Informatics Lab
Met Office
covid19@informaticslab.co.uk

Maria Athanassiadou, Theo McCaie, Kevin Donkers

COVID-19 Response Pangeo: NWP data set

20 April 2020 (original)
5 May 2020 (updated with wind data)

OVERVIEW

This document describes the dataset of weather parameters, created from the Met Office Numerical Weather Prediction Unified Model (UM). The dataset was put in place in response to requests from epidemiologists and researchers in general to help investigate potential links between weather and COVID-19.

The initial database was constructed for a 3 month period: 1/1/2020 - 31/3/2020, with data from the Global UM only. Subsequently, it has been augmented with more data; more meteorological parameters and data from the high resolution UM over the UK (UKV). The dataset is now updated weekly with the latest Met Office data, both from the Global UM and high resolution UKV.

The document provides the provenance of the dataset. It documents:

- It documents the original data from the UM
- Describes the methodologies used to process the data
- Provides a list of the types of data available to end user

It is a living document that is updated as new data, new parameters and new processed data, become available.

UM data

Global Gridded data

The global gridded data are from operational forecast runs of the Met Office UM, that have been archived in the Met Office MASS archives.

For each day, there are 4 model initialisation times: at 00Z, 06Z, 12Z, 18Z. At these times, the UM is initialised from millions of observations, in a Data Assimilation cycle and then runs forward in

time. As the model runs, files with many meteorological parameters are outputted from these runs at (usually) hourly intervals and are stored in MASS.

The original files from MASS are in pp format. Each pp file contains a lot of parameters on a regular latitude longitude global grid of ~12km resolution. From each one of the initialisation times, pertinent files (with labels 000, 003, 006; see Table 1) are extracted and the original pp files are processed, to create a continuous dataset, as illustrated in Figure 1. The processed files are in nc format. Each file contains hourly values of a single parameter (i.e., temperature), for one day. The values at 00Z, 06Z, 12Z and 18Z, are analysis, the hours in between are from forecasts.

Table 1: Summary of operational global UM files. *yyyy*=year, *mm*=month, *dd*=day

Filename	Model initialisation time	Validity times in file
***filename_yyymmdd*_00_000.pp**	**From the 00Z run**	**Initialisation T0 = 00Z**
***filename_yyymmdd*_00_003.pp**		**T+1, T+2, T+3 i.e, 1, 2, 3 hours after T0**
***filename_yyymmdd*_00_006.pp**		**T+4, T+5, T+6 i.e., 4, 5, 6 hours after T0**
***filename_yyymmdd*_06_000.pp**	**From the 06Z run**	**Initialisation T0 = 06Z**
***filename_yyymmdd*_06_003.pp**		**T+1, T+2, T+3 i.e, 1, 2, 3 hours after T0**
filename_yyymmdd_06_006.pp		**T+4, T+5, T+6 i.e., 4, 5, 6 hours after T0**
filename_yyymmdd_12_000.pp	**From the 12Z run**	**Initialisation T0 = 12Z**
filename_yyymmdd_12_003.pp		**T+1, T+2, T+3 i.e, 1, 2, 3 hours after T0**
filename_yyymmdd_12_006.pp		**T+4, T+5, T+6 i.e., 4, 5, 6 hours after T0**
filename_yyymmdd_18_000.pp	**From the 18Z run**	**Initialisation T0 = 18Z**
filename_yyymmdd_18_003.pp		**T+1, T+2, T+3 i.e, 1, 2, 3 hours after T0**
filename_yyymmdd_18_006.pp		**T+4, T+5, T+6 i.e., 4, 5, 6 hours after T0**

Each day file (containing hourly gridded data) is 451Mb.

