Atmospheric Research 197 (2017) 1–20

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

Invited review article

Global precipitation measurements for validating climate models

F.J. Tapiador[a,*], A. Navarro[a], V. Levizzani[b], E. García-Ortega[c], G.J. Huffman[d], C. Kidd[d,e], P.A. Kucera[f], C.D. Kummerow[g], H. Masunaga[h], W.A. Petersen[i], R. Roca[j], J.-L. Sánchez[c], W.-K. Tao[d], F.J. Turk[k]

[a] University of Castilla-La Mancha (UCLM), Department of Environmental Sciences, Institute of Environmental Sciences, Toledo, Spain
[b] National Council of Research, Institute of Atmospheric Sciences and Climate (CNR-ISAC), Bologna, Italy
[c] Institute of Environment, University of León, Spain
[d] NASA-Goddard Space Flight Center, Greenbelt, MD, USA
[e] University of Maryland, College Park, MD, USA
[f] National Center for Atmospheric Research, Boulder, CO, USA
[g] Colorado State University, Ft. Collins, CO, USA
[h] Institute for Space-Earth Environmental Research, Nagoya University, Japan
[i] NASA-Marshall Space Flight Center, Huntsville, AL, USA
[j] OMP/LEGOS, Toulouse, France
[k] NASA-Jet Propulsion Laboratory, Pasadena, CA, USA

ABSTRACT

The advent of global precipitation data sets with increasing temporal span has made it possible to use them for validating climate models. In order to fulfill the requirement of global coverage, existing products integrate satellite-derived retrievals from many sensors with direct ground observations (gauges, disdrometers, radars), which are used as reference for the satellites. While the resulting product can be deemed as the best-available source of quality validation data, awareness of the limitations of such data sets is important to avoid extracting wrong or unsubstantiated conclusions when assessing climate model abilities. This paper provides guidance on the use of precipitation data sets for climate research, including model validation and verification for improving physical parameterizations. The strengths and limitations of the data sets for climate modeling applications are presented, and a protocol for quality assurance of both observational databases and models is discussed. The

Abbreviations: 4D, Four-dimensional; 1DD, GPCP One-Degree Daily; AIRS, Atmospheric Infrared Sounder; ANN, Artificial neural networks; APHRODITE, Asian Precipitation - Highly-Resolved Observational Data Integration Towards Evaluation; AR5, IPCC Fifth Assessment Report; ARM, Atmospheric Radiation Measurement; ATBD, Algorithm Theoretical Basis Document; BIAS, Bristol-NOAA InterActive Scheme; C3VP, Canadian CloudSat/CALIPSO Validation Program; CaPPM, Cloud and Precipitation Processes Mission; CDR, Climate Data Records; CESM, Community Earth System Model; CFAD, Contoured-Frequency-by-Altitude Diagram; CFSR, Climate Forecast System Reanalysis; CHIRPS, Climate Hazards group Infrared Precipitation with Station; CMAP, CPC Merged Analysis of Precipitation; CMORPH, CPC MORPHing technique; CORDEX, Coordinated Regional Climate Downscaling Experiment; COSMO, Consortium for Small-scale Modeling; CPC, Climate Prediction Center; CRM, Cloud Resolving Model; CRU, Climate Research Unit; CSH, Convective-Stratiform Heating; CST, Convective-Stratiform Technique; DPR, Dual-frequency Precipitation Radar; DSD, Drop Size Distribution; ECA, European Climate Assessment; ENSO, El Nino-Southern Oscillation; ERA-I, European Centre for Medium-Range Weather Forecast (ECMWF) interim reanalysis; GCE, Goddard Cumulus Ensemble; GCM, General Circulation Model/Global Climate Model; G-CRM, Global Cloud Resolving Model; GHCN, Global Historical Climatology Network; GMI, GPM Microwave Imager; GPCC, Global Precipitation Climatology Centre; GPCP, Global Precipitation Climatology Project; GPI, Global Precipitation Index; GPM, Global Precipitation Measurement mission; GPROF, Goddard profiling algorithm; G-SDSU, Goddard Satellite Data Simulator Unit; GSMAP, Global Satellite Map Product; GV, Ground validation; HH, Hydrometeor Heating; HR-GCM, High Resolution General Circulation Model; HRPP, High-Resolution Precipitation Products; iFloodS, Iowa Flood Studies; IMERG, Integrated Multi-satellitE Retrievals for GPM; IPCC, Intergovernmental Panel on Climate Change; ipHex, Integrated Precipitation and Hydrology Experiment; IPSL, Institut Pierre Simon Laplace; IR, Infrared radiation; ITCZ, Intertropical Convergence Zone; JCR, Journal Citation Reports; JRA25, Japanese 25-year Reanalysis; LH, Latent Heat; LMDZ, Laboratoire de Météorologie Dynamique Zoom; MERRA, Modern-Era Retrospective analysis for Research and Applications; MP, Microphysics; NRL, Naval Research Laboratory; NU-WRF, NASA-Unified Weather Research and Forecasting; OLYMPEX, Olympic Mountain Experiment; PBL, Planetary Boundary Layer; PDF, Probability distribution function; PERSIANN, Precipitation Estimation from Remote-Sensed Information using ANN; PMIR, Passive Microwave-InfraRed; PMW, Passive Microwave; PR, Precipitation Radar; PRESTORM, Preliminary Regional Experiment for STORM-Central; PRH, Precipitation Radar Heating; QA, Quality Assurance; QC, Quality Control; RAMS, Regional Atmospheric Modeling System; RCM, Regional Climate Model; REFAME, Rain Estimation using Forward Adjusted-advection of Microwave Estimates; RSS, Remote Sensing Systems; SBM, Spectral bin microphysics; SG, Satellite-Gauge; SLH, Spectral Latent Heating; SSM/I, Special Sensor Microwave Imager; SSMIS, Special Sensor Microwave Imager/Sounder; TAMSAT, Tropical Applications of Meteorology using SATellite data and ground-based observations; TMI, TRMM Microwave Imager; TMPA, TRMM Multi-Satellite Precipitation Analysis; TMPI, Threshold-Matched Precipitation Index; TOVS, TIROS Operational Vertical Sounder; TRMM, Tropical Rainfall Measuring Mission; VIS, Visible; WRF-SBM, Weather Research and Forecasting – Spectral Bin Microphysics

* Corresponding author.
E-mail address: Francisco.Tapiador@uclm.es (F.J. Tapiador).

http://dx.doi.org/10.1016/j.atmosres.2017.06.021
Received 7 April 2017; Received in revised form 21 May 2017; Accepted 19 June 2017
Available online 20 June 2017

paper helps elaborating the recent IPCC AR5 acknowledgment of large observational uncertainties in precipitation observations for climate model validation.

1. Introduction

Precipitation is a major element in the Earth's hydrological cycle and at the same time is tied dynamically to the atmospheric circulation by redistributing the latent heating through the troposphere. Precipitation thus serves as a critical linkage between the global water and energy cycles.

Among the key questions in this outstanding research topic is how much global precipitation has been changing over time in association with global warming. It is known that precipitable water has increased with temperature nearly as rapidly as predicted from the Clausius-Clapeyron eq. (6–7% K^{-1}), while the rate of global precipitation change is expected to be only a few % K^{-1} at best (e.g., Allen and Ingram, 2002). The lower amplitude increase in precipitation may be understood in terms of the energy budget constraint that latent heating must be balanced primarily by the atmospheric radiative cooling (Mitchell et al., 1987).

The fact that water vapor increases faster than precipitation on a global scale suggests that as a whole the Earth's hydrological cycle slows down in the warming climate. On the other hand, future climate projections also imply that on a regional scale precipitation intensifies where it is already moist in the present climate (IPCC, 2014). As such, a full understanding of the nature of precipitation under delicate balance (or short-term imbalance) in the global water and energy budget remains a major challenge. We are therefore in urgent need of long term, continuous and accurate measurements of global precipitation to better document how the climate system behaves and better prepare for the future climate change (cfr. Michaelides et al., 2009; Michaelides, 2013a, 2013b, 2014, 2016).

However, although precipitation measurements are often considered as the “truth” to validate models against, it is important to be aware that measurements have their own uncertainties of different kinds. Rain gauge analyses such as the Global Precipitation Climatology Centre (GPCC) product (Becker et al., 2013; Schneider et al., 2014, 2017), for example, have spatial representativeness issues since the ground stations are highly inhomogeneously distributed over land and are totally absent over oceans (Kidd et al., 2017).

Satellite data products are superior to gauge products in spatial coverage over the globe but are subject to retrieval errors and biases. Merged data products such as Global Precipitation Climatology Project (GPCP; Adler et al., 2016; Huffman et al., 2009) and Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP; Xie and Arkin, 1997) have been among the most extensively used products for model validation purposes. Their use in climatological studies is in constant growth as the temporal coverage of the data sets increases (Kidd, 2001). In those products, multiple satellite and gauge measurements are combined so as to maximize the spatial and temporal sampling, but retrieval errors are generally even more difficult to track down in the merged products owing to the complexity of the algorithm.

The purpose of this paper is to discuss how currently available precipitation data sets may be used to validate climate models, to illustrate the uncertainties and limitations of the products and simulations, and to propose a common set of standards for both reference data sets and climate models in order to avoid pitfalls and issues arising from different practices between the observational and modeling communities.

2. Data sets of global precipitation

For climate-scale comparisons of the precipitation component of the hydrological cycle, complete, global precipitation data sets are required and much effort is spent on providing climatologically-sound data sets, with particular attention paid to avoid possible inconsistencies in such products. However, note that at present no single-source of global precipitation measurements exists (Michaelides et al., 2009). Many single-source data sets exist that provide climate-scale precipitation products, some of which are combined to provide multi-sourced global precipitation products.

In terms of coverage, surface data per se essentially refers to land-only measurements (including islands). Even over the land areas there is great variation in the availability and density of the observations (Kidd et al., 2017), which affects the representativeness of the measurements. Over the oceans, the few islands that provide measurements do not adequately represent the precipitation over the surrounding oceans or even accurately represent the immediate oceanic surroundings. Some land areas are now covered by surface-based radar networks; these regions tend to have also adequate gauge measurements, but gauges don't adequately capture the spatial variability of rainfall that radar can provide. Satellite data sets, although touted as ‘global’ are usually *nearly* global, typically being limited to 60°S to 60°N due to the extent of the available satellite observations, or to the limitations in the retrieval schemes.

2.1. Surface-based data sets

2.1.1. Rain gauge-based products

A great number of instruments are designed to provide in situ measurements of precipitation. The most common and longest-serving is the rain gauge. Gauges designed for measuring precipitation (rainfall and snowfall) represent the fundamental, de facto standard of precipitation measurements across the globe. Sevruk and Klemm (1989) and New et al. (2001) put the number of gauges worldwide at more than 150,000, while Groisman and Legates (1995) estimated the number of ‘different’ gauges to be as many as 250,000. While it is certain that many gauges exist, these numbers depend largely upon their construction principles (i.e., what is considered as a *valid* gauge), their density and or gauge record; in particular, not all gauges have operated continuously or simultaneously. Indeed, the number of gauges available at a particular temporal resolution, for a specific period, or with a certain temporal latency depends greatly upon regional/national data policies. The reader can find an up-to-date appraisal of gauge coverage and distribution in Kidd et al. (2017).

Despite the impressive number of gauges, their availability and therefore representativeness across the Earth's surface is highly variable (see Kidd et al., 2017). The vast majority of gauges over the Earth's surface are concentrated in populated regions. Over the oceans very few gauges exist with most being ‘coastal’ and not necessarily representative of the open ocean. Furthermore, while the number of gauges that report daily accumulations of precipitation might be considered adequate, gauges that report sub-daily precipitation (critical for extreme pluvial events) are very limited in number (Kidd et al., 2017).

The basic rain gauge has a number of limitations. Sevruk and Klemm (1989) noted more than 50 different types of gauge design (whether aerodynamic or not) with different orifice size and heights above ground. At the gauge-scale, the ‘capture’ of precipitation by a rain gauge is affected by the wind flow around the orifice (Duchon and Essenberg, 2001). Turbulence induced by the wind-gauge interaction interrupts the flow across the gauge orifice affecting light precipitation the most but to some extent also heavy rainfall (Duchon and Biddle, 2010), resulting in an under-catch at low intensities and higher wind

speeds (Nešpor et al., 2000).

Although pit gauges are generally deemed to produce as close to the correct 'surface' precipitation as possible, these are difficult to maintain and are therefore extremely few in number. Errors may also arise due to physical problems (such as blockages of the gauge orifice and evaporation of precipitation from the funnel/collector) or electronic/recording issues (problems with data loggers, or written records). Most gauges have difficulty in accurately recording snowfall since most are primarily designed for the collection and measurement of rain; the lower fall speed of snow compounds wind-orifice effects and thus affects the (in)efficiencies of the gauges (Goodison et al., 1998; Rasmussen et al., 2012). While rainfall can be usually measured to within 10–20% in one-minute intervals (Vuerich et al., 2009), wind-effects may result in less than 25% of the snowfall being caught (Goodison et al., 1998).

At the local scale, the exposure of the gauge within the local environment is critical to the spatial representativeness of the gauge measurement. Standards set by the WMO are designed to ensure consistency among gauge measurements to reduce inherent errors caused by siting or exposure (WMO, 2008). However, the auto-correlation length of precipitation is surprisingly small particularly for instantaneous precipitation (Habib et al., 2001); even gauges in close proximity exhibit a good degree of variability (Ciach, 2003). Accumulating precipitation over time increases the correlation length (Bell et al., 1990), although the correlation length is dependent upon the meteorology of the precipitation event and the local topography.

The conversion of the gauge information into the final precipitation product can lead to differences, not only related to the number of gauges available. A number of schemes exist to take the point measurements of the gauges and provide a spatial estimate of precipitation. Although a simple task at first glance, particularly where gauge density is high, the use of different schemes and criteria may lead to different end results. Interpolation schemes such as kriging, are used to try and reduce processing errors, although the interpolation scheme itself may introduce biases in the final product.

A more "subtle" aspect needs also to be considered regarding the timing of the gauge observations, which notably differs between the various national networks. This sometimes creates untagged multi-day accumulations with clear implications for the comparisons with other sources of data (e.g., Viney and Bates, 2004).

While gauge data are generally used "as is", errors and uncertainties associated with such precipitation measurements are reasonably well understood and corrections or further quality control measures can be applied. However, despite errors, rain gauges remain arguably the most accurate instrument by which to measure rainfall at the surface.

A number of *global* daily gauge data sets exist, as summarized in Tapiador et al. (2012) (Tables 1–5), that provides precipitation products at varying temporal and spatial resolutions. Despite the number of gauge products, it should be noted that many gauges contribute to one or more gauge products and thus such data products are not necessarily independent. Indeed, much of the variations in the gauge analyses originate in the quality control and the point-spatial conversion methodology.

The Global Precipitation Climatology Centre (GPCC; Becker et al., 2013) provides perhaps the foremost repository of global precipitation data derived from gauges. The GPCC gridded product is generated from data obtained from 180 institutions, providing a total of about 85,000 gauge locations having provided observations at least once since the start of the data set in 1901. The Full (long-term or climatological) GPCC analysis is designed to provide a high-quality product. To ensure continuous records of precipitation from any single station, the GPCC imposes a 10-year minimum constraint. This limits the number of available stations as of 2013 (2015) to 67,298 (75,165) for the best month, or 67,149 (75,033) for the worst. The total is 65,335 (73,586) stations across all 12 months of the year (Becker et al., 2013; Schneider et al., 2015). This Full product is updated about every 2 years.

Other global data sets include the CPC Gauge-Based Analysis of Global Daily Precipitation (Xie et al., 2010) and the Global Historical Climatology Network (GHCN; Menne et al., 2012), both of which provide daily gridded precipitation products derived from meteorological observations worldwide. The number of available gauges in the CPC data set varies considerably by year (and by region/year) with a maximum number of precipitation observations of just over 30,000 stations, about half of which are in the US. The GHCN collects precipitation data from approximately 7500 gauges while snow depth is measured from about 17,000 stations, again virtually all in the US.

The Climate Research Unit (CRU, at the University of East Anglia) gauge product (Mitchell and Jones, 2005) aims to provide a consistent precipitation data set at a resolution of 0.5° × 0.5° globally (excluding Antarctica) from 1901 to 2001 (New et al., 2000). The number of gauges used varies over time, ranging from a minimum of 4957 in 1901 to a maximum of 14,579 in 1981. To interpolate these gauge data to the 0.5° × 0.5° grid, a minimum 30-year record is required for a particular gauge.

Although the products mentioned above provide similar global monthly precipitation totals (e.g., Chen et al., 2002), these products tend to differ at regional scales due to contributions from different gauges, particularly where the gauge density is low such as over tropical Africa and tropical South America (Qian et al., 2006) and differences in the analysis procedures. For example, issues in validation are expected in the CRU product over the Amazon between 1901 and 1921 when only one station was available for this large region.

While not being global, regional data sets, such as the APHRODITE product (Yatagai et al., 2012) based over SE Asia, the China Gauge-based Daily Precipitation Analysis (CGDPA; Shen and Xiong, 2016), the European Climate Assessment (ECA) for Europe (Klein Tank et al., 2002), and the database described by Liebmann and Allured (2005) for South America are often able to obtain a greater number of regional gauges through local sources and tailor the final product to particular region characteristics, such as topography. For example, due to the complex topography in the domain of the APHRODITE product, topographic information is used to better constrain the interpolation of the gauge information to the final gridded product.

2.1.2. The role of radars and disdrometers

Other surface-base instruments, such as radars and disdrometers are also widely used for measuring precipitation, but lack the coverage and longevity of record necessary for wide-area analysis. However, these are included for completeness and in recognition that they may provide useful information on precipitation intensities and characteristics that could be useful in validation climate models locally or regionally.

