

# Deep Neural Network High Spatiotemporal Resolution Precipitation Estimation (Deep-STEP) Using Passive Microwave and Infrared Data

VESTA AFZALI GOROOH,<sup>a</sup> ATA AKBARI ASANJAN,<sup>b,c</sup> PHU NGUYEN,<sup>a</sup> KUOLIN HSU,<sup>a</sup> AND SOROOSH SOROOSHIAN<sup>a,d</sup>

<sup>a</sup> Center for Hydrometeorology and Remote Sensing, The Henry Samueli School of Engineering, Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, California

<sup>b</sup> Universities Space Research Association, Mountain View, California

<sup>c</sup> NASA Ames Research Center, Moffett Field, California

<sup>d</sup> Department of Earth System Science, University of California, Irvine, Irvine, California

(Manuscript received 8 October 2021, in final form 10 February 2022)

**ABSTRACT:** Recent developments in “headline-making” deep neural networks (DNNs), specifically convolutional neural networks (CNNs), along with advancements in computational power, open great opportunities to integrate massive amounts of real-time observations to characterize spatiotemporal structures of surface precipitation. This study aims to develop a CNN algorithm, named Deep Neural Network High Spatiotemporal Resolution Precipitation Estimation (Deep-STEP), that ingests direct satellite passive microwave (PMW) brightness temperatures (T<sub>b</sub>) at emission and scattering frequencies combined with infrared (IR) T<sub>b</sub> from geostationary satellites and surface information to automatically extract geospatial features related to the precipitable clouds. These features allow the end-to-end Deep-STEP algorithm to instantaneously map surface precipitation intensities with a spatial resolution of 4 km. The main advantages of Deep-STEP, as compared to current state-of-the-art techniques, are 1) it learns and estimates complex precipitation systems directly from raw measurements in near-real time, 2) it uses the automatic spatial neighborhood feature extraction approach, and 3) it fuses coarse-resolution PMW footprints with IR images to reliably retrieve surface precipitation at a high spatial resolution. We anticipate our proposed DNN algorithm to be a starting point for more sophisticated and efficient precipitation retrieval systems in terms of accuracy, fine spatial pattern detection skills, and computational costs.


**KEYWORDS:** Precipitation; Satellite observations; Neural networks; Deep learning

## 1. Introduction

Clouds and precipitation are key components of Earth’s hydrological cycle, yet there is a lack of deep understanding of their physics and dynamics. Reliable cloud and precipitation measurements are useful for understanding the current state of the global water resource, as well as for determining the characteristics of these processes over time and space (Funk et al. 2015; Behrangi et al. 2011; Boucher et al. 2013). In recent decades, the availability of meteorological satellite technologies has provided a remedy for the shortcomings of ground-based observations. These technologies have also brought a sustained global-scale source of information that promotes high-resolution cloud detection and precipitation estimation more efficiently and accurately (Sun and Tang 2020; Houze 2014). Satellite precipitation retrieval schemes offer a diverse range of applications from near-real-time high-resolution estimates for flood warning systems and short-term weather predictions, to long-term climate data for the monitoring of global trends (Sorooshian et al. 2000; Damberg and AghaKouchak 2014; Nguyen et al. 2016).

At the time of writing, the spaceborne atmospheric radars, active microwave (AMW), have only been carried on a few

low-Earth-orbiting (LEO) satellites. Precipitation and cloud radar technologies pioneered by the Precipitation Radar (PR) on board the Tropical Rainfall Measuring Mission (TRMM) satellite, and the Cloud Profiling Radar (CPR) on the *CloudSat* satellite and further advanced by Dual-Frequency Precipitation Radar (DPR) on the Global Precipitation Measurement (GPM) mission *Core Observatory* (GPM-*CO*) satellite. These AMW sensors provide high-quality and precise cloud and precipitation features but are limited to narrow swath widths. These observations are useful for climatological and event-based investigations, and they can be used for calibration and validation in the near-real-time cloud and precipitation retrieval algorithms (Kidd and Levizzani 2011; Tang et al. 2017; Gou et al. 2018). There are a variety of operational passive radiometers on geostationary/geosynchronous Earth orbit (GEO), mostly cover the visible (VIS) and infrared (IR) portion of the electromagnetic spectrum, flying on the international constellation of meteorological satellites such as Advanced Baseline Imager (ABI) on the U.S. Geostationary Operational Environmental Satellite (GOES)-R series. High spatiotemporal and spectral resolution data streams from GEO-based passive sensors, with 0.5–4-km spatial resolution and 5–30-min global refresh rates, meet the requirements for near-real-time monitoring of cloud distributions and precipitation processes. However, these measurements are mainly reliant on cloud-top properties with limited information on the vertical structure of the atmosphere, the precipitating cores of the clouds, and the underneath precipitation-generating mechanisms. The current operational LEO-based passive microwave (PMW) radiometers

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

Corresponding author: Vesta Afzali GorooH, vafzalig@uci.edu

DOI: 10.1175/JHM-D-21-0194.1

© 2022 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) ([www.ametsoc.org/PUBSReuseLicenses](#)).

such as the international network of sensors in the GPM mission (Hou et al. 2014) measure the upwelling radiation of the Earth's surface (brightness temperatures;  $T_b$ ) in emission ( $<37$  GHz) and scattering channels ( $>37$  GHz) (Kummerow 2020). Although PMW imagers and sounders provide more direct measurements for instantaneous precipitation estimation compared to VIS/IR readings, their temporal resolutions are far coarser than GEO-based radiometers. The existing GEO and LEO satellite constellations are discussed in detail by Kidd et al. (2021).

There are several space–time gridded satellite-based surface precipitation products that are widely used by scientific communities and operational engineering applications. Although different research groups have been attempting to develop an optimum algorithm for accurate and global-scale precipitation estimation, the task remains elusive mostly because of the trade-off between sampling frequency of VIS/IR, PMW, and AMW sources of information. Depending on the final user, the representation of precipitation phenomena also varies; hydrologists mostly require the gridded dataset for the intensity and occurrence of “surface” precipitation, but atmospheric scientists are more interested in characteristics of vertical profile values (Kidd et al. 2018). Several precipitation retrieval algorithms primarily deploy VIS/IR measurements as their main input such as the widely used Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS; Hong et al. 2004) and PERSIANN Dynamic Infrared–Rain Rate (PDIR; Nguyen et al. 2020). These algorithms predefine the cloud-top features based on coldness, geometry, and texture, differentiating various cloud types using an unsupervised classification technique in order to link surface precipitation rates with IR brightness temperatures ( $T_b$ s). Although IR-based algorithms provide high-resolution spatiotemporal data about precipitation on a quasi-global scale, and in near–real time, their performance is limited to indirect relations between the tops of the cloud  $T_b$ s (IR measurements) and surface precipitation rates. There also exists some precipitation retrievals that utilize PMW reading as their dominant input, such as the NASA's Goddard profiling algorithm (GPROF; Kummerow et al. 2015). The GPROF algorithm retrieves both PMW-based instantaneous surface precipitation and vertical rainfall structures, using a dictionary-based Bayesian inversion method. An a priori database is generated which contains several million sets of PMW  $T_b$ s along with associated surface rainfall and hydrometeor profiles, as well as surface types, 2-m temperatures, and total precipitable water (TPW). This database is subcategorized by temperature, TPW, and 14 main surface type classes, defined over oceans and various types of land surfaces. A conditional probabilistic Bayesian scheme retrieves precipitation data by comparing satellite observations to radiances in the (appropriate) database. In general, PMW-based precipitation retrievals are more accurate than IR-based estimates, due to their reliance on direct relations between PMW  $T_b$ s and vertical properties of atmosphere and hydrometeors. However, these products are sparse and less frequent when compared to IR retrievals. Also, precipitation retrievals from spaceborne PMW rely on point-wise low-resolution measurements without considering neighborhood

properties. The NASA Integrated Multi-satellitE Retrievals for the GPM mission (IMERG) is one of the well-known precipitation algorithms, utilizing intercalibrated PMW and PMW-calibrated IR precipitation estimates with multiple ancillary data to provide consistent global estimates in early, late, and final runs (Huffman et al. 2020). The IMERG relies on a combination of previously established satellite precipitation algorithms. The orbital PMW precipitation retrievals from the GPM constellation are mostly processed by the GPROF algorithm. Then, they are merged with recalibrated PERSIANN-CCS estimates using the Climate Prediction Center (CPC) morphing–Kalman filter (CMORPH-KF; Joyce and Xie 2011) quasi-Lagrangian time interpolation scheme. The IMERG scheme requires a large number of variables, high computational power, and data storage to derive half-hourly,  $0.1^\circ$  spatial resolution surface precipitation products (Skofronick-Jackson et al. 2018).

There is a growing interest in using data-driven machine learning techniques in a variety of studies such as hydrologic modeling (Hsu et al. 2002; Yu et al. 2006; Raghavendra and Deka 2014; Yaseen et al. 2015; Kratzert et al. 2019; Ardabili et al. 2020; Hu et al. 2019; Jha and Sahoo 2015; Afzaal et al. 2019), climate modeling (O'Gorman and Dwyer 2018; Krasnopolsky and Fox-Rabinovitz 2006; Rasp et al. 2018), cloud detection and classification (Xie et al. 2017; Cai and Wang 2018; Jeppesen et al. 2019; Goroooh et al. 2020), precipitation estimation and forecasting (Pan et al. 2019; Miao et al. 2019; Xiao and Chandrasekar 1997; Behrangi et al. 2009; Kuligowski and Barros 1998; Hsu et al. 1997; Akbari Asanjan et al. 2018), and many more. Among the multitude of machine learning methods, deep neural networks (DNNs) have provided the flexibility necessary to analyze massive remotely sensed datasets. Furthermore, DNNs have offered exciting opportunities to lessen the limitations of pure, process-oriented physical models, and to address the relationship between nontrivial, lagged long-distance, high-dimensional, and multiscale geophysical variables (Reichstein et al. 2019). While neural network–based approaches for precipitation estimation have come a long way since early efforts at the turn of the twenty-first century, understanding and making ultimate use of multisource remote sensing data is ongoing.

One of the first attempts at deploying the neural networks for precipitation estimation was the study conducted by Sanò et al. (2016) using a combination of LEO-based PMW with multiple auxiliary data as inputs to a two-layer pixel-wise neural network. Tang et al. (2018) also presented a neural network to establish relationships between precipitation data from merged spaceborne radars, PMW, IR, and environmental data, for high-latitude regions. Chen et al. (2020) designed a deep multilayer perceptron system to produce the rainfall information, using IR data and PMW-based retrievals as inputs, for an urban-scale area over the Dallas–Fort Worth metroplex. Pfreundschuh et al. (2018) introduced quantile regression neural networks (QRNNs) to estimate quantiles of the posterior distribution for PMW Bayesian retrieval algorithms, such as GPROF. QRNNs mitigates the requirement of a lookup in a large retrieval database, and it offers a simple integration of ancillary data into the surface precipitation retrieval algorithms. Although these studies have

shown satisfactory performance using multiple sources of information for precipitation retrieval, the fully connected pixel-wise structure of their algorithms establish relations between output pixels (precipitation) and input pixels (satellite remotely sensed information), which are insufficient for observing variations in spatial neighborhoods. The importance of using nonlocal approaches in satellite precipitation retrieval has been discussed in detail by [Foufoula-Georgiou et al. \(2020\)](#). [Tao et al. \(2018\)](#) utilized the Stacked Denoising Autoencoder (SDAE) to extract cloud features from GEO-based IR imagery, and ([Sadeghi et al. 2019](#)) introduced convolutional neural network (CNN) algorithms for deriving the surface precipitation from bispectral GEO satellite imagery. [Hayatbini et al. \(2019\)](#) also introduced a data-driven framework for near-real-time precipitation estimation using conditional generative adversarial networks (cGANs) and CNNs for multispectral information from GEO imagers. These studies and many more ([Z. Li et al. 2021](#); [Moraux et al. 2019](#)) use state-of-the-art deep learning (DL) techniques to automatically extract precipitation related features, but they are limited to one type of remotely sensed dataset as input to their models (e.g., solely IR or PMW or radar measurements). CNN architectures are a quickly growing branch of DNN, characterized by their ability to automatically extract spatial features from raw image data, as opposed to traditional machine learning approaches that rely on “hand-crafted” features. CNNs use convolutional transformation and learnable filters that allow sharing weights across the spatial data. This is a unique advantage of CNNs that in conjunction with pooling operations increases the efficiency of neural networks by reducing the number of hyperparameters in the model. CNNs learn local neighborhood features as well as a large field of view in images without requiring a huge amount of training data. Based on this literature review, no study was found that explores advanced and powerful CNNs using raw spaceborne PMW and IR information (multisensor multisatellite measurements) as inputs to derive high-resolution surface rainfall estimates over a large area.

Our research reported in this manuscript is designed to explore whether “end-to-end” DNNs are capable of mitigating the aforementioned common weaknesses in satellite precipitation products (e.g., imperfect data integration between PMW and IR data) and the need for a large number of physical and environmental variables in process-oriented precipitation retrieval algorithms. It should be noted that end-to-end learning in DNNs refers to training a complex learning system by one model that bypasses the intermediate multistages usually present in traditional pipeline designs. The specific objectives of this study are 1) to explore the potential of CNNs to develop an end-to-end high spatial resolution near-real-time precipitation estimation framework using spatial features related to precipitation from PMW measurements combined with IR imageries and 2) to demonstrate the capability of the developed DNN multisensor multisatellite algorithm compared to well-known operational process-oriented products such as IMERG and GPROF. In this paper, we are focusing on developing a CNN-based satellite precipitation retrieval algorithm using GPM Microwave Imager (GMI) measurements combined with IR Tbs over the United States, as a proof of concept.

