Global Precipitation Forecasts by Merging Extrapolation-Based Nowcast and Numerical Weather Prediction with Locally Optimized Weights

SHUNJI KOTSUKI

RIKEN Center for Computational Science, and RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program, and RIKEN Cluster for Pioneering Research, Kobe, Japan

KENTA KUROSAWA

RIKEN Center for Computational Science, Kobe, Japan

SHIGENORI OTSUKA

RIKEN Center for Computational Science, and RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program, and RIKEN Cluster for Pioneering Research, Kobe, Japan

KOJI TERASAKI

RIKEN Center for Computational Science, Kobe, Japan

TAKEMASA MIYOSHI

RIKEN Center for Computational Science, and RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program, and RIKEN Cluster for Pioneering Research, Kobe, and Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan, and Department of Atmospheric and Oceanic Science, University of Maryland, College Park, College Park, Maryland

(Manuscript received 19 September 2018, in final form 24 March 2019)

ABSTRACT

Over the past few decades, precipitation forecasts by numerical weather prediction (NWP) models have been remarkably improved. Yet, precipitation nowcasting based on spatiotemporal extrapolation tends to provide a better precipitation forecast at shorter lead times with much less computation. Therefore, merging the precipitation forecasts from the NWP and extrapolation systems would be a viable approach to quantitative precipitation forecast (QPF). Although the optimal weights between the NWP and extrapolation systems are usually defined as a global constant, the weights would vary in space, particularly for global QPF. This study proposes a method to find the optimal weights at each location using the local threat score (LTS), a spatially localized version of the threat score. We test the locally optimal weighting with a global NWP system composed of the local ensemble transform Kalman filter and the Nonhydrostatic Icosahedral Atmospheric Model (NICAM-LETKF). For the extrapolation system, the RIKEN's global precipitation nowcasting system called GSMAp_RNC is used. GSMAp_RNC extrapolates precipitation patterns from the Japan Aerospace Exploration Agency (JAXA)'s Global Satellite Mapping of Precipitation (GSMAp). The benefit of merging in global precipitation forecast lasts longer compared to regional precipitation forecast. The results show that the locally optimal weighting is beneficial.

1. Introduction

Quantitative precipitation forecast (QPF) is an important goal of numerical weather prediction (NWP). Over the past decades, precipitation forecasts by NWP models have been remarkably improved owing to the increased number of atmospheric observations, improved NWP models, and advanced data assimilation

Denotes content that is immediately available upon publication as open access.

Corresponding author: Shunji Kotsuki, shunji.kotsuki@riken.jp

(DA) methods. Precipitation nowcasting based on spatiotemporal extrapolation (hereafter, simply “extrapolation”) is also known to be a computationally feasible QPF method, which estimates future precipitation patterns by extrapolating the latest pattern of observed precipitation. It is known that extrapolation statistically outperforms NWP at shorter lead times. However, the forecast skill of extrapolation drops more quickly than that of NWP for longer lead time (e.g., [Sun et al. 2014](#)). Hence, optimal forecasts can be achieved by merging precipitation forecasts from NWP and extrapolation systems ([Browning 1980](#); [Golding 1998](#); [Bowler et al. 2006](#); [Sun et al. 2014](#)). For example, [Bowler et al. \(2006\)](#) succeeded in improving convective-scale precipitation forecasts up to 5 h by merging NWP and extrapolation.

Merging precipitation forecasts have been mainly explored at regional scales using ground-based radar observations. Recently, German Weather Service (Deutscher Wetterdienst: DWD) has launched the Seamless Integrated Forecasting System (SINFONY) project, which harmonizes precipitation forecasts of DWD’s regional NWP and extrapolation systems (R. Potthast, personal communication 2019). A few approaches have been proposed to merge different precipitation forecasts. The simplest approach computes a weighted mean of two forecasts. Another approach assimilates extrapolated radar reflectivity at a future time into NWP models ([Sokol and Zacharov 2012](#)).

In recent years, many aerospace and space agencies, such as the Japan Aerospace Exploration Agency (JAXA) and National Aeronautics and Space Administration (NASA), have been retrieving global surface precipitation mainly based on satellite-borne passive microwave radiometers ([Huffman et al. 2007, 2014](#); [Kubota et al. 2007](#); [Ushio et al. 2009](#)). Among them, JAXA has produced Global Satellite Mapping of Precipitation (GSMaP). [Otsuka et al. \(2016\)](#) applied spatiotemporal extrapolation to the GSMaP data. The extrapolation system, a.k.a. GSMaP RIKEN Nowcast (GSMaP_RNC), is routinely operated to provide global precipitation forecasts up to 12 h ([Otsuka et al. 2019](#), manuscript submitted to *J. Meteor. Soc. Japan*).

This study aims to improve global precipitation forecasts by merging precipitation fields from NWP and extrapolation systems. [Beck et al. \(2017\)](#) merged gauge, satellite and reanalysis data to produce global precipitation data. However, no study has conducted merged global precipitation for forecasts, to the best of the authors’ knowledge. As the NWP component, we use a global DA system that comprises the Non-hydrostatic Icosahedral Atmospheric Model (NICAM; [Satoh et al. 2008, 2014](#)) and the local ensemble transform

Kalman filter (LETKF; [Hunt et al. 2007](#)). The GSMaP_RNC is used as the extrapolation component. This study adopts the simplest merging approach: a weighted mean of precipitation forecasts by NICAM-LETKF and GSMaP_RNC.