Surface-based radars now provide good quality spatial measurements over most of the US/Canada, Europe, Western Russia, Japan, Korea, Australia, and New Zealand. Weather radars offer frequent spatial observations of precipitation over relatively large areas in near real time for time-critical applications, such as flash floods, etc. They can provide useful information to modelers on the spatial and temporal variability of precipitation together with rain intensity distribution. However, radars do have a number of limitations, such as range effects, blockages of the radar beam, and imprecise backscatter/precipitation rate (Z-R or Z-S) relationships. Calibration against surface rain gauges together with improvements in radar technology have reduced some of the errors associated with the Z-R relationships, allowed the phase of precipitation to be retrieved and the variability of the DSD to be assessed (e.g., Chandrasekar et al., 2012). Polarimetric measurements allow significant steps forward in precipitation estimation (e.g., Bringi and Chandrasekar, 2001). At the same time, dual-polarization radar measurements of precipitation are influenced by size, shape, orientation, and phase of the hydrometeors, and thus these measurements are a powerful tool for identifying hydrometeor characteristics along with precipitation intensity (e.g., Herzegh and Jameson, 1992; Vivekanandan et al., 1999).

Disdrometers are designed to retrieve key physical properties of precipitation particles, such as the drop size distribution and the particle shape, across a range of precipitation types (Chang et al., 2009; Adirosi et al., 2016; Thurai et al., 2016), temporal scales (Tokay and Short, 1996), and spatial scales (Tapiador et al., 2010; Jaffrain et al., 2011). However, there are known biases due to the small measurement area of the instruments (Tapiador et al., 2017), with single disdrometers severely incorrectly estimating instantaneous rain rate (up to 70%) and underestimating the mean diameter of the rain drop size distribution. For climate-scale verification, however, these instruments have the potential to investigate the representation of precipitation particle types within models and also to help understanding the precipitation processes that generate these particles.

2.2. Satellite estimates

Satellite estimates of remotely-sensed precipitation have been available over much of the globe for almost four decades and have the potential to be available on a truly global scale. Critically, satellite estimates have a distinct advantage for assessing precipitation over data-sparse land and ocean regions. Satellite observations from visible, infrared, and in particular, passive and active microwave systems are used to generate precipitation estimates using a number of techniques (Kidd and Huffman, 2011), although techniques differ in performance regionally and temporally. Simple techniques based upon the calibration of the observations using surface (or satellite) radar may yield useful results, although greater insight into the precipitation characteristics may be obtained through the use of more complex physical modeling techniques (e.g., Kummerow et al., 2000, 2001, 2007).

2.2.1. Infrared-based methods

Observations in the (thermal) infrared channels (IR) allow cloud top temperatures to be retrieved, and through a simple relationship between cloud top temperatures and surface precipitation may be used to retrieve precipitation. However, the actual relationship is considerably more complicated than this simple information can describe. Nevertheless, techniques such as the Global Precipitation Index (GPI; Arkin and Meisner, 1987), the Convective/Stratiform (CST; Adler and Negri, 1988), the Auto-Estimator (Vicente et al., 1998), and the Hydro-Estimator (Scofield and Kuligowski, 2003) have proved effective at large-scale precipitation estimation. Some visible/infrared fused techniques have utilized artificial neural networks (ANN), and for example, the Precipitation Estimation from Remote-Sensed Information using ANN (PERSIANN; Sorooshian et al., 2000) uses multi-source information from satellite and surface data sets to establish, and update the relationship between surface precipitation and IR observations. Tapiador et al. (2004) also used ANN to merge PMW with IR data. The usefulness of observations in the visible (VIS) spectrum is limited by solar radiation and so these tend to be combined with IR observations. Several new global datasets dwelling on IR observations have been proposed recently with the aim of supporting flood and drought studies. Among them, it is worth mentioning the Climate Hazards group Infrared Precipitation with Stations (CHIRPS; Funk et al., 2015). Local datasets are also available and widely used such as, for example, the Tropical Applications of Meteorology Using Satellite Data and Ground-Based Observations (TAMSAT; Tarnavsky et al., 2014).

2.2.2. Microwave-based methods

Passive microwave observations provide a more direct measure of precipitation since the upwelling microwave energy is linked more directly to the precipitation-sized hydrometeors. While simple techniques, such as frequency differences highlighting the scattering from frozen hydrometeors may be used, more complex techniques using radiative transfer calculations can be used to build a database of simulated observations and their associated geophysical parameters (Elsaesser and Kummerow, 2008). Spaceborne precipitation radars have been flown on a number of missions. The Tropical Rainfall Measuring Mission (TRMM; Kummerow et al., 1998) carried the Ku-band Precipitation Radar (PR), CloudSat carries a W-band cloud radar (Stephens et al., 2002), while the Global Precipitation Measurement mission (GPM; Hou et al., 2014; Skofronick-Jackson et al., 2017) carries the Ku/Ka band Dual-frequency Precipitation Radar (DPR). These radars provide not only a measure of surface precipitation, but also the vertical distribution of precipitation in the atmosphere; the disadvantage is that the global coverage is limited, requiring significant temporal averaging to provide useful (i.e., sample-limited) products. Combining multi-satellite information has proved useful, as evidenced by the A-train (Stephens et al., 2002), allowing complementary atmospheric observations to be exploited.

A number of satellite-derived global precipitation products are available (e.g., Kidd and Levizzani, 2011), although data availability and retrieval accuracy typically limits these products to the latitude band 60°S to 60°N. Furthermore, estimates from any single sensor are infrequent (typically less than twice a day from low Earth orbiting satellites) and may be deemed to be only a snapshot of precipitation, posing a severe drawback for documenting the evolution of precipitation. The relative high frequency sampling of geostationary visible/infrared sensors is negated by the indirectness of these observations. The ability to exploit a number of different precipitation-oriented sensors greatly improves the number of samples, while incorporating the more direct, but sample-poor microwave estimates with the less direct, but sample-rich infrared observations promises much improved precipitation estimates.

A number of techniques have been devised to combine information from both visible/infrared and microwave observations into Level-3 (i.e., gridded daily time scale or finer) products. The Bristol-NOAA InterActive Scheme (BIAS; Barrett et al., 1987) replaced VIS/IR estimates of precipitation with those derived from the passive microwave; cloud development and movement was then used interactively to advect the estimates over time. Adler et al. (1993) calibrated IR cloud top temperatures with PMW-derived estimates to correct for regional-scale variations in the cloud-top temperature/rain rate relationships, although these calibrations were generated at relatively coarse resolutions (2.5° × 2.5° monthly). Since most general circulation models (GCM) assessments use Level-3 products, it is important to understand the procedures of how these are created.

2.2.3. IR-PMW combined methods

Current combined IR–PMW techniques fall into two main categories. The first rely upon the calibration of the IR observations by PMW estimates: The NRL-Blended technique (Turk et al., 2000) and the Passive Microwave-InfraRed (PMIR) technique (Kidd et al., 2003) use a moving spatial and temporal window to generate a local relationship between the fast-refresh IR observations and the precipitation estimates sourced from the Level-2 (i.e., instantaneous swath-level) PMW observations. The second methodology relies upon the PMW to provide the estimate of precipitation and the IR data to track the movement of the precipitation between adjacent PMW estimates. These are usually termed advection morphed or Lagrangian time-interpolation schemes. Examples of advection morphed schemes include the Climate Prediction Center (CPC) Morphing (CMORPH; Joyce et al., 2004), the Global Satellite Map Product (GSMaP; Kubota et al., 2007) and the Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman et al., 2015), while the Rain Estimation using Forward Adjusted-advection of Microwave Estimates (REFAME; Behrangi et al., 2010) is an example of a Lagrangian scheme. The PMW estimates may be derived from a number of techniques. The IR information is used to derive motion vectors along which to advect the PMW estimates. These motion vectors may be generated through correlation or mesh-based techniques (Bellerby, 2006) or physically-based approaches (Tapiador, 2008). Additional steps are used in CMORPH and IMERG to 'morph' the estimates between the starting and ending PMW estimates. The most recent merged

techniques provide precipitation estimates at resolutions down to 10 km, 30 min over 60°S–60°N.

A number of mature satellite-based precipitation techniques incorporate surface precipitation data sets, allowing high spatial and temporal resolution precipitation products to be generated (e.g., Huffman et al., 2009) using surface gauge measurements to provide the anchor points. Algorithms that emphasize homogeneous input and processing are referred to as Climate Data Records (CDR), while those that give the best short-interval estimate are referred to as High-Resolution Precipitation Products (HRPP).

The TRMM Multi-satellite Precipitation Analysis (TMPA; Huffman et al., 2007) provides 3-hourly 0.25° × 0.25° resolution precipitation data from 50°N–50°S 1998–present, and consists of four products: a merged-microwave product, a microwave-calibrated IR product, a combined calibrated-IR/merged-microwave product, and a rain gauge-adjusted product. The analysis incorporates as many satellite-based precipitation estimates as possible calibrated to a single sensor. The near-real time product (3B42RT) is calibrated against the Goddard Profiling (GPROF) algorithm (Kummerow et al., 1996) TMI data (2A12RT), while the post-realtime research-grade product (3B42) is calibrated against the TMI/PR estimates (2B31; Haddad et al., 1997). Other PMW estimates are then histogram-matched against these products, before being merged into 3-hourly windows centered on the nominal meteorological observation times (i.e., 0000, 0300, 0600 … 2100 UTC). These combined PMW estimates are the used to calibrate the IR observations over overlapping 3° × 3° regions. These IR estimates are used to fill gaps in the PMW coverage in each 3-hourly time step. A final bias adjustment against the monthly GPCC monitoring gauge analyses is carried out to generate a gauge adjusted 3-hourly 0.25-degree product (3B43; Huffman et al., 2010).

As part of the Global Precipitation Climatology Project (GPCP) several products are generated and designed to address Climate Data Record standards. A merged Satellite-Gauge (SG) precipitation analysis (Huffman et al., 1997, 2010; Adler et al., 2003) provides complete global monthly 2.5° × 2.5° estimates from 1979 to the present. The product uses the precipitation estimates from the SSM/I and SSMIS passive microwave sensors to calibrate geostationary IR observations over 40°N–40°S. Poleward of this latitudinal band the passive microwave estimates are combined with those generated from TOVS or AIRS. Over land the satellite precipitation is then adjusted against the full GPCC rain gauge analysis.

The GPCP One-Degree Daily (1DD) precipitation analysis (Huffman et al., 2001, 2010), available from October 1996 forward, uses a Threshold-Matched Precipitation Index (TMPI) between 40°N–40°S to produce instantaneous precipitation from the geo-IR observations. A regionally-varying IR threshold is matched against the fractional coverage of the SSM/I and SSMIS-derived GPROF precipitation estimates. The rain rate of the IR pixel is then computed so that the TMPI monthly precipitation equals that of the corresponding SG monthly precipitation total. Outside 40°N–40°S, precipitation estimates from the TOVS and AIRS sensors are adjusted by the precipitation occurrence of GPROF at 40° latitude.

The GPCP 2.5° × 2.5° pentad precipitation analysis (Xie et al., 2003) uses the satellite-gauge product to adjust the pentad CPC Merged Analysis of Precipitation (CMAP) pentad product so that the overall magnitude of the product matches at the monthly scale but with the sub-monthly variability of the pentad CMAP product.

The CPC Merged Analysis of Precipitation (CMAP; Xie et al., 1996) generates 2.5° × 2.5° monthly global precipitation derived from rain gauges and satellite precipitation estimates. The method of combining the gauge and/or satellite information uses a maximum likelihood based upon weights inversely proportional to the square of the random error of the individual sources. Over land gauge analysis is used as the reference, while over the oceans the Pacific Atoll data set is used.

2.2.4. Known issues in using precipitation databases for validation

Issues with satellite data products relate primarily to the indirectness and the temporal sampling of the estimates. Although PMW estimates relate to well-founded knowledge of radiation-hydrometeor interactions, significant variability occurs due to the usefulness of the observations over different surfaces. Over land, only high frequency channels are useful, and these sense frozen hydrometeors, which means they sense the ice at the top of the precipitating system rather than the surface precipitation. Over the ocean, both high and low frequencies are used, where the latter respond to liquid hydrometeors. Gauge information to provide a calibration reference may be used to correct regional and seasonal variations, although the accuracy of the gauge information is critical.

For proper assessment of GCM's, it is important to understand the processes that go into the production of global-scale satellite-based Level-3 precipitation datasets. The quality of the Level-3 products mentioned above is driven by continuous, short revisit microwave-based observations. These data originate from a variety of different sensors and satellites, with different spectral and spatial resolutions, and quality over land and water surfaces (Turk et al., 2016). Given the varying revisit time and different sensitivity to precipitation intensity, the error characteristics associated with merging these many disparate datasets are not well known (Maggioni et al., 2016). One of the fundamental contributors to the overall error Level-3 characteristics is therefore the properties of the Level-2 precipitation algorithms that are applied to the different microwave data sources (e.g., Elsaesser and Kummerow, 2015). Many of these techniques are built upon Bayesian estimation schemes which, while appropriate, are fully dependent upon the quality and accuracy of a-priori precipitation estimates provided by the GPM DPR (Grecu et al., 2009). For example, assumptions made for the drop size distribution characteristics of the GPM radar-based estimates, therefore “trickle down” and influence the radiometer-based precipitation retrievals, and eventually the Level-3 merged products that get used for GCM model verification. As the quality of the GPM Level-2 products improves, further improvement should result in the Level-3 merged products.

In summary, as with any data sets, precipitation products have a number of error sources, although, due to the nature of these products, quantifying these is problematic. Understanding the provenance of the input data and the processing scheme is crucial to understanding the particularities of each precipitation product to ensure that the correct conclusions on model performance are provided.

With some exceptions, the uncertainties in these data sets are seldom considered in climate model validation, rendering the exercise prone to justified criticisms. Also, it is worth mentioning that the majority of the merged products, either CDR or HRPP, were never intended to be used for trends as their input data sources drift.

3. Direct verification of climate model precipitation outputs

Most of the databases in the previous section have been compared with climate models for a variety of purposes, for instance for uncertainty analyses. In the case of RCM, the majority of such exercises are aware of the difference between reanalysis-driven and GCM-driven RCMs, as the overriding influence is the differences in the lateral boundary conditions, and validation in only meaningful for the resulting aggregated climatologies and not for the temporal sequence of events.

The models can be validated and verified at several temporal and spatial scales. Direct comparisons of highly aggregated data (such as the mean annual precipitation in latitudinal bands) are useful as a first approximation, but they do not sufficiently prove that a climate model correctly represents the climate as even early energy balance models such as Paltridge's (1975) compared favorably with observations at those scales. Neither is reproducing the annual cycle in mid latitudes a major challenge, as models can be easily tuned to favor the main mode

of variability, and even models without microphysics parameterizations can simulate quite well the seasonal cycle of precipitation. Coarse aggregations (e.g. 5° × 5°) are not especially useful either as many basins are smaller than that or a major driver of precipitation research is hydrological applications at local scales.

The minimum set of statistics for validating precipitation in climate models are the annual and seasonal comparisons, the analyses of bias and correlations (usually through scatterplots or Taylor diagrams), the use of probability distribution functions (PDFs), latitudinal and longitudinal transects, and time-latitude or longitude diagrams. Spatial indexes such as Geary's or Moran's have also been used (Tapiador, 2010) to gauge spatial decorrelation. More sophisticated plots, such as the Hovmöller diagram, are useful for model outputs, but seldom suited for comparing with small or non-homogenous observational data sets.

Grid point to grid point comparison is a natural test for validating climate models since most applications of GCM outputs in mitigation and adaptation are most useful locally. The spatial variability of the precipitation makes instantaneous comparisons difficult, but climatologies are fully comparable. Fig. 1 shows a comparison of a standard Community Earth System Model (CESM) simulation of precipitation for present climate with several observational data sets, and for the same period (2000–2015). It is apparent that there is an overall consensus albeit differences appear (Table 1). The well-known issue of incorrectly having a split ITCZ appears in the model, but the observational databases agree on the major features of the global precipitation climatology. Fig. 2 shows the corresponding scatter plots for several comparisons between model and observational databases.

A tenet of such direct validation is that the spatial and temporal scales at which the comparison is performed represents a limit for judging a model ('scope principle'). One can hardly claim that a model with a spatial/temporal resolution of 25 km/6 h performs well unless the validation has been performed at such scales and reasonable scores have been found. In other words, the ability of the CESM model to reproduce the main precipitation climatology at 1° resolution in Figs. 1 and 2 does not necessarily imply that the model is suitable at, for instance, daily scale and 25 km spatial resolution.

Favorable scores at such coarse aggregations only show that the model can be safely used at those scales. Any claim of potential applicability or positive assessment on the model ability at 25 km/6 h resolution should be backed by a validation at that scale or at finer scale (say 5 km/1 h). Otherwise, unsubstantiated claims are likely to appear. This is important as sometimes RCMs are coupled with, for instance, crop models with the intent of providing detailed projections of yields in future climate conditions. If the climate model has not been validated at the intended working resolution it is doubtful that the results can be of any practical use, notwithstanding the existence of additional, stronger issues to question such practice (cfr. Oettli et al., 2011).

There are a number of possible ways in which climate models can be compared to observations. Satellite estimates are convenient in that, like the climate models themselves, they tend to be global and homogeneous. Critical in any such endeavors, of course, is that the "observations" are in fact trusted, and do not themselves have biases, trends or artifacts that would cloud such comparisons.

Here, we discuss three separate topics: absolute magnitudes, spatial variability, and trends plus temporal variability. These can be viewed as the first order outputs of the climate models and tests their fidelity in a macroscopic sense. More complex, but also potential metrics could be framed around process-related verification as might be related to climate models' ability to capture the diurnal cycle of precipitation, the proper partition of convective versus stratiform precipitation or finer scale triggers such as orography or coastal influences on convection.

A final, but perhaps even more complex form of verification could be framed around societal question such as a climate models' ability to capture regional trends in precipitation, reproduce drought indices, or capture trends in regional precipitation extremes. As these latter properties would require higher resolution in-situ observations for verification that have their own uncertainties, we limit ourselves here only to the first category – the physical evaluation of macroscopic precipitation properties.

