

Master's thesis

Forecasting River Discharge for Hydropower Production in Central Himalaya, Nepal

Jacob Qvam Skavang

Hydrology and Glaciology
60 ECTS study points

Department of Geosciences
Faculty of Mathematics and Natural Sciences

Spring 2023



Jacob Qvam Skavang

**Forecasting River Discharge for
Hydropower Production in Central
Himalaya, Nepal**

Supervisors:
Olga Silantyeva
Lena Merete Tallaksen
Kristoffer Aalstad

Abstract

Write an abstract

Abstract

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Aim and objectives	1
1.3	Scope	1
1.4	Thesis outline	1
2	Background	3
3	Study area	5
3.1	Geography	5
3.1.1	Hypsograph	6
3.1.2	Land use	6
3.2	Climate	6
3.2.1	Temperature	6
3.2.2	Prcipitation	6
3.2.3	Wind	6
3.2.4	Relative humidity	6
3.2.5	Radiation	6
3.2.6	Snow	6
3.2.7	Hydrology	6
4	Data and methods	9
4.1	Meteorological forcing data	9
4.1.1	Reanalysis and regional climate data	9
4.2	Topographical Data	11
4.3	Land Cover Data Sets	11
4.4	Validation datasets	11
4.4.1	Observed river discharge	11
4.4.2	Snow Products	12
4.5	Pre-processing of data	12
4.5.1	TopoScale	12
4.5.2	Catchment Discretization Technique	12
4.6	Hydrological Modeling Framework	12
4.6.1	The architecture of Shyft	12
4.6.2	Conceptual model	15
4.6.3	Spatial interpolation	16
4.6.4	Hydrological model	16
4.6.5	Model simulation	16
4.6.6	Model calibration	16

Contents

4.6.7	Water balance estimation	16
4.6.8	Model validation	16
4.6.9	Model performance evaluation	16
5	Results	19
5.1	Forcing data analysis	19
5.2	Evaluation of discharge simulation using different datasets	19
5.3	Water balance analysis	19
6	Discussion	21
7	Conclusions and outlook	23
A	Downloading data from the CDS API	25

List of Figures

3.1	The Budhi Gandaki catchment	5
4.1	Shyft arcitechture. The figure is an adaptation of the figure in [5], and was made using Miro.	13
4.2	Overview of the main Shyft API types. API types used for simulation are shown to the left of the blue dashed line, while calibration API types are found to the right of the blue dashed line. The figure shows an example of the PTSSK model stack (Priestly-Taylor-Skaugen-Snow-Kirchner). The figure is an adaptation of the figure in [5], and was made using Miro. . . .	14
4.3	Evalutation of multiple configurations. The figure is made using Miro, and it is an adaptation of the figure in [5]	17
4.4	Model Stacks components. The figure is made using Miro, and it is an adaptation of the figure in [5]	17

List of Figures

List of Tables

4.1	Summary of the forcing datasets	9
4.2	WFDE5 variables	10
4.3	SRTM 1 Arc-Second Global	11
4.4	Forcing data and validation data	16
4.5	Model calibration parameters	16

List of Tables

Preface

Here comes your preface, including acknowledgments and thanks.

Preface

Chapter 1

Introduction

1.1 Motivation

The water supply from the Himalayan and Tibetan Plateau is of great importance for millions of people [4]. The water from this region provides drinking water, supports agricultural demands, is used for hydropower generation and other agro-economic activities [22]. The countries in the High Mountain Asia (HMA) region have a huge potential for developing hydropower making the shift towards a greener economy more feasible. The Bundhi-Gandaki catchment is located in the Gorkha district of Nepal. The catchment has a high mean annual precipitation of 1495 mm and a substantial spatial variation in elevation [8]. For example, the lowest point is at 479 meters above sea level (m.a.s.l.) at Arughat hydrological station and the highest at Manaslu mountain 8163 m.a.s.l. The combination of high precipitation rate and steep gradients in elevation makes the region of superior interest for hydropower. The installed capacity is 1200 megawatt (MW) with an average energy generation of 3383 gigawatt-hours (GWh). However, the hydropower potential is dependent on the climatic conditions in precipitation, evaporation, temperature and snow/ice in the catchment [10]. Climate change has serious implications for hydropower production [7]. Changing rainfall pattern and increased temperatures will affect power generation. The retreat of glaciers, expansion of glacial lakes and changes in the seasonality and intensity of rainfall is identified as factors that will affect the power generation in the future [7]. The region is also threatened by Glacial Lake Outburst Floods (GLOFs), that can cause devastating floods downstream [7]. In addition, the extremely heterogeneous topography of the catchment makes it challenging to get accurate measurements of meteorological variables [24]. The scarcity of data makes the discharge predictions, which lies the foundation for hydropower, prone to errors. It is impossible to make observations at all high levels, especially for snow. For this reason, satellite observations and reanalysis data sets important for decision-making processes. Furthermore, different precipitation patterns on the leeward and windward sides of the catchment needs to be taken into account in simulations.

