MOHID-LAND PRE-CONFIGURED HYDROLOGICAL MODEL FOR THE PEDRO DO RIO WATERSHED: READY-TO-RUN SIMULATIONS AND RESULTS

1. Introduction

The dataset presented here provides a fully configured version of the MOHID-Land hydrological model, specifically designed for the Pedro do Rio watershed, situated in the mountainous region of Petrópolis, Rio de Janeiro, Brazil. This preconfigured model includes all the necessary setup files and configurations, enabling users to easily run simulations and obtain results for various scenarios (S1 to S6). These scenarios are detailed in the paper "Enhancing River Flow Predictions in MOHID-Land Through Integration of Gridded Soil Data and Hydraulic Parameters Using the MOHID Soil Tool".

This dataset is designed to facilitate research by offering a ready-to-run hydrological model, ensuring reproducibility, transparency, and productivity in simulations related to river flow dynamics. By providing a standardized setup, this dataset ensures that simulations can be consistently reproduced, facilitating the comparison of results and enhancing the credibility of the research. The transparent configuration allows other researchers to easily understand and replicate the setup, ensuring the integrity and consistency of the research findings.

The MOHID-Land model is structured by year, and each simulation should be executed sequentially for each year to ensure the proper flow of results. The simulation output is organized by scenario and year, with results stored in the Results folder. Detailed logs of each simulation can be accessed in the Logs directory, offering insight into the process.

This dataset aims to serve as a valuable resource for researchers and practitioners in the field of hydrological modeling, particularly those focused on watershed management and river flow predictions. By offering preconfigured simulations and relevant datasets, this tool minimizes setup time, allowing users to focus on analysis and interpretation.

2. Pre-configured MOHID-Land model inputs

2.1. Digital Elevation Model

Topography in MOHID-Land is represented through the implementation of elevation data via a Digital Elevation Model (DEM). A DEM can be defined as a structured model of rows and columns (matrix), georeferenced, which contains elevation records for each pixel in the grid, with a given spatial resolution. Topography is crucial as it provides the model with (i) the delineation of the watershed, (ii) the drainage network, and (iii) the terrain slope.

There are several radar mapping products available with different spatial resolutions. The elevation input data for this study was sourced from the 22s435 sheet of the Topodata project, which covers the entire study area and is available for download on the INPE website (http://www.dsr.inpe.br/topodata/). This DEM is provided in GeoTIFF format and is a rescaling of the Shuttle Radar Topography Mission (SRTM) data to a spatial resolution of 30 m. The original SRTM data, aimed at providing global elevation data, had a resolution of 90 m (VALERIANO & ROSSETI, 2011).

For the topography construction, a regular grid with $\Delta x = \Delta y = 0.002°$ (~200m) was used, forming a mesh of 160 rows and 200 columns, totaling 32,000 cells within the domain. This value of 0.002° was chosen to achieve a spatial resolution that ensures the basin shape and drainage network resemble real-world data. The Topodata DEM was then interpolated to the defined grid, generating the topography to be implemented in the model.

After the topography generation, the "Remove Depression" tool was applied, which mathematically corrects terrain irregularities to allow for the creation of flow accumulation lines, forming the drainage network of the model. Thus, the final **Topograph_ND_ND.dat** file is generated.

The next step involves the delineation of the watershed and the drainage network. It is important to note that once the watershed is delineated, the model performs calculations only within its boundaries, disregarding cells outside the watershed area.

For watershed delineation and drainage network construction, the Delineate Basin tool was used. This tool requires the user to provide an outlet point and then traces the upstream topography to construct the watershed. The outlet was set at the location of the Pedro do Rio stream gauge station, managed by the Brazilian National Water and Basic Sanitation Agency (ANA) (station code 58405000), with the objective of obtaining historical streamflow data for model calibration and validation.

The minimum accumulation area required to generate runoff is defined by a Threshold value (in square meters), which is user-defined. The logic behind this parameter is that the smaller the threshold, the more detailed the drainage network will be, as smaller contributing areas will be considered for generating flow. In this project, a Threshold of 2,000,000 m^2 was assigned, an empirically determined value aimed at preserving the location and number of channels in the actual drainage network.

As a result, a Strahler order 4 drainage network was generated, consisting of 776 nodes, while the watershed itself was delineated with a spatial resolution of 200 m. Thus, the final **DrainageNetwork.dnt** and **Delineation.xy** files are generated.

The terrain slope is calculated based on the elevation and grid spacing. This information is later used to define the spatially distributed soil thickness within the watershed, where thickness decreases as slope increases. In this regard, the user defines an estimated soil depth, and the model then computes the depth for each cell accordingly.

The available files can be found on the following paths:

- **/GeneralData/DigitalTerrain/Topograph_ND_ND.dat**
- **/GeneralData/DigitalTerrain/Delineation.xy**
- **/GeneralData/DigitalTerrain/DrainageNetwork.dnt**

2.2. Cross-sections

For the calculation of streamflow, which is obtained by multiplying the flow velocity in the channel by the cross-sectional area, MOHID-Land requires the implementation of cross-sections for each node of the drainage network.

