Evaluation of the RainFARM Statistical Downscaling Technique Applied to IMERG over Global Oceans Using Passive Aquatic Listener In Situ Rain Measurements

JANICE L. BYTHEWAY,[a] ELIZABETH J. THOMPSON,[a] JIE YANG,[b] AND HAONAN CHEN[c]

[a] *NOAA/Physical Sciences Lab, Boulder, Colorado*
[b] *Applied Physics Lab, University of Washington, Seattle, Washington*
[c] *Department of Electrical Engineering, Colorado State University, Fort Collins, Colorado*

(Manuscript received 6 June 2023, in final form 16 August 2023, accepted 11 September 2023)

ABSTRACT: High-resolution oceanic precipitation estimates are needed to increase our understanding of and ability to monitor ocean–atmosphere coupled processes. Satellite multisensor precipitation products such as IMERG provide global precipitation estimates at relatively high resolution (0.1°, 30 min), but the resolution at which IMERG precipitation estimates are considered reliable is coarser than the nominal resolution of the product itself. In this study, we examine the ability of the Rainfall Autoregressive Model (RainFARM) statistical downscaling technique to produce ensembles of precipitation fields at relatively high spatial and temporal resolution when applied to spatially and temporally coarsened precipitation fields from IMERG. The downscaled precipitation ensembles are evaluated against in situ oceanic rain-rate observations collected by passive aquatic listeners (PALs) in 11 different ocean domains. We also evaluate IMERG coarsened to the same resolution as the downscaled fields to determine whether the process of coarsening then downscaling improves precipitation estimates more than averaging IMERG to coarser resolution only. Evaluations were performed on individual months, seasons, by ENSO phase, and based on precipitation characteristics. Results were inconsistent, with downscaling improving precipitation estimates in some domains and time periods and producing worse performance in others. While the results imply that the performance of the downscaled precipitation estimates is related to precipitation characteristics, it is still unclear what characteristics or combinations thereof lead to the most improvement or consistent improvement when applying RainFARM to IMERG.

KEYWORDS: Atmosphere-ocean interaction; Precipitation; Satellite observations; Downscaling

1. Introduction

Over 75% of global rain falls over Earth's oceans (Trenberth et al. 2007), and this precipitation plays an important role in the surface freshwater buoyancy flux and sea surface salinity (Ramesh Kumar and Shulz 2002; Prakash et al. 2012, 2018), which couples the ocean to the atmosphere, and contributes to heat processes in driving ocean circulations (Kucera and Klepp 2022). Therefore, reliable estimates of oceanic precipitation at high spatial and temporal resolution are necessary to better understand the global water cycle and climate system (Prakash et al. 2018).

Although improvements in observation quality and record-length have resulted in changes to oceanic precipitation estimates over the years (Trenberth et al. 2007), in situ measurements of oceanic precipitation remain sparse, being generally limited to gauges on a relatively small number of buoy arrays or research vessels, many of which are confined to relatively high-precipitation tropical regions (Maggioni et al. 2016; Sapiano and Arkin 2009; Wu and Wang 2019; Prakash et al. 2018). The quality of gauge precipitation estimates is impacted by undercatch due to high winds, platform motion, turbulence around the ship structure, intermittent data outages, and even buoy vandalism (Wu and Wang 2019; Yang et al. 2015; Maggioni et al. 2016; Kucera and Klepp 2022). Observations from the Ocean Rainfall and Ice-phase precipitation measurement Network (OceanRAIN) aimed to overcome the shortcomings of ship-based gauges and tropics-focused instrument placement by deploying optical disdrometers to measure precipitation on eight ships that traversed the global oceans from June 2010 through December 2018 (Klepp 2015; Klepp et al. 2018; Kucera and Klepp 2022). The OceanRAIN dataset provides more geographic coverage than stationary buoy arrays but still represents a limited number of data points. Recently, 12 years of observations from passive aquatic listeners (PALs)—hydrophones capable of detecting the unique frequency of raindrops falling with different intensities on the ocean surface—were reprocessed and made available for use (Bytheway et al. 2023; Yang et al. 2015). PALs have been deployed on moorings and drifting Argo buoys over the global oceans. While their data have not been widely used outside of field campaigns, these newly available oceanic PAL rain datasets create an opportunity to validate satellite rain retrievals over oceans (Li et al. 2023) and test postprocessing methods, such as downscaling, to improve satellite estimates.

While in situ measurements are sparse over the oceans, satellite multisensor precipitation products such as Integrated Multi-satellitE Retrievals for GPM (IMERG) provide global or quasi-global precipitation estimates at relatively high spatial and temporal resolution. Many studies have evaluated these products over ocean using regional in situ datasets,

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JHM-D-23-0090.s1.

Corresponding author: Janice L. Bytheway, janice.bytheway@noaa.gov

DOI: 10.1175/JHM-D-23-0090.1

surface-based radars located on islands or near coasts, or lower resolution climate precipitation products (e.g., Sapiano and Arkin 2009; Maggioni et al. 2016; Wu and Wang 2019; Prakash et al. 2018; Kucera and Klepp 2022; Bytheway et al. 2023). These evaluations have found that satellite multisensor precipitation estimates underestimate or miss light and warm rain (Maggioni et al. 2016; Wu and Wang 2019), with the largest rainfall rate errors found in the latitude band between 0° and 15° (Kucera and Klepp 2022).

Although nominally available at resolutions on the order of several kilometers and hourly to subhourly, satellite multisensor precipitation estimates are typically most reliable when averaged to coarser spatial and temporal resolution (Tan et al. 2017; Bytheway et al. 2023). However, over ocean, even relatively light (~5 mm h^{-1}) precipitation events can stabilize the near-surface ocean under weak winds, even amid surface cooling (Thompson et al. 2019; Iyer and Drushka 2021a,b, so-called ocean fresh layers, freshwater lenses, or rain layers). Under these stabilizing conditions, rain can dramatically change sea surface salinity (Drushka et al. 2016, 2019; Reverdin et al. 2020). Regardless of ocean stability, rain also creates substantial sensible heat fluxes to the atmosphere (up to 200 W m^{-2} for about 100 mm h^{-1} instantaneous 1-min rain rates, or 40 mm h^{-1} 10-min average rain rates, depending on air temperature and humidity conditions; Gosnell et al. 1995). It is significant that fine-scale, short-lived, and even light rain rates can impact the ocean so dramatically, especially since smaller short-lived rain events are the most common occurrences of rain over oceans (e.g., Trivej and Stevens 2010). Rain events that are impactful on the ocean often occur on scales smaller than the nominal resolution of many satellite-based precipitation products, which in turn are offered at finer resolution than that at which they most accurately represent the precipitating field (e.g., daily, Tan et al. 2017; Wu and Wang 2019). As such, in order to use satellite multisensor precipitation estimates for oceanic applications, the subgrid-scale variability will need to be simulated (Kucera and Klepp 2022).

One way to obtain higher resolution estimates of oceanic precipitation from satellite observations is through downscaling techniques, which derive high-resolution precipitation estimates from a coarser resolution precipitation field, with a goal to reliably represent finer-scale intensities, temporal and spatial variability, and physical consistency of precipitation regardless of region or season (Maraun et al. 2010). Downscaling techniques can be broadly divided into dynamical and statistical methods, which can both be divided into a number of subcategories.

Dynamical downscaling techniques use a high-resolution regional climate model (RCM) or numerical weather prediction (NWP) model nested within a lower resolution general circulation model (GCM). The GCM provides the initial and lateral boundary conditions, while the RCM accounts for high-resolution topographic data, land–sea contrasts, surface characteristics, and other components of the Earth system (He et al. 2016; Xu et al. 2019; American Meteorological Society 2015). Dynamical downscaling techniques, although physically based, are often computationally expensive, susceptible to systematic biases from the GCM and RCM, and sensitive to domain size, location, surface type, and RCM vertical and horizontal resolutions (He et al. 2016; Xu et al. 2019; Tapiador et al. 2020). While the grid size of the RCMs is smaller than that of the parent GCMs, it is often still insufficient for many hydrologic applications (Tapiador et al. 2020). An overview of dynamical downscaling, including many examples of dynamical downscaling techniques, can be found in Tapiador et al. (2020).

Statistical downscaling techniques typically rely on empirical relationships between large- and small-scale variables, obtaining fine-scale fields by relating coarse-scale observations to smaller-scale environmental data like topography or vegetation (Wilby et al. 1999; Chen et al. 2018; He et al. 2016; Abbasian et al. 2020; Gutmann et al. 2022). Statistical downscaling models for precipitation typically produce an ensemble of random precipitation fields that satisfy the large-scale constraints (i.e., average rainfall over an area) and are consistent with known statistical properties of the fine-scale rainfall distribution. These precipitation fields are neither physically based nor should they be considered to be the "actual" small-scale rainfall field. Rather, they are plausible representations that provide an estimate of fine-scale variability while maintaining the appropriate statistical relationships between the coarse- and fine-scale precipitation (Brussolo et al. 2008; Rebora et al. 2006a; D'Onofrio et al. 2014; Ferraris et al. 2003; Foufoula-Georgiou et al. 2014). There are many different types of statistical downscaling techniques, descriptions, and examples of which can be found in Wilby et al. (1999), Kang and Ramírez (2010), Ferraris et al. (2003), He et al. (2016), Chen et al. (2018), and references therein.

Statistical downscaling techniques have the benefit over dynamical downscaling techniques of being fast and computationally efficient and are often able to provide estimates of uncertainty and probability of extremes (He et al. 2016; Abbasian et al. 2020). Challenges of statistical downscaling techniques include the ability to maintain consistency of the downscaled processes with the coarse-scale field, representing processes not present in the historical data record from which empirical relationships are derived, and cases lacking a strong empirical relationship between large- and small-scale processes (He et al. 2016; Xu et al. 2019).

The Rainfall Autoregressive Model (RainFARM) is a stochastic statistical downscaling model developed specifically to produce an ensemble of downscaled precipitation features on the order of 1 km and a few minutes in time from mesoscale precipitation features on the order of a few tens of kilometers in space and hours in time. RainFARM has been tested and shown successful results when applied to both observational data and model output in several different regions (Rebora et al. 2006a,b; D'Onofrio et al. 2014; Brussolo et al. 2008, 2009). RainFARM stands out from other statistical downscaling techniques because the downscaled rainfall fields are calculated based solely on information from the coarse-scale field. Note, RainFARM neither requires nor uses any information about the underlying surface or other atmospheric characteristics.

In this study, we test RainFARM's ability to harvest fine-scale precipitation estimates from IMERG over global oceans. Since the validity of the input field is critical to ensuring valid downscaling output, we apply the RainFARM downscaling technique to coarsened daily rainfall estimates from IMERG, which are known to be more accurate compared to observations than those from IMERG at native resolution. We compare oceanic PAL in situ rain measurements to IMERG estimates at native resolution, coarsened resolution, and downscaled from coarsened resolution. Section 2 summarizes the RainFARM stochastic downscaling model. Section 3 describes the IMERG and PAL datasets used in the study, as well as the domains and time periods selected for evaluation. Results from the comparisons to the PAL data are presented in section 4, with discussion and conclusions in section 5.