## 2. Data sources

In this research several datasets are utilized for input, calibration, benchmarking, and comparison of our DNN model to operational products for precipitation estimation purposes. These products come from a variety of sources, including passive sensors on LEO and GEO satellite platforms, ground-based radars each with its strengths and weaknesses. The emphasis will be on the use of a combination of multiple satellite-based observational systems, and the primary regional focus of the experiments will be over the eastern contiguous United States (CONUS) and some parts of North Atlantic Ocean near coastal regions (110°–60°W and 20°–50°N). Data sources used for this study are detailed in the following subsections.

### a. PMW information

In our study, the analysis relies on Tbs derived only from GMI (L1CGMI version 5), as the proof of concept, which contains two types of swaths ([Draper et al. 2015](#); [Hou et al. 2014](#)) with slightly different conical scanning geometries, swath S1 has a width of approximately 931 km and swath S2 has a width of approximately 825 km. The distance between along-track scans is about 13.5 km and along-scan pixel separation for low and high frequencies are 5.7 and 5.1 km, respectively. [Table 1](#) shows spatial resolution specifications for GPM-CO GMI, and we can see low-frequency channels have larger fields of view compared to high-frequency channels. Due to the coarse resolution and channel availability of 10 GHz in the current GPM constellation radiometers (frequencies less than 18-GHz bands are only available from AMSR2 and GMI sensors) bands 1 and 2 are excluded from our experiments (for details, see [Kidd et al. 2021](#); [Skofronick-Jackson et al. 2017](#)).

### b. IR images

The National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) merged longwave IR product at the half-hourly temporal and 4-km spatial resolution is used for precipitation estimation purposes ([Janowiak et al. 2001](#)). The CPC dataset is a combination of IR Tb (~10.7- $\mu\text{m}$  central wavelengths) from several international GEO-based satellites (e.g., *GMS-5*, *GOES-8*, *Meteosat-7*).

### c. Surface types

The values of Tool to Estimate Land Surface Emissivities at Microwave Frequencies (TELSEM) surface class index is used as ancillary inputs to our CNN algorithms. The surface types are derived from monthly emissivity climatology and are currently used in version 5 of the NASA Precipitation Measurement Mission (PMM) GPROF algorithm ([Prigent et al. 2003](#); [Aires et al. 2011](#)). In the GPROF surface classification scheme the surface types are numbered as ocean/large inland water (1), sea ice (2), decreasing vegetation covered (3–7), decreasing snow covered (8–11), inland water/rivers/estuaries (12), coastlines (13), and ocean/sea ice boundary (14). For detailed specifications of the surface classes used GPROF algorithm see [NASA \(2018\)](#). In our investigations, to allow the easier interpretation of the performance of the

TABLE 1. Characteristics of GMI PMW bands. Bands are identified by their frequency (GHz) and their polarization (V is vertical polarization, H is horizontal polarization).

No. band	Swath S1					Swath S2							
	1	2	3	4	5	6	7	8	9	10	11	12	13
Central frequency (GHz)	10.6 V	10.6 H	18.7V	18.7H	23.8V	36.6V	36.6H	89V	89H	166V	166H	183.3 ± 3V	183.3 ± 7V
Spatial resolution (km) <sup>a</sup>	19.4 × 32.1	19.4 × 32.1	10.9 × 18.1	10.9 × 18.1	9.7 × 16	9.4 × 15.6	9.4 × 15.6	4.4 × 7.2	4.4 × 7.2	4.1 × 6.3	4.1 × 6.3	3.8 × 5.8	3.8 × 5.8

<sup>a</sup> Represents the instantaneous fields of view for the mean orbit altitude of 407 km.

proposed model, the surface types are reformed in the way that sea ice land and snow surface classes along with the samples with low confidence (quality flags equal to 2 or 3) are excluded from our investigations.

*d. Precipitation products*

The National Severe Storms Laboratory (NSSL) Multi-Radar Multi-Sensor (MRMS) is one of the most reliable and high spatiotemporal resolution precipitation datasets over the CONUS that is widely used in literature as reference data (Zhang et al. 2016; Kirstetter et al. 2012, 2015). The WSR-88 Doppler radar data are integrated with more than 7000 gauge observations, atmospheric environmental data, and satellite data to generate MRMS quantitative precipitation products. The half-hourly gauge-adjusted MRMS precipitation product and the 2-min GPM Ground Validation MRMS (GV-MRMS) precipitation rates with 0.01° spatial resolution, provided by NASA Global Hydrology Resource Center (GHRC) Marshall Distributed Active Archive Center (DAAC), are used as the benchmarks for training and validation of our proposed models. In our preliminary experiments, high-quality GV-MRMS samples that are not affected by terrain blockage (Maddox et al. 2002) are selected. Standard PMW-derived GPROF V05 surface precipitation estimates from GMI sensor (2AGPROFGMI, GPROF hereafter; Kummerow et al. 2015) and IMERG V06 Early Run (3IMERGHH; Huffman et al. 2020) product (IMERG hereafter) are used for independent assessments. Undoubtedly, using Level 2 (swath-level) GPROF and Level 3 IMERG products will cause uncertainty in the comparison results, but these two products are selected as the most suited and well-known operational products in our investigations.

Different datasets in this investigation have different spatial resolutions and all are rescaled into gridded precipitation products through bilinear interpolation while the nearest-neighbor approach is used for resampling conical scanning GMI footprints. To maintain consistency, we use the bilinear interpolation method for upscaling and downscaling. It should be noted that, while potential errors from bilinear interpolations in the upscaling procedure could occur, their effect on our analysis is insignificant and does not impact our overall assumption.

**3. Methodology**

*a. Data integration strategy for precipitation estimation*

The training data for our proposed DNN algorithms consist of geospatial information, surface types, IR Tbs from geostationary satellites, as well as PMW Tbs from GMI on the GPM-CO satellite, all of which are collocated with GV-MRMS observations. Our data preprocessing pipeline is explained in four steps as below:

- 1) Extraction of GMI orbital tracks over the CONUS. PMW Tbs, scan times (UTC), and granular scans from GMI orbital tracks (L1CGMI files) over the study area are extracted. Of these scans, the relevant information which we utilize are geospatial locations, footprint latitude and longitude, the number of swath scans in the granule, and

- the swath pixels per scan. These data are fetched from NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) data archive on NASA's Earth Observing System Data and Information System (EOSDIS).
- 2) Collocation of GMI orbital tracks with the closest IR and GV-MRMS data. The intersection of high temporal resolution GV-MRMS readings (2 min) with PMW footprints within the time window of orbital tracks is found. The closest IR image (30-min temporal resolution) is also matched to the overpass of GMI scans to generate a stack of the IR and PMW measurements and their corresponding GV-MRMS observations.
  - 3) Orbital swaths to gridded data conversion and resampling. Gridded PMW Tb datasets are retrieved from orbital/swath information with the nearest neighbor approach (swath to grid). A data generation pipeline is applied that converts the native GMI swath data (stored in HDF5) to raster data (GeoTIFFs) using  $k$ -dimensional tree ( $k$ -d tree; Bentley 1990) for fast nearest-neighbor search over a large area (110°–60°W and 20°–50°N). For this purpose, “Pyresample” and “GDAL” packages are implemented using Python. All the generated gridded datasets including IR and GV-MRMS precipitation rates are resampled to 4-km gridded data in this step.
  - 4) Data patch generation for training DNN models. The collocated GV-MRMS precipitation rates, GMI Tbs, and IR Tbs samples (in our study for 2017 and 2018) are organized into  $2.5^\circ \times 2.5^\circ$  data patches (64 pixels  $\times$  64 pixels). These patches of data are normalized and partitioned into the training, validation, and test periods. No-rain pixels are identified using a  $0.1 \text{ mm h}^{-1}$  threshold. To mitigate the potential for imbalanced distributions of patches and to ensure sufficient samples are supplied as inputs (rain events are minority events), rainy data patches (i.e., at least one pixel of  $2.5^\circ \times 2.5^\circ$  data patches have a rain rate of more than  $0.1 \text{ mm h}^{-1}$ ) and nonrainy patches (i.e., there is not any rainy pixel  $2.5^\circ \times 2.5^\circ$  data patches) are represented uniformly in our training dataset (about 79 400 data patches). It should be noted that our DNN algorithms, due to their convolutional and pooling layers, are flexible enough to process inputs of variable sizes. Therefore, this step is used for training the CNNs from scratch, whereafter they can accept inputs of any size to estimate the precipitation rates.

Extreme memory resource utilization is one of the biggest challenges in training DNNs with a large number of datasets. The train-on-batch and generally on-the-fly learning techniques are utilized for the efficient and fast training, mainly implemented by Keras high-level libraries and the TensorFlow platform. Furthermore, this training scheme can be used to update a pretrained DNN model when new sets of data are available. In our investigations, datasets are iteratively generated in parallel by multiple CPUs and then directly fed to the GPU to train the DNN algorithms from scratch.

#### b. The architecture of precipitation estimation algorithm

CNN algorithms usually consist of a cascade of convolutional layers that perform convolutional operations on the

data from the previous layers (Long et al. 2015). U-Net-like models are popular types of deep CNNs that consist of a pair of encoders and decoders for end-to-end learning and show an impressive potential for segmentation and regression problems, even with a few numbers of training data (Ronneberger et al. 2015; Goodfellow et al. 2016). In classical deep CNN algorithms, a fully connected layer is usually designed at the output layer, but the U-Net algorithms consist of fully convolutional blocks; the encoder operates on the inputs to extract fine-scale spatial features (downsampling step) then the decoder constructs the high-level output patterns and pixel values (upsampling step and regression). A convolutional layer is designed to use the feature maps and complete an image-to-image regression process in the last layer. Each encoder consists of a filter bank with learnable parameters (e.g., kernels and weights), then batch normalization and elementwise rectified linear unit (ReLU) activation function  $[\max(0, x)]$  are used to generate the feature maps (Nair and Hinton 2010). The kernel's receptive field's size plays a key role in capturing the neighborhood information, especially when the spatial resolution is high. Batch normalization removes the internal covariate shifts and trivial variance of data batch distributions in the layers. Following that, max-pooling is applied to reduce the dimensions and achieve translation invariance over small spatial shifts in the layer's input (maximum element in a window). After multiple downsampling in the encoder part of the algorithm, robust features are captured, but there is a loss of spatial resolution in data/image especially in complex boundaries and gradients. So, it is important to keep the detailed features before each downsampling and append them in the upsampling step of the network. In the decoder, the feature maps are convolved with decoder filter banks, followed by batch normalization, ReLU activation function, and upsampling layers.

The unique property of the U-Net-like models is the skip connections that append the fine-scale features from the encoder blocks to the high-level feature maps in decoder blocks that help the model to learn more efficiently and converge much faster in the training stage. The skip connections also alleviate the vanishing gradient challenge in neural networks (Ehsani et al. 2021). The connections in the U-Net model copy and concatenate the original feature from each encoder block to the corresponding decoder block that is effective to avoid “loss of resolution” initiated by subsampling parts. Furthermore, U-Net-like architecture as a fully convolutional algorithm is flexible in input size. In our study, the DNNs use  $2.5^\circ \times 2.5^\circ$  data patches and are compatible with any input size which is very helpful for regional- and global-scale analysis (Akbari Asanjan et al. 2018). Recently, for precipitation estimation, some extensions from the CNN architecture have been proposed, with the main difference that they are in the skip connections. Sadeghi et al. (2020) show the superior performance of a U-Net architecture against classical CNN architectures without shortcuts for near-real-time IR-based precipitation estimation. Sun and Tang (2020) also recently investigated the performance of an attention-based network, AU-Net, for learning spatial and temporal mappings from coarse-resolution to fine-resolution precipitation products and proved the potential of

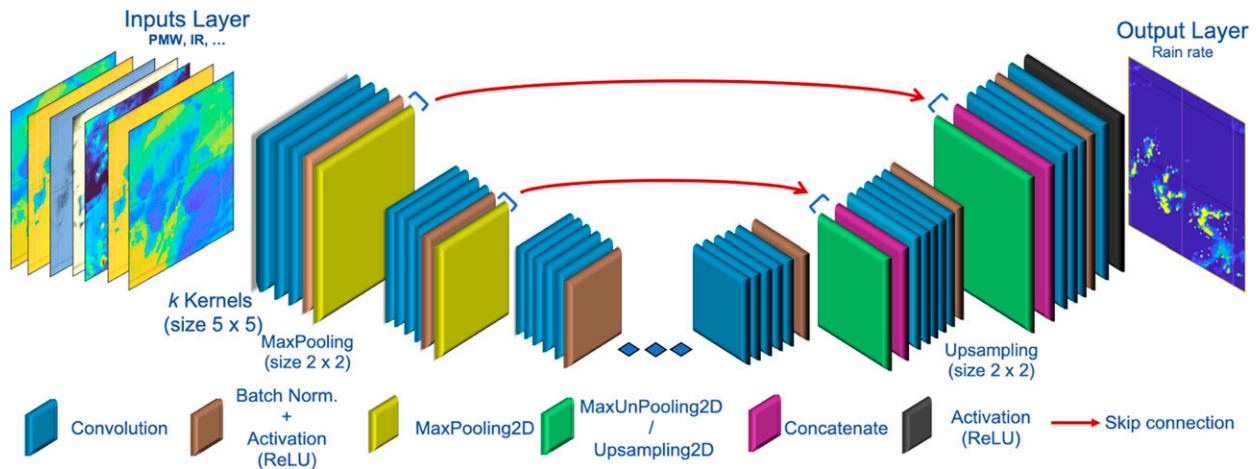


FIG. 1. Illustration of convolutional neural networks, U-Net-like architecture, the encoder operates on the input data to extract robust features (at left), then the decoder constructs the output patterns and pixel values (at right).

modifying skip connections in U-Net architecture for down-scaling satellite precipitation. Figure 1 shows the schematic architecture of the CNN models in this study. In our preliminary analysis, more than 200 DNNs with different U-Net-like network structures and different input scenarios were tuned, and we report the six top well-performed DNNs (Fig. 2), as the proof of concept, in section 4. Since emission channels are sensitive to moderate/heavy precipitation rates and scattering channels are responsive to light precipitation rates over land (J. Li et al. 2022), all GMI's channels, except channels at 10.6V/H GHz, are included in our experiments.

