PRIOGRID Ingestion: CRU Climate Data

Garrett Benz

Processing Steps

FUNCTION: *load_cru_temperature*.

1. **Summary:** The *load_cru_temperature()* function reads a specified NetCDF filename from a list in a text file and manages the full download and decompression process if needed. It loads monthly CRU temperature data into a raster stack, automatically extracting only the temperature layers and assigning date-aware names to each layer based on the start year encoded in the filename.

2. **Utility:** This function is foundational for reproducible climate data workflows. It handles remote file access and prepares the data in a standardized format for time-series analysis. Its automated date annotation and layer filtering make it plug-and-play for further operations such as extraction, summarization, or visualization over spatial units like PRIO-GRID.

FUNCTION: *validate_date_range*.

1. **Summary:** The *validate_date_range()* function checks whether a user-specified date range (in year and month) falls within the range of available raster layer dates. It converts all dates to year-month format and returns the validated start and end dates.

2. **Utility:** This function ensures temporal alignment between user-defined analysis windows and the available raster dataset. By catching errors early, it prevents misaligned extractions and improves the robustness of climate data processing pipelines.

FUNCTION: *build_ym_table*.

1. **Summary:** The *build_ym_table()* function generates a data frame of all year-month combinations between a defined start and end date. Each row includes a corresponding layer name matching the naming convention used in the loaded temperature raster.

2. **Utility:** This function standardizes the temporal structure required for looping through raster layers. By mapping month-year combinations to raster layer names, it serves as a reliable intermediary for time-series extraction and visualization across climate datasets.

FUNCTION: *extract_cru_to_pg*.

1. **Summary:** The extract_cru_to_pg() function iterates over monthly raster layers to extract zonal mean temperature values for each PRIO-GRID cell. It returns a tidy data frame indexed by grid ID, year, and month, with corresponding climate values.

2. **Utility:** This function provides a spatially resolved, time-stamped summary of climate data at the PRIO-GRID level. By translating raster information into a tidy format, it enables direct integration with statistical and modeling workflows that rely on tabular data inputs.

FUNCTION: *process_cru_to_pg*.

1. **Summary:** The *process_cru_to_pg()* function provides a complete pipeline for loading, validating, and extracting CRU temperature data for a specified date range using PRIO-GRID polygons. It encapsulates file handling, temporal validation, index mapping, and spatial summarization in a single call.

2. **Utility:** This high-level wrapper streamlines climate data preprocessing for policy or conflict modeling applications. It reduces manual error by coordinating multiple processing steps into a single function, making it well-suited for batch operations, automated pipelines, or reproducible research environments.