DOI: 10.1111/j.1467-8659.2011.01942.x
Eurographics / IEEE Symposium on Visualization 2011 (EuroVis 2011)
H. Hauser, H. Pfister, and J. J. van Wijk
(Guest Editors)

Volume 30 (2011), Number 3

Probabilistic Marching Cubes

Kai Pöthkow, Britta Weber and Hans-Christian Hege

Zuse Institute Berlin (ZIB), Berlin, Germany

Abstract

In this paper we revisit the computation and visualization of equivalents to isocontours in uncertain scalar fields. We model uncertainty by discrete random fields and, in contrast to previous methods, also take arbitrary spatial correlations into account. Starting with joint distributions of the random variables associated to the sample locations, we compute level crossing probabilities for cells of the sample grid. This corresponds to computing the probabilities that the well-known symmetry-reduced marching cubes cases occur in random field realizations. For Gaussian random fields, only marginal density functions that correspond to the vertices of the considered cell need to be integrated. We compute the integrals for each cell in the sample grid using a Monte Carlo method. The probabilistic ansatz does not suffer from degenerate cases that usually require case distinctions and solutions of ill-conditioned problems. Applications in 2D and 3D, both to synthetic and real data from ensemble simulations in climate research, illustrate the influence of spatial correlations on the spatial distribution of uncertain isocontours.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Viewing algorithms I.6.8 [Simulation and Modeling]: Types of Simulation—Monte Carlo, Visual G.3 [Mathematics of Computing]: Probability and Statistics—Probabilistic algorithms

1. Introduction

In many applications volume data that is affected by uncertainty, for example due to the nature of the measurement processes or uncertain simulation parameters, is acquired and processed. In this paper we will consider the visualization of scalar fields in a d-dimensional domain M that are discretely sampled and that are fraught with uncertainty. A standard technique for visualizing crisp scalar fields, i.e. fields that are not afflicted with uncertainty, is to depict level sets, which under certain regularity conditions are $d-1$–dimensional isocontours. For uncertain scalar fields several methods have been suggested to compute and visualize fuzzy equivalents to crisp isocontours. Recently a method to compute and visualize the positional uncertainty of isocontours in uncertain input data has been suggested by Pöthkow and Hege [PH10]. In this work the discretely sampled uncertain scalar field was modeled by a discrete random field, assuming that the random variables are spatially *un*correlated.

In this paper we use the same mathematical model, but consider the more general case where *arbitrary spatial correlations* are taken into account. The correlation structure is one of the *essential properties* of a random field and it has to be considered in order to compute *accurate* results. An important category of data representing uncertainty are ensemble data; they are commonly used, e.g., in climate research and weather forecasting.

We assume that the data have been sampled on nodes of some mesh. In order to reveal the spatial probability for the occurrence of an isocontour of a given isolevel at some spatial location, we compute probabilities that 1-, 2-, ..., $d-$dimensional mesh entities are crossed by an isocontour. The proposed procedure can be applied to any type of mesh, both structured and unstructured. For definiteness and for simplicity we exemplify the method using rectangular regular grids. In this case the types of integrals to be computed correspond to the well-known symmetry-reduced marching cubes cases.

2. Related Work

Publications on uncertainty visualization can be roughly classified into those considering uncertainty quantified by scalar values and those describing the uncertainty by ensemble data or parametric distributions. An introduction to

uncertainty visualization giving a detailed classification of uncertainty and describing visualization techniques was presented by Pang *et al.* [PWL97] and Griethe *et al.* [GS06]. MacEachren *et al.* [MRH*05] presented a review of several approaches for the representation of uncertainty of geospatial data and for improving decision making when dealing with uncertainty.

Volume rendering methods for visualization of uncertainty modeled by a scalar function have been proposed for a variety of applications. In medical applications, uncertainty can often be quantified by the class probability of a probabilistic segmentation. Kniss *et al.* [KUS*05] presented a volume rendering approach that allows the user to interactively explore the class probabilities of segmentations and uncertainty of surface boundaries by deferring the classification decision to the rendering stage. Uncertainty of tissue classification in medical volume data was also visualized by animation using fuzzy time-dependent transfer functions for direct volume rendering by Lundström *et al.* [LLPY07]. Saad *et al.* used shape and appearance knowledge to evaluate and visualize segmentation uncertainty of medical data sets [SHM10]. Praßni *et al.* [PRH10] used the uncertainty of probabilistic segmentation algorithms as a cue for the improvement of the segmentation results in a semi-automatic work flow.

The visualization of ensemble data was addressed by Potter *et al.* [PWB*09] and Sanyal *et al.* [SZD*10]. Both developed an interactive visualization tool for weather forecasts and climate simulations. Sanyal *et al.* also evaluated its efficiency in cooperation with domain experts. Luo *et al.* [LKP03] proposed methods to apply standard visualization techniques on ensemble data. They also described operators to manipulate the data and specific visualization methods. Djurcilov *et al.* [DKLP02] proposed to employ volume rendering with a specially designed transfer function, taking into account the variance of ensembles. For multivariate data where uncertainty is modeled by probability distributions, Feng *et al.* [FKLT10] proposed methods to display variants of scatter plots and parallel coordinate plots that take the probability density functions of the data into account. As an extension of vector field topology, Otto *et al.* [OGHT10] introduced spatial distributions of critical points and topological segmentations in an uncertain vector field and displayed them using height fields.

Several publications addressed the uncertainty of position, shape, and size of surfaces. We first consider papers where uncertainty modeled by a scalar value is visualized. Pang *et al.* [PWL97] created fat surfaces by displaying two surfaces that enclose the volume in which the true (but unknown) surface is located. Rhodes *et al.* [RLBS03] used color and texture mapping on isosurfaces to indicate areas of high data uncertainty but did not specify the mathematical model representing the uncertainty. Pöthkow and Hege [PH10] characterized the sensitivity of isocontours using the numerical condition and displayed it using colormapping. Grigoryan and Rheingans [GR04] used point primitives for rendering uncertain surfaces that are displaced along the normals of crisp isosurfaces in a distance proportional to the uncertainty, random numbers, and a user-defined scale factor.

Parametrized distributions describing the uncertainty of geospatial data were discussed by Zehner *et al.* [ZWK10]. They proposed to combine isosurfaces with additional geometry to indicate the positional uncertainty and show spatial confidence intervals. The distribution parameters were computed from ensemble data and colormapped to mean isosurfaces. Pöthkow and Hege [PH10] formulated equivalents to isocontours in uncertain data for which they assumed statistically independent distributions. They defined a level crossing probability for each point in a continuous domain and displayed the spatial distribution around the mean isosurface by volume rendering. This visualizes the estimated spatial spread of a crisp average isosurface. Pauly *et al.* [PMG04] presented a formulation of likelihood and confidence maps that describe the possible surface reconstructions from point cloud data for the whole domain. Instead of describing the uncertainty of a single surface, they displayed certainty measures by a cut through a volume rendering. Although no explicit distribution for values is given, this work is most closely related to the previous one, since it also shows a spatial distribution of the most probable surface. In this paper, the idea of [PH10] will be extended to the more general case where spatial correlations are considered.

