Influence of Rain Gauge Density and Temporal Resolution on the Performance of Conditional Merging Method

Dao, Duc Anh*, Kim, Soohyun**, Kim, Tae-Woong***, and Kim, Dongkyun****

Abstract

This study aims to evaluate the performance of conditional merging (CM) technique to merge radar and in-situ rain gauge data. The merging method contains the following steps: (1) the ordinary kriging (OK) technique is applied to obtain the gauge-only and radar-only kriged rainfall fields; (2) the correction field is calculated by subtracting the radar-only kriged rainfall field from the original radar field; (3) the composite rainfall is derived by adding the correction field to the gauge-only kriged field. We investigated 33 10-min rainfall fields; there were 22 instances where the CM-based rainfall fields were more accurate than the gauge-only fields. The correlation coefficients of the CM and the gauge-only OK methods varied between 0.63 and 0.83, and 0.01 and 0.74 respectively. The performance of both methods improved with greater gauge density and coarser temporal resolution. Furthermore, the OK and CM techniques’ performances were more sensitive to gauge density and temporal resolution respectively. We also found that the addition of gauge significantly improved the performance of both methods when the existing gauge density is approximately 1 gauge/100 km². However, this effect is absent when the density reaches 4 gauges/100 km².

Key words: Ordinary Kriging, Conditional Merging, Radar, Rain Gauge

1. Introduction

The rainfall data with high spatial and temporal resolution are crucial for accurate hydrologic modelling (Kalinga and Gan, 2006; Cole and Moore, 2008; Li and Zhang, 2008; Langella et al., 2010; Li and Shao, 2010; Gebregiorgis and Hossain, 2012; Berne and Krajewski, 2013; Xu et al., 2013; Yao et al., 2014; Zhang et al., 2016).

In-situ gauge and weather radar are the two most widely applied measurement techniques to observe rainfall. The
In-situ rain gauges provide accurate measurements at point locations, but its accuracy decreases with the decrease gauge spatial density especially when rainstorms have significant spatial variability (Hughes, 2006; Beesley et al., 2009; Renard et al., 2010; Kim et al., 2019). A variety of interpolation methods have been developed to produce the spatial rainfall estimates based on in-situ gauge data, including ordinary kriging (Bárdossy, 1993), kriging with external drift (Hudson and Wackernagel, 1994; Goovaerts, 2000), cokriging (Bárdossy and Lehmann, 1998; Lloyd, 2005), nearest neighbor (Isaaks and Srivastava, 1989), spline fitting techniques (Hutchinson, 1998). On the other hand, radar measurement can provide rainfall estimate over a large spatial coverage with high spatio-temporal resolution (Qi et al., 2013; Zhang et al., 2016). However, the radar rainfall data is subject to greater uncertainty compared to the in-situ gauge rainfall data because it infers the rainfall intensity from the microwave backscatter signal (Wilson and Brandes, 1979; Goudenhoofdt and Delobbe, 2009; Shen et al., 2010; Wu and Zhai, 2012; Qin et al., 2014; Karbalaaee et al., 2017).

The techniques of merging ground and radar data have been proposed to overcome the drawbacks of both measurement methods (Krajewski, 1987; Burlando et al., 1996; Schiemann et al., 2011; Pettazzi et al., 2012). The idea of merging the two different rainfall estimates was first proposed by Krajewski (1987). Then, the approach was further improved with various algorithms (Creutin et al., 1988; Seo, 1998; Sinclair and Pegram, 2005; Chumchean et al., 2006; Velasco-Forero et al., 2009; Berndt et al., 2014; Sideris et al., 2014). Haberlandt (2007) employed the kriging with external drift method to spatially interpolate the hourly rain gauge data using the drift information extracted from the radar data. Li and Shao (2010) developed a nonparametric kernel interpolation method to combine the rain gauge and TRMM satellite data, which does not rely on the stationarity assumption as kriging methods. The kriging with radar error method was developed by Ehret (2003), and it was called the Conditional Merging (CM) by Sinclair and Pegram (2005). The method adjusts the gauge-only kriged rainfall field using the information on the spatial structure of the rainstorms that is derived from the radar data.

