DOI: 10.1111/cgf.14543
Eurographics Conference on Visualization (EuroVis) 2022
R. Borgo, G. E. Marai, and T. Schreck
(Guest Editors)

Volume 41 (2022), Number 3

An Interactive Approach for Identifying Structure Definitions

Natalia Mikula[1], Tom Dörffel[2], Daniel Baum[1] and Hans-Christian Hege[1]

[1] Zuse Institute Berlin (ZIB), Germany[†]
[2] Freie Universität Berlin, Germany[‡]

Abstract

Our ability to grasp and understand complex phenomena is essentially based on recognizing structures and relating these to each other. For example, any meteorological description of a weather condition and explanation of its evolution recurs to meteorological structures, such as convection and circulation structures, cloud fields and rain fronts. All of these are spatiotemporal structures, defined by time-dependent patterns in the underlying fields. Typically, such a structure is defined by a verbal description that corresponds to the more or less uniform, often somewhat vague mental images of the experts.

However, a precise, formal definition of the structures or, more generally, of the concepts is often desirable, e.g., to enable automated data analysis or the development of phenomenological models. Here, we present a systematic approach and an interactive tool to obtain formal definitions of spatiotemporal structures. The tool enables experts to evaluate and compare different structure definitions on the basis of data sets with time-dependent fields that contain the respective structure. Since structure definitions are typically parameterized, an essential part is to identify parameter ranges that lead to desired structures in all time steps. In addition, it is important to allow a quantitative assessment of the resulting structures simultaneously. We demonstrate the use of the tool by applying it to two meteorological examples: finding structure definitions for vortex cores and center lines of temporarily evolving tropical cyclones.

Ideally, structure definitions should be objective and applicable to as many data sets as possible. However, finding such definitions, e.g., for the common atmospheric structures in meteorology, can only be a long-term goal. The proposed procedure, together with the presented tool, is just a first systematic approach aiming at facilitating this long and arduous way.

CCS Concepts

• ***Computing methodologies*** → *Modeling methodologies;* • ***Human-centered computing*** → *Scientific visualization; Visual analytics;* • ***Applied computing*** → *Earth and atmospheric sciences;*

1. Introduction

In order to understand a phenomenon hidden in data, we need to recognize structures and relationships between them. Structures are characterized by defining qualities and features. In this work, we deal with structures for which a mental image does exist but no mathematical definition. For the sake of clarity, we distinguish between the semantics of a concept, i.e., anything that defines and characterizes it, and the linguistic term used to denote it.

In all areas of daily life as well as in all sciences, concepts are needed for thinking and communicating. This can be explained very nicely using the example of meteorology: The description of a weather situation and its development over time makes use of technical terms such as 'pressure systems' (high or low), 'fronts' (warm, cold or rain), 'clouds', 'precipitation', 'jet streams' and so on. Of course, in technical language, much more differentiating terms are used. For example, there is a hierarchical classification system for clouds that distinguishes more than 30 types of clouds by classifying them according to altitude and vertical shape, instability or convection activity, and specific structural details [WMO].

One could counter that in strictly rational science such mental constructs are dispensable; for example, even a very complex meteorological state is *completely* described by a set of fields, which is governed by a set of equations. In that sense, given these fields in appropriate resolution, there is nothing more to say. What this counterargument overlooks is that humans need both categorizing and characterizing concepts of structures, and associated linguistic terms, in order to think and talk about the phenomena as well as to formulate hypotheses and theories about them. Therefore, such mental constructs are fundamental building blocks of human knowledge and are used practically everywhere.

It is an essential part of science to sharpen concepts, to differentiate them more precisely, and to categorize them more meaningfully. Yet, for most concepts, only verbal descriptions exist that are not very precise. Just consider questions like these: "Are we in the

† www.zib.de, email: {mikula, baum, hege}@zib.de
‡ www.fu-berlin.de, email: tom.doerffel@fu-berlin.de

fourth *wave* of an epidemic?", or "Is there a *vortex* in a given flow and what is its spatial extent?", or questions about the presence of complex, multifaceted conditions, for example in medicine. With regards to the first question, although everyone has an idea of an 'epidemic wave', different mathematical criteria are conceivable that capture the mental construct more precisely (c.f. Sect. 3.2). Also, everyone has a vivid idea of what a flow vortex is; nevertheless, even decades of research have not led to a full agreement on what is the best criterion for determining the presence and extent of a flow vortex (c.f. Sect. 4).

However, if data science (including data visualization) wants to answer questions on the basis of data, precise and formalized definitions of such concepts are required. Especially in view of the fact that the increasing amount of data requires more and more automatic processing, which also calls for machine-understandable definitions of concepts. Precise mathematical concept definitions are also necessary for creating phenomenological models, i.e., mathematical models that contain such concepts as building blocks.

The question is: How can one build a bridge from the world of mental concepts, which are at best described verbally, to the world of precise, formal definitions? To do this, one first needs a deep and ideally also mathematically oriented insight into the problem domain. This often enables experts to suggest a first guess for a formal definition. But, given a suggestion for such a definition, how do you assess its quality and how do you improve it? Here, visualization helps: it allows the comparison of structures that result from data and experimentally formulated formal definitions with the mental image of that structure. This enables an iterative improvement of formal structure definitions, considering also the mental images of different people or even an entire expert community (c.f. [HHK11]).

The long-term goal (which might take decades to achieve) is to find definitions that are as 'objective' as possible and that work for as many data sets as possible. We give advice on which principles should be observed in this regard (see Sect. 3.1). An essential role plays the selection of suitable indicator quantities that show the presence of a structure. Another problem is that only in exceptional cases definitions can be found that are parameter-free. For example, many definitions contain threshold values. How should these be determined?

In this paper, we describe a first attempt to a *systematic approach* for finding such definitions in a visually supported, interactive way. The presented tool makes it possible to compare various definitions, indicator quantities, computational methods, and parameter ranges and thus allows us to achieve suitable definitions in an iterative process. It can be used both for a quick, pragmatic search for a suitable definition for specific data sets, as well as for the long-term goal of finding a definition that is as universal as possible.

We demonstrate the use of the tool by applying it to two meteorological examples: finding structure definitions for the vortex cores and centerlines of temporally evolving tropical cyclones.

Overall the paper presents a systematic approach and workflow as well as a prototypical interactive tool that enables expert users

- to find formal definitions of temporal structures, so that their properties correspond for all time steps as closely as possible to those of the mental image;
- to evaluate the suitability of different variables that indicate the presence of a structure;
- to narrow down the intervals of the parameters in a structure definition such that the structure properties match the mental image, and the variance of the resulting structures becomes as small as possible.

The proposed tool utilizes parallel coordinates with time as one variable to simultaneously depict parameter intervals and structure attributes for selected time intervals. In contrast to previous approaches for interactive structure extraction (e.g., using brushing & linking), it directly supports the evaluation of all influencing factors, namely structure definitions (including computational methods for structure extraction), type of indicator quantities, and parameter intervals.

The paper is organized as follows: In Sect. 2, we present related work. Sect. 3 focuses on describing the overall systematic approach as well as the prototype tool we implemented. In Sect. 4, we apply our approach and tool to two use cases from meteorology. Finally, Sect. 5 concludes the paper with a discussion of our findings and potential future work.

2. Related work

Concepts are the building blocks of thoughts and are necessary for the representation of knowledge as well as for mental processes like categorization, inference, decision making, learning and communication. This view has become widely accepted in cognitive science and philosophy of science, see e.g., [Gär04, ZG15]. In this paper we perceive concepts as mental constructs (as opposed to abstract objects or skills) and we use a naive notion of concepts based on concrete examples, thus avoiding complex philosophical questions (see, e.g., reference [ML21] for an introduction to this topic).

The extraction of structures from data has a long history in data visualization. The structures are often mathematical features. Most prominent examples are mathematical features of spatial or spatiotemporal fields, like extremal structures, level sets, critical points and skeletons of stable manifolds. Therefore, the term 'feature extraction' has become established in visualization. When dealing with application-oriented structures that carry semantics, such as jet streams, precipitation fields or tropical storms, it is better to speak of 'structures' or 'concepts'. These are characterized by certain properties, which may include also mathematical features. This terminology corresponds to that used in many other sciences, but differs from that traditionally used in visualization, where application-oriented structures are often also referred to as 'features'.

