

# A Gradient-Based Comparison Measure for Visual analysis of Multifield Data

Suthambhara Nagaraj<sup>1</sup>, Vijay Natarajan<sup>1,2</sup>, and Ravi S. Nanjundiah<sup>3</sup>

<sup>1</sup>Department of Computer Science and Automation, Indian Institute of Science, Bangalore

email: {suthambhara,vijayn}@csa.iisc.ernet.in

<sup>2</sup>Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore

<sup>3</sup>Centre for Atmospheric and Oceanic Sciences, Indian Institute of Science, Bangalore

email: ravi@caos.iisc.ernet.in

## Abstract

*We introduce a multifield comparison measure for scalar fields that helps in studying relations between them. The comparison measure is insensitive to noise in the scalar fields and to noise in their gradients. Further, it can be computed robustly and efficiently. Results from the visual analysis of various data sets from climate science and combustion applications demonstrate the effective use of the measure.*

Categories and Subject Descriptors (according to ACM CCS): I.3.6 [Computer Graphics]: Computer Graphics—Methodology and Techniques

## 1. Introduction

Data from present day simulations and observations of physical processes often consists of multiple scalar and vector fields. Studying the interactions between the fields is pivotal to understanding the underlying phenomenon.

Single scalar fields are typically studied using techniques like isosurfacing, direct volume rendering and contour trees [HJ04, SML06, PWH01, CSvdP04, CSA00]. When visualizing multiple scalar fields, the above methods can be used separately on each field and visualized side by side or as overlays. The relationships and interactions that exist between the fields are often not captured by such methods. Simultaneous visualization of all the fields facilitates the understanding of interactions and relationships between them. This can be accomplished by employing a comparative approach to capture the relationships between variables.

We present a new gradient-based comparison measure for scalar fields that is applicable on an arbitrary number of scalar fields defined on a manifold. The measure captures the extent of alignment of the gradient vectors at a point. The distribution of the measure over the domain provides key insights into the interaction between input fields. The measure satisfies various desirable mathematical properties, can be computed efficiently, and is practically useful for studying relationships between multiple scalar fields. We are a team of

visualization researchers and a climate scientist who worked together to apply this measure for analyzing a hurricane simulation data set and a global climate simulation data set. The analysis helps explain various known meteorological and climatic phenomena. We also demonstrate the effective use of an aggregated version of the measure to the study of a combustion simulation data set.

The main contributions of this paper are :

- A new multifield comparison measure to capture interactions between multiple scalar fields defined on an  $n$ -dimensional domain,
- Theoretical results that establish the robustness of the measure by showing its insensitivity to noise in the scalar fields,
- An algorithm to compute the measure efficiently, and
- Real world applications to demonstrate the effectiveness of the measure in studying interactions between scalar fields in physical phenomena and an extension to vector fields.

The rest of the paper is organized as follows. We describe previous work in Section 2. In Section 3, we define the multifield comparison measure and prove its robustness and other properties. We motivate the use of the measure and explain its working in Section 4. Computation of the measure is described in Section 5. We describe several applications of the

measure in Section 6. In Section 7, we discuss the limitations of the multifield comparison measure and its insensitivity to noise in a real world data. We conclude the paper in Section 8.

## 2. Related Work

A popular approach to visualizing multiple fields is to combine them into a single value and then render the combined volume [CS99, BPRS98]. Woodring et al. [WS06] propose that the data fields should be rendered together within the same space for user comparison. They use set operators to combine the different fields into a single field that extracts the interesting portions of the data. These set operators can either combine the color values of the input fields or directly apply the operation in data space. Though combining volumes shows important parts of the data, the interactions between the different variables that are of importance to the domain scientists are not captured. For multifield time varying data, Lee et al. [LS09] propose a linear time algorithm to extract trend relationships among variables based on studying the change of variables over time and how these changes are related among different variables. Features in multifield data have been extracted using techniques like scatter plots [BW08] and variation density plots [NN10].

Multifield data have also been studied using statistical methods. One important work in this area uses the local statistical complexity [JWSK07] to identify features which may exhibit the same behavior in the future. Features are identified as complex if the probability that they occur again is low. In a later work, Jänicke et al. [JBTS08] improve the accuracy and efficiency of computing the local statistical complexity.

The relationship between the different scalar fields is particularly captured with the help of correlation measures. Sauber et al. [STS06] use two different techniques to compare different scalar fields at a point. One of them uses the alignment of gradients of the fields and also their magnitudes as a criterion to measure similarity. When the number of fields exceed two, pairwise similarity is computed and the least value is considered. This would detect regions where two of the fields are highly correlated. An obvious limitation of this approach is that two fields with low correlation would result in the other fields of the data to be ignored. In the same paper, the authors also describe a local correlation coefficient to detect linear dependencies between the scalar fields. The advantage of this method is its insensitivity to scaling of the data fields. It also has the same limitation as the first approach. Gosink et al. [GAJ07] also use correlation fields to study the interactions between the different variables in multi-field data. The inner product of the gradients of two fields of interest is computed over principle level sets of a third field. They use this approach to study combustion in methane and hydrogen. A limitation with using the inner product of the gradients is that only two fields can be compared.

Edelsbrunner et al. [EHN04] also employ a gradient-based approach to measure relationships between scalar fields. In their work, they introduce a measure to compare multiple scalar fields both locally at a point as well as over a region of the domain. In the case of three dimensional Euclidean space and two fields, they show that the measure at a point reduces to the length of cross product of the gradients of the fields. This measure, though useful, has a limitation that the number of scalar fields that can be compared cannot exceed the dimension of the domain.

In this paper, we also explore a gradient-based approach to compare scalar fields locally at a point. However, our method is not limited by the number of fields that can be compared unlike previous approaches. Our method also extends to time-varying scalar fields and to vector fields. Further, the measure is provably robust to noise in the input fields.

## 3. Multifield Comparison Measure

In this section, we introduce a gradient-based comparison measure for multiple scalar functions. The measure is defined as the norm of a matrix comprising the gradient vectors of the different functions. We first define the matrix norm before defining the measure and listing and proving its properties.