The initial cross-sections for the model were obtained from a monitoring campaign commissioned by the Piabanha River Basin Committee and carried out by the company Hydroscience between 2019 and 2021. For the main channel of the Piabanha River, cross-section measurements were conducted in four quarterly campaigns between August 2019 and May 2020. Meanwhile, the cross-sections of tributaries of the Piabanha River were measured in two campaigns, in January 2021 and June 2021. The report detailing the cross-sections includes information on wetted area, maximum depth, and section width for the measurement day, as well as a schematic representation with elevation levels defining the river channel. The cross-sections were estimated through visual interpretation of this schematic, considering the entire available channel at each measurement point.

MOHID-Land allows for the insertion of cross-section dimensions either according to Strahler order, where each order in the modeled drainage network receives the same dimensions, or based on the drained area. In this study, the drained area method was chosen for cross-section implementation. The cross-sections for each node between measured sections were interpolated based on available field data. In total, 13 cross-sections were measured within the experimental watershed, and their respective drainage areas were calculated.

The model allows for the insertion of either rectangular or trapezoidal cross-sections. However, given that trapezoidal cross-sections more closely resemble natural watercourses, this shape was chosen for implementation. The dimensions of the cross-sections, estimated based on visual interpretation of each schematic, are presented in Table 1. Since a direct relationship between the wetted area of the cross-section and the drainage area was not observed for all measured points, only the points where this relationship was maintained were selected (points: PCF1, ARR1, CDD1, P1, P2 and P4, according to Hydroscience (2020, 2021) reports).

Table 1. Measured cross-sections available in the experimental basin as a function of drained area.

Point	**River**	**Cross-Section area (m²)**	**Height (m)**	**Top width (m)**	**Bottom width (m)**	**Drainage area (km²)**
PLT1	Palatino	9.20	2.00	7.20	2.00	5357.00
PCF1	Poço Ferreira	**5.60**	**2.00**	**4.60**	**1.00**	**6322.00**
QTD1	Quitandinha	2.55	1.00	3.60	1.50	7966.00
JAC1	Jacó	3.25	1.00	4.00	2.50	12140.00
ARR1	Araras	**7.10**	**2.00**	**5.10**	**2.00**	**12599.00**
CRV1	Carvão	2.33	1.00	3.65	1.00	17007.00
ITM1	Itamarati	5.10	2.00	3.60	1.50	32787.00
CDD1	Cidade	**10.70**	**2.00**	**8.70**	**2.00**	**34399.00**
P1	Piabanha	**43.00**	**5.00**	**11.20**	**6.00**	**49178.00**
STA1	Santo Antônio	9.50	2.00	7.50	2.00	93165.00
P2	Piabanha	**50.00**	**5.00**	**14.00**	**6.00**	**103628.00**
P3	Piabanha	65.00	5.00	20.00	6.00	264562.00
P4	Piabanha	**62.50**	**5.00**	**19.00**	**6.00**	**396687.00**

P4, being located near the outlet of the experimental basin, was designated as the cross-section for the basin's outlet, with the total drainage area of the experimental basin assigned to this section (424.10 km²). Conversely, the smallest cross-section corresponds to a drainage area of 2 km², which was the threshold used in the basin delineation process. Therefore, the final cross-sections assigned in the model are presented in Table 2.

Table 2. Trapezoidal cross-sections assigned in the model for each drained area.

Cross- Section area (m²)	**Height (m)**	**Height (m)**	**Height (m)**	**Height (m)**
3.00	1.50	3.00	1.00	2.00
5.60	2.00	4.60	1.00	6.32
7.10	2.00	5.10	2.00	12.60
10.70	2.00	8.70	2.00	34.40
43.00	5.00	11.20	6.00	49.18
50.00	5.00	14.00	6.00	103.45
62.50	5.00	19.00	6.00	414.06

The cross-sections attributed in the **DrainageNetwork.dnt** file can be found at the following path: /**GeneralData/DigitalTerrain/DrainageNetwork.dnt**.

2.3. Land Use

In MOHID-Land, the land use and land cover map is essential for two key processes: (i) calculating surface runoff by providing flow resistance and (ii) supplying the model with surfaces that contain vegetation and therefore promote infiltration. In this study, the land use and land cover map provided by the MapBiomas project (SOUZA et al., 2020) for the year 2021 was used, available for download at https://brasil.mapbiomas.org/colecoes-mapbiomas/. MapBiomas utilizes imagery from the Landsat mission, with a spatial resolution of 30 meters, classified through a supervised approach using the Random Forest algorithm on the Google Earth Engine platform, achieving an average accuracy of 89% (SOUZA et al., 2020).

2.3.1. Manning's Roughness Coefficient (Overland Coefficient)

The determination of Manning's coefficient is related to flow resistance, where higher values correspond to greater roughness and, consequently, higher resistance to flow. Conversely, lower values indicate smoother surfaces that allow for faster flow velocities.

Since Manning's coefficient is estimated based on the surface characteristics of the watershed, the values were assigned according to the coefficient ranges for each land use type, following the classifications by Chow (1959) and Soliman et al. (2022), in combination with the land use and land cover survey for the region produced by MapBiomas for the year 2021. The classes were then interpolated to match the spatial resolution of the model grid, ensuring a spatialized roughness coefficient.

The correlation between land use classes identified from MapBiomas mapping and their assigned Manning’s coefficient is presented in Table 3.

Table 3. Assigned Manning’s Coefficients for Each MapBiomas (2021) Land Use class.

MapBiomas Code	**Land Use**	***Manning***
3	Formação Florestal	0,160
15	Pastagem	0,038
21	Mosaico de Usos Agrícolas	0,045
24	Área Urbanizada	0,040
29	Afloramento Rochoso	0,030

Source: Adapted from Chow (1959) and Soliman et al. (2022).

