## Cyclone Tracking Description - Alex Crawford

Original Algorithm: Crawford, A. D., and M. C. Serreze, 2016: Does the summer Arctic Frontal Zone influence Arctic Ocean cyclone activity? *Journal of Climate*, **29**, 4977–4993, doi:10.1175/JCLI-D-15-0755.s1.

Latest Update: Crawford, A.D., K.E. Alley, A.M. Cooke, and M.C. Serreze, 2020: [Synoptic Climatology of Rain-on-Snow Events in Alaska.](https://journals.ametsoc.org/doi/abs/10.1175/MWR-D-19-0311.1) *Mon. Wea. Rev.,* **148**, 1275–1295,<https://doi.org/10.1175/MWR-D-19-0311.1>

## 1. Input Parameters

The algorithm described here is designed to work with any of several atmospheric reanalyses with spatial resolution finer than 2°×2° latitude/longitude and a temporal resolution of 6-hourly or finer. In order to achieve the needed flexibility, it incorporates eleven different parameters that can be used to tune the algorithm for different data sources or different research questions (**Table 1**).

|  |  |  |
| --- | --- | --- |
| **Table 1**. Cyclone detection and tracking parameters that can be used to tune the algorithm and the value used for each in this study. The first set of variables define cyclone center detection, the second set defines cyclone area and multi-center cyclone (MCC) detection, and the third defines precipitation detection and cyclone tracking. | | |
| **Parameter** | **Value Used** | **Explanation** |
| kSize | 1 grid cell | Half-length of each side of the kernel used to identify whether a grid cell is a local minimum |
| nanThresh | 0.25 | Maximum fraction of neighboring grid cells with no data allowed for a grid cell to be considered a candidate minimum |
| d\_slp / d\_dist | 7.5 hPa  (1000 km)-1 | Minimum average pressure difference at a certain distance measured from a SLP minimum required for that SLP minimum to be considered a center |
| max\_elev | 1500 m | All elevations above this threshold are masked before analysis; no centers can be detected at elevations above this threshold |
| contint | 2 hPa | Contour interval used when searching for the last closed contour around a cyclone center; used to define cyclone areas and MCCs |
| mcctol | 0.5 | For multiple centers to be grouped as a MCC, this is the maximum allowed ratio of unshared area (defined by closed contours) around the lowest pressure center to shared area for all centers |
| mccdist | 1000 km | Maximum distance allowed between the primary center of a MCC and any secondary center |
| pMin | 1.5 mm day-1 | Minimum precipitation rate (scaled to time step) used to determine contiguous precipitation areas |
| rPrecip | 250 km | Minimum radius for cyclone area (additive with algorithm’s cyclone area detection) |
| maxSpeed | 150 km hr-1 | Defines the search radius for extending cyclone tracks |
| red | 0.75 | Modifies the projection of a cyclone center’s propagation between two time steps, accounting for the tendency for cyclone propagation to slow with age |

**Additional Notes:**

a. To run the algorithm, your SLP (and precipitation, if using) inputs must be in an equal-area projection. Any format that can be read as a numpy array will do.

b. You also need an elevation, latitude, and longitude field of the same projection, grid cell size, and extent as the SLP inputs. Any format that can be read as a numpy array will do. Starting in version 11, I’m using a more sophisticated distance measurement, so it also requires two additional projection-related files: one for the x distance and one for the y distance in each grid cell. This is to make the algorithm more accurate at lower latitudes, but in the Arctic, the equal-area grid cells are close enough to square that it does not really matter.

c. The main script for running the algorithm is *C3\_CycloneDetection\_11\_1.py*. In my workflow, this is often the third step (after downloading and reprojection). That’s where the prefix “C3” comes from. The suffix: “11\_1” means that the code is version 11.1. This is *not* the same version used in Crawford and Serreze (2016). That version was 10\_3. The main updates since the 10\_3 version are:

1. As stated above, more sophisticated distance measurements
2. A more flexible multi-center cyclone scheme that a) allows for a greater variety of situations and b) has a *slightly* more accurate area calculation
3. More refinement to the identification of splitting and merging events that decreases that number of false positives
4. A fix to a bug in the “regenesis” code that mixed up cyclone IDs for a small percentage of cases
5. The addition of a track matching function for assessing the impact of using different parameters or inputs.
6. Everything is now written for Python 3 and pandas 0.24.

