

Data Description

The “**African Power Pool Analysis**” folder contains the following:

1. Description of Data (PDF)
2. Power Plant Characteristics Spreadsheet (csv)
3. Usable Capacity (folder)
4. Complementarity Assessment (folder)

The following section will describe each of the previous items.

1. [Description of Data Document \(PDF\)](#)

This PDF (current file) contains the description of all the data included in the folder.

2. [Power Plant Characteristics Spreadsheet \(csv\)](#)

This document contains all the information of the Hydropower Plants included in the African Power Pool analysis. Each hydropower plant has a full name (“power_plant_name”) and a name that is used in the csv files with the time series of usable hydropower and shapefiles (“file_name”). Table 1 presents the main fields included in the CSV file.

Table 1 – Power Plant Characteristics Data Dictionary

Field Name	Description	Type	Unique	Example
power_plant_name	Complete name of the power plant	Factor	Yes	Tana Beles
file_name	Name of the power plant used for results files	Factor	Yes	tana_beles
lon	Longitude coordinate of power plant location	Numeric	No	11.68
lat	Latitude coordinate of power plant location	Numeric	No	37.21
power_mw	Installed capacity of the power plant in Megawatts (MW)	Integer	No	460
country	Country in which the power plant is located	Factor	No	Ethiopia
river	River in which the power plant is located	Factor	No	Blue Nile
basin	Basin in which the power plant is located	Factor	No	Nile
primary_region	Primary power pool to which the country belongs to	Factor	No	Eastern Africa Power Pool
secondary_region	Secondary power pool to which the country belongs to	Factor	No	-
design_flow	Design operating streamflow of the power plant in meters cube per second (m ³ /s)	Numeric	No	1489
effective_height	Design height of the power plant for output calculation in meters (m)	Numeric	No	35
max_vol_m3	Maximum volume of adjacent reservoir if present in Million meters cube (MMC)	Numeric	No	9120
max_area_km2	Maximum area of adjacent reservoir if present in square kilometers (km ²)	Numeric	No	3200
dam_height_m	Height of adjacent reservoir if present in meters (m)	Numeric	No	35
reservoir_operations	If power plant is an impoundment power plant the value is set to 1, if it is run-of-river then value set to 0.	Integer	No	1

3. Usable Capacity (folder)

This folder contains 4 sub-folders titled “1. Power Plant Level”, “2. Country Level”, “3. Power Pool Level”, and “4. Dates”. The first three folders contain CSV files with the results of the simulations of usable hydropower capacity (MW). The “1. Power Plant Level” folder contains the time series for both the simulations as run-of-river for all 87 hydropower plants and the reservoir simulations for the 37 reservoir power plants. The “2. Country Level” folder contains the time series for the aggregated country data (considering reservoir and run-of-river power plants). The “3. Power Pool level” contains the aggregated time series for the “primary_region” of each power plant as specified in 2. The “Dates” folder contains 2 CSV files with the corresponding dates for the historical simulation run (“dates_historical_available_capacity.csv”) and the projections/future runs (“dates_projections_available_capacity.csv”). We define usable capacity (MW) as the maximum monthly capacity in MW, constrained by the power plant's installed capacity, the simulated streamflow can maintain for a specific time frame. The time series correspond to the historical run of the model (1950-2005) and the projection runs of the model under two emission scenarios (RCP 4.5 2006-2099, and RCP 8.5 2006-2099). RCP 4.5 represents a mid-emission scenario with atmospheric CO₂ concentrations decreasing after 2040, leading to an increase of 1.8 °C of surface mean temperature by 2100¹. RCP 8.5 represents an extreme scenario with atmospheric CO₂ concentrations increasing throughout the century leading to an annual mean temperature increase of 3.7 °C¹. Each time series file will have the following format: “power_”, followed by the “file_name” from 2 of a power plant, then the scenario (blank, rcp45, or rcp85), the full time frame name (1950-2005 or 2006-2099), and the “.csv” extension (e.g. “power_tana_beles_rcp45_2006_2099.csv”). Tables 2 and 3 present the data dictionary for time series files. These files should be paired with their corresponding dates from the “Dates” folder for analysis purposes. For the analysis and data visualization we typically subset the time series data to 4 time frames: one historical time frame (1970-2005), and three projection time frames (2010-2039, 2040-1069, and 2070-2099).

