A new methodology for pixel-quantitative precipitation nowcasting using a pyramid Lucas Kanade optical flow approach

Yu Liu, Du-Gang Xi, Zhao-Liang Li, Yang Hong

Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100101, PR China
Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, PR China
The PLA Information Engineering University, Zhengzhou 450001, PR China
Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, PR China
ICube, UDS, CNRS, 300 Blvd Sebastien Brant, BP10413, 67412 Illkirch, France
Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, PR China

SUMMARY

Short-term high-resolution Quantitative Precipitation Nowcasting (QPN) has important implications for navigation, flood forecasting, and other hydrological and meteorological concerns. This study proposes a new algorithm called Pixel-based QPN using the Pyramid Lucas–Kanade Optical Flow method (PPLK), which comprises three steps: employing a Pyramid Lucas–Kanade Optical Flow method (PLKOF) to estimate precipitation advection, projecting rainy clouds by considering the advection and evolution pixel by pixel, and interpolating QPN imagery based on the space–time continuum of cloud patches. The PPLK methodology was evaluated with 2338 images from the geostationary meteorological satellite Fengyun-2F (FY-2F) of China and compared with two other advection-based methods, i.e., the maximum correlation method and the Horn–Schunck Optical Flow scheme. The data sample covered all intensive observations since the launch of FY-2F, despite covering a total of only approximately 10 days. The results show that the PPLK performed better than the algorithms used for comparison, demonstrating less time expenditure, more effective cloud tracking, and improved QPN accuracy.

1. Introduction

Short-term high-resolution Quantitative Precipitation Forecasting (QPF), which refers to the forecasting of future precipitation within a very short time (i.e., 0–2 h), is important for a number of hydro-meteorological applications such as flash-flood warnings and navigation safety (Ganguly and Bras, 2003; Afshar et al., 2010).

Extrapolation-based Quantitative Precipitation Nowcasting (QPN) algorithms, which extract information from current observations (e.g., radar and satellite images), are reportedly capable of producing reliable forecasts, especially within a few hours of the analysis time (Dixon and Wiener, 1993; Johnson et al., 1998; Germann and Zawadzki, 2002, 2004; 2006; Chiang et al., 2006; Vila et al., 2008; Zahraei et al., 2012, 2013; Sokol and Pesice, 2012). These forecasts can complement the QPF of Numerical Weather Prediction (NWP) models (Golding, 1998; Ganguly and Bras, 2003; Wilson et al., 2004; Sokol, 2006; Sokol and Pesice, 2012; Liang et al., 2010) because a QPF model has the limitation of a short lead time (0–6 h) that stems from the significant computational costs, incomplete data assimilation during initialization, and incompatible scales of precipitation structures (Zahraei et al., 2012, 2013).

Currently, several extrapolation-based nowcasting algorithms have been developed for hydrological applications. For example, the system for Convection Analysis and Nowcasting (SCAN) developed by the Meteorological Development Lab (MDL) (Smith et al., 1998), the Warning Decision Support System Integrated Information (WDSS-II) of the National Severe Storm Laboratory (NSSL) (Lakshmanan et al., 2007), and the Auto-Nowcaster (ANC) of NCAR (the National Center for Atmospheric Research) (Wilson et al., 1998) are three examples in the U.S.A. The Nowcasting and Initialization for Modeling Using Regional Observation Data System (NIMROD) (Golding, 1998) and the Generating Advanced Nowcasts for Deployment in Operational Land surface Flood...
forecasts (GANDOLF) have played critical roles in England (Pierce et al., 2000).

Pixel- and object-based algorithms are two QPN categories. Object-based algorithms, which consider storm events as individual objects (Johnson et al., 1998; Dixon and Wiener, 1993), always have difficulty tracking small-scale thunderstorms and rapidly developing cloud patches that are identified as consistent clouds of some type from satellite images due to their complex and flexible characters. Pixel-based algorithms, which are the main focus of this study, treat clouds from a pixel perspective; they have been used extensively in short-term, high-resolution QPN (Grecu and Krajewski, 2000; Mecklenburg et al., 2000; Germann and Zawadzki, 2002; Montanari et al., 2006; Van-Hull et al., 2008; Berenguer et al., 2011; Zahraei et al., 2012, 2013).

An extrapolation-based QPN algorithm usually consists of tracking and forecasting (extrapolation) processes, as suggested by Austin and Bellon (1974). Various attempts have been made to estimate advection based on cloud tracking. The Maximum Correlation Method (MCM), which identifies the most matching locations between two successive images to estimate storm velocity (Smythe and Znic, 1983; Tuttle and Foote, 1990; Laroche and Zawadzki, 1995), is the most widely used approach. It is simple, but the correlation surfaces frequently display diffuse or multiple optima. The minimum velocity unit obtained by the MCM is one pixel, while the cloud always moves with sub-pixel accuracy. Thus, it may introduce large uncertainty into QPN with a long lead time and coarser spatial resolution scans (e.g., from a geostationary satellite). New sub-pixel approaches have been proposed, such as Variational Echo Tracking (VET) (Germann and Zawadzki, 2002), coupled hierarchical-tracking approaches and mesh-based models of image deformation in Pixel-Based Nowcasting (PBN) (Zahraei et al., 2012). In addition to the accuracy issues, the time expenditure of these methods is another important limiting factor in real-time hydro-meteorological applications.

