

## Correcting Position Error in Precipitation Data Using Image Morphing

le Coz, Camille; Heemink, Arnold; Verlaan, Martin; ten Veldhuis, Marie-claire; van de Giesen, Nick

**DOI**

[10.3390/rs11212557](https://doi.org/10.3390/rs11212557)

**Publication date**

2019

**Document Version**

Final published version

**Published in**

Remote Sensing

**Citation (APA)**

le Coz, C., Heemink, A., Verlaan, M., ten Veldhuis, M., & van de Giesen, N. (2019). Correcting Position Error in Precipitation Data Using Image Morphing. *Remote Sensing*, 11(21), 1-27. Article 2557. <https://doi.org/10.3390/rs11212557>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**






Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

## Article

# Correcting Position Error in Precipitation Data Using Image Morphing

Camille Le Coz <sup>1,\*</sup>, Arnold Heemink <sup>2</sup>, Martin Verlaan <sup>2,3</sup> and Marie-claire ten Veldhuis <sup>1</sup>  
and Nick van de Giesen <sup>1</sup>

<sup>1</sup> Faculty of Civil Engineering and Geosciences, Water Resources, Delft University of Technology, 2628 CN Delft, The Netherlands; J.A.E.tenVeldhuis@tudelft.nl (M.-c.t.V.); N.C.vandeGiesen@tudelft.nl (N.v.d.G.)

<sup>2</sup> Faculty of Electrical Engineering, Mathematics and Computer Science, TU Delft, 2628 CN Delft, The Netherlands; A.W.Heemink@tudelft.nl (A.H.); Martin.Verlaan@deltares.nl (M.V.)

<sup>3</sup> Deltares, 2600 MH Delft, The Netherlands

\* Correspondence: c.m.l.lecoz-1@tudelft.nl

Received: 24 September 2019; Accepted: 28 October 2019; Published: 31 October 2019



**Abstract:** Rainfall estimates based on satellite data are subject to errors in the position of the rainfall events in addition to errors in their intensity. This is especially true for localized rainfall events such as the convective rainstorms that occur during the monsoon season in sub-Saharan Africa. Many satellite-based estimates use gauge information for bias correction. However, bias adjustment methods do not correct the position errors explicitly. We propose to gauge-adjust satellite-based estimates with respect to the position using a morphing method. Image morphing transforms an image, in our case a rainfall field, into another one, by applying a spatial transformation. A benefit of this approach is that it can take both the position and the intensity of a rain event into account. Its potential is investigated with two case studies. In the first case, the rain events are synthetic, represented by elliptic shapes, while the second case uses real data from a rainfall event occurring during the monsoon season in southern Ghana. In the second case, the satellite-based estimate IMERG-Late (Integrated Multi-Satellite Retrievals for GPM) is adjusted to gauge data from the Trans-African Hydro-Meteorological Observatory (TAHMO) network. The results show that the position errors can be corrected, while preserving the higher spatial variability of the satellite-based estimate.

**Keywords:** precipitation estimation; satellite-based precipitation; gauge data; IMERG; TAHMO; morphing; field displacement

## 1. Introduction

Precipitation is an important variable in weather and climate research and many other applications. Precipitation data are needed as input for hydrological models, for flood and drought monitoring or for water management in agriculture or power generation. However, estimating precipitation accurately is difficult because of its high spatial and temporal variability. This is especially true for sub-Saharan Africa, where most of the rainfall is produced during the monsoon season by convective rainstorms, which are very localized [1,2].

Rain-gauges are the most direct way to measure precipitation. However, the gauge networks in Africa are not dense enough to derive high resolution precipitation estimates. Indeed, the rain-gauge distribution is sparse in many African regions and their number has been decreasing in recent decades [3]. During the same period, many efforts have been made to derive precipitation estimates from satellite data. Satellites do not measure precipitation directly but have the advantage of covering

large areas. This is especially interesting for Africa where gauge networks are sparse and there are also almost no radar observations available.

There is an increasing number of satellite-based rainfall products, providing rainfall estimates at different spatial and temporal resolutions. Most rainfall products use additional sources of data, such as gauge estimates, for bias correction. Bias correction methods focus on correcting the intensity. However, the intensity is not the only possible error in precipitation. Rainfall events are coherent moving systems and, in the case of convective rainstorms, they are also very localized. This can lead to errors in the estimation of the position and shape of the rain events beside the errors in their intensity. For some applications, such as hydrological modeling [4,5], flash flood warnings [6] or data assimilation in a numerical weather model [7,8], detecting the correct location of the rain events can be as important as their intensity.

The position errors in weather forecast models, including precipitation, have been taken into account in the field of forecast verification. Several spatial verification approaches have been developed [9,10]. They can be divided into four categories: neighborhood, scale-decomposition (e.g., References [11–13]), object- (or feature-)based (e.g., References [14–16]) and field deformation. In this study, we focus on a method belonging to the latter category. We now give an overview of field deformation method used for weather-related variables. Field deformation methods are based on a spatial mapping or displacement that makes a field (e.g., forecast) more similar to a target field or observation. The deformation is determined by minimizing a cost function. The Feature Calibration and Alignment technique (FCA [17–19]) is one of these methods. FCA has also been used for correcting position errors in cloud or water vapor related fields in the framework of data assimilation. For instance, References [18,20] corrected position error in a numerical weather model background fields using integrated water vapor measurements from satellite. In Reference [21], the FCA is used as a preprocessing step of an ensemble-based variational assimilation scheme for (satellite) brightness temperature. Reference [22] tested this method with several types of observations—integrated water vapor, lower level pressure, brightness temperature and simulated radar reflection. Other feature alignment techniques have been developed and used in data assimilation schemes, such as Reference [23] (for simulated radar observation), References [24,25] (for some idealized cases). The FCA technique has been applied directly to rainfall data in Reference [19]. They corrected rainfall estimates derived from SSM/I data with ground-based radar estimates. They illustrated the performance of their approach for different types of rainfall events, such as Hurricane Andrew, a squall line in Oklahoma and coastal rainfall in Australia.

Some field deformation methods for spatial verification originate from image processing, such as the optical flow techniques developed in Reference [26,27] or in Reference [28] and evaluated in References [29,30]. Image warping has also been used in data assimilation frameworks. Reference [31] assimilated integrated water vapor from satellite to improve a numerical weather model forecast. However, this method requires the manual selection of pairs of points to perform the image warping. Reference [32] combined image morphing with an ensemble Kalman filter for a wild fire model. They use an automatic registration technique that only requires two fields to derive the displacement field, without any manual specification needed. Using the same morphing and registration method, a morphing fast Fourier transform (FFT) EnKF for radar precipitation is described in Reference [33]. However, this morphing FFT EnKF is not implemented and applied to rainfall data.

This present study investigates the use of the morphing approach for the position correction of rainfall estimates, using the approach proposed by References [32,33]. While the goal of Reference [33] was to derive a method to assimilate radar precipitation into a numerical weather model, we aim to correct the position error of satellite-based precipitation estimates using gauge measurements. We apply the morphing approach to real precipitation data, namely the (non-gauge adjusted) IMERG-Late estimates and the new Trans-African Hydro-Meteorological Observatory (TAHMO) gauge network.

The morphing and automatic registration methods, including the case of irregularly spaced observations, are described in Section 2. The morphing approach is applied to two cases. The first case uses synthetic rainfall events represented by ellipses (Section 3.1). The second case is a real rainfall event occurring in southern Ghana during the monsoon season (Section 3.2). Both the convergence of the automatic registration and the performance of the warping are examined in Section 4. The results of the two cases are compared and discussed in Section 5, before the conclusion in Section 6.

## 2. Methodology

In this section, we define the image registration and morphing processes, before focusing on the implementation of an automatic registration procedure. We use the framework described by Reference [33].

### 2.1. Definitions

Let  $u$  and  $v$  be two signals (or images) defined on a domain  $D \subset \mathbb{R}^2$  and  $T : (x, y) \in \mathbb{R}^2 \mapsto (T_x(x, y), T_y(x, y)) \in \mathbb{R}^2$  be a mapping function. The goal of **image registration** is to determine a spatial mapping  $T$  such that,  $\forall (x, y) \in D$ ,

$$\begin{aligned} v(x, y) &\approx u \circ (I + T)(x, y) \\ &\approx u[(I + T)(x, y)] \\ &\approx u[x + T_x(x, y), y + T_y(x, y)] \end{aligned} \quad (1)$$

where  $I$  is the identity function.

There can be several mappings  $T$  that meet the requirement  $v \approx u \circ (I + T)$ . Especially in areas without rainfall, the mapping  $T$  is not unique. We define three criteria to characterize one optimal mapping:

$$T \approx 0 \quad (2)$$

$$\nabla T \approx 0 \quad (3)$$

$$\nabla \cdot T \approx 0 \quad (4)$$

That is, the optimal mapping has to be as small, smooth and divergent-free (i.e., it is not shrinking or expanding the field) as possible.

Several approaches have been used to define the optimality of the mapping. For the FCA method applied to precipitation, Reference [19] use smoothness and barrier conditions. Contrary to our condition on the magnitude (Equation (2)), their barrier does not impact small scale displacements. Using the FCA for data assimilation, References [20,22] added two more constraints, one on the magnitude and one on the divergent. Reference [25] did not use any magnitude or barrier approach and only had constraints on the gradient and the divergence. Our constraints on the magnitude and on the smoothness are the same as those used in Reference [32]. Constraints on the divergence were used in several similar field distortion methods [20,22,25]. Thus, we also added a third constraint on the divergence in order to observe its impact. A short sensitivity study on the impact of these three coefficient is presented in Appendix B.

**Image warping** is the distortion of an image based on a spatial transformation of the domain. Warping can be used to transform an image into another one by using the spatial mapping  $T$  obtained from the registration method. The mapping  $T$  is gradually applied to the original image  $u$  as follows:

$$u_{\text{warp}(\lambda)} = u \circ (I + \lambda T) \quad 0 \leq \lambda \leq 1 \quad (5)$$

Warping works well when the residual  $v - u \circ (I + T)$  is small, which is not the case when the images  $u$  and  $v$  have different intensities for example. It is a spatial transformation. It only acts on the



coordinates, it does not modify the intensity of the image  $u$ . On the other hand, **Cross-dissolving** only acts on the intensity. It fades two images  $u$  and  $v$  into each other:

$$u_{\text{diss}(\lambda)} = u + \lambda(v - u) \quad 0 \leq \lambda \leq 1 \quad (6)$$

**Image morphing** combines warping and cross-dissolving to account for both the spatial distortion and the difference in intensity:

$$u_{\text{morph}(\lambda)} = (u + \lambda r) \circ (I + \lambda T) \quad 0 \leq \lambda \leq 1 \quad (7)$$

where  $r$  is the residual:

$$r = v \circ (I + \lambda T)^{-1} - u \quad 0 \leq \lambda \leq 1 \quad (8)$$

With this formula of  $u_{\lambda}$ , we obtain  $u_{\text{morph}(0)} = u$  and  $u_{\text{morph}(1)} = v$ .

## 2.2. Automatic Registration

The spatial mapping  $T$  used for the image morphing is determined by the image registration. Several registration methods are available. However, many of them require to define manually a set of corresponding points from the images  $u$  and  $v$ . We are interested in an automatic registration procedure that only needs the images  $u$  and  $v$  as inputs without any extra specifications. This requires the images to be similar enough for the automatic registration procedure to work.

We use the method described by Reference [33] based on the minimization of a cost function  $J$  with respect to the mapping  $T$ . The cost function can be divided in two terms (Equation (9)). The first one ( $J_o$ ) represents the mapping error between the displaced original signal  $u \circ (I + T)$  and the target signal  $v$ . The second one ( $J_b$ ) is a background term that consists of the three criteria for ‘optimal’ mapping given in Equations (2)–(4). These three criteria are used as weak constraints.

$$\begin{aligned} J(T) &= J_o(T) + J_b(T) \\ J_o(T) &= \|v - u \circ (I + T)\| \\ J_b(T) &= C_1 \|T\| + C_2 \|\nabla T\| + C_3 \|\nabla \cdot T\| \end{aligned} \quad (9)$$

where  $C_1$ ,  $C_2$  and  $C_3$  are three coefficients determined empirically and  $\|\cdot\|$  is the  $L^2$ -norm.

The minimization problem is solved iteratively, for  $T$  defined on increasingly fine grids. The iterative approach has two advantages. It helps reduce the computational cost and avoids the local minima problem (see below).

In our application, the domain  $D$  is rectangular. It can be represented by different uniform grids. The regular  $n_x \times n_y = n$  grid on which  $u$ ,  $v$  and  $u_{\text{morph}(\lambda)}$  are given is called the *pixel grid*  $D_n$ . The mapping function  $T$  is defined on a set of coarser grids  $D_i$  ( $i = 1, \dots, I$ ), called *morphing grids*. It is then represented by two gridded arrays (one for  $T_x$  and one for  $T_y$ ). The grids  $D_i$  are uniform  $(2^i + 1) \times (2^i + 1) = m_i$  grids (for  $i = 1, \dots, I$ ) covering the domain  $D$ . For  $i = 1, \dots, I$ , the mapping  $T$  discretized on  $D_i$  is noted  $T_i$ .

