Scale-Insensitive Pointwise Feature Tracking on Unstructured Grids

Paul A. Ullrich and Colin Zarzycki

Paul A. Ullrich, Department of Land, Air and Water Resources, University of California, Davis, One Shields Ave., Davis, CA 95616. Email: paullrich@ucdavis.edu

Correspondence to: Paul A. Ullrich
(paullrich@ucdavis.edu)

Abstract

This paper describes the new pointwise feature tracking module in the TempestExtremes package.

1 Introduction

5 Tracking of tropical cyclones, extratropical cyclones and AEWs.

- GFDL tracker uses latitude-longitude arcs which can cause issues at high latitudes - cite Bosler paper, indicate algorithm does not currently support unstructured grids

Extratropical cyclones

Manual counts of cyclones were performed by Petterssen (1956) in the Northern hemisphere
10 from 1899-1939, and latter binned by Klein (1957) to determine the spatial distribution of such storms. These techniques were later refined by Whittaker and Horn (1982) by accounting for cyclone trajectories. A similar survey in the Southern hemisphere was performed by Taljaard (1967) for July 1957 - December 1958. Manual tracking and characterization of cyclones was also performed by Akyildiz (1985) using ECMWF forecast data for the 1981/82 winter.

15 One of the first automated detection and tracking for extratropical cyclones was developed by Williamson (1981) using nonlinear optimization to fit cyclonic profiles to anomalies in the 500-mb geopotential height field. Storms were then tracked over a short forecast period using the best fit to the cyclone's centerpoint.

Counts of cyclones neglecting the cyclone trajectory were automatically generated from cli-
20 mate model output for both hemispheres by Lambert (1988) using local minima in 1000-hPa geopotential height. This method had some shortcomings, including mischaracterization of local lows due to Gibbs' ringing and topographically-driven lows. To overcome these problems, Alpert et al. (1990) proposed an additional minimum threshold on the local pressure gradient. Similarly, Le Treut and Kalnay (1990) detected cyclones in ECMWF pressure data using a local
25 minima in the sea-level pressure that must also be 4 mb below the average sea-level pressure of neighboring grid points, and must persist for three successive 6- or 12- hour intervals.

Closed contours were first employed by Bell and Bosart (1989), who used a 30m 500 mb geopotential height contour to identify closed circulation centers. Their approach used radial arms generated at $15°$ intervals over a great circle distance of $2°$ and required that geopotential heights rise by at least 30m along each arm. Unfortunately, one shortfall of this approach is that it does not guarantee the presence of a closed contour at high model resolution.

Murray and Simmonds (1991) extracted low pressure centers from interpolated GCM data using local optimization, based on earlier work in Rice (1982).

Feature tracking on the sphere was revisited by Hodges (1995), which extended tracking algorithms designed for Cartesian geometry Hodges et al. (1994) that were built from image processing techniques.

Storm tracks and African easterly waves were tracked by Hodges et al. (2003).

König et al. (1993) uses minima in the 1000 hPa geopotential height to identify cyclones with associated maxima in the 850 hPa vorticity.

Tropical cyclones

Nonetheless, the definition of an optimal objective criteria for atmospheric features has eluded development, suggesting that there may be no singular criteria capable of both perfect detection and zero false positive rate. It is widely acknowledged that weaker tropical storms are difficult to track, and the observational record of these less-intense, short-lived storms is questionable (Landsea et al., 2010).

Automated tropical cyclone tracking has received particular attention throughout the literature. Published techniques generally rely on identifying TC candidates as near-surface (generally 850hPa) local vorticity maxima or as local sea level pressure minima. To eliminate extratropical cyclones and weak cyclonic depressions, it is further required that the candidate be associated with a nearby warm core and be associated with a minimum threshold on surface winds for at least 1-3 days. The definition of a "warm core" varies between modeling centers, including such options as air temperature anomaly on pressure surfaces (Vitart et al., 1997; Zhao et al., 2009; Murakami et al., 2012), geopotential thickness (Tsutsui and Kasahara, 1996) and decay of vorticity with height (Bengtsson et al., 2007; Strachan et al., 2013). Additional filtering of candidate storms over topography or within a specified latitudinal range may be required. To

3

better match observations, additional geographical, model or feature-dependent criteria may be applied (Camargo and Zebiak, 2002; Walsh et al., 2007; Murakami and Sugi, 2010; Murakami et al., 2012).

