, each of which represents an individual NAM rainfall runoff model with its own meteorological input and parameter set (generated from land use observations and terrain models). The high resolution makes it possible to generate parameters based on physical characteristics of an area. It thus eases the possibility to calibrate the model, looking a small scale.

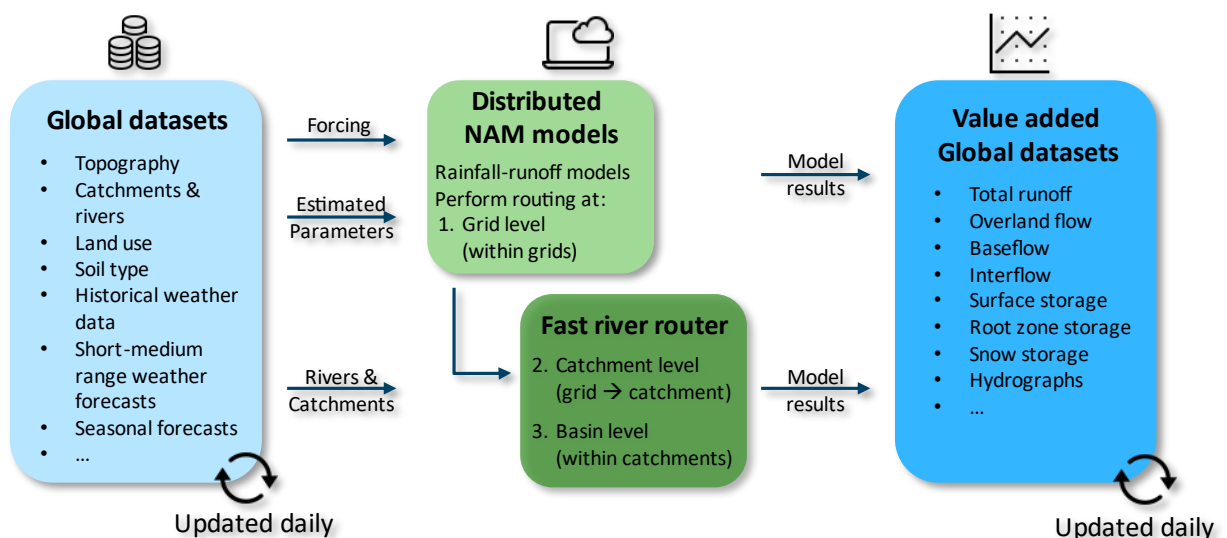


Figure 3.6 Conceptual diagram of DHI's Global Hydrological Model. The two green boxes describe the three water routing steps.

The Upper Hong-Thai Binh forecasting model will include different rainfall datasets to produce different simulations: hindcast, medium range forecast and seasonal forecasts. The rainfall data is retrieved from different sources, namely:

⁸ <https://www.dhigroup.com/data-portals/global-hydrological-model>

- ERA5: is the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of the global climate. Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using data assimilation.
- GFS: sourced from the National Centers for Environmental Prediction of the National Oceanic and Atmospheric Administration (NCEP NOAA), is a deterministic rainfall forecast for the next 16 days based on the NCEP Global Forecast System (GFS) Model.
- CFS: also source from NCEP NOAA, the Climate Forecast System (CFS) is a long-range numerical weather prediction and a climate model run to bridge weather and climate timescales. This dataset provides a 20-member ensemble forecast with 9-month lead time.

A description of the data retrieval process is available in section 3.4.

Figure 3.7 overleaf presents a diagram of the operational setup and interlinkage between hindcast and forecasting simulations. As illustrated, the historical simulations range from 1997 to ca. 7 days before real time. The hindcast has a timestep of 1 day. The latest model status will be extracted to hotstart the medium range forecast with 15 days lead time and at a 3-hourly timestep. During this simulation, the model status at real-time will be extracted (00:00 am of the simulation day) to provide hotstart for the seasonal forecast with 9-month lead time and at a daily timestep.

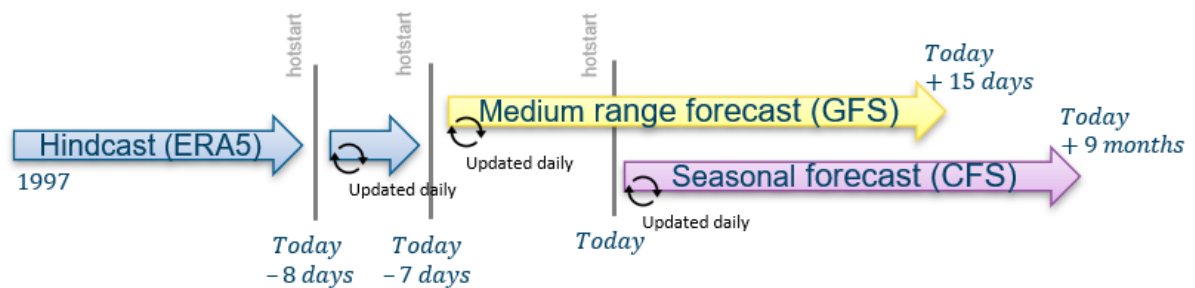


Figure 3.7 Time coverage for the operational setup of the Hong-Thai Binh GHM setup

Forecasts can be produced every day for NAWAPI's management support. The model simulates not only river discharge, but also rainfall-runoff modelling results, such as overland flow and soil moisture.

Figure 3.8 shows the difference in resolution between the natural catchment boundaries and the GHM model cells. It compares the size of the 10 by 10 km NAM model tiles, defined by the coloured corner points, and the green catchment polygons. The runoff from each catchment is computed as the area weighted runoff of the model cells intersecting it.

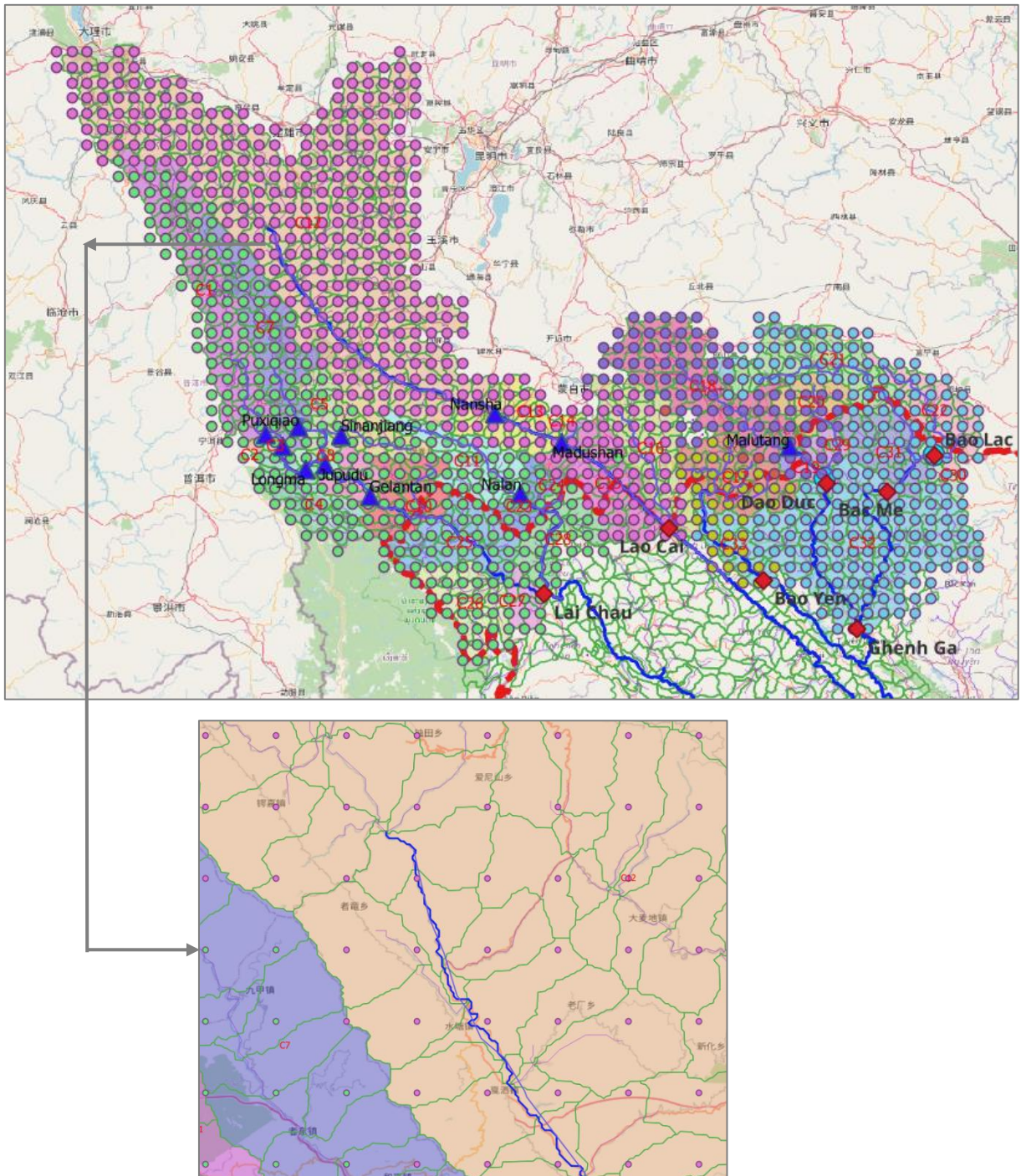


Figure 3.8 The resolution of the Upper Hong-Thai Binh Basin rainfall-runoff model, illustrated as corners of the single squared NAM model tiles, coloured according to the gauging station to which the tile contributes.