The signals  $u$  and  $v$ , and so the observation term  $J_o$  of the cost function, are discretized on the pixel grid  $D_n$ . The background term  $J_b$  is discretized on the morphing grid  $D_i$ . We use the second order central scheme except at the boundaries where the first order backward or forward schemes are used. We use bilinear interpolation to estimate the value of  $u$  and  $v$  on the distorted grid (e.g.,  $u \circ (I + T)$ ) and to interpolate  $T$  on the different morphing grids  $D_i$ .

The finest morphing grid  $D_I$  does not need to be the same as the pixel grid  $D_n$ . On the contrary, it is computationally advantageous when the morphing grid  $D_I$  has a much coarser resolution. When the number of nodes  $m_i$  of the morphing grids is much smaller than the number of nodes  $n$  of the pixel grid,

solving the minimization problem on the set of morphing grids  $D_i$  is less computationally expensive than to solve it for  $T$  defined on the high resolution pixel grid  $D_n$ .

### Algorithm

The algorithm iterates over the morphing grids  $D_i$  ( $i = 1, \dots, I$ ), starting on the coarsest 3-by3 grid  $D_1$ , until it reaches the finest morphing grid  $D_I$ . For each iteration, the three main steps are similar to those in Reference [32] and are illustrated in Figure 1.

1. **Smoothing** of the images  $u$  and  $v$ : the images are smoothed by convolution with a 2D-Gaussian

$$G_{2D}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right),$$

where  $\sigma = 0.05 / (2^{2i} + 1)$ . The finer the grid  $D_i$  is, the narrower the Gaussian is. Thus, for small  $i$ , the fine features are ignored and the focus is given to the large-scale ones. When  $i$  increases, more and more fine features are taken into account. This way,  $T_i$  for small  $i$  will make the larger features match. Then, for increasing  $i$ , more and more detailed images are matched.

The cost function  $J$  is often non-convex with respect to  $T$  and so can have several local minima. The smoothing combined with the hierarchy of grids reduce the local minima problem. They ensure that the large-scale features are fitting first, hence avoiding local minima.

After the smoothing, the two fields are normalized such that their maximum is the same. The images obtained after smoothing and normalization are noted as  $\tilde{u}_i$  and  $\tilde{v}_i$ .

2. **Initialization**: solving the minimization problem on grid  $D_i$  requires a first guess  $T_i^{\text{fg}}$ . For  $i = 1$ ,  $T_1^{\text{fg}}$  is set to zeros, that is, no deformation. For  $i = 2, \dots, I$ , the mapping  $T_{i-1}^*$  obtained by solving the minimization problem on grid  $D_{i-1}$  is interpolated into the grid  $D_i$  and used as the first guess  $T_i^{\text{fg}}$ .
3. **Optimization**: The actual minimization problem to be solved is based on the smoothed fields, that is,  $J_o(T) = \|\tilde{v}_i - \tilde{u}_i \circ (I + T)\|$ . Contrary to Reference [32], we solved the minimization problem for all the nodes at the same time.

There is a number of inequality constraints on this minimization problem, due to our requirements of invertibility. An iterative barrier approach is used to transform this constrained minimization problem into an unconstrained one [34,35]. In the barrier approach, the minimization is applied to a penalized cost function  $J_p(T) = J(T) + \beta \sum_h C_h(T)$ , where  $C_h$  are the constraint functions and  $\beta$  the barrier coefficient (over which we iterate when the constraints are not respected). The constraints and the minimization method are described with more details in Appendix A.

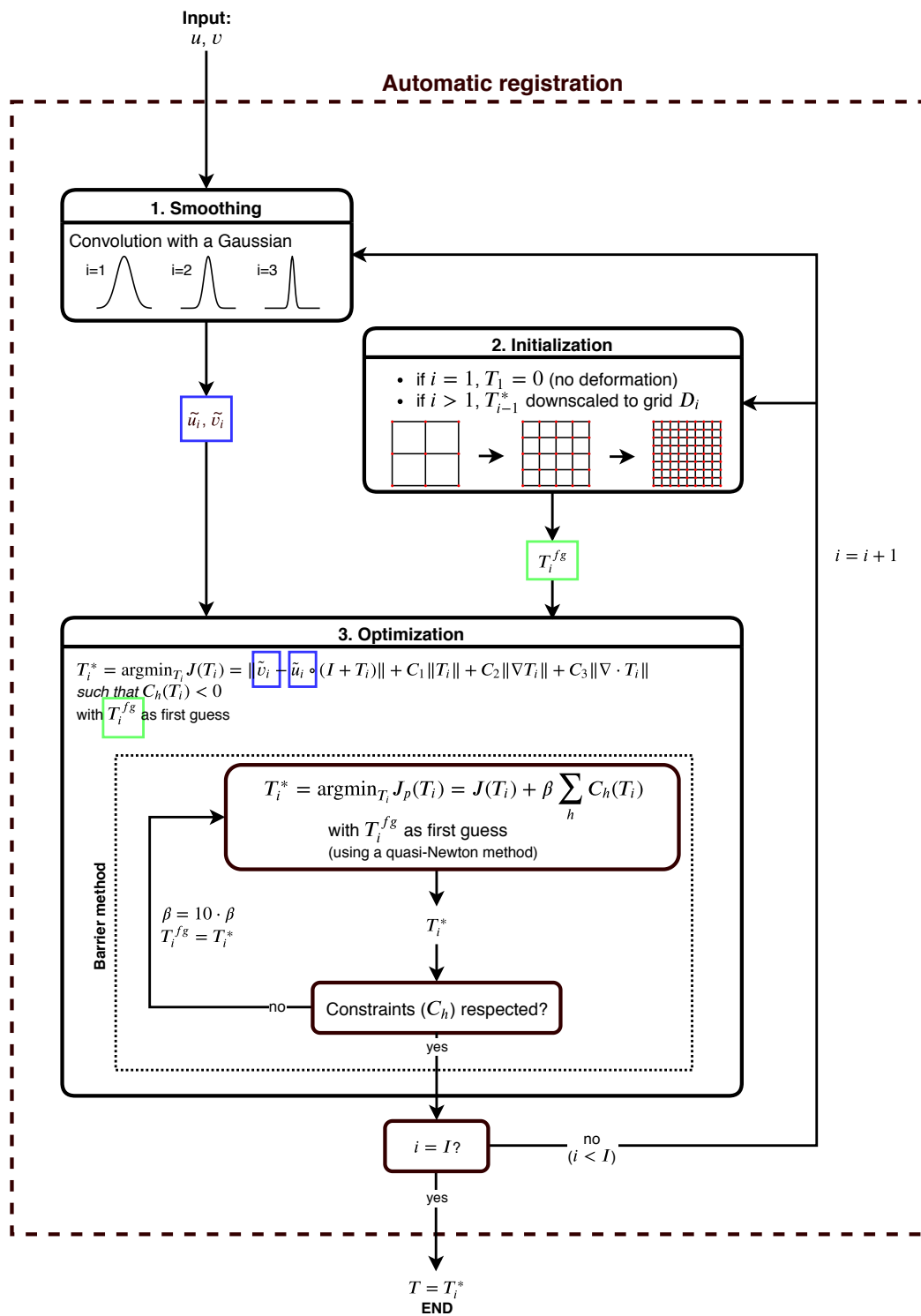


Figure 1. Algorithm for the automatic registration.

The python scripts for the automatic registration and the morphing are available online (Supplementary <https://github.com/clecoz/precipitation-morphing.git>). The scripts permit to reproduce the results for the synthetic case described below.

### 2.3. Dealing with Irregularly Spaced Observations

The automatic registration algorithm described above assumes that both signals  $u$  and  $v$  are on the same regular grid. However, in practice, one might deal with irregularly spaced observations, such as rain-gauge data.

In such a case, the observations are interpolated on the same regular grid, using kriging (details about the kriging are given in Section 3.2). In the remainder of the article, we will refer to the gauge interpolation as “kriging”, while “interpolation” will refer to the bi-linear interpolation used in the automatic registration and morphing. The cost function  $J$  from Equation (9) is modified to take the unequal coverage of the domain into account. A mask function  $M$  is added in the first term of  $J$ :

$$J(T) = \|M \cdot (v - u \circ (I + T))\| + C_1 \|T\| + C_2 \|\nabla T\| + C_3 \|\nabla \cdot T\| \quad (10)$$

where  $\cdot$  is the element-wise matrix multiplication. The mask function is defined such that it is equal to 1 in a given perimeter around the observations and zero everywhere else. So, the difference  $v - u \circ (I + T)$  for the grid points far from any observation does not weigh in the cost function  $J$ .

## 3. Study Cases

### 3.1. Synthetic Case for Algorithm's Validation

A synthetic case is used to investigate the convergence of the automatic registration algorithm and to validate the morphing. The synthetic precipitation fields are generated on a regular  $n_x \times n_y$  grid, with  $n_x = n_y = 65 (= 2^6 + 1)$ . The synthetic rainfall events are represented by ellipses and are added to a zero precipitation field (Figure 2). The fields  $u$  and  $v$  have two rainfall events each. However, they are at different positions. The distortion between the fields include both a rotation and a shear. The events also differ by their intensity and shape. The event in the lower right corner has the same shape in both fields  $u$  and  $v$ , but has a difference of intensity of 5 mm/h. The event in the upper-left corner has the same intensity in both  $u$  and  $v$ , but the ellipse has a different inclination.

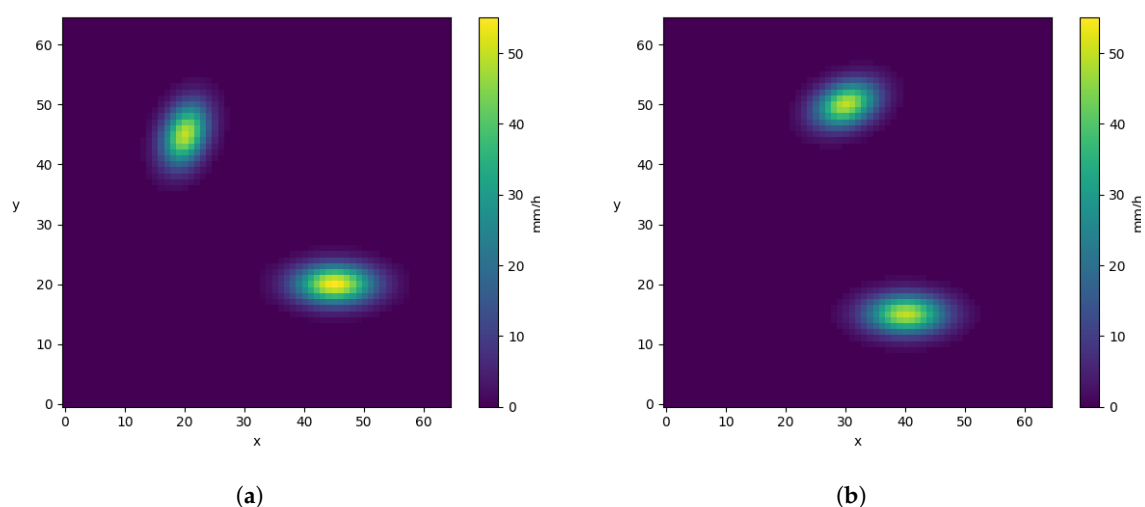


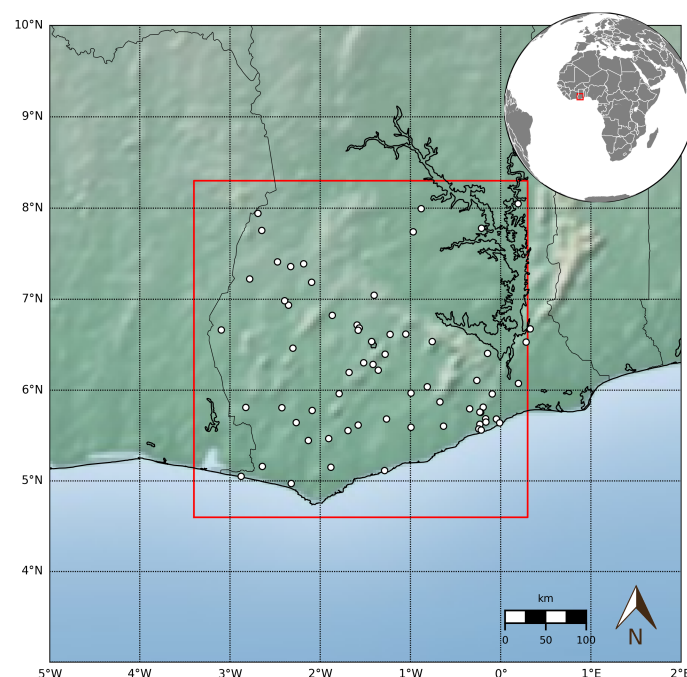
Figure 2. Precipitation fields  $u$  and  $v$  for the synthetic case, shown respectively in (a) and (b).

