TempestExtremes: A Framework for 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

Automated pointwise feature tracking is an algorithmic technique for identification and tracking of meteorological features such as extratropical cyclones, tropical cyclones and tropical easterly waves, and has emerged as an important and desirable data processing capability in climate
5 science. Software tools for feature tracking – typically referred to as "trackers" – have been employed throughout the literature to answer pressing scientific questions on anticipated changes in atmospheric features under climate change. This paper describes a new open-source software framework for pointwise feature tracking that is applicable to a wide array of climate datasets using either structured or unstructured grids. To enable support for a wide array of detection
10 schemes, a suite of algorithmic kernels have been developed that capture the core functionality of algorithmic tracking routines from throughout the literature. A review of efforts related to pointwise feature tracking from the past three decades is included. Selected results using both reanalysis datasets and unstructured grid simulations are provided.

1 Introduction

15 Automated pointwise feature tracking has emerged as an important and desirable data processing capability in climate science. Software tools for feature tracking – typically referred to as "trackers" – have been employed throughout the literature to answer pressing scientific questions on anticipated changes in atmospheric features under climate change. Exploration of the literature on trackers exposes a vast breadth of potential techniques (see Appendices A1, A2
20 and A3) that have been applied to climate datasets with varied spatial resolution and temporal frequency. Nonetheless, the definition of an optimal objective criteria for key atmospheric features has eluded development, and ambiguity in the formal definition of these features suggests that there may be no singular criteria capable of both perfect detection and zero false positive rate. Further, as observed by Walsh et al. (2007) for tropical cyclones and Neu et al. (2013) for
25 extratropical cyclones, feature tracking schemes can produce wildly varying results depending on the specific choice of threshold variables and values. Consequently, we argue that conclu-

2

sions drawn from these trackers should use an ensemble of detection thresholds and variables. To this end, it is the goal of this paper to develop and present a unified framework that enables a variety of tracking procedures applicable across effectively arbitrary spatial resolution and temporal frequency.

Most algorithmic trackers of pointwise features (such as cyclones and eddies) share a common detection procedure:

1. Identify an initial set of candidate points by searching for local extrema. Local extrema can be further isolated, for instance, by requiring that the local extrema be sufficiently anomalous when contrasted with their neighbors. For most cyclonic structures, either minima in the sea level pressure field or maxima in the absolute value of the relative vorticity are used.

2. Eliminate candidate points that do not satisfy a prescribed set of thresholds. For instance, tropical cyclones typically require the presence of an upper-level warm core that is sufficiently near the sea level pressure minima that defines the storm center.

3. Connect candidate points together in time to generate feature paths, eliminating paths that are of insufficient length or do not meet additional criteria.

The procedure described above fits into a general framework known as MapReduce (Dean and Ghemawat, 2008), which is a combination of a Map(), an embarrassingly parallel candidate identification procedure applied to individual time slices, and a Reduce(), which stitches candidates across time to build feature tracks. A key advantage of employing this framework is that substantial work has been undertaken to understand optimal strategies for parallelization of MapReduce-type algorithms (e.g. Prabhat et al. (2012)) in order to mitigate bottlenecks associated with I/O and load balancing. TempestExtremes currently implements a simple parallelization strategy via MPI, although future work on this issue is forthcoming.

The development of a framework for feature tracking that is robust across essentially arbitrary datasets requires some additional considerations. In particular, we expect that our framework should:

3

...use great-circle arcs for all distance calculations. This avoids issues associated with latitude-longitude distance that emerges near the poles.

...support structured and unstructured grids. This eliminates the need for post-processing of large native-grid output files and enables detection and characterization simultaneous with the model execution.

...not contain hard-coded variable names, so as to ensure robust applicability across reanalysis datasets and applicability to a variety of problems.

...allow for easy intercomparison of detection schemes by enabling detection criteria and thresholds that are compactly specified on the command line.

Well-known automated software trackers include Kevin Hodges' TRACK code (Hodges, 2015) and the Geophysical Fluid Dynamics Laboratory (GFDL) TSTORMS package (?) Others?. Both of these software packages have been used extensively throughout the literature for studies examining pointwise features in the atmosphere, but do not completely satisfy our four requirements above.

Although the actual procedures used for feature tracking vary throughout the literature, a review of this material reveals several core algorithms (kernels) must be exposed to the user. Based on our analysis, the most commonly employed kernels are as follows:

- Computation of anomalies in a data field from a spatially averaged mean.

- Identification of local extrema in a given 2D data field (for instance, sea level pressure minima).

- Determination of whether a closed contour exists in a data field around a particular point.

- Determination of whether, in the neighborhood of a particular point, a data field satisfies a given threshold.

- Stitching of candidates from sequential time slices to build feature tracks.

4

It is desirable that any generalization these kernels is insensitive to the spatial resolution of the input data; that is, criteria that require averaging or searching over a discrete number of grid points around each candidate (which is common throughout the early tracking literature) is incompatible with scale insensitivity since detection would be dependent on the spatial resolution
5 of the data. Note that identification of local extrema is not a resolution-insensitive procedure, since the number of extrema will often scale with the number of spatial data points, and so initial identification is typically augmented with the more scale-insensitive closed contour criteria (see section 2.5).

The remainder of the paper is organized as follows: Section 2 describes the algorithms and
10 kernels that have been implemented in the TempestExtremes software framework. Selected examples of tropical cyclone, extratropical cyclone and tropical easterly wave detection are then provided in section 3, followed by conclusions in section 4. The appendices provide a review of relevant literature on pointwise feature trackers of extratropical cyclones (appendix A1), tropical cyclones (appendix A2) and tropical easterly waves (appendix A3). A technical guide to the
15 use of the TempestExtremes tools DetectCyclonesUnstructured and StitchNodes is provided in appendices A4 and A5.

1.1 Code and Data Availability

The open-source software described in this manuscript has been released as part of the TempestExtremes software package, and is available for use under the Lesser GNU Public License
20 (LGPL). All software can be obtained from GitHub at:

https://github.com/ClimateGlobalChange/tempestextremes

2 TempestExtremes Algorithms and Kernels

This section describes the building blocks that have been utilized in constructing our detection and characterization framework. As mentioned earlier, in order to avoid sensitivity of the
25 detection scheme to grid resolution, great-circle-distance has been employed throughout. In

5

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

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

Algorithmically, this calculation is implemented as gcdist(p,q) for given graph nodes p
5 and q.

2.1 Efficient neighbor search using k-d Trees

Three-dimensional ($k = 3$) k-d trees (Bentley, 1975) are used throughout our detection code using the implementation of Tsiombikas (2015). Although k-d trees use 3D straight-line instance instead of great-circle-distance, we utilize the observation that straight-line and great-circle dis-
10 tance maintain the same ordering for points confined to the surface of the sphere. In particular, we utilize three key functions made available by the k-d tree implementation:

K = build_kd_tree(P) constructs a k-d tree K from a point set P.

q = kd_tree_nearest_neighbor(K, p) locates the nearest neighbor q to point p using the k-d tree K.

15 S = kd_tree_all_neighbors(K, p, dist) locates all points that are within a distance dist of a point p within the k-d tree K.

A key advantage of k-d trees is their relatively efficient $O(n \log n)$ construction time and $O(\log n)$ average time nearest neighbor search.

2.2 Unstructured grid specification

20 For purposes of determining connectivity of the unstructured grid, we require the specification of a graph such as the one depicted in Figure 1. The connectivity information is stored textually as an adjacency list via a variable-length comma-separated variable file. The total number of nodes (N) is specified at the top of the file, followed by N lines containing the longitude (lon),

latitude (lat), associated area, number of adjacent nodes, and finally a 1-indexed list of all adjacent nodes, such as depicted below:

```
<total number of nodes>
<lon>,<lat>,<area>,<# adj. nodes>,<first adj. node>,..,<last adj. node>
...
```

2.3 Computing a spatial averaged mean

Many existing tracking algorithms use either a spatially-averaged mean field or an anomaly field computed against the spatially-averaged mean (Haarsma et al., 1993; Bengtsson et al., 1995). The mean operation (implemented in TempestExtremes as _MEAN() in the variable specification) is computed on unstructured grids via graph search (see Algorithm 1). Anomaly from the mean can then be computed in conjunction with the _DIFF operator (see appendix A4.2).