For this technical assistance, the area covering the Hong-Thai Binh River Basin in both China and Vietnam has been extracted and included in the operational system. The model has been calibrated on the upper parts of the river basin which has further improved the transboundary flow calculation.

The input to the rainfall-runoff model is the following:

1. Parameters derived for the AOI during calibration

2. Historical reanalysis data from ERA5, including daily resampled precipitation, 6-hourly resampled temperature (2m) and data for deriving daily potential evaporation (see section 3.5).
3. Deterministic medium range forecasts will be retrieved from the Global Forecast System (GFS) as 3-hourly precipitation and temperature.
4. Furthermore, daily resampled precipitation and 6-hourly temperature of the Climate Forecast System (CFS) will be used as seasonal forecast with 20 ensemble members and 9-month lead time.
5. Flow records from Vietnamese gauging stations close to the border will be used to calibrate the model and rainfall from Vietnamese ground stations.
6. Images of land use and soil types.
7. Locations of main reservoirs in China.
8. Assessed climate change factors for the Chinese parts of the basin.

The GHM provides both spatially distributed simulation at the grid level and at the natural catchment level is available.

NAM runoff at the grid level is extracted. NAM discharge as well as runoff at the catchment level are also extracted.

These GHM outputs are used in the MHB model as input to the reservoirs, making it possible to assess their operation rules. These outputs are also used in the MHB model to account for the intermediate catchments between the lowest reservoir in each river branch and the border to Vietnam.

After inclusion in the forecasting model the transfer functions have been adjusted (calibrated) to provide the best match to the transboundary flow observations. More information is provided on calibration in section 4.

3.4 Data retrieval and processing

Precipitation, evaporation, and temperature datasets are needed as inputs to DHI's GHM. Recent historic, current, and forecasts of each of these three types of data are required for real-time and forecasting operation. Routines have been developed to download ERA5 reanalysis data, see Figure 3.9, and NCEP forecasts operationally. DHI's frequencies for data retrieval are daily (ERA5 and CFS) and 6-hourly (GFS).

As mentioned previously, DHI's GHM's model domain is divided into a series of **tiles** each covering 10x10 degrees and containing 10,000 **cells** of 0.1x0.1 degrees. In addition, it allows the use of DFS2 format (MIKE by DHI software file format for time varying two-dimensional datasets) as a file format to define model inputs.

For this reason, conversion scripts were developed to resample any dataset to the resolution needed and store the data for each tile as DFS2.

Using a range of Python scripts to retrieve reanalysis data ERA5⁹ (using the Climate Data Store (CDS) Application Program Interface (API) - CDSAPI) and GFS data, or C# for retrieving CFS, all data is consistently shaped and stored in NetCDF-format, before being converted to GHM tiles. In the latter step, 321 DFS2-files are produced, where 1 is used for Hong-Thai-Binh's GHM (1 for each of the 20 ensemble members for CFS).

⁹ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels>

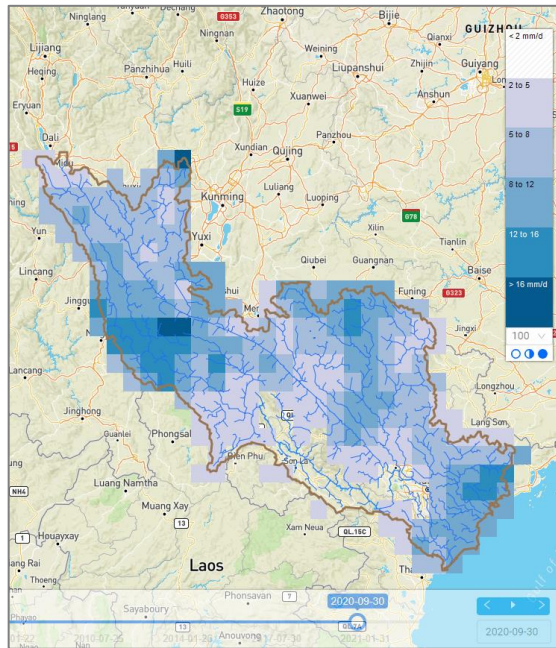


Figure 3.9 ERA5 Precipitation product covering the study area.

CFS is processed using an internal C# code. Using the latest 20 forecasts provided by NCEP, ensemble members are built matching the forecasts' time dimension.

A DHI software tool called the Flooddrought Toolbox¹⁰ was applied by the team. It is an internal toolbox for downloading and processing spatial data. Being based on python libraries such as xarray, rasterio and the like, it is optimized on retrieving global data with time dimension and using a variety of data sources. The command line interfaced toolbox, provides consistently shaped data within one dataset on the dimensions time, latitude, and longitude, which is stored in NetCDF-format.

The resulting global NetCDF-files are converted using python scripts based on the libraries xarray¹¹ and mikeio¹².

3.5 Priestly-Taylor Potential Evaporation approach

A Priestley-Taylor Potential Evaporation (PT PET) dataset is used in DHI's GHM's real-time and forecasting operation, described in McMahon et al. (2013). Python scripts have been implemented to calculate PT PET based on ERA5 reanalysis data with resolution 0.25° by 0.25°. The computed dataset is depicted in Figure 3.10.

While the real-time dataset is being updated daily, all forecast is using the monthly long time mean of the years 2000-2020.

PT PET uses historical reanalysis data from ERA5, including daily resampled 2 m dewpoint temperature, relative humidity, solar radiation downwards and 6-hourly resampled temperature (2 m) as input data. It furthermore uses static data of potential evaporation data from the TerraClimate¹³ dataset and the FAO dataset on ecological zones (FAO, 2001). And produces a daily, historical dataset with global coverage.

¹⁰ [https://github.com/DHI-GRAS/flooddrought_\(Restricted access\)](https://github.com/DHI-GRAS/flooddrought_(Restricted%20access))

¹¹ <https://docs.xarray.dev/en/stable/>

¹² <https://dhi.github.io/mikeio/>

¹³ The gridded TerraClimate datasets are available at 1/24th degree spatial resolution from 1958-2015. Hourly values were resampled to 4-times daily values, and yearly averages from the most recent historical data are used during forecasting.

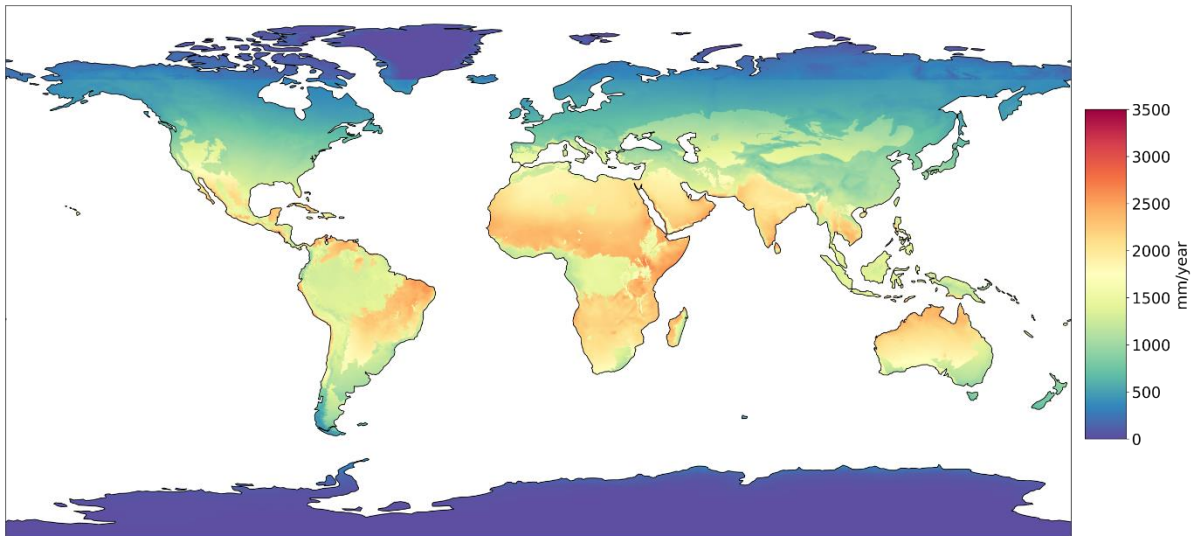


Figure 3.10 Yearly average evaporation from the Priestley-Taylor evaporation dataset