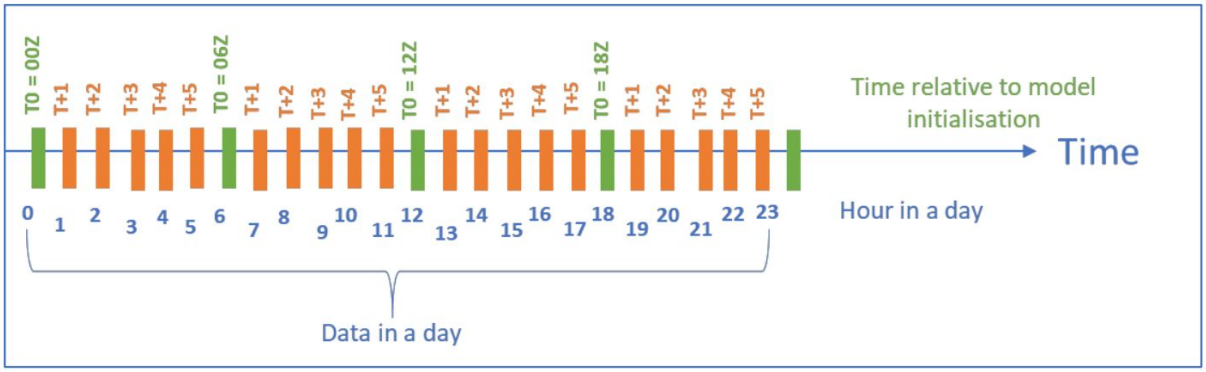

Figure 1: Schematic diagram of data in a day. Same for all parameters.

List of meteorological fields:

- t1o5m = Temperature at 1.5m [K]
- sh = Specific Humidity [Kg/Kg]
- sw = Shortwave radiation (surrogate for sunshine) [W/m2]
- precip = Precipitation (=rain+snow) [kg/m-2 s-1]. Multiply by 3600 to obtain [mm/hr]
- pmsl = Mean sea level pressure [Pa]
- windspeed = Wind Speed [m s-1]
- windgust = Gust of Wind [m s-1]
- oro = Orography [m]

List of hourly parameters:

- hourly_pmsl = Mean sea level pressure, on the hour, every hour.
- hourly_t1o5m = Temperature at 1.5m, on the hour, every hour. For example, the temperature at 11Z on a day, is the temperature at 11Z (instantaneous from model time step).
- hourly_sh = Specific Humidity, same times as for temperature
- hourly_sw = Shortwave radiation, same times as for temperature
- hourly_precip = One hour mean (from the model) with time stamp at half hour. For example, the precipitation at 11Z on a day is the mean precipitation (averaged over all time steps in the model), from 10Z to 11Z, time stamp 10:30Z.
- hourly_windspeed = Wind speed every hour, on the hour
- hourly_windgust = Wind gust every hour, on the hour

Day averaged files are also produced by taking the 24hr mean, max, min of a field. These are still gridded data, on a 12 km x 12 km uniform latitude-longitude grid, same to the hourly data used to generate them.

List of daily parameters:

- daily_t1o5m_mean = Day meanTemperature at 1.5m
- daily_t1o5m_max = Day maxTemperature at 1.5m
- daily_t1o5m_min = Day minTemperature at 1.5m
- daily_sh_mean = Day mean Specific Humidity
- daily_sh_max = Day max Specific Humidity
- daily_sh_min = Day min Specific Humidity
- daily_sw_mean = Day mean for short wave radiation (surrogate for sunshine)
- daily_sw_max = Day max for short wave radiation (surrogate for sunshine)
- daily_precip_mean = Day mean of precipitation
- daily_precip_max = Day max, of precipitation
- daily_windspeed_max = Day maximum of wind speed
- daily_windspeed_mean = Day mean of wind speed
- daily_windspeed_min = Day minimum of wind speed
- daily_windgust_min = Day minimum of wind gust
- daily_windgust_max = Day maximum of wind gust
- daily_windgust_mean = Day mean of wind gust

Each daily file is 19Mb.

A note on time metadata

In the original pp data, there are 2 time parameters: forecast_period & forecast_reference_time. The forecast_reference_time is the time of model initialisation (00Z, 06Z, 12Z,18Z). The forecast_period is the time from the “parent” forecast_reference_time (initialisation). As a result, in the metadata in the hourly files, time (forecast_period, forecast_reference_time); a 2D array, rather than 0-23hr. When the time averages are taken, to create the day averaged files, the times

metadata can be confusing; but the data are daily mean, daily max and daily min of the hourly values.

