Accepted Manuscript

Ensemble-based characterization of uncertain environmental features

Rafał Wójcik, Dennis McLaughlin, Seyed Hamed Alemohammad, Dara Entekhabi

PII: S0309-1708(14)00066-9
DOI: http://dx.doi.org/10.1016/j.advwatres.2014.04.005
Reference: ADWR 2186

To appear in: Advances in Water Resources

Please cite this article as: Wójcik, R., McLaughlin, D., Alemohammad, S.H., Entekhabi, D., Ensemble-based characterization of uncertain environmental features, Advances in Water Resources (2014), doi: http://dx.doi.org/10.1016/j.advwatres.2014.04.005

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Ensemble-based characterization of uncertain environmental features

Rafał Wójcik*, Dennis McLaughlin, Seyed Hamed Alemohammad, Dara Entekhabi

*Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

This paper considers the characterization of uncertain spatial features that cannot be observed directly but must be inferred from noisy measurements. Examples of interest in environmental applications include rainfall patterns, solute plumes, and geological features. We formulate the characterization process as a Bayesian sampling problem and solve it with a non-parametric version of importance sampling. All images are concisely described with a small number of image attributes. These are derived from a multidimensional scaling procedure that maps high dimensional vectors of image pixel values to much lower dimensional vectors of attribute values. The importance sampling procedure is carried out entirely in terms of attribute values. Posterior attribute probabilities are derived from non-parametric estimates of the attribute likelihood and proposal density. The likelihood is inferred from an archive of noisy operational images that are paired with more accurate

*Corresponding author

Email addresses: rwojcik@mit.edu (Rafał Wójcik), dennism@mit.edu (Dennis McLaughlin), hamed_al@mit.edu (Seyed Hamed Alemohammad), darae@mit.edu (Dara Entekhabi)

Preprint submitted to Elsevier April 5, 2014
ground truth images. Proposal samples are generated from a non-stationary multi-point statistical algorithm that uses training images to convey distinctive feature characteristics. To illustrate concepts we carry out a virtual experiment that identifies rainy areas on the Earth’s surface from either one or two remote sensing measurements. The two sensor case illustrates the method’s ability to merge measurements with different error properties. In both cases, the importance sampling procedure is able to identify the proposals that most closely resemble a specified true image.

1. Introduction

In many fields there is a need to characterize uncertain spatial features that cannot be observed directly. Examples of particular interest in environmental applications include characterization of surface rainfall in ungaged areas [59], tracking of subsurface solute plumes, and identification of geological features such as groundwater aquifers, mineral deposits, and oil reservoirs [26, 14, 6]. Remote sensing measurements such as passive and active satellite microwave, geodetic, or seismic observations can often provide useful but imperfect information about uncertain features. The statistical attributes of these noisy measurements can be estimated by comparing them to more accurate but less readily available "ground truth" measurements. For example, ground-based weather radar and rain gage data can serve as ground truth for an evaluation of errors in satellite microwave measurements that have greater coverage but may be less accurate. In subsurface flow borehole measurements can serve as local ground truth for seismic or other remotely sensed data.
The diverse measurements collected in a typical environmental assessment can be conveniently combined in a Bayesian framework that considers all available sources of information. In a feature-based application images are characterized by vectors of appropriate variables (e.g. pixel values or feature attributes). The Bayesian approach conditions a prior distribution of the uncertain true image/vector on a set of noisy measured images/vectors, yielding a posterior distribution that characterizes the uncertainty remaining after the measurements are taken into account.

Many researchers have investigated methods for using noisy measurements to characterize complex environmental features. One option is to identify a point estimate (a single image) that optimizes an appropriate deterministic performance objective (e.g. a least-squares measure of misfit between the estimate and a measured image). These methods can be viewed as Bayesian a posteriori estimators (i.e. estimators of the posterior mode) if Gaussian assumptions are adopted. Feature-based optimization methods often use level-set techniques to characterize irregular and/or disconnected feature boundaries [32, 3]. Level set optimization methods are flexible and popular in the image processing community but they do not generate posterior distributions of uncertain feature variables. Posterior distributions are needed for probabilistic applications such as ensemble forecasting, risk assessment, extreme value analysis, and stochastic control.

Ensemble methods such as Markov Chain Monte Carlo (MCMC) and importance sampling provide a probabilistic characterization that yields samples from the Bayesian posterior [17, 7, 1, 44]. However, these methods often rely on parametric prior and measurement error distributions (e.g. the
multivariate Gaussian distribution) and are generally not applied to feature characterization problems. This reflects the fact that very large sample sizes are needed to properly characterize the non-parametric multivariate probability densities of high-dimensional image vectors \([12, 65]\). Spatial features such as geological facies, solute plumes, and rain storms are usually difficult to describe with parametric probability models. The multipoint statistical approach used in many geological applications [refs] provides an alternative that generates realistic non-parametric prior or proposal samples from training images. However, multipoint methods are generally unable to sample from the Bayesian posterior distribution.

An ideal approach for characterizing uncertain features would combine the level set method’s ability to handle complex geometries, the general Bayesian probabilistic framework provided by MCMC and importance sampling, and the non-parametric samples generated by multi-point statistical techniques. This paper describes an approximate method for integrating these different capabilities in a single characterization procedure. It adopts a non-parametric importance sampling approach with proposal samples generated from training images and measurement error samples generated from archived data. The computational limitations of ensemble sampling are partially circumvented by representing each image with a small vector of feature attributes that provides a more concise and efficient description than a classical pixel-based description. Since it is difficult to specify in advance a set of universal attributes that adequately characterize complex features we use a multidimensional scaling technique to derive the attributes directly from image pixel values.
In the following sections we first formulate an attribute-based approach to the feature characterization problem, using importance sampling to generate approximate probability-weighted samples from the Bayesian posterior. We then consider how multidimensional scaling can identify image attributes from pixel-based image vectors and how the priors and likelihoods used in importance sampling can be generated from archived attribute data. The general concepts are illustrated with a virtual experiment based on real rainfall measurements. The paper concludes with a discussion of conceptual and computational issues that identify directions for future research.

2. Image based importance sampling

Importance sampling is a procedure that generates samples from a posterior probability distribution that is related through Bayes theorem to a prior distribution and a likelihood function. Following the approach outlined above, we formulate the feature-based importance sampling problem in terms of a small number of distinctive image attributes. In particular, we define a random vector \( \hat{x} \) composed of true image attributes. The true image is observed by one or more sensors that produce noisy images, which we call current operational measurements to distinguish them from the archived operational measurements discussed below. The current measurement from sensor \( r \) is described by an attribute vector \( \hat{z}_r \). Attributes for all of the measurements are assembled in the global measurement vector \( \hat{z} \). Section 3.4 describes how the image attributes are derived.