For a survey on the substantial history of structure extraction and visualization, see reviews on geometric-topological objects [MLP*10, PPF*11], on coherent set detection [HFB*17], on feature tracking in meteorological contexts [CBJ*14, CDJ15, VMN*18, EMB*21], and on the related problem of image segmentation [EBJS17]. Most approaches in visualization refer to structures defined by single scalar or vector fields, e.g., level sets or ridges of Lyaponov exponents, or application-related structures, like centerlines of vortices [SWH05a] or jet streams [KHS*18].

Data visualization often focuses on isolating application-defined regions of interest in spatiotemporal data. This is done by interactively identifying subregions in a typically multidimensional space spanned by data attributes that best characterize the structures of interest [JH18]. Related procedures have been developed for image segmentation and volume rendering, where transfer functions map data characteristics to optical qualities, like color and opacity, to highlight regions of interest [KKH01, LKG*16].

An essential prerequisite for finding suitable structure definitions with visual control is interactivity. Here we can build on the research thread that began with the work by Doleisch et al. [DGH03]. In this work, the definition of subregions in a multivariate range space is greatly facilitated by multiple coordinated views (histograms, scatter plots, parallel coordinates, or function graphs), representing different variables side by side. When data points are interactively selected ('brushed') in one view, the associated data items are immediately highlighted in all linked views. This 'linking' is particularly powerful when attribute views are combined with 3D views, showing the corresponding preimages in the domain. This technique can also be extended to time-dependent data [DMG*04, DHGK06] and has been applied to simulation data of Hurricane Isabel [DMH04]. While in their work all structures are defined by ranges of multivariate quantities, it is not concerned with finding more general and also parameter-dependent *definitions* of structures.

A tool for interactive exploration of sets of parameter-dependent 3D geometries has been presented in [BHGK14]. However, this work also did not aim at obtaining structural definitions. To the best of the authors' knowledge, there is no tool to date that supports the explicit goal of obtaining mathematical definitions of mentally and verbally given structures by using visual data analysis techniques.

Visual comparison of similar, but different spatial data has been more deeply explored in the context of ensemble visualization, see, e.g., [KBVH17] and references therein.

The difficulty of finding a generally accepted, precise definition of concepts that is as objective as possible (i.e., independent of the person observing) is best illustrated by the example of flow vortices. Data visualization, together with flow research, has contributed a great deal to this. There is not enough space here, to sketch the nearly half-century-long research thread; instead we refer to the current state of research [THR*21] and the papers cited therein. In recent work [vLHD*21], a formal approach to the definition and persistence of meteorological structures has been taken. This is exactly the kind of research we want to support and facilitate with the proposed procedure and interactive tool.

3. Structure definition finder

In the following, we first describe the proposed general procedure to identify structure definitions based on the user's mental image, followed by an illustrative example. Then, we present the prototypical interactive tool that we apply to concrete examples in Sect. 4.

3.1. Methodological approach

The proposed systematic approach is graphically represented in Fig. 1. It depicts an iterative workflow that allows the user to identify and further narrow down appropriate structure definitions as more data or new ideas about structure attributes are added. The workflow starts with a verbal characterization of the structure that we want to define more precisely in mathematical terms.

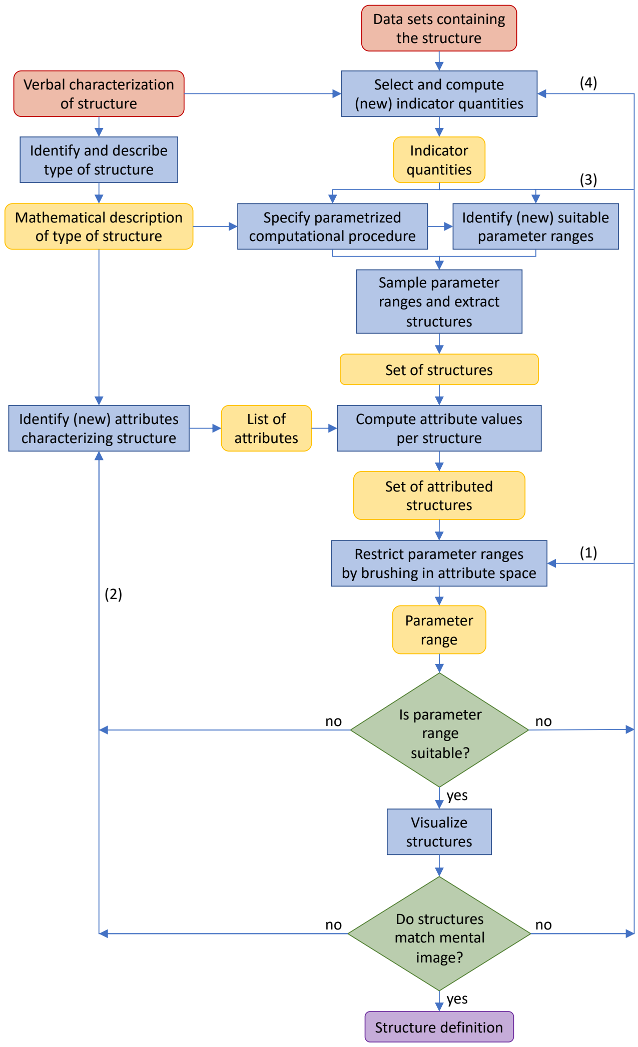

Figure 1: *Workflow for finding structure definitions, as described in Sect. 3.1. Red denotes inputs, yellow - intermediate results, blue - tasks to be carried out, green - decisions, and purple - the final result.*

Based on the verbal characterization and the knowledge of mathematical structures, the user can describe the structure using mathematical terms, e.g., as a point, a line, an area, a tree, a graph, etc. The mathematical type of structure already restricts the class of mathematical methods for structure identification. Eligible are, for example, methods of geometric data analysis (e.g., extraction of level sets or extremal structures), shape analysis (e.g., extraction of shapes that resemble a prototypical shape), cluster analysis and classification (e.g., computation of a hierarchical clustering), topological data analysis (e.g., computation of persistence diagrams), and time series analysis (e.g., identification of typical temporal patterns), as well as combinations of these. The class of procedures is

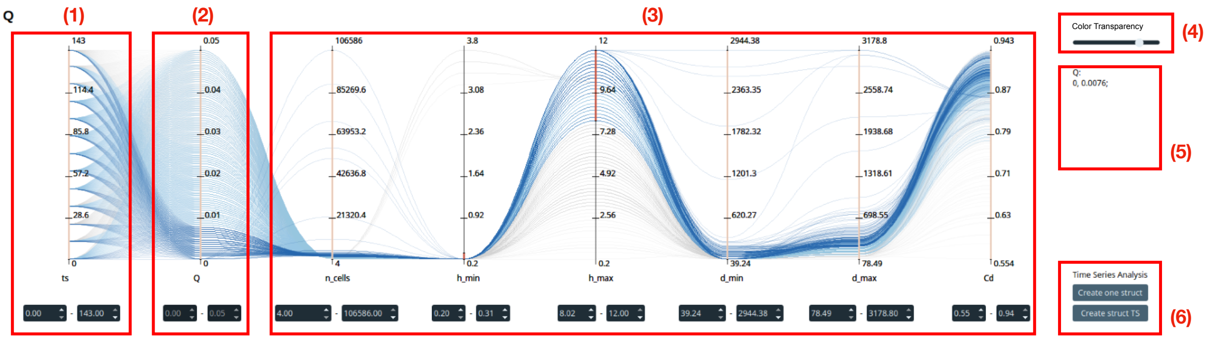

Figure 2: *Parallel coordinates window and its main areas: (1) time axis; (2) parameter space; (3) attribute space; (4) color/transparency settings; (5) quantitative representation of selected parameter intervals; and (6) buttons for interactive exploration of selected structures.*

relatively limited and it would be an interesting endeavor to characterize and systematize them in more detail (for a first approach in flow visualization, see [WST*07]). The next step is the identification of attributes, in our case, geometric and topological attributes, which are suited to describe the structure more or less completely and accurately. Of particular interest are attributes that allow the user to distinguish structures matching her/his mental image from structures that do not. This part of the workflow is shown on the left side of Fig. 1.

