ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

IAC Institute for Atmospheric and Climate Science

Jan Engelmann

Tropical Cyclone Tracking with varying parameter thresholds

Semester Thesis

Institute for Atmospheric and Climate Science - Atmospheric Physics
Swiss Federal Institute of Technology (ETH) Zurich

Supervision

Bernhard Enz
Prof. Dr. Ulrike Lohmann

November 2020

Contents

Abstract

Tropical Cyclone (TC) tracking in simulation data requires parameter thresholds that specify the expected intensity of these characterising variables. This results in assumptions for a specific climate model that may or may not lead to a successful tracking of TCs. With the purpose of being able to run the algorithm with different parameter assumptions, an existing algorithm was optimised and a 60-fold speedup reached. This enabled experimentation with numerous different threshold combinations. It was found that correctly adjusting the warm core criterion is of central importance since it balances the unwanted tracking of noise with the desired early discovery of tropical cyclones. Furthermore, including a requirement in regards to the minimum sea surface temperature during TC genesis made the algorithm more robust. The analysis of the vorticity criterion suggested further investigation of its effectiveness. Finally, matching tracks across threshold combinations significantly improved the understanding of parameter interplay and the tracking of TCs across their lifetime.

Chapter 1

Introduction

Tropical cyclones (TCs), also known as hurricanes or typhoons, are storms of extreme nature in many regards. Not only are they the most deadly and expensive natural catastrophes in the United States, but also their physics is quite challenging with many open questions remaining[1]. However, writing them off merely as a complex threat to civilisation would be over-simplistic. Research has shown that TCs play a crucial role in the global heat balance and moisture circulation[2][3].

1.1 Impact on society

While tropical cyclones form and intensify above the ocean, they have the largest impact on society during landfall. The damage happens due to a combination of strong winds and catastrophic storm surges. On average, hurricanes inflict normalised damages of about $10-billion/year in the United States[4]. A single strong storm can cause thousands of deaths. Hurricane Katrina in 2005, for example took the lives of over 1200 people [5]. Due to these enormous implications for society and opportunities to save lives and money, active research is happening on tropical cyclone impact reduction. With an unsure impact of climate change on tropical cyclone frequency but an expected increase in their intensity and a trend of urbanisation on the American east-coast, improving the understanding of tropical cyclones is of great importance.

1.2 Underlying Physics

Tropical cyclones can be classified using the Saffir-Simpson wind scale. It is defined by the maximum wind observed in the TC. Due to their characteristic structure, the maximum wind usually occurs at the eyewall. Using the maximum wind as a means to categorise TCs is motivated by the strong correlation between wind speeds and the inflicted damage[6]. The exact categorisation can be seen in Table 1.1 and the occurrence frequency of the different categories is displayed in Fig. 1.1. Tropical depressions (TD) are tropical low pressure systems that have a windspeed of less than 17 m/s. When they intensify above this threshold they are called tropical storms (TS). At this stage, they are assigned a name for easier communication between the different meteorological institutes and with the public. These tropical systems may already exhibit a physical structure comparable to that of a TC as described below but do not necessarily have to. Once a tropical storm intensifies to a maximum sustained wind speed beyond 33 m/s it is called a TC and a category is assigned.

Tropical cyclones rotate around a pressure minimum (the eye) which is enclosed by the eye wall. While the air is almost still in the eye, the maximum wind speeds is measured in the eye wall. The rotation around the center happens in different bands of updraft below the cloud that are alternated with rain bands. Finally, the warm core results in an anti-cyclonic outflow at the top of the storm. This structure is depicted in Fig. 1.2 and together with the requirement of thermal wind balance leads to the characteristic warm core structure. In order for a storm to develop, a

Tropical cyclones	
category	wind speed [m/s]
1	33–42
2	43–49
3	50–58
4	59–70
5	≥ 70

Other tropical low pressure systems	
name, category	wind speed [m/s]
tropical depression, -1	≤ 17
tropical storm, 0	18–32

Table 1.1: Simpson scale defined by 1-minute maximum sustained winds [6]. The category number of the other tropical low pressure systems was assigned by the author and will be used in the results section 3

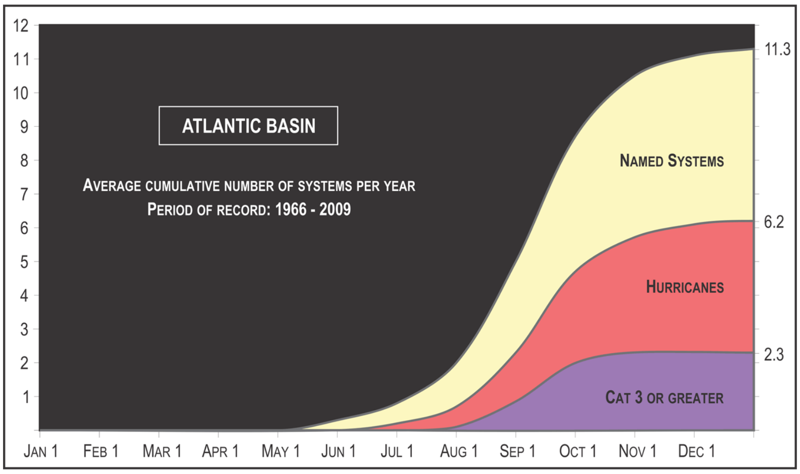


Figure 1.1: Average cumulative number of storms in the Atlantic. Named systems are mostly tropical storms but can also be tropical depressions.[7]

number of TC genesis criteria have to be met. Not all of them have to be satisfied but they do offer a good indicator for the probability of storm formation. The criteria as summarised in [9] are as follows:

- sea surface temperature (SST) above 26.5°C to at least a depth of 50m
- sufficiently moist mid-troposphere for deep convection
- Appreciable moisture flux at the ocean-air interface to sustain a conditionally unstable thermodynamic environment [10]
- A distance of at least 5° from the equator, so that the Coriolis effect is strong enough to initiate the cyclone's rotation. [11]
- A pre-existing weather disturbance with sufficient vorticity and convergence, e.g. a tropical easterly wave.
- Low vertical wind shear between the surface and the upper troposphere

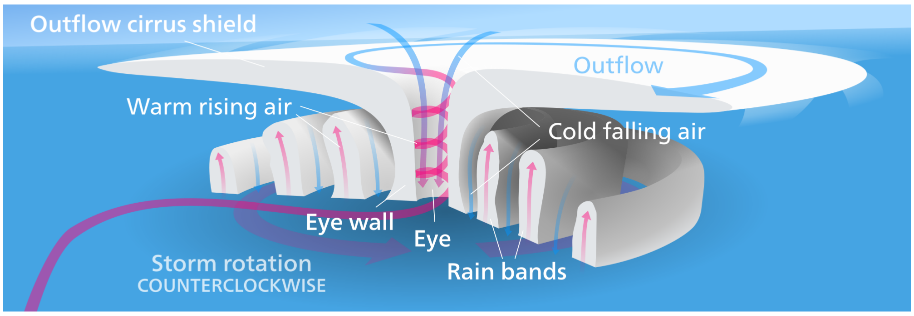


Figure 1.2: Structure of a tropical cyclone in the Northern hemisphere [8]

1.3 Previous work on TC Tracking

A large body of work exists on tracking meteorological phenomena. In this context tracking can be understood as following the same physical object in time.
For instance an effort was conducted to compare different extratropical cyclone tracking algorithms in [12]. While extratropical cyclones differ significantly from tropical cyclones for example in regards to their occurrence frequency and physical structure, there are some common themes that apply to both tracking endeavours. The study found that the 15 compared algorithms disagree strongly on the total number of cyclones and the detection of weak cyclones, and agree best for stronger cyclones. It will be shown in Sec. 3 of this report that these findings also apply to the tracking of TCs.
Another work compared the predictions of two TC tracking schemes for several idealised climate simulations [13]. The authors find a strong dependence of the results on the used parameter thresholds. They furthermore distinguish between algorithms using experimentally motivated parameter thresholds and those using deviations from the surrounding mean. They find that deviation conditions are more universal in respect to the climate model that they are applied to. The algorithm used in our report uses a mix of both types.
The algorithm of this paper was inspired by the one in [14]. In the paper the authors discuss tracking of TCs in ERA-40 reanalysis data. The algorithm was adjusted for use with ICON data in the following ways: The sea level pressure (SLP) minimum criterion and vorticity threshold are retained. The original algorithm requires SLP minima to be 250 km apart and to exhibit a minimum vorticity of $5\times10^{-5}\, s^{-1}$ (Equivalent to **slpdis** and **vormin** from Sec. 1.3). Furthermore, the vertical wind shear criterion is disregarded because TCs can also exist despite wind shear and the more definitive warm core criterion could be implemented to to the higher resolution of the ICON simulation. The minimum required lifetime is relaxed from 36 to 18 hours in the interest of tracking shorter-lived storms. Finally no special criterion over land is used but a minimum SST of 24 °C required.