2.4 Extrema detection

For purposes of computational efficiency, candidate points are initially located by identifying local extrema in a given field (for instance SLP) via find_all_minima (Algorithm 2). Candidates are then eliminated if they are "too close" to stronger extrema (Algorithm 3) (e.g. Pinto et al. (2005)). The initial search field is specified to TempestExtremes either via the --searchbymin or --searchbymax command line argument. The merge distance used in merge_candidates_minima is specified via the --mergedist command line argument.

2.5 Closed contour criteria

Although a first pass at candidate points may be made by looking for local extrema (comparing against all neighboring nodes), this criteria is not robust across model resolution. That is, the

distance between a node and its neighbors decreases proportional to the local grid spacing, and so does not definite "physical" criterion. Consequently, we instead advocate for a *closed contour criteria* to define candidate nodes. Closed contours were first employed by Bell and Bosart (1989), who used a 30m 500 mb geopotential height contour to identify closed circulation cen-
5 ters. 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, the use of radial arms to define the closed contour is again sensitive to model resolution, since it has the potential to only sample as many neighbors as radial arms employed.

Here we propose an alternative closed contour criteria that is largely insensitive to model
10 resolution that uses graph search to ensure that all paths along the unstructured grid from an initial location p0 lead to a sufficiently large decrease (or increase) in a given field G. This criteria is illustrated in Figure 2, and is implemented in Algorithm 4 and 5 (for closed contours around local maxima). The closed contour criteria is implemented in TempestExtremes via the command line argument --closedcontourcmd. An analogous command line ar-
15 gument --noclosedcontourcmd is also provided, which has similar functionality but discards candidates that satisfy the closed contour criteria (this may be desirable, for instance, to identify cyclonic structures that do not have a warm core).

2.6 Thresholding

Additional threshold criteria may be applied at the Map() stage in order to further eliminate
20 undesirable candidates. For example, a common threshold criteria requires that a field G satisfy some minimum value within a distance dist of the candidate, as implemented in Algorithm 6.

2.7 Stitching

The basic track stitching procedure (which represents the Reduce() stage in MapReduce) is implemented in Algorithm 7 using the output from the Map() procedure at each time level
25 (stored in set array P[1..T]), a maximum great-circle-distance between nodes, dist, and a maximum gap size, maxgap. Here, gap size refers to the maximum number of sequential non-

detections that can occur before a path is considered terminated. This argument is useful, for instance, for tracking tropical storms that temporarily weaken before strengthening into tropical cyclones.

For simplicity, k-d trees are constructed at each time level in order to maximize the efficiency
5 of the search. Each candidate pair (time, node) can only be used in one path, and so construction simply requires exhausting the list of available candidates. Once paths have been constructed, additional criteria can be applied – for instance, minimum path length or additional criteria based on minimum path length or minimum distance between the start and endpoints of the path (see Appendix A5). Thresholds based on field values may also be applied, *e.g.* sea level
10 pressure must be below a particular value for at least 8 time steps of each track.

3 Selected examples

Several selected examples of the feature detection tool are now provided. The first three examples use data from the NCEP Climate Forecast System Reanalysis (CFSR), available at 0.5 degree global resolution with 6-hourly output from 1979-present (Saha et al., 2010). The
15 remaining example uses a custom variable-resolution simulation with CESM (Zarzycki and Jablonowski, 2014) on both the native grid data and the regridded latitude-longitude grid data.

3.1 Tropical cyclones in CFSR

Our first example employs TempestExtremes for tropical cyclones. The command line we use to detect tropical cyclone-like features in CFSR is provided below. Three-dimensional (time
20 + 2D space) hyperslabs of CFSR data have been extracted, with TMP_L100 corresponding to 400hPa air temperature, and U_GRD_L100 and V_GRD_L100 corresponding to 850hPa zonal and meridional velocities. Candidates are initially identified by minima in the sea level pressure (PRMSL_L101), and then eliminated if a smaller minimum exists within a great-circle-distance of 2.0 degrees. The closed contour criteria is then applied, requiring an increase in SLP of at
25 least 200Pa over a distance of 4 degrees away from the candidate node, and a decrease in 400hPa

9

air temperature of 0.4K within 8 degrees of the node within 1.1 degrees of the candidate with maximum air temperature. Since CFSR is on a structured latitude-longitude grid, the output format is $i, j, lon, lat, psl, maxu, zs$, where i, j are the longitude and latitude coordinates within the dataset, lon, lat are the actual longitude and latitude of the candidate, psl is the SLP at the candidate point (equal to the maximum SLP within 0 degrees of the candidate), $maxu$ is the vector magnitude of the maximum 850 hPa wind within 4 degrees of the candidate, and zs is the topographic height at the candidate point.

```
    ./DetectCyclonesUnstructured
      --in_data "$uvfile;$tpfile;$hfile" --out $outf
      --searchbymin PRMSL_L101 --mergedist 2.0
      --closedcontourcmd "PRMSL_L101,200.,4,0;
        TMP_L100,-0.4,8.0,1.1"
      --outputcmd "PRMSL_L101,max,0;
        _VECMAG(U_GRD_L100,V_GRD_L100),max,4;
        HGT_L1,max,0"
```

All outputs from DetectCyclonesUnstructured are then concatenated into a single file containing candidates at all times (pgbhnl.dcu_tc_all.dat). Candidates are then stitched in time to form paths, with a maximum distance between candidates of 8.0 degrees (great-circle-distance), consisting of at least 8 candidates per path, and with a maximum gap size of 2 (most consecutive timesteps with no associated candidate). Because localized low-pressure regions that are unrelated to tropical cyclones can form as a consequence of topographic forcing, we also require that for at least 8 time steps the underlying topographic height (zs) be at most 100 meters. The associated command line for StitchNodes is:

```
    ./StitchNodes
      --in pgbhnl.dcu_tc_all.dat
      --out pgbhnl.dcu_tc_stitch.dat
      --format "i,j,lon,lat,psl,maxu,zs"
      --range 8.0 --minlength 8 --maxgap 2
      --threshold "zs,<=,100.0,8"
```

Once the complete set of tropical cyclone paths has been computed, total tropical cyclone counts over each 2° grid cell are plotted in Figure 3. Overall the results show very good agreement with reference fields (Citation?).

3.2 Extratropical cyclones in CFSR

For our second example, we are interested in tracking extratropical cyclone features. The command line we have used to detect cyclonic features without the characteristic warm-core of tropical cyclones (here referred to as extratropical cyclones) is given below. The command is identical to the TC detection configuration specified in section 3.1, except requiring that the feature does not possess a closed contour in the 400hPa temperature field (no warm core).

```
./DetectCyclonesUnstructured
   --in_data "$uvfile;$tpfile;$hfile" --out $outf
   --searchbymin PRMSL_L101 --mergedist 2.0
   --closedcontourcmd "PRMSL_L101,200.,4,0"
   --noclosedcontourcmd "TMP_L100,-0.4,8.0,1.1"
   --outputcmd "PRMSL_L101,max,0;
      _VECMAG(U_GRD_L100,V_GRD_L100),max,4;
      HGT_L1,max,0"
```

Stitching is similarly analogous to section 3.1, except using a slightly more strict criteria on the underlying topographic height. The topographic filtering proved necessary in order to adequately filter out an abundance of topographically-driven low pressure systems, particularly in the Himalayas region. The command line used for stitching is given below:

```
./StitchNodes
   --in pgbhnl.dcu_tc_all.dat
   --out pgbhnl.dcu_tc_stitch.dat
   --format "i,j,lon,lat,psl,maxu,zs"
   --range 8.0 --minlength 8 --maxgap 2
   --threshold "zs,<=,70.0,8"
```

11

Once the complete set of extratropical cyclone paths has been computed, total extratropical cyclone density over each $2°$ grid cell is plotted in Figure 4. Although not extensively verified, the qualitative density of extratropical cyclones is well within the range of results from different trackers, as given by Neu et al. (2013).