The second input to the workflow is data (typically discretely-sampled functions) containing the structure for which a definition is to be found. Sometimes, one does not work directly on the original input data, e.g., the velocity field, but on derived data, like Q or Ω (c.f. Sect. 4). For the sake of generality, let us call all data that indicate the presence of the desired structure 'indicator quantities', be it the original or derived data. These indicator quantities are usually well-established quantities in the respective application field and they can be limited by using application-based higher-level principles, e.g., invariances under transformations or invariance to the choice of physical dimensions. The latter is achieved by choosing dimensionless quantities. In order to extract the structure from one or more indicator quantities, computational procedures are needed. Similarly to the choice of indicator quantities, procedures should be selected that are well-established in the application field. Usually, a procedure depends on parameters that will influence the appearance of the extracted structure. An algorithmic structure definition thus consists of a computational procedure, one or more indicator quantities, and suitable computational parameters (or parameter ranges). For a recent survey on the definition, extraction and tracking of persistent structures in meteorology, see reference [vLHD*21].

Given potential computational procedures as well as indicator quantities for extracting the structure of interest, the task of finding a structure definition reduces to sampling the whole space of possibilities, including computational procedures, indicator quantities, and computational parameters, and comparing the outcomes with the user's mental image. Visually assessing all sampled structures would be very time-consuming. Therefore, one aims at reducing the total number of samples to those that are most likely to match the mental image. Here, the structure attributes come into play, which again can be highly application-specific. They are computed for each of the sampled structures and can then be used for filtering.

The restriction of the parameter space is done in the next step of the workflow, by interactively brushing the attribute space. As result, potential computational procedures, indicator quantities, and computational parameters are identified that lead to structures matching the selected structure attributes. Finally, the structures extracted in this way are visualized and visually assessed. Hence, visual inspection plays the final role of quality control. Since we have narrowed down the parameter space by brushing in the attribute space, visual assessment is reduced to a small set of structures. It could reveal the need to (1) further explore the parameter and attribute space, (2) identify other attributes that better describe the mental image, (3) identify other computational parameters and/or procedures, and/or (4) identify further indicator quantities (c.f. Fig. 1). If only a refinement of the attribute selection is necessary, after some iterations one might be done. Otherwise, new attribute descriptors and/or computational procedures will have to be implemented.

Finding a generalized structure definition requires a set of data containing the structure. This could be time steps of a single time series, as in the applications presented here, a data collection of a similar type, or even a set of time series. Ideally, the resulting structure definition should then work for all data in the given set.

The final step of finding a structure definition is to translate back the algorithmic definition given by the computational procedure and its parameters into precise mathematical expressions. For example, applying the Marching Cube algorithm to an indicator field is equivalent to applying a threshold operator to a scalar field. The result of this back translation is a mathematical structure definition consisting of mathematical expressions and, if these are parameterized, suitable parameter values or ranges.

The derivation of a structure definition from a mental image requires, in general, *interactive* exploration of the parameter and attribute space. However, for many applications, it is unfeasible to assume that all parts of the presented workflow (Sect. 3.1 and Fig. 1) are fully interactive. Hence, preprocessing may be required, in particular for the computation of the indicator quantities, the set of structures, and the computation of the structure attributes. Once this is done and all information has been precomputed, an interactive exploration tool is needed that allows one to browse the attribute space and visualize the matched structures in real time.

In the next section, we give an illustrative example of using the workflow to find a structure definition of epidemic waves before we describe a prototypical interactive exploration tool that we use in our applications (Sect. 4).

3.2. A first example: Epidemic waves

To illustrate this rather abstract description, we now outline the procedure using a concrete example, namely the aforementioned concept of an 'epidemic wave' (EW). For the sake of clarity, we will discuss only the essential steps and the major questions to be answered. Furthermore, we will aim for a simple solution. In particular, we will not use insights from mathematical modeling of infection events with ODEs/PDEs [BCCF19] or agent-based models [WZSC21].

An epidemic prevails if a disease caused by a pathogen spreads very rapidly to a large number of hosts in a population. Then infection rates, i.e., cases of infection per day, are elevated and their time course often resembles a crest of a wave—which lead to the metaphor of the EW.

If we consider that the course of infection can differ in different age groups, e.g., because of different contact behavior, and that at one point in time several pathogen mutants with different effects may be active, it becomes clear that the overall epidemic event is composed of several subwaves. The subwaves can be shifted in time, potentially leading to complex wave forms of the overall signal and impeding the separation of successive wave crests. The overall effect is reflected in cumulative, location-dependent infection rates, and for regions (cities, countries, etc.) in regionalized mean infection rates. For simplicity, let us consider just the last quantity. The mathematical structure of an EW is then a certain pattern in a scalar time series.

The question now is what would be a good indicator variable to determine whether or not a wave is present at a particular time. An important requirement is that this variable is independent of factors under which the definition should be invariant. For example, in the present case, the definition should be independent of population size. The most basic candidate for an indicator quantity therefore is the regionalized mean infection rate normalized by the size of its population. This dimensionless quantity allows comparison between different regions.

A more advanced approach would be to base the definition, as in reference [ZMGW21], on the estimated effective reproductive number R, i.e., the average number of people infected by a single infectious individual—and, if a mathematical model of the epidemic event is available, on further information the model provides.

The next question is what features of the time series could be used to base the extraction of periods in which an EW is prevalent. This is where the mental image particularly comes into play. One can start, e.g., from the somewhat vaguely formulated key criteria in reference [ZMGW21]: "1) an epidemic wave constitutes some upward and/or downward periods; 2) the increase in an upward period or the decrease in a downward period have to be substantial by sustaining over a period of time to distinguish them from an uptick, a downtick, reporting errors, or volatility in new cases."

The next step is to design and implement an extraction method based on well-defined mathematical operations. The resulting procedure, which will later be applied to time series data, will not be parameter-free.

Then, attributes of the extracted structures are defined. We need attributes that indicate how well requirements 1) and 2) are fulfilled but others can also be added, e.g. measures for the 'waviness', 'mountainousness', and 'pronouncedness' of an extracted structure, or features like strength of growth or decay, width and steepness of crest, etc. There are no limits to the imagination here; anything related to the mental image could potentially be useful.

The next step is to sample the parameter space of the method while applying it to (many) data sets. When analyzing time series of epidemic events, multiple waves can often be identified, e.g., in the COVID-19 pandemic so far five waves in many countries. Considering data from different regions and from different epidemics could lead to a more generally applicable definition of an EW.

Then, the results, i.e., the extracted waves and their attributes, are visually inspected. By interactively constraining the attributes, one can then identify the parameter ranges that lead to results that match the mental image as closely as possible.

The mathematical description of the extraction procedure together with the selected parameter ranges then represent the formal definition of the EW concept.

3.3. Prototypical interactive exploration tool

The prototypical tool, implemented in the visualization system Amira [SWH05b], consists of two components: a procedure for the extraction of parameterized structures, and an interactive tool for the visual exploration of the parameter and attribute spaces.

The extraction component uses computational methods available in the system and is applied in a precalculation step. The interactive exploration component is based on parallel coordinates (PC). We chose PCs because (1) they are easy for the user to interpret (each extracted structure corresponds to exactly one identifiable curve), (2) scalability with respect to dimensions is good, and (3) subintervals for each variable can be easily defined. Should the selection of more complex regions in the parameter-attribute space be required, one must choose other, maybe less scalable techniques, such as scatter-plot matrices, where more complex subselections are possible for every two dimensions. However, these would also lead to more complex definitions, which we are not seeking in this first work on the subject. Most closely related to our prototypical tool is probably the work by Beham et al. [BHGK14] for exploring the parameter space of a geometry generator.