While there are many satellite products available, there are relatively few classes of products that the community has converged upon. Following Ebert et al. (1996), it is generally accepted that passive microwave sensors have better instantaneous skill at sensing rainfall than their VIS or IR counterparts. This is because microwaves are attenuated only weakly by clouds and can thus sense the broad column of hydrometers (e.g., Wilheit et al., 1977). Although VIS/IR methods are generally limited to cloud top information (e.g., Platnick et al., 2003) their availability from geostationary satellites allows much better sampling of the relatively infrequent precipitation events. This led to better correlations of precipitation at daily and monthly scales (Ebert et al.,

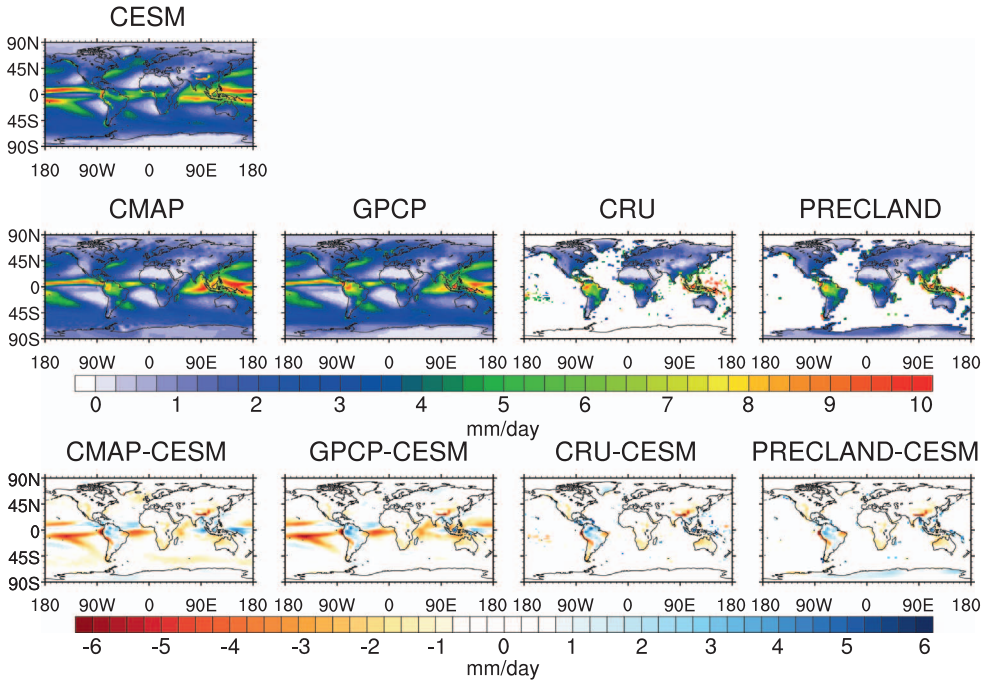

Fig. 1. A comparison of the global annual mean (2000–2015) of precipitation as simulated by the CESM model (top) and four different observational databases. The third row depicts the differences between the observations (second row) and the model (top row). The GCM simulation includes Carbon and Nitrogen cycles.

Table 1
Cross-correlations of the CESM precipitation estimates with observational data in Fig. 2 (2000–2015 climatology).

Total annual precipitation (at 2.5° spatial resolution)	CESM		
	Global	NH	SH

CMAP	# Points = 13,824
	y = 0.817x + 0.207
	R^2 = 0.747
		# Points = 6912
		y = 0.903x + 0.111
		R^2 = 0.818
			# Points = 6912
			y = 0.725x + 0.342
			R^2 = 0.671

GPCP	# Points = 13,824
	y = 0.787x + 0.339
	R^2 = 0.737
		# Points = 6912
		y = 0.83x + 0.368
		R^2 = 0.798
			# Points = 6912
			y = 0.747x + 0.312
			R^2 = 0.679

PRECLAND	# Points = 6561
	y = 0.89x + 0.249
	R^2 = 0.708
		# Points = 3865
		y = 0.947x + 0.048
		R^2 = 0.741
			# Points = 2696
			y = 0.838x + 0.476
			R^2 = 0.688

CRU	# Points = 4759
	y = 0.94x + 0.139
	R^2 = 0.681
		# Points = 3673
		y = 0.958x + 0.136
		R^2 = 0.687
			# Points = 1086
			y = 0.946x − 0.001
			R^2 = 0.576

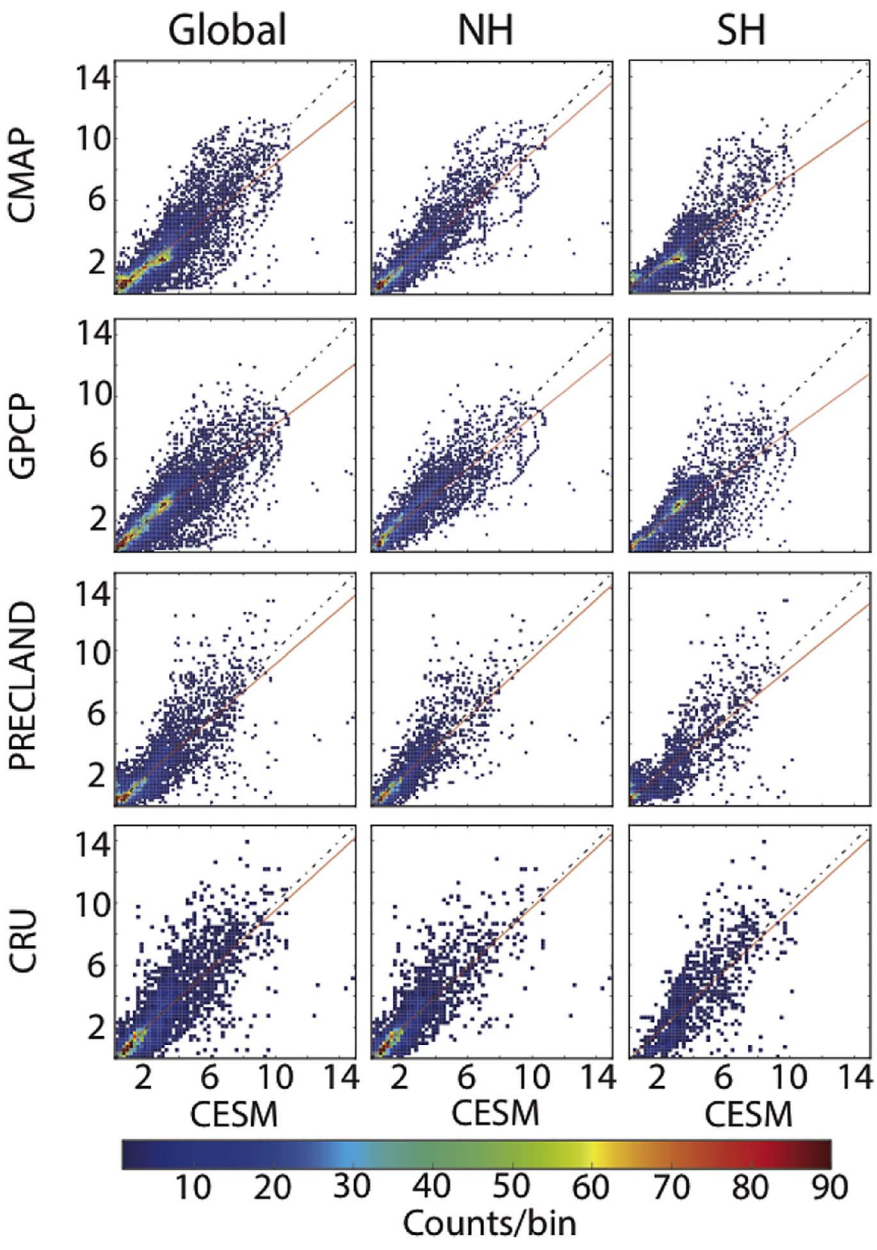

Fig. 2. Scatter plots comparing the annual mean (2000–2015) at every grid point for different ground reference data (GPCP, CRU, PRECLAND and CMAP) and a GCM simulation (CESM model); for the whole planet (left column), the Northern hemisphere (middle) and the Southern hemisphere (right).

1996) than the more precise but infrequent microwave observations from low Earth orbiting satellites. The general approach followed by algorithm developers has thus been to merge the instantaneous microwave estimates with more frequent infrared estimates to overcome sampling limitations.

As mentioned above, because rain gauges are readily available over land and are still considered to be the reference standard against which satellite estimates are evaluated, most producers of multi-sensor, multi-satellite products have also incorporated rain gauge estimates into their products. This is generally done through some form of optimal interpolation scheme to ensure that areas with many gauges do not differ significantly from their measurements while the satellite has greater weight in areas with few or no gauges. Adjustments are usually made at monthly or climatological time scales, although climatological adjustment factors can be applied to instantaneous estimates.

The TMPA (Huffman et al., 2007), GSMAP (Kubota et al., 2007), and CMORPH (Joyce et al., 2004), IMERG (Huffman et al., 2015), GPCP (Adler et al., 2003; Huffman et al., 2009) and CMAP (Xie et al., 1996) are all well-known examples of such techniques that blend microwave with infrared sampling and gauges for bias corrections. The differences among these products thus stem from the subtleties employed to merge the infrared and passive microwave as well as the bias adjustment to gauges. These subtleties are beyond the scope of this paper but have to be considered when comparing climate models with observations. Instead, and as a first approximation to the problem we focus here primarily on areas of broad agreement among these products as a measure of where observations can be considered robust (Hegerl et al., 2015).

In this section, we focus on individual sensor estimates along with the GPCP and CMAP products described above. While TMPA, GSMAP, CMORPH and IMERG are all high-quality products that may also be used in some of these applications (Kim et al., 2017; Li et al., 2017a; Guo et al., 2016), they are designed primarily to produce the highest possible quality snapshot rainfall product. These products blend all available information at any given time to optimize sampling and no effort is generally made to homogenize the time series. This can cause variations in the product over time as old satellites drop out and new, generally more capable satellites are introduced into the time series.

3.1. Absolute magnitudes

The TRMM mission did much to bring together different estimates of precipitation in the tropics. These products have been transferred to GPM which also shows good agreement in the tropics. Fig. 3 shows the current zonal mean estimates of precipitation from GPM's Ku radar (Iguchi et al., 2000), passive microwave radiometer (Kummerow et al., 1996, 2001), GPCP (Huffman et al., 2009), and CMAP (Xie et al., 1996) estimates in the tropical band [35°N–35°S]. Agreement is good over both ocean and land as shown in panel's b and c respectively. The good agreement over land is particularly interesting in that GPCP and CMAP include rain gauge adjustments while the GPM Ku radar and radiometer estimates are independent of gauges. Notwithstanding that algorithms were designed with knowledge of reference values, the good agreement would suggest that tropical precipitation is relatively well measured and that bias adjustments from gauge analyses are relatively minor at global scales. The agreement is good even in areas of large rain gauge concentrations (e.g., US, Europe, Japan and China) where the optimal interpolation essentially reproduces the rain gauge estimate within a few percent of the mean rain (Wang et al., 2014).

Matters become more complicated when higher latitudes are included. Both active and passive sensors face unique challenges related to the shallow and often very light precipitation common at high latitudes. Passive microwave sensors are relatively skilled at retrieving the total water content over oceans. However, small drops associated with drizzle at high latitudes offer little contrast with cloud droplets at microwave frequencies. As such, it has always been difficult to relate the observed emission signal into clouds and precipitation. Hilburn and Wentz (2008) used observations along the west coast of the United States to determine the cloud threshold that corresponded to the correct frequency of precipitation as observed on the ground. Further emission was then partitioned between cloud water and precipitation to optimize correlations with surface observations of precipitation.

Other methods such as GPROF (Kummerow et al., 1996, 2001) have used cloud resolving model simulations to perform the partition between cloud water and actual precipitation but the method relies on limited high latitude simulations and may not be universally applicable (Elsaesser and Kummerow, 2015).

Active sensors on TRMM and GPM have limited sensitivity to drizzle. While the GPM radar can detect precipitation down to approximately 0.2 mm h^{-1} (Hou et al., 2014), this lower threshold corresponds to assumed tropical drop size distribution and not to the very small drizzle drops that are often reported at high latitudes. CloudSat, in contrast, has enough sensitivity to detect all clouds and drizzle (Stephens et al., 2008) but cannot detect echoes near the surface when the attenuation becomes too large, as in an area of heavy rain. CloudSat

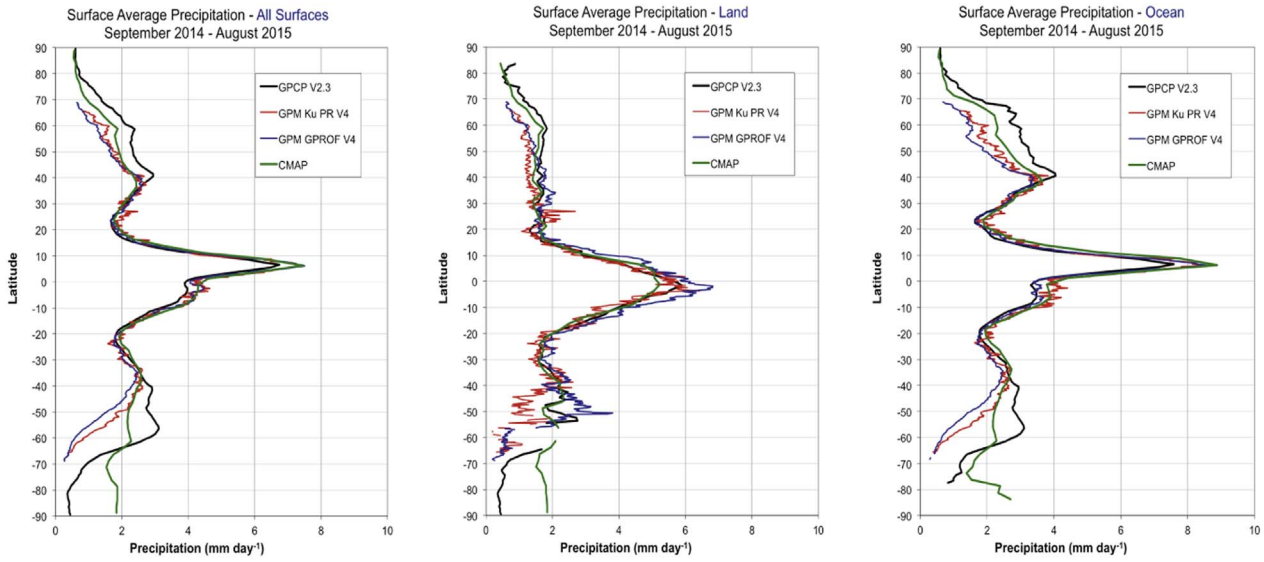

Fig. 3. Zonal means for GPM Ku PR V4, GPM GPROF V4, GPCP and CMAP for Sept. 2014–Aug. 2015. Panel (a) is for all surfaces; panel (b) is for land only; while panel (c) is for oceans only.

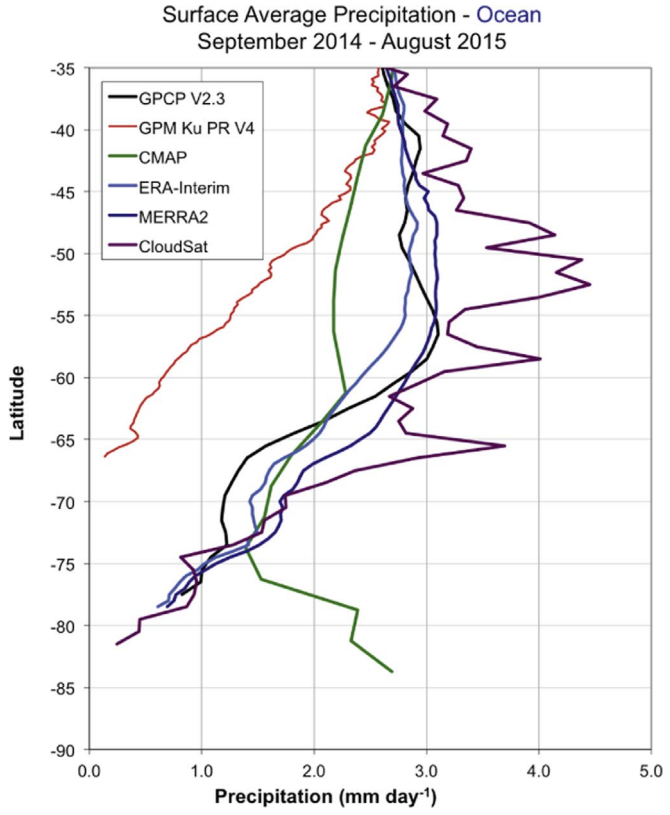

Fig. 4. Zonal ocean means precipitation from −35 to −90 S (i.e. Antarctic shelf) for 6 estimates: GPCP, GPM- Ku radar, CMAP, ERA and MERRA-2, and CloudSat.

nevertheless represents perhaps the best estimate of climatological precipitation at these high latitudes.

Fig. 4 focus on the southern hemisphere to show how the GPM radar, CloudSat, GPCP and CMAP all compare over high latitude southern oceans. ERA-Interim (Dee et al., 2011) and MERRA-2 (Gelaro et al., 2017) are also included for reference. The CloudSat zonal means are still a bit noisy due to the limited sampling of the nadir-only radar but the general trend can already be discerned. The southern oceans are shown to illustrate a worst-case scenario as there are no gauges against which the global merged products such as GPCP may calibrate.

Those issues have to be considered when validating climate models. Given the extent of tuning that models have there is a danger of overfitting parameters to observations that are known to present its own uncertainties, biases and limitations. For example, the number of preferred conically-scanning microwave radiometers is expected to decrease in the coming decade as the legacy SSMIS sensors reach end-of-life (Huffman et al., 2015). To maintain global coverage, this implies more reliance upon the large number of cross-track scanning, coarser resolution passive microwave sounders. While these relatively abundant observations are more indirectly related to near-surface precipitation (Kidd et al., 2016), they could also provide measures of the depth and intensity of convection (Haddad et al., 2017).