The file containing the Manning coefficient assigned based on land use for each cell of the 2D domain can be found at the following path: **/GeneralData/Manning/Manning.dat**

2.3.2. Vegetation

The 2021 MapBiomas land use and land cover map was used to identify vegetation classes. Three vegetation classes were identified: Forest Formation, Pasture, and Agricultural Use Mosaic. The MapBiomas IDs were maintained and correlated with the IDs from the MOHID vegetation database. For classes that do not correspond to vegetation, a value of 0 was assigned in the database (indicating the absence of vegetation). The correlation is shown in Table 4.

Table 4. Correlation between vegetation classes for MapBiomas and MOHID database IDs.

MapBiomas Code	**MOHID *database***	**Land Use**
3	8	Formação Florestal
15	12	Pastagem
21	1	Mosaico de Usos Agrícolas
24	0	Área Urbanizada
29	0	Afloramento Rochoso

For the simulation of vegetation growth and soil water uptake by roots, a set of characteristics must be specified for each vegetation class present in the basin. Therefore, a literature review was conducted to establish some of the characteristics of the classes mapped by MapBiomas, as presented in Table 4.

The file **VegetationID.dat**, which contains the IDs of the MapBiomas classes and indicates the vegetation present in each cell of the 2D domain, can be found at the following path: **/GeneralData/Vegetation/VegetationID.dat**

MOHID-Land allows studies at the spatial resolution of agricultural plots, enabling detailed assessments of water demand and crop water balance. However, since the objective of this study is to evaluate hydrological behavior at the watershed scale, agricultural activity was generalized to a relevant crop within the basin. As the land use and land cover map does not specify the type of crop present in the "Agricultural Use Mosaic" class, data from IBGE (2022) for the municipality of Petrópolis were consulted to identify the predominant crop in the basin. The analysis revealed a high incidence of vegetable production, with tomatoes standing out, totaling 625 tons. Consequently, tomato cultivation was prioritized in this study to represent the agricultural area of the basin. It is important to note that for watershed-scale studies aimed at assessing the hydrological behavior of crops in average terms, another relevant crop could be used, with its water demand being adjusted during the calibration process.

Regarding pasture, Martins et al. (2007) indicate that grasses of the *Brachiaria* species are predominant in Brazil, accounting for more than 80% of the total volume of seeds traded. The widespread use of this species is attributed to its good adaptability and ability to cover poor and acidic soils. Therefore, *Brachiaria* was used to generalize the pasture areas in the study basin.

Additionally, a significant portion of the basin is covered by the Atlantic Forest.

Once the vegetation corresponding to each land use class was identified, the development parameters for each vegetation type were defined, as described in Table 5.

Table 5. Crop development parameters for each vegetation class in the watershed

Crop	**Variable**	**Value**	**Source**
Tomato	Base Temperature	5°C	Clemente e Boiteux (2012)
	Optimal Temperature	20°C	
	Growth Cycle	120 days	
	Root Depth	0.50m	
	Planting Date	February	
	Harvest Date	May	
	Maximum LAI	3.3*	Perin (2021)
	Canopy Height	2.3m*	
	Initial Kc	0.60	Allen *et al.*, (1998)
	Mid-season Kc	1.15	
	Late-season Kc	0.90	
	Base Temperature	15°C	Santos *et al.*, (2008)

Pasture	Optimal Temperature	30°C	
	Root Depth	40cm	Beloni *et al.*, (2016)
	Maximum LAI	5.7 **	Molan (2004)
	Canopy Height	40cm	
	Planting Date	October	Zimmer *et al.* (1983)
	Initial Kc	0.40	Allen *et al.*, (1998)
	Mid-season Kc	1.05	
	Late-season Kc	0.85	
Forest	Base Temperature	0°C	Biudes *et al.*, (1998)
	Optimal Temperature	28°C	
	Root Depth	5m	Canadell *et al.*, (1996)
	Canopy Height	20m	Castro (2008)
	Maximum LAI	6	Pezzopane *et al.*, (2022)
	Minimum LAI	3	
	Initial Kc	0.95***	Allen *et al.*, (1998)
	Mid-season Kc	1.00***	
	Late-season Kc	1.00***	

*Considering the Italian tomato variety for summer planting.
** LAI is defined as the total leaf area per unit of soil area. Therefore, the value presented here refers to a well-developed pasture with no water deficit.
*** The Kc value for the Palm tree is considered, a typical tree of the Atlantic Forest biome.

The file **VegetationParameters.dat** contains the correlation between the MOHID-Land database ID and the land use classes defined in the VegetationID.dat file. It can be found at the following path: **/GeneralData/Vegetation/VegetationParameters.dat.**

The file **GrowthDatabase.dat,** which contains, among other parameters, the values for the parameters defined in Table 5, can be found at the following path: **/GeneralData/Vegetation/GrowthDatabase.dat.**

The values for the Feddes coefficients of soil water uptake for each crop were consulted in the literature and are described as shown in Table 6.

Table 6. Feddes Coefficients for soil water uptake for each crop.