The PC view is shown in Fig. 2. A single line in the PC view corresponds to one structure sampled from the parameter space of one time step, accompanied with a set of computed attributes. Hence, in our application, the high-dimensional space that is visualized using parallel coordinates consists of the time dimension, the dimensions of the parameter space, and the dimensions of the attribute space. The view's first axis represents time (Fig. 2 (1)). Brushing on the time axis could be done to restrict the considered time range to a subrange. However, since in our applications we are interested in parameters that are suitable for the whole time

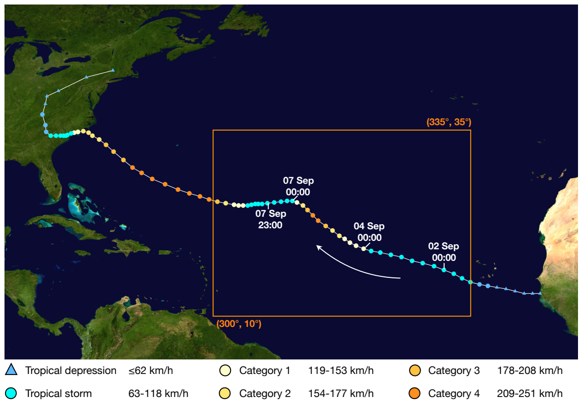

Figure 3: *Track map of Hurricane Florence 2018 (basemap and track taken from [Wik18]). The points show the location of the storm at 6-hour intervals. The color represents the storm's maximum sustained wind speeds according to the Saffir–Simpson scale (see legend). The orange box defines the area of interest.*

range, we do not use this functionality. The second axis represents the parameter space (Fig. 2 (2)). The fact that this space is only one-dimensional in our example applications is a special case; for higher-dimensional parameter spaces, additional axes would be added. The parameter axes display only the result of interactive brushing in the attribute space (Fig. 2 (3)). Brushing is done on each attribute axis separately. The selections from the individual brushing on all the attribute axes are concatenated to obtain one final selection. Note that for the concatenation of the selection, currently only the logical AND operation is supported. As result of the brushing, the lines corresponding to the selected attribute values are colored, by default in a light blue. Transparency and color of the lines can be adapted (Fig. 2 (4)). The selected lines are analyzed w.r.t. their parameter values and parameter intervals are identified that are valid over the whole selected time range. These ranges are shown on the right sidebar (Fig. 2 (5)) and the corresponding lines are highlighted in dark blue to distinguish them from the rest of the selected lines. Lines that are not selected are shown in gray. The PC view is connected with a direct visualization that allows simultaneous investigation of the selected structures (Fig. 2 (6)). Please see the accompanying video given in the supplementary material.

4. Example applications

In this section, we give a clear and concise example of our approach by demonstrating the procedure for specific structures, namely spatiotemporal geometric structures of tropical cyclones (TCs), i.e., their vortex cores and centerlines. The example is motivated by a new theory of TC intensification, according to which the inclination of the centerline, in interaction with atmospheric heat patterns that adopt some spatial orientation to the inclined centerlines, is a key parameter for TC intensification [PMOK12, DKGN20, DPK*21].

The analysis is made on the basis of ERA5 reanalysis data of the Hurricane Florence 2018. The data is given on a regular horizontal grid with resolution of $0.25° \times 0.25°$, i.e., about 30 km [HBB*18a, HBB*18b]. In all our visualizations, the data are vertically scaled with a factor of 100 to improve the perception of the structures. The time frame for the analysis is chosen to cover three different phases of the TC evolution: (1) stages of intensification; (2) full maturity; (3) weakening. The track of Hurricane Florence and the region of interest are shown in Fig. 3.

4.1. TC vortex core regions

What mental images related to TCs do meteorologists have in mind? In the meteorological literature, TCs are described as weather systems associated with strong winds and precipitation in the tropics [Mar03]. The flow- and thermodynamics-related structures are vaguely characterized as follows [Hou09]:

- The flow is essentially horizontal along closed streamlines (primary circulation).
- The horizontal (circular) flow structure exhibits a certain degree of symmetry.
- Surface-pressure possesses a local minimum at the center with surrounding closed isobars.
- Wind speeds are elevated (>33 m/s: hurricane strength).
- Convection is organized (cloudless eye, eyewall, outer rainbands).
- The core region is warmer than its surroundings.
- The storm's life-cycle takes place in the tropics.

Extraction and tracking algorithms based on such descriptions are characterized by a large number of parameters [PSUS05], many of which are not traceable to physical principles. Especially the mental image of a TC in terms of its outer bound is hard to grasp since the transition to the larger-scale atmosphere is smooth. For the core region, characterized by the strongest winds, however, there is the mental image of the flow being most affected by vortical motions. For extraction of the core region, focusing on the major physical phenomenon, namely the *meso-scale vortex*, seems to be a better strategy. If necessary, thermodynamical attributes can be considered in a second step.

Atmospheric vortices can be characterized by high local vorticity, more rigid-body rotation than stretching or shearing, low local pressure, closed or spiraling streamlines or pathlines, and coherent motion of neighboring fluid particles [Epp17]. Although still "no single definition of a vortex is currently universally accepted, despite the fact that fluid dynamicists continue to think in terms of vortices" [CBA05], there is general agreement that vortices are best characterized by dimensionless indicator quantities that signal their presence and local strength and are formed from tensor invariants of the velocity gradient ∇u (see, e.g., [SNU16]).

4.1.1. Indicator quantities for vortex core extraction

There is an abundance of methods for the extraction of vortex cores. Even those based on decomposing the velocity gradient ∇u into strain-rate tensor $A = \frac{1}{2}(\nabla u + \nabla u^\mathsf{T})$ and vorticity tensor $B = \frac{1}{2}(\nabla u - \nabla u^\mathsf{T})$ may provide different results.

- The widely used Q-criterion [HWM88] identifies a vortex as a "connected fluid region with a positive second invariant of ∇u" [Kol07] , i.e.,

$$Q = \frac{1}{2}(\|B\|^2 - \|A\|^2) > 0. \tag{1}$$

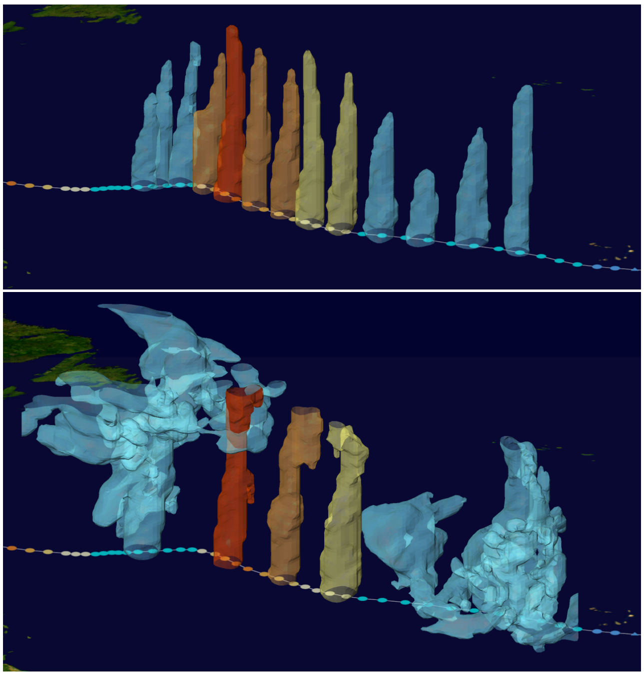

Figure 4: *Vortex regions extracted using* $Q > 0.02$ *(top) and* $\Omega > 0.52$ *(bottom). The top image shows extracted vortex regions for every 12 hours from the time frame of 02-07 September. For* Ω*, the following time steps were chosen: 02 September 00:00, 04 September 00:00, 05 September 00:00, 06 September 00:00, 07 September 23:00. Colors represent the intensity of TC according to the Saffir-Simpson scale (see legend in Fig. 3).*

- Ω was introduced to represent the ratio of vortical deformation over the whole deformation inside a vortex core [LWYD16], i.e.,

$$\Omega = \frac{\|B\|^2}{\|B\|^2 + \|A\|^2}.$$

To avoid division by 0, it was proposed [DWC*18] to add a small positive number $\varepsilon_\Omega = 0.001(\|B\|^2 - \|A\|^2)_{\max}$ to the denominator such that the criterion becomes

$$\Omega = \frac{\|B\|^2}{\|B\|^2 + \|A\|^2 + \varepsilon_\Omega} > 0.5. \quad (2)$$

The authors claim that this method "is pretty universal and does not need much adjustment in different cases and the iso-surfaces of $\Omega = 0.52$ can always capture the vortices properly" [LWYD16]. However, in order to capture the vortex core region, further adjustment of the threshold is needed. Ω is both dimensionless and normalized, resulting in values that are restricted to the interval $[0, 1]$.