Chapter 2

Data and Methods

The purpose of this project was to evaluate the parameter thresholds for tracking TCs in ICON simulations. In the following section, the settings used for generating the simulation data and the algorithm itself will be explained.

2.1 Simulation Data

ICON Model

The Icosahedral Nonhydrostatic Model (ICON) has been developed by the German Weather Service (DWD), the Max Planck Institute for Meteorology and several partner institutes[15]. As the name implies, the grid is in a first step generated by mapping the earths surface to an Icosahedron (platonic solid with 20 equilateral triangles as faces). The faces are then split up into smaller triangles in order to achieve the desired resolution. The model delivers all typical meteorological quantities on this grid. It uses the fully compressible and non-hydrostatic version of the Euler equations for the fundamental transport processes. Physical processes which cannot be resolved with the grid size are then parametrised by using complex functions that take into account the grid box mean values of the model variables. Physical processes which need to be treated in this way include solar and thermal radiation, cloud microphysics and turbulent transfer above the earth's surface[16].

Analysed simulation output

The North Atlantic basin with its surrounding area (between 120 °W–15 °W and 0 °N–70 °N) was simulated from the beginning of May until the beginning of December. The output data are saved every six hours. Two different kinds of time-lagged ensembles were used. Both used ERA5 reanalysis data from the year 2013 for the initial weather state and the 6-hourly boundary conditions. The reference ensemble (ref) used the snapshot of the current time step as boundary conditions. The second ensemble (rm) used monthly rolling mean boundary conditions based on the average of the same reanalysis data during the timespan 15 days before and after the current timestep. The members within each of these two ensembles were created by varying the initial weather state of the simulation. Namely each of the first 10 days of May were used as initial conditions for a separate simulation run. The output data from each run is mapped from the icosahedral to a Latitude-Longitude grid to facilitate distance calculations. Finally, this results in 20 separate simulation runs that can be analyzed for TCs.

2.2 Algorithm

An existing algorithm implemented by Bernhard Enz and inspired by [14] was improved in regards to runtime, robustness and readability. The reduced runtime was important to be able to run the

algorithm on the same simulation data but with different threshold values that decide whether a TC was detected or not. By comparing the different results, reasonable thresholds and the importance of the different criteria were determined.
The algorithm consists of two steps. In the first step, the TC candidates are found for each time-step without taking into account if a storm already existed previously. Afterwards all entries are analysed and those that qualify as nearest neighbours temporally and spatially are connected to TC tracks. Both steps will be explained in the following sections.

2.2.1 TC candidate search

TC candidates are found by searching the simulated domain using several criteria. They are summarised in Tab. 2.1 and will be explained in detail in the following sections. Before diving into

criterion	**parameters**
sea level pressure minimum	slpdis
minimal vorticity threshold	vormin
warm core criterion	temdif, temdis

Table 2.1: Each criterion can be adjusted by changing its characterising parameter

the functioning of the algorithm, a quick description of the steps taken to speed up the existing algorithm will follow.

Speeding up the candidate search

The speed-up was achieved by removing almost all Python loops from the algorithm. It performs operations like minimum-finding, comparison of all array elements with a threshold value, calculating means, and masking parts of an array. All of these operations are efficiently implemented in NumPy [17] and its scientific computing ecosystem. Since most of NumPy internal functions are implemented in the C/C++ programming language they make better use of the computing hardware and therefore run much faster than native Python code which needs to be interpreted while it is being run. The use of highly optimised low level code is called vectorisation. Furthermore, the Python multiprocessing library was used to distribute the processing of different simulation runs to different processes.
Finally, the output data format was changed to Pandas dataframes which come with rich statistical functionality for quick analysis of the resulting tracks.
The fundamental remodelling and speed-up of the code amounts to the largest part of the time invested in this research project.

Sea Level Pressure Minimum

As outlined in Sec. 1.2, TCs are low pressure systems. In fact some of the lowest pressures on earth were measured inside the eyes of TCs. Therefore the first step to finding TC candidates for a specific time-step is to locate the sea level pressure minima. Here a hundredfold speed-up was achieved by replacing the previous manual minimum-finding algorithm with a vectorised version from the image processing library scikit-image [18]. This function finds the local minima and requires them to be a certain distance apart. The stronger minimum is kept if two candidates are within the specified distance **slpdis**. The resulting minima are then further analysed.

Minimal Vorticity Threshold

If the vorticity at a pressure minimum is below the minimum threshold **vormin** (see Tab. 2.2), it is discarded as a potential TC candidate.

Warm Core Criterion

The last qualifying characteristic is the warm core structure. The temperature at the height of 300 hPa in the pressure minimum is compared to the average temperature of the surrounding area at the same height. The form and strength of this warm core requirement can be adjusted by using the parameters **temdif** and **temdis** as specified in Tab. 2.2. The former corresponds to the required temperature difference between the center of the storm and its environment. Only storms that have a warm core that is at least **temdif** degrees warmer than the environment, are kept as valid TC candidates. The latter specifies the side-length of the square with the storm in the middle which is defined as the environment.

Saving of TC information

At the end of each time-step the remaining candidates are saved with their corresponding date, time, position, maximum windspeed and sea level pressure in the center.

2.2.2 Creating TC tracks from previously found TC candidates

The TC tracks are created by comparing the positions of the candidates of each time step with the locations of active TCs from the previous time-step. At the beginning of each new time-step, previously active TCs that were interrupted are archived if they lived for at least 18 hours and deleted otherwise. Then for each entry it is checked if a previous TC is within its maximum travel distance. This distance is set to 72 kilometres since a TC with a top speed of 20 m/s can travel this far within 6 hours. If several entries could have been reached by the same active TC, the entry with the lower sea level pressure is added to the track and the other entries are discarded. If no active TC within range of the entry is found, a new active TC is created. Each TC is finally assigned a unique integer for identification.

2.2.3 Varying parameter thresholds

Naturally, the results of the algorithm depend on the choice of the four parameters from Tab. 2.1. In order to answer the research question of this project, the algorithm was run for different combinations of these parameters. Specifically, all combinations of the values in Tab. 2.2 were used.

parameter	unit	values
slpdis	km	50, 100, 150, 200
vormin	1/s	10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}
temdif	K	0.5, 0.75, 1, 1.25, 1.5
temdis	km	50, 100, 200, 300, 400

Table 2.2: List of all parameters which results in 400 (=4*4*5*5) combinations

In summary, the algorithm is run with every parameter set on all 20 simulation runs.

2.3 Tracking Data Analysis

2.3.1 Algorithm output data format

The output of the algorithm is a pandas dataframe with one row per unique TC per timestep. Its columns are described in Tab. 2.3.

information	columns	description
timestamp	date	time and date of the TC entry
position	lon_idx, lat_idx, lon, lat	position of the found TC in longitude/latitude grid indices and coordinates
intensity	maxwind, curr_cat, cat	maximum wind within a 100km distance, snapshot category at the current time, maximum category of the corresponding TC
parameter combination	param_id	id of the parameter combination used for tracking
unique TC identifier	tc_id	unique TC identifier across all parameter combinations and simulation runs
simulation run	mem, exp	variables specifying the member and boundary conditions of the analysed simulation run
genesis sea level temperature	genesis_sst	genesis sea level temperature of the particular TC, identical across all time entries of a TC.

Table 2.3: description of the columns of the output pandas dataframe [19]

Chapter 3

Results

3.1 Filtering out noise

After the tracking and stitching steps are completed, the algorithm results can still be improved by filtering the results. A large improvement was achieved by requiring a certain sea surface temperature at TC genesis.

Sea Surface Temperature Criterion

As outlined in Sec. 1.2, the TCs need warm ocean water as an energy source, when they form. It has been shown that the large majority has an SST over 25.5°C [20]. Therefore it is expected that no reasonable TCs are filtered out when requiring a genesis SST of at least 24°C. However, as can be seen in Fig. 3.1, a large part of the unwanted tracks in the North of the domain are removed.