d. All scripts related to this algorithm use a custom Python module and custom Python objects. For ease of use, always store the module script in the same directory as the other scripts you are using or in the appropriate site-packages folder. For version 11\_1, the module script is *CycloneModule\_11\_1.py*.

e. The Python modules I used for version 10.3 are as follows:

Python 2.7.11 (including *os* and *copy*)

GDAL 2.0.0 (including *osgeo*, *gdal*, *gdalconst*, and *gdalnumeric*)

NumPy 1.10.4 (1.8 also compatible)

SciPy 0.17 (0.14 also compatible)

Pandas 0.14 (0.12 also compatible, but code runs 25-50 times slower with Pandas 0.15)\*

cPickle 1.71

Crawford, A. D., and M. C. Serreze, 2016: Does the summer Arctic Frontal Zone influence Arctic Ocean cyclone activity? *Journal of Climate*, **29**, 4977–4993, doi:10.1175/JCLI-D-15-0755.s1.

f. The Python modules I used for version 11.1 are as follows:

Python 3.7.1

GDAL 2.4.1

NumPy 1.16.4

SciPy 1.2.1

pandas 0.24.2

basemap 1.2.1 (reprojection and plotting only)

matplotlib 3.0.3. (reprojection and plotting only)

Crawford, A.D., K.E. Alley, A.M. Cooke, and M.C. Serreze, 2020: [Synoptic Climatology of Rain-on-Snow Events in Alaska.](https://journals.ametsoc.org/doi/abs/10.1175/MWR-D-19-0311.1) *Mon. Wea. Rev.,* **148**, 1275–1295,<https://doi.org/10.1175/MWR-D-19-0311.1>

## 2. Outputs

The output from this algorithm is two-fold. First, the synoptic information for each SLP field is stored in a customized cyclone field object, which is readable in Python. These objects contain information regarding cyclone location, area, intensity, and associated precipitation. **Table 2** lists the full set of recorded characteristics. Second, a list of cyclone track objects is saved for each month. Cyclones that exist during two months are grouped with the month in which they experience lysis.

|  |  |  |
| --- | --- | --- |
| **Table 2.** Characteristics recorded for each cyclone observation time. | | |
| **Variable** | **Units** | **Description** |
| **Location & Propagation** | | |
| x, y | grid cells | Column and row in EASE2 grid (from upper-left) of cyclone center |
| Dx, Dy | grid cells | Propagation of cyclone across the EASE2 grid since last observation |
| long, lat | -180 to +180°E  -90 to 90°N | Longitude and latitude of cyclone center |
| u, v | km hr-1 | Zonal and meridional propagation velocity since last observation |
| uv | km hr-1 | Propagation speed since last observation |
| **Cyclone Identification** | | |
| id | -- | Unique ID for the cyclone center in the instantaneous cyclone field |
| pid | -- | Unique ID of the lowest pressure cyclone center in a MCC in the instantaneous cyclone field |
| tid | -- | Unique ID of the cyclone center track for the given month |
| ftid | -- | Former track ID of the cyclone center track, only relevant if it existed in the prior month |
| ptid | -- | Track ID of the primary center in a MCC |
| sid | -- | Unique ID for the track of a cyclone system for the given month |
| **Cyclone Size & Intensity** | | |
| p\_cent | Pa | SLP at cyclone center |
| p\_edge | Pa | SLP at cyclone edge (last closed isobar) |
| area | 104 km2 | \* Area enclosed by last closed isobar |
| radius | 102 km | \* Radius of a circle with the same area as cyclone |
| depth | Pa | Edge pressure – central pressure |
| DpDr | Pa / 102 km | \* Depth / radius |
| DpDt | Pa / day | Deepening rate (scaled by latitude) |
| DsqP | Pa / 104 km2 | \* Laplacian of central pressure (∇2p) |
| **Other** | | |
| type | 0, 1, 2 | 1 = primary center, 2 = secondary center (in a MCC), 0 = this row is only present for calculating propagation (used during splits, merges, and lysis events) |
| centers | # | Number of centers in the cyclone system; if a secondary center of a MCC, set to 0 |
| time | days | Days since 1 Jan 1900 0000 UTC |
| Ege, Ely, Emg, Esp, Erg | 0, 1, 2, or 3 | Records whether the cyclone experienced genesis (ge), lysis (ly), merging (mg), splitting (sp), or regenesis (rg); 0 = no event, 1 = center-only, 2 = area-only, 3 = both center and area involved in event |

\*If the grid cell size is 100 km. Adjust as needed if using a different resolution.