- The kinematic vorticity number [Tru53] $W_k = \|B\|/\|A\|$ is another promising quantity for the determination of vortex properties [SNU16]. If the rotation rate prevails over the strain rate, W_k is larger than 1. Early circulations that have a potential to develop into stronger vortices could be captured with a lower threshold of W_k (e.g., W_k slightly smaller than 1). On the other hand, a higher threshold of W_k focuses on already developed strong vortices. To avoid non-physical noise, in our calculation we use an experimentally determined $\varepsilon_{W_k} = 0.03(\|B\| - \|A\|)$. The vortex core is then identified by

$$W_k = \frac{\|B\|}{\|A\| + \varepsilon_{W_k}} > 1. \quad (3)$$

Similar to Ω, W_k is dimensionless, but it is not normalized. Hence, values can potentially fall into the interval $[0, \infty]$.

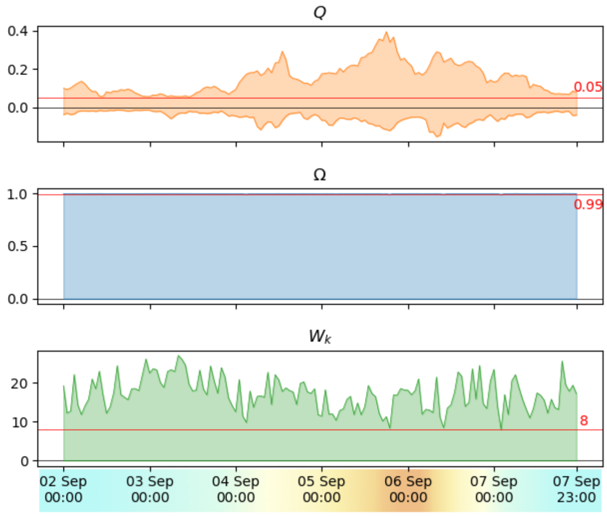

Figure 5: *Minimum and maximum values for three chosen indicator quantities in the time frame of 02-07 September 2018. Red line represents the upper limit of analysis intervals. Color gradient in the axis area represents the intensity of TC (see legend in Fig. 3).*

All three indicator quantities are Galilei-invariant. They characterize vortices, but their application may result in different vortex regions. The inequalities in Eq. 1-3 provide only theoretical thresholds. For real-case analyses, these values need to be adjusted to capture the appropriate structures. Fig. 4 shows vortex regions extracted using $Q > 0.02$ and $\Omega > 0.52$. Although at the highest hurricane intensity (yellow and orange structures), both quantities can capture the structures of the vortex core well, there are some issues during the intensification and weakening phases. Structures extracted with $Q > 0.02$ shrink up to 4.6 km while regions defined with $\Omega > 0.52$ do not represent TC cores correctly. In Fig. 5, the value intervals for three indicator quantities over the time frame of interest are shown (Fig. 3). Ω lies always in $[0, 1]$ while the extremum values for Q and W_k change over time. For the further analysis, parameter values that do not exist for the entire time window are excluded. The upper limits of the selected parameter intervals are depicted by a red line in Fig. 5.

4.1.2. Geometrical attributes of vortex regions

A vortex core region is associated with a dense cylinder-like structure. Based on this mental image and the examples shown above, the following attributes were chosen to describe the geometry of the extracted vortex core regions:

n_{cells}	number of voxels
h_{min}	minimum height
h_{max}	maximum height
d_{min}	minimum diameter
d_{max}	maximum diameter
C_d	compactness value [Bri08]

The *number of voxels* describes the overall size of the structure. *Minimum* and *maximum heights* give the lowest and the highest levels where the structure is defined. These attributes help to ensure that the extracted structure starts at a very low level and reaches a large height. The diameters of the structure were computed for each individual horizontal slice using the formula $d = \sqrt{dX_{max}^2 + dY_{max}^2}$, where dX_{max} and dY_{max} are the maximum distances the structure spreads in x and y directions, respectively. The derived value d approximates the diameter of a circle enclosing all selected voxels in the slice. It can be interpreted as the radius of the circumscribed circle. The *diameter* attributes are derived as minimum and maximum values of d over all horizontal slices. *Diameter* and *compactness value* control the width of the structure and its shape. The latter was introduced by Bribiesca [Bri08] and is defined as $C_d = A_c / A_{c_{max}}$, where A_c is the contact surface area and $A_{c_{max}}$ the maximum contact surface area. For a solid composed of n voxels, the approximation $A_{c_{max}} \approx 3(n - (n)^{\frac{2}{3}})$ can be used. The measure of discrete compactness is dimensionless, has values from 0 to 1 and is maximized by a cube.

4.1.3. Analysis and results

Three indicator quantities from Sect. 4.1.1 were computed based on the velocity field u for 13 time steps in time interval 02-07 September 2018. For each indicator quantity and each time step, sample structures were derived using 100 evenly distributed values from the appropriate parameter intervals, i.e.,

- $Q \in [0, 0.05]$
- $\Omega \in [0.5, 0.99]$
- $W_k \in [1, 8]$

Properties describing each structure were computed as discussed in Sect. 4.1.2. To inspect the 3D structures defined by properties selected via the PC view, isosurfaces of the smallest and biggest structures are visualized, i.e., structures corresponding to the upper and lower limit of the derived parameter interval, respectively.

PC views with the desired attributes for the three indicator quantities are shown in Fig. 6. The vortex core should start at the first presented level 0.2 km and have at least 8 km height. To limit the horizontal spread, structures with maximum diameter exceeding 600 km and with small compactness value, i.e., $C_d < 0.8$, were excluded. Lines representing structures meeting all the constraints are shown in the PC view by two shades of blue. Dark blue highlights the interval that would lead to the derivation of structures with the desired properties for every time step included in the analysis.

The procedure leads to the following structure definitions for TC cores based on the indicator quantities Q, Ω and W_k:

$$C_Q = \{x \in \mathbb{R}^3 \,|\, Q(x) \;\; > \tau_Q;\; \tau_Q \;\; \in [0.002, 0.0076]\},$$
$$C_\Omega = \{x \in \mathbb{R}^3 \,|\, \Omega(x) \;\; > \tau_\Omega;\; \tau_\Omega \;\; \in [0.7574, 0.8316]\},$$
$$C_{W_k} = \{x \in \mathbb{R}^3 \,|\, W_k(x) > \tau_{W_k}; \tau_{W_k} \in [1.4949, 1.9899]\}.$$

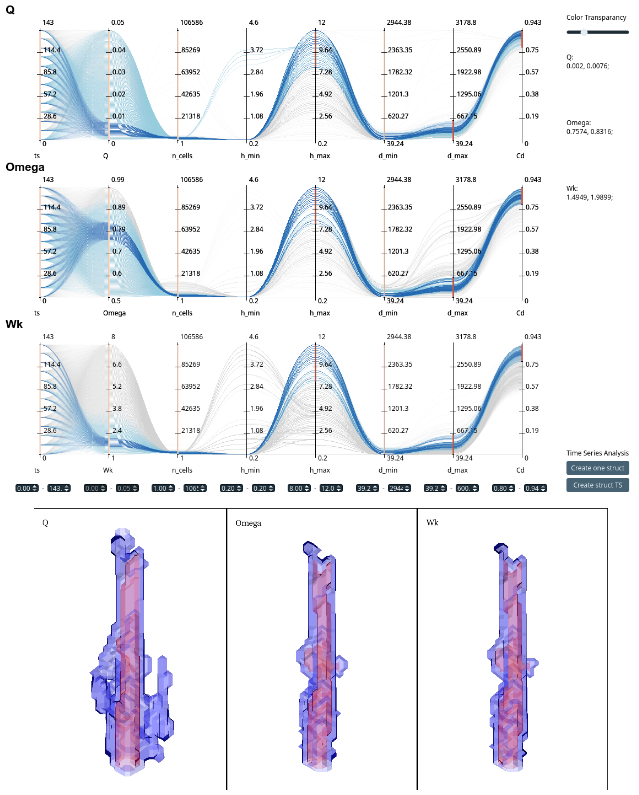

Figure 6: *Parallel coordinates plot window and structures extracted by selecting proper parameter values for three indicator quantities Q, Ω and W_k. Red and blue isosurfaces show the smallest and biggest structures corresponding to the largest and smallest parameter values in the identified parameter interval, respectively.*

The obtained definitions correspond to the ERA5 reanalysis data of Hurricane Florence 2018. Further analyses of, e.g., ERA5 reanalysis data of other hurricanes are necessary in order to get structure definitions that are more generally applicable.

Some resulting structures found for Q and Ω for the whole time window are shown in Fig. 7. Both indicator quantities provide comparable vortex core regions; those based on Q are somewhat smoother.

