Copula-Based Modeling of TMI Brightness Temperature With Rainfall Type

J. Indu and D. Nagesh Kumar

***Abstract*—Overland rain retrieval using spaceborne microwave radiometer offers a myriad of complications as land presents itself as a radiometrically warm and highly variable background. Hence, land rainfall algorithms of the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) have traditionally incorporated empirical relations of microwave brightness temperature (Tb) with rain rate, rather than relying on physically based radiative transfer modeling of rainfall (as implemented in the TMI ocean algorithm). In this paper, sensitivity analysis is conducted using the Spearman rank correlation coefficient as benchmark, to estimate the best combination of TMI low-frequency channels that are highly sensitive to the near surface rainfall rate from the TRMM Precipitation Radar (PR). Results indicate that the TMI channel combinations not only contain information about rainfall wherein liquid water drops are the dominant hydrometeors but also aid in surface noise reduction over a predominantly vegetative land surface background. Furthermore, the variations of rainfall signature in these channel combinations are not understood properly due to their inherent uncertainties and highly nonlinear relationship with rainfall. Copula theory is a powerful tool to characterize the dependence between complex hydrological variables as well as aid in uncertainty modeling by ensemble generation. Hence, this paper proposes a regional model using Archimedean copulas, to study the dependence of TMI channel combinations with respect to precipitation, over the land regions of Mahanadi basin, India, using version 7 orbital data from the passive and active sensors on board TRMM, namely, TMI and PR. Studies conducted for different rainfall regimes over the study area show the suitability of Clayton and Gumbel copulas for modeling convective and stratiform rainfall types for the majority of the intraseasonal months. Furthermore, large ensembles of TMI Tb (from the most sensitive TMI channel combination) were generated conditional on various quantiles (25th, 50th, 75th, and 95th) of the convective and the stratiform rainfall. Comparatively greater ambiguity was observed to model extreme values of the convective rain type. Finally, the efficiency of the proposed model was tested by comparing the results with traditionally employed linear and quadratic models. Results reveal the superior performance of the proposed copula-based technique.**

***Index Terms*—Copula, quantile regression, river basin, Tropical Rainfall Measuring Mission (TRMM).**

Manuscript received June 16, 2013; revised September 5, 2013; accepted October 2, 2013. The work of D. N. Kumar was supported by IBM through IBM Faculty Award 2012 and the Ministry of Earth Sciences, Government of India, through the project MOES/ATMOS/PP-IX/09.

The authors are with the Department of Civil Engineering, Indian Institute of Science, Bangalore 560 012, India.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2013.2285225

I. Introduction

OVER the years, microwave radiometers have proven to be valuable tools for the quantitative estimation of precipitation from space with its cloud-penetrating capability. The passive and active sensors on board the Tropical Rainfall Measuring Mission (TRMM) [successful program by National Aeronautics and Space Administration (NASA) and Japan's space agency Japan Aerospace Exploration Agency], namely, the TRMM Microwave Imager (TMI) and Precipitation Radar (PR), have taken unprecedented satellite images of the Earth's weather for the past 16 years. The combination of a spaceborne radar (i.e., PR) and a radiometer (i.e., TMI) on the same space platform has advanced microwave rainfall retrieval techniques considerably due to the increased understanding in the transfer of microwave radiation through clouds and hydrometeors (precipitation-sized particles).

The brightness temperature (Tb) registered by a downward viewing spaceborne radiometer has an indirect relationship with the rainfall rate dependent on background emissivity. Hence, the underlying physics for rainfall retrieval is different for land (overland algorithms) and oceans (overocean algorithms). The ocean surfaces offer a radiometrically cold background enabling the warm randomly polarized emission from rainfall to be easily distinguished from ocean surface emission. Land, on the contrary, exists as a radiometrically warm and highly unpolarized background. The highly nonhomogeneous land surface background tends to have varying emissivity values which add clutter to the emission from rainfall, thereby making it extremely difficult to detect rainfall signals. Hence, land rainfall retrieval algorithms using passive microwaves (PMWs) have been traditionally dependent on empirical relationships utilizing the ice scattering property at 85.5-GHz Tb [1]–[7]. The quantitative assessment of PMW rainfall indicates that, although retrieval over ocean surfaces is performed with acceptable accuracy, overland retrieval, based on ice scattering at an 85-GHz microwave frequency channel, continues to remain ambiguous. Some of the well-documented algorithms for PMW rainfall retrieval over land regions are summarized as follows. Rainfall detection using polarization-corrected temperature (PCT) was first introduced by Spencer *et al.* [2]. This index relied on a linear combination of the vertically and the horizontally polarized Tb at an 85.5-GHz channel to extract the ice scattering signature and obtain a continuous precipitation field [3]. Statistical regression-based algorithms were developed based on rain indices like the PCT and scattering index (SI). The underlying concept of SI was first proposed by Grody [4] to estimate the ice scattering signature of raining clouds by

subtracting the nonscattering portion of Tb from the observed 85-GHz Tb. This approach was used to derive empirical relations with rainfall [5], [6] and has been used in the current TRMM facility overland rain retrieval algorithm. Algorithms were developed relating rainfall rates to the difference between 19- and 85-GHz Tb by Liu and Curry [10]. Prabhakara *et al.* [11] developed a global model to estimate the overland rain rate from TMI observations utilizing the index of optical depth, the difference in Tb values of 19- and 37-GHz channels, and the horizontal gradient of Tb values in the 85.5-GHz channel. Studies using artificial neural networks were conducted to relate TMI Tb from low-frequency microwave channels with respect to the near surface rain rates of PR, over different oceanic and land regions of India [12], [13]. Several regional models were developed linking TMI Tb with PR rainfall rate. Dinku and Anagnostou [14] performed an empirical modeling of TMI land rain rates for the summer seasons of 2000 to 2002 for four different convective tropical regions. Their algorithm consisted of multichannel-based rain screening and convective/stratiform rain classification followed by the fitting of a nonlinear (linear) regression for the rain rate retrieval of stratiform (convective) rain regimes. In their study, it was observed that, among the four geographic regions considered, the Ganga Brahmaputra Meghna river basin has a significant difference between global and regional calibrations. Gopalan *et al.* [7], using the University of Utah level 1 precipitation feature database, developed rain rate relationships with respect to the 85-GHz Tb. Their study proposed a robust cubic polynomial model of TMI rain rate with Tb at 85 GHz for convective storms while a linear model was developed for stratiform storms. Aonashi *et al.* [15] developed an overland empirical algorithm based on ice scattering signals wherein Tb from a 37-GHz channel was utilized as a scattering correction factor to the 85-GHz scattering signatures. The factor was applied to overcome the saturation of the 85-GHz scattering signature during heavy rainfall. These observations lead to strive for a regional model over the Indian landmass.

All the algorithms highlighted the fundamental dependence of overland rainfall with ice scattering signatures at the 85-GHz frequency channel. However, these retrieval techniques suffer from a major disadvantage of being inherently empirical in nature due to the unknown phase, density, size, distribution, orientation, and shape of ice particles within the sampling volume [16], [17]. Moreover, the increased wetness of the land surface during rainfall tends to alter the emissivity of land which can be misinterpreted as due to ice scattering and, thereby, as rainfall signature. As the frozen hydrometeors have an indirect relationship with surface rainfall that varies significantly from region to region, the algorithms solely relying on ice scattering fail to detect rainfall from clouds that lack ice particles [18], [19]. Due to these problems, the transfer function linking microwave Tb with rainfall rates is not well understood in the scattering regime for overland rainfall retrieval.

Keeping these in mind, the main objective of this paper is to develop an integrated regional model to estimate the joint variability of PR rainfall with respect to TMI Tb over the midsize basin of Mahanadi, India. For modeling, this study uses low-frequency channel combinations of TMI instead of relying on the 85.5-GHz high-frequency channel. Channel combinations were selected based on their increased sensitivity to overland rainfall for the study area. The relationship between TMI channel Tb and rainfall is highly uncertain and nonlinear in nature. Copula theory is well known to characterize complex hydrological variables as well as aid in their uncertainty modeling by ensemble generation [20]–[22]. Hence, for this study, the Tb–rainfall rate relationship is modeled using copula theory. Section II presents the description of TRMM data products used in this study. The study region chosen is explained in Section III. The proposed methodology is outlined in Section IV. Section V summarizes the results of the proposed method applied on Mahanadi basin, India. Finally, the key conclusions are outlined in Section VI.

II. TRMM Data Products

TRMM is a joint mission between the National Aeronautics and Space Administration and the Japan Aerospace Exploration Agency to monitor and study the tropical rainfall. Launched in 1997 into a near circular orbit, it has two instruments operating in the microwave spectrum, namely, TMI and PR. A detailed description of the TRMM sensor package is available in [23]. To summarize, the passive instrument TMI measures Tb at five different frequencies (10.65, 19.35, 21.3, 37.0, and 85.5 GHz) using both horizontal (H) and vertical (V) polarizations except for the 21.3-GHz channel which is measured in just the vertical polarization. Hereinafter, these channels will be referred to as 10 V, 10 H, 19 V, 19 H, 21 V, 37 V, 37 H, 85 V, and 85 H, respectively. The concept of effective field of view (EFOV) is introduced wherein the EFOV in the cross-track (CT) direction represents the results of one integration time period or "one sample." When compared with the instantaneous field of view (IFOV), the EFOV-CT appears to be "artificially narrow" because the EFOV in the down track (DT) direction is taken to be same as the IFOV-DT [23]. The major instrument characteristics of TMI are tabulated in Table I [23]. The active instrument PR, operating at a frequency of 13.8 GHz (Ku-band), is capable of the following: 1) providing the 3-D structure of rainfall, particularly of the vertical distribution; 2) obtaining quantitative rainfall measurements over land as well as over ocean; and 3) improving the overall accuracy of TRMM precipitation retrieval by the combined use of active (PR) and passive (TMI) sensor data. In this paper, we use version 7 TMI 1B11 data for Tb data, PR 2A21 for surface flag data, PR 2A25 for near surface rainfall rate (NSR) data, and PR 2A23 for rain-type data.