3. Mathematical Model

3.1. Terminology

We consider a scalar field $g : M \to \mathbb{R}$ on a compact domain $M \subset \mathbb{R}^d$ that has been discretely sampled on nodes $\{\mathbf{x}\}_{i \in I}$ of a grid with data values $\{Y\}_{i \in I}$, where $I = \{1, \ldots, n\}$ labels the sample points. To refer to grid entities like nodes, edges, faces and volume cells, we use the term η-*cell*: a 0-cell is a *vertex*, a 1-cell is an *edge*, a 2-cell is a *polygon*, a 3-cell is a *polyhedron*, and so on. The sampling grid is a d-dimensional grid composed of d-cells that discretize a d-dimensional geometric domain in $\mathbb{R}^d$.

3.2. Isosurfaces for scalar fields

3.2.1. Iso-Contours and Iso-Regions

We recall a few basic definitions and facts, c.f. [Mil63], [PT88], [Mat02]. Assume that the data values are interpolated in M by a smooth function y (C^1 or higher). A point $\mathbf{p}$ in M is called a *critical point* of y if $\nabla y_{\mathbf{p}} = 0$. Other points of M are called *regular points* of y. Given a real number ϑ we call $y^{-1}(\vartheta)$ the ϑ-*level* of y, and we say it is a *critical level* (and that ϑ is a *critical value* of y) if it contains at least one critical point of y. Other real numbers ϑ are called *regular values* of y and the corresponding levels $y^{-1}(\vartheta)$ are called

regular levels. From the inverse function theorem it follows that for a regular value ϑ, $y^{-1}(\vartheta)$ is a smooth, codimension one submanifold of M, which we then call *isocontour* (if it is non-empty). For a critical value ϑ the corresponding critical level $y^{-1}(\vartheta)$ is not manifold. At a saddle point with value ϑ, connected components of the critical level touch. Maxima and minima with value ϑ are isolated points in the ϑ-level.

A critical point $\mathbf{p}$ is called *non-degenerate* if the Hessian $H_{\mathbf{p}}y$ is non-singular, i.e. $det\, H_{\mathbf{p}}y \neq 0$. From the Morse Lemma it follows that non-degenerate critical points are *isolated*. If *all* critical points of function y are non-degenerate (and thus are isolated), y is called *Morse function*. If a function contains *degenerate* critical points, the level set $y^{-1}(\vartheta)$ can be 0- to d-dimensional, since points with $\nabla y_{\mathbf{p}} = 0$ can form arbitrary regions in the domain. We will refer to regions of degenerated points as *plateaus*.

3.2.2. Computational Problems

The foregoing might look like mathematical sophistry, but it is algorithmically relevant. Depending on the interpolation and reconstruction method, the resulting isosurface can vary. In case trilinear interpolation is used, marching cubes type algorithms have to deal with ambiguities. An explanation of these cases and literature on dealing with these ambiguities is given, for example, in [NY06]. Computing level sets for non-Morse functions at a value ϑ containing plateaus is even more complicated. Marching cubes type algorithms that assume $y^{-1}(\vartheta)$ to be a $d-1$-dimensional surface, inevitably fail. An extension of the marching cubes algorithm dealing with degenerate critical levels has been suggested by Weber et al. [WSH03]. Even if the level set contains no critical points, problems occur if it is close to a plateau: the condition number $\|\nabla y(\mathbf{x})\|^{-1}$ is then large, i.e. the computation of the isocontour is ill-conditioned, *independently* of the algorithm [PH10]. The computed results therefore are not reliable and the visual impression of the computed contour can be misleading.

3.3. Basic Probabilistic Setting

We assume that the data values $\{Y\}_{i\in I}$ at the sampling points $\{\mathbf{x}\}_{i\in I}$ are random variables with probability density distributions f_i, means $\mu_i = \mathrm{E}(Y_i)$, finite variances $\sigma_i^2 = \mathrm{E}(Y_i-\mu_i)^2$ and covariances $\mathrm{Cov}(Y_i,Y_j) = \mathrm{E}((Y_i-\mu_i)(Y_j-\mu_j))$ for all $i, j \in I$. Let the true but unknown values of the field being described by the function $h(\mathbf{x})$. We assume then that $h(\mathbf{x}_i) = \mu_i$. Thus, we model the input data as parameter-discrete random field with state space $\mathbb{R}$ whose means coincide with the true values at the sample points.

Given N realizations of the random field $\{y\}_{i\in I}^{j=1,\ldots N}$, i.e. a sample of random fields, the sample means (or empirical means) $\hat{\mu}_i = \frac{1}{N}\sum_{k=1}^{N} y_i^k$ and the entries of the sample covariance matrix $\widehat{\mathrm{Cov}}_{i,j} = \frac{1}{N-1}\sum_{k=1}^{N}(y_i^k-\hat{\mu}_i)(y_j^k-\hat{\mu}_j)$ for all $i, j \in I$ are unbiased estimates of the means $\mathrm{E}(Y_i)$ and covariances $\mathrm{Cov}(Y_i,Y_j)$, respectively.

3.4. Joint Distribution Functions

For a vector $\mathbf{Y}$ of n random variables $\{Y\}_{i\in I}$ the variances and covariances can be represented by a covariance matrix

$$\Sigma = [\mathrm{Cov}(Y_i,Y_j)]_{i=1,2,\ldots,n;j=1,2,\ldots,n}.$$

In case $\mathbf{Y}$ conforms to a multivariate Gaussian distribution $\mathbf{Y} \sim \mathcal{N}_n(\mu,\Sigma)$ with $\mu = [\mathrm{E}(Y_1),\mathrm{E}(Y_2),\ldots,\mathrm{E}(Y_n)]$, it can be uniquely described by a joint probability density function

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}}\exp\left(-\tfrac{1}{2}(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)\right)$$

with $\mathbf{y}\in\mathbb{R}^n$.