Finding optimal weights is essential for merged forecasts. Regional applications usually use spatially uniform weights. However, optimal weights may vary in space due to spatially inhomogeneous forecast skills of NWP and extrapolation systems. Our first scientific question is if optimal weights vary in space for global applications. For that purpose, we introduce a spatially localized version of a threat score (critical success index) to optimize spatially distributed weights. Our second question is if merged forecasts are beneficial for global precipitation forecasts. The basic idea of locally optimal weighting is not necessarily limited to global applications so that we will discuss its potential applicability to regional precipitation forecasts.

This paper is organized as follows. [Section 2](#) describes the GSMaP_RNC and NICAM-LETKF systems, and how their forecasts are combined. [Section 3](#) details the experimental design. The results are presented and discussed in [sections 4](#) and [5](#). Finally, a summary is presented in [section 6](#).

2. Method

a. GSMaP_RNC system

The GSMaP_RNC extrapolates the latest pattern of the near-real-time version of GSMaP (a.k.a. GSMaP_NRT). The GSMaP_NRT covers 60°S–60°N with a spatial resolution of $0.1^\circ \times 0.1^\circ$. Estimating motion vectors of patterns plays an important role in extrapolation systems. A commonly used technique is the cross-correlation method, so-called optical flow, which estimates motions using two (or more) consecutive images. However, the optical flow occasionally produces erroneous motions due to observation errors. [Otsuka et al. \(2016\)](#) proposed the use of LETKF with 20 members to obtain reliable and smooth motions. [Otsuka et al. \(2016\)](#) successfully improved global precipitation forecasts compared to extrapolation without the LETKF. The GSMaP_RNC conducts hourly updated precipitation nowcasting up to 12 h.

b. NICAM-LETKF system

Deterministic forecasts are performed by the NICAM from analysis ensemble mean of the NICAM-LETKF ([Terasaki et al. 2015](#); [Terasaki and Miyoshi 2017](#); [Kotsuki et al. 2017a,b, 2019a](#)). We use the NICAM with the sixth grid division level, corresponding to the horizontal resolution of 112 km, and 38 vertical layers up to

FIG. 1. Schematic of the (a) global threat score (GTS), and (b) local threat score (LTS).

about 40 km. The Arakawa–Schubert scheme (Arakawa and Schubert 1974) and Berry’s parameterization (Berry 1967) are employed for cumulus parameterization and large-scale condensation schemes, respectively.

The NICAM-LETKF assimilates satellite radiances of the Advanced Microwave Sounding Unit-A (AMSU-A) and GSMaP precipitation in addition to conventional observations from the National Centers for Environmental Prediction (NCEP)’s operational system (a.k.a. NCEP PREPBUFR). Because the resolution of NICAM (112 km) is coarser than that of the GSMaP (~ 10 km), we apply the nearest-neighbor average (also known as budget interpolation; Accadia et al. 2003) to project precipitation forecasts by the NICAM into the GSMaP grid points.

c. Local threat score (LTS)

Here we introduce a spatially localized version of the threat score [hereafter local threat score (LTS)]. First, we explain a standard global threat score (GTS; Fig. 1a). Table 1 shows a contingency table summarizing events of a forecasting system. Based on the 2×2 combinations of observations (raining or not) and forecasts (raining or not) with a given precipitation threshold, precipitation events are partitioned into four categories: true positive (TP), false negative (FN), false positive (FP), or true negative (TN). Namely, each precipitation forecast is categorized to be TP = 1, or FP = 1, or FN = 1, or TN = 1. The GTS is given by

$$\text{GTS}(t) = \frac{\text{TP}_G(t)}{\text{TP}_G(t) + \text{FP}_G(t) + \text{FN}_G(t)}, \quad (1)$$

$$\text{TP}_G(t) = \sum_i \text{TP}_i(t) S_i, \quad (2)$$

$$\text{FP}_G(t) = \sum_i \text{FP}_i(t) S_i, \quad \text{and} \quad (3)$$

$$\text{FN}_G(t) = \sum_i \text{FN}_i(t) S_i, \quad (4)$$

where t , i , and S_i denote the verification time, i th grid point, and the area of the i th grid point (m^2), respectively. Subscript G denotes the global domain. As indicated by Eq. (1), the standard GTS evaluates the forecast skill averaged over the spatial domain at every time step, not averaged in the temporal domain.

