2.3. Climate model validation

Francisco Tapiador[1] and Vincenzo Levizzani[2]

[1]University of Castilla-La Mancha (UCLM), Spain

[2]National Research Council of Italy-Institute of Atmospheric Sciences and Climate (CNR-ISAC)

As of 2020, climate models include traditional Global Circulation/Climate Models (GCMs), Regional Climate Models (RCMs) and, more recently, Earth System Models (ESMs). All these models are different from Numerical Weather Prediction (NWP) models in that they are intended to provide estimates of the climate rather than meteorological forecasts: they produce "climates", that is, mean values and other statistical moments of the distribution, but not realistic sequences of the actual weather. The World Meteorological Organization (WMO)'s standard length for constructing climatologies is 30 years, with the standard period being the 1960–1990 interval (the "historical climate"). More recently, the 1980–2010 period is used, which is more into the "satellite-era" of meteorological observations.

Climate model outputs are generically named "simulations". They are deemed "projections" when they aim to gauge the climates of the future based on a set of assumptions of anthropogenic forcings and therefore ultimately of social behavior (scenarios). Working in hindsight or to derive estimates of past geological eras, climate models produce "present-climate climatologies" and "paleoclimates" respectively. By going back in time, they can benefit from ancillary data and historical information to fine-tune the simulations, but in the case of experiments on future climate change, such advantage is naturally absent, so different challenges appear.

Validation of climate models is mostly circumscribed to simulations of the present-day climatologies (say 1960–1990) and comparison with meteorological observations. While paleoclimates can also be validated, the absence of an instrumental record obliges us to rely on proxies, which are indirect. Assuming the "rosy" and ceteris paribus assumptions (Smith, 2002), simulations of future climates are useful to understand our changing climate. Nonetheless, the first step to trust such simulations is ensuring models provide a faithful picture of current climate.

Model developers routinely compare their partial results in the developing phase with observations and adjust their models accordingly (Voosen, 2016). While this is unavoidable, it can hardly be considered as scientific validation. First, validation requires independence both in data not used to develop the model, and in the group of people who do the validation. However important self-validation by the developers is for model consistency, it is not a substitute for proper scientific scrutiny by independent, unrelated teams. Secondly, there is a set of protocols and standards that make validation a field in itself. Quality Control (QC) standards have been proposed for this field to address the increasingly pressing requirement as climate becomes more and more interwoven 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, and that includes the validation of climate model outputs.

A major issue in the field of validating present climates is that in the development stage, climate models are tuned to current conditions. The empirical values and assumptions implied in the procedure may vary in the future due to ongoing human emissions and land use changes, and therefore must be fully documented.

DOI: 10.13021/gewex.precip.2.3

Quality Control is now an integral part of climate model validation, as it has long been in the remote sensing field where International Organization for Standardization (ISO) requirements are strictly followed. Failure to do so may result in catastrophic mission failures. International standardization techniques are slowly permeating the procedures of climate validation. It needs to be noted that providing confidence that quality requirements will be fulfilled (e.g., ISO 9000) implies not only that the product is suited for the specific purpose it was conceived in the first place (a dataset or a climate model in our case), but also that the product has been created following a well-defined set of rules and methods that builds confidence in the whole production process. Quality Assurance (QA) procedures are designed to minimize errors and mistakes, setting double-blind evaluations and sanity checks and providing a traceable flow of several stages of the process of generating the product.

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 datasets, 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 datasets, in the same way that metadata and technical notes perform for pure observational datasets. 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 the 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.

Validating climate model outputs of prognostic variables such as temperature is a difficult exercise per se and becomes even more problematic when dealing with diagnostic variables such as precipitation. Precipitation has been considered “the ultimate test” for validating models for a number of reasons (Tapiador et al., 2019). However, it is difficult to find homogeneous precipitation datasets over reasonably long time periods that cover the whole globe. These are hard to find among ground-based observing systems, which are obviously limited to land areas. Indeed, until 2010 validations of climate models were done by the same teams that develop the software and for grid-point, rain gauge measurements only. Moving to independent, satellite and gauge-satellite combined data, Tapiador (2010) (see Figure 2.3.1) relied on increased availability of robust and public datasets such as those described in Tapiador et al. (2017). The latest published Intergovernmental Panel on Climate Change (IPCC) report, the Fifth Assessment Report (AR5) (IPCC, 2015), acknowledged that the detection and attribution of regional precipitation changes had generally focused on continental areas using in situ data because observational coverage over oceans was limited to a few island stations, although model-data comparisons over continents also illustrated large observational uncertainties. The report also noted that available satellite datasets that could supplement oceanic studies are short and their long-term homogeneity is still unclear, and that accordingly they have not yet been used for detection and attribution of changes. The IPCC concluded in 2014 that “continuing uncertainties in climate model simulations of precipitation make quantitative model/data comparisons difficult (e.g., Stephens et al., 2010), which also limits confidence in detection and attribution.”

