Automated Invariant Alignment to Improve Canonical Variates in Image Fusion of Satellite and Weather Radar Data

JACOB S. VESTERGAARD

DTU Informatics, Technical University of Denmark, Kongens Lyngby, Denmark

ALLAN A. NIELSEN

DTU Space, Technical University of Denmark, Kongens Lyngby, Denmark

(Manuscript received 22 December 2011, in final form 22 August 2012)

ABSTRACT

Canonical correlation analysis (CCA) maximizes the correlation between two sets of multivariate data. CCA is applied to multivariate satellite data and univariate radar data to produce a subspace descriptive of heavily precipitating clouds. A misalignment, inherent to the nature of the two datasets, was observed, corrupting the subspace. A method for aligning the two datasets is proposed to overcome this issue and render a useful subspace projection. The observed corruption of the subspace gives rise to the hypothesis that the optimal correspondence between a heavily precipitating cloud in the radar data and the associated cloud top registered in the satellite data is found by a scale, rotation, and translation invariant transformation together with a temporal displacement. The method starts by determining a conformal transformation of the radar data at the time of maximum precipitation for optimal correspondence with the satellite data at the same time. This optimization is repeated for an increasing temporal lag until no further improvement can be found. The method is applied to three meteorological events that caused heavy precipitation in Denmark. The three cases are analyzed with and without using the proposed method. In all cases, the use of prealignment shows significant improvements in the descriptive capabilities of the subspaces, thus supporting the posed hypothesis.

1. Introduction

Having two multivariate datasets describing different characteristics of the same scene often poses the question of how to use the two sets in a joint analysis. The combination of several sources of image data is tractable for a variety of applications, for example, sharpening of images, substitution of missing data, segmentation, and change detection. A review of existing methods for the fusion of multisource imagery is given by Schowengerdt (2007). We consider here the challenges of determining a suitable subspace projection of the original variables for the purpose of image fusion.

The best choice of subspace depends very much on the data and the application. For a single set of variables, the methods range from well-known linear transformations (Jolliffe 2002; Green et al. 1988) over nonorthogonal subspaces (Bell and Sejnowski 1995) to highly nonlinear (kernel) methods (Schölkopf et al. 1998; Nielsen 2011; Breiman and Friedman 1985). However, data fusion is inherently a task involving multiple datasets and the question is how desirable properties can be extracted by simultaneous use of these data. Canonical correlation analysis (CCA) answers this by maximizing the correlation between projections of the two datasets (Hotelling 1936). However, if for some reason, the two datasets are not well aligned (geometrically, temporally, or otherwise), the subspace projection produced by CCA is corrupted. Such a misalignment was observed in several cases between weather radar imagery and satellite imagery and a statistically based method correcting for this misalignment is proposed.

While describing the same physical scene, the radar imagery contains reflectance of water droplets and the satellite imagery describes reflectance of the cloud tops, for different wavelengths of electromagnetic radiation.

Corresponding author address: Jacob S. Vestergaard, Technical University of Denmark, Asmussens Allé, DTU Informatics, 2800 Kongens Lyngby, Denmark.
E-mail: jsve@imm.dtu.dk

DOI: 10.1175/JAMC-D-12-05.1

Hence, two different characteristics of the same physical scene are available.

The expected benefit of fusing weather radar and satellite data by use of CCA is to determine a subspace projection maximizing the correlation between the two datasets, whereby accentuating the cloud tops visible in the optical satellite data causing heavy rain. To ensure that the emphasis is on heavy rain, the analysis includes only the area of an isolated powerful squall.

Initial experiments revealed inconsistent subspaces and low correlations when using CCA to maximize correlations between radar data at time t_0 and satellite data at the same time. The observed misalignment is probably due to the different points of view of the two sensors; that is, the radar is situated on the ground, observing precipitation as it occurs below the cloud base, whereas the satellite is in a geostationary orbit and registering cloud tops. Furthermore, cloud tops of heavily precipitating clouds often have a larger spatial extent, due to the forming of an anvil, causing a lack of correspondence and further corrupting the produced subspace.

Therefore, a hypothesis emerged that the observed squall at time t_0 has greater correspondence with a similar phenomenon in the satellite data at time t_i, $i \leq 0$. This can be of importance when using satellite data as part of a nowcasting system. It was further hypothesized that the observed phenomenon in the two data sources is related by a geometric transformation.

The relations included in the hypothesis are rooted in the underlying physical situation. The geometric transformation allows for 1) stretching to account for the formation of an anvil and outflow and 2) rotation and translation to adjust for the nonuniform airflow in the embedded clouds and any misalignments due to the two sensors' different viewpoints. When heavy convection results in an anvil, the precipitating clouds are concealed from the satellite; wherefore, a time shift is included in the model to properly relate the images through the changes.

The method presented investigates—and verifies—this hypothesis and determines the temporal displacement and the optimal geometric transformation of the convective area in order to produce a subspace descriptive of heavy precipitation events. The benefits of an improved subspace are, for example, an easier segmentation of these types of clouds or a more compact (a single component) representation of the satellite data, which will be highly correlated with the information provided by the radar data.

This relies on the availability of radar imagery at time t_0, where a spatially isolated, heavy precipitation event is known to occur. The flow of the method is to move back in time through a discrete time series of satellite imagery, in each time step finding the geometric transformation $T^{(i)}$ of the area of interest that maximizes the correspondence with the satellite imagery at time t_i, $i \leq 0$, until further temporal displacement results in a decrease in the correspondence.

The cases analyzed are presented in section 2 together with a description of the available data sources. The method proposed is presented in detail in section 3 and an evaluation of the results is given in section 4.

