

# User's Guide for ERA5 Data Processing

Amir A. Aliabadi and Mohsen Moradi

Atmospheric Innovations Research (AIR) Laboratory

School of Engineering, University of Guelph, Guelph, Canada

<http://www.aaa-scientists.com>

[aliabadi@uoguelph.ca](mailto:aliabadi@uoguelph.ca)

This document is typeset using L<sup>A</sup>T<sub>E</sub>X

January 16, 2023

## 1 Weather Files and ERA5 Reanalysis Data Products

Sustainable land use development requires an understanding of future climate and weather conditions. Urban Physics Models (UPMs) are often used to investigate urban development alternatives under different future climate scenarios [Pörtner et al., 2022]. The Vertical City Weather Generator (VCWG) is an example computationally-fast UPM that predicts temporal and vertical variation of meteorological variables in the urban environment, building envelope temperatures, and temporal variation of building performance metrics, such as indoor air temperature, indoor specific humidity, building thermal and electricity loads, and natural gas and electricity consumptions [Moradi et al., 2021, Aliabadi et al., 2021, Moradi et al., 2022]. UPMs often require weather files (for instance in EnergyPlus Weather (EPW) or International Weather for Energy Calculations (IWECC) formats) for each region with at least hourly time resolution. For instance VCWG takes weather files in the EPW format.

EPW files can be generated using the ERA5 dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF). The ERA5 dataset provides the required variables for the EPW file format at an hourly resolution. The spatial horizontal resolution of the ERA5 dataset is 31 km. Quality-checked monthly updates of ERA5 dataset are available since 1979 until present, and they are published within three months of real time. ERA5 dataset combines historical observations into global estimates using advanced modeling and data assimilation systems [Hersbach et al., 2020]. For generation of weather files, it may be preferred to use reanalysis data products, as opposed to pure observations, since reanalysis data are available at greater spatial and temporal resolutions with fewer data gaps. In a sense, reanalysis data can be regarded as *ground truth* for situations where the quality of weather observations are poor with possibly low spatial or temporal resolutions. This guide provides the step by step instructions to retrieve information from ERA5 data products and format it according to the EPW convention.

## 2 ERA5 Data Download

This guide prepares EPW weather files associated with surface level forcing. Once the specific latitude and longitude for forcing are selected, data must be retrieved from two “Reanalysis” products of ERA5<sup>123</sup>. These are “ERA5” and “ERA5Land”. ERA5Land contains most of the data at high spatial resolution, except for a few radiative terms, which need to be downloaded from ERA5 at lower spatial resolution. These variables are listed in Table 1. The ERA5 data are downloaded in NetCDF format and assembled into EPW files.

Table 1: ERA5 variables used to prepare EPW files.

Data Product	Variable Names
ERA5Land	2m Dew Point Temperature [ $^{\circ}\text{C}$ ]
ERA5Land	2m Temperature [ $^{\circ}\text{C}$ ]
ERA5Land	Soil Temperature [ $^{\circ}\text{C}$ ]
ERA5Land	Wind [ $\text{ms}^{-1}$ ]
ERA5Land	Pressure [Pa]
ERA5Land	Precipitation [mm]
ERA5	Total Sky Direct Solar Radiation at Surface [ $\text{Wm}^{-2}$ ]
ERA5	Surface Solar Radiation Downwards [ $\text{Wm}^{-2}$ ]
ERA5	Surface Thermal Radiation Downwards [ $\text{Wm}^{-2}$ ]

In the first step a raw EPW file should be accessible for formatting purposes. This file is in the main directory “rawEPW.epw”. Then the user should make a directory for a given city, e.g. “Guelph”. Then two sub directories should be further defined as “ERA5” and “ERA5Land”. The user should register in the ERA5 website.

To download ERA5 data, a specific web page can be visited <sup>4</sup>. This is the “ERA5 hourly data on single levels from 1959 to present”. Depending on the number of variables chosen, one may not be able to download all data for a year in a single file. One may have to split the data into multiple files. “Product type” should be “Reanalysis”. Click on “Radiation and Heat”. The user should pick the following variables “Total sky direct solar radiation at surface”, “Surface solar radiation downwards”, and “Surface thermal radiation downwards”. Click these three terms. Then select the “Year” of interest. Click the “Months”, or all of the months. If the data is too large, the website will prevent you from a single download. In that case, you should download fewer months separately. Select all “Days” and all “Times”. Avoid leap years. In “Geographical area”, click “Subregion extraction”. Select a conservative range, plus/minus one degree that contains the region of interest. The footprint of the data should be large enough to contain the site of interest. Format should be “NetCDF”. Then the user submits a form and the website puts the request in a queue. At a later time one may log in the website and view the “request page”. When the data is ready, it must be downloaded within a few hours, or else the link will not be active any more. After download is complete, copy the files in the folder “ERA5”.