This study considers that optimal weights vary in space for global applications. Here, we introduce the LTS to optimize spatially distributed weights. Patil et al. (2001) found the local low dimensionality of atmospheric dynamics. We considered that precipitation forecast skills could also be localized in space. The LTS is given by

$$\text{LTS}_i = \frac{\text{TP}_{L,i}}{\text{TP}_{L,i} + \text{FP}_{L,i} + \text{FN}_{L,i}}, \quad (5)$$

$$\text{TP}_{L,i} = \sum_{t \in T} \sum_{j \in D_i} \text{TP}_j(t) S_j, \quad (6)$$

$$\text{FP}_{L,i} = \sum_{t \in T} \sum_{j \in D_i} \text{FP}_j(t) S_j, \quad \text{and} \quad (7)$$

$$\text{FN}_{L,i} = \sum_{t \in T} \sum_{j \in D_i} \text{FN}_j(t) S_j, \quad (8)$$

where T is the temporal sampling period defined later (section 3) and D_i is the local sampling domain around the i th grid point. Following Patil et al. (2001), a surrounding $\pm 5^\circ \times \pm 5^\circ$ domain is used as D_i at the i th grid point. Here, we took a time average over the period T for robust estimates of LTS. In this study, we use the

TABLE 1. Contingency table representing event observation and forecast. Here, N is the number of total samples (TP + FN + FP + TN).

Events	Observation		
	Yes	No	Total
Forecast Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
Total	TP + FN	FP + TN	N

FIG. 2. Workflow of locally optimal merging of precipitation forecasts by NICAM (NWP) and GSMaP_RNC (nowcasting).

LTS to evaluate the forecast skill averaged in time and in a surrounding area at every grid location (Fig. 1b).

d. Merged precipitation based on LTS

Figure 2 shows the workflow of the locally optimal merging of precipitation forecasts. This study combines two precipitation forecasts as follows:

$$\text{MERGE}_{i,FT}(t) = \text{LW}_{i,FT} \text{NICAM}_i(t) + (1 - \text{LW}_{i,FT}) \text{GSMaP_RNC}_i(t), \quad (9)$$

where MERGE, NICAM, and GSMaP_RNC denote the precipitation intensities (mm h^{-1}) for merged forecasts, NICAM, and GSMaP_RNC, respectively. Here, $\text{LW}_{i,FT}$ is the local weight defined at the i th grid point and forecast time of FT (h). Analytically optimizing LW is a difficult

problem because of the unique functional characteristics of TS. Therefore, this study takes the brute-force manual optimization selecting optimal weights from 11 candidates (from 0.0 to 1.0 with a 0.1 increment) over a training period. First, we compute global maps of LTS by Eq. (5) for merged precipitation forecasts over the training period with 11 candidates (step 1 of Fig. 2). Second, the optimal weights are selected at each grid point so that the LTS over the training period is maximized (step 2 of Fig. 2).

Globally uniform weights (GW) can also be optimized with GTS in Eq. (1). The merged precipitation forecasts with GW and LW are compared in section 5.

3. Experimental settings

Terasaki et al. (2019) performed a long-term NICAM-LETKF experiment to investigate stability of the system.

FIG. 3. Optimized global weights as a function of FT (h) sampled over the 1-yr training period from September 2014 to August 2015. NICAM weighs more as GW increases.

FIG. 4. Spatial patterns of optimized local weights for FTs (a) 2, (b) 6, and (c) 10 h, sampled over the 1-yr training period. Warmer (colder) colors represent that NICAM (GSMaP_RNC) weighs more than GSMaP_RNC (NICAM). Missing values are grayed out.

The current study performs the NICAM forecasts at every 0000, 0600, 1200, and 1800 UTC from analysis means of the NICAM-LETKF experiment performed by Terasaki et al. (2019). Extrapolation data are provided by the GSMaP_RNC version 11 developed in July 2017 (Otsuka et al. 2019, manuscript submitted to *J. Meteor. Soc. Japan*). While GSMaP_RNC is updated every hour, this study only considers the precipitation forecasts initialized at every 0000, 0600, 1200 and 1800 UTC when the NICAM forecasts are available.

Precipitation forecasts are validated against the standard version of GSMaP (a.k.a. GSMaP_MVK), version 7. For areas where the passive microwave radiometers observations are not available during the 1-h observation time window, the GSMaP_MVK supplements such unobserved regions by moving precipitation patterns

based on a motion tracking algorithm. Such supplemented areas are not used for verification. A threshold value of 0.5 mm h^{-1} is used to compute threat scores.

The training and test periods are from September 2014 to August 2015 and from September 2015 to August 2016, respectively. Hereafter, precipitation forecasts of merged forecasts with GW and LW are referred to be MERGE_GW, and MERGE_LW, respectively.

4. Results

a. Optimized weights in the training period

Figure 3 shows the optimized GW as a function of FT. GSMaP_RNC weighs more than NICAM up to FT = 4 h, and less than NICAM after FT = 6 h. Figure 4 shows the spatial patterns of the optimized LW. No LW is

FIG. 5. Zonal-mean observed and 6-h forecasted precipitation (mm day^{-1}) for NICAM (blue), GSMaP_RNC (red), GSMaP_MVK (yellow), GSMaP_NRT (purple), and IMERG (green) and GPCP version 1.3 (light blue), averaged over the 1-yr test period. Black dashed lines represent 40°S and 40°N , respectively.

computed at grids where precipitation events are hardly observed or forecasted. Here we defined the effective sample ratio (ESR) by $\text{ESR} = 1 - [\text{TN}_{L,i}/(\text{TP}_{L,i} + \text{FP}_{L,i} + \text{FN}_{L,i} + \text{TN}_{L,i})]$, and used 0.01 as the threshold of ESR for computing LW. Again, NICAM weighs more as FT extends. NICAM obtains higher weights near the north–south boundaries of the GSMaP domain (around 60°S and 60°N in Fig. 4). This may be because the present

algorithm of GSMaP_NRT (version 6.0 algorithm) estimates precipitation intensity only for rain, but not for snow (JAXA 2017a). Therefore, the GSMaP_NRT tends to lack observations and make GSMaP_RNC less skillful near the boundaries. Consequently, NICAM obtains higher weights near the boundaries.