The objective of this study is to evaluate the performance of the CM method. There have been various studies evaluating the performance of this technique (Haberlandt, 2007; Kim et al., 2007; Chappel et al., 2013; Baik et al., 2016; Sun et al., 2016; Chao et al., 2018). However, those studies used hourly and daily rainfall data, which do not meet the requirement of temporal resolution and cause the uncertainty in flood simulation in urban area (Rafieeinasab et al., 2015). In contrast, this study employed the gauge rainfall data with the temporal resolution of 10-minutes, which is commonly used in urban flood analysis. In addition, we can investigate the effect of temporal resolution on the performance of rainfall estimation in the range of sub-hourly timescales, which was not done in previous studies. The influence of gauge density on performance was also considered in this study.

2. Methodology

2.1 Study area and rainfall data

The study area is Seoul, the capital of Korea. The city is highly urbanized, covering an area of 605 km² with 239 sub-watersheds (Fig. 1). The study used 1-minute rainfall from 29 rain gauges (green dots in Fig. 1) and 10-minute temporal.
resolution radar provided by the Korea Meteorological Administration (KMA). The detail information of 29 stations is described in Table 1. The temporal periods investigated in this study is between 2010/09/21 12:30 and 2010/09/21 17:50 which totals up to 9,280 gauges measurements and 33 radar images. Fig. 2 shows the 10-minute hyetograph and the accumulated rainfall of the average areal rainfall over Seoul for the rainfall event period. The rain event was caused by the collision of the Siberian High and the North Pacific High, which caused localized rainfall with great spatial variability. The accumulated rainfall over the event period is the second highest in the record history of Seoul.

![Fig. 2. The 10-minute Hyetograph and the Accumulated Rainfall of the Event of the September 2010 that was Investigated in this Study](image)