3.2. Spatial variability

While issues related to the detection thresholds of active and passive sensors, as well as attenuation at high rain rates (particularly for W-band radars) require some mitigation or at least merging of different satellite estimates, spatial variability is generally quite consistent among different satellite estimates. This is particularly true for TRMM and GPM that sample across the entire diurnal range. To illustrate this behavior, Fig. 5 shows an annual mean map for GPCP and the GPM Ku radar products. While a difference map would largely highlight the

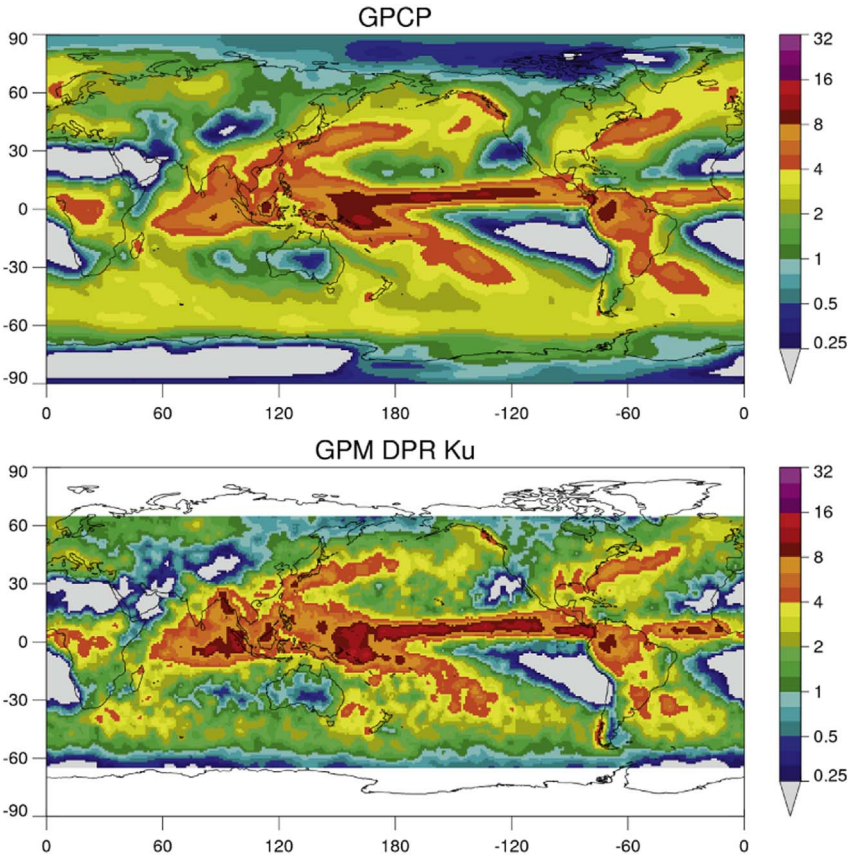

Fig. 5. Annual mean daily rainfall (mm/day) for the GPCP and GPM Ku-based radar algorithm (V4).

smaller sampling of the radar product compared to the multisatellite-based GPCP, the very good agreement is clearly visible. Other products have similar agreements (Katsanos et al., 2016; Retalis et al., 2016). This can be attributed largely to accumulations resulting largely from moderate rainfall (1–10 mm h^{-1}) that are retrieved quite well by most sensors.

3.2.1. Trends and temporal variability

It is important to differentiate trends and variability at the global scale, which may be due to changes in the overall climate system, from temporal changes at regional scales that tend to follow the general circulation and are closely related to the spatial variability. The latter, characterized by such effects as interannual variability, or the manifestation of El Nino Southern Oscillation (ENSO) in the Central Pacific, are easily detectable in the various satellite records as illustrated by time series of TRMM radar, radiometer, GPCP and CMAP times series of precipitation over a 20° × 20° area centered over the central Pacific [10°S to 10°N; 170°E to 170°W] as shown in Fig. 6.

In contrast to regional variability, trends and variability at global scales are more difficult to assess and there is less agreement among satellite products. This difficulty arises because the global trend and variability are the residue of the relatively large, opposing-sign variations at the regional scale. A comparison among GPCP, CMAP, the TRMM passive microwave product and the RSS products (Hilburn and Wentz, 2008) illustrates this fact. Fig. 7 shows these products averaged over the tropical oceans for a 15-year period corresponding to the TRMM era.

In terms of overall magnitudes, the agreement among GPCP, the TRMM product and the RSS product are quite good while the CMAP product produces 10% more precipitation. If land is included (not shown), the agreement between GPCP and CMAP improves as both products rely on more or less the same gauges for bias adjustments. However, CMAP also has more pronounced variability over this time period that is not seen in the other products. The GPCP time series appears to have the least variability of the remaining three products. This may be related to the nature of the radiometer scheme used in the GPCP algorithm itself as different algorithms can have different sensitivities to changes in precipitation system structure and morphology.

A possible explanation for these differences can be found in the TRMM record which also shows disagreement between its radar and radiometer estimates when averaged over tropical oceans as shown in Fig. 8 By using a well-characterized dual-polarized radar at Kwajalein, it is possible to explain the differences in terms of the frequency of occurrence of unique precipitation states, defined as shallow, isolated deep convection, and organized convection. When viewed by convective organization, the ENSO related biases can be explained as mere shifts of deep isolated convection to more organized convection (largely in the Central Pacific) as the ENSO persists. This sensitivity of microwave sensors (both active and passive) remains an issue that the GPM dual frequency radar is in the process of untangling.

At this time, it is somewhat problematic to use these global scale *trends and variations* to validate climate models as there is still some disagreement among the satellite products themselves. However, looking at the envelope of the observational products, models that lie outside such envelope need to be treated with care. In particular, ENSO moves around the precipitation without driving large global changes, according to the satellites. We note similar difficulties with rain gauge only products, where wind corrections can sometimes alter the conclusions about precipitation trends at the global scale (e.g., Metcalfe et al., 1997; Ungersböck et al., 2001).

4. Validation of the physics

Apart from performing direct validation against precipitation databases, there are other strategies for using observations to improve climate models. In this respect, the limitations of models include the precise modeling of the ENSO phenomenon (Neale et al., 2008), the representation of the diurnal cycle of rainfall (Betts and Jakob, 2002) or the frequency of occurrence of high- and low- intensity rainfall events (Sun et al., 2006). All these issues have been identified long ago, still persist, and are directly related with the representation of subgrid-scale processes such as convection. Note that the validation of precipitation physics may be premature for GCMs, but is included here as the GCMs are quickly reaching resolutions where cloud permitting parameterizations become a reality (e.g., Bretherton, 2015).

A major research topic to improve the water cycle modeling and addressing such issues is comparison of modeled and observed latent heat (LH). Latent heat release is a consequence of phase changes between the vapor, liquid, and frozen states of water, which cannot be measured/detected using present observational instruments. The

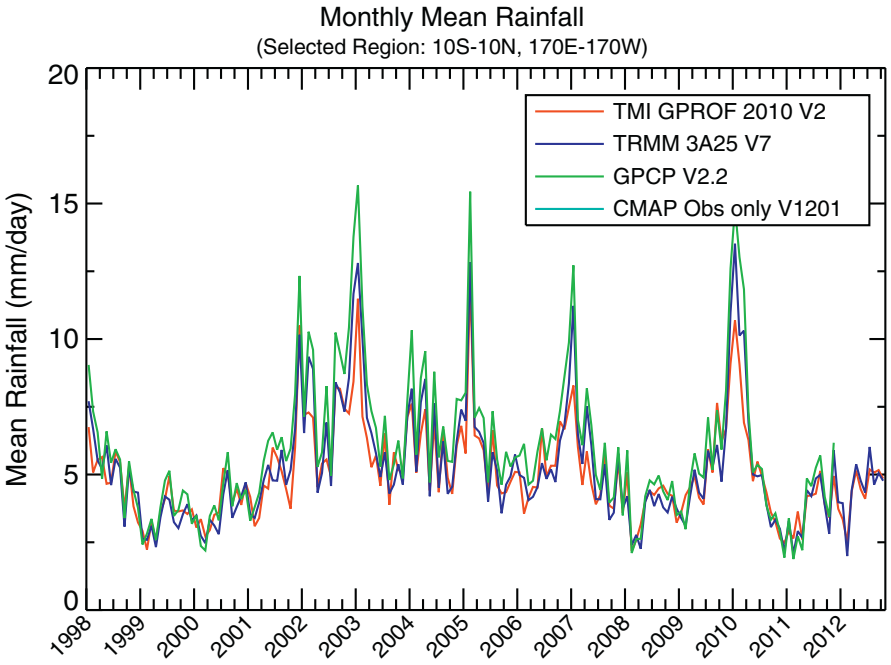

Fig. 6. Time series of TRMM radiometer (TMI GPROF 2010, V2), radar (3A25, V7), GPCP V2.2 and CMAP (observation only, V1201) times series of precipitation over a 20 × 20° area centered over the central Pacific [10°S to 10°N; 170°E to 170°W].

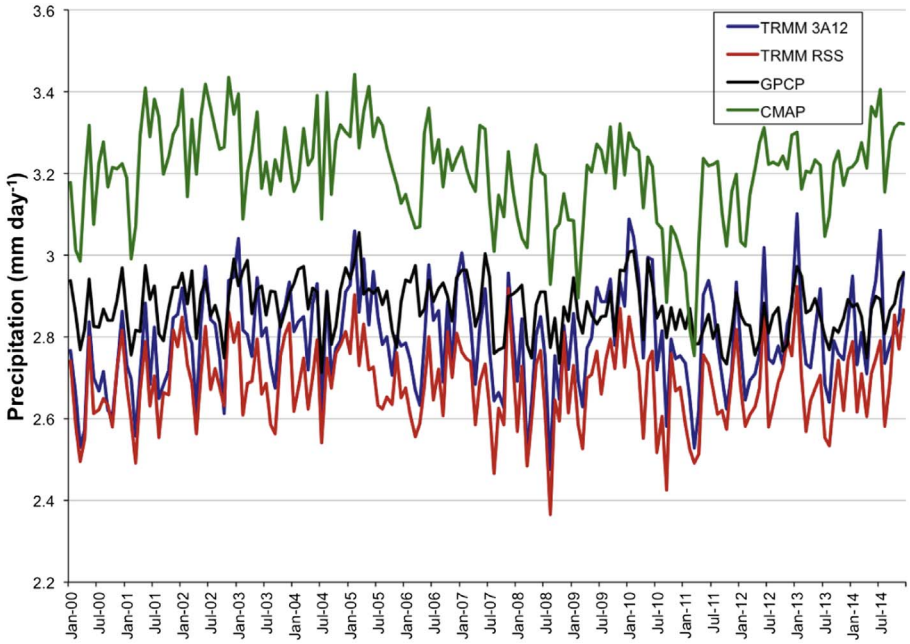

Fig. 7. Time series of tropical oceanic precipitation from the TRMM mission radiometer product (TRMM 3A12), Remote Sensing Systems (RSS), GPCP and CMAP.

vertical distribution of LH has a strong influence on the atmosphere, controlling large-scale tropical circulations, exciting and modulating tropical waves, maintaining the intensities of tropical cyclones, and even providing the energetics of midlatitude cyclones and other mid-latitude weather systems (Li et al., 2017b; Melcón et al., 2017).

Another direction is in the field of microphysics (MP). This is the framework through which to understand the links between interactive water vapor, aerosol, cloud and precipitation processes. Global measurements of microphysics are important to avoid overfitting the models to specific places when tuning the empirical parameters, which is the standard procedure to adjust models to observations (Voosen, 2016).

A third way of dealing with evaluating climate model physics is setting up dedicated ground validation (GV) campaigns designed to provide multisource, complementary information on precipitation processes. Here the combined use of databases from instrumented research aircraft, radars, and satellites is invaluable, but complicated logistics, costs and technical difficulties both limit and make extremely valuable existing campaigns such as iFloodS (Ryu et al., 2016), ipHex (Barros et al., 2014) or OLYMPEX (Houze et al., 2017).

4.1. The use of latent heat (LH) measurements to evaluate models

The launch of the TRMM satellite in November 1997 provided a much needed and accurate measurement of rainfall as well as the ability to estimate the four-dimensional (4D) structure of latent heating (LH) over the global tropics (Simpson et al., 1988, 1996).

The success of TRMM made it possible to have another major NASA precipitation measuring mission, the GPM mission. GPM is considered by NASA to be the centerpiece mission of its Global Water & Energy Cycle research program. Cloud Resolving Models (CRMs) have been identified as being a valuable tool for algorithm developers and is

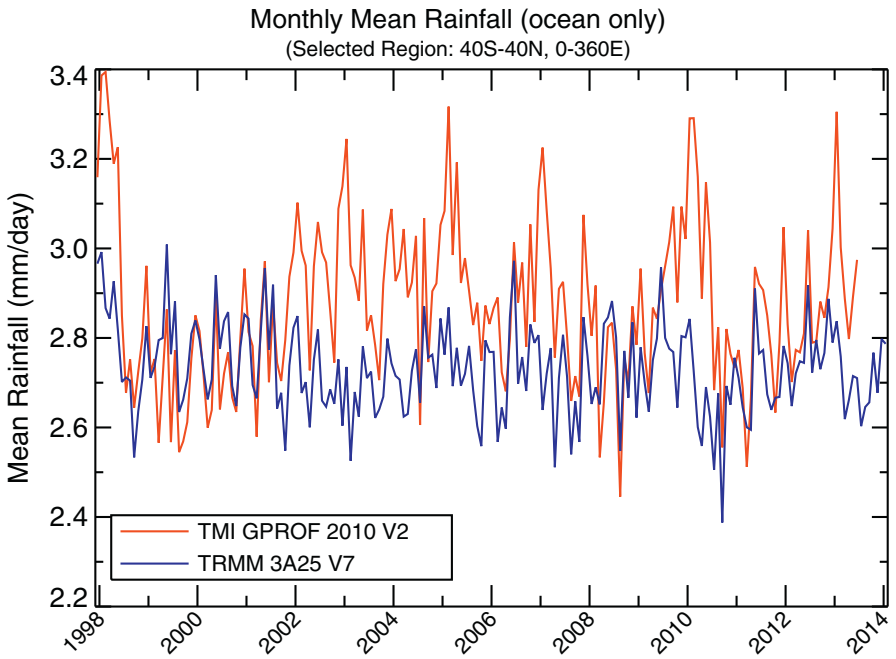

Fig. 8. Time series of TRMM radiometer (TMI GPROF 2010 V2), and radar (3A25, V7), averaged over global tropical oceans.

considered a key component for one of the major GPM ground validation (GV) sites. In addition, CRMs are one of the most important tools used to establish quantitative relationships between diabatic heating and rainfall. Thus, simulated data from the Goddard Cumulus Ensemble (GCE) have been used extensively in TRMM for the development of both rainfall and heating retrieval algorithms (Simpson et al., 1996; Tao et al., 2006).

Five different TRMM LH algorithms designed for application with satellite-estimated surface rain rate and precipitation profile inputs have been developed, compared, validated, and applied for over two decades (Tao et al., 2001, 2006, 2016b). They are the: (1) Goddard Convective-Stratiform Heating (CSH) algorithm, (2) Spectral Latent Heating (SLH) algorithm, (3) Goddard Trained Radiometer (TRAIN) algorithm, (4) Hydrometeor Heating (HH) algorithm, and (5) Precipitation Radar Heating (PRH) algorithm. The strengths and weaknesses of each algorithm are discussed in Tao et al. (2006).

Ling and Zhang (2011) compared the heating profiles between TRMM retrieved (CSH, SLH and TRAIN) and global re-analyses (ERA-1, JRA25 and CFSR). All heating data exhibit three longitudinal maxima but with different amplitudes; for example, heating over South America and Africa is much stronger in three models (CSH, SLH, and CFSR) than the others. Heating is obviously weaker over the Maritime Continent than over the eastern Indian Ocean and western Pacific in some data (e.g., Q1, TRAIN LH, ERA-I Q1, and JRA25 Q1) but not so in others. Among all, TRAIN has the largest low-level heating over the east Pacific, which might be an overestimate owing to shallow convection (Grecu et al., 2009). Low-level heating over the eastern Pacific is also present with smaller amplitudes in Q1 from ERA-I and LH from CFSR.

The distribution of boundary heating of the LH from CFSR is almost the same, and it may also be related to precipitating marine stratus clouds over the ocean (VanZanten et al., 2005). The two heating peaks are more obvious in some TRMM and other reanalysis data, such as LH from CSH and SLH and Q1 from ERA-I (Fig. 9). The double peak in the heating field has previously been observed and discussed (Zhang and Hagos, 2009; Takayabu et al., 2010). It is reasonable to say that the upper peak is related to precipitation by cold (ice or mixed phase) clouds and the lower one to precipitation by warm (liquid phase) clouds. LH in TRAIN and CFSR and Q1 in JRA25 do not have any obvious double-peak structure.

Ling and Zhang (2011) also pointed out that the discrepancies among the heating data sets are not merely between the TRMM and reanalysis data sets or between LH and Q1. Differences within the TRMM and reanalysis products, respectively, and within various products of LH or Q1 are no less than those between the TRMM and reanalysis data and between LH and Q1. These differences reflect our current level of estimating diabatic heating fields: we may get some basic properties of the heating field (e.g., longitudinal locations of maxima) correct, but there are many details with large uncertainties. These uncertainties should by no means stop us from cautiously using the currently available heating products to provide as much information as they may credibly provide.

4.2. Helping in the development of microphysics (MP) schemes

CRMs with advanced microphysical schemes have been used to study the interactions between aerosol, cloud and precipitation processes at high resolution. These processes play a critical role in the global water and energy cycle. Validation of CRMs with observational databases is important both to ascertain the fidelity of the outputs and to improve the models.