The NICAM tends to overproduce weak precipitation in regions where shallow convective systems are dominant

FIG. 6. Spatial patterns of precipitation forecast bias relative to GPCP version 1.3 at FT = 6 (mm day^{-1}) for (a) NICAM, (b) GSMaP_RNC, (c) MERGE_LW, and (d) MERGE_GW averaged over the 1-yr test period. Red and blue colors represent overestimated and underestimated precipitation compared to GPCP, respectively. Black dashed lines represent 40°N and 40°S .

FIG. 7. Spatial patterns of observed and forecasted precipitation (mm h^{-1}) for (a),(e),(i) GSMaP_MVK, (b),(f),(j) MERGE_LW, (c),(g),(k) GSMaP_RNC, and (d),(h),(l) NICAM. The precipitation patterns at (top) 1400, (middle) 1800, and (bottom) 2200 UTC 17 Dec 2015. Forecasted precipitation is shown for FT 2 h in (b), (c), and (d); FT 6 h in (f), (g), and (h); and FT 10 h in (j), (k), and (l), initiated at 1200 UTC 17 Dec 2015. Numbers at the top-right of (b), (c), (d), (f), (g), (h), (j), (k), and (l) are LTS averaged over the presented domain.

(Kotsuki et al. 2017a, 2018). In those areas, NICAM is less skillful and GSMaP_RNC weighs more than NICAM even at FT = 10 h (Fig. 4c; e.g., off the west coast of Australia, Peru, and Africa).

b. Precipitation bias

Here we discuss biases in the precipitation forecasts. Zonal-mean 6-h precipitation forecasts of NICAM and GSMaP_RNC are compared with four observation-based data (Fig. 5): GSMaP_NRT, GSMaP_MVK, Integrated Multisatellite Retrievals for GPM (IMERG version 05; Huffman et al. 2014) and Global Precipitation Climatology Project (GPCP) version 1.3. GSMaP families (GSMaP_NRT, GSMaP_MVK, and GSMaP_RNC) and IMERG provide hourly precipitation, and GPCC provides daily precipitation. GSMaP families show significantly lower precipitation than the other data (NICAM, IMERG, and GPCP) in the high latitudes (i.e., 60° – 40° S and 40° – 60° N) (cf. the appendix). While GSMaP_MVK provides more precipitation at the high latitudes than GSMaP_NRT, NICAM forecasts provide even more precipitation at the high latitudes, closer to GPCP. It should be noted that IMERG, another satellite-based precipitation product, shows overproduced precipitation compared to GPCP in the high latitudes.

Figure 6 compares the biases in precipitation forecast relative to GPCP. NICAM underestimates precipitation in the tropics (Fig. 6a). In contrast, GSMaP_RNC overestimates precipitation in the tropics (Fig. 6b) due to the overestimated precipitation of GSMaP_NRT (Fig. 5). Consequently, MERGE_LW and MERGE_GW show reduced precipitation bias in the tropics (Figs. 6c and 6d). In the high latitudes, NICAM shows good agreement with GPCP while GSMaP_RNC underestimates precipitation (Figs. 6a and 6b). Compared to MERGE_GW, MERGE_LW shows a smaller precipitation bias in the high latitudes owing to the higher weights on NICAM (Fig. 4b).

GPCP is commonly considered the best global precipitation reference. GSMaP families and IMERG show relatively higher errors compared to GPCP on the daily scale (Fig. 5). Therefore, it would still be challenging to estimate accurate precipitation in the high latitudes with higher temporal resolution. In the following subsections, quantitative verification is performed for hourly precipitation forecasts by NICAM, GSMaP_RNC, MERGE_GW and MERGE_LW. Due to the large uncertainties in hourly observed precipitation (GSMaP families and IMERG) in the high latitudes, we use data only in the lower latitudes (40° S– 40° N) for the quantitative verification.

FIG. 8. Time series of monthly mean LTS of 6-h precipitation forecasts for NICAM (blue), GSMaP_RNC (black), MERGE_LW (red), and MERGE_GW (green). Domains are shown for (a) NH (20° – 35° N), (b) TR (20° S– 20° N), and (c) SH (35° S– 20° S).

c. Precipitation forecasts in the test period

Figure 7 shows an example of observed and forecasted precipitation patterns over the Northwest Pacific Ocean on 17 December 2015. NICAM produces spatially smoother precipitation than GSMaP_RNC. Also, the numerical diffusion of GSMaP_RNC results in spatially smoother precipitation as FT extends (Figs. 7g and 7k). GSMaP_RNC produces a suspicious precipitation shape on a southwestern edge of the precipitation system at FT = 10 h (Fig. 7k; 20° – 30° N, 120° – 144° E). The spatiotemporal extrapolations occasionally produce such suspicious precipitation patterns because of the simple image-based extrapolation without using physical equations. On the other hand, the suspicious precipitation pattern disappears in MERGE_LW (Fig. 7j). MERGE_LW keeps fine precipitation patterns compared to NICAM. In this specific example, MERGE_LW results in the highest LTS over the presented domain.

