# A novel method for objective identification of 3-D potential vorticity anomalies

Christoph Fischer[1,2], Andreas H. Fink[2], Elmar Schömer[1], Roderick van der Linden[2], Michael Maier-Gerber[2], Marc Rautenhaus[3], and Michael Riemer[4]

[1]Institute of Computer Science, Johannes-Gutenberg University, Mainz, Germany
[2]Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, Karlsruhe, Germany
[3]Regional Computing Centre, Visual Data Analysis Group, University of Hamburg, Hamburg, Germany
[4]Institute for Atmospheric Physics, Johannes-Gutenberg University, Mainz, Germany

**Correspondence:** Christoph Fischer (christoph.fischer@uni-mainz.de)

**Abstract.** Potential vorticity (PV) analysis plays a central role in studying atmospheric dynamics and in particular in studying the life cycle of weather systems. The three-dimensional (3-D) structure and temporal evolution of the associated PV ~~anomalies~~features, however, are not yet fully understood. An automated technique to objectively identify 3-D PV ~~anomalies~~ features can help to shed light on 3-D atmospheric dynamics in specific case studies, as well as facilitate statistical evaluations within climatological studies. Such a technique to identify PV ~~anomalies~~ features fully in 3-D, however, does not yet exist. This study presents a novel algorithm for the objective identification of PV anomalies along the dynamical tropopause in gridded data, as commonly output by numerical simulation models. The algorithm is inspired by morphological image processing techniques and can be applied to both two-dimensional (2-D) and 3-D fields on vertically isentropic levels. The method maps input data to a horizontally stereographic projection and relies on an efficient computation of horizontal distances within the projected field. Candidates for PV anomaly features are filtered according to heuristic criteria, and feature description vectors are obtained for further analysis. The generated feature descriptions are well suited for subsequent case studies of 3-D atmospheric dynamics as represented by the underlying numerical simulation~~, or for generation of climatologies of feature characteristics~~. We evaluate our approach by comparison with an existing 2-D technique, and demonstrate the full 3-D perspective by means of a case study of an extreme precipitation event that was dynamically linked to a prominent subtropical PV anomaly. The case study demonstrates variations in the 3-D structure of the detected PV anomalies that would not have been captured by a 2-D method. We discuss further advantages of using a 3-D approach, including elimination of temporal inconsistencies in the detected features due to 3-D structural variation, and elimination of the need to manually select a specific isentropic level on which the anomalies are assumed to be best captured. These advantages, as well as the suitability of the implementation to process big data sets, also open applications for climatological analyses. The method is made available as open-source for straightforward use by the atmospheric community.

# 1 Introduction

Weather systems and extreme weather events result from nontrivial three-dimensional (3-D) interactions in the atmosphere. The potential vorticity (PV) perspective on atmospheric dynamics provides an often-used conceptual framework to understand such interactions (as reviewed by Hoskins et al., 1985). A key component of this framework is that PV is materially conserved in the absence of nonconservative processes ~~, a reasonable approximationfor large-scale atmospheric flow.~~ such as, e.g., latent heat release in clouds, longwave radiative cooling, or turbulent mixing. In the absence of strong latent heat release, it turns out that nonconservative PV modification is comparatively small and that material conservation of PV provides a very good first approximation. When considered on isentropic levels, the temporal evolution of PV at a given location is hence governed by quasi-horizontal advection only, which typically yields a rather smooth PV evolution.

Of particular importance to the PV perspective is the tropopause, as it separates air masses with typically low PV values in the troposphere from typically high-PV air in the stratosphere[1]. A strong PV gradient across the tropopause has motivated the definition of the so-called dynamical tropopause as an upper-level PV isosurface in the literature. Typically used PV values range between 1.5 - 4 PVU, with PV = 2 PVU being an often-used standard value (e.g., Morgan and Nielsen-Gammon, 1998).

The tropopause slopes from higher isentropic levels in the tropics towards lower isentropic levels at the poles. On a given isentropic level, poleward excursions of the tropopause denote an anomaly of low PV, whereas equatorward excursions denote an anomaly of high PV. A relatively smooth meridional undulation of the tropopause signifies quasi-linear Rossby wave dynamics. In the context of extreme weather, however, much attention has been given to highly non-linear PV features, in particular zonally narrow and meridionally extended "tongues" of high PV that intrude equatorwards (e.g., Massacand et al., 1998), known in the literature as PV streamers (PVSs). These elongated anomalies can have complex life cycles and behavior: They are often linked to Rossby wave breaking (RWB) events (e.g., McIntyre and Palmer, 1983; Thorncroft et al., 1993), interact with surface cyclones (e.g., Galarneau Jr et al., 2015; Bentley et al., 2017; Maier-Gerber et al., 2019), and may detach from the main stratospheric high-PV reservoir on a given isentropic level, forming so-called PV cutoffs. These cutoffs can diabatically decay, be reabsorbed by the main reservoir, undergo further dynamical processes, and can be associated with extreme weather events (Portmann et al., 2021).

In the present study, we consider the objective identification PV structures. Objective identification of atmospheric features from numerical simulation output has proven beneficial for a number of applications in both atmospheric research and operational meteorology. Typical uses include statistical analysis such as the computation of climatologies of feature occurrence (e.g., Limbach et al., 2012; Dawe and Austin, 2012) and operational weather forecasting purposes (e.g. Hewson and Titley, 2010) including ensemble forecast analysis (e.g. Hewson and Titley, 2010; Rautenhaus et al., 2015a). Recently, combination of state-of-the-art interactive 3-D visualization techniques (Rautenhaus et al., 2018, provide a comprehensive survey) with 3-D detection of atmospheric features opened the door for comprehensive case studies of 3-D atmospheric dynamics (e.g., Rautenhaus et al., 2015a; Kern et al., 2018, 2019; Bader et al., 2019). With respect to objective identification of PV struc-

---

[1]PV is typically positive on the northern hemisphere and negative on the southern hemisphere. Low PV air here refers to low absolute values and high PV air to high absolute values. For notational convenience we consider the northern hemispheric situation only.

tures, Papin et al. (2020) recently provided an overview of identification techniques for PVSs. They classified the methods into techniques based on the reversal of the meridional PV gradient, and techniques based on distance thresholds along a 2-PVU contour. For example, a 2-D identification technique that has found wide acceptance and use, and may thus effectively serve as a state-of-the-art benchmark technique, has been introduced by Wernli and Sprenger (2007). This technique is based on a scale analysis, according to which the 2-PVU contour of a PVS contains points that are far apart in the direction along the contour, yet relatively close together regarding their great circle distance. Wernli and Sprenger (2007) applied this technique to generate a climatology from ERA-15 reanalysis data, which Kunz et al. (2015) later extended to a 33 year time period.

All techniques discussed by Papin et al. (2020) operate on 2-D data only. Furthermore, gradient-reversal techniques only consider a subset of PVSs (often specifically focusing on RWB events), which are sensitive to small changes in the input data. To comprehensively study the atmospheric dynamics associated with a weather event of interest, however, it is often important to consider the full 3-D structure of the atmosphere, including related PV features. For instance, Bithell et al. (1999) put effort into analyzing and visualizing the 3-D structure of atmospheric PV. They showed a wide spectrum of 3-D PV anomalies (PVAs, see below) that are difficult to analyze (or even to track over time) in 2-D vertical or horizontal (isentropic) cross-sections alone. They further demonstrated that the evolution of wave-related features along the tropopause may cover an extensive vertical range that is difficult to describe in the framework of a 2-D analysis. Therefore, Bithell et al. concluded that the 3-D perspective is a useful tool for the comprehensive study of the evolution of the dynamical tropopause and related weather systems.

We note that in 2-D analysis, the concepts of both PVSs and PV cutoffs have been widely established. They are clearly defined and widely used. However, a 3-D anomaly stretching over multiple isentropes can exhibit both types, PVS and PV cutoff, depending on the considered level. Therefore, this 2-D classification is challenging for evaluations, and especially does not hold for an analysis of 3-D features. In this study and for 3-D analysis, we hence name these features PV anomalies (PVAs). These anomalies can exhibit characteristics of both streamers and cutoffs when considered on a single isentropic level only. 3-D cutoffs, i.e. fully isolated areas of stratospheric air surrounded by tropospheric air, are only rarely a result of detachment from the main reservoir and mostly appear in the lower troposphere due to latent heat release (e.g., Bennetts and Hoskins, 1979).

An identification of 3-D PV structures has previously been performed for subsets of anomalies only, although the importance of their vertical structure has been demonstrated in the literature (e.g. Portmann et al., 2021; Lamarque and Hess, 1994). For example, Škerlak et al. (2015) investigated the 3-D structure of one subset of PVAs, namely tropopause folds and their link to extreme weather events. These folds are defined as structures with multiple tropopause crossings in the vertical direction. Portmann et al. (2018, 2021) focused on cutoffs that are identified independently on multiple isentropic levels and are subsequently stacked vertically using an overlap and proximity heuristic. Regarding a 3-D field, ~~cutoffs~~ these can be thought of as stalactites attached to the stratospheric PV reservoir. Portmann et al. (2021) found that cutoffs can be potentially complex 3-D features, which can intensify on higher levels while decaying on lower levels, resulting in a vertical (cross-isentropic) displacement of these features. Therefore, tracking PVSs on a chosen 2-D isentropic level is not always sufficient to analyze an event during its entire life cycle. As investigated by Bithell et al. (1999), stalactites and tropopause folds are only a subset of

possible anomaly types in a 3-D PV field. They classified the complex structures they encountered as *tubes*, *spirals*, *stalactites*, *folds*, among others, demonstrating complex scenarios that are not discernible within a 2-D perspective. Existing approaches used in 2-D identification lack the concepts of 3-D cohesiveness and information regarding extent, genesis and evolution of these structures. For instance, using the complete 3-D field offers the possibility to consider long time periods without the need to adjust a selected isentropic level to changes occurring between different seasons.

In summary, while many previous studies contributed to identifying different aspects of PV features on single isentropic levels, the objective identification of fully 3-D PV structures is still an open challenge and motivates our work. We expect that a 3-D identification method will yield the opportunity for novel analyses, including new aspects within climatological studies and comprehensive case studies of 3-D atmospheric dynamics. In the article at hand, we hence present a novel algorithm for the identification of PV anomalies, which can be applied not only to 2-D, but also to 3-D fields. In the latter case, the algorithm operates on the complete 3-D fields instead of on individual isentropic levels. To the best of our knowledge, this is the first approach aiming for a strategy that identifies the full spectrum of PVAs in 3-D. We use image processing techniques for the identification, more specifically numerical solutions that base on morphological operators. These operators, which originate from analytical geometry, have been adapted to suit the requirements of a meteorological application. This adaptation mainly revolves around using physical units to ensure interpretability, while considering the properties of the used projection. Parameters of the identification are adjustable and the algorithm is usable on different resolutions~~and projections~~. A low-dimensional and human-readable feature vector is computed for each identified PVA consisting of quantifiable measures, e.g., centroid, intensity or best-fit geometry. This memory-efficient feature representation can serve as a basis for statistical analyses or as potentially relevant predictors. The identified 3-D PVA features are also well suited for studying atmospheric dynamics by means of interactive 3-D visual analysis (IVA; cf. Rautenhaus et al., 2018). For example, reducing atmospheric processes of importance to concise visual depictions facilitates combination of multiple aspects of atmospheric dynamics in a comprehensive 3-D display well suited for rapid exploration of the considered numerical simulation data (e.g., Rautenhaus et al., 2015a; Kern et al., 2019). Here, we demonstrate such analysis by incorporating the identified PVA features into the 3-D meteorological visualization framework "Met.3D" (Rautenhaus et al., 2015b), which we use to shed light on the 3-D structure of PVAs encountered during an extreme precipitation event previously investigated by Van der Linden et al. (2017).

This article is structured as follows. Section 2 introduces the basic principle of the algorithm. In Sect. 3, a distance measure required for the method is introduced. The identification algorithm is described for 2-D data in Sect. 4, and compared to the method of Wernli and Sprenger (2007) in Sect. 5. Section 6 introduces the generalized 3-D algorithm, which we apply in Sect. 7 for a case study, discussing the potential of 3-D PVA analysis. The paper is concluded in Sect. 8.

## 2 Basic principle of the algorithm

The identification algorithm is motivated by so-called morphological image processing techniques. These operations, often used on binary data, process images based on the characteristics of their shape (Dougherty, 2020). Structuring elements (mostly matrices used as masks) are used to decide on how pixels in the image are altered based on their surroundings. Morphological
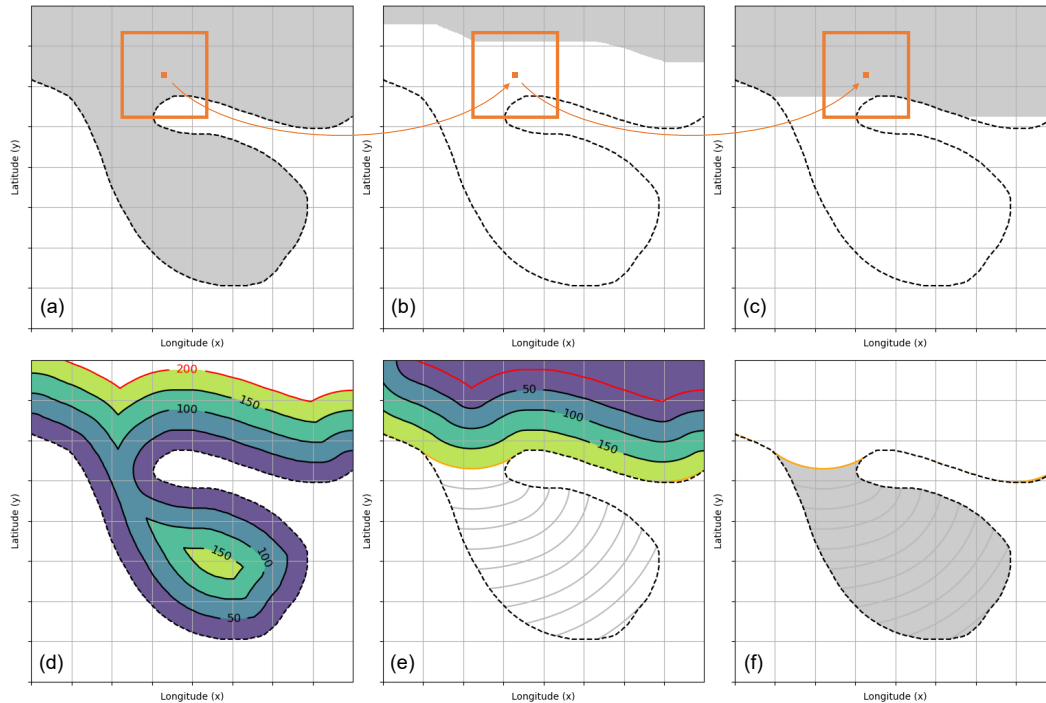
**Figure 1.** A sketch of our identification strategy applied to a PVS-like structure. (a-c): The process of a morphological opening (erosion followed by dilation) is applied to the gray structure using a square mask (orange). This removes thin structures depending on the mask's size. (d-f): Our adapted strategy, which uses distances instead of a binary mask. First, the distance emerging from the outer boundary is computed. Then, starting at a set distance (red line), the same distance is added back, respecting the boundaries of the domain. Areas not reconstructed are identified as anomalies.

operations are often used for morphological noise suppression or shape analysis. In the following, we first consider the 2-D case to illustrate our approach. Afterwards the method will be expanded to 3-D. To make use of morphological techniques, anomalies in the PV field (e.g., intrusions, cutoffs etc.) can be thought of as morphological "noise". A "noise-free" PV field then resembles an idealized state of the atmosphere, where the height of the tropopause is symmetrical around the poles.