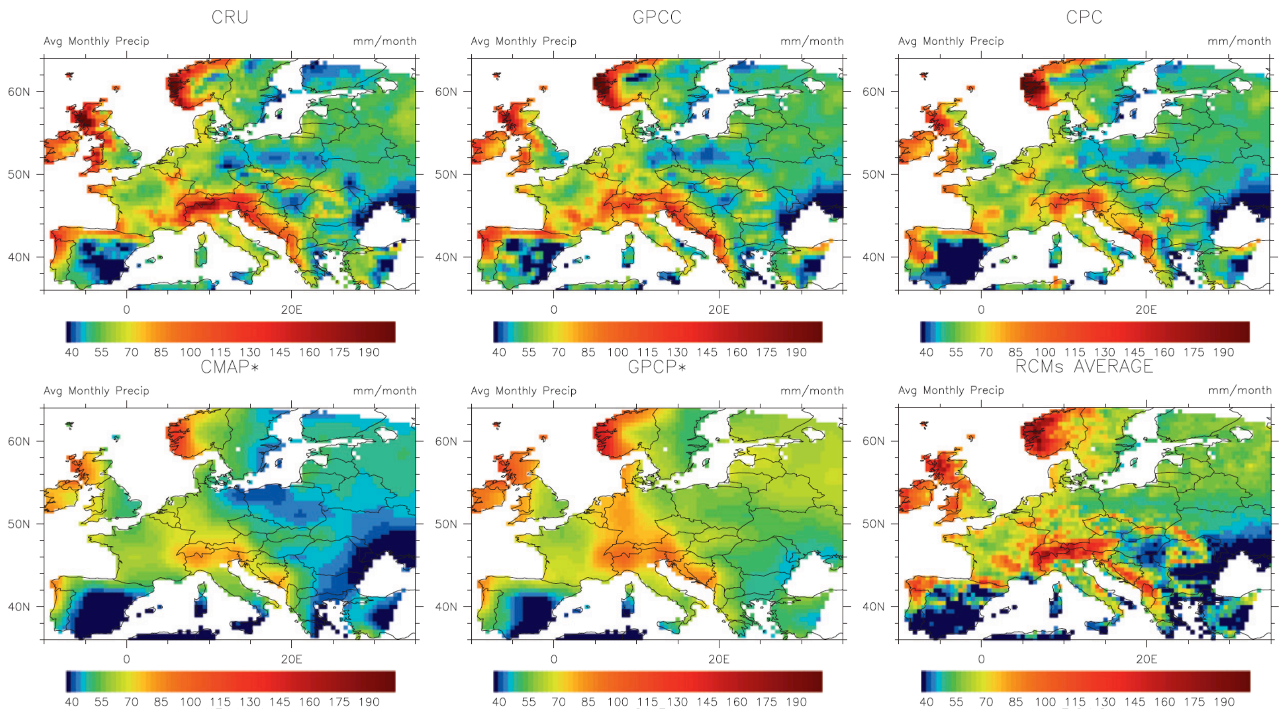

Figure 2.3.1. An example of the first validation of regional climate outputs using several precipitation datasets instead of gauge-only data. From Tapiador, 2010

Tapiador et al. (2017) first conducted a thorough analysis of the existing satellite-based precipitation datasets in view of their application for climate model validation. The authors provide guidance on the use of precipitation datasets for climate research, including model validation and verification for improving physical parameterizations. Strengths and limitations of the datasets for climate modeling applications are presented, underlining that not all datasets are suitable for this purpose. The checklist of items that must be considered in the field of validation of precipitation outputs from climate models includes several crucial points that we will now try to separate between facts and recommendations.

2.3.1. Facts

1. Rain gauges provide pointwise measurements that may not be fully representative of the area, especially for large areas with few observations (for example, 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 that have diverse natures (beam blockage, attenuation, anomalous propagation, etc.).
4. Precipitation (solid, liquid and mixed phase) has a large spatial and temporal variability.
5. Satellite estimates are indirect and have limited temporal sampling.
6. Satellite estimates over land, coast and ocean are derived using different methods and assumptions.