Figure 8 shows the time series of monthly mean LTS for the 6-h precipitation forecasts. The first year is the training period, and the second year is the test period. Monthly mean LTSs are computed for three regions: the Northern Hemisphere (NH; 20° – 35° N), tropics (TR; 20° S– 20° N), and Southern Hemisphere (SH; 35° S– 20° S). Again, we consider a surrounding $\pm 5^{\circ} \times \pm 5^{\circ}$ domain to compute the LTS. At FT = 6 h, GSMaP_RNC tends to outperform NICAM in all regions. GSMaP_RNC and NICAM are more skillful in winter than in summer. NICAM shows larger seasonal variations than GSMaP_RNC. At FT = 6 h, merged precipitation forecasts outperform GSMaP_RNC and NICAM over the two years. GSMaP_RNC and NICAM-LETKF ran stably over the two years, and the merged forecasts were also stable over the experimental period.

Figure 9 compares the spatial patterns of LTS over the test period. The LTS generally decreases along with FT (Figs. 9a,d,g). The LTS over the ocean tends to be higher than that over land. Also, the precipitation forecasts in the

FIG. 9. Spatial patterns of LTS relative to GSMaP_MVK for FTs (a),(b),(c) 2, (d),(e),(f) 6, and (g),(h),(i) 10 h, sampled over the 1-yr test period. (left) The LTSs for MERGE_LW. Differences in LTS (center) between MERGE_LW and GSMaP_RNC and (right) between MERGE_LW and NICAM (LTS of MERGE_LW minus LTS of MERGE_GW). Magenta (green) in (b), (e), and (h) and in (c), (f), and (i) indicates MERGE_LW is better (worse) than GSMaP_RNC and NICAM. Missing values are grayed out.

storm track regions (35° – 20° S and 20° – 35° N) tend to be more skillful than those in the tropics. At FT = 2 h, MERGE_LW is almost equivalent to GSMaP_RNC and is clearly better than NICAM (Figs. 9b and 9c). At FT = 6 h, MERGE_LW is more skillful over the storm track regions than GSMaP_RNC (Fig. 9 e), and over the ocean than NICAM (Fig. 9 f). At FT = 10 h, MERGE_LW is clearly better than GSMaP_RNC over the globe, and slightly better than NICAM over the ocean (Figs. 9h and 9i).

Precipitation forecasts are also verified with another index: mean absolute difference (MAD) relative to the GSMaP_MVK (Fig. 10). Here, a local MAD is computed in a manner similar to LTS:

$$\text{MAD}_i = \frac{\sum_t \sum_{j \in D_i} |\text{obs}_j - \text{fcst}_j| S_j}{\sum_t \sum_{j \in D_i} S_j}, \quad (10)$$

FIG. 10. Spatial patterns of local MAD (mm h^{-1}) relative to GSMaP_MVK for FTs (a),(b),(c) 2, (d),(e),(f) 6, and (g),(h),(i) 10 h, sampled over the 1-yr test period. (left) The MADs for MERGE_LW. Differences in MAD (center) between MERGE_LW and GSMaP_RNC and (right) between MERGE_LW and NICAM, respectively. Magenta (green) in (b), (e), and (h) and in (c), (f), and (i) indicates MERGE_LW is better (worse) than the GSMaP_RNC and NICAM. Missing values are grayed out.

FIG. 11. Spatial patterns of difference in LTS relative to GSMaP_MVK between MERGE_LW and MERGE_GW (LTS of MERGE_LW minus LTS of MERGE_GW) for FTs (a) 2, (b) 6, and (c) 10 h, sampled over the 1-yr test period from September 2015 to August 2016. Magenta (green) represents that MERGE_LW is better (worse) than MERGE_GW. Missing values are grayed out.

where obs and fcast denote the observed and forecasted precipitation, respectively. In contrast to LTS, the larger MAD, the worse the skill. In terms of MAD, MERGE_LW is more skillful than NICAM over the globe and almost equivalent to GSMaP_RNC at FT = 2 h (Figs. 10b and 10c). MERGE_LW outperforms GSMaP_RNC in the tropics at FT = 6 and 10 h (Figs. 10b,e,h). MERGE_LW is more skillful than NICAM over the ocean at FT = 6 and 10 h (Figs. 10f and 10i), except for slight degradation over the Indian and west Pacific Oceans.

In general, MERGE_LW outperforms NICAM and GSMaP_RNC. Therefore, merging NWP and extrapolation is beneficial for global precipitation forecasts.