The interest on this topic lies in that there are still many uncertainties associated with various microphysics schemes. In part, this reflects the fact that microphysical processes cannot always be measured (or observed) directly. Herein cloud properties, which can be

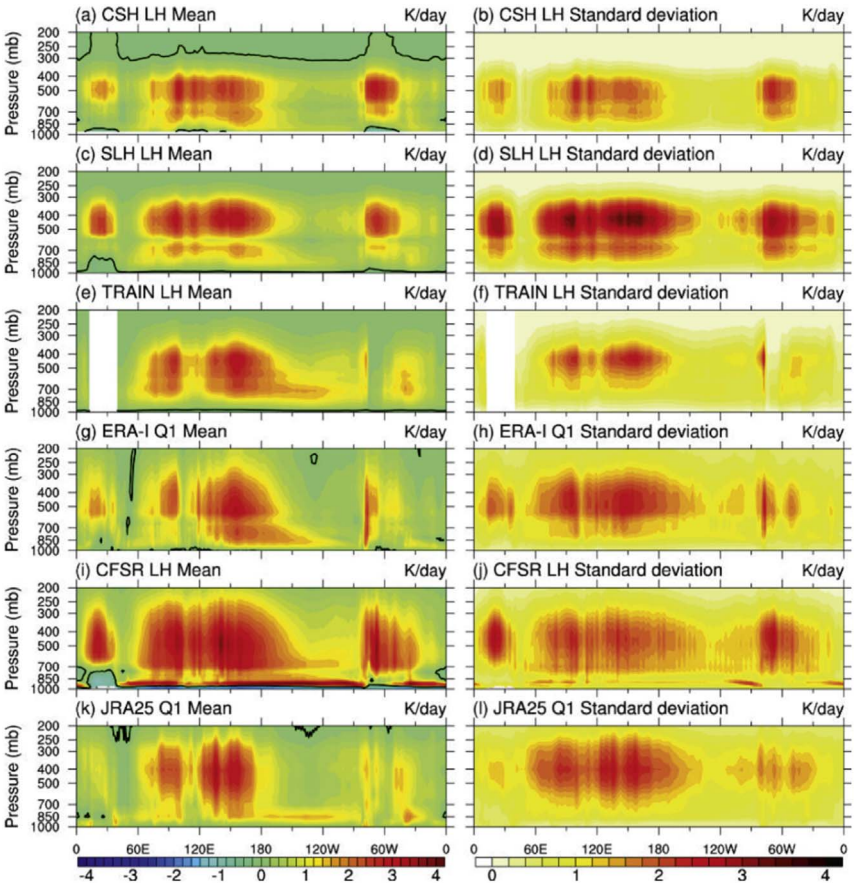

Fig. 9. A comparison of latent heat (LH) retrieval from (a) Goddard Convective-Stratiform Heating (CSH) algorithm, (b) Spectral Latent Heating (SLH) algorithm, (c) Goddard Trained Radiometer (TRAIN) algorithm and observations: ERA-I Q1 mean, CFSR LH mean and JRA25 Q1 mean. The good performances illustrate the usefulness of observations to improve the physics within climate models.

estimated, have been used to validate model results.

The spectral bin microphysical (SBM) schemes represent the most sophisticated representations of microphysical processes. They generally perform better in simulating realistic cloud properties and surface precipitation compared with bulk microphysical schemes (Li et al., 2010). SBM schemes have helped to improve the bulk scheme (Lang et al., 2014; Tao et al., 2016a, also RAMS microphysics used bin scheme to parameterize their cloud activation).

However useful, SBM schemes are not perfect. For example, in Fig. 10 we show nine years of TRMM PR and 85 GHz TMI observations of squall systems during late spring/early summer over the central US compared against a SBM simulated squall line case (PRESTORM campaign of 1985; Li et al., 2010). The figure shows improvements of a spectral bin microphysical scheme using long-term TRMM satellite observations (fully described in Li et al. (2010)). The columns are for three different instruments: the ground-based C-band radar (first column), the TRMM Precipitation Radar (second column), and the TRMM microwave imager at 85 GHz (third column). Note that the ground-based radar refers to the same case as the model simulation, whereas TRMM PR and TMI are a 9-year composite of TRMM observations at the same location (the Southern Great Plains) and period (late spring/early summer).

Comparisons against a surface C-band radar (first column) and the TRMM PR radar (second column) show an over estimation of radar reflectivity in the original scheme (second row). To improve the simulated radar reflectivity profiles, the temperature dependence of the collection efficiency between ice-phase particles, especially those of the plate-type, was modified. This improvement reduced the coalescence of various ice-phase particles and produced smaller aggregates, resulting in better radar CFADs comparisons in the stratiform region, as shown in the third row. Note that the SBM MP scheme is more direct (and realistic) than the bulk MP parameterizations used in GCMs. Uncertainties in these can be expected to be larger.

With the increase of computer power, one of the fastest developments in the modeling community is the high-resolution global cloud-resolving models (G-CRM). The spatial resolutions of these global models are rapidly approaching those of the traditional CRMs and thus may soon replace traditional limited-area Regional Climate Models (RCMs), which by their nature present several limitations.

In spite of the demonstrated past ability of RCMs to derive climatologies suitable for geographical applications (Tapiador et al., 2011), and of existing efforts to coordinate RCMs ensembles at regional scale (such as CORDEX) escalating computing resources are making RCMs redundant as coupled GCMs and G-CRMs can now be run at high spatial and temporal resolutions. Besides, there are strong conceptual reasons that favor global modeling and variable-resolution, convection-permitting models such as COSMO (cfr. Nuissier et al., 2016), namely RCM issues in providing correct large-scale atmospheric circulation across a region (Trenberth, 2007); and the critical dependence on the type of driving data, the climate variable, and the region used (Di Luca et al., 2016). Moreover, some RCMs perform poorly (correlation − 0.2) in describing key hydrological variables such as the dependence between the number of snowfall days and temperature (Pons et al., 2016), and in modeling evaporation/precipitation feedbacks (Lucarini et al., 2007) raising serious concerns about its use to inform policies.

To conclude this section, it is worth mentioning that a prospective NASA satellite mission called the Cloud and Precipitation Processes Mission (or CaPPM) is currently being proposed. Such a mission could provide global estimations of cloud and precipitation properties, which are needed to evaluate and improve dynamical and microphysical parameterizations, and the feedbacks between them in both CRMs and global G-CRM.

4.3. The role of ground validation (GV) campaigns

On one hand, CRMs provide simulated 4D data sets for improving the performances of rainfall, snowfall and LH retrieval algorithms. On the other hand, CRMs (and G-CRMs) need to be validated and improved by observations. To leverage resources, ground validation (GV) sites are set. These are intended to collect high quality, detailed information and data from several sources during the periods of interest. They are a third way of improving GCM physics.

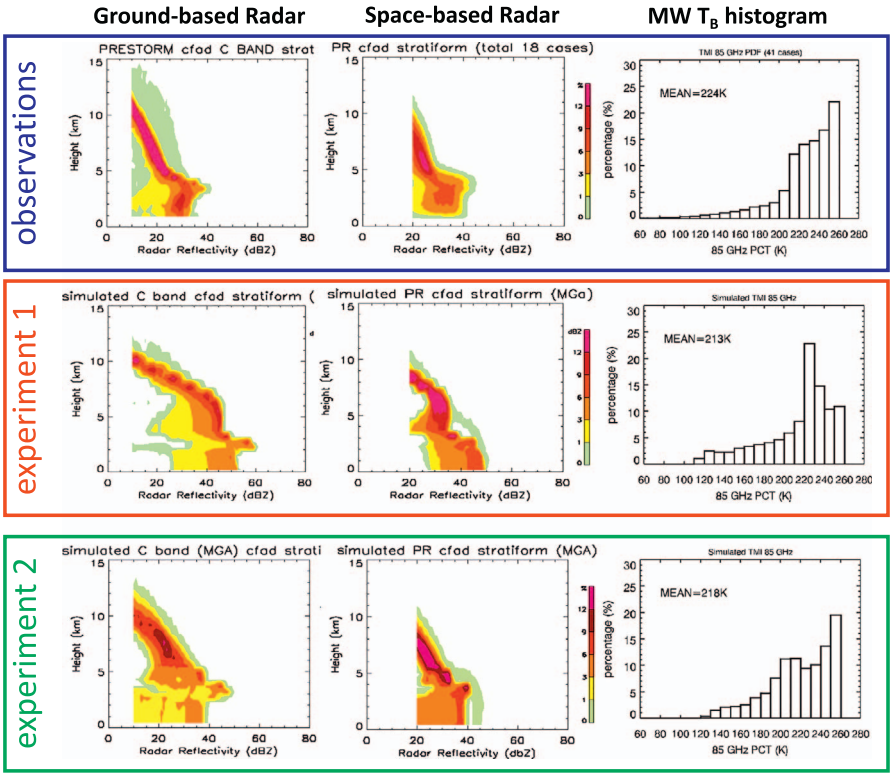

Fig. 10. An example of the improvements and sensitivity of a spectral bin microphysical scheme using observations. The top row is the observations; the middle row is the original SBM simulations; the third row is the improved SBM simulations by modifying the size threshold separating aggregates from graupel, and the ice crystal collection of cloud droplets. Each column represents one instrument. The left column is the surface C-band radar reflectivity CFADs in stratiform region during the PRESTORM field campaign; the middle column is the CFADs of the averaged TRMM PR observations for all the squall cases identified over a 9-year period; the right column is the distributions of TRMM TMI 85 GHz polarization corrected brightness temperature for the same cases as the middle column. Elaborated from Li et al. (2010).

In addition to collecting datasets conducive to physically validating GPM precipitation retrieval algorithms, some of their major objectives include the use of CRMs in precipitation science such as: (1) testing the fidelity of CRM simulations via intensive statistical comparisons between simulated and observed cloud properties and LH fields for a variety of case types, (2) establishing the limits of CRM space-time integration capabilities for quantitative precipitation estimates, and (3) supporting the development and refinement of physically-based GMI (GPM Microwave Imager), DPR and DPR-GMI combined retrieval algorithms using ground-based GPM GV Ku-Ka band radar and CRM simulations. Both GCE (Tao et al., 2014) and NU-WRF (Peters-Lidard et al., 2015) were used for GPM GV campaigns in real time forecasts (Tao et al., 2013; Wu et al., 2016) and for validating and improving the performance of model simulations (Shi et al., 2010; Iguchi et al., 2012a, 2014).

Fig. 11 exemplifies the need of such exercises to improve climate models. It shows a case that used GPM GV to validate and improve the performance of simulating snow events during of the GV field campaign (C3VP). The upper part of the figure shows horizontal distribution of vertically maximum C-band radar reflectivity (in dB) over an area between the Georgian Bay of Lake Huron and Lake Ontario at 1200 UTC on January 20, 2007 derived from King-City radar measurements (right) and simulated by the WRF-SBM and Goddard Satellite Data Simulator Unit (G-SDSU; left).

The G-SDSU (Matsui et al., 2014) computes satellite-consistent radiances or backscattering signals from simulated atmospheric profiles and condensates consistent with the microphysics. Fig. 11 shows how a band of high reflectivity forming from the south edge of the Georgian Bay and extending to Lake Ontario side. This band was due to a snowstorm formed by the interaction between cold dry air blowing over the lake and heat/moisture supplied from the relatively warm water surface of the lake (lake-effect snowstorm). The lower part of the figure shows the scatter plot diagrams between airborne-instrument-based bulk density (vertical axis) and bulk effective radius (horizontal axis) of snow particles sampled in the lake-effect snowstorm case. The left panel shows the scatter plot derived from actual aircraft measurements on the day (red dots) and those simulated in the two simulations of the WRF-SBM, i.e. the control run including snow riming effects (blue dots) and a test run excluding snow riming effects (green dots); both control and test runs employed the Mellor-Yamada-Janjic PBL (Planetary Boundary Layer) scheme. The right panel shows the scatter plots simulated in a WRF-SBM run employing Medium Range Forecast PBL scheme (green dots) as well as that from the aircraft measurements (red dots).

The broad distribution of the red dots in vertical directions indicates that the snowstorm is composed of non-rimed low-density and rimed dense snow particles. The WRF-SBM control run (blue dots) was not able to simulate non-rimed low-density particles included in the measurement plots. This bias is related to the strength of heat/moisture fluxes from the lake surface through the PBL process to form the lake-effect snowstorm.

The existence of the relationship between the riming process and the PBL process is demonstrated by results of the sensitivity test employing the different PBL scheme (green dots in the left panel). Once again, the case illustrates the many nuances in using observations to validate and improve climate models.

These exercises illustrate not only the challenges faced while validating climate models with observations, but also the potential and usefulness of such endeavors. The examples are numerous. Thus for instance, and regarding just another component of the hydrological cycle, Lin (2014) illustrated the value of the estimates of subgrid-scale humidity variability to improve the representation of clouds in GCMs.

5. Discussion: On the need of quality control (QC) standards for climate model validation

As seen above, validating precipitation estimates from climate models requires a good knowledge of the details of the observational databases in terms of their limitations, applicability and uncertainties. In order to make such comparisons as transparent as possible, and to drive model improvements in models using a physical approach instead of relying on statistical tuning, it is necessary to establish a set of well-defined protocols and procedures. These have to be objective, transparent, traceable and easy to apply for everyone involved in this emerging field.

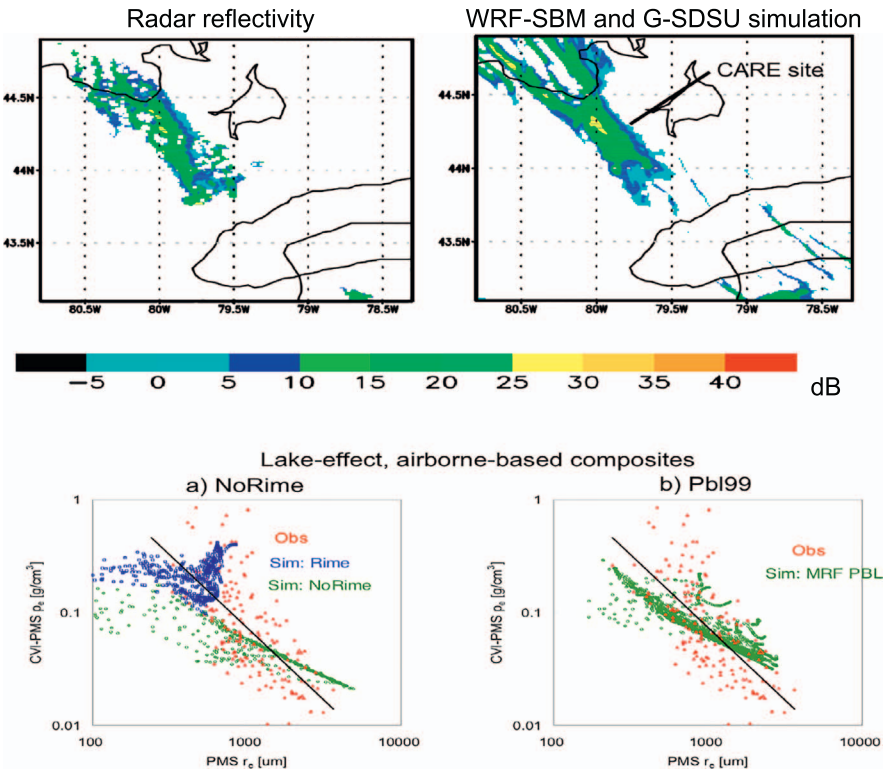

Fig. 11. (top) Horizontal distribution of vertically maximum C-band radar reflectivity (dB) at 1200 UTC on January 20, 2007 derived from King-City radar measurements (upper left) and simulated using WRF-SBM and G-SDSU (upper right). (bottom) Scatter plot diagrams between airborne-instrument-based bulk density (vertical axis) and bulk effective radius (horizontal axis) of snow particles sampled in the lake-effect snowstorm case; the plots are derived from aircraft measurements (red dots) and those simulated in the three types simulations for WRF-SBM, i.e. the control run including snow riming effects (blue dots), a test run excluding snow riming effects (green dots on lower-left side), and a run employing Medium Range Forecast PBL scheme in place of the Mellor-Yamada-Janjic PBL scheme (green dots on lower-right side). From Iguchi et al., 2012b. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Besides, the use of climate models for a variety of societal needs such as dam design, urban water management and actuary activities, has motivated the development of quality criteria so that final users, as well as scientists, are fully informed of the capabilities and limitations of the products.

Both precipitation databases and outputs from climate models should be considered within quality control (QC) standards. This is an increasingly pressing requirement as climate becomes more and more interweaved with activities of mitigation and adaptation to global warming. The public and the decision makers demand that the science behind policies is traceable, transparent and auditable.

A note on terminology is in order as the meaning of 'quality' differs between communities. In industry 'quality' is precisely defined by the quality assurance (QA) concept. The word is not addressed to the intrinsic value of the product but to 'providing confidence that quality requirements will be fulfilled' (e.g., ISO 9000). Thus, QA implies not only that the product is suited for the specific purpose it was built (a data set or a climate model in our case) but also that the product has been created following a well-defined set of rules and methods than builds confidence in the whole production process. QA procedures are designed to minimize errors and mistakes, setting double-blind evaluations, sanity checks and providing a traceable flow of the several stages of the process of generating the product. It is worth clarifying that passing QA does not imply a good performance of the product. All it means is that the product conforms to standards. Another issue worth mentioning here is the topic of convergence. Convergence is not intrinsically a good estimate of observational product or model performance since all the developers are equally aware of the range of sensible values for their product and model outputs.

To achieve a QA-standard each step of the production process has to be clearly defined and subject to auditing. This is not a problem for most merged precipitation data sets since these are carefully designed products whose science can be traced back to an Algorithm Theoretical Basis Document (ATBD). The ATBDs are the cornerstone of the confidence in merged precipitation data sets, in the same way that metadata and technical notes perform for pure observational data sets. They provide the rationale of the many decisions taken over the process of developing the product, and allow users to trace back each step, also permitting duplication of the product by another party.

