Reproducibility package

Assessing the value of impact-based forecasting for early action to reduce flood impacts

A case study for the 2024 Brazil floods

Contents

1. Overview
2. Instructions for Replicators
3. Data Availability
4. List of Exhibits

Overview

This package contains the code and data to reproduce the output of the Policy Research Working Paper "Assessing the value of impact-based forecasting for early action to reduce flood impacts: A case study for the 2024 Brazil floods" by Roy Pontman, Jeroen C.J.H. Aerts, Paolo Avner, Jens de Bruijn, Rafaella Gouveia Loureiro Oliveira, Bramka Arga Jafino, Hans de Moel, Moussa Sidibe, and Tim Busker.

It is important to emphasize that the entire developed model framework is publicly available on GitHub: https://github.com/GEB-model/GEB, where for this study the following branch is created: "Roy_Brazil_floods_meteorological_warnings". This framework contains the used models, downloads the raw input data, processes the input data and visualizes the first model outputs such as the evaluation of the meteorological forecasts (Figure 2) and the validation of the simulated flood maps (Figure 5, 6 and A.1). The architecture of the model is documented in detail on the corresponding GitHub repository. In the present work, we have introduced additional directories, such that the resulting package is organized as follows:

```
Reproducibility package/
|
├── models/                                          # Used models
|   ├── rio_sinos_med_forecast/                      # Similar folder structure as rio_sinos_sev_forecast but for medium flood depth conditions
|   └── rio_sinos_sev_forecast/                      # Sinos model with results for severe flood depth conditions
|       └── base/
|           ├── input/                               # Input data directory with processed data
|           ├── cache/                               # Input data directory with RAW data
|           ├── simulation_root/                     # SFINCS initialisation and stored model parameters during spin up and forecasting
|           ├── output/                              # Model output folder
|               ├── flood_prob_exceedance_maps/      # Output folder for Exceedance flood probability maps (ensemble forecasts)
|               ├── action_maps/                     # Output folder for AbM part with the actions for each household stored for each scenario
|               ├── building_fraction_hit/           # Output folder for the fraction of buildings hit in each admin unit for each scenario
|               ├── damage_maps/                     # Output folder for damage maps derived from the flood maps
|               ├── flood_maps/                      # Output folder for all flood map simulations
|               ├── report/                          # Output folder for user defined model parameters within the model.yml
|               ├── warning_logs/                    # Output folder for where the warnings are logged per forecast for each household
|               └── evaluate/                        # Output folder for the validation of flood maps and evaluation of meteorological forecasts
|           ├── GEB.log                              # Log file of the GEB model simulations
|           ├── model.yml                            # Model configuration file, settings to run the model
|           ├── build.log                            # Log file of building the GEB model
|           ├── update.yml                           # Model configuration file, settings to update certain model inputs
|           └── build.yml                            # Model configuration file, settings to build the GEB model
├── output analysis scripts/
|   ├── general input files/                         # Needed input files for output analysis
|   ├── RP_Processing_Model_Results_And_Figures.py   # Python script that further processes the model outputs into figures presented in the paper
|   ├── output/                                      # Output with data and figures presented in the paper
|       ├── warnings/                                # Output folder for warned administrative units per warning type
|       └── action_maps/                             # Output folder for analysed action maps per scenario
|   └── environment_requirements.lock                # Python environment txt file with dependencies
└── README.pdf                                       # This README file
```

Instructions for Replicators

Prerequisites

The model implementation and data processing pipeline are executed in Python. To employ the GEB model, follow the usage instructions provided in the previously referenced GitHub repository. For reproducibility of the analysis, install all required dependencies by creating a python environment with your preferred package manager. The use of the UV package manager (https://docs.astral.sh/uv/getting-started/installation/) and Visual Studio Code (VS Code) is strongly recommended for this study.

It is essential that these commands are issued from the correct working directory within the VS Code terminal. Create a new virtual environment (.venv) in the project directory `./output analysis scripts` by executing `***uv venv***` in the integrated terminal of VS Code. Subsequently, activate the virtual environment using either `***source .venv/bin/activate***` (for Linux) or `***.venv\Scripts\activate***` (for Windows), depending on the operating system. To reproduce the software environment used in the study, install all dependencies with their exact versions specified in the provided `./output analysis scripts/environment_requirements.lock` file by running:

`***uv pip sync environment_requirements.lock***`.

To enable use of this environment within Jupyter notebooks, first install the ipykernel package:

`***uv pip install ipykernel***`, and then register and label the kernel with:

`***python -m ipykernel install --user --name flood_env --display-name Output_Analysis_PRWP***`.