In our study, the Tensorflow (2.2.0) platform and an adaptive learning rate method are used and employ the first and second moments of the gradient to adapt the learning rate for each weight of the neural network as its optimizer (Adam; adaptive moment estimation). The cost function used in training the CNN model is the mean square error (MSE), the size of 2D kernels is  $5 \times 5$ ; the max-pooling and upsampling are performed with  $2 \times 2$  windows with the same padding option (output has the same dimension as the input). Data preprocessing, training, and parameter tuning of the model are carried on Google Colab and an NVIDIA Quadro P6000 GPU. Training time depends on the model size, batch size, and the total number of epochs along with the availability of CPUs for the custom data generator. After sensitivity analysis and optimization of hyperparameters, 0.001, 32, and 75 are selected for learning rate, batch size, and the number of epochs, respectively. The experiments are carried out for 2017 and 2018. July 2017, November 2018, and Hurricane Harvey (23–30 August 2017) were excluded for an independent test

purpose and then 80% and 20% of the dataset are utilized for training and validation of the model, respectively. The reason for selecting months of July and November is the GPM GVMRMS data availability (as the reference dataset provided by NASA GHRC) with high-quality index values and to include relatively warm and cold precipitation patterns in test periods.

### c. Performance measurements

To assess the performance of satellite precipitation estimation algorithms, the common categorical evaluation metrics including probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) are used which are defined as

$$\text{POD} = \frac{H}{H + M}, \quad (1)$$

$$\text{FAR} = \frac{F}{F + H}, \quad (2)$$

$$\text{CSI} = \frac{H}{H + F + M}, \quad (3)$$

where  $H$  (hit) indicates that both the precipitation model and reference observation detect the event,  $M$  (miss) identifies events captured by reference data but missed by the developed model, and  $F$  (false alarm) indicates events captured by reference but not confirmed by the model. Following the concept of categorical indices, we define volumetric hit index (VHit), volumetric false alarm ratio (VFAR), and volumetric critical success index (VCSI) as follows:

$$\text{VHit} = \frac{\sum_{i=1}^n [P_i | (P_i > \text{thr} \ \& \ \text{MRMS}_i > \text{thr})]}{\sum_{i=1}^n [P_i | (P_i > \text{thr} \ \& \ \text{MRMS}_i > \text{thr})] + \sum_{i=1}^n [\text{MRMS}_i | (P_i \leq \text{thr} \ \& \ \text{MRMS}_i > \text{thr})]}, \quad (4)$$

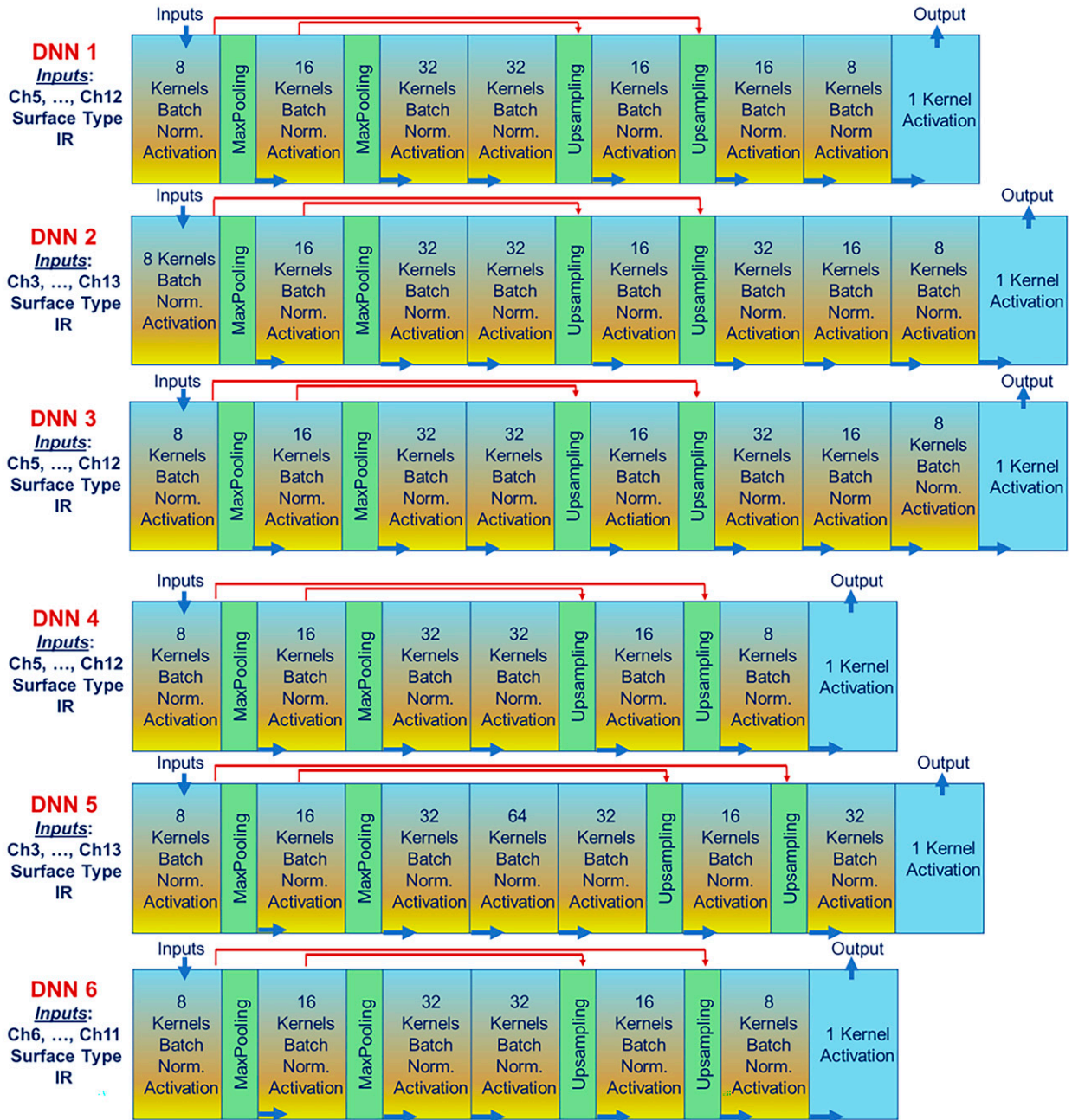


FIG. 2. Details of proposed DNN models. Models' names (DNN 1–DNN 6) with their input layers are located on the left-hand side of each panel. The red arrays indicate the skip connections in U-Net architecture.

$$VFAR = \frac{\sum_{i=1}^n [P_i | (P_i > thr \ \& \ MRMS_i \leq thr)]}{\sum_{i=1}^n [P_i | (P_i > thr \ \& \ MRMS_i > thr)] + \sum_{i=1}^n [P_i | (P_i > thr \ \& \ MRMS_i \leq thr)]}, \tag{5}$$

$$VCSI = \frac{\sum_{i=1}^n [P_i | (P_i > thr \ \& \ MRMS_i > thr)]}{\sum_{i=1}^n [P_i | (P_i > thr \ \& \ MRMS_i > thr)] + \sum_{i=1}^n [P_i | (MRMS_i \leq thr \ \& \ MRMS_i > thr)] + \sum_{i=1}^n [P_i | (P_i < thr \ \& \ MRMS_i \leq thr)]}, \tag{6}$$

where  $P_i$  represents  $i$ th satellite precipitation estimates,  $MRMS_i$  refers to corresponding reference observations, and  $n$  is the total number of data. The “thr” is the threshold above which the VHI, VFAR, and VCSI are calculated. These categorical metrics are explained by [AghaKouchak and Mehran \(2013\)](#). Continuous verification metrics such as Pearson correlation coefficient (CORR), multiplicative bias (BIAS), and root-mean-square error (RMSE) are also used to evaluate the performance precipitation estimates:

$$\text{CORR} = \frac{1}{n} \frac{\sum_{i=1}^n (P_i - \bar{P})(MRMS_i - \overline{MRMS})}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2} \sqrt{\sum_{i=1}^n (MRMS_i - \overline{MRMS})^2}}, \quad (7)$$

$$\text{BIAS} = \frac{1}{n} \sum_{i=1}^n \frac{P_i}{MRMS_i}, \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - MRMS_i)^2}. \quad (9)$$

#### d. Permutation feature importance

After selecting the best DNN algorithm (described in [section 3b](#)), we name our proposed model Deep Neural Network High Spatiotemporal Resolution Precipitation Estimation (Deep-STEP). In the last part of our investigations, we examine which model’s inputs have the biggest impact on the performance of the proposed precipitation estimation model. The inspection approach is based on the study by [Breiman \(2001\)](#) can be summarized in four main steps:

- 1) We consider the mean square error (MSE) score from the Deep-STEP model as the metric of importance on the testing and validation dataset (i.e., *benchmark MSE*).
- 2) In turn, each predictor is shuffled without changing other predictors (inputs) and target precipitation rates.
- 3) Then, the altered datasets are given to the Deep-STEP and the MSE scores are retrieved (i.e., *permuted MSE*).
- 4) The difference between the benchmark score and the permuted score for each altered input set is calculated.

The highest score difference is associated with the input set with corruption applied to the most important predictor. Thus, the highest score difference highlights which input contributes the most to the Deep-STEP algorithm. To rank the importance of inputs, the percentage of a performance change is defined as

$$\text{Performance change (\%)} = \frac{\text{benchmark MSE} - \text{permuted MSE}}{\text{benchmark MSE}} \times 100. \quad (10)$$

## 4. Results and discussions

The proposed algorithms in our investigations are inspired by U-Net-like architectures originally from medical image segmentation studies that have been modified for pre-

cipitation estimation application (regression problem). We tuned various combinations of hyperparameters (number of kernels, number of hidden layers, etc.) with different input datasets and learning variables to select the final algorithm of our set of architectures. After sensitivity analysis and optimization of hyperparameters, we selected 0.001 and 32 for learning rate and batch size, respectively. The performances of the six DNN precipitation models (named DNN 1–DNN 6) are reported for data covering the month of July 2017, as the proof of concept. [Figure 2](#) indicates the designed models while [Fig. 3](#) shows their corresponding performances at half-hourly temporal scales with reference to GV-MRMS data. These plots demonstrate the common statistical metrics for precipitation rates more than  $0.1 \text{ mm h}^{-1}$  over the study area. IMERG and GPROF products show 0.36 and 0.49 correlation coefficients with the GV-MRMS precipitation observations, respectively, whereas all our proposed DNNs show higher correlation values (more than 0.53). GPROF has the highest rate in the probability of detection (0.82) followed by DNN 2 and DNN 1 models. In terms of FAR, DNN 2 and DNN 6 achieve the best performance among all the models, 0.45 and 0.41, respectively. In terms of CSI, which can be considered an overall detection skill, again DNN 2 and DNN 6 architectures perform better compared to others, despite slight overestimation in the total amount of precipitation by 23% and 19%, respectively. The results for volumetric metrics demonstrate that 91% of the volume of observed precipitation is correctly detected by DNN 2, similar to the GPROF algorithm with VHit rates of 92%. VFAR shows that false precipitation rates in the proposed models are below 17% of the total volume of rainfall with respect to GV-MRMS reference data. DNN 2 is also superior to all algorithms in terms of VCSI (0.79) that provides the general volumetric performance of models associated with the volume of hit, false, and missed components of capturing precipitation. It should be noted that several experiments (input combinations, architectures, etc.) were assessed in the test period to choose the best model that reveals the best agreement with reference data over our study area. Statistics are just reported for July 2017 and a comparison between various DNN structures and different input combinations reveal that DNN 2 consistently exhibited better performance throughout the whole test period. DNN 2, hereafter referred to as Deep-STEP, utilizes PMW Tbs (from 19V/H to  $183 \pm 3\text{V}$  and  $183 \pm 8\text{V}$  GHz GMI channels), IR Tbs ( $\sim 10.7 \mu\text{m}$ ), and surface type information as inputs of the model.