2. Data

Three scenarios of extreme weather in Denmark are analyzed, all categorized by a meteorologist as thunderstorms in warm air masses and producing heavy precipitation. The three cases are treated and categorized in relation to Danish weather standards; thus, the extremity of the downpour should be seen in relation to this geographical region. A brief summary of each scenario will be given.

(i) 16 July 2007

A heavy rainfall hit Jutland during the night of 16–17 June, where thunderstorms developed in advance of a cold front moving in from the southwest. Downpour intensities above 50 mm h^{-1} were recorded with lightning frequencies of up to 50 flashes per minute.

(ii) 11 August 2007

Warm and moist air over the eastern part of Denmark and southern Sweden developed into thunderstorms with downpour intensities of up to 60 mm h^{-1}. Locally, cloudbursts of 2-h duration were reported, causing flooded roads in several locations.

(iii) 20 August 2007

A multicell convective system developed in the warm moist air coming from northern Germany and maturing over southwestern Denmark. Extreme downpour intensities of approximately 53 mm in only 10 min were recorded, causing damage to roads and train tracks, as well as local flooding. The extremity of this scenario brought the meteorological synoptic situation under scrutiny (Nielsen 2008).

For each scenario treated, the two data sources are separate time series of satellite data and radar data, respectively. The two time series have the same temporal extents, but not the same sampling frequencies. The specific properties and origin of the data sources are given below.

a. Satellite data

Multispectral image data from the *Second Meteosat Second Generation* (*MSG-2*, renamed *Meteosat-9*) weather satellite, launched in December 2005, are made available by the Danish Meteorological Institute (DMI). The Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board *MSG-2* has a spatial resolution of 3 km at nadir and a temporal resolution of 15 min. Twelve spectral channels (three VIS, one near-IR, and eight IR) are available with characteristics shown in Table 1. The reflectance has an interpretation as brightness temperatures (Müller 2010). The spatial resolution for latitudes corresponding to Denmark is increased to 6–8 km × 4 km, due to the satellite's perspective from the geostationary orbit.

Only the eight infrared channels are used here, as two of the three analyzed weather systems occur at nighttime, rendering the visible and near-IR wavelengths useless.

TABLE 1. Characteristics of channels in the multispectral satellite dataset from SEVIRI. [Table from Schmetz et al. (2002).]

Channel No.		Characteristics of spectral band (μm) λ_{cen}	λ_{min}	λ_{max}	Main gaseous absorber or window
1	VIS0.6	0.635	0.56	0.71	Window
2	VIS0.8	0.81	0.74	0.88	Window
3	NIR1.6	1.64	1.50	1.78	Window
4	IR3.9	3.90	3.48	4.36	Window
5	WV6.2	6.25	5.35	7.15	Water vapor
6	WV7.3	7.35	6.85	7.85	Water vapor
7	IR8.7	8.70	8.30	9.10	Window
8	IR9.7	9.66	9.38	9.94	Ozone
9	IR10.8	10.80	9.80	11.80	Window
10	IR12.0	12.00	11.00	13.00	Window
11	IR13.4	13.40	12.40	14.40	Carbon dioxide
12	HRV	Broadband (~0.4–1.1)			Window/ water vapor

b. Radar data

The radar imagery provided by DMI is a fusion of data from the institute's five operational weather radars in Denmark. The images have a ground sampling distance of 1 km and a temporal resolution of 10 min. It was found that an accumulation of the radar reflectance eased the task of identifying the convective area of interest; hence, the reflectance values referred to are actually the total reflectivity in each pixel within ±30 min of the sampling time; that is, the value of a given pixel in the radar imagery at time t_i is given as

$$\mathrm{dB}Z_i = 10 \times \log_{10}\left(\sum_{j=i-3}^{i+3} \frac{z_j}{1\,\mathrm{mm}^6\,\mathrm{m}^{-3}}\right), \tag{1}$$

where z_i is the radar reflectivity at time t_i. This accumulation makes the spatially isolated cloud easier to segment.

Prior to any analysis, the two data sources are projected to a common grid covering Denmark of size 400 × 500 pixels, where each pixel corresponds to 2 km × 2 km. Hence, a pixel-to-pixel correspondence is established.

A prerequisite for the following methodology is identification of a point in time of maximum precipitation. This time will be denoted as t_0 and will be used as a starting point for the algorithm.

3. Method

The multispectral satellite data will be considered as a set of $p = 8$ stochastic random variables and arranged in a matrix $\mathbf{X} \in \mathbb{R}^{N\times p}$. The radar data are considered as a single stochastic variable, that is, univariate, and arranged in a vector $\mathbf{y} \in \mathbb{R}^N$:

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_p^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Np} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}. \tag{2}$$

A superscript will be added to the notation to indicate the time when the satellite data were relevant; that is, $\mathbf{X}^{(i)}$ is then the satellite data matrix at time t_i. As only radar reflectance of a single squall at time t_0—and transformations of this—are used, this notation is not introduced for the radar data. Here, N is the number of pixels considered in the analysis and will be introduced later when examining the images.

a. Canonical correlation analysis

CCA is a method for analyzing the relations between two sets of variables. It was first introduced by Hotelling (1936) and is described in most textbooks on multivariate statistical analysis (see, e.g., Anderson 1984; Wackernagel 1995).