Finally, in VS Code, select the newly created kernel via the “Select Kernel” interface to ensure that notebooks are executed within this configured environment.

To enable use of this environment without interactive notebooks, select the interpreter with ctrl+shift+P in VS Code. Furthermore, it is essential that this reproducibility package be stored and executed on a local workstation or high-performance computing (HPC) infrastructure, rather than on cloud-based platforms such as One Drive or Google Drive.

General

Executing the model

Assuming that the GEB framework has been configured for the region with the build.yml, the GEB model must be executed in two stages. First, the command "**geb spinup**" must be executed, followed by the execution of "**geb run**". Depending on the configuration defined in the model.yml file, individual model segments can be selectively activated or deactivated. An example of the used model configuration is provided in "./models/rio_sinos_sev_forecast/base/model.yml". In this configuration, all available processing segments are activated, enabling the full execution chain required to obtain the complete set of outputs, ranging from flood maps to household-level action maps for the IbW severe scenario.

Please note that once all flood maps have been generated for the various forecast ensembles, this model component can be deactivated to reduce computational effort and overall runtime. This can be done by flagging (including hashtags before the lines of code) the following lines within the "model.py", a script part of the GEB model:

> *# self.multiverse(forecast_issue_datetime=dt, return_mean_discharge=True) # run the multiverse for the current timestep*

Subsequently, the different warning and response scenarios can be simulated more efficiently in a post-processing step by changing the agent settings within the model.yml.

Output files

The model automatically stores simulation inputs and outputs in the following directories: "./models/model/base/input" for input data and "./models/model/base/output" for output data. In this context, the specific "model" instance may correspond to either "rio_sinos_med_forecast" or "rio_sinos_sev_forecast". In general, the model input files are created automatically when building the model as further specified in the GitHub repository. These two models, provided within the package, are employed to generate distinct scenarios corresponding to medium and severe flood depths. In this configuration, only a subset of the model parameters differs between the two cases, whereas most of the input variables and output variables remain consistent. More specifically, the probability threshold differs together with the flood depths of interest.

The primary output files from the GEB framework consist of the different flood maps (Deterministic and exceedance probability), the meteorological warning files, the household action maps and the damage maps. These outputs are created by running the model. It should be noted that the report presents results only for two forecast lead times, namely the 7-day and 2-day lead times. Nonetheless, the model generates simulations for all intermediate lead times between 7 days and 1 day.

Primary output files from the model:

The deterministic flood maps: The deterministic flood maps are derived from distinct meteorological and hydrological boundary conditions obtained from individual ensemble members or reanalysis products (ERA5). For each forecast lead time, separate maps are generated for every ensemble member as well as for the ERA5 baseline. These products are stored within an xarray dataset, which is defined by the following dimensions:

XARRAY PARAMETER	DESCRIPTION	SIZE (GRID CELLS)
Y	Longitude dimension	2344
X	Latitude dimension	3892

20240426T000000 - 20240507T000000	Array variable of the maximum flood depth [m] in each grid cell (x,y)	

The flood exceedance probability maps: The flood exceedance probability maps quantify the exceedance of the likelihood of flooding, derived from the ensemble of 50-member flood simulations stored in the "./flood_maps" directory. The probability threshold (set in the model.yml) and the flood depth threshold jointly determine whether the exceedance criterion is met for a given grid cell. These maps are generated for each forecast lead time and are stored in the "./ flood_prob_exceedance_maps" directory as an xarray dataset characterized by the following dimensions:

XARRAY PARAMETER	DESCRIPTION	SIZE (GRID CELLS)
Y	Longitude dimension	2344
X	Latitude dimension	3892
PROB_EXCEEDANCE_MAP _RANGE1_STRATEGY1	Array variable of grid cells exceed a flood depth [m] with a certain probability	

Furthermore, Figure 5 and Figure 6 are created by executing the following command line in the terminal `***geb evaluate --method hydrology.evaluate_hydrodynamics***`. With this command the created flood maps are validated against the observed flood map specified in the model.yml. Depending on whether the probability_maps parameter is set on True within the evaluate_hydrodynamics function. This function can be found within the hydrology.py of the evaluate folder within the GEB framework and determines whether the flood maps associated with specific exceedance probabilities are validated or that individual flood maps are validated (e.g. ensemble members, ERA5-based simulations, etc.). It is important to note that Figure 5 is generated with the probability_maps parameter set to `True`, whereas Figure 6 is produced with the probability_maps parameter set to `False`. Furthermore, when the `use forecasts` option is set to `False` in the `model.yml` configuration file, the ERA5 simulation presented in Appendix A is subjected to validation. Figure 5 presents a compilation of two validation maps. For each forecast, the corresponding validation map is stored in the directory:

"./models/rio_sinos_med_forecast/base/output/evaluate/hydrodynamics/forecasts/forecast_20240426T000000/probability_map/prob_exceedance_map_range1_strategy1_validation_floodextent_plot.png"

, where the timestamp "20240426T000000" specifies the forecast instance under consideration. The compiled figure as presented in the report can be found in "./output analysis scripts/output".