African Easterly Waves

A similar manual study to Akyildiz (1985) was performed by Reed et al. (1988) to detect and track African easterly waves.

2 Algorithms

In terms of regular latitude-longitude coordinates, great circle distance between points (λ_1, φ_1) and (λ_2, φ_2) is defined via the symmetric operation

$$r(\lambda_1, \varphi_1; \lambda_2, \varphi_2) = a \arccos\left(\sin\varphi_1 \sin\varphi_2 + \cos\varphi_1 \cos\varphi_2 \cos(\lambda_1 - \lambda_2)\right). \tag{1}$$

For nearest neighbor search `nearest_neighbor(P, f)`, we employ a three-dimensional KD tree (Bentley, 1975) using the implementation of Tsiombikas (2015). A key advantage of KD trees is their relatively efficient $O(\log n)$ nearest neighbor search.

2.1 MapReduce

At high horizontal resolutions, climate datasets become increasingly unwieldy to store and analyze. Simultaneously, single-core computing systems have effectively reached peak performance, and so supercomputing systems have been forced to adopt massively parallel infrastructure. To address the need for feature tracking in multi-decadal high-resolution ensemble datasets, algorithms based on MapReduce have been developed (Prabhat et al., 2012; Zarzycki and Jablonowski, 2014). In this case, candidates are first isolated on individual time slices (in an embarrassingly parallel manner), and then stitched across time to build trajectories. For detection codes based on MapReduce, parallel performance is usually bottlenecked by file system access times (I/O).

4

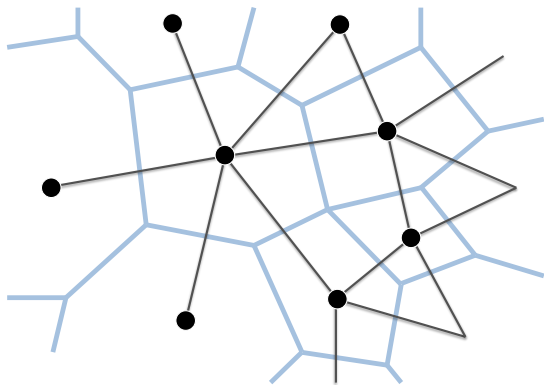## 2.2 Unstructured Grid Specification

Fig. 1. An example unstructured grid.

2.3 Extrema Detection

```
set P = find_all_minima(field G)
  for each cell f
    is_minima[f] = true
    for each neighbor cell v of f
      if G[v] < G[f] then
        is_minima[f] = false
    if is_minima[f] then
      insert f into P
```

```
set Q = merge_candidates_minima(field G, set P, dist)
  K = build_kd_tree(P)
  for each candidate p in P
    retain_p = true
    N = all_neighbors(K, p, dist)
    for all q in N
      if (G[q] < G[p]) then retain_p = false
    if retain_p then insert p into Q
```

2.4 Closed Contour Criteria

```
point pmax = find_max_near(point p, field G, maxdist)
  set visited = {}
  set tovisit = {p}
  pmax = p
  while tovisit is not empty
    q = remove point from tovisit
    if (q in visited) then continue
    add q to visited
    if (gcdist(p,q) > maxdist) then continue
    if (G[q] > G[pmax]) then pmax = q
```

```
closed_contour(point p, field G, dist, maxdist, thresh)
  p0 = find_max_near(p, G, maxdist)
  set visited = {}
  set tovisit = {p0}
  while tovisit is not empty
    q = remove point from tovisit
    if (q in visited) then continue
    add q to visited
    if (gcdist(p0,q) > dist) then return false
    if (G[p0] - G[q] < thresh) then
      add all neighbors of q to tovisit
  return true
```