2. The RainFARM downscaling model

The RainFARM downscaling model produces an ensemble of possible realizations of the high-resolution rainfall field from a smooth field at low spatial and/or temporal resolution. These realizations are not a physically accurate depiction of high-resolution precipitation, but rather a representation of precipitation processes with appropriate statistical properties. Full details of the RainFARM downscaling model can be found in Rebora et al. (2006a). Here, we provide a brief synopsis of RainFARM, along with several features and caveats that will influence both our application of RainFARM to IMERG precipitation estimates and our interpretation of the evaluation exercises.

RainFARM is based on a nonlinear transformation of a linearly correlated stochastic field. A key feature of RainFARM is the extrapolation of the large-scale power spectrum to smaller scales. A power-law functional form is assumed for the extrapolation, with random Fourier phases at unresolved scales. The large-scale spectral information is propagated to smaller scales while maintaining the total rainfall volume, linear correlation structure, and spatial distribution of large-scale rainfall patterns. Successful application of RainFARM requires that the coarse-scale field provides sufficient information to estimate the slopes of the spatial and temporal power spectra and that the downscaling be performed only on the range of scales characterized by the approximate linear scaling behavior (Brussolo et al. 2009; Rebora et al. 2006a,b). In this way, RainFARM produces estimates of fine-scale precipitation fields without any additional surface or environmental data.

Beginning with a mesoscale precipitation field $P(X, Y, T)$, P is defined on scales $L_0 \leq (X, Y) \leq L_{\max}$ and $T_0 \leq T \leq T_{\max}$, where $L_{\max}$ and $T_{\max}$ are the size of the domain and length of record, respectively. The terms L_0 and T_0 are the scales at which the field P can be considered a reliable estimate, which is not necessarily the nominal resolution of the mesoscale rain field. The goal is to disaggregate P into the stochastic field $r(x, y, t)$ with resolutions $l_0 < L_0$ and $t_0 < T_0$, where r is a member of an ensemble of n possible realizations of the unknown fine-scale precipitation field (Rebora et al. 2006a).

In calculating the spatial spectral slopes of P, the fields are assumed to be spatially isotropic (i.e., slopes are the same in the x and y directions; Rebora et al. 2006a). While the spectral slopes are calculated directly from the large-scale behavior of the precipitation field, they are likely conditional, that is, they depend on the synoptic pattern of the precipitation field (Rebora et al. 2006b). It is important to calculate the spectral slopes using not necessarily the native resolution of the field to be downscaled, *but the scale at which the representation of the field is reliable*. Tan et al. (2017), Wu and Wang (2019), and Bytheway et al. (2023) all found that comparisons between satellite-based precipitation estimates and in situ observations agree best at coarser spatiotemporal scales, particularly daily. For this reason, we coarsen IMERG from its native resolution to several lower spatial and temporal resolution fields as described in section 3b to investigate the sensitivity of the downscaled results to starting resolution and whether spatial or temporal downscaling is most effective in providing estimates of the fine-scale variability of oceanic precipitation.

3. Data

a. *IMERG V6 final run*

1) DATASET DESCRIPTION

IMERG provides global estimates of rain rate at 30-min intervals at 0.1° (~10 km) grid spacing. Retrieved precipitation estimates from passive microwave (PMW) satellites in the global precipitation measurement (GPM) constellation are intercalibrated using the GPM core observatory as a reference (Huffman et al. 2018). Motion vectors calculated using numerical model output are used to advect precipitating features during time steps without PMW overpasses (Tan et al. 2019). Additional microwave-calibrated rainfall estimates from infrared radiances following the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS; Hong et al. 2004) algorithm are also included. Precipitation estimates over land are calibrated using the Global Precipitation Climatology Center (GPCC) monthly precipitation dataset following Huffman et al. (2007). No gauge-based bias correction or calibration is performed over oceans due to scarcity of in situ gauge data (G. Huffman 2022, personal communication).

IMERG precipitation estimates are available at three levels of latency to serve different user needs. IMERG-Early and IMERG-Late are available 6 and 18 h after valid time, respectively, for near-term uses such as hazard forecasting and water management. IMERG-Final becomes available approximately 4 months after valid time and includes both monthly gauge adjustments to reduce bias over land and more PMW data than the Early and Late runs due to data delivery delays (Tan et al. 2017). In this study, we use the IMERG-Final product.

2) PREPARATION FOR DOWNSCALING AND EVALUATION

As described, the field to be downscaled using RainFARM must be at a resolution at which the data are considered to be

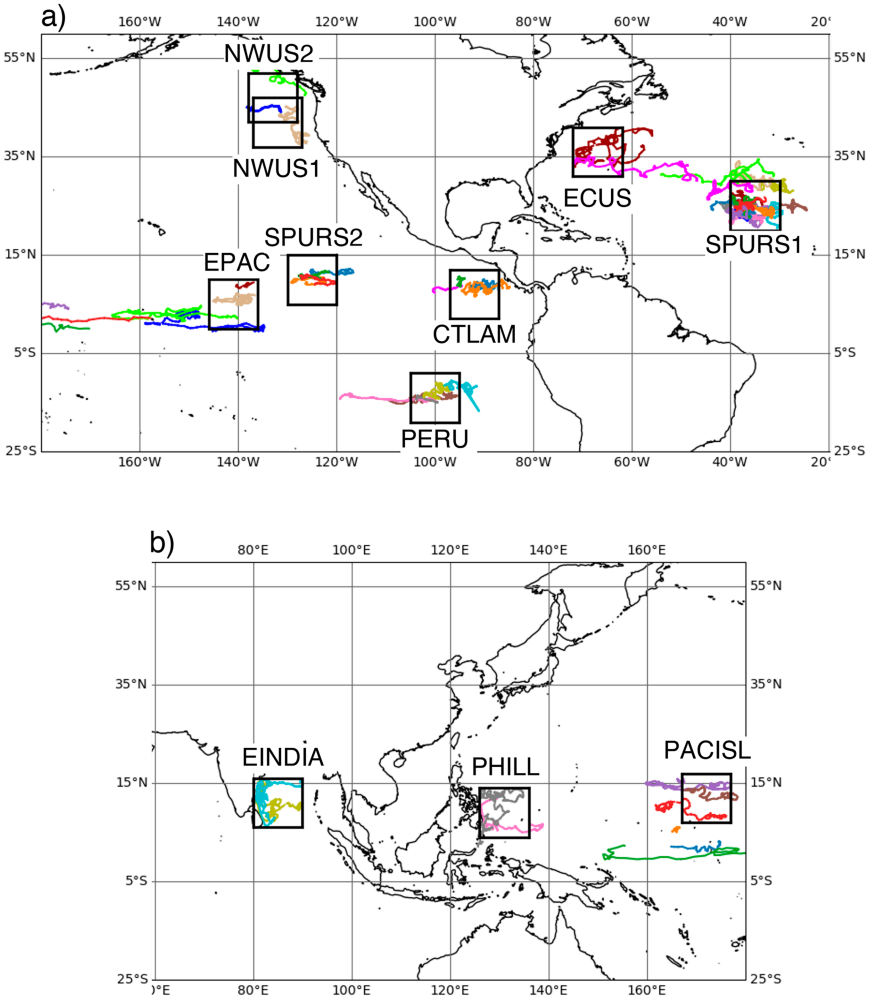

FIG. 1. Trajectories of PALs deployed between 2011 and 2018 in (a) the eastern Pacific and Atlantic Oceans and (b) western Pacific and Indian Oceans. Different color trajectories are used for visibility. The 10° × 10° domains where downscaling is performed are also shown and are described in Table 1.

reliable. The low-resolution field must also be of sufficient size to provide enough information for the extrapolation of power spectra to smaller scales. Therefore, we perform the downscaling in domains of 10° × 10° (see Fig. 1) and, prior to downscaling, average IMERG from its native 0.1° 30-min resolution to both 0.1° daily and 0.5° daily average rain rates. This results in domains of 100 × 100 grid boxes at 0.1° grid spacing and 20 × 20 grid boxes at 0.5° grid spacing. Spatial and temporal power spectra and their associated slopes are calculated monthly to account for seasonal changes in the precipitation characteristics in each domain.

Five downscaling experiments are performed, each producing monthly ensembles of 100 high-resolution precipitation fields. From both the 0.5° and 0.1° daily rain rates, we temporally downscale to both 3- and 1-hourly rain rates. Additionally, the 0.5° fields are spatially downscaled to 0.1° daily precipitation. Observations from the in situ PALs are used to evaluate both the downscaled fields and IMERG fields coarsened to the same resolutions as produced by the downscaling. In this way, we determine whether an ensemble of downscaled realizations can represent the variability of fine-scale precipitation, and whether information from the spatial and temporal power spectra of the coarse-scale field can produce a downscaled precipitation field with similar or improved comparisons to PAL observations as IMERG coarsened to the same resolution.

A variety of evaluation metrics are used to evaluate the coarsened and downscaled precipitation fields, including bias, correlation coefficient, critical success index (CSI), false alarm ratio (FAR), frequency bias, mean absolute error, mean error, probability of detection (POD), and root-mean-square error (RMSE). Contingency table statistics (POD, FAR, and CSI) are calculated using scale-dependent thresholds determined using Eq. (1) from Tan et al. (2017). Evaluation statistics are calculated monthly for each month when multiple PALs are available in a given domain and only for IMERG

TABLE 1. Characteristics of the 11 domains evaluated in this study.