### 3.1. Matrix Norm

Let  $A$  be a  $m \times n$  matrix of real numbers. The *norm* of the matrix  $A$ , denoted as  $\|A\|$ , is defined as

$$\|A\| = \max_{\|x\|=1, x \in \mathbb{R}^n} \|Ax\|,$$

where  $\|x\|$  represents the Euclidean norm of vector  $x$  [HJ85]. We list four properties of the matrix norm that we will use later to prove key properties of the comparison measure. In particular, if  $A$  and  $B$  are matrices of real numbers, then

1.  $\|A\| > 0$  if  $A \neq 0$  and  $\|A\| = 0$  iff  $A = 0$ .
2. For  $\alpha \in \mathbb{R}$ ,  $\|\alpha A\| = |\alpha| \|A\|$ .
3.  $\|A + B\| \leq \|A\| + \|B\|$  and  $\|A - B\| \geq \left| \|A\| - \|B\| \right|$
4.  $\|AB\| \leq \|A\| \|B\|$ .

### 3.2. Comparison Measure

Let  $\mathbb{M}$  be a compact Riemannian manifold of dimension  $n$ . Let  $(x_1, x_2, \dots, x_n)$  be a local coordinate system such that the unit tangent vectors form an orthonormal basis with respect to the Riemannian metric. Let  $F = \{f_1, f_2, f_3, \dots, f_m\}$  be a set of smooth functions defined on the manifold. The derivative at a point  $p \in \mathbb{M}$  is written as a matrix of partial derivatives,

$$dF(p) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{bmatrix}$$

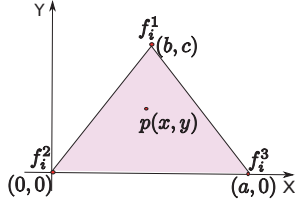


Figure 1: Piecewise-linear function defined on a triangle in a 2D mesh.

We define the *multifield comparison measure*  $\eta_p^F$  at point  $p$  as the norm of the matrix  $dF(p)$ ,  $\eta_p^F = \|dF(p)\|$ . The measure  $\eta_p^F$  satisfies three important properties: symmetry, coordinate system independence and stability. We now state and prove these properties.

**Symmetry.** The measure is independent of the permutation of the functions in  $F$ . The proof follows directly from the definition.

**Coordinate system independence.** Let  $(x'_1, x'_2, \dots, x'_n)$  denote a second orthonormal local coordinate system. Let  $J$  denote the Jacobian matrix that transforms the second coordinate system to the first. Since  $J$  represents a transformation between orthonormal coordinate systems,  $|Jv| = |v|$  for  $v \in \mathbb{R}^n$ . This implies that  $\|J\| = 1$ . We first observe that if  $dF'(p)$  is the derivative at  $p$  for the new coordinate system, then,

$$dF'(p) = dF(p)J.$$

This implies that  $\|dF'(p)\| = \|dF(p)J\| \leq \|dF(p)\| \|J\|$  by applying Property 4 of the matrix norm. Since  $\|J\| = 1$ , we have  $\|dF'(p)\| \leq \|dF(p)\|$ . Similarly, we can prove that  $\|dF(p)\| \leq \|dF'(p)\|$  by considering the Jacobian that transforms the first coordinate system to the second. This implies that the matrix norms are equal independent of the coordinate system.

**Stability.** We prove robustness of the measure when the scalar functions in the set  $F$  are piecewise-linear functions defined on a triangle mesh. We first observe that the derivative in the interior of a triangle is well defined and is a constant. To simplify the description, we prove stability for a 2D triangle mesh representing a surface. Extension to higher dimensions is straightforward.

Consider a triangle in the mesh with coordinates as shown in Figure 1. The function value at any interior point  $p(x, y)$  is given by

$$f_i^p = (A_1 f_i^1 + A_2 f_i^2 + A_3 f_i^3)x + (B_1 f_i^1 + B_2 f_i^2 + B_3 f_i^3)y,$$

where the constants  $(A_1, A_2, A_3, B_1, B_2, B_3)$  depend only on  $a, b$ , and  $c$ . Consider a perturbation  $\bar{F} = \{\bar{f}_1, \dots, \bar{f}_m\}$  of the functions in the set  $F$ , where  $\bar{f}_i = f_i + \varepsilon_i$  and  $\varepsilon_i$  assumes small values. The partial derivatives satisfy the following re-

lationship:

$$\frac{\partial \bar{f}_i}{\partial x} - \frac{\partial f_i}{\partial x} = (A_1 \varepsilon_i^1 + A_2 \varepsilon_i^2 + A_3 \varepsilon_i^3)$$

and

$$\frac{\partial \bar{f}_i}{\partial y} - \frac{\partial f_i}{\partial y} = (B_1 \varepsilon_i^1 + B_2 \varepsilon_i^2 + B_3 \varepsilon_i^3).$$

Therefore, the difference between the derivative matrices is

$$\begin{aligned} d\bar{F}(p) - dF(p) &= \begin{bmatrix} A_1 \varepsilon_1^1 + A_2 \varepsilon_1^2 + A_3 \varepsilon_1^3 & B_1 \varepsilon_1^1 + B_2 \varepsilon_1^2 + B_3 \varepsilon_1^3 \\ \vdots & \vdots \\ A_1 \varepsilon_m^1 + A_2 \varepsilon_m^2 + A_3 \varepsilon_m^3 & B_1 \varepsilon_m^1 + B_2 \varepsilon_m^2 + B_3 \varepsilon_m^3 \end{bmatrix} \\ &= \begin{bmatrix} \varepsilon_1^1 & \varepsilon_1^2 & \varepsilon_1^3 \\ \vdots & \vdots & \vdots \\ \varepsilon_m^1 & \varepsilon_m^2 & \varepsilon_m^3 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{bmatrix}. \end{aligned}$$

Using Property 3 of the matrix norm,

$$|\eta_p^{\bar{F}} - \eta_p^F| \leq \|d\bar{F}(p) - dF(p)\|.$$

Let  $\varepsilon = \sup_{1 \leq i \leq m, p \in \mathbb{M}} |\varepsilon_i(p)|$ . Using Property 4 of the matrix norm to rewrite  $\|d\bar{F}(p) - dF(p)\|$ , we get

$$\begin{aligned} |\eta_p^{\bar{F}} - \eta_p^F| &\leq \|d\bar{F}(p) - dF(p)\| \\ &\leq \left\| \begin{bmatrix} \varepsilon_1^1 & \varepsilon_1^2 & \varepsilon_1^3 \\ \vdots & \vdots & \vdots \\ \varepsilon_m^1 & \varepsilon_m^2 & \varepsilon_m^3 \end{bmatrix} \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{bmatrix} \right\| \\ &\leq \sqrt{\sum_{i=1}^m \sum_{j=1}^3 (\varepsilon_i^j)^2} \left\| \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{bmatrix} \right\| \\ &\leq (\sqrt{3m})\varepsilon \left\| \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \end{bmatrix} \right\|. \end{aligned}$$

The above equation indicates that a finite change in the functions results in a bounded change in the multifield comparison measure. The amount of change additionally depends on the size of the triangle.

