Surv Geophys (2014) 35:765–783
DOI 10.1007/s10712-013-9264-9

Downscaling Satellite Precipitation with Emphasis on Extremes: A Variational ℓ_1-Norm Regularization in the Derivative Domain

E. Foufoula-Georgiou · A. M. Ebtehaj · S. Q. Zhang · A. Y. Hou

Received: 1 April 2013 / Accepted: 5 November 2013 / Published online: 11 December 2013

Abstract The increasing availability of precipitation observations from space, e.g., from the Tropical Rainfall Measuring Mission (TRMM) and the forthcoming Global Precipitation Measuring (GPM) Mission, has fueled renewed interest in developing frameworks for downscaling and multi-sensor data fusion that can handle large data sets in computationally efficient ways while optimally reproducing desired properties of the underlying rainfall fields. Of special interest is the reproduction of extreme precipitation intensities and gradients, as these are directly relevant to hazard prediction. In this paper, we present a new formalism for downscaling satellite precipitation observations, which explicitly allows for the preservation of some key geometrical and statistical properties of spatial precipitation. These include sharp intensity gradients (due to high-intensity regions embedded within lower-intensity areas), coherent spatial structures (due to regions of slowly varying rainfall), and thicker-than-Gaussian tails of precipitation gradients and intensities. Specifically, we pose the downscaling problem as a discrete inverse problem and solve it via a regularized variational approach (variational downscaling) where the regularization term is selected to impose the desired smoothness in the solution while allowing for some steep gradients (called ℓ_1-norm or total variation regularization). We demonstrate the duality between this geometrically inspired solution and its Bayesian statistical interpretation, which is

E. Foufoula-Georgiou (✉)
Saint Anthony Falls Laboratory, Department of Civil Engineering, University of Minnesota, Minneapolis, MN, USA
e-mail: efi@umn.edu

A. M. Ebtehaj
Saint Anthony Falls Laboratory, Department of Civil Engineering, School of Mathematics, University of Minnesota, Minneapolis, MN, USA
e-mail: ebteh001@umn.edu

S. Q. Zhang · A. Y. Hou
NASA Goddard Space Flight Center, Greenbelt, MA, USA
e-mail: sara.q.zhang@nasa.gov

A. Y. Hou
e-mail: arthur.y.hou@nasa.gov

equivalent to assuming a Laplace prior distribution for the precipitation intensities in the derivative (wavelet) space. When the observation operator is not known, we discuss the effect of its misspecification and explore a previously proposed dictionary-based sparse inverse downscaling methodology to indirectly learn the observation operator from a data base of coincidental high- and low-resolution observations. The proposed method and ideas are illustrated in case studies featuring the downscaling of a hurricane precipitation field.

Keywords Sparsity · Inverse problems · ℓ_1-norm regularization · Non-smooth convex optimization · Generalized Gaussian density · Extremes · Hurricanes

1 Introduction

Precipitation is one of the key components of the water cycle and, as such, it has been the subject of intense research in the atmospheric and hydrologic sciences over the past decades. While it still remains the most difficult variable to accurately predict in numerical weather and climate models, its statistical space–time structure at multiple scales has been extensively studied using several approaches (e.g., Lovejoy and Mandelbrot 1985; Lovejoy and Schertzer 1990; Kumar and Foufoula-Georgiou 1993a, b; Deidda 2000; Harris et al. 2001; Venugopal et al. 2006a, b; Badas et al. 2006). These studies have documented a considerable variability spread over a large range of space and timescales and an organization that manifests itself in power law spectra and more complex self-similar structures expressed via nonlinear scaling of higher-order statistical moments (e.g., Lovejoy and Schertzer 1990; Venugopal et al. 2006a). Stochastic models of multi-scale rainfall variability have been proposed based on inverse wavelet transforms (Perica and Foufoula-Georgiou 1996), multiplicative cascades (Deidda 2000), exponential Langevin-type models (Sapozhnikov and Foufoula-Georgiou 2007), among others.

The small-scale variability of precipitation (of the order of a few kms in space and a few minutes in time) is known to have important implications for accurate prediction of hydrologic extremes especially over small basins (e.g., Rebora et al. 2006a, b) and for the prediction of the evolving larger-scale spatial organization of land–atmosphere fluxes in coupled models (Nykanen et al. 2001). This small-scale precipitation variability, however, is not typically available in many regions of the world where coverage with high-resolution ground radars is absent or in mountainous regions where spatial gaps are present due to radar blockage. It is also missing from climate model predictions that are typically run at low resolution over larger areas of the world. As a result, methods for downscaling precipitation to enhance the resolution of incomplete or low-resolution observations from space or numerical weather/climate model outputs continue to present a challenge of both theoretical and practical interests.

To date, multiple passive and active ground-based (i.e., gauges and radars) and spaceborne sensors (i.e., geostationary, polar and quasi-equatorial orbiting satellites) exist that overlappingly measure precipitation with different space–time resolutions and accuracies. Sparsely populated networks of rain gauges provide relatively accurate point measurements of precipitation continuously over time, while ground-based radars detect precipitation in fine enough spatiotemporal scales (e.g., $\sim$6 min at 1 $\times$ 1 km) but over limited areal extents. The ground-based radar data are among the most accurate and high-resolution estimates of spatial rainfall. However, this source of information is subject to

various shortcomings such as instrumental errors, beam blockage by orographic features, and overshooting range effects (Krajewski and Smith 2002). The only civilian active spaceborne Tropical Rainfall Measuring Mission-Precipitation Radar (PR) sensor (TRMM-PR) provides high-resolution reflectivity of rainfall fields (i.e., $\sim 4 \times 4$ km) over a narrow band in the tropics with relatively low temporal revisiting frequency compared to the other passive spaceborne sensors of lower resolution. The forthcoming Global Precipitation Measuring (GPM) Mission, a constellation of nine satellites, promises to deliver observations of high precision precipitation and cloud dynamics at a global scale (3-h revisiting time) and over varying resolutions and create opportunities for improving climate modeling and hazard prediction at local scales (Flaming 2004).

Precipitation observations from space are especially valuable in regions where no ground observations are available either from rain gauges or from ground radars, such as over the oceans or in underdeveloped regions of the world. It is over these regions, however, that some extreme tropical storms develop for which high-resolution information would provide important means for hazard prediction and warning as well as detailed information on extremes, which could be used in nested models or in a data assimilation setting. These tropical storms have distinct geometrical and statistical structures, as shown below, posing extra demands on the methodologies of precipitation downscaling, data fusion, and data assimilation.

As an illustrative example, Fig. 1 shows a snapshot of the two-dimensional rainfall intensity patterns and the three-dimensional structure of precipitating clouds for typhoon Neoguri, the first typhoon of the 2008 season in the western Pacific Ocean, on April 17, 2008, as observed by the TRMM-PR and the TRMM Microwave Imager (TMI). One notices the geometrically structured precipitation bands embedded within the larger two-dimensional storm system and the localized "towers" of high-intensity rainfall spatially embedded within lower-intensity rainfall background. These localized high-intensity cells and the steep sporadic gradients of precipitation intensity in such a storm are more clearly demonstrated via a one-dimensional cross section as shown in Fig. 2. Specifically, Fig. 2b

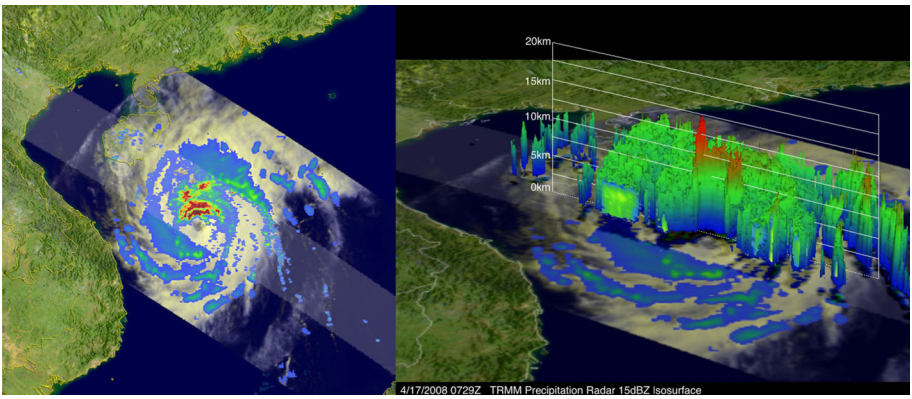

Fig. 1 *Left panel* rainfall pattern of typhoon Neoguri in the western Pacific Ocean, on April 17, 2008. The *dark red* bands indicate regions of the most intense rain. Rainfall rates in the inner swath are from TRMM's-PR, while in the outer swath from the TRMM Microwave Imager (TMI); *Right panel* the three-dimensional structure of precipitating clouds for typhoon Neoguri as observed by the TRMM-PR. This figure illustrates the need for a downscaling scheme that has the ability to reproduce steep rainfall gradients embedded within the storm. *Source*: NASA's Earth Observatory, available online through the TRMM extreme event image archives (http://trmm.gsfc.nasa.gov)