125     Fig. 1a-c shows the process of a so-called morphological opening applied to a binary image, which could resemble a PVS. An opening is defined as an erosion followed by a dilation of the data. For the erosion (Fig. 1a-b), a structuring element (orange) is used as mask. This mask is shifted over the binary input image. Each pixel $(x, y)$ is kept as part of the domain only if all pixels in the structuring element surrounding $(x, y)$ are also part of this domain. Thus, the erosion removes all areas along edges depending on the size of the mask. As seen in the figure, it especially destroys elongated and filament-like structures.

130     Following the erosion, the dilation step (Fig. 1b-c) reconstructs areas of the domain. Using the same structuring element, a pixel $(x, y)$ is added back to the domain if any of the surrounding pixels defined by the structuring element is part of the
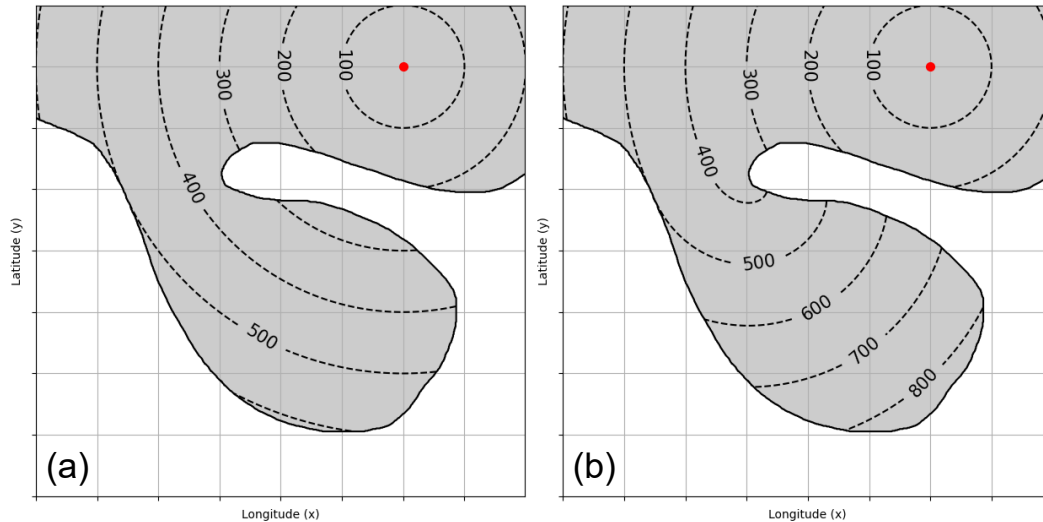
**Figure 2.** A visual example (same PVS-like structure as in Fig. 1) showing the shortest distance measure required for the algorithm. (a): Starting at the red dot, the direct spherical distance is shown, which can be calculated easily. (b): The distance following a well defined domain (e.g., PV field) is computed using a Fast ~~marching method~~ Marching Method (~~FFM~~FMM) as required by our strategy, see Sect. 3.2 for details.

domain. As seen in Fig. 1c, this reconstructs the general shape of the original object, but filament-like structures and other forms of noise are not reconstructed. These non-reconstructed areas can be extracted and interpreted as PVSs.

This approach uses an abstract concept of binary masks in discrete environments, and care is required to maintain inter-
135   pretability and to design the algorithm with meaningful parameters. Fig. 1d-e shows an adaptation using Euclidean distances instead of a mask. Masks are typically square, and results based on these masks would give a distorted distance measure. The use of Euclidean distances in this Cartesian field results in an uniform effect along all directions. The adapted strategy works as followed: First, we measure distances emerging from the boundary (e.g., tropopause) into the domain, as seen in Fig. 1d. Then, ~~we choose a distance~~ all points with a specific distance from the boundary based on a given parameter are extracted
140   (e.g., the red line in Fig. 1d ~~), and only keep areas as part of the domain that are further away from the boundary than the given parameter~~with a distance of 200). A comparable parameter for the upper part of the figure would be the size of the structuring element. Then, starting from the newly defined boundary (red line), distances are measured outwards. ~~Again, the result is at the given parameter~~ Points with the previously specified distance from the red boundary form a new boundary (orange line in Fig. 1e-f). Areas ~~further away than the value of the parameter~~ with a higher distance are identified as anomalies.
145   Using this concept in real-world environments unfortunately is non-trivial. On the one hand, the distance measure required must follow the domain instead of using a direct spherical distance, as seen in Fig. 2. Most algorithms that satisfy this require-ment suffer from metrical errors induced by the discrete nature of algorithms. Furthermore, the real world closely resembles
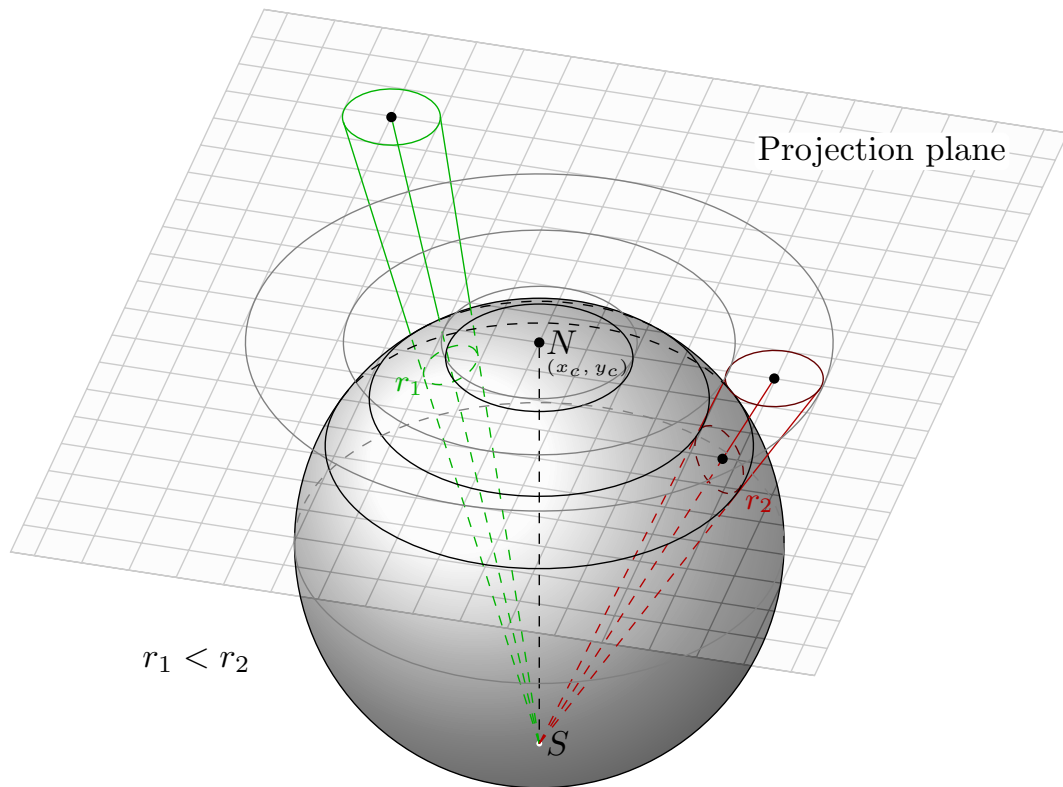
**Figure 3.** Depiction of the stereographic projection used in our identification strategy. This figure shows that circles on the sphere are mapped to circles on the projection plane. Therefore, this projection is conformal, a vital property for this strategy. Note that despite having the same shape, the sizes of the circles on the sphere vary by distance from the projection center.

a sphere and cannot be approximated using a Cartesian grid. Instead we have to consider the distortions of the projection our algorithm is working in. The next section describes how distances following the field (as in Fig. 2b) are computed.

## 3   Solving the distance measure problem

### 3.1   Stereographic projection

For the identification algorithm described in this study we choose a pole-centered stereographic projection (Fig. 3). Its point of projection is on the surface of the sphere opposite to the tangent plane where the projection is created. Thus, we can use the South Pole as point of projection to have the view centered around the North Pole. This creates a singularity at the south pole while minimizing the distortions in the northern hemisphere (Snyder, 1987).

We choose this type of projection due to several advantages:

1. The PV field encircles the earth and is centered around the poles. Using the pole as center of projection leads to a clear depiction of the features to be detected.

2. Contrary to equirectangular projections, this stereographic projection avoids singularities at the North Pole. PV intrusions over poles can occur, and their analysis is highly difficult in projections where a singularity is involved. Additionally, the stereographic projection solves issues handling the antimeridian boundary.

3. A stereographic projection is conformal, and the only known true perspective projection with this property (Snyder, 1987). A conformal projection preserves angles, thus, two lines on a sphere will intersect in the same angle as in their projection. As a result, infinitesimal shapes on a sphere are conserved in this projection. All Tissot's indicatrices are circles, therefore perfect circles on the sphere will always map to perfect circles in the projection, generally of a different area (see Fig. 3). This property will be useful to efficiently compute distances along the earth's surface regarding this projection later on. Note that map projections cannot be both area-preserving and conformal at the same time (Pearson, 1990).

The main drawback of this strategy is the introduction of a singularity at the projection center, i.e. the opposing pole. To avoid this problem while mapping global data, we create one stereographic projection centered around the North Pole and one centered around the South Pole, both extending towards the equator. The singularity at the pole shifts to a boundary at the equator. ~~The sign of PV changes~~ Since the dynamical tropopause changes its sign at the equator ~~, and therefore it is questionable to identify PVAs over this boundary anyways~~ and generally lacks proper definition at lower latitudes, this area is not considered as a region of interest for this present study.

## 3.2 Computing distances within the projection

As introduced in the previous section, our novel algorithm requires a way to calculate distances within the projection following a given field, as illustrated in Fig. 2. Here, we can use the conformal property of the projection. Since circles on the surface of a sphere are projected to circles in the stereographic projection, we can assign a distance field $\Delta s(x, y)$ for all points, resp. pixels $(x, y)$, in the projected field. Each pixel in the field $\Delta s$ contains the radius of a circle on the sphere in km that projects onto a circle of radius 1 pixel around $(x, y)$ in the projection. For example in Fig. 3, the center of green and red unit circle on the projection plane are assigned values $\Delta s(x, y) = r_i$, where $r_i$ is the radius of the corresponding circle on the sphere. Thus, based on the conformal properties, ~~this distance~~ $\Delta s(x, y)$ can be expressed as a scalar ~~value. The scalar field $\Delta s(x, y)$ then can field and~~ be calculated using just the distortion factor at every point. For a sphere, according to Snyder (1987), the distortion factor $h$ at a point $(x, y)$ in a stereographic projection can be calculated by

$$h = \cos(\frac{\theta(x, y)}{2})^{-2}, \tag{1}$$

where $\theta(x, y)$ denotes the angle from the center of the projection at $(x, y)$. $h$ denotes the scale factor in all directions. For example, projecting the northern hemisphere with the North Pole as projection center yields a factor of 1 at the pole ($\theta = 0$) as expected, and a factor of 2 at the equator ($\theta = \frac{\pi}{2}$). Towards the South Pole, $h$ becomes infinite. The distance field $\Delta s(x, y)$

can simply be calculated by the pixel radius in km at the projection center $(x_c, y_c)$ (e.g., the pole, see Fig. 3), divided by the distortion factor:

$$\Delta s(x,y) = \frac{\Delta s(x_c, y_c)}{h}. \tag{2}$$

$\Delta s(x_c, y_c)$ denotes the radius on the sphere that projects onto a circle on the projection plane centered around $(x_c, y_c)$ with a radius of one pixel. Since the pole is the projection center, the pole of the sphere touches the projection plane in $(x_c, y_c)$. By mapping one neighboring pixel of the pole to their regarding latitude and longitude coordinates, $\Delta s(x_c, y_c)$ can be obtained by computing the spherical distance between the pole and this neighboring pixel. Due to the conformal property of the projection, any of the neighboring pixels yield the same distance.

As next step, we define the distance map $u(x)$ containing the minimal distance from a well defined boundary $\Gamma$ to all $x$, where the distance $u(\Gamma)$ from the boundary to the boundary is zero. This problem is similar to the computation of signed distance functions or solving path planning problems. Popular algorithms to generate this field $u$ starting at $\Gamma$ contain variations of the Dijkstra algorithm (Dijkstra et al., 1959). Starting at the boundary, this construction is performed by evolving a level set to the destination based on a distance map (Petres et al., 2005).

However, most of these approaches suffer from metrication errors. They construct the path on segments parallel to the grid dimensions. Directions invariant to the grid orientation will always lead to errors induced by the $L_1$ norm (see Cohen and Kimmel, 1997). To get consistent results for the identification independent of the location of a PVS, these inaccuracies should be avoided.

This shortest path problem can also be formulated as path space integral. We search the shortest distance $u$ emerging from any point in the boundary $\Gamma$ to $x$ with respect to a given cost function $\Delta s$:

$$u(x) = \min_{P_{\Gamma,x} \gamma \in \Gamma} \int_{\gamma}^{x} \Delta s(z)\,dz. \tag{3}$$

Generally, this formulation can be used with arbitrary cost functions, but replacing it with our distance field $\Delta s$ yields a cost equal to the length of the path. This is a result of using a distance map which contains unit pixel distances in the projection. For uniform $\Delta s$, i.e., a field without distortions, the formulation collapses to a simple euclidean distance expression.

