



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich



MASTER'S THESIS

Evaluating the Performance of Weather Forecasting Models in Predicting Mediterranean Cyclones

Department of Earth Sciences, ETH Zurich
Institute for Atmospheric and Climate Science

Dr. Emmanouil Flaounas, Institute for Atmospheric and Climate Science, ETH Zurich
(main advisor)

Dr. Franziska Aemisegger, Institute for Atmospheric and Climate Science, ETH
Zurich

Submitted by

Larissa Ott
14-704-589

Zurich, March 24, 2020

Abstract

Cyclones strongly influence weather and climate and can be associated with extreme weather phenomena. A useful framework to study the formation and development of cyclones is potential vorticity (PV), according to which upper-level PV anomalies, often in the form of a PV streamer, induce cyclogenesis. A good forecasting system is expected to credibly represent these anomalies. For testing a forecasting model's capability to adequately reproduce PV streamers, object based measures present a valuable verification method. An example for such an object based approach is the SAL method. This method identifies objects within the forecasted and a reference field and evaluates them based on three components for structure (S), amplitude (A) and location (L). By using an object based measure like SAL the structure of the forecasted field is taken into account and specific shortcomings of a model can be identified. The SAL method was designed for the evaluation of precipitation forecasts. In this study, we test if the SAL method can also be applied to PV fields and investigate relationships between forecasted PV streamers. This is done for the ensemble forecast from the European Centre for Medium-Range Weather Forecasts for Medicane Zorbas, a tropical-like Mediterranean cyclone that occurred in September 2018, where ensemble members showed considerable differences in their forecasted PV streamers and cyclones. The tests revealed that the SAL method in its original definition is not suitable for the application on PV fields. Only the amplitude component works as expected. For the evaluation of the PV streamer's location two new components were defined. One component is based on the overlap between the objects of the forecasted and reference field and is able to capture the location displacement of forecasted PV objects. Investigations on the relationship between the location of PV streamer and cyclone lead to insights about what characteristics of the PV streamer influence the cyclone location. The important characteristic is directionality. It was found that the location of the most eastern part of the PV streamer has an influence on cyclone location and it makes a difference whether a PV streamer was displaced to the east or to the west in the forecast. In order to quantify this a second component was defined that evaluates the location based on its potential to produce a well forecasted cyclone in terms of location. With this additional component the relationship between the location of PV streamer and cyclone can be shown. This study presents a new set of three components for the application on PV forecasts that evaluate the amplitude, the location and additionally the potential for the forecasted cyclone location. In a future study, the proposed framework should be applied systematically to ensemble forecasts of different Mediterranean cyclones.

Abbreviations

ECMWF European Centre for Medium-Range Weather Forecasts

IFS Integrated Forecasting System

MSE Mean Squared Error

PV Potential Vorticity

PVU Potential Vorticity Unit

QPF Quantitative Precipitation Forecast

RMSE Root Mean Squared Error

SLP Sea Level Pressure

UTC Coordinated Universal Time

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 4 |
| 1.1 | PV framework | 4 |
| 1.2 | Forecast evaluation | 5 |
| 1.3 | Objectives | 6 |
| 2 | Data Sets and Methods | 7 |
| 2.1 | Test Case | 7 |
| 2.2 | Data | 7 |
| 2.3 | Methods | 7 |
| 2.3.1 | Cyclone location and intensity | 7 |
| 2.3.2 | SAL Method | 7 |
| 3 | Results and Discussion | 11 |
| 3.1 | Applying SAL to PV field | 11 |
| 3.2 | Modifications | 19 |
| 3.3 | Definition of new components | 20 |
| 3.3.1 | Overlap component | 20 |
| 3.3.2 | Size difference | 24 |
| 3.3.3 | Eastern component | 25 |
| 4 | Conclusions and Outlook | 32 |
| | References | 35 |
| | Appendix A | 37 |
| | Appendix B | 41 |

1 Introduction

Cyclones are systems of rotating wind around a center of low atmospheric pressure and they strongly influence weather and climate. Our region of interest, the Mediterranean, is a region of frequent cyclogenesis (Pettersen 1956). Intense Mediterranean cyclones are often associated with extreme weather phenomena, such as heavy rain, strong winds and storm surges (Lionello et al. 2006). Accurate forecasting of cyclones is essential for preventative measures to be taken and it is important to understand the processes that govern the formation and development of cyclones.

1.1 PV framework

A useful framework to study cyclogenesis is the one of potential vorticity (PV). It was first introduced by Rossby (1939) and its general form and tendency equation derived by Ertel (1942):

$$q = \frac{1}{\rho} \boldsymbol{\eta}_a \cdot \nabla \theta \quad (1)$$

$$\frac{dq}{dt} = \frac{1}{\rho} (\boldsymbol{\eta}_a \cdot \nabla \dot{\theta} + \nabla \times \mathbf{F} \cdot \nabla \theta) \quad (2)$$

Where q is PV, ρ the density, $\boldsymbol{\eta}_a$ the three-dimensional absolute vorticity vector, θ the potential temperature, $\dot{\theta}$ the heating expressed as the material change in θ and \mathbf{F} the three dimensional frictional acceleration. The two fundamental aspects of PV are its conservation and invertibility properties. The former states that PV is conserved along adiabatic, inviscid flow, which can be seen from equation 2 while the latter states that the wind and mass fields can be recovered from the PV field through inversion if balance and boundary conditions are given (Hoskins et al. 1985). Using this framework Hoskins et al. (1985) proposed the baroclinic scenario of cyclogenesis, according to which cyclones form due to a preceding upper-level PV anomaly over a pre-existing low-level baroclinic region. This process is illustrated by figure 1. The upper-level anomaly, usually in the form of a PV streamer, induces a cyclonic wind field, which reaches down to the surface. At surface level this leads to thermal advection, creating a warm temperature anomaly downstream of the upper level PV anomaly. The temperature anomaly then induces its own cyclonic wind field at the ground, leading to the formation of a cyclone located to the east of the upper-level PV anomaly.

Cyclone intensification happens when the two anomalies interact and phase-lock. This mechanism is typical for cyclones in the storm track region of the mid-latitudes, but the most intense Mediterranean cyclones show these characteristics as well (Flaounas et al. 2015).

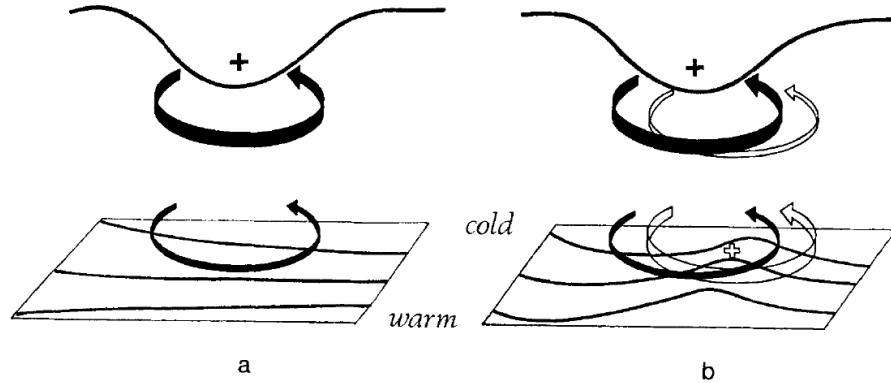


Figure 1: Schematic example of PV induced cyclogenesis after Hoskins et al. 1985.

1.2 Forecast evaluation

The atmosphere is a chaotic system: small changes in the initial conditions can lead to vastly different outcomes. Even small uncertainties in measurements could therefore lead to significantly erroneous forecasts (Lorenz 1963). For Numerical Weather Prediction (NWP) ensemble forecasts are widely used in order to account for the atmospheric system’s sensitivity to uncertainties. For the generation of these ensemble members, the forecast is run several times with slightly perturbed initial conditions. Often the aim is to maximise the ensemble spread to give an indication of the predictability of the atmosphere. Ensemble forecasts can also be used to generate probabilistic forecasts and averaged over a reasonable period, the ensemble mean is more accurate than a single deterministic forecast produced from the same model (Inness and Dorling 2012). Nevertheless, ensemble forecasts can still be subject to inherent model uncertainties or produce erroneous forecasts. Therefore it is important to verify the model’s capability to produce accurate forecasts. There are different ways a forecast can be deemed a “good” forecast. In this study we are concerned with forecast *quality*, which refers to the correspondence of a forecast with observations (Murphy 1993).

This forecast quality can be assessed through a variety of verification measures. Traditional measures to evaluate forecasts of continuous variables include the widely used Mean Squared Error (MSE), or its square root the Root Mean Squared Error (RMSE) (Jolliffe and Stephenson 2012). However, for using such measures for forecasts on a spatial field the forecasts are compared to the observations grid point by grid point and do not account for the inherent structure of the objects. This leads to the double-penalty problem that arises when a model e.g. of a Quantitative Precipitation Forecasts (QPF) predicts the precipitation field correctly but slightly misplaced. The forecast will get penalized twice, once for predicting a precipitation object where there was none and once for the object that was not predicted in the location where there was one. Therefore, a model that forecasts no precipitation at all will get better scores even though it performs worse (Rossa et al. 2008). Analogous to precipitation fields this also applies to PV fields.

In order to avoid this problem several alternative methods have been developed, for example object based measures that identify features in the forecasted field and evaluate them based on different attributes (Casati et al. 2008). One of these is the SAL Method, proposed by Wernli et al. (2008). This is an object-based quality measure that was developed for the evaluation of QPFs. This method identifies objects within the chosen domain but does not match them between forecast and observation. The forecasted fields are then evaluated based on three components: Structure (S), amplitude (A) and location (L). This measure tries to replicate a visual judgement of the spatial structures. With the application of three components the specific shortcomings of the forecasting model can be identified.

1.3 Objectives

In this Master Thesis, the capacity of forecasting models to predict PV anomalies is investigated, most importantly the model's representation of the upper level PV-streamer, as it plays a major role for cyclogenesis. Through this we expect to increase our understanding of where the model's weaknesses lie in terms of forecasting cyclones. For this purpose, the application of the SAL method is tested for the use on PV fields. The method was originally designed for the evaluation of QPFs and within the framework of this project we will apply and if needed adapt the method in order to be applied for forecasted PV fields as well.

In this Master Thesis we aim to answer the following question:

- Can the SAL method be used to evaluate forecasts of PV streamers?

For this purpose, the ensemble forecasts of a PV streamer for a test case will be assessed with the SAL method.

2 Data Sets and Methods

2.1 Test Case

The methods in this study will be performed on a test case: the Medicane (Mediterranean tropical-like cyclone) Zorbas, which occurred in September 2018. The cyclone formed on the 27.09.2018 at 12 UTC preceded by the presence of an upper-level PV streamer, that reached into the central Mediterranean. The forecast of Medicane Zorbas makes an ideal case for testing evaluation methods, since the ensemble forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) showed considerable differences between the forecast members, which lead to uncertainties in the position and structure of the forecasted cyclone (Portmann et al. 2020).

2.2 Data

The forecasts to be evaluated in this study are the operational Integrated Forecasting System (IFS) ensemble forecasts from ECMWF of Medicane Zorbas. In particular the forecasts of the 27.09.2018 at 12 UTC are investigated, at the time where in most ensemble members the PV cut off was forming. These forecasts were initialised at the 24.09.2018 at 00 UTC and therefore have a lead time of 84 hours. For the evaluation of the PV streamer the PV field at 320 K is used with a resolution of $0.5^\circ \times 0.5^\circ$. Information about the cyclone is taken from the sea level pressure (SLP) field at the same resolution. As a reference the ERA5 reanalysis data set is used for the 27.09.2018 at 12 UTC as well. Again the PV field at 320 K and the SLP fields are used with a resolution of $0.5^\circ \times 0.5^\circ$. For both the reference and forecasts a Mediterranean domain is defined, ranging from 5°W to 40°E in longitude and from 25°N to 50°N in latitude.

2.3 Methods

2.3.1 Cyclone location and intensity

Information about the cyclone location and intensity is extracted manually for all 50 members from the SLP data set for the 27.09.2018 at 12 UTC. The cyclone location used corresponds to the grid point with the lowest SLP value inside the visually detected low pressure system in the Mediterranean domain. In one member (4) no cyclone could be detected. The SLP value at the grid point used for the cyclone location serves as a measure for the cyclone intensity.

2.3.2 SAL Method

For the evaluation of the forecasted PV streamers, we compare the forecasted PV fields of the 50 members from the ECMWF forecast to the era5 reanalysis which serves as a reference field. The evaluation will be done with the SAL Method as described by Wernli et al. (2008). The method works as follows: The first step in the procedure is to

define the domain of interest, in our case the Mediterranean region. For the calculation of the S and L parameters, coherent PV objects will be identified. For this we need to define a threshold value until which neighboring grid points are included in the PV object.

As the final step the three components are calculated. The amplitude component A is the normalized difference of the domain-averaged precipitation values:

$$A = \frac{D(R_{\text{mod}}) - D(R_{\text{obs}})}{0.5[D(R_{\text{mod}}) + D(R_{\text{obs}})]} \quad (3)$$

Where R represents the PV field and $D(R)$ the domain average of R . A is therefore not dependent on the structure of the fields but measures the accuracy of the total PV fields. The possible values for A range between $[-2 \ 2]$, where 0 means a perfect forecast. Positive values mean that the PV values in the forecast were overestimated while negative values indicate underestimation.