### Table 1. The Information of Gauge Stations

<table>
<thead>
<tr>
<th>No.</th>
<th>ID</th>
<th>Name</th>
<th>Latitude</th>
<th>Longitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>400</td>
<td>Gangnam</td>
<td>37.5134</td>
<td>127.0467</td>
</tr>
<tr>
<td>2</td>
<td>401</td>
<td>Seocho</td>
<td>37.4887</td>
<td>127.0156</td>
</tr>
<tr>
<td>3</td>
<td>402</td>
<td>Gangdong</td>
<td>37.5555</td>
<td>127.1450</td>
</tr>
<tr>
<td>4</td>
<td>403</td>
<td>Songpa</td>
<td>37.5115</td>
<td>127.0967</td>
</tr>
<tr>
<td>5</td>
<td>404</td>
<td>Gangseo</td>
<td>37.5739</td>
<td>126.8295</td>
</tr>
<tr>
<td>6</td>
<td>405</td>
<td>Yangcheon</td>
<td>37.5296</td>
<td>126.8782</td>
</tr>
<tr>
<td>7</td>
<td>406</td>
<td>Dongsan</td>
<td>37.5995</td>
<td>127.0295</td>
</tr>
<tr>
<td>8</td>
<td>408</td>
<td>Dongdaemoon</td>
<td>37.5846</td>
<td>127.0604</td>
</tr>
<tr>
<td>9</td>
<td>409</td>
<td>Jungrang</td>
<td>37.5855</td>
<td>127.0868</td>
</tr>
<tr>
<td>10</td>
<td>411</td>
<td>Mapo</td>
<td>37.5517</td>
<td>126.9292</td>
</tr>
<tr>
<td>11</td>
<td>412</td>
<td>Seodaemun</td>
<td>37.5705</td>
<td>126.9408</td>
</tr>
<tr>
<td>12</td>
<td>413</td>
<td>Gwangjin</td>
<td>37.5416</td>
<td>127.0805</td>
</tr>
<tr>
<td>13</td>
<td>414</td>
<td>Seongbuk</td>
<td>37.6117</td>
<td>126.9994</td>
</tr>
<tr>
<td>14</td>
<td>415</td>
<td>Yongin</td>
<td>37.5204</td>
<td>126.9761</td>
</tr>
<tr>
<td>15</td>
<td>416</td>
<td>Eunpyeong</td>
<td>37.6106</td>
<td>126.9339</td>
</tr>
<tr>
<td>16</td>
<td>417</td>
<td>Geumcheon</td>
<td>37.4655</td>
<td>126.9002</td>
</tr>
<tr>
<td>17</td>
<td>418</td>
<td>Hangang</td>
<td>37.5249</td>
<td>126.9391</td>
</tr>
<tr>
<td>18</td>
<td>419</td>
<td>Junggu</td>
<td>37.5524</td>
<td>126.9874</td>
</tr>
<tr>
<td>19</td>
<td>421</td>
<td>Seongdong</td>
<td>37.5472</td>
<td>127.0389</td>
</tr>
<tr>
<td>20</td>
<td>423</td>
<td>Guro</td>
<td>37.494</td>
<td>126.8270</td>
</tr>
<tr>
<td>21</td>
<td>424</td>
<td>Gangbuk</td>
<td>37.6397</td>
<td>127.0258</td>
</tr>
<tr>
<td>22</td>
<td>509</td>
<td>Gwanak</td>
<td>37.4528</td>
<td>126.9502</td>
</tr>
<tr>
<td>23</td>
<td>510</td>
<td>Yeongdeungpo</td>
<td>37.5271</td>
<td>126.9071</td>
</tr>
<tr>
<td>24</td>
<td>511</td>
<td>Gongchondong</td>
<td>37.555</td>
<td>126.6905</td>
</tr>
<tr>
<td>25</td>
<td>540</td>
<td>Goyang</td>
<td>37.6413</td>
<td>126.8774</td>
</tr>
<tr>
<td>26</td>
<td>541</td>
<td>Namyangju</td>
<td>37.6500</td>
<td>127.1800</td>
</tr>
<tr>
<td>27</td>
<td>569</td>
<td>Guri</td>
<td>37.5994</td>
<td>127.1312</td>
</tr>
<tr>
<td>28</td>
<td>589</td>
<td>Neunggok</td>
<td>37.6205</td>
<td>126.7943</td>
</tr>
<tr>
<td>29</td>
<td>590</td>
<td>Gwacheon</td>
<td>37.4263</td>
<td>126.9887</td>
</tr>
</tbody>
</table>

2.2 Conditional merging method

The conditional merging technique (Ehret, 2003; Sinclair and Pegram, 2005) assumes that the gauge provides the true source of spatial interpolation and the radar data provides auxiliary information on rainfall spatial variability to improve the accuracy (Goudenhoofdt and Delobbe, 2009; Rabiei and Haberlandt, 2015). The method has been proven to produce accurate spatial fields of various environmental variables (Sinclair and Pegram, 2005; Goudenhoofdt and Delobbe, 2009; Berndt et al., 2014; Baik et al., 2016; Kim et al., 2016).

Regarding to kriging method, it is assumed that a theoretical semivariogram model is to be fitted to an empirical one based on the normally distributed precipitation dataset. An appropriate estimate of the semivariogram value, \( \gamma(h) \), for point pairs separated by lag distance \( h \) is a measure representing the spatial variability of a regionalized variable \( z \) and is given by

\[
\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (z(u_i) - z(u_i + h))^2
\]