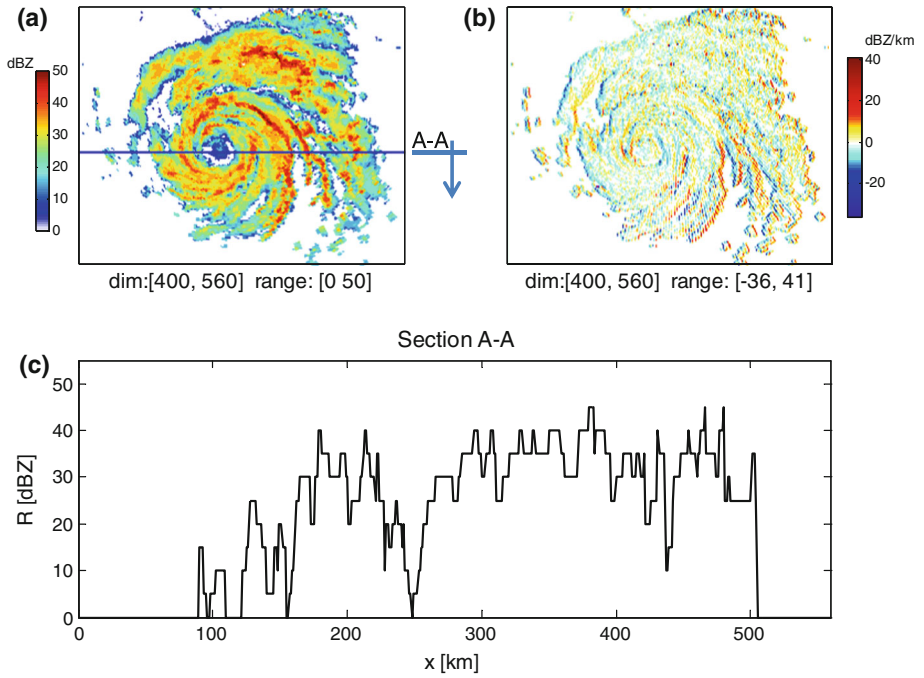

Fig. 2 **a** A high-resolution (HR) snapshot of hurricane Claudette, 07-15-2003, 11:51:00 UTC as monitored by NEXRAD station over Texas at resolution 1 × 1 km and **b** the field of the computed horizontal first-order derivative using the Sobel filter. A horizontal cross section through the storm is shown in (**c**). One observes how the particular geometrical structure of hurricane precipitation projects itself onto an almost piece-wise linear one-dimensional function with sporadic large gradients embedded within regions of almost constant rainfall

demonstrates how the typical circular bands of high rainfall intensity manifest themselves into an almost piece-wise linear structure in the 1D cross section. How is this geometrical structure to be reproduced in downscaling lower resolution and noisy observations of tropical storms, say available at 10-km resolution, down to 1- or 2-km resolution products?

Moving from a geometrical description to a statistical description, we note that coherent precipitation intensity areas (similar intensity in nearby pixels) will result in almost zero values in a derivative space, while the abrupt changes in rainfall intensity (large gradients and discontinuities) will project as high values. In other words, we expect to see a probability distribution in the derivative space that has a large mass close to zero and a few large positive and negative values. Figure 3a shows the histogram of the derivatives of precipitation intensities of hurricane Claudette in the horizontal (zonal) direction (computed via a redundant orthogonal Haar wavelet transform, which is equivalent to using a first-order difference discrete approximation). It is obvious that this histogram is considerably different than a Gaussian probability distribution function (PDF) with a larger mass around zero (capturing the large number of nearby pixels with similar intensity) and much heavier tails than Gaussian (capturing the occasional very steep gradients). How can such a statistical structure be explicitly incorporated in a precipitation downscaling scheme, specifically for hurricanes and tropical storms?

The purpose of this paper is to present a new framework for precipitation downscaling casting the problem as a discrete inverse problem and solving it via a variational

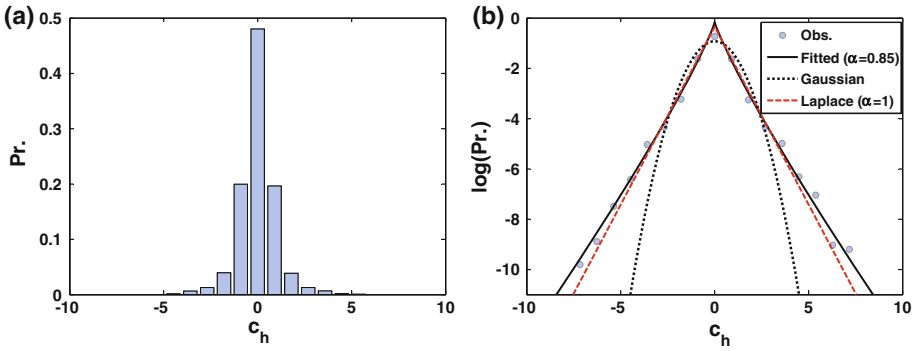

Fig. 3 **a** Histogram of the derivatives in the *horizontal direction* of the hurricane snapshot shown in Fig. 2. The derivative coefficients are obtained by the Sobel operator that produces a second-order discrete approximation of the field derivative. **b** Same histogram plotted on a log-probability scale showing the empirical PDF (*circles*), the fitted generalized Gaussian PDF with parameter $\alpha = 0.85$, the Gaussian PDF ($\alpha = 2.0$), and the Laplace density ($\alpha = 1.0$) for comparison. Note that the assumption of a Laplace density for the rainfall derivatives is theoretically consistent with the proposed ℓ_1-norm variational downscaling framework

regularization approach, which imposes constraints on the specific degree of smoothness (regularity) of the precipitation fields. The proposed regularization is selected to allow the preservation of large gradients while at the same time impose the desired smoothness on the solution. The paper is structured as follows. In Sect. 2, the need for regularization is explained with special emphasis on a total variation regularization scheme (ℓ_1-norm in the derivative space) in order to reproduce steep gradients and to preserve the heavy-tailed structure of rainfall. In this Section, the statistical interpretation of the variational ℓ_1-norm regularization is also explained. In particular, it is elucidated that the downscaled rainfall fields obtained via ℓ_1-norm regularization in the derivative domain is equivalent to the Bayesian maximum a posteriori (MAP) estimate with a Laplace prior distribution in the precipitation derivatives, a special case of the generalized Gaussian distribution $p(x) \propto \exp(-\lambda|x|^{\alpha})$ with $\alpha = 1$ (Ebtehaj and Foufoula-Georgiou 2011). Section 3 presents insights into the problem of an unknown downgrading observation operator or kernel that "converts" the high-resolution rainfall to the lower-resolution observations and discusses an alternative methodology, dictionary-based sparse precipitation downscaling (SPaD), developed in (Ebtehaj et al. 2012). In Sect. 4, we present a detailed implementation of our variational downscaling (VarD) methodology in a tropical (hurricane) storm and compare the results of VarD with those of the SPaD method. Finally, concluding remarks and directions for future research are presented in Sect. 5.