The location component L consists of two parts that add together for the total location component ($L = L_1 + L_2$). The first term measures the accuracy of the PV distribution by calculating the normalized distance between the centres of mass of the observed and the forecasted fields:

$$L_1 = \frac{|x(R_{\text{mod}}) - x(R_{\text{obs}})|}{d} \quad (4)$$

Here $x(R)$ stands for the center of mass and d is the largest distance between two boundary points of the considered domain. The values range between $[0 \ 1]$, where 0 means that there is no difference in centre of mass between the forecasted and the reference field, but it is not necessarily a perfect forecast. A situation where the reference shows one object and the forecast two objects, but both displaced in a way that the centre of mass aligns with the centre of mass of the reference still yields a 0. This is why there is a second part of the L-component that takes the location of individual objects into account. For the calculation of L_2 the different PV objects are identified and considered separately. For every object the integrated amount R_n of all PV values within the object is calculated. Then for every object the weighted averaged distance between the centres of mass of the individual object x_n and the total PV field x is calculated:

$$r = \frac{\sum_{n=1}^M R_n |x - x_n|}{\sum_{n=1}^M R_n} \quad (5)$$

Where M denotes the number of objects identified. From this L_2 is calculated from the difference of r in the reference and the forecasted field:

$$L_2 = 2 \left[\frac{|r(R_{\text{mod}}) - r(R_{\text{obs}})|}{d} \right] \quad (6)$$

This component also ranges from [0 1], where 0 means either that the averaged distance between the objects and the centre of mass is the same for both reference and forecast, or that both fields contain only one object. The combined component L therefore ranges between [0 2].

Lastly the structure component S is calculated by comparing the volume of the PV objects. The scaled volume V_n is computed for every object as the sum of all values within the object scaled with the maximum value found within the object ($V_n = R_n/R_n^{max}$). In a next step for both fields the weighted mean volume $V(R)$ is calculated as follows:

$$V(R) = \frac{\sum_{n=1}^M R_n V_n}{\sum_{n=1}^M R_n} \quad (7)$$

This volume is weighted by the integrated amount of PV within the objects and it contains information about size and shape of the PV objects. Finally the S-component is calculated, analogous to A as the normalized difference of the scaled Volume of all objects V:

$$S = \frac{V(R_{mod}) - V(R_{obs})}{0.5[V(R_{mod}) + V(R_{obs})]} \quad (8)$$

The S-component also ranges between [-2 2], with 0 standing for a perfect forecast, positive values indicate that forecasted objects were too large or too peaked in structure, negative values the opposite.

A schematic with idealized examples in figure 2 shows how SAL values correspond to different combinations of objects in forecast and observations.

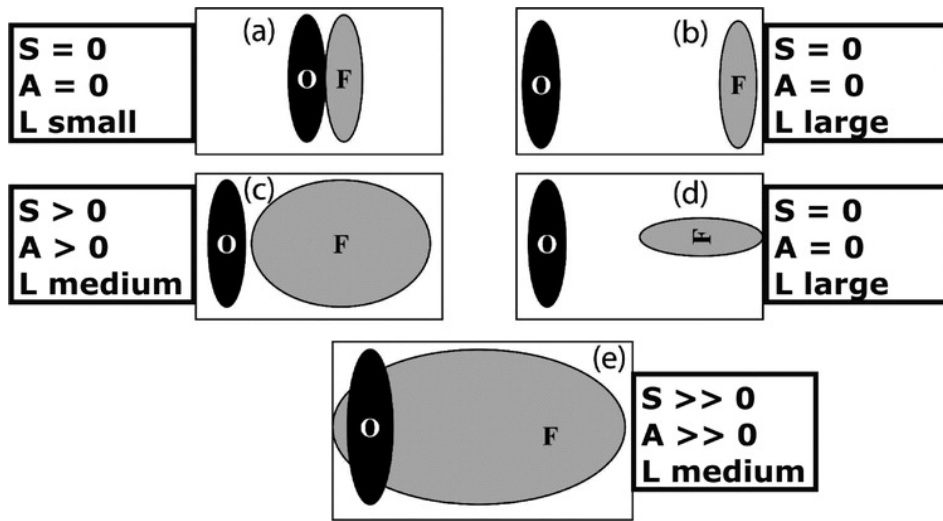


Figure 2: Schematic example of SAL measures for different combinations of forecast and observation fields (Wernli et al. 2008).

3 Results and Discussion

3.1 Applying SAL to PV field

When applying the SAL method to the PV field, the first step is to choose a threshold above which values in the PV field are assigned to an object. Since a value of 2 PVU is often used as a dynamical definition of the tropopause it makes sense to set the threshold to 2 PVU in order to outline objects of stratospheric high-PV air from tropospheric low-PV air. A visual representation of the SAL values for the 50 ensemble members applied to PV with a threshold of 2 PVU can be seen in figure 3. Each dot corresponds to one of the 50 members from the operational IFS ensemble forecasts for Medicane Zorbas from ECMWF, compared to the reference field from ERA5. Their position in the x-direction corresponding to the S-component and their position in the y-direction to the A-component. The L-component is represented with colour. The box ranges between $[-2 \ 2]$ for the S- and A-component, showing the whole possible range for both components. The colour range for the L-component however, is adapted to the range of values within the 50 members.

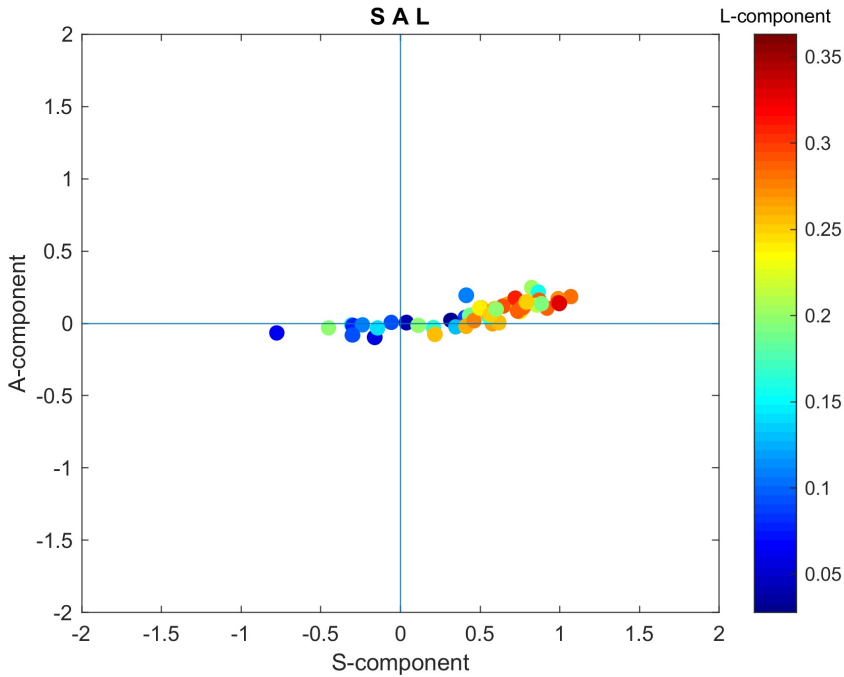


Figure 3: SAL plot for all 50 members

From the distribution of the dots we can deduce that the A-component for all members is fairly good, as the values are close to zero. However, since most of them are located on the positive side, more cases seem to over-, rather than underestimate the PV values. The range for the S-component is larger and contains more positive values as well, which

means the objects are overall larger or flatter. The S- and A-component seem correlated (correlation coefficient: 0.7984). As mentioned by Wernli et al. (2008) this is probably due to the fact that it is hard to get higher average values with objects that are too small or too peaked. Therefore it is rare to get high (low) values for the A-component while simultaneously getting low (high) values for the S-component. The L-component seems to be correlated with the S-component. Its range is also relatively small, as all values are below 0.12. This is because the values for the L-component depend on the domain size. As our domain encompasses the whole Mediterranean, small L scores can already mean large errors.

In order to better understand the behaviour of the SAL method when applied to the PV field, some examples will be discussed. The reference PV field (above a threshold of 2 PVU) is shown in figure 4a. In figure 4bc the PV fields of members 1 and 43 are shown respectively. These were the two of the 50 members which got the best SAL scores, in the sense that for these members all three components were below a certain threshold ($-0.1 < S < 0.1$ and $-0.1 < A < 0.1$ and $L < 0.1$). The red contours indicate the position of the reference PV streamer for direct comparison.

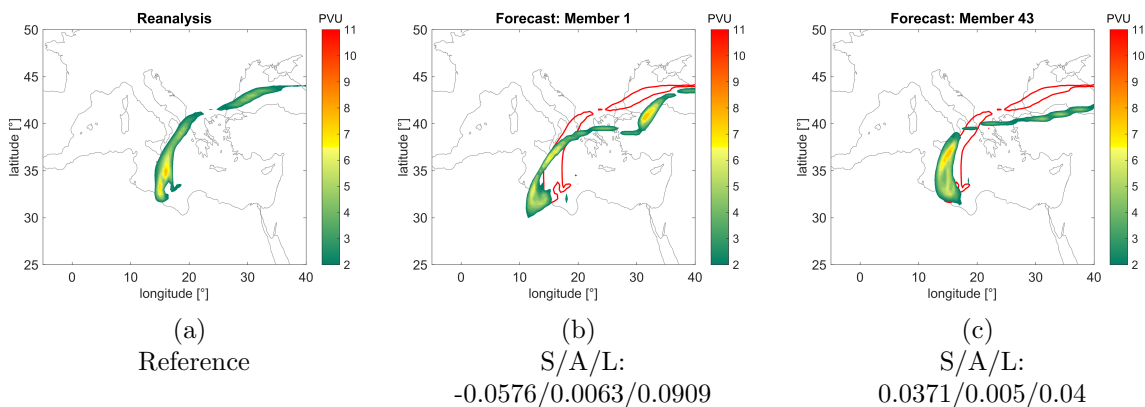


Figure 4: PV streamers for the reference field and members 1 and 43. PV field shown above threshold of 2 PVU, for the forecasted PV fields, the outline of the reference PV objects is shown in red.

Structure component S

As mentioned before, the S-component is a measure of how peaked or flat an object is, as well as a measure for its size. Figure 5ab shows the members with the most positive (1.0683) and the most negative (-0.7767) S score. Member 2 received the highest S score due to the fact that the object is much larger than the reference objects. Member 4, the forecast with the most negative score for the S-component, shows smaller objects than the reference. As the S-component is calculated by weighting the mean of the volume of all objects, it is the average size of the identified objects that is important. Therefore member 2 is not only the most positive one because of its size, but additionally because it is only one large object that is compared to a reference that consists of mostly two smaller objects. In contrast to that we have member 4 where the PV streamer

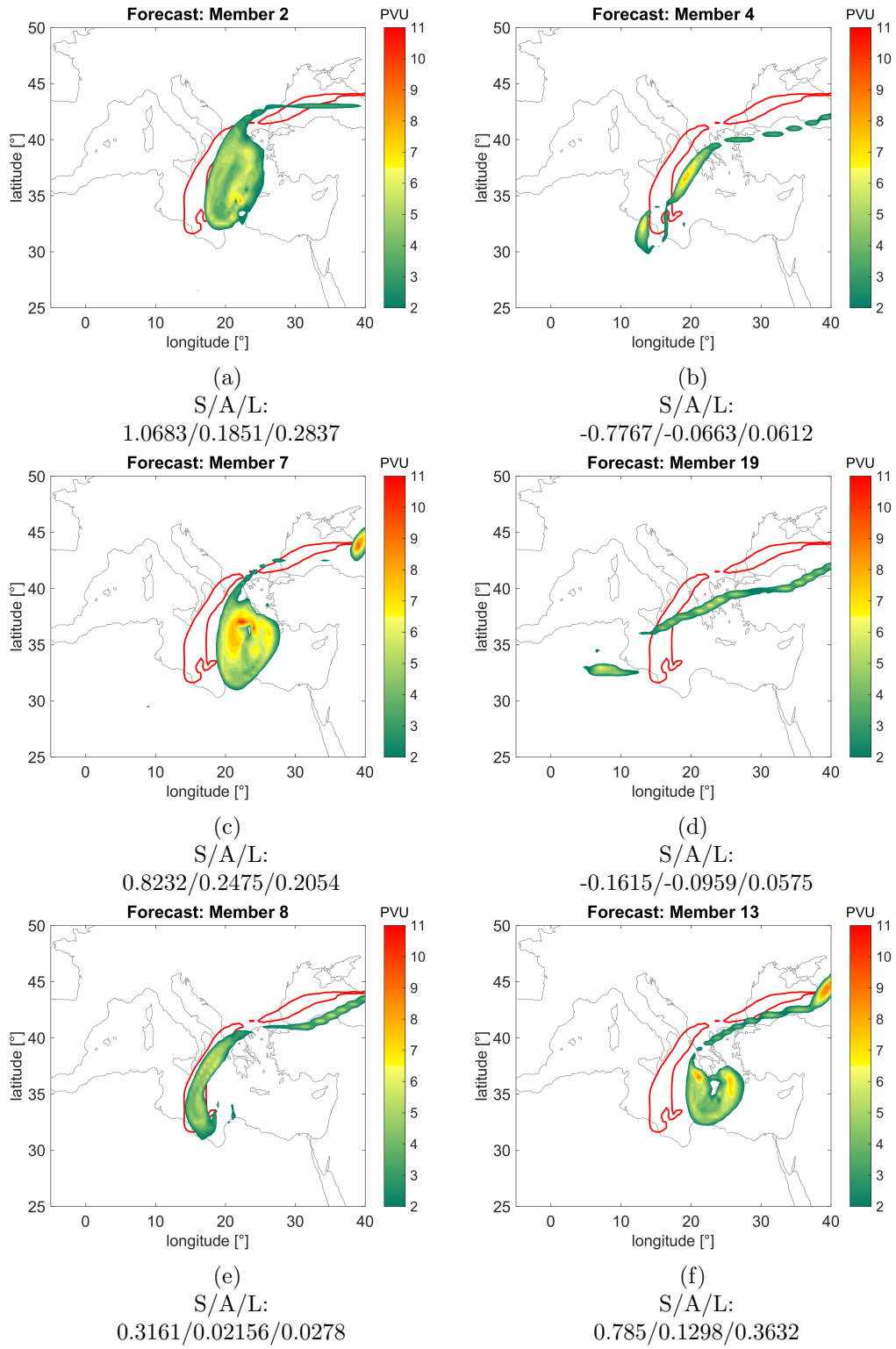


Figure 5: PV fields for some example members above threshold of 2 PVU, reference PV objects in red

disintegrated into several small objects, resulting in a negative S-component. The best S score was achieved by member 43 (figure 4c) which consists of a comparable number of objects of similar size to the reference. How peaked or flat an object is seems to play a minor role for the PV field.