Similarly, Figure 6 is stored in the directory:

"./models/rio_sinos_med_forecast/base/output/evaluate/hydrodynamics/forecasts/forecast_20240502T000000/member_50/20240426T000000-20240507T000000_forecast_performance_spread.png". This figure is consistently stored in the final ensemble member of the most recent available forecast initialization date, which for this study is 2 May 2024.

The meteorological warning files: Using the `evaluate` function of GEB, meteorological warnings are derived from the ECMWF forecast data stored in the model's input directory. The warnings based on the ECMWF ensemble forecasts are stored for each forecast in the following directory: "./evaluate/forecasts/rainfall_warnings/timeseries_forecast". Within the `rainfall_warnings` directory, each forecast is stored according to a standardized naming. For example, the forecast issued for 1 May

is represented by the file `timeseries_20240501T000000`. Here, the issued warnings are stored for the ERA5 baseline and the control- and ensemble forecast as a zarr file:

XARRAY PARAMETER	DESCRIPTION	SIZE
Y	Longitude dimension	16
X	Latitude dimension	21
TIME	Time dimension (time steps)	11
PR_KG_PER_M2_PER_S _20240501T000000	Array variable of rainfall warnings in grid cells with warning_level: 0=No Warning, 1=Orange, 2=Purple	

Moreover, Figure 2 is stored in the directory "./evaluate/forecasts" under the file name "ERA5_vs_ECMWF_Control_&_Probabilistic_Cumulative_Precipitation_Forecasts.png". This figure is generated automatically upon execution of the terminal command `***geb evaluate --method meteorological_forecasts.evaluate_forecasts***`, provided that the function `evaluate_precipitation_ forecasts` remains active (i.e., not commented out) within the file `meteorological_forecasts.py` located in the `evaluate` directory of the GEB model.

The household action maps: Dependent on the agent settings within the model.yml, the model simulates different scenarios. Within the agent settings, the warning_response parameter determines whether the AbM is activated or not by setting it True or False. Based on a combination of settings, the model simulates the scenarios as specified in the paper. The most important setting is the warning_type, specifying whether a General Flood Warning is issued (FW) or an Impact-based Warning (IbW):

MODEL SETTING	MEDIUM FLOOD DEPTH		SEVERE FLOOD DEPTH	
WARNING TYPE	Impact_based	Flood_general	Impact_based	Flood_general
PROBABILITY THRESHOLD	0.3	0.3	0.6	0.6
AREA THRESHOLD	-	0.1	-	0.1
BUILDING THRESHOLD	0.1	-	0.1	-
COMMUNICATION EFFICIENCY	0.78	0,78	0.78	0,78
EVACUATION_LEAD_TIME _THRESHOLD	24	24	24	24
RESPONSE RATES	0.41	0.36	0.51	0.36

By repeatedly executing this component of the model with varying parameter configurations, the different scenarios presented in the paper are generated and stored in the './models/model/output/action_maps' directory. These maps encode information on the actions of each household within the catchment area, including instances in which households do not undertake any action. In addition, the model produces log files that document, at the household level, the issuance of warnings, the receipt of these warnings, and the subsequent actions undertaken in response. Action maps are stored in a Geoparquet file format, and the most relevant attributes (columns) of an action map are:

GEODATAFRAME COLUMNS	DESCRIPTION
GEOMETRY	Location of the household
POSTCODE	Postcode of the admin unit in which the household is located
WARNING_REACHED	Whether a warning reached the household
WARNING_LEVEL	Which warning level reached the household
EVACUATED	Whether the household evacuated after a warning

ELEVATED_POSSESSIONS	Whether the household protected their possessions after a warning

Damage maps: The damage maps represent the spatial distribution of modelled impacts on residential building structures and their contents resulting from simulated flood depths. As a baseline, these maps are generated using the ERA5-driven flood simulations under the assumption of no household-level protective actions. In addition, alternative damage maps are produced that incorporate the effects of household adaptation measures, leading to reduced simulated damages. For each scenario stored in the `action_maps`, the corresponding damage maps are also provided. The damage maps provide information at the individual-building scale and are stored in a GeoPackage (GPKG) file format:

GEODATAFRAME COLUMNS	**DESCRIPTION**
GEOMETRY	Location of the building
POSTCODE	Postcode of the admin unit in which the building is located
OBJECT_TYPE	Whether the households living in the building took protective actions
MAXIMUM_DAMAGE	The maximum damage per m^2
DAMAGE	The simulated damage of the building
ACTION_LEAD_TIME	The timing of households' decisions to undertake mitigation actions substantially influences the extent to which their overall environmental impact can be reduced
ELEVATED_POSSESSIONS	Whether households decided to protect their possessions inside the building

The aggregated damage to residential contents and building structures is quantified by loading the damage maps corresponding to the various scenarios, together with the ERA5 baseline, within the script `RP_Processing_Model_Results_And_Figures.py`. In this section of the code, a comma-separated values (CSV) file, entitled `Tab5_damage_reduction_all_scenarios.csv`, is generated to store all computed total damage estimates.

Reproducing Manuscript Results

Users may reproduce all results using one of the following approaches: 1) Employ the provided model simulations directly to regenerate all results (recommended). 2) Clone the `geb` repository, use the existing models, rerun the model simulations, and subsequently regenerate all results or 3) Develop an entirely new model, perform all corresponding simulations, and then reproduce all results.

Use the existing model simulations to recreate all the results

Based on the primary output files, some outputs are further processed to retrieve the results as presented in the report. To create all processed outputs, execute the following script from the directory: ./output analysis scripts/RP_Processing_Model_Results_And_Figures.py. This script (1) generates figures presented in the report and (2) can be used to fill tables. Note that, at each processing stage within the script, all available forecasts are included by default, even when the analysis in the report focuses on specific forecast dates. However, it is possible to restrict the execution of the processing steps such that they are applied exclusively to a selection of forecast dates. This improves the runtime slightly. The runtime to reproduce all the processed results for all the forecast dates will take up to 12-14 hours, tested on a Microsoft Surface Laptop 6 with 16 GB RAM. The total storage required (the entire reproducibility package) is up to 21 GB, of which about 1GB is for the output analysis folder.

First, the DEM and land cover maps are produced (few minutes run time), after which the warnings are generated, visualized, and validated (8-10 hours run time). Subsequently, the household action maps are evaluated for the different scenarios (with computational runtimes of approximately 3–4 hours), the damage maps with household actions are analysed (few minutes run time) and an initial threshold sensitivity analysis is conducted (with runtimes on the order of a few minutes). Given the long runtime, it is recommended to execute the script in separate steps by commenting out sections that have been successfully executed. These outputs are stored in the directory `./output analysis scripts/output`. Also, we included all the presented figures of the report in this directory as a png file.

Clone the `geb` repository, use the existing models, rerun the model simulations, and subsequently regenerate all results

To reproduce the model simulations, the `geb` repository must be cloned from the designated GitHub source, and the corresponding `geb` Python package must be installed. The detailed installation and usage procedures are provided in the manual included in the repository and should be followed carefully. Once the repository has been cloned and the package has been correctly installed, the pre-existing model configurations can be employed to execute the simulations. Please note that the current simulations are executed on a high-performance computing (HPC) system with substantially greater working and storage memory resources. Replicating all model simulations on a standard local laptop (16GB RAM) may therefore require computation times on the order of several weeks. Nevertheless, the model simulations can be conducted by following the procedure outlined in the "Executing the Model" section.

Develop an entirely new model, perform all corresponding simulations, and then reproduce all results

To recreate the model, the repository must be cloned, and the Python package must be installed. In contrast to the reproduction of the model simulations, this procedure additionally requires recompilation of the complete model in this study region. For this purpose, the instructions provided in the GitHub manual must be followed. Also, note that compilation of the model on a HPC system typically requires several days and that the GitHub is continuously evolving. Consequently, when an updated model is developed a few years later, its architecture can be refined, which may lead to different results. Furthermore, constructing the model on a local laptop may require up to several days of computation.

Data Availability

This section outlines where and how the data supporting the findings of the study can be accessed and used.

Summary of Data Availability

• ☒ All data are publicly available.

• ☐ Some data cannot be made publicly available.

• ☐ No data can be made publicly available.