2 Precipitation Downscaling as a Regularized Inverse Problem

2.1 Basic Concepts in the Continuous Space

Consider the true state (or signal) $f(t)$ that is not known but is observed indirectly via a measuring device, which imposes a smoothing on the original state and returns the observation $g(s)$. Let $f(t)$ and $g(s)$ relate via the following linear transformation:

$$\int_0^1 K(s,t)f(t)\mathrm{d}t = g(s)\,0,\ t \le 1, \tag{1}$$

where $K(s, t)$ is a known kernel, which downgrades the true state by damping its high-resolution components and making it smoother. The problem of recovering $f(t)$ knowing the observation $g(s)$ and the kernel $K(s, t)$ is a well-studied inverse problem, known as the Fredholm integral equation of the first kind. Inverse problems are by their nature ill-posed, in the sense that they do not satisfy at least one of the following three conditions: (1) existence of a solution, (2) uniqueness of the solution, and (3) stability in the solution, i.e., robustness to perturbations in the observation. It can be shown that the above inverse problem is very sensitive to the observation noise, since high frequencies are amplified in the inversion process (so-called inverse noise) and they can easily spoil and blow up the solution (see Hansen 2010). In this sense, even a small but high-frequency random perturbation in $g(s)$ can lead to a very large perturbation in the estimate of $f(t)$. This is relevant to the problem of reconstructing small-scale features in precipitation fields (downscaling) from low-resolution noisy data, when the noise can be of low magnitude but high frequency, e.g., discontinuities in overlapping regions of different sensors or instrument noise.

Therefore, naturally, if we define the distance between the observations and the true state by the following residual Euclidean norm:

$$R(f) = \left\| \int_0^1 K(s,t)\, f(t)\ \mathrm{d}t - g(s) \right\|_2, \tag{2}$$

then minimizing $R(f)$ alone does not guarantee a unique and stable solution of the inverse problem. Rather, additional constraints have to be imposed to enforce some regularity (or smoothness) of the solution and suppress some of the unwanted inverse noise components leading to a unique and more stable solution. Let us denote by $S(f)$ a smoothing norm, which measures the desired regularity of $f(t)$. Then, obtaining a unique and stable solution to the inverse problem amounts to solving a variational minimization problem of the form

$$f(t) = \underset{f}{\operatorname{argmin}} \left\{ R(f)^2 + \lambda^2 S(f) \right\}, \tag{3}$$

The value of λ (called the regularization parameter) is chosen as to provide a balance between the weight given to fitting the observations, as measured by the magnitude of the residual term $R(f)$, and the degree of regularity of the solution measured by the smoothing norm $S(f)$. Common choices for $S(f)$ are ℓ_2-norms of the function $f(t)$ or its derivatives, i.e.,

$$S(f) = \left\| f^{(d)} \right\|_2^2 = \int_0^1 \left| f^{(d)}(t) \right|^2 \mathrm{d}t,\ d = 0, 1, \ldots \tag{4}$$

where $f^{(d)}$ denotes the dth order derivative of f. Another smoothing norm of specific interest in the present study is the ℓ_1-norm of the gradient of f, that is,

$$S_{\mathrm{TV}}(f) = \|f\|_1 = \int_0^1 \left| f^{(1)}(t) \right| \mathrm{d}t, \tag{5}$$

known as the Total Variation (TV) of the function $f(t)$. Both the $S(f)$ and $S_{\mathrm{TV}}(f)$ norms yield robust solutions with desired regularities but the $S_{\mathrm{TV}}(f)$ penalizes local jumps and isolated

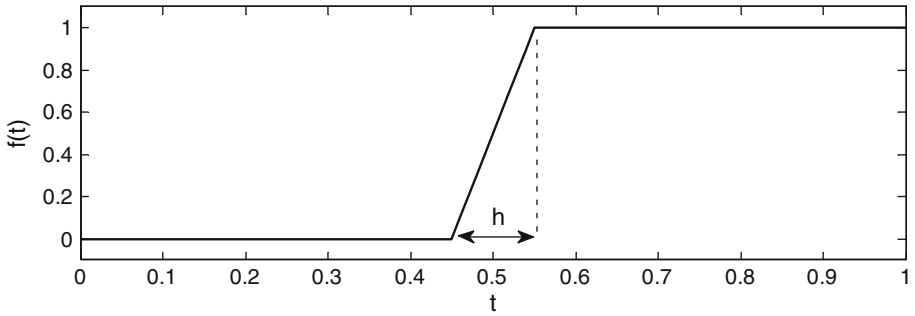

Fig. 4 A piecewise linear function $f(t)$ with a slope $f^{(1)} = 1/h$ at the non-horizontal part. As it is easily shown (see text), for this function, the ℓ_1(total variation)-norm $\left\|f^{(1)}\right\|_1$ is constant and independent of h while the ℓ_2-norm $\left\|f^{(1)}\right\|_2^2 = 1/h$ goes to infinity as h goes to zero (i.e., for a very steep gradient). As a result, the ℓ_2-norm solutions do not allow steep gradients, while the ℓ_1-norm does

singularities in a quite different way than the ℓ_2-norm of $S(f)$. It is important to demonstrate this point as it plays a key role in the proposed downscaling scheme.

Let us consider a piecewise linear function:

$$f(t) = \begin{cases} 0, & 0 \le t < \frac{1}{2}(1-h) \\ \frac{t}{h} - \frac{1-h}{2h}, & \frac{1}{2}(1-h) \le t \le \frac{1}{2}(1+h) \\ 1, & \frac{1}{2}(1+h) < t \le 1 \end{cases}, \tag{6}$$

as shown in Fig. 4. It can be shown that the smoothing norms associated with the ℓ_1 and ℓ_2-norms of $f^{(1)}(t)$ satisfy:

$$\left\|f^{(1)}\right\|_1 = \int_0^1 \left|f^{(1)}(t)\right| \mathrm{d}t = \int_0^h \frac{1}{h} \mathrm{d}t = 1 \tag{7}$$

while

$$\left\|f^{(1)}\right\|_2^2 = \int_0^1 \left|f^{(1)}(t)\right|^2 \mathrm{d}t = \int_0^h \frac{1}{h^2} \mathrm{d}t = \frac{1}{h}. \tag{8}$$

It is observed that the TV smoothing norm $S_{\mathrm{TV}}(f) = \left\|f^{(1)}\right\|_1$ is independent of the slope of the middle part of $f(t)$ while the smoothing ℓ_2-norm is inversely proportional to h and, as such, it severely penalizes steep gradients (when h is small). In other words, the ℓ_2-norm of $f^{(1)}$ will not allow any steep gradients and will produce a very smooth solution. Clearly, this is not desirable in solving an inverse problem associated with the reconstruction of small-scale details in precipitation fields, such as in the hurricane storm shown in Fig. 2.

2.2 Discrete Representation

Writing Eq. (1) in a discrete form, the problem of downscaling amounts to estimating a high-resolution (HR) state, denoted in an m-element vector as $\mathbf{x} \in \mathbb{R}^m$, from its low-resolution (LR) counterpart $\mathbf{y} \in \mathbb{R}^n$, where $m \gg n$. It is assumed that this LR counterpart

relates to the high-resolution (HR) state via a linear downgrading (e.g., a linear blurring and/or downsampling[1]) operator $\mathbf{H} \in \mathbb{R}^{m \times n}$ as follows:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \tag{9}$$

where $\mathbf{v} \sim \mathcal{N}(0, \mathbf{R})$ is a zero-mean Gaussian error with covariance $\mathbf{R}$. Due to the fact that the dimension of $\mathbf{y}$ is less than that of $\mathbf{x}$, the operator $\mathbf{H}$ is a rectangular matrix with more columns than rows and thus solving problem (9) for $\mathbf{x}$ is an ill-posed inverse problem (an under-determined system of equations with many solutions). As discussed above, we seek to impose a proper regularization to make the inverse problem well posed.

Following the developments presented above in a continuous setting and replacing $f^{(1)}$ with a discrete approximation derivative operator $\mathbf{L}$, the choice of the smoothing ℓ_2-norm regularization for $S(\mathbf{x})$ becomes $\|\mathbf{L}\mathbf{x}\|_2^2$ while for the ℓ_1-norm becomes $\|\mathbf{L}\mathbf{x}\|_1$, where in discrete space $\|\mathbf{x}\|_2^2 = \Sigma_{i=1}^m |x_i|^2$ and $\|\mathbf{x}\|_1 = \Sigma_{i=1}^m |x_i|$.