CCA maximizes the correlation ρ between linear combinations of two sets of variables (e.g., $\mathbf{X} \in \mathbb{R}^{N\times p}$ and $\mathbf{Y} \in \mathbb{R}^{N\times q}$):

$$\rho = \mathrm{corr}(\mathbf{X}\mathbf{a}, \mathbf{Y}\mathbf{b}) = \frac{\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{12}\mathbf{b}}{\sqrt{\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{11}\mathbf{a}}\sqrt{\mathbf{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}\mathbf{b}}}, \tag{3}$$

where $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$ are the dispersion matrices of $\mathbf{X}$ and $\mathbf{Y}$, respectively, and $\boldsymbol{\Sigma}_{12}$ is the covariance matrix of $\mathbf{X}$ and

$\mathbf{Y}$. This can be done by solving the generalized eigenvalue problem

$$\rho^2 = \frac{\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\mathbf{a}}{\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{11}\mathbf{a}} = \frac{\mathbf{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\Sigma}_{12}\mathbf{b}}{\mathbf{b}^{\mathrm{T}}\boldsymbol{\Sigma}_{22}\mathbf{b}}, \tag{4}$$

where the eigenvectors $\mathbf{a}_1, \ldots, \mathbf{a}_p$ with corresponding eigenvalues $\rho_1^2 \geq \cdots \geq \rho_p^2$ are the desired projection directions for $\mathbf{X}$.

Within this context, CCA will be used to assess which spectral bands in the satellite data are correlated with the radar data and thereby determine a subspace projection of the satellite data that maximizes the correlation with the radar data.

1) Simplification for univariate case

As the radar data are univariate, $q = 1 \Rightarrow \mathbf{Y} = \mathbf{y}$, the projection direction $\mathbf{a}$ can be shown to be

$$\mathbf{a} = \frac{1}{\rho}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\sigma}_{12}b = \frac{\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\sigma}_{12}}{\sqrt{\boldsymbol{\sigma}_{12}^{T}\boldsymbol{\Sigma}_{11}^{-1}\boldsymbol{\sigma}_{12}}}, \tag{5}$$

whereby solving an eigenvalue problem is avoided (Vestergaard 2011). Here, the covariance between $\mathbf{X}$ and $\mathbf{y}$ is denoted by $\boldsymbol{\sigma}_{12}$ to emphasize that this is a vector when $\mathbf{y}$ is univariate. The projection of $\mathbf{X}$ onto $\mathbf{a}$ is called the (first) canonical variate.

2) Visual inspection of eigenvectors

Because of the more complex covariance relations in CCA compared to, for example, principal component analysis, it is often not feasible to visually inspect the eigenvectors directly. Instead, the correlations between the original variables and the projected variables are considered. These variables can be calculated efficiently from the covariance matrices, which are already estimated to obtain the eigenvector. First, we recall the correlation between a single variable $\mathbf{x}_i$ and the canonical variate to be

$$\mathrm{corr}(\mathbf{x}_i, \mathbf{Xa}) = \frac{\mathrm{cov}(\mathbf{x}_i, \mathbf{Xa})}{\sigma_{x_i}\sqrt{\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{11}\mathbf{a}}},$$

where $\mathrm{cov}(\mathbf{x}_i, \mathbf{Xa})$ denotes the covariance between $\mathbf{x}_i$ and $\mathbf{Xa}$, and σ_{x_i} is the standard deviation of $\mathbf{x}_i$. For all p variables the correlation can be written as

$$\mathrm{corr}(\mathbf{X}, \mathbf{Xa}) = \mathbf{D}^{-1}\boldsymbol{\Sigma}_{11}\mathbf{a}(\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}_{11}\mathbf{a})^{-1/2}, \tag{6}$$

where $\mathbf{D}$ is a matrix with diagonal elements $D_{ii} = \sigma_{x_i}$ and zeros off-diagonal.

b. *Aligning with invariance to scale, rotation, and translation*

It is hypothesized that a conformal alignment of a squall isolated in the radar data can increase the correspondence with the satellite imagery. Such an alignment can be formulated as a transformation of the homogeneous image coordinates $\mathbf{z}_h = [x_i, y_i, 1]^{\mathrm{T}}$ using a rotation, a scaling, and a translation matrix:

1) counterclockwise rotation with the angle θ,

$$\mathbf{z}_r = \mathbf{R}\mathbf{z}_h = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix};$$

2) scaling with s_x in the x direction and s_y in the y direction,

$$\mathbf{z}_s = \mathbf{S}\mathbf{z}_h = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}; \quad \text{and}$$

3) translation with (t_x, t_y),

$$\mathbf{z}_t = \mathbf{T}\mathbf{z}_h = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}.$$

Note that x and y are used as image coordinates in this context.

Collecting the parameters in a single set and constraining to isotropic scaling (i.e., $s \equiv s_x = s_y$), the transformation $\mathcal{T}$ of the homogeneous coordinates $\mathbf{z}_h$, given the parameters $\boldsymbol{\theta} = [\theta, s, t_x, t_y]$, is

$$\hat{\mathbf{z}}_h = \mathcal{T}(\mathbf{z}_h \,|\, \boldsymbol{\theta}) = \mathbf{TSR}\mathbf{z}_h, \tag{7}$$

where we combine the three separate transformations into one by matrix multiplication. Nearest-neighbor interpolation is used to fill in missing values after the transformation. The number of observations N in the analysis will be the number of pixels having an intensity value after transformation of the squall; that is, it varies between different transformations.