The optical flow method, a computer vision method, estimates the motion of objects in a visual scene when projected onto a two-dimensional plane (Aubert et al., 1999). It can track objects on a sub-pixel level, and it has been widely applied in many areas for anticipated accuracy improvements, e.g., motion detection, object segmentation, time-to-collision and focus-of-expansion calculations, motion-compensated encoding, and stereo-disparity measurements (Beauchemin and Barron, 1995). The method can be divided into two main categories: local differential and global variational solutions. Since the first introduction by Bowler et al. (2004), the optical flow method has been applied in QPN studies, such as the Multi-scale Optical-flow by Variational Analysis (MOVA) scheme in the Hong Kong Short-range Warning of Intense Rainstorms in Localized System (SWIRLS) (Cheung and Yeung, 2012) and the Horn–Schunck Optical flow (HSOF) in Severe Weather Automatic Nowcast system (SWAN) of China (Han et al., 2008). Most of those studies are based on radar data, which always has challenges when applied to satellites because QPN of satellite tends to demand a higher accuracy in tracking methods on the sub-pixel level for the small movement speed of clouds with coarser spatial resolutions. Further, the smoother spatial characteristics of satellite products tend to make tracking cloud movement more difficult in the overlap region of two consecutive images because of the lack of obvious tracking signs compared with an equivalent terrestrial radar product. In addition, most applied optical flow methods are based on global variational solutions. They have some limitations in real-time QPN applications due to their computationally expensive nature despite the improved accuracy of the advection field. The local variational method is computationally more efficient and has exhibited good accuracy in various applications. However, the only application of the local variational method was the Lucas–Kanade optical flow method (LKOF) in the Cumulonimbus Tracking and Monitoring system (Cb-TRAM) (Zinner et al., 2008), which was only concerned with Cb cloud tracking rather than QPN.

Projection discontinuity is an important problem that is commonly confronted in the forecasting process of pixel-based QPN. A low-pass spatial filter, which averages the advection field of each rainy pixel with its surrounding window, is often used for addressing this problem. However, the method loses some information of the estimated precipitation advection and may also fail to describe the rapidly developing character of rainy cloud patches, especially for small-scale storms and rapidly expanding storms, due to the pre-smoothing process. Therefore, this study proposed to directly conduct an extrapolation process without using a low-pass spatial filter and add a new process of precipitation interpolation to address the projected discrete points produced in the extrapolation process.

The main contributions of this study can be summarized as follows: (1) the introduction of a typical local differential optical flow method, the Pyramid Lucas Kanade Optical Flow method (PLKOF), to track precipitation advection; (2) the projection of rainfall by considering precipitation advection and evolution using a linear pixel extrapolation method; (3) the presentation of a spatial interpolation method based on the spatio-temporal characteristics of cloud patches; (4) the establishment a QPN model using the geostationary meteorological satellite Fengyun-2F (FY-2F) to provide large-scale QPN information; and (5) the estimation of the performance of the proposed method with a comparison of two other advection-based methods, i.e., PMC (pixel-based QPN using MCM and PHS (pixel-based QPN using HSOF). The MCM is the most widely used method in cloud tracking, and HSOF is a newly adopted optical method.

In this study, Section 2 describes the implementation of the proposed model. The applied datasets and cases are shown in Section 3. Their performances are presented in Section 4. Section 5 concludes the results, discusses the limitations of the current implementation, and outlines possible future research.

2. Methodology

A traditional QPN algorithm consists of tracking and forecasting (extrapolation) processes, as suggested by Austin and Bellon (1974). Considering that pixel-based extrapolation methods always produce projection discontinuities, as presented in Fig. 1 (extrapolation), this study suggests adding a new spatial interpolation process to assign data to non-data rainy pixels.

Thus, the proposed PPLK (Pixel-based short-term QPN using PLKOF) comprises three processes, as presented in Fig. 1: pixel-based tracking, rainfall extrapolation, and spatial interpolation. Using the pixels with a black center as a sample to demonstrate the implementation of the PPLK, the steps are as follows: (1) Pixel-based tracking: extracting the pixel advection between \( t - \Delta t \) and \( t \) on the pixel level using PLKOF; (2) extrapolation: obtaining QPN images with the projection of rainy clouds considering both the precipitation advection and change in rainfall intensity pixel-by-pixel; and (3) interpolation: assigning data to non-data rainy pixels with the inverse distance method on the assumption that the topological relationships between cloud cells do not change between two consecutive times of \( t \) and \( t + \Delta t \).

2.1. Pixel-based tracking

Advection is a key element in storm movement and QPN. This study compared three advection-extracting methods: MCM, HSOF, and PLKOF. The first method is the most widely used in
QPN (e.g., Smythe and Zrnic, 1983; Tuttle and Foote, 1990; Laroche and Zawadzki, 1995). The second method was recently introduced (Cheung and Yeung, 2012; Han et al., 2008). The third method is the proposed method in this study.

2.1.1. MCM

Cloud-pixel advection between two consecutive images is derived by a maximum correlation measure to identify the most matching locations. First, the whole domain of the image is divided into a number of equally sized windows. Then, the patterns of pixels within a window are compared in a given area of the subsequent image to identify the potential corresponding locations as determined by the highest cross-correlations. To reduce computation time, a search radius that determines the maximum possible speed of the rainfall systems should be assigned, and it should consider the spatial resolution of the scans.

2.1.2. HSOF

The basic idea of optical flow is to derive a continuous motion field from temporal and spatial derivatives of two subsequent images (Sonka et al., 1993).