4.2. TC core line

Structure definitions often build upon each other. The second example demonstrates this situation. A TC core line is a structure that describes position and course of a TC. It is located in or close to the center of the TC vortex and represents another mental concept that is not explicitly described by data fields. Representation of the TC domain in cylindrical coordinates that are centered at the vortex core line is very useful for studying TC structure and dynamics (see [PMOK12, DPK*21]). Quantitative results, however, highly depend on a TC core line.

Information about core line computation, used geometrical attributes for parameter and attribute space exploration, as well as analysis and results are described in the Suppl. Mat., Sect. S1.1.

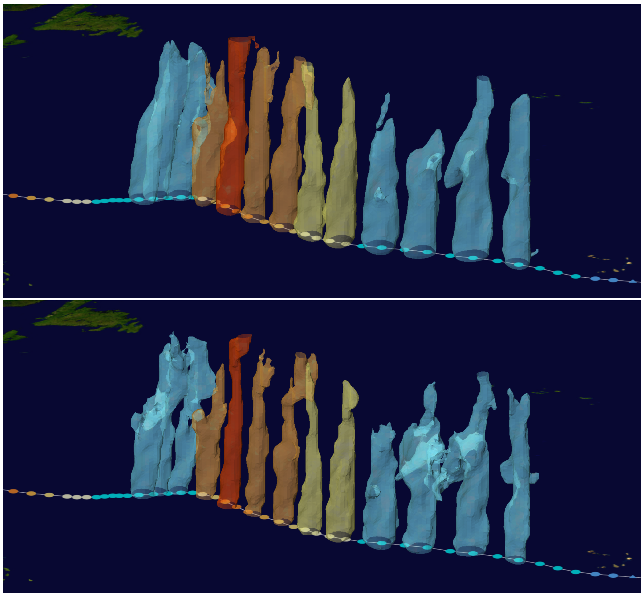

Figure 7: *Vortex regions showing the average structure from the derived intervals for Q (above) and Ω (below). Time steps at every 12 hours of the time frame of 02-07 September are shown. Colors represent the intensity of TC according to the Saffir-Simpson scale (see legend in Fig. 3).*

4.3. Performance and computational costs

The computationally most expensive part of the workflow is the preprocessing step. Extracting 100 structures sampled from the initial parameter interval took on average 4.38 sec for each time step. Memory costs are 1.5 MB per vortex region structure and only 17.3 KB per center line. In total, for each indicator quantity, $1.5MB * \#structures(=100) * \#timesteps(=13) \approx 2GB$ memory was used. Computation of attributes took around 0.6 sec per structure. Attributes were stored in CSV format, which costs less than 20 KB of memory per indicator quantity. The calculations were performed on a 8-core Intel i9-9900K machine with 3.6 GHz CPU and 16 GB RAM. Interaction with the PC view works in real time.

4.4. Assessment by domain scientists

During the development of the tool, we repeatedly discussed with researchers from meteorology and mathematical modeling. After presenting the workflow and implemented tool, they gave us the following feedback:

TC core regions: The extracted boundaries look reasonable, correspond to the experts' expectations, and are useful for further visual analysis. For better judgment and understanding of the structures, it would be helpful to additionally visualize physical context (e.g., the ground pressure or the cloud field).

TC core lines: The extracted lines look reasonable. According to the expert, it is of great value to get a handle of the variability under varying parameters. The major criterion for judging their validity is consistency with the mathematical model that describes the hurricane intensification [PMOK12]. This requirement is met. The helical structures discernible for some of the core lines raised great interest, as they could indicate interesting physical phenomena.

General feedback: The domain experts appreciated the elaboration of the methodological core of structure definition finding and its transformation into a practically executable workflow. The implemented semi-automated approach met with great approval, since (1) *all* possibilities are analyzed at the same time by systematic sampling the parameter space; (2) an *overview* for *all* resulting structures in the time+parameter+attribute space is provided by the PC view; (3) the set of reasonable structures can *iteratively be narrowed down* by constraining attribute values with 3D view control; and (4) understanding is greatly improved by the linked PC and 3D views. They furthermore appreciated that the structure extraction is based on well-defined, fundamental mathematical operations and that dimensionless indicator quantities have been used, enabling comparison of physically different situations.

They agreed that the next step should be to apply the method to other hurricanes—first to reanalysis data and then also to high-resolution simulation data to capture effects of smaller-scale processes as well. They pointed out that if the tool is used for very large data sets, one might end up with more structures than one can look at in 3D. They agreed that this is mitigated by first reducing the number of structures in the PC view before looking at them in 3D (still, more can be done here, see Sect. 5).

They made the conceptually helpful remark that difficulties in finding a single definition that works for all time steps and all hurricanes could give clues to *different physical processes* underlying the structures. Furthermore, they pointed out that the choice of attributes for the selection of structures must be well considered, because once the attribute ranges are restricted, the variability of the attributes can only be studied within these restricted ranges. The best choice of attributes could therefore also depend on the meteorological question.

They showed great interest in using the tool themselves, to bring in deeper physical knowledge and find structure definitions for other common meteorological structures; furthermore, to generate attribute statistics for many structures in order to derive, e.g., evolution equations. The possibility of using this tool and the visualization of other physical fields to examine the *relationship between structures and the underlying processes* met with particular interest. This could lead to more differentiated structure definitions.

5. Discussion and future work

In many scientific fields, important concepts are described purely verbally. From such descriptions, as well as from many individual examples, schematic drawings, photographs, and data visualizations, a mental image emerges that is shared by a specialist community. What is often missing is a precise formal definition; this becomes even more necessary the more we make data-based decisions and act data-driven.

In this paper, we describe a principled approach that makes it possible to move from mental images to precise formal definitions. It entails a general, abstract workflow and a matching interactive tool. We outline the approach using the concept of 'epidemic wave' as an example, and demonstrate it in detail using the example of two concepts from meteorology that build on each other, namely core regions and core lines of tropical cyclones.

The basic idea is to perform the following sequence of steps: (1) identify the mathematical type of the structure; (2) identify variables that indicate the presence or absence of the structure in question in the given data; (3) create a parameterized mathematical algorithm to extract the structure from the data; (4) attribute the extracted structures, with properties that play a role in the mental image; (5) restrict the attribute space under visual control so that the remaining structures match the mental image; (6) determine the parameter ranges leading to these matched structures; and (7) translate the algorithmic definition of the computational method and its parameter ranges back into a mathematical/formal definition.

In the associated prototypical interactive tool, we use parallel coordinates and the paradigm of linking & brushing on the attribute space together with a direct 3D visualization of the selected structures. This allows the user to verify whether the obtained structure definition indeed matches the mental image. Furthermore, it enables the user to narrow the structure definition further and further. In order to identify structure definitions that are applicable to all time steps of time-dependent data (a common scenario in the analysis of meteorological phenomena), it is essential to display the structure attributes for all time steps simultaneously.

An inherent limitation of our approach is that a potentially large number of structures must be extracted for different parameters and time steps and all necessary attributes characterizing the structures must be computed. However, this can usually be done automatically in a preprocessing step. Once this is done, our approach is interactive.

In order to test the proposed approach, we used a prototypical tool to identify definitions for two structures from meteorology: vortex core regions and core lines of tropical cyclones. We do not claim to have found the 'best, universal' method for extraction of these specific structures. Rather, we demonstrate that it is possible to transform mental images of structures into precise mathematical definitions through interactive, visually supported exploration. The motivation on the meteorological side was that the vortex core line is a major ingredient of the asymptotic theory on strongly tilted TCs [PMOK12, DPK*21]. The TC core line determines the origin of a tilted cylindrical coordinate system that simplifies the assessment of both, symmetric and asymmetric structures. Further evaluations of this theory therefore require a robust definition of the TC core line. The present work provides such a definition and also allows to estimate statistically the uncertainty comprised in the ensemble. Along this line, we will apply our prototype tool to further TC data, e.g., Hurricanes Isabel (2003), Earl (2010), Eduoard (2014), and Fiona (2016). The question we will ask is whether we can find similar parameter ranges across different TC data including higher-resolution data created with the ICON simulation model [ZRRB15].

The tool was very well received by meteorologists and a number of further ideas emerged on how it could be improved and used practically in meteorological research. One idea is to improve the scalability towards very large datasets by performing a cluster analysis in the attribute space and then continuing at the level of cluster representatives. Other future work will include the application and extension of our tool to more precisely define further meteorological structures such as convective cells and cold fronts. This will show us limitations of our current prototype and allow us to generalize it further. On the meteorological side, it will enable us to analyze process-structure relationships, which could then even lead to more physically sound structure definitions.