The dataset is calculated via the methods described in (McMahon, 2013) and the Priestley-Taylor formula:

$$E_{PT} = \alpha_{PT} \left[\frac{\Delta}{\Delta + \gamma} \frac{R_N}{\lambda} - \frac{G}{\lambda} \right]$$

where E_{PT} [$mm d^{-1}$] is the Priestley-Taylor evaporation, α_{PT} [-] is the Priestley-Taylor constant, Δ [$kPa \text{ } ^\circ C^{-1}$] is the slope of the saturation vapor pressure curve, γ [$kPa \text{ } ^\circ C^{-1}$] is the psychrometric constant, R_N [$MJ m^{-2} d^{-1}$] is net radiation, λ [$MJ m^{-3}$] is the latent heat of vaporization, and G [$MJ m^{-2} d^{-1}$] is the soil heat flux density.

The Priestley-Taylor evaporation model is a simplified version of the Penman-Monteith model, where the vapor pressure deficit and vapor pressure resistance are combined into the empirical constant α_{PT} . α_{PT} was selected based on the methods described by Suttleworth (1993) and varies spatially based on ecological zones described by FAO. A value of 1.74 was selected for the following zones: subtropical desert, subtropical dry forest, temperate desert, temperate steppe (except North America), tropical desert, tropical dry forest, tropical moist forest, and tropical shrubland. A value of 1.1 was selected for tropical forests due to the humid conditions. A value of 1.26 was assigned to all remaining zones.

A time series for the parameter Δ was calculated for the daily mean temperature in each grid cell. A temporally constant value for γ was selected for each grid cell based on elevation from the GMTED2010¹⁴ (Danielson, 2011). A time series for R_N was calculated for each grid cell based on the dew point temperature. A spatially and temporally constant value of $2.54 MJ m^{-2} d^{-1}$ for λ was selected, as this is the latent heat of vaporization for water at 20°C. G was assumed to be zero. The Priestley-Taylor equation is valid up to latitude 65°N. Values from the TerraClimate dataset are used from 65°N to 80°N. Yearly averages are used in forecasting.

3.6 Forecast bias-adjustment

A range of factors can cause errors in climate models. Those errors and biases can in models lead to unrealistic results, if the data is used uncorrected. There are multiple methods developed to overcome large biases in climate models. For correcting the forcing data of the GHM, DHI uses the statistical downscaling method quantile mapping.

¹⁴ https://www.usgs.gov/coastal-changes-and-impacts/gmted2010?qt-science_support_page_related_con=0#qt-science_support_page_related_con

This bias correction method happens in two steps:

1. Training: based on the cumulative distribution functions of reference and simulated data for the same period, adjustment factors are mapped.
2. Adjustment: appliance of adjustment factors from step 1 on the cumulative distribution functions of a simulation period.

Bias correction is performed for all forecasts of the GHM and for the most influential datasets as GHM input. As ERA5 has been identified to be the dataset, the GHM performs best on, it has been chosen to be the reference data for global bias correction of GHM forcing data. Furthermore, historical analysis data could be retrieved from NCEP for both GFS and CFS. Figure 3.11 shows archived GFS forecasts, which have been bias corrected using the approach described. The timeseries shown is the mean areal timeseries for cells being part of the GHM for Hong Thai Binh basin.

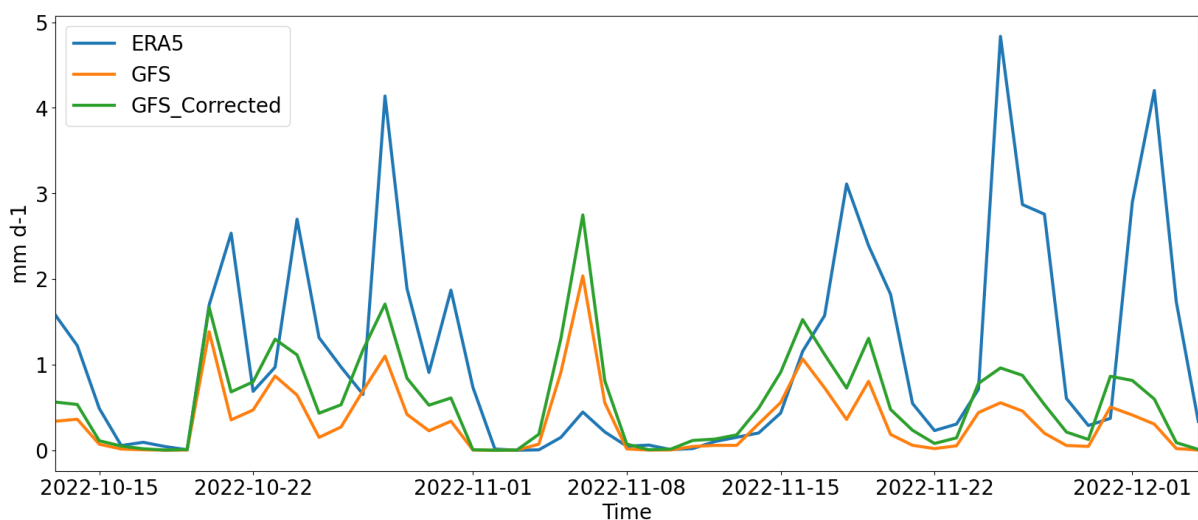


Figure 3.11 GFS precipitation before and after bias correction for the period 2022-10-13 - 2022-12-05, compared to the reference dataset of ERA5 precipitation

Table 3-1 shows the curve alignment of the archived GFS forecast with and without bias correction. Having a value of RMSE closer to 0 in the bias-corrected forecast indicates, that the precipitation can better be predicted using bias-correction. With R-square being closer to 1, the accuracy of the bias-corrected forecast is also considered to be better.

Table 3-1 Curve fitting parameters of archived GFS forecasts before and after bias-correction.

Indicator	GFS	GFS Corrected
RMSE	1.38	1.22
R-square	-10.54	-3.69

For the bias correction, Python scripts have been developed using xclim¹⁵. Based on xarray and dask, it is optimized on computing indices from large climate datasets. The forecasted data is bias corrected on global scale, before being processed to GHM forcing data as described in chapter 3.4.

¹⁵ <https://xclim.readthedocs.io/en/stable>

4 Model Calibration and Validation

4.1 Calibration summary

The modelling system has been calibrated against flow records from five Vietnamese flow stations which were selected according to the criteria of being located close to the national border and (together) monitoring all transboundary river flows.

The locations of the five Gauging stations are shown in Figure 4.1.

In general, the calibration was performed on flow data from 2015 to 2020 (incl.). This period has been selected as late in the observation series as possible to best reflect the present climate and the present level of flow regulation. For some of the gauging stations the records stopped earlier than 2015, in which cases calibration has been carried out on data from 2008-2013.

For the stations calibrated on the 2015-2016 period, model validation has been made on period 2008-2013 period.

The model calibration was based on simulations without any meteorologic ground data from China. On one hand, this makes the calibration more uncertain, on the other hand the EO observations provide an independent and very stable data source. This is a very important advantage for the final real time forecasting system, therefore it has been given priority in the development of the technology.

Hydrological GHM model and the MHB model were calibrated in combination: the GHM model providing inflows to the reservoirs and the intermediate river reaches while MHB simulating the regulation and the river flows down to the gauging stations.

As further commented upon below the calibration on the main rivers Da (at Lai Chau) and Hong (at Lao Cai) are good. These two stations represent 75% of the total (gauged) flow. Larger discrepancies are found for some of the eastern river which may stem from uncertainties in either the meteorological input or in the gauged flow.