3.5. Probabilities of Classes of Realizations

Given a level value ϑ we want to compute probabilities for *classes of realizations of the random field* that are characterized by the fact that $m\,(\leq n)$ random variables Y_i are constrained to subsets S_i. Since the n random variables representing the random field are possibly correlated, we have to integrate the n-dimensional density function $f_{\mathbf{Y}}(y_1,\ldots,y_n)$. Assuming that the n random variables have been ordered such that the constrained random variables are the first m ones, we have to compute integrals of the form

$$\mathrm{Prob}(Y_1\in S_1,\ldots,Y_m\in S_m) = \int_{S_1}\mathrm{d}y_1\ldots\int_{S_m}\mathrm{d}y_m\int_{\mathbb{R}}\mathrm{d}y_{m+1}\ldots\int_{\mathbb{R}}\mathrm{d}y_n\; f_{\mathbf{Y}}(y_1,\ldots,y_n). \quad (1)$$

This means, we have to marginalize the variables $y_{m+1},\ldots,y_n$ and to compute the remaining m-dimensional integral. Marginalization in general might be difficult. However, for multivariate Gaussian distributions we can utilize the nice property that marginalized distributions are again Gaussian distributions with the 'right' means and covariances:

$$\begin{aligned}&\int_{-\infty}^{\infty}\mathrm{d}y_{m+1}\ldots\int_{-\infty}^{\infty}\mathrm{d}y_n\,\frac{1}{(2\pi)^{n/2}\det(\Sigma)^{1/2}}\\ &\qquad\exp\left(-\tfrac{1}{2}(\mathbf{y}-\mu)^T\Sigma^{-1}(\mathbf{y}-\mu)\right)\\ &= \frac{1}{(2\pi)^{m/2}\det(\tilde{\Sigma})^{1/2}}\exp\left(-\tfrac{1}{2}(\tilde{\mathbf{y}}-\tilde{\mu})^T\tilde{\Sigma}^{-1}(\tilde{\mathbf{y}}-\tilde{\mu})\right)\\ &=: f_{\tilde{\mathbf{Y}}}(y_1,\ldots,y_m) \quad (2)\end{aligned}$$

where $\tilde{\mathbf{Y}}$ is the reduced m-dimensional random vector and $\tilde{\mathbf{y}}$, $\tilde{\mu}$ and $\tilde{\Sigma}$ are the quantities $\mathbf{y}$, μ and Σ with the $n-m$ rows/columns *deleted* that correspond to the *marginalized variables* $y_{m+1}\ldots y_n$. Plugging Eq. (2) into Eq. (1) yields

$$\mathrm{Prob}(Y_1\in S_1,\ldots,Y_m\in S_m) = \int_{S_1}\mathrm{d}y_1\ldots\int_{S_m}\mathrm{d}y_m\, f_{\tilde{\mathbf{Y}}}(y_1,\ldots,y_m). \quad (3)$$

In the following we consider only Gaussian fields and use Eq. (3) as basic equation. Note, that we do not require the covariances to be defined by some analytic correlation function (e.g., exponential or linear).

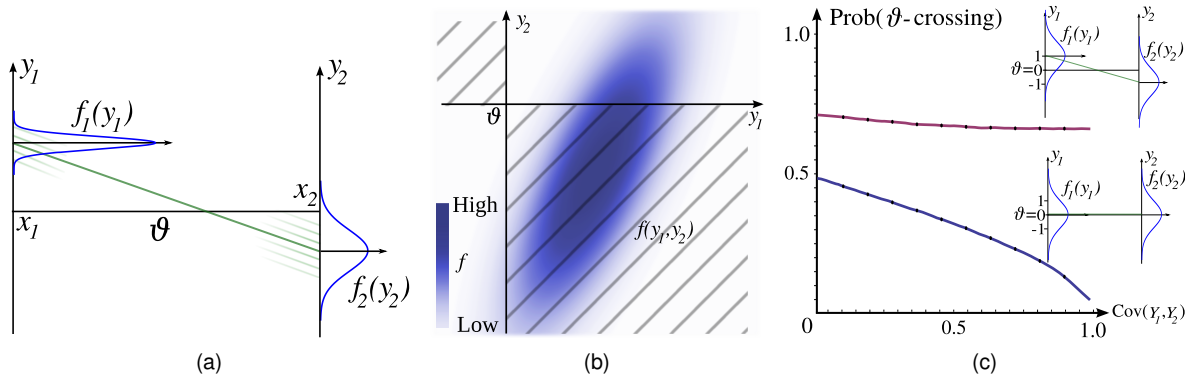

Figure 1: *Example for the 1D case (edge): (a) The marginal distributions at the grid points are shown in blue. Exemplarily, one realization of the linear interpolant is shown (green solid line); other realizations with lower probability are indicated (green transparent lines). In the depicted case the ϑ-level crossing probability is relatively high. In (b) a density plot of the joint distribution with a correlation coefficient of 0.75 is displayed. The quadrants constituting the integration domain for the computation of the level crossing probability are indicated by the hatched grey area. In (c) the impact of changing covariance between two adjacent grid points on the level crossing probability is shown for two pairs of input distributions with $\sigma_i = 1$. As the covariance increases the probability decreases slightly for the first case and significantly for the second case.*

3.6. Level Crossing Probabilities

Any realization $\{y\}_{i\in I}$ of $\mathbf{Y}$ defines a grid function. For any grid function imagine an extension to a C^0 function $g_{\{y\}}$ that is defined in the continuous domain and that interpolates between the sample points $\{\mathbf{x}_i\}_{i\in I}$ such that in each η–cell c ($\eta \leq d$) the extremal values are taken at the vertices of c. Examples for such interpolations are linear interpolation for simplicial cells and $\eta-$linear interpolation for $\eta-$dimensional hexahedral cells.

Let $\tilde{I} \in I$ be the set of indices of the vertex points of cell c. Then cell c crosses the ϑ-level of $g_{\{y\}}$ if and only if in the set of differences $(y_i - \vartheta)_{i\in\tilde{I}}$ both signs occur. Equivalently, cell c does not cross the ϑ-level of $g_{\{y\}}$, if and only if all differences $(y_i - \vartheta)_{i\in\tilde{I}}$ have the same sign.

We want to compute the probability that a η-cell c of the d-dimensional sample grid ($\eta \leq d$) crosses the ϑ-level of interpolated realizations of the random variables $\{Y\}_{i\in I}$. We call this the *ϑ-level crossing probability* of cell c and denote it by $\text{Prob}_c(\vartheta\text{-crossing})$. In order to compute this probability we have to integrate the joint density function of the random variables $\{Y\}_{i\in\tilde{I}}$ over sets $\{y_i \in \mathbb{R} \,|\, y_i \leq \vartheta\}$ and $\{y_j \in \mathbb{R} \,|\, y_j \geq \vartheta\}$ using Eq. (3).

Alternatively we can compute the probability

$$\text{Prob}_c(\vartheta\text{-crossing}) = 1 - \text{Prob}_c(\vartheta\text{-non-crossing}), \qquad (4)$$

which in cells of dimension greater then one is less expensive to calculate.