3.3 Tropical easterly waves in CFSR

Tropical easterly waves are our third example of a pointwise feature that has been assessed in the tracking literature. In this example, tracking is performed in the northern hemisphere (southern hemisphere) using maxima (minima) in the 600hPa relative vorticity field. Since CFSR only provides absolute vorticity, relative vorticity must first be extracted by taking the difference between absolute vorticity and the planetary vorticity (the Coriolis parameter). This is done on the command line via `_DIFF(ABS_V_L100,_F())`, where `ABS_V_L100` is the CFSR absolute vorticity variable and `_F()` is a built-in function for computing the Coriolis parameter (equal to $f = 2\Omega \sin \phi$). In the northern hemisphere, we follow Thorncroft and Hodges (2001) and isolate tropical easterly wave features by requiring a drop of relative vorticity equal to 5×10^{-5} s^{-1}, and have discarded detections outside of the latitudinal range $[0, 30N]$. The command line used is as follows:

```
./DetectCyclonesUnstructured
  --in_data "$uvfile;$hfile" --out $outf
  --searchbymax "_DIFF(ABS_V_L100(0),_F())" --mergedist 2.0
  --closedcontourcmd "_DIFF(ABS_V_L100(0),_F()),-5.e-5,4,0"
  --outputcmd "ABS_V_L100(0),max,0;_DIFF(ABS_V_L100(0),_F()),max,0;
      HGT_L1,max,0"
  --minlat -35.0 --maxlat 35.0
```

Northern hemisphere tropical easterly wave paths are constructed using a maximum distance of $3°$ great-circle-distance between subsequent detections, a minimum path length equal to 8 sequential detections and no allowed gaps. We further require that the easterly waves have a distance of at least $10°$ between start and endpoints and are present in the northern hemisphere ($\phi \geq 0°$) for at least 8 timesteps. The command line is as follows.

```
   ./StitchNodes
     --in pgbhnl.dcu_aew_nh_all.dat --out pgbhnl.dcu_aew_nh_stitch.dat
     --format "i,j,lon,lat,relv,zs"
     --range 3.0 --minlength 8 --maxgap 0
5    --min_endpoint_dist 10.0
     --threshold "lat,<=,25.0,8;lat,>=,0.0,8"
```

An analogous procedure is applied in the southern hemisphere, except searching on minima in the relative vorticity field and limiting the latitudinal range in StitchNodes to [25S,0] for at least 8 timesteps. Counts of total (northern hemisphere plus southern hemisphere) tropical
10 easterly waves within each 2° grid volume are given in Figure 5, showing heavy wave activity throughout the Atlantic and Pacific basins. These results are very similar to other reported easterly wave densities, such as Belanger et al. (2014) and Thorncroft and Hodges (2001), except for (a) the substantially enhanced tropical easterly wave count reported over eastern Africa (which could be eliminated by filtering over topography) and (b) essentially no observed wave
15 activity off of the western coast of South America. Nonetheless, it is well known that easterly wave climatology varies strongly across reanalysis datasets and exhibits sensitivity to the choice of tracking scheme (Hodges et al., 2003).

3.4 Tropical cyclones in a simulation with CESM

For our final example, we assess the differences in tropical cyclone character obtained from
20 native and regridded datasets. Using the variable-resolution option in the Community Earth System Model (VR-CESM) to refine the northern hemisphere to 0.25° resolution, a simulation of a complete hurricane season (September - January) has been performed. With the high-order spectral element dynamical core used to solve the fluid equations in the atmosphere, VR-CESM has been demonstrated to be effective in simulating tropical cyclone-like features (Zarzycki and
25 Jablonowski, 2014; Zarzycki et al., 2014). However, even at the relatively fine global resolution of 0.25°, the eye of the tropical cyclone is only partially resolved. Since VR-CESM uses an unstructured mesh with degrees of freedom stored at spectral element Gauss-Lobatto (GL) nodes, data is typically analyzed only after being regridded to a regular latitude-longitude mesh of ap-

proximately equal resolution. The regridding procedure has the potential to smear out grid-scale features.

For this example, we use the high-order regridding package TempestRemap (Ullrich and Taylor, 2015; Ullrich et al., 2016) for remapping the native spectral element output to a regular latitude-longitude grid with 0.25° grid spacing. For purposes of determining connectivity on the variable-resolution spectral element mesh, we connect GL nodes along the coordinate axis of each quadrilateral element (see Figure 6). DetectCyclonesUnstructured is then applied to both the native grid data and the regridded data on the regular latitude-longitude mesh (using the configuration specified in section 3.1) and tropical cyclones categorized by maximum wind speed (Colin: more here). The results of this analysis are depicted in Figure 7. As expected, the native grid output produces almost identical tracks, but more powerful tropical cyclones overall (with many tropical cyclones dropping down by a full category as a consequence of the remapping procedure).

4 Conclusions

Automated pointwise feature trackers have been frequently and successfully employed over the past several decades to extract useful information from large climate datasets. With spatial and temporal resolution increasing rapidly in response to increasingly powerful computational resources, climate datasets have grown increasingly unwieldy and so there has been a growing need for such large dataset processing tools. This paper has outlined a framework for pointwise feature tracking (TempestExtremes) that exposes a suite of generalized kernels drawn from the literature on trackers of the past several decades. This framework is sufficiently robust to be applicable to many climate and weather datasets, including data on unstructured grids. We expect such a framework would be useful for isolating uncertainties that emerge from particular parameter choices in tracking schemes, or to compute optimal threshold values for detecting pointwise features in, e.g. reanalysis data. Future development plans in TempestExtremes include the construction of analogous kernels for tracking areal features (blobs), such as clouds or atmospheric rivers.

14

A1 A Review of Extratropical Cyclone Tracking Algorithms

This appendix reviews the existing literature on extratropical cyclone tracking, which has been one of the earliest and most common instances of both manual and automated feature tracking. Manual counts of cyclones were performed by Petterssen (1956) in the Northern hemisphere 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.

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 climate 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 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. Murray and Simmonds (1991) extracted low pressure centers from interpolated GCM data using local optimization, based on earlier work in Rice (1982). These original papers primarily sought minimum in the SLP field or looked for maxima in the Laplacian of the SLP field.

Several modern extratropical cyclone detection algorithms remain in use, having built on this earlier work. Short descriptions of many of these schemes are given here. Some of these algorithms use the notion of a local neighborhood or periphery, as defined in Figure 8.

16

– Serreze et al. (1993); Serreze (1995): Assessed $\sim$ 381-400 km arctic dataset for extratopical cyclone behavior. Search on SLP for local minima at least 2 mb higher than neighbors. Tracking is performed with a maximum search distance of 1400km per 12 hour period.

– Sinclair (1994, 1997): Assessed 2.5° ECMWF data over the southern hemisphere. Search on local minima in the 1000 hPa geostrophic vorticity field (computed from the Laplacian of the 1000 hPa geopotential), adjusted for topography and presence of heat lows (see paper for details), satisfying $\zeta_g < -2 \times 10^{-5}\text{s}^{-1}$.

– Blender et al. (1997): Assessed T106 ECMWF analyses ($\sim$ 125 km). Search on local minima in the 1000 hPa geopotential height field. Require a positive mean gradient in the 1000 hPa geopotential height field in a 1000×1000 km^2 area around each candidate. Tracking is performed using nearest-neighbor search with a maximum displacement velocity of 80 km/h, eliminating cyclones with tracks shorter than 3 days.

