Nature-based solutions to mitigate climate extremes impacting the water-food-ecosystem nexus within the Rocha River Basin (Cochabamba - Bolivia)

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: N(ew data) or E(xisting data)	Indicate: D(igital) or P(hysical)	Indicate: Audiovisual Images Sound Numerical Textual Model SOftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
Climate data sets	Meteorological stations data for variables like precipitation, temperatures	E	D	N	.xlsx	<1GB	
Climate grids	Gridded datasets for climate variables, precipitation, temperatures	E	D	N	.nc	<100 GB	
Hydrometric and soil moisture data	Daily datasets of river waterlevels, flow velocities. Soil moisture data collected from sensors at 2 meters depth to account for infiltration rates	N	D	N	.xlsx, .csv	<1 GB	
Vegetation indexes	Monthly time series of vegetation indexes calculated for the basin	N	D	I	.tif, .xlsx	<100 GB	
Hydrological models	Numerical models for water balance components of the basin	N	D	M	.tif, .xlsx	<100 GB	
Wells water levels	Datasets collected from fieldwork by measuring groundwater levels	N	D	N	.xlsx	<1GB	
Land Use and Land Cover Maps	Maps showing land cover classes changing over time	N	D	M	.tiff	<100 GB	
Nature Based Solutions Maps	Potential zones where NBS should be applied in the basin	N	D	M	.tiff, .nc	< 100GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Climate datasets collected from SENAMHI BOLIVIA: https://senamhi.gob.bo/index.php/inicio

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven

or UZ Leuven privacy register number (G or S number).

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- Yes

Data sets collected from ENDE CORANI may be used for research purposes only. Due to the confidential status of a hydropower company.

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

For each of the datasets mentioned above, the approach is described as follows:

Climate data sets	**Methodology**: Description of the data sources (e.g., weather stations, satellite-derived), data acquisition protocols, quality control procedures, and any homogenization steps. **Variables & Units**: Temperature (°C), precipitation (mm). **Documentation Includes**: Metadata file (e.g., in XML or JSON) with station locations, time coverage. Data processing logs (e.g., outlier treatment, gap filling). README with definitions of variables, data format (e.g., csv, xlsx). **Processing**: Notebooks/scripts for quality checks and trends analysis. **Shared**: Clean datasets with accompanying README text file and metadata.

Climate grids	**Methodology**: Gridded datasets such as ERA5, CHIRPS, GMET will be downloaded to assess the spatial distribution of precipitation and temperature patterns in the basin. The spatial and temporal resolutions will be tested and validated using statistical metrics. **Variables & Units**: Same as above, but spatially continuous. **Documentation Includes**: Grid specifications (e.g., resolution, extent, projection). Description of performance statistics (RMSE, bias) compared with point observations. README explaining data structure, units, and cell value interpretations. **Processing**: Scripts for grid calculations and validation plots. **Shared**: NetCDF/GeoTIFF files + README text file.
Hydrometric and soil moisture data	**Methodology**: River gauge measurement techniques, velocimeters, soil moisture sampling and measurements, maintenance logs. **Variables & Units**: Water level (m), discharge (m³/s), soil moisture (mm). **Documentation Includes**: Station metadata (location, altitude, river name, catchment area). Measurement frequency and instrumentation details. Data QA/QC procedures. **Processing**: Rating curves and time series for modelling validation **Shared**: Datasets time series including hydrometric handbook or summary.
Vegetation indexes	**Methodology**: Source satellite/sensor (e.g., MODIS, Landsat, Sentinel), computation of NDVI/NDWI, preprocessing (cloud masking, atmospheric correction). **Variables & Units**: NDVI (unitless, -1 to 1), NDWI, etc. **Documentation Includes**: Data acquisition frequency. Scripts/methods used for index calculation in GGE. Projection info and data format description. **Processing**: Scripts for temporal analysis and anomaly detection in GGE and Python. **Shared**: GeoTIFF files, processing workflow, and README text file.
Hydrological models	**Methodology**: Model type (e.g., SWAT, Modflow), input datasets used, calibration and validation methods. **Variables & Units**: Simulated flows (m³/s), soil moisture contents (%), infiltration rates (mm/day). **Documentation Includes**: Model setup description, input/output file formats. Calibration parameters and performance metrics. Model assumptions and limitations. **Processing**: Model configuration files and scenario testing notes. **Shared**: Input/output datasets, copy of the calibrated models
Wells water levels	**Methodology**: Measurement techniques (e.g., manual dip meter, pressure transducers), frequency, QA/QC process. **Variables & Units**: Groundwater level (m below ground level or m.a.s.l). **Documentation Includes**: Well metadata (location, depth, aquifer type). Time series data structure and units. Calibration procedures and equipment specs. **Processing**: Measurements and sensor level loggers evaluation. **Shared**: Datasets file, well inventory sheet and README text file.
Land Use and Land Cover Maps	**Methodology**: Classification method (e.g., supervised, unsupervised), source imagery, classification techniques used such as K means, random forest, etc. **Variables & Units**: Land cover class codes (nominal). **Documentation Includes**: Classification accuracy assessment and confusion matrix. Legend and class definitions. Projection info and scale/resolution. **Processing**: Codebook for classes, training data logs, scripts in GGE and Python, QGIS. **Shared**: GeoTIFF/Vector files + classification report + metadata