The rows in the matrix  $dF(p)$  represent the gradient vectors of the function. It is therefore easy to see (using property 3 from section 3.1) that a bounded change in the gradients will also result in a bounded change in the comparison measure. The multifield comparison measure is therefore robust with respect to perturbations in the scalar functions as well as their gradients.

#### 4. Analyzing Synthetic Functions

We describe the motivation for the definition of the multifield comparison measure by considering the case of one and two analytic functions. Next, we show how the comparison

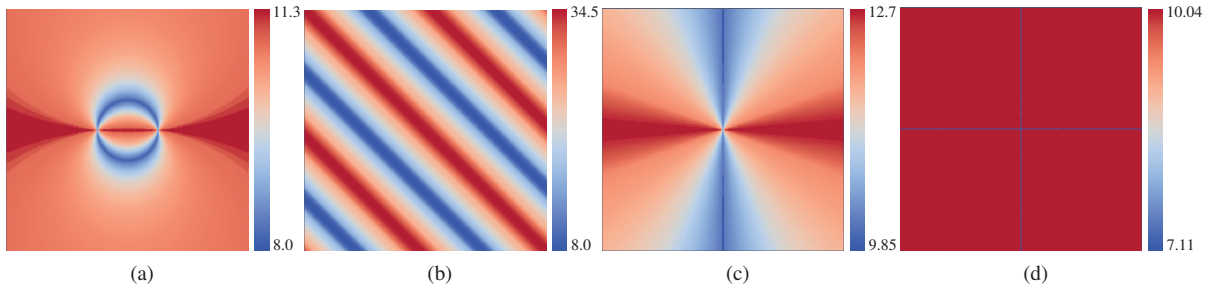


Figure 2: Multifield comparison measure  $\eta^F$  computed for synthetic functions defined on a 2D grid with the center as origin. (a) Two functions  $f_1(x, y) = \sqrt{(x-0.25)^2 + y^2}$  and  $f_2(x, y) = \sqrt{(x+0.25)^2 + y^2}$ . The measure  $\eta^F$  attains high values on the Jacobi set and low values where the gradients are orthogonal. (b) The sinusoidal function  $f_1(x, y) = \sin(3(x+y))$  and the linear function  $f_2(x, y) = y$ . (c) Three functions  $f_1(x, y) = \sqrt{x^2 + y^2}$ ,  $f_2(x, y) = \frac{1}{2}(\sqrt{3}x + y)$ , and  $f_3(x, y) = \frac{1}{2}(-\sqrt{3}x + y)$ . (d) One hundred different scalar functions, whose gradient vectors have unit magnitude and directions are chosen uniformly at random at points on the two axes and are chosen to be some constant at remaining points on the plane.

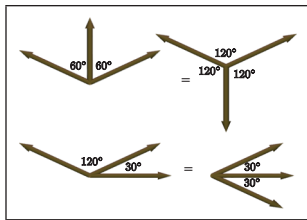


Figure 3: Two pairs of equivalent configurations of gradients of three functions described in Figure 2c. Gradient vectors subtend an angle of  $120^\circ$  at points along the  $Y$ -axis (top) and are more closely aligned with each other at points along the  $X$ -axis.

measure naturally extends to a larger number of scalar functions. We also show that the comparison measure can be used to capture the variation in gradient vectors over time for time varying scalar fields and to capture variation in time-varying vector fields.

**One / two scalar functions.** In the special case of a single scalar function, the comparison measure at a point  $p$  is equal to the maximum rate of change of the function, which is essentially the length of the gradient vector at  $p$ . The gradient vector and its length have been used to study the behavior of a function with respect to the domain [GNP\*06]. The proposed measure  $\eta^F$  is a natural extension of the notion of gradients and their relationship to multiple functions.

In the case of two smooth scalar functions, the gradients along with their mutual alignment is an indicator of the relationship between them [EHNPO4, GAJ07, STS06]. Given two non-zero gradients, the multifield comparison measure  $\eta^F$  assumes the highest value when the gradients are parallel or anti-parallel. This set of points where the gradients align is called the Jacobi set [EH02]. The Jacobi set has been previously used to study the relationship between scalar

functions [EHNPO4]. The comparison measure assumes a minimum when the gradients are orthogonal. Orthogonality of the gradients indicates mutual independence of the functions. Figure 2a shows the computed comparison measure for two paraboloids,  $f_1(x, y) = \sqrt{(x-0.25)^2 + y^2}$  and  $f_2(x, y) = \sqrt{(x+0.25)^2 + y^2}$  defined on a 2D grid. The Jacobi set is the line  $y = 0$  shown in dark red. The dark blue circle joining the centers of the paraboloids is the set of points where the gradients are orthogonal. Figure 2b depicts another example, a sinusoidal function  $f_1(x, y) = \sin(3(x+y))$  and a linear function  $f_2(x, y) = y$ . The comparison measure assumes high values at the Jacobi set (shown in bright red) and the set of points with orthogonal alignment of gradients has low values (shown in blue). We note that when the gradient of a function is replaced with its negative, the measure remains the same. Two different configurations of the gradients yield the same comparison measure if it is possible to make a transition from one to another by replacing gradients with their negatives.

**Multiple / time-varying scalar functions.** Consider three functions  $f_1(x, y) = \sqrt{x^2 + y^2}$ ,  $f_2(x, y) = \frac{1}{2}(\sqrt{3}x + y)$ , and  $f_3(x, y) = \frac{1}{2}(-\sqrt{3}x + y)$ . The multifield comparison measure (see Figure 2c) is minimum along the  $Y$ -axis. The gradient vectors at points on the  $Y$ -axis subtend an angle of  $120^\circ$  with each other, whereas they are more aligned at points on the  $X$ -axis, see Figure 3. Figure 2d shows the comparison measure computed for hundred different scalar functions defined on a grid. All gradients have unit magnitude and the direction of the gradients for points on the  $X$  and  $Y$  axis are chosen uniformly at random. The gradients of all scalar functions are chosen to be equal at other points on the plane. We observe that the values of the comparison measure on the two axes are low compared to the values elsewhere on the grid. This indicates that given a set of gradient vectors with fixed magnitudes, the measure takes high values where the directions are more “coherent”.