– Lionello et al. (2002): Assessed a T106 ($\sim$ 125 km) ECHAM-4 dataset. Search on local minima in the SLP field. Tracking requires using previous cyclone velocity to delineate a prediction region, and tracks are discarded if they do not continue into the prediction region.

– Zolina and Gulev (2002): Assessed T106 ($\sim$ 125 km) and T42 ($\sim$ 300 km) datasets. Search on local minima in the SLP field.

– Pinto et al. (2005): Assessed T42 ($\sim$ 300 km) regridded NCEP reanalysis, regridded onto a 0.75° grid by cubic spline interpolation. Search on local minima in pressure field with maxima of quasi-geostrophic relative vorticity, computed from the Laplacian of pressure, within 1200km. Cyclones over topography above 1500m are removed. Require quasi-geostrophic relative vorticity $> 0.1hPa/(°\text{lat})$ and retain only the strongest detections within 3 degrees. Cyclone tracking requires a prediction velocity and search following Murray and Simmonds (1991).

17

– Benestad and Chen (2006): Assessed 2.5° ERA40 data. Search uses multiple least-squares regression to estimate the values of the coefficients of a Fourier approximation followed by 1D search in north-south and east-west directions (in effect smoothing the SLP field).

– Simmonds et al. (2008): Assessed several 2.5° datasets over the arctic. Search on local
5   minima in the Laplacian of pressure, rejecting cyclones over topography above 1000m and requiring the presence of a nearby pressure minima. Identified lows must satisfy a Laplacian with value $> 0.2 hPa/(°\text{lat})^2$ over a radius of 2°. Tracking uses a probability estimate using a predicted position.

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

A2 A Review of Tropical Cyclone Tracking Algorithms

More recently, and as higher resolution climate data has become available, extratropical cyclone tracking techniques have been modified in order to support tropical cyclone tracking. To
15 eliminate "false positives" associated with extratropical cyclones and weak cyclonic depressions, many schemes require 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),
20 geopotential thickness (Tsutsui and Kasahara, 1996) and decay of vorticity with height (Bengtsson et al., 2007a; Strachan et al., 2013). Additional filtering of candidate storms over topography or within a specified latitudinal range may be required. To better match observations, additional geographical, model or feature-dependent criteria may be applied (Camargo and Zebiak, 2002; Walsh et al., 2007; Murakami and Sugi, 2010a; Murakami et al., 2012). It is widely acknowl-
25 edged 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).

A tabulated overview of the thresholds utilized by many of these schemes can be found in Walsh et al. (2007), along with several proposed guidelines on detection schemes. We extend this tabulation with the following short descriptions of many published schemes.

- Broccoli and Manabe (1990): Assessed a R15 ($\sim$ 600 km) and R30 ($\sim$ 300 km) dataset. Search on PSL with 1.5 mb local min (R15) or 0.75 mb local min (R30), with local surface wind velocity > 17 m/s, latitude $< 30°$. Cyclones are tracked over a range of 1200 km / day.

- Wu and Lau (1992): Assessed a $7.5°$ longitude $\times$ $4.5°$ latitude dataset. Search on minimum 1000mb geopotential height, requiring positive 950mb relative vorticity, negative 950mb divergence, positive 500mb vertical velocity, latitude $< 40.5°$, 200mb minus 1000mb layer thickness must be locally maximal and exceed by 60 m the average layer thickness within 1500km west to east, and 950mb wind must be > 17.2 m/s locally. Cyclones are tracked over a range of $7.5°$ longitude or $9°$ latitude per day.

- Haarsma et al. (1993): Assessed a $\sim$ 300 km dataset. Search on local minimum PSL and require 850mb relative vorticity $> 3.5 \times 10^{-5}$s^{-1}, and temperature anomaly at 250mb $\Delta T250 > 0.5K$, at 500mb $\Delta T500 > -0.5K$, and $\Delta T250 - \Delta T850 > -1.0K$, where the anomaly is computed against a $15° \times 15°$ spatial mean around the center of the storm. Cyclones are tracked for a minimum of 3 days.

- Bengtsson et al. (1995, 1996): Assessed a T106 ($\sim$ 125 km) dataset. Search on 850mb relative vorticity $> 3.5 \times 10^{-5}$s^{-1}. Require a 850mb wind maximum > 15 m/s, local SLP minimum, and mean 850mb wind $>$ mean 300mb wind within 7x7 grid points around candidate. Further require temperature anomaly sum $\Delta T700 + \Delta T500 + \Delta T300 > 3$ K and $\Delta T300 > \Delta T850$ where the anomaly is computed against a 7x7 gridpoint average centered on candidate. Cyclones are tracked for a minimum of 1.5 days.

- Krishnamurti et al. (1998): Assessed a T42 ($\sim$ 300 km) climate dataset. Similar to Bengtsson et al. (1995, 1996), except using a 4x4 grid point region for 850mb wind maximum, SLP minimum and temperature mean. Cyclones are tracked for at least 1 day.

- Sugi et al. (2002): Assessed a T106 (~ 125 km) climate dataset. Tracking criteria similar to Bengtsson et al. (1995). Search is conducted for local PSL minima that are at least < 1020 hPa. Cyclones are tracked for a minimum of 2 days.

- Oouchi et al. (2006): Assessed a 20 km dataset using a similar technique to Sugi et al. (2002). PSL at storm center must be 2 hPa lower than mean over 7x7 grid box and require that storm center latitude $< 45°$ with an initial position $< 30°$. Near the storm require relative vorticity at 850 hPa must be $> 3.5 \times 10^{-5} \text{s}^{-1}$, maximum wind speed at 850hPa must be > 15 m/s, and the maximum wind speed at 850hPa is larger than at 300hPa. Further require temperature anomaly sum $\Delta T700 + \Delta T500 + \Delta T300 > 2$ K near storm. Cyclones are tracked for at least 1.5 days.

- Murakami and Sugi (2010b): Assessed four datasets with resolutions from TL95 (~ 180 km) to TL959 (~ 20 km). Cyclone identification similar to Oouchi et al. (2006) with a resolution-dependent relative vorticity criteria.

- Murakami et al. (2012): Assessed four datasets with resolutions from 20 km to 60 km. Cyclone identification similar to Oouchi et al. (2006) with a resolution-dependent relative vorticity and temperature anomaly criteria. Temperature anomalies are computed against a $10° \times 10°$ grid box. Additional filtering is applied in the North Indian Ocean requiring maximum wind speed be within 100-200km of storm center. Tracking incorporates a maximum gap size of 1 (a single time step failure).

- Vitart et al. (1997, 1999); Vitart et al. (2001): Assessed a T42 ($2.8°$, ~ 300 km) dataset. Search on 850mb relative vorticity maxima $> 3.5 \times 10^{-5} \text{s}^{-1}$ with a nearby PSL minimum. Must possess a warm core within $2°$ latitude defined as a local average 500mb to 200mb temperature maximum with a decrease of 0.5 K in all directions within $8°$. Must possess a local maximum in 200mb - 1000mb thickness with a decrease of 50m in all directions within $8°$. When tracking, the minimum distance between storms is 800 km / day, tracks must last at least 2 days and the maximum wind velocity within $8°$ of the storm center must be 17 m/s for at least 2 (not necessarily consecutive) days.

20

– Tsutsui and Kasahara (1996): Assessed a T42 ($2.8°$, ~ 300 km) dataset. Search on minima in 1000mb geopotential height field, with at least an average drop of 20 m among neighboring points, and a further 20 m drop of average among neighboring points from periphery. Require average local 900mb vorticity to be cyclonic, average local 900mb divergence to be negative, average local 500mb vertical velocity to be upward, 200mb minus 1000mb layer thickness maximum among neighbors is greater than any value in periphery, and average local 200mb zonal wind velocity is less than 10 m/s or local points contain at least one point with easterly velocity. Require the latitude $< 40°$, topographic height underlying candidates should be < 400 m, one local point must have a 900mb wind speed of at least 17.2 m/s, and one local point must exceed 100 mm/d over at least one day. Cyclones are tracked for a minimum of 2 days.