Cohen and Kimmel (1997) explored that this path formulation satisfies the Eikonal equation

$$||\nabla u(x)|| = \Delta s(x). \tag{4}$$

This first-order nonlinear differential equation is used widely in scientific applications, mainly in wave propagation problems as a front propagation approach. In these cases, the cost function on the right-hand side is typically expressed as a propagation velocity at every point in the domain. To our knowledge, we are the first to formulate this equation in a sense to compensate for distorted distances in a projection instead, leading to a formulation for distances from a given boundary for conformal projections. The field of shortest distances $u$ along the projection plane contains the distances from a given boundary to points in this field. Intuitively, the gradient of $u$ can be thought of a measure that is anti-proportional to the distortions defined by the distance field $\Delta s$.

Solving the above differential equation yields a far more accurate result than graph based approaches on a discrete field, resolving metrication errors. There are approaches solving the Eikonal equation, mainly the *Fast Marching Method* (FMM) introduced by Sethian (1996). Starting at known distances in the distance field, i.e. at the boundary where the distance is zero, the algorithm takes advantage of the fact that this field can iteratively be built from the boundary outward since the cost function $\Delta s$ is strictly positive. We refer the interested reader to Petres et al. (2005) for a more in-depth description and evaluation about this algorithm.

## 4 Identification technique in 2-D

Based on the above introduced strategies, we outline our identification technique for 2-D PVSs on a given isentropic level. Figure 4 illustrates the individual steps; pseudo-code of the algorithm is shown in Algorithm 1.

---

**Algorithm 1** Pseudocode of the identification process, both for 2-D and 3-D application. 2-D and 3-D identification differ with respect to distance measure (scale analysis for 3-D) and filtering strategy (additional area based filtering for 3-D); cf. Sect. 6.

---

**Input:** $d_{PV}$: PV field, $w$: width threshold

$d_{PV}^{proj} \leftarrow$ PROJECTION($d_{PV}$)        ▷ Project data to stereographic grid (Fig. 4b)

$\Omega \leftarrow$ GETDOMAIN($d_{PV}^{proj} > 2$ PVU)        ▷ Set stratospheric air mass as domain (Fig. 4c)

$\Gamma_1 \leftarrow$ GETBOUNDARY($d_{PV}^{proj} = 2$ PVU)        ▷ Determine dynamical tropopause (Fig. 4c)

$d_{inner} \leftarrow$ DISTANCES($\Gamma_1, \Omega, \Delta s$)        ▷ Compute distances from boundary into domain (Fig. 4d)

$\Gamma_2 \leftarrow$ GETBOUNDARY($d_{inner} = \frac{w}{2}$)        ▷ Extract boundary where distance equals half the width threshold $\frac{w}{2}$ (Fig. 4d, e)

$d_{outer} \leftarrow$ DISTANCES($\Gamma_2, \Omega, \Delta s$)        ▷ Compute distances from $\Gamma_2$ into domain (Fig. 4f)

objects $\leftarrow$ LABEL($d_{outer} > \frac{w}{2}$)        ▷ Extract and label areas with distance greater than $\frac{w}{2}$ (Fig. 4g)

**return** FILTERED(objects)        ▷ Filter anomalies by length or other heuristics (Fig. 4h)

---

### 4.1 Data

For the 2-D analysis, we use isentropic levels from both the ERA-5 reanalysis (Hersbach et al., 2020) data interpolated to $0.5° \times 0.5°$ in both latitude and longitude, and data from the Subseasonal-to-seasonal (S2S) prediction database (Vitart et al., 2017), as indicated in the corresponding figures. The latter contains forecasts out to 45 days and is available daily on a regular latitude-longitude grid with a grid spacing of $1.5° \times 1.5°$. The data is remapped to a stereographic grid using the Climate Data Operators (CDO; Schulzweida, 2019), hence we note that input grids other that regular longitude-latitude but supported by CDO work equally well.

### 4.2 Computation of the distance field

Before executing the identification strategy, the distance field $\Delta s$ (as introduced in Sect. 3) has to be computed (see Fig. 4a). This field is of importance for the core functionality of the identification, as well as to produce metrics for each PVS. Since the
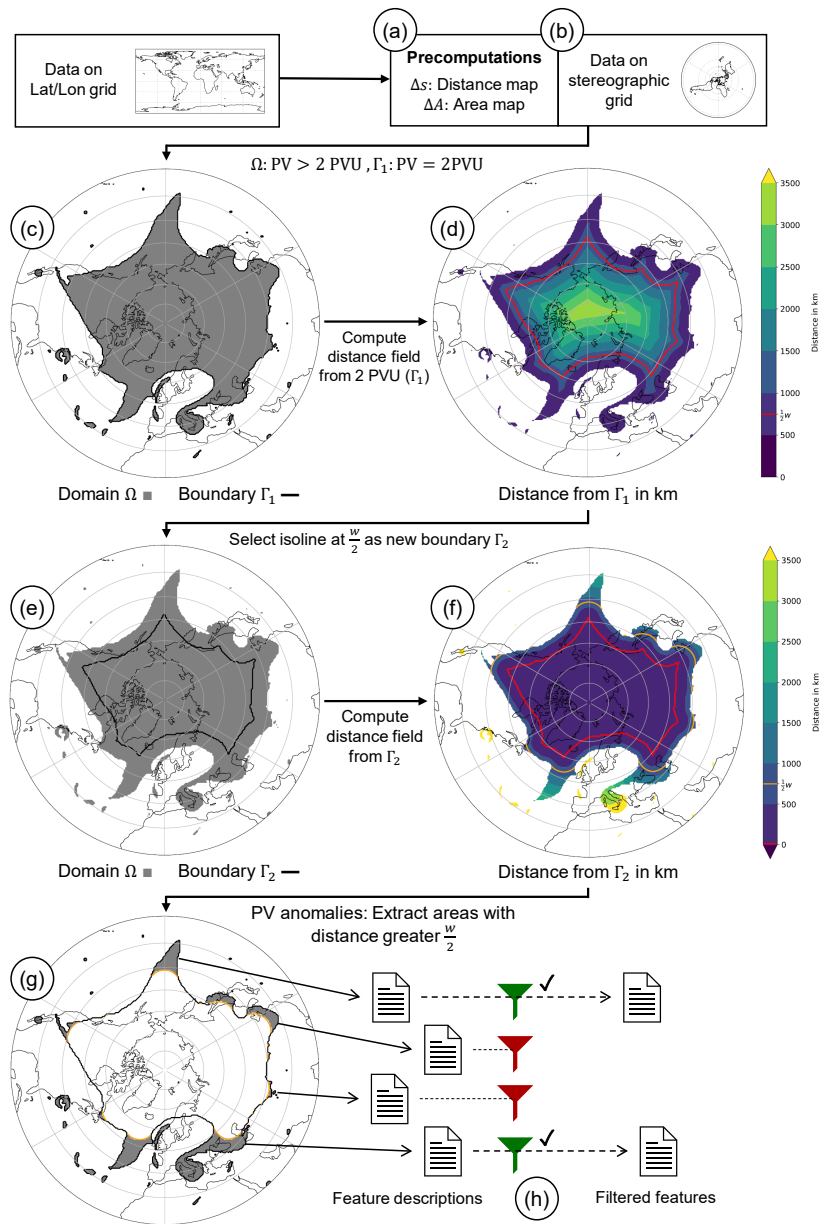
**Figure 4.** Illustration of the algorithm to identify PVSs. (a, b): Input data is mapped to a stereographic grid; precomputations are performed. (c, d): The distance field from the 2-PVU boundary is computed (cf. Sect. 3). Points at a distance equal to half the width threshold $\frac{w}{2}$ are denoted by the red isoline. (e, f): From this isoline, distances outwards are computed. Again, the isoline corresponding to $\frac{w}{2}$ is drawn with respect to the new distance field (orange). (g): Regions with a distance greater than $\frac{w}{2}$ from (f) are extracted. They form anomalies in the field. (h): The anomalies are labeled and filtered, yielding a set of identified PVSs. Visualized is the 335 K isentrope from the ERA5 reanalysis at 7 September 2016, 00 UTC.

values of the distance field do not depend on the input data itself, this can be done as a precomputation step, and cached for future use.

In this study, we use a polar centered stereographic grid with a projection pixel spacing of 75 km and a total size of $340 \times 340$ pixels. This configuration covers one hemisphere completely. A more fine spacing (e.g., 50 km with a size of $510 \times 510$ pixels) increases memory consumption and run time significantly. The resolution of the projection should be chosen to match the resolution of the input data, i.e. the pixel footprints should cover similar areas. For our ERA5 analyses, the mentioned resolution is in that regard sufficient.

Furthermore, as additional metric, the *area map* $\Delta A$ is computed in this step as well. This map contains for each pixel in the projection the spherical area it projects. It can be approximated by the square of the distance map $\Delta A = (\Delta s)^2$. The area map can be used after the identification to calculate the centroid of a given PVS more accurately by taking area distortions into account.

### 4.3 Identification algorithm

Inputs for the algorithm are the field of PV data, and the width threshold $w$. It will become clear that this parameter leads to an identification of PVS up to a maximum width of $w$ km. In our study, we use a value of $w = 1500$ km, a width threshold used widely in existing techniques like the one by Wernli and Sprenger (2007). More input parameters are not required for the 2-D identification. In a first step, the stratospheric body of air is extracted from the projected data set using the previously introduced threshold of 2 PVU. This results in a binary field as seen in Fig. 4c.

The next step is called *erosion*, which is based on the respective morphological operation and encapsulates the main idea of the strategy, as introduced in Sect. 2. The domain $\Omega$ is defined as the stratospheric air mass, and the boundary $\Gamma_1$ as the dynamical tropopause, as shown in Fig. 4c. Starting from the boundary, the ~~distances~~ distance from this boundary ~~(dynamical tropopause)~~ to every point in the stratospheric air mass is computed by solving Eq. 4 using the FMM as outlined in Sect. 3. Figure 4d illustrates this step and shows the color coded distance field. These are the minimum distances from the boundary $\Gamma_1$ to each grid point, respecting the boundaries of the PV field.

After that, areas that are $\geq \frac{w}{2}$ km away from the boundary are extracted, ~~this area resembles~~ these areas form the "inner core". Higher values for the width threshold $w$ would remove more areas along the tropopause and thus smoothing the inner core further. Generally, this inner core may contain multiple disjunct areas. In these cases we pick the biggest one as the core to proceed with, defined by its area. The outer ~~boundary of the identified inner PV core is defined as new boundary $\Gamma_2$ (redline contour of the previously defined inner core (red/black contour in Fig. 4d~~) and the original PV field is kept as the domain $\Omega$/e) is defined as boundary $\Gamma_2$. This boundary is the starting point for the next operation.

The *dilation* operation resembles in morphology the second basic operator, as introduced in Sect. 2. The contour of the previously defined inner core is used as input boundary, and distances outward with respect to the domain $\Omega$ are computed. Keeping this domain is necessary to measure distances following the stratospheric domain with respect to the distance map (as illustrated in Fig. 2). The distances can be seen in Fig. 4f. ~~Areas of the dilated field with a distance $\geq \frac{w}{2}$ are extracted again (orange line in Fig. 4f). For a smooth PV field or a small value for $w$, this converges to the inverse operation to the erosion.~~

275 ~~However generally, the erosion destroys~~ While filament-like structures ~~and cutoffs, which cannot be reconstructed that way (also illustrated in Fig. 1d-f). The distance field will "grow" into the PVSs, providing an accurate measure for the length of a PVS~~ are assigned meaningful distances, note that cutoffs cannot be reached by the algorithm in this step because they are not spatially connected to the stratospheric reservoir. Their distance is set to infinity, making them easily recognizable later. Feature descriptions for these objects can be computed, nonetheless.

280 These extracted areas resemble the identified PVS (shaded areas in Fig. 4g). Since $\frac{w}{2}$ km have been eliminated along the contour (from all sides) during the erosion step, anomalies with a maximum width of $w$ km are identified. Figure 5 shows outputs of the algorithm for different values of $w$. Higher values for this parameter lead to the identification of wider anomalies, as well to further extension of anomalies towards the main reservoir. The identification of cutoffs is not sensitive to parameter changes.

285 Finally, the identified anomalies are filtered. We separate the filtering into two steps.

1. *Area based filtering.* This step can be skipped for the 2-D identification, but will be of importance for the 3-D case as outlined later on. Areas identified by the dilation step (Fig. 4g) can be excluded, for example based on spatial information and surroundings.

2. *Object based filtering.* After the area based filtering, the identified anomalies are labeled as PVS by cohesiveness, and
290 are referred to as objects hereafter. For each object, a *feature vector* is computed consisting of various characteristics (see below). Then, objects can be filtered based on heuristics defined on these characteristics (Fig. 4h). Initially, our approach identifies both reasonably sized anomalies as well as very small ones (see e.g., Fig. 4g). Thus, we only keep PVSs that satisfy a certain length. We use a length threshold of 1000 km, which is based on the 2000 km threshold used by Wernli and Sprenger (2007) for the minimum length along the contour. Depending on the use case, other metrics can be used
295 for filtering as well.

## 4.4 Feature Vectors

Here, some elementary measures for each of the identified PVSs are introduced. This list is not exhaustive and can be extended with more metrics. For an identified object $\mathcal{O}$ containing the pixels (or area) $(x,y) = \mathbf{x} \in \mathcal{O}$, we compute:

– The object's area $A_{\mathcal{O}}$ as integrated pixel area using the precomputed area map $\Delta A$:

$$A_{\mathcal{O}} = \int_{\mathcal{O}} \Delta A(\mathbf{x}) d\mathbf{x} \approx \sum_{(x,y) \in \mathcal{O}} \Delta A(x,y). \tag{5}$$

Besides as filtering metric, the object area is a component of some of the subsequent mentioned metrics.