Thus, the solution (HR state $\mathbf{x}$) can be obtained by solving the following regularized weighted least squares minimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_{\mathbf{R}^{-1}}^2 + \lambda S(\mathbf{x}) \right\}, \tag{10}$$

It is clear that the smaller the value of λ, the more weight is given to fitting the observations (often resulting in data over-fitting), while a large value of λ puts more weight into preserving the underlying properties of the state of interest $\mathbf{x}$, such as large gradients. The goal is to find a good balance between the two terms. Currently, no closed form method exists for the selection of this regularization parameter and the balance has to be obtained via a problem-specific statistical cross validation (e.g., Hansen 2010). Note that the problem in (10) with $S(\mathbf{x}) = \|\mathbf{L}\mathbf{x}\|_1$ is:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_{\mathbf{R}^{-1}}^2 + \lambda \|\mathbf{L}\mathbf{x}\|_1 \right\}, \tag{11}$$

that is, a non-smooth convex optimization problem as the regularization term is non-differentiable at the origin. As a result, the conventional iterative gradient methods do not work and one has to use greedy methods (Mallat and Zhang 1993) or apply the recently developed non-smooth optimization algorithms such as the iterative shrinkage thresholding method (Tibshirani 1996), the basis pursuit method (Chen et al. 1998, 2001), the constrained quadratic programming (Figueiredo et al. 2007), the proximal gradient-based methods (Beck and Teboulle 2009), or the interior point methods (Kim et al. 2007). In this work, we have adopted the method suggested by Figueiredo et al. (2007).

2.3 Geometrical Versus Statistical Interpretation of the ℓ_1-Norm Regularized Downscaling

As was discussed in the introduction, the motivation for introducing a new downscaling framework lies in the desire to reproduce some geometrical but also some statistical features of precipitation fields. Specifically, the question was posed as to how a downscaling scheme could be constructed that can reproduce both the abrupt localized gradients

[1] Here, by downsampling, we mean to reduce the sampling rate of the rainfall observations by a factor greater than one.

and also the characteristic probability distribution of the precipitation intensity gradients such as that displayed in Fig. 3a.

It can be shown that the solution of (10) obtained via ℓ_2-norm regularization (i.e., $S(\mathbf{x}) = \|\mathbf{Lx}\|_2^2$) is equivalent to the Bayesian maximum a posteriori (MAP) estimator where the transformed variable $\mathbf{Lx}$ is well explained by a Gaussian distribution. On the other hand, considering $S(\mathbf{x}) = \|\mathbf{Lx}\|_1$, the ℓ_1-norm regularized solution of (10), i.e., the solution of Eq. (11), is the MAP estimator where $\mathbf{Lx}$ is well explained by the multivariate Laplace distribution (the generalized Gaussian family with $\alpha = 1$). In other words, the ℓ_1-regularization implicitly assumes that the probability of $\mathbf{Lx}$ goes as $\exp\left(-\lambda\|\mathbf{Lx}\|_1\right)$ (Lewicki and Sejnowski 2000; Ebtehaj and Foufoula-Georgiou 2013). We note that for the storm of Fig. 2, the estimated tail parameter α is 0.85 (see Fig. 3), which denotes that the pdf of $\mathbf{Lx}$ goes as $\exp\left(-\lambda\|\mathbf{Lx}\|_\alpha^\alpha\right)$, where $\|\mathbf{x}\|_\alpha^\alpha = \Sigma_{i=1}^m |x_i|^\alpha$. This value of α implies that the Laplace distribution ($\alpha = 1$) is only an approximation of the true distribution of the analyzed precipitation (see Fig. 3b for comparison), making thus the proposed ℓ_1-norm regularization solution only an approximate solution in a statistical sense. Finding a solution via regularized inverse estimation that satisfies a prior probability for ($\mathbf{Lx}$) with $\alpha < 1$ requires solving a non-convex optimization, which may suffer from local minima and may be hard to solve for large-scale problems. For this reason, we limit our discussion to the ℓ_1-regularization recognizing the slight sub-optimality of the solution for precipitation applications but also its superiority relative to the Gaussian assumption about the rainfall derivatives.

3 Working with an Unknown Downgrading Operator (H)

In the above formulation of the downscaling problem as an inverse problem, the downgrading operator $\mathbf{H}$ is assumed to be linear and known a priori. A mathematically convenient form for the downgrading operator is to assume that it can be represented via a linear convolution followed by downsampling. In other words, one may assume that the low-resolution (LR) observation is obtained by applying an overlapping box (weighted) averaging over the HR field and keeping one observation only, typically at the center, per averaging box (downsampling). However, the downgrading operator is not generally known in practice and its characterization might be sensor-dependent. Also often, this operator is highly nonlinear (e.g., the relationship between the radiometer-observed brightness temperature and the precipitation reflectivity observed by the radar) and its linearization may introduce large estimation errors. This nonlinearity may also pose severe challenges from the optimization point of view and may give rise to a hard non-convex problem with many local minima (Bertsekas 1999).

To deal with the problem of an unknown downgrading operator, Ebtehaj et al. (2012) proposed a dictionary-learning-based methodology that allows to implicitly incorporate the downgrading effect via statistical learning without the need to explicitly characterize the downgrading operator. In this methodology, the downgrading operator is being learned via a dictionary of coincidental HR and LR observations (e.g., in practice, TRMM-PR, and ground-based NEXRAD or TMI and NEXRAD). The methodology is explained in detail by Ebtehaj et al. (2012) and is only briefly summarized herein.

In simple terms, the idea is to reconstruct a HR counterpart of the LR rainfall field based on learning from a representative data base of previously observed coincidental LR and HR rainfall fields (e.g., TRMM-PR and NEXRAD observations). As is evident, due to different

underlying physics, the shape and patterns of rainfall intensities, viewed in a storm-scale field of view, might be drastically dissimilar. However, the small-scale patterns of rainfall when viewed over smaller windows might be repetitive and "similar" within different regions of the same storm or within different storms. Therefore, the central idea is to (a) collect a representative set of coincidentally observed LR and HR rainfall fields, with some similarities in their underlying physics; (b) zoom down into small-scale patterns (patches) of the given LR rainfall field; (c) for each patch, find few but very similar LR patches in the collected data base; (d) for those similar LR patches, obtain the corresponding HR patches in the data base and then reconstruct the HR counterparts of the LR patch of interest based on an optimality criterion; and (e) repeat this procedure for all possible patches and obtain a HR estimate for the observed LR rainfall field.

To be more specific, let us consider that the representative training set of N coincidental pairs of LR and HR rainfall fields are denoted by $\left\{\mathcal{Z}_l^i\right\}_{i=1}^N$ and $\left\{\mathcal{Z}_h^i\right\}_{i=1}^N$, respectively. As previously explained, for each patch $\mathbf{y}_l$ of the given LR rainfall field, we need to find a few very similar patches in $\left\{\mathcal{Z}_l^i\right\}_{i=1}^N$, where similarity is defined in terms of localized rainfall fluctuations and not in the mean values of the rainfall patches. To this end, all of the LR fields are projected (i.e., $\mathcal{Z}_l^i \rightarrow \left(\mathcal{Z}_h^i\right)'$) onto a redundant orthogonal basis (called feature space) to capture the rainfall local fluctuations including horizontal and vertical edges (i.e., zonal and meridional) and curvatures. This was performed by Ebtehaj et al. (2012) via an undecimated orthogonal Haar wavelet, which basically performs a high-pass filtering in each direction using first- and second-order differencing. Then, all of the constituent patches of the transformed LR fields in the data base were extracted, vectorized in a fixed order, and then stored as columns of a matrix $\mathbf{\Psi}$, the so-called empirical LR-dictionary.

Clearly, for each coincidental pair $\left(\mathcal{Z}_l^i, \mathcal{Z}_h^i\right)$, a set of "residual fields" can be formed by subtracting the LR fields from their HR counterparts via $\mathcal{R}_h^i = \mathcal{Z}_h^i - \mathbf{Q}\mathcal{Z}_l^i$, where $\mathbf{Q}$ is a readily available interpolation operator (e.g., a nearest-neighbor or bilinear, bicubic interpolator). Notice that, these residual fields contain the rainfall variability and high-frequency (fine spatial-scale) components that are not captured by the LR sensor and need to be recovered. Therefore, all of the constituent patches $\mathbf{r}_h$ of the residual fields can also be collected, vectorized in a fixed order, and then stored in the columns of a matrix Φ, the so-called HR-dictionary. Note that, by the explained construction, the empirical LR and HR dictionaries share the same number of columns while there is a one-to-one correspondence between them. In other words, while the columns of the $\mathbf{\Psi}$ contain LR rainfall features, the columns of the Φ contain the corresponding HR residuals, needed for the reconstruction of the HR field.

The premise is that the local variability of any LR patch $\mathbf{y}_l$, denoted by $\mathbf{y}_l'$, in any storm can be well approximated by a linear combination of the elements of the LR dictionary as follows:

$$\mathbf{y}_l' = \mathbf{\Psi}\mathbf{c} + \mathbf{v}, \tag{12}$$

where $\mathbf{c}$ is the vector of representation coefficients in the LR dictionary and $\mathbf{v} \sim \mathcal{N}(0, \mathbf{R})$ denotes the estimation error that can be well explained by a Gaussian density.