7. Merged precipitation databases are not intended for trend analyses as sensor drifts are present over limited time spans.
8. Many of the techniques used in Level-2 products are built upon Bayesian estimates (that is, they require an a priori estimate).
9. The quality of Level-3 precipitation products is driven by microwave observations and therefore is dependent on their availability and quality.
10. There are significant latitudinal differences in the satellite and ground-based estimates in terms of known biases and uncertainties.
11. The error characteristics resulting from the merging of disparate datasets are not well known.
12. There are known uncertainties in the estimation of diabatic heating fields that affect how models represent some precipitation processes.
13. Model outputs that have been bias-corrected or that are the results of model output statistic techniques cannot be validated.
14. Series derived from global circulation model (GCM)-driven regional climate models (RCMs) cannot be directly compared with time series of observations.
15. High-resolution global cloud-resolving models (G-CRM) are becoming better suited than RCMs to inform policies and advance our knowledge of the physics of precipitation.
16. End-to-end characteristics of the satellite-based retrieval process are not yet fully understood.
17. There is less agreement among satellite products in trends and variability at global scale than in regional variability.

2.3.2. Recommendations

i. Uncertainty figures from ground radars should be considered when assimilating data.
ii. Challenges posed to the validation of precipitation by its spatiotemporal variability need to be considered, paying attention when using precipitation datasets in model validation.
iii. The indirect character of satellite estimates and their limited temporal sampling should be considered in the comparisons.
iv. The tuning of the models with specific datasets must be considered for ensuring a truly independent validation.
v. Parameterizations must be validated with data not used in their development and tuning.
vi. Global measurements of microphysics are important to avoid overfitting models to empirical parameters.
vii. Ground validation campaigns are essential for improving the representation of precipitation in models.

viii. "Scope principle": a model cannot claim performances at better resolutions that those at which it has been validated.

ix. Blending methods in deriving global precipitation products involves subtleties than must be considered in any validation process.

x. Parameters and techniques used in the estimation process using satellites and rain gauges may not be universally applicable, both in space and time.

xi. The precise measurement of shallow and very light precipitation still represents a scientific challenge and more research is needed in this direction.

xii. While precipitation is a key variable to validate models, there is no agreement on the reference to be compared with. More research and targeted observations are required to fill this gap.

xiii. Public auditing of model code and precipitation database algorithms is required if models are used for policy-making and societal applications other than pure research.

xiv. Every aspect of model and database development should be subject to QC methods and be fully traceable, transparent and auditable.

xv. Models must be independently validated by scientists not involved in their development or belonging to the same research network.

xvi. Users should be made fully aware of the confidence level that can be attributed to model outputs and observational databases.

When considering the need for validation campaigns that produce new insights for a correct representation of precipitation processes in the models, under-represented areas and processes come first. This is particularly true for tropical forests. In this latter case, the problem is still to make sure that the correct microphysics is understood for its inclusion into the precipitation estimation algorithms prior to conducting any meaningful validation exercise. Recently the project Cloud Processes of the Main Precipitation Systems in Brazil: A Contribution to Cloud-Resolving Modeling and to the Global Precipitation Measurement (CHUVA), held in Brazil, has substantially contributed to improving our level of understanding in this direction (Machado et al., 2014). Another key example is the need for characterizing the relation of mid-latitude frontal precipitation mechanisms and their modification by terrain to rainfall estimation uncertainties. The Olympic Mountain Experiment (OLYMPEX) (Houze et al., 2017) assessed satellite measurements made by the GPM along the northeastern Pacific coastline. At the same time, warm rain processes are still not completely understood, and the Integrated Precipitation and Hydrology EXperiment (IPHEX) sought to characterize warm season orographic precipitation regimes and the relationship between precipitation regimes and hydrologic processes in regions of complex terrain (Erlingis et al., 2018).

There are a number of considerations that need to be made before approaching satellite precipitation datasets for model validation. Validation can be performed on the precipitation model means, which are those most used for applications, or on other first-order statistics. These include the modeling of the ENSO phenomenon (Neale et al., 2008), the representation of the diurnal cycle of rainfall (Betts and Jakob, 2002) and the frequency of occurrence of high- and low-intensity rainfall events (Sun et al., 2006). Validation can also be performed on more physical quantities such as latent heat (LH) and in the intricacies of the microphysics of precipitation in models.