– The length $l_{\mathcal{O}}$ of a streamer, which is defined as the maximum distance $d$ ~~emerging $\Gamma_2$~~ present in the anomaly $\mathcal{O}$ in the dilation step, see Fig. 4f, minus $\frac{w}{2}$:

$$l_{\mathcal{O}} = \max_{\mathbf{x} \in \mathcal{O}} d(\mathbf{x}) - \frac{w}{2}. \tag{6}$$
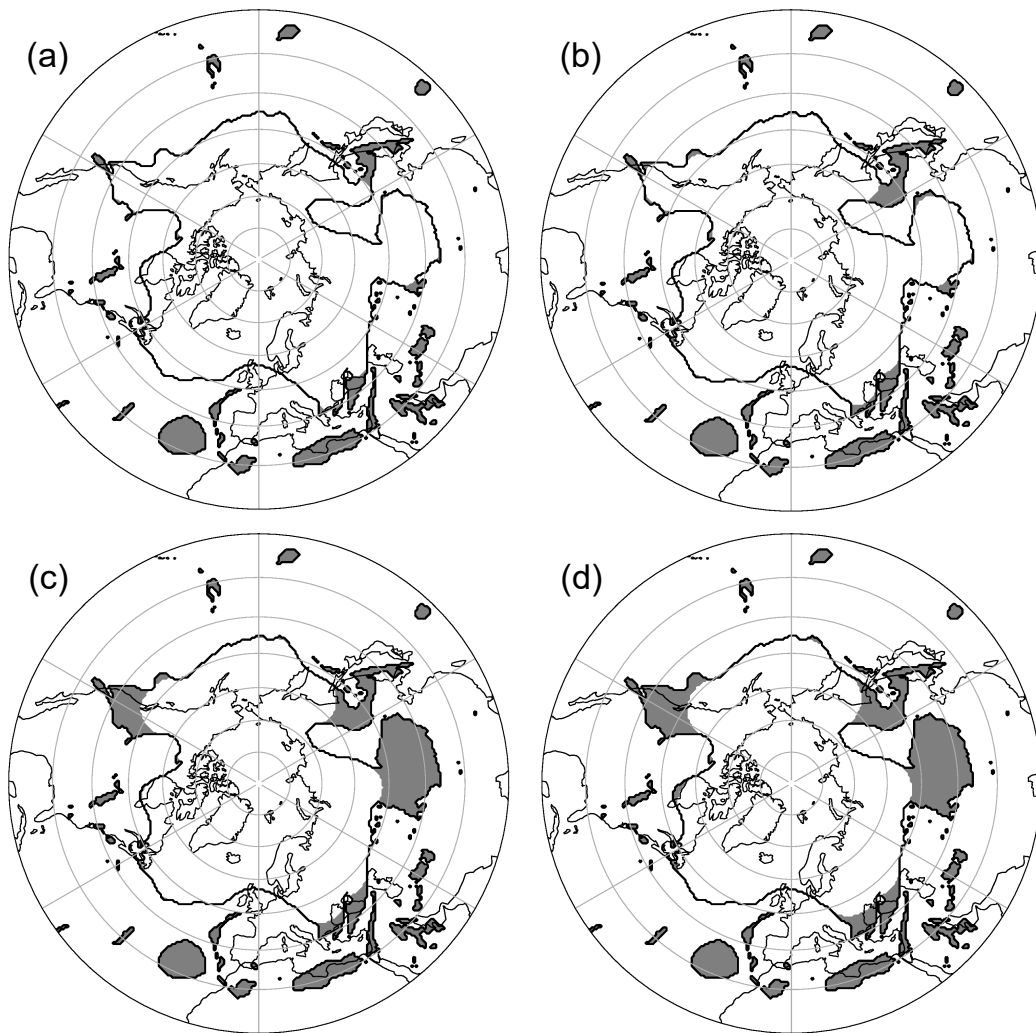
**Figure 5.** Sensitivity of the identification process to the width threshold $w$. Here, $w$ equals (a) 500 km; (b) 1000 km; (c) 1500 km; (d) 2000 km; using the example of the 330 K isentrope from the ERA5 reanalysis at 14 July 2015, 00 UTC. Results are not filtered to highlight the effect of the parameter. Our default configuration is (c).

305 ~~This is a very useful metric since our~~ Our identification strategy yields a useful and robust length measure, while ~~most~~ other identification strategies only measure the length along the 2-PVU contour. The length is used as a filtering parameter, as further elaborated in Sect. 5, and well suited for statistical analyses.

– The object's average PV, weighted by the area measure at each point:

$$PV_{\mathcal{O}} = A_{\mathcal{O}}^{-1} \int\limits_{\mathcal{O}} PV(\mathbf{x})\Delta A(\mathbf{x})d\mathbf{x} \tag{7}$$

310 – The object's centroid as weighted average of the position over the area $\Delta A$ and the PV itself:

$$C_{\mathbf{x}} = (\int\limits_{\mathcal{O}} \Delta A(\mathbf{x})PV(\mathbf{x})d\mathbf{x})^{-1} \int\limits_{\mathcal{O}} \mathbf{x}\Delta A(\mathbf{x})PV(\mathbf{x})d\mathbf{x} \tag{8}$$

The centroid gives a quantifiable estimate of the object's center. For example, in Fig.6b the centroid of an object lies in the intersection of the green main axes. Their tracks can be used to investigate case studies or to perform big data analyses. In the 3-D case, it also reveals the vertical movement of the anomaly, as outlined in the case study later on.

315 – The main axes of the object. Evaluating the eigenvalues and eigenvectors of the covariance matrix of the second order moments define the main axes of the shape (see Mukundan and Ramakrishnan, 1998). Image moments themselves are quantifiable metrics to evaluate visual information present in an image, like shape, position and orientation. Main axes form the best-fitting ellipsoid to the data, see Fig. 6. Furthermore, the major main axes can be seen as the general orientation of the object, thus the horizontal tilt of a PVS in 2-D data sets. These computations can be applied in 3-D as 320 well, but interpretation gets increasingly difficult.

## 5 Evaluation and comparison in 2-D

Before extending the idea to 3-D data sets, the introduced algorithm will be evaluated. For this purpose, the identification strategy by Wernli and Sprenger (2007); Sprenger et al. (2013, 2017) (hereafter abbreviated by WS07) will serve as benchmark. Their strategy is well established, used in various follow-up work, and considered as state-of-the-art.

325 The PVS identification strategy by WS07 is motivated by the thin and elongated nature of PVSs. They identify the outermost 2-PVU contour, and connect points along it with a spacing of about 30-50 km. This step returns a polygon resembling the outermost dynamical tropopause for the selected isentropic level. Then, the vertices of this polygon are pairwise compared regarding their direct spherical distance $d$ and length along the contour $l$. For small $d$ and big $l$ ($d < 1500$ km and $l > 2000$ km according to Sprenger et al. (2013)~~, also with a set threshold for their aspect-ratio $l/d$~~), these points are regarded as streamer 330 base points, and the enclosed shape by the polygon vertices in between the base points ~~form~~ forms a PVS.

Comparing the algorithm parameters, our width threshold $w$ has a similar interpretation to the $d$ in the work by WS07, so we set $w = d$ to evaluation purposes. This choice is also confirmed by climatological analyses we conducted, showing an inordinate increase in detected structures for smaller values of $w$. While their strategy has the additional $l$ parameter (length
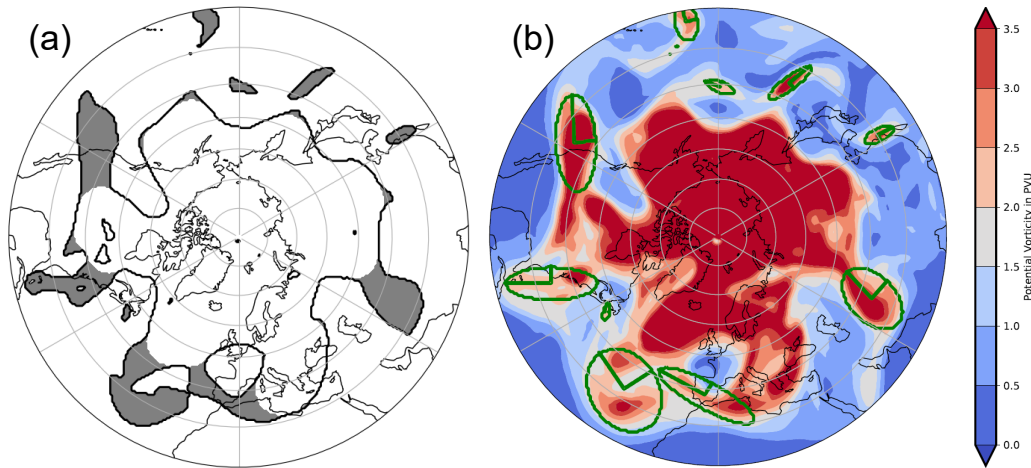
**Figure 6.** 200–500 hPa layer-averaged PV based on the S2S prediction database (Vitart et al., 2017) initialized and valid at 7 May 1998, 00 UTC. (a) PVSs identified using the default configuration without filtering the results; in (b) filtered PVSs using a length threshold of 1000 km, approximated by best-fit ellipses, with raw PV values shaded in the background. Ellipses are defined by their computed main axes.
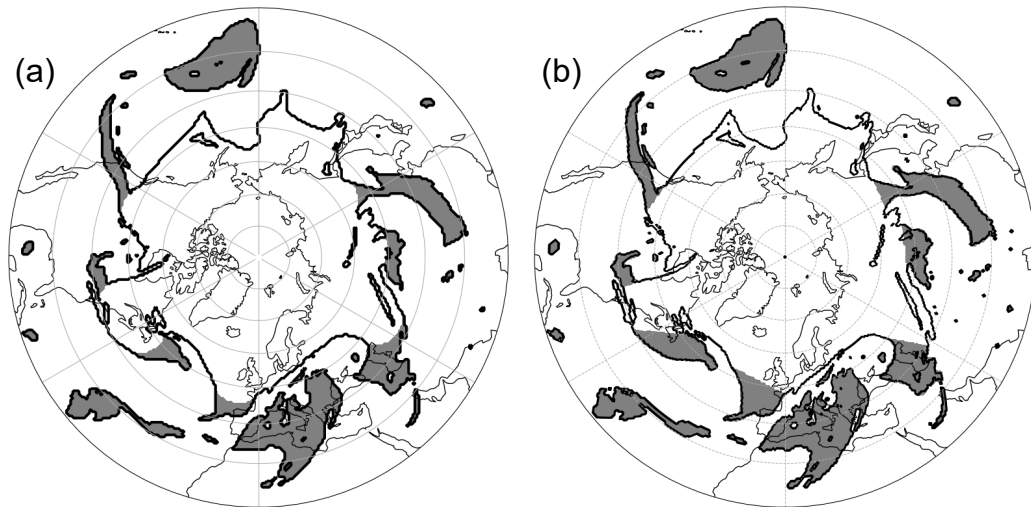
**Figure 7.** Comparison of the identified PVSs based on the 340 K isentrope from the ERA5 data set at 24 July 2015, 00 UTC. (a) Results of our strategy using the default configuration; and (b) by the algorithm from Wernli and Sprenger (2007). The anomaly over East Asia will be investigated in the case study later on.

along contour), we do not set length restrictions in the first step, but filter the identified PVSs later based on a length threshold. 335 This filtering parameter can be refined and changed depending on the use case.
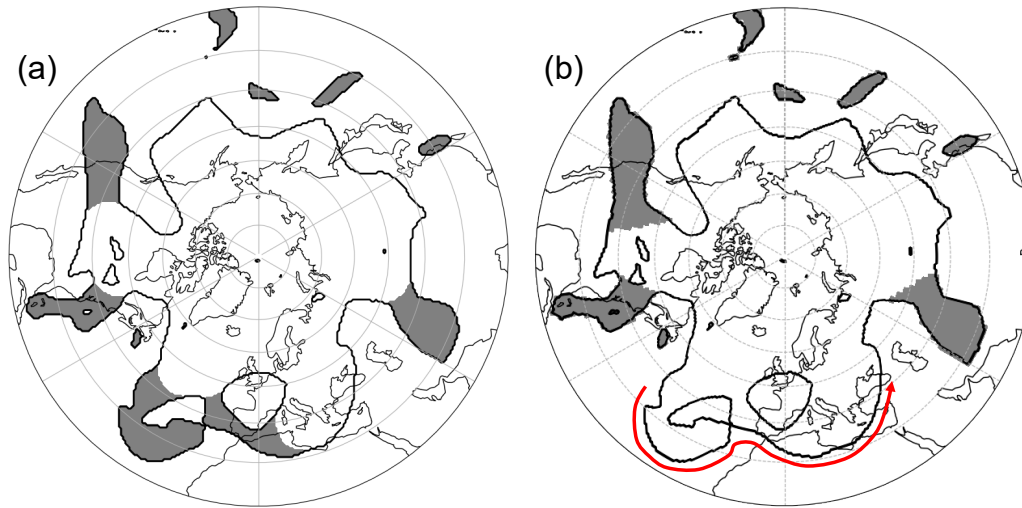
**Figure 8.** Same as Fig. 7, but with the data set used in Fig. 6. The red arrow indicates the outermost 2-PVU contour, which is used to identify anomalies in the work by Wernli and Sprenger (2007).

Figure 7 shows the identification results in a 2-D PV field for both techniques. This example represents one of the predominant instances where we expect similar results. As shown in the figure, both techniques indeed identify the same structures as streamers. These also resemble the structures an educated user would identify as such. However, the shapes slightly differ: While the method by WS07 (Fig. 7b) identifies streamer base points and connects them in a straight line, our approach naturally

340 creates roundings at these spots. These roundings reveal the boundary on what is considered as an anomaly in the field by our algorithm. The amount of detected anomalies is similar to the climatological results presented in their study, as supported by analyses we conducted.

Figure 8 on the other hand shows a more complex example demonstrating the improvements gained by considering the full PV field compared to only the outermost 2-PVU contour. Over Europe and the Eastern Atlantic, two PV intrusions (one

345 cyclonic and one anti-cyclonic) are being observed at this specific time step.

These structures are not identified by the approach from WS07~~, ours on the other hand is capable of identifying them~~. By only ~~considering~~ tracing the outermost 2-PVU boundary ~~, their algorithm only spot a broad trough~~ (red arrow in Fig. 8b), the inner state of the PV domain is not considered. However, this inner state might reveal important information on the large-scale flow in this area. By just following this red arrow, this results in a broad PV trough being spotted from the Central Atlantic

350 all the way to the Black Sea~~without taking into account the tropospheric air encapsulated in the stratospheric reservoir. With respect to our strategy, the dilation step does not reconstruct these areas~~. This wide trough does not fulfill the required thresholds for a streamer. The strategy we employ also considers the inner state of the reservoir, including all troposphere-stratosphere boundaries in the domain. The erosion step starts at these inner boundaries as well. Therefore, these elongated structures are not reconstructed in the following dilation step, identifying them thereafter. In this specific case, the two anomalies are spatially

355 connected. Our strategy therefore considers them as one merged object. Further heuristics are required to distinguish the origin of these merged anomalies in order to separate them, but this issue is beyond the scope of this work.

