A new event-based error decomposition scheme for satellite precipitation products

Runze Li1, Clement Guillette1, Efis Foufoula-Georgiou1,2, Pierre-Emmanuel Kirtzetter1,4
1Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California
2Department of Earth System Science, University of California, Irvine, Irvine, California
Hydrometeorology and Remote Sensing Laboratory, University of Oklahoma, Norman, Oklahoma
3NOAA/NWS National Weather Service, Norman, Oklahoma

Background

1. Traditionally, the mean absolute error of satellite precipitation products is only decomposed into “Hit precision”, “False precision”, and “Hit bias”, which is not enough to get in-depth information on error sources.
2. Precipitation naturally exists in the form of “events” leading to strong temporal correlation. Considering this event information will provide a better understanding of satellite error sources.
3. From an event-based view, satellite error can be essentially attributed to three causes:
   - “wrong start time and end time when an event is detected”
   - “wrong precipitation rate in the overlapped part of a detected event”
   - “an event that is totally missed or falsely detected”
4. As shown below, start and end time shifts between the MRMS and IMERG events widely exist, which is an important contribution to the satellite product bias.

Result 1: Error distribution map

- By more detailed classification of error types, distinct spatial patterns of different components are found, which suggests different responses of each error component to the different climatological precipitation dynamic features.
- For example, the delayed event start time is more severe in the northeast than in the southeast, while it is the opposite for the intensity underestimation.
- Overestimation resulting from the isolated falsely detected events is much higher in the northwest than in other regions, while the strongest lags in event end time happen in the southeast.

Conclusion remarks

1. This study develops a new error decomposition scheme for satellite precipitation products, which makes further attribution of the satellite errors by adding precipitation event information, rather than treating precipitation at each time step as independent in the traditional method. Different event-based error components could therefore be further linked with different precipitation dynamic causes.
2. The national mean start and end time shifts in IMERG are ~31 and ~10 minutes, respectively, and contribute to about 30% of the precipitation amount bias and 40% of the precipitation frequency bias, indicating the important contribution of event information.
3. The dominant sources of errors differ significantly among regions and seasons, with isolated falsely detected events in the Mountain West but displaced event start or end times in the Northeast, isolated falsely detected events in summer but displaced event start or end times in winter.
4. The “morph” sources tend to prolong the event duration both at the start and end time, while the IR sources have a large number of isolated falsely detected events.

Data and Method

1. Data
   - GPM IMERG V08B half-hourly Final Run
   - CONUS
   - GV/MRMS gauge–radar data
   - 2018-2020

2. Location/time domain
   - GV/MRMS gauge–radar data

3. Method
   - Definition:
     - Precipitation event: Continuous series of half-hourly rain rates no less than 0.3 mm/h in each pixel.
     - Hit event: A pair of events in MRMS and IMERG with at least one time step overlap.
     - Isolated event: An event in MRMS or IMERG with no overlapping time step with any event in the other dataset.
   - Event-based error decomposition scheme:

Result 2: National and regional statistics

- For the whole CONUS, the different start and end times of the “hit events”, the “isolated events”, and the different intensities in the overlapped hours of the “hit events” contribute to about 30%, 20%, and 50%, respectively, of both the underestimation amount and overestimation amount.
- The different start and end times of the “hit events” and the “isolated events” contribute to about 40% and 60%, respectively, of both the total underestimation frequency and overestimation frequency.
- National misses resulting from the advanced event start times contribute respectively to the total amount and frequency underestimations about 3 and 2.5 times more than those resulting from the delayed event start times in IMERG, while false alarms resulting from the advanced event start times contribute respectively to the total amount and frequency overestimations 4 and 2.5 times more than those resulting from the delayed event end times in IMERG.
- The error compositions vary greatly among different regions, e.g., the isolated events contribute to about 70% of total false cases in Mountain West, and the advanced event start times account for over 30% of total false cases in Southeast.

Result 3: Seasonal statistics

- The key issue for the amount underestimation lies in hit intensity differences in summer but lies in misses of all types in winter.
- The key issue for the frequency overestimation lies in displaced event start or end times in summer but lies in falsely detected isolated events in winter.

Result 4: Diurnal statistics

- Generally the contribution of each component to the total error is stable across the day, albeit with small but abrupt variations.
- Some components could still have relatively large variations, e.g., the False Alarm-Start accounts for about 37% of the total false alarms at 14:00 but 19% at 03:00; the Hit-Positive accounts for about 59% of the total amount overestimation at 02:00 but 41% at 24:00.

Result 5: Statistics by data sources

- The amount bias in IR-derived data is largely attributed to false alarms rather than hit bias, while the isolated events in IMERG could contribute to almost 90% of the total false alarms.
- The “morph” data tend to prolong the event duration both at the start and end time.
- The delays of event start time are more severe when SAPHIR is the source, contributing to the total false alarms twice as much as for the other PMW sources.

Concluding remarks

- The different start and end times of “hit events” and “isolated events” contribute to the total error at each time step as independent in the traditional method. Different event-based error components could therefore be further linked with different precipitation dynamic causes.
- The national mean start and end time shifts in IMERG are ~31 and ~10 minutes, respectively, and contribute to about 30% of the precipitation amount bias and 40% of the precipitation frequency bias, indicating the important contribution of event information.
- The dominant sources of errors differ significantly among regions and seasons, with isolated falsely detected events in the Mountain West but displaced event start or end times in the Northeast, isolated falsely detected events in summer but displaced event start or end times in winter for the total false alarms.
- The “morph” sources tend to prolong the event duration both at the start and end time, while the IR sources have a large number of isolated falsely detected events.