Given a single time varying scalar field, we construct the set  $F$  of multiple scalar functions with one function corresponding to each time step. The multifield comparison measure in this case measures the variation of the scalar function over time. We extend the measure to compare multiple vector fields or analyze the variation in time-varying vector fields by replacing each row in the derivative matrix  $dF(p)$  with the input vector at the point  $p$ .

## 5. Computation

Evaluating the multifield comparison measure at a point requires the solution to a maximization problem. In this section, we describe how this computation can be reduced to the faster evaluation of the maximum eigenvalue of a positive semi-definite matrix.

**Maximum eigenvalue computation.** From the definitions of the multifield comparison measure and the norm of a matrix, we have

$$\eta_p^F = \left( \max_{x \in \mathbb{R}^n, \|x\|=1} x^T (dF(p))^T (dF(p)) x \right)^{\frac{1}{2}}.$$

We rewrite the matrix product  $(dF(p))^T (dF(p))$  as  $U^T \Lambda U$ , where  $U$  is an orthogonal matrix and  $\Lambda$  is a diagonal matrix consisting of the eigenvalues of  $(dF(p))^T (dF(p))$  as entries in its diagonal. This follows from the spectral theorem from linear algebra [HK71]:

$$\eta_p^F = \left( \max_{x \in \mathbb{R}^n, \|x\|=1} x^T U^T \Lambda U x \right)^{\frac{1}{2}}.$$

Since the orthogonal matrix  $U$  represents a length preserving and invertible transformation, we can write the above expression as

$$\begin{aligned} \eta_p^F &= \left( \max_{x \in \mathbb{R}^n, \|x\|=1} x^T \Lambda x \right)^{\frac{1}{2}} \\ &= \max \{ \sqrt{\lambda} : \lambda \text{ is a diagonal element of } \Lambda \} \\ &= \max \{ \sqrt{\lambda} : \lambda \text{ is an eigenvalue of } (dF(p))^T (dF(p)) \}. \end{aligned}$$

For piecewise linear functions defined on a triangle mesh, the derivative matrix  $dF(p)$  is constant within a triangle and can be computed by choosing a local coordinate system. When the data is available over a structured grid and linearly interpolated along each coordinate axis, the difference between neighboring points in each axis direction can be used to approximate the partial derivatives at sample points and hence compute  $\eta^F$ . We note that following an approach similar to proving stability for piecewise linear functions, we can show that the comparison measure computed using this approximation for the gradient is also robust to noise in the input.

**Analysis.** The size of the  $n \times n$  matrix  $(dF(p))^T (dF(p))$  depends only on the dimension of the domain. Therefore, the time taken for computing the measure also depends only on

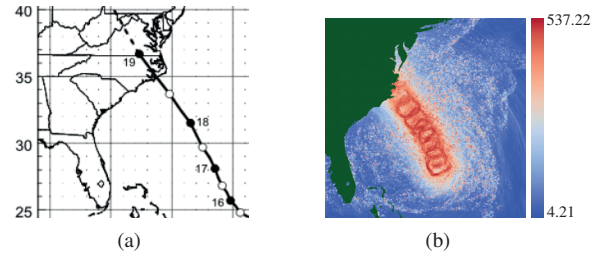


Figure 4: (a) The hurricane track released by the US National Hurricane Center (source: <http://www.nhc.noaa.gov>). The track relevant to the period of simulation is between point 17 and point 19 when the hurricane struck the coast. (b) Multifield comparison measure computed for nine pressure fields. The region in red with high values of the comparison measure corresponds to the trace of the eye of the hurricane. Land is shown in green.

the dimension of the domain and is, in particular, independent of the number of fields  $m$ .

## 6. Applications

We use the multifield comparison measure to study various real-world data from weather modeling, climate simulations, and combustion simulations. Observations on the combustion data were compared with prior work described in the literature. First, we study a simulation of the hurricane Isabel. Next, we apply our multifield comparison measure on a global wind pattern data set. Finally, we study a combustion simulation data set by aggregating the multifield comparison measure over the domain at each time step. The data in the following applications varies in the dimensionality of the domain and the number of fields – two scalar fields defined on a 3D domain, multiple time-varying scalar fields defined on 2D and 3D domains, and time-varying vector fields. In all cases the analysis is based on the proposed multifield comparison measure  $\eta^F$ . We perform experiments on data sets obtained using a variety of climate models. In Section 6.2, we describe results from a single climate model. The supplementary material includes figures from our experimental results on three additional models, which demonstrate the applicability of the proposed method.

### 6.1. Isabel Hurricane

Hurricane Isabel struck the west Atlantic region in September 2003. A simulation of the phenomenon was performed on a  $600 \times 600 \times 100$  grid corresponding to a physical volume of  $2139\text{km} \times 2004\text{km} \times 19.8\text{ km}$  over 48 time steps corresponding to 48 simulated hours [WBK04]. Several quantities were computed by the simulation. The scalar fields relevant to our experiment are pressure (Pf) and the horizontal wind velocity components (Uf and Vf).



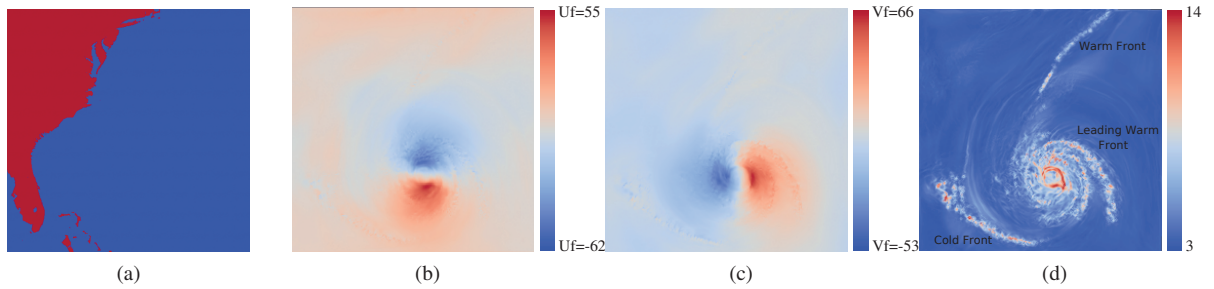


Figure 5: Fronts in Hurricane Isabel at hour 10. (a) Region of simulation. Land mass is shown in red. (b) Volume rendering (top view) of horizontal wind speed  $U_f$ . (c) Volume rendering (top view) of horizontal wind speed  $V_f$ . (d) Volume rendering (top view) of multifield comparison measure  $\eta^F$  computed for  $U_f$  and  $V_f$  showing the rainbands at different fronts. The location of the fronts is not available from the individual scalar fields  $U_f$  and  $V_f$ .