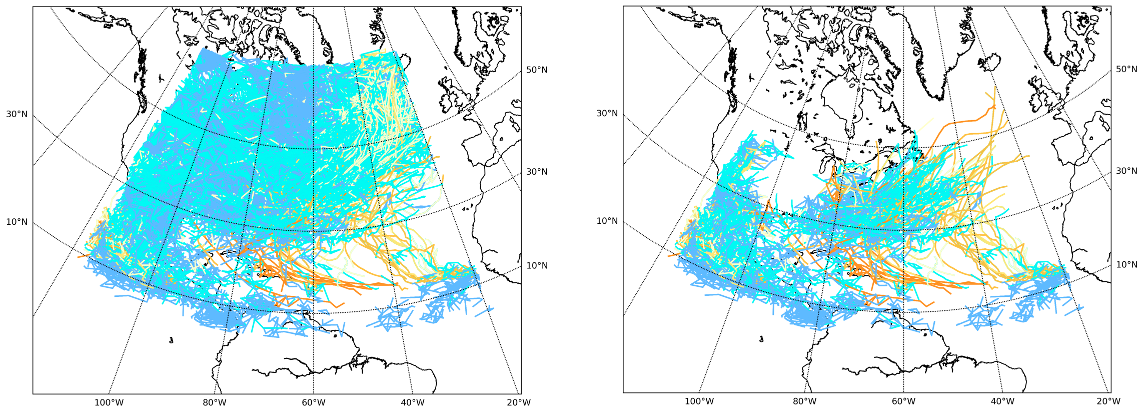


Figure 3.1: Comparison of all tracks without and with the SST criterion on the left and right

Analysing geographically unreasonable tracks

Even with the application of the SST-criterion, unreasonable TC tracks remain. For instance the tracks over Wyoming shown in Fig. 3.2 should not be so frequent. Interestingly these TCs develop over the Great Salt Lake. To determine the parameters responsible for this, the 20 parameter combinations that account for the large majority of these, share the common feature that they all correspond to the same weak warm core criterion. They had a **temdif** of 0.5 °C and a **temdis** of 400 km. Therefore if only a very low temperature difference is required for an area that can be larger than smaller-sized TCs, low pressure systems that do not correspond to tropical cyclones are tracked. While this may not be surprising, it does emphasise the importance of a well-trimmed warm core criterion.

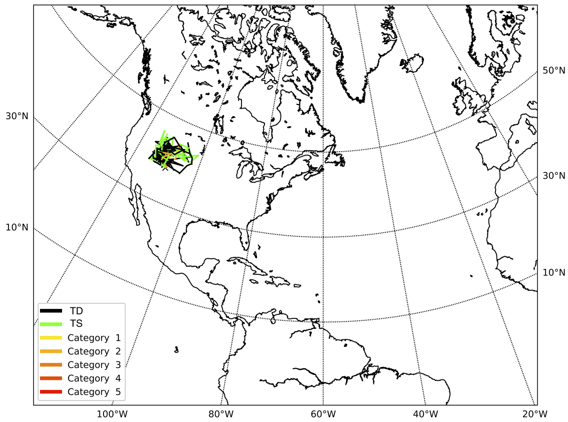


Figure 3.2: Set of unreasonable tracks over Wyoming and the surrounding states.

3.2 Validating Results

In Sec. 3.1, it was found that with a sufficiently strong warm core condition the tracks appear in reasonable areas. Before comparing the algorithm output for different parameter combinations, it still remains to be shown that the produced tracks actually follow TCs and not other low pressure systems. For this purpose, ten different storms were randomly chosen. For these storms the radial, tangential and vertical wind and the sea level pressure were visualised in the azimuthal mean. It was found that all storms qualitatively exhibit the physically expected structure that was described in Sec. 1.2. The resulting plots for a representative storm can be seen in Fig. 3.3.

3.3 Variation of the Warm Core Criterion strength

With the aim of understanding the impact of different warm core criteria strengths, the resulting cyclone distributions for a range of different **temdif** values were compared. As can be seen in Fig. 3.4, a weaker warm core criterion leads to a distribution with more lower-intensity storms. When comparing with the absolute counts, it can be seen that this is the result from weaker storms being tracked for lower **temdif**s. Logically, for TCs of category 2 and upwards, no difference in the counts is observed. Furthermore for **temdif**s larger than 1 K, no tropical depressions which correspond to category -1 are tracked. It can therefore be concluded that the warm core criterion can be used very efficiently to filter out noise if only strong TCs are of interest.

3.4 Comparison of the Warm Core Criterion and the Vorticity Threshold

In order to compare the importance of the warm core criterion with the vorticity threshold, only the TC tracks that were found using the parameter combinations from Tab. 3.1 were analysed. These combinations were chosen because they correspond to different relative strengths of the two criteria.
It was expected that the vorticity threshold should influence the distribution of tracked TCs only if a weak warm core criterion is applied. This was hypothesised because most low pressure systems with strong warm cores should exhibit the necessary minimum vorticity while weak warm core systems might not have a strong enough rotating motion. However, as can be seen in Fig. 3.5, even

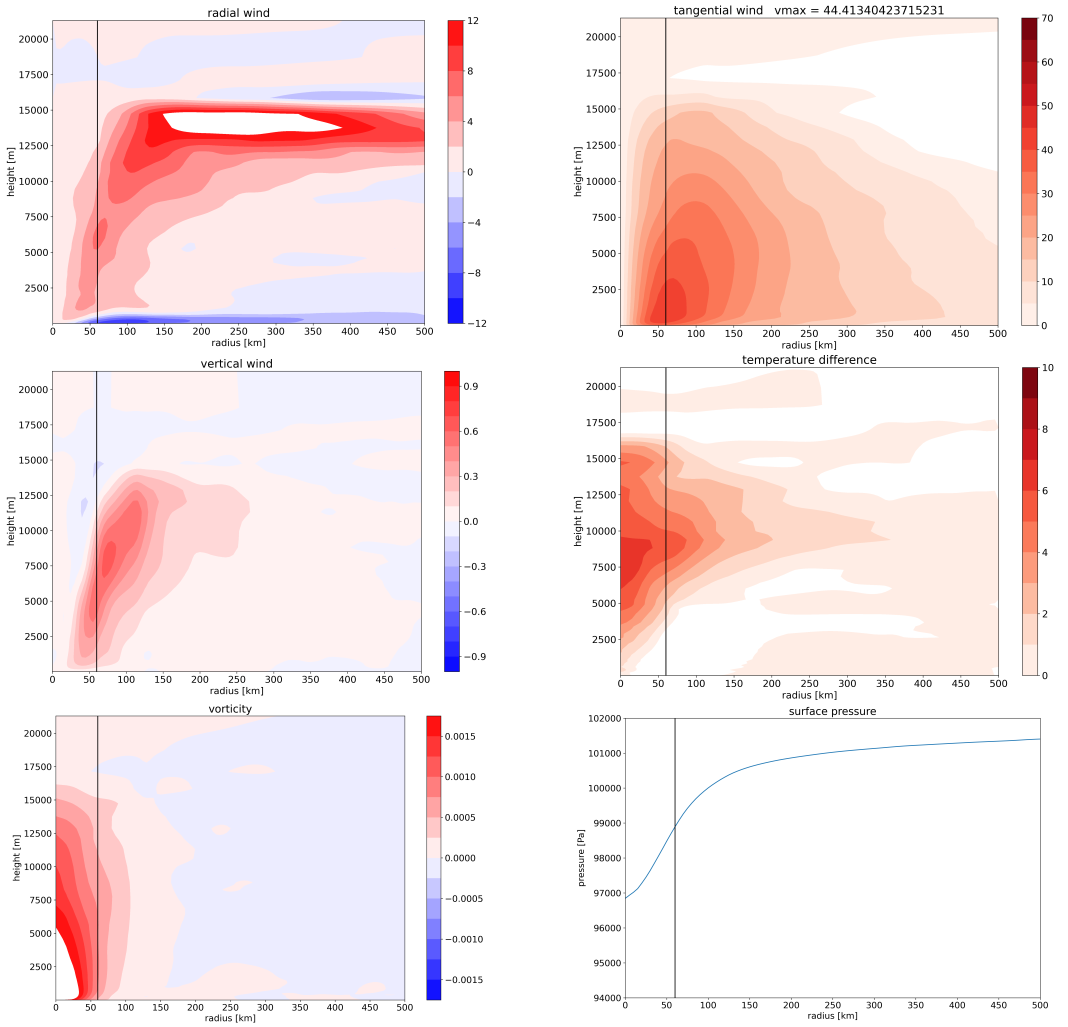


Figure 3.3: Azimuthal mean plots. From left to right and top to bottom: radial wind, tangential wind, vertical wind, temperature difference, vorticity and sea level pressure. The vertical black line marks the radius of maximum wind. The TC (99554) is category 4 and its track can be seen in Fig. 3.8.

parameter	**unit**	**values**
slpdis	m	100000
vormin	1/s	1e-6, 1e-5, 1e-4
temdif	K	0.5, 1 ,1.5
temdis	m	200000

Table 3.1: Parameter combinations used for the comparison

with a very weak warm core criterion does the vorticity threshold not influence the storm intensity distributions. A comparison for different values of **temdif** and an analysis of the impact of the vorticity criterion on the storm lifetime can be found in the Appendix.