1. **Latent heat release**. It is a consequence of phase changes between the vapor, liquid and frozen states of water, which cannot be measured or detected using present observational instruments. The vertical distribution of LH has, however, 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 midlatitude weather systems (Li et al., 2017). The launch of the Tropical Rainfall Measuring Mission (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 LH over the global tropics (Simpson et al., 1988, 1996). The success of TRMM made it possible to have another major precipitation measuring mission from the National Aeronautics and Space Administration (NASA), the Global Precipitation Measurement (GPM) mission. GPM is considered by NASA to be the centerpiece mission of its Global Water & Energy Cycle research program. On the modeling side, Cloud resolving models (CRMs) have been identified as being a valuable tool for algorithm developers and are 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) model 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 reanalyses [(ERA-I, Japanese 25-year ReAnalysis (JRA-25) and Climate Forecast System Reanalysis (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 in the others. Heating is weaker over the Maritime Continent than over the eastern Indian Ocean and western Pacific in some data [for example, apparent heat source Q_1 (Q_1), TRAIN LH, ERA-I Q_1, and JRA25 Q_1], 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 and Stevens, 2005). 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 Q_1 in JRA25 do not have any obvious double-peak structure. Ling and Zhang (2011) also pointed out that the discrepancies among the heating datasets are not merely between the TRMM and reanalysis datasets or between LH and Q_1. Differences within the TRMM

and reanalysis products, respectively, and within various products of LH or Q_1 are no less than those between the TRMM and reanalysis data and between LH and Q_1. These differences reflect our current level of estimating diabatic heating fields: we may get some basic properties of the heating field (for example, 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.

2. **Microphysics of precipitation**. 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). 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 in this topic lies in the 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 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 used the microphysics bin scheme from the Regional Atmospheric Modeling System (RAMS) to parameterize their cloud activation]. However useful, SBM schemes are not perfect, though they are more direct (and realistic) than the bulk MP parameterizations used in GCMs. Uncertainties in these can be expected to be larger.

There are also additional considerations to be made to validate climate models with precipitation datasets drawing on recent results. Apart from time span, spatial resolution and calibration quality of the data, there are very important subjects that need attention:

1. First, we need to consider the way the retrieval of precipitation was conducted. As stated by Stephens and Kummerow (2007), precipitation retrievals from space are highly sensitive to the specific radiative and microphysical model used in the retrieval process. Identifying a cloud as precipitating is not a trivial exercise and can result in large errors that make the dataset almost useless. This is the reason why most recent datasets including observations from passive and active sensors are necessary for improving cloud and precipitation retrievals (Levizzani et al., 2020a, 2020b).

2. The second crucial aspect concerns the precipitation phase. When validating climate models, precipitation type should be known, but this is a kind of knowledge that is far from being totally achieved. For the time being, only data from the MODerate resolution Imaging Spectroradiometer (MODIS) were used to verify climate model outputs such as in the case of Matiu et al. (2020) within the European Coordinated Regional Downscaling Experiment (EURO-CORDEX).

3. However, failure to accurately predict the location, magnitude and frequency of precipitation, including snowfall, heavily impacts climate modeling (for example, Field and Heymsfield, 2015). This is why a number of studies have recently started using observations and climate models to identify deficiencies in the actual modeling and validation approaches. Fowler et al. (2020), for example, show results that underscore the importance of evaluating clouds, their optical properties, and the top-of-the-atmosphere radiation budget in addition to precipitation when performing mesh refinement global simulations. Heymsfield et al. (2020) have been the first to produce a global view of the precipitation process partitioning, using a combination of satellite and global climate modeling data (Figure 2.3.2). They showed that significant differences between satellite- and model-based results are found and the reasons require investigations far more complex than the simple traditional surface precipitation differences. Note also that increasing temperatures may also imply increasing melting level height, which obviously impacts surface precipitation phase and intensity and thus model verification (Prein and Heymsfield, 2020). At the same time, care must be taken in improving frozen hydrometeors representation in the climate models (especially at the regional scale) since large discrepancies (up to 5 times) are found between modeled and observed brightness temperatures in the microwave that may undermine the value of intercomparison results (for example, Rysman et al., 2018).
4. Precipitation intensity at the ground is thus not sufficient to characterize the changing climate. The mean and the other statistical moments are just a first step in validation. Trenberth et al. (2003) argued that advancing understanding and the ability to model and predict the character of precipitation is vital and requires new approaches to examining data and models. The timing, duration and intensity of precipitation can be explored via the diurnal cycle (Betts and Jakob, 2002), whose correct simulation in models remains an unsolved challenge of vital importance in global climate change. This can only be done with truly global datasets such as those derived from satellite observations. Here, reanalyses are expected to play a dominant role in the near future.
5. Observational uncertainty quantification is essential for climate studies, climate model evaluation and statistical post-processing. Recently, Tang et al. (2020) have shown variable performance of the IMERG product with respect to other precipitation datasets and reanalyses. Prein and Gobiet (2017) have in turn demonstrated that differences between global precipitation datasets have the same magnitude as precipitation errors found in regional climate models.
6. Finally, precipitation is the most important process for a deeper understanding of climatic changes, but it is linked to several other processes within the water cycle and thus a combined use of precipitation datasets with other datasets (for example, soil moisture, sea surface temperature, evapotranspiration, wind fields, etc.) is unavoidable when assessing climate model outputs (Levizzani and Cattani, 2019).