The locations of the calibration stations are shown on Figure 4.1, and the details on the model calibration and its results are given in the sections that follow.

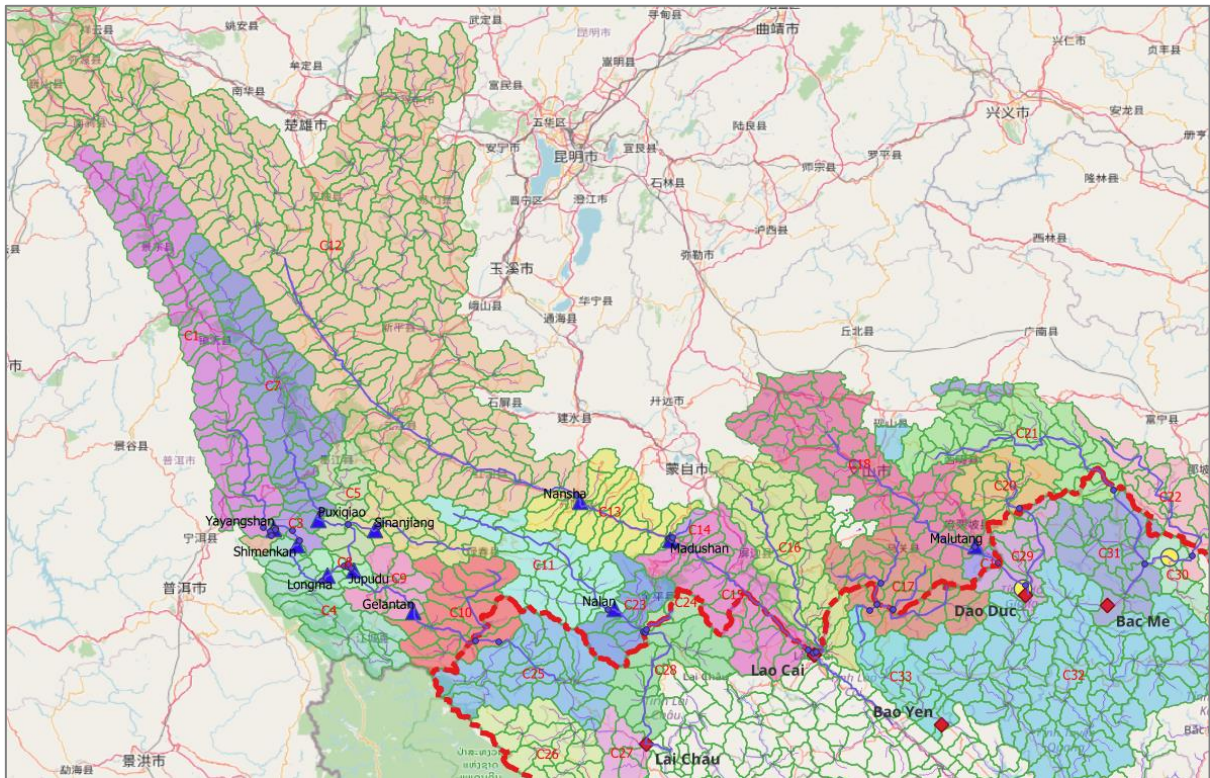


Figure 4.1 Locations of the Vietnamese flow gauging stations (red Diamonds) used for calibration and the two rain gauges (yellow circles) that was not used in the model.

4.2 Observation series and calibration/ validation periods

Meteorological observations

Meteorological ground observations have only been available from two stations at Dao Duc and Bao Lac. These two locations are quite far from the parts of the basin in China that generate the bulk of transboundary inflow to Vietnam. The data from these two stations have therefore not been used in the calibration.

Due to the relatively small surface areas of the reservoirs as compared to their throughflow, the direct rainfall and evaporation of the reservoirs has not been specifically accounted for in the model.

Flow observations used in the calibration

The locations of the five gauging stations from which the flow observations were used for the model calibration are shown in Figure 4.1 above. As the figure indicates these five gauging stations (together) monitor the discharge from all the upstream Chinese parts of the basin.

Data from two additional stations, one at Bao Lac (close to the border at the easternmost part of the basin) and one at GhenH Ga (a more downstream location deep within the Vietnamese territory) were available but not used in the calibration due to its large catchment fraction inside Vietnam and the possible implications of introducing the extensive water use in these parts of the basin.

It should be noted that the important Lai Chau station on the Da River has been relocated several times due to implications on flow observations from construction works. It is now located downstream of the new Lai Chau Reservoir, which has been active since 2016 and just upstream

of the reservoir lake of the large Son La reservoir. Although the station description claims that it was moved to keep it free of back water effects it is not clear if this is actually the case, as there seems to be a shift in the recorded flow regime from the older to the newer part of its record.

Therefore, the records used for calibration of the Da river are the records of inflow to the new reservoir back calculated from recordings of flow releases and volumes of the reservoir, a record that spans from 2016 to end of 2020.

The model validation of the Da basin is performed on older data from the Lai Chau gauging station (20008-2013).

Calibration procedure for the Upper Hong Thai Binh basin

Both regulation and hydrological parameters influence the calibration, but the hydrological parameters have been found to have the largest impact - at least on the main rivers.

A search in the NAWAPI MHB model of the Lower Hong Thai Binh Basin did not detect water abstractions in the same order as the river flows. Hence, it was decided to consider the water use as being a part of the basin characteristics i.e. embedded in the hydrological parameters.

The global GHM model from which the tiles covering the basin area were extracted, are already pre-calibrated and the NAM parameters pre-set using transfer functions that converts physical parameters, such as soil types, terrain slopes, vegetation etc. into the normal NAM parameters listed in Table 4-1. Hence, our point of departure is an already calibrated rainfall-runoff model.

Table 4-1 List of NAM parameters adjusted during the calibration.

Parameter	Unit	Description
U_{max}	mm	Maximum water content in the surface storage. U_{max} can be interpreted as including the water content in the interception storage, in surface depression storages, and in the uppermost few cm's of the soil
L_{max}	mm	Maximum water content in the lower zone storage. L_{max} can be interpreted as the maximum soil water content in the root zone available for the vegetative transpiration
CQOF	--	Overland flow runoff coefficient. CQOF determines the distribution of excess rainfall into overland flow and infiltration
TOF	--	Threshold value for overland flow. Overland flow is only generated if the relative moisture content in the lower zone storage is larger than TOF
TIF	--	Threshold value for interflow. Interflow is only generated if the relative moisture content in the lower zone storage is larger than TIF
TG	--	Threshold value for recharge. Recharge to the groundwater storage is only generated if the relative moisture content in the lower zone storage is larger than TG
CK_{IF}	seconds	Time constant for interflow from the surface storage. It is the dominant routing parameter of the interflow because $CK_{IF} \gg CK_{1,2}$
$CK_{1,2}$	seconds	Time constant for overland flow and interflow routing. Overland flow and interflow are routed through two linear reservoirs in series with the same time constant $CK_{1,2}$
CK_{BF}	seconds	Baseflow time constant. Baseflow from the groundwater storage is generated using a linear reservoir model with time constant CK_{BF}

Next, each NAM model cell was classified according to their highest contribution to each five sub-basins. A script was used to introduce a set of calibration factors (one set for each sub basin) allowing to increase or decrease each of NAM parameters (listed in Table 4-1) for all the NAM cells affiliated to that subbasin by a certain factor. This means that all cells in one sub basin will have the same parameter adjusted by the same factor (within certain maximum and minimum limits for each parameter).

In addition to the parameter adjustments, the precipitation and evaporation input were adjusted by a factor shared by all cells within the sub basin.

The calibration routine has proven quite efficient for users experienced with the calibration of the NAM model code which is available in the MIKE models.

Calibration of the flow regulation model MIKE HYDRO Basin

The abstractions from the upstream reservoirs were adjusted during the calibration in order to maintain as high extraction as possible during the dry season. When excess water is available (i.e. when the reservoir is at or above the flood control level) all excess water will either be turbinated or spilled. For the downstream flow forecasting it does not matter if the water passes the turbines or if it is spilled directly to the downstream river.

4.3 Calibration results

General

Table 4-2 provides an overview of the calibration through key statistics. It is seen that good calibrations have been obtained for at Lai Chau on the Da River and at Lao Cai on the Hong River providing 50% and 25% of the total flow observed at all the five gauging sites, respectively.

Table 4-2 Statistics for validation of the calibration for the five Sub basins.