Tsutsui (2002): Assessed a T42 ($2.8°$, ~ 300 km) dataset. Search is performed similar to Tsutsui and Kasahara (1996), but with a simplified criteria. PSL is required to be less than the local average minus 2 hPa, and local average must be less than periphery average minus 2hPa. Layer thickness between 200hPa and 700hPa, denoted by Z, must satisfy $Z_0 + \max(Z_{\pm1\Delta}) > 2\max(Z_{\pm2\Delta})$, where $Z_{\pm1\Delta}$ denotes immediate neighbors and $Z_{\pm2\Delta}$ denotes the periphery.

– Walsh (1997); Walsh and Watterson (1997); Walsh and Katzfey (2000): Assessed a 125 km regional climate dataset over Australia. Required 850 hPa relative vorticity $> 2.0 \times 10^{-5}\text{s}^{-1}$, temperature anomaly sum $\Delta T700 + \Delta T500 + \Delta T300 > 0$ K and $\Delta T300 > \Delta T850$, with anomaly computed against the mean over a region 2 grid points north/south and 13 grid points east/west. Further require 10m surface wind > 10 m/s and 850 hPa tangental wind speed > 300 hPa tangental wind speed. Cyclones are tracked for a minimum of 2 days.

– Nguyen and Walsh (2001): Similar to Walsh and Watterson (1997), assessed a 125 km regional dataset over Australia. Vorticity requirement was changed to 850 hPa relative vorticity $> 1.0 \times 10^{-5}\text{s}^{-1}$ with PSL minimum within 250km. Also required mean wind speed in 500km $\times$ 500km region at 850 hPa was larger than at 300 hPa and a mean

tangental wind speed within a radius of $1°$ and $2.5°$ greater than 5 m/s. Cyclones were tracked for a minimum of 1 day, with relaxed criteria after this time.

– Walsh et al. (2004): Assessed a 30 km dataset using a similar tracking strategy to Nguyen and Walsh (2001). The temperature anomaly was computed against a 1200km longitude $\times$ 400km latitude region, and the mean wind speed was computed over a 800km $\times$ 800km region around the storm. Further required that $V10 \geq 17$ m/s near storm.

– McDonald et al. (2005): Assessed a $2.5°$ latitude by $3.75°$ longitude dataset. Search on local maxima of 850hPa relative vorticity with magnitude greater than $5 \times 10^{-5}\text{s}^{-1}$ with initial latitude $< 30°$. Temperature anomalies must satisfy $\Delta T300 > 0$ along the track, $\Delta T300 > 0.5$ K for any two points along the track and $\Delta T300 > \Delta T850$ for any two points along the track, where the anomaly was computed against a $15° \times 15°$ mean. Cyclones are tracked for a minimum of 2 days.

– Bengtsson et al. (2007b): Assessed T63, T213 and T319 datasets. Required 850hPa relative vorticity minus 250hPa relative vorticity exceed $6 \times 10^{-5}\text{s}^{-1}$, that 850 hPa relative vorticity $> 6 \times 10^{-5}\text{s}^{-1}$ and that relative vorticity be positive for all levels between 850 hPa and 250 hPa. Only northern hemisphere cyclones were preserved (latitude $< 60°$). Cyclones are tracked for a minimum of 1 day.

– Knutson et al. (2007); Zhao et al. (2009): Assessed a ~ 50 km dataset. Search on absolute 850hPa relative vorticity maxima $> 1.6 \times 10^{-4}\text{s}^{-1}$ within $6° \times 6°$ areas with a local minimum in SLP within $2°$ of the detection. Further require a maximum in average temperature between 300 hPa and 500 hPa within $2°$ of the detection that is 1K warmer than the local mean. Cyclones are tracked for a minimum of 3 days, with a maximum search radius of 400km per 6 hours, and requiring that at least 3 days have a maximum surface wind speed greater than 17 m/s.

– Strachan et al. (2013): Assessed datasets from ~ 60km to ~ 270 km. At T63 resolution, required that 850hPa relative vorticity attain a threshold of $> 6 \times 10^{-5}\text{s}^{-1}$, and relative

22

vorticity at 500 hPa and 200 hPa be positive. Further required that the relative vorticity difference between 850 hPa and 200 hPa $> 6 \times 10^{-5} s^{-1}$. Cyclones are tracked for a minimum of 1 day.

– Zarzycki and Jablonowski (2014): Assessed a ~ 28km dataset. Search on SLP minimum with latitude $< 45°$ and absolute 850hPa relative vorticity maxima $> 1.0 \times 10^{-4} s^{-1}$ within $4°$. Require a local maximum 500-200 hPa average temperature within $2°$ of storm center which decreases by at least 0.8 K at a radius of $5°$ in all directions. Cyclones are tracked for a minimum of 2 days, with a maximum search radius of 400km per 6 hours, and requiring that at least 2 days have a maximum surface wind speed is greater than 17 m/s within $4°$ of the candidate.

A3 A (Short) Review of Tropical Easterly Wave Tracking Algorithms

Tropical easterly waves are featured more sparsely within the literature, but are nonetheless an important pointwise feature in climate datasets. Pointwise tracking is complementary to statistical techniques which typically examine the variability, for instance, in the African Easterly Jet (AEJ) (*i.e.*, Ceron and Gueremy (1999)). The first manual study that identified and tracked African easterly wave was performed by Reed et al. (1988) using positive relative vorticity anomalies. This strategy was also applied by Thorncroft and Hodges (2001), Hodges et al. (2003) and Serra et al. (2010). Other studies have used curvature vorticity anomalies Belanger et al. (2014); Bain et al. (2014), streamfunction (Berry et al., 2007) and meridional velocity anomalies Skinner and Diffenbaugh (2014).

A4 Software Documentation: DetectCyclonesUnstructured

```
Usage: DetectCyclonesUnstructured <parameter list>
Parameters:
  --in_data <string> [""]
  --in_data_list <string> [""]
  --in_connect <string> [""]
```

23

```
     --out <string> [""]
     --out_file_list <string> [""]
     --searchbymin <string> [""] (default PSL)
     --searchbymax <string> [""]
5    --minlon <double> [0.000000] (degrees)
     --maxlon <double> [0.000000] (degrees)
     --minlat <double> [0.000000] (degrees)
     --maxlat <double> [0.000000] (degrees)
     --topofile <string> [""]
10   --maxtopoht <double> [0.000000] (m)
     --mergedist <double> [0.000000] (degrees)
     --closedcontourcmd <string> [""] [var,delta,dist,minmaxdist;...]
     --noclosedcontourcmd <string> [""] [var,delta,dist,minmaxdist;...]
     --thresholdcmd <string> [""] [var,op,value,dist;...]
15   --outputcmd <string> [""] [var,op,dist;...]
     --timestride <integer> [1]
     --regional <bool> [false]
     --out_header <bool> [false]
     --verbosity <integer> [0]

20   --in_data <string>
         A list of input datafiles in NetCDF format, separated by semi-colons.

     --in_data_list <string>
         A file containing the --in_data argument for a sequence of processing operations (one
         per line).

25   --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.
```

24

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

`--out_file_list <string>`
A file containing the `--out` argument for a sequence of processing operations (one per line).

`--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.

`--minlon <double>`
The minimum longitude for candidate points.

`--maxlon <double>`
The maximum longitude for candidate points.

`--minlat <double>`
The minimum latitude for candidate points.

`--maxlat <double>`
The maximum latitude for candidate points.

`--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.