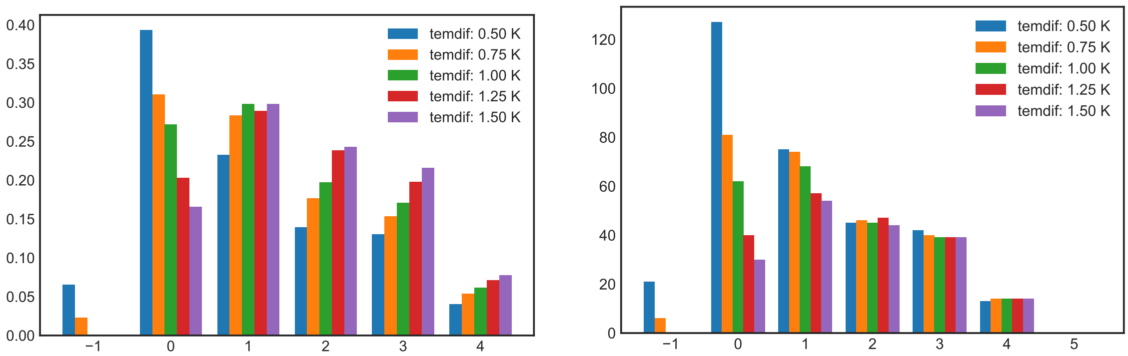


Figure 3.4: Maximum TC category histograms for different **temdif** parameters. Each histogram on the left has unit norm. The right plot shows the absolute counts of TCs. The X-axis describes the TC categories as defined in Tab 1.1. Comparing the normalised distributions on the left with the absolute counts on the right shows the good agreement of the algorithm for different values of **temdif** on strong TCs and the higher frequency of lower category storms for the lower **temdif** values.

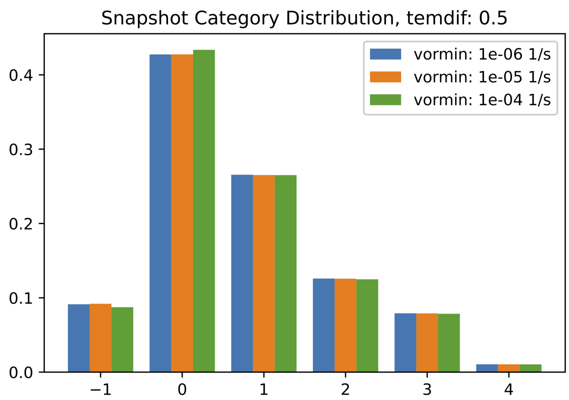


Figure 3.5: Snapshot category distribution for a weak warm core criterion and several vorticity thresholds. The Y-axis shows the normalised counts.

3.5 TC Genesis regions

A direct consequence of the observations from Sec. 3.4 is that with a weaker warm core criterion the TCs can be found already when they have not had much time to intensify. This is relevant because from best track data it is expected that most TCs form in the main development region (MDR) which lies roughly between $10-20$ °N and $20-80$ °W. However, when checking the genesis locations and density in Fig. 3.6, it can be seen that the formation of TCs in ICON is not limited to the MDR. When using a weaker warm core criterion, the areas of highest TC density shift more towards the expected region. This can be interpreted as the successful detection of TCs in earlier development phases.

3.6 Track occurrence frequency

It has been shown that the choice of parameters affects typical distributions like the lifetime and intensity of the found TCs. With the purpose of assessing the effect of different warm core criteria on the length of the tracks, the point occurrence frequency was visualised. This quantity is the frequency of a specific location at a certain time being recognised as part of a TC track. Specifically,

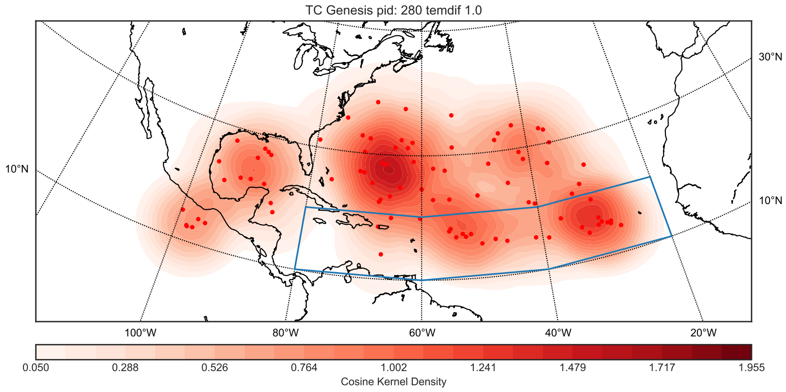


Figure 3.6: Genesis spots and density for a specific parameter combination.

it evaluates to 1 if all parameter combinations determine a point to be a TC and to 0.5 if only the half of them do. The tracks shown in Fig. 3.7 are determined using the parameter combinations from Tab. 3.2.

parameter	**unit**	**values**
temdif	K	0.5, 1 ,1.5
temdis	m	200000
slpdis	m	100000
vormin	1/s	10^{-5}

Table 3.2: Parameter combinations used for occurrence frequency plots in Fig. 3.7.

It can be observed that the weaker warm core criterion tracks TCs that are found by only a small part of the parameter combinations. Furthermore, the ends of TC tracks that are found in all of the three plots are longer for the lower values of temdif. Therefore, the weaker warm core criterion enables tracking in earlier development phases. This observation will be made more precise in the following section.

3.7 Matching TC tracks across parameter combinations

So far the effects of different parameter variations have ben thoroughly analysed. However, it still remains to be seen if different parameter combinations identify the same TCs. In order to achieve this, different tracks were matched according to the following procedure:

- the track of a certain TC was chosen as a base track
- the point of maximum wind of this track was found
- match tracks found in the same simulation run created with different parameter combinations if: they have an entry at the same time like the maximum wind entry of the base track and at a position within 70 km

Finally, all tracks, their start- and end-points, the TC intensity and the occurrence frequency within the family of tracks were plotted. This analysis was performed on 100 tracks sampled from all parameter combinations and for 54 tracks that were found using the original parameter combination. A resulting plot from the first selection can be seen in Fig. 3.8.