The general procedure to compute such probabilities can be applied to any type of mesh entity, for example to arbitrary polyhedral cells in grids of arbitrary dimension d. In the following we consider exemplarily edges, rectangles and cuboids (duals of voxels) in such grids. Obviously the procedure can be extended to η-simplices or arbitrary η-polyhedra with $\eta \leq d$.

3.6.1. Edges (1-cells)

For a scalar field in one or more dimensions we consider two random variables Y_0, Y_1 that are associated with adjacent grid points x_1, x_2. Consider the random vector $\mathbf{Y} = [Y_1, Y_2]$ where the joint probability distribution is described by a bivariate Gaussian PDF $f_{\mathbf{Y}}(y_1, y_2)$ with $y_1, y_2 \in \mathbb{R}$, see Fig. 1.

The ϑ-level crossing probability is given by

$$\begin{aligned}\text{Prob}_c(\vartheta\text{-crossing}) = \\ &= \text{Prob}(Y_1 \leq \vartheta, Y_2 > \vartheta) \quad + \text{Prob}(Y_1 > \vartheta, Y_2 \leq \vartheta) \\ &= \int_{y_1\leq\vartheta} \mathrm{d}y_1 \int_{y_2>\vartheta} \mathrm{d}y_2\, f_{\mathbf{Y}}(y_1,y_2) + \int_{y_1>\vartheta} \mathrm{d}y_1 \int_{y_2\leq\vartheta} \mathrm{d}y_2\, f_{\mathbf{Y}}(y_1,y_2)\end{aligned} \qquad (5)$$

Alternatively:

$$\begin{aligned}\text{Prob}_c(\vartheta\text{-non-crossing}) = \\ &= \text{Prob}(Y_1 \leq \vartheta, Y_2 \leq \vartheta) \quad + \text{Prob}(Y_1 > \vartheta, Y_2 > \vartheta) \\ &= \int_{y_1\leq\vartheta} \mathrm{d}y_1 \int_{y_2\leq\vartheta} \mathrm{d}y_2\, f_{\mathbf{Y}}(y_1,y_2) + \int_{y_1>\vartheta} \mathrm{d}y_1 \int_{y_2>\vartheta} \mathrm{d}y_2\, f_{\mathbf{Y}}(y_1,y_2).\end{aligned} \qquad (6)$$

Since the four quadrants $\{(y_1,y_2)|y_1 \leq \vartheta \text{ and } y_2 \leq \vartheta\}$, $\{(y_1,y_2)|y_1 \leq \vartheta \text{ and } y_2 > \vartheta\}$, $\{(y_1,y_2)|y_1 > \vartheta \text{ and } y_2 \leq \vartheta\}$ and $\{(y_1,y_2)|y_1 > \vartheta \text{ and } y_2 > \vartheta\}$ are disjoint and their union is $\mathbb{R}^2$ we can read off Eq. (4).

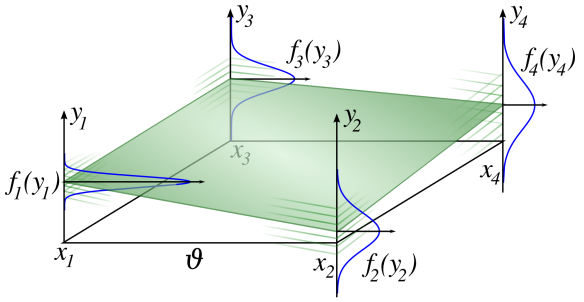

Figure 2: *Example for the computation of a level crossing probability in 2D: The marginal distributions at the grid points are shown in blue. Exemplarily, one realization of the bilinear interpolant is shown in green (the particular one, where all random variables take the value of their means). The ϑ-level crossing probability of the interpolant is relatively low in this specific case.*

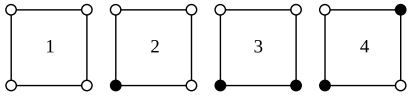

Figure 3: *Four distinct configurations for the marching squares algorithm. The other configurations can be constructed by inverting, rotating and mirroring the grid points. The integrals of the probabilistic formulation correspond to these cases.*

In Fig. 1c the impact of changing covariance between two adjacent grid points on the level crossing probability is shown for two pairs of input distributions with unit standard deviation. As the covariance increases the crossing probability slightly decreases for the first case and decreases significantly for the second case. Recall that high covariance $\mathrm{Cov}(Y_1,Y_2)$ means that, for example, a positive deviation from the mean of a realization of Y_1 also implies a positive deviation of a realization of Y_2. For that reason the effect of decreasing probability is larger for the second case while for the first case the effect is smaller because of the different mean values of the Gaussians.

3.6.2. Rectangles (2-cells)

For a scalar field in two or more dimensions we consider $\mathbf{Y} = [Y_1,Y_2,Y_3,Y_4]$ at the grid points $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$ that are the corners of a pixel, see Fig. 2. The joint probability distribution is described by a four-dimensional Gaussian PDF $f_{\mathbf{Y}}(y_1,y_2,y_3,y_4)$ with $y_1,y_2,y_3,y_4 \in \mathbb{R}$.

Using integrals over $f_{\mathbf{Y}}$ we can compute probabilities for the different cases of the *marching squares* algorithm, see Fig. 3. Probabilities for the four distinctive cases are:

$$\mathrm{P}_{\vartheta,1} = \int_{(y_1>\vartheta \wedge y_2>\vartheta \wedge y_3>\vartheta \wedge y_4>\vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \;\; f_{\mathbf{Y}}(y_1,y_2,y_3,y_4) \quad (7)$$

$$\mathrm{P}_{\vartheta,2} = \int_{(y_1\le\vartheta \wedge y_2>\vartheta \wedge y_3>\vartheta \wedge y_4>\vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \;\; f_{\mathbf{Y}}(y_1,y_2,y_3,y_4) \quad (8)$$

$$\mathrm{P}_{\vartheta,3} = \int_{(y_1\le\vartheta \wedge y_2\le\vartheta \wedge y_3>\vartheta \wedge y_4>\vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \;\; f_{\mathbf{Y}}(y_1,y_2,y_3,y_4) \quad (9)$$

$$\mathrm{P}_{\vartheta,4} = \int_{(y_1\le\vartheta \wedge y_2>\vartheta \wedge y_3\le\vartheta \wedge y_4>\vartheta)} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \;\; f_{\mathbf{Y}}(y_1,y_2,y_3,y_4) \quad (10)$$

The remaining 12 cases can be constructed by rotating and mirroring the grid points.