As introduced, we filter PVSs that do not satisfy a length threshold of 1000 km, which is comparable with the 2000 km contour length used in WS07. Note that our identification strategy does not have to be re-executed when changing the length threshold, instead the filtering step returns a different subset of identified streamers. Furthermore, our length calculation yields
360 a precise measure for the streamer, from the reservoir base to its tip. Contour length techniques can suffer from inconsistencies introduced by noise. All in all, our strategy generally leads to additional identified streamers, especially in complex environments.

For a set test case, the computation of anomalies takes a similar amount of time. Regarding our strategy, this time also includes the computation of feature vectors (e.g., best fit geometry and centroids) for all detected structures. Nevertheless,
365 due to different approaches in preprocessing, different programming languages and different use of parallelism, a quantifiable comparison of the run times proves to be difficult. For our algorithm, 85% of the computational time is required to project the data onto a stereographic grid. Analyzing bigger data sets, data level parallelism (e.g., Single Instruction Multiple Data (SIMD) processing) can highly improve the efficiency of projecting data from multiple isentropic levels and multiple time steps. Projecting a 3-D data set with 50 isentropic levels takes only twice the time compared to a single isentropic level,
370 signifying a very high speedup.

Similar to WS07, we currently do not consider the origin of a PVA as a filtering strategy. For example high PV of non-stratospheric origin (e.g., in tropical cyclones (Molinari et al., 1998) or related to deep moist convection (Weijenborg et al., 2017)) could also be identified as a PVS or cutoff. More sophisticated time series and algorithms are needed to distinguish the origin of a PV filament, e.g., using specific humidity thresholds as suggested by Škerlak et al. (2015). Our concept of feature
375 descriptions can give some insight and provide quantified values for each identified anomaly, but this problem goes beyond the scope of this work.

# 6 Extending the principle to three dimensions

## 6.1 Data

For the 3-D identification, ERA-5 data on model levels (here we use levels 40-137) are interpolated to isentropes with a vertical
380 spacing of 2 K. We consider the region between 290 K and 380 K. The vertical spacing of 2 K proved to be a reasonable trade-off between computation time and disk usage on one hand, and representing the vertical structure of the PV field on the other hand. Our experiments showed that coarser spacings (e.g., 5 K) hide dynamically relevant and fine features in the vertical structure, while a finer spacing would increase memory and computational cost significantly without adding much more relevant structures. Using pressure or model levels would be an alternative worth consideration, since it is natively
385 available in higher vertical resolution as model output. Therefore, it would not require the interpolation on isentropic levels, but would reduce the interpretability of the PV structures, especially when evaluating them using cross-sections or development over time.
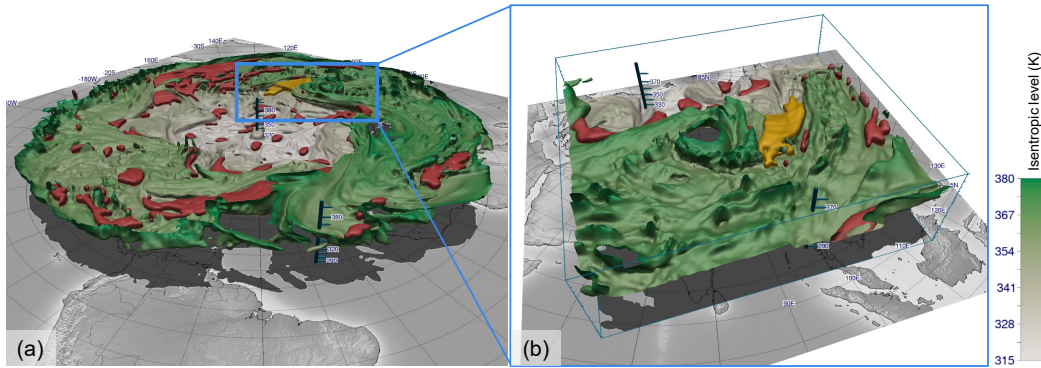
**Figure 9.** 3-D visualizations of the dynamical tropopause (defined by the 2-PVU threshold), which is shaded by isentropic height. Red and yellow are the identified anomalies. The yellow one is emphasized for the case study later on. (a) The entire northern hemisphere is visualized; in (b) focused view of southeastern Asia, which is investigated in the case study in Sect. 7.

## 6.2 Strategy

~~There are multiple ways~~

There are several possibilities to extend the algorithm introduced in Sect. 4 ~~introduced algorithm~~ to 3-D. A straightforward idea consists of applying the 2-D identification to each isentropic layer individually and stacking the results on top of each other to create 3-D objects (e.g., Portmann et al. (2021) for cutoffs). ~~Applying~~ However, applying our algorithm in ~~a similar way to 3-D data sets would be trivial, but would not support the idea of including the information of the~~ such fashion would not include all information available in the full 3-D ~~shape around each~~ environment at a given point. For example, the use of thresholds (e.g., minimum contour length or aspect-ratio) in a 2-D identification strategy can lead to artifacts when applied to ~~fine resolved~~ multiple vertical levels: A PVS might be ~~just~~ above a given ~~threshold and therefore being identified on~~ identification threshold on two specific isentropic levels, but ~~not~~ below this threshold on a level in between ~~where the structure is just below this set threshold. This generates so-called popping artifacts~~ these two, therefore not identifying this structure on this intermediate level. This leads to unwanted vertical gaps in the detected structure.

Figure 9a shows a 3-D visualization of the tropopause (defined by the 2-PVU isosurface) and the identified anomalies. Like for the previous 2-D visualizations, we use a stereographic map projection to emphasize the circulating nature of the flow. The tropopause is shaded by isentropic height. The air mass below (above) this isosurface is classified as tropospheric (stratospheric) air. As expected, the dynamical tropopause has a lower height towards the North Pole (beige shaded area), and a sharp gradient is situated in the mid-latitudes where the jet stream is located (Koch et al., 2006). Visible holes in the isosurface resemble tropopause levels beyond our visualization domain of 290 K to 380 K. The red and yellow regions (latter emphasized for the case study) are the identified PVAs. As visible in Fig. 9 and supported by the video supplement, identified anomalies can manifest in various shapes. Anomalies are typically located in valleys, folds, or generally along depressions of the tropopause

boundary. Because the algorithm works in a broader sense to identify irregularities along the boundary, it does not distinguish between the type of anomaly. Especially for our case study, we use the East Asian region shown in Fig. 9b to analyze the event.

410      Our algorithm is designed in a manner that it can be executed on 2-D and 3-D data in a similar way. Note that the basic ~~ideas for~~ morphological operations (erosion and dilation~~(,~~ see Sect. 2) can be applied independently on the number of dimensions~~and can therefore be used the same way. Existing identification techniques (see an overview of them in Papin et al., 2020) rely on following the 2-PVU boundary, and searching points along this boundary meeting certain criteria. However~~. This requires the computation of distances in 3-D ~~, there is no unique direction to follow the dynamical tropopause, making such techniques~~
415    ~~not applicable. In our study on the other hand, the 2-D boundary Γ introduced in Sect. 4 can be extended to a 3-D boundary (isosurface) and handled the same way.~~

     ~~The distance measure introduced in the 2-D identification needs to be expanded to another dimension to be applicable to a 3-D field~~space. Therefore, we need to revise the distance measure due to mismatching units in the horizontal domain (km) and vertical domain (K). An exact distance approach would compute the height of every grid point in the atmosphere given
420    its current state ~~,~~ or based on a standard atmosphere model. Due to too high computational effort, we keep isentropic levels as vertical distance measure, and combine the mismatching units in the horizontal ~~(euclidean distance) and vertical (potential temperature) direction~~ and vertical dimension motivated by a scale analysis. ~~In a true-scale visualization, 3-D PV anomalies would be very flat. Bithell et al. (1999) used a vertical scaling factor of about 100 in their z-axis to achieve clearly visible 3-D structures in the PV field. We approximate~~ One challenge in computing meaningful distance is posed by the different horizontal
425    and vertical scales, and in our case also different units, of the PVAs in the atmosphere. Typically we find the horizontal scale of PVAs ~~in the magnitude~~ to be on the order of 1000 km, and their vertical scale ~~in the magnitude~~ on the order of 10 K. ~~As a result, we stretch~~ To achieve meaningful 3-D visualization, Bithell et al. (1999) used a vertical scaling factor of about 100 along the z-axis to achieve distinctly visible 3-D structures along the tropopause. To achieve similar order of magnitude in all three spatial dimensions, we follow Bithell et al. and scale the vertical dimension by ~~factor~~ a factor of 100 km K$^{-1}$ ~~to achieve
430    a similar scaling like Bithell et al. (1999), and use an euclidean measure~~when computing these stretched distances. This helps ~~our algorithm to "work" along all dimensions simultaneously rather than focusing on the horizontal dominant level. For the algorithm, we use the same~~, then use a Euclidean distance measure. With respect to the width threshold $w$ ~~as before~~, the same value as in 2-D is used.

     Figure ~~9 shows an example of~~ 10 shows the main steps of the 3-D ~~visualizations of the~~ algorithm. The 3-D erosion step
435    is illustrated in Fig. 10a, analogous to the 2-D case in Fig. 4d. As in 2-D, the domain $\Omega$ is the stratospheric air mass, and the boundary $\Gamma_1$ is the dynamical tropopause defined by the 2-PVU ~~contour and the identified anomalies. For visualization purposes, we use the stereographic projection to emphasize the circulating nature of the flow. The isosurface represents an isovalue of~~threshold. Thus, adding the vertical dimension, the domain $\Omega$ becomes the three-dimensional stratospheric air mass, and the boundary $\Gamma_1$ becomes the 2 PVU ~~(tropopause) and is shaded by isentropic height. This is analogous to the boundary~~
440    ~~Γ in the~~isosurface, separating stratospheric and tropospheric air. Distances are measured emerging from the boundary $\Gamma_1$ into the stratosphere using the same algorithm as in 2-D ~~ease. As expected, the tropopause has a lower height towards the North Pole (beige shaded area), and a sharp gradient is situated in the mid-latitudes where the jet stream is located (Koch et al., 2006)~~
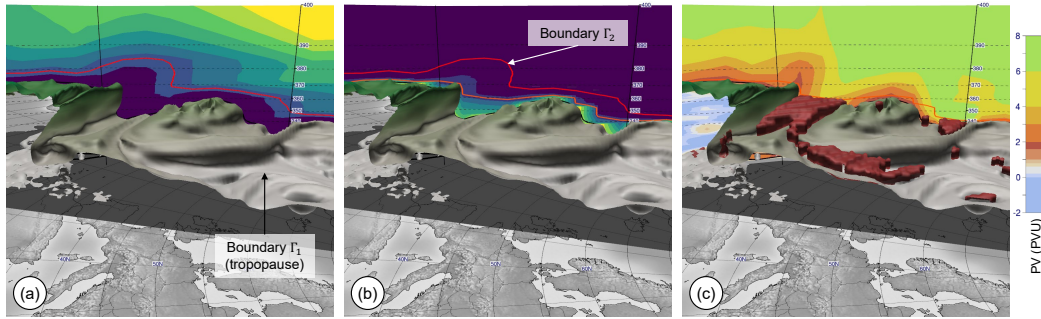
**Figure 10.** The identification method introduced in Fig. 4 for 2-D extended to 3-D~~visualizations of~~ . The tropopause over the North Atlantic (2 PVU isosurface) is displayed along with a vertical cross section showing: (a) distances from the dynamical tropopause (~~defined by the 2-PVU threshold~~boundary $\Gamma_1$) ~~, which is shaded by isentropic height~~towards the stratosphere (see Fig.~~Red~~ 4d for the 2-D equivalent and ~~yellow are~~ the ~~identified anomalies~~color map); (b) distances measured outwards from the red boundary $\Gamma_2$ into the stratospheric domain (see Fig.~~The yellow one is emphasized~~ 4f for the ~~case study later on. (a~~2-D equivalent and the color map)~~The entire northern hemisphere is visualized~~; ~~in (b~~c) ~~focused view~~the Potential vorticity. Additionally shown are the identified anomalies on top of ~~southeastern Asia, which is investigated in~~ the ~~case study~~tropopause (red objects). Note that all distances are measured in ~~Sect~~3-D, while the visualizations only show a 2-D cross section. ~~7~~Displayed on all panels is the ERA5 reanalysis at 24 July 2015, 00 UTC.

~~. Visible holes in the isosurface resemble tropopause levels beyond our domain of 290~~but applied along all dimensions. This results in a 3-D distance field (same color mapping as in 2-D, c.f. Fig. 4). Figure 10a shows the distance along a vertical cross section. As in 2-D, we define the new boundary $\Gamma_2$ as all points that have a distance of $\frac{w}{2}$ ~~K to 380~~km from the boundary (red contour in Fig. 10a, note that in 3-D space this becomes a 3-D isosurface). The isosurface $\Gamma_2$ resembles the inner core and lies above the dynamical tropopause, bounded by a smoother boundary than the tropopause itself. Distances outwards are computed starting at $\Gamma_2$ (Fig. 10b, cf. Fig. 4f for 2-D), using the same distance measure. Distances here are constrained by the domain $\Omega$ and grow into the anomalies. Then, as in 2-D, another boundary is defined at a distance of $\frac{w}{2}$ ~~K. The red and yellow regions (latter emphasized for the case study)are being identified as PVAs. Especially for our case study later on, we use the zoomed-in area around East Asia (~~km (orange contour in Fig. 10b, cf. Fig. 4f). Areas farther away than $\frac{w}{2}$ km from $\Gamma_2$ are extracted. These are the areas identified as anomalies. As shown in Fig. ~~9b) to analyze the event.~~10c, these anomalies are volumetric objects. Stratospheric air masses isolated from the stratospheric reservoir (3-D cutoffs) are identified similar to the 2-D process. These areas cannot be reached by the distance computation starting at $\Gamma_2$.