The objective of importance sampling is to generate a set of samples \( \hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{N_x} \) of \( \hat{x} \) from the posterior probability density \( p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}) \). This
density can be expressed, through Bayes theorem, as:

$$p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}) = \hat{c} p_{\hat{Z}|\hat{X}}(\hat{z}|\hat{x}) p_{\hat{X}}(\hat{x})$$  \hspace{1cm} (1)$$

where $p_{\hat{Z}|\hat{X}}(\hat{z}|\hat{x})$ is the likelihood function, $p_{\hat{X}}(\hat{x})$ is a prior probability density that does not depend on the measurements, and $\hat{c}$ is a proportionality constant selected to insure that the posterior density integrates to unity. The prior density quantifies our prior uncertainty about the true image while the likelihood function describes the effects of measurement error. Since it is difficult to sample directly from the posterior density we work instead with a more convenient set of equally likely random samples from a proposal density $q(\hat{x})$. Unlike the prior, the proposal density can depend on the current measurements. We use the multipoint statistical techniques described below to generate proposal samples. An approximate posterior probability distribution can be derived by appropriate weighting of these proposal samples. It is convenient to illustrate the process by considering the following equivalent expressions for the conditional expectation of $\hat{x}$ over $p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z})$ (for a given $\hat{z}$):

$$E_{p}[\hat{x}|\hat{z}] = \int \hat{x} p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}) d\hat{x} = \int q(\hat{x}) \left[ \hat{x} \frac{p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z})}{q(\hat{x})} \right] d\hat{x} = E_{q} \left[ \hat{x} \frac{p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z})}{q(\hat{x})} \right]$$  \hspace{1cm} (2)$$

This mean can be estimated from the $N_x$ proposal samples:

$$E_{q} \left[ \hat{x} \frac{p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z})}{q(\hat{x})} \right] \approx \frac{\hat{c}}{N_x} \sum_{i=1}^{N_x} \frac{p_{\hat{Z}|\hat{X}}(\hat{z}|\hat{x}_i)p_{\hat{X}}(\hat{x}_i)}{q(\hat{x}_i)} = \int \hat{x} \left[ \frac{\hat{c}}{N_x} \sum_{i=1}^{N_x} \frac{p_{\hat{Z}|\hat{X}}(\hat{z}|\hat{x}_i)p_{\hat{X}}(\hat{x}_i)}{q(\hat{x}_i)} \delta(\hat{x} - \hat{x}_i) \right] d\hat{x}$$  \hspace{1cm} (3)$$

where $\hat{x}_i$ is sampled from $q(\hat{x})$. The final bracketed expression in (3) is a discrete approximation to the posterior density $p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z})$ that appears in
the second term in (2). This expression can be re-written as:

\[ p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}) \approx \sum_{i=1}^{N_x} w_i \delta(\hat{x} - \hat{x}_i) \] (4)

where:

\[ w_i = \frac{\hat{c} \cdot p_{\hat{Z}|\hat{X}}(\hat{z}|\hat{x}_i)p_{\hat{X}}(\hat{x}_i)}{q(\hat{x}_i)} \] (5)

and \( \hat{c} \) is selected such that \( \sum_i w_i = 1 \). Equation (4) tells us that the proposal image attribute vectors \( \hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{N_x} \) can be interpreted as samples from the posterior density if they are assigned the weights \( \hat{w}_1, \hat{w}_2, \ldots, \hat{w}_{N_x} \) computed in (5) rather than the equal weights that apply when they are treated as samples from the proposal density. So the proposal and posterior samples have the same values but different probabilities. Note that, in the special case where the proposal density is the same as the prior the \( p_{\hat{X}}(\hat{x}_i) \) and \( q(\hat{x}_i) \) terms in (5) cancel and the weights only depend on the likelihood. However, it is usually best to draw proposal samples from a distribution that depends on the current measurement rather than from the prior distribution, which does not.

3. Generating the information needed for importance sampling

The information sampling approach outlined in Section 2 needs a set of proposals \( \hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{N_x} \), the likelihood function \( p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}) \), and the prior \( p_{\hat{X}}(\hat{x}) \) and proposal \( q(\hat{x}) \) probability densities. The prior density can be specified directly in terms of attributes since it does not depend on measurements. The process for generating proposals and for obtaining the likelihood and proposal density functions involves several steps that start with high-dimensional
pixel-based image vectors and concludes with the low-dimensional attribute-based importance weight calculations of (5). The overall flow of information is summarized in Fig.1 and the steps are described in more detail in the following subsections.

3.1. Data sources

It is essential to rely on low-dimensional attribute vectors in an image-based importance sampling procedure. But it is more convenient to work with high-dimensional vectors of pixel values when generating proposals and compiling the information needed to construct attribute-based probability density functions. To clarify the process it is convenient to define the various images and associated pixel and attribute vectors used in the importance sampling procedure:

1. **True image**: One image described by a vector \(x_T\) with \(N_p\) pixel values. The true image is unknown but assumed to be adequately approximated by one or more of the proposals.

2. **Current operational measurements**: \(N_z\) noisy images derived from \(N_z\) different operational sensors that all observe the same true image. These sensors produce measurements that are lower quality (lower resolution, less accurate, etc.) than the ground truth measurements (see below) but are more readily available, more convenient, or with better coverage. The current operational measurements are described by the vectors \(z_1, z_2, \ldots, z_{N_z}\), each with \(N_p\) pixel values, or by smaller vectors \(\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_{N_z}\), each with \(N_a\) attribute values. The pixel-based current measurement vectors for all of the \(r = 1, \ldots, N_z\) sensors are assembled in the \(N_pN_z\)-dimensional composite measurement vector \(z\) while
the corresponding attribute-based vectors are assembled in the $N_aN_z$-dimensional composite measurement vector $\hat{z}$.

3. **Proposals**: $N_x$ images described by the vectors $x_1, x_2, \ldots, x_{N_x}$, each with $N_p$ pixel values, or by $N_x$ smaller vectors $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{N_x}$, each with $N_a$ attribute values. These proposals are candidate images that are ranked by the importance sampling procedure according to the probability that they correspond to the true image.