[Table 2](#) demonstrates the precipitation estimation statistics of our best DNN model (Deep-STEP) benchmarked against MRMS and shows a valuable comparison between IMERG, GPROF, and Deep-STEP algorithms for July 2017 and November 2018. A threshold of  $0.1 \text{ mm h}^{-1}$  is used to distinguish between rainy and nonrainy events. In July 2017 (warm month), Deep-STEP shows the best performance compared to IMERG in both categorical and volumetric indices, while GPROF slightly outperforms Deep-STEP in terms of POD, BIAS, and RMSE by 0.11, 0.21, and  $1 \text{ mm h}^{-1}$ , respectively. Overall, Deep-STEP, which employs an end-to-end CNN, performs substantially better compared to well-known operational algorithms in the warm testing period. The evaluations



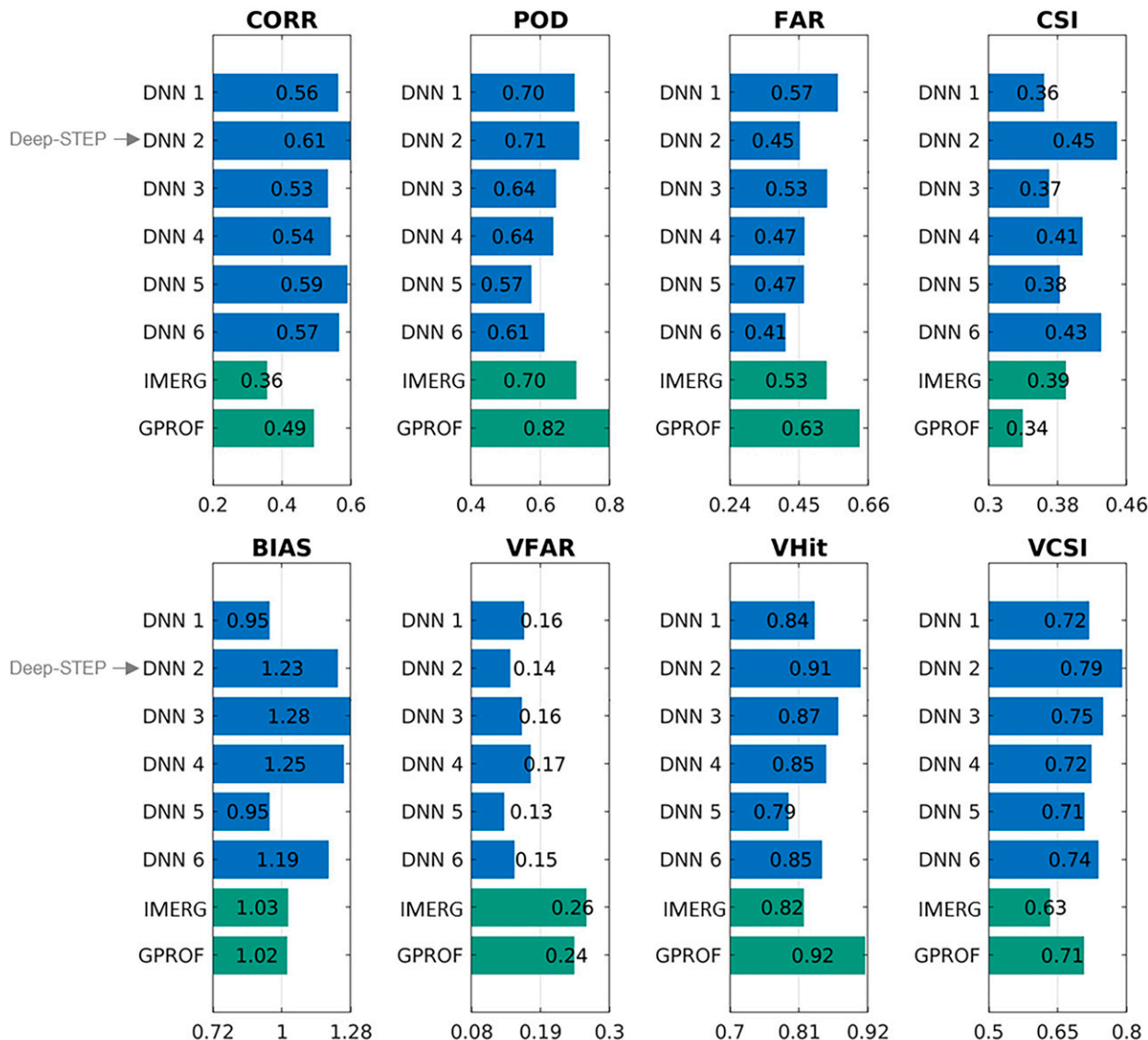


FIG. 3. Statistical indices for precipitation models (named DNN 1–DNN 6) for choosing the best DNN model (Deep-STEP).

conducted over November 2018, show that retrieving precipitation is challenging for satellite algorithms in the cold testing period (November 2018). However, Deep-STEP retrievals show a high correlation with the reference data (CORR = 0.71) along with relatively low BIAS and low RMSE, 0.93 and 0.80, respectively. Concurrently, IMERG’s CORR of 0.48, BIAS of 1.20, and RMSE of 1.23 mm h<sup>-1</sup> show that this model experiences more severe overestimations compared to GPROF’s CORR, BIAS, and RMSE values of 0.53, 1.08, and 1.05 mm h<sup>-1</sup>, respectively. Table 2 manifests that there are negligible differences in categorical detection skills of precipitation algorithms, while Deep-STEP is superior to other products in general volumetric indices with VCSI and VFAR values of 0.72, and 0.11, respectively.

Despite the usefulness of the general evaluations presented in the previous section, assessing precipitation retrievals for different surface classes at high temporal resolution shows the improvements obtained by our developed model (Deep-STEP) with respect to operational algorithms (IMERG and GPROF).

*a. Performance of precipitation algorithms over different surface types*

Figures 4 and 5 show scatterplots of each surface precipitation product (y axis) versus GV-MRMS observations (x axis) for July 2017 and November 2018, respectively. These plots demonstrate the precipitation rates between the satellite-based estimates and the reference data for precipitation rates

TABLE 2. Statistical performance of the retrievals in July 2017 and November 2018: correlation (CORR), root-mean-square error (RMSE), multiplicative bias (BIAS), probability of detection (POD), false alarm ratio (FAR), critical success index (CSI), volumetric hit index (VHit), volumetric false alarm ratio (VFAR), and volumetric critical success index (VCSI). Best performance is highlighted in bold font in each test period.

Metrics/model	July 2017			November 2018		
	Deep-STEP	GPROF	IMERG	Deep-STEP	GPROF	IMERG
CORR	<b>0.61</b>	0.49	0.36	<b>0.71</b>	0.53	0.48
POD	0.71	<b>0.82</b>	0.70	0.51	<b>0.68</b>	0.55
FAR	<b>0.45</b>	0.63	0.53	0.46	0.55	<b>0.41</b>
CSI	<b>0.45</b>	0.34	0.39	0.35	0.36	<b>0.39</b>
BIAS	1.23	<b>1.02</b>	1.03	<b>0.93</b>	1.08	1.20
RMSE (mm h <sup>-1</sup> )	1.20	<b>1.19</b>	1.38	<b>0.80</b>	1.05	1.23
VCSI	<b>0.79</b>	0.71	0.63	<b>0.72</b>	0.71	0.70
VFAR	<b>0.14</b>	0.24	0.26	<b>0.11</b>	0.19	0.17
VHit	0.91	<b>0.92</b>	0.82	0.78	<b>0.86</b>	0.82

more than 0.1 mm h<sup>-1</sup> over the study area. It should be mentioned that the GPROF algorithm uses GPM DPR products in its a priori database over all surface types (14 surface classes mentioned in section 2), except for snow and ice-covered areas, where GV-MRMS snowfall retrievals are used. In our investigations, the snow and ice surface types are not considered in training the Deep-STEP model. Thus, the precipitation retrievals with assigned snow and ice surface conditions (GPM surface classifications used in GPROF products) and their corresponding (time and space matched) precipitation estimates from GPROF, IMERG, and Deep-STEP were removed from our assessments.

Figure 4a shows that Deep-STEP and GPROF are 0.55 correlated to the GV-MRMS observations over the ocean and large inland water bodies, while IMERG shows less correlation with the value 0.36 for this surface class. In terms of multiplicative bias, the precipitation algorithms perform similarly over oceans and large inland water bodies with values ranging from 1.04 to 1.09 in the warm month (July 2017). We can see that the average BIAS values significantly increase for IMERG and GPROF retrievals in the cold month (November 2018) over this class. Over inland water bodies/rivers/wetland category, Deep-STEP outperforms IMERG and GPROF products in the warm month (Fig. 4b) and the Deep-STEP's performance is also notable compared to other products during the cold month (Fig. 5b). Figures 4c and 5c clearly demonstrate that Deep-STEP is skillful in capturing precipitation over complex coastal regions for both warm and cold months. Since coastal surfaces are not homogenous nor captured well with the coarse resolution PMW footprints (mixed land and water pixels), IMERG and GPROF do not perform skillfully over these regions. It should be noted that the scatterplots of GPROF estimates versus GV-MRMS product show bimodal features over land and water boundaries surface type. However, such distribution is not vivid for IMERG and Deep-STEP retrievals. A similar bimodal distribution is also reported by Utsumi et al. (2020) and You et al. (2020). In our investigations, the overlap of IR images and various PMW channels enhances both the spatial resolution and the accuracy of precipitation retrievals (Petty and Bennartz 2017).

It is worth mentioning that precipitating clouds over large water bodies in low-frequency PMW channels (<37 GHz) show radiometric "warming" parts (i.e., a higher equivalent blackbody Tb, compared to the precipitation-free background) against the radiometrically "cold" water background. High-frequency channels (>37 GHz) respond to the amount of precipitation-sized hydrometeors (depth and density) through a radiometric "cooling" (i.e., lower Tbs with respect to the background) owing to the scattering of the upwelling emission from the rain. Therefore, by using low- and high-frequency PMW channels with high-resolution IR images, our model has performed very well over water bodies. Over our study region, the areas of densely vegetated land sharply increase eastward (from 95° to 60°W longitude), and in this part of the assessments, we regrouped the four decreasing vegetation covered surface types (types 3–6) into vegetated land (types 5 and 6) and arid land (types 3 and 4) groups. As shown in the scatterplots in Figs. 4d and 4e, Deep-STEP overestimates the precipitation rates with CORR values of 0.64 and 0.67 over vegetated and arid land classes, respectively. At the same time, GPROF (with lower CORR values than Deep-STEP) slightly underestimates and overestimates the precipitation over vegetated and arid surfaces, respectively. In November 2018, precipitation retrievals over vegetated and arid land surface types (Figs. 5d,e) are generally more correlated to observations compared to July 2017. Specifically, CORR values increase by 0.11, 0.15, and 0.20 for Deep-STEP, GPROF, and IMERG retrievals over vegetated land class, respectively. By examining the precipitation scatterplots over arid surface type for November 2018, one can see that Deep-STEP and IMERG slightly underestimate light rainfall events while they overestimate the moderate and heavy precipitation rates (>1 mm h<sup>-1</sup>). GPROF shows a relatively poor performance (mostly underestimation) over arid areas with the general multiplicative bias of 0.88 that was expected due to the low variations in radiometric Tbs over land.

The performance of high spatiotemporal resolution satellite precipitation algorithms in detecting spatial patterns are important in disaster management, flood warning systems,

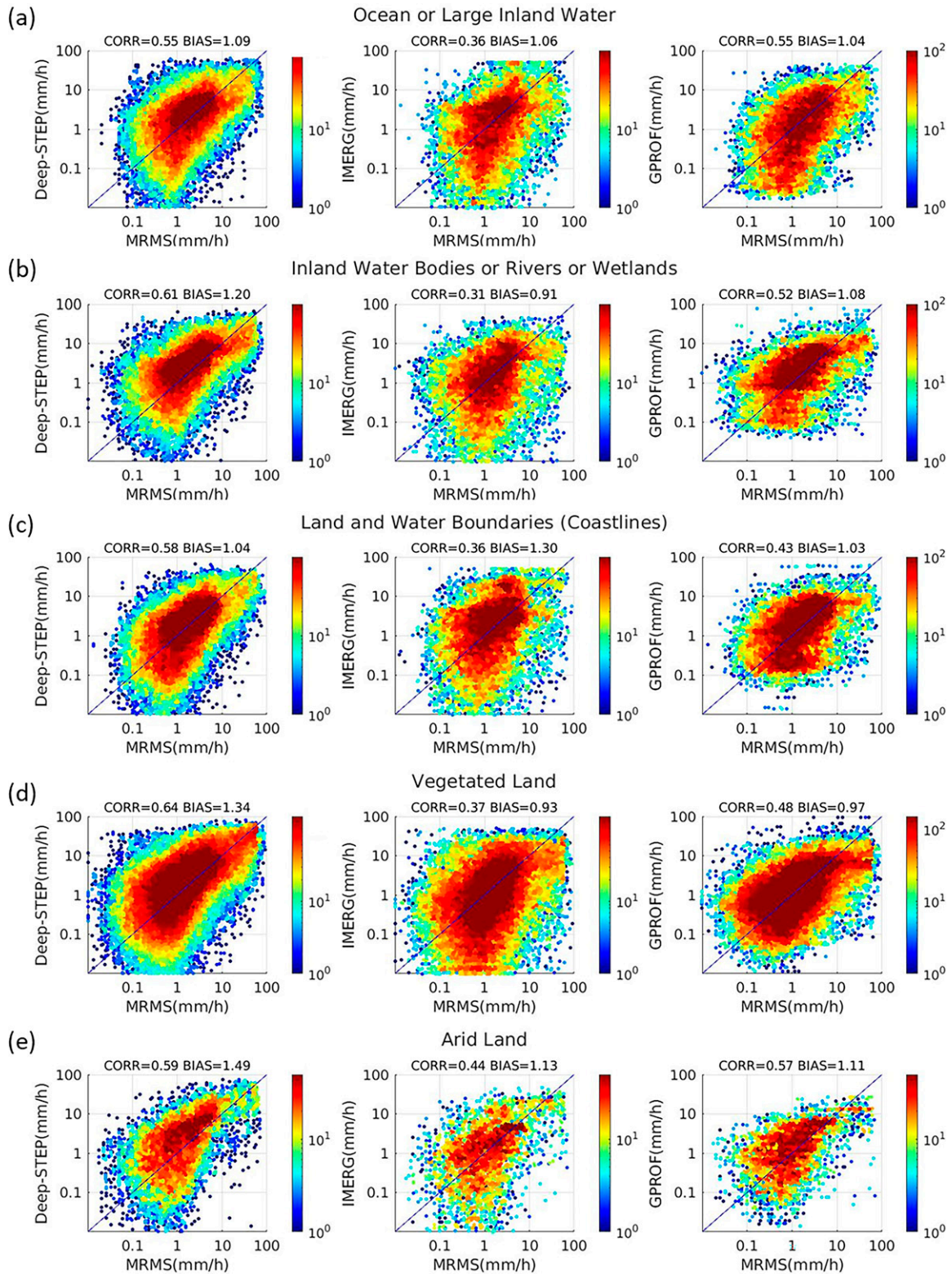


FIG. 4. Scatterplot of surface precipitation retrievals for Deep-STEP, IMERG, and GPROF vs GV-MRMS precipitation rates over (a) ocean/large inland water bodies (surface type 1), (b) inland water bodies/rivers/wetlands (surface type 12), (c) coastlines and land/water boundaries (surface type 13), (d) vegetated land, and (e) arid land (decreasing vegetation covered: surface type 3 = Amazon-like, surface type 7 = Sahara Desert-like), July 2017. The color bar displays the number of cases.