1.2 Aim and objectives

1.3 Scope

1.4 Thesis outline

Chapter 1. Introduction

Chapter 2

Background

Chapter 2. Background

Chapter 3

Study area

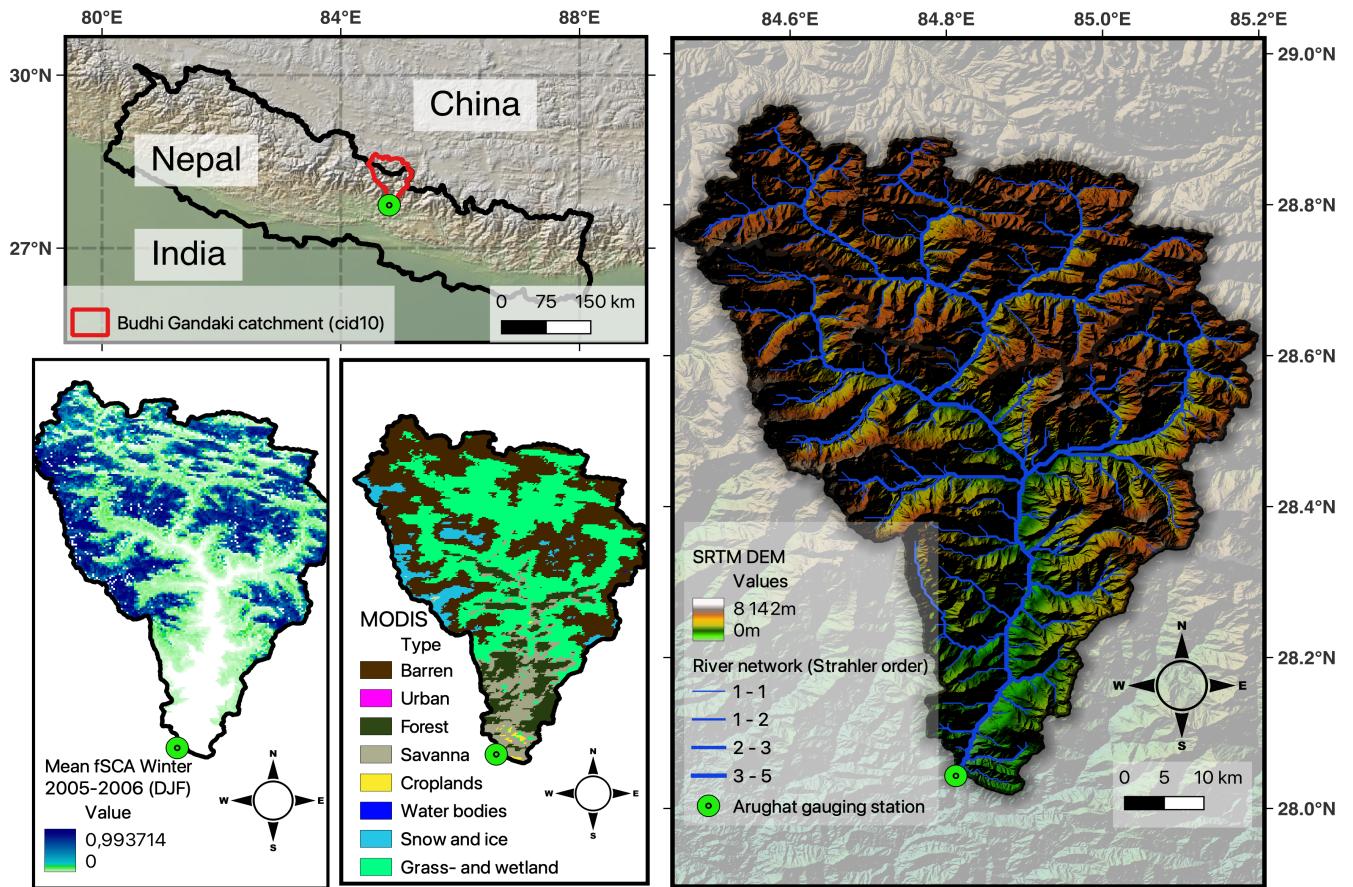


Figure 3.1: The Budhi Gandaki catchment

3.1 Geography

The Budhi Gandaki catchment is located in central Nepal (between $27^{\circ}50'$ and $29^{\circ}00'$ N latitudes, and $84^{\circ}30'$ and $85^{\circ}10'$ E longitudes [17]. The catchment is part of the larger Narayani River Catchment originating in Tibet to the north, and is encompassing parts of the Gorkha and Dhading districts of Nepal [8]. The Budhi Gandaki catchment

is one of the main tributaries of the Trishuli River in Nepal which in turn is a major tributary to the Gandaki River that meets the Ganges in India [15]. The Budhi Gandaki catchment is about 113 km long and between 15 and 30 km wide [17]. The catchment area upstream of the gauging station at Arughat bazaar is 3871.38 km^2 [15]. About 56 % of the catchment area is in Nepal and 44 % lies in China. The catchment elevation ranges from green hills (479 m.a.s.l.) at the gauging station in the south to the snow covered Manaslu mountain peak (8029 m.a.s.l.) [15]. The catchment contains several climatic zones due to the extreme elevation changes within the catchment. The climate changes from tropical (below 1000 m.a.s.l.) to sub-tropical zone (1000 – 2000 m), temperate zone (2000–3000 m), subalpine zone (3000–4000 m) and alpine zone (4000–5000 m) [15]. The zone above 5000 m is characterized by permanent snow and no human habitation.

3.1.1 Hypsograph

3.1.2 Land use

3.2 Climate

3.2.1 Temperature

The mean temperature ranges from 6 °C during winter to 35 °C in summer measured near the Budhi Gandaki Hydropower Project proposed dam site [8].

3.2.2 Precipitation

Long-term annual catchment precipitation (1983-2012) is 1301 mm (maximum of 278 mm in July and minimum of 11 mm in November [17]). The annual precipitation varies from greater than 2500 mm at Arughat gauging station to less than 700 mm in the Tibetan part of the catchment [17]. The hydrological regime is greatly influenced by the seasons [2]. The four seasons are defined as winter (December to February), pre-monsoon (March to May), monsoon (June to September) and post-monsoon (October to November) [1, 29]. The catchment receives about 80 % of the annual precipitation during the monsoon season [15]. The maximum relative humidity is also reached during the monsoon season. The rainfall decreases from west to east during this period, while the greatest rainfall intensity is occurring on the south facing slopes of the mountains. The catchment is among the tributaries to the Narayani river that are monsoon fed (those in the middle and high mountain regions), in contrast to the glacier and snow melt fed (those originating in a higher Himalayan region) [2].

3.2.3 Wind

3.2.4 Relative humidity

3.2.5 Radiation

3.2.6 Snow

3.2.7 Hydrology

The observed stream flow at Arughat seem to closely follow the precipitation pattern [18]. The mean annual flow near the Budhi Gandaki Hydroelectric Project (BGHEP) dam site is estimated at 222 m/s [9].

3.2. Climate

Percentage of Seasonal PPT, show pie chart

Show annual trends

Monthly average rain days, temperature

Radiation duration

Rating curve

Maximum discharge pr month

Chapter 3. Study area

Chapter 4

Data and methods

4.1 Meteorological forcing data

Forcing data are used to drive the hydrological model. Shyft needs precipitation, relative humidity, temperature, wind speed and incoming shortwave radiation as input data. Table 4.1 gives a summary of the forcing data used in this study.

Table 4.1: Summary of the forcing datasets

Forcing dataset	Data period	Spatial resolution	Temporal resolution	n	Reference
WFDE5	1990-2019	0.5° × 0.5°	Hourly	9	[6]
ERA5 + TopoScale	1999-2015	0.5° × 0.5°	Hourly		[13]

4.1.1 Reanalysis and regional climate data

WFDE5

The bias-adjusted ERA5 reanalysis data for impact studies (WFDE5) is generated through applying the WATCH Forcing Data methodology to surface meteorological variables from the ERA5 reanalysis [6]. Reanalysis data is a combination of model data and worldwide observations into a globally complete and consistent dataset using the laws of physics [6]. The WFDE5 dataset contains 11 variables with an hourly temporal resolution on a regular longitude-latitude 0.5 degree grid. The dataset has a global coverage, but is only defined at land and lake points. The WFDE5 dataset is spanning from January 1979 to the end of 2019, and has been adjusted using an elevation correction and monthly-scale bias based on Climatic Research Unit (CRU) data (for temperature, diurnal temperature range, cloud-cover, wet days number and precipitation fields). The datasets used for monthly bias-correction are CRU TS4.03 from CRU [12] from 1979 to 2019 for all variables, and the GPCCv2018 full data product [27] for rainfall. The dataset is distributed by the Copernicus Climate Change Service (C3S) through the Climate Data Store (CDS) as monthly netCDF files that are downloaded as zip-files or compressed tar files [28]. NetCDF files are self-describing data that are machine-independent. The netCDF format supports creation, access, and sharing of array-oriented scientific data. The data can be downloaded via the CDS API or through the C3S Climate Data Store.