The optimal transformation of the image coordinates, and thus the alignment of the squall, can be formulated as a nonlinear minimization problem:

$$\boldsymbol{\theta}_\star = \arg\min_{\boldsymbol{\theta}} \langle \mathcal{D}\{X, \mathbf{y}[\mathcal{T}(\mathbf{z}_h \,|\, \boldsymbol{\theta})]\}\rangle, \tag{8}$$

where $\mathcal{D}$ is some function describing the dissimilarity between the satellite imagery and the transformed radar data. Precisely which function to choose will be discussed below. For minimizing differences between pixel intensities, Eq. (8) can be formulated as a linear minimization problem (Eldén 2007, 122–128) but minimization of the correlation measures discussed below is a nonlinear optimization problem.

A bounded version of the simplex algorithm by D'Errico (2012) was used to solve the optimization problem in Eq. (8). This implementation allows for defining upper and lower bounds for the parameters, which is a desirable feature in this context. The bounds used were $\boldsymbol{\theta}_{\mathrm{LB}} = [-45°, 0.5, -20, -20]^{\mathrm{T}}$ and $\boldsymbol{\theta}_{\mathrm{UB}} = [45°, 1.5, 20, 20]^{\mathrm{T}}$, allowing for a rotation between $\pm 45°$, a scaling between 0.5 and 1.5 of the original size, and a translation of ± 20 pixels (40 km) in each direction. These are bounds on the optimization between time steps and not for the entire optimization.

DEFINING A DISSIMILARITY MEASURE

A requirement for the choice of the dissimilarity function is the capability of comparing a number of univariate observations (radar data) with the corresponding multispectral observations (satellite data) and evaluating to a scalar reflecting the difference between the two sets:

$$\mathcal{D}: (\mathbb{R}^{N\times p}, \mathbb{R}^{N}) \mapsto \mathbb{R}.$$

The two sets of observations are not measured on the same scale; wherefore, for instance, the simple absolute pixel-wise difference does not make sense. Two different applicable correlation measures are discussed below. Although correlation is a measure of similarity, a large negative correlation will equal large correspondence in this case, since low brightness temperatures in the satellite data will correspond to heavy rain (i.e., high reflectance in the radar data).

(i) Cross correlation

The cross-correlation function is defined under the hypothesis of second-order stationarity; that is, the mean and the autocovariance function do not depend on location, and are usually used for comparison of a spatially lagged variable with another variable (Wackernagel 1995). Here, the transformation is extended from a spatial lag—a translation—to also include rotation, scaling, and a temporal lag. For a single of the variables $\mathbf{x}_i$, the cross correlation with a transformation of $\mathbf{y}$ is

$$\rho_i = \mathrm{corr}\{\mathbf{x}_i, \mathbf{y}[\mathcal{T}(\mathbf{z}_h \mid \boldsymbol{\theta})]\} = \frac{\mathrm{cov}\{\mathbf{x}_i, \mathbf{y}[\mathcal{T}(\mathbf{z}_h \mid \boldsymbol{\theta})]\}}{\sigma_{\mathrm{xi}}\sigma_y},$$

where σ_{xi} is the standard deviation of $\mathbf{x}_i$ and σ_y is the standard deviation of $\mathbf{y}[\mathbf{T}(\mathbf{z}_h \mid \boldsymbol{\theta})]$. Thus, the vector of cross correlations is $\boldsymbol{\rho} = [\rho_1, \rho_2, \ldots, \rho_p]^{\mathrm{T}}$. To obtain a scalar expression, the dissimilarity in terms of cross correlation can either be calculated as an average of p correlations or as the correlation between the first principal component of $\mathbf{X}$ and the transformed radar data, where the latter was found to be appropriate within this context.

(ii) Multiple correlation coefficient

The multiple correlation coefficient (MCC) is a measure of correlation between a single variable from one set of variables $\mathbf{X}_1$ and the optimal linear combination of a second set of p variables $\mathbf{X}_2$. For the ith variable in $\mathbf{X}_1$, the MCC is defined as

$$\rho_{i|1,\ldots,p} = 1 - \frac{\det\boldsymbol{\Sigma}_i}{\sigma_i \det\boldsymbol{\Sigma}_{22}},$$

where $\boldsymbol{\Sigma}_i$ is defined as

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \sigma_i^2 & \boldsymbol{\sigma}_i^{\mathrm{T}} \\ \boldsymbol{\sigma}_i & \boldsymbol{\Sigma}_{22} \end{bmatrix},$$

with σ_i^2 being the variance of the ith variable, $\boldsymbol{\sigma}_i$ being the ith column in the covariance matrix of $\mathbf{X}_2$ and $\mathbf{X}_1$, and $\boldsymbol{\Sigma}_{22}$ being the dispersion matrix of $\mathbf{X}_2$.

The MCC has an obvious resemblance to the correlation achieved by a CCA when having a univariate first set of variables. Using the MCC as an optimization measure would therefore correspond to repeatedly performing a CCA for every possible transformation of the radar data and using the correlation from this analysis as an objective value. Given that we aim to improve the CCA subspace, this approach would seem reasonable. However, experiments show that using MCC as an objective value does not provide as much improvement in the subspace compared to using cross correlation, as described above. This is probably due to the need to estimate an optimal linear transformation while determining the correlation, causing a more fluctuant objective value and in turn a more difficult optimization problem. Thus, the cross-correlation measure is used as the dissimilarity measure in Eq. (8).

c. Determining temporal displacement

The method proposed for determining the temporal displacement will be outlined step by step in this section and presented in algorithmic form.

The first step is to select the convective area of interest in the radar imagery. In the cases treated here, it was found adequate to perform a simple thresholding and subsequently to select the largest connected component. The spatial coordinates corresponding to this area will be denoted $\mathbf{z}$, such that $\mathbf{y}(\mathbf{z})$ corresponds to a single isolated squall in the radar data at time t_0. We chose to use $\tau = 35$ dBZ as a lower threshold for this segmentation, which is a popular choice, suggested by, for example, Mecikalski et al. (2010).