A. Collocation Strategy

The 1B11 data product provides Tb measured at EFOV with horizontal resolutions varying with frequency (5 km $\times$ 7 km for 85 GHz to 10 km $\times$ 63 km for 10 GHz). Due to the varying spatial resolutions along the DT and CT directions, for each of the TMI channels among themselves and with that of PR (4.3 km $\times$ 5 km) data, collocation was performed as the initial step. Several studies have approached collocation by spatial resolution enhancement [24], [25]. In this study, the resolution of low-frequency channels (10 V, 10 H, 19 V, 19 H, 21 V, 37 V, and 37 H) is increased by linear interpolation technique

TABLE I
TMI INSTRUMENT CHARACTERISTICS

Channel Frequency (GHz)	10.65	10.65	19.35	19.35	21.30	37.00	37.00	85.50	85.50
Polarization	V	H	V	H	V	V	H	V	H
Instantaneous field of view (IFOV)-Down track (DT) (km)	59.0	60.1	30.5	30.1	27.2	16.0	16.0	6.7	6.9
IFOV-Cross track (CT) (km)	35.7	36.4	18.4	18.2	16.5	9.7	9.7	4.1	4.2
Integration time per sample (ms)	6.60	6.60	6.60	6.60	6.60	6.60	6.60	3.30	3.30
Effective field of view (EFOV)-CT (km)	9.1	9.1	9.1	9.1	9.1	9.1	9.1	4.6	4.6
EFOV-DT (km)	63.2	63.2	30.4	30.4	22.6	16.0	16.0	7.2	7.2
EFOVs per scan	104	104	104	104	104	104	104	208	208

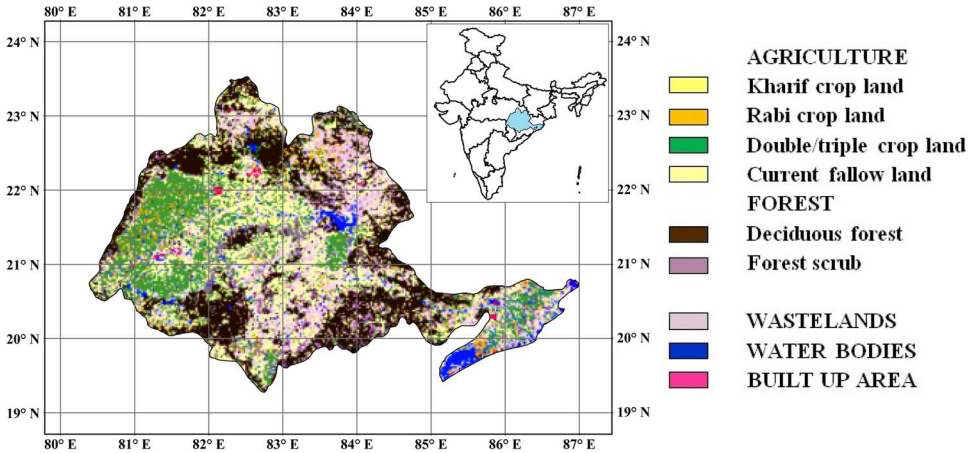

Fig. 1. Land use/land cover map of Mahanadi basin (for year 2010).

to match it with the resolution of 85 V channels. Collocation is performed by using the geolocation information from the TRMM PR and TMI data set, to assign a TMI pixel at the 85 V resolution as the "nearest neighbor" for every PR pixel in an orbit, using (1), shown at the bottom of the page. Here, D_i refers to the distance between each of the ith TMI pixels from a given PR pixel. This process makes available three to four PR pixels as the nearest neighbors for every TMI pixel within a PR swath [7]. As a result, for every high-resolution TMI 85 V pixel, corresponding PR pixels, near surface rain rates, and rain type were estimated. To extract all the pixels lying over the land region, surface-type information from the PR data product was used. The rainfall type (convective and stratiform) represented by each of these collocated overland pixels was estimated by utilizing the storm-type information present in the TRMM 2A23 data product. The NSR data archived within the PR 2A25 product refer to the rainfall rate at the lowest point in the clutter-free region, estimated using radar reflectivity–rainfall (Z-R) relationship. Understanding this relationship over the complex Indian terrain is a big challenge. This study uses a four-year data period from 2008 to 2012 to examine the variability of collocated TMI channel combinations with rainfall types from PR. The selection of the data period was based on previous works in the literature which conducted TMI Tb–rainfall related studies [26]–[29].

III. STUDY REGION

The study region chosen for this work is the basin of Mahanadi river, India, situated between latitudes 19° N to

$$D_i = \sqrt{(\text{Latitude}_{\text{PR}} - \text{Latitude}_{\text{TMI},i})^2 + (\text{Longitude}_{\text{PR}} - \text{Longitude}_{\text{TMI},i})^2} \quad (1)$$

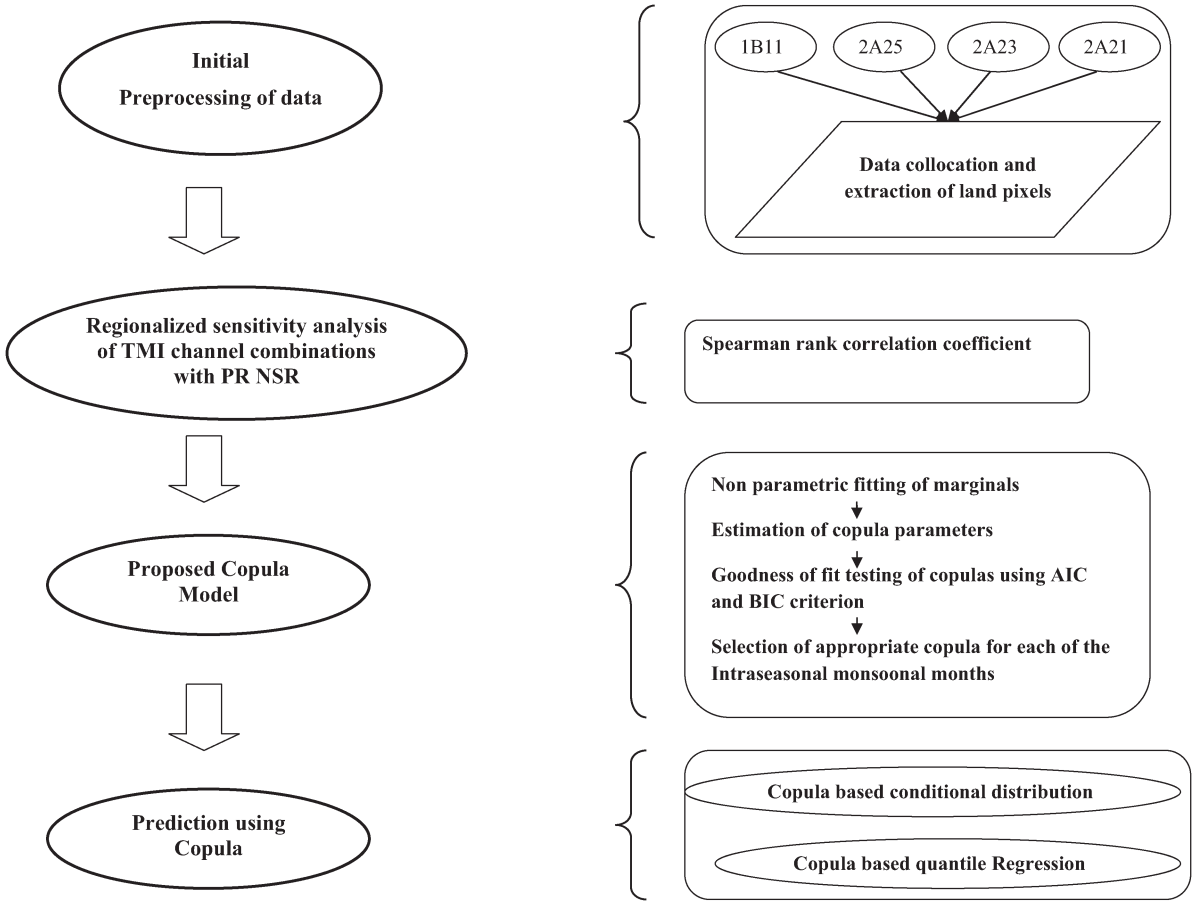

Fig. 2. Flowchart summarizing the methodology.

24° N and longitudes 80° E to 87° E (see Fig. 1). The physiographic classification of the basin comprises the hilly regions of the northern plateau and Eastern Ghats, the delta of coastal plains, and the central interior region traversed by the river Mahanadi and its tributaries. The land use/land cover of the basin consists of forest, cropland, grasslands, etc. (see Fig. 1). The basin receives heavy to very heavy rainfall when monsoon depressions from the Bay of Bengal move northwestward slightly south of their normal track. Mahanadi basin is notorious for being subjected to frequent flooding every year. Nearly 91% of the annual precipitation (600 to over 1600 mm) for the basin occurs from June to September (JJAS), also known as the summer monsoon months, which is responsible for influencing the agricultural output from the basin. Even a small variation of this seasonal rainfall can have an adverse impact on the economy. Analysis of PMW data to accurately estimate rainfall over a hydrologically variant basin such as Mahanadi stresses on the proper estimation of the dependence between PMW frequency channels and rainfall intensities. Within the microwave spectrum, a predominantly vegetative land surface background, such as that observed in the basin, will contribute volume scattering since the microwave radiation can arise from below and within the canopy. It is thus a very complex and challenging task to discern the atmospheric contribution to the upwelling Tb. An analysis of the response of TMI channel frequencies in the presence of rainfall types over the complicated land background of the basin will greatly aid in future studies pertaining to rainfall modeling, extremes in precipitation, flood forecasts, weather forecasting, etc.

IV. Proposed Methodology

The purpose of this paper is to study the dependence of TMI Tb (from TMI channel combinations) with PR NSR types. A flowchart summarizing the methodology adopted in this paper is shown in Fig. 2. The data products from 1B11, 2A21, 2A25, and 2A23 are subjected to the initial data processing procedure consisting of collocation, extraction of overland pixels, and division into convective and stratiform pixels. This procedure is applied for nearly 1397 orbits passing over the study region between the monsoonal months of 2008 to 2012. For the overland regions of the study area, a total of 9216 and 26 139 data points are extracted to represent convective and stratiform rain types, respectively.