The large dependence on the number of objects identified makes the component susceptible to the camel effect, where a small change in the threshold can change the number of objects identified. With a change in threshold from 1.5 to 2 PVU there is a systematic change of the S-component to higher values in most of the forecast members. The reason for that was the change in the number of objects in the reference PV field, as shown in figure 6.

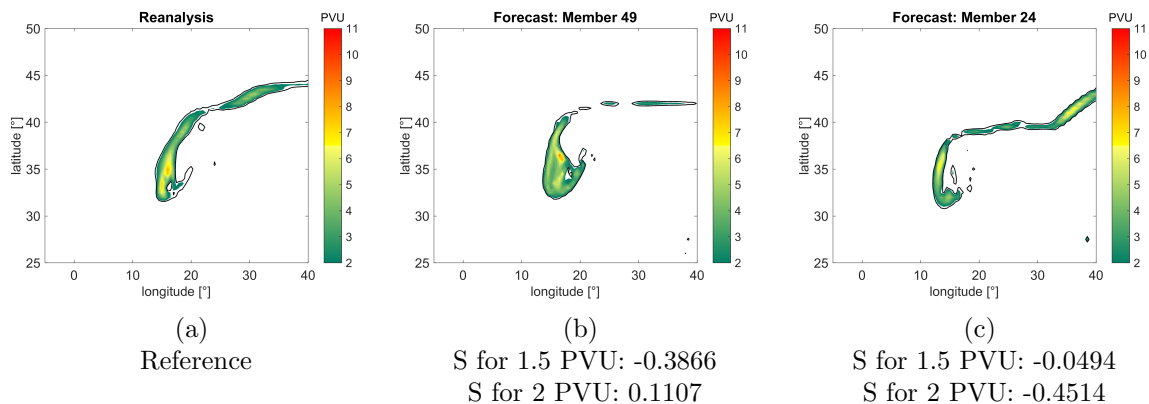


Figure 6: PV fields for the reference and two forecast members in color above a threshold of 2 PVU, PV outline for threshold of 1.5 PVU in black

The PV field at a threshold of 2 PVU is shown in color, while the black outline shows the PV objects at a threshold of 1.5 PVU. Between these two thresholds the reference PV field splits up into several objects while most members like member 49 still contain the same number of objects. This leads to a more positive S-component in all these members because the average weighted volume of the reference field decreased but not in the forecasts. One of the exceptions was member 24 (figure 6c). This member also splits up into more objects between these threshold levels and therefore does not receive a higher score for the S-component. On the contrary, as it splits up into more objects than the reference, the score gets even lower.

Amplitude component A

Figure 5cd shows the members with the most positive (0.2475) and negative (-0.0959) A-component. The most positive one, member 7, shows very high PV values. Due to the smooth transition from low to high PV values, high values in an object also mean high average values and therefore higher overestimation of amplitude. The member with the most negative score, member 19 displays a very thin PV structure. Members where we could expect a good A score consist of objects of similar size and with a similar maximum value to the reference objects.

Location component L

Figure 5ef shows member 8 with the best L score (0.0179) and member 13 with the worst score for L (0.3631). From a visual perspective this makes sense, as member 13 is displaced quite far to the east while member 8 overlaps well with the reference object. What we would expect for the L-component, is that it would correlate well with the location error of the forecasted cyclone. However, the plot in Figure 7 shows no clear relationship (correlation coefficient: 0.3993). The lines are dividing the plot visually into four boxes, based on the distribution of the members within the plot. The categories correspond to cases of "good" and "bad" L scores, as well as "good" and "bad" Location errors for the cyclone. Figure 8 shows the average PV field for each of these boxes, where the plot in the top left corresponds to the box at the top left of the plot etc. From these fields we can deduce what type of PV field leads to which score. The box on the lower left (figure 8c) corresponds to cases with a low L score and a low Location error. The average PV field for this category overlaps quite well with the reference PV streamer (red contour), but there are also some displaced members. The box on the upper right (8b) on the other hand shows the opposite case: members with a high L score as well as a high location error. For these forecast members we expect the location of the PV streamer to be quite bad, which seems to be the case, as the PV streamer is displaced to the east of the reference streamer and barely overlaps with it at all. The average cyclone location is displaced to the east as well. These two boxes represent the expected outcome. The problem lies in the other two categories.

The box in the lower right (8d) shows cases that have a low cyclone location error but still receive a high score for the L-component. The average PV field indicates that this is not due to the fact that the PV streamer was badly located, as it shows a similar distribution as the box to the left. This means that cases with a good PV location are still able to get high L values. The box in the upper left (8a) presents a similar but opposite situation. The cases in this category show a high cyclone location error but receive a low L score, even though the location of the PV streamer is displaced as well, as the average PV field shows. These plots show that the L-component is not able to differentiate between situations of high or low displacement, even though they are visually different.

To better understand this behaviour we take a closer look at the calculation of the L-component. As described in the methods section, the component consists of two parts L_1 and L_2 , of which both use centre of mass as an indicator for location. L_1 measures the total displacement of the identified objects by evaluating the distance between centres of mass of the PV field, while L_2 measures the weighted averaged distance of individual objects. Figure 9 shows a few example members displayed with their centres of mass, as well as their values for L, L_1 and L_2 . Members 28 and 29 (figure 9ab) receive a similar value for the L-component (with a score of 0.2673 and 0.2605 respectively). But from a visual perspective we would expect member 29 to receive a much better score, since the PV objects overlap with the reference objects very well, while in member 28 the PV structures are displaced to the east, not overlapping with the reference objects at all. Although the L-component returns the same values for both members, their values

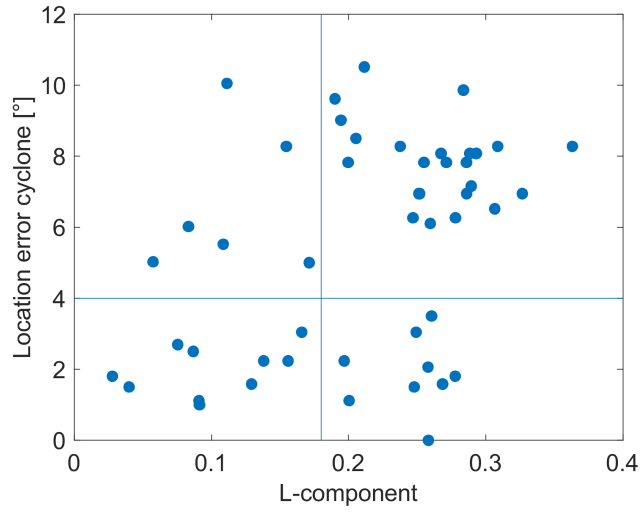


Figure 7: Relationship between Location error of the cyclone and the L-component

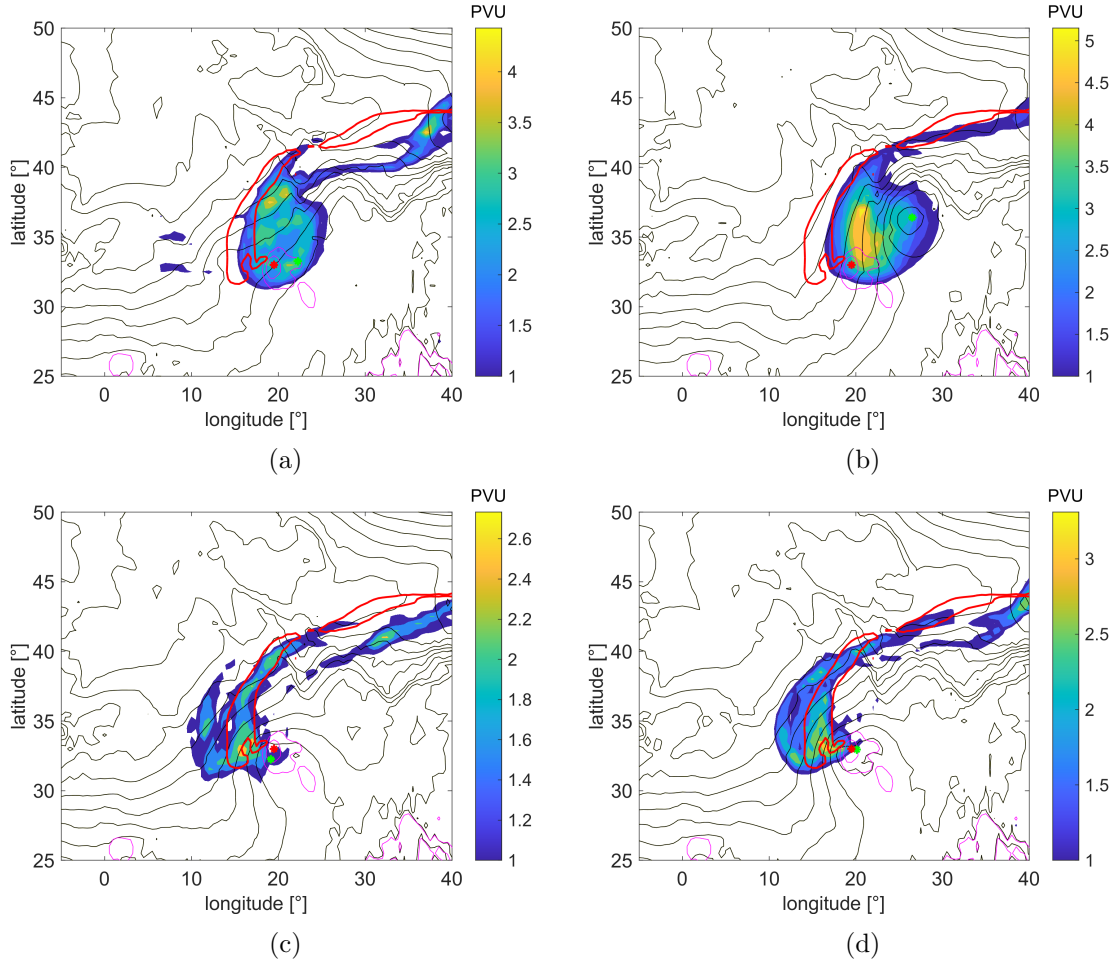


Figure 8: Average PV fields, reference PV objects (red), average SLP contours (black), lowest SLP contours of reference (magenta), average cyclone location (green star) and reference cyclone location (red star).