~~Generally, the~~ The same filtering process as in 2-D is applied~~.~~: The goal is primarily to keep as many detected features as possible. The user can then decide which of these features are of importance for a specific use case and filter these by their feature descriptions accordingly. The centroid ~~can be~~ of an anomaly is computed in 3-D (see Eq. 8), while we replace the area measure with a volume measure (in km$^2$K, or approximated in km$^3$ using a standard atmosphere model). Using the centroid in 3-D, the vertical evolution of an anomaly can be analyzed. The case study in Sect. 7 investigates an anomaly with a clear shift in vertical position over time. Some of the feature descriptions introduced in Sect. 4.4 are not computed in 3-D due to
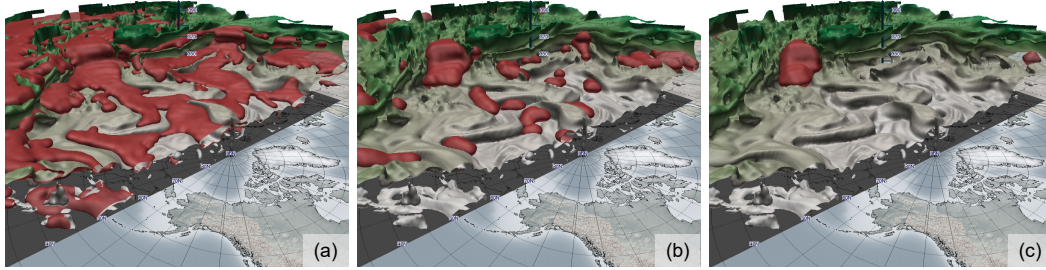
**Figure 11.** Effect of filtering anomaly candidates with respect to the extent threshold $b$. Areas of anomalies with a vertical extent smaller than $b$ are omitted from the identification, this leads to clear and separated features. (a) $b = 0\,\mathrm{K}$ (no filtering); (b) $b = 6\,\mathrm{K}$; (c) $b = 12\,\mathrm{K}$. As specified in Sect. 6, our default configuration is 6 K.

misleading interpretations introduced by the scale analysis. For example, the length measure implicitly contains the applied scale factor of $100\,\mathrm{km\,K^{-1}}$. As stated above, this scaling is required for the 3-D functionality. ~~For analyses on these features,~~ Instead, for analyses we suggest to use measures that are easy to interpret, like the object's bounding box or volume. Typical use cases for filtering involve setting thresholds for position, size, or values of other fields in the same area (e.g., humidity).

465 However, in 3-D, we have to introduce an area based filtering strategy (cf. Sect. 4.3). Fig. 11a shows the raw and unfiltered results of our identification applying the strategy described in this section. ~~Thus, despite their distinct nature on 2-D isentropic levels, these~~ Despite the distinct nature of isentropes in 2-D ~~PVS are connected~~, the 3-D display reveals that these are actually connected dynamically to each other ~~via~~ in 3-D ~~structures. Actually, this~~. This behavior is expected from a globally coherent flow structure. However, it is desired to have clear and distinct anomalies ~~, so they can be described~~ in order to describe these

470 individually for statistical analyses. Therefore, we filter the identified anomaly areas to get a set of detached and separated objects. We discard identified areas with a vertical extension of smaller than $b = 6\,\mathrm{K}$. This specific value has been chosen as a good compromise of grasping clear features, while separating them at areas of weak vertical extent. Illustrated in Fig. 11b, c are identification results using different values of the extent threshold $b$. Figure 11b shows our default configuration.

Existing identification techniques (see an overview of them in Papin et al. (2020)) rely on following the 2-PVU boundary,
475 and searching points along this boundary meeting certain criteria (e.g., Wernli and Sprenger (2007)). However, in 3-D, there is no unique direction to follow the boundary (the dynamical tropopause), making such techniques not applicable. Our strategy does not rely on a spatial tracking of a contour, but works in a broader sense, where the introduced domain $\Omega$ and boundaries $\Gamma_1, \Gamma_2$ can be extended to a 3-D perspective.

### 6.3 Implementation Details

480 Our feature identification strategy is implemented in a Python framework for general identification and tracking of meteorological structures. This framework is part of the *enstools* Python package (Redl et al., 2018), which provides routines for meteorological data processing. Identified features can be saved to disk in an efficient binary format or as humanly readable JSON format. The algorithm itself relies on the CDO Python wrapper (Schulzweida, 2019) for projections. The *scikit-fmm*

toolkit (Furtney, 2019) provides an implementation for the FMM algorithm, which has been adapted to fit our requirements. Our software is open-source and can be accessed as outlined in the code availability section.

On a test system (AMD Ryzen 5 2600X), the identification process for a 3-D data set with 46 isentropes (290 K to 380 K) and a horizontal grid spacing of $0.5° \times 0.5°$ for the northern hemisphere takes 5.9 s, where projecting the data takes about 23% of the total time. The computation time heavily depends on what features are being computed for the individual objects (e.g., main axes, volume). Compared to the 2-D identification, where one isentropic level took about 1 s to process, this 3-D identification reveals a big speedup regarding the amount of data being processed.

## 7   Case Study: Precursors of an extreme precipitation event affecting northeastern Vietnam

We evaluate our identification method along a case study. The selected extreme precipitation event affected northeastern Vietnam in late July and early August 2015. This event has been analyzed in detail by Van der Linden et al. (2017), particularly focusing on the predictability of the event in ensemble forecasts. According to Van der Linden et al. (2017), predictability of the event was strongly related to the presence and location of an upper-level trough in ensemble forecasts. The trough was diagnosed using 200 hPa geopotential height and upper-tropospheric PV fields. Here, we want to take a closer look into the precursors of this trough, and to analyze whether our novel technique can help to improve the analysis of the dynamics leading to the event. Different configurations of the algorithm are considered and our intent is to shed light on the 3-D structures. Visualizations are made using the open-source visualization framework Met.3D (Rautenhaus et al., 2015b). The visualized data is based on the ERA5 reanalysis as introduced in Sect. 6.1, using a 6-hourly temporal resolution. To assess the rainfall associated with the event, the 6-hourly rainfall totals centered around the respective time steps and based on the satellite-based NASA Global Precipitation Measurement (GPM) Integrated Multisatellite Retrievals for GPM (Huffman et al., 2015) product were computed. This data is available at a horizontal grid spacing of $0.1° \times 0.1°$.

Between 25 July and 3 August 2015, record-breaking precipitation amounts of more than 1000 mm were observed at the coast of northeastern Vietnam. According to the analysis by Van der Linden et al. (2017), intense and long-lasting rainfall was mainly related to a surface low over the Gulf of Tonkin, which slowly moved westwards. Van der Linden et al. (2017) provided evidence that the formation and westward movement of the surface low was related to a subtropical trough that was located over southern China slightly to the northwest of Vietnam. Figure 12a, b shows the 200 hPa isosurface in an isentropic field. The isosurface is shaded by geopotential height. It also shows the location of the identified trough axis by Van der Linden et al. (2017), using a 2-D analysis. In Fig. 12b, the identified PVA has been added. Clearly, the geopotential anomaly coincides with the position of the trough axis and the PVA. Fig. 12c shows for the same point in time a corresponding vertical cross-section, which will be used for the PV analysis in this case study. The location of the PVA matches the PV intrusion over southeastern Asia.
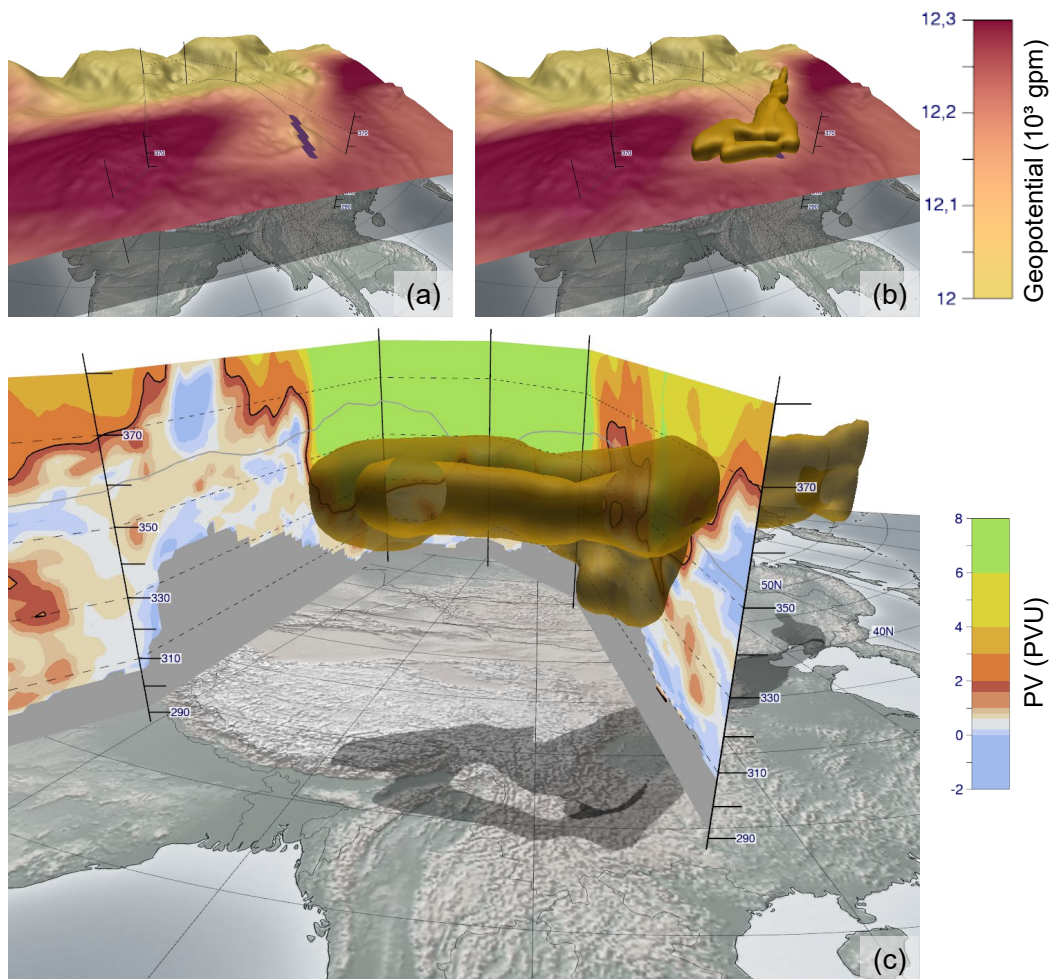
**Figure 12.** Comparison of the location of the anomalies defined by the 200-hPa geopotential height and by the dynamical tropopause, at 25 July 2015, 12 UTC. (a) Shown is the 200 hPa isosurface on isentropic levels, shaded by geopotential height. The purple line is the trough identified by Van der Linden et al. (2017) in a 2-D environment, cf. their Fig. 7. (b) Same as (a), but the identified PVA has been added. Its location clearly coincides with the geopotential trough and the identified trough axis. (c) A vertical cross-section of the area along the section indicated in (a, b), and the identified PVA. Shaded on the cross-section is the PV, where the black contour highlights the 2-PVU boundary. The gray line is the 200-hPa contour, thus a cut of panel (a).

## 7.1 PV analysis

To examine the precursors of the extreme event, we visualize the 3-D PV environment over eastern Asia for five different time steps in Fig. 13. These panels provide snap-shots, starting four days prior to the extensive rainfall. For an additional depiction of the event based on an extended time frame, also including intermediate time steps, we provide an animation of the synoptic evolution of the PVA over time in a supplementary video at Fischer et al. (2021b).

At 21 July 2015, a wide trough is located over eastern China, which is below the thresholds for our identification strategy in 2-D and 3-D to be detected as a PVA. Over the following days, a high tropopause area is moving from the Arabian Peninsula towards the Tibetan Plateau, further illustrated in the supplementary animation (Fischer et al., 2021b). In the visualization, this anomaly is characterized by a "hole" in the 2-PVU isosurface, and a high tropopause on the western side of the vertical cross-sections in Fig. 13a-c. Holes in the tropopause are heights beyond our domain that has an upper limit of 380 K. From 22 July 2015, 00 UTC onward, our algorithm identifies a PVA related to the event over Vietnam (shaded in yellow in Fig. 13). The high tropopause over the Tibetan Plateau shows to be stationary over multiple days, leading to intense advection of high PV air southwards along its eastern flank. This leads to an intensification of the identified anomaly until 24 July (Fig. 13a-c). After this, the influence of the Tibetan high becomes weaker, cutting the intrusion of high PV air from the north. The identified structure partially cuts off horizontally from the main reservoir (Fig. 13c). During this process, the centroid of the PVA moves southward from around 30°N over mainland China to around 20°N just west of Vietnam. This is clearly visible by the PVA moving southwards along the cross-section from Fig. 13a-c. The period afterwards coincides with the most intense rainfall over the coast of northeastern Vietnam as indicated by the 6-hourly precipitation rate (cp. Fig. 5 in Van der Linden et al. (2017), and emphasized blue boxes in Fig. 13b-d). Over the following days some dynamical coherence to the main tropopause fold can still be observed (Fig. 13d, e). Finally, on the first days of August the PVA dissipates (see Fischer et al., 2021b).

Comparison with the results of Van der Linden et al. (2017), who used 200-hPa geopotential height and 500–200-hPa vertically averaged PV to detect the subtropical trough (their Fig. 7), clearly illustrates the advantage of our novel 3-D approach. Using vertically averaged PV, and single-level geopotential height, the trough could not be identified after 25 July, and 29 July, respectively, by Van der Linden et al. (2017). Since high precipitation amounts were also observed after 29 July and in more western parts of the country (not shown), a continuing influence of the PVA seems reasonable.

## 7.2 Algorithm evaluation

As outlined in Sect. 6, two parameters in this algorithm for 3-D cases are used, the width threshold $w$ (see Sect. 4) and the extent threshold $b$ to separate PV anomalies into clear disjunct features. For this case study, we use the default configuration of $w = 1500\,\mathrm{km}$ and $b = 6\,\mathrm{K}$.