<sup>1</sup><https://cds.climate.copernicus.eu/>

<sup>2</sup><https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5>

<sup>3</sup><https://www.shinyweatherdata.com/>

<sup>4</sup><https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview>

To download ERA5Land data, a specific web page can be visited<sup>5</sup>. This is the “ERA5-Land hourly data from 1950 to present”. Click “Download data”. From the “Temperature” box, choose “2m dewpoint temperature”, “2m temperature”, and “Soil temperature” only on levels 1, 2, and 3. You can also choose variables from “Radiation and heat”, although these variables are more reliable from the “ERA5” dataset discussed before. Choose “Surface thermal radiation downwards” and “Surface solar radiation downwards”. Click all variables in the box for “Wind, Pressure and Precipitation”. From the box “Vegetation”, pick both variables “Leaf area index, high vegetation” and “Leaf area index, low vegetation”. Even though EPW datasets do not require leaf area index, this information can be used later to set the Leaf Area Density (LAD) profile (the vertical integral of which is Leaf Area Index (LAI)) in the specific latitude and longitude of the site of interest. Select the required “Year” and “Month”. Select “all” of the “Day” and “Time” options. If download is prevented, split the months. Avoid leap years. From “Geographical area” select the “sub-region extraction” and specify the latitude and longitude of the site plus/minus one degree that contains the region of interest. The footprint of the data should be large enough to contain the site of interest. Format should be “NetCDF”. Then the user submits a form and the website puts the request in a queue. At a later time one may log in the website and view the “request page”. When the data is ready, it must be downloaded within a few hours, or else the link will be not active any more. After download is complete, copy the files in the folder “ERA5Land”.

### 3 Conversion of ERA5 Dataset to EPW Format

Now open python file “EPWGenerator\_ERA.py” in a python editor. It is ideal to run this code using python version 3 or later. You need to download new packages: “xarray” v0.19. The function “write\_epw” performs most of the formatting, and does not need to be changed. Most of the settings are specified after this function.

For example, for Toronto settings, let us create an EPW file for January to April 2015. Set the latitude and longitude as “lat\_rural = 43.67967” and “lon\_rural = -79.86907”. Set the time zone difference from GMT by “GMT = -5”. To begin with use “RawEPW\_file = 'rawEPW.epw' ” and “NewEPW\_file = 'ERA5\_Toronto\_Jan\_Apr2015.epw' ” for the generation of the first EPW file. The content of the months January to April will be updated with the new information. Always the content from the “NewEPW\_file” are appended to the “RawEPW\_file”, so for appending of consecutive periods to an existing EPW file, keep updating the “RawEPW\_file” name to the latest file created in the last step. To specify the source NetCFD files, set the variables: “ERA5land\_file = 'Toronto/ERA5Land/ERA5Land\_Jan\_Feb\_Mar\_Apr\_2015.nc' ” and “ ERA5\_file = 'Toronto/ERA5/ERA5\_Jan\_Feb\_Mar\_Apr\_2015.nc' ”. The variable “epw\_precision = 1” specifies the number of decimal points, except for precipitation, which has a higher precision in mm per hour. Do not change this. Variable “timeInitial = 8” specifies the starting row, in which the data should be overwritten in the EPW file. Make sure the number is consistent with the starting month. For cities to the west of GMT time zone (i.e. west of U.K.), the numbers should not be modified, but for cities to the east of GMT time zone (i.e. east of U.K.), GMT should be added to this variable.

For the first 4-month period. Set these variables

---

<sup>5</sup><https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land?tab=overview>

StartingTime\_ERA5 = '2015-01-01',  
EndingTime\_ERA5 = '2015-04-30',  
RawEPW\_file = 'rawEPW.epw',  
NewEPW\_file = 'ERA5\_Toronto\_Jan\_Apr2015.epw',  
ERA5land\_file = 'Toronto/ERA5Land/ERA5Land\_Jan\_Feb\_Mar\_Apr\_2015.nc',  
ERA5\_file = 'Toronto/ERA5/ERA5\_Jan\_Feb\_Mar\_Apr\_2015.nc', and  
timeInitial = 8. Then run the code.