### 3.2. Southern Ghana Case

In this case, we apply the automatic registration and the warping to real precipitation data. Our goal is to gauge-adjust a satellite-based estimate with respect to the location of the rain events. We assume the gauge measurements to be more accurate, but the IMERG has a higher spatial variability that the gauges cannot reproduce because of their network's density. Thus, our goal is to keep the spatial variability of IMERG, while correcting the position mismatch with respect to the gauge. For this real case, we will only look at the warping for this case. The morphing would make the warped field similar to the gauges, even in areas without gauges.

The study area is a square domain over southern Ghana encompassing the Ghanaian cocoa region. This domain has been chosen because of the particular high density of the Trans-African Hydro-Meteorological Observatory (TAHMO) network in this area (Figure 3).

Southern Ghana has two rainy seasons. The main one extends from March to mid-July and the second one occurs during September and October. We chose a rainfall event during the main monsoon season. More precisely, we selected one hour during this event (from 18:00 to 19:00 of 22 April 2018). Given the spatial resolution of IMERG ( $0.1^\circ$  lat/lon), hourly accumulation is in good agreement with the rainfall spatial and temporal variability [5]. Longer time scale would make it more difficult to identify the individual events and very few precipitation datasets are available at a sub-daily scale.



**Figure 3.** Study domain (red rectangle) and the TAHMO stations available within the domain (white dots).

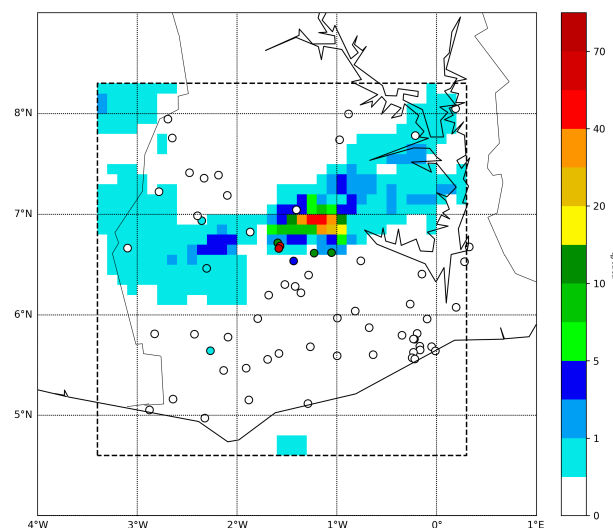
#### 3.2.1. Precipitation Datasets

IMERG (Integrated Multi-satellite Retrievals for GPM) is a high resolution global precipitation product produced by NASA as part of the Global Precipitation Measurement (GPM) mission [36–38]. IMERG merges several satellite estimates from infrared, passive-microwave and satellite-radar. Three versions are available at half-hourly and  $0.1^\circ$  lat/lon resolution: the Early, Late and Final runs. The Final run is gauge-adjusted at monthly scale with the GPCC (Global Precipitation Climatology Centre) gauge product. This product is not available for the Early and Late runs, which have

respectively 5 h and 15 h latency. Instead, they are climatologically adjusted to the Final run. So, they indirectly incorporate past gauge data through this climatological adjustment, but are independent of recent rainfall measurements. The warping will be applied to IMERG-Late run, on a  $37 \times 37$  grid points area corresponding to the study domain.

The Trans-African Hydro-Meteorological Observatory (TAHMO [39]) initiative aims to develop a dense network of 20,000 hydro-meteorological stations in sub-Saharan Africa (equivalent to one station each 30 km). These low-cost stations measure the standard meteorological variables, including precipitation, at a high temporal resolution (5 min). The current TAHMO network contains over 500 stations, mainly in West and East Africa. Data from 66 TAHMO stations are available within the study area for the selected hour.

IMERG-Late estimates and TAHMO measurements for the selected hour, i.e., for the 22 April 2018 between 18:00:00 and 19:00:00, can be seen in Figure 4. Both show one main rainfall event, with a maximum of 45 mm/h according to IMERG-Late. TAHMO recorded a slightly higher maximum of 53.45 mm/h at a gauge located South-West of IMERG's peak.



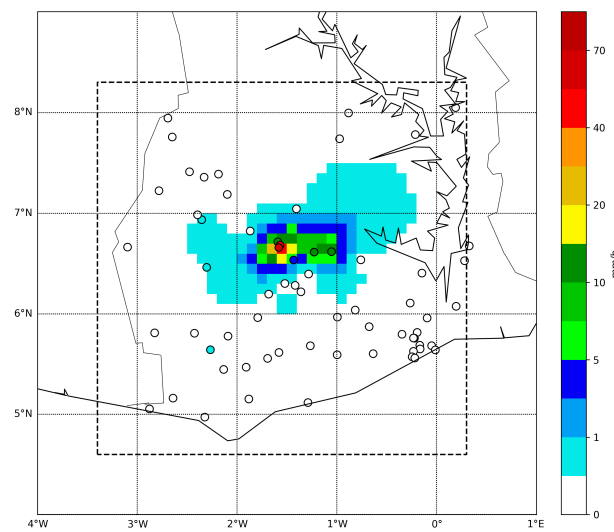
**Figure 4.** IMERG-Late (background) and TAHMO (circles) accumulated rainfall between 18:00:00 and 19:00:00, within the study domain (dotted line).

### 3.2.2. Data Pre-Processing

TAHMO measurements are given on an irregular grid, while the automatic registration algorithm assumes that all the data are defined on a regular grid. The TAHMO measurements were kriged on the same  $0.1^\circ$  lat/lon grid as IMERG estimates (Figure 5). We used ordinary kriging with a square root transform [40]. The ordinary kriging was done with *PyKrig*, a kriging toolkit for python (<https://pypi.org/project/PyKrig/>). We choose an exponential variogram with a sill of  $1.0 \text{ (mm}^2/\text{h}^2)$ , a range of  $1.5^\circ$  and a nugget of  $0.01 \text{ (mm}^2/\text{h}^2)$ .

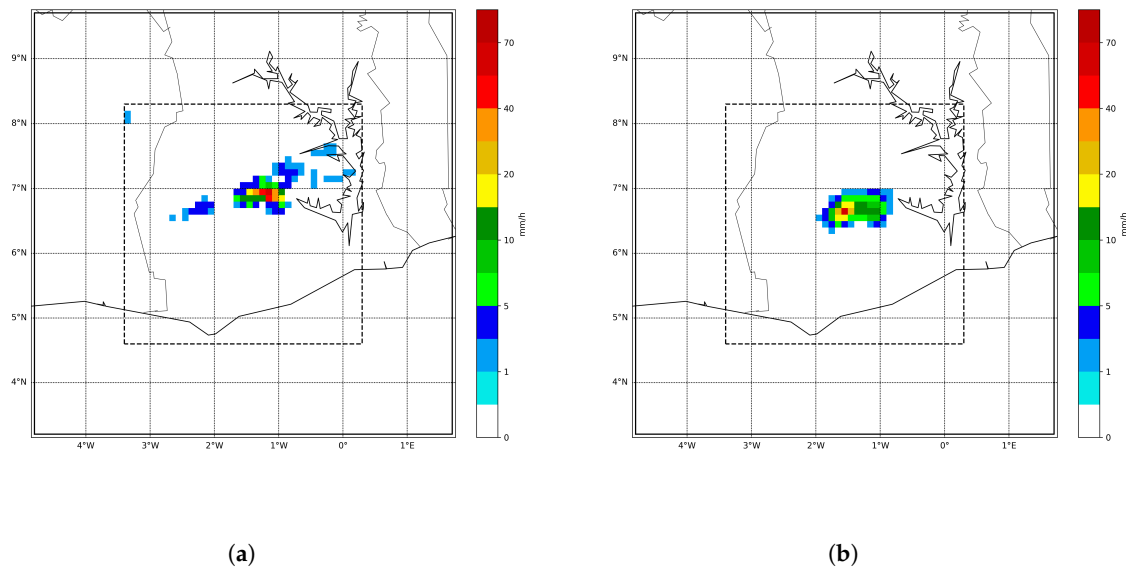
Contrary to the previous cases, this case uses real noisy data. Hence, some pre-processing (beside the kriging) is necessary before applying the automatic registration. First, light precipitation ( $<0.1 \text{ mm/h}$ ) is removed from both fields. Lighter precipitation is in general more difficult to detect, hence subject to large uncertainties. Second, we add a zero-precipitation padding area around the study domain. This area is added to avoid problems near the boundary due to the minimization constraints. The extended domain corresponds to a  $n_x \times n_y$  pixel grid, with  $n_x = n_y = 65$ . Finally, we normalized

the two fields such that the maximum rainfall is equal to 50 mm/h for both. Figure 6a,b show the fields  $u$  (IMERG-Late) and  $v$  (TAHMO) after the three pre-processing steps described above. The study domain is delimited by the dotted line, while the extended domain including the zero-padding area is delimited by the solid line. The extended domain corresponds to a regular  $n_x \times n_y$  grid, with  $n_x = n_y = 65$ . These modified precipitation fields will be used as inputs for the automatic registration procedure to find the mapping  $T$ . Once the mapping  $T$  is found, it will be applied to the original fields  $u$  (Figure 4) on the extended domain.



**Figure 5.** TAHMO measurements (circles) and TAHMO kriged (background) within the study domain (dotted line) for the selected event.





**Figure 6.** Fields  $u$  (IMERG-Late) and  $v$  (TAHMO) after pre-processing and used as inputs. The dotted line delimits the study domain and the solid line the extended one, with the zero-padding area in-between. (a) Field  $u$  (IMERG) after pre-processing. (b) Field  $v$  (TAHMO) after pre-processing.

## 4. Results

### 4.1. Synthetic Cases

The coefficients  $C_1$ ,  $C_2$  and  $C_3$  in the cost function  $J$  are set empirically to  $C_1 = 1$ ,  $C_2 = C_3 = 10$ . We will first look at the convergence of the minimization problem on each domain  $D_i$  for  $i = 1, \dots, I$ , with  $I = 5$  (so that  $m_I = 33 \times 33 < n = n_x \times n_y = 65 \times 65$ , see Section 3). Then, we will look at the performance of the automatic registration and morphing procedure.

The coefficients  $C_1$ ,  $C_2$  and  $C_3$  influence the value of the cost function, and so the convergence. However, their impact on the mapped field is limited and mainly affect the area with no or low precipitation. For example, multiplying the coefficients by two will give similar results in terms of mean absolute error and root mean square error (see Appendix B).

#### 4.1.1. Convergence

For each domain  $D_i$  ( $i = 1, \dots, 5$ ), we investigate the convergence of both the inner loop (from the barrier approach, that is, the iterations on the coefficient  $\beta$ ) and of the outer loop (from the L-BFGS-B minimization method) for the synthetic case (Figure 2).

The value of the cost function before and after minimization for each step  $i$ , as well as the number of iterations needed by the L-BFGS-B method and the barrier approach to reach convergence, are shown in Table 1. The iterations on the coefficient  $\beta$  are performed only if some constraints are violated. For the ideal case described above, iterations over  $\beta$  are required for grids  $D_2$  and  $D_5$  (with a maximum of 3 iterations for  $D_3$ ). The number of iterations needed by the L-BFGS-B method varies depending on the domain  $D_i$  and the coefficient  $\beta$ . It is interesting to observe that less iterations are necessary for increasing  $\beta$ . Domain  $D_3$  necessitates the highest number of iterations and  $D_1$  the lowest (1308 and 81 iterations respectively). These numbers are specific to this case and will vary depending on the input fields  $u$  and  $v$ , and on the coefficients  $C_1$ ,  $C_2$  and  $C_3$ . However, they show that convergence may require a high number of iterations and that this number is difficult to predict.

**Table 1.** Optimization results after each step  $i = 0, \dots, 5$  for the synthetic case. The number of iterations needed for the barrier approach ( $\beta$  iterations) and for the L-BFGS-B method are given separately. The total number of iterations correspond to the sum of the L-BFGS-B iterations for each  $\beta$  iteration. The cost function  $J_p$  is evaluated before and after optimization (i.e., for the first guess  $T_i^{\text{fg}}$  and the ‘optimal’ grid  $T_i^*$ ). The latter is also separate into three terms, the mapping error ( $J_m$ ), the background error ( $J_b$ ) and the penalization term (not shown here because of its value close to zero).

$i$	1	2	3	4	5
<b>Number of iterations</b>					
- Total	81	996	1308	502	443
- $\beta$ iter.	1	3	1	1	2
- L-BFGS-B iter. for each $\beta$	81, $\beta = 1$	956, $\beta = 1$ 36, $\beta = 10$ 4, $\beta = 100$	1308, $\beta = 1$	502, $\beta = 1$	263, $\beta = 1$ 180, $\beta = 10$
<b>Cost function</b>					
$J_p(T_i^{\text{fg}})$	213.169	54.722	40.568	24.278	16.371
$J_p(T_i^*)$	30.058	18.408	19.193	14.790	13.191
$J_o(T_i^*)$	27.552	16.352	17.847	13.399	11.734
$J_b(T_i^*)$	2.507	2.056	1.345	1.391	1.457

The reduction of the cost function is more important on the coarser domains. It is divided by 7 on domain  $D_1$ , by 3 on domain  $D_2$  and by 2 on domain  $D_3$  (Table 1). This is probably due to better first guesses, since the displaced grid obtained on a domain  $D_i$  is used to initialize the finer grid  $D_{i+1}$ . The observation error term  $J_o$  of the cost function is reduced greatly after optimization on the four coarsest domains (from 27.552 after optimization on domain  $D_1$  to 13.399 on domain  $D_4$ ). However, it increases from domain  $D_2$  to  $D_3$ , before decreasing again on domain  $D_4$ . This increase could be due to the smoothing which is narrower for increasing  $i$ , making the position error more apparent in the mapping error. This could also be due to the background error term ( $J_b$ ), which larger decrease is also from domain  $D_2$  to  $D_3$ .