**Additional Notes:**

**Type:** Whenever I calculate track-wide statistics (e.g., maximum deepening rate, average intensity, etc.), I do not use any row with type == 0. These rows are used to help identify splitting and merging events and indicate a projection of the cyclone track to before (after) it first (last) appears. Most cyclone data tables should end with one such row with type == 0 and Ely > 0 (a lysis event).

**Depth**: For single-center cyclones (SCC) and primary centers in multi-center cyclones (MCC), the depth will always be a multiple of 200 Pa (based on the contour interval used for area detection). However, for secondary centers of a MCC (centers which are not the lowest pressure in the system), the depth will probably not be a multiple of 200 Pa because the edge is defined based on the pressure of a different center.

**Area:** The minimum area for a cyclone system is 1 grid cell (100 km \* 100 km = 10,000 km2). If a cyclone has the minimum area, then the central pressure and edge pressure are identical, so the depth is 0 Pa.

**Empty Cells**: Empty cells occur in the first and last rows. Nothing wrong; there’s just no data for that variable at that time.

**Time:** Whether leap days are included or not depends on the inputs – so be careful with climate models. MERRA and ERA-Interim both have leap days included.

**Directory Structure:** I tend to use many directories to keep data organized, and I used geotiffs as my output. This is not an efficient file format, but it’s a friendlier format for non-coders than netCDF. If you want to use a similar system that I use for storing output, look at the structure in the test dataset. You of course may prefer to change directory names or styles, in which case all editing of filepaths can be done on the top-level scripts.

**BBox:** I use the “BBox##” subdirectories for storing any subset of the data I might want. (I use the term “BBox” for “bounding box”, even though sometimes “subset” would be more appropriate. There’s no reason why you can’t change the name, but if I share any code with “BBox” in it, you might need to change that code, too.)

**Unique Track Identifiers**: there are several id numbers used in this algorithm, and they can be confusing to look at. For most purposes, if you’re interested in system tracks only, you’re going to be interested in the “sid” (system id). If you want to ignore the multi-center cyclone aspect and just look at cyclone center tracks, the “tid” (track id) is what you need. The “ftid” (former track id) and “ptid” (parent track id in a multi-center cyclone) are only needed if you are looking at cyclone interactions. In all cases, the id number is unique to the month under consideration. For a universally unique identifier, you need to combine the year, month, and id number. Note that cyclones are always stored with the month in which they experience lysis.

## 3. Common Post-Processing Tasks

A. Output from an algorithm like this invariably includes features that are not of interest. To limit the output, it is customary to only look at storms satisfying certain criteria. Common choices are listed below, but note that applying these must be done in post-processing. The main script will not do it.

1) Lifespan >= 24 hr (4 time steps)

2) Track Length >= 100 km (1 grid cell)

3) Minimum Elevation <= 500 m\*

\* This means the cyclone can spend most of its lifetime over high elevations, but there has to be at least one observation of the system over elevations lower than 500 m. This differs from the max\_elev variable used as an input. Whereas the input parameter determines whether a grid cell can be considered a cyclone center, the post-processing variable determines whether an entire track should be kept. However, the two are related. If max\_elev = 500 m, having a minimum elevation <= 500 m (or 600 m or 700 m, etc.) criterion is redundant and pointless.

B. The output from this algorithm has a separate track for each cyclone center, which means some storms will be represented multiple times because they are multi-center cyclones. To limit the dataset to only one track per system, you must run the *C3\_SystemDetection\_11\_0.*py script *after* running the *C3\_CycloneDetection\_11\_1.py* script.

C. I have a suite of scripts for subsetting, aggregating, and analyzing results from the algorithm. The following three are my main workhorses, so those are the ones I share on github:

* C4 = Aggregation of cyclone characteristics by month/season, including climatologies; the output is mostly geotiffs, which can be used for map-making and in GIS programs
* C5 = Subsetting of cyclone output by location and other characteristics
* C10 = Summarize basic cyclone track statistics; each storm stored as row in a CSV file

## 4. Test Dataset

To test one month of data, use the ERA-Interim SLP from August 2016 (in the “Test Data” folder) as your inputs. I’ve already re-projected these to a 100 km by 100 km EASE2 grid (Northern Hemisphere). Compare your results to the I’ve included. The five accessory files you will need are also included. “etopo1” is the DEM; then there’s latitude, longitude, x distance, and y distance.