Domain name	Description	No. of PALs	Temporal coverage	Spatial coordinates
SPURS-2	NASA SPURS-2 field campaign	4 drifter, 2 mooring	Nov 2016–Jul 2018	120°–130°W, 5°–15°N
SPURS-1	NASA SPURS field campaign	15 drifter	Nov 2012–Jun 2016	30°–40°W, 20°–30°N
PERU	West of coastal Peru	5 drifter	Jan–Mar 2017	95°–105°W, 9°–19°S
EPAC	Eastern Pacific	4 drifter	Jun 2011–14	136°–146°W, 0°–10°N
CTLAM	West coast of Central America	4 drifter	Dec 2013–Jan 2017	87°–97°W, 2°–12°N
PACISL	Pacific Islands	3 drifter	Jul 2012–Mar 2015	167°–177°E, 7°–17°N
PHILL	Philippines	2 drifter	Feb 2012–Jul 2013	126°–136°E, 4°–14°N
EINDIA	Bay of Bengal	2 drifter	Sep 2011–Feb 2015	80°–90°E, 6°–16°N
NWUS1	West Coast of United States—Washington/Oregon	2 drifter	Dec 2011–Mar2015	127°–137°W, 37°–47°N
NWUS2	West Coast of United States/Canada	3 drifter	Dec 2011–Mar 2015	128°–138°W, 42°–52°N
ECUS	East Coast of United States	2 drifter	Nov 2012–Jan 2015	62°–72°W, 31°–41°N

grid boxes collocated with PALs based on the average PAL location during the averaging period (i.e., 1, 3, 24 h). Because RainFARM is designed to maintain the overall statistical representation and total water of the large-scale precipitation field, statistical descriptors (e.g., mean rain rate) and associated performance statistics (bias) of the downscaled precipitation fields will not change significantly from that of the large-scale field. However, because the downscaling will inherently change the spatial and temporal distributions of the precipitation, metrics such as POD, FAR, RMSE, and correlation coefficient may differ between the large-scale and downscaled fields. We also aggregate the statistics for all Januarys, all Februarys, etc., in domains having more than one full year of PAL observations in order to evaluate the performance of RainFARM over the seasonal cycle, as well as aggregating the statistics based on phase of the El Niño–Southern Oscillation (ENSO) index, monthly mean rain rate, and binned spatial and temporal spectral slopes.

b. PAL

PALs have been deployed on drifting Argo floats and stationary moorings across the global oceans, frequently as part of field projects (Bingham et al. 2019; Lindstrom et al. 2017, 2019; Riser et al. 2019; Yang et al. 2023). The PAL records a time series of ambient noise and converts it to an acoustic spectrum, from which dominant noise sources such as rain, wind, biological (e.g., marine mammals), and anthropogenic sources (e.g., shipping, oil drilling) can be identified. Data points classified as either rain or wind are then used to estimate either rain rate or wind speed. That is, wind and rain retrievals are exclusive to eliminate the chance of rain noise contaminating wind speed estimates. Rain-rate and wind speed estimates from PALs mounted below the surface on stationary moorings compare favorably with collocated surface rain measurements on minute time scales (Ma and Nystuen 2005; Nystuen 2001; Yang et al. 2015; Riser et al. 2019).

PALs on drifting Argo floats are deployed at 1-km depth, collecting rain and wind samples every 2–9 min. Every ~9.5 days, the float rises to the surface to collect an ocean vertical profile during ascent and telemeter the collected data via the Iridium satellite network. PALs have also been deployed on a number of research moorings, typically at the same 1000-m depth as the Argo floats. When deployed at 1-km depth, the circular surface listening area being sampled by the PAL has a diameter of 5 km (i.e., ~5 times the depth where the instrument is located). This 5-km areal precipitation estimate is well suited for comparison with gridded precipitation estimates from satellites (Bytheway et al. 2023; Yang et al. 2015), although the PAL's areal footprint is still smaller than the grid spacing of IMERG, at both native and coarsened spatial resolutions. This spatial-scale mismatch may impact the results, particularly for IMERG coarsened to 0.5°.

Individual PALs typically remain deployed for 1–4 years, and therefore, the number of PALs available at any given time varies. Recently, 12 years of PAL data were made available in NetCDF format with uniform 1-min time steps (Thompson et al. 2023). Figure 1 illustrates the locations of PALs over the period 2011–18, along with 11 selected 10° × 10° domains having periods of record with multiple PAL observations. Descriptions of each domain, including domain names, temporal extent, and the number of PALs available for comparison, are given in Table 1.

4. Results

The quasi-global nature of the selected domains, differing temporal availability of PALs, and differing seasonal and dynamical characteristics of precipitation in the domains result in a variety of outcomes when RainFARM is applied to IMERG in each domain and time period. As such, we will focus on generalized outcomes from the application of RainFARM to IMERG, providing examples from individual domains to illustrate the results.

First, we show that RainFARM is performing as expected—that is, that the downscaled fields maintain the statistical characteristics of the large-scale precipitation field. Figure 2 compares the downscaled precipitation ensemble to IMERG at both native and coarsened resolutions and the PAL mounted on the SPURS-2 Central Mooring (10.05°N, 125.03°W) for November 2016. The left column of Fig. 2 shows results for 0.5° daily rain rates spatially downscaled to 0.1° daily rain rates, while the right column shows temporally downscaled realizations from daily to hourly rain rates at 0.1° grid spacing.

When downscaling spatially, RainFARM provides a more variable ensemble of precipitation estimates compared to the

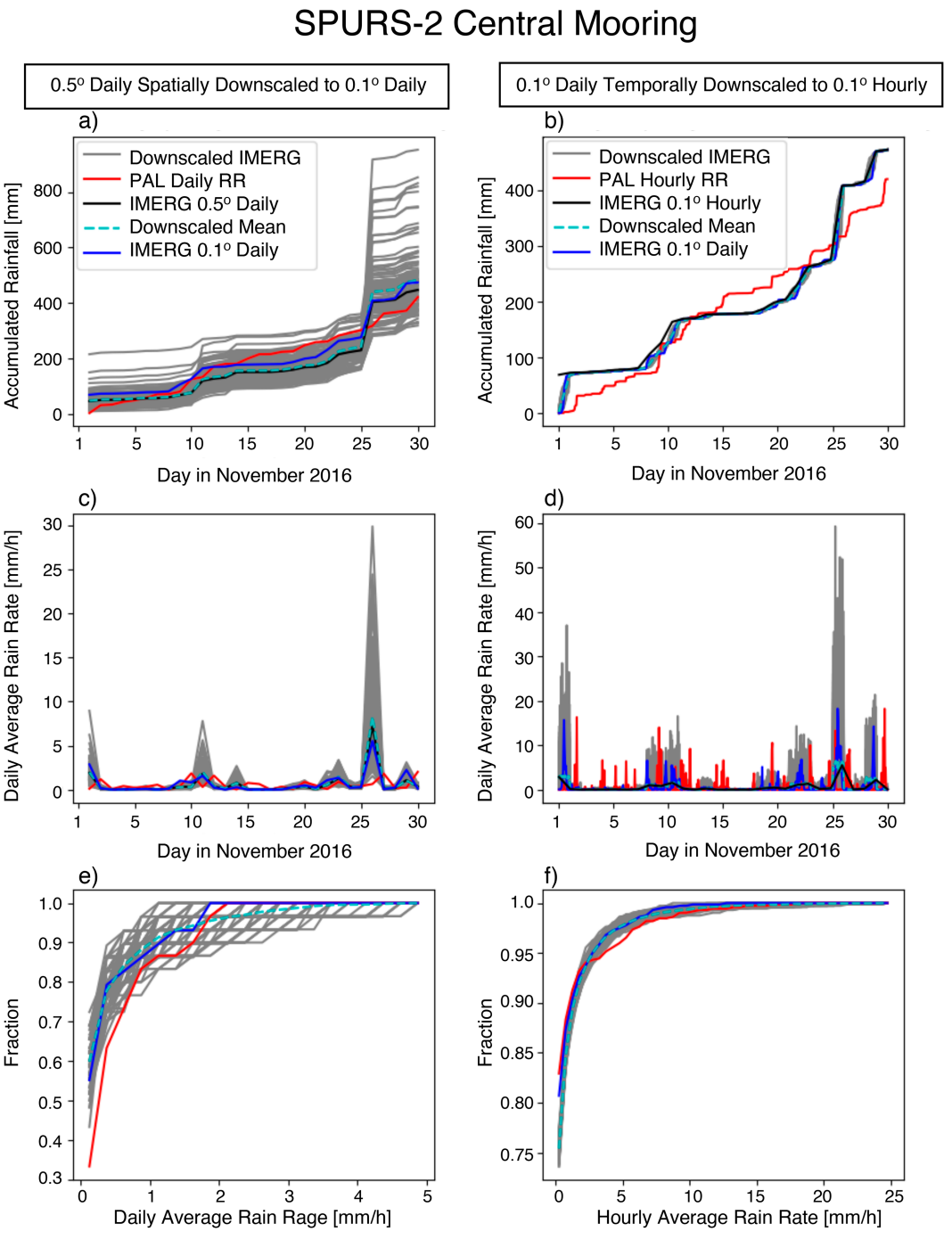

FIG. 2. (left) Spatially and (right) temporally downscaled IMERG ensembles (gray) compared to the SPURS-2 Central Mooring PAL (red), IMERG coarsened prior to downscaling (black), and IMERG coarsened to the same resolution as the downscaled fields (blue) for November 2016 in the SPURS-2 domain. Dashed cyan lines represent the mean of the downscaled ensembles. (a),(b) Accumulated monthly precipitation. (c),(d) Average rain rate at each time step. (e),(f) CDFs of rain rate.

temporally downscaled realizations (Figs. 2a,b). At the end of November 2016, accumulated precipitation estimates from the spatially downscaled ensemble range from 300 mm to more than 800 mm, while the ensembles of realizations from the temporally downscaled experiments are all near 500 mm. In either case, the mean of the downscaled realizations (dashed cyan lines) remains very close to the accumulated precipitation estimate from the coarsened field that was

downscaled (black lines), although the average of the temporally downscaled realizations more closely matches the IMERG coarsened to the same resolution (0.1° 1 h, blue line in Fig. 2b). Because the downscaling algorithm distributes the daily rainfall from the 0.1° daily field over the 24 h in the day according to the extrapolated spectral slope of the daily precipitation field, agreement between the ensemble mean rain and the IMERG 0.1° hourly rain field indicates that the extrapolated spectral information is representing the fine-scale temporal variability of the rain well.

Accumulated precipitation from IMERG at 0.5° daily is lower than the accumulated daily precipitation from the PAL until 25 November, when IMERG captured an intense precipitation event with a daily rainfall of ~150 mm (Fig. 2a). IMERG coarsened to 0.1° daily estimates higher daily precipitation than the PAL early in the month, appears to miss several events midmonth, and returns to overestimates in the wake of the intense precipitation event on the 25th. This event was not captured by the PAL, although the PAL does indicate an intense precipitation event the next day. This highlights the difficulty of comparing gridded satellite precipitation estimates to in situ measurements in highly heterogeneous precipitating fields, even when both observations have relatively similar spatial resolutions (approximately 10-km IMERG versus 5-km PAL). However, when IMERG is coarsened to a resolution at which the precipitation estimates can be considered somewhat reliable prior to downscaling, the spatially downscaled ensemble produced by RainFARM envelopes the in situ measurement. The reduced variability in the temporally downscaled realizations does not provide similar enveloping of the estimates, instead maintaining more consistency with the starting field (Fig. 2b).

Figures 2c and 2d show average rain rates from PAL and both coarsened and downscaled IMERG at each time step. Similar to the prior example, for both the spatially and temporally downscaled fields, the average of the downscaled realizations again closely matches the value of the coarsened fields that were downscaled. The mean of the spatially downscaled realizations also compares well with IMERG averaged to 0.1° daily (Fig. 2c); however, the temporally downscaled ensemble mean does not compare as well with IMERG 0.1° hourly rain rates (Fig. 2d). Figure 2d also shows timing mismatches between PAL and IMERG at hourly temporal scale, with IMERG frequently missing precipitation that is observed by the PAL, or indicating precipitation when the PAL does not, again highlighting the challenges of comparing satellite and surface-based measurements in regions of highly variable precipitation, since the location of the PAL within the IMERG grid box can influence the evaluation statistics. While Fig. 2c also shows that the spatially downscaled realizations mostly envelope the PAL estimates, there are also several time periods when they do not. The temporally downscaled ensemble (Fig. 2d) shows more overlap with PAL at individual hours than the accumulated rainfall time series in Fig. 2b.