4. **Archived operational measurements**: $N_z$ sets of noisy images. Each set contains $N_u$ images derived from a particular sensor. The images for sensor $r$ (where $r = 1, \ldots, N_z$) are described by vectors $u_{1,r}, u_{2,r}, \ldots, u_{N_u,r}$, each with $N_p$ pixel values, or by smaller vectors $\hat{u}_{1,r}, \hat{u}_{2,r}, \ldots, \hat{u}_{N_u,r}$, each with $N_a$ attribute values. All of the $u_{j,r}$ vectors are assembled in the $N_pN_uN_z$-dimensional composite measurement vector $u$ while all of the $\hat{u}_{j,r}$ vectors are assembled in the $N_aN_uN_z$-dimensional composite measurement vector $\hat{u}$.

5. **Archived ground truth measurements**: $N_u$ images derived from a reference sensor with smaller errors than the operational sensors. These images are described by the vectors $v_1, v_2, \ldots, v_{N_u}$, each with $N_p$ pixel values, or by $N_u$ smaller vectors $\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_{N_u}$, each with $N_a$ attribute values. All of the $v_j$ vectors are assembled in the $N_pN_u$-dimensional composite measurement vector $v$ while all of the $\hat{v}_j$ vectors are assembled in the $N_aN_u$-dimensional composite measurement vector $\hat{v}$.

The archived images from a given operational sensor are paired with corresponding ground truth images from the reference sensor (each pair has a
particular subscript $j$). These paired measurements may be used to estimate measurement error probability densities, as discussed in Section 3.3. This process of estimating error properties from archived measurements is not essential to the importance sampling approach but, if feasible, provides a convenient way to develop non-parametric probabilistic measurement error models.

3.2. Proposal generation

Importance sampling works much better in feature characterization applications if the proposals it uses are realistic images that share important features with the unknown true image. At the same time, they must be sufficiently different from one another to properly represent uncertainty. It is possible to balance these requirements by generating conditional realizations from an appropriate training image (or images).

There are a number of ways to generate random but structured images from training images. One such procedure is the multipoint statistical generator [28] provided in the SNESIM package [57, 58, 53] used in our example. When archived ground truth measurements are available it is convenient to use them as proposal training images since they reveal a range of features observed in the past. In other situations where such archival measurements are not available it may be sufficient to generate a training image manually, for example by sampling and transforming features observed in the current measurement image. Figure 2c shows some typical unconditional realizations generated from SNESIM using the archived ground truth measurement of Fig. 2a as a training image. This figure relies on some of the images from our rainfall example, which adopts a categorical (binary) description.
of intensity (i.e. red indicates rain and blue indicates no rain). Similar categorical characterizations are convenient in subsurface applications, where different categories represent different geological facies. The multi-point approach is sufficiently general to handle training images with continuously varying intensities, although realization generation in such cases is usually more computationally demanding. In the example, note that the red areas in the realizations are approximately the size and shape of the red areas in the training image but they can occur anywhere in the image domain.

We can modify unconditional realizations to obtain non-stationary behavior if all realizations are constrained to reproduce intensity values specified in particular pixels. These could be values from a current measurement image, a training image, or some other source. The resulting non-stationary image has different ensemble probabilities in different locations. This is illustrated in Fig. 2d, where the realizations are still derived from the training image of Fig. 2a but also constrained by intensity values at specified pixels (black dots) in the current operational measurement of Fig. 2b. The constraining pixels are distributed randomly over the domain, constituting 1% of the total number of pixels. The red and blue areas in the resulting realizations tend to occur in the general vicinity of the corresponding areas in the current measurement image. There is still variability among realizations but not as much as in the unconditional case. The SNESIM multipoint image generation procedure provides a convenient way to construct proposals for importance sampling. We can generate a range of proposals, some that look very much like the measurement and some that look much different, by varying the fraction of pixels used to constrain the unconditional realizations.
This is illustrated in Fig. 3, again using images from the rainfall example introduced in Section 3.1. Here the fraction of constraining points varied from 1% to 25%. The resulting ensemble in Fig. 3c provides a diverse set of proposals, some close to the current measurement and others that are quite different. The true image is shown in Fig. 3a for comparison.

There are many variants on the proposal generation procedure outlined above. The SNESIM algorithm and the conditioning approach described above are adequate for the applications considered here. In other applications with rather different features or continuous intensities other methods may be appropriate. The important point is that we can generate “realistic” proposal images in a manner compatible with our application, without ever specifying an explicit probability model for the high-dimensional pixel vectors used to characterize these images. This approach provides considerable flexibility and makes it more likely that the importance sampling proposals will be similar to the unknown true image.

Once the proposals are generated as high-dimensional vectors of pixel values they can be transformed through the multi-dimensional scaling procedure into low-dimensional vectors of image attributes suitable for importance sampling. Then, kernel density methods [50, 56, 61] or mixture models [55, 22, 35] can be used to fit a continuous $N_a$-dimensional multivariate probability density $q(\hat{x})$ to the proposal attribute vectors $\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_{N_a}$. The posterior importance sampling weights in (5) depend on the values of this density obtained at the proposal attribute vectors. Note that a proposal density is needed only in the low-dimensional attribute space, not in the pixel space.
3.3. Measurement errors

In Bayesian sampling applications likelihood functions are typically derived from models of the measurement error process. The most common model is to assume that measurements are the sum of a known function of the true (hidden) variable and a random noise term. When expressed in terms of image attributes this model may be written:

$$\hat{z} = H\hat{x} + \hat{\epsilon}$$

(6)

where $H$ is an $N_z N_a \times N_a$ matrix that relates the measurement and true image attributes and $\hat{\epsilon}$ is an $N_z N_a$-dimensional random error vector described by the multivariate probability density $p_{\hat{\epsilon}}(\hat{\epsilon})$. In the feature characterization problems considered here the measurement operator is linear but the general approach also accommodates nonlinear operators.

Equation (6) implies that the attribute-based likelihood for any given proposed true value $\hat{x} = \hat{x}_i$ may be written as:

$$p_{\hat{z}|\hat{x}}(\hat{z}|\hat{x}_i) = p_{\hat{\epsilon}}(\hat{z} - H\hat{x}_i)$$

(7)

Although an additive error can always be defined as the difference between a noisy measurement and an accurate ground truth image, an additive measurement model is useful only if the errors do not depend on the true image. In this case, the error statistics can be specified without knowing the true feature. We use the commonly adopted additive error assumption here because measurement attribute errors in our rainfall example do, indeed, appear to be independent of the true value, based on comparisons of the archival operational and ground truth attribute values. In other applications measurement attribute errors may not be additive and different measurement error
models should be considered. In any case, these error models will rely on the attribute vectors obtained from archived operational and ground truth measurements. If available, the archived operational and ground truth measurements mentioned in Section 3.1 provide a good basis for constructing a realistic non-parametric measurement error density. Figure 4 shows some typical pairs of archived ground truth measurements (Fig. 4a) and corresponding archived operational measurements (Fig. 4b) taken from our rainfall example. The difference images obtained by subtracting each ground truth measurement from its operational counterpart are shown in Fig. 4c. Corresponding images are located in the same positions in the figure.