After considering all possible caveats in using satellite precipitation datasets for climate model validation, nonetheless we register their increasing use in this field.

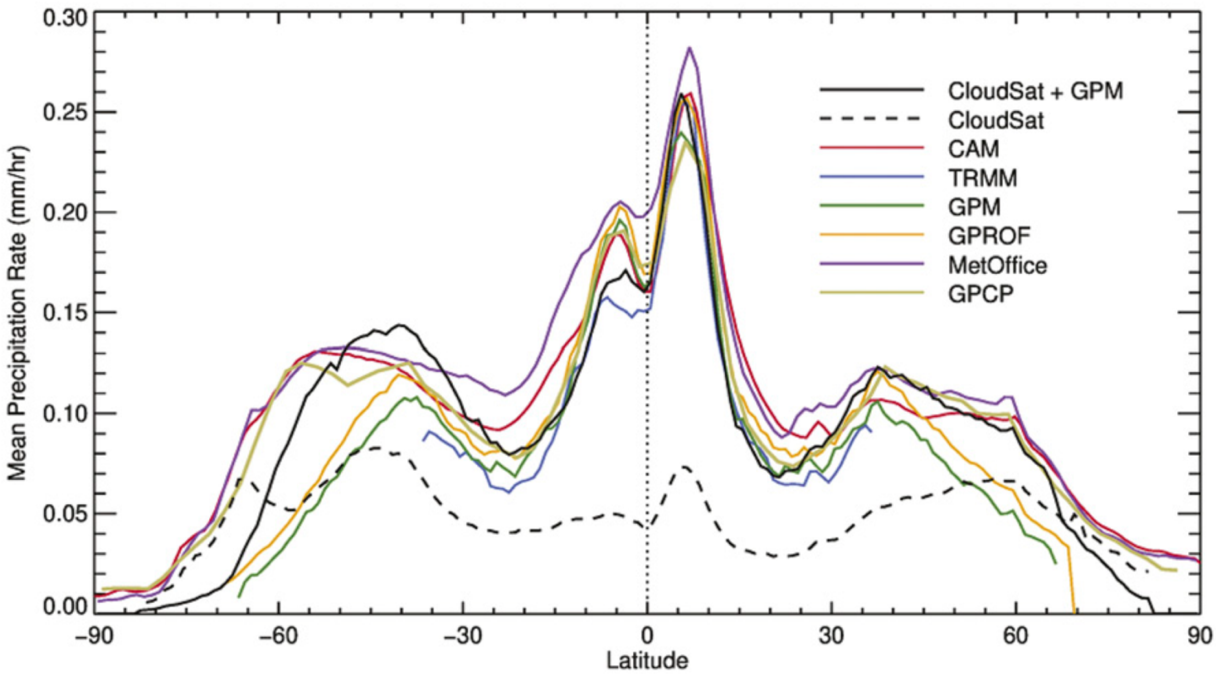

Figure 2.3.2. Mean surface precipitation rate retrieved from CloudSat, Global Precipitation Measurement (GPM) mission, Goddard Profiling (GPROF) algorithm, and Tropical Rainfall Measurement Mission (TRMM), derived from output from the Community Atmosphere Model (CAM) and the Met Office models, and from the Global Precipitation Climatology Project (GPCP) product. These are for land and ocean areas. [from Heymsfield et al. 2020; courtesy American Meteorological Society]. Note that model results are not independent from the satellite observations.