A final demonstration of RainFARM preserving the statistics of the large-scale rainfall is shown in the cumulative probability distribution functions (CDFs) of rain rates from PAL, IMERG, and RainFARM (Figs. 2e,f). The small number of daily data points used to create the CDFs from individual ensemble members (30) results in the jumpy, stepwise appearance of the CDFs in Fig. 2e. For both spatial and temporal downscaling, the mean of the downscaled realizations compares closely to the CDF of rain rates from the coarsened IMERG fields at the same resolutions. The spatially downscaled realizations also broaden the CDF to include more frequent higher rain rates than the coarsened IMERG. Thus, the application of RainFARM may provide precipitating ensemble members that capture extremes not found in the coarsened data.

While the ensemble of downscaled realizations successfully captures the statistical properties of the precipitating field and some ensemble members might somewhat resemble the observations, the individual ensemble members are not expected to provide physically realistic representations of the precipitating field. Figure 3a shows IMERG daily rainfall at 0.5° grid spacing for 16 May 2014 off the coast of Central America (CTLAM). Figures 3b and 3c show two individual realizations of the spatially downscaled precipitation field at 0.1° daily resolution. The individual 0.5° grid boxes are still evident in the high-resolution fields, with the rainfall from each 0.5° grid box somewhat randomly distributed into the 0.1° grid boxes in a disorganized and often unphysical manner. Where one would generally expect to see gradients from heavy to light precipitation, the downscaled realizations have more randomly distributed grid boxes with heavy average rain rates sometimes adjacent to grid boxes with no precipitation. Additionally, RainFARM does not change the size of the input precipitating area. The algorithm cannot produce precipitation where the coarse-scale field does not indicate the presence of precipitation, nor will it produce zeros in grid cells where the coarse-scale field indicates nonzero precipitation, although a threshold can be assigned such that any downscaled grid box with rain rate below that threshold is set to 0.

Figures 3d and 3e show a single ensemble member from the field in Fig. 3a temporally downscaled to 0.5° hourly resolution at consecutive time steps (1000 and 1100 UTC, respectively). Spatially, these fields are more physically realistic than the 0.1° fields shown in Figs. 3b and 3c. They also show plausible representations of the temporal evolution of the precipitating field from one hour to the next, with some areas intensifying and others weakening. While individual hourly rain maps, and even a few consecutive hours, may appear physically consistent with observed spatial distributions of precipitation, a loop of these hourly downscaled fields over a longer period of time reveals less realistic behavior, showing stationary precipitating fields with grid boxes that appear to randomly increase or decrease in intensity.

Before presenting conclusions based on evaluation metrics, we first describe the organization of Figs. 4 and 5. Results shown are for downscaled ensembles starting from IMERG daily fields at 0.5° spatial resolution, with evaluation metrics for spatially downscaled realizations shown in the left column (top row in Fig. 5) and for temporally downscaled realizations in the right column (bottom row of Fig. 5). Each panel is divided into five columns representing a different evaluation metric: correlation coefficient (Corr), POD, FAR, CSI, and

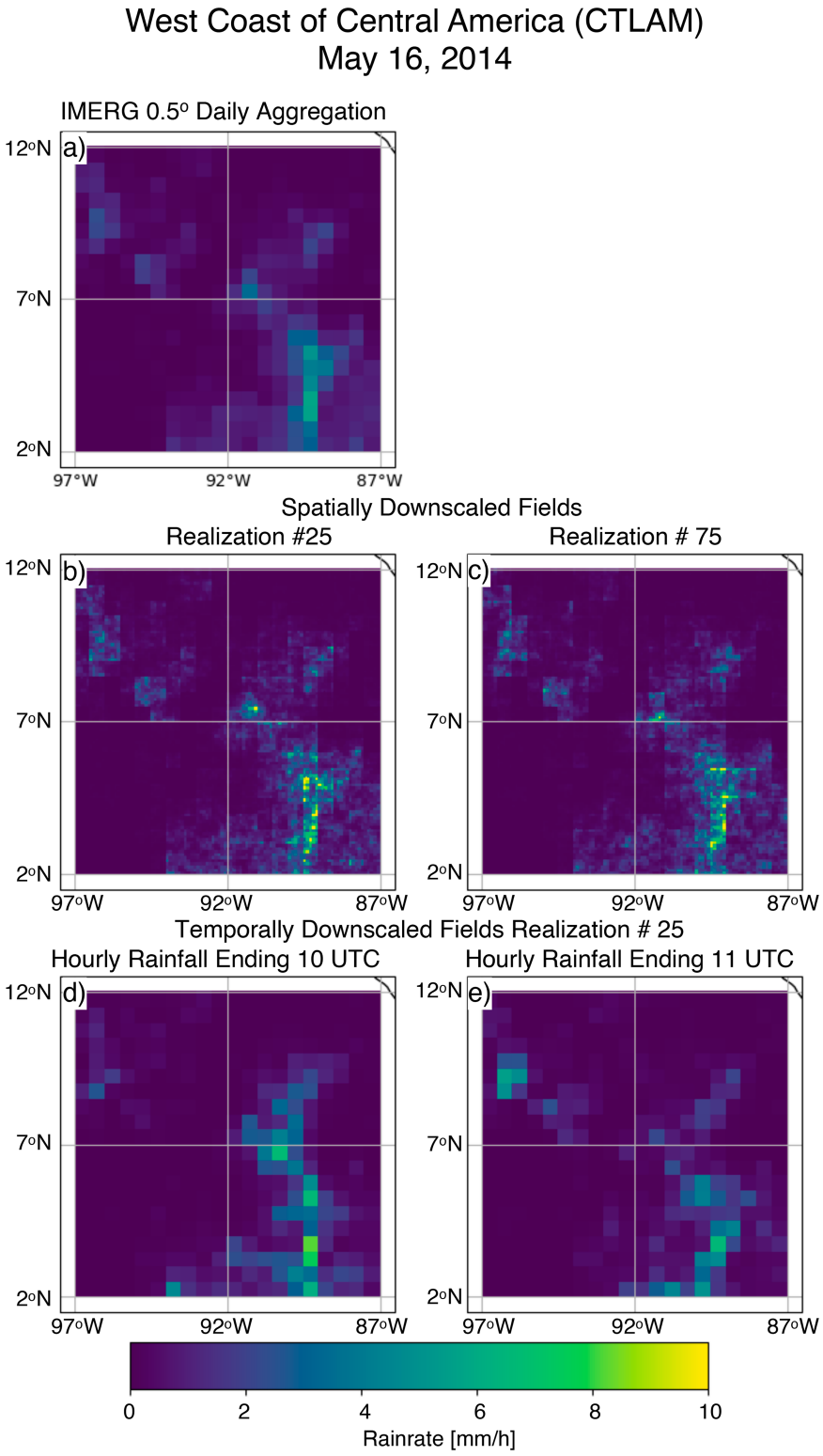

FIG. 3. (a) The 16 May 2014 IMERG daily average rain rate at 0.5° grid spacing in the CTLAM domain. (b),(c) Spatially downscaled realizations of the field shown in (a) at 0.1° grid spacing. (d) Temporally downscaled realizations from ensemble member 25 for 1000 UTC at 0.5° hourly resolution. (e) As in (d), but for 1100 UTC.

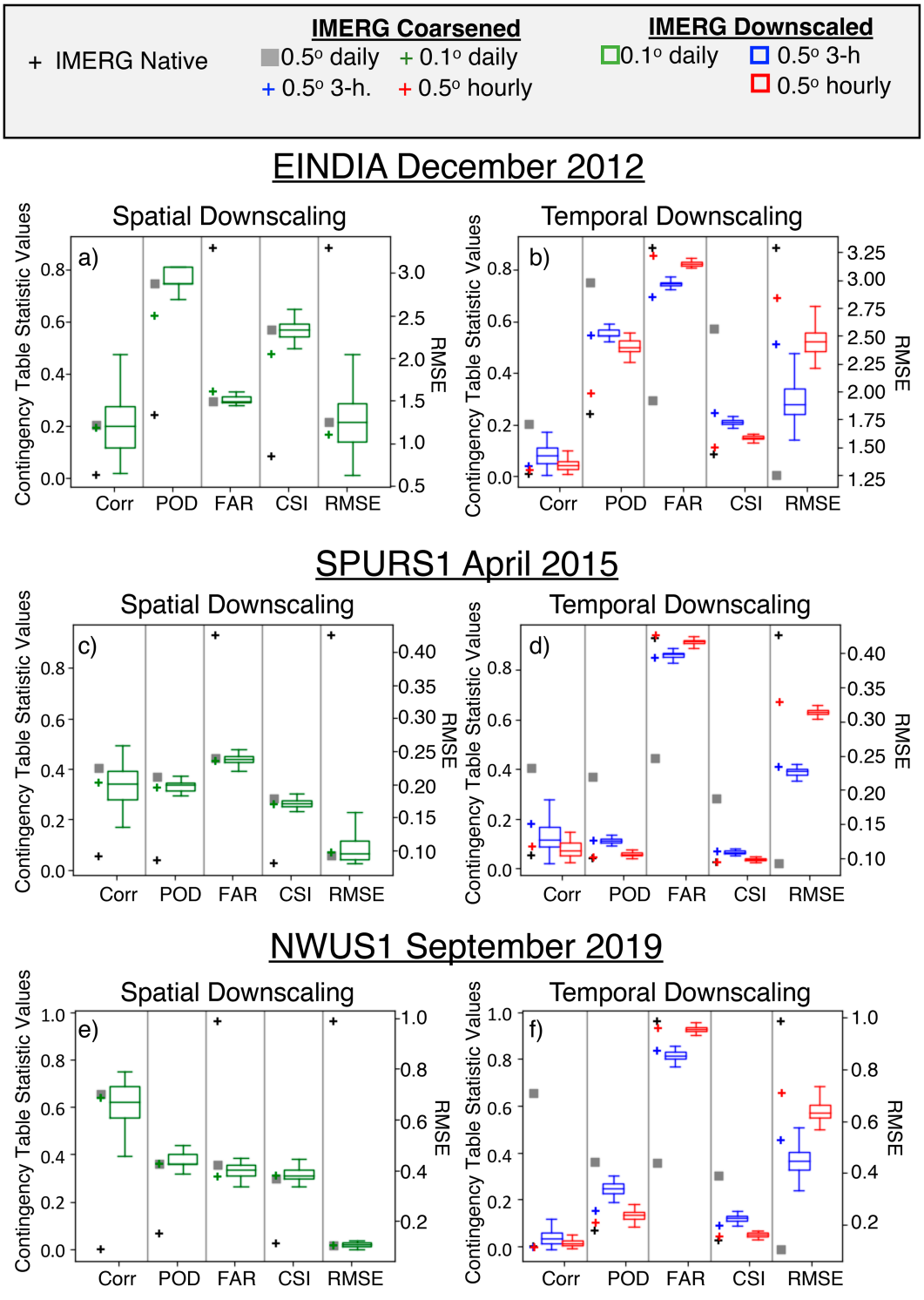

FIG. 4. Evaluation statistics for IMERG at 0.5° daily downscaled (left) spatially and (right) temporally for (a),(b) December 2012 in EINDIA; (c),(d) April 2015 in SPURS-1; and (e),(f) September 2019 in NWUS1. Data points in the leftmost portion of each column represent IMERG at native (black cross), 0.5° daily (gray square), 0.1° daily (green cross), 0.5° 3 h (blue cross), and 0.5°1 h (red cross). Boxplots show median, IQR, and 1.5 times IQR of values calculated for 100-member RainFARM-produced ensembles.