The algorithm moves backward through time from t_0, corresponding to the time of maximum convection. For each temporal lag t_i, the correspondence between the satellite data $\mathbf{X}^{(i)}$ and the original radar data $\mathbf{y}$ is maximized using the alignment described in section 3b. The optimal set of parameters $\boldsymbol{\theta}_\star^{(i)}$ will be used as starting point for the optimization in the next step (i.e., for time t_{i-1}). This is based on the assumption that the optimal transformation at time t_{i-1} is close (in parametric space) to the optimal transformation at time t_i. This is intuitive since a deep convective cloud is expected to exhibit a smooth development in the satellite data.

The stopping criterion for the method is the demand for a decrease in dissimilarity; that is, if

$$\mathcal{D}\langle \mathbf{X}^{(i)}, \mathbf{y}\{\mathcal{T}[\mathbf{z}_h \,|\, \boldsymbol{\theta}_\star^{(i)}]\}\rangle > \mathcal{D}\langle \mathbf{X}^{(i+1)}, \mathbf{y}\{\mathcal{T}[\mathbf{z}_h \,|\, \boldsymbol{\theta}_\star^{(i+1)}]\}\rangle\,,$$

the algorithm is stopped and the optimal correspondence is achieved by a temporal displacement of $i + 1$ lags and an alignment with the parameter set $\boldsymbol{\theta}_\star^{(i+1)}$. Hence, one step past the optimal displacement is required to realize that this is the solution. The method is formulated as pseudocode in algorithm 1 in the appendix.

We propose using this alignment method prior to analyzing the two datasets using CCA. To summarize, rather than the simpler

$$\mathrm{corr}[\mathbf{X}^{(0)}\mathbf{a}, \mathbf{y}]\,, \tag{9}$$

we will use the expression

$$\mathrm{corr}\{\mathbf{X}^{(i^\star)}\mathbf{a}, \mathbf{y}[\mathcal{T}(\mathbf{z}_h \,|\, \boldsymbol{\theta}_\star^{i^\star})]\}\,, \tag{10}$$

which is maximized.

In the following section, a comparison will be given of subspaces yielded by CCA with and without aligning the data using the proposed method.

4. Results

The three scenarios presented in section 2 are all treated with the proposed method. For comparison, they are also analyzed using CCA without any alignment at all. The results will be compared by (i) visually inspecting the first canonical variate (i.e., the satellite imagery projected onto the direction determined by CCA) and (ii) inspecting the correlation structure of the original spectral bands with this projection, calculated from Eq. (6).

To properly interpret the visualizations of the canonical variates, one should consider the contrast between

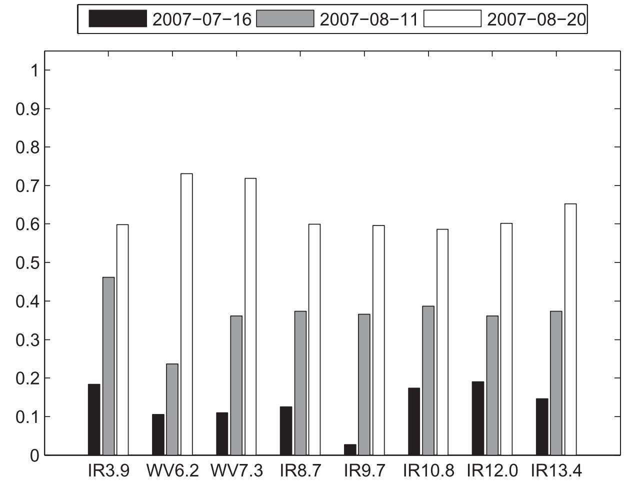

FIG. 1. Correlation structure from CCA without using the proposed alignment method.

clouds carrying heavy precipitation and other clouds. A large contrast, accentuating the heavy precipitating clouds, would be expected from a descriptive subspace. The actual scale on the image intensities is not important. For the three cases treated here, the areas of heavy precipitation are known from weather reports.

The first result presented is the one originally driving the hypothesis that an initial alignment of the two datasets is needed: the convective area of interest is segmented in the radar data at time t_0 and the canonical correlation analysis is performed between this area and the satellite data also at time t_0. In Fig. 1 a bar plot of the correlations between the canonical variates and the original data is shown. Performing a CCA on three separate scenarios representing the same meteorological phenomenon, it is expected that the correlation structures of the resulting subspaces from each of these separate analyses are comparable and that each of the subspaces are descriptive of a precipitating cloud. However, as Fig. 1 shows, the correlation structures are very different between scenarios, and the first canonical variates (Figs. 2–4) do not provide a satisfactory visualization of data accentuating precipitating clouds; the cloud tops are not distinguishable from their surroundings.