Crop	Variable	Value	Source
Tomato	Feddes h1	-0.1	Hydrus1D*
	Feddes h2	-0.25	
	Feddes h3	-15	
	Feddes h4	-80	
Pasture	Feddes h1	-0.1	Hydrus1D*
	Feddes h2	-0.25	

	Feddes h3	-8	
	Feddes h4	-80	
Forest	Feddes h1	0	Grinevskii (2011)
	Feddes h2	-1	
	Feddes h3	-3.3	
	Feddes h4	-150	

Database from the model for water movement in soil, HYDRUS 1D (SIMUNEK et al., 1998).

2.4. Soil

The soil hydraulic parameters were calculated using the MOHID Soil Tool available at: https://github.com/dhiegosales/MOHID-SOIL-TOOL.

The input data as well as the outputs from the MOHID Soil Tool, which serve as inputs for MOHID-Land, can be accessed in the dataset: *Supplementary Soil Dataset for Enhancing River Flow Predictions in MOHID-Land Through Integration of Gridded Soil Data and Hydraulic Parameters Using the MOHID Soil Tool* available at: https://doi.org/10.5281/zenodo.14914611.

The ID corresponding to the soil type for each cell in the layer can be accessed (file .dat) in the dataset directory: **/GeneralData/PorousMedia/**.

The hydraulic coefficients of the Van Genuchten model for each cell in the 3D soil domain can be found in the file **PorousMedia_1.dat**, located in the **/data/** directory for the corresponding simulation and year.

2.5. Drainage network

For the drainage network, i.e., in the 1D domain, the Manning's coefficient is assigned to all the nodes in the network with a single value, using the keyword GLOBAL_MANNING. The initial value assigned was 0.035, as per the studies by Tavares et al. (2019) and Sales et al. (2021), for studies conducted in basins near the study region.

The value of the Global Manning coefficient can be changed in the file ***DrainageNetwork_1.dat***, located in the **/data** directory in the corresponding simulation and year.

2.6. Atmosphere

Since MOHID does not explicitly model the atmosphere, climate variables need to be imposed on the model. This section aims to describe the procedures used to select the atmospheric data sources utilized in the research.

All atmospheric variables are referenced through the **Atmosphere_1.dat** file, which is located in the **/data** directory corresponding to each simulation and year.

2.6.1. Solar radiation, relative humidity, wind speed, and temperature

Given that MOHID requires atmospheric data to calculate evapotranspiration, historical data on solar radiation, relative humidity, wind speed, and temperature were sought from available stations. Among the available operators, only INMET has series for these variables, with stations located near the study watershed, as described in Table 7. Despite the relative proximity to the Duque de Caxias station, it does not adequately represent the climate conditions of the region due to the considerable difference in altitude and the orographic barrier provided by the Serra do Mar Mountain range. On the other hand, the Paty do Alferes station is of recent installation and has limited historical data available.

Table 7. INMET automatic climatological stations.

Code	Station	Latitude	Longitude	Elevation (m)	Period
A603	Duque de Caxias - Xerém	-22.5897	-43.2822	22	2002-2023
A616	Pico do Couto	-22.4647	-43.2914	1777	2006-2023
A618	Teresópolis Parque Nacional	-22.4486	-42.9869	981	1992-2023
A624	Nova Friburgo Salinas	-22.3347	-42.6769	1070	2010-2023
A625	Três Rios	-22.0983	-43.2083	295	2016-2023
A629	Carmo	-21.9386	-42.6008	293	1990-2023
A637	Paty do Alferes - Avelar	-22.3472	-43.4177	508	2022-2023

The evaluation of the historical series from INMET stations showed large periods of data gaps, with no data for the period of interest (2006-2022), limiting their use as MOHID requires valid hourly data. An alternative for obtaining valid atmospheric parameter series is the use of global models. Among them, ERA5 provides hourly atmospheric data on a 0.25° x 0.25° grid, operational since 1950, making it a very versatile data source (HERSBACH et al., 2020). This dataset was chosen based on its already satisfactory results in simulations with MOHID-Land, as described in Oliveira et al. (2020) and Oliveira et al. (2023).

Regarding the use of global models, Horton (2022) states that processed reanalysis procedures are data sources for areas with scarce station coverage. Among global models, ERA5 stands out for its high spatial resolution globally and its ability to assimilate more data than its predecessor, ERA-Interim, with the inclusion of terrestrial radar data and new satellite sensors, as well as improved observation operators, allowing for better model-observation comparisons.

Supporting the good performance of the model, Yilmaz (2023), when evaluating surface temperature trends in Turkey, found a high degree of consistency between the ERA5 model results and the measured observations, suggesting that it can be used as a substitute for observations.

Given that precipitation is the most relevant atmospheric variable, and that other variables are related to the evapotranspiration process, this work considers the ERA5 model suitable to provide the atmospheric parameters for MOHID-Land.

Thus, a Python algorithm was developed for the use of the ERA5 API and subsequent processing of the historical data series, from 01/01/2006 to 31/12/2022. The daily variables in .netCDF format acquired were: total cloud cover [fraction], 2-meter temperature [K], 2-meter dewpoint temperature [K], surface solar radiation [Jm^{-2}], 10-meter wind u-component [ms^{-2}], and 10-meter wind v-component [ms^{-2}] (HERSBACH et al., 2023). Since the Pedro do Rio watershed is almost entirely within one ERA5 model pixel, the coordinates of the centroid for pixel were used to extract time series data for all variables. The values were then generalized for the entire watershed.