The “optical flow method” used in QPN is based on the brightness constancy assumption, which is essentially the same as the assumptions used in other methods, such as Lagrangian persistence in VET (Laroche and Zawadzki, 1995; Germann and Zawadzki, 2002).

The optical flow method is based on the following mathematical constraint equation:

\[ I_x u_t + I_y v_t + I_t = \nabla I \cdot \mathbf{v} + \frac{dI}{dt} \tag{1} \]

where \( I(x, y, t) \) is the rainfall intensity at the pixel \((x, y)\) at time \(t\). \( v_m = (u, v) \) is the quantity to be derived, and \( u \) and \( v \) are the advection-field components in the \( x \) direction (west–east) and the \( y \) direction (north–south), respectively. \( I_x \), \( I_y \), and \( I_t \) are the partial derivatives of \( x \), \( y \), and \( t \) for rainfall intensity, respectively, and \( \frac{dI}{dt} \) is the change in rainfall intensity observed in the pixel \((x, y)\). The optical flow method assumes no changes in the rain intensity in consecutive images \( \frac{dI}{dt} = 0 \). Because there are two unknowns, \( u \) and \( v \), but only one constraining equation, there is no unique solution to Eq. (1). To address this aperture problem, additional constraints were proposed, resulting in different algorithms and solutions to the optical flow constraint equation (see Aubert et al. (1999) for a review). Described below are two algorithms relevant to the present study.

The classical approach was pioneered by Horn and Schunck (1981), who formulated the optical flow constraint as a global variational problem by minimizing the following cost function:

\[ J = J_0 + \alpha \times J_{HS} \tag{2} \]

where

\[ J_0 = \int \int [I_t + I_x u + I_y v]^2 \, dx \, dy \tag{3} \]

and \( J_{HS} \) is a global constraint on the smoothness of the gradient of the optical flow field:

\[ J_{HS} = \int \int (|\nabla u|^2 + |\nabla v|^2) \, dx \, dy \tag{4} \]

Here, \( \alpha \) is a tunable parameter that determines the weight of the smoothness term.

![Fig. 1. The proposed PPLK algorithm (pixel-based short-term quantitative precipitation nowcasting (QPN) using the Pyramid Lucas–Kanade optical flow method (PLKOF)). The pixel \( n \) in the image at time \( t - \Delta t \) and forecasted \( t + \Delta t \) are \( a_n \) (e.g., \( a_1, a_2 \ldots a_{64} \)), \( b_n \), and \( c_n \), respectively. Yellow, purple and pink pixels are precipitation pixels at times \( t - \Delta t \), \( t \) and \( t + \Delta t \), respectively. Blue pixels and green pixels were situated in a cloud patch and on the boundary of a cloud patch, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)](image-url)
2.1.3. PLKOF

The aperture problem of the optical flow constraint mentioned previously can also be resolved locally, e.g., by the local differential method introduced by Lucas and Kanade (1981). In this method, the optical flow of the point \( q \) is assumed to be constant in a neighborhood of size \( \Omega \) and satisfies the following linear equations:

\[
L_i(q) \cdot u + L_j(q) \cdot v = -L_k(q) \quad \text{where} \quad q \in \Omega
\]  

(5)

Here, \( L_i, L_j, L_k \) is the corresponding partial derivative evaluated at point \( q \). Adapting a least-square principle and applying weights to different points in the neighborhood through Gaussian convolution, Eq. (5) can be solved with the following solution:

\[
\begin{align*}
\begin{bmatrix} u \\ v \end{bmatrix} &= \begin{bmatrix} K_0 + (L_i I_x) & K_0 + (L_j I_y) \\ K_0 + (L_i I_y) & K_0 + (L_j I_y) \end{bmatrix}^{-1} \begin{bmatrix} -K_0 + (L_i I_x) \\ -K_0 + (L_j I_y) \end{bmatrix} \end{align*}
\]  

(6)

where the operator \( * \) denotes convolution and \( K_0 \) is a Gaussian kernel with standard deviation \( \rho \) (Cheung and Yeung, 2012).

Although the local scheme is fast compared with the global one, it has limitations in tracking fast-moving objects with advection at more than one pixel due to the optical flow constraint equation. To address this problem, the image pyramid technique (Burt, 1981) was implemented; it can detect large-displacement and small-displacement motions using a "Gaussian image pyramid" multi-resolution data structure. In addition, it can also enhance the robustness of the optical flow method against noise, which was common in the rain intensity field due to the pre-smoothing of the Gaussian convolution.

The implementation of the PLKOF is as follows:

(a) Build pyramid: establish pyramid multi-resolution images of \( m + 1 \) levels (level 0, 1, ..., \( m \)) with FY-2F rainfall estimation images. Level m is the top of the pyramid. Level 0 is the bottom, which is also the original rainfall estimation image. In this study, 4 levels of pyramid images with spatial resolutions of 0.1°, 0.2°, 0.3°, and 0.4° from the bottom to the top were used. Their box sizes started at the full domain.

(b) Estimate the derivative of a rainy pixel \((k, l)\) with respect to the \( u_{nl}(k, l) \) in the \( x \) direction (west–east) and \( v_{nl}(k, l) \) in the \( y \) direction (north–south) for the top of the pyramid (level \( m \)) using LKOF with a Gaussian kernel before processing.