Catchment	Calibration						Validation					
	Qobs Type	Period	Qsim	Qobs	Water balance Qsim/Qobs	R2	Qobs Type	Period	Qsim	Qobs	Water balance Qsim/Qobs	R2
Unit	-	Range	Cumecs	Cumecs	Fraction	-	-	Range	Cumecs	Cumecs	Fraction	-
Lai Chau	Res inflow	2016-2020	796	735	1.08	0.81	River GS	2008-2013	999	953	1.05	0.73
Lao Cai	River GS	2015-2020	429	424	1.01	0.61	River GS	2015-2020	339	388	0.87	0.65
Dao Duc	River GS	2015-2020	138	135	1.03	0.33	River GS	2015-2020	106	126	0.84	0.18
Bac Me	River GS	2008-2013	129	199	0.65	0.41						
Bao Yen	River GS	2008-2013	384	153	2.51	-6.01						

The water balances at the two locations are both reasonable (less than 10 % deviation) without adjustment in the meteorological input and the with quite high coefficient of determination (R2), a frequently used goodness-of-fit measure that ideally should be close to 1. Also, the model performance during the validation period is quite convincing for these two stations with acceptable water balances and good coefficients of determination.

From the average annual simulated and observed flows illustrated in Figure 4.2 overleaf it is seen that the model performs well during both the wet and the dry seasons.

The Dao Duc basin has a relatively large water balance error in the calibration period but performs substantially better during the validation with quite convincing result.

For the remaining two subbasins the water balances are poor. These are both small basins (according to total area and according to the basin area located in China). The fact that the discharge in one basin is significantly overestimated while the flow in the other is underestimated

might suggest certain problems with either the stations, missing information on other regulations or with the meteorological input.

Since the flow records for these two sub basins do not have data for the 2015-2020 period, the model has been calibrated on the 2008-2013 period, the period that has been used for validation of the other subbasins. No validation has been made of the calibration of these two stations.

Graphical comparisons of simulated and observed flow

In Figure 4.2 a comparison of the simulated and observed average monthly flow at the five gauging stations is shown. The figure gives an overview of the average annual and seasonal model performance at the various gauging stations.



Figure 4.2 Comparison of the simulated and observed average monthly flow showing the average annual and seasonal model performance at the various gauging stations (in m³/s).

It is noted that the comparison at the two main stations Lai Chau on the Da River and Lao Cai on the Hong mainstream, shows a good seasonal performance during the calibration as well as validation period. The Dao Duc station performs well during the validation period after having over-estimated the wettest three months during the calibration period.

The two small stations Bac Me and Bao Yen do not perform well. Neither have large areas upstream of the border. It should nevertheless be investigated if differences between the simulated and observed runoff at these stations may be explained by spatial rainfall variations, extensive water use in the basin or by other issues.

Figure 4.3 compares the daily simulated and observed inflow to the Lai Chau reservoir. While almost all the high flow peaks are detected by the model, the observed flow series include many lower peaks. We find it impressive that the ERA5 rainfall product seems to capture the majority of these larger flow peaks even though, it is actually a numerically simulated meteorological product. We find it likely that the large number of lower peaks, stems from regulation of the flows at the upstream reservoirs. As described earlier, the operation policies are not very obvious at the upstream dams, and it is likely the varying day to day demand for power is a more important factor for the day-to-day operation of the facilities, than the reservoir storage and the season. Hence the model cannot be expected to capture these operational smaller variations.

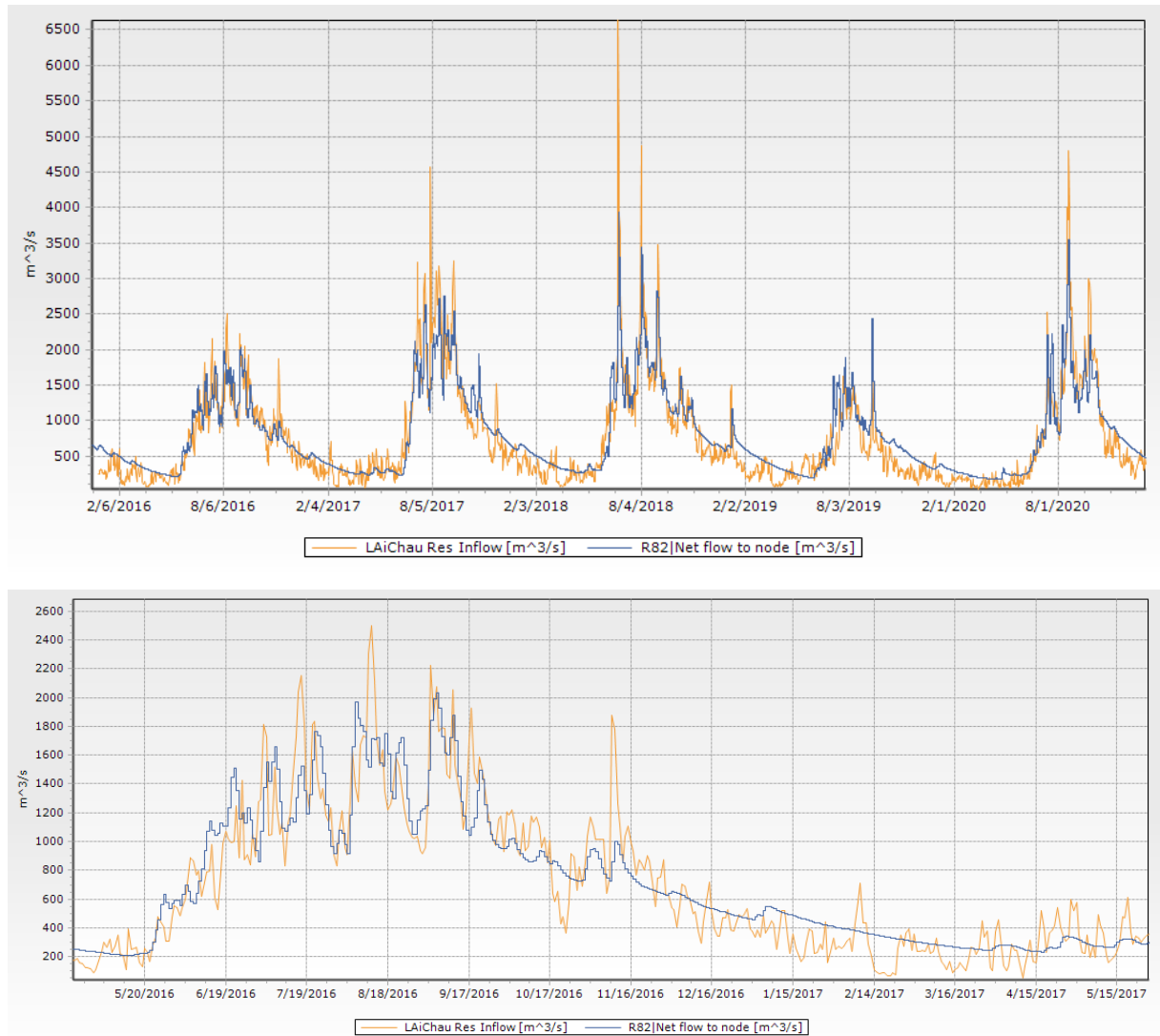


Figure 4.3 Simulated hydrograph (blue curve) and observed flow hydrograph (orange line) of inflows to the new Lai Chau Reservoir.

Figure 4.4 shows the same plots as Figure 4.3 but for the Lao Cai station on the main river. The same flow pattern as described above, is also found at this location. The simulation results are good although the baseflow recessions are not quite as good as in the Da river.