An analogous algorithm no_closed_contour() is also provided, which has similar functionality but discards candidates that satisfy the closed contour criteria.
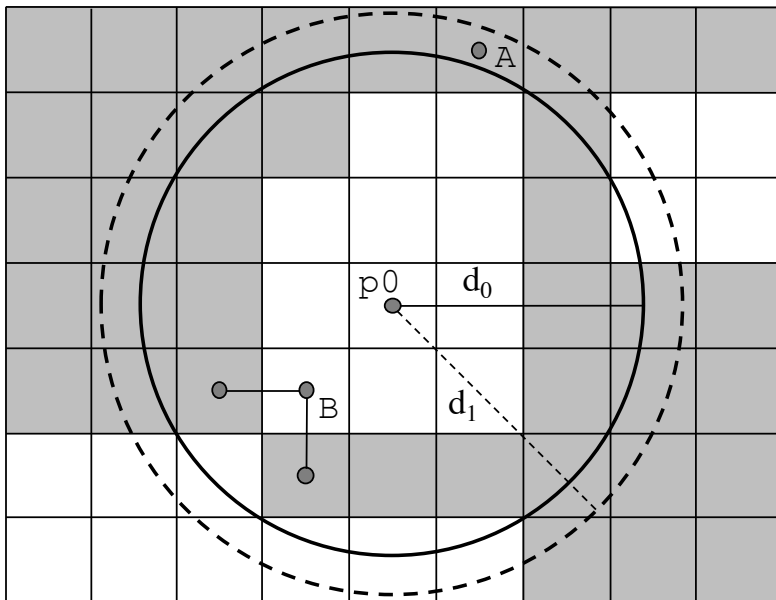

Fig. 2. An illustration of the closed contour criteria. Cells shaded in white (gray) satisfy (do not satisfy) the threshold of the field value at p0. Since only edge-neighbors are included, B constitutes a boundary to the interior of the closed contour. Because A lays outside the solid circle, the contour with distance d_0 is not a closed contour, but the dashed contour with distance d_1 is a closed contour.

2.5 Thresholding

```
threshold(point p, field G, dist, thresh)
  p0 = find_max_near(p, G, dist)
  if (G[p0] < thresh) then
    return false
  else
    return true
```

2.6 Stitching

```
path set S = stitch_nodes(P[1..T], dist)
  for each time level t = 1..T
    K[t] = build_kd_tree(P[t])
  for each time level t = 1..T
    while P[t] is not empty
      initialize empty path s
      p = remove next candidate from P[t]
      add p into s
      for time level u = t+1..T
        q = kd_tree_nearest_neighbor(K[u], p)
        if (q in P[u]) and (gcdist(p,q) < dist) then
          add q into s
          remove q from P[u]
          p = q
        else break
      add s into S
```

3   Validation

3.1   Tropical Cyclones

3.2   Extratropical Cyclones

3.3   African Easterly Waves

4   Conclusions

Acknowledgements.

A1 Software Documentation: DetectCyclonesUnstructured

```
Usage: DetectCyclonesUnstructured <parameter list>
Parameters:
  --in_data <string> [""]
  --in_connect <string> [""]
  --out <string> [""]
  --searchbymin <string> [""] (default PSL)
  --searchbymax <string> [""]
  --maxlat <double> [0.000000] (degrees)
  --minlat <double> [0.000000] (degrees)
  --topofile <string> [""]
  --maxtopoht <double> [0.000000] (m)
  --mergedist <double> [0.000000] (degrees)
  --closedcontourcmd <string> [""] [var,dist,delta,minmaxdist;...]
  --noclosedcontourcmd <string> [""] [var,dist,delta,minmaxdist;...]
  --thresholdcmd <string> [""] [var,op,value,dist;...]
  --outputcmd <string> [""] [var,op,dist;...]
  --timestride <integer> [1]
  --regional <bool> [false]
  --out_header <bool> [false]
  --verbosity <integer> [0]

    --in_data <string>
    The input datafile in NetCDF format.

    --in_connect <string>
    A connectivity file, which uses a vertex list to describe the graph structure of the input
    grid. This parameter is not required if the data is on a latitude-longitude grid.
```

14

`--out <string>`
The output file containing the filtered list of candidates in plain text format.

`--searchbymin <string>`
The input variable to use for initially selecting candidate points (defined as local minima). By default this is "PSL", representing detection of surface pressure minima. Only one of `searchbymin` and `searchbymax` may be set.