By analyzing a sample of 100 storms over Texas, it was documented by Ebtehaj et al. (2012) that the vector of representation coefficients $\mathbf{c}$ in the LR dictionary is very sparse. In other words, any desired local rainfall variability in the given LR field can be approximated by a linear combination of only a few columns of the LR empirical dictionary (of the order

of 3–5 elements). To impose this sparsity (called "group sparsity") in solving (12) for $\mathbf{c}$, the solution needs to be constrained via an ℓ_1-norm regularization as follows:

$$\hat{\mathbf{c}} = \underset{\mathbf{c}}{\operatorname{argmin}} \left\{ \frac{1}{2} \left\| \mathbf{y}_l' - \mathbf{\Psi c} \right\|_{\mathbf{R}^{-1}}^2 + \lambda \|\mathbf{c}\|_1 \right\}. \tag{13}$$

Using the representation coefficients obtained from (13), one can recover the corresponding residual fields (the details missed by the LR sensor) as follows:

$$\hat{\mathbf{r}} = \mathbf{\Phi} \hat{\mathbf{c}}. \tag{14}$$

Having the estimated residual fields, the HR patch can be obtained as $\hat{\mathbf{x}} = \mathbf{Q}\mathbf{y}_l + \hat{\mathbf{r}}$. Applying the same estimation methodology for all of the patches of the given LR rainfall field, we can recover the entire HR rainfall field (see Ebtehaj et al. 2012). The most important implication of the above framework is that we characterized the pair of $(\mathbf{\Psi}, \Phi)$ empirically without explicit access to the structure of the downgrading operator $\mathbf{H}$, which is the main advantage of this dictionary-based rainfall downscaling method versus the previously explained approach. Since advantage was taken of the rainfall group sparsity (and also implicitly of the sparsity of the precipitation fields themselves), the dictionary-based downscaling methodology was termed SPaD.

4 Results from a Case Study

To demonstrate the proposed downscaling methodology, we have chosen a specific tropical storm, hurricane Claudette, which occurred in July 2003. Claudette began as a tropical wave in the eastern Caribbean on July 8, 2003 and moved quickly westward to the Gulf of Mexico. It remained a tropical storm until just before making landfall in Port O'Conner, Texas, when it quickly strengthened to a category 1 hurricane. Although Claudette produced moderate rainfall across southern Texas, peaking at approximately 6.5 inches (165 mm), it maintained a tropical storm intensity for over 24 h after landfall with winds gusting to 83 mph (134 km/h) at Victoria Regional Airport, Texas. The storm caused excessive beach erosion and damages estimated at 180 million dollars. For this storm, we have available data from a NEXRAD station in Houston, Texas, for which a snapshot at 11:51:00 (UTC) on July 15, 2003 is shown in Fig. 2.

The issues we want to examine here are the following: (1) the ability of the proposed variational downscaling (VarD) scheme to reproduce the steep gradients in precipitation intensities as evidenced by reproducing the tails of the PDF of intensity gradients; (2) the effect of an unknown kernel (smoothing and downsampling operation imposed on the true HR field by a sensor) on the downscaling scheme performance using the proposed methodology; (3) a comparison of the VarD method with a local dictionary-based methodology based on sparse representation (SPaD) as discussed in the previous section; and (4) insights into the ability of the proposed VarD methodology and SPaD to reproduce not only the extreme gradients but also the extreme rainfall intensities, i.e., the tails of the rainfall intensity probability distribution functions (PDFs).

The original HR data at 1×1 km (Fig. 5a) were downgraded to 8×8 km LR observations via a coarse-graining filter consisting of a simple box averaging of size 8×8 followed by downsampling with a factor of 8 (i.e., keeping one observation per box of 8×8 km). The resulting LR field is shown in Fig. 5b and is considered to be the field that would be available to us from a satellite sensor. Figure 5c, d shows the results of

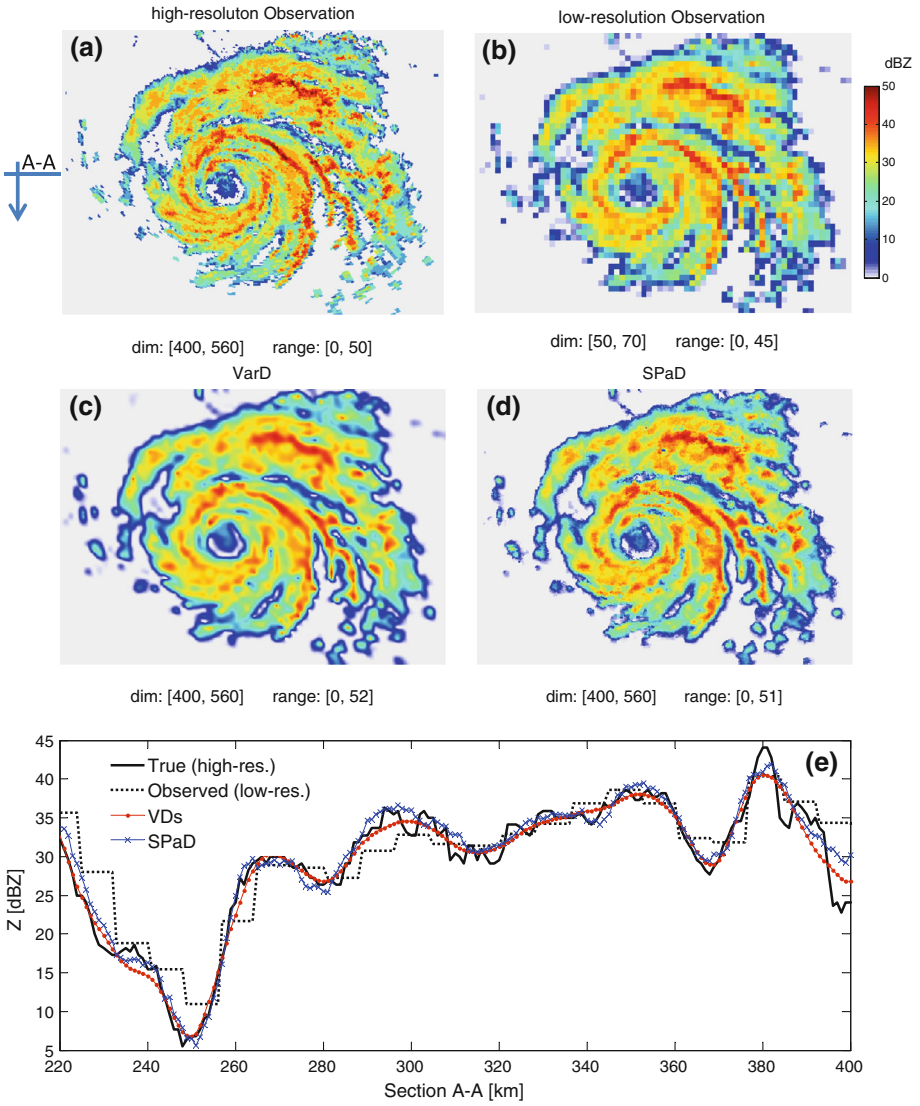

Fig. 5 **a** Original HR base reflectivity snapshot at resolution 1 × 1 km over TX (hurricane Claudette, 08-16-2003, UTC 11:51:00); **b** The synthetic LR observation obtained by coarse graining of the field up to scale 8 × 8 km (smoothing with an average filter of size 8 × 8 followed by downsampling by a factor 8); **c** result of the downscaled field at resolution 1 × 1 km using the variational downscaling (VarD) method; and **d** results of the dictionary-based sparse precipitation downscaling (SPaD) method at resolution 1 × 1 km; **e** intensities averaged over a bandwidth of 8 km centered at a cross section A-A in (**a**), displaying the true HR field, the LR coarse-grained field (observations), and the two downscaled fields

downscaling the 8 × 8 km field to 1 × 1 km resolution using the VarD and SPaD methodologies with $\lambda \cong 0.05 \left\| \mathbf{L}^{-\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{y} \right\|_{\infty}$ in the original formulation of the problem (11), where $\|\mathbf{x}\|_{\infty} = \max(|x_1|, \ldots, |x_m|)$. Note that in all of our experiments, we empirically found that $0 < \lambda \leq 0.10 \left\| \mathbf{L}^{-\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{y} \right\|_{\infty}$ works well for rainfall downscaling in both

methods, while it can be shown that the solution of problem (11) is zero for all $\lambda \geq \left\| \mathbf{L}^{-\mathrm{T}} \mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{y} \right\|_{\infty}$.