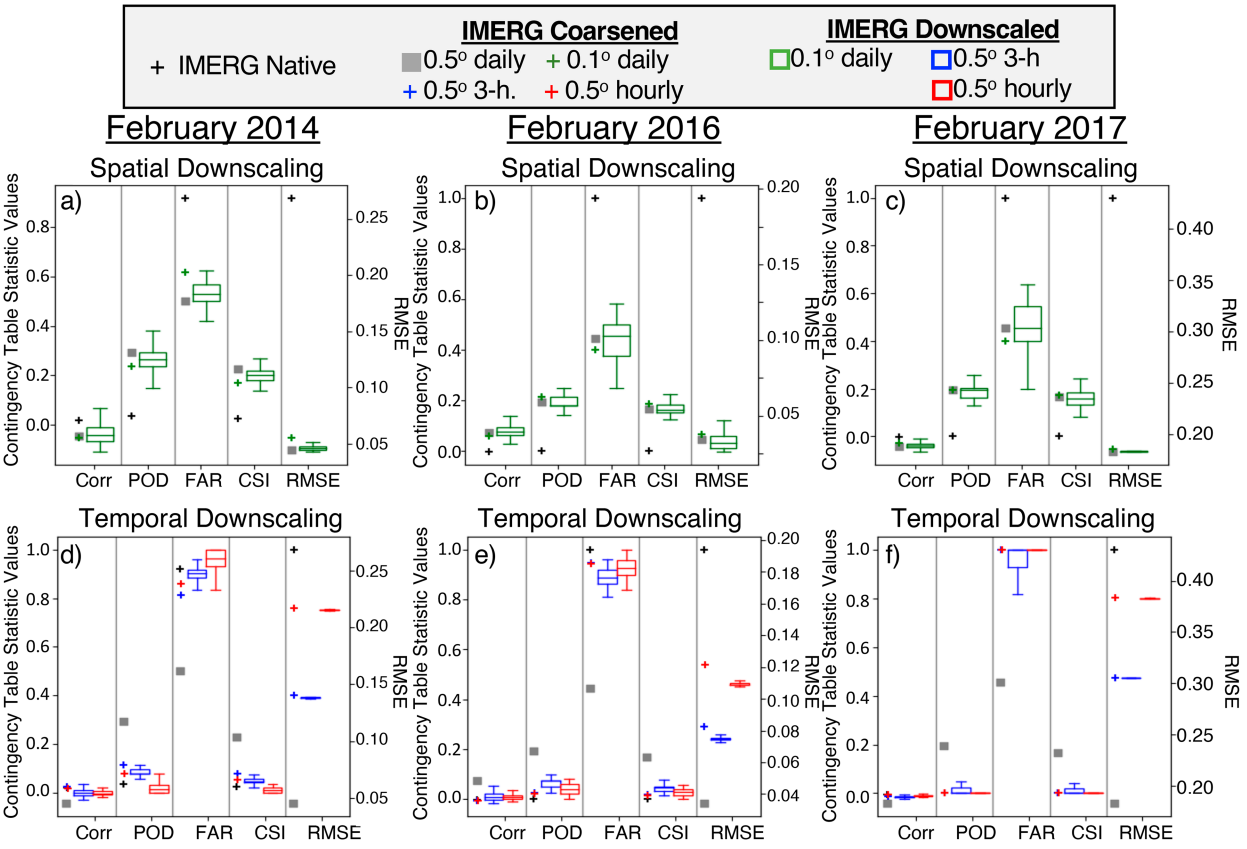

FIG. 5. (top) As in the left panels of Fig. 4 and (bottom) as in the right panels of Fig. 4, but for the PERU domain in (a),(d) February 2014; (b),(e) February 2016; and (c),(f) February 2017.

RMSE. In each column, a series of individual data points (crosses or square) in the leftmost portion represents metrics calculated for coarsened IMERG fields. All panels have a gray square representing the IMERG at 0.5° daily (i.e., the field being downscaled) and a black cross representing IMERG at native (0.1° 0.5 h) resolution. The remaining crosses represent IMERG coarsened to the resolutions produced by RainFARM: 0.1° daily (green) and 0.5° at 3-h (blue) and 1-h (red) time steps. Spatial downscaling is only performed on the 0.5° daily fields. Results from temporally downscaling 0.1° daily IMERG fields are not shown for brevity and clarity of the figures and unless otherwise noted were similar to the results shown for 0.5° fields. Boxplots show the distribution of the evaluation statistics calculated for the 100-member high-resolution ensemble produced by RainFARM, with the median denoted by the horizontal line within the box, the box itself indicating the interquartile range (IQR), and whiskers indicating 1.5 times the IQR. Boxplots are color matched to the crosses representing the coarsened IMERG fields (i.e., green for 0.1° daily, blue for 3 hourly, and red for hourly), and this color scheme will be used to represent the same resolution fields throughout the remainder of the manuscript.

The RainFARM downscaling algorithm produces an ensemble of precipitation scenarios that aim to characterize the expected variability of the fine-scale field, as informed by the coarse field. We compare evaluation metrics for the coarsened IMERG field, the downscaled ensembles, and IMERG coarsened to the same resolution as the downscaled ensembles. If the coarsened IMERG is inconsistent with the in situ observations, the power spectra are not expected to provide the necessary information to estimate higher resolution rainfall. In this case, a simple coarsening of IMERG from native resolution to an intermediate resolution should provide a better comparison to PAL than the downscaled IMERG results. Conversely, if the coarsened IMERG field compares favorably to PAL, the spatial and temporal power spectra *may* provide additional guidance on the expected fine-scale rainfall variability, such that coarsening then downscaling provides an ensemble that compares more favorably to PAL than coarsening alone. It is also possible for coarsened IMERG to compare favorably to PAL, but for the spatial and/or temporal power spectra to not follow linear scaling behavior, which would also potentially lead to worse comparisons to PAL by the downscaled ensemble.

Because precipitation characteristics vary in space and time, RainFARM produced a variety of results. In some cases, the entire distribution of evaluation statistics for the downscaled ensemble is shifted toward better or worse comparison to PALs than IMERG coarsened to the same resolution, and in many cases, the downscaled ensembles compare

more favorably to PAL in some metrics and less favorably in others. Several examples of the downscaled realizations providing improved, degraded, or relatively unchanged comparisons to PALs than those of the coarsened IMERG fields are shown in Fig. 4 for December 2012 in the Bay of Bengal (EINDIA, Figs. 4a,b), April 2015 in the SPURS-1 field campaign region (Figs. 4c,d), and September 2019 over U.S. West Coast near Washington and Oregon (Figs. 4e,f).

For the spatially downscaled realizations in the East India domain (Fig. 4a), the median of the ensemble evaluation statistics is nearly equal to the values calculated for IMERG coarsened to 0.5° daily (i.e., the field that was downscaled). For POD, FAR, and CSI, the IQR of the ensemble evaluation statistics is shifted toward better values than IMERG coarsened to the same resolution. The median ensemble correlation coefficient (~0.2) is nearly equal to that obtained when comparing PAL to IMERG coarsened to 0.1° daily, while the ensemble RMSE distribution is shifted toward higher values than IMERG at the same resolution. This suggests that coarsening then downscaling provides a more reliable estimate of the fine-scale spatial variability, although the fine-scale intensity representation may be less reliable (i.e., contingency table statistics improve, but metrics based on intensity are the same or worsen). The temporally downscaled realizations in EINDIA (Fig. 4b) behave differently, with slightly increased correlation coefficients compared to IMERG coarsened to the same resolution and RMSE distributions that are entirely lower than the coarsened fields. For POD, FAR, and CSI, the entire distribution of the 0.5° 1-h ensemble is shifted toward improved comparison to PALs than the IMERG coarsened to the same resolution. Downscaling to 0.5° 3 h provides very slightly improved POD, while FAR and CSI are both negatively impacted by the downscaling (Fig. 4b).

In the SPURS-1 domain, the spread of values from the downscaled ensembles is smaller than that calculated in EINDIA. For most metrics, and both spatial and temporal downscaling, the median value of the ensemble of downscaled fields is the same as that calculated for IMERG coarsened to the same resolution (Figs. 4c,d). This is especially true for spatially downscaled fields. For temporally downscaled fields, ensembles show almost no change from the coarsened IMERG values for POD and CSI, and very small changes in correlation, RMSE and FAR. In this domain, coarsening then downscaling provides little benefit over simply coarsening IMERG to slightly lower resolution.

Turning to an extratropical domain, compared to coarsened IMERG, the spatial downscaling in NWUS1 produces very little change in CSI and RMSE (Fig. 4e), a very slight improvement in POD, and overall lower correlation coefficient and higher FAR. Where the reliability of the spatially downscaled ensemble is mostly reduced or unchanged compared to coarsened IMERG, temporal downscaling in this domain results in slight improvements for all five of the calculated metrics (Fig. 4f). This suggests that the spatial spectral slopes do not provide much information about the fine-scale variability, while the temporal power spectrum can inform a precipitation time series that is more consistent at PAL sites than IMERG coarsened to the same resolution.

Statistic	Downscaling	2014	2016	2017
Corr	Spatial			
	Temporal 3h			
	Temporal 1h			
POD	Spatial			
	Temporal 3h			
	Temporal 1h			
FAR	Spatial			
	Temporal 3h			
	Temporal 1h			
CSI	Spatial			
	Temporal 3h			
	Temporal 1h			
RMSE	Spatial			
	Temporal 3h			
	Temporal 1h			

FIG. 6. Change in evaluation metrics from downscaled ensembles compared to IMERG coarsened to the same resolution for spatially and temporally downscaled fields in the PERU domain for February 2014, 2016, and 2017. Green indicates better performance, while red indicates degraded performance and yellow indicates little to no change.