The operational measurements can have different numbers of distinct features (identified by the red areas) than the ground truth measurements, with different sizes and locations. Nevertheless, they clearly resemble their ground truth counterparts. The measurement differences depend on the true image, contrary to the independent additive error assumption typically made in Bayesian statistics. For example, note that the top left image, where the red areas in both measurements lie in the bottom half of the frame. The differences for this image are also confined to the bottom half, indicating a correlation with the true image. As mentioned above, such correlations sometimes decrease or even disappear when the images are described in terms of attribute rather than pixel values.

We can estimate the attribute measurement error probability density from paired vectors of archived operational and ground truth image attributes (e.g. from attribute vectors derived from the images shown in Figures 4a and Fig. 4b). When the attribute measurement errors are additive the differences
\( \hat{u}_{j,r} - \hat{v}_j \) between the archived attribute vectors for sensor \( r \) can be treated as random samples from the population of actual measurement attribute errors described by \( p_{\hat{\epsilon}}(\hat{\epsilon}) \). Kernel density or mixture models [50, 56, 49, 22] can be used to fit a continuous \( N_a \)-dimensional probability density \( p_{\hat{\epsilon}}(\hat{\epsilon}) \) to the \( \hat{u}_{j,r} - \hat{v}_j \) samples. We can then combine (5) and (7) to obtain an importance weight expression that depends directly on \( p_{\hat{\epsilon}} \):

\[
\begin{align*}
    w_i &= \frac{\hat{c}}{N_x} \frac{p_{\hat{\epsilon}}(\hat{\epsilon}_1 - H_1 \hat{x}_i)p_{\hat{\epsilon}}(\hat{\epsilon}_2 - H_2 \hat{x}_i) \ldots p_{\hat{\epsilon}}(\hat{\epsilon}_{N_z} - H_{N_z} \hat{x}_i)}{q(\hat{x}_i)} \\
    q(\hat{x}_i) &= p_{\hat{x}}(\hat{x}_i)
\end{align*}
\]

(8)

When there is good reason to believe that the measurement errors from different sensors are independent \( p_{\hat{\epsilon}} \) factors into the product of \( N_z \) densities, one corresponding to each sensor:

\[
p_{\hat{\epsilon}}(\hat{\epsilon}) = p_{\hat{\epsilon}_1} (\hat{\epsilon}_1) p_{\hat{\epsilon}_2} (\hat{\epsilon}_2) \ldots p_{\hat{\epsilon}_{N_z}} (\hat{\epsilon}_{N_z})
\]

(9)

Each of these individual error densities can be estimated separately from the operational measurement attribute vectors for the corresponding sensor. An estimate of \( p_{\hat{\epsilon}}(\hat{\epsilon}) \) can then be constructed directly from (9), giving the following special case of (8) for independent sensor errors:

\[
\begin{align*}
    w_i &= \frac{\hat{c}}{N_x} \frac{p_{\hat{\epsilon}_1}(\hat{\epsilon}_1 - H_1 \hat{x}_i)p_{\hat{\epsilon}_2}(\hat{\epsilon}_2 - H_2 \hat{x}_i) \ldots p_{\hat{\epsilon}_{N_z}}(\hat{\epsilon}_{N_z} - H_{N_z} \hat{x}_i)}{q(\hat{x}_i)} \\
    q(\hat{x}_i) &= p_{\hat{x}}(\hat{x}_i)
\end{align*}
\]

(10)

where \( p_{\hat{\epsilon}_r}(\hat{\epsilon}_r - H_r \hat{x}_i) \) is the likelihood associated with sensor \( r \).

### 3.4. Data-driven high-dimensional scaling

Sections 3.2 and 3.3 indicate that we can identify the proposal and measurement error terms in the posterior weight expression of (8) if we can derive the attribute vectors \( \hat{x}_i, \hat{\epsilon}_r, \hat{u}_{j,r} \), and \( \hat{v}_j \) for the proposal samples, current measurements, and archived measurements, respectively. All of these images are
available in a pixel-based form, described by the $x_i$, $z_r$, $u_{j,r}$ and $v_j$ vectors. In order to obtain the desired attribute vectors we need a transformation that maps any given image’s pixel vector into a corresponding attribute vector.

Each image can be viewed as a point in an $N_p$-dimensional pixel space, with coordinates equal to the pixel values, or as a corresponding point in a much smaller $N_a$-dimensional attribute space, with coordinates equal to the attribute values. The posterior weight derived for a given image/point in the attribute space also applies to the corresponding image/point in the $N_p \gg N_a$ pixel space. Since there are many pixels but only a few attributes the transformation from pixel to attribute space will generally yield only an approximate representation of the original image. But if the attributes are carefully chosen and the pixel-based description is has significant redundancy the approximation may be adequate for importance sampling purposes.

We formulate the transformation problem by defining composite vectors and that include all of the known pixel and attribute-based vectors defined in Section 3.1:

$$y = [x_1, \ldots, x_{N_x}, z_1, \ldots, z_{N_z}, u_{1,1}, \ldots, u_{N_u,1}, u_{1,2}, \ldots, u_{N_u,2}, u_{1,N_z}, \ldots, u_{N_u,N_z}, v_1, \ldots, v_{N_v}]$$

$$\hat{y} = [\hat{x}_1, \ldots, \hat{x}_{N_x}, \hat{z}_1, \ldots, \hat{z}_{N_z}, \hat{u}_{1,1}, \ldots, \hat{u}_{N_u,1}, \hat{u}_{1,2}, \ldots, \hat{u}_{N_u,2}, \hat{u}_{1,N_z}, \ldots, \hat{u}_{N_u,N_z}, \hat{v}_1, \ldots, \hat{v}_{N_v}]$$

The pixel-based vector $y$ has dimension $N_y = N_I N_p$ while the attribute-based vector $\hat{y}$ has dimension $N_\hat{y} = N_I N_a$, where $N_I = N_x + N_z + N_z N_u + N_u$ is the total number of images to be transformed. We represent the simultaneous transformation of the entire set of $N_I$ images from pixel to attribute space by $T(\cdot)$:

$$\hat{y} = T(y) \quad (11)$$
Our task is to find a transformation that retains as much relevant information as possible when mapping from pixel to attribute values. An approach suggested by (8) indicates that the weight $w_i$ given to the proposal $\hat{x}_i$ depends on the difference between the points $\hat{z}$ and $H\hat{x}_i$ in the attribute space. If the measurement error probability density indicates that large errors are unlikely the proposal will be given less weight when $\hat{z}$ and $H\hat{x}_i$ are further apart. It is, therefore, reasonable to expect the transformation $T(\cdot)$ to preserve, as much as possible, the distances between any two images/points in the pixel space when they are mapped to the attribute space. That is, points which are close (far apart) in the attribute space should also be close (far apart) in the pixel space.