The proposed method is applied to the three scenarios and the results are summarized in Table 2. All transformation parameters listed are with respect to the radar data at time t_0. For the first scenario (16 August 2007), the optimal transformation is found as a rotation of approximately 14°, with no change in scale, and a translation of (−6.5, 19.4) pixels, corresponding to (13.0, 38.8) km to the west and north. No temporal displacement was found necessary for this scenario as $t_{i*} = t_0$. The transformation for the second scenario (11 August

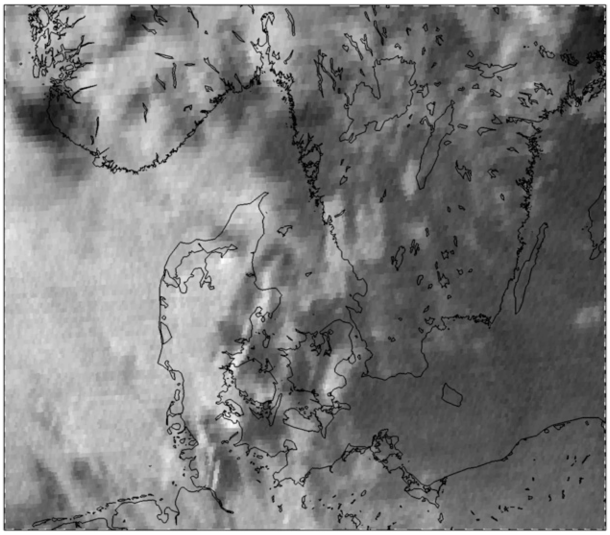

FIG. 2. The first canonical variate at time t_0 on 16 Jul 2007. No prior alignment of radar data is performed. Image intensities stretched to mean ± 3 standard deviations.

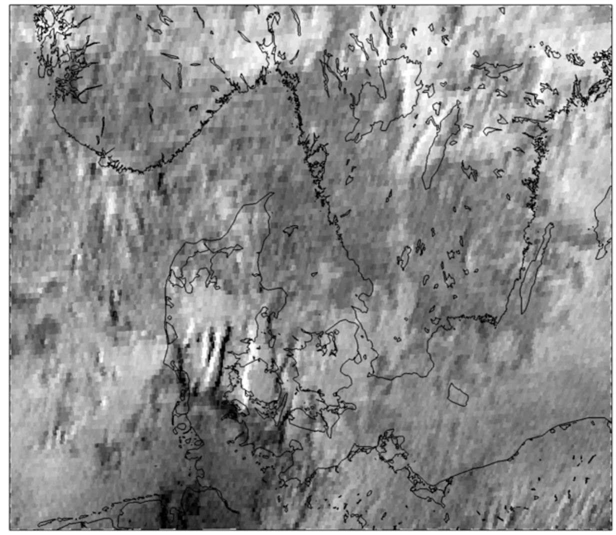

FIG. 4. As in Fig. 2, but for 20 Aug 2007.

2007) is found to be optimal for a temporal displacement of one lag and a transformation containing primarily a rotation (−28.6°) and a translation of (14.8, −2.6) pixels (i.e., to the east). The optimization yields a final cross correlation of −0.298, which is seen to be significantly smaller in magnitude than that of the two other scenarios (−0.575 and −0.728). Here, it is worth remembering that a large negative value is preferred, as high values in the radar data and low values in the satellite data are associated with heavy rain.

The final scenario (20 August 2007) exhibits the largest transformation, as a temporal displacement of four time lags, corresponding to 40 min, is found optimal. From the chosen parameters in each lag, it is seen that an approximately 45° rotation is found to be optimal together with a displacement to the north. A small increment in scale to $s = 1.1$ is found to be appropriate at time t_0, decreasing from there. This is due to the rapid development of the cloud-top area over this period: at time t_0 the cloud base, causing precipitation, is much smaller than the associated cloud top. When going back in time from there, the cloud top decreases in area, and at time t_{-4} the best correspondence with the precipitation in the radar data at time t_0 is found, using a scale of 0.6.

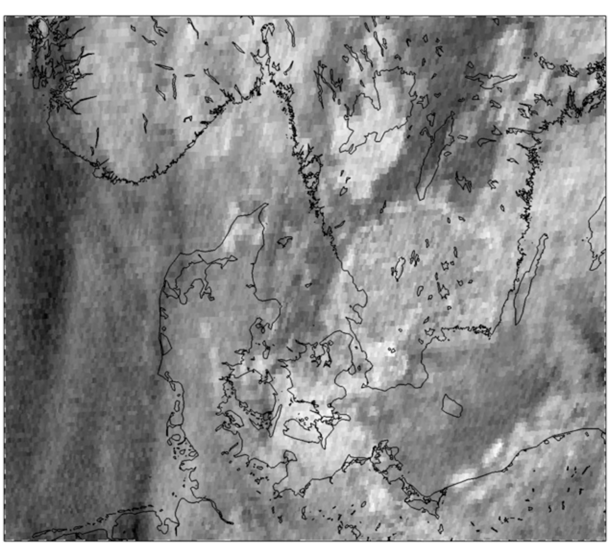

FIG. 3. As in Fig. 2, but for 11 Aug 2007.

Correlation structures for the three scenarios, when the proposed alignment method is used, are shown in Fig. 5. Comparing with Fig. 1, the difference is apparent: the three scenarios' correlations now exhibit similar structures and have in general increased correlations. This supports the hypothesis that three similar meteorological situations should also show similarities when

TABLE 2. Optimal parameters and resulting function value (cross correlation) for each step back in time for each scenario. The optimal temporal displacement is the final lag displayed.