Figure 4 highlights the inconsistency of RainFARM's ability to represent fine-scale variability between domains, and sometimes, inconsistency between spatial and temporal downscaling results in the same domain. For a given location and season, one might expect similar precipitation characteristics year to year, and therefore similar capacity for RainFARM to produce reliable estimates of fine-scale variability, with exceptions for areas strongly impacted by decadal oscillations such as ENSO. Figure 5 shows results from the month of February over 3 years (2014, 2016, 2017) in the PERU domain. Results are not shown for February 2015 because rain rates in the downscaled fields at most resolutions did not reach the thresholds necessary for calculation of contingency table statistics. A summary of the benefit of downscaling for each metric in each year is shown in Fig. 6, with green boxes indicating improved representation of fine-scale variability over coarsened IMERG, red indicating that the downscaling produced worsened comparisons to PAL observations (i.e., the distribution

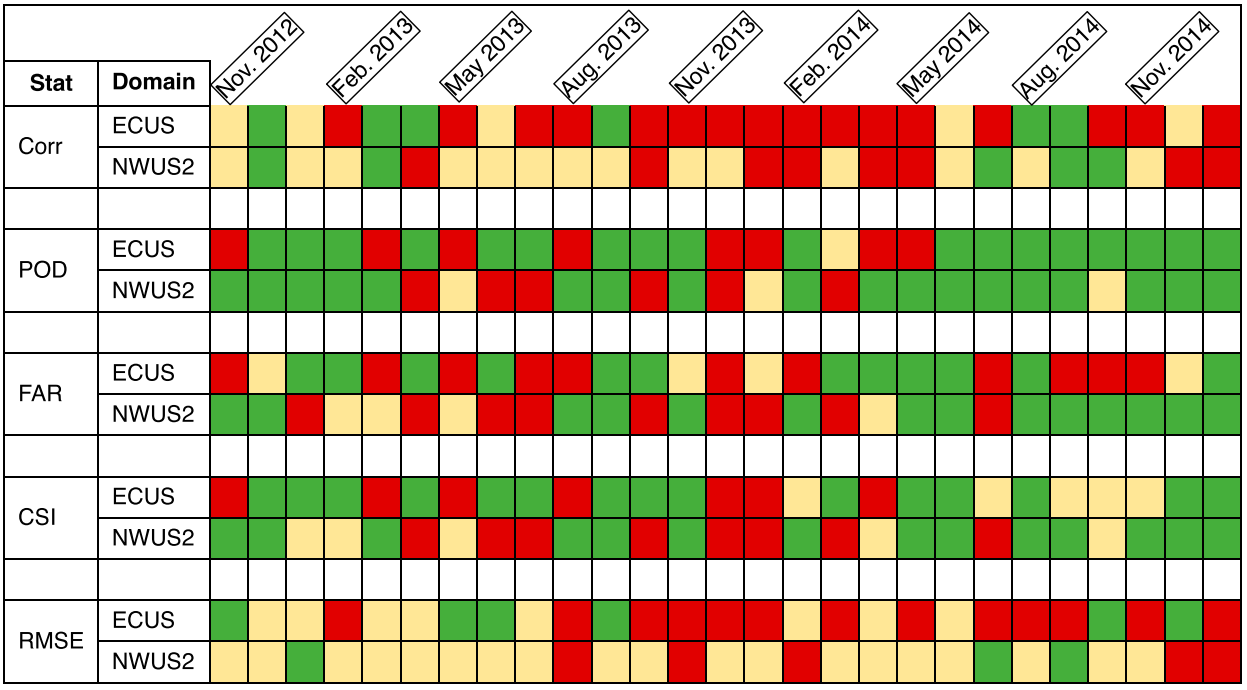

FIG. 7. As in Fig. 6, but for spatial downscaling in the ECUS and NWUS2 domains for each month that both domains had PAL deployments.

of ensemble evaluation metrics is shifted toward better or worse values, respectively, than IMERG at the same resolution), and yellow indicating relatively little change in performance (i.e., the ensemble median is approximately equal to coarsened IMERG with ensemble spread not skewed toward either better or worse values). In 2014, spatial downscaling resulted in overall improved validation statistics, but temporal downscaling produced no change or worsened comparisons with the PALs. In 2016, temporal downscaling resulted in improved statistics, while spatial downscaling results were mixed—the correlation coefficient and RMSE were improved, but contingency table statistics were worse. In 2017, spatially downscaled ensembles overall compared less favorably to PAL than did IMERG coarsened to the same resolution, while temporally downscaled ensembles did not make much difference, except for contingency table statistics at 3 h.

We have shown that, when comparing IMERG to PAL, downscaling does not consistently improve the representation of the high-resolution precipitation from domain to domain, or for the same season in a single domain. To further illustrate the inconsistency of downscaling results, we now discuss RainFARM's performance in two domains that should exhibit similar precipitation. Both the ECUS and NWUS2 domains are located along the extratropical storm tracks. One might expect that while precipitation patterns are not identical, they might exhibit similar spatial and temporal spectral rain characteristics since much of the precipitation is the result of strong dynamical forcing by extratropical cyclones. Figures 7–9 summarize the overall changes in monthly evaluation statistics that result from applying RainFARM downscaling to coarsened IMERG versus coarsening IMERG only, employing the same "stoplight" color scheme used in Fig. 6 for spatial, temporal to 3 h, and temporal to 1 h downscaling, respectively. Full time series of the change in the distributions of the evaluation statistics are available in the online supplemental material.

In Fig. 7, there are clear differences in the spatially downscaled results between the two domains, particularly in correlation and RMSE. Whereas spatially downscaled fields in the ECUS domain most often compare less favorably to PAL than coarsened IMERG, in the NWUS2 domain spatial downscaling most often results in little to no change in these two metrics. For the contingency table statistics, there are often months where downscaling has the same impact on evaluation metrics in both domains, but there does not appear to be a pattern (e.g., seasonal cycle) as to when there is agreement and when there is not.

The variable impact of downscaling between the two domains is more obvious for temporally downscaled ensembles shown in Figs. 8 and 9 for 3- and 1-h time scales, respectively. While both domains exhibit mostly improved POD when temporally downscaled, the remaining comparison metrics in the ECUS domain worsened in most months for the remaining statistics, while the downscaled validation statistics in the NWUS2 domain frequently exhibit improvement over validation statistics from IMERG coarsened to the same resolution.

5. Discussion and conclusions

In this paper, we described the application of the RainFARM statistical downscaling technique to rain-rate estimates from IMERG. IMERG fields were coarsened to lower spatial and temporal resolution prior to downscaling, and downscaling

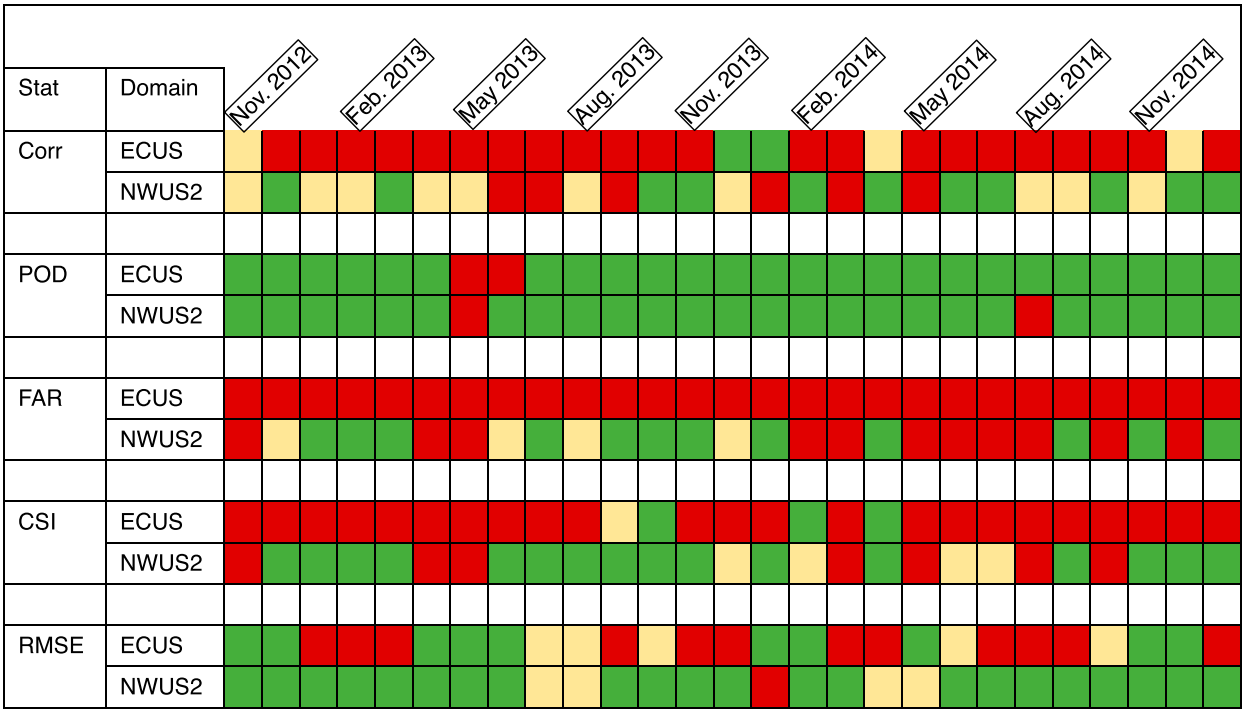

FIG. 8. As in Fig. 7, but for temporal downscaling to 0.5° 3 h.

was applied in both space and time. The ensembles of downscaled precipitation fields were compared to in situ precipitation observations from PALs in 11 domains over the global oceans. IMERG was also coarsened to the same resolutions that were produced by the downscaling to determine whether the process of coarsening then downscaling produces estimates of fine-scale rainfall variability that compare more favorably to the PALs than coarsening IMERG alone. The

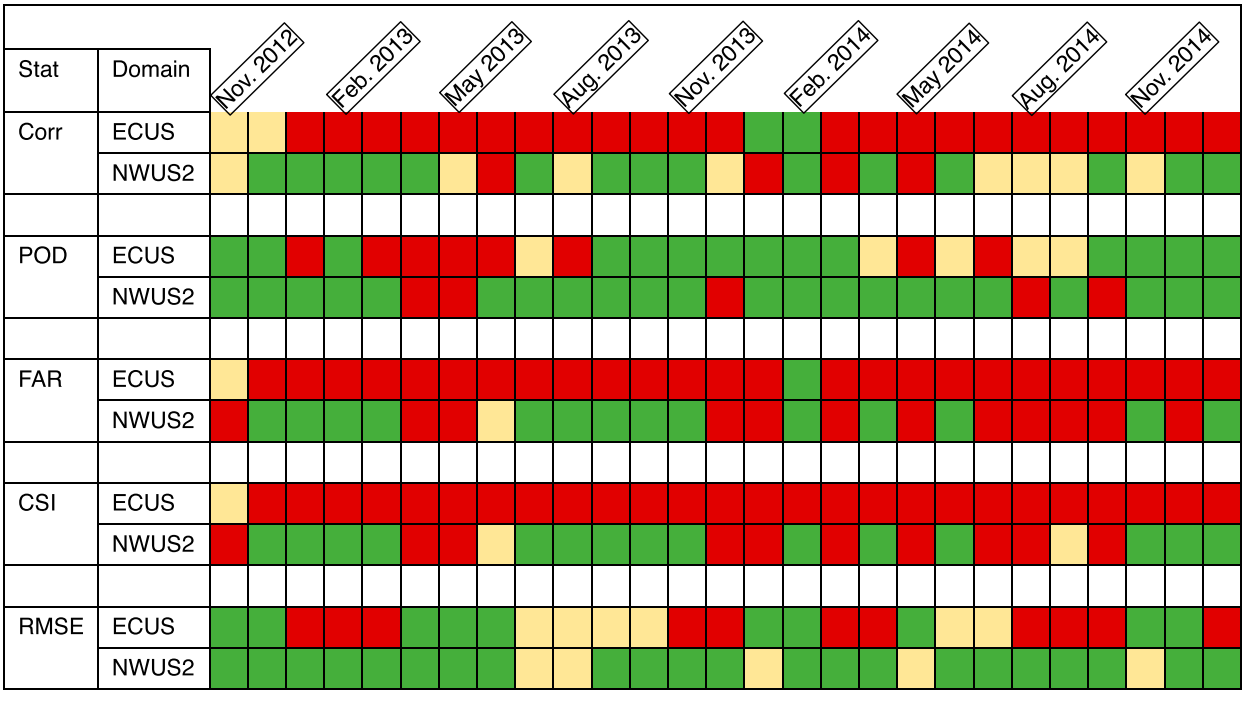

FIG. 9. As in Fig. 7, but for temporal downscaling to 0.5° 1 h.