There are various ways to derive distance-preserving transformations from high to low-dimensional vector spaces. All of these rely on quantitative definitions of the distance between pairs of points in a given space. In general, it is both possible and desirable to use different distance (or similarity) measures in the two spaces. We have found that a Euclidean measure is generally satisfactory in the low-dimensional attribute space. Classical distance measures, such as the Minkowski, Mahalanobis or Chebyshev [16] do not seem to work as well in the high-dimensional pixel space. In particular, such measures have trouble detecting differences between categorical images with scattered multiple features. Other distance measures (or, more generally, similarity indices) that focus on image overlap seem better suited for binary, pixel-based images. One example is the Simple Matching Coefficient or SMC [21, 18]. The SMC distance between two pixel-based binary images is:

$$d_{SMC} = 1 - \frac{n_{11} + n_{00}}{n_{01} + n_{10} + n_{11} + n_{00}}$$  (12)
where \( n_{11} \) is the number of pixels that have a value of 1 in both images, \( n_{01} \) is the number of pixels that have value 0 in the first and and 1 in the second image, \( n_{10} \) is the number of pixels that have value 1 in the first and 0 in the second image and \( n_{00} \) is the number of pixels that have value 0 in both images. The value of this distance is 0 if the images coincide and 1 (the largest possible value) if they do not overlap anywhere (i.e. they have no pixels in common). The SMC distance emphasizes differences in the location and size rather than the shape of features found in two images. Two other examples, that are applicable for both binary and non-negative real-valued vectors, are the generalized Jaccard and Tanimoto distances [34, 45, 51]. The generalized Jaccard distance between \( N_p \) element vectors \( X \) and \( Y \) of non-negative pixel values for two images is:

\[
d_{J\text{ac}} = \begin{cases} 
1 - \frac{\sum_{i=1}^{N_p} \min(X_i, Y_i)}{\sum_{i=1}^{N_p} \max(X_i, Y_i)} & \text{if } \sum_{i=1}^{N_p} \max(X_i, Y_i) > 0 \\
0 & \text{if } \sum_{i=1}^{N_p} \max(X_i, Y_i) = 0
\end{cases} 
\]  

(13)

The corresponding Tanimoto distance (which is not a metric because it does not satisfy the triangle inequality) is:

\[
d_{T\text{an}} = \begin{cases} 
\frac{X \cdot Y}{X \cdot X + Y \cdot Y - X \cdot Y} & \text{if } X \cdot X + Y \cdot Y - X \cdot Y > 0 \\
0 & \text{if } X \cdot X + Y \cdot Y - X \cdot Y = 0
\end{cases} 
\]  

(14)

where \( X \cdot Y \) represents the dot product between vectors \( X \) and \( Y \). These two distances have the same value when the images are binary, as they are in the example of Section 4. Although no distance measure is suitable for every application, we obtain good results in our rainfall example by using a SMC in the pixel space and a Euclidean distance in the attribute space.
Non-metric multi-dimensional scaling methods (see e.g. [37, 43]) can be used to derive approximate distance preserving transformations between pairs of points \((s, t)\) in high and low-dimensional spaces. These methods are usually designed to find the complete set of image attribute vectors \(\hat{y}\) that minimize a "stress function" of the following general form:

\[
\zeta = \sum_{s < t} F \left[ |d(y_s, y_t) - \hat{d}(\hat{y}_s, \hat{y}_t)| \right] \quad s, t = 1, \ldots, N
\]  

(15)

where, for our application, \(d(y_s, y_t)\) is the Generalized Jaccard distance between points \(y_s\) and \(y_t\) in the pixel space, \(\hat{d}(\hat{y}_s, \hat{y}_t)\) is the Euclidean distance between the corresponding points \(\hat{y}_s\) and \(\hat{y}_t\) in the attribute space. The function \(F(\cdot)\) is a specified stress measure. Some classical examples are Sammon’s mapping [54, 42, 41] and recent variants [5, 52, 4, 13]. Here we use a stress function proposed in [43] that is designed to work well when the high dimension space is very large. This function, referred to as data-driven high-dimensional scaling (DD-HDS), is:

\[
\zeta_{\text{DD-HDS}} = \sum_{s < t} \left[ |d^{st}(Y_s, Y_t) - \hat{d}^{st}(\hat{Y}_s, \hat{Y}_t)| \left( 1 - \int_{-\infty}^{\min(d^{st}(Y_s, Y_t), \hat{d}^{st}(\hat{Y}_s, \hat{Y}_t))} f(u, \mu, \sigma) \, du \right) \right]
\]  

(16)

where \(f(u, \mu, \sigma)\) is the probability density function of a Gaussian random variable \(u\) with mean \(\mu\) and standard deviation \(\sigma\). It is suggested in [43] to choose these parameters in a way that they adapt to the effective distribution of distances \(d_{s,t}^{st}\) for \(s, t\) in the original, high-dimensional space. One option is to use:

\[
\mu = \text{mean}_{1 \leq s < t \leq N} (d^{st}(Y_s, Y_t)) - 2(1 - \lambda) \text{std}_{1 \leq s < t \leq N} (d^{st}(Y_s, Y_t))
\]  

(17)

\[
\sigma = 2\lambda \text{std}_{1 \leq s < t \leq N} (d^{st}(Y_s, Y_t))
\]  

(18)
where the mean and standard deviation (std) are taken over the distribution of distances between all pairs of data in the original space and $\lambda$ is a positive user-defined parameter (usually to be taken between 0.1 and 0.9) that controls the relative emphasis given to large vs. small distances in high-dimensional space.

In general, the stress must be minimized with a numerical optimization algorithm that gives the optimum set of image points $\hat{y}$ in the attribute space. Here we use the original implementation of DD-HDS given in [43], which is based on an algorithm described in [24], where it is called the “Force Directed Placement” paradigm [48].