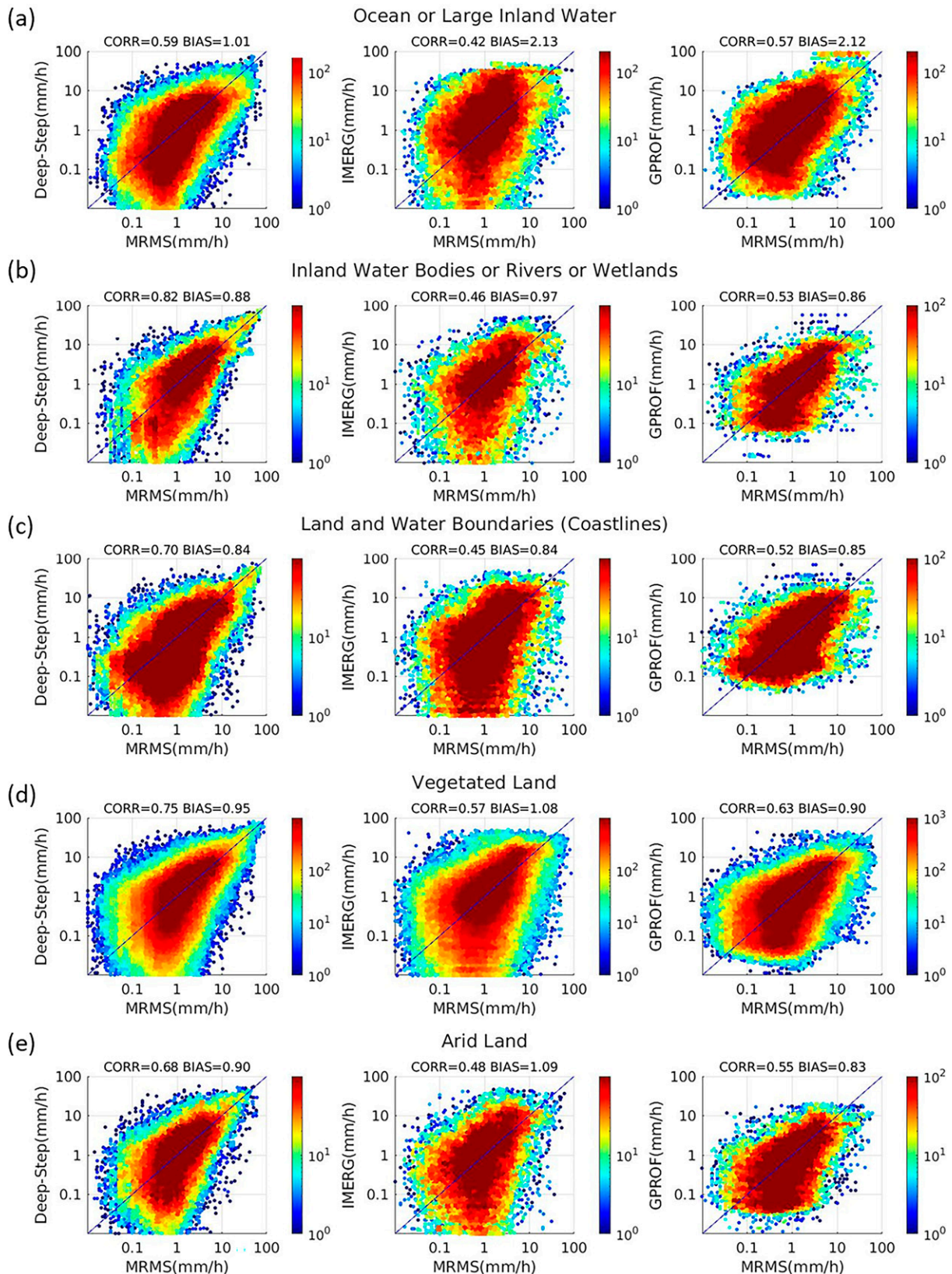


FIG. 5. As in Fig 4, but for November 2018.

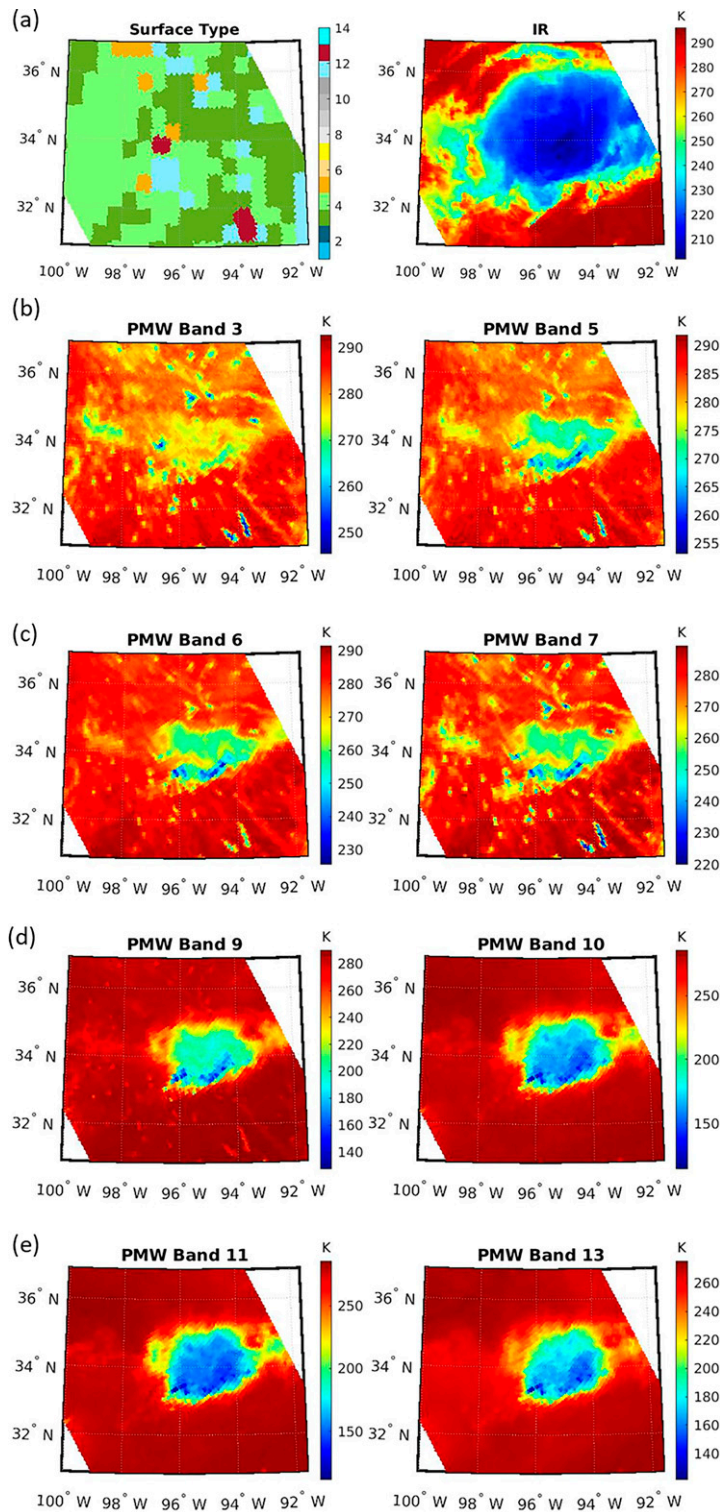


FIG. 6. Case study (I) over Oklahoma, northern Texas, and eastern Arkansas and Louisiana at 1240 UTC 5 Jul 2017: (a) surface type and infrared (IR) information, (b) 19V GHz (PMW band 3) and 23V GHz (PMW band 5), (c) 37V GHz (PMW band 6) and 37H GHz (PMW band 7), (d) 89H GHz (PMW band 9) and 166V GHz (PMW band 10), and (e) 166H GHz (PMW band 11) and 183H GHz (PMW band 13) channels from GMI.

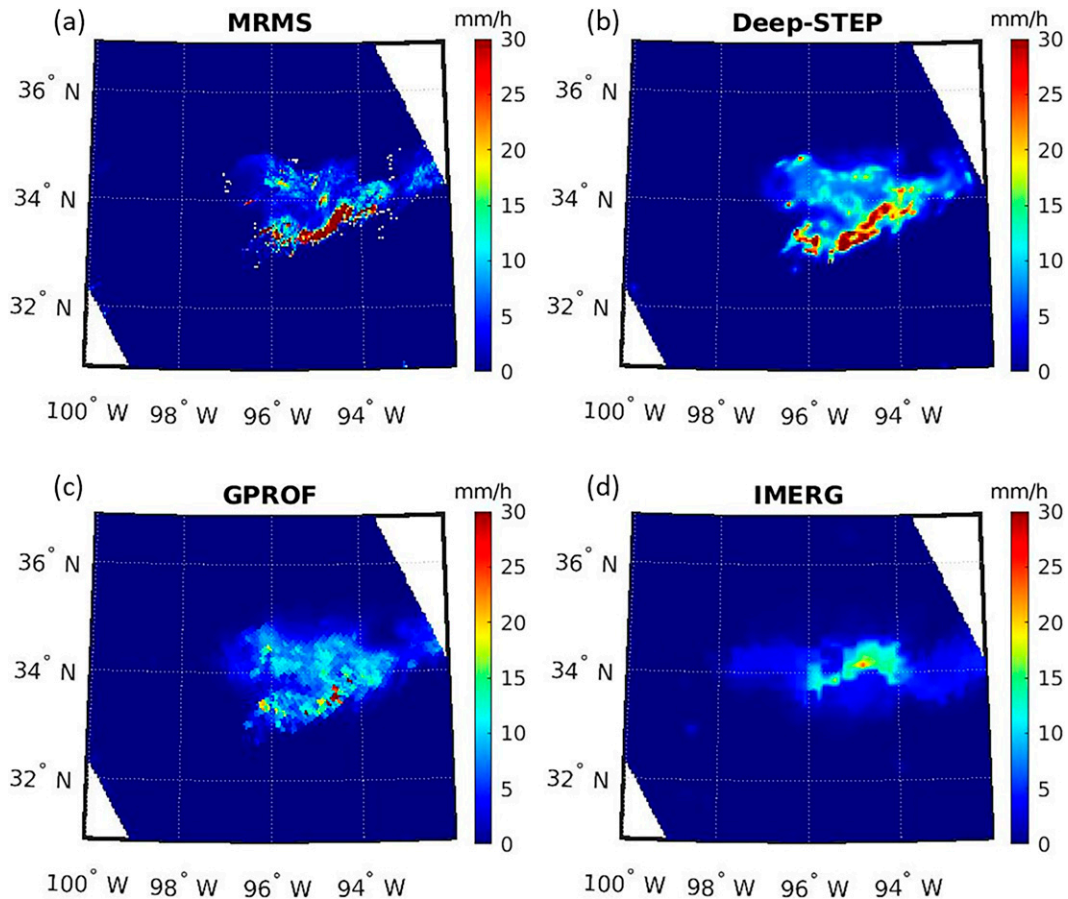


FIG. 7. Surface precipitation retrievals for case study (I) from (a) GV-MRMS, (b) Deep-STEP, and (c) GPROF, and (d) IMERG, from 1230 to 1300 UTC 5 Jul 2017, over Oklahoma, northern Texas, and eastern parts of Arkansas and Louisiana.

short-term weather predictions, and many more applications. With the aforementioned assessments over different surface types, we now examine two extreme precipitation events over our study area.

#### b. Visualization of precipitation retrievals over event I

Figure 6 presents a visualization of input information for the Deep-STEP model during an extreme precipitation event on 5 July 2017. This event was captured by the GPM GMI radiometer while occurring over Oklahoma, northern parts of Texas, and eastern parts of Arkansas and Louisiana states. Figure 6a shows the surface types (retrieved from GPM surface classification scheme) and IR image ( $\sim 10 \mu\text{m}$ ) from GOES ABI. The surface types over this area are mainly vegetated land (surface type 3 and 4 from GPM surface classification scheme) and low cloud-top temperatures (emitted radiances) with cloud boundaries can be seen in the longwave IR image. The blue color in the IR image shows colder cloud tops and the possibility of rainfall, but it cannot exactly illustrate the location of precipitation beneath the clouds. Figures 6b and 6c illustrate the 19V, 23V, 37V, and 37H GHz equivalent blackbody Tbs

over this area. We can see that over land due to the radiometrically warm background, the low frequencies alone are not sufficient to detect precipitating clouds. However, high-frequency PMW measurements more effectively discriminate between the presence of rainy regions and land surface without precipitation (Figs. 6d,e). Figure 7 displays a visual comparison between precipitation estimates from Deep-STEP, GPROF, and IMERG algorithms and GV-MRMS reference data. Deep-STEP provides a realistic representation of the spatial patterns of surface precipitation rates. Meanwhile, similar patterns are vaguely observed by GPROF due to limitations of coarse resolution PMW footprints. IMERG misses the entire spatial pattern of this event and underestimates the heavy precipitation rates over  $25 \text{ mm h}^{-1}$ . IMERG also does not perform well in deriving the spatial patterns and the amount of lighter rain rates ( $< 25 \text{ mm h}^{-1}$ ) compared to other algorithms. Deep-STEP proves its ability to derive precipitation rates over land by fusing IR images, emission, and scattering PMW frequencies. Deep-STEP offers the capability of data fusion and automatic neighborhood feature extraction (with CNNs) to enhance the spatial resolution of PMW-based precipitation retrievals over land.