The first thing that strikes the eye is the lack of different tracks. Due to the SLP minimum-finding-procedure, all parameter combinations find the same locations and only very rarely do the tracks not fall on top of each other. What differs strongly, however, are the start- and end-points

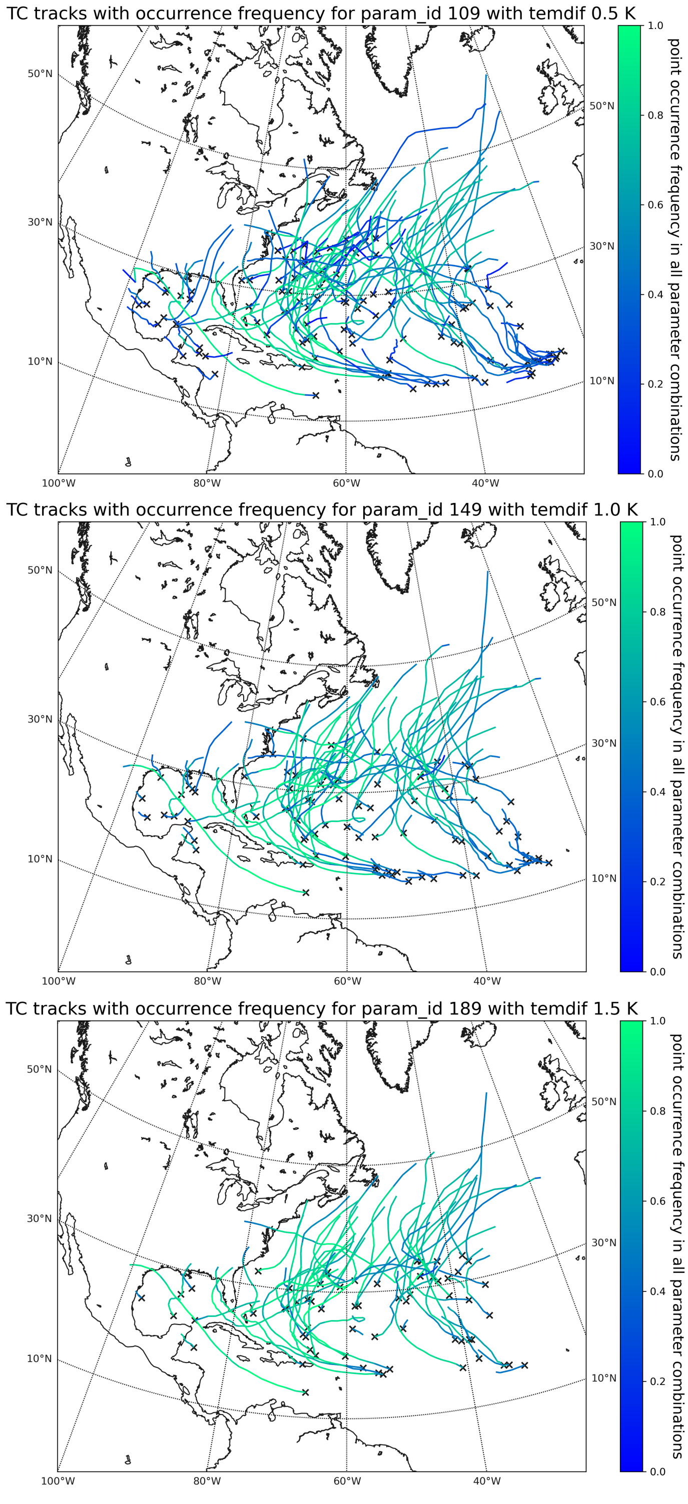
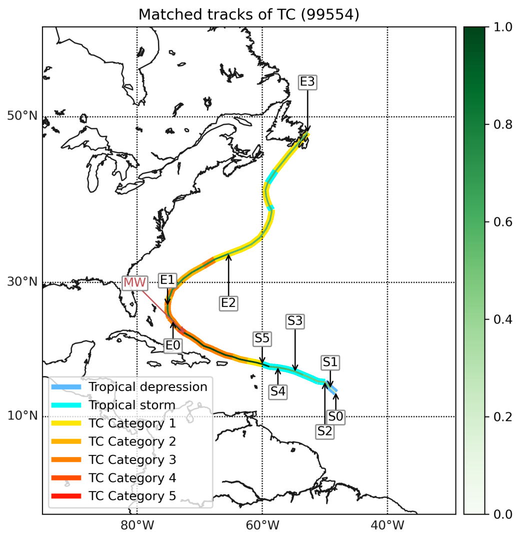


Figure 3.7: Tracks from all ensemble members for the parameter combinations in Tab. 3.2.

of the tracks. Only between the start-point S5 and the end-point E0 is the track found by all parameter combinations. This is the duration of highest intensity, as expected.
When considering the start-points S0–S3 and their corresponding parameter means on the right of the figure, it can be seen that the required temperature difference from the environment (temdif)



Parameter means for the individual starting-points

	temdif	temdis	vormin	slpdis	count
S0	5.83e-01	3.67e+05	3.7e-05	1.25e+05	36
S1	8.12e-01	3.e+05	3.7e-05	1.25e+05	48
S2	1.21e+00	3.e+05	3.70e-05	1.25e+05	72
S3	1.38e+00	2.e+05	3.70e-05	1.25e+05	24
S4	5.e-01	1.e+05	3.70e-05	1.25e+05	12
S5	9.22e-01	1.88e+05	6.39e-04	1.25e+05	128

Parameter means at the point of maximum wind

	temdif	temdis	vormin	slpdis	count
MW	9.5e-01	2.48e+05	2.78e-04	1.25e+05	320.0

Parameter means for the individual end-points

	temdif	temdis	vormin	slpdis	count
E0	8.75e-01	8.75e+04	2.78e-04	1.25e+05	64
E1	5.e-01	1.e+05	2.78e-04	1.25e+05	16
E2	1.e+00	3.e+05	1.00e-03	1.25e+05	60
E3	1.e+00	3.e+05	3.7e-05	1.25e+05	180

Figure 3.8: Tracks matching the category 4 TC (99554). The inner colour shows how many parameter combinations found the specific track segment and the background colour is set according to the current TC intensity. The start- (S1,S2,. . .), end-points (E1, E2,. . .) and the point of maximum wind (MW) are labeled.
The tables on the right display the mean values of the parameters that correspond to the labeled points. The count specifies how many parameter combinations lead to tracks ending or starting at the given location. Since the tracks were matched using the point of maximum wind (MW), its mean values are those of the family of tracks. Its count shows how many of the 400 parameter combinations found the collection of tracks. The colour coding is dependent on the maximum and minimum values of each table column and differs between tables.

is monotonically increasing while the size of the area considered as environment is decreasing. This can clearly be interpreted as an increasingly strict warm core criterion. Once again, it can be concluded that a weaker warm core criterion leads to tracking of TCs in earlier development phases.
While the temperature difference required for S4 is quite low with 0.5 K, the very restrictive temdis of 100 km leads to a later tracking of the TC. This can be understood when considering the plot of the temperature anomaly of the same TC in Fig. 3.9 at the given time. At 300 hPa which corresponds to roughly 9 km no sufficiently strong temperature anomaly can be determined when only considering the area 100 km away from the TC center.
The last starting point (S5) occurs at the time when the storm intensifies to a category one TC. This late detection is due to the high vorticity threshold which is more than an order of magnitude larger than for the other starting points.
A similar analysis can be performed for the endpoints. We will now look a bit more closely at the parameters corresponding to the different start points.

3.8 Parameter feature engineering

The previous discussion suggests that a more general criterion might be reached through algebraic combination of different parameters. Specifically, a single warm-core criterion would be a more elegant characterisation of this physical TC property. A first naive approach was taken using Eq. 3.1.

$$\mathrm{warmcore}_i = \frac{\mathrm{temdif}_i}{\max_i\{\mathrm{temdif}_i\}} - 3 \times \frac{\mathrm{temdis}_i}{\max_i\{\mathrm{temdis}_i\}} \tag{3.1}$$

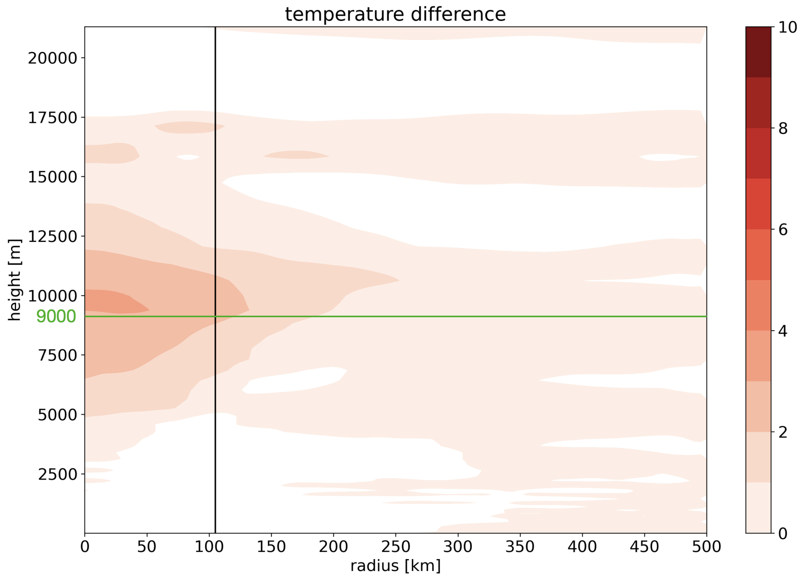


Figure 3.9: Temperature anomaly of the TC (99554) at the point S4. The height of 9km is marked in green.