4. Rainfall example

4.1. Experimental design

We illustrate the key aspects of our feature-based importance sampling procedure with an example that uses noisy satellite measurements to characterize rainfall patterns. For this example we consider binary images that have only two intensity levels (rain or no-rain), primarily to focus on spatial intermittency. Irregular intermittent features are commonly encountered in environmental applications but are difficult to describe with traditional continuous random field models. When intensities are categorized into a finite number of levels (e.g. two) all of the probabilities used in the importance sampling procedure are discrete and there is a finite (but very large) number of possible true images. This simplifies both the computation and visualization of results. However, it should be noted that the methods described in the preceding sections are sufficiently general to deal with continuously vari-
able intensities or with intermittent features that have variable intensities within disconnected areas. We intend to examine these more general cases in the future.

The hydrologic and meteorological communities have devoted considerable attention to the characterization and quantitative forecasting of uncertain rainfall events. Recent advances in remote sensing of rainfall have provided an opportunity to merge data from multiple sensors with different error properties. Examples of multi-sensor rainfall characterization algorithms include [23, 60, 39, 29, 9, 63], and [19]. Algorithms that have been run operationally to produce global rainfall estimates include [64, 31, 38, 40, 30], and [36]. Most of these efforts provide deterministic products and do not attempt to quantify estimation errors or to rank different possible outcomes. Exploratory studies that explicitly account for uncertainty in satellite-based rainfall estimates include [11, 27, 2, 10]. These studies typically rely on parametric measurement error models that may not adequately describe the complex errors that can affect the geometry as well as the intensity of noisy spatial measurements. The approach taken in our example and summarized in the preceding sections provides probabilistic products based on non-parametric measurement error models.

It is possible to check the performance of a feature characterization procedure either by processing actual measurements or by conducting a virtual experiment where a known "true feature" is used to derive synthetic measurements. In the latter case posterior realizations can be compared directly to the known truth. When evaluating new algorithms it is often advantageous to first conduct a controlled virtual experiment to test the concept and
then to assess practical feasibility with real measurements. In this paper we
describe the virtual component of a longer-term research program concerned
with rainfall characterization.

In order to make the virtual experiment realistic we derive true features,
proposals, and synthetic measurements from images in the NEXt Generation
RADdar (NEXRAD-IV) dataset. NEXRAD measurements, which are avail-
able for most of the continental US, are collected by the National Weather
Service ground-based WSR-88D radar network [25] and processed by the
National Centers for Environmental Prediction (NCEP). NEXRAD rainfall
intensity is derived from radar backscatter with the procedure described by
[8]. The NEXRAD data used here are arranged on a latitude-longitude grid
of 0.25° by 0.25° pixels covering 31.625° to 47.625° N and 108.625° to 80.625°
W at 12 hour intervals over the period Jul. 2003 through Dec. 2010. The
set of archived ground truth images (pixel vectors \(v_1, v_2, \ldots, v_{N_u}\)) consists
of 2773 images from different 16 by 16 pixel (4° by 4°) sections of the large
NEXRAD grid. We designate one of the ground truth images to be the true
image (pixel vector \(x\)). This true image is set aside and no longer included
in the archived ground truth set.

Although NEXRAD is generally believed to provide high quality measure-
ments of rainfall it is not available globally. Global rainfall intensities must
be inferred from low orbit satellite sensors such as AMSU [62] and SSMI [20]
which typically detect rainfall through observations of microwave brightness
or cloud top temperature. Low orbit microwave observations are generally
less accurate than NEXRAD and are available only at intermittent times.
The synthetic operational measurements used in our virtual experiment are
intended to imitate the general characteristics of these noisy observations.

The current operational measurement for each of the $N_z$ sensors is a single 16 by 16 pixel random SNESIM realization with the common true image used as a training image. The realization’s deviation from the true image represents measurement error. We consider two options with either $N_z = 1$ current measurement from Sensor 1 or $N_z = 2$ current measurements from Sensors 1 and 2. The Sensor 1 measurement is constrained by a randomly selected set of 25% of the training image pixel intensities while the somewhat more accurate Sensor 2 measurement is constrained by 20% of the training image pixel values.

We use the techniques of Section 3.2 to generate $N_x = 240$ pixel-based proposals composed of three sets of 80 proposals constrained by, respectively, 1%, 5%, and 25% of the image pixels (see Figure 3). When $N_z = 1$ all 240 proposals are constrained by values from the current Sensor 1 measurement. When $N_z = 2$ the first 120 proposals are constrained by the Sensor 1 measurement and the remaining 120 proposals are constrained by the Sensor 2 measurement.

In our example we obtain a set of $N_u = 300$ paired archived operational and ground truth synthetic measurements using a procedure similar to the one used to generate the current operational measurement. Each archived operational measurement is a SNESIM realization that uses a particular ground truth image as a training image. The realization is constrained by intensity values at 25% or 20% of the training image pixels for measurements from Sensors 1 and 2, respectively. When we use only Sensor 1 all of the archived operational measurements are assumed to be from this sensor. When we use
both Sensors 1 and 2 half of the archived measurements are from each sensor.

The archived operational measurement and the ground truth measurement used as its training image form a distinct pair. The differences between the mapped measurement attribute values define an additive attribute error for the pair, as described in Section 3.3 and shown in Figure 4. The probability \( p_{\hat{E}}(\hat{e}) \) is estimated by fitting a Gaussian mixture to errors derived from archival operational measurements that are similar to the current operational measurement, as measured by the Jaccard distance in the pixel space. When mapping image pixel values to attribute values we specify only \( N_a = 2 \) attributes so we can plot non-parametric density fits to the proposal density and likelihood functions as contours in two dimensions. These design choices yield a set of 841 images to be mapped from the 256-dimensional pixel space to the 2-dimensional attribute space. The mapping is performed with the DD-HDS technique described in Section 3.4.

4.2. Results

The results of the image mapping for the single current measurement case are shown in Fig. 5a, which distinguishes archived ground truth measurements (blue), archived operational measurements (green), a single current operational measurement (black), and the proposals (red). The attributes that form the axes for this plot are data-dependent and do not have any particular geometric interpretation. Their primary role is to provide a concise image description that keeps images that are close (far apart) in the pixel space close (far apart) in the attribute space. It is apparent that the proposals tend to lie in the upper half of the diagram, with a subset relatively close to the current operational measurement (these are the proposals that
are most constrained by the current measurement). The archived ground truth and operational measurements tend to be centered around the middle of the diagram, suggesting that they correspond to a diverse set of historical true images.