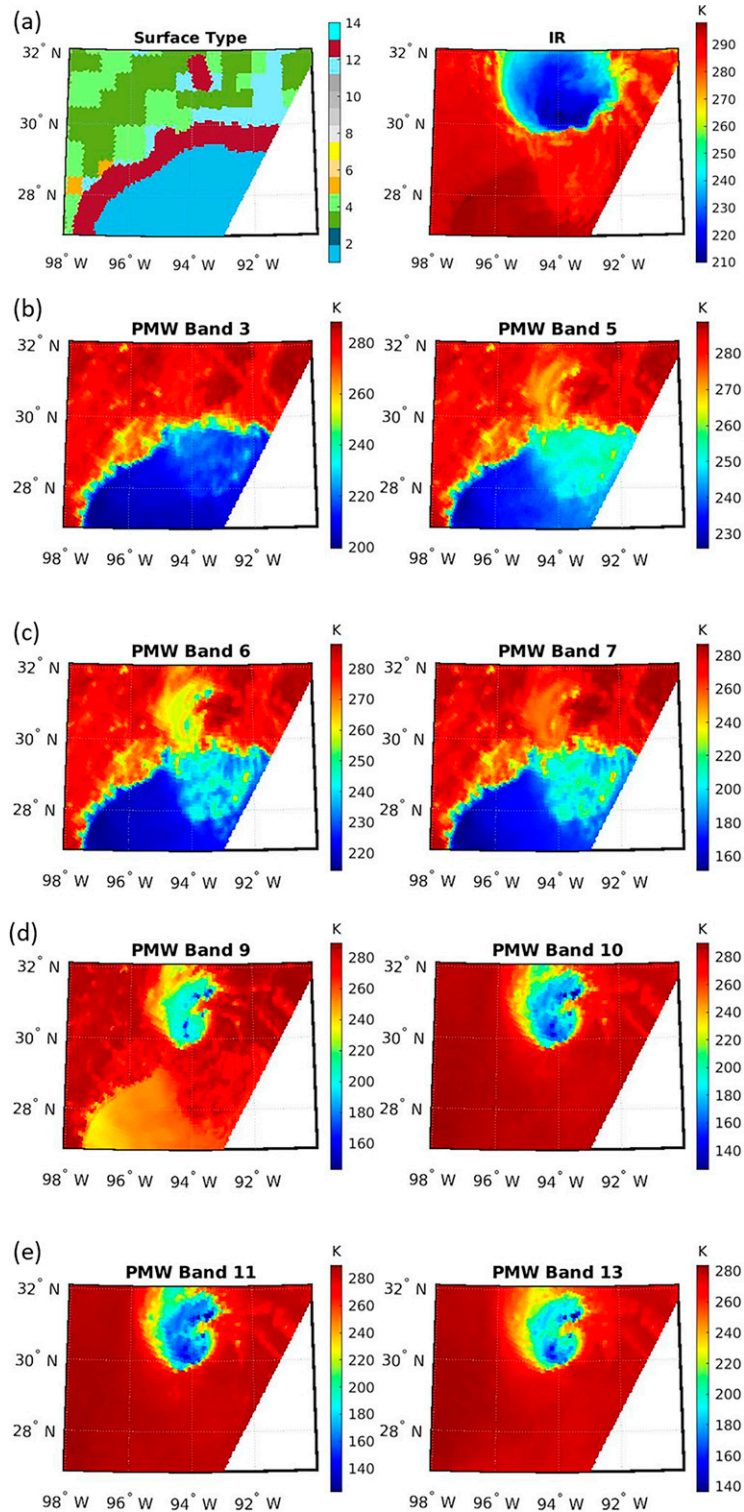


FIG. 8. Case study (II) Hurricane Harvey over the northern Gulf States at 1046 UTC 30 Aug 2017: (a) surface type and infrared (IR) information, (b) 19V GHz (PMW band 3) and 23V GHz (PMW band 5), (c) 37V GHz (PMW band 6) and 37H GHz (PMW band 7), (d) 89H GHz (PMW band 9) and 166V GHz (PMW band 10), and (e) 166H GHz (PMW band 11) and 183H GHz (PMW band 13) channels from GMI.

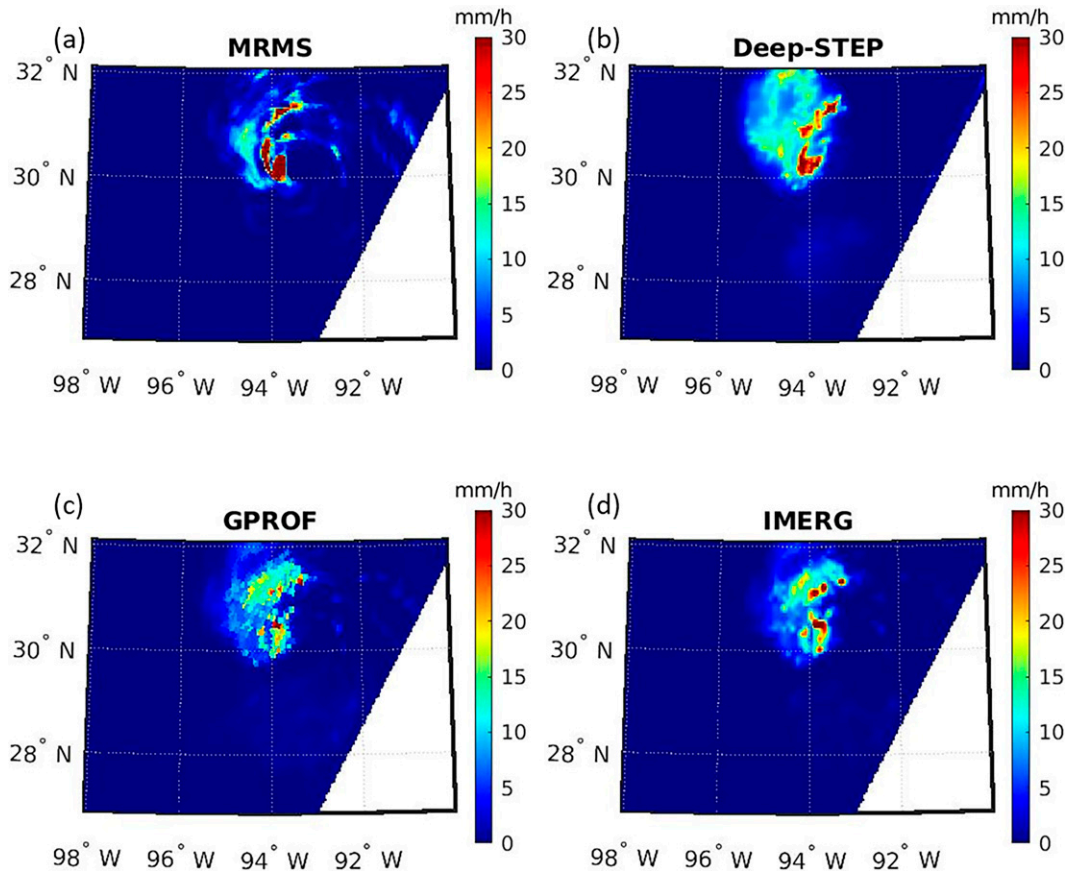


FIG. 9. Surface precipitation retrievals for case study (I) from (a) GV-MRMS, (b) Deep-STEP, (c) GPROF, and (d) IMERG, from 1030 to 1100 UTC 30 Aug 2017, over Hurricane Harvey.

### c. Visualization of precipitation retrievals over event II

Figure 8 displays samples for IR Tbs, PMW Tbs from 19V to 166H GHz channels along with surface type information as inputs of the Deep-STEP over Hurricane Harvey. Hurricane Harvey stalled over Texas on 26 August and moved toward the east until its landfalls on 30 August 2017. GPM-CO satellite overpassed this event over the coastline (near Cameron, Louisiana) on 30 August at 1357 UTC. The IR Tbs clearly captures the cold cloud tops in this convective system and the IR channel allows us to observe rainy clouds directly. GMI PMW observations vividly illustrate the difference between the radiometrically warm land and cold ocean in the low-frequency channels (Figs. 8b,c). The low-frequency PMW channels are sensitive to small- and medium-sized droplets while the presence of hydrometeors and precipitating parts of the clouds (low Tbs) can be more clearly observed in GMI high-frequency channels (Figs. 8d,e) over land and coastal areas. We can see that from emission frequencies ( $37 \text{ GHz} >$ ) channels, energy leaving from the land surface is about the same magnitude as from precipitating parts of the clouds, whereas the scattering frequencies ( $37 \text{ GHz} <$ ) more clearly distinguish precipitation signatures from land. By comparison, the precipitation signals from clouds are distinctly revealed over the

ocean in both emission and scattering channels. Figure 9 presents the spatial extent of Deep-STEP precipitation retrievals over this event compared to the reference GV-MRMS dataset, IMERG, and GPROF products. Deep-STEP adequately captures the spatial patterns of high surface precipitation rates ( $>25 \text{ mm h}^{-1}$ ), but it slightly overestimates the precipitation rates less than  $10 \text{ mm h}^{-1}$  over northern parts of the hurricane. GPROF and IMERG do not detect spatial variability of precipitation, especially over the center of the hurricane. Clearly, GPROF and IMERG's coarse resolution retrievals limit their abilities to recognize the fine spatial details of rainfall over the coastal surface types, while Deep-STEP takes advantage of fusing IR images with PMW information to retrieve precipitation in high spatial resolution.

Last, we would like to mention that after obtaining the coincident precipitation estimations from Level 3 IMERG Early Run and standard Level 2 GPROF, we can see that IMERG retrievals are different from GPROF in both visual and statistical comparisons. Considering that near-real-time (NRT) GPROF is one of the merged algorithms in the IMERG framework, the uncertainty of IMERG product at early run is very high at half-hourly temporal resolution (Huffman et al. 2020). In general, the discrepancies between IMERG and GPROF retrievals during a half-hour window



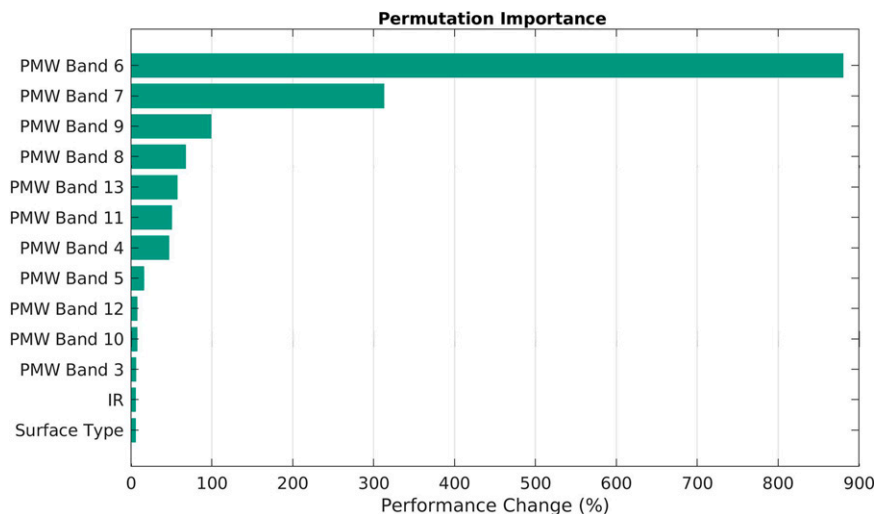


FIG. 10. Permutation importance for Deep-STEP. Performance change indicates the dependency of the Deep-STEP on the corresponding input.

might be caused by different PMW sensors overpasses at one spatial location, gridding, intercalibration, and morphing the PMW data in the IMERG system.

#### d. Permutation feature importance

After evaluating the performance of the Deep-STEP model, we investigate the importance of inputs to the model. As described in section 3d, our approach is to break the relationship between each input and the target precipitation to determine how much the model is reliant on the input. When an input is shuffled, the MSE score increases (bad performance of the model) which results in a large difference between permuted MSE value and the benchmark score. Larger gap between the original model MSE and altered input MSE indicates the high dependency of the model on the corresponding shuffled input. The benchmark MSE for our study period is  $1.57 \text{ mm}^2 \text{ h}^{-2}$  and by corrupting the structure of each input, we observe a boost in the MSE values. Figure 10 shows the permuted MSE as the percentage variation from the benchmark MSE. Among the inputs, the PMW 36.6V/H GHz channels, as emission frequency bands, followed by 89V/H GHz channels, as scattering frequency bands, are the most important inputs for the Deep-STEP model with more than 65% change in the performance. The results indicate that the model is also dependent on PMW Tbs from  $183.3 \pm 7\text{V}$ , and 166H GHz, and 18.7H GHz bands with about 57%, 50%, and 47% performance change, respectively.

It should be noted that in our investigations the inputs are intercorrelated. Thus, when a single input (highly correlated features) is shuffled, the Deep-STEP extrapolate to unknown regions of the feature space, and the model might overemphasize the performance variations (e.g., PMW band 6 and 7 in Fig. 10). This part of the study is an example for DNN interpretations and there exist advanced permutation importance approaches to get a better understanding of physical connections and interpretability of models (Hooker et al. 2021).