In the examples shown, only a one-dimensional parameter space was needed. However, the approach and the presented tool also support multi-dimensional parameter spaces. For these the main problem is sampling the parameter space and dealing with the potentially very large number of extracted sample structures. Therefore, more elaborate sampling methods than the one used here may be needed.

Our prototype implementation implicitly concatenates the attribute selection using the logical AND operation, thus creating a single high-dimensional attribute interval. However, the support of more complicated logical expressions might be needed for other structures. For this, the development of an expression editor alongside the parallel coordinates view will be explored. Another important aspect that we would like to study in more detail is the sensitivity of the structure of interest with respect to parameter changes, in analogy to studies in iso-contour extraction [PH10].

With regards to structure definitions there are limits in principle. On the one hand, mental images are fuzzy and their verbal descriptions are often incomplete. On the other hand, there is no predefined unique path when translating mental images into formal definitions. In our framework, this manifests itself in three ways: (1) It is not clear a priori, which structural indicators are most appropriate; (2) It is not clear, which mathematical representation is most appropriate for a structure; (3) Parameters of the structure representation can often only be fixed with finite precision. The proposed framework with its flexibility and built-in refinement capabilities aims to mitigate these problems as well as by allowing for long-term competition for ever sharper definitions.

We view the proposed framework as an initial approach and hope that it will open the door to improved approaches that lead to the development of more precise definitions for an increasingly broad range of concepts that are relevant to data-driven analyses, decisions, and actions.

Acknowledgments

We would like to thank our cooperation partners from applied mathematics and meteorology Rupert Klein, Annette Müller, Peter Nevir, Stephan Pfahl, and Lisa Schielicke, for their exemplary interdisciplinary cooperation and their willingness to share their knowledge and ideas with us. Furthermore, we would like to thank Nataša Đurđevac Conrad and Tim Conrad for their description of the expert mental image of an epidemic wave, valuable hints to literature, and their interest in an epidemiologically and mathematically founded definition. Natalia Mikula and Tom Dörffel have been supported by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project Number 235221301, Project (C06) "Multi-scale structure of atmospheric vortices".

References

[BCCF19] Brauer F., Castillo-Chavez C., Feng Z.: *Mathematical models in epidemiology*, vol. 69 of *Texts in Applied Mathematics*. Springer, 2019. 5

[BHGK14] Beham M., Herzner W., Gröller E., Kehrer J.: Cupid: Cluster-based exploration of geometry generators with parallel coordinates and radial trees. *IEEE Trans. Vis. Comput. Graph. 20* (12 2014), 1693–1702. doi:10.1109/TVCG.2014.2346626. 3, 5

[Bri08] Bribiesca E.: An easy measure of compactness for 2D and 3D shapes. *Pattern Recognition 41*, 2 (2008), 543–554. doi:10.1016/j.patcog.2007.06.029. 8

[CBA05] Chakraborty P., Balachandar S., Adrian R. J.: On the relationships between local vortex identification schemes. *J. Fluid Mech. 535* (2005), 189–214. doi:10.1017/S0022112005004726. 6

[CBJ*14] Clark A. J., Bullock R. G., Jensen T. L., Xue M., Kong F.: Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Weather Forecast. 29*, 3 (2014), 517–542. doi:10.1175/WAF-D-13-00098.1. 2

[CDJ15] Cai H., Dumais Jr R. E.: Object-based evaluation of a numerical weather prediction model's performance through forecast storm characteristic analysis. *Weather Forecast. 30*, 6 (2015), 1451–1468. doi:10.1175/WAF-D-15-0008.1. 2

[DGH03] Doleisch H., Gasser M., Hauser H.: Interactive Feature Specification for Focus+Context Visualization of Complex Simulation Data. In *Eurographics / IEEE VGTC Symposium on Visualization* (2003), pp. 239–248. doi:10.2312/VisSym/VisSym03/239-248. 3

[DHGK06] Doleisch H., Hauser H., Gasser M., Kosara R.: Interactive focus+context analysis of large, time-dependent flow simulation data. *Simul. 82*, 12 (2006), 851–865. doi:10.1177/0037549707078278. 3

[DKGN20] Dörffel T., Klein R., Gopolakrishnan S. G., Nolan D. S.: Stabilization of tropical cyclones against vertical wind shear by asymmetric diabatic heat release. In *34th Conference on Hurricanes and Tropical Meteorology Virtual Meeting, Proceedings* (2020). URL: https://ams.confex.com/ams/34HURR/mediafile/Manuscript/Paper373672/ExtendedAbstract.pdf. 6

[DMG*04] Doleisch H., Mayer M., Gasser M., Wanker R., Hauser H.: Case study: Visual analysis of complex, time-dependent simulation results of a diesel exhaust system. In *6th Joint Eurographics - IEEE TCVG Symposium on Visualization, VisSym 2004, Konstanz, Germany, May 19-21, 2004* (2004), pp. 91–96, 343. doi:10.2312/VisSym/VisSym04/091-096. 3

[DMH04] Doleisch H., Muigg P., Hauser H.: Interactive visual analysis of hurricane Isabel with SimVis. *IEEE Visualization 2004 Contest Entry* (2004). URL: http://vis.computer.org/vis2004contest/vrvis/. 3

[DPK*21] Dörffel T., Papke A., Klein R., Ernst N., Smolarkiewicz P. K.: Dynamics of tilted atmospheric vortices under asymmetric diabatic heating. *Theor. Comput. Fluid Dyn.* (2021), 1–43. doi:10.1007/s00162-021-00591-x. 6, 8, 10

[DWC*18] Dong X.-R., Wang Y., Chen X.-P., Dong Y., Zhang Y., Liu C.: Determination of epsilon for Omega vortex identification method. *J. Hydrodyn. 30* (07 2018), 1–9. doi:10.1007/s42241-018-0066-x. 7

[EBJS17] El-Baz A., Jiang X., Suri J. S.: *Biomedical Image Segmentation*. CRC Press, 2017. 2

[EMB*21] Engelke W., Masood T. B., Beran J., Caballero R., Hotz I.: Topology-based feature design and tracking for multi-center cyclones. In *Topological Methods in Data Analysis and Visualization VI*. Springer, 2021, pp. 71–85. doi:10.1007/978-3-030-83500-2_5. 2

[Epp17] Epps B.: Review of vortex identification methods. In *55th AIAA Aerospace Sciences Meeting* (2017), p. 0989. doi:/10.2514/6.2017-0989. 6

[Gär04] Gärdenfors P.: *Conceptual spaces: The geometry of thought*. MIT press, 2004. 2

[HBB*18a] Hersbach H., Bell B., Berrisford P., Biavati G., Horányi A., Sabater M., J., Nicolas J., Peubey C., R. R., Rozum I., Schepers D., Simmons A., Soci C., Dee D., Thépaut J.-N.: ERA5 hourly data on pressure levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* (2018). doi:10.24381/cds.bd0915c6. 6

[HBB*18b] Hersbach H., Bell B., Berrisford P., Biavati G., Horányi A., Sabater M., J., Nicolas J., Peubey C., R. R., Rozum I., Schepers D., Simmons A., Soci C., Dee D., Thépaut J.-N.: ERA5 hourly data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* (2018). doi:10.24381/cds.adbb2d47. 6

[HFB*17] Hadjighasem A., Farazmand M., Blazevski D., Froyland G., Haller G.: A critical comparison of Lagrangian methods for coherent structure detection. *Chaos 27*, 5 (2017), 053104. doi:10.1063/1.4982720. 2

[HHK11] Hege H.-C., Hotz I., Kasten J.: Distillation and visualization of spatiotemporal structures in turbulent flow fields. *Journal of Physics: Conference Series 318*, 6 (dec 2011), 062009. doi:10.1088/1742-6596/318/6/062009. 2

[Hou09] Houze R. A.: Clouds in tropical cyclones. *Mon. Wea. Rev. 138* (2009), 293–344. doi:10.1175/2009MWR2989.1. 6