With the goal of relating the two parameters temdif and temdis, both were first max-normalised. Furthermore, due to their physical meanings, the new warm-core criterion has to be proportional to temdif and inversely proportional to temdis. Finally, the factor three on the temdis term was determined through trial and error. Fig. 3.10 shows a plot of the relationship between the vorticity and the new warm-core criterion for the different start-points. Although the clustering is in no way satisfactory, it can be seen that the means of the warm-core criterion of points S0-S4 are increasing. Furthermore, with a combination of the minimum vorticity threshold with the new warm-core criterion S5 can be easily separated from the other locations. Further feature engineering might prove fruitful.

3.9 Analysis of an interrupted track

Originally, the parameters from Tab. 3.3 were usually used to track TCs. Consequently, 58 plots

parameter	**unit**	**values**
temdif	K	1
temdis	km	200
slpdis	km	100
vormin	1/s	10^{-5}

Table 3.3: Original parameter combination

with matched tracks were created from base TCs that were found using the original parameters. They can be found in the supplementary information. One peculiarity of these plots is that occasionally the point of maximum wind of the base TC is not the actual point of maximum wind of the family of tracks. An exemplary incident is depicted in Fig. 3.11. The original parameter combination only tracks the TC up to E1. This is unfortunate since the actual point of maximum wind is much later. The original parameters therefore loose the TC before it intensifies to category four. Most likely, the TC would have been found later again. However, the track is interrupted and incorrectly identified as two different TCs. This shows the great potential of the parameter variation approach. By allowing different thresholds and matching the tracks in retrospective, the

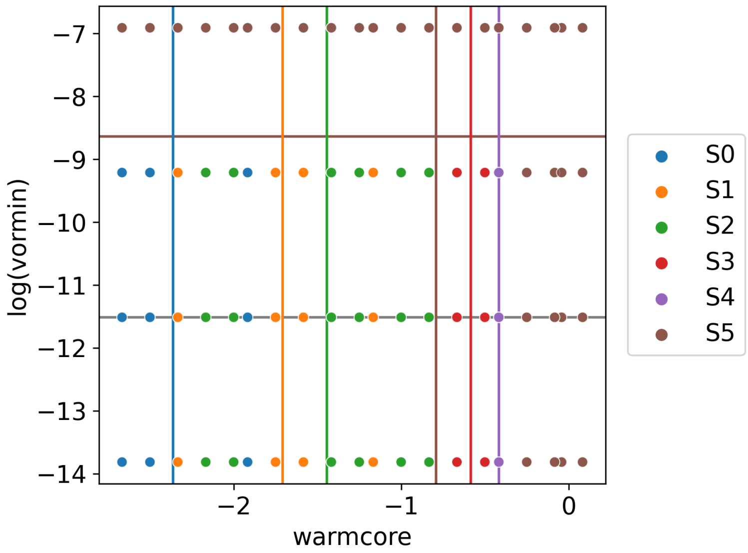
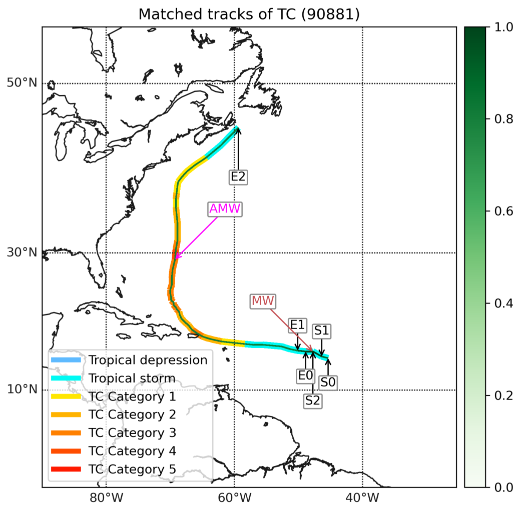


Figure 3.10: Parameter scatterplot between the minimum vorticity and the warm core criterion from Eq. 3.1. S1-S5 are the starting points of the TC with tc_id 99554. The horizontal and vertical lines show the respective parameter means of the start-points. The horizontal grey line is the mean of start-points S0-S4.



Parameter means for the individual starting-points

	temdif	temdis	vormin	slpdis	count
S0	7.78e-01	3.22e+05	3.7e-05	1.25e+05	108
S1	1.12e+00	2.76e+05	3.89e-05	1.29e+05	34
S2	1.27e+00	2.55e+05	3.99e-05	1.32e+05	22

Parameter means at the point of maximum wind

	temdif	temdis	vormin	slpdis	count
MW	9.15e-01	3.04e+05	3.78e-05	1.27e+05	164.0

Parameter means for the individual end-points

	temdif	temdis	vormin	slpdis	count
E0	5.e-01	1.e+05	4.33e-05	1.4e+05	10
E1	1.27e+00	2.55e+05	3.99e-05	1.32e+05	22
E2	8.86e-01	3.27e+05	3.7e-05	1.25e+05	132

Figure 3.11: An interrupted track where the actual radius of maximum wind (AMW) is different from the radius of maximum wind (MW) of the base TC used for the track matching.

TC is tracked over its entire lifetime. The next step should be to use these techniques automatically on all tracks and to filter out tracks that not sufficiently many parameters agree on.

Chapter 4

Conclusion

This project significantly improved the speed and efficiency of a TC tracking algorithm. For a single ensemble member and parameter combination the speedup is sixtyfold. However, because the new algorithm is I/O bound, it really wins over the previous one when many parameter combinations are tested on the same simulation run. A tracking run including 600 parameter combinations and 20 members is 30000-times faster than running the previous algorithm for each parameter combination and member separately.
An exploratory data analysis revealed interesting aspects of the different tracking criteria as outlined before. The capability to quickly try different parameter combinations and the already produced tracking data empowers a wide range of further analysis steps. One idea would be to define a metric that measures the quality of the produced tracking data. This metric could be based on statistics from reanalysis data.
A very illuminating endeavour was the matching of tracks across parameter combinations. This lead to an improved understanding of the influence of each parameter and their interplay on the result. Furthermore, it showed the way to robust tracking across TC development phases.
On the technical side, the use of a high level library made for parallel computation could lead to an even better use of the computing resources. From initial screenings, the Dask library [21] seems promising. Especially the reading of data could be improved, since now the 20 different processes that track TCs in the different simulation runs access different files on the same disk. Using one process that reads data and feeds it to other process that analyse it would reduce the disk handover time between processes.
Finally, combining the ideas presented in this conclusion, would lead to a procedure that can be robustly applied to different simulation and reanalysis data requiring only little calibration.

Appendix A

Additional TC Genesis Plots

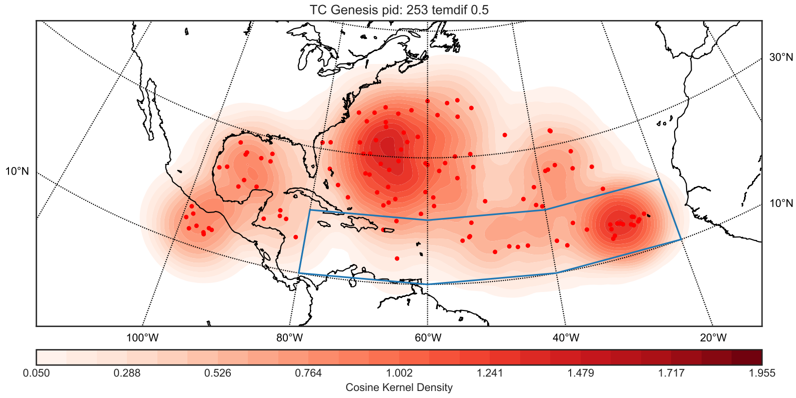


Figure A.1: Genesis spots and density for temdif = 0.5.