Figure 5b shows contours of a Gaussian mixture approximation for the proposal probability density $q(\hat{x})$, fit to the red proposal samples. The highest probability region is the dense cluster near the current measurement. Figure 5c shows contours of a Gaussian mixture approximation for the measurement error probability density $p_{\hat{E}}(\hat{\epsilon})$ fit to the blue differences $\hat{v}_j - \hat{v}_j$ between archived operational and ground truth measurements. The Gaussian mixture estimation algorithm is described in [22]. The likelihood value for a particular proposal $\hat{x}_i$ is obtained by evaluating (7) at the black dots, which represent the differences $\hat{z} - \hat{x}_i$ between the single current operational measurement and the proposal. The posterior weight $w_i$ for each proposal is obtained from (8), with the relatively uninformative prior density $p_{\hat{X}}(\hat{x}_i)$ assumed to be uniform throughout the attribute plane, covering the range -1.5 to 1.5 for each attribute.

Figure 6 shows the spectrum of proposal weights, ordered from largest to smallest, for this example. Note that a relatively small number of samples have higher weights but no single sample dominates. Figures 7c and 7d show, respectively, the 15 most probable and 15 least probable proposals in the complete set of 240. The true image and current measurement image are shown for reference in Figs. 7a and 7b, respectively. It is apparent that the higher weighted proposals in Fig. 7c look more like the true image. The variation among these proposals give a visual indication of the uncertainty
in the true image. This uncertainty reflects both the variability built into the proposal population and the measurement error assessment inferred from archived measurements.

The Demartine’s diagram [43, 15] shown in Fig. 8 provides a way to examine the quality of the multi-dimensional scaling transformation from pixel to attribute spaces. This diagram plots the attribute space distance vs the pixel space distance for all possible pairs of mapped images, indexed by (s, t). If the DD-HDS transformation preserved distances between pairs perfectly, the points in Fig. 8 would fall on a narrow curve, with larger pixel space distances always giving larger attribute space distances. In fact, the points in the figure form a broad cloud that trends generally from the lower left (where both distances are small) to the upper right (where both distances are large). Since there are many pairs the cloud is shaded to indicate the density of points in a given area (darker colored areas have more points). The plot indicates that there are pairs where the distances in the two spaces are not compatible, although most pairs fall along the center of the cloud. A better transformation would be characterized by a narrower cloud with a dark color. It is possible to obtain better results by using fewer images or more attributes since it becomes more difficult to preserve distances when there are more pairs and the attribute space is small.

In the context of our application, [46, 47] and [33] show that when $N_I$ images in a given pixel space with any specified metric are mapped into a smaller $N_a$-dimensional Euclidean space the multiplicative distortion in the set of mapped points is $O(\min[N_I^{2} \log^{2} N_I, N_I])$ where distortion is measured in terms of a stress functions or performance criterion such as (16). So
the mapping error increases as \( N_I \) increases for fixed \( N_a \) and as \( N_a \) decreases for fixed \( N_I \). In addition, the optimization algorithms used to minimize the multi-dimensional scaling stress function such as (16) become impractical, or at least very expensive, for data sets containing more than a few thousand points since the stress computation effort increases quadratically with \( N_I \).

These factors encourage the use of small proposal ensembles and archived measurement data sets and large numbers of attributes. But the importance sampling process becomes less able to properly assign posterior weights when the proposal size decreases and the non-parametric likelihood derivation become less reliable when the number of archived data points decreases. Also, both importance sampling and likelihood derivation need more images to maintain good performance when the number of attributes is increased. So there is a clear tension between the need to obtain a more accurate and efficient multi-dimensional scaling transformation (fewer images, more attributes) and a more accurate importance sampling procedure (more images, fewer attributes).

Figure 9 shows the likelihoods obtained for two different current measurements, refered to as the measurement 1 and measurement 2, generated with different constraint procedures from the same true image (25% constraining pixels for measurement 1 and 20% constraining pixels for measurement 2). Note the distinctly different shapes of the individual likelihoods, reflecting the different ways the image attributes are affected by errors in the two measurements. When the product of these likelihoods is inserted into (9) we obtain the results shown in Fig. 10. Figure 10a shows the single true image (the same as in Fig. 7a) while Figs. 10b and 10c show the two current
measurements used to characterize this image. The most probable and least probable proposal sets look broadly similar to those obtained in Fig. 7 with one measurement but they are not identical. Note that some of the proposals from Fig. 7 have reappeared in the most probable set, with different ranks. The two-measurement result presented here is an example of image fusion, the combination of different measurements and archival information to obtain an improved characterization of hidden features.

5. Conclusions and future research directions

Importance sampling provides a useful framework for interpreting and merging observations of complex geometrical features. It accommodates general models of image structure and measurement error and is relatively easy to implement. However, importance sampling needs adequate ensemble sizes, accurate measurement error models, and realistic feature proposals to be useful in practical applications. When images are described with high-dimensional vectors of pixel values a very large ensemble of proposed images is needed to provide a representative sample of all possible combinations. In such cases importance sampling is not practical. The image characterization approach proposed here describes images with low-dimensional image attribute vectors that are derived by applying a multi-dimensional scaling transformation to pixel-based images. The transformation is not specified in advance but inferred from a large set of proposals and archived measurements. Importance sampling in a properly constructed attribute space is feasible with a moderate number of proposal samples.

In order to properly represent measurement errors our approach relies on
non-parametric probability densities estimated from an image archive rather than hypothesized parametric densities. A non-parametric approach is better able to deal with unconventional measurement errors such as misregistered images or artifacts near feature boundaries. This increases the applicability of the method and insures that the posterior probabilities properly reflect the effects of measurement uncertainty. Moreover, it enables the image characterization procedure to make best use of information from different sensors with different strengths and weaknesses.

We also use a non-parametric approach to generate proposals. These proposals are constructed with a multipoint statistical algorithm that relies on training images to convey information about distinctive feature properties. The proposals are constrained, to various degrees, by current measurements to insure that they are able to capture non-stationarities. Since proposals generated in this way are realistic they yield better results than proposals drawn from hypothesized parametric densities.

The primary limitation of an attribute-based approach to importance sampling is the inevitable loss of information that occurs when high-dimensional images are described with much lower-dimensional attribute vectors. In particular, as the number of proposal and archive samples increases it becomes harder to preserve in the attribute space distances between the mapped images. This reduces the performance benefits that would otherwise be achieved by increasing the sample size. An importance sampling procedure that relies on multi-dimensional scaling must compromise between the need to keep the sample size small for accurate mapping of images and the need to keep the sample size large for accurate representation of attribute probabilities.
Future research in this area should build on the inherent advantages of the importance sampling approach, especially flexibility and generality, while addressing its limitations. It would be useful to examine alternative dimensionality reduction options that perform better than multi-dimensional scaling as sample sizes are increased. These alternatives will likely need to go beyond distance preservation when selecting image attributes, while remaining data-driven. Other topics worth further effort include an assessment of alternative similarity measures in the pixel and attribute spaces and an investigation of the benefits to be obtained by increasing the number of attributes.