In the RCMs realm, satellite precipitation datasets were widely used during the Coordinated Regional Climate Downscaling Experiment (CORDEX) and the Coupled Model Intercomparison Project (CMIP) both sponsored by the World Climate Research Programme (WCRP) of the WMO. Target areas have been mostly Africa and Asia with special attention to the Tropics using a range of RCMs. Before that, data from the Ensemble-based Predictions of Climate Changes and their Impacts (ENSEMBLES) and Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects (PRUDENCE) projects were used to compare satellite datasets with RCM outputs. Biases of the single model depending on the African region and season were identified during CORDEX while simulating the West African summer monsoon (Akinsanola et al., 2015). In a previous study during CORDEX-Africa, Nikulin et al. (2012) showed that a multimodel average generally outperforms any individual simulation, showing biases of similar magnitude to differences across a number of observational datasets. At the same time, the authors confirmed that a common problem in the majority of the RCMs is that precipitation is triggered too early during the diurnal cycle with differences among the models as first suggested by Dai (2006). More recently, Wu et al. (2020) results show that improvements in the ability of RCMs to simulate precipitation in Africa compared to their driving reanalysis in many cases are simply related to model formulation and not necessarily to higher resolution. Such model formulation-related improvements are strongly model dependent.

The precipitation datasets have proved instrumental also in identifying deficiencies in the simulation of the CMIP5 models. Figure 2.3.3 shows the climatologies of 40 CMIP5 models and several precipitation reference data. Over the Tropics, model output and satellite precipitation dataset intercomparisons represent a substantial added value in identifying strengths and

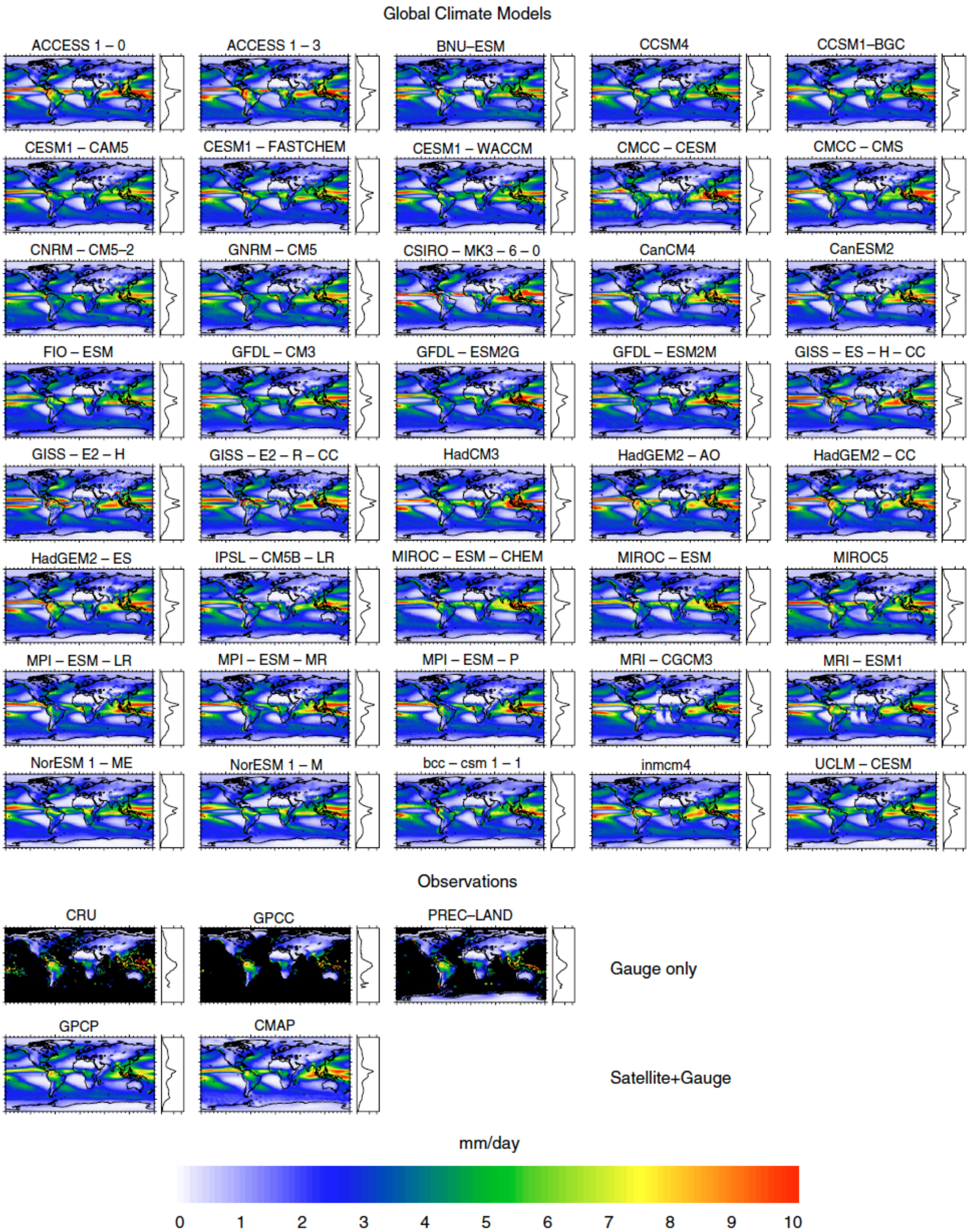

Figure 2.3.3. An example of the use of satellite-derived precipitation datasets in order to validate 40 Global Climate Models (GCMs) and Earth System Models (ESMs). The data represents the 1980–2000 climatology of the mean precipitation. From Tapiador et al. (2018)