`--searchbymax <string>`
The input variable to use for initially selecting candidate points (defined as local maxima). Only one of `searchbymin` and `searchbymax` may be set.

`--maxlat <double>`
The maximum absolute latitude for candidate points. Candidates at higher latitudes are discarded.

`--minlat <double>`
The minimum absolute latitude for candidate points. Candidates at lower latitudes are discarded.

`--topofile <string>`
An auxiliary file containing topographic information.

`--maxtopoht <double>`
If `--topofile` is specified, discard candidates over topography higher than specified by this parameter.

`--mergedist <double>`
Merge candidate points with distance (in degrees) shorter than the specified value. Among two candidates within the merge distance, only the candidate with lowest `searchbymin` or highest `searchbymax` value will be retained.

15

`--closedcontourcmd <cmd1>;<cmd2>;...` Eliminate candidates if they do not have a closed contour. Closed contour commands are separated by a semi-colon. Each closed contour command takes the form `var,dist,delta,pivotdist`. These arguments are as follows.

5   `var <variable>` The variable used for the contour search.

     `dist <double>` The great-circle distance (in degrees) from the pivot within which the closed contour criteria must be satisfied.

     `delta <double>` The amount by which the field must change from the pivot value. If positive (negative) the field must increase (decrease) by this value along the 10   contour.

     `pivotdist <double>` The distance away from the candidate to search for the pivot. If `delta` is positive (negative), the pivot is a local minimum (maximum).

`--noclosedcontourcmd <cmd1>;<cmd2>;...`
As `closedcontourcmd`, except eliminates candidates if a closed contour is present.

15   `--thresholdcmd <cmd1>;<cmd2>;...` Eliminate candidates that do not satisfy a threshold criteria (there must exist a point within a given distance of the candidate that satisfies a given equality or inequality). Threshold commands are separated by a semi-colon. Each threshold command takes the form `var,op,value,dist`. These arguments are as follows.

20   `var <variable>` The variable used for the contour search.

     `op <string>` Operator that must be satisfied for threshold (options include $>$, $>=$, $<$, $<=$, $=$, $!=$).

     `value <double>` The value on the RHS of the comparison.

     `dist <double>` The great circle distance away from the candidate to search for 25   a point that satisfies the threshold (in degrees).

16

--outputcmd <cmd1>;<cmd2>;... Include additional columns in the output file. Output commands take the form var,op,dist. These arguments are as follows.

var <variable> The variable used for the contour search.

op <string> Operator that is applied over all points within the specified distance of the candidate (options include max, min, avg, maxdist, mindist).

dist <double> The great circle distance away from the candidate wherein the operator is applied (in degrees).

--timestride <integer>
Only examine discrete times at the given stride (by default 1).

--regional
When a latitude-longitude grid is employed, do not consider longitudinal boundaries to be periodic.

--out_header
Output a header describing the columns of the data file.

--verbosity <integer>
Set the verbosity level (default 0).

A1.1 Variable Specification

Quantities of type <variable> include both NetCDF variables in the input file (for example, "Z850") and simple operations performed on those variables. By default it is assumed that NetCDF variables are specified in the .nc file as

        float Z850(time, lat, lon)   or   float Z850(time, ncol)

for structured latitude-longitude grids and unstructured grids, respectively. If variables have no time variable, they have the related specification

17

```
                    float Z850(lat, lon)   or   float Z850(ncol)
```

If variables include an additional dimension, for instance,

```
    float Z(time, lev, lat, lon)   or   float Z(time, lev, ncol)
```

they may be specified on the command-line as `Z(<lev>)`, where the integer index `<lev>`
corresponds to the first dimension (or the dimension after `time`, if present).

Simple operations on variables are also supported, including

`_VECMAG(<variable>, <variable>)` 2-component vector magnitude,

`_PLUS(<variable>, <variable>)` Pointwise sum of variables,

`_DIFF(<variable>, <variable>)` Pointwise difference of variables.

The following are valid examples of `<variable>` type,

`_VECMAG(U850, V850)`   and   `_DIFF(U(3),U(5))`.