Table 4.2 shows the names and units of the near-surface variables from WFDE5 used in this study.

Table 4.2: WFDE5 variables

Variable name	Description	Units	Data for monthly bias correction
Wind	10 m wind speed	m s^{-1}	Nil
Tair	2 m air temperature	K	1)
PSurf	Pressure at the surface	Pa	Nil
SWdown	Downward shortwave radiation flux	W m^{-2}	2)
Rainf	Rainfall flux	$\text{kg m}^{-2} \text{s}^{-1}$	3)
Snowf	Snowfall flux	$\text{kg m}^{-2} \text{s}^{-1}$	3)
Qair	2 m specific humidity	kg kg^{-1}	Nil

1) CRU TS4.03 temperature and diurnal temperature range

2) CRU TS4.03 cloud and effects of inter-annual changes in atmospheric loading

3) ERA5 ratio of rainfall/precipitation and snowfall gauge correction

Shyft requires the units of the forcing data to be converted into standard input units. Temperature was converted from Kelvin (K) to Celsius degrees ($^{\circ}\text{C}$) by simply subtracting 273.15 from the absolute temperature. Similarly, the surface pressure was converted from Pa to hPa by multiplying with 0.01. Precipitation was converted to mm h^{-1} by summing the snowfall flux ($\text{kg m}^{-2} \text{s}^{-1}$) and the rainfall flux ($\text{g m}^{-2} \text{s}^{-1}$), and then multiplying the sum with 3600 (s h^{-1}).

The relative humidity was calculated using the MetPy package in Python [21]. The relative humidity function in this package calculates relative humidity from total atmospheric pressure (hPa), air temperature ($^{\circ}\text{C}$) and specific humidity (kg kg^{-1}). The formula is based upon [30] and [25].

The units of wind (m s^{-1}) and global radiation (W m^{-2}) were not converted.

$$RH = \frac{q}{(1 - q)w_s} \quad (4.1)$$

where RH is relative humidity as unitless ratio, q is the specific humidity and w_s is the saturation mixing ratio. The saturation mixing ratio is the ratio between the density of water vapor and the density of dry air [14].

ERA5

The ERA5 dataset is the fifth generation reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) [13]. The dataset provides hourly estimates for a large number of atmospheric, ocean-wave and surface quantities. The data is available from 1950 to present, and has been re-gridded to a regular lat-lon grid of 0.25 degrees. The dataset is updated daily, and can be downloaded in either GRIB or netCDF format from the CDS API or the C3S Climate Data Store.

4.2 Topographical Data

The SRTM 1 Arc-Second Global elevation data from the NASA's Shuttle Radar Tomography Mission (NASA-SRTM) provides a worldwide coverage of void filled data filled with a resolution of 1 arc-second (~ 30 meters) [23]. The Digital Elevation Model (DEM) data can be downloaded from the coverage map in the USGS Earth Explorer. Each GeoTIFF file covers 1 degree tiles, and is about 25 MB each.

Table 4.3 shows the DEM data that was downloaded for this study.

Table 4.3: SRTM 1 Arc-Second Global

Projection	Geographic
Horizontal datum	WGS84
Vertical Datum	EGM96 (Earth Gravitational Mode)
Vertical Units	Meters
Spatial resolution	1 Arc-second (~ 30 m)
Raster size	1 degree tiles
Tiles	N25-31 E82-88
C-band wavelength	5.6 cm

The DEM files were merged into one DEM file in QGIS. This DEM file was then opened in Python using the Rasterio package which reads and writes GeoTIFF files, and provides a Python API based upon Numpy N-dimensional arrays and GeoJSON. The coordinate points from the forcing data were then sampled in the DEM file to get the elevations needed for the Shyft input data.

4.3 Land Cover Data Sets

The Land Cover Classification System (LCCS) Land Cover Map Fine Resolution V2.3 Global dataset from the GlobCover Portal provides global composites and land cover maps [3]. The GlobCover products have been processed by ESA and by the Université Catholique de Louvain using input observations from the 300 m MERIS sensor on board on the ENVISAT satellite mission. The land cover maps covers December 2004 - June 2006 and January - December 2009. The surface reflectance mosaic products are projected in a Platé-Carré projection (WGS84 ellipsoid) with a $1/360^\circ$ pixel resolution. The land cover classes as defined by a set of classifiers.

4.4 Validation datasets

4.4.1 Observed river discharge

Daily stream flow data from the period 2000-2015 is obtained from The Department of Hydrology and Meteorology, DHM. The flow gauging station is located at Arguhat bazaar (485 m.a.s.l) at 28.043611° N and 84.816389° E. The station has been operating since 28 November 1963.

4.4.2 Snow Products

The High Mountain Asia (HMA) snow reanalysis data set from the NASA Snow and Ice Data Center (NSIDC) covers the period 1 October 1999 to 30 September 2017 with a daily temporal resolution and 16 arc-seconds x 16 arc-seconds spatial resolution [19].

4.5 Pre-processing of data

4.5.1 TopoScale

4.5.2 Catchment Discretization Technique

4.6 Hydrological Modeling Framework

Shyft is an open-source cross-platform hydrological toolbox built to provide a computation framework for spatially distributed hydrologic models [5]. The platform is developed by Statkraft AS, the largest generator of renewable energy in Europe, in cooperation with the research community at the University of Oslo. The Shyft software is developed for operational, regional-scale hydropower inflow forecasting, and it aims to i) provide a flexible forecasting toolbox for operational environments, ii) facilitate computationally efficient calculations of hydrologic response at the regional scale, iii) enable the use of multiple hypothesis to quantify forecast uncertainties, iv) allow for multiple model forcing configurations, and v) promote expeditious implementations of the findings from research into operational modeling.

One of the strengths of Shyft is that multiple models may be built through the creation of hydrologic algorithms from a library of well known routines or by the construction of new routines [5]. The user of Shyft can get access to all of the components of the framework via Python through an application programming interface (API), while still having high computational performance as the algorithms are implemented in modern C++. This API set-up enables rapid use of different model configurations and selection of an optimal forecast model.