RainFARM algorithm was chosen for its publicly available source code and demonstrated prior ability to produce ensembles of precipitation fields that are statistically consistent with the large- and small-scale observed precipitation fields.

Evaluations of the downscaled precipitation ensembles showed inconsistent results in all domains, with some months showing more consistency with the in situ data and others showing worsened or unchanged comparisons to PAL than IMERG coarsened to the same resolution. Comparisons to IMERG at coarsened resolution are a significant benchmark since IMERG tends to show better performance relative to in situ observations at increased spatial and temporal scales (e.g., daily, degree-wide). In some cases, spatial downscaling provided improved estimates of precipitation over coarsened IMERG, while in other cases, temporal downscaling produced more favorable results. In others still, both spatial and temporal downscaling produced either improved or worsened evaluation metrics than coarsened IMERG. The ability of the RainFARM algorithm to produce improved estimates of fine-scale precipitation compared to coarsened IMERG depends both on the reliability of the coarse-scale (0.5° daily) precipitation estimate, and on the spatial and temporal power spectra exhibiting a scaling behavior such that information from large scales can be extrapolated to fine scales. These results suggest that there are conditions under which the spatial and temporal power spectra from coarsened IMERG provide sufficient information to better estimate fine-scale variability via downscaling, and there does not appear to be a systematic indication that precipitation regime influences whether spatial or temporal downscaling will work better, or at all. Because the coarse-scale precipitation field needs to contain enough data for the RainFARM algorithm to adequately calculate the spectral slopes, we chose to perform the downscaling on each individual month. It is possible that downscaling on individual precipitating events could impact the outcome of the downscaling, but this is both beyond the scope of this work and would require a denser validation dataset. Additionally, the use of a constant or different scale-dependent threshold for computing contingency table statistics may produce different results. Sensitivity analysis to the threshold used for evaluation was not performed as part of this study.

Spatially downscaled realizations provide a more varied ensemble, but the spatial distribution of the precipitation field is very unrealistic. Single maps of temporally downscaled realizations look more physically realistic but produce ensembles with smaller variability, and ensemble members also have no memory of the spatial distribution of precipitation from previous time steps, resulting in an unrealistic time series of precipitation fields when viewed as a loop. For most of the simulations, the spread of the evaluation metrics calculated with the downscaled ensemble envelopes the value of the same metric calculated for IMERG coarsened to the same resolution. In those cases, coarsening then downscaling provides an estimate of the uncertainty of the coarsened IMERG field, which could be a benefit in some applications, but does not necessarily provide a better or worse estimate of the precipitation at a given point.

D'Onofrio et al. (2014) applied RainFARM to RCM output and showed that, when compared to a dense gauge network, the downscaled ensemble fields compared more favorably to the gauge data than the coarse resolution model output. In this work, we find that RainFARM cannot reliably produce such improvements over the ocean, likely due to two characteristics of the previous work that we were unable to recreate in this study.

First, the validity of the field being downscaled is very important, as a more reliable coarse field will produce a more reliable downscaled ensemble. Previous studies have started with coarsened radar data or physically based atmospheric model output that already have relatively reliable representations of precipitation. In this study, the starting data, IMERG, is known to be unreliable at its native resolution, so it was coarsened to lower resolutions that are known to be more reliable (Tan et al. 2017; Bytheway et al. 2023). However, the quality of these estimates still leaves much to be desired. In fact, Guilloteau et al. (2021) found that while IMERG temporal power spectra compared well to a radar-based multisensor reference over the southeastern United States at scales down to 4 h, spatial power spectra comparisons became poor at scales finer than 250 km. Thus, even going up to 0.5° may not be sufficient to obtain a reliable enough estimate of precipitation from IMERG for RainFARM to be successful. This points to the need for continued work on satellite precipitation algorithms to improve the accuracy of these products at their native resolution.

Second, the D'Onofrio et al. (2014) study had a dense sensor network for comparison, whereas most of the domains considered in this study contained fewer than five PALs for comparison. The ability to gain improvements in the evaluation metrics is strongly dependent on how RainFARM distributes rainfall from a coarse scale into multiple higher resolution parcels (either grid box or time steps) given the extrapolated spectral information. For example, for spatial downscaling, we began with a relatively large 0.5° square grid box, which was distributed into 25 grid boxes of 0.1°. As shown in Fig. 3, the spatial distribution of precipitation in these smaller grid boxes is physically unrealistic, and the performance of a given ensemble member depends on the rain rate assigned to the 0.1° box containing the PAL. While this same issue also applies to previous studies, the use of a denser sensor network can provide more points for comparison and a clearer picture of how downscaling changes the performance of the precipitation estimate ensemble. It is anticipated that with a denser in situ sensor network, we would find more consistent or more explainable results in each domain or between domains.

In addition to the results shown herein, downscaled realizations were also compared to PAL based on seasons (wet versus dry or warm versus cold), ENSO phase, and based on precipitation characteristics (monthly mean rain rate and spatial and temporal spectral slopes). In several domains, there was a clear relationship between factors impacting storm characteristics and performance of the downscaled realizations (e.g., improvement over coarsened IMERG in wet season but not dry, or for a given ENSO phase). However, like

the results already discussed, the conditions for realizing the benefits of downscaling were not consistently present between domains, and an overall description of storm characteristics that produce reliable downscaled realizations from IMERG using RainFARM remains to be found.

The need for high-resolution, accurate estimates of precipitation over ocean remains crucial to our understanding and monitoring of oceanic processes and ocean–atmosphere coupling. Air–sea fluxes, ocean freshening, ocean turbulence, and ocean cooling are all impacted by the instantaneous rain rate and operate on scales much faster and smaller than daily, 1° satellite estimates of rain that have proven accuracy. Currently, precipitation estimates based on retrievals from satellite observations are the best way to continuously monitor global precipitation over oceans, but these estimates are in need of continued improvement and refinement. While the algorithms that produce these products are almost always undergoing development and improvement, each update is a resource-intensive undertaking that can take several years to complete. In the meantime, it remains worthwhile to pursue computationally inexpensive postprocessing techniques such as more advanced downscaling or machine learning techniques to try to improve precipitation estimates over the oceans, while recognizing that these techniques may be fundamentally limited by the representativeness of the input coarse-scale field.

Acknowledgments. This work was funded by NASA PMM Grant 18-PMMST18-0018. JLB acknowledges additional support by NOAA Cooperative Agreements NA17OAR4320101 and NA22OAR4320151. EJT was supported by the NOAA Physical Sciences Laboratory. The authors thank Jost von Hardenberg for providing the RainFARM code and for help interpreting early results and Mimi Hughes for internal review of this manuscript. The authors also thank the Argo group, led by Stephen C. Riser, at the School of Oceanography and the engineering team at the Applied Physics Laboratory, University of Washington, and funding agencies of NOAA and NASA for their continual support since 2008.

Data availability statement. IMERG V6 data are available via NASA Earthdata Search: https://search.earthdata.nasa.gov/search?q=GPM_3IMERGHH_06 (requires registration). RainFARM is available in Julia language from https://github.com/jhardenberg/RainFARM.jl. PAL data are available from NASA https://dx.doi.org/10.5067/GPMGV/PAL/DATA101.

REFERENCES

Abbasian, M. S., A. Abrishamachi, M. R. Najafi, and S. Moghim, 2020: Multi-site statistical downscaling of precipitation using generalized hierarchical linear models: A case study of the imperiled Lake Urmia basin. *Hydrol. Sci. J.*, **65**, 2466–2481, https://doi.org/10.1080/02626667.2020.1810255.

American Meteorological Society, 2015: Regional climate model. Glossary of Meteorology, https://glossary.ametsoc.org/wiki/Regional_climate_model.

Bingham, F. M., V. Tsontos, A. deCharon, C. J. Lauter, and L. Taylor, 2019: The SPURS-2 eastern tropical Pacific field campaign data collection. *Oceanography*, **32** (2), 142–149, https://doi.org/10.5670/oceanog.2019.222.

Brussolo, E., J. von Hardenberg, L. Ferraris, N. Rebora, and A. Provenzale, 2008: Verification of quantitative precipitation forecasts via stochastic downscaling. *J. Hydrometeor.*, **9**, 1084–1094, https://doi.org/10.1175/2008JHM994.1.

——, ——, and N. Rebora, 2009: Stochastic versus dynamical downscaling of ensemble precipitation forecasts. *J. Hydrometeor.*, **10**, 1051–1061, https://doi.org/10.1175/2009JHM1109.1.

Bytheway, J. L., E. J. Thompson, J. Yang, and H. Chen, 2023: Evaluating satellite precipitation estimates over oceans using passive aquatic listeners. *Geophys. Res. Lett.*, **50**, e2022GL102087, https://doi.org/10.1029/2022GL102087.

Chen, Y., J. Huang, S. Sheng, L. R. Mansaray, Z. Liu, H. Wu, and X. Wang, 2018: A new downscaling-integration framework for high-resolution monthly precipitation estimates: Combining rain gauge observations, satellite-derived precipitation data and geographical ancillary data. *Remote Sens. Environ.*, **214**, 154–172, https://doi.org/10.1016/j.rse.2018.05.021.

D'Onofrio, D., E. Palazzi, J. von Hardenberg, A. Provenzale, and S. Calmanti, 2014: Stochastic rainfall downscaling of climate models. *J. Hydrometeor.*, **15**, 830–843, https://doi.org/10.1175/JHM-D-13-096.1.