Statement about Rights

• ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

• ☐ I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package

Data sources

To construct the complete model framework, multiple data sources are utilized. As GEB is an open-source model, it is publicly accessible to all interested users. All input files and their corresponding data sources are systematically catalogued below, as used in the project branch hosted on GitHub: https://github.com/GEB-model/GEB/blob/Roy_Brazil_floods_meteorological_warnings/geb/build/data_catalog/__init__.py

It is important to note that certain catalogue keys correspond to the same underlying source (e.g., GADM, MERIT Basins, MERIT Hydro, and the GLOBGM variables); accordingly, these were consolidated into a single data-source entry. In addition, where available, we recorded the versions of the datasets used and interpreted the access date as the time at which the models included in the reproducibility package were built. It is also important to note that not all data sources are utilized in the study configuration; rather, they are automatically configured by default during the construction of GEB.

1. ISIMIP CO2

Filename(s)	CO2_ppm.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	isimip_co2
Source: "name"	ISIMIP CO2
Source: "author"	Matthias Büchner, Christopher P.O. Reyer
Source: "url"	https://data.isimip.org
Source: "license"	CC BY-SA 4.0
Source: "paper_doi"	10.1007/s10584-011-0156-z

2. ERA5

Filename(s)	All meteorological variables {meteorological_variable}.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	era5
Source: "url"	https://data.earthdatahub.destine.eu/era5/reanalysis-era5-land-no-antartica-v0.zarr
Source: "name"	ERA5
Source: "author"	ECMWF

3. ECMWF Forecasts

Filename(s)	ecmwf_{forecast_date}_{forecast_model}_{forecast_resolution}_{forecast_horizon}h_{forecast_timestep_hours}h.grb
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	ecmwf_forecasts

Source: "name"	ECMWF Forecasts Operational Archive
Source: "author"	ECMWF
Source: "url"	https://www.ecmwf.int/en/forecasts/access-forecasts/access-archive-datasets
Source: "license"	https://www.ecmwf.int/en/forecasts/accessing-forecasts/licences-available

4. SoilGrids

Filename(s)	{variable}_{depth}_mean.vrt
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	soilgrids
Source: "name"	SoilGrids
Source: "author"	ISRIC - World Soil Information
Source: "license"	CC BY 4.0
Source: "url"	https://soilgrids.org
Source: "version"	2020
Source: "paper_doi"	10.5194/soil-2020-65

5. WHYMAP

Filename(s)	whymap.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	why_map
Source: "name"	WHYMAP
Source: "author"	BGR (Federal Institute for Geosciences and Natural Resources)
Source: "license"	CC BY-SA
Source: "url"	https://www.whymap.org/whymap/EN/Maps_Data/Gwr/gwr_node_en.html

6. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide

Filename(s)	lowder_farm_size_distribution.xlsx
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	lowder_farm_size_distribution
Source: "name"	The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide
Source: "author"	Lowder et al. (2016)
Source: "paper_doi"	10.1016/j.worlddev.2015.10.041
Source: "license"	CC BY-NC-ND 4.0
Source: "url"	https://doi.org/10.1016/j.worlddev.2015.10.041

7. GEBCO

Filename(s)	gebco.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	gebco
Source: "name"	GEBCO
Source: "author"	Weatherall et al (2020)
Source: "license"	https://www.gebco.net/data-products/gridded-bathymetry/terms-of-use

Source: "url"	https://www.gebco.net/
Source: "paper_doi"	10.5285/a29c5465-b138-234d-e053-6c86abc040b9

8. Global Runoff Data Centre

Filename(s)	GRDC.zip
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	GRDC
Source: "name"	Global Runoff Data Centre
Source: "author"	Global Runoff Data Centre
Source: "license"	https://grdc.bafg.de/downloads/policy_guidelines.pdf

9. Global Map of Irrigation Areas – groundwater

Filename(s)	global_irrigation_area_groundwater.asc
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	global_irrigation_area_groundwater
Source: "name"	Global Map of Irrigation Areas
Source: "author"	Siebert et al. (2010)
Source: "paper_doi"	10.5194/hess-14-1863-2010
Source: "license"	CC BY 4.0
Source: "url"	https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/

10. Global Map of Irrigation Areas – surface water

Filename(s)	global_irrigation_area_surface_water.asc
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	global_irrigation_area_surface_water
Source: "name"	Global Map of Irrigation Areas
Source: "author"	Siebert et al. (2010)
Source: "paper_doi"	10.5194/hess-14-1863-2010
Source: "license"	CC BY 4.0
Source: "url"	https://www.fao.org/aquastat/en/geospatial-information/global-maps-irrigated-areas/latest-version/