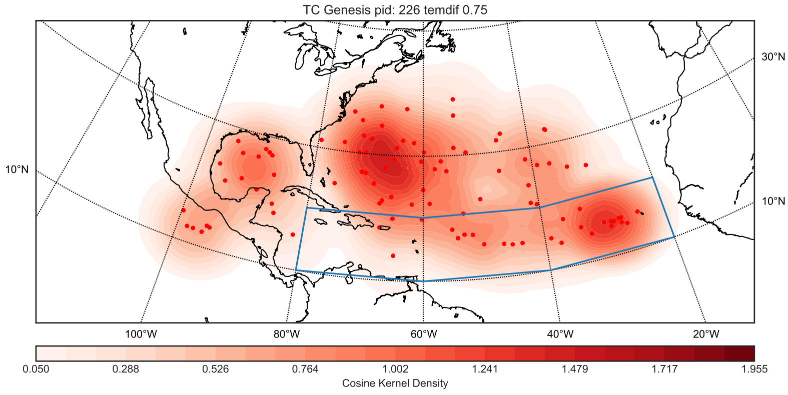


Figure A.2: Genesis spots and density for temdif = 0.75.

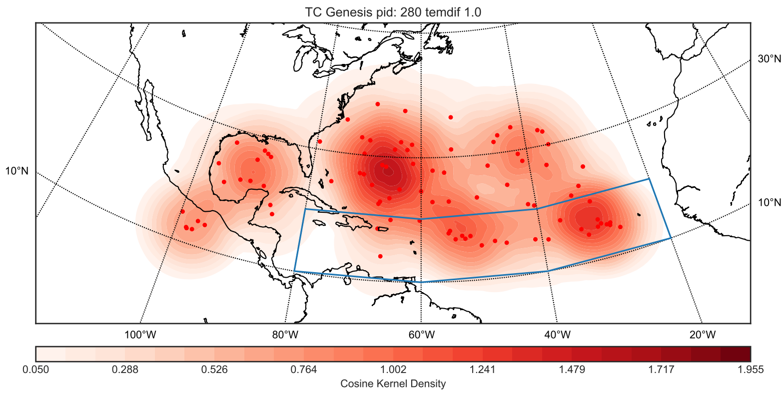


Figure A.3: Genesis spots and density for temdif = 1.0.

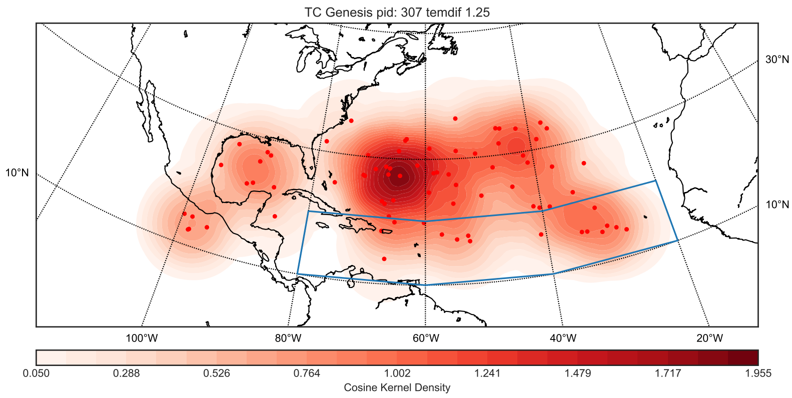


Figure A.4: Genesis spots and density for temdif = 1.25.

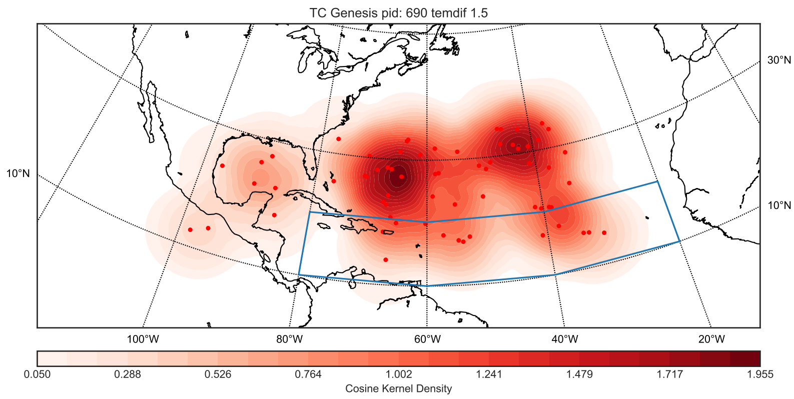
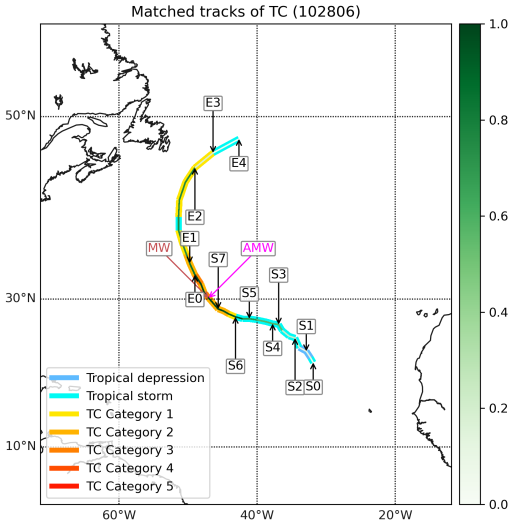


Figure A.5: Genesis spots and density for temdif = 1.5.

Appendix B

Further matched tracks plots



Parameter means for the individual starting-points

	temdif	temdis	vormin	slpdis	count
S0	6.25e-01	4.e+05	3.70e-05	1.25e+05	24
S1	5.e-01	3.e+05	3.70e-05	1.25e+05	12
S2	1.e+00	3.67e+05	3.7e-05	1.25e+05	36
S3	1.03e+00	2.44e+05	3.70e-05	1.25e+05	108
S4	1.5e+00	2.e+05	3.70e-05	1.25e+05	12
S5	8.75e-01	8.75e+04	3.70e-05	1.25e+05	48
S6	9.5e-01	2.48e+05	1.00e-03	1.25e+05	80
S7	1.5e+00	1.e+05	2.78e-04	1.25e+05	16

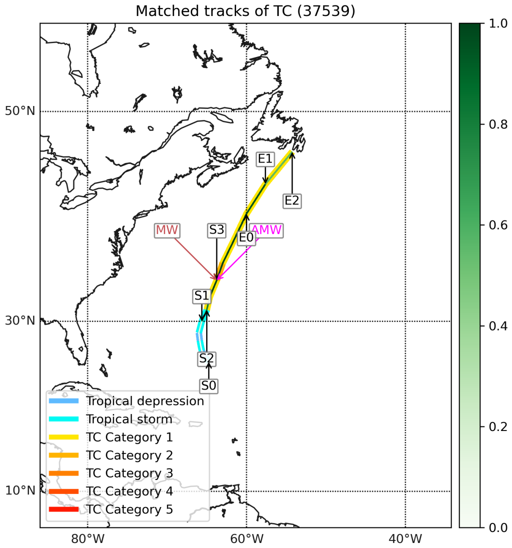
Parameter means at the point of maximum wind

	temdif	temdis	vormin	slpdis	count
MW	9.76e-01	2.40e+05	2.78e-04	1.25e+05	336.0

Parameter means for the individual end-points

	temdif	temdis	vormin	slpdis	count
E0	1.08e+00	8.33e+04	2.78e-04	1.25e+05	48
E1	8.75e-01	1.e+05	2.78e-04	1.25e+05	32
E2	1.02e+00	2.59e+05	3.51e-04	1.25e+05	196
E3	8.75e-01	3.75e+05	3.7e-05	1.25e+05	48
E4	5.e-01	4.e+05	3.70e-05	1.25e+05	12

Figure B.1: Matching plot of TC (102806), Category 3.