As discussed before, VarD assumes the downgrading operator **H** to be known. In our case, we used as **H** the "true" operator, i.e., the same operator we used to coarse grain the HR (1 × 1 km) reflectivity field to the LR (8 × 8 km) one. It is observed that the VarD downscaled field has a smoother appearance than the original field (it does not have the 1 × 1 km pixelized appearance of the original HR field), which is not unexpected given that the ℓ_1-regularization promotes smoothness in the solution while allowing for some steep gradients as demonstrated in the illustrative example of Fig. 4. A one-dimensional cross section shown in Fig. 5e confirms this observation and shows that the downscaled field is much closer to the true field compared to the LR field.

Suppose now that the true filter **H** is not known and only the LR field is given without guidance as to what "filtering" the sensor did to the HR field to return the LR observations. As discussed in the previous section, and in Ebtehaj et al. (2012), we demonstrated that this filter can be "learned" implicitly and locally using coincidental high- and low-resolution images available for a number of similar storms. In that study, a sample of 100 HR summer storms over Texas was used to construct a set of coincidental LR storms (using again a simple box averaging and a downsampling operator). This hundred storm sample was then used to compute the LR and HR dictionaries, which formed the basis of the SPaD method as explained in the previous section. This same dictionary was used herein to recover the 1 × 1 km HR rainfall field of the Claudette storm from 8 × 8 km observations. The results are shown in Fig. 5d.

In general, it is expected that the SPaD method will outperform the VarD method when the operator **H** is not known at all or is locally varying, due, for example, to instrument range effects or cloud interference or different performance of an instrument in low- versus high-resolution rainfall intensities. However, it is noted that, since in our data base the LR and HR fields relate to each other with a simple box averaging operator **H** (by construction), we expect that the dictionary-learning SPaD downscaling will perform comparably to the VarD method. Extra information in SPaD will be gained by the localized nature of the estimation methodology, which might reproduce extra high-frequency (small-scale) features, obtained from the available dictionaries that may not be recovered in the VarD approach.

To more quantitatively compare the two downscaled fields to the true underlying HR field and to each other, we compare in Fig. 6 the PDF of the derivatives in the horizontal direction in terms of their q–q plot (quantiles of the variable of interest vs. standard normal quantiles). We observe that both methods are able to reproduce the heavy tails of the PDF of the precipitation gradients, which are much thicker than those of the Gaussian PDF, and thus, both methods are able to reproduce high gradients in the HR recovered field. VarD is seen to slightly outperform SPaD in reproducing high positive gradients, not surprisingly since, in VarD, the **H** operator was customized to this specific storm, while, in SPaD, information from a suite of other storms was also used.

Turning our attention to the preservation of the statistics of the precipitation field itself, we show in Fig. 7 the comparison of the PDFs of the LR rainfall field with that of the true HR field and the downscaled fields. We recall that although the preservation of the thicker-than-Gaussian (Laplace) tails in the PDF of precipitation intensity gradients is explicitly incorporated in the ℓ_1-norm VarD downscaling methodology, no explicit preservation of the extreme rainfall intensities themselves is accounted for. However, it is clear from Fig. 7 that VarD performs satisfactorily in reproducing extreme rainfall intensities in the

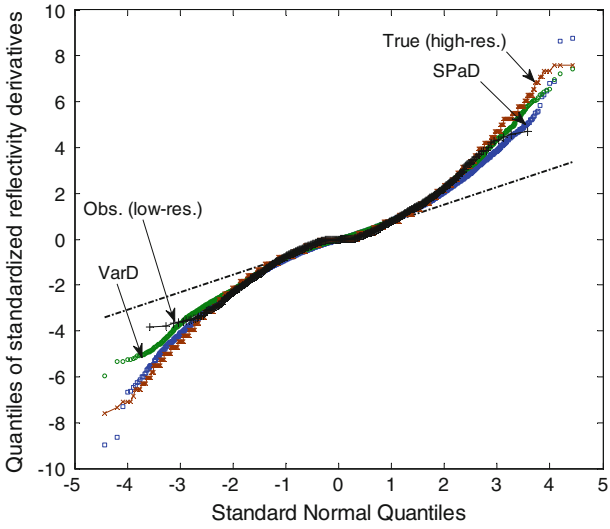

Fig. 6 The quantiles of the standardized horizontal (zonal) derivatives ($\nabla_x Z$) of the true HR field (*red* ×), the LR observation (black +), the VarD downscaled field (*green* o), and the SPaD downscaled field (*blue* □) versus standard normal quantiles. The broken *straight line* represents quantiles of the Gaussian distribution. It is observed that both downscaling methodologies are able to reproduce the extreme reflectivity gradients

downscaled field and is able to enhance substantially the tails of the low-resolution rainfall fields. One of the reasons for reproducing extreme rainfall intensities is that typically extreme gradients are collocated with high rainfall intensities. This was observed and documented by Perica and Foufoula-Georgiou (1996) and is also documented for the Claudette storm in Fig. 8. So, indirectly, VarD is bound to preserve satisfactorily the tails of the PDF of precipitation intensities. From Fig. 7, it is apparent that SPaD outperforms VarD in preserving extreme rainfall. This is attributed to the fact that, in SPaD, the operator is learned directly on the precipitation intensities, and not on the gradients, allowing thus for a more direct reconstruction of extreme intensities, provided that such extremes are available in the data base.

Table 1 presents a comparison of the downscaling methodologies in terms of several quantitative metrics: the mean square error: $\mathrm{MSE} = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2 / \|\mathbf{x}\|_2^2$, the maximum absolute error: $\mathrm{MAE} = \|\mathbf{x} - \hat{\mathbf{x}}\|_1 / \|\mathbf{x}\|_1$, the peak signal-to-noise ratio: $\mathrm{PSNR} = 20 \log_{10}[\max(\hat{\mathbf{x}})/\mathrm{std}(\mathbf{x} - \hat{\mathbf{x}})]$, and the Kullback–Leibler divergence: $\mathrm{KLD}(p_{\mathbf{x}} || p_{\hat{\mathbf{x}}}) = \Sigma_i \ln[p_{\mathbf{x}}(i)/p_{\hat{\mathbf{x}}}(i)] p_{\mathbf{x}}(i)$ or relative entropy metric, where $p_{\mathbf{x}}(i)$ and $p_{\hat{\mathbf{x}}}(i)$ are the discrete probabilities of the true and estimated rainfall, respectively. The KLD is a non-negative measure that represents a relative degree of closeness of two PDFs in terms of their entropy, while smaller values signify a stronger degree of similarity. It can be seen from Table 1 that both downscaling methods produce HR fields that are closer to the true field compared to the LR field and that the VarD and SPaD methods considerably outperform the "naïve" simple downscaling methods such as the result obtained by the bicubic interpolation scheme. SPaD is seen to outperform VarD in terms of the entropy metric (smaller KLD value) further speaking for the better reproduction of very extreme rainfall intensities.