11. GLObal Groundwater Model (GLOBGM)

Filename(s)	hydraulic_conductivity_globgm.nc specific_yield_aquifer_globgm.nc water_table_depth_globgm.tif head_lower_layer_globgm.tif head_upper_layer_globgm.tif recession_coefficient_globgm.nc thickness_confining_layer_globgm.nc total_groundwater_thickness_globgm.nc dem_globgm.nc
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	hydraulic_conductivity_globgm specific_yield_aquifer_globgm water_table_depth_globgm head_lower_layer_globgm head_upper_layer_globgm recession_coefficient_globgm

	thickness_confining_layer_globgm total_groundwater_thickness_globgm dem_globgm
Source: "url"	https://geo.public.data.uu.nl/vault-globgm/research-globgm%5B1669042611%5D/original/input/version_1.0/
Source: "author"	Verkaik et al. (2024)
Source: "paper_doi"	10.5194/gmd-17-275-2024
Source: "license"	CC BY 4.0

12. Global Data Lab

Filename(s)	gdl_regions_v4.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	GDL_regions_v4
Source: "name"	Global Data Lab
Source: "author"	Radboud University
Source: "license"	https://globaldatalab.org/termsofuse/
Source: "url"	https://globaldatalab.org/mygdl/downloads/shapefiles/

13. World Bank Inflation Data

Filename(s)	wb_inflation_rate.csv
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	wb_inflation_rate
Source: "name"	World Bank Inflation Data
Source: "author"	The World Bank
Source: "license"	CC BY 4.0

14. World Bank – Official exchange rate (LCU per US$, period average)

Filename(s)	world_bank_price_ratio.csv
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	world_bank_price_ratio
Source: "name"	Official exchange rate (LCU per US$, period average)
Source: "author"	The World Bank
Source: "license"	CC BY 4.0
Source: "url"	https://data.worldbank.org/indicator/PA.NUS.PPPC.RF

15. World Bank Local Currency Unit per US Dollar Data

Filename(s)	world_bank_LCU_per_USD.csv
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	world_bank_LCU_per_USD
Source: "name"	World Bank Local Currency Unit per US Dollar Data
Source: "author"	The World Bank
Source: "license"	CC BY 4.0
Source: "url"	https://data.worldbank.org/indicator/PA.NUS.FCRF

16. ESA WorldCover

Filename(s)	classification.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	esa_worldcover_2021
Source: "name"	ESA WorldCover
Source: "author"	European Space Agency (ESA)
Source: "version"	v200
Source: "license"	CC BY 4.0

17. HydroLAKES

Filename(s)	hydrolakes.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	hydrolakes
Source: "name"	HydroLAKES
Source: "author"	Arjen Haag
Source: "version"	v10.0
Source: "license"	CC BY 4.0

18. GADM

Filename(s)	gadm_level0.parquet gadm_level1.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	GADM_level0 GADM_level1
Source: "name"	GADM
Source: "author"	GADM
Source: "version"	4.1
Source: "license"	https://gadm.org/license.html

Citation: GADM. GADM.

19. MERIT Basins

Filename(s)	merit_basins_catchments.parquet merit_basins_rivers.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	merit_basins_catchments merit_basins_rivers
Source: "name"	MERIT Basins
Source: "author"	Lin et al.
Source: "version"	v0.7
Source: "license"	CC BY-NC-SA 4.0

20. MERIT Hydro

Filename(s)	merit_hydro_dir.zarr merit_hydro_elv.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	merit_hydro_dir merit_hydro_elv
Source: "name"	MERIT Hydro
Source: "author"	Yamazaki et al.
Source: "version"	2019
Source: "license"	CC BY 4.0 or ODbL 1.0

21. FABDEM

Filename(s)	fabdem.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	fabdem
Source: "name"	FABDEM
Source: "author"	Hawker et al. (2022)
Source: "version"	1-2
Source: "license"	CC BY-NC-SA 4.0
Source: "url"	https://data.bris.ac.uk/data/dataset/25wfy0f9ukoge2gs7a5mqpq2j7

Source: "paper_doi"	10.1088/1748-9326/ac4d4f

22. MERIT-SWORD

Filename(s)	merit_sword.zarr
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	merit_sword
Source: "name"	MERIT-SWORD
Source: "version"	v0.4
Source: "license"	CC BY-NC-SA 4.0
Source: "url"	doi.org/10.5281/zenodo.14675925

23. SWORD

Filename(s)	sword.gpkg
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	sword
Source: "name"	SWORD
Source: "version"	v16
Source: "license"	CC BY 4.0
Source: "url"	doi.org/10.5281/zenodo.14727521

24. OpenBuildingMap

Filename(s)	open_building_map.parquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	open_building_map
Source: "name"	OpenBuildingMap
Source: "author"	Oostwegel et al. (2025)
Source: "version"	1
Source: "license"	CC BY-NC-SA 4.0
Source: "url"	https://dataservices.gfz-potsdam.de/panmetaworks/showshort.php?id=45829b80-e892-11ef-914a-f12b0080820d
Source: "paper_doi"	https://doi.org/10.5880/GFZ.LKUT.2025.002