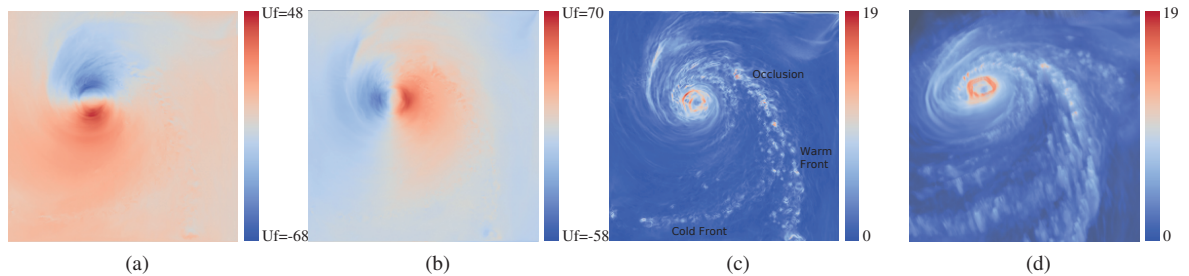


Figure 6: Fronts in Hurricane Isabel at hour 40. (a) Volume rendering (top view) of horizontal wind speed  $U_f$ . (b) Volume rendering (top view) of horizontal wind speed  $V_f$ . (c) Volume rendering (top view) of multifield comparison measure  $\eta^F$  computed for  $U_f$  and  $V_f$  showing the rainbands at different fronts. The cold front leads the warm front resulting in an occlusion. (d) Volume rendering from a different viewpoint.

**Hurricane track.** The horizontal pressure gradient can be observed to be high near the eye of the hurricane throughout the simulation. We study the pressure field at every fifth time step defined on a 2D slice corresponding to altitude 1500m. Figure 4b shows the multifield comparison measure computed for the nine pressure fields. The red circular regions that correspond to high values of the comparison measure correspond to the hurricane track. Figure 4a shows the track provided by the US National Hurricane Center for reference. The results are similar when we compute the multifield comparison measure for all 48 pressure fields. One of the horizontal pressure gradients assumes a high value at the eye compared to the rest and hence the comparison measure is higher compared to regions far from the path of the eye.

**Rainbands and fronts.** Cloud structures associated with an area of rainfall, called rainbands, occur mainly at boundaries separating two masses of air of different densities and temperatures, called fronts. The leading edge of the cooler mass of air is called the cold front and the leading edge of a warm air mass is called the warm front. The turbulence of the horizontal wind velocity is high near rain bands. We study the fronts by computing the multifield comparison measure for the pair of 3D scalar fields  $U_f$  and  $V_f$ , where the 3D do-

main corresponds to the volume in the altitude range 1500m-5800m.

First, we compute the multifield comparison measure for the fields  $U_f$  and  $V_f$  in the  $10^{th}$  time step. Figure 5 shows the result of our experiment as volume rendered images with the view point located above the volume. In particular, Figure 5d shows the location of two warm fronts and a cold front. This information about fronts cannot be extracted from the two functions individually (see Figures 5b and 5c). The comparison measure successfully captures the relationship between the fields at the fronts. The warm front leads the cold front in the  $10^{th}$  time step. The precipitation structures indicated by light red spots on the rain bands on the land mass were responsible for heavy rainfall in Florida. Next, we compute the multifield comparison measure for the fields  $U_f$  and  $V_f$  in the  $40^{th}$  hour of simulation. We observe that the warm front at the north has disappeared, see Figures 6c and 6d. The previously leading warm front is overtaken by the cold front resulting in an occlusion.

**Value of Study.** Both structures discussed above are key to a comprehensive study of the hurricane. The track of a hurricane or a cyclone generated from a forecast is helpful in predicting the areas susceptible to severe weather. Fronts often

give valuable information about severe weather to the forecaster. Rainbands at cold fronts are often strong in nature and can be responsible for heavy thunder storms. Typically, occlusion fronts are associated with thunder storms and their passage results in the reduction of humidity.

## 6.2. Global Wind Patterns

Prevailing winds are winds that blow in a dominant direction at a particular point. Movements in the Earth's atmosphere affect these winds. In regions of mid-latitudes, the winds blow from west to the east and are known as westerlies. The winds found in the tropics near the equator are easterlies or trade winds. Figure 7a shows the different prevailing winds on earth. We study wind patterns on earth using a climate simulation of 50 years between 1960 and 2009 [RBB\*03]. The data is available for 600 time steps corresponding to each month over the period of simulation. Each time step is a 3D grid with resolution corresponding to  $1^\circ \times 1^\circ \times 16\text{plev}$  (pressure elevations) on earth. Pressure elevations correspond to pressures varying from 1000 hPa on the surface to 30 hPa in the upper atmosphere.

**Westerlies and trade winds.** The wind velocity at a point on the grid is a vector quantity. We use the monthly wind velocities as rows in the derivative matrix and compute the matrix norm for 600 vector fields. The norm  $\eta^F$  measures the variation of the wind velocities over a time period of 50 years. Figure 7b shows the distribution of the computed comparison measure over a surface corresponding to pressure elevation 925 hPa. Comparing with wind patterns in Figure 7a, we see that the measure assumes high values in regions that lie in the path of prevailing winds, particularly the westerlies found in the regions surrounding Antarctica, the region of hurricanes in Atlantic, the cyclone prone region between Madagascar and Australia, and the trade winds across the Atlantic sea traveling towards the Caribbean sea. The distribution of the comparison measure over the isobar for pressure level 300 hPa, which corresponds to approximately 30000 feet above sea level, is shown in Figure 7c. The values of the comparison measure are higher compared to Figure 7b because friction and other effects can cause the wind flow at 925 hPa to be less steady than at higher levels such as 300 hPa. We note that the comparison measure assumes high values over the temperate regions corresponding to the westerly jet. This is a semi-permanent feature of the mid-latitudes. Many regions in the tropics undergo a seasonal reversal of wind (called the monsoons). Lower values of the comparison measure over the tropics indicates unsteadiness and corresponds to a seasonal reversal in wind pattern over this part of the world.