(c) Assume the level \( m = 1, m = 2, ..., 0 \) is level \( i \), and estimate the LKOF of pixel \((k, l)\) at level \( i \):

\[
u_i(k, l) = u_i(k, l) + u_i(k, l)
\]  

(7)

\[
u_i(k, l) = v_i(k, l) + v_i(k, l)
\]  

(8)

where \( u_i(k, l) \) and \( v_i(k, l) \) are the compensation optical flows of pixel \((k, l)\) at level \( i \). They are twice the compensation LKOF estimated at level \( i + 1 \) using the bilinear interpolation method. \( u_i(k, l) \) and \( v_i(k, l) \) are the standard LKOFs of the compensated image of level \( i + 1 \); repeat this step until the estimation of the optical flow field for all pixels at level 0 (original image) is finished.

(d) Precipitation advection quality control

Considering the determination of precipitation advection is a challenging issue in remote sensing sequences because precipitation clouds are deformable in nature. In addition, it has been observed that clouds commonly evolve through a sequence of pulse-like events and are subject to splits and mergers during their lifecycles, which may complicate the tracking. Thus, the forward and backward precipitation advection of pixels, which was estimated forward from time \( t - 1 \) to \( t \) and backward from time \( t \) to \( t - 1 \), is not always consistent. Therefore, quality control of the precipitation advection is conducted immediately after each computation step. Normally, the ultimate precipitation advection is estimated based on the average of the forward and backward precipitation advection.

Considering the probable maximum wind speed of 150 m/s, cloud advection is not more than 10 pixels in FY-2F-based quantitative precipitation estimation images with a spatial resolution of 0.1° and a temporal resolution of 6 min. If the difference in the forward and backward precipitation advection estimated previously is greater than 10 pixels, then the smaller one was chosen because the unreasonably large difference may stem from the complex non-linear development of cloud patches, such as the birth, rapid expansion, and dissipation of cloud patches.

2.2. Rainfall extrapolation

As for the semi-Lagrangian method (Turner et al., 2004; Germann and Zawadzki, 2002, 2004), the advected value should be redistributed to the neighboring pixels in a forward scheme, and interpolation is required in a backward scheme. Both redistribution and interpolation always result in a loss of power at small scales (Germann and Zawadzki, 2002). Thus, this paper attempted to use a simple linear-based method, which does not need to redistribute and interpolate the advected value despite the fact that it does not allow for large-scale rotational motion unlike the semi-Lagrangian method. The complement of the linear-based rain projection (extrapolation) considered both the precipitation advection and evolution.

(a) Project the cloud forward and backward using the linear extrapolation method and the precipitation advection obtained in Section 2.1 on the pixel level. In a forward scheme we start at purple pixels \((t)\) and advect them downstream to the purple pixels \((t + 1)\), whereas in a backward scheme we move upstream and determine the original purple pixels \((t)\) that would end up at yellow grid point pixels \((t - 1)\).

(b) Estimate the precipitation evolution on the assumption that the rainfall rate varied linearly and can be calculated based on the observed and the backward forecasted precipitation at time \( t \) according to step (a).

(c) Forecast the rainfall intensity at time \( t + 1 \) pixel-by-pixel by considering the precipitation advection and evolution based on step (a) and step (b). It can be described as below:

\[
\begin{align*}
P_{t = n \Delta t} (x_i + n \Delta x, y_j + n \Delta y) &= \max \{ \min [P_t (x_i, y_j)] + n \Delta P, \text{Threshold}_{\text{max}} - \text{Threshold}_{\text{min}} \} \\
\end{align*}
\]  

(9)

\[
(x_i + n \Delta x, y_j + n \Delta y) = \left( x_i, y_j \right) + n \Delta (x_i, y_j)
\]  

(10)

where \( P_{t = n \Delta t} \) is the predicted rain rate at \( t + n \Delta t \), \( (x_i + n \Delta x, y_j + n \Delta y) \) is the predicted location of pixel \((x_i, y_j)\). The \( \text{Threshold}_{\text{max}} \) and \( \text{Threshold}_{\text{min}} \) are the maximum and minimum possible rain rate of 50 mm/h and 0 mm/h.

2.3. Spatial interpolation

Considering the space–time continuum of cloud patches, this study interpolated the QPN images obtained in Section 2.2 with the assumption that the topological relationships of the cloud pixels do not change in consecutive images. The implementation process is shown as follows.

(a) Identify pixels in the extrapolation image that are neighboring in the previous image (e.g., pink pixels with a black center in Fig. 1 (extrapolation): \( C_{15}, C_{22}, C_{43}, \) and \( C_{46} \).)
(b) Determine whether the non-data pixels (white pixels) in Fig. 1 (extrapolation) are located in the square with the four pink pixels as vertices. Thus, four white pixels centered with a black point ($C_{28}$, $C_{29}$, $C_{36}$, and $C_{37}$) are found.

(c) Estimate the value of the identified pixels in step (b) using the four pink pixels ($C_{19}$, $C_{22}$, $C_{43}$, and $C_{46}$) and the inverse distance method. Then, mark them as blue as presented in Fig. 1 (extrapolation).

(d) Repeat steps (b and c) until all white pixels surrounded by four neighboring pink pixels are found. Thus, the blue pixels in Fig. 1 (interpolation) are identified and estimated. Those pixels are situated in a cloud patch.

(e) Estimate the value of the white pixels surrounded by three neighboring pink pixels in a similar way as used with the four points presented in steps (b–d) but only using three points; then, mark them as green in Fig. 1 (interpolation). Those estimated pixels are situated on the boundary of a cloud patch.

### Table 1

Information for 2338 imageries of 8 periods, including time, length, and spatial coverage.