`--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,delta,dist,minmaxdist`. These arguments are as follows.

> `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 contour.
>
> `minmaxdist <double>` The distance away from the candidate to search for the minima/maxima. 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.

`--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.

> `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 a point that satisfies the threshold (in degrees).

26

--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 assume 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).

A4.1 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

27

```
          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 operators are also supported, including

`_ABS(<variable>)` Absolute value of a variable,

`_AVG(<variable>, <variable>)` Pointwise average of variables,

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

`_F()` Coriolis parameter,

`_MEAN(<variable>, <distance>)` Spatial mean over a given radius,

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

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

For instance, the following are valid examples of `<variable>` type,

`_MEAN(PSL,2.0)`, `_VECMAG(U850, V850)` and `_DIFF(U(3),U(5))`.

A4.2 MPI Support

The `DetectCyclonesUnstructured` executable supports parallelization via MPI when the `--in_data_list` argument is specified. When enabled, the parallelization procedure simply distributes the processing operations evenly among available MPI threads.

A5 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).
```

29

`--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 satisfy this criteria.

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

Algorithm 1. Compute the spatial mean value of a field G over a region of radius dist using graph search on an unstructured grid.

```
field F = mean(field G, dist)
  for each node p
    total_area = 0
    F[p] = 0
    visited = []
    tovisit = [p]
    while visited is not empty
      q = remove node from tovisit
      add q to visited
      F[p] = F[p] + G[q] * area[q]
      total_area = total_area + area[q]
      for each neighbor s of q
        if (gcd(p,s) < dist) and (s is not in visited) then
          add s to tovisit
    F[p] = F[p] / total_area
```

Algorithm 2. Locate the set of all nodes P that are local minima for a field G (for instance, SLP) defined on an unstructured grid. The procedure for locating maxima is analogous.

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

Algorithm 3. Given a field G defined on an unstructured grid and a set of candidate points P, remove candidate minima that are within a distance dist of a more extreme minimum, and return the new candidate set Q.

```
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 = kd_tree_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
```

Algorithm 4. Find the node pmax containing the maximal value of the field G within a distance maxdist of the node p. An analogous procedure find_min_near is provided for locating nodes containing minimal values of the field.

```
node pmax = find_max_near(node p, field G, maxdist)
  set visited = {}
  set tovisit = {p}
  pmax = p
  while tovisit is not empty
    q = remove node 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
```

Algorithm 5. Determine if there is a closed contour in field G of magnitude thresh around the point p0, defined by p0 = find_max_near(p, G, maxdist), within distance dist. That is, along all paths away from p0, the field G must drop by at least thresh within distance dist. The closed contour criteria is depicted in Figure 2. An analogous procedure is defined for closed contours around minima.

```
closed_contour_max(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
```

Algorithm 6. Determine if a candidate node p satisfies the requirement that there exists another node p0 within distance dist of p with G[p] > thresh.

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

Algorithm 7. Determine all feature paths S, given array of candidate nodes $P[1..T]$ and maximum great-circle distance between nodes at subsequent time levels dist.

```
path set S = stitch_nodes(set array P[1..T], dist, maxgap)
  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
      gap = 0
      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 if (gap < maxgap) then
          gap = gap + 1
        else
          break
      add s into S
```
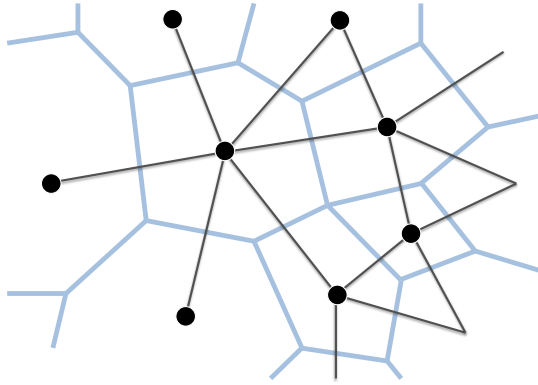
Fig. 1. An example adjacency graph describing an unstructured grid (blue lines), where nodes are co-located with volume centerpoint locations (solid circles) and edges connect adjacent volumes.
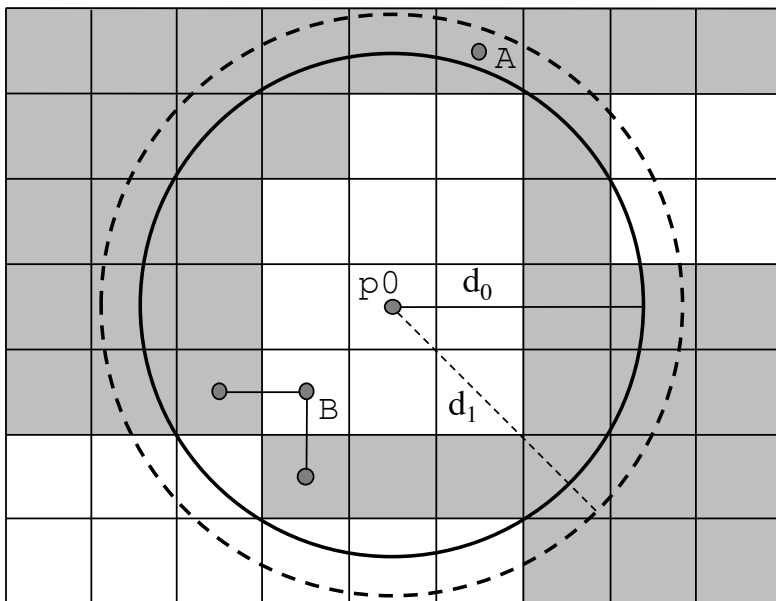
Fig. 2. An illustration of the closed contour criteria. Nodes shaded in white (gray) satisfy (do not satisfy) the threshold of the field value at p0. Since only edge-neighbors are included, B constitutes a boundary to the interior of the closed contour. Because A lays outside the solid circle, the contour with distance d_0 is not a closed contour, whereas the dashed contour with distance d_1 does satisfy the closed contour criteria.
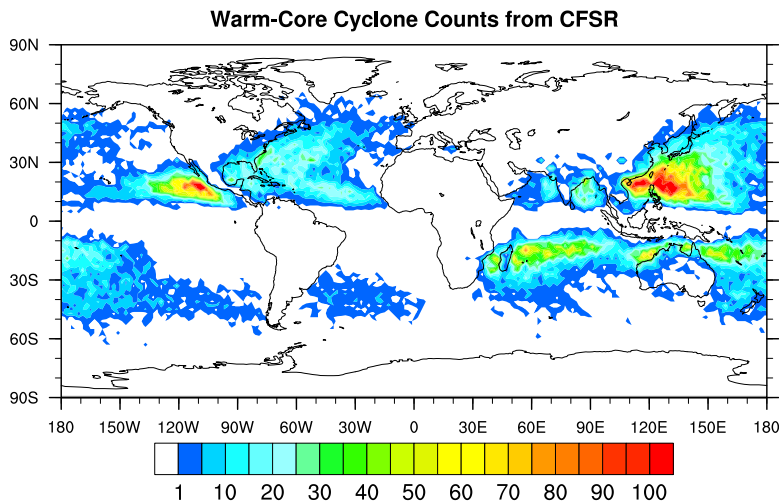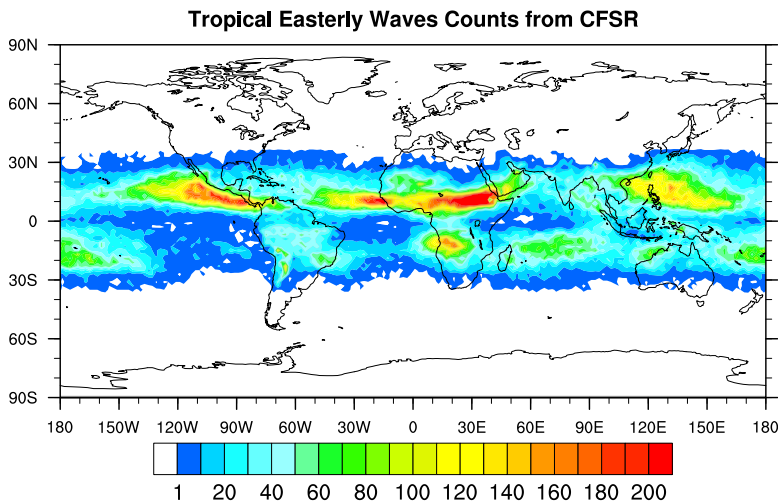
Fig. 3. Tropical cyclone counts within each $2° \times 2°$ grid cell over the period 1979-2010 obtained using the procedure described in section 3.1.
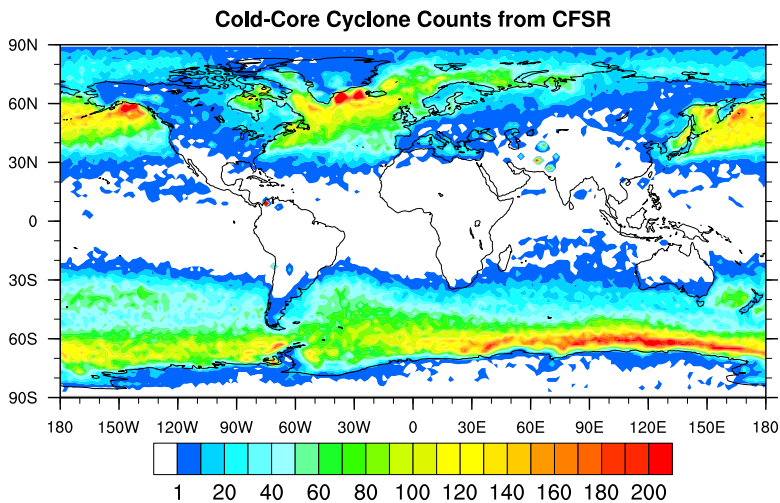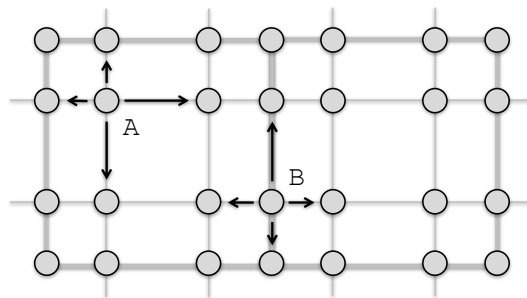
Fig. 4. Extratropical cyclone counts within each $2° \times 2°$ grid cell over the period 1979-2010 obtained using the procedure described in section 3.2.