25. OpenStreetMap

Filename(s)	{asset}.geoparquet
Last access date / date of model build	26 February 2026
GEB catalogue key(s)	open_street_map
Source: "name"	OpenStreetMap
Source: "author"	OpenStreetMap contributors
Source: "license"	ODbL 1.0
Source: "url"	https://www.openstreetmap.org/copyright

With respect to the processing of the results stored in the output analysis scripts directory, only the observed flood extent file, recorte_mancha_guaiba.shp, and the administrative units file, RS_setores_CD2022_epsg4326.geoparquet, are derived from external data sources that differ from those documented for the remaining input datasets.

Da Silva, L. A., Cuartas, L. A., Gonçalves, D. A., & DOS SANTOS ALVALÁ, R. C. (2024). MAPEAMENTO DAS ÁREAS IMPACTADAS PELA INUNDAÇÃO DE MAIO DE 2024 NA REGIÃO HIDROGRÁFICA DO GUAÍBA (ISSN: 2764-9040). ABRHidro. https://anais.abrhidro.org.br/job.php?Job=16818

Instituto Brasileiro de Geografia e Estatística (IBGE). (2023). Censo Demográfico 2022: Tipos de domicílios, moradores e densidade demográfica - Primeiros resultados [Demographic Census 2022: Household types, residents, and demographic density - First results]. https://biblioteca.ibge.gov.br/visualizacao/livros/liv102011.pdf

List of Exhibits

The provided code reproduces:

- ☐ All numbers provided in text in the paper
- ☐ All tables and figures in the paper
- ☒ Selected tables and figures in the paper, as explained and justified below

The table below provides a list of all exhibits in the manuscript. Checkmarks indicate those exhibits that are reproduced by this reproducibility package. Figure outputs are stored in `./output analysis scripts/output/`. The table data reproduced by the code are shown as command-line print outputs in the VS Code terminal or as csv files within the corresponding directory. All processed outputs from the model simulations are generated with the script that can be found in the following directory: `./output analysis scripts/RP_Processing_Model_Results _And_Figures.py`.

For the fifth column in Table 1 we derived the 36% from Marengo et al. (2025), the other percentages are based on assumptions as described in the manuscript:

Marengo, J. A., Alcantara, E. H., Moraes, O. L., Martinez, R., Seluchi, M., Alvalá, R. C., Dolif, G., & Goncalves, D. (2025). Early Warning Services for Disaster Risk Reduction in Brazil: The experience of CEMADEN during the floods of Rio Grande do Sul of May 2024. International Journal Of Disaster Risk Reduction, 126, 105645. https://doi.org/10.1016/j.ijdrr.2025.105645

For the third row in Table F.1, we used outputs from the following literature:

Governo Do Estado – Rio Grande Do Sul (2025). MUPRS: Mapa Unico Plano Rio Grande. https://mup.rs.gov.br/. Last access: 17 December 2025.

World Bank. (2025b). Disaster Risk Finance Diagnostic — Brazil. International Bank for Reconstruction and Development. Retrieved December 17, 2025, from https://documents1.worldbank.org/curated/en/099011226122036005/pdf/P508343-6274ad72-a6be-40fe-aa1e-3e19bcb722e4.pdf

	EXHIBIT NAME	OUTPUT	FOOTNOTES
[X]	**Figure 1**	Fig1_rio_sinos_catchment_overview_map.png	[1]
[X]	**Figure 2**	Fig2_ERA5_vs_ECMWF_Control_&_Probabilistic_Cumulative_Precipitation_Forecasts.png	[2]
[]	**Figure 3**	Fig3_Methodological framework.png	[3]
[X]	**Figure 4**	Fig4_rio_sinos_med_forecast_landcover_map.png	
[X]	**Figure 5**	Fig5_spatial_flood_probability__map_validation.png	[4]
[X]	**Figure 6**	Fig6_flood_forecast_performance_metrics_of_full_ensemble.png	[5]
[X]	**Figure 7**	Fig7_IbW_warning_generation_system.png	[6]
[X]	**Figure 8**	Fig8_integrated_warning_system_all_forecasts_20240426_20240501_medium_severe.png	
[X]	**Figure 9**	Fig9_spatial_warning_validation.png	[7]
[X]	**Figure 10**	Fig10_actions_over_time_white_cumulative.png	[8]
[X]	**Figure 11**	Fig11 warned outcomes bar evaluation.png	[9]
[X]	**Figure 12**	Fig12_evacuation_performance_analysis_sev.png	[10]
[X]	**Figure 13**	Fig13_comparison_severe_IbW_realistic_vs_perfect.png	[11]
[]	**Table 1**	n/a	[12]
[]	**Table 2**	n/a	[13]
[X]	**Table 3**	Prints in the command line	
[X]	**Table 4**	n/a	[14]
[X]	**Table 5**	n/a	[15]

