README file for the "CHASE-PL – Future Hydrology data set (CPL-FH)"

Content: projections of daily natural streamflow and monthly water balance components simulated with SWAT model for the entire Vistula and Odra basins (VOB), for the reference time period 1974-2000, and two future horizons, 2024-2050 and 2074-2100, under RCP 4.5 and RCP 8.5.

Publisher: Warsaw University of Life Sciences (WULS-SGGW), Poland.

Data citation: Piniewski, M., Szcześniak, M., Kardel, I. (2017) CHASE-PL – Future Hydrology data set (CPL-FH). Warsaw University of Life Sciences WULS-SGGW. Dataset. DOI: 10.4121/uuid:301b610c-b1f7-42e0-a317-7be83ca5941c

Contacts: Mikołaj Piniewski **(**m.piniewski@levis.sggw.pl)

Data description

The data set consists of three parts: (1) model inputs; (2) raw model outputs; (3) aggregated model outputs. the model inputs allow the experienced users to reproduce the outputs, or to create the new ones by adjusting settings or parameters; the aggregated model outputs are useful for various users not interested in daily projections based on single climate models, but in the impact indicators, such as the relative changes in mean seasonal runoff measured by different statistics of the climate model ensemble. These indicators are stored in GIS format which enables immediate viewing of spatial variability of projected impacts in different variables.

1. Model inputs

The zipped directory of SWAT input files *Txtinout.zip* contains the full collection of all files necessary to run the SWAT project of the VOB. It contains the climate input files: precipitation files (*.pcp), temperature files (*.tmp), weather generator files (*.wgn), sub-basin files (*.sub) and snow files (*.sno), all related to the calibrated and validated model that runs for the period 1951-2013. There is another zipped file (*ClimateScenarios.zip*) that contains modified climate inputs that were constructed based on the RCM outputs with the help of the ArcSWAT GIS interface. In total, 45 groups of files originating from nine climate models (Table 1), three time horizons (reference period 1971-2000, near future 2021-2050, and far future 2071-2100) and two RCPs (4.5 and 8.5) are included. The directory naming follows a convention: *R_cmZ_YYYY-YYYY*, where:

R can be "ref", meaning reference data, "rcp45" or "rcp85" meaning different Representative Concentration Pathways, RCP 4.5 and RCP 8.5, respectively;

cmZ – defined GCM-run-RCM combination, where Z refers to the codes from Table 1.

YYYY-YYYY defines the beginning and ending year of the simulation period, i.e. either 1971-2000, or 2021-2050, or 2071-2100.

Execution of the model using the input data from the *Txtinout* directory should start with setting all necessary parameters in the *file.cio* configuration file. The user can adjust the simulation start and end dates, output print time step (day, month, year), the range of variables printed, etc. Four SWAT executable files are available: for 32-bit and 64-bit operation systems, and a release and debug versions.

Table 1 List of available GCM-run-RCM combinations composing the multi-model ensemble.

Code	GCM	RCM
01	CNRM-CERFACS-CNRM-CM5	CLMcom-CCLM4-8-17
02	CNRM-CERFACS-CNRM-CM5	SMHI-RCA4
03	ICHEC-EC-EARTH	CLMcom-CCLM4-8-17
04	ICHEC-EC-EARTH	SMHI-RCA4
05	ICHEC-EC-EARTH	KNMI-RACMO22E
06	ICHEC-EC-EARTH	DMI-HIRHAM5
07	IPSL-IPSL-CM5A-MR	SMHI-RCA4
08	MPI-M-MPI-ESM-LR	CLMcom-CCLM4-8-17
09	MPI-M-MPI-ESM-LR	SMHI-RCA4

For running the model with one of 45 climate scenario data from the *ClimateScenarios.zip* file, the user needs to replace all the files in the *Txtinout* directory with the respective files from the directory storing input data pertaining to a given scenario and rerun the executable file.

For editing large output files the use of a free text editor for Windows EditPad Lite is recommended.

2. Raw model outputs

The data set contains the simulation output from SWAT for 46 different model runs

- The calibrated and validated SWAT model run for the historical period 1954-2013 (cf. Piniewski *et al*., 2017);
- Model runs forced with nine different bias-corrected RCM data for the reference period 1974-2000 (Ref);
- Model runs forced with nine different RCM data under RCP 4.5 for the near future (NF), i.e. 2024-2050;
- Model runs forced with nine different RCM data under RCP 8.5 for NF;
- Model runs forced with nine different RCM data under RCP 4.5 for the far future (FF), i.e. 2074-2100;
- Model runs forced with nine different RCM data under RCP 8.5 for FF.

Two types of output variables can be distinguished: sub-basin-level variables (water balance) and reach-level variables (streamflow, cf. Table 2). The former were aggregated to the monthly scale,

due to a large number of variables and space limitations, and the latter were stored in the original daily scale.

Model output from the calibrated and validated SWAT run constitutes an upgrade of the CPL-NH data set in the CHASE-PL collection of data sets on 4TU (Piniewski *et al*., 2015). The updated data set comes from the same forcing (Berezowski *et al*., 2015) but uses slightly different parametrization than the one before.

Table 2 The list of variables in output files.