It is worth presenting here some extra insight into the effect of a misdiagnosed observation filter $\mathbf{H}$ on the downscaled field. As shown in the illustrative example of Fig. 9, when the observation operator is smoother (a Gaussian filter) as compared to the operator used in the VarD downscaling (a box average filter), the downscaled field exhibits a

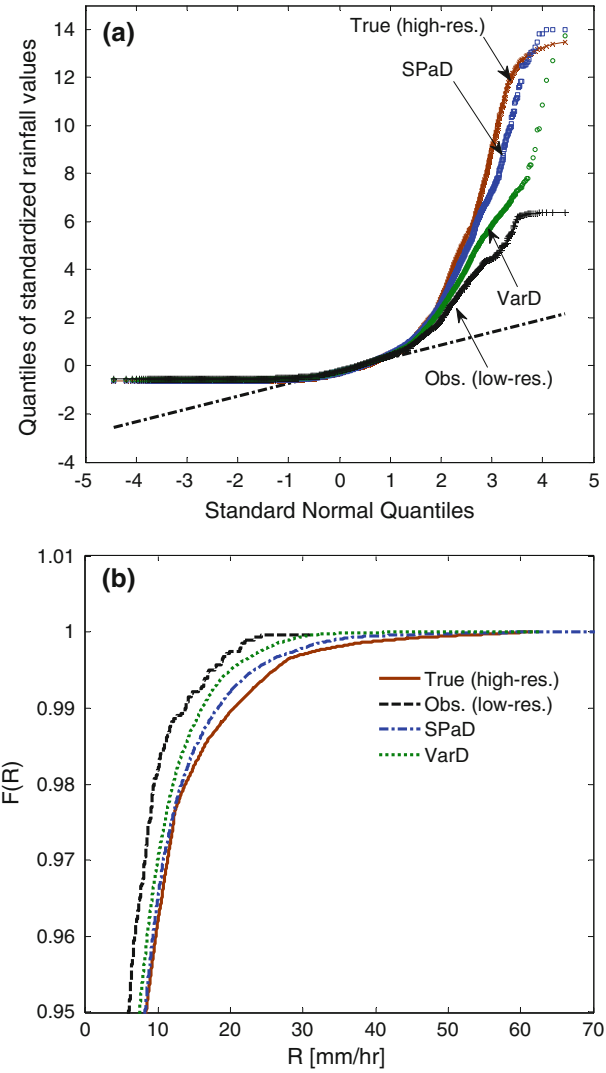

Fig. 7 **a** The quantiles of the standardized recovered rain (i.e., $Z = 300R^{1.4}$) versus standard normal quantiles. The rainfall quantile values are only for the positive rainy part of the fields and are standardized by subtracting the mean and dividing by the standard deviation. **b** Comparison of the cumulative distribution of rainfall intensities focusing on extremes. Both plots show how both the VarD and SPaD downscaling methodologies reproduce extreme rainfall intensities not present in the observed LR fields

blockiness coming from the mismatch between the assumed and true filters. In fact, this blockiness provides a qualitative diagnostic of the filter mismatch, in that it picks up the fact that the underlying true observation filter (the Gaussian in this case) was smoother than the one used for recovery. Apart from the visual inspection of the downscaled field, Fig. 9 (caption) provides the comparison metrics that show the underperformance of this downscaled field relatively to the one obtained using the correct filter (compare values with those in Table 1). The possibility of developing a methodology to learn properties (e.g., smoothness and nonlinearity) of the underlying observation filter in the case that no coincidental LR and HR data sets are available to apply the dictionary-based methodology is appealing and warrants further exploration.

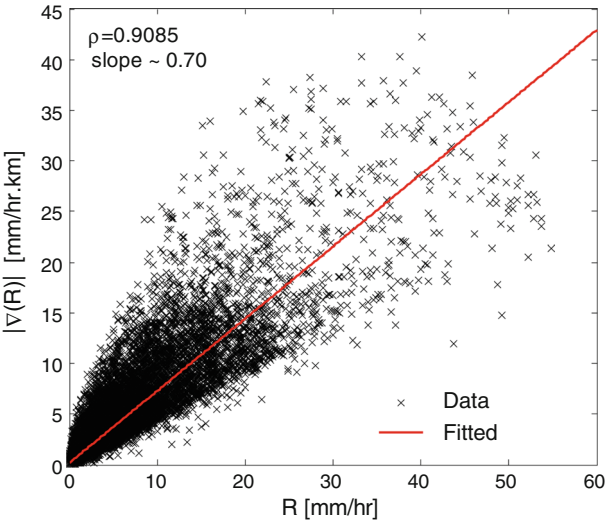

Fig. 8 Magnitude of precipitation gradients $|\nabla(R)| = \sqrt{(\nabla_x R)^2 + (\nabla_y R)^2}$ versus precipitation intensity for the nonzero pixels of the Claudette storm at resolution of 1 × 1 km showing that high precipitation gradients are mostly collocated with high precipitation intensities. Only pixels for which the gradient was at least 20 % of the local precipitation intensity were considered

Table 1 Error statistics obtained by comparing the HR precipitation reflectivity image of Hurricane Claudette (true) with the LR one, the downscaled fields via Bicubic interpolation, the VarD, and the SPaD methodologies (see text for definition of these metrics)

	Quality metrics			
	MSE †	MAE	PSNR	KLD
Low. res.	0.305	0.260	17.834	0.089
Bicubic	0.275	0.246	18.742	0.113
VarD	0.194	0.172	22.539	0.065
SPaD	0.209	0.177	22.015	0.044

† *MSE* mean squared error, *MAE* mean absolute error, *PSNR* peak signal-to-noise ratio, *KLD* Kullback–Leibler divergence

5 Concluding Remarks

The problem of downscaling climate variables remains of interest as more spaceborne observations become available and as the need to translate low-resolution (LR) climate predictions to regional and local scales becomes essential for long-term planning purposes. Of special interest are downscaling schemes that can accurately reproduce not only overall statistical properties of rainfall but also specific features of interest, such as extreme rainfall intensities and abrupt gradients. In this paper, such a precipitation downscaling scheme was introduced using a formalism of inverse estimation and solving the (ill-posed) inverse problem by imposing certain constraints that guarantee stability and uniqueness of the solution while also enforcing a certain type of smoothness that allows for some abrupt gradients. Mathematically, this inverse problem is solved via what is called an ℓ_1-norm or total variation regularization. We showed the equivalence of the proposed total variation regularized solution to a statistical maximum a posteriori (MAP) Bayesian solution, which has a Laplace prior distribution in the derivative domain. We demonstrated the

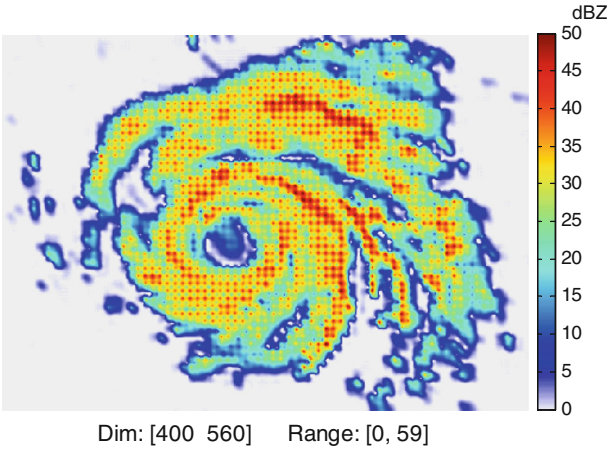

Fig. 9 VarD result for downscaling precipitation reflectivity from scale 8 × 8 to 1 × 1 km with a "wrong" observation operator. In this experiment, the imposed observation operator was a Gaussian filter of size 8 × 8 with standard deviation 2 while in downscaling, we assumed a uniform average filter of the same size. It is clear from the result that the quality of downscaling is blocky and is severely deteriorated because of the misspecification of the observation operator in the downscaling scheme. The selected quantitative measures are as follows: MSE = 0.244; MAE = 0.220; PSNR: 22.0; and KLD = 0.075 (see Table 1)

performance of the proposed downscaling scheme on a tropical storm and concluded that it was able to capture adequately both the extremes of rainfall intensities and gradients.

A practical challenge faced in applying the proposed methodology is that the observation operator (which relates the true unknown HR field to the LR observations) might not be known. In fact, it might be even changing locally due to sensor properties as affected, for example, by range or precipitation intensity and composition. If coincidental high- and low-resolution fields are available in a data base, the data-driven dictionary-based methodology introduced by Ebtehaj et al. (2012) offers promise and, although more computationally intensive, it might offer advantages in capturing more faithfully local details and extremes. However, a lot more work is needed to understand the sensitivity of the dictionary-based methodology to the selection of a data base from environments different than the storm of interest, as well as when the observation filter relates nonlinearly to the underlying field as is the case in problems of retrieval, i.e., estimation of precipitation intensity from radiances recorded by the TRMM microwave imager.