Drushka, K., W. E. Asher, B. Ward, and K. Walesby, 2016: Understanding the formation and evolution of rain-formed fresh lenses at the ocean surface. *J. Geophys. Res. Oceans*, **121**, 2673–2689, https://doi.org/10.1002/2015JC011527.

——, ——, A. T. Jessup, E. J. Thompson, S. Iyer, and D. Clark, 2019: Capturing fresh layers with the surface salinity profiler. *Oceanography*, **32** (2), 76–85, https://doi.org/10.5670/oceanog.2019.215.

Ferraris, L., S. Gabellani, N. Rebora, and A. Provenzale, 2003: A comparison of stochastic models for spatial rainfall downscaling. *Water Resour. Res.*, **39**, 1368, https://doi.org/10.1029/2003WR002504.

Foufoula-Georgiou, E., A. M. Ebtehaj, S. Q. Zhang, and A. Y. Hou, 2014: Downscaling satellite precipitation with emphasis on extremes: A variational ℓ_1-norm regularization in the derivative domain. *Surv. Geophys.*, **35**, 765–783, https://doi.org/10.1007/s10712-013-9264-9.

Gosnell, R., C. W. Fairall, and P. J. Webster, 1995: The sensible heat of rainfall in the tropical ocean. *J. Geophys. Res.*, **100**, 18 437–18 442, https://doi.org/10.1029/95JC01833.

Guilloteau, C., E. Foufoula-Georgiou, P. Kirstetter, J. Tan, and G. J. Huffman, 2021: How well do multisatellite products capture the space-time dynamics of precipitation? Part I: Five products assessed via a wavenumber–frequency decomposition. *J. Hydrometeor.*, **22**, 2805–2823, https://doi.org/10.1175/JHM-D-21-0075.1.

Gutmann, E. D., J. J. Hamman, M. P. Clark, T. Eidhammer, A. W. Wood, and J. R. Arnold, 2022: En-GARD: A statistical downscaling framework to produce and test large ensembles of climate projections. *J. Hydrometeor.*, **23**, 1545–1561, https://doi.org/10.1175/JHM-D-21-0142.1.

He, X., N. W. Chaney, M. Schleiss, and J. Sheffield, 2016: Spatial downscaling of precipitation using adaptable random forests. *Water Resour. Res.*, **52**, 8217–8237, https://doi.org/10.1002/2016WR019034.

Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J.*

Appl. Meteor., **43**, 1834–1853, https://doi.org/10.1175/JAM2173.1.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, https://doi.org/10.1175/JHM560.1.

——, and Coauthors, 2018: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). Algorithm Theoretical Basis Doc., version 5.2, 35 pp., https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V5.2_0.pdf.

Iyer, S., and K. Drushka, 2021a: The influence of preexisting stratification and tropical rain modes on the mixed layer salinity response to rainfall. *J. Geophys. Res. Oceans*, **126**, e2021JC017574, https://doi.org/10.1029/2021JC017574.

——, and ——, 2021b: Turbulence within rain-formed fresh lenses during the SPURS-2 experiment. *J. Phys. Oceanogr.*, **51**, 1705–1721, https://doi.org/10.1175/JPO-D-20-0303.1

Kang, B., and J. A. Ramírez, 2010: A coupled stochastic space-time intermittent random cascade model for rainfall downscaling. *Water Resour. Res.*, **46**, W10534, https://doi.org/10.1029/2008WR007692.

Klepp, C., 2015: The oceanic shipboard precipitation measurement network for surface validation—OceanRAIN. *Atmos. Res.*, **163**, 74–90, https://doi.org/10.1016/j.atmosres.2014.12.014.

——, and Coauthors, 2018: OceanRAIN, a new in-situ shipboard global ocean surface-reference dataset of all water cycle components. *Sci. Data*, **5**, 180122, https://doi.org/10.1038/sdata.2018.122.

Kucera, P. A., and C. Klepp, 2022: Evaluation of high-resolution satellite precipitation over the global oceans. *Precipitation Science: Measurement, Remote Sensing, Microphysics and Modeling*, S. Michaelides, Ed., Elsevier, 305–322, https://doi.org/10.1016/B978-0-12-822973-6.00008-1.

Li, Z., E. J. Thompson, A. Behrangi, H. Chen, and J. Yang, 2023: Performance of GPCP products over oceans: Evaluation using passive aquatic listeners. *Geophys. Res. Lett.*, **50**, e2023GL104310, https://doi.org/10.1029/2023GL104310.

Lindstrom, E. J., and Coauthors, 2017: Autonomous multi-platform observations during the salinity processes in the upper-ocean regional study. *Oceanography*, **30** (2), 38–48, https://doi.org/10.5670/oceanog.2017.218.

——, J. B. Edson, J. J. Schanze, and A. Y. Shcherbina, 2019: SPURS-2: Salinity processes in the upper-ocean regional study 2—The eastern equatorial Pacific experiment. *Oceanography*, **32** (2), 15–19, https://doi.org/10.5670/oceanog.2019.207.

Ma, B. B., and J. A. Nystuen, 2005: Passive acoustic detection and measurement of rainfall at sea. *J. Atmos. Oceanic Technol.*, **22**, 1225–1248, https://doi.org/10.1175/JTECH1773.1.

Maggioni, V., P. C. Meyers, and M. D. Robinson, 2016: A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *J. Hydrometeor.*, **17**, 1101–1117, https://doi.org/10.1175/JHM-D-15-0190.1.

Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.*, **48**, RG3003, https://doi.org/10.1029/2009RG000314.

Nystuen, J. A., 2001: Listening to raindrops from underwater: An acoustic disdrometer. *J. Atmos. Oceanic Technol.*, **18**, 1640–1657, https://doi.org/10.1175/1520-0426(2001)018<1640:LTRFUA>2.0.CO;2.

Prakash, S., C. Mahesh, and R. M. Gairola, 2012: Observed relationship between surface freshwater flux and salinity in the North Indian Ocean. *Atmos. Oceanic Sci. Lett.*, **5**, 163–169, https://doi.org/10.1080/16742834.2012.11446984.

——, M. R. Ramesh Kumar, S. Mathew, and R. Venkatesan, 2018: How accurate are satellite estimates of precipitation over the North Indian Ocean? *Theor. Appl. Climatol.*, **134**, 467–475, https://doi.org/10.1007/s00704-017-2287-2.

Ramesh Kumar, M. R., and J. Shulz, 2002: Analysis of freshwater flux climatology over the Indian Ocean using the HOAPS data. *Remote Sens. Environ.*, **80**, 363–372, https://doi.org/10.1016/S0034-4257(01)00302-9.

Rebora, N., L. Ferraris, J. von Hardenberg, and A. Provenzale, 2006a: RainFARM: Rainfall downscaling by a filtered autoregressive model. *J. Hydrometeor.*, **7**, 724–738, https://doi.org/10.1175/JHM517.1.

——, ——, ——, and ——, 2006b: Rainfall downscaling and flood forecasting: A case study in the Mediterranean area. *Nat. Hazards Earth Syst. Sci.*, **6**, 611–619, https://doi.org/10.5194/nhess-6-611-2006.

Reverdin, G., A. Supply, K. Drushka, E. J. Thompson, W. E. Asher, and A. Lourenço, 2020: Intense and small freshwater pools from rainfall investigated during Spurs-2 on 9 November 2017 in the eastern tropical Pacific. *J. Geophys. Res. Oceans*, **125**, e2019JC015558, https://doi.org/10.1029/2019JC015558.

Riser, S. C., J. Yang, and R. Drucker, 2019: Observations of large-scale rainfall, wind and sea surface salinity variability in the eastern tropical Pacific. *Oceanography*, **32**(2), 42–49, https://doi.org/10.5670/oceanog.2019.211.

Sapiano, M. R. P., and P. A. Arkin, 2009: An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeor.*, **10**, 149–166, https://doi.org/10.1175/2008JHM1052.1.

Tan, J., W. A. Petersen, P.-E. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeor.*, **18**, 307–319, https://doi.org/10.1175/JHM-D-16-0174.1.

——, G. J. Huffman, D. T. Bolvin, and E. J. Nelkin, 2019: IMERG V06: Changes to the morphing algorithm. *J. Atmos. Oceanic Technol.*, **36**, 2471–2482, https://doi.org/10.1175/JTECH-D-19-0114.1.

Tapiador, F. J., A. Navarro, R. Moreno, J. L. Sanchez, and E. Garcia-Ortega, 2020: Regional climate models: 30 years of dynamical downscaling. *Atmos. Res.*, **235**, 104785, https://doi.org/10.1016/j.atmosres.2019.104785.

Thompson, E. J., J. N. Moum, C. W. Fairall, and S. A. Rutledge, 2019: Wind limits on rain layers and diurnal warm layers. *J. Geophys. Res. Oceans*, **124**, 897–924, https://doi.org/10.1029/2018JC014130.

——, H. Chen, J. Yang, and J. L. Bytheway, 2023: High-resolution precipitation and wind measurements from passive aquatic listeners (PALs). NASA Global Hydrometeorology Resource Center DAAC, accessed 19 April 2023, https://doi.org/10.5067/GPMGV/PAL/DATA101.

Trenberth, K. E., L. Smith, T. Qian, A. Dai, and J. Fasullo, 2007: Estimates of the global water budget and its annual cycle using observational and model data. *J. Hydrometeor.*, **8**, 758–769, https://doi.org/10.1175/JHM600.1.

Trivej, P., and B. Stevens, 2010: The echo size distribution of precipitating shallow cumuli. *J. Atmos. Sci.*, **67**, 788–804, https://doi.org/10.1175/2009JAS3178.1.

Wilby, R. L., L. E. Hay, and G. H. Leavesley, 1999: A comparison of downscaled and raw GCM output: Implications for climate

change scenarios in the San Juan River basin, Colorado. *J. Hydrol.*, **225**, 67–91, https://doi.org/10.1016/S0022-1694(99)00136-5.

Wu, Q., and Y. Wang, 2019: Comparison of oceanic multisatellite precipitation data from Tropical Rainfall Measurement Mission and Global Precipitation Measurement Mission datasets with rain gauge data from ocean buoys. *J. Atmos. Oceanic Technol.*, **36**, 903–920, https://doi.org/10.1175/JTECH-D-18-0152.1.

Xu, Z., Y. Han, and Z. Yang, 2019: Dynamical downscaling of regional climate: A review of methods and limitations. *Sci. China Earth Sci.*, **62**, 365–375, https://doi.org/10.1007/s11430-018-9261-5.

Yang, J., S. C. Riser, J. A. Nystuen, W. E. Asher, and A. T. Jessup, 2015: Regional rainfall measurements using the passive aquatic listener during the SPURS field campaign. *Oceanography*, **28** (1), 124–133, https://doi.org/10.5670/oceanog.2015.10.

——, J. A. Jystuen, S. C. Riser, and E. I. Thorsos, 2023: Open ocean ambient noise data in the frequency band of 100 Hz–50 kHz from the Pacific Ocean. *JASA Express Lett.*, **3**, 036001, https://doi.org/10.1121/10.0017349.