Relative humidity is not provided by ERA5 and was calculated from the relationship between the partial water vapor pressure and the saturation pressure as follows:

$$RH = \frac{e}{es} \tag{1}$$

$$e = 6.1078\, e^{\left(\frac{17.1 * T_d}{235 + T_d}\right)} \tag{2}$$

$$e_s = 6.1078\, e^{\left(\frac{17.1 * T}{235 + T}\right)} \tag{3}$$

where: RH is the relative humidity [fraction]; T_d is the dewpoint temperature [K]; T is the air temperature [K]; e is the partial water vapor pressure [hPa]; e_s is the saturation vapor pressure [hPa].

The air temperature at 2m was converted to Celsius as follows:

$$T[C] = T[K] - 273.15 \tag{4}$$

The surface solar radiation (RSS) was converted to Wm^{-2} as follows:

$$RSS[Wm^{-2}] = \frac{RSS[Jm^{-2}]}{3600} \tag{5}$$

The wind velocity profiles for the u and v components were converted from 10m to 2m, as per the equation (ALLEN *et al.*, 1998):

$$u_2 = u_{10} \frac{4.87}{ln(67.8*10)-5.42} \tag{53}$$

where: u_2 is the wind speed at 2m for the u and v components [m/s]; u_{10} is the wind speed at 10m for the u and v components [m/s].

The paths for the time series of all atmospheric variables, for the period from January 1, 2006, to December 31, 2022 in *.srm* format, can be found in the dataset.

- **/GeneralData/BoundaryConditions/01_Timeseries/CloudCover.srm**
- **/GeneralData/BoundaryConditions/01_Timeseries/RelativeHumidity.srm**
- **/GeneralData/BoundaryConditions/01_Timeseries/SolarRadiation.srm**
- **/GeneralData/BoundaryConditions/01_Timeseries/Temperature.srm**
- **/GeneralData/BoundaryConditions/01_Timeseries/WindComponentU.srm**
- **/GeneralData/BoundaryConditions/01_Timeseries/WindComponentV.srm**

2.6.2. Precipitation

Daily precipitation data were gathered from thirty-nine stations across the watershed, sourced from four institutions: ANA, CPRM, Rio de Janeiro State Environmental Institute (INEA), and the National Center for Natural Disaster Monitoring and Alerts (CEMADEN). Subsequently, the station timeseries were grouped into fifteen synthetic stations using a clustering method based on the median of neighboring stations. The set of all pluviometric stations used and the description of which synthetic station each monitoring station belongs to are described in Table 8.

Table 8. Selected Pluviometric Stations.

Operator	Station	Name	Latitude	Longitude	Synthetic Station
INEA	2243315	Alto da Serra	-22.5124	-43.1721	Morin
INEA	2243323	Araras	-22.4339	-43.2552	Araras
INEA	2243324	Barão do Rio Branco	-22.4903	-43.1852	Liceu
INEA	2243318	Correias	-22.4450	-43.1424	Correas

INEA	2243321	Independência	-22.5477	-43.2092	Quitandinha
INEA	2243319	Itaipava	-22.4056	-43.1028	Parque Petrópolis
INEA	2243329	LNCC	-22.5303	-43.2171	Quitandinha
INEA	2243332	Morin	-22.4903	-43.1856	Morin
INEA	2243331	Quitandinha	-22.5197	-43.213	Quitandinha
INEA	2243328	Samambaia	-22.4577	-43.1400	Correas
INEA	2243325	Cuiabá	-22.3794	-43.0681	Cuiabá
CEMADEN	330390602A	Alto da Serra	-22.5300	-43.1709	Morin
CEMADEN	330390605A	Araras	-22.4270	-43.2490	Araras
CEMADEN	330390614A	Ciep Brizolão137	-22.4540	-43.1430	Correas
CEMADEN	330390624A	Estrada da Cachoeira	-22.3530	-43.0949	Itaipava
CEMADEN	330390606A	Estrada do Cantagalo	-22.3739	-43.0459	Cuiabá
CEMADEN	330390611A	Itaipava	-22.3880	-43.1319	Itaipava
CEMADEN	330390622A	Itaipava2	-22.3690	-43.1120	Itaipava
CEMADEN	330390612A	Independência2	-22.5481	-43.2089	Quitandinha
CEMADEN	330390620A	Morin	-22.5270	-43.1610	Morin
CEMADEN	330390619A	Nogueira	-22.4179	-43.1220	Parque Petrópolis
CEMADEN	330390618A	Rua Amazonas/Quitandinha	-22.5290	-43.2220	Quitandinha
CEMADEN	330390617A	Rua Araruama/Quitandinha	-22.5200	-43.1930	Quitandinha
CEMADEN	330390603A	Rua Paraná/Quitandinha	-22.5200	-43.2100	Quitandinha
CEMADEN	330390603A	Vale do Cuiabá	-22.4019	-43.0469	Cuiabá
CEMADEN	2243439	Corrêas	-22.4611	-43.0989	Correas
CEMADEN	330390626A	Vale do Cuiabá2	-22.3359	-43.0469	Cuiabá
CEMADEN	330390609A	Vila Constância	-22.4010	-43.0970	Parque Petrópolis
EIBEX	2243287	Esperança	-22.5108	-43.1825	Quitandinha
EIBEX	2243289	Liceu Carlos Chagas	-22.4872	-43.1772	Liceu
EIBEX	2243288	Morin	-22.5167	-43.1689	Morin
EIBEX	2243286	Parque Petrópolis	-22.4053	-43.1333	Parque Petrópolis
EIBEX	2243302	Rocio 2	-22.4775	-43.2581	Rocio
EIBEX	2243291	Sítio das Nascentes	-22.4686	-43.1025	Nascentes
ANA	2243301	Vila açu	-22.4614	-43.095	Vila Açu
ANA	2243268	Binguen	-22.5047	-43.2247	Binguen
ANA	2243010	Itamarati	-22.4853	-43.1492	Itamarati
ANA	2243011	Rio da Cidade	-22.4381	-43.1703	Rio da Cidade
ANA	2243012	Pedro do Rio	-22.3325	-43.1361	Pedro do Rio