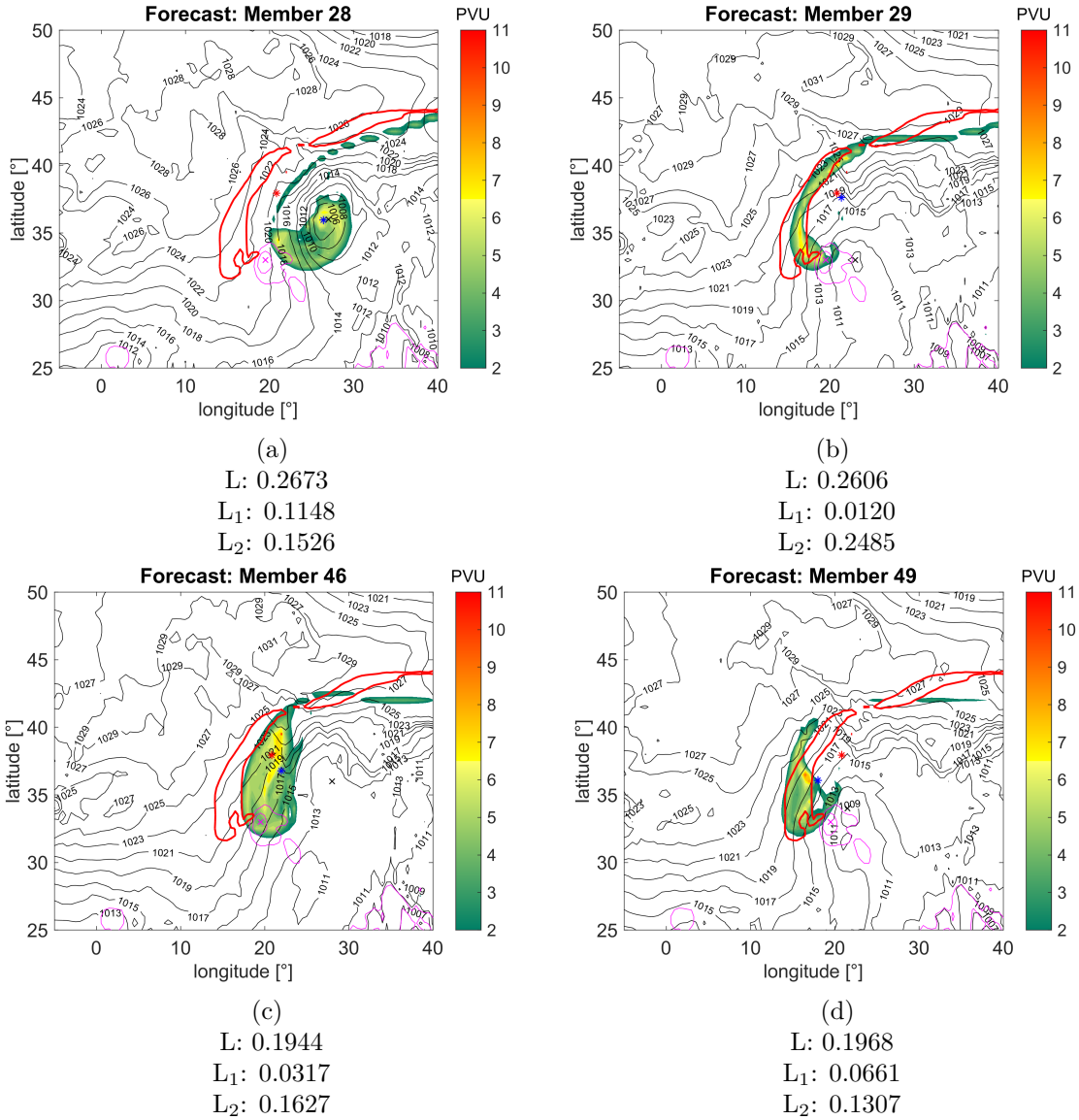


Figure 9: PV fields above 2 PVU with SLP contours (black), cyclone centre (black x), reference PV objects above 2 PVU (red), lowest reference SLP contours (magenta), reference cyclone centre (magenta x), centre of mass (blue star) and reference centre of mass (red star).

for L_1 and L_2 are different. The centre of mass for member 29 is located in the vicinity of the centre of mass of the reference PV field, resulting in a low score for L_1 , while the centre of mass in member 28 is displaced to the east, which leads to a higher score. Only considering L_1 would lead to the expected result. The reason why both members end up having the same score is the opposite tendency in the L_2 score. Member 28 receives a low score for L_2 because the distribution of objects relative to the centre of mass of the total field is rather similar to the reference field. Member 29 on the other hand, receives a high score due to the fact that it mostly consists of one large PV object, rather than two similar sized objects like in the reference field. Having one large object leads to a small average distance as the centre of mass of the large object is near to the total centre of mass while in the case of two similar sized objects, there are two objects with centres of mass more distant to the total centre of mass, which results in a larger difference in the average weighted distance between centres of mass and therefore in a higher value for L_2 . This example shows that the L-component depends on the number of identified objects and can lead to cases where it is not able to differentiate between visually different cases.

Another example of two members that receive similar values for L despite showing visually different PV fields, as well as a large difference in cyclone location are members 46 and 49 (figure 9cd) with scores of 0.1944 and 0.1968 respectively. Despite the main object in member 46 being displaced completely to the east while the object in member 49 is overlapping with the reference, the value for L_1 is lower for member 46. This is due to the fact that the centre of mass of the reference field is influenced by two similar sized objects and therefore lies outside of the PV streamer. Both members in this example consist of mainly one large PV object and therefore display a centre of mass that is closer to this object. This leads to a smaller differences of centres of mass in the member with an object that is displaced, rather than the one that overlaps. L_2 is higher in member 46 because of a similar reason. As both of the members consist mostly of one large object, the weighted average distance to the total centre of mass is smaller than in the reference for both cases, but as the total centre of mass for member 46 lies within the object, the distance is even shorter and therefore receives a higher value for L_2 . From these examples follows, that the L-component does not adequately reflect the location of the PV objects as intended.

3.2 Modifications

Due to the inadequate application of the SAL method to forecasts of PV fields, some attempts at modifying the original method were made. However, the modifications did not improve the SAL method, therefore only a short overview of the modifications is shown.

S- & A-components

When applied to the PV field, the S-component mainly evaluates the size difference between the forecasted and the reference objects. However, the A-component is also directly tied to size and two components showing the same characteristics are not needed. Together with the S-component's susceptibility to the camel effect, we decided to discard the use of the S-component. The A-component is kept but with a minor alteration: Instead of using the mean PV value over the whole domain only values above the set threshold are taken into account. This does not alter the results greatly but makes more sense conceptually, since the distribution of low PV values is not of great interest.

Standardisation and Normalisation

Other attempts in modifying the SAL values consisted of standardising and normalising the PV fields before the application of the components. Standardising shifts the distribution of all values within the domain, such that the mean is zero and the standard deviation 1, while normalisation scales the distribution between the values of [0 1]. These processes only affect the A-component, which is the only component that does not need improvement. These methods were therefore discarded.

Modifications of L-component

Several modifications to the L-components were tried. They are explained shortly, some more details about them can be found in the appendix. The first approach to modify the L-component was to weight the PV exponentially in order to give more weight to high PV values. As mentioned in the last section, the centre of mass is not representative for the location of the actual PV cut off. However, the weighting method does not work, because it is sensitive to high PV values that were found outside of the cut off. Another approach was the varying of the threshold. An objective approach to find a threshold that would yield a better correlation between the L-component and the cyclone location error lead to the result that better correlations were found at very low thresholds, because low PV objects were dragging the centre of mass towards the centre for all members. Higher thresholds had a similar effect as the weighting. These methods were therefore discarded as well. Since none of the modifications for the L-component yielded a successful result, we came to the conclusion that the definition of an entirely new component for location might be necessary.

3.3 Definition of new components

This section deals with the definition and application of new components that are able to adequately assess the forecasted PV fields in terms of location. The aim for this components is to address the problems of the original L-components when applied to PV fields. First, the behaviour of the components should resemble that of a visual judgement. Second, they should be able to differentiate between the cases where we would expect the cyclone error to be high or low, since these cases also were visibly different.

3.3.1 Overlap component

The first new component to be introduced is the L_o -component. Instead of taking the centre of mass as a measure for location this component uses overlap. The original method compares a location that often lies outside of the PV streamer itself and is highly influenced by grid points outside the cut off and are less interesting in terms of cyclone development. By using overlap as a measure, only locations with high PV values are compared.

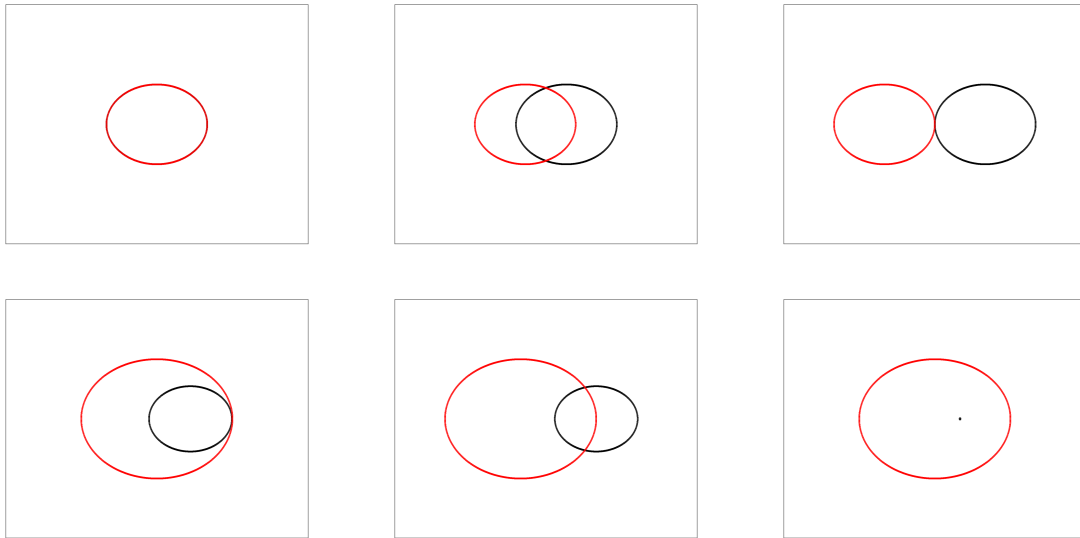


Figure 10: Idealized cases for forecast and reference pairs, with the red and black contour showing the objects identified by a threshold in either field.

The method for using overlap in a component was first conceptualised with ideal cases. For cases such as shown in the top row of figure 10, where we have objects of the same size, the application of a component using overlap is straightforward: An object in the reference field completely overlapping with the object in the forecasted field means a perfect score. When the objects are not touching at all they receive the worst possible score. Every case in between with some parts of the objects overlapping will correspond

to a score proportional to the amount of overlapping grid points compared to the objects total size. With scores ranging between $[0\ 1]$, where 0 means a perfect score, the worst case would translate to a score of 1 and two objects that would overlap by half would receive 0.5 as a score. However, it is not as straightforward if the objects are of different sizes, such as shown in the bottom row of figure 10. In the case shown at the bottom left, the black object is entirely enclosed inside the red object. From the perspective of the black circle it is a perfect overlap, but the red circle is only partly covered. The L_o -component deals with this by using the average size of the objects in the reference and forecasted fields.

It is calculated as follows:

$$L_o = \left| \frac{n_{\text{overlap}}}{0.5[n_{\text{ref}} + n_{\text{for}}]} - 1 \right| \quad (9)$$

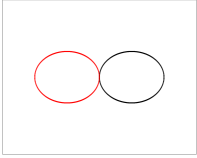
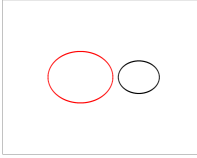
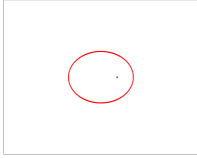
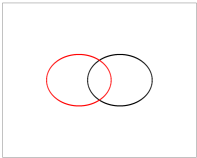
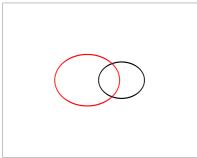
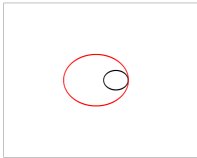
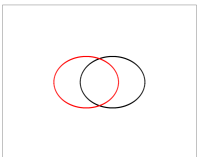
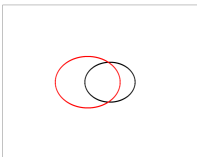

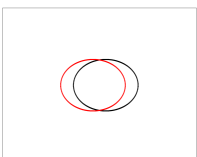
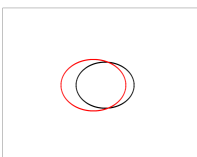

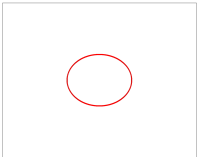
Where n_{overlap} is the number of overlapping grid points, n_{ref} is the number of grid points within objects of the reference field and n_{for} the number of grid points within objects in the forecasted field. With this method, individual objects are not identified. The numbers of grid points used in the calculation, are all grid points above the set threshold, whether they belong to the same object or not.

The formula is constructed to range between $[0\ 1]$ with a perfect forecast resulting in a score of 0, same as in the original SAL components. By using the average of n_{ref} and n_{for} in the denominator the score does not depend on whether the reference or the forecasted objects are larger. A case like the one at the bottom left of figure 10, where a smaller object is fully enclosed within a larger one, would therefore receive a score accounting for the size of both objects. If the smaller object would be half the size of the larger one, from the perspective of the larger object it is overlapped only by half which would translate to a score of 0.5, while from the perspective of the small object the overlap is complete, translating to a perfect score of 0. Using the average size of both objects the resulting score is a $0.\bar{3}$. If the size difference becomes too large, as in the case in figure 10 (bottom right), the score will become 1 even though they fully overlap.

For a better understanding of what a given score of L_o would mean, more examples of idealized cases are shown in table 1. For each score there is a range of possibilities in terms of overlap and size differences. The column to the left shows the cases for objects of the same size. A score of 0.5 would mean that half of their grid points are overlapping. Since 0 means a perfect forecast, a score of 0.25 would mean that the objects overlap by 75 %, if there is no size difference. The column on the right shows the case for each score with the maximum possible size difference, which is the case when the objects overlap completely. The middle column shows a case in between, where there is a size difference and only partial overlapping. The size difference between the objects is indicated with a factor underneath each example. A factor of 0.6 means that the area of the black circle is 0.6 times the area of the red circle. A score of zero can only be achieved for a perfect overlap, so the size difference naturally has to be the

same in both fields as well.