Since the Shyft toolbox is built for operational, region-scale hydropower inflow forecasting, some of its key design principles vary from what may be prioritized in a pure research environment [5]. Shyft is well-documented open source code and follows the latest code standards, due to the strict requirements regarding software security, testing and code quality in an operational environment. One of the key concepts in the Shyft philosophy is that "data should live at the source". This implies that users of Shyft are encouraged to write their own repositories that connect directly to the source data. This is done to ensure that the data is always up-to-date and that the data is not stored in multiple places. Another key design principle of Shyft facilitates flexible entry points for novel algorithms, while also testing these in parallel with operational runs. This also allows for flexible and exploratory research.

4.6.1 The architecture of Shyft

The Shyft platform is distributed through three GitLab repositories: shyft, shyft-data and shyft-docs. The shyft repository contains the source code for the Shyft platform, while the shyft-data repository contains the test-data needed for the full runs and demo. The shyft-docs repository contains the end-user documentation for the Shyft platform.

Additionally, there is also a Docker repository that contains the Docker build recipes for the Shyft eco-system.

The Shyft platform is built on top of C++ and Python [5]. Basic data structures, hydrologic algorithms and models in the Shyft core is written in C++ to provide high computational efficiency [31]. The C++ code is then compiled into a Python extension module, which is then imported into Python. The Shyft API which is coded in Python is the main interface for the user to interact with the Shyft platform. Figure 4.1 shows the architecture of the Shyft platform.

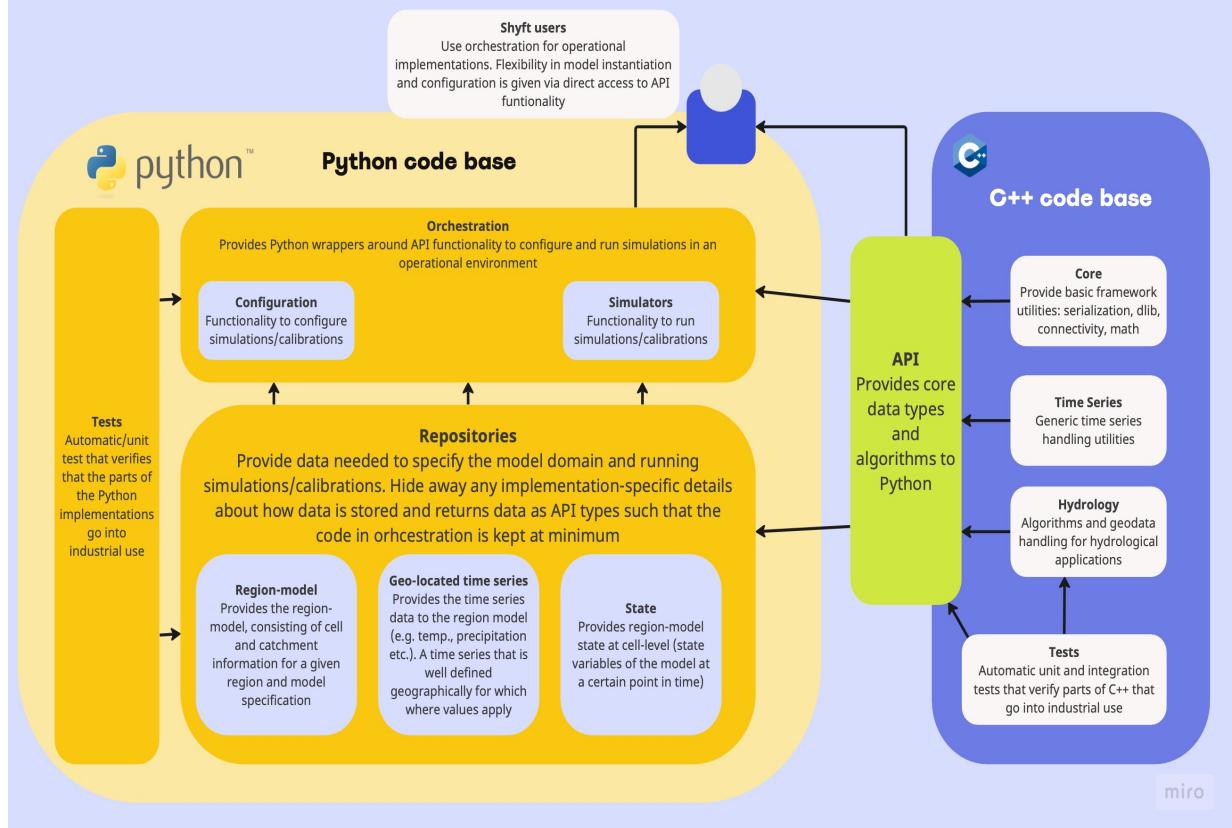


Figure 4.1: Shyft architecture. The figure is an adaptation of the figure in [5], and was made using Miro.

C++ core

The C++ code base contains several separate code folders [5]. The core folder handles framework-related functionality such as serialization; the hydrology folder includes hydrological algorithms and the time series folder handles generic time series.

Python API

The Shyft API is providing the relevant Shyft Core implementations that are needed to configure and utilize models in Python [5]. Figure 4.2 gives an overview of the main Shyft API types. The idea behind the Python API is that users can write pure Python code without being exposed to the C++ code base. The API allows the user to configure and run a model and access data at various levels. Once a model is initiated, the data is stored at a random access memory of the computer, and the user can access the data