The results provided in our rainfall example indicate that attribute-based importance sampling is a promising option for characterizing complex categorical images. Ultimately, one of the most significant aspects of the importance sampling approach is the perspective it provides, by presenting a range of possible image alternatives that gives an accessible and informative picture of uncertainty.

References


[38] T. Kubota, S. Shige, H. Hashizume, K. Aonashi, N. Takahashi, S. Seto, Y.N. Takayabu, T. Ushio, K. Nakagawa, K. Iwanami, M. Kachi,


List of Figures

1 Information flow for feature characterization with importance sampling .................................................. 42
2 Unconditional and conditional realizations generation procedure: (a) Training image; (b) Current measurement and 1% random constraining pixels (black dots); (c) Unconditional realizations simulated using the training image in (a); (d) conditional realizations simulated using the training image in (a) and random constraining pixels in (b). ........................................... 43
3 (a) True image. (b) Current operational measurement. (c) Examples of conditional proposal realizations $x_i$ (12 realizations for each fraction of constraining pixels). The fractions of constraining pixels are displayed in bold in the left panel. Each image in (a),(b) and (c) is 16x16 pixels. ................................................. 44
4 (a) Examples of NEXRAD-IV rainfall signatures used as ground-truth images in this study. Red indicates rain and blue no rain. The dimension of each image is 16x16 pixels. (b) Examples of synthetic historical measurements generated using conditional simulation described in Section 3.2. Each synthetic measurement corresponds to the ground-truth image in (a). So for example the image in upper left panel corresponds to the image in upper left panel in (a). (c) the difference images obtained by subtracting each ground truth image in (a) from its operational counterpart in (b). Blue color represents areas with no error, green represents false alarms (measurement is rainy but not the ground-truth) and yellow represents misses (ground-truth is rainy but not the measurement). The black line shows the boundary of the ground-truth rain. ................................................. 45
5 (a) 2-D attribute space obtained by DD-HDS mapping of 256-D rain support data set. (b) Estimate of proposal pdf (black contour lines); red dots represent proposal sample mapped into the attribute space; prior is assumed to be uniform. (c) Estimate of the likelihood (black contour lines), archived operational measurement errors (blue dots) and current operational measurement errors (black dots). ..................................................... 46
6 Posterior weights sorted in descending order .......................................................... 47
(a) True image. (b) Current operational measurement. (c) 20 proposal realizations with the highest posterior weights. (d) 20 proposal realizations with the smallest posterior weights.

(dy – dx) diagram for the DD-HDS mapping in Fig. 5(a). Heat color scale represents the number of images pairs corresponding to SMC and Euclidean distances indicated on the x and y axes, respectively. Images that are close in the pixel space generally map to points that are close in the attribute space.

(a) Estimate of the likelihood for measurement 1 in Fig.10b (black contour lines). Estimate of the likelihood for measurement 2 in Fig.10c (black contour lines). Archived operational measurement errors are represented by blue dots and current operational measurement errors are represented as black dots.

(a) True image. (b) Current operational measurement 1. (c) Current operational measurement 2. (d) 20 proposal realizations with the highest posterior weights. (e) 20 proposal realizations with the smallest posterior weights.
Figure 1: Information flow for feature characterization with importance sampling
Figure 2: Unconditional and conditional realizations generation procedure: (a) Training image; (b) Current measurement and 1% random constraining pixels (black dots); (c) Unconditional realizations simulated using the training image in (a); (d) conditional realizations simulated using the training image in (a) and random constraining pixels in (b).
Figure 3: (a) True image. (b) Current operational measurement. (c) Examples of conditional proposal realizations $x_i$ (12 realizations for each fraction of constraining pixels). The fractions of constraining pixels are displayed in bold in the left panel. Each image in (a), (b) and (c) is 16x16 pixels.
Figure 4: (a) Examples of NEXRAD-IV rainfall signatures used as ground-truth images in this study. Red indicates rain and blue no rain. The dimension of each image is 16x16 pixels. (b) Examples of synthetic historical measurements generated using conditional simulation described in Section 3.2. Each synthetic measurement corresponds to the ground-truth image in (a). So for example the image in upper left panel corresponds to the image in upper left panel in (a). (c) the difference images obtained by subtracting each ground truth image in (a) from its operational counterpart in (b). Blue color represents areas with no error, green represents false alarms (measurement is rainy but not the ground-truth) and yellow represents misses (ground-truth is rainy but not the measurement). The black line shows the boundary of the ground-truth rain.
Figure 5: (a) 2-D attribute space obtained by DD-HDS mapping of 256-D rain support data set. (b) Estimate of proposal pdf (black contour lines); red dots represent proposal sample mapped into the attribute space; prior is assumed to be uniform. (c) Estimate of the likelihood (black contour lines), archived operational measurement errors (blue dots) and current operational measurement errors (black dots).
Figure 6: Posterior weights sorted in descending order
Figure 7: (a) True image. (b) Current operational measurement. (c) 20 proposal realizations with the highest posterior weights. (d) 20 proposal realizations with the smallest posterior weights.
Figure 8: \( dy - dx \) diagram for the DD-HDS mapping in Fig. 5(a). Heat color scale represents the number of image pairs corresponding to SMC and Euclidean distances indicated on the x and y axes, respectively. Images that are close in the pixel space generally map to points that are close in the attribute space.
Figure 9: (a) Estimate of the likelihood for measurement 1 in Fig.10b (black contour lines). Estimate of the likelihood for measurement 2 in Fig.10c (black contour lines). Archived operational measurement errors are represented by blue dots and current operational measurement errors are represented as black dots.
Figure 10: (a) True image. (b) Current operational measurement 1. (c) Current operational measurement 2. (d) 20 proposal realizations with the highest posterior weights. (e) 20 proposal realizations with the smallest posterior weights.
• We considered uncertain spatial features inferred from noisy measurements
• Bayesian state filtering problem was solved by importance sampling
• Feature descriptors were derived using non-linear multidimensional scaling
• Proposal ensembles were simulated by a multi-point statistical algorithm
• We estimated rainy areas on the Earth’s surface using remote sensing measurements