Table 1: Range of possibilities for different scores of L_o

| L_o | no size difference | in between | maximum size difference |
|-------|---|--|---|
| 1 |  |  Factor: 0.4 |  Factor: 0.0001 |
| 0.75 |  |  Factor: 0.5 |  Factor: 0.142857 |
| 0.5 |  |  Factor: 0.6 |  Factor: 0.333333 |
| 0.25 |  |  Factor: 0.8 |  Factor: 0.6 |
| 0 |  | | |

When applied to the PV field of our forecasts of Medicane Zorbas, the resulting relationship of the L_o -component with the cyclone location can be seen in figure 11a). Here instead of choosing a threshold of 2 PVU, the results were averaged for a range of thresholds from 1.5 to 5 PVU. By taking an average of several thresholds, changing behaviour at different thresholds is accounted for. Including higher thresholds until 5 PVU also has a weighting effect, as it shows how well areas of high PVU overlap. Also depicted are the average PV fields of the members within the boxes of the scatter plot. As figure 11c shows, the L_o -component succeeds in attributing low scores to cases with low cyclone error. Also the visual judgement for the PV streamer location matches the lower scores. Unfortunately, for high values of L_o , the component is not able to differentiate between high and low cyclone location errors either.

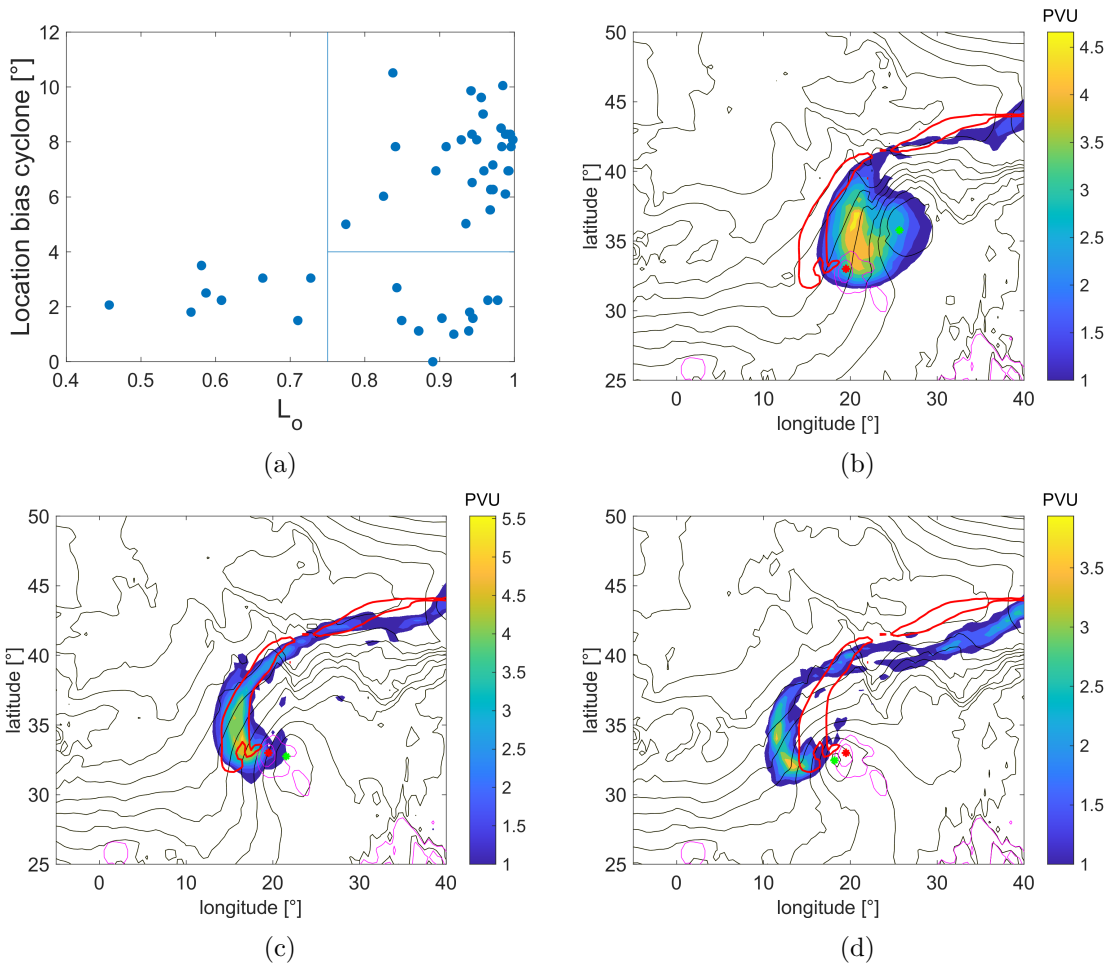


Figure 11: a) Relationship between L_o -component and cyclone location error b-d) averaged PV fields for the members in corresponding box of the scatter plot, average SLP contours (black), 2 PVU outline of reference field (red), lowest SLP contours of reference (magenta), average cyclone location of forecast members (green star) and cyclone location of reference (red star).

3.3.2 Size difference

In order to address the problem of the L_o -component, the application of a new component was tested that evaluates size difference. The method succeeded in differentiating the cases where a low or high cyclone location error could be expected, that L_o could not. However, this was only due to a peculiarity within our test case and since the A-component already is a measure for size the method was discarded in the end. Nevertheless it gave us interesting insights, that lead to the localisation of the problem and finally to a new component that manages to address this issue.

The before mentioned peculiarity concerns the high correlation between the location and the size of a forecasted PV streamer. In the ECMWF ensemble forecast of Medicane Zorbas, forecasted PV streamers that were displaced farther from the reference also were larger in size. This correlation can be seen in figure 12a) that shows the relationship between the cyclone location error and a component L_s , that evaluates size difference (details on the component can be found in the appendix). In figure 12bc the averaged PV fields for low (12b) and high scores of L_s (12c) are shown. A high score relates to a high size difference between the objects in the forecast and reference. These figures illustrate the correlation between size and location well. Interestingly there is a third element in this relationship: The direction. Not only are all the large forecasted PV streamers displaced the farthest, they are also all displaced to the east. This will be discussed more in the following section.

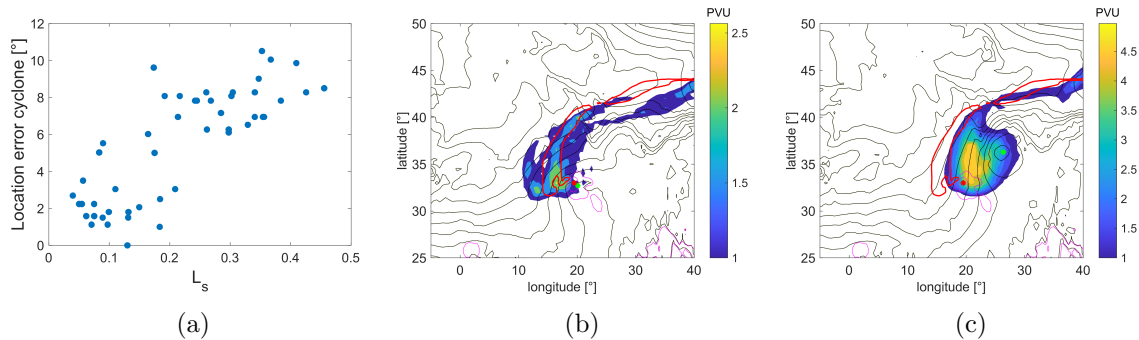


Figure 12: a) Relationship between size component and cyclone location error and average PV and SLP fields for members with b) low size difference and c) high size difference.

3.3.3 Eastern component

In order to address the problem of the overlap component, we need to understand why it occurs. Conceptually the component makes a lot of sense and there is valuable information gained from an L_o score. Especially for low scores, the component matches the visual evaluation of the PV streamer locations and forecasted cyclones also show a low location error. It is only cases with a higher score in L_o where the relationship is not clear anymore. But as figure 11 already showed, even though the L_o -component is not able to differentiate between cases of high and low cyclone location error, there is a clear distinction between these members. Visually they do describe a different situation, where we also would expect a different cyclone location. So we need a component that is able to capture this difference. The most obvious difference is the size, but since this might be case dependent this is not a reliable characteristic. In figure 13 we have an example of two members, 14 and 37, that got almost the same score for L_o (0.8950 and 0.9031), but a great difference in cyclone location error.

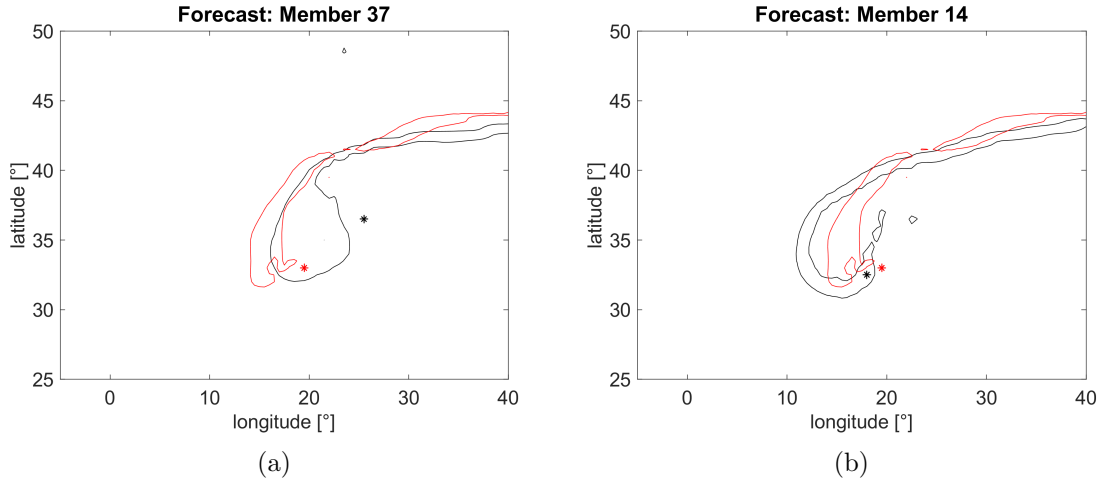


Figure 13: 2 PVU contours for members 14 and 37 (black), 2 PVU contour of reference PV field (red) and corresponding cyclone locations (star in same colour).

They both overlap very little with the reference PV streamer, but the potential for a good cyclone location is much greater in member 14. This is because it is also important what part of the PV streamers overlap or are in a close location, as well as in what direction the streamer is displaced. Figure 14 shows a sketch that shows the importance of the displacement direction. The outline of the reference PV streamer is shown in black, while the forecasted object is outlined in red. The two situations are exactly the same in terms of size difference, the distance of the displacement and the amount of overlap. The only difference is that in one case the object is displaced to the west, and in the other to the east. Their corresponding cyclones are denoted with a star of the same colour. Because the cyclone generally forms to the east of the PV streamer, in the case of a displacement to the west the location error is much smaller than in the case with an eastward displacement.

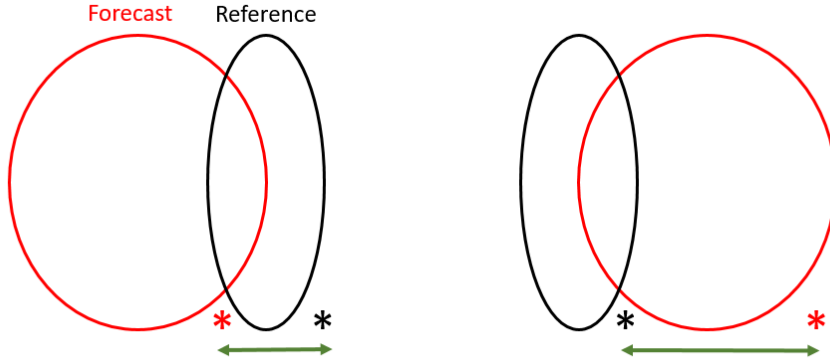


Figure 14: An illustration showing two cases of a reference (black) and forecasted (red) PV streamer pair with the location of their cyclones (depicted with stars of same colour).

In order to differentiate between these cases we need a component that takes east- or westward orientation into account, which will be done with the eastern component L_e . This component selects the eastern part of the PV streamer from the forecasted and the reference fields and evaluates their distance. A large streamer, such as shown in the sketch in figure 14 would therefore receive a worse score for the case on the right, as the eastern parts of the two streamers are farther apart.

The way the component works is as follows: For the selection of the eastern part of the cut off, first the cut off must be selected separately from the streamer, so that the more eastern parts inside the streamer are not selected as well. This is done by selecting grid points only up to a maximum latitude, based on the distribution of grid points above the threshold within the domain. In our method we use the median of all latitudinal coordinates of the grid points above a threshold of 2 PVU. From the remaining grid points individual objects are identified and only the one with the largest amount of grid points will be selected. From this object only grid points above a minimum longitude are selected, which is the median of the longitudinal coordinates from the remaining grid points. Here a higher quantile could be used instead of the median, but with the median the selection is more robust to peculiar eastern borders especially of thin PV objects. Figure 15 shows which parts of the PV streamers would be selected for the example cases of members 14 and 37 and the reference.