Where \( u_i \) and \( u_i + h \) refer point pairs with the distance \( h \), and \( N(h) \) is the number of point pairs with the same lag distance ; \( z(u_i) \) and \( z(u_i + h) \) indicate the observed variable at locations \( u_i \) and \( u_i + h \), respectively, \( i \) is the number of data values. The set of estimates with each lag distance is designated as the empirical semivariogram which will be used to fit theoretical semivariogram by adopting a weighted least squares strategy to seek the best variogram parameters (Robinson and Metternicht, 2006). In this study, we used the isotropic exponential model (Eq. (2)) which was fitted to experimental variogram.
\[ \gamma_h = C_0 + C_1 \left[ 1 - \exp \left( -\frac{3h}{a} \right) \right] \]  

(2)

Here, \( a \) is range, \( C_1 \) the sill and \( C_0 \) the nugget effect.

Fig. 3 describes the CM technique applied in this study. This technique has the following procedures:

(a) One-minute ground-observed rainfall data is converted into 10-minute format to be synchronized with the measurement interval of the radar data. Then, gauge rainfall data were interpolated using Ordinary Kriging technique to obtain the gauge-only kriged rainfall field.

(b) Collect the radar rainfall data.

(c) The radar pixel values at the gauge locations were picked up and apply Ordinary Kriging technique to obtain the radar-only kriged rainfall field.

(d) Compute the correction field by subtracting the radar-only kriged rainfall field (c) from the original radar field (b).

(e) The composite rainfall was derived by adding correction field (d) to gauge-only kriged rainfall field (a).

2.3 Performance validation

The study used the leave-one-out cross-validation (CV) method to validate performance of two spatial interpolation method including the CM and the OK method (Schuurmans and Bierkens, 2007; Erdin, 2009; Kim et al., 2013, 2016; Kim et al., 2019). In this method, one of the rain gauges is assumed to be non-existent. Then, the remaining values are used to generate rainfall estimate at the point that was assumed non-existent. This set of process is repeated for all gauges. Lastly, the observed and the estimated values are compared.

2.4 Scenarios for gauge density and temporal resolution

The previous studies on the spatial interpolation method have concluded that the density of gauge network and temporal resolution are important to the performance of spatial interpolation (Krajewski, 1987; Goudenhoofdt and Delobbe, 2009; Berndt et al., 2014; Baik et al., 2016). This study analyzed this effect through 25 scenarios changing the number of available gauges (6, 12, 18, 24, and 29 gauges) and temporal resolution (10, 20, 40, 60, and 80 minute).

In detail, given one case of gauge density, we create 5 scenarios of temporal resolution and it is applied for all case to generate totally 25 scenarios.

The network density scenarios were generated by randomly selecting the rain gauges, in which the scenario 1 represents to the use all available rain gauge data (Fig. 1). In 4 remaining cases, the number of gauges used to estimate rainfall was gradually reduced (Fig. 4).
3. Results and discussions

3.1 The performance of conditional merging method

Fig. 5(a) compares the observed (x) and the estimated (y) rainfall produced by the OK and the CM method for time snapshot of 12:50 PM, September 21, 2010. The green stars and red circles represent the OK method and the CM method, respectively. For this specific time, the CM method showed the higher correlation coefficient (r=0.83) than the OK method (r=0.66).

Fig. 5(b) compares the correlation coefficient for all 33 10-minute time steps investigated in this study. The green diamonds and red circles correspond to the CM and OK method, respectively. The orange polygon indicates the period that CM method outperforms the OK method, while the blue polygon indicates the opposite case. There are 22 times that CM method shows the greater correlation coefficient value compared to 11 times of the OK method. Moreover, the area of blue polygon is substantially smaller than that of orange polygon, indicating that the CM method remarkably improved the rainfall estimation.

3.2 The influence of gauge density and temporal resolution

Fig. 6 compares the correlation coefficient and the RMSE of the CM and the OK techniques varying with different gauge density values.