To investigate further the impact of each domain  $D_i$  on the automatic registration procedure, we examine the mean absolute error (MAE) of the morphed ( $u_{\text{morph}} = u_{\text{morph}(\lambda=1)}$ ) and warped ( $u_{\text{warp}} = u_{\text{warp}(\lambda=1)}$ ) signals for different values of  $I$  (shown in Table 2). The MAE of the morphed signal is divided by around 4.2 from  $I = 1$  to  $I = 2$  and by 2.6 from  $I = 2$  to  $I = 3$ . Domains  $D_4$  and  $D_5$  have a smaller impact on the MAE (MAE divided by around 1.3). Similar trend can be observed for the MAE of the warped signal, but with a smaller decrease. The MAE is divided by 1.9 from  $I = 1$  to  $I = 2$  and from  $I = 2$  to  $I = 3$ . Increasing  $I$  from 4 or 5 has only a minor effect on the MAE of the warped signal.

**Table 2.** MAE (in mm/h) of the morphed ( $u_{\text{morph}}$ ) and warped ( $u_{\text{warp}}$ ) signals obtained at different steps  $i$ , for the synthetic case.

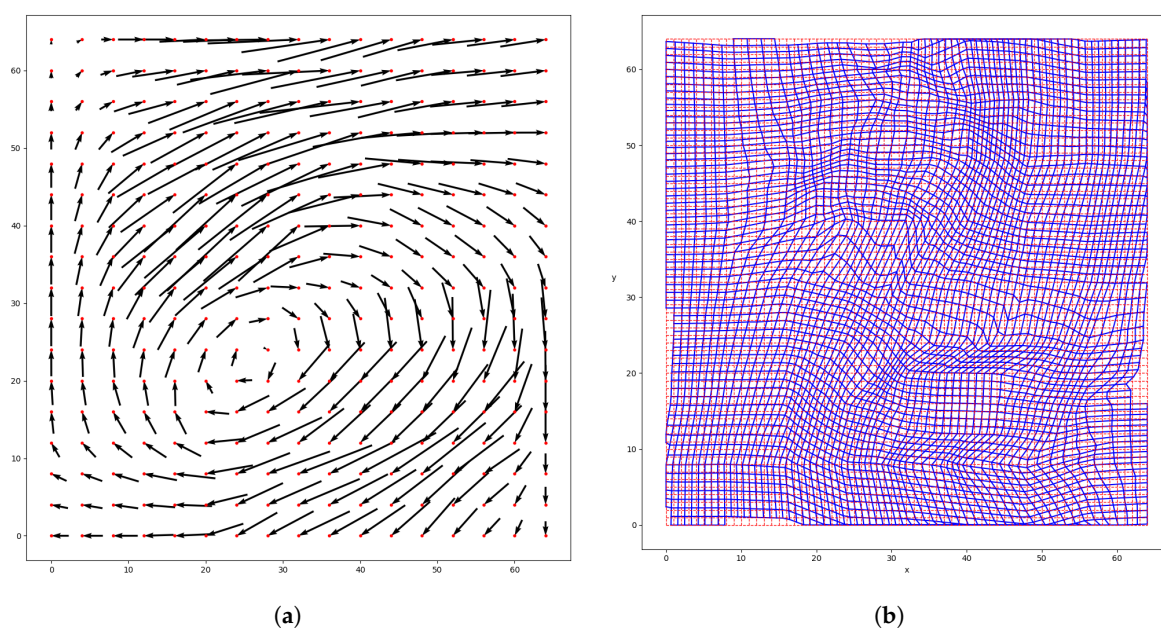
$i$	1	2	3	4	5
<b>MAE(<math>u_{\text{morph}} - v</math>)</b>	0.678	0.161	0.062	0.050	0.036
<b>MAE(<math>u_{\text{warp}} - v</math>)</b>	0.531	0.281	0.147	0.137	0.133

The number of variables to optimize on a domain  $D_i$  increases with increasing  $i$ . So, solving the minimization problem becomes more computationally expensive. Moreover, we saw that domain  $D_5$  has a smaller impact on the cost function and on the MAE of the morphed signals. Thus, in order to limit the computational cost, the finer morphing grid will be set to  $I = 4$  in the following validation section.

#### 4.1.2. Validation

In this subsection, we are looking at the performance of the automatic registration and of the morphing procedure on the synthetic case for  $I = 4$ .

From the automatic registration procedure, we obtain a mapping function  $T$  that will be used to distort the image in the morphing process. The mapping  $T$  is shown in Figure 7a. The corresponding grid distortion is illustrated in Figure 7b which shows the regular  $n_x \times n_y$  grid (on which  $u$  and  $v$  are defined) before and after the mapping has been applied. By comparing with the signals  $u$  and  $v$  (in Figure 2), one can see that the mapping was able to take into account both the rotation and the shear between the events. The biggest distortions occur around the two rainy peaks, while the distorted grid is more regular and closer to the original grid in the areas further away from the peaks. This is due to the first regulation term of the cost functions ensuring that the transformation  $T$  is as “small” as possible. The second regulation term is responsible for the smooth transition between these areas. The impact of the third regulation term is more difficult to see directly. The third term penalizes grid cell to shrinkage or expansion. Due to the locations of the rainfall events in the signals  $u$  and  $v$ , almost the entire domain is influenced by the mapping  $T$ . Large distortions near the boundary can raise some inconsistencies in the distorted grid, such as in the one in the lower-right corner in Figure 7b. A way to avoid such discrepancies near the boundary would be to add a padding area around the domain filled with zero precipitation.

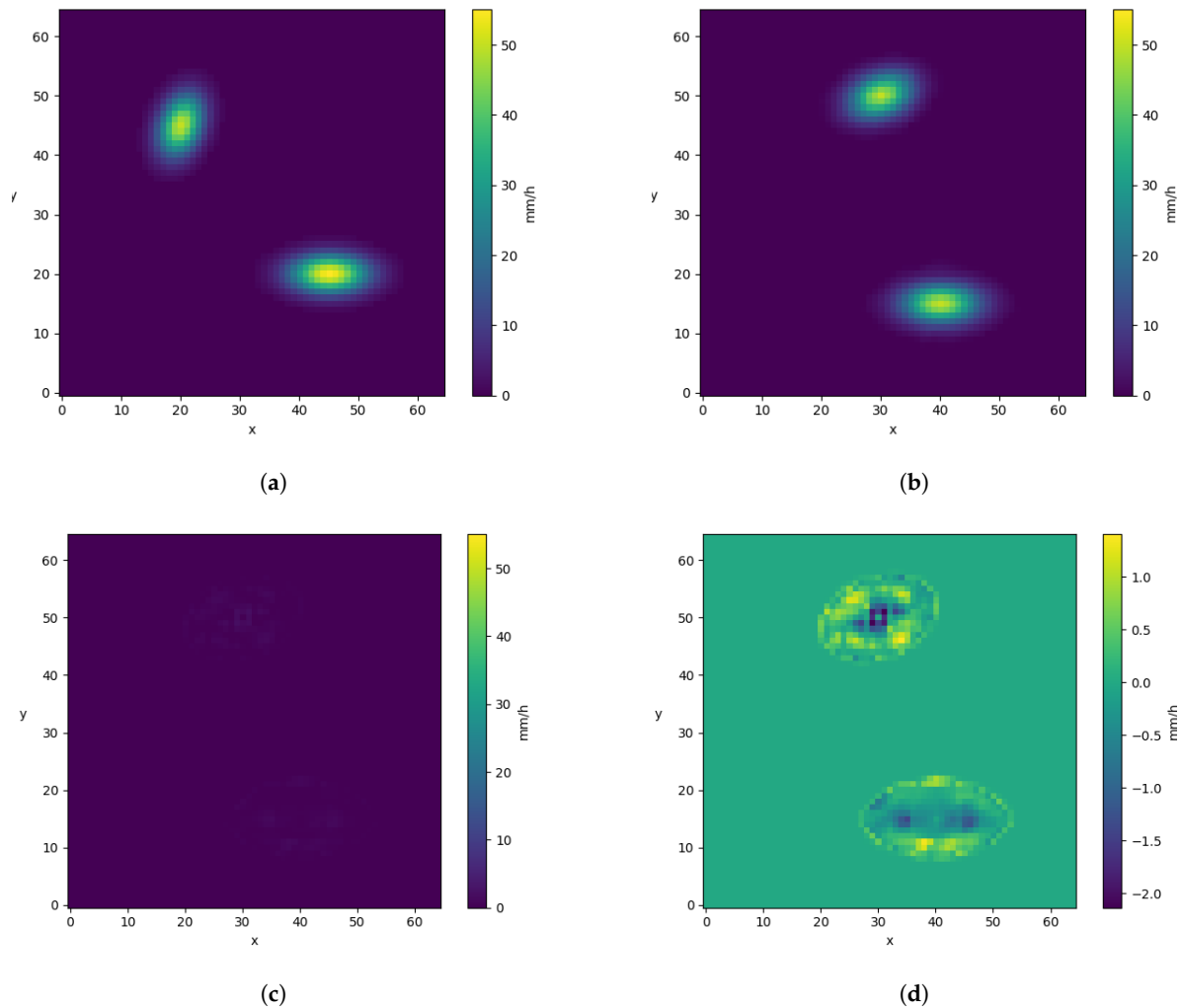


**Figure 7.** (a) Mapping  $T$  obtained from the automatic registration (with  $C_1 = 0.1$  and  $C_2 = C_3 = 1$ ) and (b) its effect on the pixel grid, for the synthetic case. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .

Once the mapping  $T$  is obtained, it can be applied to the signal  $u$ . One can see the impact of the mapping by comparing the morphed signal  $u_{\text{morph}}$  (Figure 8b) to the original signal  $u$  (Figure 8a). The error between the morphed signal and the signal  $v$  is shown in Figure 8c. Figure 8c shows that the absolute error is small compared to the intensity of the rain events. This is also reflected by the low MAE in Table 1. Figure 8d shows that the errors have a clear delimitation, corresponding to the rain events of the target field  $v$ . For this synthetic case, the center of the two events in  $u_{\text{morph}}$  and  $v$  are corresponding exactly with respect of the position.

It is interesting to note that, while the lower event had an error on the intensity and an error on its location, the error of the morphed signal is lower around its peak than for the second (upper) events. The error of the warped field is close to 5 mm/h near this peak (not shown here). This shows

the advantage of morphing over warping when there is an intensity error. The upper event does not have any intensity error, but it has a shape error (i.e., the ellipses have different inclinations). The largest error in the morphed field  $u_{\text{morph}}$  occurs around its peak. The top is underestimated, while the surroundings are overestimated. This pattern would suggest an error due to the linear interpolation used for the morphing. This underestimation is greatly reduced when using a 2d spline instead of the current bi-linear interpolation (not shown here).



**Figure 8.** Morphing results: (a) original signal  $u$ , (b) morphed signal  $u_{\text{morph}}$ , (c) absolute error between the morphed signal and the target signal  $v$  and (d) error between the same two signals (using a different scale). (a)  $u$ . (b)  $u_{\text{morph}}$ . (c)  $|u_{\text{morph}} - v|$ . (d)  $u_{\text{morph}} - v$ .

#### 4.2. Southern Ghana Case

The results described in this section are obtained for  $I = 4$ , with the regulation coefficients empirically set to  $C_1 = 0.1$  and  $C_2 = C_3 = 1$  (these coefficients are the same as for the synthetic cases). We define the mask function  $M$  such that  $M = 1$  at the grid points where the kriging variance  $\sigma_{\text{kriging}} < 0.5 \cdot \text{sill} = 0.5 \text{ mm}^2/\text{h}^2$  and  $M = 0$  otherwise. The padding area is thus masked too.

##### 4.2.1. Convergence

We first look at the convergence of the automatic registration procedure when applied to this real (noisy) case. Table 3 shows the number of iterations (nit) needed by the L-BFGS-B method to converge for each  $\beta$ , as well as the cost function before and after the minimization. Domain  $D_4$  (for  $\beta = 1$ ) necessitates the highest number of iterations and  $D_1$  the lowest (1463 and 10, respectively). Iterations

on the coefficient  $\beta$  are preformed only on domain  $D_4$ . Contrary to the synthetic case, the reduction of the cost function is important on the finer morphing grid. It is divided by two on  $D_2$  and  $D_3$  and by five on  $D_4$ . In this case, the background term  $J_b$  is higher than the mapping error term  $J_m$  on all domains  $D_i$ . The cost function after optimization  $J_p(T_i^*)$  increases from domain  $D_1$  to  $D_4$ , mostly due to the background term  $J_b$ . The important increase of the mapping error  $J_m(T_i^*)$  on domain  $D_4$  is mostly due to the smoothing that reveals sharper features.