Scenario	Lag	θ	s	t_x	t_y	f
16 Jul 2007	t_0	14.1	1.0	−6.5	19.4	−0.575
11 Aug 2007	t_0	−39.7	1.1	19.3	−2.2	−0.292
	t_{-1}	−28.6	1.0	14.8	−2.6	−0.298
20 Aug 2007	t_0	44.6	1.1	−0.9	11.5	−0.579
	t_{-1}	44.8	1.0	−2.1	10.5	−0.584
	t_{-2}	44.9	0.8	0.6	13.0	−0.608
	t_{-3}	43.6	0.6	3.0	16.9	−0.616
	t_{-4}	44.7	0.6	2.8	16.2	−0.728

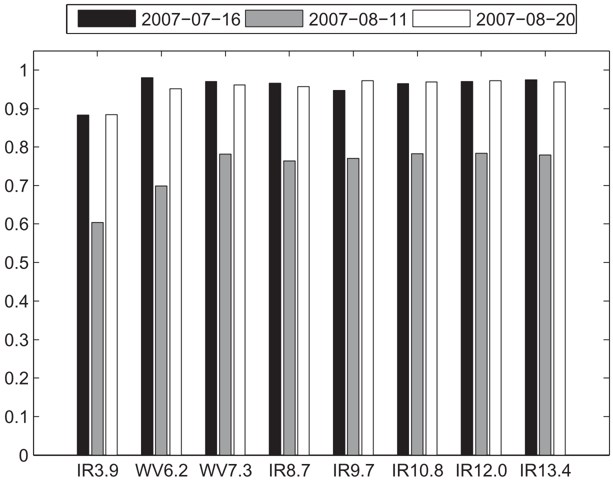

FIG. 5. Correlation structure from CCA when applying the proposed alignment method.

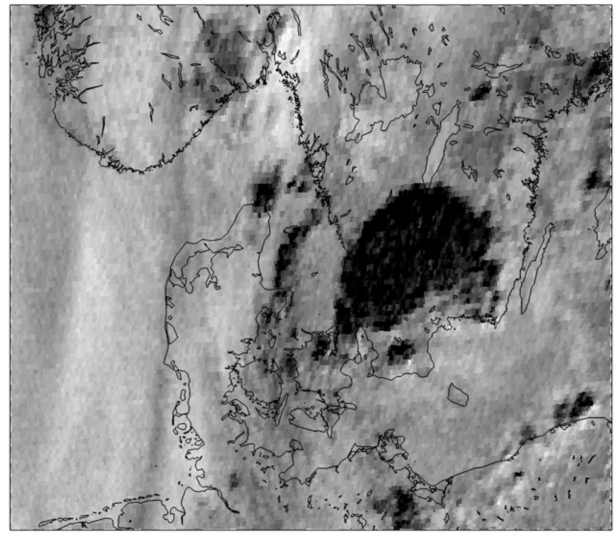

FIG. 7. As in Fig. 6, but for 11 Aug 2007.

analyzing them using CCA. The average correlation over all bands and all scenarios has increased from 0.38 to 0.89 by performing the prealignment, providing a quantitative evaluation of the method.

The individual analyses yield the canonical variates shown in Figs. 6–8. The overwhelming improvement caused by the alignment method is clearly visible: the information contained in the satellite data having correspondence with the radar data is now dominant, making the precipitating cloud of interest much easier to distinguish from the background. Even the scenario with the smallest final objective value (11 August 2007) has been significantly improved (cf. Figs. 3 and 7). These results show that the transformation is of paramount importance.

It is worth noting that an improved subspace could also be achieved without allowing for a temporal displacement. This can be of importance if applying the method to scenarios, where the cloud of interest is less isolated from surrounding heavy precipitation. The method is sensitive to such cases, as it can possibly ‘‘switch’’ to another cloud in the satellite data, if a higher correspondence can be achieved.

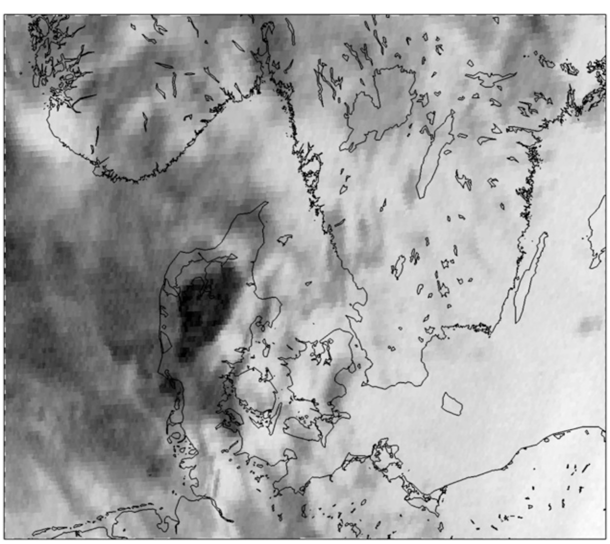

FIG. 6. The first canonical variate at time t_0 on 16 Jul 2007. Prior to CCA, the radar data have been aligned using the proposed method. Image intensities are stretched to mean ± 3 standard deviations.

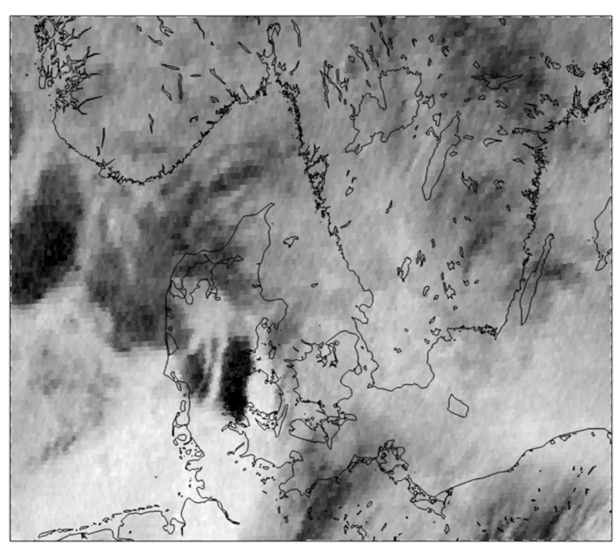

FIG. 8. As in Fig. 6, but for 20 Aug 2007.