<table>
<thead>
<tr>
<th>No. time period</th>
<th>Time (ymmdddhhmm)</th>
<th>No. of imageries</th>
<th>Length (h)</th>
<th>Spatial coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>201307170600–201307181554</td>
<td>341</td>
<td>34.1</td>
<td>60°–160°E, 12°–50°N</td>
</tr>
<tr>
<td>T2</td>
<td>201307240400–201307251154</td>
<td>319</td>
<td>31.9</td>
<td>60°–160°E, 25°–50°N</td>
</tr>
<tr>
<td>T3</td>
<td>201308010412–201308030354</td>
<td>478</td>
<td>47.8</td>
<td>60°–160°E, 10°–50°N</td>
</tr>
<tr>
<td>T4</td>
<td>201308051200–201308071154</td>
<td>480</td>
<td>48</td>
<td>60°–160°E, 5°–50°N</td>
</tr>
<tr>
<td>T5</td>
<td>201308120400–201308121148</td>
<td>80</td>
<td>8</td>
<td>60°–160°E, 10°–50°N</td>
</tr>
<tr>
<td>T6</td>
<td>201309200400–201309200936</td>
<td>55</td>
<td>5.5</td>
<td>60°–160°E, 10°–50°N</td>
</tr>
<tr>
<td>T7</td>
<td>201309220900–201309230330</td>
<td>128</td>
<td>12.8</td>
<td>60°–160°E, 10°–50°N</td>
</tr>
<tr>
<td>T8</td>
<td>201310050306–201310071454</td>
<td>457</td>
<td>45.7</td>
<td>60°–160°E, 18°–50°N</td>
</tr>
</tbody>
</table>

Fig. 2. Consecutive images of estimated precipitation for 3 periods (T1, T2 and T8). T1, T2, and T8 are the same as in Table 1. Δt = 4 h. The estimated precipitation images for T1 (a) and (b); T2 (c) and (d); and T3 (e) and (f) at 08:00 and 12:00 UTC. The units of the x and y directions are °E and °N, respectively.
2.4. Verification procedure

To demonstrate the precipitation advection derived by the various cloud-tracking methods, this study analyzed two aspects: a shift of 2 pixels to the left (west) from the original image and from two consecutive observed images (precipitation estimation products). The first approach helps to estimate the ability of cloud tracking without cloud deformation, while the latter shows the practical precipitation advection extracted.

The performance of the PPLK was evaluated with two indexes: computational cost (time) and precision. The first index is demonstrated by the time consumption, and the latter is indicated by the relative bias (Bias), coefficient of correlation (Corr), normalized mean square error (NMSE), probability of detection (POD), false-alarm ratio (FAR), and critical success index (CSI). These evaluation indices are defined in the following manner:

\[
\text{Bias} = \frac{\sum_{i=1}^{N} P(i) - \sum_{i=1}^{N} \bar{P}(i)}{\sum_{i=1}^{N} P(i)} \times 100\%
\]

\[
\text{Corr} = \frac{\sum_{i=1}^{N} (P(i) - \bar{P}) \times \sum_{i=1}^{N} (\bar{P}(i) - \bar{P})}{\sqrt{\sum_{i=1}^{N} (P(i) - \bar{P})^2 \times \sum_{i=1}^{N} (\bar{P}(i) - \bar{P})^2}}
\]

\[
\text{NMSE} = \frac{\sum_{i=1}^{N} |P(i) - \bar{P}(i)|^2}{N}
\]

\[
\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}}
\]

\[
\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}}
\]

\[
\text{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}}
\]

where \(P(i)\) and \(\bar{P}(i)\) are the forecasted and observed rainfall intensity of pixel \(I\); \(N\) is the number of pixels in the prediction domain.

Fig. 3. Cloud advection of a convective storm (06:00 UTC on 17 July 2013) (a) derived by PLKOF for a shift of 2 pixels to the left (west) from the original image; (b) the same as (a) but for two consecutive observed images; (c) and (e) are the same as (a) but using the MCM and HSOF, respectively; (d) and (f) are the same as (b) but using the MCM and HSOF, respectively.
\( \overline{P} \) and \( \overline{P} \) are the average forecasted and observed precipitation, respectively; and \( \text{hits}, \text{misses}, \text{and false alarms} \) are the number of hits, failures, and false alarms with a rainfall threshold of 0.1 mm/h, respectively. The best QPN algorithm should produce Bias, NMSE and FAR = 0 and Corr, POD and CSI = 1.

### 3. Data and case studies

The PPLK model was implemented to predict the rainfall rate in the next 2 h using the geostationary meteorological satellite Fengyun-2F (FY-2F), which is the fourth operational geostationary meteorological satellite developed by China. FY-2F has a visible image and four infrared images available hourly in normal periods and half hourly in the flood season. It also has flexible capabilities to scan target areas with a high temporal resolution, such as when observing typhoons, strong convection and other weather disasters, with a frequency of 6–12 min/scan.

In this study, 2338 FY-2F precipitation estimation images of 8 periods during 2013 (Table 1), provided by the NMSC (National Satellite Meteorological Center of Chinese Meteorological Agency: http://www.nsmc.cma.gov.cn/NewSite/NSMC/Home/ Index.html), were used for applying and verifying the proposed algorithm. They had 0.01 spatial and 6 min temporal resolutions and covered a rectangular region between 60°E–160°E and 0°N–50°N, depending on the precipitation system, with the task of observing the main participation system occurring in China.