Reputable climate models also have the equivalent to the ATBD in form of model documentation describing the physics of the dynamical core, the numerical methods employed, the parameterizations, and the empirical choices used to fine-tune the model. A numerical model lacking such information can be considered 'in house', 'internal code' or 'research code' but ought not be used to inform public policies or to derive consensus conclusions such as those of the IPCC.

Documentation, however, is just a first step for building confidence in data sets and models. Additional requirements can be proposed to reinforce credibility. Table 2 gathers three tiers of those possible requirements divided into basic, extended and full.

The basic requirements are very simple. In addition to being documented, the products (databases and models) should have been described in scientific journals. The standard criteria for quality here is peer-review and indexing such as in Journal Citation Reports (JCR). Journals in the top of the categories are obviously more credible than those in the bottom, but as a basic requirement, publication suffices. It is important however that the model or the algorithm is fully described: cases in which the model is used with others and provides similar performances are not enough. The algorithm or the modeling strategy has to be submitted to public scrutiny in order to evaluate the applicability of the product for a certain application.

The second basic requirement refers to more in-depth scientific discussion. Proprietary systems and 'in house' algorithms are useful for preliminary climate research but can provide only a minimum confidence until the data are public and can be confronted into the peer-review and post-publication processes. Algorithms and models falling under the 'basic requirements' tag indicate little more than the existence of the product, but the tier is nonetheless relevant as some climate models and data sets do not even satisfy this simple specification.

The extended level of confidence includes more requirements. Extended confidence is tied to how the product compares with others. Data or models yielding radically different results from similar products, or models just replicating the annual cycle should be taken with caution. Replicability is one of the cornerstones of science. Within this tier, results from case studies are a plus as they allow evaluating performances that can remain hidden to statistical comparisons.

Full confidence includes additional, stronger requirements. The probability of bugs in the code decreases exponentially with the number of eyes scrutinizing the code. A community of users also reduces to almost zero the chance of malpractice, incompetence or fraud. It is therefore sensible that proprietary models whose source code no one has ever seen and which are used by just a few people in the world, are not put on the same tier as another used by a community of hundreds of developers and whose code is downloadable for anyone to check.

To achieve full confidence, it is also important to be able to evaluate the degree of novelty of the product. Some allegedly independent or new models are actually 'avatars' of a well-known public code which has been tailored. There is nothing wrong in these practices as long as there is a reason for doing so and the procedure is transparent. However, using such models within an ensemble may not contribute added value to the experiment because of the lack of independence and

Table 2
Basic, extended and full set of requirements of data and climate models involved in validation.

Tier	Requirements	
	Validation data	Climate models
Tier 1: Basic confidence	ATBD available	Full technical documentation of the model core is available (incl. dynamical core, numerical methods, parameterization, and empirical coefficients used for tuning)
	Algorithms or data set creation methodology has been described in peer-reviewed literature	The model has been fully described in a peer-reviewed journal
Plus		
Tier 2: Extended confidence	Products are publicly and freely available	Outputs are publicly and freely available
	There are published comparisons with other data set	Aggregated results compare reasonably well with pre-existing models
	Case studies are available	Benchmarking and case studies are available
Plus		
Tier 3: Full confidence	The code for the algorithms is publicly available/replicable	Source code is publicly available
		The results of the simulations are fully replicable by a third party
	The data set has been successfully used by several independent groups and the results have been published in reputable journals	The model has been validated by several independent groups and the results have been published in reputable journals

diversity, a quality very much sought when building an ensemble. Just reporting the parameterization type is not enough: most parameterizations contain a choice of values for key variables. Since in the development phase models are evaluated in specific locations, no one can assure that the parameterization would work in another place. Besides, the same parameterization can be wired in the model in a different way depending on programming practices or computing resources, and it is always helpful to have the code examined by a fresh pair of eyes. Again, the more people have gone through thousands of lines of code, the better for the community.

The top requirements for quality assurance are in the line of the outputs or products having been used and replicated by a third party. A model used by several independent groups and whose results are routinely published in top journals cannot be in the same category as other models developed by a few people and used just by them. Ideally, the validation should be as independent as possible; preferably by another group. Consequently, the QA for those models must be clearly different. At this point, it is worth mentioning that once science ventures into providing input for policies only models with the higher confidence should be used.

Examples of full confidence level data sets using those criteria would include the GPCP database, on the data side; and the CESM on the model side. CESM satisfies all the requirements described above. Many versions of the code are routinely run in several independent research centers, and there is a community improving the model and reporting results, study cases and the unavoidable bugs featuring in complex and large projects.

The same applies to databases such as GPCP, which anyone with adequate technological resources can replicate, and that have been examined by a large community. Such dissemination builds up confidence in the data to be the best possible at the time, and therefore suitable for been used to derive public policies.

6. Conclusions

The validation of climate models is critical to improve the physics of such tools and to gauge its ability to provide useful input to decision makers. However, observational, reference data sets present a series of uncertainties and limitations. Awareness of those is a must to avoid pitfalls and extract wrong or unsubstantiated advice on the spatial and temporal distribution of precipitation. Also, there is a lack of standards, best practices and quality assurance methods for both observation data sets and models.

This paper has elaborated over both aspects. Table 3 gathers the central points that require attention in the task of validating precipitation outputs from climate models.

A main conclusion is that since there are no “perfect” data sets that are optimal in both sampling and retrieval error, it would be always desired to use multiple precipitation products as the reference to verify model simulations so the audience can grasp the magnitude of measurement uncertainties. The IPCC AR5 acknowledgment of large observational uncertainties in precipitation observations for climate model validation still applies, and the more we discover about precipitation physics the more questions arise. Comparisons considering the uncertainties and limitations of the reference data are rare and cumbersome, but nonetheless necessary.

Another major conclusion is that there is the need for setting up quality standards to ensure products are judiciously used. Simple schemes such as the three-tiers scheme we propose can help to understand how suited for a specific application a model or data set is. Thus, a regional precipitation data set might be helpful in the model development stages but may not be that useful for extracting conclusions at global scale over a full climatological period. Here, topics such as the habit of validating at a spatial or temporal resolution and then claim performances at better resolutions are highly relevant and deserve full attention.

As of 2017, it is safe to state that the future of climate modeling points towards high-resolution, coupled, convection-permitting, full-physics GCMs projects lead by large institutions. Modeling efforts are directed towards variable resolution GCMs and HR-GCMs. Consequently, validation efforts should be directed towards improving the physics sustaining such models, and here global precipitation databases can play a determinant role. Physical validation, more complex and costly than direct validation, and often involving setting up

Table 3
A checklist of known issues that must be considered in the field of validation of precipitation outputs from climate models.

1. Rain gauges provide pointwise estimates that may be not fully representative of the area, especially for large areas with a few observations (e.g., the Amazon basin).
2. Rain gauges have known technical limitations and biases and the spatial distribution/length record of the instruments is highly variable.
3. Ground radars are characterized by many sources of uncertainty (i.e., beam blockage, attenuation) that should be taken into account.
4. Precipitation (solid, liquid and mixed phase) has a large spatial and temporal variability making its validation both challenging and important using precipitation datasets in model validation.
5. Satellite estimates are indirect and have limited temporal sampling, and this should be considered in the comparisons.
6. Satellites estimates over land, coast and ocean are derived using different methods and assumptions.
7. Merged precipitation databases are not intended for trend analyses as sensors drift and/or are available over limited time spans.
8. Many of the techniques used in Level-2 products are built upon Bayesian estimates (i.e., they require a prior estimate).
9. The quality of Level-3 precipitation products is driven by microwave observations and therefore is dependent on their availability and quality.
10. The error characteristics resulting from the merging of disparate datasets are not well known.
11. All climate models are tuned to observations, and this must be considered for ensuring a truly independent validation.
12. Global measurements of microphysics are important to avoid overfitting models to empirical parameters.
13. There are known uncertainties in the estimation of diabatic heating fields that affect how models represent some precipitation processes.
14. Model outputs that have been bias-corrected or that are the results of model output statistic techniques cannot be validated.
15. Series derived from GCM-driven RCMs cannot be directly compared with time series of observations.
16. High-resolution global cloud-resolving models (G-CRM) are becoming best suited than RCMs to inform policies and advance our knowledge of the physics of precipitation.
17. Parameterizations can only be validated with data not used in their development and tuning.
18. ‘Scope principle’: a model cannot claim performances at better resolutions that those at which it has been validated.
19. Blending methods in deriving global precipitation products involves subtleties than must be considered in any validation process.
20. There are significant latitudinal differences in the satellite and ground based estimates in terms of known biases and uncertainties.
21. Parameters and techniques used in the estimation process using satellites and rain gauges may not be universally applicable, both in space and time.
22. Ground validation campaigns are essential for improving the representation of precipitation in models.
23. End-to-end characteristics of the satellite-based retrieval process are not yet fully understood.
24. There is less agreement among satellite products in trends and variability at global scale than in regional variability.
25. The precise measurement of shallow and very light precipitation still represents a scientific challenge.
26. While precipitation is a key variable to validate models, there is not agreement in the reference to be compared with. More research and targeted observations are required to fill this gap.
27. Public auditing of model code and precipitation databases algorithms is required if models are used for policy-making and societal applications other than pure research.
28. Every aspect of model and database development should be subject to QC methods and be fully traceable, transparent and auditable.
29. Models must be independently validated by scientists not involved in their development or belonging to the same research network.
30. Users should be made fully aware of the confidence level that can be attributed to model outputs and observational databases.

dedicated ground validation campaigns are important to advance the field.

Until the end-to-end characteristics of the satellite-based precipitation retrieval process are fully understood and parameterized in a dynamic fashion (dynamic since the overall errors are aggregated from always-changing satellite orbital patterns), it will be difficult to *quantitatively* compare GCM models and satellite precipitation data beyond more than daily precipitation totals, threshold scores and precipitation patterns. The future lies in extracting the information content within the multiple active and PMW data, to assess physical processes (cfr. Michaelides et al., 2015) interior to the GCM models (Skofronick-Jackson et al., 2017).

Validation of climate models is full of subtleties and details that quite often are only obvious to data set developers. Thus, for instance it is premature to use global scale trends and variations from observations to validate climate models as there is still some disagreement among the satellite products themselves. Given the extent of tuning that models have there is a danger of overfitting parameters to observations that are known to present its own uncertainties, biases and limitations. The application of those tuned models to a future climate under other set of emissions and forcings may not yield robust conclusions and thus be counterproductive.

A major reason for being prescriptive in the area of climate model validation and verification is reinforcing accountability. From the moment science is used to inform policies it can become political and therefore subject to increased public scrutiny. That means increased levels of quality assurance. Therein, awareness of the specifics of model tuning is critical to further societal advance. In an era of information, transparency and public responsibility, it is important to base policies on nothing but the best information. Both modelers and precipitation data set producers should make their codes as public as possible so everyone can check the assumptions made, evaluate the performances, and examine the empirical assumptions. It would be also advisable requiring that validation is done independently from model developers.

Finally, it is worth mentioning the validation must be performed with independent data not previously used for model development. While this may seem obvious, the fact is too often forgotten. A bias-corrected model cannot be fairly validated. On this respect, ongoing missions such as GPM provide a wealth of continuously updated information and planned missions such as CaPPM can serve to test hypotheses in the extremely demanding task of precisely quantifying precipitation over the globe in 4D.

Acknowledgments

Funding from projects CGL2013-48367-P, CGL2016-80609-R (Ministerio de Economía y Competitividad), UNCM08-1E-086 (Ministerio de Ciencia e Innovación), and CYTEMA (UCLM) is gratefully acknowledged by FJT (Tapiador) and AN. AN acknowledges FPU13/02798 grant and VL is grateful for support from the EU under grant agreement 603608 (eartH2Observe).

References

Adirosi, E., Baldini, L., Roberto, N., Gatlin, P., Tokay, A., 2016. Improvement of vertical profiles of raindrop size distribution from micro rain radar using 2D video disdrometer measurements. Atmos. Res. 169, 404–415. http://dx.doi.org/10.1016/j.atmosres.2015.07.002.

Adler, R.F., Negri, A.J., 1988. A satellite infrared technique to estimate tropical convective and stratiform rainfall. J. Appl. Meteorol. 27, 30–51. http://dx.doi.org/10.1175/1520-0450(1988)027<0030:ASITTE>2.0.CO;2.

Adler, R.F., Negri, A.J., Keehn, P.R., Hakkarinen, I.M., 1993. Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data. J. Appl. Meteorol. 32, 335–356. http://dx.doi.org/10.1175/1520-0450(1993)032<0335:EOMROJ>2.0.CO;2.

Adler, R.F., Huffman, G.J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., Nelkin, E., 2003. The version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). J. Hydrometeorol. 4, 1147–1167. http://dx.doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2.

Adler, R., Sapiano, M., Huffman, G., Bolvin, D., Wang, J.-J., Gu, G., Nelkin, E., Xie, P., Chiu, L., Ferraro, R., Schneider, U., Becker, A., 2016. New Global Precipitation Climatology Project monthly analysis product corrects satellite data drifts. GEWEX News 26 (4), 7–9.

Allen, M.A., Ingram, W.J., 2002. Constraints on future changes in climate and the hydrologic cycle. Nature 419, 224–232. http://dx.doi.org/10.1038/nature01092.

Arkin, P., Meisner, B., 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982–84. Mon. Weather Rev. 115, 51–74. http://dx.doi.org/10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2.

Barrett, E.C., Kidd, C., Bailey, J.O., 1987. The use of SMMR data in support of the Bristol/NOAA interactive scheme (BIAS) for satellite improved rainfall monitoring. Annual Rep. U.S. Dept. of Commerce, Cooperative Agreement NA86AA-H-RA001. (77 pp).

Barros, A.P., Petersen, W., Schwaller, M., Cifelli, R., Mahoney, K., Peters-Lidard, C., Shepherd, M., Nesbitt, S., Wolff, D., Heymsfield, G., Starr, D., Anagnostou, E., Gourley, J.J., Kim, E., Krajewski, W., Lackman, G., Lang, T., Miller, D., Mace, G., Petters, M., Smith, J., Tao, W.-K., Tsay, S.-C., Zipser, E., 2014. NASA GPM-Ground Validation: Integrated Precipitation and Hydrology Experiment 2014 Science Plan. NASAhttp://dx.doi.org/10.7924/G8CC0XMR.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., Ziese, M., 2013. A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. Earth Syst. Sci. Data 5, 71–99. http://dx.doi.org/10.5194/essd-5-71-2013.

Behrangi, A., Imam, B., Hsu, K., Sorooshian, S., Bellerby, T.J., Huffman, G.J., 2010. REFAME: rain estimation using forward-adjusted advection of microwave estimates. J. Hydrometeorol. 11, 1305–1321. http://dx.doi.org/10.1175/2010JHM1248.1.

Bell, T.L., Abdullah, A., Martin, R.L., North, G.R., 1990. Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. J. Geophys. Res. 95, 2195. http://dx.doi.org/10.1029/JD095iD03p02195.

Bellerby, T.J., 2006. High-resolution 2-D cloud-top advection from geostationary satellite imagery. IEEE Trans. Geosci. Remote Sens. 44, 3639–3648. http://dx.doi.org/10.1109/TGRS.2006.881117.

Betts, A.K., Jakob, C., 2002. Evaluation of the diurnal cycle of precipitation, surface thermodynamics, and surface fluxes in the ECMWF model using LBA data. J. Geophys. Res. 107, 8045. http://dx.doi.org/10.1029/2001JD000427.

Bretherton, C.S., 2015. Insights into low-latitude cloud feedbacks from high-resolution models. Phil. Trans. R. Soc. A 373 (2054). http://dx.doi.org/10.1098/rsta.2014.0415.

Bringi, V.N., Chandrasekar, V., 2001. Polarimetric Doppler Weather Radar: Principles and Applications, 1st Ed. Cambridge Univ. Press, Cambridge (664 pp).

Chandrasekar, V., Wang, Y., Chen, H., 2012. The CASA quantitative precipitation estimation system: a five year validation study. Nat. Hazards Earth Syst. Sci. 12, 2811–2820. http://dx.doi.org/10.5194/nhess-12-2811-2012.

Chang, W.Y., Wang, T.-C.C., Lin, P.-L., 2009. Characteristics of the raindrop size distribution and drop shape relation in typhoon systems in the western pacific from the 2D video disdrometer and NCU C-band polarimetric radar. J. Atmos. Ocean. Technol. 26, 1973–1993. http://dx.doi.org/10.1175/2009JTECHA1236.1.

Chen, M., Xie, P., Janowiak, J.E., Arkin, P.A., Chen, M., Xie, P., Janowiak, J.E., Arkin, P.A., 2002. Global land precipitation: a 50-yr monthly analysis based on gauge observations. J. Hydrometeorol. 3, 249–266. http://dx.doi.org/10.1175/1525-7541(2002)003<0249:GLPAYM>2.0.CO;2.

Ciach, G.J., 2003. Local random errors in tipping-bucket rain gauge measurements. J. Atmos. Ocean. Technol. 20, 752–759. http://dx.doi.org/10.1175/1520-0426(2003)20<752:LREITB>2.0.CO;2.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healey, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-interim reanalysis: configuration and performance of the data assimilation system. Quart. J. Roy. Meteor. Soc. 137, 553–597. http://dx.doi.org/10.1002/qj.828.

Di Luca, A., Argüeso, D., Evans, P., de Elía, R., Laprise, R., 2016. Quantifying the overall added value of dynamical downscaling and the contribution from different spatial scales. J. Geophys. Res. 121. http://dx.doi.org/10.1002/2015JD024009.

Duchon, C.E., Biddle, C.J., 2010. Undercatch of tipping-bucket gauges in high rain rate events. Adv. Geosci. 25, 11–15. http://dx.doi.org/10.5194/adgeo-25-11-2010.

Duchon, C.E., Essenberg, G.R., 2001. Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields. Water Resour. Res. 37, 3253–3263. http://dx.doi.org/10.1029/2001WR000541.