weaknesses of the approaches of the various models to predict water cycle changes. For example, when considering one of the key convective areas of the planet, the Congo basin in Central Africa, there appears to be little agreement as to the distribution and quantity of rainfall across the basin with datasets differing by an order of magnitude in some seasons

(Washington et al., 2013). Higher-resolution satellite data can surely help in disentangling persisting uncertainties in the area. Over the ocean, the satellite precipitation datasets have contributed to find that precipitation in CMIP5 models is overestimated in most areas (Yang et al., 2018). This is consistent with the previous results of Hirota and Takayabu (2013), who concluded that a proper representation of the sensitivity of deep convection to humidity and higher resolution of the ocean models with better equatorial trades are important for reducing the double ITCZ and the cold tongue biases.

Another key aspect of climate model studies is the need for verifying improvements among different versions of the models and of their ensembles. An example of the application of satellite precipitation datasets is provided by Kumar et al. (2014), whose results show little change in the central tendency, variability, uncertainty of historical skills or consensus across CMIP3 and CMIP5. At the same time, there are regions and seasons where significant changes, performance improvements and even degradation in skills are suggested. This fact clearly demonstrates the potential of using satellite-derived datasets at the global as well as the regional scales. In fact, Pathak et al. (2019) have found that over the south Asian region, some of the convective and large-scale precipitation biases are common across CMIP5 model groups, emphasizing that although on a global scale the bias patterns may be sufficiently different to cluster the models into different groups, regionally, it may not be true.

Substantial work is still needed in examining the ability of both satellite observations and models in capturing extremes, droughts and floods. While observations need to be done at very high resolution and for long time periods, posing problems for using the available global datasets, models need to better represent convection. Kendon et al. (2019) argue that with a more accurate representation of convection in the models, projected changes in both wet and dry extremes over Africa may be more severe than they are actually predicted. Such conclusions need, however, to be linked to the way we conduct intercomparisons and numerical experiments. Recently, Yosef et al. (2020) have concluded that the use of percentile-based indices, such as those of the Expert Team on Climate Change Detection and Indices (ETCCDI), gives very different results when choosing a base period that included records from the last two decades (e.g., 1981–2010, 1988–2017) over Israel. At the same time, Alexander et al. (2019) argue that to advance the use of satellite precipitation data in the applications on extremes differences between data products, limitations in satellite-based estimation processes, and the inherent challenges of scale need to be better understood. Several efforts are on their way in recent times to come to the production of global datasets able to account for extremes (for example, Beck et al., 2020). We will most surely witness an intense activity in this field because it is crucial for climate change studies and to make sure that global precipitation datasets contain the necessary information for all types of events, especially the most severe ones that should be linked to changed climatic conditions.

Alongside research for clarifying the number and intensity of extremes per se, one more subject has come up in recent times. The understanding of short-duration rainfall extremes is also crucial, but data are often subject to errors and inhomogeneities and these events are poorly quantified in projections of future climate change. Consequently, knowledge of the processes contributing to intense, short-duration rainfall is less complete compared with those on daily timescales as argued by Blenkinsop et al. (2018) who launched the project INTENSE to overcome this lack of knowledge via sub-daily gauge dataset collection. Satellite datasets should help in this direction as they acquire higher temporal resolution.

2.3.3. References

Akinsanola, A.A., K.O. Ogunjobi, I.E. Gbode and V.O. Ajayi, 2015: Assessing the capabilities of three regional climate models over CORDEX Africa in simulating West African summer monsoon precipitation. *Advances in Meteorology*, 935431, https://doi.org/10.1155/2015/935431.

Alexander, L.V., H.J. Fowler, M. Bador, A. Behrangi, M.G. Donat, R. Dunn, C. Funk, J. Goldie, E. Lewis, M. Rogé, S.I. Seneviratne and V. Venugopal, 2019: On the use of indices to study extreme precipitation on sub-daily and daily timescales. *Environmental Research Letters*, **14**, 125008, https://doi.org/10.1088/1748-9326/ab51b6.