The precipitation estimation from FY-2F was based on a statistical relationship between the gauge rain and multi-channel infra-red light temperature (IR1, 10.3–11.3 \( \mu \)m; IR2, 11.5–12.5 \( \mu \)m and WV 6.3–7.6 \( \mu \)m) which was established firstly for FY-2C and then used in other FY series satellites, such as FY-2D/E. There was significant intrinsic uncertainty in the FY-2F infra-red-based precipitation product as in other geostationary satellites. Because other high frequency observations data (e.g., radar and gauge data) at the large scale needed for the verification of satellite data were not available and the comparison with other geostationary satellites was difficult for different spatial and temporal scales, this paper mainly focused on the methodology and did not perform the verification of the FY-2F precipitation product.

Considering the co-existence of several storms in a single image, the images covered almost all types of storms in various situations, scales, structures, life cycles, and moving characters. Among these images, the storms in the second period (T2) and eighth period (T8), as shown in Figs. 2b and c, respectively, were relatively small scale and fast moving with complicated development processes.

### 4. Verification and results

#### 4.1 Precipitation advection

Fig. 3 is the precipitation advection of a strong convection sample derived by PLKOF vs. MCM and HSOF at 06:00 UTC on 17 July 2013. The first column (Fig. 3a, c, e) shows the precipitation advection results assuming that there was a shift of 2 pixels from the right (east) to the left (west) for the image of T1; The second column (Fig. 3b, e, f) was the same as the first column but for two consecutive images of FY-2F (T1 and T2).

The results show that PLKOF (Fig. 3a) and MCM (Fig. 3c) tracked clouds without deformation because the derived precipitation advection was the same as the original set, which was 2 pixels to the west. The HSOF method (Fig. 3e) usually underestimated the cloud velocity, especially on the cloud boundary and in the small rain-rate areas. The cloud motion direction always deviated from the actual direction, although its general direction was correct.

The cloud motion of two consecutive observed images show that the PLKOF (Fig. 3b) provides a better representation of the cloud motion than the MCM (Fig. 3d) and the HSOF (Fig. 3f) with respect to the spatial reasonability of the precipitation advection surfaces, i.e., smoother and fewer cloud-direction-opposing motion vectors. The cloud-advection surfaces of the MCM frequently displayed diffuse or multiple optima. Compared with the MCM, the HSOF improved the reasonability of cloud tracking because the main direction of cloud motion can be easily identified despite the existence of some unreasonable opposite-direction motion vectors in adjacent areas.

#### Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Cloud tracking</th>
<th>Extrapolation and interpolation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scope (pixels)</td>
<td>Time consumption</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MCM</td>
<td>5 × 5</td>
<td>0.38</td>
<td>0.89</td>
</tr>
<tr>
<td>PHS</td>
<td>25 × 25</td>
<td>2.67</td>
<td>0.04</td>
</tr>
<tr>
<td>PPLK</td>
<td>50 × 50</td>
<td>24.51</td>
<td>0.06</td>
</tr>
</tbody>
</table>

The cloud motion of two consecutive observed images show that the PLKOF (Fig. 3b) provides a better representation of the cloud motion than the MCM (Fig. 3d) and the HSOF (Fig. 3f) with respect to the spatial reasonability of the precipitation advection surfaces, i.e., smoother and fewer cloud-direction-opposing motion vectors. The cloud-advection surfaces of the MCM frequently displayed diffuse or multiple optima. Compared with the MCM, the HSOF improved the reasonability of cloud tracking because the main direction of cloud motion can be easily identified despite the existence of some unreasonable opposite-direction motion vectors in adjacent areas.

#### Table 3

<table>
<thead>
<tr>
<th>Lead time (min)</th>
<th>Rainfall event</th>
<th>Bias (%)</th>
<th>Corr</th>
<th>NMSE (mm/h)</th>
<th>POD</th>
<th>FAR</th>
<th>CSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 T1</td>
<td>10.82</td>
<td>0.82</td>
<td>1.41</td>
<td>0.89</td>
<td>0.37</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>30 T2</td>
<td>12.16</td>
<td>0.71</td>
<td>0.46</td>
<td>0.79</td>
<td>0.44</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>30 T3</td>
<td>6.24</td>
<td>0.85</td>
<td>1.66</td>
<td>0.90</td>
<td>0.33</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>30 T4</td>
<td>7.85</td>
<td>0.84</td>
<td>1.40</td>
<td>0.90</td>
<td>0.36</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>30 T5</td>
<td>7.68</td>
<td>0.86</td>
<td>1.97</td>
<td>0.91</td>
<td>0.35</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>30 T6</td>
<td>3.87</td>
<td>0.90</td>
<td>1.72</td>
<td>0.93</td>
<td>0.26</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>30 T7</td>
<td>8.85</td>
<td>0.85</td>
<td>1.47</td>
<td>0.91</td>
<td>0.34</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>30 T8</td>
<td>9.13</td>
<td>0.82</td>
<td>0.50</td>
<td>0.85</td>
<td>0.39</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>8.63</td>
<td>0.84</td>
<td>0.80</td>
<td>0.87</td>
<td>0.37</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>

The time consumption of cloud tracking was doubled as the ultimate precipitation advection was estimated based on forward and backward precipitation advection.

** The search scope of MCM.
*** This study divided the satellite imagery into several sub-regions to reduce the amount of computation. \( Pn \) (e.g., \( P1, P2 \ldots Pn \)) refers to the time consumption of QPN + n min lead time.