Variable code	Spatial unit	Description	Units
PCP	Sub-basin	Total amount of precipitation falling on the sub-basin during time step	mm
SNOM	Sub-basin	Amount of snow or ice melting during time step	mm
PET	Sub-basin	Potential evapotranspiration from the sub-basin during the time step	mm
ET	Sub-basin	Actual evapotranspiration from the sub-basin during the time step	mm
SW	Sub-basin	Soil water content . Amount of water in the soil profile at the end of the time period.	mm
PERC	Sub-basin	Water that percolates past the root zone during the time step (mm).	mm
SURQ	Sub-basin	Surface runoff contribution to streamflow during time step	mm
GWQ	Sub-basin	Groundwater contribution to streamflow (mm). Water from the shallow aquifer that returns to the reach during the time step.	mm
WYLD	Sub-basin	Water yield (mm H_2O). The net amount of water that leaves the sub-basin and contributes to streamflow in the reach during the time step.	mm
FLOW	Reach	Natural discharge of water in the reach	m^3/s

Raw model output data are stored as comma separated values (*.csv* format), 46 files for sub-basin variables and 46 files for the reach variable, i.e. streamflow. All files follow the naming convention *X_subbasin.csv* or *X_reach.csv*, where *X* can attain one of 46 values:

"Calibration_1954-2013", in the case of the calibrated and validated model for the period 1954-2013, or *"R_cmZ_YYYY-YYYY"*, where:

- *R* can be "ref", meaning reference data, "rcp45" or "rcp85" meaning different Representative Concentration Pathways, RCP 4.5 and RCP 8.5, respectively;
- *cmZ* – defined GCM-run-RCM combination, where *Z* refers to the codes from Table 1;
- *YYYY-YYYY* defines the beginning and ending year of the simulation period, i.e. either 1974-2000, or 2024-2050, or 2074-2100.

All files have similar, simple structure: one column with sub-basin/reach ID, one column with (daily or monthly) date, and the subsequent columns with variables from Table 2 and respective values.

3. Aggregated model outputs

The aggregated model outputs include the multi-model ensemble (cf. Table 1) 5-th percentiles, medians, and the 95-th percentiles of the relative changes in the multi-annual or multi-seasonal mean values of analysed variables (cf. Table 2) between respective future horizons and the reference period. The statistics are calculated for all sub-basins and reaches, for two future horizons (always with respect to the reference period) under both RCPs. The data are stored in eight shapefiles according to the following convention: *subbasin_aggreg_R_YYYY-YYYY_wrt_1974-2000.shp* for the water balance data pertaining to SWAT sub-basins, and *reach_aggreg_R_YYYY-YYYY_wrt_1974-2000.shp* for the flow data pertaining to SWAT reaches, where:

- *R* can be “rcp45” or “rcp85” meaning different Representative Concentration Pathways, RCP 4.5 and RCP 8.5, respectively;
- *YYYY-YYYY* defines the beginning and ending year of the future projection horizon, i.e. 2024-2050, or 2074-2100.

Each shapefile stores geographic vector data of 2,633 SWAT sub-basins (polygons) or reaches (polylines) as well as a list of attributes, among which are the standard attributes computed by ArcSWAT in the SWAT project creation phase and the ensemble statistics related to particular variables from Table 2. The names of these fields are *Var_X_time*, where *Var* denotes the variable codes from Table 2, *X* denotes one of three possible ensemble statistics: "L" for low change, i.e. 5-th percentile; "M" for median change, i.e. 50-th percentile; and "H" for high change, i.e. 95-th percentile, and *time* denotes a temporal aggregation: "Ann" for annual, and "DJF", "MAM", "JJA" and "SON" for corresponding seasons (winter, spring, summer, autumn, respectively).

The aggregated model outputs can be also visually analysed in the interactive geoportal of the CHASE-PL project: http://climateimpact.sggw.pl (section Maps - Impact).

Acknowledgements: Support of the project CHASE-PL (Climate change impact assessment for selected sectors in Poland) of the Polish–Norwegian Research Programme operated by the National Centre for Research and Development (NCBiR) under the Norwegian Financial Mechanism 2009-2014 in the frame of Project Contract No. Pol Nor/200799/90/2014 is gratefully acknowledged.

References

Berezowski, T., Piniewski, M., Szcześniak, M., Kardel, I., Michałowski R. (2015) CHASE-PL Forcing Data: Gridded Daily Precipitation & Temperature Dataset 5 km (CPLFD-GDPT5).

Warsaw University of Life Sciences WULS-SGGW. Dataset. http://dx.doi.org/10.4121/uuid:e939aec0-bdd1-440f-bd1e-c49ff10d0a07

Piniewski, M., Szcześniak, M., Kardel, I., Berezowski, T., Okruszko, T., Srinivasan, R., Schuler, D.V., Kundzewicz, Z. W. (2017) Hydrological modelling of the Vistula and 1 Odra river basins using SWAT. *Hydrol Sci. J.* (In press)

Piniewski, M., Szcześniak, M., Kardel, I., Berezowski, T. (2015) CHASE-PL – Natural Hydrology dataset (CPL-NH). Warsaw University of Life Sciences WULS-SGGW. Dataset. http://dx.doi.org/10.4121/uuid:b8ab4f5f-f692-4c93-a910-2947aea28f42