**Table 3.** Optimization results after each step  $i = 0, \dots, 4$  for the southern Ghana case. The number of iterations needed for the barrier approach ( $\beta$  iterations) and for the L-BFGS-B method are given separately. The total number of iterations correspond to the sum of the L-BFGS-B iterations for each  $\beta$  iterations. The cost function  $J_p$  is evaluated before and after optimization (i.e., for the first guess  $T_i^{\text{fg}}$  and the 'optimal' grid  $T_i^*$ ). The latter has also been separated into three terms, the mapping error ( $J_m$ ), the background error ( $J_b$ ) and the penalization term (not shown here because of its value close to zero).

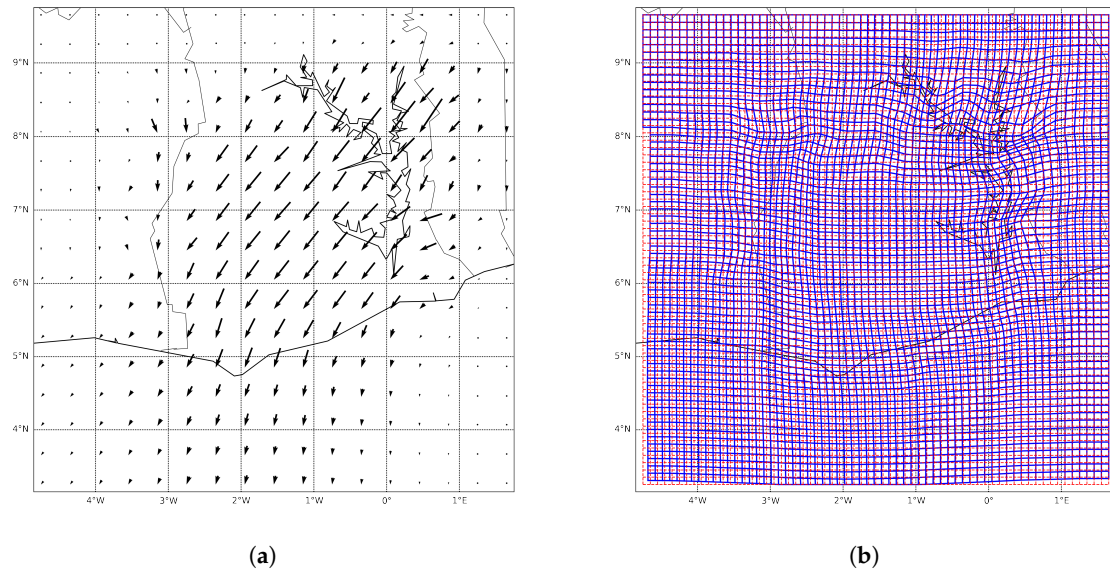
$i$	1	2	3	4
<b>Number of iterations</b>				
- Total	10	60	433	1463
- $\beta$ iter.	1	1	1	5
- L-BFGS-B iter. for each $\beta$	10, $\beta = 1$	60, $\beta = 1$	433, $\beta = 1$	1430, $\beta = 1$ 16, $\beta = 10$ 9, $\beta = 100$ 4, $\beta = 1000$ 4, $\beta = 10,000$
<b>Cost function</b>				
$J_p(T_i^{\text{fg}})$	0.0516	0.0938	0.1068	0.3602
$J_p(T_i^*)$	0.0421	0.0424	0.0483	0.0679
$J_m(T_i^*)$	0.0168	0.0058	0.0060	0.0252
$J_b(T_i^*)$	0.0253	0.0376	0.0422	0.0427

#### 4.2.2. Validation

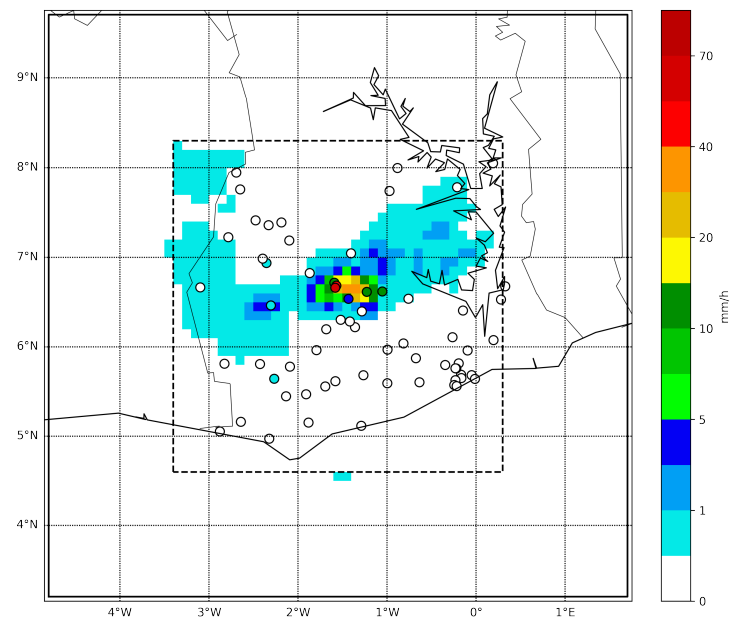
The mapping  $T$  obtained from the automatic registration and its effect on the pixel grid  $D_n$  is shown in Figure 9. The field is shifted toward the South-West. The deformations are more important in the center and in the South of the domain, while being very small or null near the boundary. The area near the boundary corresponds to the padding area which is filled with zeros. The regulation terms are thus dominating the cost function in this area, especially the first one that ensure the mapping to be as small as possible. In contrast, the first term of the cost function is dominant in the center of the domain where the rainfall event is located. The second and third regulation terms ensure that the transition between these two areas is smooth.

This mapping  $T$  is then applied to the field  $u$ , the satellite estimate from IMERG, to correct the location of the rain event. Figure 10 shows the warped field  $u_{\text{warp}}$  and the TAHMO measurements. One can see that the location of the rainfall event is corresponding more to the gauge data than before the warping (Figure 4). We define the center of the event by the grid cell with the maximum precipitation. Using the kriged gauge field  $v$  as the truth, we compute the position error of the event's center before and after the warping. It decreases from 55,365 km to 22,096 km. This remaining error of 22,096 km is due to an error in the longitude. Indeed, the maximum rainfall is at the correct latitude, but is two grid cells to the East of the actual peak. This error in longitude can be explained by the internal structure of the event. By comparing the pre-processed fields in Figure 6, one can see that the peak is located in the eastern part of the event for IMERG ( $u$ ), but in the West for the kriged gauge ( $v$ ). This difference could be due to the kriging and gauge network density (no gauge measurement is available near the peak of the warped field).





**Figure 9.** (a) Mapping  $T$  obtained from the automatic registration (with  $C_1 = 0.1$  and  $C_2 = C_3 = 1$ ) and (b) its effect on the pixel grid, for the southern Ghana case. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .



**Figure 10.** Warped signal  $u_{\text{warp}}$  (background) and TAHMO measurements (circles).

To investigate further the automatic registration, we compute the MAE and RMSE between the warped field and the target one, using the mappings  $T_i$  obtained at each step  $i$  (Table 4). Applying the mapping  $T_1$ , defined on the coarsest morphing grid, already greatly reduces the MAE and the RMSE compared to the original errors. Increasing the resolution of the morphing grid, further decreases the MAE and RMSE, except for the RMSE for  $i = 2$ . While the cost function is divided by five on domain  $D_4$  (see Table 3), the reduction of the error from  $D_3$  to  $D_4$  is much smaller.

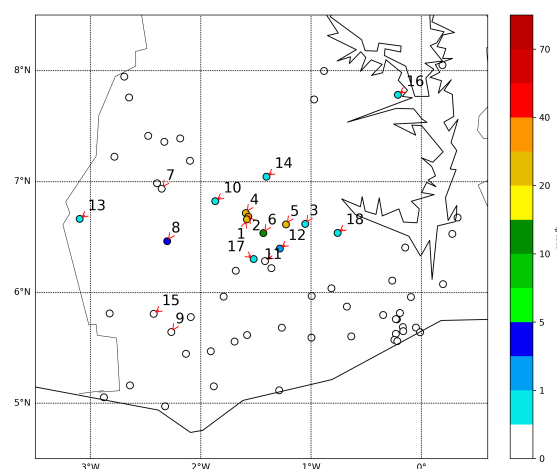
The target field  $v$  has been obtained by kriging the gauge data. Hence, some interpolation errors were introduced, especially in the areas far from the gauges. Comparison to the whole field  $v$  is not

representative, since it contains some large uncertainties. In the second part of Table 4, the warped field  $u_{\text{warp}}$  is estimated at the station locations (shown in Figure 11 for  $T_4$ ) and is directly compared to the gauge measurements. As previously, a large part of error reduction is occurring on domain  $D_1$ . In total (i.e., for  $i = I = 4$ ), the RMSE has been divided by almost two, and the MAE by 1.5.

**Table 4.** MAE and RMSE of the warped ( $u_{\text{warp}}$ ) signal compared to the kriged TAHMO field ( $v$ ) and to the gauge measurements, obtained at different steps  $i$ , for the southern Ghana case.

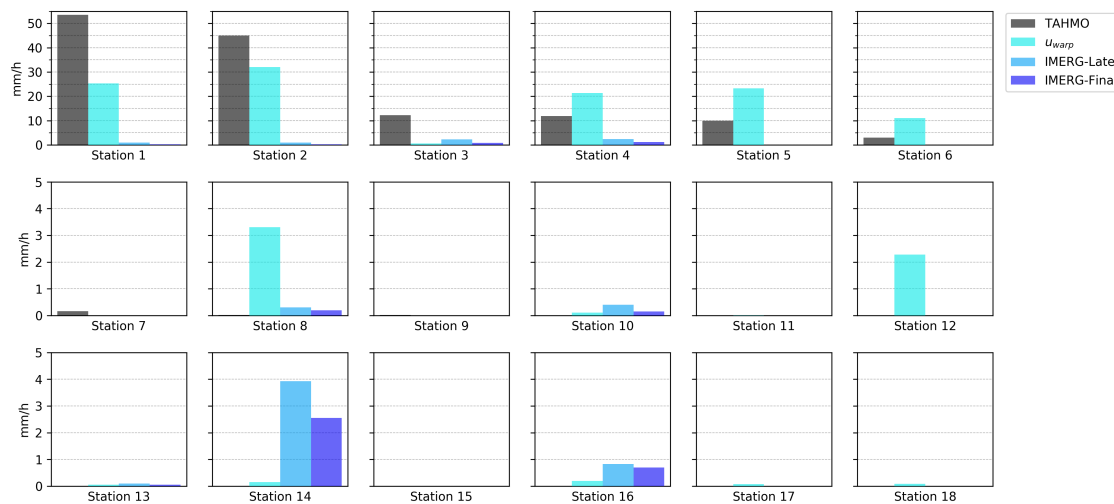
$i$	Before	1	2	3	4
<b>Kriged field (<math>u_{\text{warp}} - v</math>)</b>					
- MAE (mm/h)	0.1748	0.0968	0.0936	0.0886	0.0865
- RMSE (mm/h)	1.6476	0.9171	0.9496	0.8976	0.8565
<b>Gauge measurements</b>					
- MAE (mm/h)	2.0687	1.4365	1.4434	1.3849	1.3619
- RMSE (mm/h)	8.7053	5.5057	5.4984	4.9349	4.6690

Figure 11 shows the warped field at the station locations. The stations at which the gauge data and the warped field disagree are numbered. At the other stations, both TAHMO and the warped field estimate zero precipitation. The precipitation amounts at the 18 numbered stations according to TAHMO, IMERG-Late ( $u$ ), IMERG-Final and the warped field ( $u_{\text{warp}}$ ) are shown in Figure 12. They are ordered in decreasing order of precipitation with respect to TAHMO measurements. For stations 1 and 2 (with highest rainfall intensity), the warped signal is much closer to the measurements than IMERG-Late and IMERG-Final. However,  $u_{\text{warp}}$  underestimates the intensity and estimates more rainfall at station 2 than 1 (the opposite of TAHMO). This underestimation can be explained by two factors. First, the maximum of IMERG-Late  $u$  on our domain ( $=45$  mm/h) is lower than the one recorded by TAHMO. Second, we used linear interpolation for the warping. The rainfall at station 3 is underestimated by the three satellite estimates. At the stations 4 and 5, both IMERG-Late and IMERG-Final underestimate while  $u_{\text{warp}}$  overestimates. IMERG misses the precipitation at station 6, while  $u_{\text{warp}}$  overestimates by almost a factor four. The remainder of the stations has very low or no precipitation according to TAHMO. The three satellite estimates are similarly low (less than 1 mm/h) at these stations, with a few exceptions: IMERG-Late and IMERG-Final at station 14 and  $u_{\text{warp}}$  at stations 8 and 12. The overestimation by  $u_{\text{warp}}$  at station 8 corresponds to a second lower peak present in the original field  $u$  but not in the target field  $v$ .