5. Discussion

a. Advantage of localized weights

Here we compare merged precipitation forecasts with local and spatially uniform weights. Figure 11 shows the spatial patterns of the difference in LTS between MERGE_LW and MERGE_GW. The MERGE_LW

tends to outperform MERGE_GW, suggesting that locally optimized weighting be beneficial for global precipitation forecasts. Relatively large improvements are seen in off the west coast of Australia, off the coast of the Arabian Peninsula and the United States of America. On the other hand, some regions are degraded by the locally optimized weighting maybe due to insufficient training and test periods. Figure 12 shows the spatial patterns of ESR. A small ESR corresponds to less frequent precipitation events. For example, an ESR of 0.01 means that precipitation is observed or forecasted approximately once in 100 h. MERGE_LW underperforms MERGE_GW in the region where ESR is relatively small (Fig. 11; Fig. 12; $\text{ESR} \leq 0.04$). Extending the training period would stabilize LW in such regions.

Figure 13 summarizes the mean precipitation forecast scores of NICAM, GSMaP_RNC, MERGE_LW, and MERGE_GW as a function of FT. Here, spatially averaged LTS and local MAD are presented (Figs. 13a and 13b). GSMaP_RNC is better than NICAM up to FT = 7 h for LTS, and up to FT = 4 h for MAD. MERGE_GW and MERGE_LW outperform both GSMaP_RNC and NICAM after 4 h for LTS. The benefit of merging in

FIG. 12. Spatial patterns of ESR (see the definition in the main text) to optimize LW for FTs (a) 2, (b) 6, and (c) 10 h in the 1-yr training period. Solid and dashed white contours represent 0.01 and 0.04, respectively.

global precipitation forecast lasts longer compared to regional precipitation forecast (Bowler et al. 2006; 1–5 h). MERGE_LW shows slightly higher LTS than the MERGE_GW. From the viewpoint of MAD, MERGE_LW and MERGE_GW outperform both NICAM and GSMaP_RNC at any FTs. MAD shows no clear difference between MERGE_LW and MERGE_GW. However, MERGE_LW results in less biased precipitation in high latitudes than the MERGE_GW (Figs. 6c and 6d). Overall, the three verification scores (LTS, MAD, and bias) suggest that the locally optimized weights provide an improvement over the globally optimized weights.

b. Implications to regional systems

The optimal weights would vary in space when forecast skills of NWP and extrapolation differ spatially;

therefore, the locally optimal weighting may be beneficial to regional systems with spatially inhomogeneous geographical features and observing networks. For spatially finer NWP and extrapolation systems, it is important to consider errors in precipitation location. Spatially localized version of the fraction skill score is a candidate to optimize LW in regional applications. Also, forecast skills of regional systems could differ more significantly than those of global systems depending on precipitating events such as typhoons, localized convective rains, and extratropical storms. Optimizing event-dependent LW would improve merged precipitation forecasts further in regional applications. In addition, the optimal size of local sampling domain would also depend on precipitating events. Those possible extensions are within the scope of our future studies.

FIG. 13. Mean precipitation forecast scores averaged over the 1-yr test period for (a) LTS, and (b) MAD (mm h^{-1}) as a function of FT (h). Global-mean LTS and MAD are computed for a domain of 35°S – 35°N . Blue, black, red and green lines represent the NICAM, GSMaP_RNC, MERGE_LW, and MERGE_GW, respectively.

6. Summary

This study aimed to improve precipitation forecasts by combining NWP and extrapolation systems with locally optimal weights. We used the GSMaP_RNC and NICAM-LETKF as extrapolation and NWP components. We proposed the local threat score (LTS), and optimized the spatially distributed local weights (LW) to maximize the LTS. We demonstrated that merging precipitation forecasts was beneficial even for global applications. The benefit of merging in global precipitation forecast lasts longer compared to regional precipitation forecast. The merged forecasts with LW (MERGE_LW) outperformed the merged forecasts with the spatially uniform global weight (MERGE_GW), suggesting that using the locally optimized weights provides better precipitation forecasts than using the globally optimized weights. The locally optimal weighting may also be beneficial to regional systems since forecast skills of NWP and extrapolation would differ spatially due to spatially inhomogeneous geographical features and observing networks.

There still remains a room for further improvements. Increasing the spatial resolution of NICAM would improve forecast skills and produce finer precipitation patterns. It is also important to improve the DA system and NICAM model for more skillful precipitation forecasts. In recent years, DA has become a promising tool to estimate model parameters in addition to model states (Ruiz et al. 2013; Schirber et al. 2013). Kotsuki et al. (2018) estimated a model parameter in NICAM with DA to improve precipitation forecasts. We will explore further improvements of the NICAM-LETKF system.

This study simply merged precipitation forecasts by GSMaP_RNC and NICAM. The state-of-the-art machine learning techniques potentially improve merged forecasts further. Those possible extensions will be further explored in our future studies. Recently, the near-real-time NICAM-LETKF system started running on JAXA's second-generation supercomputer system (JSS2; Kotsuki et al. 2019b). It is an important future direction to provide the merged precipitation forecasts routinely based on the success of present study.