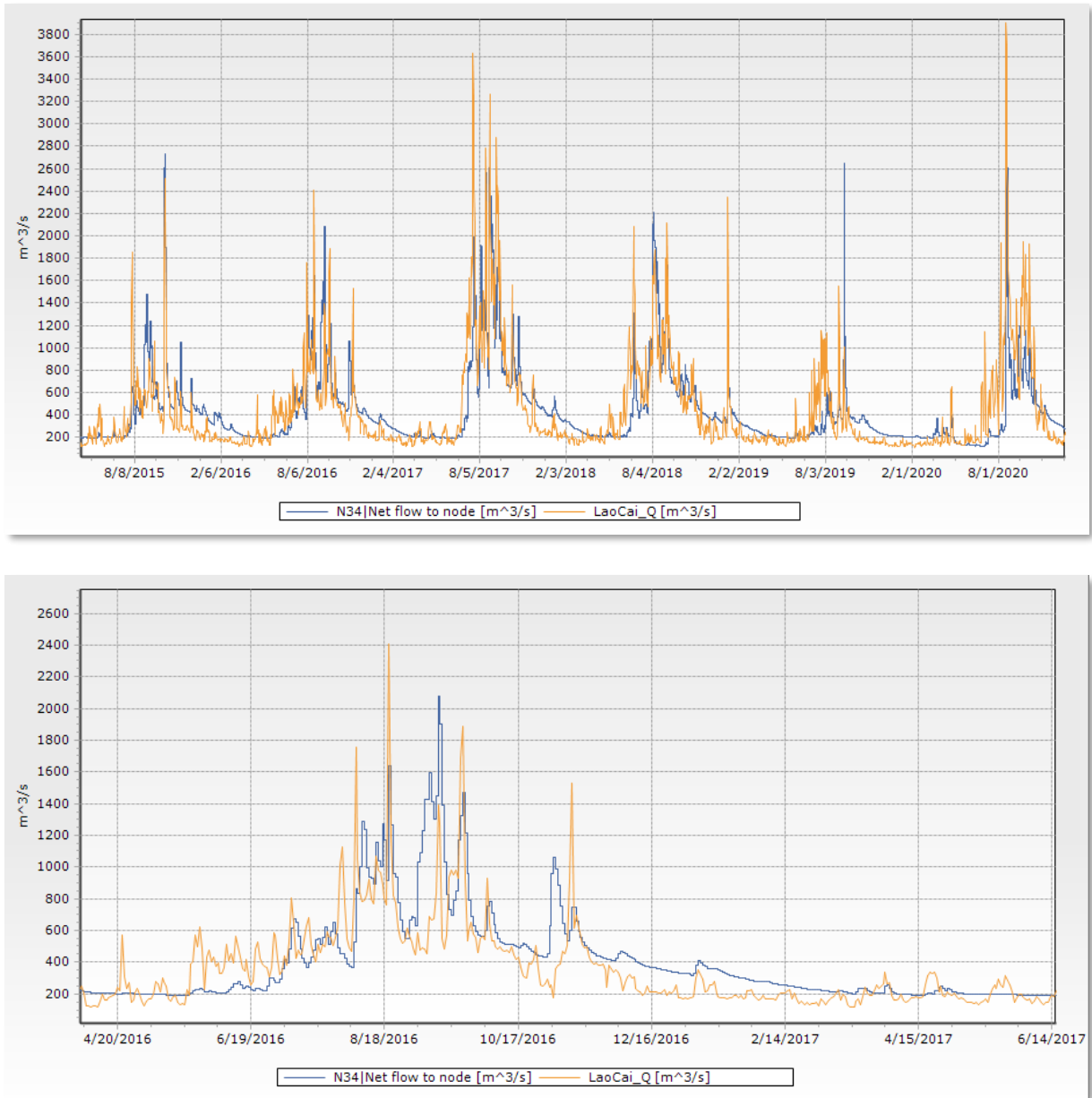


Figure 4.4 Simulated and observed cross boundary flow hydrographs at the Lao Cai gauging station in Vietnam.

5 Results Presentation and Reporting

5.1 Results presentation

With each system simulation, the forecasts are stored in the database, including all MIKE HYDRO Basin input and output data. Key outputs are daily discharge at border locations. For ease of inspection by the operator, it is recommended that favourite plots as well as maps are produced together with the CEWAFO system operators to ensure their reporting or monitoring needs are met. Maps are easily saved with the preferred locations for which discharge in m³/s can be plotted.

In addition to the above, the following indicators will be automatically calculated and updated with every run: seasonal river flow, monthly river flow and 10-day river flow. Should additional indicators be required, we will assist the system operators with the configuration.

5.2 Operational report

To inform the operators of the seasonal forecasting system, a reporting framework was set up. This automatically sends the operator(s) an email with a description of the performance of each of the jobs run. An exemplary can be seen in Figure 5.1. The email starts with a time tag to indicate when the report was generated, followed by the name of the MIKE Workbench database as the amount of free disk space. Next, two tables are included with the result of system runs, namely:

- Jobs table: indication of number of jobs executed in the last 48 hours, job name, time of completion of run, whether it was executed successfully or returned an error.
- Simulations table: name of the scenario the run corresponds to (the system runs seasonal and mead), time of completion of run, approval status and time of forecast (tof).

The detail of the system will be captured in the technical manual to be produced during Output 5 of the technical assistance.



Figure 5.1 Report sent via email, listing the last job instances and simulations in a database.

6 Climate Change

Climate change should be considered when planning water resources allocation to ensure robust decision making for long-term mitigation of flood and drought impacts under a future changing climate. In this project, we have estimated climate delta change factors for the entire Hong Thai Binh basin including the upstream catchment areas located in China. Climate change estimates are produced nationally for the catchments within Vietnam. Therefore, using global sources, data has been made available for the remaining catchment areas in China. This data may be applied by NAWAPI together with their hydrological models to carry out their assessment of the impacts of climate change for the different scenarios representing a range of possible futures.

6.1 Future scenarios

The Intergovernmental Panel on Climate Change (IPCC) released its Sixth Assessment Report (AR6) from Working Group 1 on the physical science basis of climate change in August 2021. The report shows that human-induced climate change is already affecting many weather and climate extremes in every region across the globe, and there is increased evidence of observed changes in extremes such as heatwaves, heavy precipitation, and droughts.

In the report, a set of five new emissions' scenarios are used to explore the climate response to a range of greenhouse gas, land use and air pollutant futures. The emissions' scenarios cover the range of possible future development of anthropogenic drivers of climate change found in the literature and represent a broader range of futures than was assessed in the previous IPCC Fifth Assessment Report (AR5) published in 2014.

The five emissions scenarios can be understood in terms of a combination of two pathways, the Representative Concentration Pathways (RCPs) which set pathways for greenhouse gas concentrations and the amount of warming that could occur by the end of the century (as used in the previous IPCC AR5 report) and the new Shared Socioeconomic Pathways (SSPs) with a variety of socio-economic mitigation and adaptation challenges that set the stage on which reductions in emissions will – or will not- be achieved.

The five scenarios are named SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5 where the first digit refers to the SSP storyline for the socio-economic mitigation and adaptation challenges represented, and the second and third digits are the RCP climate forcing, as described below:

SSP1-1.9 = SSP 1 and RCP 1.9

SSP1-2.6 = SSP 1 and RCP 2.6

SSP2-4.5 = SSP 2 and RCP 4.5

SSP3-7.0 = SSP 3 and RCP 7.0

SSP5-8.5 = SSP 5 and RCP 8.5

The SSP storylines are based on narratives describing alternative socio-economic developments, including sustainable development, regional rivalry, inequality, fossil-fuelled development, and middle-of-the-road development (O'Neill et al., 2015, and Riahi et al., 2017) as illustrated in Figure 6.1. SSP2 represents a baseline scenario where the world follows a path in which social, economic, and technological trends do not shift markedly from historical patterns.

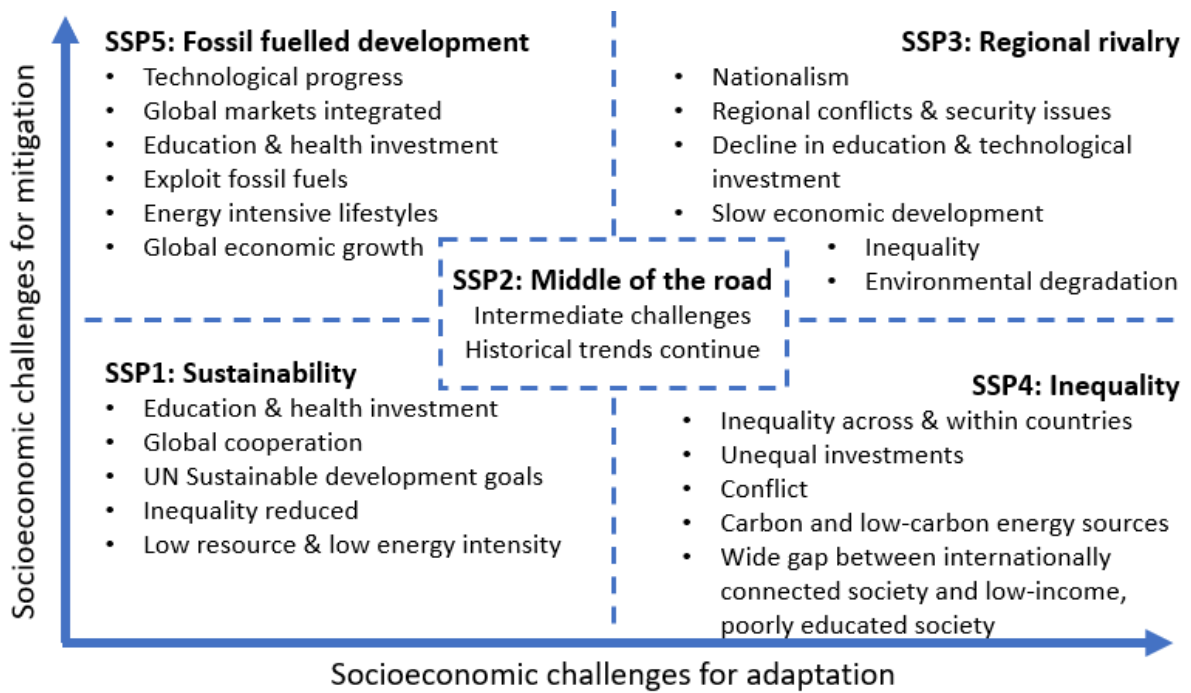


Figure 6.1 A summary of the Shared Socioeconomic Pathways (SSPs) that are used to form the latest IPCC AR6 climate scenarios (O'Neill et al., 2015).