A. Sensitivity Analysis

Overland rainfall retrieval algorithms using PMW techniques are based on the ice scattering phenomenon at the 85 V channel. Recent studies by You *et al.* [28] using three years (1998 to 2000) of TRMM orbital data examined the correlation coefficients between PR NSR and TMI Tb using 81 channels (9 TMI channels + 72 channel combinations from the 9 TMI channels) for the overland regions of tropics. The study came out with 20 channels that were highly sensitive to NSR. This

study uses these 20 channel combinations (hereinafter referred to also as TMI channels). The TMI channel sensitivities for PR NSR over the study area are examined using the Spearman rank correlation coefficient. The dependence between variables is usually measured using Pearson's coefficient of correlation (CC). However, this parameter only models linear dependence, which, in some cases, may not exist [29]. The relationship of TMI Tb with NSR from PR is highly nonlinear in nature. Hence, traditional means of correlating both variables using the Pearson product moment correlation coefficient might not be appropriate. Therefore, the following analysis makes use of the more robust Spearman rank correlation coefficient. If x and y denote the ranks of data pairs, n denotes the sample size, and $\overline{x}$ and $\overline{y}$ denote the means of x and y, the Spearman's rank correlation coefficient is defined as

$$r_{xy} = \frac{\frac{1}{n-1}\sum_{i=1}^{i=n}(x_i-\overline{x})(y_i-\overline{y})}{\left[\frac{1}{n-1}\sum_{i=1}^{i=n}(x_i-\overline{x})^2\right]^{\frac{1}{2}} * \left[\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2\right]^{\frac{1}{2}}}. \quad (2)$$

Due to the inherent uncertainties of TMI channels with rainfall rate, the transfer function/joint density of these sensitive channels for each of the JJAS months was modeled against rainfall types (convective and stratiform) from PR using copula theory.

B. Copula Theory

Natural events like rainfall often result due to the joint behavior of several mutually dependent random variables that require a multivariate approach to analyze and study both the hydrological and the meteorological phenomenon. Hence, for this study, copula theory is proposed to model the joint variability between TRMM Tb (from passive sensor TMI) and NSR (from active sensor PR). In the field of water resources engineering, although the theory of copulas has been successfully applied in various applications involving flood frequency analysis and soil moisture-related studies [30]–[33], its development is still in a nascent stage.

The theory of copulas first introduced by Sklar [34] is used to obtain the joint distribution of two continuous random variables once their marginal distributions are known/estimated. For a bivariate case, the Sklar theorem [34] is stated as follows:

Let $H_{X,Y}(x,y)$ be a 2-D joint distribution function with marginal distributions as $F_X(x)$ and *$G_Y(y)$. Then, there exists a copula C such that, for all x, y, $F_X(x)$, and $G_Y(y) \in R$*

$$H_{X,Y}(x,y) = C\left[F_X(x), G_Y(y)\right]. \quad (3)$$

Pertaining to this study, the statement means that, if X and Y are continuous random variables representing TMI Tb (K) and PR rainfall rates (mm/h), $F_X(x)$ and $G_Y(y)$ represent their corresponding marginal distribution functions, and $H_{X,Y}(x,y)$ denotes their joint distribution function, then there exists a copula (C) that joins/couples the joint distribution function $(H_{X,Y}(x,y))$ to their corresponding 1-D marginal distribution functions ($F_X(x)$ and $G_Y(y)$) [34]. (The usual convention adopted in probability theory is being followed here wherein an uppercase letter denotes a random variable and a lowercase letter refers to the value taken by the corresponding random variable.) This definition of copula theory can also be represented in another way. Let $F_X(x)$ and $G_Y(y)$ represent the cumulative distribution functions (*cdfs*) of the variables X and Y with $H_{X,Y}(x,y)$ as their joint *cdf*. Since the three functions $F_X(x)$, $G_Y(y)$, and $H_{X,Y}(x,y)$ lie in the interval [0, 1], each pair of variables (x,y) can be represented in terms of a point $(F_X(x), G_Y(y))$ within the unit square [0, 1] $\times$ [0, 1], and this ordered pair, in turn, corresponds to a number $H_{X,Y}(x,y)$ in [0, 1] [35]. As such, in copula theory, the calculation of joint distribution is divided between calculating the individual marginal *cdf* of both the variables and calculating the copula function (C). The dependence relationship is entirely determined by the copula, while the scaling and shape are entirely determined by the marginals.

1) Estimation of Nonparametric Marginal Distributions: The flexibility of copula theory is that any type of marginal distributions can be joined/coupled to obtain their joint distribution. The marginal distributions can be estimated either by fitting any of the parametric distributions (e.g., normal, exponential, etc.) or by using the nonparametric approach of kernel density. This study uses the nonparametric kernel density-based technique [36] to estimate the marginal distributions of TMI Tb and PR rainfall rate, owing to its flexibility in capturing the scale-free dependence pattern between both the variables. The kernel density technique involves a weighted moving average of the empirical frequency distribution of the sample [37], [38]. The kernel density-based estimation of probability density function (pdf) pertaining to hydrologic variables can be found in [39]–[41], etc. The kernel density-based technique computes the statistical distribution using histograms that are estimated using the relation [42]

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} k\left(\frac{x-x_i}{h}\right) \quad (4)$$

where k denotes the kernel function, h is the bandwidth/smoothing parameter used for smoothing the shape of the estimated pdf, and x_i is the ith observation. For this study, out of the different types of kernel functions, the Epanechnikov kernel function is used, which is given as

$$\begin{aligned} k(x) &= 0.75(1-x^2), \qquad |x| \le 1 \\ k(x) &= 0, \qquad \text{otherwise.} \end{aligned} \quad (5)$$

2) Estimation of Copula Parameters: A bivariate copula captures the scale-free dependence structure between two variables X and Y which implies that the manner in which X and Y "move together" is modeled regardless of the scale in which each variable is measured. For a bivariate copula, the scale of dependence is described by its copula parameters which are then used to determine the joint distribution and simulate the marginals [33]. Generally, correlation measures are used to summarize the information in copula. As the usual Pearson linear product moment correlation depends on the marginal distributions, it is not a desirable measure of association for non-normal multivariate distributions. Two standard nonparametric correlation measures, namely, Spearman's correlation and Kendall's correlation, are widely used which can be expressed solely in terms of the copula function [43].

TABLE II
ARCHIMEDEAN COPULAS WITH THEIR GENERATOR FUNCTIONS ALONG WITH KENDALL'S TAU

FAMILY	$C_\theta(u,v)$	$\varphi_\theta(t)$	$\theta \in$	τ
Clayton	$[\max(u^{-\theta}+v^{-\theta}-1,0)]^{\frac{1}{\theta}}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$[-1,\infty)\setminus 0$	$\frac{\theta}{\theta+2}$
Frank	$-\frac{1}{\theta}\ln\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right)$	$-\ln\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$(-\infty,\infty)\setminus 0$	$1-\frac{4}{\theta}[D_1(-\theta)-1]^a$
Gumbel	$\exp(-[(-\ln u)^\theta+(-\ln v)^\theta]^{\frac{1}{\theta}})$	$(-\ln t)^\theta$	$[1,\infty)$	$\frac{\theta-1}{\theta}$

D_1 is first order Debye function

This study uses three types of Archimedean [44] copulas, namely, Clayton, Gumbel, and Frank, for modeling. The general structure of the Archimedean copula is given by

$$C_\phi(u,v) = \phi^{-1}\left(\phi(u)+\phi(v)\right) \qquad \text{for } u,v \in (0,1] \tag{6}$$

where ϕ denotes the copula generator which is a convex decreasing function with domain [0, 1] and range [0, ∞), ϕ^{-1} is the inverse copula generator, and u and v are the marginal distributions of the variables X and Y (i.e., $u = F_X(x)$ and $v = F_Y(y)$). The functional forms of the three common Archimedean families of copulas are shown in Table II.

For the Archimedean family, this study estimates copula parameters by calculating Kendall's rank correlation for the copula and for the data. All the three types of Archimedean copula used in this study are one-parameter copula, wherein the strength of the dependence between the two variables increases with an increase in the copula parameter (θ). The relationship between Kendall's rank correlation coefficient τ and ϕ for an Archimedean copula family is given by

$$\tau = 1 + 4\int_0^1 \frac{\phi(u)}{\phi'(u)}\,du \tag{7}$$

where $\phi'(u)$ represents the derivative of $\phi(u)$ with respect to u (marginal distribution). Kendall's τ can also be mathematically expressed in terms of copula function (C) according to (8) [35] as

$$\tau = 4\int C(u,v)dC(u,v) - 1. \tag{8}$$

Kendall's rank correlation coefficient denotes a nonparametric measure of association between two variables. If (x_1,y_1), $(x_2,y_2),\ldots(x_n,y_n)$ denote the n pairs of both the random variables (TMI Tb and PR NSR for our study), Kendall's τ estimates the difference between the probability of concordance and the probability of discordance for the pairs of random variables using the relationship

$$\tau = \frac{n_C - n_D}{n(n+1)/2} \tag{9}$$

where n denotes the sample size, and n_C and n_D are the number of concordant and discordant pairs in the sample. Two pairs (x_1,y_1) and (x_2,y_2) are known to be concordant if $(x_i - x_2)(y_1-y_2) > 0$ and discordant if $(x_i-x_2)(y_1-y_2) < 0$. All the n sample data are being compared pairwise, resulting in $n(n+1)/2$ comparisons.

A crucial factor in dependence modeling is the tail dependence which will be different for different families of Archimedean copulas. The tail dependence determines the association between the extreme values of two random variables and depends only on their copula. The Gumbel copula [45] is usually used for asymmetrical tail dependence structure [46], i.e., it exhibits higher correlation in the right tail. If X and Y are two variables, the upper tail dependence using the Gumbel copula models the probability that Y exceeds a given threshold given that X has already exceeded that threshold. The Clayton copula has a lower tail dependence or tighter concentration of mass in the left tail. The Clayton copula is considered appropriate to model the probability that Y is below a threshold, given that X is already below that threshold. The Frank [47]–[49] copula is the only Archimedean copula family which is radially symmetrical (i.e., symmetric about the main diagonal and antidiagonal of its domain). A good overview of tail dependence for various families of Archimedean copulas can be found in [50] and [51].