To give an understanding of identifying 3-D structures over their 2-D counterpart, Fig. 14 illustrates two different time steps of the case study. In Fig. 14a-c, a time step early in the development of the event is shown. The yellow anomaly can be identified as an equatorward intrusion of stratospheric air in the 3-D view. Considering identification on 2-D levels (340 K and 355 K isentropes in Fig. 14a), the structure is not identified on the 355 K isentrope due to its broad shape, but the anomaly is
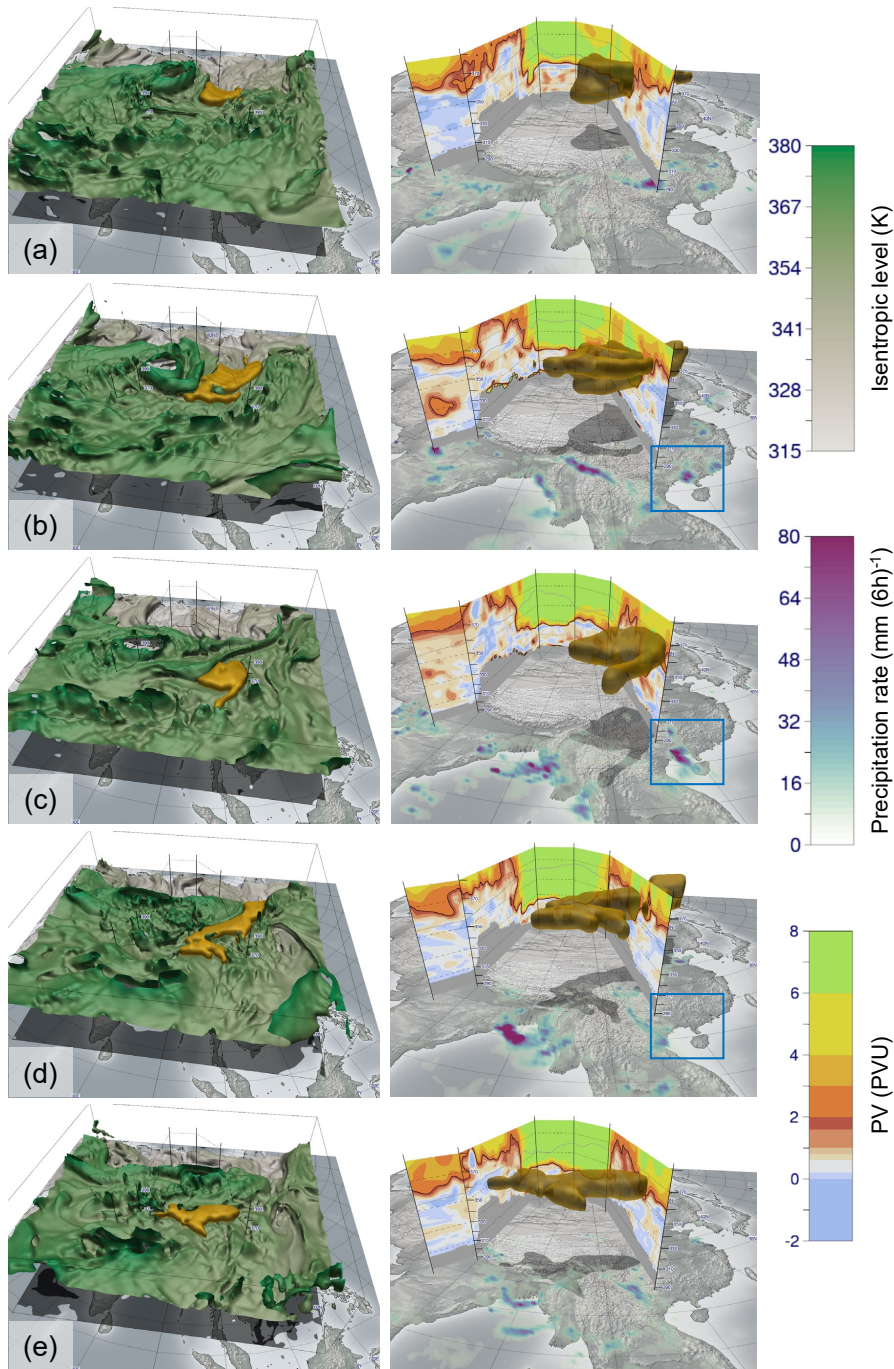
**Figure 13.** The synoptic-dynamic development associated with the event. Here, five time steps are visualized: (a) 22 July, 12 UTC; (b) 24 July, 00 UTC; (c) 27 July, 00 UTC; (d) 30 July, 12 UTC; (e) 31 July, 12 UTC. Respectively on the left, a 3-D visualization of the dynamical tropopause shaded by isentropic height, and the identified anomaly of interest is shown. On the right, the cross-section as displayed in Fig. 12c is shown. The blue box in (b-d) emphasizes the torrential rainfall triggering the flood. An animation with an extended time frame is provided at Fischer et al. (2021b).
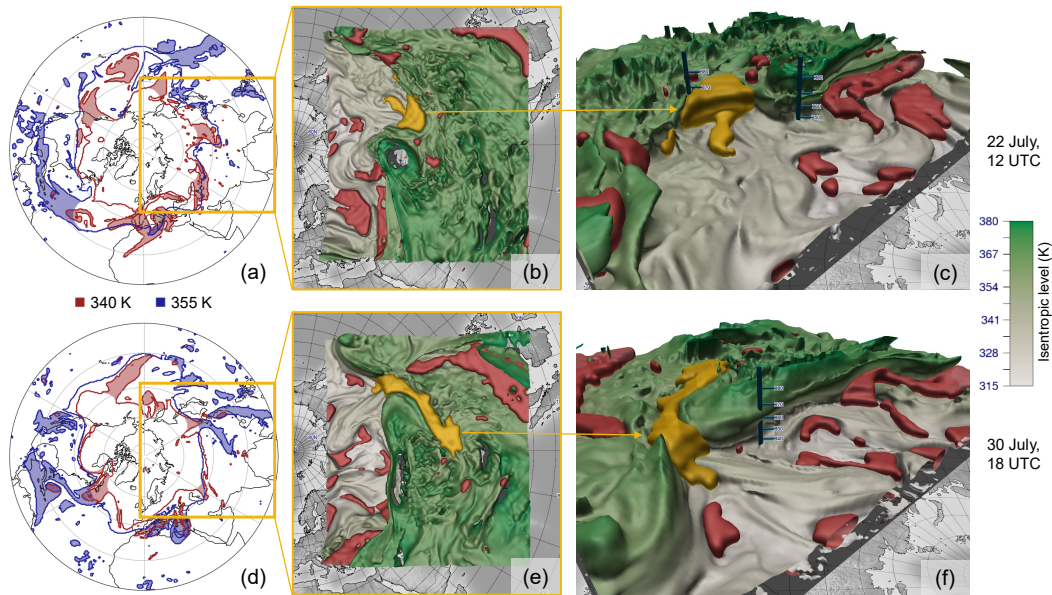
**Figure 14.** Comparison of the 340 K and 355 K isentrope, and the corresponding 3-D analysis for two time steps throughout the case study. (a): The 2-D analysis at 22 July, 12 UTC, where the anomaly is identified on 340 K (red), but not on 355 K (blue); (b): top-down view of the 3-D analysis; and (c): side view of the 3-D anomaly. (d-f): Same as (a-c), but 8 days later. The anomaly is now identifiable on 355 K, but not anymore on 340 K. Therefore, a 2-D analysis of the event considering a single isentrope is insufficient.

clearly distinct on the 340 K level. Because our 3-D technique (Fig. 14b, c) takes the whole 3-D neighborhood into account, the identified object covers both selected levels. Just looking at single 2-D isentropes individually is not sufficient to grasp the vertical structure of the PV object. Later in the event (Fig. 14d-f) the same anomaly that is clearly identifiable in 3-D can
550 be detected on the 355 K level, but not anymore on the 340 K level using the 2-D identification. ~~Therefore, the PVA moved vertically upwards throughout the event, making it impossible to automatically analyze and track when using a single isentropic level.~~ This also falls into place when analyzing the height of the 3-D PVA's centroid ~~. During genesis, the centroid~~ (position weighted by volume and PV, see Sect. 4.4). During genesis, the centroid of the PVA has a height of 345 K, and in later stages of the event, it is situated at over 360 K. Therefore, the PVA moved vertically upward throughout the event, making it impossible
555 to automatically analyze and track when considering a single isentropic level.

The case study shows as well that distinguishing between the concepts of a PVS and PV cutoff in 2-D is not applicable to the 3-D view. The vast majority of cutoffs are attached vertically to the main stratospheric reservoir, while PVS are connected typically both horizontally and vertically. Furthermore, a ~~clear~~ structure could often be interpreted as a streamer or a cutoff depending on the isentropic cross-section one is investigating. Visually, it can become unclear at all, which type a 3-D struc-
560 ture more belongs to. This supports our suggested nomenclature to take a more abstract approach and name these intrusions anomalies (PVAs).

## 8 Conclusions

In this study, a novel automated identification strategy for PVAs along the dynamical tropopause has been presented. It is based on sequential distance measures in the PV field to extract anomaly-like structures. The method is capable of identifying these structures in both 2-D and 3-D data sets robustly. It is based on image processing operations, which are able to detect disturbances in a multidimensional field, and has been adapted to suit the needs and requirements of the present application. Specifically, the adaptation keeps the interpretability of the process by using physical units, and compensates the distortions introduced by the particular flattened projections used. Here, we use the Eikonal equation and reinterpret it in a novel fashion to efficiently compute accurate distances in the distorted field, taking advantage of the properties of the projection. The strategy has proven well scalable for processing big data sets. The presented algorithm takes a width parameter $w$ as input, which controls the maximum width of anomalies being identified. Identified anomalies are assigned feature vectors containing metrics for the specific object, e.g., centroid, bounding box, intensity or best-fit ellipsoid. The objects are then filtered based on these metrics contained in the feature vectors.

A drawback of existing 2-D identification techniques is that their decisions are solely based on the outermost 2-PVU contour for a given isentropic level. It is shown that this hides important information about the structure of the field in some cases. By applying our algorithm, which considers the whole field, we therefore identify additional features not being easily captured by most algorithms described in the literature. The algorithm can be executed in the same fashion for 2-D and 3-D data sets since the individual image processing operations are independent of the number of dimensions. One main issue of the 3-D identification consists of clearly identifying separate anomalies in the 3-D PV field. We therefore added a heuristic that separates anomalies by detaching them at points of weakest vertical extent for a clear separation. This is important for meaningful feature descriptions afterwards.

In 3-D analysis, the algorithm provides an independent configuration for all seasons and use cases, and is therefore suitable for climatologies. In 2-D analysis on the other hand, care has to be taken of choosing a suitable isentrope for analysis during a season or for a specific event, or depending on which regimes one is interested in. Also more interesting features can be detected by considering the full 3-D neighborhood instead of focusing on a single 2-D level. The 3-D structures are more cohesive and consistent in space and time compared to their 2-D counterparts. ~~Popping artifacts, which appear in 2-D identification when~~ Furthermore, vertical gaps introduced by stacking results of ~~individual levels,~~ the 2-D identification are prevented.

A case study has been investigated to show up advantages of the 3-D analysis. A trough detected using single-level geopotential height or vertically averaged PV as in Van der Linden et al. (2017), which was instrumental in the evolution of the presented extreme event, would dissipate too early when compared with the results using the novel approach. Comparisons of 2-D and 3-D identification results reveal that the 3-D object is present both over a broader vertical range and longer time period. Especially the longer period coincides more closely with the observed rainfall than the 2D analysis by Van der Linden et al. (2017). Therefore, the novel method provides a more detailed view of the dynamical forcings of the extreme event by taking into account the vertical evolution and movement. This supports conclusions by Bithell et al. (1999). They emphasized that

595  without a full 3-D view of the developing system the extent of features, such as tropopause folds, and their depth of penetration into the troposphere, are difficult to follow.

We put effort in describing the identified object in an abstract geometric fashion, similar to ellipses in 2-D. There, most structures can be well approximated by one (or in degenerative cases a few) simple geometric shapes. These are also clearly interpretable, e.g., the main axis yields the tilt of the streamer. For 3-D objects, we considered using subsets of quadrics

600  ~~(second order surfaces, see Krivoshapko and Ivanov, 2015)~~ (second order surfaces, see Krivoshapko and Ivanov (2015)) but recognized that the extension of 2-D fitting to 3-D data is a highly non-trivial task. To explore the geometrical structure, but also to exploratorily analyze the presented algorithm and the data set itself, interactive 3-D visualizations proved to be an essential tool for comprehensive depictions. Derived displays, such as cross-sections, are vital to highlight dynamical processes in complex environments and to perform coherent conclusions.

605  In the future, this technique can be applied both to more individual use cases and to big data sets. Analyses and climatologies help to exploratorily examine the properties and behavior of PVAs in certain 3-D environments. Computed feature descriptions could be a base to find correlations and clusters in data, classifying anomalies into distinct categories. Since the algorithm is well scalable, it can be applied to ensemble forecasts as well. However, the performance bottleneck for big data analysis is the computation of high resolution isentropic data out of model levels. Therefore, processing large data sets might require

610  a trade-off by using a different vertical dimension even though this leads to a more challenging interpretation, especially when evaluating horizontal cross-sections. Although this study focuses on the identification on anomalies along the dynamical tropopause, the algorithm in a broader sense can be applied to identify anomalies along any 2-D or 3-D spatial boundary. Therefore, other identification tasks in meteorology can also benefit from these ideas and strategies.

Further research is required on certain aspects of this identification, which might lead to improved results. For individual

615  use cases, the tracking of these anomalies could be done by hand, whereas for climatologies automated tracking techniques are required. Ideas for further work include simple overlap tracking (e.g., Limbach et al., 2012), or a Lagrangian view (e.g., Portmann et al., 2021). The feature descriptions, which have been computed for each PVA, might serve as base for similarity measures and be used for tracking purposes.

The automated 3-D PV identification, tracking, and its application to many meteorological cases opens an avenue to study,

620  which characteristics of the PV objects are related to the genesis and improved forecast using statistical and machine learning approaches. For example, Maier-Gerber et al. (2021) successfully used ensemble mean and standard deviation of vertically averaged upper-level PV forecasts in the Gulf of Mexico as a predictor for subseasonal statistical-dynamical forecasting of tropical cyclone occurrence. Here, a 3-D PV object related approach holds the promise of further improvements. These abstracted objects can also facilitate combinations with other displays, e.g., other identified atmospheric features, to explore

625  further connections.

*Code and data availability.* The implementation of our framework is available at https://gitlab.physik.uni-muenchen.de/Christoph.Fischer/enstools-feature (where it will be updated in future studies) and archived at https://doi.org/10.5281/zenodo.5638561 (Fischer et al., 2021a). It is real-

ized as part of the open source framework *enstools*, available at https://github.com/wavestoweather/enstools. The data sets used in this study (ERA5 reanalysis, S2S) are publicly available at the relevant cited sources.