The level crossing probability for a pixel can be computed by considering the *complement* of the cases where *no* level crossing occurs:

$$\mathrm{Prob}_c(\vartheta\text{-crossing}) = 1 - \int_{\substack{(y_1\le\vartheta \wedge y_2\le\vartheta \wedge y_3\le\vartheta \wedge y_4\le\vartheta) \\ \vee(y_1>\vartheta \wedge y_2>\vartheta \wedge y_3>\vartheta \wedge y_4>\vartheta)}} \mathrm{d}y_1 \int \mathrm{d}y_2 \int \mathrm{d}y_3 \int \mathrm{d}y_4 \;\; f_{\mathbf{Y}}(y_1,y_2,y_3,y_4) \quad (11)$$

3.6.3. Cuboids (3-cells)

For a scalar field in three or more dimensions we consider 8 random variables located at the corners of a cuboid, whose joint probability function is an 8-dimensional Gaussian PDF. Of the $2^8 = 256$ cases, we have 254 cases with crossing (comprised of 14 distinct marching cubes cases) and 2 cases without crossing. The simplest way to compute level crossing probabilities again is to use Eq. (4) and compute probabilities that *no* level crossing occurs, analogously to Eq. (11).

4. Implementation

4.1. Estimation of Means, Variances and Covariances

We compute the multivariate normal distributions for each 3-cell in 3D and 2-cell in 2D of our mesh. The expected values are estimated for each node. The covariance matrix for each distribution consists of the variances of covariances of the four or eight neighboring nodes when considering rectangles or cuboids, respectively (c.f. Sect. 3.5). We compute the covariances from the ensemble values for the 3- and 2-cells using the formula $\widehat{\mathrm{Cov}}_{i,j} = \frac{1}{N-1}\sum_{k=1}^{N}(y_i^k-\hat{\mu}_i)(y_j^k-\hat{\mu}_j)$, where $\hat{\mu}_i$ and $\hat{\mu}_j$ are the respective arithmetic means. In case the spatial resolution of the dataset is too low we resample it by interpolating the ensemble values before we compute the statistical quantities.

4.2. Computing Class Probabilities

From the multivariate normal distributions associated with each grid cell we compute level crossings by Monte-Carlo sampling. Uncorrelated samples conforming to a uniform

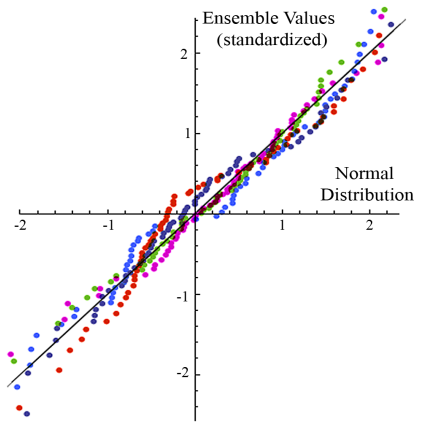

Figure 4: *Assessment of normality for 5 randomly chosen distributions from the temperature field ensemble using a Q-Q-Plot. The distributions do not show severe deviations from the normal distribution, i.e. small differences compared to a linear shape.*

distribution are generated and converted to normally distributed values using the Box-Muller transform [BM58]. These samples are adjusted to the multivariate normal distribution by applying a Cholesky decomposition to the covariance matrix and multiplying the samples with the lower triangle matrix [Gen04, p. 197]. When considering edges we need a vector of two realizations. For rectangles and cuboids we need vectors of 4 and 8 realizations, respectively. The realizations are inspected for agreement with the cases described in Sect. 3.6. From the ratio of samples that agree to the respective cases to those that don't we compute the level crossing probability for each grid cell. Using pseudocode the algorithm can be summarized as follows:

```
for each cell c {
    L_c ← CholeskyDecomposition(Σ_c)
    #crossings ← 0
    for 1...#samples {
        y ← random numbers y_1 ... y_m ~ U(0,1)
        y ← BoxMullerTransform(y)
        y ← L_c y + μ_c
        if(crossing_ϑ(y)) #crossings ← #crossings + 1
    }
    Prob_c ← #crossings/#samples
}
```

5. Results

To illustrate its essential properties we apply the algorithm to synthetic datasets. We show the method's effectiveness for real world data by employing it on an ensemble dataset from climate research. The computations were performed on an Intel Xeon X5550 2.66 GHz system.

5.1. Synthetic Datasets

Fig. 5 shows uncertain isolines for a synthetic 2D dataset. The expected values of the input data (with a low amount of noise) correspond to a sine pattern on the left side of the image that gradually blends to a plateau on the right, see, shown in Fig. 5a. The variances and covariances are constant. The probabilities in the grid of 1.000×250 pixels were computed using 4.000 samples per pixel in 225 seconds. In Fig. 5b the level crossing probabilities for $\vartheta = 0$ are mapped to color while the crisp isoline in the mean value field is shown in black. Uncertain isosurfaces in a synthetic 3D dataset are displayed in Fig. 6. The expected values of the input data are given by the simple analytic formula $\mu(x,y,z) = (\cos(7x) + \cos(7y) + \cos(7z))\exp(-4.5r)$, where $r = \sqrt{x^2+y^2+z^2}$. The variances are constant in the 3 images. A global correlation coefficient (for each pair of vertices) is varied to study the influence of correlation and set to 0 in 6a, to 0.65 in 6b and to 0.95 in 6c. The probabilities in the grid of $256 \times 256 \times 128$ voxels were computed using 1.600 samples each in ≈ 45 minutes for each result. The level crossing probabilities ($\vartheta = 0.013$) are displayed using DVR. A crisp mean isosurface is shown in white.

5.2. Climate Data

We employ the algorithm on daily average hindcast data from the DEMETER project [Pal04] and compare our results to those presented in [PH10]. The dataset contains results of 7 different climate models and 9 different sets of simulation parameters forming an ensemble with 63 members. This representation accounts for both model uncertainty and uncertainty of the input data. The means, variances and covariances are computed from a temperature field ensemble for Feb 20th, 2000. To check whether the assumption of Gaussian distributions of the data values is appropriate we assess normality using a Q-Q plot [Job91, p. 63]. An example is shown in Fig. 4 where quantiles for 5 randomly chosen distributions from the temperature field ensemble are displayed.