3) Fitting Most Appropriate Copula: After calculating the parameters of each copula, it is necessary to decide which copula family best represents the dependence structure between the variables of interest (TMI Tb and NSR in our case). For our study, the choice of an appropriate copula for each of the monsoonal months was based on the measures of the Akaike information criterion (AIC) and Bayesian information criterion (BIC). Both these measures are adopted because of their capability to describe the tradeoff between bias (or accuracy) and variance (complexity) in model construction. AIC conducts model comparison based on the concept of information entropy while BIC estimates the model fit from the perspective of decision theory. In an absolute sense, the AIC/BIC values of a single model are unable to convey information about how well that particular model fits the data. The measures of AIC and BIC find meaning only when compared with several model fits. AIC measures the relative goodness of fit of a statistical model. If k is the number of parameters in the copula and L is the maximized value of the likelihood function (for a copula family), then the expressions for AIC and BIC are given as

$$AIC = 2k - 2\ln(L) \tag{10}$$
$$BIC = -2\ln(L) + k\ln(N). \tag{11}$$

The copula having the highest value for the log likelihood or the lowest values for AIC and BIC measures is chosen to represent the dependence structure in a better manner.

To summarize, the complete copula methodology can be listed in the form of the following steps: 1) Fitting marginal distributions ($F_X(x)$ and $G_Y(y)$) for each of the random variables X and Y, using techniques described well in the statistical literature. This study used a kernel density-based nonparametric approach to estimate marginals of both the variables. 2) Fitting the copula after estimating the appropriate copula function (C). For the Archimedean family of copulas, the relationship between Kendall's τ and copula generator ϕ reduces the copula fitting step to estimating the Kendall's τ from the data and solving (7).

The predictive potential of copula theory is explained below using conditional density and quantile regression.

C. Copula-Based Conditional Distribution

Once the copula-based joint distribution is estimated, i.e., $F_X(x)$, $G_Y(y)$, and $C(u,v)$ are obtained, both unconditional and conditional random samples can be generated from this distribution using Monte Carlo simulations [33]. Ensemble generation using conditional distribution is especially important for our study because of the uncertainty in both the variables of interest (i.e., TMI Tb and PR NSR) and because of the need to estimate the conditional probability distribution of Tb values, given the various quantiles of rainfall. The conditional distribution of a random variable (X), conditioned on an observed variable (Y), can be estimated using (12) [32] as

$$C_{X/Y=y}(x) = \frac{\partial}{\partial y} C_{X,Y}(x,y)/Y = y. \tag{12}$$

This means that, once the copula family that best describes the dependence between X and Y is known, its partial derivative with respect to one of the variables yields the conditional distribution, given that variable.

D. Quantile Regression Based on Copula Theory

Regression functions are the most widely used tools for describing multivariate relationship. For random variables X (representing TMI Tb) and Y (representing NSR from PR), the regression curve $y = E(Y/x)$ specifies a "typical" (mean) value of Y for each value of X and vice versa. However, $E(Y/x)$ and $E(X/y)$ are parametric and thus do not have simple expressions in terms of distribution functions and copulas [26]. Let X and Y be continuous random variables with joint distribution function $H_{X,Y}$ and marginal distribution functions F_X and G_Y with copula C. Then, $u = F_X(x)$ and $v = G_Y(y)$ are uniform (0, 1) random variables with joint distribution function C. If p denotes the pth quantile, then

$$\begin{aligned} p &= P[Y \le y/X = x] \\ &= P\left[v \le G(y)/u = F(x)\right] = \left.\frac{\partial C(u,v)}{\partial u}\right|_{\substack{u=F(x)\\ v=G(y)}} \end{aligned} \tag{13}$$

is used to find the pth quantile regression curve [52] $y_p = y_p(x)$ of Y on X. The above relations can be used to generate observations on Y simply by evaluating this expression and replacing p with different quantile values.

The steps to find the quantile regression curve using the Clayton copula are given as follows:

1) Fixing the conditional probability of Y given $X = x$ at some p so that

$$\frac{\partial C(u,v)}{\partial u} = p. = \left(1 + u^{\theta}(v^{-\theta} - 1)\right)^{\frac{1+\theta}{\theta}}$$

(using the function for the Clayton copula from Table II).

2) Solving for v, we have $v = ((p^{-\theta/(1+\theta)} - 1)u^{-\theta} + 1)^{-1/\theta}$, which will give different relationships between u and v for different values of p. Using this expression, we can obtain the conditional quantile function conditional on X as

$$y = F_Y^{-1}\left(\left[\left(p^{\frac{-\theta}{1+\theta}} - 1\right) F_X(x)^{-\theta} + 1\right]^{\frac{-1}{\theta}}\right). \tag{14}$$

It is worth noting that the proposed approach can be applied with any copula function with the expression for C varying with each copula type as shown in Table II. This enables to easily derive parametric families of conditional quantile functions from parametric copula functions [12]. Further details regarding copula theory and associated derivations can be obtained in [53]–[55].

V. Results

A. Sensitivity Analysis

For each of the JJAS months, the Spearman rank correlation coefficients were calculated between the 20 TMI channels and NSR of PR. Fisher's test was carried out to estimate the statistical significance between correlation differences. From Table III, it can be seen that, for stratiform rainfall, the channels of 21 V–37 V (for June and July) and 19 V–37 V (for August and September) were found to represent the highest correlation with NSR. On the other hand, the channels of 19 H–37 V (for June) and 19 V–37 V (for July, August, and September) were obtained as most sensitive to overland convective rainfall. It can be seen from Table III that, in comparison with the 85 V channel, these channels explain more variability in NSR. Statistical tests indicate that the difference between the correlation coefficients of 85 V and 19 V–37 V (21 V–37 V) is significant at 99% confidence level. The CC values of convective-type rainfall are 0.22, 0.35, 0.33, and 0.26 for JJAS. As scattering is primarily caused by frozen ice hydrometeors aloft, the emitted signal by liquid rain drops gets substantially blocked due to intense scattering. Hence, the measured Tb values are indirectly related to rain mass instead of rainfall below the cloud base. Therefore, any attempt to correlate Tb values with heavy rainfall below the cloud base results in low correlation values as observed for the study region. In the case of stratiform rainfall, the corresponding CC values are quite higher, i.e., 0.65, 0.53, 0.58, and 0.59, indicating good correlation. The high correlations of 19 V–37 V (21 V–37 V) channels suggest that, over the basin, the vertical distribution of hydrometeors is dominated by the bottom heavy liquid water which is more directly related to NSR. Another plausible reason for better correlation is the benefit offered by the subtraction of Tb from two channels, thereby reducing the uncertainty induced by surface emissivity variation. Moreover, the change in land surface emissivity (owing to surface wetness) has less impact on 19 V–37 V as the emissivity for 19 V and 37 V channels varies in a similar fashion. The channel combinations of 19 V, 37 V, and 22 V can be related with the differential optical depth of the atmosphere and have potential to represent the response of hydrometeors of different kinds within a column of the atmosphere [11]. This indicates that the channels of 19 V–37 V, 21 V–37 V, and 19 H–37 V not only contain information about rainfall wherein liquid water drops are the dominant hydrometeors but also aid in surface noise reduction over a predominantly vegetative

TABLE III
RESULTS OF SENSITIVITY ANALYSIS OVER THE STUDY REGION

Serial No:	TMI Channel Combinations	CONVECTIVE				STRATIFORM			
		June	July	August	September	June	July	August	September
1.	**21V-37V**	0.19	0.33	0.30	0.22	**0.65**	**0.58**	0.57	0.57
2.	**19V-37V**	0.20	**0.35**	**0.33**	**0.26**	0.61	0.53	**0.59**	**0.60**
3.	21V-85V	0.16	0.30	0.29	0.22	0.42	0.47	0.49	0.52
4.	19V-85V	0.17	0.31	0.29	0.23	0.43	0.47	0.50	0.53
5.	19H-37H	0.20	0.33	0.31	0.24	0.54	0.39	0.49	0.55
6.	19H-85V	0.18	0.31	0.30	0.23	0.44	0.43	0.49	0.53
7.	19H-85H	0.18	0.31	0.30	0.24	0.45	0.45	0.48	0.52
8.	19V-85H	0.17	0.31	0.30	0.23	0.44	0.47	0.49	0.53
9.	10V-85V	0.17	0.33	0.28	0.24	0.37	0.30	0.45	0.51
10.	21V-85H	0.16	0.30	0.29	0.22	0.43	0.47	0.48	0.52
11.	10V-85H	0.17	0.29	0.28	0.24	0.38	0.32	0.45	0.51
12.	**19H-37V**	**0.22**	0.33	0.31	0.24	0.53	0.39	0.46	0.54
13.	37H-85V	0.15	0.28	0.28	0.21	0.33	0.40	0.42	0.48
14.	10V-37V	0.17	0.24	0.26	0.24	0.34	0.15	0.38	0.51
15.	37H-85H	0.15	0.29	0.28	0.22	0.35	0.42	0.42	0.49
16.	37V-85V	0.14	0.28	0.28	0.21	0.31	0.38	0.41	0.47
17.	85V	-0.19	-0.31	-0.30	-0.24	-0.45	-0.51	-0.50	-0.53
18.	10H-85H	0.17	0.27	0.27	0.22	0.35	0.19	0.35	0.50
19.	10H+85H	0.17	0.27	-0.30	0.22	0.33	0.17	0.34	0.50
20.	85H+85V	-0.19	-0.31	0.29	-0.24	-0.45	-0.51	-0.50	-0.53

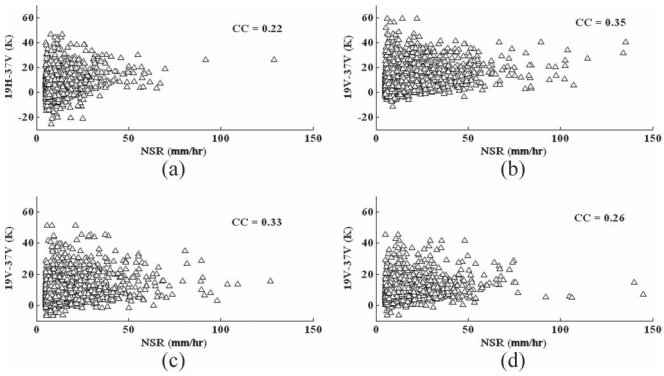

Fig. 3. Relationship of near surface rain rate for convective rain type with respect to sensitive channel combinations for intraseasonal months. (a) June. (b) July. (c) August. (d) September.