Since the future direction of our investigation is the globe implementation of the Deep-STEP model by using PMW sensors in the GPM constellation, a broader analysis will be needed for important inputs selection.

## 5. Conclusions and future directions

The present study explores the application of CNN architecture to ingest coarse spatial resolution PMW footprints with IR Tbs and surface type for high spatial resolution surface precipitation estimation. Emission and scattering channels from the GMI sensor on board the GPM-CO satellite are directly integrated with GEO-based IR images, and of our set of architectures, we select the best CNN algorithm (referred to as Deep-STEP). Deep-STEP is trained with GPM GV-MRMS surface precipitation rates. Our proposed end-to-end Deep-STEP model is trained using advanced on-the-fly learning techniques that mitigates the significant computational requirements. The CNN model architecture automatically extracts cloud features and translates them into surface precipitation maps at 4-km spatial resolution. The proposed model is investigated over the eastern CONUS and its precipitation retrievals are compared with the well-known GPROF and IMERG products. We evaluated the performance of these algorithms based on continuous verification metrics (e.g., correlation coefficient, multiplicative bias), categorical and volumetric indices (e.g., CSI, FAR, VCSI, VHit). It should be noted that our model is trained and tested with respect to GV-MRMS surface precipitation rates as reference. Thus, Deep-STEP is at an advantage compared to the GPROF and IMERG products that are trained with GPM AMW retrievals but evaluated against the GV-MRMS product.

The results are assessed over different surface types, and the permutation feature importance of our proposed

precipitation estimation model is also examined. The main findings are summarized as follow:

- The proposed Deep-STEP model shows a promising opportunity for developing an end-to-end multisensor multisatellite surface precipitation retrievals algorithm. According to our experiments over half-hourly temporal scales, retrieved precipitation from the proposed Deep-STEP algorithm shows higher correlation values with the reference precipitation product compared to GPROF and IMERG models. Specifically, Deep-STEP precipitation estimates during cold months are more highly correlated with reference precipitation rates compared to warm months.
- Based on our analysis, Deep-STEP outperforms GPROF and IMERG in terms of overall verification metrics. Though Deep-STEP slightly overestimates the total amount of precipitation during warmer test periods, it marginally underestimates precipitation rates in colder test periods. In general, the Deep-STEP model shows significant improvements due to its leveraging of two different sources of passive remotely sensed data (PMW and IR spectrums).
- Our comprehensive evaluations show that the Deep-STEP algorithm is superior to GPROF and IMERG algorithms in capturing high spatiotemporal resolution precipitation rates over complex surface types such as coastal regions while it slightly overestimates light and moderate precipitation rates over arid and vegetated surfaces.
- Using low spatial resolution PMW footprints with many auxiliary variables limits the performance of process-oriented satellite precipitation algorithms in capturing patterns of precipitation. The performance of our proposed Deep-STEP model over extreme precipitation events highlights the promises of extracting relevant fine-scale precipitation features by fusing coarse-resolution PMW information with IR images.
- Last, the results from the permutation feature importance analysis support the idea that the main source of information for retrieving precipitation rates in our proposed framework comes from PMW sensors. Our investigations show that PMW channels at 36.6V/H and 89V/H GHz frequencies are the most important inputs to the Deep-STEP algorithm, and by incorporating IR images they collectively enhance the spatial detection of precipitation.

Overall, we show the effectiveness of fusing a set of Tbs from coarse-resolution PMW footprints with IR images and using an end-to-end algorithm with an automatic neighborhood feature extraction approach to capture the fine spatial patterns of precipitation events. However, a broader statistical analysis for spatial and temporal variations is needed to draw more comprehensive conclusions. We anticipate our CNN-based model to be a starting point for more sophisticated and efficient high spatial resolution precipitation retrieval systems, both in terms of accuracy and in computational cost. It is important to acknowledge the fact that while Deep-STEP offers modest performance improvements over the state-of-the-art highly specialized operational algorithms, these consistent improvements should be considered from the

perspective of their relative computational cost-effectiveness. Our model offers improved performance at a computational cost that is many orders of magnitude lower than operational satellite precipitation algorithms, and this translates to a highly effective, highly accessible system that is capable of achieving better or even similar results. The future extension of the Deep-STEP will be mainly training and validating our proposed scheme using the GPM DPR precipitation product as a reference dataset on a global scale. The Deep-STEP should be applied on GPM PMW sensors, both imagers and sounders, to alleviate the limited spatiotemporal data coverage from LEO satellites and to provide global-scale high spatiotemporal resolution precipitation retrievals (e.g., 4-km spatial and 3-hourly temporal resolution). One of the challenges for global implementation will be adapting Deep-STEP to each radiometer in the GPM constellation, due to variation of PMW channels, incidence angles, spatial and temporal resolutions, and many more. Addressing this challenge requires further investigation. Moreover, a detailed analysis on the importance of PMW spectral windows (similar to our experiments in section 4d) from the revolutionized constellation of small radiometers (Stephens et al. 2020), which mostly operate at high PMW frequencies could lead to interesting studies in the future.

*Acknowledgments.* This research is supported by Future Investigators in NASA Earth and Space Science and Technology (NNH19ZDA001N-FINESST), NASA Grant 80NSSC19K0726, MASEEH fellowship, U.S. Department of Energy (DOE Prime Award DE-IA0000018), California Energy Commission (CEC Award 300-15-005), NOAA/NESDIS/NCDC (Prime Award NA 09NES4400006 and NCSU CICS and Subaward 2009-1380-01), National Oceanic and Atmospheric Administration (ST133017 CQ0058) with Riverside Technology, Inc., NVIDIA Academic Hardware Grant Program (donation of Quadra P6000 GPU), and University of California (4600010378 TO15 Am 22). Assistance provided by the University of California, Irvine (UCI) office of vice-chancellor for research of graduate students, UCI Graduate and Postdoctoral Scholar Resource Center, and Mr. Andrew Gregory Hansen is greatly appreciated.

*Data availability statement.* The datasets used in this study are publicly available through Goddard Earth Sciences Data and Information Services Center (GES DISC; <https://disc.gsfc.nasa.gov/>) and NASA Global Hydrology Resource Center (GHRC) Marshall Distributed Active Archive Center (DAAC; [https://ghrc.nsstc.nasa.gov/hydro/#/?\\_k=x77cch](https://ghrc.nsstc.nasa.gov/hydro/#/?_k=x77cch)).

## REFERENCES

- Afzaal, H., A. A. Farooque, F. Abbas, B. Acharya, and T. Esau, 2019: Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning. *Water*, **12**, 5, <https://doi.org/10.3390/w12010005>.
- AghaKouchak, A., and A. Mehran, 2013: Extended contingency table: Performance metrics for satellite observations and climate model simulations. *Water Resour. Res.*, **49**, 7144–7149, <https://doi.org/10.1002/wrcr.20498>.

- Aires, F., C. Prigent, F. Bernardo, C. Jiménez, R. Saunders, and P. Brunel, 2011: A Tool to Estimate Land-Surface Emissivities at Microwave frequencies (TELSEM) for use in numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **137**, 690–699, <https://doi.org/10.1002/qj.803>.
- Akbari Asanjan, A., T. Yang, K. Hsu, S. Sorooshian, J. Lin, and Q. Peng, 2018: Short-term precipitation forecast based on the PERSIANN system and LSTM recurrent neural networks. *J. Geophys. Res. Atmos.*, **123**, 12 543–12 563, <https://doi.org/10.1029/2018JD028375>.
- Ardabili, S., A. Mosavi, M. Dehghani, and A. R. Várkonyi-Kóczy, 2020: Deep learning and machine learning in hydrological processes climate change and Earth systems a systematic review. *Engineering for Sustainable Future. INTER-ACADEMIA 2019*, A. Várkonyi-Kóczy, Eds., Lecture Notes in Networks and Systems, **123**, Springer, 52–62, [https://doi.org/10.1007/978-3-030-36841-8\\_5](https://doi.org/10.1007/978-3-030-36841-8_5).
- Behrangi, A., K. L. Hsu, B. Imam, S. Sorooshian, and R. J. Kuliowski, 2009: Evaluating the utility of multispectral information in delineating the areal extent of precipitation. *J. Hydrometeorol.*, **10**, 684–700, <https://doi.org/10.1175/2009JHM1077.1>.
- , B. Khakbaz, T. C. Jaw, A. AghaKouchak, K. Hsu, and S. Sorooshian, 2011: Hydrologic evaluation of satellite precipitation products over a mid-size basin. *J. Hydrol.*, **397**, 225–237, <https://doi.org/10.1016/j.jhydrol.2010.11.043>.
- Bentley, L., 1990: K-d trees for semidynamic point sets. *SCG '90: Proceedings of the Sixth Annual Symposium on Computational Geometry*, Association for Computing Machinery, 187–197, <https://doi.org/10.1145/98524.98564>.
- Boucher, O., and Coauthors, 2013: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 571–657.
- Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5–32, <https://doi.org/10.1023/A:1010933404324>.
- Cai, K., and H. Wang, 2018: Cloud classification of satellite image based on convolutional neural networks. *Proc. IEEE International Conf. on Software Engineering and Service Sciences, ICSESS*, Beijing, China, Institute of Electrical and Electronics Engineers, 874–877, <https://doi.org/10.1109/ICSESS.2017.8343049>.
- Chen, H., V. Chandrasekar, R. Cifelli, and P. Xie, 2020: A machine learning system for precipitation estimation using satellite and ground radar network observations. *IEEE Trans. Geosci. Remote Sensing*, **58**, 982–994, <https://doi.org/10.1109/TGRS.2019.2942280>.
- Damberg, L., and A. AghaKouchak, 2014: Global trends and patterns of drought from space. *Theor. Appl. Climatol.*, **117**, 441–448, <https://doi.org/10.1007/s00704-013-1019-5>.
- Draper, D. W., D. A. Newell, F. J. Wentz, S. Krimchansky, and G. M. Skofronick-Jackson, 2015: The Global Precipitation Measurement (GPM) Microwave Imager (GMI): Instrument overview and early on-orbit performance. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing*, **8**, 3452–3462, <https://doi.org/10.1109/JSTARS.2015.2403303>.
- Ehsani, M. R., A. Zarei, H. v. Gupta, K. Barnard, and A. Behrangi, 2021: Nowcasting-Nets: Deep neural network structures for precipitation nowcasting using IMERG. arXiv, 41 pp., <https://arxiv.org/abs/2108.06868>.
- Foufoula-Georgiou, E., and Coauthors, 2020: Advancing precipitation estimation, prediction, and impact studies. *Bull. Amer. Meteor. Soc.*, **101**, E1584–E1592, <https://doi.org/10.1175/BAMS-D-20-0014.1>.
- Funk, C., A. Verdin, J. Michaelsen, P. Peterson, D. Pedreros, and G. Husak, 2015: A global satellite-assisted precipitation climatology. *Earth Syst. Sci. Data*, **7**, 275–287, <https://doi.org/10.5194/essd-7-275-2015>.
- Goodfellow, I., Y. Bengio, and A. Courville, 2016: *Deep Learning*. MIT Press, 775 pp.
- Gorooh, V. A., S. Kalia, P. Nguyen, K. Lin Hsu, S. Sorooshian, S. Ganguly, and R. R. Nemani, 2020: Deep Neural Network Cloud-Type Classification (DeepCTC) model and its application in evaluating PERSIANN-CCS. *Remote Sens.*, **12**, 316, <https://doi.org/10.3390/rs12020316>.
- Gou, Y., Y. Ma, H. Chen, and Y. Wen, 2018: Radar-derived quantitative precipitation estimation in complex terrain over the eastern Tibetan Plateau. *Atmos. Res.*, **203**, 286–297, <https://doi.org/10.1016/j.atmosres.2017.12.017>.
- Hayatbini, N., and Coauthors, 2019: Conditional Generative Adversarial Networks (cGANs) for near real-time precipitation estimation from multispectral GOES-16 satellite images—PERSIANN-cGAN. *Remote Sens.*, **11**, 2193, <https://doi.org/10.3390/rs11192193>.
- 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. Climatol.*, **43**, 1834–1853, <https://doi.org/10.1175/JAM2173.1>.
- Hooker, G., L. Mentch, and S. Zhou, 2021: Unrestricted permutation forces extrapolation: Variable importance requires at least one more model, or there is no free variable importance. *Stat. Comput.*, **31**, 82, <https://doi.org/10.1007/s11222-021-10057-z>.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Houze, R. A., 2014: Atmospheric dynamics. *Int. Geophys.*, **104**, 25–46, <https://doi.org/10.1016/B978-0-12-374266-7.00002-0>.
- Hsu, K., X. Gao, S. Sorooshian, and H. Gupta, 1997: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.*, **36**, 1176–1190, [https://doi.org/10.1175/1520-0450\(1997\)036<1176:PEFRSI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2).
- , H. Gupta, X. Gao, S. Sorooshian, and B. Imam, 2002: Self-organizing linear output map (SOLO): An artificial neural network suitable for hydrologic modeling and analysis. *Water Resour. Res.*, **38**, 1–17, <https://doi.org/10.1029/2001WR000795>.
- Hu, R., F. Fang, C. C. Pain, and I. M. Navon, 2019: Rapid spatio-temporal flood prediction and uncertainty quantification using a deep learning method. *J. Hydrol.*, **575**, 911–920, <https://doi.org/10.1016/j.jhydrol.2019.05.087>.
- Huffman, G. J., and Coauthors, 2020: Integrated multi-satellite retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG). *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, **67**, Springer, 343–353, [https://doi.org/10.1007/978-3-030-24568-9\\_19](https://doi.org/10.1007/978-3-030-24568-9_19).
- Janowiak, J. E., R. J. Joyce, and Y. Yarosh, 2001: A real-time global half-hourly pixel-resolution infrared dataset and its applications. *Bull. Amer. Meteor. Soc.*, **82**, 205–218, [https://doi.org/10.1175/1520-0477\(2001\)082<0205:ARTGHH>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<0205:ARTGHH>2.3.CO;2).
- Jeppesen, J. H., R. H. Jacobsen, F. Inceoglu, and T. S. Toftegaard, 2019: A cloud detection algorithm for satellite imagery based on deep learning. *Remote Sensing Environ.*, **229**, 247–259, <https://doi.org/10.1016/j.rse.2019.03.039>.