**Figure 11.** Warped signal  $u_{\text{warp}}$  at the stations location. The stations that differ from the TAHMO measurements are numbered.





**Figure 12.** Rainfall (in mm/h) at 18 TAHMO stations according to TAHMO, IMERG-Late, IMERG-Final and the warped signal ( $u_{\text{warp}}$ ).

## 5. Discussion

### 5.1. Convergence

The automatic registration algorithm converges for both the synthetic and the real (Southern Ghana) case. However, some differences can be noticed when comparing the optimization results shown in Tables 1 and 3.

The minimization method (L-BFGS-B) needs more iterations for the synthetic than for the real case for all steps  $i$  except for  $i = 4$ . Iterations on the coefficient  $\beta$  were needed for step  $i = 2$  and  $i = 5$  for the synthetic case, while the real case needed it for its finer step  $i = 4$ . The real case is noisier than the synthetic one, but the displacement between the field  $u$  and  $v$  is more straightforward. Indeed, for the synthetic case, the mapping  $T$  has to describe both a rotation and a shear. For the real case, the mapping  $T$  only has to represent a translation. On the coarser grids, the translation does not violate the constraints on the barrier. On the other hand, a rotation is more likely to violate the constraints and so to require iterations on  $\beta$  on coarser grids too. This shows that the number of iterations of both the minimization and barrier method depends on the input fields  $u$  and  $v$ , especially on the mapping complexity. It also depends on the coefficients  $C_1$ ,  $C_2$  and  $C_3$ , but the influence of the chosen coefficient values is limited (results not shown here).

For the synthetic case, the decrease in the cost function is more important for the first steps. This can be explained by better first guesses for the finer grids. On the contrary, for the real case, the decrease is more important in the last steps. While the events were identical in the synthetic case, they have different shapes in the real one. On the coarser grid, these differences are masked by the strong smoothing. They become more visible on the finer grid on which there is less smoothing. The sharper features being more sensitive to small position errors results in a higher cost function. The finer morphing grids allow the mapping to take the shape difference into account, on top of the position error. The reduction of the cost function is thus becoming more important.

The intermediate mappings  $T_i$  give information about the impact of the steps  $i = 1, \dots, I$ . They were evaluated by looking at the MAE of the warped fields  $u_{\text{warp}}$  (Tables 2 and 4). In the real case, most of the MAE decrease is reached after the first iteration (divided by around 2). The decrease after the subsequent iterations is more marginal. The mapping on the coarser morphing grid  $D_1$  is already able to capture reasonably well the displacement, that is, the translation toward the South-West. Thus, the finer morphing grids induce less improvement. In the synthetic case, the MAE is improved greatly after the first iteration too (divided by 7). However, the subsequent iterations continue to decrease the MAE (divided by 3 after step 2 and by 2 after step 3). The first morphing grids are too coarse to describe

accurately the complex displacement, which combine a rotation and a shear. Hence, increasing the resolution of the morphing grids improves the mapping  $T_i$  and thus the MAE. The more complex the displacement is, the finer the morphing grid needs to be and so the higher  $I$  has to be (i.e., the more steps  $i$  we need).

The computational time increases exponentially with the number of steps  $i$ . All computations shown here were done on a personal computer. For the synthetic case, the first four steps (i.e.,  $i = 1, 2, 3$  and  $4$ ) were completed in approximately 2 min, while the fifth iteration ( $i = 5$ ) needed between 4 to 10 min (depending on the computer computational capacity). The real case requires fewer iterations than the synthetic one and so a shorter computational time ( $\sim 1$  min for  $I = 4$  and  $\sim 4$  min for  $I = 5$ ).

## 5.2. Validation

The automatic registration procedure converged, for both the synthetic and the real cases. It has been shown that the error between the two original fields was considerably reduced by applying the mapping  $T$ , even without bias adjustment (i.e., the warped field  $u_{\text{warp}}$ ). An issue encountered in the synthetic case was the grid distortion near the domain boundary (Figure 7). Some inconsistencies can appear when the rainfall events are close to the boundary. This was solved in the real case by adding a padding area, filled with zero precipitation, around the domain. This padding area enables the mapping to have a smooth transition from the largest displacement near the events to (almost) none near the new extended boundary (Figure 9a).

The automatic registration produced reasonable coordinate mapping in these two cases. However, problems can arise if the dissimilarity between the two original fields are too strong. We do not have a method to quantify this problem beforehand. However, there are some minimum conditions, such as having the same number of events in both fields or the proximity of these events. The smoothing steps of the registration algorithm can also be increased or decreased to allow the events to move further or not. In this study, we did not push to the cases to the extreme to determine a feasibility threshold. The goal of this article was to prove the applicability of registration and morphing to precipitation data. A next step would be to apply it to other cases, including different rainfall regimes.

The warping succeeded in correcting the general position error between the fields  $u$  and  $v$  in both cases. In the real case, the shape of the event is different in the original fields. One can notice that the shape of the event is slightly altered by the mapping, but the internal structure stayed similar. The peak (rainfall above 20 mm/h) is larger in the field  $u$  than in the target field  $v$ , but the event (rainfall above 1 mm/h) has a larger longitudinal spread according to  $v$  (see Figures 4 and 5). This can explain some of the intensity differences at the station location (Figure 12). The stations 4 to 6 are in the center of the event, the station 4 is especially very close to the peak (stations 1 and 2). The overestimation by the warped field is related to the larger peak in the field  $u$ . Similarly, the underestimation at station 3 and the overestimation at station 12 is due to their location near the edge of the event. Stations 3 and 12 are located close to the East edge and South edge of the event respectively and so are affected by the spread difference of the event according to  $u$  and  $v$ .

The morphing has been evaluated only for the synthetic case. Table 2 shows the added advantages of morphing over warping. The MAE of the warped signal is larger by factor 2 on domain  $D_3$  and by a factor of 2.7 on domain  $D_4$  compared to the MAE of the morphed signal. This important difference is due to the intensity difference between the lower event in  $u$  and the one in  $v$ . However, this advantage decreases when the intensity difference decreases. When the intensity is the same for the two fields, warping is more advantageous. Without difference in intensity, the residual  $r$  in the morphing formula Equation (7) is unnecessary and only add numeric errors (because of the inverse transform  $(I + T)^{-1}$  and the extra linear interpolation). The morphing was not tested for the real case because of the irregular nature of the observations. The uncertainty of the kriged field is high in large part of the domain where no gauges are available. We made assumptions on the spatial mapping through the three criteria for optimality. This allowed us to correct the position through the entire domain. However, we can not make similar assumptions for the intensity.

### 5.3. Applications

In this paper, we corrected a satellite-based precipitation estimate based on gauge measurements. This position correction could be particularly useful to pre-process satellite rainfall data for applications needing accurate rain event positioning. Image morphing can take both the position and the intensity into account but we do not recommend to correct both at the same time. The morphed estimate would then be comparable to the kriged gauge field, without any advantage of the more detailed spatial structure of the satellite observations. A two-step approach is preferred, with first the position correction using the warping and then a bias correction such as the additive-multiplicative one used by IMERG-Final. Such position correction could be particularly beneficial as a pre-processing step for hydrological modelling applications. Rainfall data is an important input for hydrological models and can have a large impact on their accuracy [41,42]. The correct positioning of rainfall events can be as crucial as their intensities, especially for the localized events.

The morphing can also be applied to rainfall fields from other sources, such as a numerical model. It can then be used for data assimilation. Two approaches are possible. The position correction can be applied as a first step before the usual data assimilation on the intensity [22,24]. It is also possible to assimilate both intensity and distortion at the same time, represented respectively by the residual  $r$  and the mapping  $T$  [32,33]. In the second case one can take full advantage of the morphing formulation. In this paper, we did not perform data assimilation as it was described in References [32,33]. Instead, we used a similar method to theirs to correct the position in a satellite-based estimate using gauge data. There are three main differences between our method and the morphing described in References [32,33]. First, they used two penalty terms to ensure the smoothness of the displacement field (based on its magnitude and gradient), we add a third penalty term based on the divergence. Second, they solved the minimization problem for one grid point at a time (i.e., they have several 1D minimization problems), while we solve it for all the grid points together (i.e., we have one multi-variable minimization problem). Finally, we extend the method to non-gridded observations. Contrary to radar data, the gauge measurements used in our study case are irregularly spaced (i.e., non-gridded). In Reference [33], the framework for assimilating radar rainfall using morphing is described but is not actually applied to real rainfall data.

The main limitation of image morphing is in fact the limitations of the automatic registration. As discussed above, it can fail if the fields are too dissimilar. It is also influenced by the three regulation coefficients  $C_1$ ,  $C_2$  and  $C_3$ . For example, in the case of a low intensity event, the regulation terms in  $J_b$  can dominate the cost function, not allowing the rain event to move. In this paper, we explore the feasibility of image morphing for position correction in precipitation estimates. However, we have not pushed to the extreme the cases to quantify its limits. This paper is meant as a proof-of-concept. The next step will be to extend the study to other cases, involving different rainfall regimes. Extreme cases should be included to determine the boundary within which the automatic registration succeeds

## 6. Conclusions

We have investigated the use of a morphing approach for the gauge-adjustment of satellite-based rainfall estimates with respect to position error. The morphing method, adapted from Reference [32], has been applied to two cases. Synthetic rainfall events, represented by ellipses, have been used to test the automatic registration and the morphing method. The second case, a convective rainfall event in southern Ghana, showed the potential of the method when applied to real, noisy precipitation data. We applied the position correction such that the gauge data were downscaled while keeping the high spatial variability of the satellite-based product. The rain events estimated by IMERG-Late were spatially shifted to match the gauge data. The morphing method can take both the intensity and the position of the rain events into account. This is an advantage compared to the traditional gauge-adjustment methods that are only looking at the intensity bias.

The automatic registration is able to represent different types of distortions. However, its performance of the registration depends on the degree of difference between the fields  $u$  and  $v$  and on the regulation coefficients. The more complex the distortion between the fields is, the more computationally expensive the registration is. For example, in the case of a simple distortion (such as a translation), it is possible to choose a smaller number of steps  $I$ . The minimization method at each step would also need fewer iterations. On the other hand, if the fields are too dissimilar, the registration can fail. The regulation coefficients also influence both the convergence and the result of the registration.

This paper explores the use of an image morphing method to correct location errors in precipitation estimates. The next step will be to extend the study to other case studies, including different rainfall regimes. It should also be pushed to more extreme cases to determine the method's limitations more precisely. For example, the regulation coefficients have been chosen empirically in this study. A next step will be to develop a more robust way to select them, for example by defining adaptive coefficients.

**Supplementary Materials:** The python scripts for the automatic registration and the morphing are available online at <https://github.com/clecoz/precipitation-morphing.git>. The scripts permit to reproduce the synthetic case shown in this article.

**Author Contributions:** Conceptualization, C.L.C., A.H., M.V., M.-c.t.V. and N.v.d.G.; methodology, C.L.C., A.H. and M.V.; software, C.L.C.; formal analysis, C.L.C.; writing—original draft preparation, C.L.C.; writing—review and editing, A.H., M.V., M.-c.t.V. and N.v.d.G.; funding acquisition, A.H. and N.v.d.G.

**Funding:** This research is supported by the TU Delft | Global Initiative, a program of the Delft University of Technology to boost Science and Technology for Global Development.

**Acknowledgments:** This research is supported by the TU Delft | Global Initiative, a program of the Delft University of Technology to boost Science and Technology for Global Development. The work leading to these results has received funding from the European Community's Horizon 2020 Programme (2014–2020) under grant agreement No. 776691 (TWIGA). The opinions expressed in the document are of the authors only and no way reflect the European Commission's opinions. The European Union is not liable for any use that may be made of the information. We acknowledge TAHMO ([www.tahmo.org](http://www.tahmo.org)) for providing the in-situ data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A. Optimization Step of the Automatic Registration Algorithm

The automatic registration is based on the minimization of the cost function  $J$  given in Equation (9). This is done in the third step of the algorithm (Figure 1). In this appendix, we give more information about this important step of the algorithm, with a focus on the constraints (Appendix A.1) and on the minimization method (Appendix A.2).

### Appendix A.1. Constraints

This minimization problem include a certain number of constraints. These constraints come from two requirements. First, we require that the nodes stay within the domain  $D$  (i.e.,  $(I + T)(x_k, y_j) \in D \quad \forall (x_k, y_j) \in D_i$ ). That is, the nodes on the boundary of the domain are allowed to move inside the domain but no nodes are allowed to leave it. These constraints can be seen as inequality constraints or as bounds. Second, the mapping  $I + T$  has to be invertible. Each node  $(x_k, y_j)$  of the grid  $D_i$  is constrained to the domain between the points  $(x_{k+1}, y_j)$ ,  $(x_k, y_{j-1})$ ,  $(x_{k-1}, y_j)$  and  $(x_k, y_{j+1})$ , in order to insure that the inverse mapping  $(I + T)^{-1}$  exists (see Reference [32]). This second requirement translates as inequality constraints.