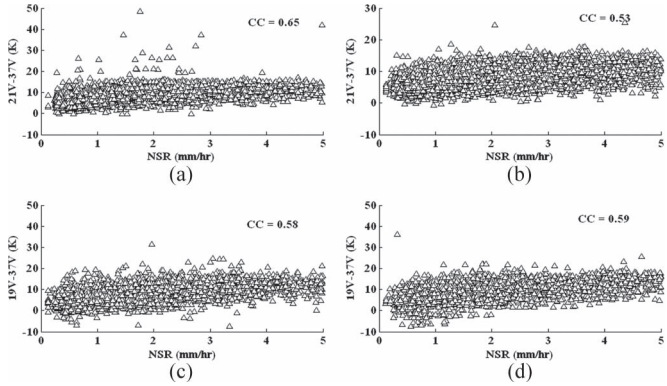

Fig. 4. Relationship of near surface rain rate for stratiform rain type with respect to sensitive channel combinations for intraseasonal months. (a) June. (b) July. (c) August. (d) September.

land surface such as that observed in the basin. Monsoon over Mahanadi basin sets in around the first week of June. Compared to the other monsoonal months, June receives comparatively less amount of rainfall. Of the total representative data for the month of June, nearly 25% constitute convective rainfall, and the rest comprise stratiform rainfall. The effect of sample size, along with the large number of outliers among the data pairs analyzed, is reflected in the rank correlation values for the rainfall types during June.

The variation of rainfall signature at the 19 V–37 V (or 21 V–37 V) channel has not been well assessed so far. From the scatter diagrams of Figs. 3 and 4, it can be observed that the data

TABLE IV
RESULTS OF COPULA MODELING FOR RAIN TYPE WITH SENSITIVE CHANNEL COMBINATION

RAINFALL TYPE	MONTH	TMI CHANNEL COMBINATION	FRANK			CLAYTON			GUMBEL		
			θ	AIC	BIC	θ	AIC	BIC	θ	AIC	BIC
CONVECTIVE	JUNE	19H-37V	1.42	-386.81	-379.60	0.28	-571.17	-563.96	**1.16**	**-640.26**	**-633.05**
	JULY	19V-37V	2.33	-1286.38	-1279.17	**0.52**	**-1708.10**	**-1700.89**	1.27	-1402.14	-1394.93
	AUGUST	19V-37V	1.98	-1068.78	-1061.57	**0.48**	**-1329.09**	**-1321.88**	1.22	-1017.62	-1010.41
	SEPTEMBER	19V-37V	1.69	-732.31	-725.10	**0.40**	**-1132.34**	**-1125.13**	1.18	-798.84	-791.63
STRATIFORM	JUNE	21V-37V	4.12	-3643.38	-3636.17	**1.01**	**-4278.35**	**-4271.14**	1.51	-3527.27	-3520.06
	JULY	21V-37V	3.87	-3320.84	-3313.63	0.81	-3011.57	-3004.36	**1.52**	**-3628.33**	**-3621.12**
	AUGUST	19V-37V	4.3	-3915.26	-3908.05	0.88	-3154.86	-3147.65	**1.60**	**-4043.98**	**-4036.77**
	SEPTEMBER	19V-37V	4.51	-4633.91	-4626.70	0.91	-3342.06	-3334.85	**1.644**	**-4697.32**	**-4690.11**

points tend to be highly populated at the lower end of rainfall for the convective rain type, indicating that different values of Tb can be associated with the same rain rates, depending on rainfall inhomogeneity. Several factors contribute to this uncertainty like the spatial resolutions of low-frequency TMI channels and the difference in viewing angles (significant for convective rainfall) between TMI and PR, to name a few. Also, for both the rainfall types, Tb values less than zero can be observed (from Figs. 3 and 4), indicative of a decrease in emissivity over tropical land caused by wet ground. The dependence structure of these channels with respect to rainfall intensities is potentially very useful in improving rainfall retrieval over land regions. However, scatter diagrams suggest that a simple relationship based on regression techniques may not represent this dependence due to the inherent uncertainties. Hence, in this study, we focus on explaining the variability of 19 V–37 V and 21 V–37 V channels with respect to NSR using copula theory.

B. Copula-Based Simulation

The first step for copula-based dependence modeling consists of fitting an appropriate marginal distribution to TMI Tb and NSR. Several parametric distributions can be used for fitting the marginals. For this study, to allow maximum flexibility in choosing the appropriate distribution, marginals of TMI Tb and NSR were modeled using a nonparametric kernel density function. Among the different dependence structures, three types of Archimedean copulas, namely, Clayton, Frank, and Gumbel, were used for bivariate modeling. The parameters and associated goodness of fit measures for different TMI Tb-NSR combinations and copula types are tabulated in Table IV. For the majority of JJAS months, the Gumbel copula is found to better describe the joint density between TMI channels and NSR when rainfall is stratiform in nature. The choice of the Gumbel copula indicates that there is strong right tail dependence in modeling Tb (of TMI channels) and stratiform rainfall (from PR) for the months of July, August, and September. It can be observed from the scatterplots [see Fig. 5(b)–(d)] that there is large clustering of Tb values for stratiform rainfall < 1.5 mm/h, indicative of uncertainty at low intensities of stratiform rainfall. For the convective rain regime, the Clayton copula was found suitable for the majority of JJAS months. This is indicative of the fact that correlation between both the variables (TMI Tb and PR rainfall) is strongest in the left tail of the joint distribution. It can be inferred that comparatively greater ambiguity exists in modeling convective rainfall. However, for the month of June, the copula family selected is different from that of the other monsoonal months for both the rainfall types. For convective rainfall of high intensities, lesser data points with comparatively low ambiguity are observed during June. Hence, the Gumbel copula, which is sensitive to the right tail of the joint distribution, was found to be suitable. For the stratiform rainfall of June, this situation is reversed wherein data points of lesser ambiguity are observed for rainfall intensities < 1 mm/h. Hence, the Clayton copula was found to appropriately represent the stratiform rainfall of June, owing to its sensitivity to the left tail of the joint distribution. For clarity in understanding the models, the pdf generated using Frank, Clayton, and Gumbel copula families for the month of June during stratiform rainfall is shown in Fig. 5.

A plausible way to account for uncertainties associated with any hydrological variable is to generate an ensemble of realizations that represent possible variability in them [27], [28].

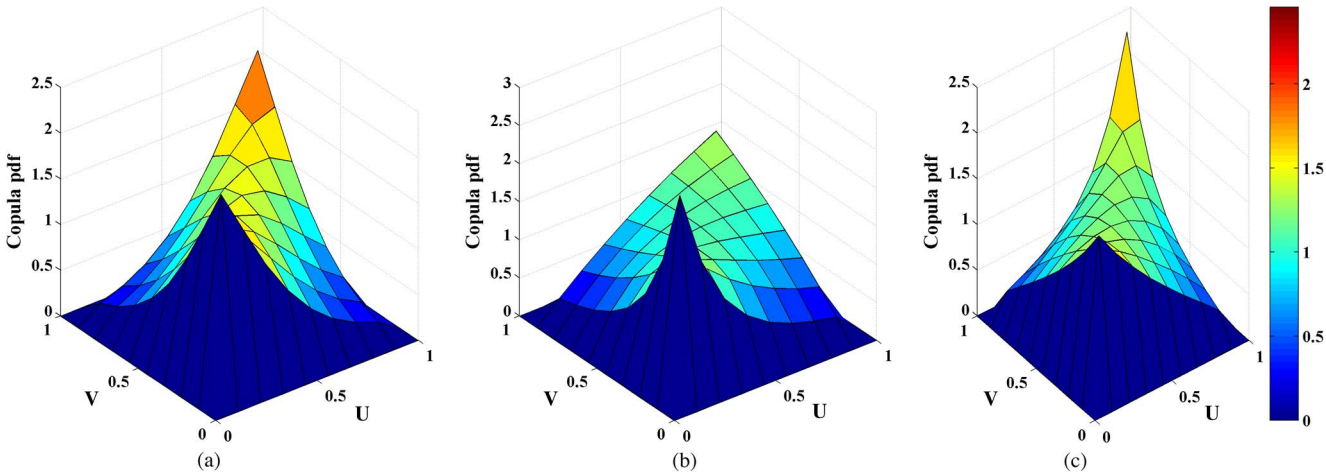

Fig. 5. PDF for the month of June generated using copula theory showing (a) Frank copula (4.12), (b) Clayton copula (1.01), and (c) Gumbel copula (1.51).

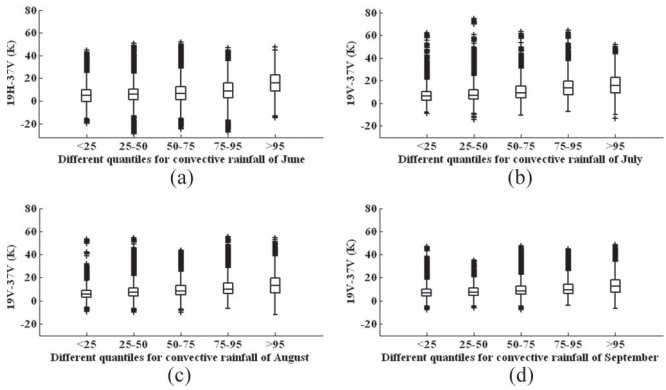

Fig. 6. Box plots of simulated Tb of sensitive channels for different quantiles of convective rainfall for (a) June, (b) July, (c) August, and (d) September.