For both methods, the correlation coefficient value (r) increases with the increase of rain gauge density, while the RMSE value increases as density decreases. The OK method was more sensitive to the gauge density compared to the CM method. This is because the OK method relies only on the gauge measurements while the CM method uses both radar and gauge data. We also found that the influence of the gauge density does not increase dramatically when gauge density exceeds 4 gauges/100 km$^2$ for both techniques. The CM method shows the higher r value and lower RMSE.
Fig. 5. The Comparison Between CM and OK Technique in Terms of Correlation Coefficient

Fig. 6. Average Performance of CM and OK Technique Over all Gauge Density Scenarios at a 10-minute Temporal Resolution

Fig. 7. Average Performance of CM and OK Technique Over all Temporal Scenarios with Station Density of 100%
value than the OK method for all 5 gauge density scenarios.

Fig. 7 compares the $r$ and RMSE value varying with different temporal resolution of source rainfall data. The plots suggest that the accuracy of both methods increases with the increase of temporal resolution. The CM method showed greater performance for all 5 cases of rainfall temporal resolution.

The Fig. 8 shows the $r$ and the RMSE values varying with gauge density and temporal resolution. The color contours were generated by placing the $r$ and RMSE values in the 2 dimensional space composed of gauge density and temporal resolution and spatially interpolating them.

In general, the $r$ value increases as the increase of gauge density and accumulation time for both cases. Note that the areas in the contour plots with steep slope represent the abrupt performance change, which may provide a guideline on the minimum gauge density and the temporal resolution of the data to acquire a given degree of accuracy. For example, look at the red arrow in Fig. 8(a). The correlation coefficient at the starting and the ending point of the arrow correspond to the hourly rainfall estimation with the gauge density of 1.3 gauges/100 km$^2$ and 2 gauges/100 km$^2$. In this range, installation of only a few gauges can dramatically increase the accuracy of rainfall estimation. The white arrow, which has the same length as the red arrow suggests that the performance improvement would be minimal for the installation of additional gauges exceeding a given threshold of gauge density. The results also show that the CM method is more influenced by the temporal resolution compared to OK method.

4. Discussion

This study demonstrates that the CM method has strengthened the accuracy of rainfall simulation, which plays
a major role in flood studies using hydrological modeling. We also found that increasing the number of gauge stations will improve the performance of methods investigated in study. However, this improvement degree is substantial when the station density is in the range of 1 to 4 gauges / 100 km². It helps decision makers to choose the optimal number of stations ensuring the accuracy of rainfall simulation with reasonable cost. Increasing the accumulation time also improves the accuracy of the CM method but reduces the applicability of its product. For example, for the study of urban flood, it requires the rainfall data with accumulation time of 15 minutes at least. Therefore, the consideration between improving the performance of CM method through increasing the accumulation time and applicability of the product should also be given.

5. Conclusions

The main objective of this study was to investigate the performance of conditional merging (CM) technique that combines the gauge and radar rainfall data varying with different in-situ gauge density and temporal resolution of input data. The control data for comparison was the rainfall field produced by the Ordinary Kriging (OK) technique that used only the in-situ gauge data. The leave-one-out cross validation method was used to quantify the performance of the CM and the OK method. The main finding of this study can be summarized as follows:

1. The CM method showed the improved performance (greater r and smaller RMSE values) for all 25 scenarios of gauge densities and temporal resolutions.

2. The performance of the CM and the OK method increased with the greater gauge density and coarser temporal resolution. The performance of the OK method was more sensitive to the gauge density than the CM method was. In contrast, temporal resolution has greater impact on the performance of CM method compared to OK method.

3. The performance of both methods did not dramatically improve exceeding the gauge density of 4 gauges/100 km². Conversely, addition of only a few gauges can dramatically improve the rainfall estimation performance of both methods when the existing gauge density is very low (~ 1 gauge/100 km²).

Acknowledgements

This research was supported by a grant (MOIS-DP-2015-05) of Disaster Prediction and Mitigation Technology Development Program funded by Ministry of Interior and Safety (MOIS, Korea).

References