Since nodes on the boundary are allowed to move inside the domain, some computations will need the value of an image outside the domain  $D$  (e.g.,  $u \circ (I + T)^{-1}$  on  $D$  or the smoothing of  $u$  and  $v$ ). To allow these computations, all the images are extended on  $\mathbb{R}^2$  by assuming that there is no precipitation outside the domain, that is,  $\forall (x, y) \notin D, \quad u(x, y) = v(x, y) = 0$ .

### Appendix A.2. Minimization Method

In Reference [32], the minimization problem was solved by optimizing  $T_i$  one node after each other (for all nodes of the grid  $D_i$ ). For each node, the coordinates  $x$  and  $y$  were updated alternatively using

a 1-D constrained non-linear optimization function. We use a different approach and optimize all the nodes at the same time. The solution does not depend on the order in which the nodes are optimized. However, the number of variables to optimize ( $2 \cdot (2^i + 1)^2$  for grid  $D_i$ ) increases exponentially with  $i$ . Constrained optimization algorithms become computationally expensive for such large numbers of independent variables. Thus, for computation efficiency, we used an iterative barrier approach ([34,35]).

In the barrier approach, the inequality constraints are added to the cost function  $J$  as penalization terms:

$$J_p(T_i) = J(T_i) + \beta \sum_h C_h(T_i) \quad (\text{A1})$$

where  $C_h$  are constraint functions and  $\beta$  is the barrier coefficient. The constraints that cannot be written as bounds are converted into constraint functions  $C_h$ , such that  $C_h(T_i) > 0$  if the constraint is violated and  $C_h(T_i) = 0$  if it is respected.

This new minimization problem does not have inequality constraints, only bounds from our first requirement. Here, we minimized the penalized cost function  $J_p$  using the limited-memory quasi-Newton method for bound-constrained problems (L-BFGS-B) method (with the Python function `scipy.optimize.minimize`).

As mentioned above, the barrier approach is iterative. First, the cost function  $J_p$  is minimized using the L-BFGS-B method with  $\beta = 1$ . If all the constraints are respected, the procedure stops. Otherwise, the iterations continue with  $\beta = 10 \cdot \beta$ . The iterations continue until all the constraints are respected or until one of the stopping criteria is reached. We set two stopping criteria: (1) the decrease of the cost function  $P$  is smaller than  $< 10^{-5}$  and (2) the root mean square difference of the grids  $D_i$  before and after is smaller  $< 10^{-5}$ .

## Appendix B. Sensitivity Study for the Regulation Coefficients $C_1$ , $C_2$ and $C_3$

The automatic registration is based on the minimization of the cost function  $J$ , which is composed of an observation term and a background term. The background term  $J_b$  consists of the three criteria defining the 'optimal' mapping  $T$ . The three criteria concern the magnitude (Equation (2)), the smoothness (Equation (3)) and the divergent (Equation (4)) of  $T$ . These three criteria can be tuned with the regulation coefficients  $C_1$ ,  $C_2$  and  $C_3$  respectively. In this section, we look at the influence of these three coefficient on the automatic registration.

For this sensitivity study, we use the synthetic case described in Section 3.1. In the result section (Section 4), we empirically set the coefficients to  $C_1 = 0.1$  and  $C_2 = C_3 = 1$ . Here, we consider four more cases. We first look at the impact of the individual coefficients, setting the two other to zeros. Then we examine the influence of their intensity by multiplying them by five. The only difference between the results shown here and the ones in Section 3.1 are the regulation coefficients.

The impact of the coefficients on the morphed and warped signals can be seen in Table A1. It shows the mean absolute error (MAE) of both signals for the four cases described above and for the original case from Section 3.1. The MAEs are staying similar. The three cases with only one coefficient have a slightly higher MAE for the morphed signal than the original case, but slightly lower for the warped signal. Multiplying the three coefficients by five have the inverse effect. The MAE of the morphed signal is lower but that of the warped signal is slightly higher. The coefficients seem to have a limited impact on the warped and morphed signal.

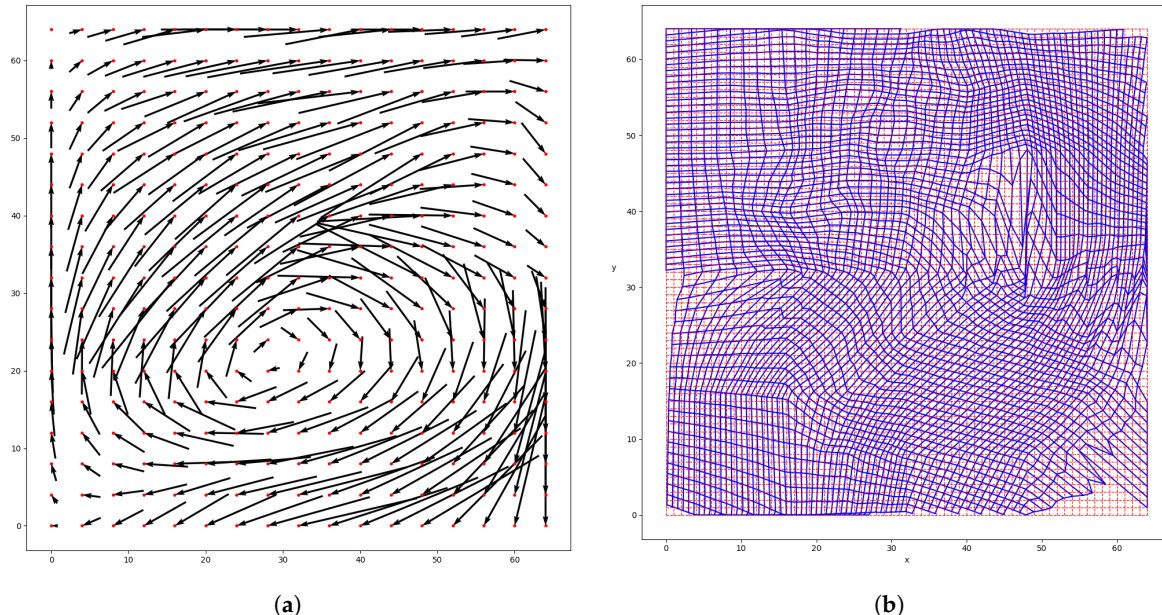


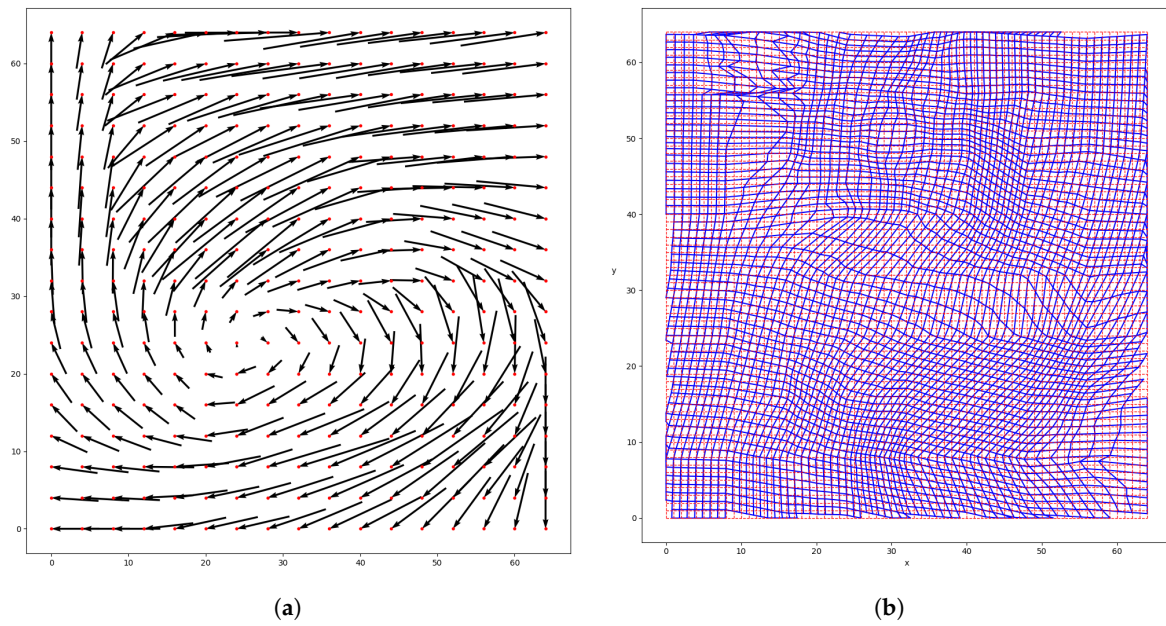
**Table A1.** MAE of the warped ( $u_{\text{warp}}$ ) and morphed  $u_{\text{morph}}$  signals compared to the target field ( $v$ ) for the different sensitivity runs (for  $I = 4$ ).

	Original	Only $C_1$	Only $C_2$	Only $C_3$	All Coef. $\times 5$
$C_1$	0.1	0.1	0	0	0.5
$C_2$	1	0	1	0	5
$C_3$	1	0	0	1	5
- MAE $u_{\text{morph}}$ (mm/h)	0.0496	0.0520	0.0535	0.0515	0.0396
- MAE $u_{\text{warp}}$ (mm/h)	0.1371	0.1363	0.1368	0.1366	0.1374

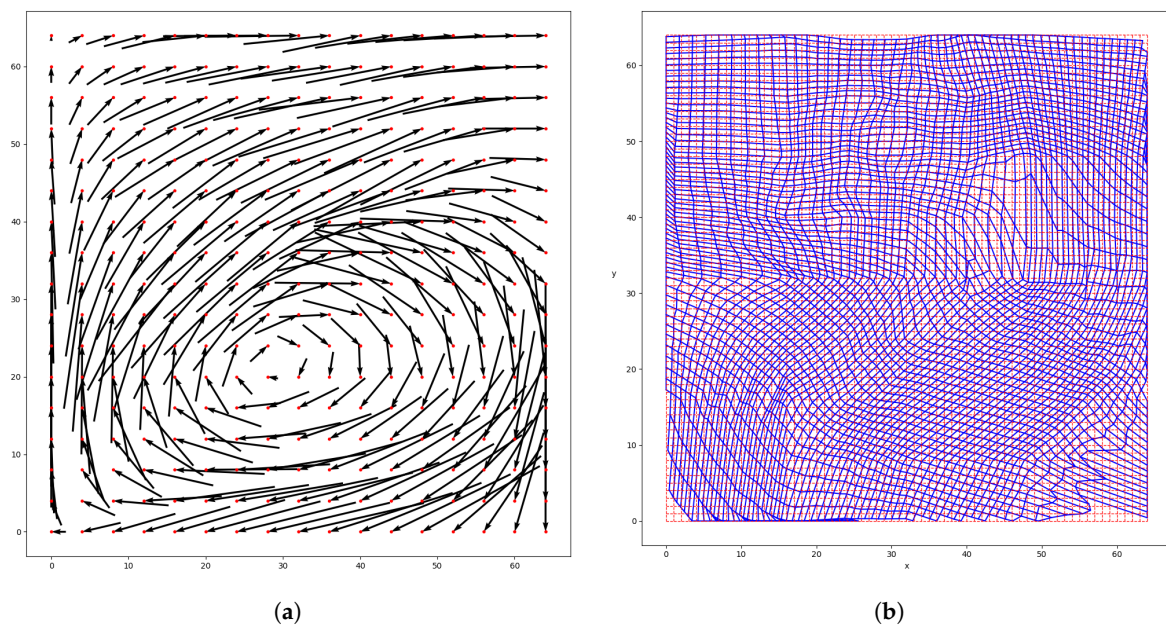
We now look at the impact of these coefficients on the mapping  $T$ . Figures A1–A4 show the mapping  $T$  and the resulting grid deformation for the four new cases. They can be compared to Figure 7 for the original case. The main differences occur in areas without precipitation. Without the smoothing constraints, larger distortion occurs near the boundary near the lower-right corner (visible in Figures A1 and A3 but not in Figure A2). Similarly, the grid appears very stretched on the right upper side. The case with only the magnitude constraint ( $C_1 = 0.1$ ) exhibits the smooth transformation (Figure A1). The divergent-free condition favours rotation-like patterns, in order to conserve the grid cell's volume. The weight of the coefficient also has an impact on the mapping  $T$ . By comparing Figures 7a and A4a, one can see that the distortion is similar over the rainy area, while the spatial displacement is smaller over the non-rainy area. The distortion is smoother and very similar to the regular one in the upper-left corner, i.e. away from the rainy peak.

To summarize, without the three weak constraints the distorted grid shows unnatural distortion. However, the regulation coefficients mainly affect the mapping  $T$  in areas without rainfall. This is reflected in the small influence it has over the morphed or warped field (Table A1).

**Figure A1.** Sensitivity study:  $C_1 = 0.1$  and  $C_2 = C_3 = 0$ . (a) Mapping  $T$  obtained from the automatic registration and (b) its effect on the pixel grid. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .

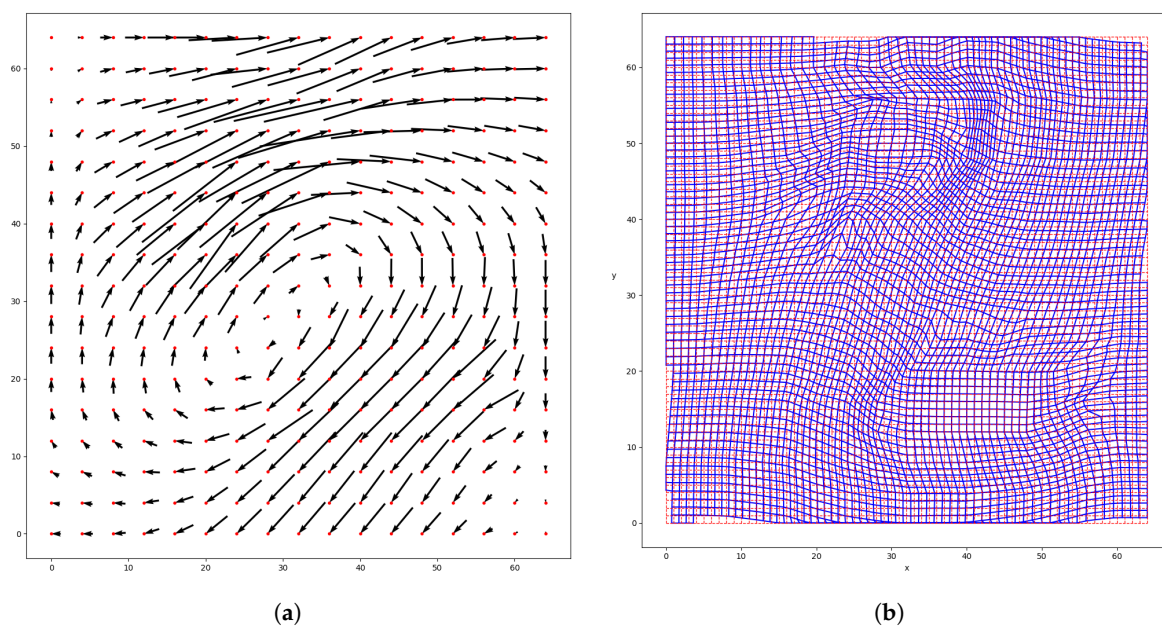


**Figure A2.** Sensitivity study:  $C_2 = 1$  and  $C_1 = C_3 = 0$ . (a) Mapping  $T$  obtained from the automatic registration and (b) its effect on the pixel grid. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .



**Figure A3.** Sensitivity study:  $C_3 = 1$  and  $C_1 = C_2 = 0$ . (a) Mapping  $T$  obtained from the automatic registration and (b) its effect on the pixel grid. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .





**Figure A4.** Sensitivity study:  $C_1 = 0.5$  and  $C_1 = C_2 = 5$ . (a) Mapping  $T$  obtained from the automatic registration and (b) its effect on the pixel grid. (a) Mapping  $T$ . (b) Pixel grid  $D_n$  before (in red) and after (in blue) distortion by the mapping  $T$ .

## References

- Engel, T.; Fink, A.H.; Knippertz, P.; Pante, G.; Bliedernicht, J. Extreme Precipitation in the West African Cities of Dakar and Ouagadougou: Atmospheric Dynamics and Implications for Flood Risk Assessments. *J. Hydrometeorol.* **2017**, *18*, 2937–2957. [\[CrossRef\]](#)
- Maranan, M.; Fink, A.H.; Knippertz, P. Rainfall types over southern West Africa: Objective identification, climatology and synoptic environment. *Q. J. R. Meteorol. Soc.* **2018**, *144*, 1628–1648. [\[CrossRef\]](#)
- Lorenz, C.; Kunstmann, H. The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. *J. Hydrometeorol.* **2012**, *13*, 1397–1420. [\[CrossRef\]](#)
- Morin, E.; Goodrich, D.C.; Maddox, R.A.; Gao, X.; Gupta, H.V.; Sorooshian, S. Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response. *Adv. Water Resour.* **2006**, *29*, 843–860. [\[CrossRef\]](#)
- Cristiano, E.; ten Veldhuis, M.C.; van de Giesen, N. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—A review. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 3859–3878. [\[CrossRef\]](#)
- Arnaud, P.; Bouvier, C.; Cisneros, L.; Dominguez, R. Influence of rainfall spatial variability on flood prediction. *J. Hydrol.* **2002**, *260*, 216–230. [\[CrossRef\]](#)
- Xiao, Q.; Zou, X.; Kuo, Y.H. Incorporating the SSM/I-Derived Precipitable Water and Rainfall Rate into a Numerical Model: A Case Study for the ERICA IOP-4 Cyclone. *Mon. Weather. Rev.* **2000**, *128*, 87–108. [\[CrossRef\]](#)
- Lopez, P. Direct 4D-Var Assimilation of NCEP Stage IV Radar and Gauge Precipitation Data at ECMWF. *Mon. Weather. Rev.* **2011**, *139*, 2098–2116. [\[CrossRef\]](#)
- Gilleland, E.; Ahijevych, D.; Brown, B.G.; Casati, B.; Ebert, E.E. Intercomparison of Spatial Forecast Verification Methods. *Weather. Forecast.* **2009**, *24*, 1416–1430. [\[CrossRef\]](#)
- Ahijevych, D.; Gilleland, E.; Brown, B.G.; Ebert, E.E. Application of Spatial Verification Methods to Idealized and NWP-Gridded Precipitation Forecasts. *Weather. Forecast.* **2009**, *24*, 1485–1497. [\[CrossRef\]](#)
- Briggs, W.M.; Levine, R.A. Wavelets and Field Forecast Verification. *Mon. Weather. Rev.* **1997**, *125*, 1329–1341. [\[CrossRef\]](#)
- Casati, B.; Ross, G.; Stephenson, D.B. A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteorol. Appl.* **2004**, *11*, 141–154. [\[CrossRef\]](#)

13. Casati, B. New Developments of the Intensity-Scale Technique within the Spatial Verification Methods Intercomparison Project. *Weather. Forecast.* **2010**, *25*, 113–143. [\[CrossRef\]](#)
14. Ebert, E.; McBride, J. Verification of precipitation in weather systems: determination of systematic errors. *J. Hydrol.* **2000**, *239*, 179–202. [\[CrossRef\]](#)
15. Davis, C.; Brown, B.; Bullock, R. Object-Based Verification of Precipitation Forecasts. Part I: Methodology and Application to Mesoscale Rain Areas. *Mon. Weather. Rev.* **2006**, *134*, 1772–1784. [\[CrossRef\]](#)
16. Marzban, C.; Sandgathe, S. Cluster Analysis for Verification of Precipitation Fields. *Weather. Forecast.* **2006**, *21*, 824–838. [\[CrossRef\]](#)
17. Hoffman, R.N.; Liu, Z.; Louis, J.F.; Grassotti, C. Distortion Representation of Forecast Errors. *Mon. Weather. Rev.* **1995**, *123*, 2758–2770. [\[CrossRef\]](#)
18. Hoffman, R.N.; Grassotti, C. A Technique for Assimilating SSM/I Observations of Marine Atmospheric Storms: Tests with ECMWF Analyses. *J. Appl. Meteorol.* **1996**, *35*, 1177–1188. [\[CrossRef\]](#)
19. Grassotti, C.; Iskenderian, H.; Hoffman, R.N. Calibration and Alignment. *J. Appl. Meteorol.* **1999**, *38*, 677–695. [\[CrossRef\]](#)
20. Nehrkorn, T.; Woods, B.K.; Hoffman, R.N.; Auligné, T. Correcting for Position Errors in Variational Data Assimilation. *Mon. Weather. Rev.* **2015**, *143*, 1368–1381. [\[CrossRef\]](#)
21. Aonashi, K.; Eito, H. Displaced Ensemble Variational Assimilation Method to Incorporate Microwave Imager Brightness Temperatures into a Cloud-resolving Model. *J. Meteorol. Soc. Jpn. Ser. II* **2011**, *89*, 175–194. [\[CrossRef\]](#)
22. Nehrkorn, T.; Woods, B.; Auligné, T.; Hoffman, R.N. Application of Feature Calibration and Alignment to High-Resolution Analysis: Examples Using Observations Sensitive to Cloud and Water Vapor. *Mon. Weather. Rev.* **2014**, *142*, 686–702. [\[CrossRef\]](#)
23. Brewster, K.A. Phase-Correcting Data Assimilation and Application to Storm-Scale Numerical Weather Prediction. Part I: Method Description and Simulation Testing. *Mon. Weather Rev.* **2003**, *131*, 480–492. [\[CrossRef\]](#)
24. Lawson, W.G.; Hansen, J.A. Alignment Error Models and Ensemble-Based Data Assimilation. *Mon. Weather Rev.* **2005**, *133*, 1687–1709. [\[CrossRef\]](#)
25. Ravela, S.; Emanuel, K.; McLaughlin, D. Data assimilation by field alignment. *Phys. D Nonlinear Phenom.* **2007**, *230*, 127–145. [\[CrossRef\]](#)
26. Keil, C.; Craig, G.C. A Displacement-Based Error Measure Applied in a Regional Ensemble Forecasting System. *Mon. Weather Rev.* **2007**, *135*, 3248–3259. [\[CrossRef\]](#)
27. Keil, C.; Craig, G.C. A Displacement and Amplitude Score Employing an Optical Flow Technique. *Weather Forecast.* **2009**, *24*, 1297–1308. [\[CrossRef\]](#)
28. Marzban, C.; Sandgathe, S. Optical Flow for Verification. *Weather Forecast.* **2010**, *25*, 1479–1494. [\[CrossRef\]](#)
29. Marzban, C.; Sandgathe, S.; Lyons, H.; Lederer, N. Three Spatial Verification Techniques: Cluster Analysis, Variogram, and Optical Flow. *Weather Forecast.* **2009**, *24*, 1457–1471. [\[CrossRef\]](#)
30. Han, F.; Szunyogh, I. A Morphing-Based Technique for the Verification of Precipitation Forecasts. *Mon. Weather Rev.* **2016**, *144*, 295–313. [\[CrossRef\]](#)
31. Alexander, G.D.; Weinman, J.A.; Schols, J.L. The Use of Digital Warping of Microwave Integrated Water Vapor Imagery to Improve Forecasts of Marine Extratropical Cyclones. *Mon. Weather Rev.* **1998**, *126*, 1469–1496. [\[CrossRef\]](#)
32. Beezley, J.D.; Mandel, J. Morphing ensemble Kalman filters. *Tellus A* **2008**, *60*, 131–140. [\[CrossRef\]](#)
33. Mandel, J.; Beezley, J.D.; Eben, K.; Jurus, P.; Kondratenko, V.Y.; Resler, J. *Data Assimilation by Morphing Fast Fourier Transform Ensemble Kalman Filter for Precipitation Forecasts Using Radar Images*; UCD/CCM Report 289; Center for Computational Mathematics, University of Colorado at Denver: Denver, CO, USA, April 2010.
34. Jongen, H.T.; Meer, K.; Triesch, E. Penalty-, Barrier-, Multiplier-, Interior Point-Methods. In *Optimization Theory*; Springer: Boston, MA, USA, 2004; pp. 153–172. [\[CrossRef\]](#)
35. Nocedal, J.; Wright, S.J. Penalty and Augmented Lagrangian Methods. In *Numerical Optimization*; Springer: New York, NY, USA, 2006; pp. 497–528. [\[CrossRef\]](#)
36. Dezfuli, A.K.; Ichoku, C.M.; Huffman, G.J.; Mohr, K.I.; Selker, J.S.; van de Giesen, N.; Hochreutener, R.; Annor, F.O. Validation of IMERG Precipitation in Africa. *J. Hydrometeorol.* **2017**, *18*, 2817–2825. [\[CrossRef\]](#)
37. Dezfuli, A.K.; Ichoku, C.M.; Mohr, K.I.; Huffman, G.J. Precipitation Characteristics in West and East Africa from Satellite and in Situ Observations. *J. Hydrometeorol.* **2017**, *18*, 1799–1805. [\[CrossRef\]](#)

38. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.; Kidd, C.; Nelkin, E.J.; Sorooshian, S.; Tan, J.; Xie, P. *NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG)*; Algorithm Theoretical Basis Document Version 5.2; NASA GSFC: Greenbelt, MD, USA, 2018. Available online: [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V5.2\\_0.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2_0.pdf) (accessed on 16 September 2019).
39. Van de Giesen, N.; Hut, R.; Selker, J. The Trans-African Hydro-Meteorological Observatory (TAHMO). *Wiley Interdiscip. Rev. Water* **2014**, *1*, 341–348. [[CrossRef](#)]
40. Soenario, I.; Sluiter, R. *Optimization of Rainfall Interpolation*; Technical Report; Ministerie van Verkeer en Waterstaat, Koninklijk Nederlands Meteorologisch Instituut: De Bilt, The Netherlands, 2010.
41. Thiemeig, V.; Rojas, R.; Zambrano-Bigiarini, M.; Roo, A.D. Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *J. Hydrol.* **2013**, *499*, 324–338. [[CrossRef](#)]
42. Beck, H.E.; Vergopolan, N.; Pan, M.; Levizzani, V.; van Dijk, A.I.J.M.; Weedon, G.P.; Brocca, L.; Pappenberger, F.; Huffman, G.J.; Wood, E.F. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 6201–6217. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).