Acknowledgments. S. Kotsuki and K. Kurosawa developed the algorithm to merge precipitation forecasts, conducted the experiments and analyzed the results; S. Otsuka and K. Terasaki performed the GSMaP_RNC and NICAM-LETKF experiments, respectively; T. Miyoshi is the PI and directed the research. The authors thank the members of Data Assimilation Research Team, RIKEN Center for Computational Science (R-CCS), T. Kubota of JAXA, Prof. M. Satoh and Dr. K. Kanemaru of the University of Tokyo for valuable suggestions and discussions. This study was partly supported by Japan Aerospace Exploration Agency (JAXA) Precipitation Measuring Mission (PMM), Advancement of meteorological and global environmental predictions utilizing observational 'Big Data' of the social and scientific priority issues (Theme 4) to be tackled by using post K computer of the FLAGSHIP2020 Project of the Ministry of Education, Culture, Sports, Science and Technology Japan (MEXT), the Initiative for Excellent Young Researchers of MEXT, and the Japan Society for the Promotion of Science (JSPS) KAKENHI Grants JP15K18128, JP18H01549, and JP16K17807. This project used the Supercomputer for

earth Observation, Rockets, and Aeronautics (SORA) at JAXA, and the K computer at R-CCS (Project IDs: ra000015, ra001011, hp150289, hp160162, hp160229, hp170178, hp170246, hp170305, hp180194, and hp180062).

APPENDIX

Weak Precipitation of GSMaP Data in the High Latitudes

The GSMaP algorithm estimates precipitation based on a lookup table between passive microwave radiometers' brightness temperature and precipitation intensity. Present lookup table at the high latitudes is constructed for every $5^\circ \times 5^\circ$ from observations of passive microwave radiometers and dual-frequency precipitation radar of the core satellite of global precipitation measurement (GPM; JAXA 2017b). Jaggy lines of GSMaP_MVK and GSMaP_NRT at the high latitudes (Fig. 11) would be derived from the $5^\circ \times 5^\circ$ lookup table. It is known that satellite-borne precipitation radar tends to underestimate weak precipitation. Kida et al. (2010) pointed out that the precipitation radar of the Tropical Rainfall Measuring Mission (TRMM) satellite did not detect weak precipitation appropriately [less than 0.10 mm h^{-1} ; cf. Fig. 6 of Kida et al. (2010)]. Weak precipitation would be underestimated in GSMaP, which relies on the lookup table developed with the dual-frequency precipitation radar.

REFERENCES

- Accadia, C., S. Mariani, M. Casaioli, A. Lavagnini, and A. Speranza, 2003: Sensitivity of precipitation forecast skill scores to bilinear interpolation and a simple nearest-neighbor average method on high-resolution verification grids. *Wea. Forecasting*, **18**, 918–932, [https://doi.org/10.1175/1520-0434\(2003\)018<0918:SOPFSS>2.0.CO;2](https://doi.org/10.1175/1520-0434(2003)018<0918:SOPFSS>2.0.CO;2).
- Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31**, 674–701, [https://doi.org/10.1175/1520-0469\(1974\)031<0674:IOACCE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2).
- Beck, H. E., A. I. Van Dijk, V. Levizzani, J. Schellekens, D. Gonzalez Miralles, B. Martens, and A. De Roo, 2017: MSWEP: 3-hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrol. Earth Syst. Sci.*, **21**, 589–615, <https://doi.org/10.5194/hess-21-589-2017>.
- Berry, E., 1967: Cloud droplet growth by collection. *J. Atmos. Sci.*, **24**, 688–701, [https://doi.org/10.1175/1520-0469\(1967\)024<0688:CDGBC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1967)024<0688:CDGBC>2.0.CO;2).
- Bowler, N. E., C. E. Pierce, and A. W. Seed, 2006: STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quart. J. Roy. Meteor. Soc.*, **132**, 2127–2155, <https://doi.org/10.1256/qj.04.100>.
- Browning, K. A., 1980: Local weather forecasting. *Proc. Roy. Soc. London A*, **371**, 179–211, <https://doi.org/10.1098/rspa.1980.0076>.
- Golding, B. W., 1998: Nimrod: A system for generating automated very short range forecasts. *Meteor. Appl.*, **5**, 1–16, <https://doi.org/10.1017/S1350482798000577>.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , D. T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, and P. Xie, 2014: GPM Integrated Multi-Satellite Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) version 4.4. PPS, NASA/GSFC, 30 pp.
- Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, <https://doi.org/10.1016/j.physd.2006.11.008>.
- JAXA, 2017a: User's guide for Global Satellite Mapping of Precipitation Microwave-IR Combined Product (GSMaP_MVK), Gauge-calibrated Rainfall Product (GSMaP_Gauge) Version 7, Reanalysis Products (GSMaP_RNL), and Gauge-calibrated Reanalysis Product (GSMaP_Gauge_RNL) Version 6. Accessed 19 September 2018, <http://shareku.eorc.jaxa.jp/GSMaP/>.
- , 2017b: Global Precipitation Measurement (GPM) GSMaP Product V04 (Algorithm version 7) Release Note (in Japanese). JAXA, 34 pp., http://www.eorc.jaxa.jp/GPM/doc/product_info/release_note_gsmavp04-v7_ja.pdf.
- Kida, S., S. Shige, T. Manabe, T. L'ecuyer, and G. Liu, 2010: Cloud liquid water path for the rain/no-rain classification method over ocean in the GSMaP algorithm. *Trans. JSASS Aerospace Tech. Japan*, **8**, 19–23.
- Kotsuki, S., T. Miyoshi, K. Terasaki, G.-Y. Lien, and E. Kalnay, 2017a: Assimilating the global satellite mapping of precipitation data with the Nonhydrostatic Icosahedral Atmospheric Model (NICAM). *J. Geophys. Res. Atmos.*, **122**, 631–650, <https://doi.org/10.1002/2016JD025355>.
- , Y. Ota, and T. Miyoshi, 2017b: Adaptive covariance relaxation methods for ensemble data assimilation: Experiments in the real atmosphere. *Quart. J. Roy. Meteor. Soc.*, **143**, 2001–2015, <https://doi.org/10.1002/qj.3060>.
- , K. Terasaki, H. Yashiro, H. Tomita, M. Satoh, and T. Miyoshi, 2018: Online model parameter estimation with ensemble data assimilation in the real global atmosphere: A case with the Nonhydrostatic Icosahedral Atmospheric Model (NICAM) and the global satellite mapping of precipitation data. *J. Geophys. Res. Atmos.*, **123**, 7375–7392, <https://doi.org/10.1029/2017JD028092>.
- , K. Kurosawa, and T. Miyoshi, 2019a: On the properties of ensemble forecast sensitivity to observations. *Quart. J. Roy. Meteor. Soc.*, <https://doi.org/10.1002/qj.3534>, in press.
- , K. Terasaki, K. Kanemaru, M. Satoh, T. Kubota, and T. Miyoshi, 2019b: Predictability of record-breaking rainfall in Japan in July 2018: Ensemble forecast experiments with the near-real-time global atmospheric data assimilation system NEXRA. *SOLA*, **15A**, 1–7, <https://doi.org/10.2151/sola.15A-001>.
- Kubota, T., H. Hashizume, N. Takahashi, S. Shige, K. Okamoto, T. Ushio, K. Aonashi, and M. Kachi, 2007: Global precipitation map using satelliteborne microwave radiometers by the GSMaP project: Production and validation. *Int. Geosci. Remote Sensing Symp. (IGARSS)*, Denver, CO, IEEE, 2584–2587, <https://doi.org/10.1109/IGARSS.2006.668>.