Ebert, E.E., Manton, M.J., Arkin, P.A., Allam, R.J., Holpin, G.E., Gruber, A., 1996. Results from the GPCP algorithm intercomparison programme. Bull. Am. Meteorol. Soc. 77, 2875–2887. http://dx.doi.org/10.1175/1520-0477(1996)077<2875:RFTGAI>2.0.CO;2.

Elsaesser, G.S., Kummerow, C., 2008. Toward a fully parametric retrieval of the nonraining parameters over the global oceans. J. Appl. Meteorol. Climatol. 47, 1599–1618. http://dx.doi.org/10.1175/2007JAMC1712.1.

Elsaesser, G.S., Kummerow, C.D., 2015. The Sensitivity of Rainfall Estimation to Error Assumptions in a Bayesian Passive Microwave Retrieval Algorithm. J. Appl. Meteor. Climatol. 54, 408–422. http://dx.doi.org/10.1175/JAMC-D-14-0105.1.

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., Husak, G., 2015. A global satellite-assisted precipitation climatology. Earth Syst. Sci. Data 7, 275–287. http://dx.doi.org/10.5194/essdd-8-401-2015.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R.,

Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A., Gu, W., Kim, G., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. 0.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Climate 0. http://dx.doi.org/10.1175/JCLI-D-16-0758.1.

Goodison, B.E., Louie, P.Y.T., Yang, D., 1998. WMO solid precipitation measurement intercomparison. In: Final Report. WMO/TD-No. 872, (212 pp).

Grecu, M., Olson, W.S., Shie, C.L., L'Ecuyer, T.S., Tao, W.-K., 2009. Combining satellite microwave radiometer and radar observations to estimate atmospheric heating profiles. J. Clim. 22, 6356–6376. http://dx.doi.org/10.1175/2009JCLI3020.1.

Groisman, P.Y., Legates, D.R., 1995. Documenting and detecting long-term precipitation trends: where we are and what should be done. Clim. Chang. 31, 601–622. http://dx.doi.org/10.1007/BF01095163.

Guo, H., Chen, S., Bao, A., Behrangi, A., Hong, Y., Ndayisaba, F., Hu, J., Stepanian, P.M., 2016. Early assessment of Integrated multi-satellite retrievals for global precipitation measurement over China. Atmos. Res. 176–177, 121–133. http://dx.doi.org/10.1016/j.atmosres.2016.02.020.

Habib, E., Krajewski, W.F., Ciach, G.J., 2001. Estimation of rainfall interstation correlation. J. Hydrometeorol. 2, 621–629. http://dx.doi.org/10.1175/1525-7541(2001)002<0621:EORIC>2.0.CO;2.

Haddad, Z.S., Smith, E.A., Kummerow, C.D., Iguchi, T., Farrar, M.R., Durden, S.L., Alves, M., Olson, W.S., 1997. The TRMM "Day-1" radar/radiometer combined rain-profiling algorithm. J. Meteor. Soc. Japan. 75, 799–809.

Haddad, Z.S., Sawaya, R.C., Kacimi, S., Sy, O.O., Turk, F.J., Steward, J., 2017. Interpreting millimeter-wave radiances over tropical convective clouds. J. Geophys. Res. 122, 1650–1664. http://dx.doi.org/10.1002/2016JD025923.

Hegerl, G.C., Black, E., Allan, R.P., Ingram, W.J., Polson, D., Trenberth, K.E., Chadwick, R.S., Arkin, P.A., Sarojini, B.B., Becker, A., Dai, A., Durack, P., Easterling, D., Fowler, H., Kendon, E., Huffman, G.J., Liu, C., Marsh, R., New, M., Osborn, T.J., Skliris, N., Stott, P.A., Vidale, P.-L., Wijffels, S.E., Wilcox, L.J., Willett, K.M., Zhang, X., 2015. Challenges in quantifying changes in the global water cycle. Bull. Am. Meteorol. Soc. 96, 1097–1115. http://dx.doi.org/10.1175/BAMS-D-13-00212.1.

Herzegh, P.H., Jameson, A.R., 1992. Observing precipitation through dual-polarization radar measurements. Bull. Am. Meteorol. Soc. 73, 1365–1374. http://dx.doi.org/10.1175/1520-0477(1992)073<1365:OPTDPR>2.0.CO;2.

Hilburn, K.A., Wentz, F.J., 2008. Intercalibrated passive microwave rain products from the Unified Microwave Ocean Retrieval Algorithm (UMORA). J. Appl. Meteorol. Climatol. 47, 778–794. http://dx.doi.org/10.1175/2007JAMC1635.1.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95, 701–722. http://dx.doi.org/10.1175/BAMS-D-13-00164.1.

Houze Jr., R.A., McMurdie, L., Petersen, W., Schwaller, M., Baccus, W., Lundquist, J., Mass, C., Nijssen, B., Rutledge, S., Hudak, D., Tanelli, S., Mace, J., Poellot, M., Lettenmaier, D., Zagrodnik, J., Rowe, A., DeHart, L., Maddaus, L., Barnes, H., 2017. Olympic Mountains Experiment (OLYMPEX). Bull. Am. Meteorol. Soc (submitted).

Huffman, G.J., Adler, R.F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., Schneider, U., 1997. The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. Bull. Am. Meteorol. Soc. 78, 5–20. http://dx.doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Huffman, G.J., Adler, R.F., Morrissey, M.M., Bolvin, D.T., Curtis, S., Joyce, R., McGavock, B., Susskind, J., 2001. Global precipitation at one-degree daily resolution from multisatellite observations. J. Hydrometeorol. 2, 36–50. http://dx.doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8, 38–55. http://dx.doi.org/10.1175/JHM560.1.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., 2009. Improving the global precipitation record: GPCP Version 2.1. Geophys. Res. Lett. 36 (L17808). http://dx.doi.org/10.1029/2009GL040000.

Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM multi-satellite precipitation analysis (TMPA). In: Hossain, F., Gebremichael, M. (Eds.), Satellite Rainfall Applications for Surface Hydrology. Springer Netherlands, Dordrecht, pp. 3–22. http://dx.doi.org/10.1007/978-90-481-2915-7_1.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., 2015. Algorithm Theoretical Basis Document (ATBD) Version 4.5: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). NASA, Greenbelt, MD, USA.

Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., Okamoto, K., 2000. Rain-profiling algorithm for the TRMM precipitation radar. J. Appl. Meteorol. 39, 2038–2052. http://dx.doi.org/10.1175/1520-0450(2001)040<2038:RPAFTT>2.0.CO;2.

Iguchi, T., Matsui, T., Shi, J.J., Tao, W.-K., Khain, A.P., Hou, A., Cifelli, R., Heymsfield, A., Tokay, A., 2012a. Numerical analysis using WRF-SBM for the cloud microphysical structures in the C3VP field campaign: impacts of supercooled droplets and resultant riming on snow microphysics. J. Geophys. Res. 117, D23206. http://dx.doi.org/10.1029/2012JD018101.

Iguchi, T., Matsui, T., Tokay, A., Kollias, P., Tao, W.-K., 2012b. Two distinct modes in one-day rainfall event during MC3E field campaign: analysis of disdrometer observations and WRF-SBM simulation. Geophys. Res. Lett. 39, L24805. http://dx.doi.org/10.1029/2012GL053329.

Iguchi, T., Matsui, T., Tao, W.-K., Khain, A., Phillips, V.T.J., Kidd, C., L'Ecuyer, T., Braun, A.A., Hou, A., Schwaller, M.R., 2014. Numerical simulations using WRF-SBM for mixed-phase precipitation and consequent bright band structure observed in the LPVEx field campaign. J. Appl. Meteorol. Climatol. 53, 2710–2731. http://dx.doi.org/10.1175/JAMC-D-13-0334.1.

IPCC, 2014. Climate change 2014: synthesis report. In: Team, Core Writing, Pachauri, R.K., Meyer, L.A. (Eds.), Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva, Switzerland (151 pp).

Jaffrain, J., Studzinski, A., Berne, A., 2011. A network of disdrometers to quantify the small-scale variability of the raindrop size distribution. Water Resour. Res. 47, W00H06. http://dx.doi.org/10.1029/2010WR009872.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5, 487–503. http://dx.doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Katsanos, D., Retalis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. Atmos. Res. 169, 459–464. http://dx.doi.org/10.1016/j.atmosres.2015.05.015.

Kidd, C., 2001. Satellite rainfall climatology: a review. Int. J. Climatol. 21, 1041–1066. http://dx.doi.org/10.1002/joc.635.

Kidd, C., Huffman, G., 2011. Global precipitation measurement. Meteorol. Appl. 18, 334–353. http://dx.doi.org/10.1002/met.284.

Kidd, C., Levizzani, V., 2011. Status of satellite precipitation retrievals. Hydrol. Earth Syst. Sci. 15, 1109–1116. http://dx.doi.org/10.5194/hess-15-1109-2011.

Kidd, C., Kniveton, D.R., Todd, M.C., Bellerby, T.J., Kidd, C., Kniveton, D.R., Todd, M.C., Bellerby, T.J., 2003. Satellite rainfall estimation using combined passive microwave and infrared algorithms. J. Hydrometeorol. 4, 1088–1104. http://dx.doi.org/10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2.

Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C., Randel, D., 2016. Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme. J. Hydrometeorol. 17, 383–400. http://dx.doi.org/10.1175/JHM-D-15-0051.1.

Kidd, C., Becker, A., Huffman, G.J., Muller, C.L., Joe, P., Skofronick-Jackson, G., Kirschbaum, D.B., 2017. So, how much of the Earth's surface *is* covered by rain gauges? Bull. Am. Meteorol. Soc. 98, 69–78. http://dx.doi.org/10.1175/BAMS-D-14-00283.1.

Kim, K., Park, J., Baik, J., Choi, M., 2017. Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. Atmos. Res. 187, 95–105. http://dx.doi.org/10.1016/j.atmosres.2016.12.007.

Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V., Petrovic, P., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int. J. Climatol. 22, 1441–1453. http://dx.doi.org/10.1002/joc.773.

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y.N., Ushio, T., Nakagawa, K., Iwanami, K., Kachi, M., Okamoto, K., 2007. Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: production and validation. IEEE Trans. Geosci. Remote Sens. 45, 2259–2275. http://dx.doi.org/10.1109/TGRS.2007.895337.

Kummerow, C., Olson, W.S., Giglio, L., 1996. A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. IEEE Trans. Geosci. Remote Sens. 34, 1213–1232. http://dx.doi.org/10.1109/36.536538.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., Simpson, J., 1998. The Tropical Rainfall Measuring Mission (TRMM) sensor package. J. Atmos. Ocean. Technol. 15, 809–817. http://dx.doi.org/10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2.

Kummerow, C., Simpson, J., Thiele, O., Barnes, W., Chang, A.T.C., Stocker, E., Adler, R.F., Hou, A., Kakar, R., Wentz, F., Ashcroft, P., Kozu, T., Hong, Y., Okamoto, K., Iguchi, T., Kuroiwa, H., Im, E., Haddad, Z., Huffman, G., Ferrier, B., Olson, W.S., Zipser, E., Smith, E.A., Wilheit, T.T., North, G., Krishnamurti, T., Nakamura, K., 2000. The Status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. J. Appl. Meteorol. 39, 1965–1982. http://dx.doi.org/10.1175/1520-0450(2001)040<1965:TSOTTR>2.0.CO;2.

Kummerow, C., Hong, Y., Olson, W.S., Yang, S., Adler, R.F., McCollum, J., Ferraro, R., Petty, G., Shin, D.-B., Wilheit, T.T., 2001. The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. J. Appl. Meteorol. 40, 1801–1820. http://dx.doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2.

Kummerow, C., Masunaga, H., Bauer, P., 2007. A next-generation microwave rainfall retreival algorithm for use in TRMM and GPM. In: Levizzani, V., Bauer, P., Turk, F.J. (Eds.), Measuring Precipitation from Space. Springer, New York, NY (745 pp).

Lang, S., Tao, W.-K., Chern, J.-D., Wu, D., Li, X., 2014. Benefits of a 4th ice class in the simulated radar reflectivities of convective systems using a bulk microphysics scheme. J. Atmos. Sci. 71, 3583–3612. http://dx.doi.org/10.1175/JAS-D-13-0330.1.

Li, X., Tao, W.-K., Matsui, T., Liu, C., Masunaga, H., 2010. Improving a spectral bin microphysical scheme using trmm satellite observations. Quart. J. Roy. Meteor. Soc. 136, 382–399. http://dx.doi.org/10.1002/qj.569.

Li, N., Tang, G., Zhao, P., Hong, Y., Gou, Y., Yang, K., 2017a. Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis IMERG in Ganjiang River basin. Atmos. Res. 183, 212–223. http://dx.doi.org/10.1016/j.atmosres.2016.07.020.

Li, J., Wu, K., Li, F., Chen, Y., Huang, Y., Feng, Y., 2017b. Effects of latent heat in various cloud microphysics processes on autumn rainstorms with different intensities on Hainan Island, China. Atmos. Res. 189, 47–60. http://dx.doi.org/10.1016/j.atmosres.2017.01.010.

Liebmann, B., Allured, D., 2005. Daily precipitation grids for South America. Bull. Am. Meteorol. Soc. 86, 1567–1570. http://dx.doi.org/10.1175/BAMS-86-11-1567.

Lin, Y., 2014. Humidity variability revealed by a sounding array and its implications for