Results for the 2 meter temperature field are shown in Fig. 7. The ensemble means are shown in 7a. The level crossing probabilities ($\vartheta = 0°$C) are mapped to color in 7b. The probabilities in the grid of 144×73 pixels were computed using 8.000 samples per pixel in 11 seconds. For comparison the relative count of crisp isolines in the 63 ensemble members crossing the respective grid cell is shown in 7c. The probabilities, computed according to [PH10] and not considering correlation, are displayed in 7d. The corresponding mean isoline is depicted in black. We combine the 2 meter data set with temperature fields of pressure levels 850, 500 and 200 hPa in the earth's atmosphere to obtain a 3D ensemble where the third coordinate represents air pressure. We compute means, variances and covariances for all hexahedral grid cells. Fig. 8 shows uncertain isosurfaces $\vartheta = 0°$C. For Fig. 8a the probabilities are computed as described in [PH10] (not considering correlation). For the

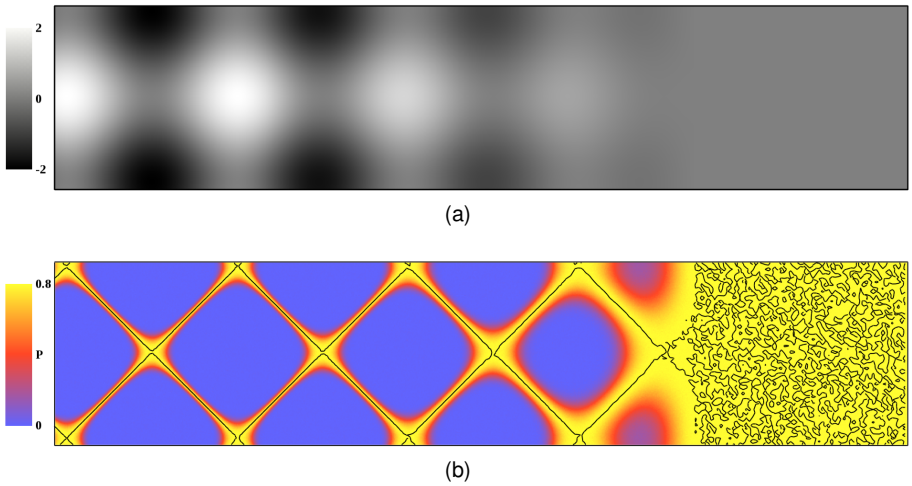

Figure 5: *Uncertain isolines for a synthetic 2D dataset. The expected values conform to a sine pattern (with a low amount of noise added) on the left that gradually approaches a plateau on the right as seen in (a). The variances and covariances are constant. In (b) the probabilities for $\vartheta = 0$ are color mapped while the crisp isoline of the expected values is shown in black. While the computation of isolines is ill-conditioned at critical points (especially plateaus) the probabilistic ansatz does not suffer from this problem and calculates high probabilities for the whole plateau.*

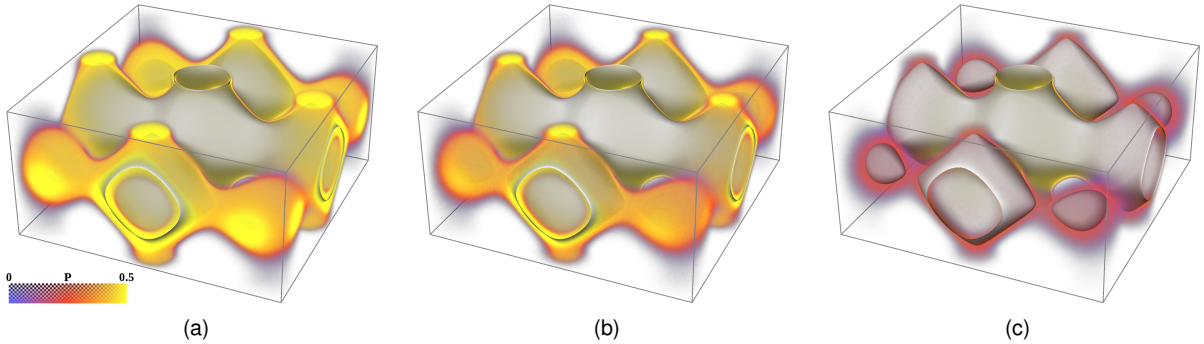

Figure 6: *Uncertain isosurfaces in a synthetic 3D dataset. The expected values are given by an analytic formula and the variances are constant. The correlation coefficient is globally set to* 0 *in (a), to* 0.65 *in (b) and to* 0.95 *in (c). The probabilities are displayed using direct volume rendering and a crisp isosurface of the expected values is shown in white. The results show that increasing correlation between the grid points (from left to right) decreases the level crossing probabilities in the proximity of the mean surface and leads to more localized spatial distributions of uncertain isocontours.*

probabilities that are displayed in Fig. 8a correlation was taken into account. The probabilities in a grid resampled to $432 \times 219 \times 68$ voxels were computed using 8.000 samples per pixel in 194 minutes.

6. Discussion

In the results the impact of spatial correlation on the level crossing probabilities is clearly visible. Fig. 6 shows that increasing correlation between the grid points decreases the probabilities surrounding the mean crisp surface which leads to thinner spatial distributions of uncertain isocontours. As we can see in Fig. 7 neglecting correlation leads to overestimation of the uncertain isolines' spatial distributions. The distribution in Fig. 7b is thinner than the distribution in Fig. 7d and it is similar to Fig. 7c which we regard as a ground truth. In Fig 8 the probabilities for the cuboid case show thinner spatial distributions compared to the method described in [PH10] that neglects correlation.

Fig. 5b illustrates that the computation of crisp isolines is ill-conditioned at critical points (especially plateaus) while the probabilistic ansatz does not suffer from this problem and calculates high probabilities for the whole plateau.