**Storm tracks.** The regions over the ocean with warm temperatures ( $> 27^\circ\text{C}$ ) are susceptible to storms. We filter out regions with lower temperatures and restrict our analysis to the months from June to November with the aim of locating storm tracks. Regions shown in blue in Figure 7e have been

filtered out. The red regions match closely with the storm tracks shown in Figure 7d. We notice that even though the west coast of South America has trade winds, storms are particularly absent due to lower temperatures. The storm prevalent regions in the Indian, Atlantic, and Pacific oceans have high values of the comparison measure.

**Value of Study.** Wind patterns give the details of the wind over a particular place. They cause various local and global phenomena and are widely studied by climate scientists. For example, the trade winds are responsible for tropical cyclones over oceans. Storm tracks generated using winds give us information on regions where storms are more probable. Computing storm tracks for long periods would require ascertaining tracks of every individual hurricane and cyclone and plotting them. Our approach simplifies this computation by considering all 600 time steps together to generate the distribution of the comparison measure.

## 6.3. Hydrogen Combustion

We study phases in the combustion of an inhomogeneous turbulent mixture of fuel and oxidizer. Hydrogen fuel at 300K is mixed with an oxidizer (21% oxygen) at 1200K. The influence of turbulence on the different phases of combustion is studied in the simulation. The compressed fuel ignites at multiple spots because of the inhomogeneity in the air-fuel ratio. Depending on the air-fuel ratio, the flame either propagates in an outward direction from the ignition spot or burns out. Further details of the computation can be found in the description of the simulation by Echehki and Chen [EC03] and in the description of a visual analysis of this data by Koegler [Koe01]. The domain of the simulation is a  $600 \times 600$  grid for 67 time steps. The species mass fractions of the fuel  $\text{H}_2$ , oxygen  $\text{O}_2$ , and intermediate  $\text{HO}_2$  are given at each grid point for all time steps.

**Phases of combustion.** We aggregate the multifield comparison measure for a time step  $t$  by computing its integral over the domain  $D$  of simulation:

$$\eta^{F,t} = \int_{x \in D} \eta_x^F dx.$$

In our study, we consider all three scalar fields for computing the comparison measure for a time step, i.e.  $F = \{\text{H}_2, \text{O}_2, \text{HO}_2\}$ . The plot of  $\eta^{F,t}$  over time is shown in Figure 8. We compare our results with a plot of the global comparison measure  $\kappa$ , introduced by Edelsbrunner et al. [EHN04], see Figure 8. The fields used for computing  $\kappa$  are  $\text{H}_2$  and  $\text{O}_2$ . The maximum number of fields that  $\kappa$  can compare over a 2D domain is two, whereas we are able to consider all three fields in our analysis.

Following the work of Koegler, the areas in the domain that eventually ignite can be considered as features. In the pre-ignition phase, the concentrations of the intermediate radicals build up in regions that have sufficient mass frac-