- Otsuka, S., S. Kotsuki, and T. Miyoshi, 2016: Nowcasting with data assimilation: A case of global satellite mapping of precipitation. *Wea. Forecasting*, **31**, 1409–1416, <https://doi.org/10.1175/WAF-D-16-0039.1>.
- Patil, D. J., B. R. Hunt, E. Kalnay, J. A. Yorke, and E. Ott, 2001: Local low dimensionality of atmospheric dynamics. *Phys. Rev. Lett.*, **86**, 5878–5881, <https://doi.org/10.1103/PhysRevLett.86.5878>.
- Ruiz, J. J., M. Pulido, and T. Miyoshi, 2013: Estimating model parameters with ensemble-based data assimilation: Parameter covariance treatment. *J. Meteor. Soc. Japan*, **91**, 453–469, <https://doi.org/10.2151/jmsj.2013-403>.
- Satoh, M., T. Matsuno, H. Tomita, H. Miura, T. Nasuno, and S. Iga, 2008: Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *J. Comput. Phys.*, **227**, 3486–3514, <https://doi.org/10.1016/j.jcp.2007.02.006>.
- , and Coauthors, 2014: The Non-hydrostatic Icosahedral Atmospheric Model: Description and development. *Prog. Earth Planet. Sci.*, **1**, 18, <https://doi.org/10.1186/s40645-014-0018-1>.
- Schirber, S., D. Klocke, R. Pincus, J. Quaas, and J. L. Anderson, 2013: Parameter estimation using data assimilation in an atmospheric general circulation model: From a perfect toward the real world. *J. Adv. Model. Earth Syst.*, **5**, 58–70, <https://doi.org/10.1029/2012MS000167>.
- Sokol, Z., and P. Zacharov, 2012: Nowcasting of precipitation by an NWP model using assimilation of extrapolated radar reflectivity. *Quart. J. Roy. Meteor. Soc.*, **138**, 1072–1082, <https://doi.org/10.1002/qj.970>.
- Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, <https://doi.org/10.1175/BAMS-D-11-00263.1>.
- Terasaki, K., and T. Miyoshi, 2017: Assimilating AMSU-A radiances with the NICAM-LETKF. *J. Meteor. Soc. Japan*, **95**, 433–446, <https://doi.org/10.2151/jmsj.2017-028>.
- , M. Sawada, and T. Miyoshi, 2015: Local ensemble transform Kalman filter experiments with the Nonhydrostatic Icosahedral Atmospheric Model NICAM. *SOLA*, **11**, 23–26, <https://doi.org/10.2151/sola.2015-006>.
- , S. Kotsuki, and T. Miyoshi, 2019: Multi-year analysis using the NICAM-LETKF data assimilation system. *SOLA*, **15**, 41–46, <https://doi.org/10.2151/sola.2019-009>.
- Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Japan*, **87A**, 137–151, <https://doi.org/10.2151/jmsj.87A.137>.