Parameter means for the individual starting-points

	temdif	temdis	vormin	slpdis	count
S0	5.e-01	4.e+05	3.70e-05	1.25e+05	12
S1	7.5e-01	3.67e+05	3.7e-05	1.25e+05	36
S2	1.03e+00	2.67e+05	3.33e-04	1.25e+05	208
S3	8.75e-01	8.75e+04	2.78e-04	1.25e+05	64

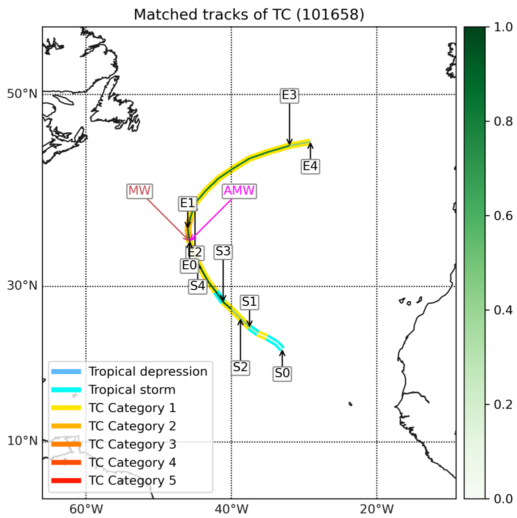
Parameter means at the point of maximum wind

	temdif	temdis	vormin	slpdis	count
MW	9.5e-01	2.48e+05	2.78e-04	1.25e+05	320.0

Parameter means for the individual end-points

	temdif	temdis	vormin	slpdis	count
E0	1.25e+00	1.e+05	2.78e-04	1.25e+05	16
E1	9.35e-01	2.e+05	6.27e-04	1.25e+05	124
E2	9.33e-01	2.93e+05	3.7e-05	1.25e+05	180

Figure B.2: Matching plot of TC (37539), Category 2.



Parameter means for the individual starting-points

	temdif	temdis	vormin	slpdis	count
S0	5.e-01	4.e+05	3.70e-05	1.25e+05	12
S1	9.64e-01	3.57e+05	3.70e-05	1.25e+05	84
S2	8.91e-01	3.06e+05	6.39e-04	1.25e+05	64
S3	1.02e+00	1.42e+05	2.78e-04	1.25e+05	192
S4	1.e+00	5.e+04	2.78e-04	1.25e+05	16

Parameter means at the point of maximum wind

	temdif	temdis	vormin	slpdis	count
MW	9.67e-01	2.24e+05	2.78e-04	1.25e+05	368.0

Parameter means for the individual end-points

	temdif	temdis	vormin	slpdis	count
E0	8.75e-01	5.e+04	2.78e-04	1.25e+05	32
E1	1.5e+00	1.e+05	2.78e-04	1.25e+05	16
E2	1.25e+00	1.e+05	2.78e-04	1.25e+05	16
E3	8.97e-01	1.91e+05	5.75e-04	1.25e+05	136
E4	9.64e-01	3.07e+05	3.7e-05	1.25e+05	168

Figure B.3: Matching plot of TC (101658), Category 2.

Appendix C

Lifetime dependence on the warm core criterion

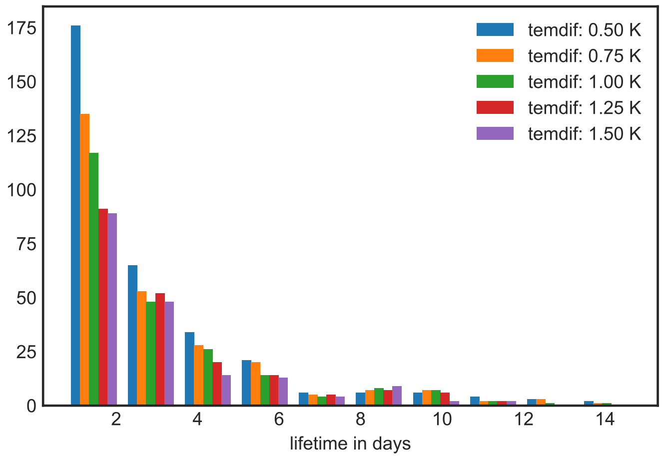


Figure C.1: Histogram showing the number of TCs with a certain lifetime in days.

Nomenclature

Acronyms and Abbreviations

ETH	Eidgenössische Technische Hochschule
TC	Tropical cyclone
TD	Tropical depression
TS	Tropical storm
DWD	German Weather Service
MPIM	Max Planck Institute for Meteorology
ICON	Icosahedral Nonhydrostatic Model developed by the DWD and the MPIM
SST	sea surface temperature
SLP	sea level pressure
MDR	main development region

Bibliography

[1] K. Emanuel, "Tropical cyclones," *Annual Review of Earth and Planetary Sciences*, vol. 31, no. 1, pp. 75–104, 2003.

[2] A. Jones, J. Haywood, N. Dunstone, K. Emanuel, M. Hawcroft, K. Hodges, and A. Jones, "Impacts of hemispheric solar geoengineering on tropical cyclone frequency," *Nature Communications*, vol. 8, 12 2017.

[3] P. Webster and G. Holland, "Role of hurricanes in the global heat balance," *AGU Fall Meeting Abstracts*, 12 2006.

[4] R. Pielke, J. Gratz, C. Landsea, D. Collins, M. Saunders, and R. Musulin, "Normalized hurricane damage in the united states: 1900–2005," *Natural Hazards Review*, vol. 9, 02 2008.

[5] I. Beven, John L., L. A. Avila, E. S. Blake, D. P. Brown, J. L. Franklin, R. D. Knabb, R. J. Pasch, J. R. Rhome, and S. R. Stewart, "Atlantic Hurricane Season of 2005," *Monthly Weather Review*, vol. 136, pp. 1109–1173, 03 2008.

[6] "Saffir-simpson hurricane wind scale." `https://www.nhc.noaa.gov/aboutsshws.php`. (Accessed on 11/08/2020).

[7] "Tropical cyclone climatology." `https://www.nhc.noaa.gov/climo/`. (Accessed on 11/10/2020).

[8] Kelvinsong, "File:hurricane-en.svg - wikimedia commons." `https://commons.wikimedia.org/wiki/File:Hurricane-en.svg`. (Accessed on 11/10/2020).

[9] U. Lohmann, F. Lüönd, and F. Mahrt, *Storms and cloud dynamics*, pp. 285–322. Cambridge University Press, 2016.

[10] M. T. Montgomery and R. K. Smith, "Recent developments in the fluid dynamics of tropical cyclones," *Annual Review of Fluid Mechanics*, vol. 49, no. 1, pp. 541–574, 2017.

[11] E. Palmén, "On the formation and structure of tropical hurricanes," 1948.

[12] U. Neu, M. G. Akperov, N. Bellenbaum, R. Benestad, R. Blender, R. Caballero, A. Cocozza, H. F. Dacre, Y. Feng, K. Fraedrich, J. Grieger, S. Gulev, J. Hanley, T. Hewson, M. Inatsu, K. Keay, S. F. Kew, I. Kindem, G. C. Leckebusch, M. L. R. Liberato, P. Lionello, I. I. Mokhov, J. G. Pinto, C. C. Raible, M. Reale, I. Rudeva, M. Schuster, I. Simmonds, M. Sinclair, M. Sprenger, N. D. Tilinina, I. F. Trigo, S. Ulbrich, U. Ulbrich, X. L. Wang, and H. Wernli, "IMILAST: A Community Effort to Intercompare Extratropical Cyclone Detection and Tracking Algorithms: ," *Bulletin of the American Meteorological Society*, vol. 94, pp. 529–547, 04 2013.

[13] M. Horn, K. Walsh, M. Zhao, S. J. Camargo, E. Scoccimarro, H. Murakami, H. Wang, A. Ballinger, A. Kumar, D. A. Shaevitz, J. A. Jonas, and K. Oouchi, "Tracking Scheme Dependence of Simulated Tropical Cyclone Response to Idealized Climate Simulations," *Journal of Climate*, vol. 27, pp. 9197–9213, 12 2014.

[14] S. Kleppek, V. Muccione, C. Raible, D. Bresch, P. Köllner-Heck, and T. Stocker, "Tropical cyclones in era40: A detection and tracking method," *Geophysical Research Letters - GEOPHYS RES LETT*, vol. 35, 05 2008.

[15] "Overview - icon :: Icosahedral nonhydrostatic weather and climate model - project management service." `https://code.mpimet.mpg.de/projects/iconpublic`. (Accessed on 11/24/2020).

[16] "Wetter und klima - deutscher wetterdienst - numerical weather prediction models." `https://www.dwd.de/EN/research/weatherforecasting/num_modelling/01_num_weather_prediction_modells/num_weather_prediction_models_node.html`. (Accessed on 11/11/2020).

[17] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del R'ıo, M. Wiebe, P. Peterson, P. G'erard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, pp. 357–362, Sept. 2020.

[18] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 6 2014.

[19] T. pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020.

[20] R. Dare and J. Mcbride, "The threshold sea surface temperature condition for tropical cyclogenesis," *Journal of Climate - J CLIMATE*, vol. 24, pp. 4570–4576, 09 2011.

[21] Dask Development Team, *Dask: Library for dynamic task scheduling*, 2016.