UKV Gridded data

The UKV is the Met Office operational limited area model, covering the UK at a high resolution grid of 1 Km x 1 km. It takes initial and boundary conditions every hour, from the Global UM and integrates forward in time to produce the weather forecast over the UK. It is on a rotated lat-lon coordinate system, to ensure a relatively uniform grid over the UK. Convection is explicit (not parameterized). More information can be found in [ref].

Although run (initialised) every hour, files are stored in MASS from only certain initialisations; 03Z, 09Z, 15Z, 21Z. The overall extraction and processing data flow is very similar to the Global data, but with a few caveats.

The UKV files in MASS are in pp format, and contain a lot of parameters on a rotated latitude longitude grid of ~1 km resolution. From each one of the initialisation times, pertinent files (with labels 000, 002, 004; see Table 2) are extracted and the original pp files are processed, to create a continuous dataset, as illustrated in Figure 2. The processed files are in nc format. Each file contains hourly values of a single parameter (i.e., temperature), for one day. The values at 03Z, 09Z, 15Z and 21Z, are analysis, the hours in between are from forecasts.

Table 2: Summary of operational UKV files. *yyyy*=year, *mm*=month, *dd*=day

Filename	Model initialisation time	Validity times in file
***filename_yyymmdd*_03_000.pp**	**From the 03Z run**	**Initialisation T0 = 03Z, T+1, T+2**
***filename_yyymmdd*_03_002.pp**		**T+3, T+4 i.e, 3 & 4 hours after T0**
***filename_yyymmdd*_03_004.pp**		**T+5, T+6 i.e., 5 & 6 hours after T0**
***filename_yyymmdd*_09_000.pp**	**From the 09Z run**	**Initialisation T0 = 09Z, T+1, T+2**
***filename_yyymmdd*_09_002.pp**		**T+3, T+4 i.e, 3 & 4 hours after T0**
filename_yyymmdd_09_004.pp		**T+5, T+6 i.e., 5 & 6 hours after T0**
filename_yyymmdd_15_000.pp	**From the 15Z run**	**Initialisation T0 = 15Z, T+1, T+2**
filename_yyymmdd_15_002.pp		**T+3, T+4 i.e, 3 & 4 hours after T0**
filename_yyymmdd_15_004.pp		**T+5, T+6 i.e., 5 & 6 hours after T0**
filename_yyymmdd_21_000.pp	**From the 21Z run**	**Initialisation T0 = 21Z, T+1, T+2**

filename_yyymmdd_21_002.pp	**T+3, T+4 i.e, 3 & 4 hours after T0**
filename_yyymmdd_21_0.pp	**T+5, T+6 i.e., 5 & 6 hours after T0**

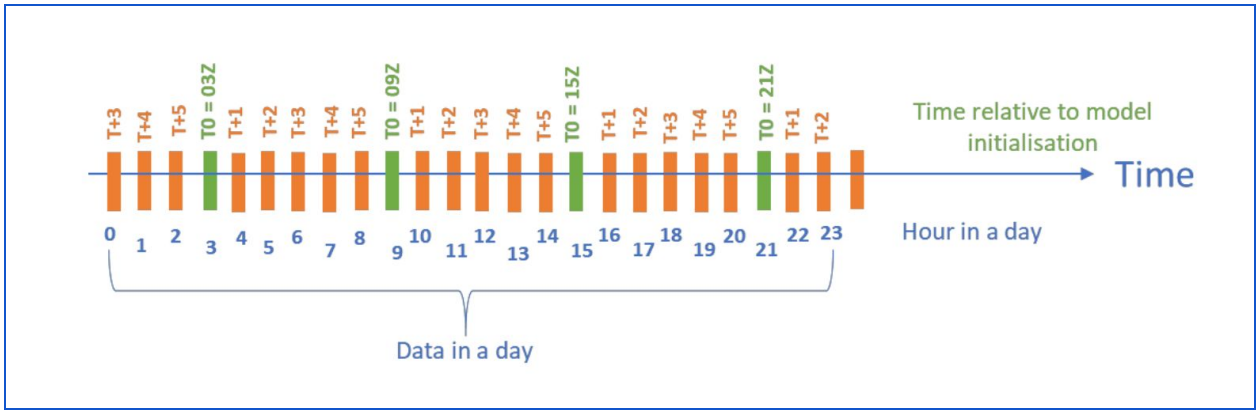

Figure 2: Schematic diagram of data in a day for the UKV model. Same for all parameters.