A2 Software Documentation: StitchNodes

```
Usage: StitchNodes <parameter list>
Parameters:
    --in <string> [""]
    --out <string> [""]
    --format <string> ["no,i,j,lon,lat"]
    --range <double> [5.000000] (degrees)
    --minlength <integer> [3]
    --min_endpoint_dist <double> [0.000000] (degrees)
    --min_path_dist <double> [0.000000] (degrees)
    --maxgap <integer> [0]
    --threshold <string> [""] [col,op,value,count;...]
    --timestride <integer> [1]
    --out_format <string> ["std"] (std|visit)
```

`--in <string>`
The input file (a list of candidates from DetectCyclonesUnstructured).

`--out <string>`
The output file containing the filtered list of candidates in plain text format.

`--format <string>`
The structure of the columns of the input file.

`--range <double>`
The maximum distance between candidates along a path.

`--minlength <integer>`
The minimum length of a path (in terms of number of discrete times).

`--min_endpoint_dist <double>`
The minimum great-circle distance between the first candidate on a path and the last candidate (in degrees).

`--min_path_dist <double>`
The minimum path length, defined as the sum of all great-circle distances between candidate nodes (in degrees).

`--maxgap <integer>`
The largest gap (missing candidate nodes) along the path (in discrete time points).

`--threshold <cmd1>;<cmd2>;...`
Eliminate paths that do not satisfy a threshold criteria (a specified number of candidates along path must satisfy an equality or inequality). Threshold commands are separated by a semi-colon. Each threshold command takes the form `col,op,value,count`. These arguments are as follows.

> `col <integer>` The column in the input file to use in the threshold criteria.

op <string> Operator used for comparison of column value (options include >, >=, <, <=, =, !=).

value <double> The value on the right-hand-side of the operator.

count <integer> The minimum number of candidates along the path that must 5 satisfy this criteria.

--timestride <integer>
Only examine discrete times at the given stride (by default 1).

References

Akyildiz, V.: Systematic errors in the behaviour of cyclones in the ECMWF operational models, Tellus
10 A, 37, 297–308, 1985.

Alpert, P., Neeman, B., and Shay-El, Y.: Climatological analysis of Mediterranean cyclones using ECMWF data, Tellus A, 42, 65–77, 1990.

Bell, G. D. and Bosart, L. F.: A 15-year climatology of Northern Hemisphere 500 mb closed cyclone and anticyclone centers, Monthly Weather Review, 117, 2142–2164, 1989.

15 Bengtsson, L., Hodges, K. I., and Esch, M.: Tropical cyclones in a T159 resolution global climate model: comparison with observations and re-analyses, Tellus A, 59, 396–416, doi:10.1111/j.1600-0870.2007.00236.x, 2007.

Bentley, J. L.: Multidimensional binary search trees used for associative searching, Communications of the ACM, 18, 509–517, 1975.

20 Camargo, S. J. and Zebiak, S. E.: Improving the Detection and Tracking of Tropical Cyclones in Atmospheric General Circulation Models, Weather and Forecasting, 17, 1152–1162, 2002.

Hodges, K.: Feature tracking on the unit sphere, Monthly Weather Review, 123, 3458–3465, 1995.

Hodges, K. et al.: A general-method for tracking analysis and its application to meteorological data, Monthly Weather Review, 122, 2573–2586, 1994.

25 Hodges, K. I., Hoskins, B. J., Boyle, J., and Thorncroft, C.: A comparison of recent reanalysis datasets using objective feature tracking: Storm tracks and tropical easterly waves, Monthly Weather Review, 131, 2012–2037, 2003.

Klein, W. H.: Principle tracks and mean frequencies of cyclones and anticyclones in the Northern Hemisphere, US Weather Bureau, 1957.

König, W., Sausen, R., and Sielmann, F.: Objective identification of cyclones in GCM simulations, Journal of Climate, 6, 2217–2231, 1993.

Lambert, S. J.: A cyclone climatology of the Canadian Climate Centre general circulation model, Journal of Climate, 1, 109–115, 1988.

Landsea, C. W., Vecchi, G. A., Bengtsson, L., and Knutson, T. R.: Impact of Duration Thresholds on Atlantic Tropical Cyclone Counts, Journal of Climate, 23, 2508–2519, doi:10.1175/2009JCLI3034.1, 2010.