In order to evaluate the distance in these eastern parts the original L-component is used, which uses the difference in centre of mass. Since we are left with only one object in every case, in fact only L_1 is used. Even though we rejected the L-component as a measure for location of the PV field, it can still be used in the context of the eastern component, since we are only left with a small part of the PV field, and therefore got rid of most that was causing the problems in the first place, like high values inside the

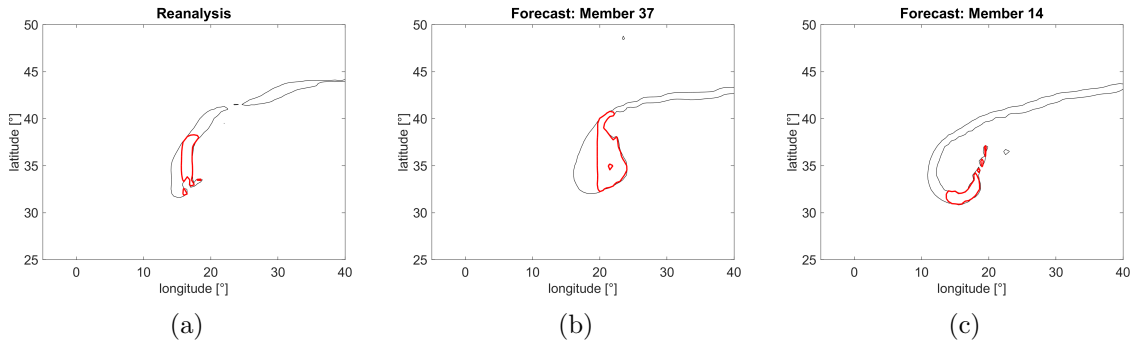


Figure 15: 2 PVU contours for reference and members 14 and 37 (black) and contour of remaining grid points after selection for L_e -component (red).

streamer and other grid points not belonging to the PV cut off.

Figure 16 shows the same plot as figure 11, which is the relationship between the L_o -component and the cyclone location error, but with the addition of the L_e component shown in colour. This plot shows that the L_e -component is able to differentiate between the two cases of high and low location error for members with a high L_o score.

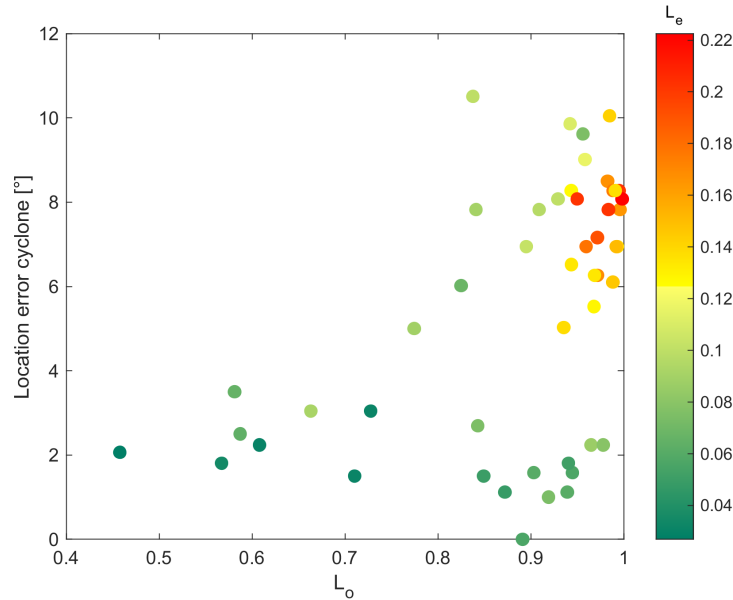


Figure 16: Relationship between the L_o -component and the cyclone location error with the L_e -component shown in colour.

In figure 17 the relationship between the L_e -component and the cyclone location error is shown. There is a correlation between them (correlation coefficient: 0.7111), however the relationship is not a perfect one. The worst cases in terms of cyclone location error are not the cases that receive the highest L_e score. Also shown are the averaged PV fields for the five members with the highest location error (figure 17b) and the five

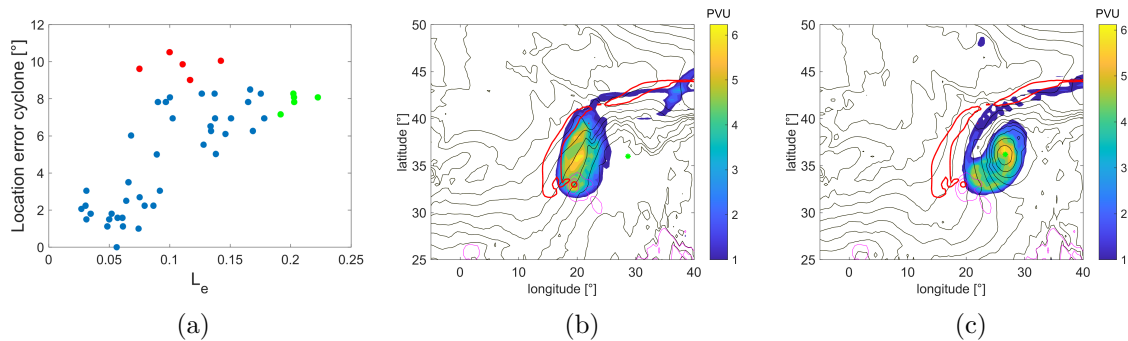


Figure 17: Relationship between L_e -component and cyclone location error (a) and averaged PV fields of members indicated in red (b) and green (c).

members with the highest score for the eastern component (figure 17c). In these plots we can see that the members with the worst cyclone locations show PV streamers that are located nearer to the reference streamer than in the members with the highest L_e score. Therefore the relationship between the location of the PV streamer and the location of the cyclone is not a perfect one to begin with and cases with a visually better location for the PV streamer can still produce cyclones that are displaced farther away. In figure 18 there are some examples of how the L_e -component works when applied to individual cases. Members 37 and 14 (figure 18ab) received a similar L_o score (difference of 0.0081) but member 14 predicted a cyclone with only a small error to the reference, while the cyclone in member 37 is displaced far to the east. This is due to the fact that in member 14 the eastern part of the PV streamer is located in the vicinity of the eastern part of the reference, even though the two structures overlap very little. The scores for the L_e -component reflect that by giving a low score for member 14 (0.0608) and a higher score for member 37 (0.1024).

A very similar situation is presented in figure 18cd. Members 12 and 36 also got very similar scores for L_o (difference: 0.003), but receive different scores for the L_e component (0.128 and 0.0521 respectively), since member 12 is larger and displaced to the east which means the eastern part of the streamer is also displaced farther away from the reference than in member 36. In member 36 on the other hand the displacement happened to the west and the eastern part of the streamer is still located in the vicinity of the reference, which also lead to the cyclone being forecasted near the reference cyclone. This shows that the L_e -component is a useful tool in differentiating between the cases that would lead to a lower or higher location error for the forecasted cyclone. Figure 19a shows the relationship between the L_e - and the L_o -component. This plot makes it clear that for low L_o values - which means a high amount of overlap between the forecasted and the reference PV field - the values for L_e are also low, which is to be expected. For L_o values higher than 0.85 the L_e -component starts differentiating between the cases that have a low amount of overlap but the eastern part of the streamer is still located close to the eastern part of the reference and the cases where there is a low amount of overlap as well as a large distance between the eastern parts of the PV

streamers, where we can expect the cyclone to be displaced farther away.

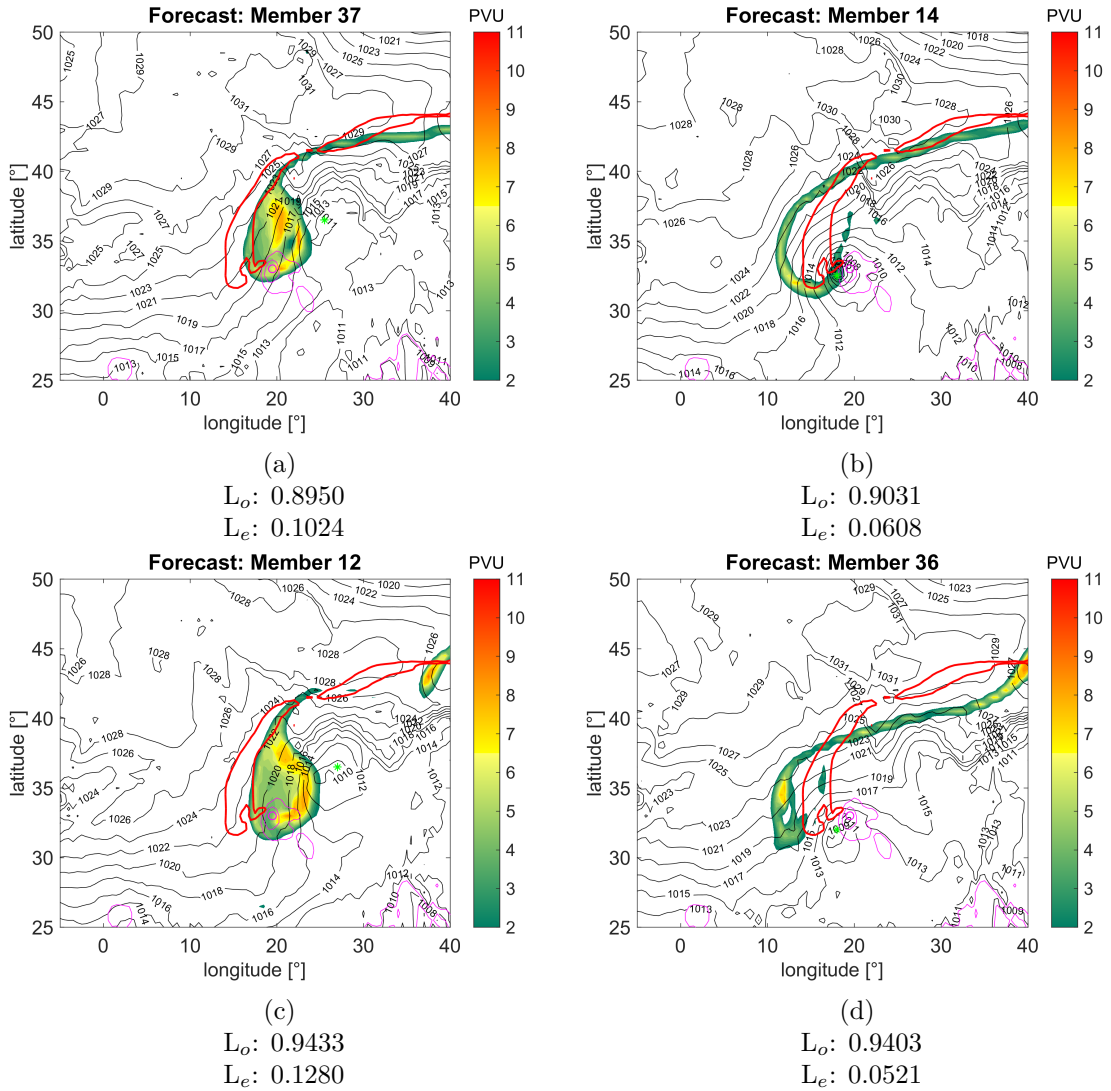


Figure 18: Examples of members where the L_e -component makes a distinction. PV fields above 2 PVU (colour), outline of reference PV field at 2 PVU (red), SLP contours (black), lowest SLP contours of the reference (magenta), cyclone location (green star) and reference cyclone location (red circle).