List of meteorological fields:

- t1o5m = Temperature at 1.5m [K]
- sh = Specific Humidity [Kg/Kg]
- sw = Shortwave radiation (surrogate for sunshine) [W/m2]
- rain = Rain [kg/m-2 s-1]. Multiply by 3600 to obtain [mm/hr]
- snow = Snow [kg/m-2 s-1].
- windspeed = Wind Speed [m s-1]
- windgust = Gust of Wind [m s-1]
- oro = Orography [m]

List of hourly parameters:

- hourly_t1o5m = Temperature at 1.5m, on the hour, every hour. For example, the temperature at 11Z on a day, is the temperature at 11Z (instantaneous from model time step).
- hourly_sh = Specific Humidity, same times as for temperature
- hourly_sw = Shortwave radiation, same times as for temperature

- hourly_rain = One hour mean (from the model) with a timestamp at half hour. For example, the precipitation at 11Z on a day is the mean precipitation (averaged over all time steps in the model), from 10Z to 11Z, time stamp 10:30Z.
- hourly_snow = One hour mean (from the model) with timestamp at half hour.
- hourly_windspeed = Wind speed every hour, on the hour.
- hourly_windgust = Wind gust every hour on the hour.

Each file, containing hourly fields for a single day, is 46Mb.

Day averaged files are also produced by taking the 24hr mean, max, min of a field. These are still gridded data, on a 1 km x 1 km rotated latitude-longitude grid, same to the hourly data used to generate them.

List of daily parameters:

- daily_t1o5m_mean = Day meanTemperature at 1.5m
- daily_t1o5m_max = Day maxTemperature at 1.5m
- daily_t1o5m_min = Day minTemperature at 1.5m
- daily_sh_mean = Day mean Specific Humidity
- daily_sh_max = Day max Specific Humidity
- daily_sh_min = Day min Specific Humidity
- daily_sw_mean = Day mean for short wave radiation (surrogate for sunshine)
- daily_sw_max = Day max for short wave radiation (surrogate for sunshine)
- daily_rain_mean = Day mean of precipitation
- daily_rain_max = Day max, of precipitation
- daily_snow_mean = Day mean of precipitation
- daily_snow_max = Day max, of precipitation
- daily_windspeed_max = Day maximum of wind speed
- daily_windspeed_mean = Day mean of wind speed
- daily_windspeed_min = Day minimum of wind speed
- daily_windgust_min = Day minimum of wind gust
- daily_windgust_max = Day maximum of wind gust
- daily_windgust_mean = Day mean of wind gust

Each daily file is 2Mb.

Shaped files

These are processed files (csv format) from the gridded data, in various country, county or administrative areas.

A lot of this type of files (csv format) both for the UK and Globally are currently available and more are generated depending on demand. Description of these will be available soon.

OTHER DATA

Air Quality data

It is envisaged that air quality data will be available on the platform soon.

REFERENCES

Global UM

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geosci. Model Dev., 12, 1909–1963,https://doi.org/10.5194/gmd-12-1909-2019, 2019.
https://www.geosci-model-dev.net/12/1909/2019/gmd-12-1909-2019.html

UKV

Tang, Y., Lean, H., and Bornemann, J.: The benefits of the Met Office variable resolution NWP model for forecasting convection, Meteorol. Appl., 20, 417–426, https://doi.org/10.1002/met.1300, 2013.

Bush, M., Allen, T., Bain, C., Boutle, I., Edwards, J., Finnenkoetter, A., Franklin, C., Hanley, K., Lean, H., Lock, A., Manners, J., Mittermaier, M., Morcrette, C., North, R., Petch, J., Short, C., Vosper, S., Walters, D., Webster, S., Weeks, M., Wilkinson, J., Wood, N., and Zerroukat, M.: The first Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL1, Geosci. Model Dev., 13, 1999–2029, https://doi.org/10.5194/gmd-13-1999-2020, 2020.
https://www.geosci-model-dev.net/13/1999/2020/