Table 2 – Time Series Data Dictionary

Field Name	Description	Type	Unique	Example
ACCESS1_0	Monthly simulated usable capacity (MW) using the corresponding global climate model from NASA's NEX-GDDP dataset. There is a total of 672 records for each GCM for the historical simulation (Jan-1950 to Dec-2005), and 1128 records for each emissions scenario (RCP 4.5 Jan-2006 to Dec-2099, and RCP 8.5 Jan-2006 to Dec-2099).	Numeric	No	217.73
bcc_csm1_1				
BNU_ESM				
CanESM2				
CCSM4				
CESM1_BGC				
CNRM_CM5				
CSIRO_Mk3_6_0				
GFDL_CM3				
GFDL_ESM2G				
GFDL_ESM2M				
inmcm4				
IPSL_CM5A_LR				
IPSL_CM5A_MR				
MIROC_ESM				
MIROC_ESM_CHEM				
MIROC5				
MPI_ESM_LR				
MPI_ESM_MR				
MRI_CGCM3				
NorESM1_M				

Table 3 – Dates Data Dictionary

Field Name	Description	Type	Unique	Example
Date	The first day of the month for each month in the simulation. If historical range goes from 1/1/1950 to 12/1/2005. If projection range goes from 1/1/2006 to 12/1/2099.	Date	No	1/1/1950
Month	The number of the month to which the usable capacity (MW) simulation corresponds to.	Integer	No	1
Year	The year to which the usable capacity (MW) simulation corresponds to.	Integer	No	1950

4. Complementarity Assessment (folder)

This folder contains two CSV files titled “comp_index_countries_current_and_planned.csv”, and “highly_comp_power_plants.csv”. This CSV files present the complementarity scores for countries and power plants. Tables 4 and 5 describe the data presented in each of the corresponding files.

Table 4 – Data dictionary for “comp_index_countries_current_and_planned.csv”

Field Name	Description	Type	Unique	Example
country_1	Complete name of country 1.	Factor	Yes	South Africa
short_name_country_1	Name of country 1 used for file names.	Factor	Yes	south_africa
country_2	Complete name of country 2.	Factor	Yes	Lesotho
short_name_country_2	Name of country 2 used for file names.	Factor	Yes	Lesotho
historical_pearson	Temporal correlation (Pearson correlation) between countries for specific period (historical, 2010-2039, 2040-2069, and 2070-2099) and scenario (RCP 4.5 and RCP 8.5). Values vary between -1 and 1.	Numeric	No	0.43
rcp45_2010_2039_pearson				
rcp45_2040_2069_pearson				
rcp45_2070_2099_pearson				
rcp85_2010_2039_pearson				
rcp85_2040_2069_pearson				
rcp85_2070_2099_pearson				
historical_amplitude	Amplitude complementarity between countries for specific period (historical, 2010-2039, 2040-2069, and 2070-2099) and scenario (RCP 4.5 and RCP 8.5). Values vary between 0 and 1.	Numeric	No	0.19
rcp45_2010_2039_amplitude				
rcp45_2040_2069_amplitude				
rcp45_2070_2099_amplitude				
rcp85_2010_2039_amplitude				
rcp85_2040_2069_amplitude				
rcp85_2070_2099_amplitude				
min_distance	Minimum distance for interconnection in kilometers (km).	Numeric	No	0
dist_score	Normalized distance score. Values vary between 0 and 1.	Numeric	No	1
historical_CI	Complementarity index between countries combining all three previous indexes (temporal, amplitude, and distance) for specific period (historical, 2010-2039, 2040-2069, and 2070-2099) and scenario (RCP 4.5 and RCP 8.5). Values vary between -1 and 1.	Numeric	No	0.08
rcp45_2010_2039_CI				
rcp45_2040_2069_CI				
rcp45_2070_2099_CI				
rcp85_2010_2039_CI				
rcp85_2040_2069_CI				
rcp85_2070_2099_CI				

Table 5 – Data dictionary for “highly_comp_power_plants.csv”

Field Name	Description	Type	Unique	Example
power_plant_1	File name for power plant 1.	Factor	Yes	roseires
power_plant_2	File name for power plant 2.	Factor	Yes	kidatu
country_1	Complete name for country where power plant 1 is located in.	Factor	No	Sudan
short_name_country_1	Short name for country where power plant 1 is located in.	Factor	No	sudan
country_2	Complete name for country where power plant 2 is located in.	Factor	No	Tanzania
short_name_country_2	Short name for country where power plant 2 is located in.	Factor	No	tanzania
connection	If interconnection between countries exist then value is equal to 1, if not value is 0.	Numeric	No	0
planned_connection	If there is a planned interconnection between countries then value is equal to 1, if not value is 0.	Numeric	No	1
min_distance	Minimum distance for interconnection in kilometers (km).	Numeric	No	0
dist_score	Normalized distance score. Values vary between 0 and 1.	Numeric	No	1
historical_CI	Complementarity index between countries combining all three indexes (temporal, amplitude, and distance) for specific period (historical, and 2070-2099) and scenario (RCP 4.5 and RCP 8.5). Values vary between 0 and -1.	Numeric	No	-0.86
rcp45_2070_2099				
rcp85_2070_2099				

References

1. Hartmann, D. L., Tank, a. M. G. K. & Rusticucci, M. *IPCC Fifth Assessment Report, Climatic Change 2013: The Physical Science Basis. IPCC AR5*, (2013).
2. Climate Analytics Group & NASA Ames Research Center. NEX-GDDP Dataset. (2018).