The plot in figure 19b shows the relationship between the L_e - and the L_c -component. The plot consists of two main clusters: A cluster of members with both high values for L_e and L_c , and a cluster for members with low values for both. In figure 19cd the averaged PV fields for both of these clusters are shown. They illustrate again the correlation between the eastern displacement and the size difference of the PV objects. There are two members belonging to the cluster with low values that receive a score for L_e , that is higher than for the rest of the members belonging to the cluster. They

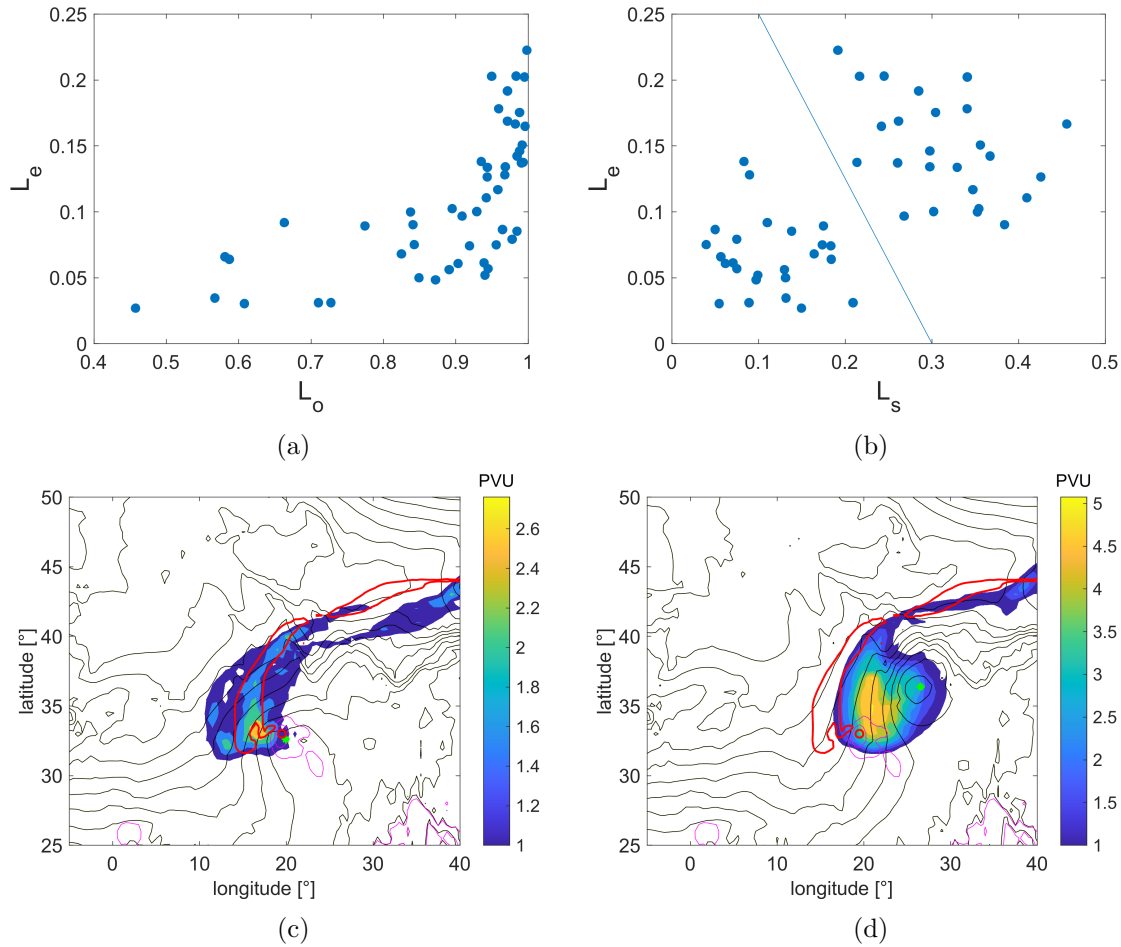


Figure 19: Relationship of L_e with L_o (a) and L_s (b) and the averaged PV fields for the indicated regions in the plot (cd).

receive a higher L_e score because these are two members with a displacement to the west farther than most other members, which makes their L_e values higher as well (members 19 and 31, also seen in figure 21).

Figure 20 shows a plot depicting the distribution of the three remaining components: The A-, L_o - and L_e -component. Figure 20b shows the whole possible range for A and L_o , while 20a shows only the part of the range where there are values. Members that receive low scores in all three components are located near the zero line for A and to the left of the plot and are dark green in colour. In this plot we see again, that A was mostly positive and close to zero. The L_o -component on the other hand shows hardly any members with very low scores, because it is hard for a real PV field to get an overlap that is close enough to get scores this low. For low L_o values, A and L_e are also relatively low. The L_e -component shown in colour illustrates that only members with a high L_o score also get high L_e score, which was already shown before. In relation to the A-component this plot shows that high L_e scores are distributed over a range

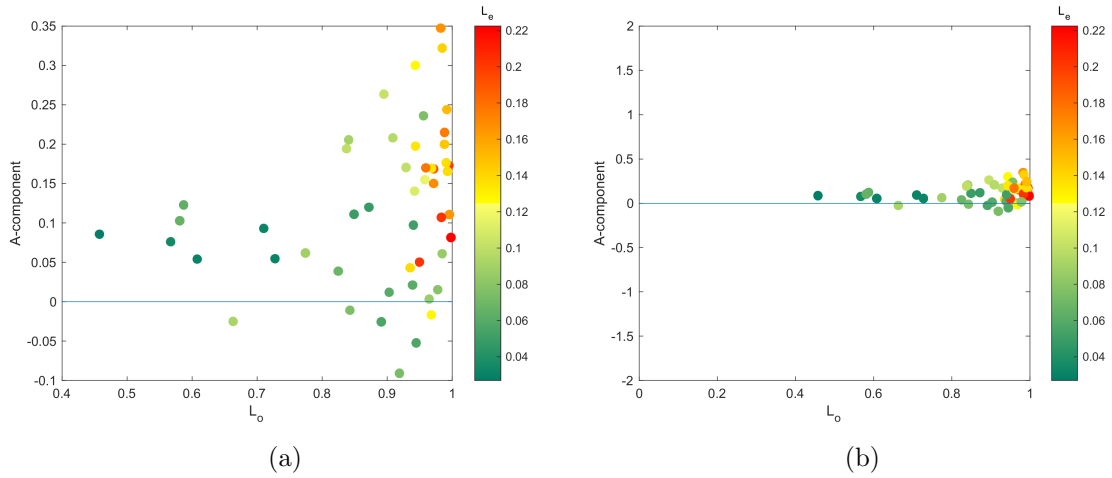


Figure 20: Plot for the three remaining components: A, L_o and L_e for the 50 members of the ECMWF forecast for medicane Zorbas. On the right (b) the whole possible range for A and L_o are depicted.

of different A values. Figure 21 shows three members that received high scores for L_e but low scores for A. Member 32 is an example where the A-component was low even though it was larger in size. Members 19 and 31 are the two members mentioned before that show higher L_e scores even though they are displaced to the west, because their displacement is larger than for other members with westward displaced PV streamers.

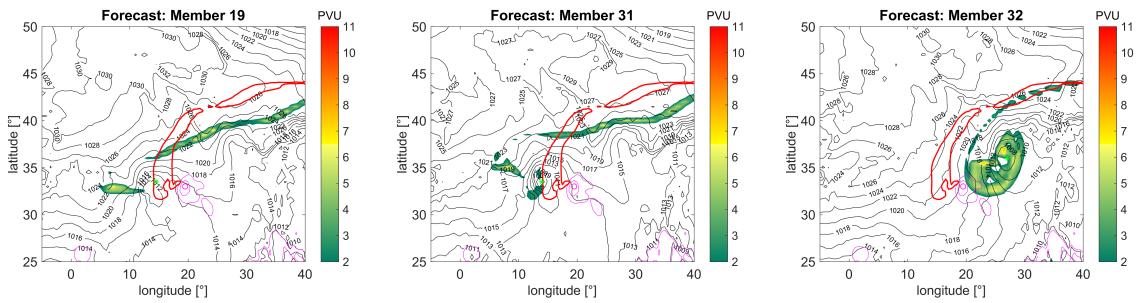


Figure 21: PV fields above 2 PVU (colour), outline of reference PV field at 2 PVU (red), SLP contours (black), lowest SLP contours of the reference (magenta), cyclone location (green star) and reference cyclone location (red circle).

4 Conclusions and Outlook

In this thesis the application of the SAL method, originally designed for QPFs, was tested on ensemble forecasts of PV fields of Medicane Zorbas from September 2018. These tests revealed that only one of the three components did work as expected and without problems: the A-component, which evaluates the amplitude of forecasted objects. The S-component was designed specifically to evaluate the structure of precipitation forecasts. Especially the steepness of the inner structure is a characteristic that is not as important in PV streamers. Applied to the PV field the component mostly evaluates the size of the objects. Since there is a similar transition from low to high PV values with a PV streamer, the A-component is also an indirect measure for size. The S-component is therefore not needed, especially because it also suffers from the camel effect when applied to PV. Lastly, the L-component did not seem to work as expected at all. Visual comparison of the PV fields with their respective results did not match. Also the component had no relationship with the predicted cyclone location, even though from visual judgement the relationship seemed evident. Therefore the SAL method, as it is originally defined, can not be applied to evaluate PV fields, with the exception of the A-component.

Therefore two other components were defined, that showed promising results to evaluate the location of forecasted PV streamers. The first one, L_o , measures the overlap between the reference and forecasted objects. This method worked in terms of matching a visual judgement of the PV field and therefore provides valuable information. But it has to be kept in mind, that it is hard for a model to produce a forecast that gets a very low score for this component, especially for higher resolutions. In order to get scores close to zero most of the grid points belonging to an object need to be forecasted perfectly. The L_o -component did not succeed in reflecting the relationship between the PV streamer location and the cyclone location, which is why a second component is needed. The problem with the cyclone location is, that an important aspect of whether the location of a PV streamer might produce a cyclone with a low location error is the direction in which the streamer is misplaced in the forecast. The same forecasted PV streamer has a different potential of producing a reasonably well forecasted cyclone if it was displaced to the east or west. Cyclogenesis usually happens to the east of the PV streamer. Since temperature decreases to the north (on the Northern Hemisphere) the cyclonic wind field induced by the upper-level PV streamer causes a positive temperature anomaly on the eastern side of the streamer. This anomaly in turn induces a cyclonic wind field at the surface. A PV streamer that was forecasted too large and is displaced to the east, will likely produce a cyclone that is displaced farther away than if the same PV streamer were displaced to the same amount to the west. The second new component, the eastern component L_e , reflects this directionality by evaluating the distance between the most eastern part of the PV streamer. With this method only the displacement of the eastern part of the PV streamer is evaluated, which seems to have the most impact on the cyclone location. It was shown that the L_e -component is able to capture the relationship between the location of the PV streamer and the cyclone location.

The relationship between the location of the forecasted PV streamer and the forecasted cyclone was expected and also evident from a visual perspective. The original L-component however, was neither capable of adequately assessing the location error of the PV streamer, nor could the relationship with the cyclone location error be shown. With the introduction of the two new components L_o and L_e both of these shortcomings were addressed and the relationship between the PV streamer and cyclone location could be quantified. It was also shown that the relationship is not strictly linear, since there are several cases with a better PV streamer location but worse cyclone location, which indicates that other factors than the PV streamer location can influence the cyclone location.

This study only evaluated the forecast of one point in time. The particular point in time that was compared, was the moment of cyclogenesis in the reference. Therefore this study concentrated on evaluating the model's capacity of producing the PV streamer and the cyclone only for this particular time. This is sufficient for testing the applicability of the SAL method, since for the evaluation of the PV field only one point in time is investigated, as opposed to QPFs where accumulated precipitation values over time are evaluated, which would not make sense for PV forecasts. However, for investigating a relationship between the PV streamer and the cyclone, it might be necessary to consider a longer period of time. This study only concentrated on one case. In a future study further testing of the components on different cases could be done. The L_o -component is defined using the number of grid points as a variable and could be sensitive to resolution. For the L_e -component more tests are also recommended, since it depends on an arbitrary definition. Grid points are selected based on a latitude and longitude determined by median values of the PV streamer's location. This selection method was defined to fit our case, but it could behave differently on another case. With other PV streamers the quantile for selecting grid points might have to be adapted or the selection method refined to represent more closely the form of the PV streamer.

Acknowledgements

I acknowledge the ECMWF for providing ERA5 reanalysis data and the operational IFS ensemble forecast data. I want to thank my supervisors Emmanouil Flaounas and Franziska Aemisegger for their great support. Thank you for the interesting discussions, your valuable input and feedback and that you were always ready to answer my questions.

References

- Casati, B., Wilson, L., Stephenson, D., Nurmi, P., Ghelli, A., Pocerlich, M., Damrath, U., Ebert, E., Brown, B., and Mason, S. (2008). Forecast verification: Current status and future directions. *Meteorological Applications: A Journal of Forecasting, Practical Applications, Training Techniques and Modelling*, 15(1):3–18.
- Ertel, H. (1942). Ein neuer hydrodynamischer Wirbelsatz. *Meteorologische Zeitschrift*, 59(9):277–281.
- Flaounas, E., Raveh-Rubin, S., Wernli, H., Drobinski, P., and Bastin, S. (2015). The dynamical structure of intense Mediterranean cyclones. *Climate Dynamics*, 44(9-10):2411–2427.
- Hoskins, B. J., McIntyre, M., and Robertson, A. W. (1985). On the use and significance of isentropic potential vorticity maps. *Quarterly Journal of the Royal Meteorological Society*, 111(470):877–946.
- Inness, P. M. and Dorling, S. (2012). *Operational weather forecasting*. John Wiley & Sons.
- Jolliffe, I. T. and Stephenson, D. B. (2012). *Forecast verification: a practitioner’s guide in atmospheric science*. John Wiley & Sons.
- Lionello, P., Bhend, J., Buzzi, A., Della-Marta, P., Krichak, S., Jansà, A., Maheras, P., Sanna, A., Trigo, I., and Trigo, R. (2006). Cyclones in the Mediterranean region: Climatology and effects on the environment. In *Mediterranean*, volume 4 of *Developments in Earth and Environmental Science*, pages 325 – 372. Elsevier.
- Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141.
- Murphy, A. H. (1993). What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather and forecasting*, 8(2):281–293.
- Petterson, S. (1956). *Weather Analysis and Forecasting: Volume I: Motion and Motion Systems*. McGraw-Hill.
- Portmann, R., González-Alemán, J. J., Sprenger, M., and Wernli, H. (2020). Mediane Zorbas: Origin of an uncertain potential vorticity streamer and impact on cyclone formation. *Weather and Climate Dynamics Discussions*, Under review.
- Rossa, A., Nurmi, P., and Ebert, E. (2008). Overview of methods for the verification of quantitative precipitation forecasts. In *Precipitation: Advances in Measurement, Estimation and Prediction*, pages 419–452. Springer.

Rossby, C. (1939). Planetary flow patterns in the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 66:68–87.