To address any missing values in the time series, the HyKit tool, developed at UNESCO-IHE (MASKEY, 2013), was employed. This tool fills the data gaps by weighing the values of synthetic stations according to their distance and elevation. To ensure uniformity across all timeseries, the double mass method (SEARCY & HARDISON, 1960) was utilized. This method involves cumulating precipitation values from each station and comparing them with the cumulated data from a reference station to assess linearity between the series. The complete methodology and dispersion of the stations can be accessed in Costa et al. (2024)

The precipitation values were then interpolated to the MOHID-Land grid to obtain spatialized rainfall. For this purpose, the MOHID system features a tool called FillMatrix, which, by receiving time series of different properties (e.g., precipitation, solar radiation, wind speed, temperature, among others) from monitoring stations, can interpolate the values of each station to the project's grid. This process estimates the values for each pixel. The tool allows interpolation using the TIN (Triangulated Irregular Network) and IDW (Inverse Distance Weighting) methods, generating an HDF5 file as output, where the property varies in both time and space.

The IDW method was chosen for interpolation due to its ability to estimate values with greater influence from nearby stations, which is suitable for the study area. The available precipitation period for the basin spans from January 1, 2006, to December 31, 2022.

The HDF5 file generated with the spatialized rainfall data is available in the dataset at the following path:

- **/GeneralData/BoundaryConditions/02_HDF/Precipitation/02_HDF5/03_58_estacoes_subbacias/Pedro_do_Rio_IDW_2006_to_2022_58_estacoes.hdf5**

3. Usage

3.1. Execution Instructions

To run the simulations, the user must have Windows 10/11 (64-bit) and extract the contents of the provided ZIP file.

a) Navigate to the exe folder.
b) Double Click on the .bat file corresponding to the desired simulation. (the user can access the online log in the directory Log and the correspondent simulation)

It is important to note that each simulation is structured by year, and the user must execute the simulation for the previous year before running the next one.

3.2. Output Organization

Simulation results are stored in the Results folder, organized by year and simulation scenario.

Log files for each simulation are also available in the directory Logs.

All data referenced in the paper correspond to the Node_Pedro do Rio, located in the Run1 folder within each year and simulation directory. The user can open this file useing Microsft Excel software or equal.

3.3. Additional instructions

- Detailed instructions for running the MOHID-Land model can be found at: http://wiki.mohid.com/index.php?title=Mohid_Landhttp://wiki.mohid.com/index.php?title=Mohid_Land
- The MOHID-Land source code is available at: https://github.com/Mohid-Water-Modelling-System/Mohid

4. Links

Repository: https://zenodo.org/records/14914466

5. Citation:

Sales, D. D, Lugon Junior, J., Costa, D. A., Neves, R. & Silva Neto, A. J. (2025). MOHID-Land Preconfigured Hydrological Model for the Pedro do Rio Watershed: Ready-to-Run Simulations and Results (1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14914466

References

ALLEN, R. G.; PEREIRA, L. S; RAES, D. & SMITH, M. Crop evapotranspiration: guidelines for computing crop water requirements. Rome: FAO, 370 p, 1998.

BELONI, T., PIOTTO, V. C., MARI, G. C., PINHEIRO, A. A., TORMENA, C. A., & CECATO, U. Root system and resistance to penetration of Mombaça grass fertilized with nitrogen and irrigated. Semina: Ciências Agrárias, 37(5), 3243-3251, 2016.

BIUDES, M. S., SOUZA, M. C., MACHADO, N. G., MORAIS DANELICHEN, V. H., VOURLITIS, G. L., & NOGUEIRA, J. S. Modelling gross primary production of a tropical semi-deciduous forest in the southern Amazon Basin. International Journal of Remote Sensing, 35(4), 1540-1562, 2014.

CANADELL, J., JACKSON, R. B., EHLERINGER, J. B., MOONEY, H. A., SALA, O. E., & SCHULZE, E. D. Maximum rooting depth of vegetation types at the global scale. Oecologia, 108, 583-595, 1996.

CASTRO, E. B. V. (org) Plano de Manejo do Parque Nacional da Serra dos Órgãos. Brasília, DF: Instituto Chico Mendes de Conservação da Biodiversidade: p. 371, 2008.

CHOW, V. T. Open-channel hydraulics. Caldwell, 1959.

CLEMENTE, F.M.V.T. & BOITEUX, L.S. (Org.). Produção de tomate para processamento industrial. 1ed. Brasília, DF: Embrapa Hortaliças, 2012, v. 1.