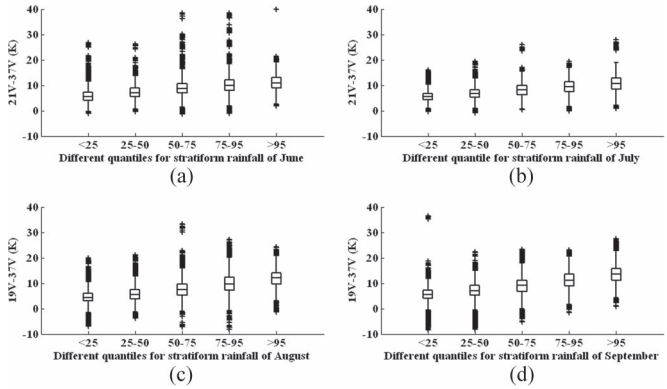

Fig. 7. Box plots of simulated Tb of sensitive channels for different quantiles of stratiform rainfall for (a) June, (b) July, (c) August, and (d) September.

In this study, the ensembles of TMI Tb were examined for different classes of rainfall (convective and stratiform) and plotted using box plots. The five different rainfall classes used in this study are based on five quantile classes of the convective and the stratiform rainfall, namely: 1) $<$ 25th; 2) 25th–50th; 3) 50th–75th; 4) 75th–95th; and 5) $>$ 95th. The data points (i.e., Tb) lying in each of these classes were simulated in sufficiently large numbers ($>$ 10 000) using the best fitting copula for each of the JJAS months. The results are shown in Figs. 6 and 7 from which it can be observed that the TMI Tb values for various channel combinations show a large number of outliers (i.e., values placed far away from the rest of the distribution) for various classes of rainfall. Positive values indicate that quantitatively larger Tb values have been registered over the vegetative background of Mahanadi basin for channels of 19 V, 19 H, and 21 V values (during the convective and the stratiform rainfall) when compared with 37 V Tb values. It can be inferred that the presence of frozen and liquid hydrometeors in the atmosphere over the study region causes scattering in the 37 V channel, thereby depressing (reducing) the Tb value registered by the TMI sensor for the 37-GHz channel. Over a vegetative background, such as that observed for the study region considered, the channels of 19 V, 19 H, and 21 V are known to be less affected by scattering and more representative of the emissivity from vegetative land surface. Hence, the values of 19 V–37 V (19 H–37 V and 21 V–37 V) result in positive values, with larger values indicating more scattering in the 37 V channel.

For convective rainfall, a large amount of negative values was observed. A possible explanation to this is the depression due to ice scattering at the 37 V channel. This is highlighted in the range of values for June (up to 40 K). For stratiform rainfall during the months of August and September, the range of Tb includes very few negative values, indicative of atmospheric liquid hydrometeors over the basin. For both the rain types, box plots show a significant overlap in the interquartile ranges for each rainfall class. This indicates that the dependence of TMI Tb with rainfall (NSR) is associated with uncertainty. The quantification of this uncertainty will not only give information about the microphysical characteristics of rainfall over the basin but also assist in hydrological and meteorological applications.

To aid in uncertainty quantification, this study generates ensembles of the convective and the stratiform rainfall for each of the aforementioned classes (based on quantile levels) using conditional distribution based on copula theory. In other words, given the TMI Tb (K), the range of rainfall (mm/h) for different quantile classes was simulated using the best fitting copula family for each of JJAS months using a three-year (2009–2011)

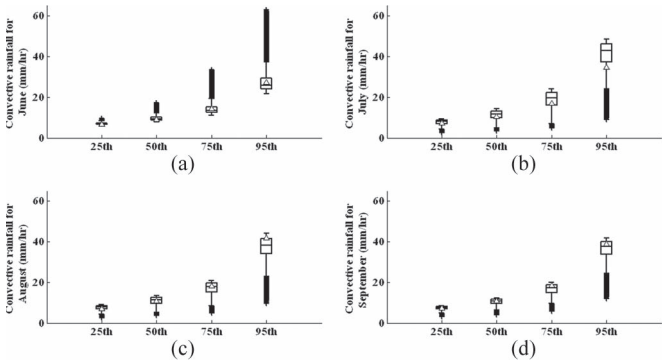

Fig. 8. Box plots of simulated values of convective rainfall for each of the 25th, 50th, 75th, and 95th quantiles for (a) June, (b) July, (c) August, and (d) September. △ indicates the observed quantile values of JJAS in 2012.

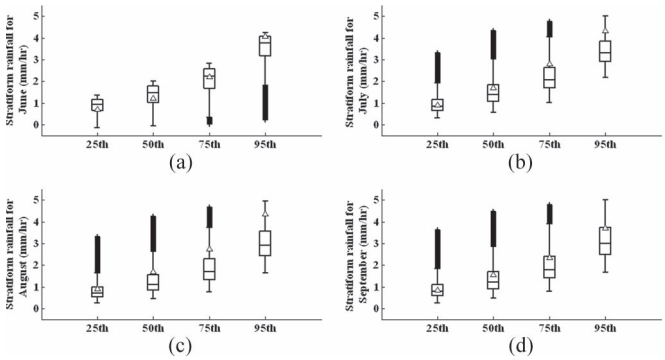

Fig. 9. Box plots of simulated values of stratiform rainfall for each of the 25th, 50th, 75th, and 95th quantiles for (a) June, (b) July, (c) August, and (d) September. △ indicates the observed quantile values of JJAS in 2012.

collocated database. In Figs. 8 and 9, the outliers in the form of high (low) values of stratiform (convective) rainfall can be attributed to the copula family selected which is Gumbel (Clayton). From Figs. 8 and 9, it can be observed that, for the month of June, the box plots depict outliers toward higher extreme (lower extreme) for convective (stratiform) rainfall. This is attributed to the selection of the Gumbel (for convective) and the Clayton (for stratiform) copula, respectively. The rainfall quantiles observed for year 2012 (shown as Δ in Figs. 8 and 9) are plotted along with the simulated quantiles generated for rainfall classes using the best fitting copula for each month. It can be seen that the observed quantiles fall well within the predicted range of their population for both convective and stratiform rainfall regimes. It can be inferred that, in spite of the large uncertainties observed in TMI Tb channels, the proposed method can be efficiently used to generate precipitation ensembles pertaining to different quantile levels for both the convective and the stratiform rainfall.

C. Copula-Based Quantile Regression

Quantile regression technique based on copula theory provides the conditional expectation of rainfall values at different quantiles, given the value of Tb. Hence, in this study, copula theory was used to generate regression curves between TMI Tb and NSR. Collocated data of three years (2009–2011)

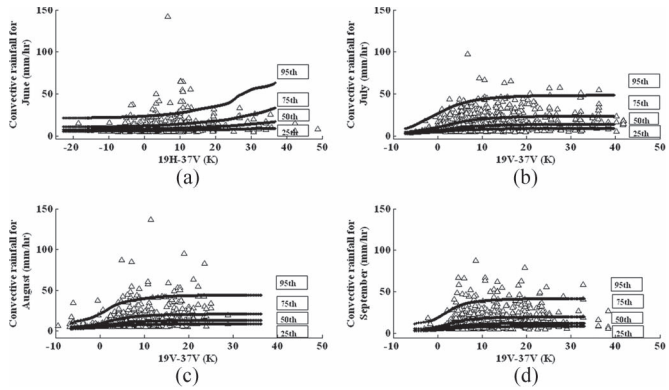

Fig. 10. Copula-based quantile regression curves for various quantiles of convective rain rate simulated and overlain on 2012 data of (a) June, (b) July, (c) August, and (d) September.

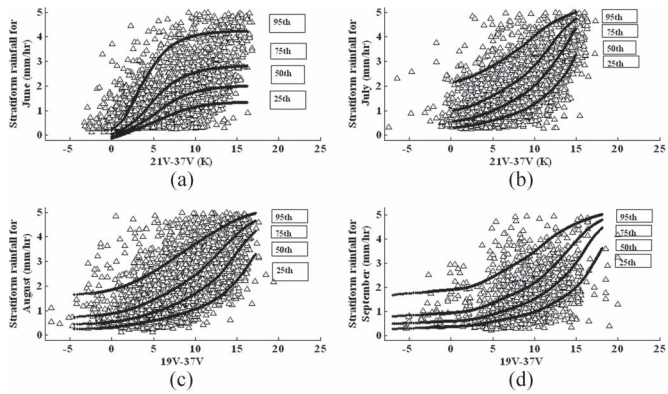

Fig. 11. Copula-based quantile regression curves for various quantiles simulated and overlain on 2012 data of (a) June, (b) July, (c) August, and (d) September.

were used to create the database of rainfall values. The curves generated for different quantile levels using this database were overlain on the JJAS data of 2012 for visual comparison (see Figs. 10 and 11). For stratiform rainfall types, the regression curves for each quantile show a steep increase with the increase in the value of channel combination (19 V–37 V and 21 V–37 V). The relation for convective rainfall increases steeply with the initial increase in Tb values, and after reaching the peak point, the curves fail to show any relation. There is relatively lower degree of association between the channels and extremes of convective precipitation. This stresses the need for modeling the extremes of convective rainfall. A clear dependence structure is observed for the quantile curves of stratiform rainfall. A steep increase in TMI Tb is observed for every unit increase in the rain rate, indicating a very strong dependence structure.