#### Table 3

<table>
<thead>
<tr>
<th>Lead time (min)</th>
<th>Rainfall event</th>
<th>Bias (%)</th>
<th>Corr</th>
<th>NMSE (mm/h)</th>
<th>POD</th>
<th>FAR</th>
<th>CSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>90 T1</td>
<td>21.59</td>
<td>0.60</td>
<td>2.56</td>
<td>0.69</td>
<td>0.51</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>90 T2</td>
<td>28.88</td>
<td>0.43</td>
<td>0.82</td>
<td>0.53</td>
<td>0.64</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>90 T3</td>
<td>12.74</td>
<td>0.63</td>
<td>3.17</td>
<td>0.69</td>
<td>0.47</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>90 T4</td>
<td>14.66</td>
<td>0.62</td>
<td>2.60</td>
<td>0.69</td>
<td>0.49</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>90 T5</td>
<td>15.60</td>
<td>0.65</td>
<td>3.70</td>
<td>0.71</td>
<td>0.47</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td>90 T6</td>
<td>5.81</td>
<td>0.73</td>
<td>3.15</td>
<td>0.77</td>
<td>0.36</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>90 T7</td>
<td>16.66</td>
<td>0.69</td>
<td>2.46</td>
<td>0.77</td>
<td>0.44</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>90 T8</td>
<td>16.61</td>
<td>0.60</td>
<td>0.85</td>
<td>0.65</td>
<td>0.51</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>17.55</td>
<td>0.60</td>
<td>2.16</td>
<td>0.67</td>
<td>0.55</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>

The time consumption of cloud tracking was doubled as the ultimate precipitation advection was estimated based on forward and backward precipitation advection.

** The search scope of MCM.
*** This study divided the satellite imagery into several sub-regions to reduce the amount of computation. \( Pn \) (e.g., \( P1, P2 \ldots Pn \)) refers to the time consumption of QPN + n min lead time.
4.2. Time consumption

Considering the FY-2F-based quantitative precipitation estimation dataset with a 0.1° spatial resolution (approximately 11.1 km × 11.1 km at the equator and approximately 5 km × 5 km at 50°N in the latitudinal direction), a 6-min temporal resolution, and a probable maximum wind speed of 150 m/s, the maximum search radius of the MCM was set to 10 pixels (360 s/scan × 150 m/s/5000 m). Then, a search radius of 5 pixels was used for tracking clouds with moderate speeds.

Due to the computationally expensive nature of HSOF, this study divided the FY-2F image into several sub-windows. Thus, two window sizes (25 × 25 and 50 × 50) were used based on the comprehensive analysis of the time expenditure and the deformation of the precipitation advection at the boundary of each window.

Table 2 presents the average time expenditure of QPN with a +120 min lead time using PPLK, PMC, and PHS for 30 images between 16:00 and 21:00 on September 4, 2013. It also shows that the average time consumed during cloud tracking is 0.26 min for PLKOF, 0.38 min and 0.63 min for the MCM with a search radius of 5 and 10 pixels, respectively, and 2.67 min and 24.51 min for the HSOF with sub-windows of 25 × 25 pixels and 50 × 50 pixels, respectively. The PPLK provides a better representation compared with the HSOF and the MCM with respect to the time efficiency.

The total time consumption of the extrapolation and interpolation was 0.89 min for a QPN with a +120 min lead time and a 6 min interval. More time is required for a longer lead time. For example, a QPN with a +5 min lead time needs 0.01 min, and a QPN with a +120 min lead time needs 0.08 min. A QPN with a longer lead time not only produces more discrete rainy points but also increases their distance. Thus, more time is needed in the interpolation process. Although the total time consumption of extrapolation and interpolation is large compared with that of cloud tracking, it can be reduced by increasing the time intervals of QPN, e.g., the time consumption of QPN for a +120 min lead time can be reduced from 0.89 min to 0.08 min with a time interval increase from 6 min to 120 min.

![Fig. 4. QPN with +30, +60, +90 and +120 min lead times for 8 selected periods with intervals of 6 min. (a) Relative bias (Bias), (b) coefficient of correlation (Corr), (c) normalized mean square error (NMSE), (d) probability of detection (POD), (e) false-alarm ratio (FAR), and (f) critical success index (CSI) of 2338 images for 8 periods. T1, T2, ..., T8 are the same as in Table 1.](image-url)
4.3. QPN results

4.3.1. QPN of PPLK

Table 3 and Fig. 4 present the 6 measures of QPN averaged over 2388 images of 8 periods using the PPLK method for +30, +60, +90, and +120 min lead times. The average Bias, Corr, NMSE, POD, FAR and CSI are 8.63%, 0.84, 0.80 mm/h, 0.87, 0.37 and 0.58 for a +30 min lead time, respectively, 12.59%, 0.72, 1.13 mm/h, 0.75, 0.42, 0.4 for a +60 min lead time, respectively, and 22.35%, 0.55, 1.62 mm/h, 0.63, 0.55, and 0.36 for a +120 min lead time, respectively.

The PPLK performs similarly well in those periods, except in T2 and T8, which comprised storms that were relatively small-scale and fast-moving and involved complicated development processes, such as numerous newly born and dissipated cloud patches, as presented in Fig. 1.

4.3.2. Comparison of the QPN method with different cloud-tracking methods

Fig. 5 presents the comparisons of PPLK vs. PMC, PHS and observations for the 6 measures (Bias, Corr, NMSE, POD, FAR and CSI) averaged over 177 images of the period T1 (06:18 UTC on 17 July to 00:00 UTC on 18 July 2013). For the optical flow-based QPN, both PPLK and PHS performed better than the PMC method with great improvements in the predictability of rainfall in terms of the 6 measures. In addition, PPLK performed better compared with the PHS with slight improvements in Corr, NMSE, FAR and CSI and obvious improvements in Bias and POD which stems from the small number of misses compared with hits and false alarms.