COSTA, D.; BAYISSA, Y.; SALES, D.; DIAS, R. M. M. S.; LUGON JUNIOR, J.; SILVA NETO, A. J.; SRINIVASAN, R. Spatial and temporal variability of precipitation in a mountainous watershed using weighted interpolation by distance and elevation. In: XII En-contro de Sustentabilidade em Projeto, 2024, Belo Horizonte. XII Encontro de Sustentabilidade em Projeto, 2024. v. 12. p. 602-610.

GHASEMIZADE, M., BARONI, G., ABBASPOUR, K., & SCHIRMER, M. Combined analysis of time-varying sensitivity and identifiability indices to diagnose the response of a complex environmental model. Environmental modelling & software, 88, 22-34, 2017.

HERSBACH, H., BELL, B., BERRISFORD, P., BIAVATI, G., HORÁNYI, A., MUÑOZ SABATER, J., NICOLAS, J., PEUBEY, C., RADU, R., ROZUM, I., SCHEPERS, D., SIMMONS, A., SOCI, C., DEE, D., THÉPAUT, JN. Dados horários ERA5 em níveis únicos de 1940 até o presente. Serviço Copernicus para Mudanças Climáticas (C3S) Armazenamento de Dados Climáticos (CDS), 2023.

HORTON, P. Analogue methods and ERA5: Benefits and pitfalls. International journal of climatology, 42(7), 4078-4096, 2022.

HYDROSCIENCE. Relatório final do monitoramento de rios da Região Hidrográfica Piabanha (RH-IV) (afluentes). HIDROSCIENCE Consultoria e Restauração Ambiental. Porto Alegre, RS, 2021.

HYDROSCIENCE. Relatório monitoramento de rios da Região Hidrográfica Piabanha (RH-IV) (rio Piabanha). Volumes 1 a 4. HIDROSCIENCE Consultoria e Restauração Ambiental. Porto Alegre, RS, 2020.

IBGE (2022). Cidades@, Petrópolis. Disponível em: https://cidades.ibge.gov.br/brasil/rj/petropolis/panorama. Acesso em 10 de novembro de 2023.

MARTINS, D., TRIGUERO, L. R. C., DOMINGOS, V. D., MARTINS, C. C., MARCHI, S. R. D., & COSTA, N. V. D. Seletividade de herbicidas aplicados em pós-emergência sobre capim-braquiária. Revista Brasileira de Zootecnia, 36, 1969-1974, 2007.

MASKEY, S. HyKit: A Tool for Grid-Based Interpolation of Hydrological Variables. User's Guide (Version 1.3), 1-6, 2013.

MOLAN, L. K. Estrutura do dossel, interceptação luminosa e acúmulo de forragem em pastos de capim-marandu submetidos a alturas de pastejo por meio de lotação contínua. 159p. Dissertação (Mestrado) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2004.

OLIVEIRA, A. R., RAMOS, T. B., PINTO, L., & NEVES, R. Direct integration of reservoirs' operations in a hydrological model for streamflow estimation: coupling a CLSTM model with MOHID-Land. Hydrology and Earth System Sciences, 27(21), 3875-3893, 2023.

OLIVEIRA, A. R., RAMOS, T. B., SIMIONESEI, L., PINTO, L., & NEVES, R. Sensitivity analysis of the MOHID-Land hydrological model: A case study of the Ulla river basin. Water, 12(11), 3258, 2020.

PERIN, L. Fisiologia da produção de tomateiro enxertado em cultivo sem solo. Orientadora: Roberta Marins Nogueira Peil. 2021. 122 f. Tese (Doutorado em Agronomia) – Programa de Pós-Graduação em Sistemas de Produção Agrícola Familiar, Faculdade de Agronomia Eliseu Maciel, Universidade Federal de Pelotas, Pelotas, 2021.

PEZZOPANE, J. E. M., AMARAL, G. C., PEZZOPANE, J. R. M., MACHUCA, M. A. H., & NÓIA JÚNIOR, R. S. Spatial and seasonal variability of photosynthetically active radiation in the understory of a semi-decidual seasonal forest. Scientia Forestalis, 50, e3855.

SALES, D. S.; LUGON JUNIOR, J.; COSTA, D. A.; SILVA NETO, A. J. MOHID SOIL TOOL - Computational Tool for Determining Soil Water Percolation Parameters (Version 4.0.3) [Computer software]. 2024. https://github.com/dhiegosales/MOHID-SOIL-TOOL. Accessed mar 05 2023.

SALES, D. S.; LUGON JUNIOR, J.; OLIVEIRA, V. P. S.; Silva NETO, A. J Rainfall input from WRF-ARW atmospheric model coupled with MOHID-LAND hydrological model for flow simulation in the Paraíba do Sul river - Brazil. Journal of Urban and Environmental Engineering (UFPB), v. 15, p. 188-203, 2021.

SANTOS, F. G., SOARES, R. D., CHAVES, M. A., SILVA, M. W. R., FRANCO, Y. L., & PINHO, B. D. Índice climático de crescimento para os capins Brachiaria brizantha cv. Marandu, Cynodon dactylon cv. Tifton 85 e Panicum maximum cv. Tanzânia e relação com a produção de massa seca. Ciência Animal Brasileira, 9(3), 627-637, 2008.

SEARCY, J. K., & HARDISON, C. H. Double-Mass Curves. In Manual of Hydrology: Part 1. General Surface-Water Techniques. USGS – Geological Survey Water-Supply – Paper 1541-B. United States Government Print Office, 1960.