- Jha, M. K., and S. Sahoo, 2015: Efficacy of neural network and genetic algorithm techniques in simulating spatio-temporal fluctuations of groundwater. *Hydrol. Processes*, **29**, 671–691, <https://doi.org/10.1002/hyp.10166>.
- Joyce, R. J., and P. Xie, 2011: Kalman filter–based CMORPH. *J. Hydrometeorol.*, **12**, 1547–1563, <https://doi.org/10.1175/JHM-D-11-022.1>.
- Kidd, C., and V. Levizzani, 2011: Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.*, **15**, 1109–1116, <https://doi.org/10.5194/hess-15-1109-2011>.
- , J. Tan, P. E. Kirstetter, and W. A. Petersen, 2018: Validation of the Version 05 Level 2 precipitation products from the GPM Core Observatory and constellation satellite sensors. *Quart. J. Roy. Meteor. Soc.*, **144**, 313–328, <https://doi.org/10.1002/qj.3175>.
- , G. Huffman, V. Maggioni, P. Chambon, R. Oki, and C. K. Gov, 2021: The global satellite precipitation constellation: Current status and future requirements. *Bull. Amer. Meteor. Soc.*, **102**, E1844–E1861, <https://doi.org/10.1175/BAMS-D-20-0299.1>.
- Kirstetter, P. E., and Coauthors, 2012: Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar–based national mosaic QPE. *J. Hydrometeorol.*, **13**, 1285–1300, <https://doi.org/10.1175/JHM-D-11-0139.1>.
- , J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur, 2015: Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resour. Res.*, **51**, 1422–1442, <https://doi.org/10.1002/2014WR015672>.
- Krasnopolsky, V. M., and M. S. Fox-Rabinovitz, 2006: Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction. *Neural Networks*, **19**, 122–134, <https://doi.org/10.1016/j.neunet.2006.01.002>.
- Kratzert, F., D. Klotz, M. Herrnegger, A. K. Sampson, S. Hochreiter, and G. S. Nearing, 2019: Toward improved predictions in ungauged basins: Exploiting the power of machine learning. *Water Resour. Res.*, **55**, 11344–11354, <https://doi.org/10.1029/2019WR026065>.
- Kuligowski, R. J., and A. P. Barros, 1998: Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Wea. Forecasting*, **13**, 1194–1204, [https://doi.org/10.1175/1520-0434\(1998\)013<1194:LPPFAN>2.0.CO;2](https://doi.org/10.1175/1520-0434(1998)013<1194:LPPFAN>2.0.CO;2).
- Kummerow, C. D., 2020: Introduction to passive microwave retrieval methods. *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 67, Springer, 123–140, [https://doi.org/10.1007/978-3-030-24568-9\\_7](https://doi.org/10.1007/978-3-030-24568-9_7).
- , D. L. Randel, M. Kulie, N. Y. Wang, R. Ferraro, S. Joseph Munchak, and V. Petkovic, 2015: The evolution of the Goddard profiling algorithm to a fully parametric scheme. *J. Atmos. Oceanic Technol.*, **32**, 2265–2280, <https://doi.org/10.1175/JTECH-D-15-0039.1>.
- Li, J., D. Yuan, Y. Sun, and J. Li, 2022: Comparing the performances of WRF QPF and PERSIANN-CCS QPEs in karst flood simulation and forecasting by coupling the Karst-Liuxihe model. *Front. Earth Sci.*, <https://doi.org/10.1007/s11707-021-0909-6>, in press.
- Li, Z., Y. Wen, M. Schreier, A. Behrangi, Y. Hong, and B. Lambriksen, 2021: Advancing satellite precipitation retrievals with data driven approaches: Is black box model explainable? *Earth Space Sci.*, **8**, e2020EA001423, <https://doi.org/10.1029/2020EA001423>.
- Long, J., E. Shelhamer, and T. Darrell, 2015: Fully convolutional networks for semantic segmentation. *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Boston, MA, Institute of Electrical and Electronics Engineers, 3431–3440, <https://doi.org/10.1109/CVPR.2015.7298965>.
- Maddox, R. A., J. Zhang, J. J. Gourley, and K. W. Howard, 2002: Weather radar coverage over the contiguous United States. *Wea. Forecasting*, **17**, 927–934, [https://doi.org/10.1175/1520-0434\(2002\)017<0927:WRCOTC>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0927:WRCOTC>2.0.CO;2).
- Miao, Q., B. Pan, H. Wang, K. Hsu, and S. Sorooshian, 2019: Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network. *Water*, **11**, 977, <https://doi.org/10.3390/w11050977>.
- Morau, A., S. Dewitte, B. Cornelis, and A. Munteanu, 2019: Deep learning for precipitation estimation from satellite and rain gauges measurements. *Remote Sensing*, **11**, 2463, <https://doi.org/10.3390/rs11212463>.
- Nair, V. and G. E. Hinton, 2010: Rectified linear units improve restricted Boltzmann machines. *27th Int. Conf. on Machine Learning*, Haifa, Israel, ICML, 8 pp., <https://openreview.net/forum?id=rkb15iZdZB>.
- NASA, 2018: Global Precipitation Measurement (GPM) Mission: GPROF2017 Version 1 and Version 2 (used in GPM V5 processing). Algorithm Theoretical Basis Doc., 65 pp., [https://gpm.nasa.gov/sites/default/files/2020-05/ATBD\\_GPM\\_V5B\\_April15\\_2018.pdf](https://gpm.nasa.gov/sites/default/files/2020-05/ATBD_GPM_V5B_April15_2018.pdf).
- Nguyen, P., and Coauthors, 2016: A high resolution coupled hydrologic–hydraulic model (HiResFlood-UCI) for flash flood modeling. *J. Hydrol.*, **541**, 401–420, <https://doi.org/10.1016/j.jhydrol.2015.10.047>.
- , E. J. Shearer, M. Ombadi, V. A. Gorooh, K. Hsu, S. Sorooshian, W. S. Logan, and M. Ralph, 2020: PERSIANN Dynamic Infrared–Rain Rate Model (PDIR) for high-resolution, real-time satellite precipitation estimation. *Bull. Amer. Meteor. Soc.*, **101**, E286–E302, <https://doi.org/10.1175/BAMS-D-19-0118.1>.
- O’Gorman, P. A., and J. G. Dwyer, 2018: Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. *J. Adv. Model. Earth Syst.*, **10**, 2548–2563, <https://doi.org/10.1029/2018MS001351>.
- Pan, B., K. Hsu, A. AghaKouchak, and S. Sorooshian, 2019: Improving precipitation estimation using convolutional neural network. *Water Resour. Res.*, **55**, 2301–2321, <https://doi.org/10.1029/2018WR024090>.
- Petty, G. W., and R. Bennartz, 2017: Field-of-view characteristics and resolution matching for the Global Precipitation Measurement (GPM) Microwave Imager (GMI). *Atmos. Meas. Tech.*, **10**, 745–758, <https://doi.org/10.5194/amt-10-745-2017>.
- Pfeundschuh, S., P. Eriksson, D. Duncan, B. Rydberg, N. Häkansson, and A. Thoss, 2018: A neural network approach to estimating a posteriori distributions of Bayesian retrieval problems. *Atmos. Meas. Tech.*, **11**, 4627–4643, <https://doi.org/10.5194/amt-11-4627-2018>.
- Prigent, C., F. Aires, and W. B. Rossow, 2003: Retrieval of surface and atmospheric geophysical variables over snow-covered land from combined microwave and infrared satellite observations. *J. Appl. Meteor. Climatol.*, **42**, 368–380, [https://doi.org/10.1175/1520-0450\(2003\)042<0368:ROSAAG>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<0368:ROSAAG>2.0.CO;2).
- Raghavendra, S. N., and C. Deka, 2014: Support vector machine applications in the field of hydrology: A review. *Appl. Soft*

- Comput.*, **19**, 372–386, <https://doi.org/10.1016/j.asoc.2014.02.002>.
- Rasp, S., M. S. Pritchard, and P. Gentine, 2018: Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci. USA*, **115**, 9684–9689, <https://doi.org/10.1073/pnas.1810286115>.
- Reichstein, M., G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, 2019: Deep learning and process understanding for data-driven Earth system science. *Nature*, **566**, 195–204, <https://doi.org/10.1038/s41586-019-0912-1>.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: Convolutional networks for biomedical image segmentation. *MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention*, N. Navab et al., Eds., Lecture Notes in Computer Science, Vol. 9351, Springer, 234–241, [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- Sadeghi, M., A. A. Asanjan, M. Farizad, P. H. U. Nguyen, K. Hsu, S. Sorooshian, and D. A. N. Braithwaite, 2019: PERSIANN-CNN: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Convolutional Neural Networks. *J. Hydrometeorol.*, **20**, 2273–2289, <https://doi.org/10.1175/JHM-D-19-0110.1>.
- , P. Nguyen, K. Hsu, and S. Sorooshian, 2020: Improving near real-time precipitation estimation using a U-Net convolutional neural network and geographical information. *Environ. Modell. Software*, **134**, 104856, <https://doi.org/10.1016/j.envsoft.2020.104856>.
- Sanò, P., G. Panegrossi, D. Casella, A. C. Marra, F. di Paola, and S. Dietrich, 2016: The new Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for the cross-track scanning ATMS radiometer: Description and verification study over Europe and Africa using GPM and TRMM spaceborne radars. *Atmos. Meas. Tech.*, **9**, 5441–5460, <https://doi.org/10.5194/amt-9-5441-2016>.
- Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Amer. Meteor. Soc.*, **98**, 1679–1695, <https://doi.org/10.1175/BAMS-D-15-00306.1>.
- , D. Kirschbaum, W. Petersen, G. Huffman, C. Kidd, E. Stocker, and R. Kakar, 2018: The Global Precipitation Measurement (GPM) mission’s scientific achievements and societal contributions: Reviewing four years of advanced rain and snow observations. *Quart. J. Roy. Meteor. Soc.*, **144**, 27–48, <https://doi.org/10.1002/qj.3313>.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSS>2.3.CO;2).
- Stephens, G., and Coauthors, 2020: The emerging technological revolution in Earth observations. *Bull. Amer. Meteor. Soc.*, **101**, E274–E285, <https://doi.org/10.1175/BAMS-D-19-0146.1>.
- Sun, A. Y., and G. Tang, 2020: Downscaling satellite and reanalysis precipitation products using attention-based deep convolutional neural nets. *Front. Water*, **2**, 22, <https://doi.org/10.3389/frwa.2020.536743>.
- Tang, G., Y. Wen, J. Gao, D. Long, Y. Ma, W. Wan, and Y. Hong, 2017: Similarities and differences between three coexisting spaceborne radars in global rainfall and snowfall estimation. *Water Resour. Res.*, **53**, 3835–3853, <https://doi.org/10.1002/2016WR019961>.
- , D. Long, A. Behrangi, C. Wang, and Y. Hong, 2018: Exploring deep neural networks to retrieve rain and snow in high latitudes using multisensor and reanalysis data. *Water Resour. Res.*, **54**, 8253–8278, <https://doi.org/10.1029/2018WR023830>.
- Tao, Y., K. Hsu, A. Ihler, X. Gao, and S. Sorooshian, 2018: A two-stage deep neural network framework for precipitation estimation from bispectral satellite information. *J. Hydrometeorol.*, **19**, 393–408, <https://doi.org/10.1175/JHM-D-17-0077.1>.
- Utsumi, N., F. J. Turk, Z. S. Haddad, P. E. Kirstetter, and H. Kim, 2020: Evaluation of precipitation vertical profiles estimated by GPM-era satellite-based passive microwave retrievals. *J. Hydrometeorol.*, **22**, 95–112, <https://doi.org/10.1175/JHM-D-20-0160.1>.
- Xiao, R., and V. Chandrasekar, 1997: Development of a neural network based algorithm for rainfall estimation from radar observations. *IEEE Trans. Geosci. Remote Sensing*, **35**, 160–171, <https://doi.org/10.1109/36.551944>.
- Xie, F., M. Shi, Z. Shi, J. Yin, and D. Zhao, 2017: Multilevel cloud detection in remote sensing images based on deep learning. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sensing*, **10**, 3631–3640, <https://doi.org/10.1109/JSTARS.2017.2686488>.
- Yaseen, Z. M., A. El-shafie, O. Jaafar, H. A. Afan, and K. N. Sayl, 2015: Artificial intelligence based models for streamflow forecasting: 2000–2015. *J. Hydrol.*, **530**, 829–844, <https://doi.org/10.1016/j.jhydrol.2015.10.038>.
- You, Y., V. Petkovic, J. Tan, R. Kroodsma, W. Berg, C. Kidd, and C. Peters-Lidard, 2020: Evaluation of V05 precipitation estimates from GPM constellation radiometers using KuPR as the reference. *J. Hydrometeorol.*, **21**, 705–728, <https://doi.org/10.1175/JHM-D-19-0144.1>.
- Yu, P. S., S. T. Chen, and I. F. Chang, 2006: Support vector regression for real-time flood stage forecasting. *J. Hydrol.*, **328**, 704–716, <https://doi.org/10.1016/j.jhydrol.2006.01.021>.
- Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, <https://doi.org/10.1175/BAMS-D-14-00174.1>.