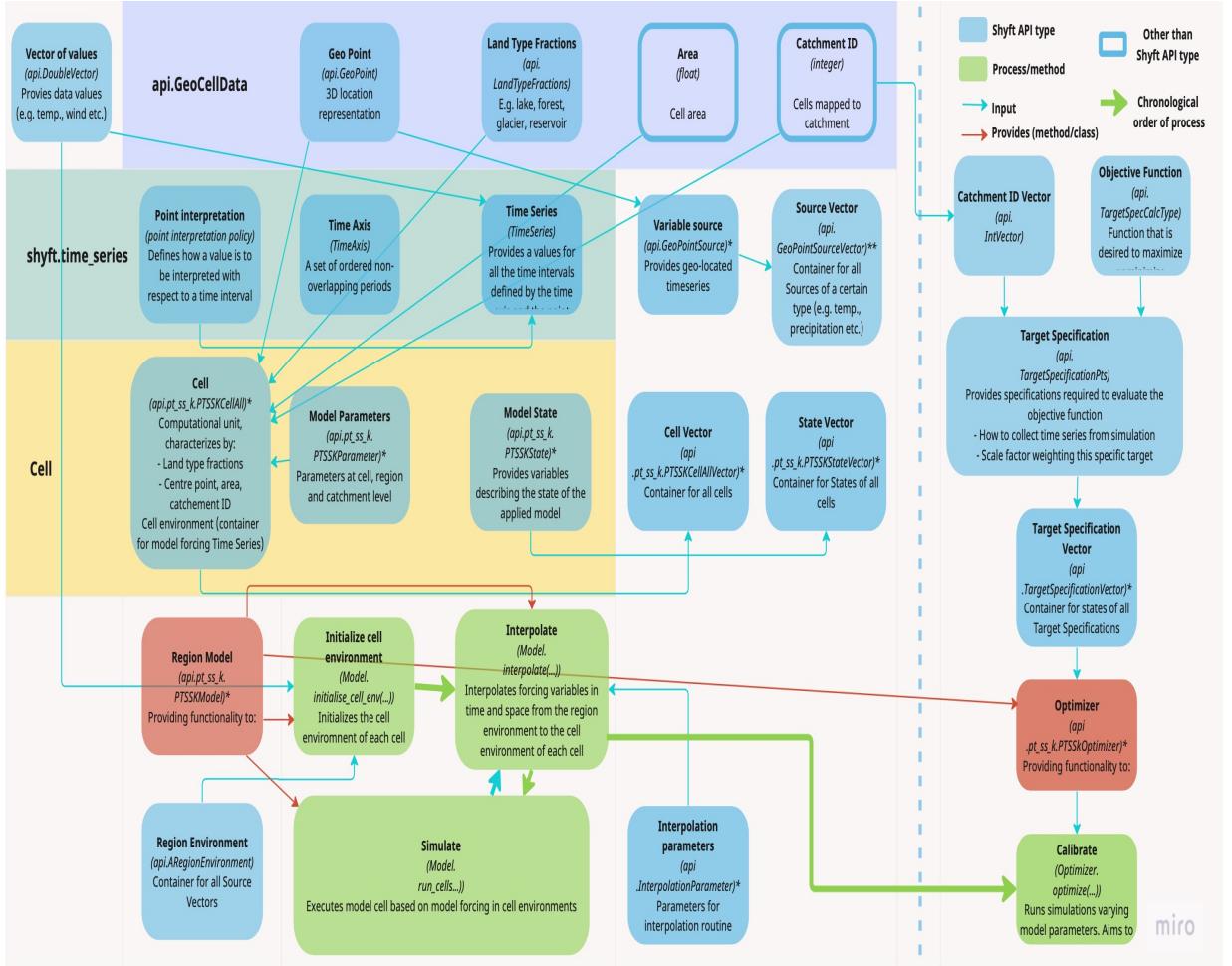


Figure 4.2: Overview of the main Shyft API types. API types used for simulation are shown to the left of the blue dashed line, while calibration API types are found to the right of the blue dashed line. The figure shows an example of the PTSSK model stack (Priestly-Taylor-Skaugen-Snow-Kirchner). The figure is an adaptation of the figure in [5], and was made using Miro.

using interactive shell commands like Jupyter Notebook or IPython. Thus, the Shyft API allows the user to interactively configure the model, provide forcing data to it, run the model, extract and plot variables. Model parameters can easily be updated once the model is instantiated in this way.

shyft.time_series is an API type that provides mathematical and statistical operations and functionality for time series [5]. Time series can be either expressions or concrete time series. Expressions are time series that are defined by a mathematical expression, while concrete time series are time series that are defined by a set of points in time. All time series have a time axis (set of non-overlapping periods), values and a point interpretation policy. The point interpretation policy defines how the value of a time series is interpreted at a given point in time, and can be either instant value, average value or sum value.

api.GeoCellData represents common constant cell properties, regardless of model and cell assemblies. The GeoCellData object contains the following properties: elevation, latitude, longitude, slope, aspect, land type, soil type, land cover fraction, soil moisture, snow cover fraction, and the cell area.

Cell contains *GeoCellData* and *TimeSeries* of model forcings (*api.GeoPointSource*) [5]. The *Cell* is also specific to the *Model* selected. The *CellVector* is a container for all the cells.

RegionModel is a container for all the *Cells* and the *ModelParameters* at region and catchment level. Since everything is vectorized, the *ModelState* collects all the states for each cell [5]. The region model provides functionality like: initialization (*Model.initialize_cell_env(...)*), interpolation (*Model.interpolate(...)*), simulation (*Model.run_cells()*), calibration (*Optimizer.optimize(...)*). The *GoalFunction* that is being minimized is based upon the *TargetSpecification* in the optimizer.

Regionenvironment contains all the *Source* vectors of a certain type (e.g. precipitation, temperature, etc.) [5].

Repositories

The necessary data for running simulations depends on the chosen hydrological model. At a minimum, most models require temperature and precipitation, but some models also require wind speed, relative humidity and shortwave radiation.

The data required to run simulations in Shyft are managed through repositories [11]. A repository is a Python based interface to data [20]. Each repository in Shyft has a specific function, a defined interface and multiple implementations of these interfaces. The repositories return data in Shyft API types in order to keep the code in Shyft orchestration at a minimum. The following interfaces are available in Shyft:

Region model repository: Presents a configured region model, without revealing underlying implementation specific details regarding the storage of model configuration and data (such as netCDF files).

Geo-located time series repository: Provides geo-located time series for meteorology- and hydrology-relevant types.

Interpolation parameter repository: Gives parameters for the interpolation method utilized in the simulation.

State repository: Stores the model states and provides them to the region model.

4.6.2 Orchestration

The orchestration layer is the composition of the simulation configuration [5]. This includes defining the model domain, selecting datasets and model algorithms, and presenting the results. The Shyft orchestration is built on top of the API functionalities and offers utilities for configuring, running and post-processing of simulation. The two main objectives of the orchestration are to offer easy entry points for users of Shyft and to utilize all the functionalities of Shyft in operational environments.

The primary model forcing variables in Shyft are precipitation, atmospheric temperature, wind speed, shortwave radiation and relative humidity.

Table 4.4: Forcing data and validation data

Input	Units	Source	Temporal resolution
<i>Input for model simulations in Shyft</i>			
Precipitation	mm h ⁻¹	WFDE5	Hourly
Temperature	°C	WFDE5	Hourly
Wind speed	m s ⁻¹	WFDE5	Hourly
Relative humidity	-	WFDE5	Hourly
Global radiation	W m ⁻²	WFDE5	Hourly
<i>Data used for validation</i>			
Discharge	m ³ s ⁻¹	DHM	Hourly
Snow cover	per pixel	MODIS	8-day

4.6.3 Spatial interpolation

4.6.4 Hydrological model

4.6.5 Model simulation

4.6.6 Model calibration

Table 4.5: Model calibration parameters

Parameter	Description and unit	Parameter used in the submodel	Lower limit	Upper limit	Sources
c_1	Outlet empirical coeff. 1 (-)	K	-8.0	0.0	[16, 26]
c_2	Outlet empirical coeff. 2 (-)	K	-1.0	1.2	[16, 26]
c_3	Outlet empirical coeff. 3 (-)	K	-0.15	-0.05	[16, 26]
Prescribed parameters					
ae scale factor	Scaling factor for AE (-)	AE	1.0	1.0	[16, 26]

'K' is the catchment response function and 'AE' is actual evapotranspiration.