Fig. 5. Tropical easterly wave counts within each $2° \times 2°$ grid cell over the period 1979-2010 obtained using the procedure described in section 3.3.

Fig. 6. An illustration of how connectivity is defined in this work for nodes on a spectral element mesh. Arrows indicate connectivity for nodes A and B.
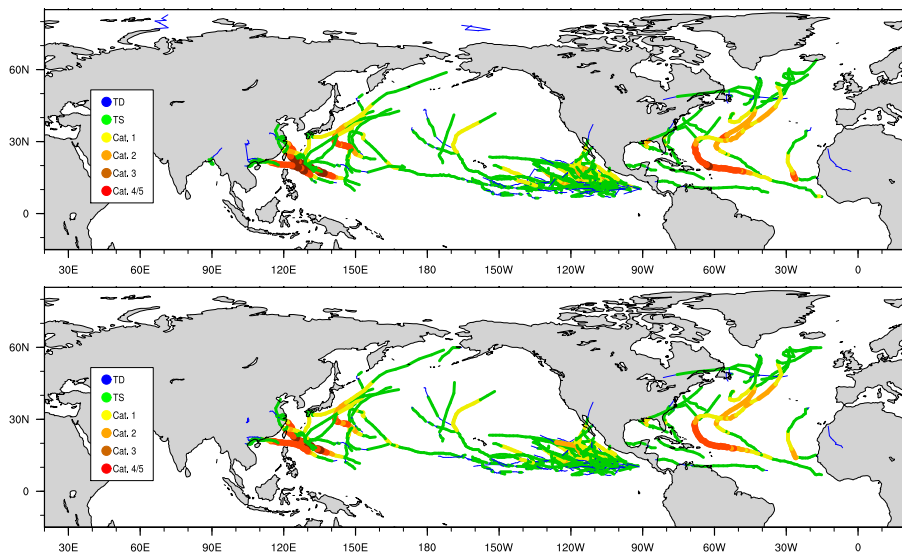
Fig. 7. Tropical cyclone trajectories and associated intensities as obtained from the simulation of a single hurricane season in CESM 3.4 using (top) native spectral-element grid data and (bottom) data regridded to a regular latitude-longitude grid with 0.25° grid spacing.
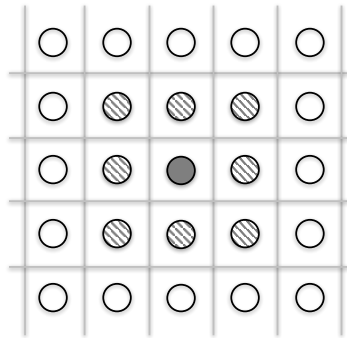
Fig. 8. The local neighborhood of a central node (shaded) typically refers to the surrounding 8 nodes (diagonal hatching). The periphery (used by Tsutsui and Kasahara (1996)), refers to the set of nodes that surround the local neighborhood (unshaded nodes).

References

Akyildiz, V.: Systematic errors in the behaviour of cyclones in the ECMWF operational models, Tellus 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.

Bain, C., Williams, K., Milton, S., and Heming, J.: Objective tracking of African easterly waves in Met Office models, Quarterly Journal of the Royal Meteorological Society, 140, 47–57, 2014.

Belanger, J. I., Jelinek, M. T., and Curry, J. A.: African Easterly Wave Climatology, Version 1, doi:10.7289/V5ZC80SX, nOAA National Centers for Environmental Information. Accessed July 21, 2016, 2014.

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.

Benestad, R. and Chen, D.: The use of a calculus-based cyclone identification method for generating storm statistics, Tellus A, 58, 473–486, 2006.

Bengtsson, L., Botzet, M., and Esch, M.: Hurricane-type vortices in a general circulation model, Tellus A, 47, 175–196, 1995.

Bengtsson, L., Botzet, M., and Esch, M.: Will greenhouse gas-induced warming over the next 50 years lead to higher frequency and greater intensity of hurricanes?, Tellus A, 48, 57–73, 1996.

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, 2007a.

Bengtsson, L., Hodges, K. I., Esch, M., Keenlyside, N., Kornblueh, L., LUO, J.-J., and Yamagata, T.: How may tropical cyclones change in a warmer climate?, Tellus a, 59, 539–561, 2007b.

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

Berry, G., Thorncroft, C., and Hewson, T.: African easterly waves during 2004-Analysis using objective techniques, Monthly weather review, 135, 1251–1267, 2007.

Blender, R., Fraedrich, K., and Lunkeit, F.: Identification of cyclone-track regimes in the North Atlantic, Quarterly Journal of the Royal Meteorological Society, 123, 727–741, 1997.

Broccoli, A. and Manabe, S.: Can existing climate models be used to study anthropogenic changes in tropical cyclone climate?, Geophys. Res. Lett., 17, 1917–1920, 1990.

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.

Ceron, J. and Gueremy, J.: Validation of the space-time variability of African easterly waves simulated by the CNRM GCM, Journal of climate, 12, 2831–2855, 1999.

Dean, J. and Ghemawat, S.: MapReduce: simplified data processing on large clusters, Communications of the ACM, 51, 107–113, 2008.