SIMUNEK, J., SEJNA, M., Van GENUCHTEN, M. T., ŠIMŮNEK, J., ŠEJNA, M., JACQUES, D., & SAKAI, M. HYDRUS-1D. Simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media, version, 2, 1998.

SOLIMAN, M., MORSY, M. M., & RADWAN, H. G. Assessment of Implementing Land Use/Land Cover LULC 2020-ESRI Global Maps in 2D Flood Modeling Application. Water, 14(23), 3963, 2022.

SOUZA Jr, C. M., Z. SHIMBO, J., ROSA, M. R., PARENTE, L. L., A. ALENCAR, A., RUDORFF, B. F., & AZEVEDO, T. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. Remote Sensing, 12(17), 2735, 2020.

TAVARES, L. P. S.; LUGON JUNIOR., J.; RODRIGUES, P. P. W.; KALAS, F. A.; WASSERMAN, J. C. Reservoir Implantation for Flood Dampening in the Macaé River Basin Using the MOHID-Land Model. Ciência e Natura, v. 41, p. 1-10, 2019.

VALERIANO, M. M.; ROSSETTI, D. F. Topodata: Brazilian full coverage refinement of SRTM data. Applied Geography (Sevenoaks), v. 32, p. 300-309, 2011.

YILMAZ, M. Accuracy assessment of temperature trends from ERA5 and ERA5-Land. Science of The Total Environment, 856, 159182, 2023.

ZIMMER, A. H.; VALE, C. B.; PIMENTEL, D. M. & SEIFFERT, N. F. Aspectos práticos ligados à formação de pastagens. EMBRAPA Gado de Corte. Circular técnica nº 12, Campo Grande, MS, 1983. Disponível em < https://old.cnpgc.embrapa.br/publicacoes/ct/ct12/ index.html>. Acesso em 12 de dezembro de 2023.

AmbHidro Technological Product Sheet

Technical Product/Technological
Technical Product/Technological
Reference Code: 03/2024

1. PRODUCT TYPE: Non-Patentable Process/Technology and Product/Material
2. PRODUCT TITLE: MOHID-Land Pre-configured Hydrological Model for the Pedro do Rio Watershed: Ready-to-Run Simulations and Results
3. AUTHORS (AMBHIDRO): Dhiego da Silva Sales, Jader Lugon Junior, David de Andrade Costa, Ramiro Neves, and Antônio José da Silva Neto
4. ALIGNMENT WITH AMBHIDRO: (X) Yes ( ) No
5. Impact Level: ( ) High (X) Medium ( ) Low
6. Demand: (X) Spontaneous ( ) By Competition ( ) Contracted
7. Research Objective: Solution to a pre-defined problem
8. Area Impacted by the Product: ( ) Economic ( ) Education ( ) Health ( ) Social (X) Environmental (X) Scientific (X) Learning
9. Type of Impact: ( ) Real (X) Potential
10. Description of Impact Type: The impact is potential, as the public availability of the results generated by the MOHID-Land Pre-configured Hydrological Model for the Pedro do Rio Watershed will provide a foundation for further research. By making all the inputs and ready-to-run simulations publicly accessible, this dataset promotes transparency and enables validation, comparison, and refinement of hydrological models. It contributes to an enhanced understanding of watershed dynamics and supports productivity in agricultural and environmental studies.
11. Replicability: (X) Yes ( ) No
12. Territorial Scope: ( ) International ( ) National (X) Regional
13. Complexity: (X) High ( ) Medium ( ) Low
14. Innovation: (X) High innovation ( ) Medium innovation ( ) Low innovation
15. Benefited Sector: Basin Committees, researchers, and scientists
16. Link with the institution PDI: (X) Yes ( ) No
17. Declare PTT link with PDI: It is related to one of the purposes of IFF, which is to promote production, development, and transfer of social technologies, emphasizing the strengthening of scientific research, innovation, and sustainable management of natural resources. This goal aligns with the following strategic objectives of the R&D plan: (i) OE5 - Develop technologies and innovative solutions according to societal demands, contributing to the development of products, processes, and services that meet community needs; (ii) OE15 - Ensure sustainable and efficient management practices of socioeconomic and environmental resources, aligned with the UN's Sustainable Development Goals (SDGs).
18. Funding: (X) Funding: SWE CNPq scholarship ( ) Cooperation ( ) No funding
19. Intellectual Property Registration: ( ) Yes (X) No
20. Registration Code: Not applicable
21. Technology Stage: ( ) Pilot ( ) In testing (X) Finalized/Implemented
22. Technology Transfer: ( ) Yes (X) No

23. NOTES ON INNOVATION, RELEVANCE, AND IMPACT:

This dataset provides a fully configured version of the MOHID-Land hydrological model, designed for the Pedro do Rio watershed in Petrópolis, Rio de Janeiro, Brazil. By offering ready-to-run simulations with preconfigured setup files, this dataset ensures reproducibility, transparency, and efficiency in hydrological modeling, particularly for river flow predictions. The transparent and standardized configuration allows for easy replication of the simulations, facilitating comparisons and enhancing the credibility of future research. With outputs organized by scenario and year, and detailed logs available, this resource enables efficient analysis and interpretation, reducing setup time and advancing research in watershed management.

Additionally, being provided with a DOI, the dataset can be properly cited, further enhancing its credibility and facilitating its recognition in academic and research contexts.

Link to the product: https://doi.org/10.5281/zenodo.14914466