Nature Based Solutions Maps	**Methodology**: Mapping framework used, source data (fieldwork, satellite), selection criteria. **Variables & Units**: Type of NBS (categorical), effectiveness score (if applicable). **Documentation Includes**: Description of each NBS type and spatial delineation approach. Criteria and indicators used for mapping (e.g., erosion control, biodiversity support). Maps including projection and data sources.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- No

 We are currently searching for possible metadata standards methods within our research group. Since the project is in collaboration with UMSS, the metadata has to be easily accessible.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Shared network drive (J-drive)
- Other (specify below)
- Large Volume Storage
- Personal network drive (I-drive)

A copy will be shared through Google Drive with UMSS researchers and collaborators. A copy will also be saved in Shared network drive (K-drive) to share with KULEUVEN researches collaborators

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution
- Personal back-ups I make (specify below)

The data will be back up with Google Drive accounts from UMSS and One Drive from KULeuven in K-drive folder.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Data stored on OneDrive and network drives from KU Leuven are only accessible to members of the research group. Data stored on online data repositories, like the one UMSS uses (Google Drive), can only be modified by the corresponding authors (ourselves).

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

There are no expected costs.

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies…).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Shared network drive (J-drive)
- Other (specify below)

Google Drive of LHUMSS research group.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

There are no expected costs.

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Yes, as open data

If access is restricted, please specify who will be able to access the data and under what conditions.

We strongly encourage open data access so data can be shared and used among researchers within the same research fields. Therefore, we do not need to specify who will be able to access the data and under what conditions.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- Yes, privacy aspects

Data from ENDE CORANI, will be treated with privacy, due to the regulations of the institution. This does not affect the rest of the

data.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- KU Leuven RDR (Research Data Repository)
- Other data repository (specify below)

We will use the Google Drive network to store the data in order to be available for the research group of UMSS.

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

What are the expected costs for data sharing? How will these costs be covered?

There are no expected costs.

Responsibilities

Who will manage data documentation and metadata during the research project?

Data documentation and metadata will be managed by my supervisors and me.

Who will manage data storage and backup during the research project?

Data storage and backup will be managed by my supervisors and me.

Who will manage data preservation and sharing?

Data preservation and sharing will be managed by my supervisors and me.

Who will update and implement this DMP?

The DMP will be updated by my supervisors and me.

Created using DMPonline.be. Last modified 12 April 2025

7 of 7