Haarsma, R. J., Mitchell, J. F., and Senior, C.: Tropical disturbances in a GCM, Climate Dynamics, 8, 247–257, 1993.

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

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

Hodges, K. I.: TRACK, http://www.nerc-essc.ac.uk/~kih/TRACK/Track.html, accessed July 8, 2016, 2015.

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.

Knutson, T. R., Sirutis, J. J., Garner, S. T., Held, I. M., and Tuleya, R. E.: Simulation of the recent multidecadal increase of Atlantic hurricane activity using an 18-km-grid regional model, Bulletin of the American Meteorological Society, 88, 1549, 2007.

Krishnamurti, T., CORREA-TORRES, R., Latif, M., and Daughenbaugh, G.: The impact of current and possibly future sea surface temperature anomalies on the frequency of Atlantic hurricanes, Tellus A, 50, 186–210, 1998.

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.

Lionello, P., Dalan, F., and Elvini, E.: Cyclones in the Mediterranean region: the present and the doubled CO_2 climate scenarios, Climate Research, 22, 147–159, 2002.

McDonald, R. E., Bleaken, D. G., Cresswell, D. R., Pope, V. D., and Senior, C. A.: Tropical storms: representation and diagnosis in climate models and the impacts of climate change, Climate Dynamics, 25, 19–36, 2005.

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

Murakami, H. and Sugi, M.: Effect of model resolution on tropical cyclone climate projections, Sola, 6, 73–76, 2010b.

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.

47

Murakami, H., Wang, Y., Yoshimura, H., Mizuta, R., Sugi, M., Shindo, E., Adachi, Y., Yukimoto, S., Hosaka, M., Kusunoki, S., et al.: Future changes in tropical cyclone activity projected by the new high-resolution MRI-AGCM*, Journal of Climate, 25, 3237–3260, 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.

Neu, U., Akperov, M. G., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R., Cocozza, A., Dacre, H. F., Feng, Y., Fraedrich, K., et al.: IMILAST: a community effort to intercompare extratropical cyclone detection and tracking algorithms, Bulletin of the American Meteorological Society, 94, 529–547, 2013.

Nguyen, K. and Walsh, K.: Interannual, decadal, and transient greenhouse simulation of tropical cyclone-like vortices in a regional climate model of the South Pacific, Journal of Climate, 14, 3043–3054, 2001.

Oouchi, K., Yoshimura, J., Yoshimura, H., Mizuta, R., Kusunoki, S., and Noda, A.: Tropical cyclone climatology in a global-warming climate as simulated in a 20 km-mesh global atmospheric model: Frequency and wind intensity analyses, ????. ? 2 ?, 84, 259–276, 2006.

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

Pinto, J. G., Spangehl, T., Ulbrich, U., and Speth, P.: Sensitivities of a cyclone detection and tracking algorithm: individual tracks and climatology, Meteorologische Zeitschrift, 14, 823–838, 2005.

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.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-Y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Delst, P. V., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., Dool, H. V. D., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins,

W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP climate forecast system reanalysis, Bulletin of the American Meteorological Society, 91, 1015–1057, 2010.

Serra, Y. L., Kiladis, G. N., and Hodges, K. I.: Tracking and mean structure of easterly waves over the Intra-Americas Sea, Journal of Climate, 23, 4823–4840, 2010.

Serreze, M., Box, J., Barry, R., and Walsh, J.: Characteristics of Arctic synoptic activity, 1952–1989, Meteorology and Atmospheric Physics, 51, 147–164, 1993.

Serreze, M. C.: Climatological aspects of cyclone development and decay in the Arctic, Atmosphere-Ocean, 33, 1–23, 1995.

Simmonds, I., Burke, C., and Keay, K.: Arctic climate change as manifest in cyclone behavior, Journal of Climate, 21, 5777–5796, 2008.

Sinclair, M. R.: An objective cyclone climatology for the Southern Hemisphere, Monthly Weather Review, 122, 2239–2256, 1994.

Sinclair, M. R.: Objective identification of cyclones and their circulation intensity, and climatology, Weather and forecasting, 12, 595–612, 1997.

Skinner, C. B. and Diffenbaugh, N. S.: Projected changes in African easterly wave intensity and track in response to greenhouse forcing, Proceedings of the National Academy of Sciences, 111, 6882–6887, 2014.

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, Journal of Climate, 26, 133–152, 2013.

Sugi, M., Noda, A., and Sato, N.: Influence of the global warming on tropical cyclone climatology: An experiment with the JMA global model., ????. ? 2 ?, 80, 249–272, 2002.

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.

Thorncroft, C. and Hodges, K.: African easterly wave variability and its relationship to Atlantic tropical cyclone activity, Journal of Climate, 14, 1166–1179, 2001.

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

Tsutsui, J.: Implications of anthropogenic climate change for tropical cyclone activity: A case study with the NCAR CCM2., ????. ? 2 ?, 80, 45–65, 2002.

49

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.

Ullrich, P. A. and Taylor, M. A.: Arbitrary-Order Conservative and Consistent Remapping and a Theory of Linear Maps: Part I, Monthly Weather Review, 143, 2419–2440, doi:10.1175/MWR-D-14-00343.1, http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-14-00343.1, 2015.

Ullrich, P. A., Devendran, D., and Johansen, H.: Arbitrary-Order Conservative and Consistent Remapping and a Theory of Linear Maps, Part 2, Monthly Weather Review, In Press, doi:10.1175/MWR-D-15-0301.1, http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-15-0301.1, 2016.

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.

Vitart, F., Anderson, J. L., and Stern, W. F.: Impact of large-scale circulation on tropical storm frequency, intensity, and location, simulated by an ensemble of GCM integrations, Journal of Climate, 12, 3237–3254, 1999.

Vitart, F., Anderson, J. L., Sirutis, J., and Tuleya, R. E.: Sensitivity of tropical storms simulated by a general circulation model to changes in cumulus parametrization, Quart. J. Roy. Meteor. Soc., 127, 25–51, doi:10.1002/qj.49712757103, 2001.

Walsh, K.: Objective detection of tropical cyclones in high-resolution analyses, Monthly Weather Review, 125, 1767–1779, 1997.

Walsh, K. and Watterson, I. G.: Tropical cyclone-like vortices in a limited area model: comparison with observed climatology, Journal of climate, 10, 2240–2259, 1997.

Walsh, K., Nguyen, K.-C., and McGregor, J.: Fine-resolution regional climate model simulations of the impact of climate change on tropical cyclones near Australia, Climate Dynamics, 22, 47–56, 2004.

Walsh, K., Fiorino, M., Landsea, C., and McInnes, K.: Objectively determined resolution-dependent threshold criteria for the detection of tropical cyclones in climate models and reanalyses, Journal of climate, 20, 2307–2314, 2007.

Walsh, K. J. and Katzfey, J. J.: The impact of climate change on the poleward movement of tropical cyclone-like vortices in a regional climate model, Journal of Climate, 13, 1116–1132, 2000.

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.

Wu, G. and Lau, N.-C.: A GCM simulation of the relationship between tropical-storm formation and ENSO, Monthly Weather Review, 120, 958–977, 1992.

Zarzycki, C. M. and Jablonowski, C.: A multidecadal simulation of Atlantic tropical cyclones using a variable-resolution global atmospheric general circulation model, Journal of Advances in Modeling Earth Systems, 6, 805–828, doi:10.1002/2014MS000352, 2014.

Zarzycki, C. M., Jablonowski, C., and Taylor, M. A.: Using Variable-Resolution Meshes to Model Tropical Cyclones in the Community Atmosphere Model, Monthly Weather Review, 142, 1221–1239, 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, Journal of Climate, 22, 6653–6678, 2009.

Zolina, O. and Gulev, S. K.: Improving the accuracy of mapping cyclone numbers and frequencies, Monthly weather review, 130, 748–759, 2002.