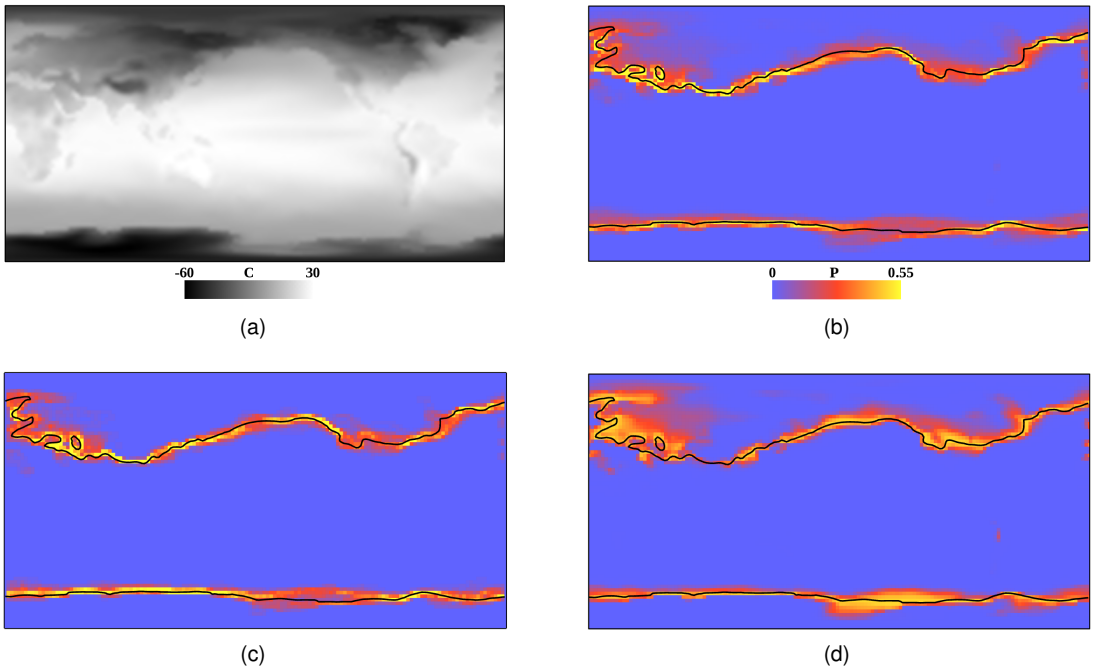

Figure 7: *Results for a 2 metre temperature field from climate simulations: The ensemble means are shown in (a). The level crossing probabilities for $\vartheta = 0^\circ C$ are colormapped in (b). For comparison the relative count of crisp isolines in the 63 ensemble members crossing the respective grid cell is shown in (c). The results computed according to [PH10] (not considering correlation) are shown in (d). While the latter result overestimates the spatial distribution of the uncertain isoline the distribution in (b) is more localized and similar to (c).*

The current implementation is limited to multivariate Gaussian distributions. However, in many applications, e.g. the climate data used in this paper, the assumption of Gaussian distributions can be considered quite suitable. The distributions shown in Fig. 4 as well as all the others that we inspected interactively do not show severe deviations from the normal distribution. The generalization to non-Gaussian distributions will not be straightforward because no simple solution (e.g. similar to Eq. (2)) for the marginalization is available. Furthermore, in most cases no direct representation of correlation that is comparable to the covariance matrix exists for non-Gaussian PDFs. For other noise models such as multiplicative noise an adapted approach will be necessary.

In Figures 6 and 8 we displayed crisp isosurfaces of the mean values because it represents the most probable locations of the isosurface. This surface is augmented with the volume rendered level crossing probabilities. This visual design corresponds to traditional 2D plots with error bars where the mean value is shown like a crisp, certain value that is augmented with the standard deviation. We believe that there is need for further research on visual designs conveying uncertainty as well as assessment of their visual effectiveness.

7. Conclusions

We proposed a method for quantification of spatial uncertainty of isocontours considering arbitrary spatial correlations of the probability distributions of the input data. We modeled the uncertain scalar field as a discrete random field for which we computed level crossing probabilities. In this paper we concentrated on Gaussian distributed data which is the most important case in practice (a theoretical explanation is given by the central limit theorem). Correlated data is the normal case not the exception. Generalizing the model of [PH10] to correlated random fields is therefore the main contribution of this work. Even if the raw input data consists of uncorrelated distributions (for example values affected by white noise) considering correlation can be necessary for derived datasets. For example, the use of convolution kernels on single realizations of the data can lead to spatially correlated distributions in the processed data.

The probabilistic procedure does not have to deal with degenerate or ambiguous cases separately as it is the case for marching cubes and related algorithms. Probabilities for the occurrence of level crossings for critical isovalues or non-Morse functions (with respect to the mean values, for example) are computed correctly without treating any special

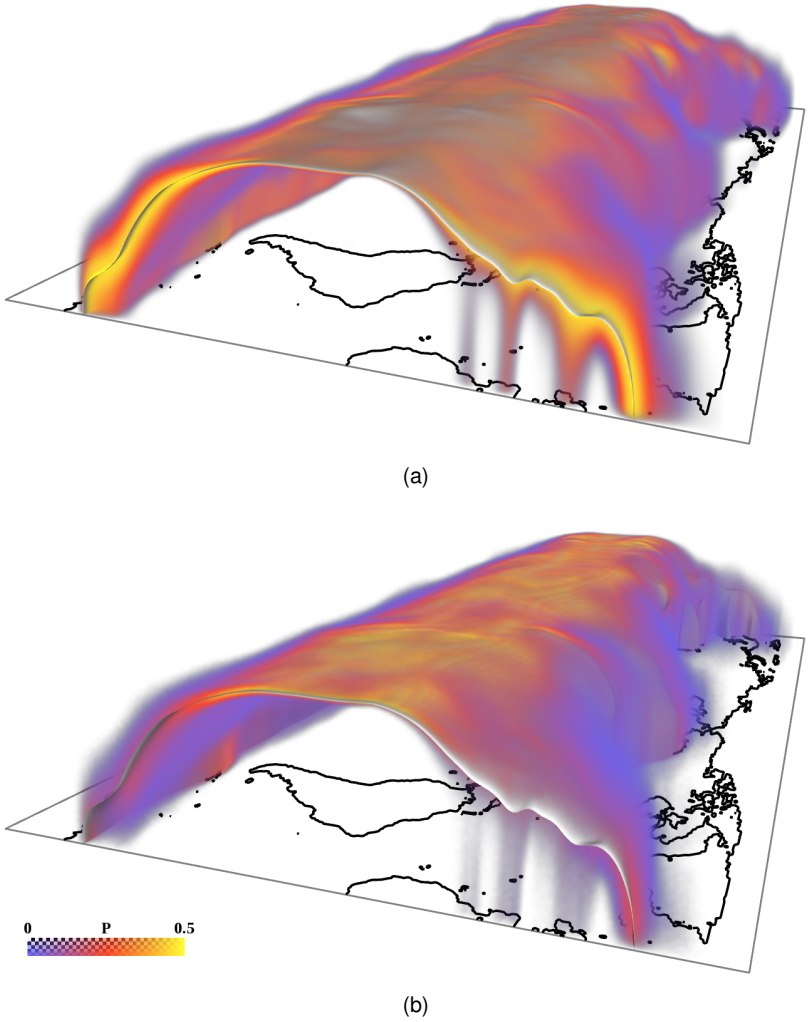

Figure 8: *Uncertain isosurfaces $\vartheta = 0^\circ C$ in a 3D temperature field. In Fig. (a) the probabilities computed using the formulation in [PH10] (not considering correlation) are shown. For Fig. (b) correlation was considered and the level crossing probabilities reveal a more localized spatial distribution of the uncertain isosurface.*

cases. This property is shared by both the approach presented here and the one from [PH10].