[1] Compiled manually in Canva by merging Fig1_countries_and_state_map.png with Fig1_rio_sinos_med_forecast_dem_map.png
[2] Created as primary output when the model simulations are evaluated (See section 'Output files')
[3] Compiled manually in Canva
[4] Created as primary output when the model simulations are evaluated (See section 'Output files') and compiled in Canva to present two forecasts.
[5] Created as primary output when the model simulations are evaluated (See section 'Output files') and compiled manually in Canva.
[6] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results') and compiled manually in Canva
[7] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results') and compiled manually in Canva
[8] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results')
[9] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results')
[10] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results') and compiled manually in Canva
[11] Created as processed output when the model simulations are processed (See section 'Reproducing Manuscript Results')

The supplementary figures and tables listed below were generated using scripts analogous to those described above.

[12] Compiled manually, with the final column showing output from the literature
[13] Compiled manually
[14] Compiled manually using the metric boxes available in the figure analysis located at `. output analysis scripts/output/action_maps/evacuation_analysis`. Please set the parameters '**action_type**' and '**metric_box_included**' inside the 'Evaluate_actions_households' function to 'evacuation' and the parameter '**metric_box_included**' to **True**. The table can be filled in using these figures.
[15] Compiled manually with for column 7 using the 'Tab5_damage_reduction_all_scenarios.csv' file which can be found in `. output analysis scripts/output`. Moreover, the other columns are a summary of all the different results presented above in the other figures and tables.

	EXHIBIT NAME	OUTPUT	FOOTNOTES
[X]	**Supplementary Figure 1**	AppendixA_ERA5_baseline_flood_extent_validation.png	[16]
[X]	**Supplementary Figure 2**	AppendixB_FW_warning_generation_system.png	[17]
[X]	**Supplementary Figure 3**	AppendixC_ibf_warning_zoomplot.png	[18]
[X]	**Supplementary Figure 4**	AppendixD_meteorological_warning_observed_floodextent.png	[19]
[X]	**Supplementary Figure 5**	AppendixE1_grouped_action_counts.png	[20]
[X]	**Supplementary Figure 6**	AppendixE2_spatiotemporal_hexbin_density_of_timing_actions.png	[21]
[X]	**Supplementary Figure 7**	AppendixG1_building_fraction_hit_r1_map_02052024	[22]
[X]	**Supplementary Figure 8**	AppendixG2_warned_outcomes_bar_evaluation.png	[23]
[X]	**Supplementary Figure 9**	AppendixG3_actions_over_time_white_cumulative.png	[24]
[X]	**Supplementary Table 1**	Table E.1	[25]
[]	**Supplementary Table 2**	Table F.1	[26]

[16] Created as output when the model simulations are evaluated (See section ‘Output files’).
[17] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[18] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[19] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva. Please specify the parameter **forecast_of_interest** in the 'create_meteorological_warnings_from_classes' function to retrieve both subplots. The parameter can be changed to 20240426T000000 and to 20240501T000000.
[20] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[21] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[22] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[23] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[24] Created as primary output when the model simulations are evaluated (See section ‘Reproducing Manuscript Results’) and compiled manually in Canva.
[25] Compiled manually with the outputs of: ` ".\output analysis scripts\output\action_maps\protection_analysis`. Please set the parameters '**action_type**' and '**metric_box_included**' inside the 'Evaluate_actions_households' function to 'protection' and the parameter '**metric_box_included**' to **True**. The table can be filled in using these figures.
[26] "Compiled manually, with the second row showing output from the GEB model, and the third shows output from the literature. The calculated damage can be derived using the damage datasets located in: " \reproducibility package\models\rio_sinos_med_forecast\base\output\action_maps\No_action_maps_ERA5”. The total damage is computed by summing the values in the damage column. The number of affected buildings is determined by counting the records for which the damage value exceeds 0. However, the numbers can also be found in the csv file: "reproducibility package\output analysis scripts\output\Tab5_damage_reduction_all_scenarios.csv". Moreover, the flooded area is derived during the simulation of the baseline and can be checked in the file: " reproducibility package\output analysis scripts\output\baseline_ERA5_flood_map\20240426T000000 - 20240507T000000_performance_metrics.txt" ".