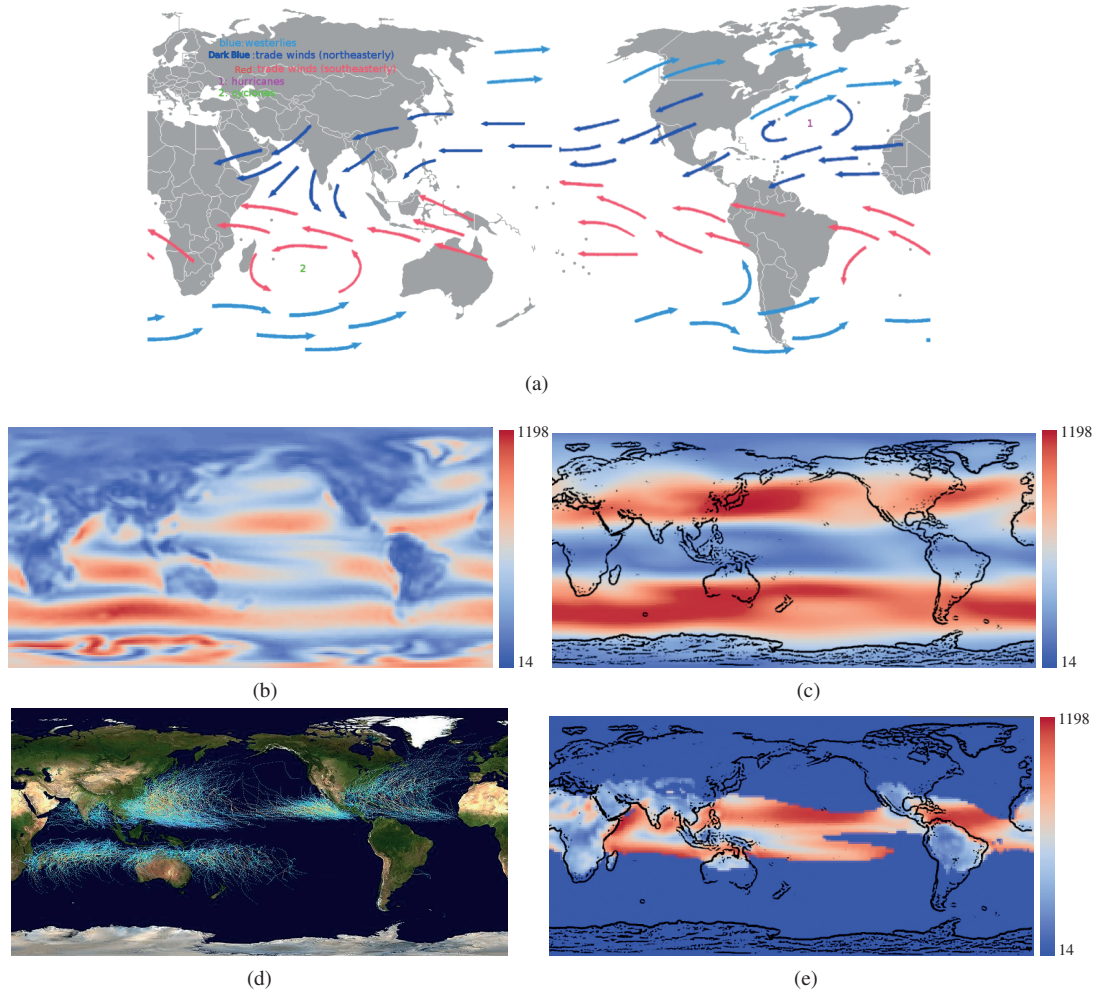


Figure 7: Multifield comparison measure  $\eta^F$  computed for wind velocities over the years 1960-2009, where the comparison is over a set of six hundred 3D vector fields. (a) Map of world showing wind patterns (source: Wikipedia) (b) Distribution of  $\eta^F$  over surface corresponding to pressure elevation 925 hPa. The dark red regions correspond to the wind patterns. (c) Distribution of  $\eta^F$  over surface corresponding to pressure elevation 300 hPa. The temperate regions exhibit higher values. (d) Storm track for the years 1985-2005 (source: Wikipedia) (e) Distribution of  $\eta^F$  after removing regions with low mean temperature ( $< 27^\circ\text{C}$ ). Red regions correspond to the storm tracks. The world map is overlaid for clarity.

tion of  $\text{H}_2$ . The number of features attains a maximum during time steps 7-14 [Koe01]. This is captured by a knee in the plot of  $\eta^{F,t}$ . The plot of  $\kappa$  does not indicate changes that happen in the pre-ignition build up of radicals because the intermediate  $\text{HO}_2$  is not considered in the computation. Ignition (time step 28) happens at areas with high radical concentrations during the ignition phase and the flame front spreads to hot enough areas with the right mix of fuel and oxidizer during the burning phase (time steps 50-55). Ignition and burning are captured by a minimum and a maximum respectively in the plot of  $\eta^{F,t}$ . These phases are also captured by  $\kappa$  because of the interaction between  $\text{H}_2$  and  $\text{O}_2$  during this time period. The beginning of the extinction

phase (approximately time step 60) where the flame begins to extinguish is also captured more clearly by a minimum in the plot of  $\eta^{F,t}$ . We believe that the reason  $\eta^{F,t}$  is able to capture more information compared to  $\kappa$  is because all three fields play a role in defining the phases.

## 7. Discussion

We now discuss some limitations of the multifield comparison measure and describe an experiment to study the sensitivity of the measure to noise in the input.

**Limitations.** The basic premise behind using the comparison measure to capture relationships in multifield real world



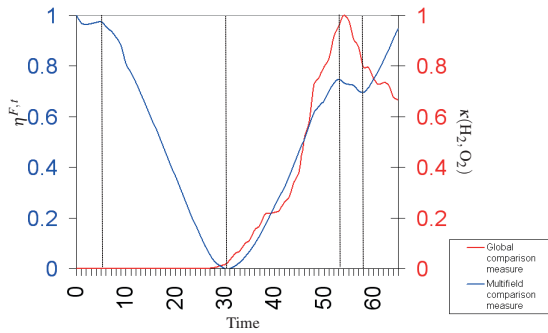


Figure 8: Analyzing phases of combustion using an aggregate  $\eta^{F,t}$  of the multifield comparison measure over the domain within a time step  $t$ . The set  $F = \{H_2, O_2, HO_2\}$ . The plot of  $\eta^{F,t}$  (blue) over time captures more phases of the combustion process compared to the plot of the global comparison measure  $\kappa(H_2, O_2)$  (red) proposed by Edelsbrunner et al. [EHNPO4]. The vertical dashed lines approximately correspond to the four phases of combustion: pre-ignition, ignition, burning, and extinction.

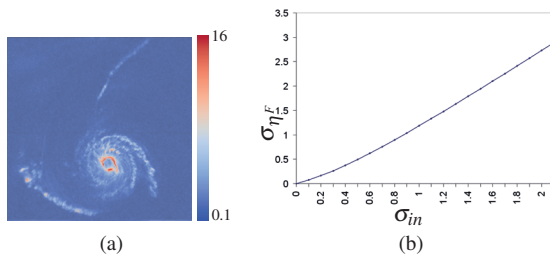


Figure 9: Stability in the comparison measure for Isabel. (a) Volume rendering (top view) of multifield comparison measure  $\eta^F$  computed after adding Gaussian noise (standard deviation = 1) to the fields  $U_f$  and  $V_f$ . The rain bands are still clearly visible. (b) Graph showing near linear relationship between the standard deviation of the noise in the input  $\sigma_{in}$ , and the mean deviation of the comparison measure  $\sigma_{\eta^F}$ .

data is the fact that the agreement among the different gradient fields can often capture interactions among fields. In many applications, where this does not hold good, using the comparison measure or other gradient based comparison measures would be less fruitful. We therefore believe that gradient based techniques such as ours complements other well known techniques like the local statistical complexity [JWSK07] and the Pearson correlation coefficient. Another drawback of the comparison measure is its sensitivity to scaling of individual fields.

**Sensitivity to noise.** We validate our claim that the comparison measure is insensitive to noise using the hurricane data described in Section 6.1. The different features present in the data like the weather fronts are still clearly visible in

the computed comparison measure field after adding a Gaussian noise (standard deviation = 1) to the input fields  $U_f$  and  $V_f$  of the 10<sup>th</sup> time step in the simulation, see Figure 9a. It should be noted that the noise we have added is large and quite unlikely to occur in real data sets. We also study the mean deviation in the comparison measure computed on input with Gaussian noise of different amplitudes. Figure 9b shows the linear relationships between the observed mean deviation of  $\eta^F$  and the noise, which indicates that the multifield comparison measure is not much sensitive to noise.

## 8. Conclusions

We have described a robust multifield comparison measure for scalar fields whose distribution over the domain plays an important role in the visual analysis of the input fields. The comparison measure is computed locally at all points of the domain as the maximum eigenvalue of a small sized matrix. We described applications of the comparison measure to study various simulation datasets from climate science and combustion studies where the data is represented using multiple 2D, 3D, or time-varying scalar fields. We used the comparison measure to study up to 600 scalar fields defined on the domain. We list three ideas for future work:

- The multifield comparison measure being sensitive to the scaling of individual fields may not be always desirable. One approach to address this issue is to scale the scalar fields or normalize their gradients. Though this method would work in some cases, it could bias the results by scaling up small insignificant gradients. A complete solution remains to be a challenge.
- Integrating the multifield comparison measure into the query-based visualization framework of Gosink et al. [GAJ07] will be an interesting task.
- Identification of important isovalues of a scalar function in multifield data is a challenging problem. A global version of the proposed comparison measure may help locate these isovalues.