D. Model Comparison With Linear and Quadratic

Traditional methods to generate Tb-NSR relations rely on linear and quadratic models. Recently, Gopalan *et al.* [7] have modeled the relationships between NSR from PR and 85 V Tb globally and came out with the result that a linear model best describes the 85 V Tb–NSR relationship for stratiform rainfall whereas a cubic model best represents the relation

TABLE V
ERROR METRICS OBTAINED USING DIFFERENT MODELS

RAINFALL TYPE	MONTH	MAE			MSE			RMSE			MARE			MAPE		
		A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
CONVECTIVE	JUNE	4.47	4.43	**2.52**	32.33	31.82	**8.49**	5.68	5.64	**2.91**	0.32	0.32	**0.22**	32.25	32.42	**22.40**
	JULY	5.16	5.39	**5.04**	41.59	41.30	**34.71**	6.44	6.42	**5.89**	0.37	0.38	**0.35**	37.59	38.37	**35.05**
	AUGUST	5.34	5.30	**3.72**	46.40	48.87	**18.29**	6.81	6.99	**4.27**	0.36	0.35	**0.28**	36.36	35.37	**28.20**
	SEPTEMBER	7.01	6.93	**5.51**	87.67	87.85	**49.59**	9.36	9.37	**7.04**	0.39	0.38	**0.31**	39.49	38.96	**31.49**
STRATIFORM	JUNE	0.67	**0.60**	0.64	0.93	**0.83**	0.87	0.96	**0.91**	0.93	0.29	**0.24**	0.26	29.82	**24.04**	26.34
	JULY	**0.65**	0.66	0.69	0.78	**0.75**	0.77	0.88	**0.86**	0.88	**0.25**	0.26	0.27	**25.94**	26.45	27.64
	AUGUST	0.66	0.65	**0.62**	0.82	0.77	**0.63**	0.90	0.87	**0.79**	0.22	0.22	**0.22**	22.45	22.60	**22.31**
	SEPTEMBER	0.46	0.41	**0.37**	0.35	0.28	**0.19**	0.59	0.53	**0.44**	0.24	0.21	**0.20**	24.79	21.72	**20.22**

A= Linear; B = Quadratic ; C= Copula
MAE=mean absolute error; MSE=mean square error; RMSE= root mean square error; MARE= mean absolute relative error; MAPE=mean absolute percentage error

during convective rainfall. Hence, regression-based relations were generated between TMI 85 V Tb and NSR data from PR for a period of three years (2009–2011). A quadratic model was used for convective rainfall while a linear model was used for stratiform rainfall as shown in the following:

$$\mathrm{NSR}_{\mathrm{Convective}} = a + b * \mathrm{Tb}_{85\,\mathrm{V}} + c * \mathrm{Tb}_{85\,\mathrm{V}}^2 \quad (15)$$

$$\mathrm{NSR}_{\mathrm{Stratiform}} = d + e * \mathrm{Tb}_{85\,\mathrm{V}}. \quad (16)$$

Here, $\mathrm{NSR}_{\mathrm{Convective}}$ and $\mathrm{NSR}_{\mathrm{Stratiform}}$ represent the rainfall rates for convective and stratiform data points, respectively, and $\mathrm{Tb}_{85\,\mathrm{V}}$ denotes the brightness temperature of the 85-GHz vertically polarized channel. It may be noted that the coefficients (a, b, c, d, and e) have different values for each rainfall type and during each of the JJAS monsoonal months. A quantitative assessment of the proposed copula-based model is conducted by comparing against the developed linear and quadratic models for each of the JJAS months. Quantiles were generated for each of the monsoonal months for both convective and stratiform rain types. The quantiles predicted using the proposed approach were compared with those generated from linear and quadratic models. Performance evaluation was conducted on the key quantile measures (25th, 50th, 75th, and 95th). Finally, performance statistics were quantified using a set of error metrics. The results are tabulated in Table V. It can be inferred from Table V that, for convective rainfall, the copula-based approach provides the least error when compared with linear/quadratic models, indicating superior performance. Even though the least number of data pairs was observed for the month of June for both the rainfall types, for the convective rain type, the copula-based approach shows superior performance, suggesting that the proposed approach has successfully modeled using a good representation of data pairs. However, for the months of June and July having stratiform rainfall, conventional models (linear and quadratic) seem to perform slightly better, even though a quantitative evaluation of errors does not show much difference in the values. This might be partly attributed to the fact that the Clayton and Gumbel copulas (selected using a three-year data period) for the June and July months have slightly fallen short in capturing new data pairs (of TMI Tb and NSR) observed in 2012. This can be attributed to the property of copulas which considers dependence structure to be constant with time [55]. Future works can be conducted allowing copulas to be time varying in nature, even though such study is in its nascent stage of development [56], [57]. However, for all the other months, the proposed copula model shows better performance when compared to the conventional models. Moreover, it can be concluded that comparatively greater ambiguity in modeling convective rainfall can be tackled by using the proposed approach.

VI. CONCLUSION

This paper has analyzed the relationship between various PMW channel frequencies of TMI with NSR from active radar (PR) using collocated version 7 orbital data products from TRMM, namely, 1B11, 2A23, 2A25, and 2A21. A new technique has been developed based on copula theory to study the dependence of TMI frequency channels with respect to rainfall types (convective and stratiform) for the land regions of Mahanadi basin, India. The present scheme conducts sensitivity

analysis to estimate the TMI channel combinations which are most sensitive to PR NSR using the robust Spearman rank correlation. Results reveal a greater sensitivity of 19 V–37 V and 21 V–37 V channel combinations to model rainfall regimes for the basin. It can be inferred that these channel combinations are more directly related to liquid rain drops and can help in characterizing the microphysical structure of hydrometeors over the basin. Furthermore, we have modeled the highly nonlinear transfer function relating sensitive TMI channel combinations with PR rainfall using copula theory. Archimedean copula-based modeling suggests that the Clayton and Gumbel copulas are well suited to represent the bivariate joint distributions of the convective and the stratiform rainfall with respect to TMI Tb for the majority of the intraseasonal months. However, for the months of June and July, conventional methods seem to perform slightly better during stratiform rainfall. The most suitable copula family for TMI channels and rainfall regimes might change from one region to another due to differences in geographical and geophysical conditions. This study is only based on data for a single region. Our approach, however, can be applied to studies in other parts of the world to select the most appropriate copula model.

Once the best fitting copula has been arrived at, we used this to simulate large realizations of channel combinations conditional on different rainfall quantile ranges ($<$ 25th, 25th–50th, 50th–75th, 75th–95th, and $>$ 95th). The results, plotted as box plots, show heavy overlap for the interquartile ranges, indicating that several values of Tb correspond to the same quantile range of rainfall. Comparatively greater ambiguity was observed to model extreme values of the convective rain type. This stresses the need for uncertainty modeling. Finally, the efficiency of the model developed was tested by comparing the results with traditionally employed linear and quadratic models. Results based on the comparison of different error metrics reveal the superior performance of the copula-based technique for the majority of the JJAS months.

The TMI land rainfall algorithm fails to detect "warm rainfall" over land due to the lack of significant ice scattering in such rainfall [7]. From this study, it can be inferred that a combination of low-resolution TMI channels successfully detects rainfall wherein liquid drops are the dominant hydrometeors. It also aids in uncertainty quantification which is helpful in hydrological and meteorological applications. Furthermore, the database of rainfall quantiles generated using the copula-based approach can be highly beneficial in rainfall modeling studies and weather applications over the basin.

ACKNOWLEDGMENT

The authors would like to thank the Goddard Distributed Active Archive Center (Goddard Earth Sciences Data and Information Services Center Distributed Active Archive Center) for providing the Tropical Rainfall Measuring Mission science data products and the two anonymous reviewers for their valuable suggestions. The first author would like to thank Dr. K. Gopalan for providing fruitful discussions for the initial part of this study. The authors would also like to thank Dr. W. Asquith for making available the copBasic package in R.

REFERENCES

[1] T. T. Wilheit, "Some comments on passive microwave measurement of rain," *Bull. Amer. Meteor. Soc.*, vol. 67, no. 10, pp. 1226–1232, Oct. 1986.
[2] R. W. Spencer, H. M. Goodman, and R. E. Hood, "Precipitation retrieval over land and ocean with the SSM/I, Part I: Identification and characteristics of the scattering signal," *J. Atmos. Ocean. Technol.*, vol. 6, no. 2, pp. 254–273, Apr. 1989.
[3] C. Kidd, "Passive microwave rainfall monitoring using polarization corrected temperatures," *Int. J. Remote Sens.*, 1998.
[4] N. Grody, "Classification of snow cover and precipitation using the special sensor microwave imager," *J. Geophys. Res.*, vol. 96, no. D4, pp. 7423–7435, Apr. 1991.
[5] R. Ferraro and G. F. Marks, "The development of SSM/I rain rate retrieval algorithms using ground based radar measurements," *J. Atmos. Ocean. Technol.*, vol. 12, pp. 755–770, 1995.
[6] N. Y. Wang, R. Ferraro, E. Zipser, and C. Kummerow, "TRMM 2A12 land precipitation product status and future plans," *J. Meteor. Soc. Japan*, vol. 87A, pp. 237–253, 2009.
[7] K. Gopalan, N. Wang, R. Ferraro, and C. Liu, "Status of the TRMM 2A12 land precipitation algorithm," *J. Atmos. Oceanic Technol.*, vol. 27, no. 8, pp. 1343–1354, Aug. 2010.
[8] R. R. Ferraro, N. C. Grody, and G. G. Marks, "Effects of surface conditions on rain identification using the DMSP-SSM/I," *Remote Sens. Rev*, vol. 11, pp. 195–209, 1994.
[9] R. R. Ferraro and G. F. Marks, "The development of SSM/I rain-rate retrieval algorithms using ground based radar measurements," *J. Atmos. Oceanic Technol.*, vol. 12, no. 4, pp. 755–770, Aug. 1995.
[10] G. Liu and J. A. Curry, "Retrieval of precipitation from satellite microwave measurements using both emission and scattering," *J. Geophys. Res.*, vol. 97, no. D9, pp. 9959–9974, Jun. 1992.
[11] C. Prabhakara, R. Iacovazzi, Jr., J.-M. Yoo, and K.-M. Kim, "A model for estimation of rain rate on tropical land from TRMM Microwave Imager radiometer observations," *J. Meteorol. Soc. Jpn.*, vol. 83, no. 4, pp. 595–609, Aug. 2005.
[12] D. K. Sarma, M. Konwar, J. Das, S. Pal, and S. Sharma, "A soft computing approach for rainfall retrieval from the TRMM Microwave Imager," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2879–2885, Dec. 2005.
[13] D. K. Sarma, M. Konwar, S. Sharma, S. Pal, J. Das, U. K. De, and G. Viswanathan, "An artificial neural network based integrated regional model for rain retrieval over land and ocean," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1689–1696, Jun. 2008.
[14] T. Dinku and E. Anagnostou, "Regional differences in overland rainfall estimated from PR-calibrated TMI algorithm," *J. Appl. Meteor.*, vol. 44, no. 2, pp. 189–205, Feb. 2005.
[15] K. Aonashi, J. Awaka, M. Hirose, T. Kozu, T. Kubota, G. Liu, S. Shige, S. Kida, S. Seto, N. Takahashi, and Y. N. Takayabu, "GSMaP passive microwave precipitation retrieval algorithm: Algorithm description and validation," *J. Meteorol. Soc. Jpn.*, vol. 87A, pp. 119–136, Mar. 2009.
[16] A. Mugnai, H. J. Cooper, E. A. Smith, and G. J. Tripoli, "Simulation of microwave brightness temperatures of an evolving hailstorm at SSM/I frequencies," *Bull. Amer. Meteor. Soc.*, vol. 71, no. 1, pp. 2–13, Jan. 1990.
[17] C. Kummerow, "The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors," *J. Appl. Meteorol.*, vol. 40, no. 11, pp. 1801–1820, Nov. 2001.
[18] R. R. Ferraro, E. A. Smith, W. Berg, and G. J. Huffman, "The tropical rainfall potential technique. Part II: Validation," *Wea. Forecasting.*, vol. 20, no. 4, pp. 465–475, Aug. 2005.
[19] G. J. Huffman, T. B. David, J. B. Eric, and B. W. David, "The TRMM Multi-satellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales," *J. Hydrometeorol.*, vol. 8, pp. 38–55, Feb. 2007.
[20] A. C. Favre, E. Adlouni, S. L. Perreault, T. Monge, and B. Bob, "Multivariate hydrological frequency analysis using copulas," *Water Resour. Res.*, vol. 40, no. 1, pp. W01101-1–W01101-12, Jan. 2004.
[21] R. Maity and D. N. Kumar, "Probabilistic prediction of hydroclimatic variables with nonparametric quantification of uncertainty," *J. Geophys. Res.*, vol. 113, no. D14, pp. D14105-1–D14105-12, Jul. 2008.
[22] A. AghaKouchak, A. Bárdossy, and E. Habib, "Copula-based uncertainty modeling: Application to multisensor precipitation estimates," *Hydrological Process.*, vol. 24, no. 15, pp. 2111–2124, Jul. 2010.
[23] C. Kummerow, W. Barnes, T. Kozu, J. Shiue, and J. Simpson, "The Tropical Rainfall Measuring Mission (TRMM) sensor package," *J. Atmos. Oceanic Technol.*, vol. 15, no. 3, pp. 809–817, Jun. 1998.
[24] N. Viltard, C. Burlaud, and C. Kummerow, "Rain retrieval from TMI brightness temperature measurements using a TRMM PR-based database," *J. Appl. Meteorol. Climatol.*, vol. 45, no. 3, pp. 455–466, Mar. 2006.