4.6.7 Water balance estimation

4.6.8 Model validation

4.6.9 Model performance evaluation

4.6. Hydrological Modeling Framework

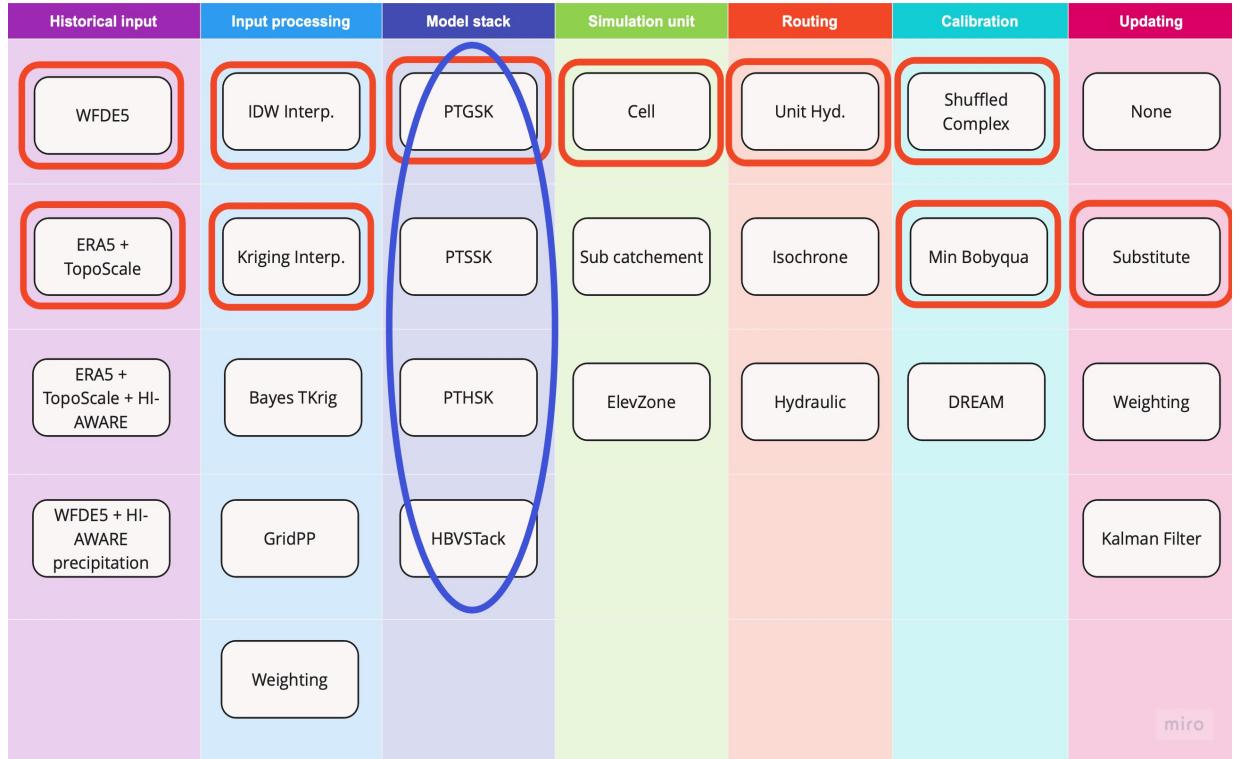


Figure 4.3: Evaluation of multiple configurations. The figure is made using Miro, and it is an adaptation of the figure in [5]

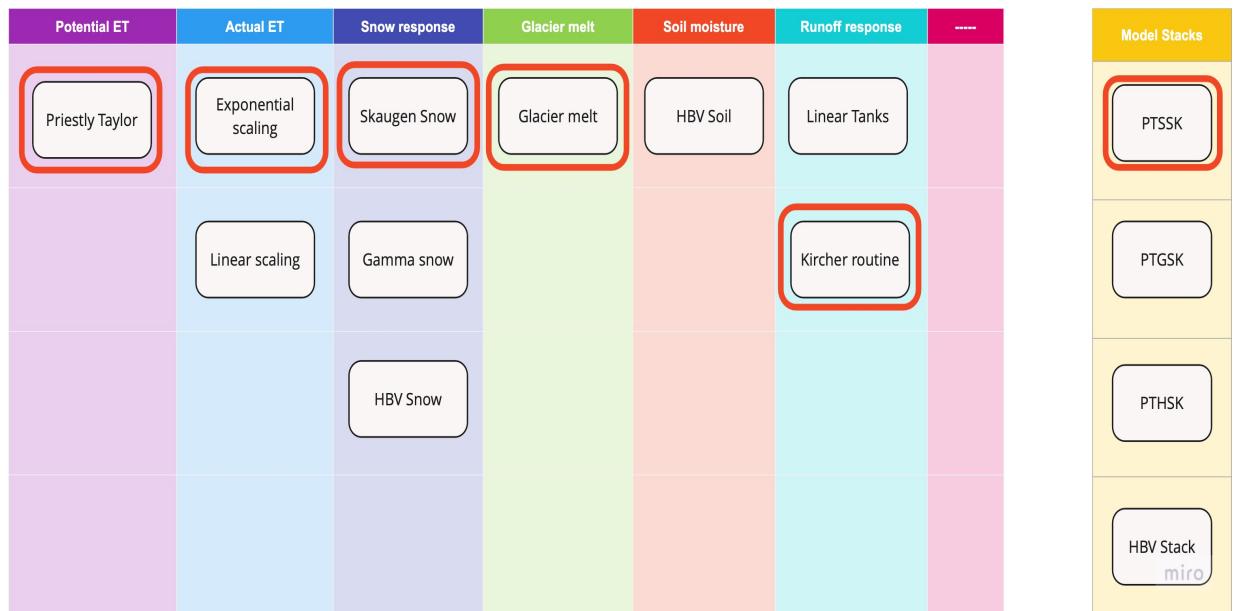


Figure 4.4: Model Stacks components. The figure is made using Miro, and it is an adaptation of the figure in [5]