The SSP narratives are combined with the RCP greenhouse gas concentration pathways, and the five resulting scenarios range from scenarios with high and very high greenhouse gas emissions with CO₂ emissions that roughly double or triple (scenario SSP3-7.0 and SSP5-8.5, respectively), to scenarios with very low and low greenhouse gas emissions with CO₂ emissions declining to net zero around 2050 and 2075 (scenario SSP1-1.9 and SSP1-2.6, respectively). A further mid-range scenario (SSP2-4.5) has intermediate greenhouse gas emissions and CO₂ emissions remaining at current levels until around 2050, as illustrated in Figure 6.2.

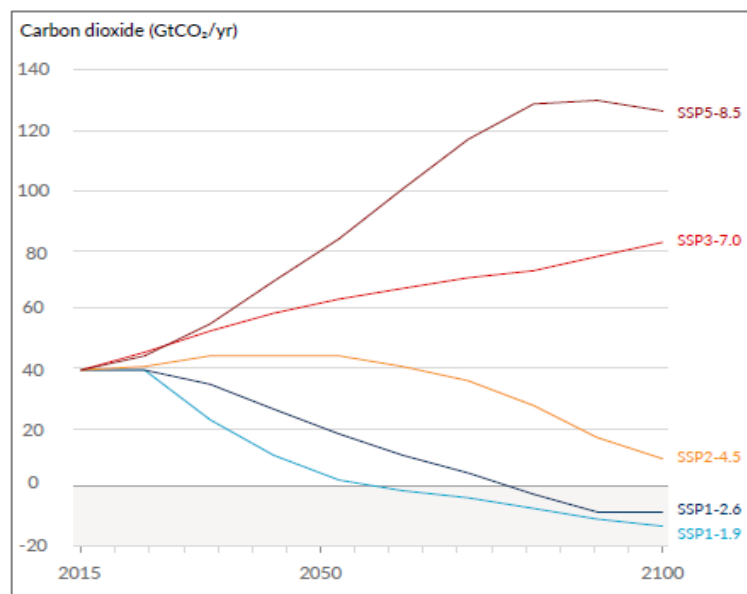


Figure 6.2 Future annual emissions of CO₂ in the five new IPCC scenarios. Future emissions cause future additional warming, with total warming dominated by past and future CO₂ emissions. Source: IPCC AR6 WG1 Summary for Policy Makers Box SPM.1.

A summary of the five scenarios is shown in Figure 6.3. In SSP1-1.9 global warming is limited to less than 1.5 °C meeting the Paris Agreement of limiting global warming to well below 2 °C.

Whilst the IPCC AR6 report from August 2021 does not assess the likelihood of the scenarios, SSP1-1.9 is considered unlikely in some studies and is mainly used to inform an assessment of the impacts of meeting the Paris Agreement. SSP1-2.6 is therefore considered a best-case scenario in some studies. In SSP5-8.5 a push for economic and social development is coupled with the exploitation of abundant fossil fuel resources and the adoption of resource and energy intensive lifestyles around the world. This scenario is considered unlikely in some studies due to the recent developments in the energy sector towards alternative energy sources.

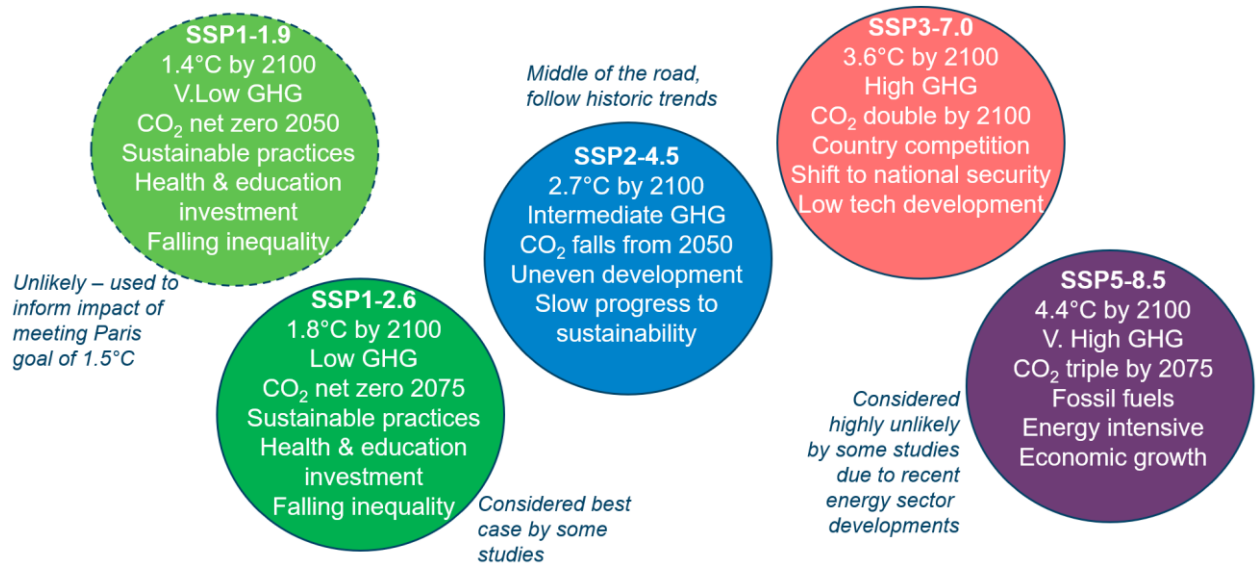


Figure 6.3 Summary of the five emissions scenarios that inform the latest IPCC AR6 report.

To inform the IPCC AR6 report, the set of five scenarios drives climate model projections of changes in the climate system which also account for solar activity and background forcing from volcanoes. Results over the 21st century are provided for the near term (2021-2040), mid-term (2041-2060) and long term (2081-2100) time periods. The climate models used to form the basis of the IPCC AR6 results are from the Coupled Model Intercomparison Project Phase 6 (CMIP6) in which multiple climate modelling groups around the world run global climate models with agreed input parameters.

In this project, we have used the outputs from these global climate models run with the five new emissions scenarios to produce data for NAWAPI to assess the impacts of climate in the entire Hong Thai Binh River basin including the Chinese territory.

6.2 Calculation of Climate Change Factors

The CMIP6 modelling group includes over 100 Global Climate Models (GCMs) from over 50 modelling centers around the world. A quality-controlled subset of CMIP6 models and data are made available through the Climate Data Store for users of the Copernicus Climate Change Service¹⁶. In this project, we used this quality-controlled subset and selected all the GCMs which have the required model output data for all five scenarios. The required data is precipitation and minimum / maximum air temperature, where the air temperature data is then used to calculate Potential Evapotranspiration (PET) using the Kay and Davies (2008) equation.

The GCMs with the available required data for all five scenarios are listed below (note that some models did not contain outputs for SSP1-1.9, but were selected because they contained outputs for all other scenarios):

- EC-Earth3-Veg-LR (Europe)

¹⁶ <https://cds.climate.copernicus.eu/cdsapp#!/dataset/projections-cmip6?tab=overview>

- CNRM-ESM2-1 (France)
- MIROC6 (Japan)
- GFDL-ESM4 (USA)
- UKESM1-0-LL (UK, South Korea, NZ)
- MPI-ESM1-2-LR (Germany) (*does not include SSP1-1.9)
- CCESS-CM2 (Australia) (*does not include SSP1-1.9)
- CanESM5-CanOE (Canada) (*does not include SSP1-1.9)
- INM-CM5-0 (Russia) (*does not include SSP1-1.9)

The outputs from these GCMs were used to calculate monthly change factors for temperature, precipitation and potential evapotranspiration (PET). The GCMs all have similar, but different, spatial resolutions of between 1 and 2 degrees latitude/longitude. The model outputs were, therefore, regridded to match the resolution of the model with the highest resolution so that further calculations could be made at a common spatial resolution of 1 degree latitude and 1.25 degrees longitude.