cloud representation in GCMs. J. Geophys. Res. 119, 10499–10514. http://dx.doi.org/10.1002/2014JD021837.
Ling, J., Zhang, C., 2011. Structural evolution in heating profiles of the MJO in global reanalyses and TRMM retrievals. J. Clim. 24, 825–842. http://dx.doi.org/10.1175/2010JCLI3826.1.
Lucarini, V., Danihlik, R., Kriegerova, I., Speranza, A., 2007. Does the Danube exist? Versions of reality given by various regional climate models and climatological data sets. J. Geophys. Res. 112, D13103. http://dx.doi.org/10.1029/2006JD008360.
Maggioni, V., Meyers, P., Robinson, M., 2016. A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. J. Hydrometeorol. 17, 1101–1117. http://dx.doi.org/10.1175/JHM-D-15-0190.1.
Matsui, T., Santanello, J., Shi, J.J., Tao, W.-K., Wu, D., Peters-Lidard, C., Kemp, E., Chin, M., Starr, D., Sekiguchi, M., Aires, F., 2014. Introducing multisensor satellite radiance-based evaluation for regional Earth System modeling. J. Geophys. Res. 119, 8450–8475. http://dx.doi.org/10.1002/2013JD021424.
Melcón, P., Merino, A., Sánchez, J.L., López, L., García-Ortega, E., 2017. Spatial patterns of thermodynamic conditions of hailstorms in southwestern France. Atmos. Res. 189, 111–126. http://dx.doi.org/10.1016/j.atmosres.2017.01.011.
Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., 2012. An overview of the Global Historical Climatology Network-daily database. J. Atmos. Ocean. Technol. 29, 897–910. http://dx.doi.org/10.1175/JTECH-D-11-00103.1.
Metcalfe, J.R., Routledge, B., Devine, K., 1997. Rainfall measurement in Canada: changing observational methods and archive adjustment procedures. J. Clim. 10, 92–101. http://dx.doi.org/10.1175/1520-0442(1997)010<0092:RMICCO>2.0.CO;2.
Michaelides, S., 2013a. Advances in precipitation science. Atmos. Res. 119, 1–2. http://dx.doi.org/10.1016/j.atmosres.2012.11.001.
Michaelides, S., 2013b. Perspectives of precipitation science: part I. Atmos. Res. 131, 1–2. http://dx.doi.org/10.1016/j.atmosres.2013.05.017.
Michaelides, S., 2014. Perspectives of precipitation science: part II. Atmos. Res. 144, 1–3. http://dx.doi.org/10.1016/j.atmosres.2015.09.025.
Michaelides, S., 2016. Perspectives of precipitation science: part III. Atmos. Res. 169, 401–403. http://dx.doi.org/10.1016/j.atmosres.2015.09.025.
Michaelides, S., Levizzani, V., Anagnostou, E., Bauer, P., Kasparis, T., Lane, J.E., 2009. Precipitation: measurement, remote sensing, climatology and modeling. Atmos. Res. 94, 512–533. http://dx.doi.org/10.1016/j.atmosres.2009.08.017.
Michaelides, S., Nastos, P., Flocas, H., 2015. Atmospheric processes in the Mediterranean. Atmos. Res. 152, 1–3. http://dx.doi.org/10.1016/j.atmosres.2014.07.027.
Mitchell, T.D., Jones, P.D., 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. Int. J. Climatol. 25, 693–712. http://dx.doi.org/10.1002/joc.1181.
Mitchell, J.F.B., Wilson, C.A., Cunnington, W.M., 1987. On CO_2 climate sensitivity and model dependence of results. Quart. J. Roy. Meteor. Soc. 113, 293–322. http://dx.doi.org/10.1002/qj.49711347517.
Neale, R.B., Richter, J.H., Jochum, M., 2008. The impact of convection on ENSO: from a delayed oscillator to a series of events. J. Clim. 21, 5904–5924. http://dx.doi.org/10.1175/2008JCLI2244.1.
Nešpor, V., Krajewski, W.F., Kruger, A., 2000. Wind-induced error of raindrop size distribution measurement using a two-dimensional video disdrometer. J. Atmos. Ocean. Technol. 17, 1483–1492. http://dx.doi.org/10.1175/1520-0426(2000)017<1483:WIEORS>2.0.CO;2.
New, M., Hulme, M., Jones, P., 2000. Representing twentieth-century space–time climate variability. Part II: development of 1901–96 monthly grids of terrestrial surface climate. J. Clim. 13, 2217–2238. http://dx.doi.org/10.1175/1520-0442(2000)013<2217:RTCSTC>2.0.CO;2.
New, M., Todd, M., Hulme, M., Jones, P., 2001. Precipitation measurements and trends in the twentieth century. Int. J. Climatol. 21, 1889–1922. http://dx.doi.org/10.1002/joc.680.
Nuissier, O., Marsigli, C., Vincendon, B., Hally, A., Bouttier, F., Montani, A., Paccagnella, T., 2016. Evaluation of two convection-permitting ensemble systems in the HyMeX Special Observation Period (SOP1) framework. Quart. J. Roy. Meteor. Soc. 142, 404–418. http://dx.doi.org/10.1002/qj.2859.
Oettli, P., Sultan, B., Baron, C., Vrac, M., 2011. Are regional climate models relevant for crop yield prediction in West Africa? Environ. Res. Lett. 6 (2011), 014008.
Paltridge, G.W., 1975. Global dynamics and climate – a system of minimum entropy exchange. Quart. J. Roy. Meteor. Soc. 101, 475–484. http://dx.doi.org/10.1002/qj.49710142906.
Peters-Lidard, C.D., Kemp, E.M., Matsui, T., Santanello, J.A., Kumar, S.V., Jacob, J.P., Clune, T., Tao, W.K., Chin, M., Hou, A., Case, J.L., Kim, D., Kim, K.M., Lau, W., Liu, Y., Shi, J., Starr, D., Tan, Q., Tao, Z., Zaitchik, B.F., Zavodsky, B., Zhang, S.Q., Zupanski, M., 2015. Integrated modeling of aerosol, cloud, precipitation and land processes at satellite-resolved scales. Environ. Model Softw. 67, 149–159. http://dx.doi.org/10.1016/j.envsoft.2015.01.007.
Platnick, S., King, M.D., Ackerman, S.A., Menzel, W.P., Baum, B.A., Riedi, J.C., Frey, R.A., 2003. The MODIS cloud products: algorithms and examples from Terra. IEEE Trans. Geosci. Remote Sens. 41, 459–473. http://dx.doi.org/10.1109/TGRS.2002.808301.
Pons, M.R., Herrera, S., Gutiérrez, J.M., 2016. Future trends of snowfall days in northern Spain from ENSEMBLES regional climate projections. Clim. Dyn. 46, 3645–3655. http://dx.doi.org/10.1007/s00382-015-2793-9.
Qian, T., Dai, A., Trenberth, K.E., Oleson, K.W., 2006. Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations. J. Hydrometeorol. 7, 953–975. http://dx.doi.org/10.1175/JHM540.1.
Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A.P., Black, J., Thériault, J.M., Kucera, P., Gochis, D., Smith, C., Nitu, R., Hall, M., Ikeda, K., Gutmann, E., 2012. How well are we measuring snow: the NOAA/FAA/NCAR winter precipitation test bed. Bull. Amer. Meteor. Soc. 93, 811–829. http://dx.doi.org/10.1175/BAMS-D-11-00052.1.
Retalis, A., Katsanos, D., Michaelides, S., 2016. Precipitation climatology over the Mediterranean Basin - validation over Cyprus. Atmos. Res. 169, 449–458. http://dx.doi.org/10.1016/j.atmosres.2015.01.012.
Ryu, Y.-E., Smith, J.A., Baeck, M.L., Cunha, L.K., Bou-Zeid, E., Krajewski, W., 2016. The regional water cycle and heavy spring rainfall in Iowa: observational and modeling analyses from the IFloodS campaign. J. Hydrometeorol. 17, 2763–2784. http://dx.doi.org/10.1175/JHM-D-15-0174.1.
Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., Rudolf, B., 2014. GPCC's new land-surface precipitation climatology based on quality-controlled in-situ data and its role in quantifying the global water cycle. Theor. Appl. Climatol. 115, 15–40. http://dx.doi.org/10.1007/s00704-013-0860-x.
Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Ziese, M., 2015. GPCC Full Data Reanalysis Version 7.0 at 0.5°: Monthly Land-surface Precipitation from Rain-gauges Built on GTS-based and Historic Data. http://dx.doi.org/10.5676/DWD_GPCC/FD_M_V7_050.
Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., Becker, A., 2017. Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). Atmosfera 8, 52. http://dx.doi.org/10.3390/atmos8030052.
Scofield, R.A., Kuligowski, R.J., 2003. Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. Weather Forecast. 18, 1037–1051. http://dx.doi.org/10.1175/1520-0434(2003)018<1037:SAOOOS>2.0.CO;2.
Sevruk, B., Klemm, S., 1989. Catalogue of national standard precipitation gauges. In: WMO Instruments and Observing Methods, Rep. No. 39, (50 pp).
Shen, Y., Xiong, A., 2016. Validation and comparison of a new gauge-based precipitation analysis over mainland China. Int. J. Climatol. 36, 252–265. http://dx.doi.org/10.1002/joc.4341.
Shi, J.J., Tao, W.-K., Matsui, T., Hou, A., Lang, S., Peters-Lidard, C., Jackson, G., Cifelli, R., Rutledge, S., Petersen, W., 2010. Microphysical properties of the January 20-22 2007 snow events over Canada: comparison with in-situ and satellite observations. J. Appl. Meteorol. Climatol. 49, 2246–2266.
Simpson, J., Adler, R.F., North, G.R., 1988. A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. Bull. Am. Meteorol. Soc. 69, 278–295. http://dx.doi.org/10.1175/1520-0477(1988)069<0278:APTRMM>2.0.CO;2.
Simpson, J., Kummerow, C., Tao, W.-K., Adler, R.F., 1996. On the Tropical Rainfall Measuring Mission (TRMM). Meteorog. Atmos. Phys. 60, 19–36. http://dx.doi.org/10.1007/BF01029783.
Skofronick-Jackson, G., Petersen, W.A., Berg, W., Kidd, C.F., Stocker, E., Kirschbaum, D.B., Kakar, R., Braun, S.A., Huffman, G.J., Iguchi, T., Kirstetter, P.E., Kummerow, C., Meneghini, R., Oki, R., Olson, W.S., Takayabu, Y.N., Furukawa, K., Wilheit, T., 2017. The Global Precipitation Measurement (GPM) Mission for science and society. Bull. Am. Meteorol. Soc. http://dx.doi.org/10.1175/BAMS-D-15-00306.1.
Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H.V., Imam, B., Braithwaite, D., 2000. Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. Bull. Am. Meteorol. Soc. 81, 2035–2046. http://dx.doi.org/10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2.
Stephens, G.L., Vane, D.G., Boain, R.J., Mace, G.G., Sassen, K., Wang, Z., Illingworth, A.J., O'Connor, E.J., Rossow, W.B., Durden, S.L., Miller, S.D., Austin, R.T., Benedetti, A., Mitrescu, C., 2002. The CloudSat mission and the A-Train: a new dimension of space-based observations of clouds and precipitation. Bull. Am. Meteorol. Soc. 83, 1771–1790 (+1742. doi:10.1175/BAMS-83-12-1771).
Stephens, G.L., Vane, D.G., Tanelli, S., Im, E., Durden, S., Rokey, M., Reike, D., Partain, P., Mace, G.G., Austin, R., L'Ecuyer, T., Haynes, J., Lebsock, M., Suzuki, K., Waliser, D., Wu, D., Kay, J., Gettelman, A., Wang, Z., Marchand, R., 2008. CloudSat mission: performance and early science after the first year in orbit. J. Geophys. Res. 113, D00A18. http://dx.doi.org/10.1029/2008JD009982.
Sun, Y., Solomon, S., Dai, A., Portmann, R.W., 2006. How often does it rain? J. Clim. 19, 916–934. http://dx.doi.org/10.1175/JCLI3672.1.
Takayabu, Y.N., Shige, S., Tao, W.K., Hirota, N., 2010. Shallow and deep latent heating modes over tropical oceans observed with TRMM PR spectral latent heating data. J. Clim. 23, 2030–2046. http://dx.doi.org/10.1175/2009JCLI3110.1.
Tao, W.-K., Lang, S., Olson, W.S., Meneghini, R., Yang, S., Simpson, J., Kummerow, C., Smith, E., Halverson, J., 2001. Retrieved vertical profiles of latent heat release using TRMM rainfall products for February 1998. J. Appl. Meteorol. 40, 957–982. http://dx.doi.org/10.1175/1520-0450(2001)040<0957:RVPOLH>2.0.CO;2.
Tao, W.-K., Smith, E.A., Adler, R.F., Haddad, Z.S., Hou, A.Y., Iguchi, T., Kakar, R., Krishnamurti, T.N., Kummerow, C.D., Lang, S., Meneghini, R., Nakamura, K., Nakazawa, T., Okamoto, K., Olson, W.S., Satoh, S., Shige, S., Simpson, J., Takayabu, Y., Tripoli, G.J., Yang, S., 2006. Retrieval of latent heating from TRMM measurements. Bull. Am. Meteorol. Soc. 87, 1555–1572. http://dx.doi.org/10.1175/BAMS-87-11-1555.
Tao, W.-K., Santanello, J.A., Chin, M., Zhou, S., Tan, Q., Kemp, E.M., Peters-Lidard, C.D., 2013. Effect of land cover on atmospheric processes and air quality over the continental United States-a NASA Unified WRF (NU-WRF) model study. Atmos. Chem. Phys. 13, 6207–6226. http://dx.doi.org/10.5194/acp-13-6207-2013.
Tao, W.-K., Lang, S., Zeng, X., Li, X., Matsui, T., Mohr, K., Posselt, D., Chern, J., Peters-Lidard, C., Norris, P.M., Kang, I.S., Choi, I., Hou, A., Lau, K.M., Yang, Y.-M., 2014. The Goddard Cumulus Ensemble model (GCE): improvements and applications for studying precipitation processes. Atmos. Res. 143, 392–424. http://dx.doi.org/10.1016/j.atmosres.2014.03.005.
Tao, W.-K., Wu, D., Lang, S., Chern, J., Peters-Lidard, C., Fridlind, A., Matsui, T., 2016a. High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: further improvements and comparisons between Goddard

microphysics schemes and observations. J. Geophys. Res. 121, 1278–1305. http://dx.doi.org/10.1002/2015JD023986.
Tao, W.-K., Takayabu, Y.N., Lang, S., Olson, W., Shige, S., Hou, A., Jackson, G., Jiang, X., Lau, W., Krishnamurti, T., Waliser, D., Zhang, C., Johnson, R., Houze, R., Ciesielski, P., Grecu, M., Hagos, S., Kakar, R., Nakamura, K., Braun, S., Bhardwaj, A., 2016b. TRMM latent heating retrieval and comparison with field campaigns and large-scale analyses, chapter 2. In: AMS Meteorological Monographs - Multi-scale Convection-Coupled Systems in the Tropics, pp. 2.1–2.234 (http://journals.ametsoc.org/toc/amsm/56).
Tapiador, F.J., 2008. A physically based satellite rainfall estimation method using fluid dynamics modelling. Int. J. Remote Sens. 29, 5851–5862. http://dx.doi.org/10.1080/01431160802029677.
Tapiador, F.J., 2010. A joint estimate of the precipitation climate signal in Europe using eight regional models and five observational datasets. J. Clim. 23, 1719–1738. http://dx.doi.org/10.1175/2009JCLI2956.1.
Tapiador, F.J., Kidd, C., Hsu, K.-L., Marzano, F., 2004. Neural networks in satellite rainfall estimation. Meteorol. Appl. 11, 83–91. http://dx.doi.org/10.1017/S1350482704001173.
Tapiador, F.J., Checa, R., de Castro, M., 2010. An experiment to measure the spatial variability of rain drop size distribution using sixteen laser disdrometers. Geophys. Res. Lett. 37. http://dx.doi.org/10.1029/2010GL044120.
Tapiador, F.J., Angelis, C.F., Viltard, N., Cuartero, F., de Castro, M., 2011. On the suitability of regional climate models for reconstructing climatologies. Atmos. Res. 101, 739–751. http://dx.doi.org/10.1016/j.atmosres.2011.05.001.
Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., García-Ortega, E., Machado, L.A.T., Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J., de Castro, M., 2012. Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104, 70–97. http://dx.doi.org/10.1016/j.atmosres.2011.10.021.
Tapiador, F.J., Navarro, A., Moreno, R., Jiménez-Alcázar, A., Marcos, C., Tokay, A., Durán, L., Bodoque, J.M., Martín, R., Petersen, W., de Castro, M., 2017. On the optimal measuring area for pointwise rainfall estimation: a dedicated experiment with fourteen laser disdrometers. J. Hydrometeorol. http://dx.doi.org/10.1175/JHM-D-16-0127.1.
Tarnavsky, E., Grimes, D., Maidment, R., Black, E., Allan, R., Stringer, M., Chadwick, R., Kayitakire, F., 2014. Extension of the TAMSAT satellite-based rainfall monitoring over Africa and from 1983 to present. J. Appl. Meteorol. Climatol. 53, 2805–2822. http://dx.doi.org/10.1175/JAMC-D-14-0016.1.
Thurai, M., Gatlin, P.N., Bringi, V.N., 2016. Separating stratiform and convective rain types based on the drop size distribution characteristics using 2D video disdrometer data. Atmos. Res. 169, 416–423. http://dx.doi.org/10.1016/j.atmosres.2015.04.011.
Tokay, A., Short, D.A., 1996. Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. J. Appl. Meteorol. 35, 355–371. http://dx.doi.org/10.1175/1520-0450(1996)035<0355:EFTRSO>2.0.CO;2.
Trenberth, K., 2007. Medicanes? In: Climate Feedback; a Blog from Nature Climate Change, (Available at: http://blogs.nature.com/climatefeedback/2007/07/medicanes_1.html).
Turk, F.J., Rohaly, G., Hawkins, J., Smith, E.A., Marzano, F.S., Mugnai, A., Levizzani, V., 2000. Meteorological applications of precipitation estimation from combined SSM/I, TRMM and geostationary satellite data. In: Pampaloni, P., Paloscia, S. (Eds.), Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere. VSP Int. Sci. Publisher, Utrecht (The Netherlands), pp. 353–363.
Turk, F.J., Haddad, Z.S., You, Y., 2016. Estimating non-raining surface parameters to assist GPM constellation radiometer precipitation algorithms. J. Atmos. Ocean. Technol. 33, 1333–1353.
Ungersböck, M., Rubel, F., Fuchs, T., Rudolf, B., 2001. Bias correction of global daily rain gauge measurements. Phys. Chem. Earth Part B 26, 411–414. http://dx.doi.org/10.1016/S1464-1909(01)00027-2.
VanZanten, M.C.M.C., Stevens, B., Vali, G., Lenschow, D.H., 2005. Observations of drizzle in nocturnal marine stratocumulus. J. Atmos. Sci. 62, 88–106. http://dx.doi.org/10.1175/JAS3611.1.
Vicente, G.A., Scofield, R.A., Menzel, W.P., 1998. The operational GOES infrared rainfall estimation technique. Bull. Am. Meteorol. Soc. 79, 1883–1893. http://dx.doi.org/10.1175/1520-0477(1998)079<1883:TOGIRE>2.0.CO;2.
Viney, N.R., Bates, B.C., 2004. It never rains on Sunday: the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high quality data set. Int. J. Climatol. 24, 1171–1192. http://dx.doi.org/10.1002/joc.1053.
Vivekanandan, J., Zrnic, D.S., Ellis, S.M., Oye, R., Ryzhkov, A.V., Straka, J., 1999. Cloud microphysics retrieval using S-band dual-polarization radar measurements. Bull. Am. Meteorol. Soc. 80, 381–388. http://dx.doi.org/10.1175/1520-0477(1999)080<0381:CMRUSB>2.0.CO;2.
Voosen, P., 2016. Climate scientists open up their black boxes to scrutiny. Science 354 (6311). http://dx.doi.org/10.1126/science.354.6311.401.
Vuerich, E., Monesi, C., Lanza, L.G., Stagi, L., Lanzinger, E., 2009. WMO field intercomparisons of rainfall intensity gauges. In: Instruments and Observing Methods. Report 99. WMO/TD-No. 1504, (96 pp.).
Wang, J.J., Adler, R.F., Huffman, G.J., Bolvin, D., 2014. An updated TRMM composite climatology of tropical rainfall and its validation. J. Clim. 27, 273–284. http://dx.doi.org/10.1175/JCLI-D-13-00331.1.
Wilheit, T.T., Chang, A.T.C.V., Rao, M.S., Rodgers, E.B., Theon, J.S., 1977. A satellite technique for quantitatively mapping rainfall rates over the oceans. J. Appl. Meteorol. 16, 551–560. http://dx.doi.org/10.1175/1520-0450(1977)016<0551:ASTFQM>2.0.CO;2.
WMO, 2008. Guide Guide to Meteorological Instruments and Methods of Observation WMO-No. 8. 978-92-63-10008-5.
Wu, D., Peters-Lidard, C., Tao, W.-K., Petersen, W., 2016. Evaluation of NU-WRF rainfall forecasts for IFloodS. J. Hydrometeorol. 17, 1317–1335. http://dx.doi.org/10.1175/JHM-D-15-0134.1.
Xie, P., Arkin, P.A., 1997. Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. Bull. Am. Meteorol. Soc. 78, 2539–2558.
Xie, P., Rudolf, B., Schneider, U., Arkin, P.A., 1996. Gauge-based monthly analysis of global land precipitation from 1971 to 1994. J. Geophys. Res. 101, 19023. http://dx.doi.org/10.1029/96JD01553.
Xie, P., Janowiak, J.E., Arkin, P.A., Adler, R., Gruber, A., Ferraro, R., Huffman, G.J., Curtis, S., 2003. GPCP pentad precipitation analyses: an experimental dataset based on gauge observations and satellite estimates. J. Clim. 16, 2197–2214. http://dx.doi.org/10.1175/2769.1.
Xie, P., Chen, M., Shi, W., 2010. CPC unified gauge-based analysis of global daily precipitation. In: Prepr. 24th Conf. Hydrology, Amer. Meteor. Soc., Atlanta, GA.
Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., Kitoh, A., 2012. APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull. Am. Meteorol. Soc. 93, 1401–1415. http://dx.doi.org/10.1175/BAMS-D-11-00122.1.
Zhang, C., Hagos, S.M., 2009. Bi-modal structure and variability of large-scale diabatic heating in the Tropics. J. Atmos. Sci. 66, 3621–3640. http://dx.doi.org/10.1175/2009JAS3089.1.