Le Treut, H. and Kalnay, E.: Comparison of observed and simulated cyclone frequency distribution as determined by an objective method, Atmosfera, 3, 1990.

Murakami, H. and Sugi, M.: Effect of Model Resolution on Tropical Cyclone Climate Projections, SOLA, 6, 73–76, doi:10.2151/sola.2010-019, 2010.

Murakami, H., Wang, Y., Yoshimura, H., Mizuta, R., Sugi, M., Shindo, E., Adachi, Y., Yukimoto, S., Hosaka, M., Kusunoki, S., Ose, T., and Kitoh, A.: Future Changes in Tropical Cyclone Activity Projected by the New High-Resolution MRI-AGCM*, J. Clim., 25, 3237–3260, doi:10.1175/JCLI-D-11-00415.1, 2012.

Murray, R. J. and Simmonds, I.: A numerical scheme for tracking cyclone centres from digital data. Part I: Development and operation of the scheme, Australian Meteorological Magazine, 39, 155–166, 1991.

Petterssen, S.: Weather analysis and forecasting. 2. Weather and weather systems, McGraw-Hill, 1956.

Prabhat, Rübel, O., Byna, S., Wu, K., Li, F., Wehner, M., and Bethel, W.: TECA: A Parallel Toolkit for Extreme Climate Analysis, Procedia Computer Science, 9, 866 – 876, doi:10.1016/j.procs.2012.04.093, proceedings of the International Conference on Computational Science, 2012, 2012.

Reed, R., Hollingsworth, A., Heckley, W., and Delsol, F.: An evaluation of the performance of the ECMWF operational system in analyzing and forecasting easterly wave disturbances over Africa and the tropical Atlantic, Monthly weather review, 116, 824–865, 1988.

Rice, J.: The Durivation of Computer-based Synophic Climatology of Southern Hemisphere Extratropical Cyclones, Unpublished B. Sc. Honours thesis. University of Melbourne, 1982.

Strachan, J., Vidale, P. L., Hodges, K., Roberts, M., and Demory, M.-E.: Investigating Global Tropical Cyclone Activity with a Hierarchy of AGCMs: The Role of Model Resolution, J. Clim., 26, 133–152, doi:10.1175/JCLI-D-12-00012.1, 2013.

Taljaard, J.: Development, distribution and movement of cyclones and anticyclones in the Southern Hemisphere during the IGY, Journal of Applied Meteorology, 6, 973–987, 1967.

21

Tsiombikas, J.: kdtree: A simple C library for working with KD-Trees, https://github.com/jtsiomb/kdtree, accessed: 2015-09-18, 2015.

Tsutsui, J.-i. and Kasahara, A.: Simulated tropical cyclones using the National Center for Atmospheric Research community climate model, Journal of Geophysical Research: Atmospheres, 101, 15 013–15 032, doi:10.1029/95JD03774, 1996.

Vitart, F., Anderson, J. L., and Stern, W. F.: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations, Journal of Climate, 10, 745–760, 1997.

Walsh, K. J. E., Fiorino, M., Landsea, C. W., and McInnes, K. L.: Objectively Determined Resolution-Dependent Threshold Criteria for the Detection of Tropical Cyclones in Climate Models and Reanalyses, J. Clim., 20, 2307, doi:10.1175/JCLI4074.1, 2007.

Whittaker, L. M. and Horn, L.: Atlas of Northern Hemisphere extratropical cyclone activity, 1958- 1977, 1982.

Williamson, D. L.: Storm track representation and verification, Tellus, 33, 513–530, 1981.

Zarzycki, C. M. and Jablonowski, C.: A multidecadal simulation of Atlantic tropical cyclones using a variable-resolution global atmospheric general circulation model, J. Adv. Model. Earth Syst., 6, 805–828, doi:10.1002/2014MS000352, 2014.

Zhao, M., Held, I. M., Lin, S.-J., and Vecchi, G. A.: Simulations of Global Hurricane Climatology, Interannual Variability, and Response to Global Warming Using a 50-km Resolution GCM, J. Clim., 22, 6653–6678, 2009.