Chapter 4. Data and methods

Chapter 5

Results

5.1 Forcing data analysis

5.2 Evaluation of discharge simulation using different datasets

5.3 Water balance analysis

Chapter 5. Results

Chapter 6

Discussion

Chapter 6. Discussion

Chapter 7

Conclusions and outlook

Chapter 7. Conclusions and outlook

Appendix A

Downloading data from the CDS API

The following Python script describes how the WFDE5 data was downloaded using the CDS API [6].

```
, , ,  
    DOWNLOAD WFDE5 data  
    How to download the near-surface-meteorological-variables  
    from the CDS API.  
    Here version 2.1 is used. Each variable is downloaded  
    and saved as a zip file with the name of each variable.  
, , ,  
  
import cdsapi  
  
# Hindu Kush-Himalayan region extent  
# (Bajracharya, SR; Shrestha, B (eds) (2011) The status of glaciers  
# in the Hindu Kush-Himalayan region. Kathmandu: ICIMOD)  
  
north = 39.31  
west = 60.85  
south = 15.95  
east = 105.04  
  
c = cdsapi.Client()  
  
  
variables = [ 'near_surface_specific_humidity' ,  
              'grid_point_altitude' ,  
              'near_surface_air_temperature' ,  
              'near_surface_specific_humidity' ,  
              'near_surface_wind_speed' ,  
              'rainfall_flux' ,  
              'snowfall_flux' ,  
              'surface_air_pressure' ,  
              'surface_downwelling_longwave_radiation' ,  
              'surface_downwelling_shortwave_radiation'  
            ]
```

```

for variable in variables:
    if variable == ('rainfall_flux' or 'snowfall_flux'):
        c.retrieve(
            'derived-near-surface-meteorological-variables',
            {   'area': [north, west, south, east],
                'variable': f'{variable}',
                'reference_dataset': 'cru_and_gpcc',
                'year': [
                    '1990', '1991', '1992',
                    '1993', '1994', '1995',
                    '1996', '1997', '1998',
                    '1999', '2000', '2001',
                    '2002', '2003', '2004',
                    '2005', '2006', '2007',
                    '2008', '2009', '2010',
                    '2011', '2012', '2013',
                    '2014', '2015', '2016',
                    '2017', '2018', '2019',
                ],
                'month': [
                    '01', '02', '03',
                    '04', '05', '06',
                    '07', '08', '09',
                    '10', '11', '12',
                ],
                'version': '2.1',
                'format': 'tgz'
            },
            f'{variable}.tgz')
    else:
        c.retrieve(
            'derived-near-surface-meteorological-variables',
            {   'area': [north, west, south, east],
                'variable': f'{variable}',
                'reference_dataset': 'cru',
                'year': [
                    '1990', '1991', '1992',
                    '1993', '1994', '1995',
                    '1996', '1997', '1998',
                    '1999', '2000', '2001',
                    '2002', '2003', '2004',
                    '2005', '2006', '2007',
                    '2008', '2009', '2010',
                    '2011', '2012', '2013',
                    '2014', '2015', '2016',
                    '2017', '2018', '2019',
                ],
                'month': [

```

```
'01', '02', '03',
'04', '05', '06',
'07', '08', '09',
'10', '11', '12',
],
'version': '2.1',
'format': 'tgz',
},
f'{variable}.tgz')
```

Appendix A. Downloading data from the CDS API

Bibliography

- [1] Bikas Chandra Bhattarai et al. “Aerosol Optical Depth Over the Nepalese Cryosphere Derived From an Empirical Model.” In: *Frontiers in Earth Science* 7 (July 5, 2019), p. 178. ISSN: 2296-6463. DOI: 10.3389/feart.2019.00178. URL: <https://www.frontiersin.org/article/10.3389/feart.2019.00178/full> (visited on 10/17/2022).
- [2] Bikas Chandra Bhattarai et al. “Evaluation of Global Forcing Datasets for Hydropower Inflow Simulation in Nepal.” In: *Hydrology Research* 51.2 (Apr. 1, 2020), pp. 202–225. ISSN: 0029-1277, 2224-7955. DOI: 10.2166/nh.2020.079. URL: <https://iwaponline.com/hr/article/51/2/202/72386/Evaluation-of-global-forcing-datasets-for> (visited on 10/17/2022).
- [3] S. Bontemps et al. *GLOBCOVER 2009 - Products Description and Validation Report*. UCLouvain & ESA Team, 2011.
- [4] Bodo Bookhagen and Douglas W. Burbank. “Toward a Complete Himalayan Hydrological Budget: Spatiotemporal Distribution of Snowmelt and Rainfall and Their Impact on River Discharge.” In: *Journal of Geophysical Research* 115.F3 (Aug. 10, 2010), F03019. ISSN: 0148-0227. DOI: 10.1029/2009JF001426. URL: <http://doi.wiley.com/10.1029/2009JF001426> (visited on 09/29/2022).
- [5] John F. Burkhardt et al. “Shyft v4.8: A Framework for Uncertainty Assessment and Distributed Hydrologic Modeling for Operational Hydrology.” In: *Geoscientific Model Development* 14.2 (Feb. 5, 2021), pp. 821–842. ISSN: 1991-9603. DOI: 10.5194/gmd-14-821-2021. URL: <https://gmd.copernicus.org/articles/14/821/2021/> (visited on 09/29/2022).
- [6] Marco Cucchi et al. “WFDE5: Bias-Adjusted ERA5 Reanalysis Data for Impact Studies.” In: *Earth System Science Data* 12.3 (Sept. 8, 2020), pp. 2097–2120. ISSN: 1866-3516. DOI: 10.5194/essd-12-2097-2020. URL: <https://essd.copernicus.org/articles/12/2097/2020/> (visited on 09/29/2022).
- [7] S. Dandekhya et al. *The Gandaki Basin – Maintaining Livelihoods in the Face of Landslides, Floods, and Drought. HI-AWARE Working Paper 9*. Kathmandu: HI-AWARE, 2017.
- [8] Rohini P. Devkota et al. “Climate Change and Adaptation Strategies in Budhi Gandaki River Basin, Nepal: A Perception-Based Analysis.” In: *Climatic Change* 140.2 (Jan. 2017), pp. 195–208. ISSN: 0165-0009, 1573-1480. DOI: 10.1007/s10584-016-1836-5. URL: <http://link.springer.com/10.1007/s10584-016-1836-5> (visited on 10/03/2022).
- [9] F. Dordé and M. L’Hostis. *Feasibility Study and Detailed Design of Budhi Gandaki HPP*. Main Report 1. Tractebel Engrineering & JADE Consult, 2015.