The presence of statistical self-similarity (scaling) in spatial rainfall, manifesting in log–log linearity in the Fourier or wavelet power spectra and also in higher-order statistical moments, has been well documented by now (see discussion in the introduction). This structure, often explained in the context of mono or multifractal formalisms, has guided the development of several stochastic downscaling methodologies (e.g., Rebora et al. 2006a, b; Perica and Foufoula-Georgiou 1996, among many others). The downscaled precipitation fields produced by these models are, by construction, respecting the rainfall scaling laws; however, they are not unique as multiple realizations of plausible high-resolution rainfall fields with the same input parameters can be produced without following a specific optimality criterion. On the other hand, the proposed downscaling methodologies produce unique high-resolution rainfall fields based on the aforementioned optimality criteria that also allow us to partially preserve the underlying non-Gaussian structure of the rainfall fields. An important question that arises then is whether statistical scaling in rainfall fields,

although not prescribed in our method, arises as an emergent property. The answer to this question is not obvious. Our preliminary results (not reported herein) demonstrate that statistical scaling indeed arises in both the ℓ_1-norm variational downscaling (VarD) and the SPaD schemes. However, the power law exponents (of the variance of the wavelet coefficients as a function of scale) and the variance of the wavelet coefficients at the smallest scale (similar to the analysis in Perica and Foufoula-Georgiou 1996) seem to be lower than those of the original fields. This might be due to the fact that, although our scheme is able to accurately capture, much better than other statistical schemes, the magnitude of the infrequent localized large gradients in precipitation fields, it might under-produce the variability of the smaller gradients, reducing thus the overall variance. This is an issue that is currently explored both from a theoretical perspective and via simulation, as in most applications one is interested to preserve both the localized extremes but also the overall variance of the smaller magnitude fluctuations.

The work presented herein falls within a larger research direction of using variational regularization approaches or equivalently, Bayesian MAP estimators with heavy-tailed priors in the derivative domain, for estimation problems in hydro-climatology, such as downscaling, multi-sensor data fusion, retrieval, and data assimilation (see Ebtehaj and Foufoula-Georgiou 2013). A relatively small number of abrupt gradients within the field of interest or heavy-tailed PDFs in the derivative domain are associated with the notion of sparsity, that is, the fact that, when the state is projected in a suitable basis, most of the projection coefficients are close to zero and only a few coefficients carry most of the state energy. Estimation problems of sparse states (posed in an inverse estimation setting or in a variational setting of minimizing a functional) require the use of ℓ_1-norm regularization, which results from imposing extra constraints on the solution to enforce sparsity. Motivated by the need to preserve sharp weather fronts in data assimilation of numerical weather prediction models, an ℓ_1-norm regularized variational data assimilation methodology was recently proposed by Freitag et al. (2012) and demonstrated in a simple setting using the advection equation for the state evolution dynamics. In Ebtehaj and Foufoula-Georgiou (2013), data assimilation in the presence of extreme gradients in the state variable was further analyzed using as illustrative example the advection–diffusion equation that forms the basis of many hydro-meteorological problems, such as those dealing with the estimation of surface heat fluxes based on the assimilation of land surface temperature (e.g., see Bateni and Entekhabi 2012). Application of these new non-smooth variational methodologies in real data assimilation problems, and also in combining data assimilation with downscaling of the state, is only in its infancy and is certain to occupy the geophysical community in the years to come.

Acknowledgments This work has been mainly supported by a NASA-GPM award (NNX10AO12G), a NASA Earth and Space Science Fellowship (NNX12AN45H), and a Doctoral Dissertation Fellowship of the University of Minnesota to the second author. The insightful comments of one anonymous referee are also gratefully acknowledged.

References

Badas MG, Deidda R, Piga E (2006) Modulation of homogeneous space-time rainfall cascades to account for orographic influences. Nat Hazard Earth Syst 6(3):427–437. doi:10.5194/nhess-6-427-2006

Bateni SM, Entekhabi D (2012) Surface heat flux estimation with the ensemble Kalman smoother: joint estimation of state and parameters. Water Resour Res 48(3). doi:10.1029/2011WR011542

Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J Imaging Sci 2(1):183–202. doi:10.1137/080716542

Bertsekas DP (1999) Nonlinear programming, 2nd edn. Athena Scientific, Belmont, MA, p 794
Chen S, Donoho D, Saunders M (2001) Atomic decomposition by basis pursuit. SIAM Rev 43(1):129–159
Chen SS, Donoho DL, Saunders MA (1998) Atomic decomposition by basis pursuit. SIAM J Sci Comput 20:33–61
Deidda R (2000) Rainfall downscaling in a space-time multifractal framework. Water Resour Res 36(7):1779–1794
Ebtehaj AM, Foufoula-Georgiou E (2011) Statistics of precipitation reflectivity images and cascade of Gaussian-scale mixtures in the wavelet domain: a formalism for reproducing extremes and coherent multiscale structures. J Geophys Res 116:D14110. doi:10.1029/2010JD015177
Ebtehaj AM, Foufoula-Georgiou E, Lerman G (2012) Sparse regularization for precipitation downscaling. J Geophys Res 116:D22110. doi:10.1029/2011JD017057
Ebtehaj AM, Foufoula-Georgiou E (2013) Variationl downscaling, fusion and assimilation of hydrometeorological states: a unified framework via regularization. Water Resour Res. doi:10.1002/wrcr.20424
Figueiredo M, Nowak R, Wright S (2007) Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. IEEE J Sel Topics Signal Process 1(4):586–597. doi:10.1109/JSTSP.2007.910281
Flaming GM (2004) Measurement of global precipitation. In: Geoscience and remote sensing symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE international, vol 2, p 918–920
Freitag MA, Nichols NK, Budd CJ (2012) Resolution of sharp fronts in the presence of model error in variational data assimilation. Q J Roy Meteor Soc. doi:10.1002/qj.2002
Hansen P (2010) Discrete inverse problems: insight and algorithms, vol. 7. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA
Harris D, Foufoula-Georgiou E, Droegemeier KK, Levit JJ (2001) Multiscale statistical properties of a high-resolution precipitation forecast. J Hydrometeor 2(4):406–418
Krajewski WF, Smith JA (2002) Radar hydrology: rainfall estimation. Adv Water Resour 25(8–12):1387–1394. doi:10.1016/S0309-1708(02)00062-3
Kim S-J, Koh K, Lustig M, Boyd S, Gorinevsky D (2007) An interior-point method for large-scale ℓ_1-regularized least squares. IEEE J Sel Topics Signal Process. 1(4):606–617. doi:10.1109/JSTSP.2007.910971
Kumar P, Foufoula-Georgiou E (1993) A multicomponent decomposition of spatial rainfall fields. 2. Self-similarity in fluctuations. Water Resour Res 29(8):2533–2544
Kumar P, Foufoula-Georgiou E (1993) A multicomponent decomposition of spatial rainfall fields. 1. Segregation of large- and small-scale features using wavelet transforms. Water Resour Res 29(8):2515–2532
Lewicki M, Sejnowski T (2000) Learning overcomplete representations. Neural Comput 12(2):337–365
Lovejoy S, Mandelbrot B (1985) Fractal properties of rain, and a fractal model. Tellus A 37(3):209–232
Lovejoy S, Schertzer D (1990) Multifractals, universality classes and satellite and radar. J Geophys Res 95(D3):2021–2034
Mallat S, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Proces 41(12):3397–3415. doi:10.1109/78.258082
Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell 11(7):674–693. doi:10.1109/34.192463
Nykanen DK, Foufoula-Georgiou E, Lapenta WM (2001) Impact of small-scale rainfall variability on larger-scale spatial organization of land-atmosphere fluxes. J Hydrometeor 2(2):105–121
Perica S, Foufoula-Georgiou E (1996) Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling. J Geophys Res 101(D21):26–347
Rebora N, Ferraris L, Von Hardenberg J, Provenzale A et al (2006) Rainfall downscaling and flood forecasting: a case study in the Mediterranean area. Nat Hazard Earth Syst 6(4):611–619
Rebora N, Ferraris L, Von Hardenberg J, Provenzale A (2006) RainFARM: rainfall downscaling by a filtered autoregressive model. J Hydrometeor 7:724–738
Sapozhnikov VB, Foufoula-Georgiou E (2007) An exponential Langevin-type model for rainfall exhibiting spatial and temporal scaling. Nonlinear Dyn Geosci :87–100
Tibshirani R (1996) Regression shrinkage and selection via the Lasso. J R Stat Soc Ser B Stat Methodol 58(1):267–288
Venugopal V, Roux SG, Foufoula-Georgiou E, Arneodo A (2006) Revisiting multifractality of high-resolution temporal rainfall using a wavelet-based formalism. Water Resour Res 42(6):6. doi:10.1029/2005WR004489
Venugopal V, Roux SG, Foufoula-Georgiou E, Arnéodo A (2006) Scaling behavior of high resolution temporal rainfall: new insights from a wavelet-based cumulant analysis. Phys Lett A 348(3):335–345