[25] A. Rapp, M. Lebsock, and C. Kummerow, "On the consequences of resampling microwave radiometer observations for use in the retrieval algorithm," *J. Appl. Meteorol. Climatol.*, vol. 48, no. 9, pp. 1981–1993, Sep. 2009.
[26] M. Grecu, W. S. Olson, and E. N. Anagnostou, "Retrieval of precipitation profiles from multiresolution, multifrequency active and passive microwave observations," *J. Appl. Meteor.*, vol. 43, no. 4, pp. 562–575, Apr. 2004.
[27] A. Mishra and R. Kumar, "Study of rainfall from TRMM Microwave Imager (TMI) observation over India," *ISRN Geophys.*, vol. 2012, pp. 1–7, 2012.
[28] Y. You, G. Liu, Y. Wang, and J. Cao, "On the sensitivity of Tropical Rainfall Measuring Mission (TRMM) Microwave Imager channels to overland rainfall," *J. Geophys. Res.*, vol. 116, no. D12, p. D12203, Jun. 2011.
[29] C. De Michele and G. Salvador, "A generalized pareto intensity-duration model of storm rainfall exploiting 2 copulas," *J. Geophys. Res.*, vol. 108, pp. 4067-1–4067-11, Jan. 2003.
[30] E. Bouye, A. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, "Copulas for finance—A reading guide and some applications," Groupe de Rech. Oper., Credit Lyonnais, Paris, France, 2000, Tech. Rep.
[31] E. W. Frees and E. A. Valdez, "Understanding relationships using copulas," *North Amer. Actuarial J.*, vol. 2, no. 1, pp. 1–25, 1998.
[32] L. Zhang and V. P. Singh, "Bivariate flood frequency analysis using the copula method," *J. Hydrologic Eng.*, vol. 11, no. 2, pp. 150–164, Mar. 2006.
[33] H. Gao, E. F. Wood, T. J. Jackson, M. Drusch, and R. Bindlish, "Using TRMM/TMI to retrieve surface soil moisture over the southern United States from 1998 to 2002," *J. Hydrometeor.*, vol. 7, no. 1, pp. 23–38, Feb. 2006.
[34] A. Sklar, *Fonctions de Repartition a n Dimensions et Leurs Marges*. Paris, France: Publ. Inst. Stat. Univ. Paris, 1959, pp. 229–231.
[35] R. B. Nelsen, *An Introduction to Copulas*, 2nd ed. New York, NY, USA: Springer-Verlag, 2006, p. 269.
[36] D. Bosq, *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction, Lecture Notes in Statistics*. New York, NY, USA: Springer-Verlag, 1998, p. 210.
[37] D. W. Scott, *Multivariate Density Estimation, Theory, Practice, and Visualization*. New York, NY, USA: Wiley, 1992.
[38] A. Sharma, "Seasonal to interseasonal rainfall probabilistic forecasts for improved water supply management: Part 3—A nonparametric probabilistic forecast model," *J. Hydrol.*, vol. 239, pp. 249–258, 2000.
[39] U. Lall, "Recent advances in nonparametric function estimation," *Rev. Geophys., Suppl.*, vol. 33, no. S2, pp. 1093–1102, Jul. 1995, U.S. Natl. Rep. Int. Union Geod. Geophys. 1991-1994.
[40] U. Lall, B. Rajagopalan, and D. G. Tarboton, "A nonparametric wet/dry spell model for resampling daily precipitation," *Water Resources Res.*, vol. 32, no. 9, pp. 2803–2823, Sep. 1996.
[41] D. G. Tarboton, A. Sharma, and U. Lall, "Disaggregation procedures for stochastic hydrology based on nonparametric density estimation," *Water Resources Res.*, vol. 34, no. 1, pp. 107–119, Jan. 1998.
[42] S. Mukhopadhyay, "A generic data-driven nonparametric framework for variability analysis of integrated circuits in nanometer technologies," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 28, no. 7, pp. 1038–1046, Jul. 2009.
[43] D. Dupuis, "Using copulas in hydrology: Benefits, cautions, and issues," *J. Hydrol. Eng.*, vol. 12, no. 4, pp. 381–393, 2007.
[44] C. H. Kimberling, "A probabilistic interpretation of complete monotonicity," *Aequationes Math.*, vol. 10, no. 2/3, pp. 152–164, 1974.
[45] E. J. Gumbel, "Bivariate exponential distributions," *J. Amer. Stat. Assoc.*, vol. 55, no. 292, pp. 698–707, Dec. 1960.
[46] R. B. Nelsen, *An Introduction to Copulas*. New York, NY, USA: Springer-Verlag, 1999.
[47] M. J. Frank, "On the simultaneous associativity of F(x, y) and x + y_F (x, y)," *Aequationes Math.*, vol. 19, pp. 194–226, 1979.
[48] R. B. Nelsen, "Properties of a one-parameter family of bivariate distributions with specified marginals," *Commun. Stat.-Theory Methods*, vol. 15, no. 11, pp. 3277–3285, 1986.
[49] C. Genest, "Frank's family of bivariate distributions," *Biometrica*, vol. 74, no. 3, pp. 549–555, Sep. 1987.
[50] H. Joe, *Multivariate Models and Dependence Concepts*. New York, NY, USA: Chapman & Hall, 1997.
[51] G. G. Venter, "Tails of copulas," in *Proc. Casualty Actuarial Soc.*, 2002, pp. 68–113, LXXXIX.
[52] R. Koenker, *Quantile Regression*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[53] C. Genest and J. MacKay, "The joy of copulas: Bivariate distributions with uniform marginals," *Am. Stat.*, vol. 40, no. 4, pp. 280–283, Nov. 1986.
[54] C. Genest and L. P. Rivest, "Statistical inference procedure for bivariate Archimedean copulas," *J. Amer. Stat. Assoc.*, vol. 88, no. 423, pp. 1034–1043, Sep. 1993.
[55] A. Patton, "Estimation of multivariate models for time series of possibly different lengths," *J. Appl. Econom.*, vol. 21, no. 2, pp. 147–173, Mar. 2006.
[56] L. Chollete, A. Heinen, and A. Valdesogo, "Modeling international financial returns with a multivariate regime switching copula," *J. Financial Econom.*, vol. 7, no. 4, pp. 437–480, 2009.
[57] C. M. Hafner and H. Manner, "Dynamic stochastic copula models: Estimation, inference and applications," *J. Appl. Econom.*, vol. 27, no. 2, pp. 269–295, Mar. 2012, forthcoming.

J. Indu received the B.Tech. degree in civil engineering from the Mar Athanasius College of Engineering, Kerala, India, in 2004 with university third rank and the M.Tech. degree in geoinformatics from Indian Institute of Technology, Kanpur, Uttar Pradesh, India, in 2008. She is currently working toward the Ph.D. degree in the Department of Civil Engineering, Indian Institute of Science, Bangalore, India.

Her current research interests include microwave remote sensing, uncertainty modeling, and nowcasting of precipitation.

D. Nagesh Kumar received the Ph.D. degree from the Department of Civil Engineering, Indian Institute of Science, Bangalore, India, in 1992.

He worked as a Boyscast Fellow at the Utah Water Research Laboratory, Utah State University, USA, in 1999. He has been a Professor with the Department of Civil Engineering, Indian Institute of Science, Bangalore, India, since May 2002. Earlier, he worked in Indian Institute of Technology, Kharagpur, West Bengal, India, and National Remote Sensing Centre, Hyderabad, India. His research interests include climate hydrology, climate change, water resources systems, artificial neural network, evolutionary algorithms, fuzzy logic, multiple criteria decision making, and remote sensing and geographic information system applications in water resources engineering. He has coauthored two text books titled "Multicriterion Analysis in Engineering and Management" published by Prentice Hall of India, New Delhi, and "Floods in a Changing Climate: Hydrologic Modeling," published by Cambridge University Press, U.K. (home page: http://civil.iisc.ernet.in/~nagesh/).