Bibliography

- [10] O. Edenhofer et al. *Renewable Energy Sources and Climate Change Mitigation: Special Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, 2011.
- [11] Martin Fowler. *Patterns of Enterprise Application Architecture*. The Addison-Wesley Signature Series. Boston: Addison-Wesley, 2003. 533 pp. ISBN: 978-0-321-12742-6.
- [12] Ian Harris et al. “Version 4 of the CRU TS Monthly High-Resolution Gridded Multivariate Climate Dataset.” In: *Scientific Data* 7.1 (Dec. 2020), p. 109. ISSN: 2052-4463. DOI: 10.1038/s41597-020-0453-3. URL: <http://www.nature.com/articles/s41597-020-0453-3> (visited on 10/17/2022).
- [13] H. Hersbach et al. “ERA5 Hourly Data on Single Levels from 1959 to Present.” In: *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* (2018). DOI: 10.24381/cds.adbb2d47.
- [14] Akira Kasahara and Masao Kanamitsu. “Weather Prediction, Numerical.” In: *Encyclopedia of Physical Science and Technology*. Elsevier, 2003, pp. 805–835. ISBN: 978-0-12-227410-7. DOI: 10.1016/B0-12-227410-5/00824-3. URL: <https://linkinghub.elsevier.com/retrieve/pii/B0122274105008243> (visited on 01/23/2023).
- [15] H B Khatri, M K Jain, and S K Jain. “Modelling of Streamflow in Snow Dominated Budhigandaki Catchment in Nepal.” In: *Journal of Earth System Science* 127.7 (Oct. 2018), p. 100. ISSN: 0253-4126, 0973-774X. DOI: 10.1007/s12040-018-1005-5. URL: <http://link.springer.com/10.1007/s12040-018-1005-5> (visited on 09/29/2022).
- [16] J. U. Lombraña. “Evaluations of Snow Simulations in SHyFT.” Trondheim: Norwegian University of Science and Technology (NTNU), 2017.
- [17] Suresh Marahatta, Laxmi Devkota, and Deepak Aryal. “Hydrological Modeling: A Better Alternative to Empirical Methods for Monthly Flow Estimation in Ungauged Basins.” In: *Journal of Water Resource and Protection* 13.03 (2021), pp. 254–270. ISSN: 1945-3094, 1945-3108. DOI: 10.4236/jwarp.2021.133015. URL: <https://www.scirp.org/journal/doi.aspx?doi=10.4236/jwarp.2021.133015> (visited on 09/29/2022).
- [18] Suresh Marahatta, Laxmi Prasad Devkota, and Deepak Aryal. “Application of SWAT in Hydrological Simulation of Complex Mountainous River Basin (Part I: Model Development).” In: *Water* 13.11 (May 31, 2021), p. 1546. ISSN: 2073-4441. DOI: 10.3390/w13111546. URL: <https://www.mdpi.com/2073-4441/13/11/1546> (visited on 09/29/2022).
- [19] Steven Margulies. *High Mountain Asia UCLA Daily Snow Reanalysis*. NASA National Snow and Ice Data Center DAAC, 2021. DOI: 10.5067/HNAUGJQXSCVU. URL: https://nsidc.org/data/HMA_SR_D/versions/1 (visited on 10/13/2022).
- [20] A. A. MASNUN. *Interfaces in Python: Protocols and ABCs*. URL: <http://masnun.rocks/2017/04/15/interfaces-in-python-protocols-and-abcs/> (visited on 02/02/2023).
- [21] Ryan M. May et al. “MetPy: A Meteorological Python Library for Data Analysis and Visualization.” In: *Bulletin of the American Meteorological Society* 103.10 (Oct. 2022), E2273–E2284. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/BAMS-D-21-0125.1. URL: <https://journals.ametsoc.org/view/journals/bams/103/10/BAMS-D-21-0125.1.xml> (visited on 01/30/2023).

- [22] M. Ménégoz, H. Gallée, and H. W. Jacobi. “Precipitation and Snow Cover in the Himalaya: From Reanalysis to Regional Climate Simulations.” In: *Hydrology and Earth System Sciences* 17.10 (Oct. 15, 2013), pp. 3921–3936. ISSN: 1607-7938. DOI: 10.5194/hess-17-3921-2013. URL: <https://hess.copernicus.org/articles/17/3921/2013/> (visited on 10/13/2022).
- [23] Earth Resources Observation and Science (EROS) Center. *Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global*. U.S. Geological Survey, 2017. DOI: 10.5066/F7PR7TFT. URL: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1-arc?qt-science_center_objects=0#qt-science_center_objects (visited on 01/23/2023).
- [24] Francesca Pellicciotti et al. “Challenges and Uncertainties in Hydrological Modeling of Remote Hindu Kush–Karakoram–Himalayan (HKH) Basins: Suggestions for Calibration Strategies.” In: *Mountain Research and Development* 32.1 (Feb. 2012), pp. 39–50. ISSN: 0276-4741, 1994-7151. DOI: 10.1659/MRD-JOURNAL-D-11-00092.1. URL: <http://www.bioone.org/doi/10.1659/MRD-JOURNAL-D-11-00092.1> (visited on 10/13/2022).
- [25] Murry L Salby. *Fundamentals of Atmospheric Physics*. Elsevier, 1996.
- [26] A. Sapkota. “Regional Modelling of the Narayani Basin in Nepal.” Trondheim: Norwegian University of Science and Technology, 2016.
- [27] Udo Schneider et al. “GPCC Full Data Monthly Version 2018.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges Built on GTS-based and Historic Data: Gridded Monthly Totals.” In: (2018). DOI: 10.5676/DWD_GPCC/FD_M_V2018_050. URL: https://opendata.dwd.de/climate_environment/GPCC/html/fulldata-monthly_v2018_doi_download.html (visited on 10/17/2022).
- [28] Copernicus Climate Change Service. *Near Surface Meteorological Variables from 1979 to 2018 Derived from Bias-Corrected Reanalysis*. ECMWF, 2020. DOI: 10.24381/CDS.20D54E34. URL: <https://cds.climate.copernicus.eu/doi/10.24381/cds.20d54e34> (visited on 10/12/2022).
- [29] Arun B. Shrestha et al. “Maximum Temperature Trends in the Himalaya and Its Vicinity: An Analysis Based on Temperature Records from Nepal for the Period 1971–94.” In: *Journal of Climate* 12.9 (Sept. 1999), pp. 2775–2786. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/1520-0442(1999)012<2775:MTTITH>2.0.CO;2. URL: [http://journals.ametsoc.org/doi/10.1175/1520-0442\(1999\)012%3C2775:MTTITH%3E2.0.CO;2](http://journals.ametsoc.org/doi/10.1175/1520-0442(1999)012%3C2775:MTTITH%3E2.0.CO;2) (visited on 10/17/2022).
- [30] J. M. Wallace. *Atmospheric Science: An Introductory Survey*. Academic Press, 1977.
- [31] Farzeen Zehra et al. *Comparative Analysis of C++ and Python in Terms of Memory and Time*. preprint. MATHEMATICS & COMPUTER SCIENCE, Dec. 21, 2020. DOI: 10.20944/preprints202012.0516.v1. URL: <https://www.preprints.org/manuscript/202012.0516/v1> (visited on 01/31/2023).