[HWM88] Hunt J., Wray A., Moin P.: Eddies, streams, and convergence zones in turbulent flows. In *Studying Turbulence Using Numerical Simulation Databases – II (Proceedings of the 1988 Summer Program)* (12 1988), Center for Turbulence Research, Stanford Univ., Calif., pp. 193–208. URL: https://ntrs.nasa.gov/citations/19890015184. 6

[JH18] Jankowai J., Hotz I.: Feature level-sets: Generalizing isosurfaces to multi-variate data. *IEEE Trans. Vis. Comput. Graph. 26*, 2 (2018), 1308–1319. doi:10.1109/TVCG.2018.2867488. 3

[KBVH17] Kolesár I., Bruckner S., Viola I., Hauser H.: A fractional cartesian composition model for semi-spatial comparative visualization design. *IEEE Trans. Vis. Comput. Graph. 23*, 1 (2017), 851–860. doi:10.1109/TVCG.2016.2598870. 3

[KHS*18] Kern M., Hewson T., Sadlo F., Westermann R., , Rautenhaus M.: Robust detection and visualization of jet-stream core lines in atmospheric flow. *IEEE Trans. Vis. Comput. Graph. 24*, 1 (2018), 893–902. doi:10.1109/TVCG.2017.2743989. 2

[KKH01] Kniss J., Kindlmann G., Hansen C.: Interactive volume rendering using multi-dimensional transfer functions and direct manipulation widgets. In *Proceedings of the Conference on Visualization '01* (2001), IEEE Computer Society, pp. 255–262. doi:10.1109/TVCG.2002.1021579. 3

[Kol07] Kolář V.: Vortex identification: New requirements and limitations. *Int. J. Heat Fluid Flow 28*, 4 (2007), 638–652. doi:10.1016/j.ijheatfluidflow.2007.03.004. 6

[LKG*16] Ljung P., Krüger J., Gröller E., Hadwiger M., Hansen C. D., Ynnerman A.: State of the art in transfer functions for direct volume rendering. *Comp. Graph. Forum 35*, 3 (2016), 669–691. doi:10.1111/cgf.12934. 3

[LWYD16] Liu C., Wang Y., Yang Y., Duan Z.: New Omega vortex identification method. *Science China Physics, Mechanics & Astronomy 59*, 8 (2016), 684711–. doi:https://doi.org/10.1007/s11433-016-0022-6. 7

[Mar03] Marks F. D.: Chapter 31: Hurricanes. In *Handbook of Weather, Climate, and Water: Dynamics, Climate, Physical Meteorology, Weather Systems, and Measurements*. John Wiley & Sons, 2003, pp. 641–676. 6

[ML21] MARGOLIS E., LAURENCE S.: Concepts. In *The Stanford Encyclopedia of Philosophy*, Zalta E. N., (Ed.), Spring 2021 ed. Metaphysics Research Lab, Stanford University, 2021. URL: https://plato.stanford.edu/archives/spr2021/entries/concepts/. 2

[MLP*10] MCLOUGHLIN T., LARAMEE R. S., PEIKERT R., POST F. H., CHEN M.: Over two decades of integration-based, geometric flow visualization. *Comp. Graph. Forum 29*, 6 (2010), 1807–1829. 2

[PH10] PÖTHKOW K., HEGE H.-C.: Positional uncertainty of isocontours: Condition analysis and probabilistic measures. *IEEE Trans. Vis. Comput. Graph. 17*, 10 (2010), 1393–1406. doi:10.1109/TVCG.2010.247. 10

[PMOK12] PÄSCHKE E., MARSCHALIK P., OWINOH A., KLEIN R.: Motion and structure of atmospheric mesoscale baroclinic vortices: dry air and weak environmental shear. *J. Fluid Mech. 701* (2012), 137–170. doi:10.1017/jfm.2012.144. 6, 8, 9, 10

[PPF*11] POBITZER A., PEIKERT R., FUCHS R., SCHINDLER B., KUHN A., THEISEL H., MATKOVIĆ K., HAUSER H.: The state of the art in topology-based visualization of unsteady flow. *Comp. Graph. Forum 30*, 6 (2011), 1789–1811. doi:10.1111/j.1467-8659.2011.01901.x. 2

[PSUS05] PINTO J. G., SPANGEHL T., ULBRICH U., SPETH P.: Sensitivities of a cyclone detection and tracking algorithm: individual tracks and climatology. *Meteorol. Zeitschrift 14*, 6 (2005), 823–838. doi:10.1127/0941-2948/2005/0068. 6

[SNU16] SCHIELICKE L., NÉVIR P., ULBRICH U.: Kinematic vorticity number–a tool for estimating vortex sizes and circulations. *Tellus A: Dyn. Meteorol. Oceanogr. 68*, 1 (2016), 29464. doi:10.3402/tellusa.v68.29464. 6, 7

[SWH05a] SAHNER J., WEINKAUF T., HEGE H.-C.: Galilean invariant extraction and iconic representation of vortex core lines. In *EuroVis 2005: Eurographics / IEEE VGTC Symposium on Visualization* (2005), The Eurographics Association, pp. 151–160. doi:10.2312/VisSym/EuroVis05/151-160. 2

[SWH05b] STALLING D., WESTERHOFF M., HEGE H.-C.: Amira: A highly interactive system for visual data analysis. In *The Visualization Handbook*, Hansen C. D., Johnson C. R., (Eds.). Academic Press, 2005, ch. 38, pp. 749–767. doi:10.1016/B978-012387582-2/50040-X. 5

[THR*21] THEISEL H., HADWIGER M., RAUTEK P., THEUSSL T., GÜNTHER T.: Vortex criteria can be objectivized by unsteadiness minimization. *Physics of Fluids 33*, 10 (2021), 107115. doi:10.1063/5.0063817. 3

[Tru53] TRUESDELL C.: Two measures of vorticity. *Arch. Ration. Mech. Anal. 2* (1953), 173–217. URL: https://www.jstor.org/stable/24900328. 7

[vLHD*21] VON LINDHEIM J., HARIKRISHNAN A., DÖRFFEL T., KLEIN R., KOLTAI P., MIKULA N., MÜLLER A., NÉVIR P., PACEY G., POLZIN R., VERCAUTEREN N.: Definition, detection, and tracking of persistent structures in atmospheric flows, 2021. arXiv:2111.13645. 3, 4

[VMN*18] VALSANGKAR A. A., MONTEIRO J. M., NARAYANAN V., HOTZ I., NATARAJAN V.: An exploratory framework for cyclone identification and tracking. *IEEE Trans. Vis. Comput. Graph. 25*, 3 (2018), 1460–1473. doi:10.1109/TVCG.2018.2810068. 2

[Wik18] WIKIPROJECT TROPICAL CYCLONES/TRACKS: Track map of hurricane florence of the 2018 atlantic hurricane season, 2018. [Online; accessed November 10, 2021]. URL: https://commons.wikimedia.org/wiki/File:Florence_2018_track.png. 6

[WMO] WMO: Cloud classification aids CL, CM and CH. https://cloudatlas.wmo.int/cloud-classification-aids-cl-cm-ch.html. [Online; accessed Dec 1, 2021]. 1

[WST*07] WEINKAUF T., SAHNER J., THEISEL H., HEGE H.-C., SEIDEL H.-P.: A unified feature extraction architecture. In *Active Flow Control*. Springer, 2007, pp. 119–133. doi:10.1007/978-3-540-71439-2_8. 4

[WZSC21] WINKELMANN S., ZONKER J., SCHÜTTE C., CONRAD N. D.: Mathematical modeling of spatio-temporal population dynamics and application to epidemic spreading. *Math. Biosci. 336* (2021), 108619. doi:10.1016/j.mbs.2021.108619. 5

[ZG15] ZENKER F., GÄRDENFORS P.: *Applications of Conceptual Spaces: The Case for Geometric Knowledge Representation*, vol. 359 of *Synthese Library: Studies in Epistemology, Logic, Methodology, and Philosophy of Science*. Springer, 2015. 2

[ZMGW21] ZHANG S. X., MARIOLI F. A., GAO R., WANG S.: A second wave? what do people mean by COVID waves? – A working definition of epidemic waves. *Risk Manag. Healthc. Policy 14* (2021), 3775–3782. doi:10.2147/RMHP.S326051. 5

[ZRRB15] ZÄNGL G., REINERT D., RÍPODAS P., BALDAUF M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Q. J. R. Meteorol. Soc. 141*, 687 (2015), 563–579. doi:10.1002/qj.2378. 10