5. Conclusions

A method has been proposed for the optimal alignment of an isolated squall, observed in weather radar imagery, to maximize correspondence—in terms of correlation—with satellite imagery. The algorithm determines the conformal transformation and temporal displacement by consecutively solving a nonlinear optimization problem for an increasing temporal lag between the two data sources. When no further increase in correspondence can be achieved, the algorithm is stopped.

The improvement in subspace yielded by canonical correlation analysis is illustrated by analysis of three cases of severe precipitation in Denmark. A quantitative evaluation was conducted, comparing the correlations achieved by CCA with and without using the proposed prealignment. An increase in average correlation over all three cases from 0.38 to 0.89 was achieved. A visual inspection of the canonical variates clearly shows the benefits of a prealignment compared to not performing the prealignment.

Provided that these three scenarios are good representatives of high-intensity precipitation in Denmark, the linear combination of the eight IR bands from *MSG-2* can be used as a fixed transformation, emphasizing heavy rain in a single component. This can be of importance for future classification or forecasting applications, for example, exploiting the extended coverage of *MSG-2* to create a surrogate for weather radars, where they cannot reach. However, a larger dataset with more cases would be needed to ensure proper validation of such a method.

Acknowledgments. We thank Dr. Thomas Bøvith and aviation meteorologist Birgitte Knudsen from the Danish Meteorological Institute (DMI) for providing the data and categorizing the events. We also thank Professor Rasmus Larsen from DTU Informatics for insightful discussions.

APPENDIX

Algorithms

Algorithm 1 maximizes correspondence with satellite data by a conformal transformation and a temporal displacement of radar data:

$\boldsymbol{\theta}_0 = (0, 1, 0, 0)$
$\mathbf{y} \leftarrow$ radar data at time t_0
$i = 1$, stop = false
while $d^{(i)} < d^{(i+1)} \vee i > -1$ do
  $i = i - 1$
  $\mathsf{X} \leftarrow$ satellite data at time t_i
  $\boldsymbol{\theta}_\star^{(i)} = \arg\min_{\boldsymbol{\theta}}\{\mathcal{D}[\mathsf{X}, \mathcal{T}(\mathbf{y} \mid \boldsymbol{\theta})]\}$ ($\boldsymbol{\theta}_0$ as starting point)
  $d^{(i)} = \mathcal{D}[\mathsf{X}, \mathsf{T}(\mathbf{y} \mid \boldsymbol{\theta}_\star)]$
  $\boldsymbol{\theta}_0 = \boldsymbol{\theta}_\star^{(i)}$
end while
return $i^\star = i + 1, \boldsymbol{\theta}_\star^{i^\star} = \boldsymbol{\theta}_\star^{(i+1)}$

REFERENCES

Anderson, T. W., 1984: *An Introduction to Multivariate Statistical Analysis*. 2nd ed. Wiley-Interscience, 686 pp.

Bell, A., and T. Sejnowski, 1995: An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.,* **7,** 1129–1159.

Breiman, L., and J. H. Friedman, 1985: Estimating optimal transformations for multiple regression and correlation. *J. Amer. Stat. Assoc.,* **80,** 580–598.

D'Errico, J., cited 2012: Bound constrained optimization using fminsearchbnd. MathWorks. [Available online at http://www.mathworks.com/matlabcentral/fileexchange/8277-fminsearchbnd.]

Eldén, L., 2007: *Matrix Methods in Data Mining and Pattern Recognition*. Fundamentals of Algorithms, Society for Industrial and Applied Mathematics, 224 pp.

Green, A. A., M. Berman, P. Switzer, and M. Craig, 1988: A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. Remote Sens.,* **26,** 65–74.

Hotelling, H., 1936: Relations between two sets of variates. *Biometrika,* **28,** 321–377.

Jolliffe, I. T., 2002: *Principal Component Analysis*. 2nd ed. Springer, 487 pp.

Mecikalski, J. R., W. M. MacKenzie, M. Koenig, and S. Muller, 2010: Cloud-top properties of growing cumulus prior to convective initiation as measured by Meteosat Second Generation. Part I: Infrared fields. *J. Appl. Meteor. Climatol.,* **49,** 521–534.

Müller, J., 2010: MSG level 1.5 image data format description. EUMETSAT Doc. EUM/MSG/ICD/105, 127 pp. [Available online at http://www.eumetsat.int/groups/ops/documents/document/pdf_ten_05105_msg_img_data.pdf.]

Nielsen, A. A., 2011: Kernel maximum autocorrelation factor and minimum noise fraction transformations. *IEEE Trans. Image Process.,* **20,** 612–624.

Nielsen, N. W., 2008: Et regnskyl af dimensioner: Skybruddet ved Gråsten i Sønderjylland den 20. august 2007 (A severe downpour: The cloudburst at Gråsten in Southern Jutland on August 20, 2007). *Vejret,* **114** (1).

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.,* **83,** 977–992.

Schölkopf, B., A. Smola, and K.-R. Müller, 1998: Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.,* **10,** 1299–1319.

Schowengerdt, R., 2007: *Remote Sensing: Models and Methods for Image Processing*. 3rd ed. Academic Press, 515 pp.

Vestergaard, J. S., 2011: Improved nowcasting of heavy precipitation using satellite and weather radar data. M.Sc. thesis, DTU Informatics, Technical University of Denmark, 177 pp.

Wackernagel, H., 1995: *Multivariate Geostatistics: An Introduction with Applications.* Springer, 256 pp.