For the second 4-month period. Set these variables  
StartingTime\_ERA5 = '2015-05-01',  
EndingTime\_ERA5 = '2015-08-31',  
RawEPW\_file = 'ERA5\_Toronto\_Jan\_Apr2015.epw',  
NewEPW\_file = 'ERA5\_Toronto\_Jan\_Aug2015.epw',  
ERA5land\_file = 'Toronto/ERA5Land/ERA5Land\_May\_Jun\_Jul\_Aug\_2015.nc',  
ERA5\_file = 'Toronto/ERA5/ERA5\_May\_Jun\_Jul\_Aug\_2015.nc', and  
timeInitial = 2888. Then run the code.

For the third 4-month period. Set these variables  
StartingTime\_ERA5 = '2015-09-01',  
EndingTime\_ERA5 = '2015-12-31',  
RawEPW\_file = 'ERA5\_Toronto\_Jan\_Aug2015.epw',  
NewEPW\_file = 'ERA5\_Toronto\_Jan\_Dec2015.epw',  
ERA5land\_file = 'Toronto/ERA5Land/ERA5Land\_Sep\_Oct\_Nov\_Dec\_2015.nc',  
ERA5\_file = 'Toronto/ERA5/ERA5\_Sep\_Oct\_Nov\_Dec\_2015.nc', and  
timeInitial = 5840. Then run the code.

Each time “EPWGenerator\_ERA.py” is run with a new start and end date, it will append the EPW file with the new information for the new period. Please make sure you do not make an EPW file for more than a year. For a given month, the soil temperature is averaged at each layer and written in the corresponding field.

## 4 EPW File Format

Note: VCWG reads data in the EPW format given a specific order. The order is strictly maintained by the number of commas and the number of lines. When the EPW file is edited, the user must make sure the number of commas and lines do not change.

After the EPW file is generated for one year, please make sure to edit the header line, e.g. for the case of Guelph as follows. The city name, province, country, latitude, longitude, and time zone difference from GMT can be changed. Leave other entries in the header, since VCWG reads content of this file in a specific order. Note that the time record in the EPW file are Standard Local Time:

“LOCATION,Guelph,ON,CAN,WYEC2-B-94238,718920,43.52360,-80.10370,-5.0,2.0”

If you recall, soil temperatures were extracted at 3 levels. In the EPW file line beginning with

GROUND TEMPERATURES, the first number specifies the number of layers. Currently set to 3. The depth of each layer is then specified with a few commas before the 12 numbers associated with each month. Currently the depths are 0.035, 0.175, and 0.64 m. Depending on the version of VCWG you are running, you need to adjust the depth of the first or all layers, for the model to work. For example, in VCWG v1.3.2 and v1.4.5 change the layer depths to 0.5, 1.0, and 1.5m, respectively.

Please also make sure the following statement is moved to a new line, after the GROUND TEMPERATURES line:

“HOLIDAYS/DAYLIGHT SAVINGS,No,0,0,0”

## 5 Note About Missing Data

Note that depending on the time zone and the snippets of data downloaded from ERA5 repository, there may be days with gaps of data in the EPW file. In this case make sure the pre-existing days with those gaps in the raw EPW file do not exhibit weather data so drastically different from the local climate of interest. For example, you do not want to analyze an Arctic city with the data gap corresponding to a desert. One remedy is to copy and past data from the previous or following day for the same hours.

## References

- [Aliabadi et al., 2021] Aliabadi, A. A., Moradi, M., McLeod, R. M., Calder, D., and Dernovsek, R. (2021). How Much Building Renewable Energy Is Enough? The Vertical City Weather Generator (VCWG v1.4.4). *Atmosphere*, 12(7):882.
- [Hersbach et al., 2020] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.
- [Moradi et al., 2021] Moradi, M., Dyer, B., Nazem, A., Nambiar, M. K., Nahian, M. R., Bueno, B., Mackey, C., Vasanthakumar, S., Nazarian, N., Krayenhoff, E. S., Norford, L. K., and Aliabadi, A. A. (2021). The Vertical City Weather Generator (VCWG v1.3.2). *Geosci. Model Dev.*, 14(2):961–984.
- [Moradi et al., 2022] Moradi, M., Krayenhoff, E. S., and Aliabadi, A. A. (2022). A comprehensive indoor–outdoor urban climate model with hydrology: The Vertical City Weather Generator (VCWG v2.0.0). *Building and Environment*, 207:108406.

[Pörtner et al., 2022] Pörtner, H.-O., Roberts, D. C., Adams, H., and Others (2022). Technical Summary. In Pörtner, H.-O., Roberts, D. C., Poloczanska, E. S., and Others, editors, *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 37–118. Cambridge University Press, Cambridge and New York.