4.3.3. Comparison of the QPN method with other methods

To demonstrate the performance of the proposed method, it was compared with other studies, such as the 7 QPN methods analyzed by Zahraei et al. (2012, 2013) using GOES-IR (infrared) and radar data and 2 QPN methods by Han et al. (2008) using radar data.

The results show that regarding the QPN with a +30 min lead time, the POD, FAR, and CSI of the PPLK were 0.87, 0.37, and 0.58, respectively, while they were 0.54–0.68, 0.37–0.49, and 0.32–0.42 for the PERCAST-GD (PERsiann-ForeCAST considering storm Growth and Decay), which was the best QPN method based on the four storms in the GOES-IR imagery by Zahraei et al. (2013). The POD, FAR, and CSI values were 0.43–0.45, 0.50–0.53, and 0.29–0.32 for the optical flow method according to the radar-based precipitation estimation data of five storms in Han et al. (2008). Considering the statistical reasonability stemming from the numerous images (2338 images) used in this study and the co-existence of numerous storms in a single image, the PPLK performed better than the comparison methods despite the use of different cases and scans.

In terms of the QPN of PPLK, for a lead time of +120 min, the POD, FAR, and CSI were 0.63, 0.55, and 0.36, respectively. The accuracy of the PPLK with a +120 min lead time was comparable with that of the PERCAST-GD (Zahraei et al., 2013) and the optical flow method (Han et al., 2008) with a +30 min lead time to an extent.

4.4. The effects of satellite resolution on PPLK

The spatial characteristics of geostationary satellite products tend to be smoother than equivalent terrestrial radar products. The effect of the satellite resolution on the performance of the QPN was discussed.

Fig. 6 presents the average Corr and CSI of all events for 0–120 min predictions vs. spatial resolution (0.1°/C176, 0.2°/C176, 0.3°/C176, 0.4°/C176, 0.5°/C176, and 0.6°/C176). The results show that the CSI and Corr of QPN decreased with coarser spatial resolutions, except the Corr of a spatial resolution between 0.1° and 0.2° which increased slightly with coarser spatial resolution. The variation in QPN of different satellite resolutions was different from those of the QPN of radars for which the coarser spatial resolutions make better predictions (Zahraei et al., 2012). The reason for the inconsistency of QPN along different spatial resolutions and between satellite and radar is that besides improving the QPN performance by reducing the image noise, coarser resolutions also decrease the accuracy of QPN for failing to track and forecast some small and metro-scale clouds that cover a few or a dozen pixels.
5. Summary and conclusions

This study proposed a new algorithm, PPLK, which utilizes the pyramid Lucas–Kanade optical flow method to estimate precipitation advection, projects rain pixels based on a linear extrapolation method that considers precipitation advection and evolution, and adds a precipitation interpolation process based on the space–time continuum of cloud patches. The performance of PPLK was evaluated using 2338 images during July–October 2013 and compared against two other advection-based models, namely, the PMC and PHS algorithms.

The results demonstrate the following:

1. The precipitation advection surfaces extracted by PLKOF are smoother and contain fewer opposite-direction cloud motion vectors compared with the MCM and HSOF.
2. The proposed PPLK method performed better than the PMC, with significantly improved QPN accuracy and slightly reduced time expenditure.
3. Compared with the PHS, PPLK has obvious advantages in real-time hydro-meteorological applications due to its high computational efficiency and slight improvement in QPN accuracy.
4. The accuracy of the PPLK with a +120 min lead time was comparable, to an extent, with the methods of Zahraei et al. (2012, 2013) and Han et al. (2008), with a +30 min lead time, despite using different cases and scans.

Because this initial part of the current research is more focused on the methodology, more complex and advanced forecasting scenarios (e.g., individual cloud patches) still need to be investigated; however, the initial results are encouraging. The proposed algorithms need to be comprehensively evaluated (e.g., in cold seasons and over several years). The authors have begun an effort to improve the algorithm by testing more scenarios and will report the findings in future publications.

As the PPLK algorithm is effective and not computationally time consuming, it could be further exploited and better utilized in real-time hydro-meteorological applications with less time expenditure, anticipated accuracy improvements and large-scale spatial coverage.

This study proved the feasibility of performing QPN based on extracting information from the rainfall estimates of satellite. Thus, further application in other current observations, such as radar and even gauged data, could be exploited. However, the rainfall estimate of satellites has significant intrinsic uncertainty stemming from inferring precipitation indirectly, and its effect on the predictability of QPN was not discussed in this study. Further work is currently underway to analyze the performance of the PPLK using original satellite infra data in addition to the estimated rainfall.

Author contribution

Conceived and designed the experiments: YL, DGX, ZLL, HY. Performed the experiments of cloud tracking: YL, DGX. Performed the experiments of rainfall extrapolation and spatial interpolation: ZLL, HY. Verify the procedure: YL, DGX, LLZ. Wrote the paper: YL, ZLL.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 41301379 and 41231170 and the China Postdoctoral Science Foundation under Grant 2013M541029. The authors would like to thank the National Weather Service Information Center of China for providing FY-2F and meteorological data. The authors greatly appreciated the careful and insightful suggestions and comments of reviewers that helped to improve the manuscript and data analysis.

References