630  *Video supplement.*  A video supplement showing 3-D visualizations of the synoptic evolution of our case study can be freely accessed at https://doi.org/10.5281/zenodo.5639001 (Fischer et al., 2021b). The visualizations are generated using Met.3D (Rautenhaus et al., 2015b).

*Author contributions.*  ES, AF and M Riemer proposed, supervised and administrated this study. CF designed the algorithm, implemented the software, performed algorithm analyses, did visualizations and wrote the publication. RL performed meteorological analyses on the case study and provided according results. M Rautenhaus contributed helpful input and comments regarding Met.3D and visualizations in general.
635  All authors provided feedback and critical review on the paper.

*Competing interests.*  The authors declare that they have no conflict of interest.

*Acknowledgements.*  The research leading to these results has been accomplished within the project C3 "Predictability of tropical and hybrid cyclones over the North Atlantic Ocean" of the Transregional Collaborative Research Center SFB/TRR 165 "Waves to Weather" funded by the German Science Foundation (DFG). This work uses S2S data. S2S is a joint initiative of the World Weather Research Programme
640  (WWRP) and the World Climate Research Programme (WCRP). The original S2S database is hosted at ECMWF as an extension of the TIGGE database. The GPM (IMERG) data were provided by the NASA/Goddard Space Flight Center, and archived at the NASA GES DISC. Thanks to the Institute for Atmospheric and Climate Science at ETH Zurich, in particular to Michael Sprenger and Heini Wernli, for sharing the source code of their PV identification technique (Wernli and Sprenger, 2007).

# References

645 Bader, R., Sprenger, M., Ban, N., Radisuhli, S., Schär, C., and Günther, T.: Extraction and Visual Analysis of Potential Vorticity Banners around the Alps, IEEE T. Vis. Comput. Gr., pp. 1–1, https://doi.org/10.1109/TVCG.2019.2934310, 2019.

Bennetts, D. A. and Hoskins, B.: Conditional symmetric instability-a possible explanation for frontal rainbands, Q. J. Roy. Meteor. Soc., 105, 945–962, https://doi.org/10.1002/qj.49710544615, 1979.

Bentley, A., Bosart, L., and Keyser, D.: Upper-Tropospheric Precursors to the Formation of Subtropical Cyclones that Undergo Tropical
650 Transition in the North Atlantic Basin, Mon. Weather Rev., 145, 503–520, https://doi.org/10.1175/MWR-D-16-0263.1, 2017.

Bithell, M., Gray, L. J., and Cox, B. D.: A Three-Dimensional View of the Evolution of Midlatitude Stratospheric Intrusions, J. Atmos. Sci., 56, 673 – 688, https://doi.org/10.1175/1520-0469(1999)056<0673:ATDVOT>2.0.CO;2, 1999.

Cohen, L. D. and Kimmel, R.: Global minimum for active contour models: A minimal path approach, Int. J. Comput. Vision, 24, 57–78, https://doi.org/10.1023/A:1007922224810, 1997.

655 Dawe, J. T. and Austin, P. H.: Statistical Analysis of an LES Shallow Cumulus Cloud Ensemble Using a Cloud Tracking Algorithm, Atmos. Chem. Phys., 12, 1101–1119, https://doi.org/10.5194/acp-12-1101-2012, 2012.

Dijkstra, E. W. et al.: A note on two problems in connexion with graphs, Numer. Math., 1, 269–271, https://doi.org/10.1007/BF01386390, 1959.

Dougherty, E.: Digital Image Processing Methods, CRC Press, 2020.

660 Fischer, C., Fink, A. H., Schömer, E., Van der Linden, R., Maier-Gerber, M., Rautenhaus, M., and Riemer, M.: A novel method for objective identification of 3-D potential vorticity anomalies - Implementation, https://doi.org/10.5281/zenodo.5638561, 2021a.

Fischer, C., Rautenhaus, M., Fink, A. H., Schömer, E., Van der Linden, R., Maier-Gerber, M., and Riemer, M.: A novel method for objective identification of 3-D potential vorticity anomalies - Visualizations using Met.3D, https://doi.org/10.5281/zenodo.5639001, 2021b.

Furtney, J.: scikit-fmm, https://github.com/scikit-fmm/scikit-fmm, 2019.

665 Galarneau Jr, T. J., McTaggart-Cowan, R., Bosart, L. F., and Davis, C. A.: Development of North Atlantic tropical disturbances near upper-level potential vorticity streamers, J. Atmos. Sci., 72, 572–597, https://doi.org/10.1175/JAS-D-14-0106.1, 2015.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E.,
670 Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hewson, T. D. and Titley, H. A.: Objective Identification, Typing and Tracking of the Complete Life-Cycles of Cyclonic Features at High Spatial Resolution, Meteorol. Appl., 17, 355–381, https://doi.org/10.1002/met.204, 2010.

Hoskins, B. J., McIntyre, M. E., and Robertson, A. W.: On the use and significance of isentropic potential vorticity maps, Q. J. Roy. Meteor.
675 Soc., 111, 877–946, https://doi.org/10.1002/qj.49711147002, 1985.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., and Tan, J.: Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation, NASA/GSFC Code, 612, 2019, 2015.

Kern, M., Hewson, T., Sadlo, F., Westermann, R., and Rautenhaus, M.: Robust Detection and Visualization of Jet-stream Core Lines in Atmospheric Flow, IEEE T. Vis. Comput. Gr., 24, 893–902, https://doi.org/10.1109/tvcg.2017.2743989, 2018.

680    Kern, M., Hewson, T., Schatler, A., Westermann, R., and Rautenhaus, M.: Interactive 3D Visual Analysis of Atmospheric Fronts, IEEE T.
Vis. Comput. Gr., 25, 1080–1090, https://doi.org/10.1109/TVCG.2018.2864806, 2019.

Koch, P., Wernli, H., and Davies, H. C.: An event-based jet-stream climatology and typology, Int. J. Climatol., 26, 283–301,
https://doi.org/10.1002/joc.1255, 2006.

Krivoshapko, S. N. and Ivanov, V. N.: The Second Order Surfaces, pp. 613–626, Springer International Publishing, Cham,
685    https://doi.org/10.1007/978-3-319-11773-7_35, 2015.

Kunz, A., Sprenger, M., and Wernli, H.: Climatology of potential vorticity streamers and associated isentropic transport pathways across PV
gradient barriers, J. Geophys. Res.-Atmos., 120, 3802–3821, https://doi.org/10.1002/2014JD022615, 2015.

Lamarque, J.-F. and Hess, P. G.: Cross-tropopause mass exchange and potential vorticity budget in a simulated tropopause folding, J. Atmos.
Sci., 51, 2246–2269, https://doi.org/10.1175/1520-0469(1994)051<2246:CTMEAP>2.0.CO;2, 1994.

690    Limbach, S., Schömer, E., and Wernli, H.: Detection, tracking and event localization of jet stream features in 4-D atmospheric data, Geosci.
Model Dev., 5, 457–470, https://doi.org/10.5194/gmd-5-457-2012, 2012.

Maier-Gerber, M., Riemer, M., Fink, A. H., Knippertz, P., Muzio, E. D., and McTaggart-Cowan, R.: Tropical Transition of Hurri-
cane Chris (2012) over the North Atlantic Ocean: A Multiscale Investigation of Predictability, Mon. Weather Rev., 147, 951 – 970,
https://doi.org/10.1175/MWR-D-18-0188.1, 2019.

695    Maier-Gerber, M., Fink, A. H., Riemer, M., Schoemer, E., Fischer, C., and Schulz, B.: Statistical–Dynamical Forecasting of Subseasonal
North Atlantic Tropical Cyclone Occurrence, Weather Forecast., 36, 2127 – 2142, https://doi.org/10.1175/WAF-D-21-0020.1, 2021.

Massacand, A. C., Wernli, H., and Davies, H. C.: Heavy precipitation on the Alpine southside: An upper-level precursor, Geophys. Res. Lett.,
25, 1435–1438, https://doi.org/10.1029/98GL50869, 1998.

McIntyre, M. E. and Palmer, T.: Breaking planetary waves in the stratosphere, Nature, 305, 593–600, https://doi.org/10.1038/305593a0,
700    1983.

Molinari, J., Skubis, S., Vollaro, D., Alsheimer, F., and Willoughby, H. E.: Potential vorticity analysis of tropical cyclone intensification, J.
Atmos. Sci., 55, 2632–2644, https://doi.org/10.1175/1520-0469(1998)055<2632:PVAOTC>2.0.CO;2, 1998.

Morgan, M. C. and Nielsen-Gammon, J. W.: Using Tropopause Maps to Diagnose Midlatitude Weather Systems, Mon. Weather Rev., 126,
2555 – 2579, https://doi.org/10.1175/1520-0493(1998)126<2555:UTMTDM>2.0.CO;2, 1998.

705    Mukundan, R. and Ramakrishnan, K. R.: Moment Functions in Image Analysis — Theory and Applications, WORLD SCIENTIFIC,
https://doi.org/10.1142/3838, 1998.

Papin, P. P., Bosart, L. F., and Torn, R. D.: A Feature-Based Approach to Classifying Summertime Potential Vorticity Streamers Linked to
Rossby Wave Breaking in the North Atlantic Basin, J. Climate, 33, 5953 – 5969, https://doi.org/10.1175/JCLI-D-19-0812.1, 2020.

Pearson, F.: Map Projections: Theory and Applications (1st ed.), CRC Press, https://doi.org/10.1201/9780203748121, 1990.

710    Petres, C., Pailhas, Y., Petillot, Y., and Lane, D.: Underwater path planing using fast marching algorithms, in: Europe Oceans 2005, vol. 2,
pp. 814–819 Vol. 2, https://doi.org/10.1109/OCEANSE.2005.1513161, 2005.

Portmann, R., Crezee, B., Quinting, J., and Wernli, H.: The complex life cycles of two long-lived potential vorticity cut-offs over Europe, Q.
J. Roy. Meteor. Soc., 144, 701–719, https://doi.org/10.1002/qj.3239, 2018.

Portmann, R., Sprenger, M., and Wernli, H.: The three-dimensional life cycles of potential vorticity cutoffs: a global and selected regional
715    climatologies in ERA-Interim (1979–2018), Weather Clim. Dynam., 2, 507–534, https://doi.org/10.5194/wcd-2-507-2021, 2021.

Rautenhaus, M., Grams, C. M., Schäfler, A., and Westermann, R.: Three-Dimensional Visualization of Ensemble Weather Forecasts – Part 2: Forecasting Warm Conveyor Belt Situations for Aircraft-Based Field Campaigns, Geosci. Model Dev., 8, 2355–2377, https://doi.org/10.5194/gmd-8-2355-2015, 2015a.

Rautenhaus, M., Kern, M., Schäfler, A., and Westermann, R.: Three-dimensional visualization of ensemble weather forecasts – Part 1: The visualization tool Met.3D (version 1.0), Geosci. Model Dev., 8, 2329–2353, https://doi.org/10.5194/gmd-8-2329-2015, 2015b.

Rautenhaus, M., Böttinger, M., Siemen, S., Hoffman, R., Kirby, R. M., Mirzargar, M., Röber, N., and Westermann, R.: Visualization in meteorology—a survey of techniques and tools for data analysis tasks, IEEE T. Vis. Comput. Gr., 24, 3268–3296, https://doi.org/10.1109/TVCG.2017.2779501, 2018.

Redl, R., Keil, C., Craig, G., Lerch, S., and Eichhorn, J.: Towards a Framework for Parallelized Post-Processing and Evaluation of Ensemble Forecasts, in: EGU General Assembly Conference Abstracts, EGU General Assembly Conference Abstracts, p. 12322, 2018.

Schulzweida, U.: CDO User Guide, https://doi.org/10.5281/zenodo.3539275, 2019.

Sethian, J. A.: A fast marching level set method for monotonically advancing fronts, P. Natl. Acad. Sci. USA, 93, 1591–1595, https://doi.org/10.1073/pnas.93.4.1591, 1996.

Snyder, J. P.: Map projections–A working manual, vol. 1395, USGS, 1987.

Sprenger, M., Martius, O., and Arnold, J.: Cold surge episodes over southeastern Brazil – a potential vorticity perspective, Int. J. Climatol., 33, 2758–2767, https://doi.org/10.1002/joc.3618, 2013.

Sprenger, M., Fragkoulidis, G., Binder, H., Croci-Maspoli, M., Graf, P., Grams, C. M., Knippertz, P., Madonna, E., Schemm, S., ?kerlak, B., and Wernli, H.: Global Climatologies of Eulerian and Lagrangian Flow Features based on ERA-Interim, B. Am. Meteorol. Soc., 98, 1739 – 1748, https://doi.org/10.1175/BAMS-D-15-00299.1, 2017.

Thorncroft, C. D., Hoskins, B. J., and McIntyre, M. E.: Two paradigms of baroclinic-wave life-cycle behaviour, Q. J. Roy. Meteor. Soc., 119, 17–55, https://doi.org/10.1002/qj.49711950903, 1993.

Van der Linden, R., Fink, A. H., Pinto, J. G., and Phan-Van, T.: The dynamics of an extreme precipitation event in northeastern Vietnam in 2015 and its predictability in the ECMWF ensemble prediction system, Weather Forecast., 32, 1041–1056, https://doi.org/10.1175/WAF-D-16-0142.1, 2017.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., Déqué, M., Ferranti, L., Fucile, E., Fuentes, M., et al.: The subseasonal to seasonal (S2S) prediction project database, B. Am. Meteorol. Soc., 98, 163–173, https://doi.org/10.1175/BAMS-D-16-0017.1, 2017.

Weijenborg, C., Chagnon, J., Friederichs, P., Gray, S., and Hense, A.: Coherent evolution of potential vorticity anomalies associated with deep moist convection, Q. J. Roy. Meteor. Soc., 143, 1254–1267, https://doi.org/10.1002/qj.3000, 2017.

Wernli, H. and Sprenger, M.: Identification and ERA-15 Climatology of Potential Vorticity Streamers and Cutoffs near the Extratropical Tropopause, J. Atmos. Sci., 64, 1569 – 1586, https://doi.org/10.1175/JAS3912.1, 2007.

Škerlak, B., Sprenger, M., Pfahl, S., Tyrlis, E., and Wernli, H.: Tropopause folds in ERA-Interim: Global climatology and relation to extreme weather events, J. Geophys. Res.-Atmos., 120, 4860–4877, https://doi.org/10.1002/2014JD022787, 2015.