Wernli, H., Paulat, M., Hagen, M., and Frei, C. (2008). SAL—a novel quality measure for the verification of quantitative precipitation forecasts. *Monthly Weather Review*, 136(11):4470–4487.

A

Modifications of the L-component

This section gives more detail about the modifications that were applied to the L-component. The aim was to modify the score in a way that matches the visual interpretation of the PV fields, as well as the ability to differentiate between cases with different errors in cyclone location.

The first approach is weighting the PV field differently. The L-component is calculated by taking the centre of mass of the PV field. For this, all grid points within the objects are weighted with their PV values. The centre of mass however, is affected by parts of the PV field that are not of great interest in the context of cyclone development. For this reason an approach was tried where PV values are weighed exponentially, in order to give more weight to high PV values. Figure A.1a shows the resulting relationship between an exponentially weighted L-component and the cyclone location error. There is no correlation (correlation coefficient: 0.0366), which is due to the cases shown in red. These are members that got a high L score despite having a low cyclone location error. In figure A.1bc one of these members, as well as the reference field are shown. While the centre of mass for the reference field is located nearer to the PV cut off due to exponential weighting, for these members in red the centre of mass is rather dragged farther away due to very high values within the PV streamer to the east. If all these cases were left out the correlation would be 0.5575 and could be promising but the method is too susceptible to cases with high PV values outside the PV cut off.

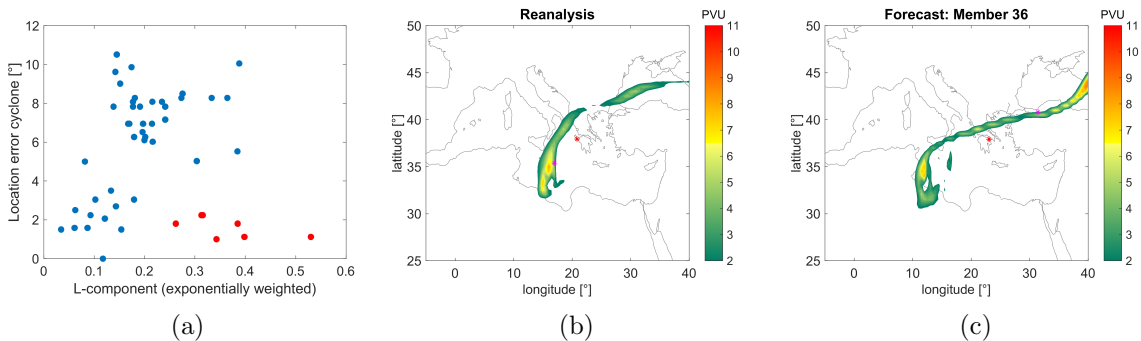


Figure A.1: Relationship between cyclone location error and exponentially weighted L-component (a) and PV fields above threshold of 2 PVU, centre of mass (red) and centre of mass for exponentially weighted PV field above 2 PVU (magenta).

The second approach is a varying of the threshold. This was done by defining the threshold based on quantiles. In the reference PV field the threshold of 2 PVU corresponds to a quantile of 0.9685, since 96.85 percent of all values within the domain are above 2 PVU. Figure A.2 shows how the correlation between the L-component and the cyclone location error behaves depending on the choice of threshold. The plot shows that the highest correlations can be achieved by low thresholds at a quantile of 0.8 or

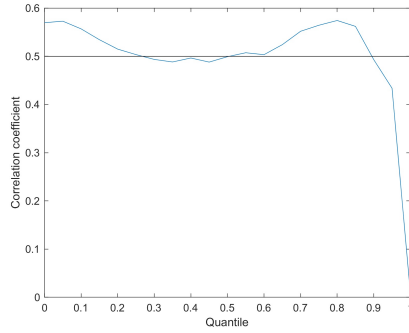


Figure A.2: Correlation coefficients for thresholds depending on quantiles.

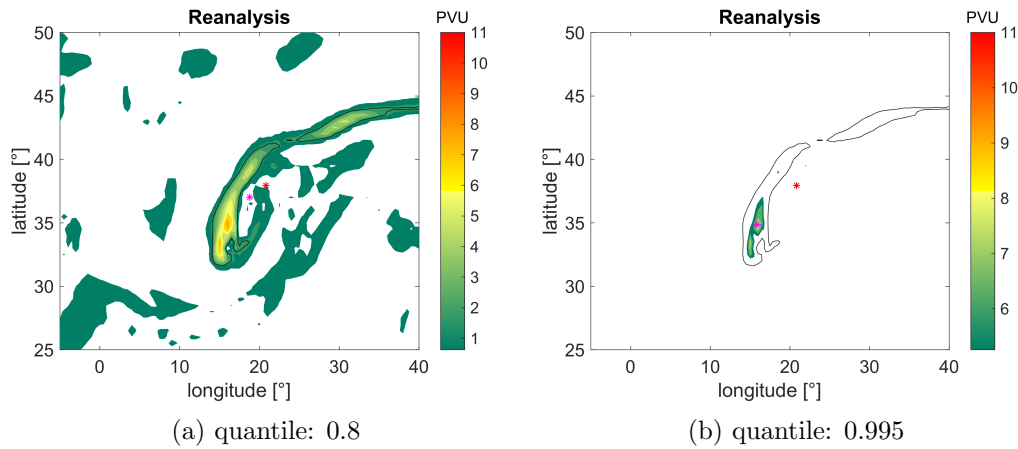


Figure A.3: PV fields above a threshold defined by a quantile, outline of the 2 PVU line (black), centre of mass (magenta) and centre of mass for the PV field above threshold of 2 PVU (red).

near zero. A quantile of 0.8 corresponds to a threshold of 0.6261 PVU for the reference field.

In figure A.3 as an example the reference field is shown for the threshold at the 0.8 quantile. Because the threshold is so low there are many low PV objects distributed over the domain. The centre of mass for these thresholds is always located more towards the centre of the domain compared to fields with a threshold at 2 PVU, because of the more balanced distribution of objects. The better correlation is therefore not due to a justified reason and this method not recommended. Thresholds higher than 2 PVU lead to an effect similar to the exponential weighting. Instead of weighting high values more, only high PV values are taken into account. Figure A.3b shows the reference field with a high threshold and centre of mass displayed.

Size difference component

In section 3.3.2 a component was mentioned that measures size difference, but was discarded. In this section this component will be explained with more detail. The size difference component L_s was developed with the purpose of having a component that could differentiate between the cases that resulted in a high cyclone location error from those with a low error. It was designed as an additional component to the overlap component L_o and as a combined component L_n it would additionally penalize cases where the objects overlap but there is a large size difference between them. The L_s -component is calculated analogous to the L_o -component, with the difference that instead of the overlapping grid points, the number of grid points of the forecasted object is in the nominator:

$$L_s = \left| \frac{n_{for}}{0.5[n_{ref} + n_{for}]} - 1 \right| \quad (10)$$

Where n_{ref} and n_{for} again stand for the number of grid points within objects in the reference and the forecasted field respectively. This component also ranges from $[0 1]$, where 0 means a perfect forecast, which in this case means that the objects contain exactly the same number of grid points over the threshold.

The combined component L_n is calculated from the overlapping and the size difference component: $L_n = L_o + L_s$. This component would range between $[0 2]$, with 0 meaning a perfect forecast in terms of size and overlap. For the ideal cases in figure 10 all the cases in the bottom row would get even higher scores than the ones from the top row, because of the additional focus on size difference. Figure A.4a shows the resulting relationship with the cyclone location error. Again the results are shown for the average over thresholds 1.5 to 5 PVU. This new component seems to be able to do the desired differentiation better, since cases with a high cyclone location error also get a high value for the L_n component and likewise cases with a low cyclone location error get low values. It is only the cases somewhere in the middle where there is still some ambiguity. However, the differentiation is only due to the size difference. In figure 12a the relationship of only the L_s -component with the location error with a clear correlation (correlation coefficient: 0.7995) was shown. From the averaged PV fields in figure 12bc we see that this clear relationship comes from the fact that all of the cases where there is a too large forecasted PV streamer, are also cases where the streamer is displaced far to the east. This is not a relationship we can rely on, since there might exist other cases where we do have a larger forecasted streamer that is not displaced or the other way around. Another problem is that the scores for the combined L_n -component become too ambiguous with higher scores. Especially scores over 1 could mean a total overlap but with some difference in size or no overlap at all, which could be considered the worst case, no matter what size difference. Also with the combination of the two components information gets lost.

For these reasons the L_n -component was discarded. It is also questionable if we need a component for the size difference, since the L_s -component correlates so well with

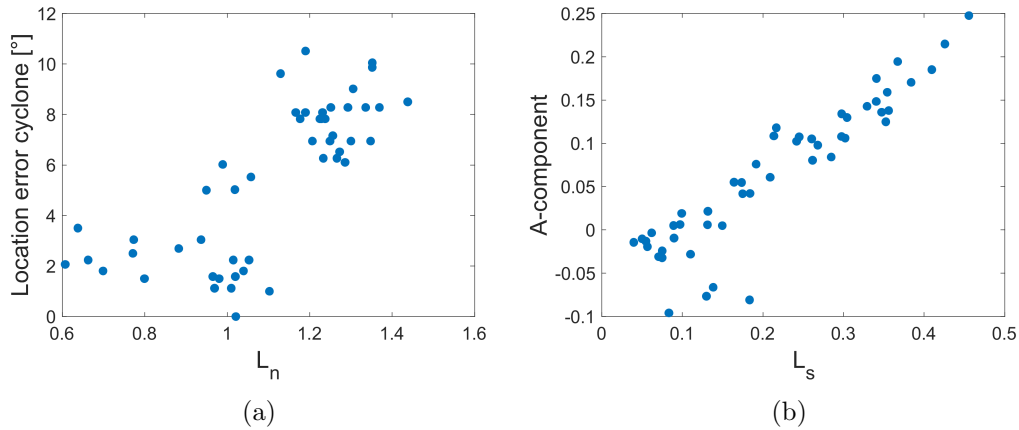


Figure A.4: Relationship between L_n -component and cyclone location error (a) and the L_s -component and the A-component (b).

the A-component (figure A.4b, correlation coefficient: 0.916). As already discussed in an earlier section, the A-component correlates well with size. The larger a PV streamer is, the larger can the maximum values get and therefore the larger is the averaged value, which leads to higher scores for the A-component. So naturally the A-component also correlates well with a component that evaluates size difference. An additional component for size is not needed and since the A-component also contains the information whether the reference or the forecasted field was larger, the L_s -component was also discarded.

B

Relationship between the PV streamer and cyclone intensity

We expected to find a correlation between the intensity of the PV streamer and the intensity of the forecasted cyclone. A forecast with too high PV values within the streamer would lead to a deeper cyclone than observed. The A-component however, did not show a correlation with the intensity error of the forecasted cyclone. On the contrary, most members showed a positive A value, meaning an overestimation of PV values, while simultaneously predicting a weaker cyclone, which can be seen in figure B.1a. Figure B.1b shows the average PV field for cases corresponding to the box in the top right of panel a. These were cases where the PV streamer was large and displaced far to the east, the cyclone intensity was generally less deep, which was against our expectations. This is not due to the fact that in this thesis only one point in time was considered, even though cyclogenesis did not occur in all members at the same time. According to Portmann et al. (2020) the members with far east displacement of the PV streamer on average showed the weakest cyclone, also when the whole cyclone development was considered.

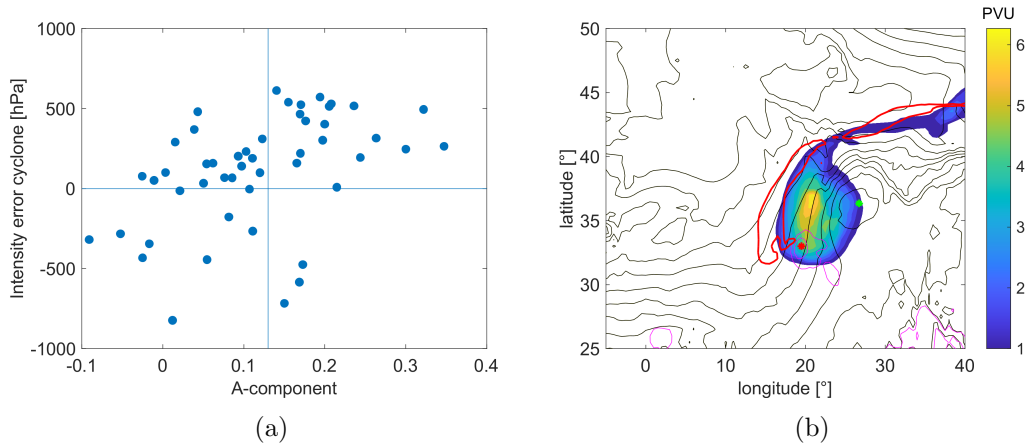


Figure B.1: Relationship between A-component and cyclone intensity error (a) and the average PV and SLP field corresponding to the top right box (b).



Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):

First name(s):

| | |
|--|--|
| | |
| | |
| | |
| | |

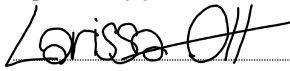
With my signature I confirm that

- I have committed none of the forms of plagiarism described in the ['Citation etiquette'](#) information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

Signature(s)

| | |
|--|--|
| |  |
| | |
| | |
| | |

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.