## Acknowledgments

This work was supported by the Department of Science and Technology, India under grant SR/S3/EECE/048/2007. Hurricane Isabel data was produced by the Weather Research and Forecast (WRF) model, courtesy of NCAR and the U.S. National Science Foundation (NSF). The climate data was part of the WCRP CMIP3 Multi-Model data repository at <https://esgcat.llnl.gov:8443>. We acknowledge the modeling groups for making their model output available for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving this data, and the WCRP's Working Group on Coupled Modelling (WGCM) for organizing the model data analysis activity. The WCRP CMIP3 multi-model dataset is supported by the Office of Science, U.S. Department of Energy. We thank all reviewers for their valuable feedback, which helped improve the paper.

## References

- [BPRS98] BAJAJ C., PASCUCCI V., RABBIOLO G., SCHIKORE D.: Hypervolume visualization: a challenge in simplicity. In *Proc. IEEE Symp. Volume Visualization* (1998), pp. 95–102. 2
- [BW08] BACHTHALER S., WEISKOPF D.: Continuous scatterplots. *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (2008), 1428–1435. 2
- [CS99] CAI W., SAKAS G.: Data intermixing and multi-volume rendering. *Computer Graphics Forum* 18, 3 (1999), 359–368. 2
- [CSA00] CARR H., SNOEYINK J., AXEN U.: Computing contour trees in all dimensions. In *Proc. ACM-SIAM Symposium on Discrete algorithms* (2000), pp. 918–926. 1
- [CSvdP04] CARR H., SNOEYINK J., VAN DE PANNE M.: Simplifying flexible isosurfaces using local geometric measures. In *Proc. IEEE Conf. Visualization* (2004), pp. 497–504. 1
- [EC03] ECHEKKI T., CHEN J. H.: Direct numerical simulation of autoignition in nonhomogeneous hydrogen-air mixtures. *Combustion and Flame* 134, 3 (2003), 169–191. 7
- [EH02] EDELSBRUNNER H., HARER J.: Jacobi set of multiple Morse functions. In *Foundations of Computational Mathematics, Minneapolis* (2002), Cambridge Univ. Press, pp. 37–57. 4
- [EHN04] EDELSBRUNNER H., HARER J., NATARAJAN V., PASCUCCI V.: Local and global comparison of continuous functions. In *Proc. IEEE Conf. Visualization* (2004), pp. 275–280. 2, 4, 7, 9
- [GAJ07] GOSINK L., ANDERSON J. BETHEL W., JOY K.: Variable interactions in query-driven visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1400–1407. 2, 4, 9
- [GNP\*06] GYULASSY A., NATARAJAN V., PASCUCCI V., BREMER P.-T., HAMANN B.: A topological approach to simplification of three-dimensional scalar functions. *IEEE Transactions on Visualization and Computer Graphics* 12, 4 (2006), 474–484. 4
- [HJ85] HORN R., JOHNSON C.: *Matrix Analysis*. Cambridge University Press, 1985. 2
- [HJ04] HANSEN C., JOHNSON C.: *Visualization Handbook*. Academic Press, 2004. 1
- [HK71] HOFFMAN K., KUNZE R.: *Linear Algebra*. Prentice Hall, 1971. 5
- [JBTS08] JÄNICKE H., BOTTINGER M., TRICOCHÉ X., SCHEUERMANN G.: Automatic detection and visualization of distinctive structures in 3d unsteady multi-fields. *Computer Graphics Forum* 27, 3 (2008), 767–774. 2
- [JWSK07] JÄNICKE H., WIEBEL A., SCHEUERMANN G., KOLLMANN W.: Multifield visualization using local statistical complexity. *IEEE Transactions on Visualization and Computer Graphics* 13 (2007), 1384–1391. 2, 9
- [Koe01] KOEGLER W. S.: Case study: application of feature tracking to analysis of autoignition simulation data. In *Proc. IEEE Conf. Visualization* (2001), pp. 461–464. 7, 8
- [LS09] LEE T.-Y., SHEN H.-W.: Visualization and exploration of temporal trend relationships in multivariate time-varying data. *IEEE Transactions on Visualization and Computer Graphics* 15 (2009), 1359–1366. 2
- [NN10] NAGARAJ S., NATARAJAN V.: Relation-aware isosurface extraction in multi-field data. *IEEE Transactions on Visualization and Computer Graphics* (2010). <http://doi.ieeecomputersociety.org/10.1109/TVCG.2010.64>. 2
- [PWH01] PEKAR V., WIEMKER R., HEMPEL D.: Fast detection of meaningful isosurfaces for volume data visualization. In *Proc. IEEE Conf. Visualization* (2001), pp. 223–230. 1
- [RBB\*03] ROECKNER E., BÄUML G., BONAVENTURA L., BROKOPF R., ESCH M., GIORGETTA M., HAGEMANN S., KIRCHNER I., KORNBLUEH L., MANZINI E., RHODIN A., SCHLESE U., SCHULZWEIDA U., TOMPKINS A.: *The Atmospheric General Circulation Model ECHAM5. PART 1: Model Description*. Tech. Rep. 349, Max Planck Institute of Meteorology, 2003. 7
- [SML06] SCHROEDER W., MARTIN K., LORENSEN B.: *The Visualization Toolkit An Object-Oriented Approach To 3D Graphics, 4th Edition*. Kitware, Inc., 2006. 1
- [STS06] SAUBER N., THEISEL H., SEIDEL H.-P.: Multifield-graphs: An approach to visualizing correlations in multifield scalar data. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 917–924. 2, 4
- [WBK04] WANG W., BRUYERE C., KUO B.: Competition data set and description in 2004 IEEE Visualization design contest, 2004. <http://vis.computer.org/vis2004contest/data.html>. 5
- [WS06] WOODRING J., SHEN H.-W.: Multi-variate, time varying, and comparative visualization with contextual cues. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 909–916. 2