The temporal resolution of the GCM output data is available at daily and monthly timesteps for a historical time period (1850 – 2014) and a projected time period (2015 – 2100), and monthly data was used in this project as this often provides a better understanding of the trends. To calculate the change factors, a baseline period of 25 years from 1990 – 2014 was selected from the historical data and used to provide the static state assumption for the current climate. Future time periods then focused on the same time periods used in the IPCC AR6 report of short term (2021 – 2040), mid-term (2041 – 2060) and long-term (2081 – 2100). Climate change factors were calculated by comparing the mean monthly temperature, PET and precipitation in the future year time periods against the baseline time period for each GCMs run with each of the scenarios.

The result is an ensemble of monthly change factors for temperature, PET and precipitation for each of the five scenarios, where the ensemble members are the different GCMs. These change factors for the Hong Thai Binh River basin have been delivered to NAWAPI, namely, the Department of Water Resources Planning and the Center for Water Resources Warning and Forecast (CEWAFO).

7 Next Steps

The following activities corresponding to the completion of Output 4 as well as Output 5 will be carried out in the next part of the TA implementation.

Activity 4.3: Report and verification of information on transboundary water resources in the Hong-Thai Binh river basin

The system performance will be tested against available relevant observations. The performance indicators may include relative bias, mean absolute relative error, skill scores and others. A report will detail the outcome of this phase. A test protocol has been outlined in Appendix A and can be used to reporting both the systems and the model's robustness.

During this period (March to October 2023), it is expected that performance of the system will be monitored every month and further adjustments/calibration will be performed, if required. The experts at CEWAFO will be utilising the system and able to suggest improvements/adjustments.

Activity 5.1: Development of the technical manual for use of the selected technologies

A first version of the Technical Manual will be issued in March 2023 and then will be further refined at the end of the project, after the test and validation period, to ensure that improvements resulting from the feedback and lessons learned through operation of the system have been integrated.

Activity 5.2: Capacity needs assessment and development of the capacity building program for use of the selected technologies

The Consultant will discuss capacity needs with NVWATER and NAWAPI over online demonstration sessions for adopting the selected technologies, a selection of trainees will be carried out, and the training program and respective materials prepared accordingly.

Activity 5.3: Training of government bodies and stakeholders for use of the selected technologies

The Consultant will conduct a training course based on the findings of Activity 5.2. It is expected that the face-to-face training to take place in March 2023.

8 References

- /1/ Danielson, J.J., & Gesch, D.B. (2011). Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073, 26 p
- /2/ Druce, D.; Tong, X.; Lei, X.; Guo, T.; Kittel, C.M.M.; Grogan, K.; Tottrup, C. An Optical and SAR Based Fusion Approach for Mapping Surface Water Dynamics over Mainland China. *Remote Sens.* 2021, 13, 1663. <https://doi.org/10.3390/rs13091663>
- /3/ Forest Resources Assessment Program, FAO. (2001). Global Ecological Zoning for the Global Forest Resources Assessment 2000. Final Report. Chapter 3: Concept and Classification. Rome, 2001.
https://www.fao.org/3/ad652e/ad652e07.htm#P412_20600
- /4/ Lehner, B., Reidy Liermann, C., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J., Rödel, R., Sindorf, N., Wisser, D. (2011): High- resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment* 9: 494–502. doi: 10.1890/100125.
www.globaldamwatch.org or <https://sedac.ciesin.columbia.edu/pfs/grand.html>
- /5/ McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., & McVicar, T. R. (2013). Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, *Hydrol. Earth Syst. Sci.*, 17, 1331–1363. <https://doi.org/10.5194/hess-17-1331-2013>. Supplementary material: <https://hess.copernicus.org/articles/17/1331/2013/hess-17-1331-2013-supplement.pdf>.
- /6/ O'Neill, B.C., E. Kriegler, K.L. Ebi, E. Kemp-Benedict, K. Riahi, D.S. Rothman, B.J. van Ruijven, D.P. van Vuuren, J. Birkmann, K. Kok, M. Levy, W. Solecki (2015). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, doi:10.1016/j.gloenvcha.2015.01.004.
- /7/ Riahi, K. et al. (2017) The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: an overview. *Glob. Environ. Change* 42, 153–168.
- /8/ Suttleworth, W.J. (1993). Chapter 4 Evapotranspiration. In D.R. Maidment (Ed.). *Handbook of Hydrology* New York: McGraw-Hill Inc.

Appendix A Test protocol – check list

The established technology will be thoroughly tested during Activity 4.3. As a recommended template for testing, a test protocol is presented below. It outlines testing requirements and strategy by including a task's component, action, and success criteria. Furthermore, all tasks are evaluated on their importance in three categories, low, medium, and high. It is expected that the success criteria of each action is checked during the review of the system.

	Component	Action	Success criteria	Importance	Checked?
Hindcast	Input data	Retrieve latest rainfall data, hindcast	Job runs successfully daily	Medium	
	Input data	Retrieve latest temperature data, hindcast	Job runs successfully daily	Medium	
	Input data	Retrieve latest potential evaporation data, hindcast	Job runs successfully daily	Medium	
	NAM execution	Run NAM module, hindcast	Job runs successfully daily	Medium	
	Routing execution	Run Routing module, hindcast	Job runs successfully daily	Medium	
	Hotstart handing	Pre-processing and post-processing, hindcast	Job runs successfully daily	Medium	
	Data postprocessing	Store data and extract model timeseries, hindcast	Job runs successfully daily	Medium	
	Data import to database	Import model timeseries to MIKE Workbench database, hindcast	Job runs successfully daily	Medium	
Medium range forecast	Input data	Retrieve latest rainfall data, gfs forecast	Job runs successfully daily	High	

	Component	Action	Success criteria	Importance	Checked?
	Input data	Retrieve latest temperature data, gfs forecast	Job runs successfully daily	High	
	Input data	Retrieve latest potential evaporation data, gfs forecast	Job runs successfully daily	High	
	NAM execution	Run NAM module, gfs forecast	Job runs successfully daily	High	
	Routing execution	Run Routing module, gfs forecast	Job runs successfully daily	High	
	Hotstart handing	Pre-processing and post-processing, gfs forecast	Job runs successfully daily	High	
	Data postprocessing	Extract model timeseries, gfs forecast	Job runs successfully daily	High	
	Data import to database	Import model timeseries to MIKE Workbench database, gfs forecast	Job runs successfully daily	High	
	MHB Model	Run model, gfs forecast	Job runs successfully daily	High	
Seasonal forecast	Input data	Retrieve latest rainfall data, cfs forecast	Job runs successfully daily	High	
	Input data	Retrieve latest temperature data, cfs forecast	Job runs successfully daily	High	
	Input data	Retrieve latest potential evaporation data, cfs forecast	Job runs successfully daily	High	

	Component	Action	Success criteria	Importance	Checked?
	NAM execution	Run NAM module, cfs forecast	Job runs successfully daily	High	
	Routing execution	Run Routing module, cfs forecast	Job runs successfully daily	High	
	Hotstart handing	Pre-processing and post-processing, cfs forecast	Job runs successfully daily	High	
	Data postprocessing	Extract model timeseries, cfs forecast	Job runs successfully daily	High	
	Data import to database	Import model timeseries to MIKE Workbench database, cfs forecast	Job runs successfully daily	High	
	MHB Model	Run model, cfs forecast	Job runs successfully daily	High	
Maintenance	Maintenance	Clean up on database and disk	Job runs successfully daily	High	
	Data base backup	Create backup and handle archive	Job runs successfully weekly	High	
	Reporting	Send email with job and simulation overview	Job runs successfully daily	Medium	
Model output	Model inputs	Timeseries extraction and accumulation	Data is processed and used by the model.	High	
	Model inputs	Naming of timeseries	Naming convention is met, and data can be found by scripts	High	

	Component	Action	Success criteria	Importance	Checked?
	Model outputs	Compare model outputs with observed data	Performance indicators are met	High	