Due to the consideration of correlation our approach leads to more accurate results compared to [PH10]. In contrast to [PH10] we associate a probability for the existence of a level crossing with each cell (of each dimension 1, ..., d) in some grid and not with every point in a continuous domain. To compute level crossing probabilities for rectangles or cuboid cells, we solve 4- or 8-dimensional integrals using a Monte Carlo method. This is computationally more expensive than the evaluation of the analytic formulas used in [PH10] (partly as approximations). In special cases where the input data contains spatially uncorrelated distributions or the correlation structure can be safely omitted the approach in [PH10] should be preferred because of greater simplicity and speed. In case a correlation structure is known and realtime computation is not required our approach should be preferred to get more accurate results. The Monte Carlo computations are embarrassingly parallel. Currently, we are working on an efficient implementation for GPUs and expect that this will improve performance significantly.

Acknowledgments

This work was supported by the German Research Foundation (DFG HE 2948/5-3) and the Max Planck Institute of Molecular Cell Biology and Genetics (Segmentation of microtubules from electron tomograms).

References

[BM58] BOX G. E. P., MULLER M. E.: A Note on the Generation of Random Normal Deviates. *The Annals of Mathematical Statistics 29*, 2 (June 1958), 610–611. 6

[DKLP02] DJURCILOV S., KIM K., LERMUSIAUX P., PANG A.: Visualizing scalar volumetric data with uncertainty. *Computers & Graphics 26*, 2 (2002), 239–248. 2

[FKLT10] FENG D., KWOCK L., LEE Y., TAYLOR R.: Matching visual saliency to confidence in plots of uncertain data. *IEEE Transactions on Visualization and Computer Graphics 16* (2010), 980–989. 2

[Gen04] GENTLE J. E.: *Random Number Generation and Monte Carlo Methods*, 2nd ed. Springer, New York, 2004. 6

[GR04] GRIGORYAN G., RHEINGANS P.: Point-based probabilistic surfaces to show surface uncertainty. *IEEE Transactions on Visualization and Computer Graphics 10*, 5 (2004), 564–573. 2

[GS06] GRIETHE H., SCHUMANN H.: The visualization of uncertain data: Methods and problems. In *SimVis* (March 2006), Schulze T., Horton G., Preim B., Schlechtweg S., (Eds.), SCS Publishing House e.V., pp. 143–156. 2

[Job91] JOBSON J.: *Applied Multivariate Data Analysis Volume 1: Regression and Experimental Design*. Springer, New York, 1991. 6

[KUS*05] KNISS J. M., UITERT R. V., STEPHENS A., LI G.-S., TASDIZEN T., HANSEN C.: Statistically quantitative volume visualization. In *Proceedings of IEEE Visualization 2005* (October 2005), pp. 287–294. 2

[LKP03] LUO A., KAO D., PANG A.: Visualizing spatial distribution data sets. In *VISSYM '03: Proceedings of the Symposium on Data Visualization 2003* (2003), Eurographics Association, pp. 29–38. 2

[LLPY07] LUNDSTRÖM C., LJUNG P., PERSSON A., YNNERMAN A.: Uncertainty visualization in medical volume rendering using probabilistic animation. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1648–1655. 2

[Mat02] MATSUMOTO Y.: *An Introduction to Morse Theory*. American Mathematical Society, 2002. 2

[Mil63] MILNOR J. W.: *Morse Theory*. Princeton Univ Press, 1963. 2

[MRH*05] MACEACHREN A., ROBINSON A., HOPPER S., GARDNER S., MURRAY R., GAHEGAN M., HETZLER E.: Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science 32*, 3 (2005), 139–161. 2

[NY06] NEWMAN T. S., YI H.: A survey of the marching cubes algorithm. *Computers & Graphics 30*, 5 (2006), 854 – 879. 3

[OGHT10] OTTO M., GERMER T., HEGE H.-C., THEISEL H.: Uncertain 2D vector field topology. *Comput. Graph. Forum 29* (2010), 347–356. 2

[Pal04] PALMER ET AL. T. N.: *Development of a European Multi-Model Ensemble System for Seasonal to Inter-Annual Prediction (DEMETER)*. Technical memorandum, European Centre for Medium-Range Weather Forecasts, Reading, England, 2004. 6

[PH10] PÖTHKOW K., HEGE H.-C.: Positional uncertainty of isocontours: Condition analysis and probabilistic measures. *IEEE Transactions on Visualization and Computer Graphics*, PrePrints (2010). doi:10.1109/TVCG.2010.247. 1, 2, 3, 6, 7, 8, 9

[PMG04] PAULY M., MITRA N., GUIBAS L.: Uncertainty and variability in point cloud surface data. In *Eurographics Symposium on Point-Based Graphics* (2004), pp. 77–84. 2

[PRH10] PRASSNI J.-S., ROPINSKI T., HINRICHS K.: Uncertainty-aware guided volume segmentation. *IEEE Transactions on Visualization and Computer Graphics 16* (2010), 1358–1365. 2

[PT88] PALAIS R. S., TERNG C.-L.: *Critical Point Theory and Submanifold Geometry*, vol. 1353, Lecture Notes in Mathematics. Springer-Verlag, 1988. 2

[PWB*09] POTTER K., WILSON A., BREMER P.-T., WILLIAMS D., DOUTRIAUX C., PASCUCCI V., JOHHSON C. R.: Ensemble-vis: A framework for the statistical visualization of ensemble data. In *IEEE Workshop on Knowledge Discovery from Climate Data: Prediction, Extremes.* (2009), pp. 233–240. 2

[PWL97] PANG A. T., WITTENBRINK C. M., LODHA S. K.: Approaches to uncertainty visualization. *The Visual Computer 13*, 8 (1997), 370–390. 2

[RLBS03] RHODES P. J., LARAMEE R. S., BERGERON R. D., SPARR T. M.: Uncertainty visualization methods in isosurface rendering. In *Eurographics 2003, Short Papers* (2003), pp. 83–88. 2

[SHM10] SAAD A., HAMARNEH G., MOLLER T.: Exploration and visualization of segmentation uncertainty using shape and appearance prior information. *IEEE Transactions on Visualization and Computer Graphics 16* (2010), 1366–1375. 2

[SZD*10] SANYAL J., ZHANG S., DYER J., MERCER A., AMBURN P., MOORHEAD R.: Noodles: A tool for visualization of numerical weather model ensemble uncertainty. *IEEE Transactions on Visualization and Computer Graphics 16* (2010), 1421–1430. 2

[WSH03] WEBER G. H., SCHEUERMANN G., HAMANN B.: Detecting critical regions in scalar fields. In *VisSym* (2003), Eurographics Association, pp. 85–94. 3

[ZWK10] ZEHNER B., WATANABE N., KOLDITZ O.: Visualization of gridded scalar data with uncertainty in geosciences. *To appear in Computers & Geoscience* (2010). 2