Beck, H.E., S. Westra, J. Tan, F. Pappenberger, G.J. Huffman, T.R. McVicar, G.J. Gründemann, N. Vergopolan, H.J. Fowler, E. Lewis, K. Verbist and E.F. Wood, 2020: PPDIST, global 0.1° daily and 3-hourly precipitation probability distribution climatologies for 1979–2018. *Scientific Data*, **7**, 302, https://doi.org/10.1038/s41597-020-00631-x.

Betts, A.K., and C. Jakob, 2002: Study of diurnal cycle of convective precipitation over Amazonia using a single column model. *Journal of Geophysical Research: Atmospheres*, **107**, D23, ACL 25-1-ACL 25-13, https://doi.org/10.1029/2002JD002264.

Blenkinsop, S., H.J. Fowler, R. Barbero, S.C. Chan, S.B. Guerreiro, E. Kendon, G. Lenderink, E. Lewis, X.-F. Li, S. Westra, L. Alexander, R.P. Allan, P. Berg, R.J.H. Dunn, M. Ekström, J.P. Evans, G. Holland, R. Jones, E. Kjellström, A. Klein-Tank, D. Lettenmaier, V. Mishra, A.F. Prein, J. Sheffield and M.R. Tye, 2018: The INTENSE project: using observations and models to understand the past, present and future of sub-daily rainfall extremes. *Advances in Science and Research*, **15**, 117–126, https://doi.org/10.5194/asr-15-117-2018.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *Journal of Climate,* **19***(*18*),* 4605–4630*,* https://doi.org/10.1175/JCLI3884.1*.*

Erlingis, J.M., J.J. Gourley, P.-E. Kirstetter, E.N. Anagnostou, J. Kalogiros, M.N. Anagnostou and W. Petersen, 2018: Evaluation of operational and experimental precipitation algorithms and microphysical insights during IPHEx. *Journal of Hydrometeorology*, **19**, 113–125, https://doi.org/10.1175/JHM-D-17-0080.1.

Field, P.R., and A.J. Heymsfield, 2015: Importance of snow to global precipitation. *Geophysical Research Letters*, **42**, 9512–9520, https://doi.org/10.1002/2015GL065497.

Fowler, L.D., M.C. Barth and K. Alapaty, 2020: Impact of scale-aware deep convection on the cloud liquid and ice water paths and precipitation using the Model for Prediction Across Scales (MPAS-v5.2). *Geoscientific Model Development*, **13**, 2851–2877, https://doi.org/10.5194/gmd-13-2851-2020.

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

Heymsfield, A.J., C. Schmitt, C-C. J. Chen, A. Bansemer, A. Gettelman, P.R. Field and C. Liu, 2020: Contributions of the liquid and ice phases to global surface precipitation: Observations and global climate modeling. *Journal of the Atmospheric Sciences*, **77**, 2629–2648, https://doi.org/10.1175/JAS-D-19-0352.1.

Hirota, N., and Y.N. Takayabu, 2013: Reproducibility of precipitation distribution over the tropical oceans in CMIP5 multi-climate models compared to CMIP3. *Climate Dynamics*, **41**, 2909–2920, https://doi.org/10.1007/s00382-013-1839-0.

Houze, R.A. Jr., L.A. McMurdie, W.A. Petersen, M.R. Schwaller, W. Baccus, J.D. Lundquist, C.F. Mass, B. Nijssen, S.A. Rutledge, D.R. Hudak, S. Tanelli, G.G. Mace, M.R. Poellot, D.P. Lettenmaier, J.P. Zagrodnik, A.K. Rowe, J.C. DeHart, L.E. Madaus, H.C. Barnes and V. Chandrasekar, 2017: The Olympic Mountains Experiment (OLYMPEX). *Bulletin of the American Meteorological Society*, **98**, 2167–2188, https://doi.org/10.1175/BAMS-D-16-0182.1.

Intergovernmental Panel on Climate Change, 2015: *Climate Change 2014: Synthesis Report*. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, ISBN:978-92-9169-143-2, 151 pp.

Kendon, E.J., R.A. Stratton, S. Tucker, J.H. Marsham, S. Berthou, D.P. Rowell and C.A. Senior, 2019: Enhanced future changes in wet and dry extremes over Africa at convection-permitting scale. *Nature Communications*, **10**, 1794, https://doi.org/10.1038/s41467-019-09776-9.

Kumar, D., E. Kodra and A.R. Ganguly, 2014: Regional and seasonal intercomparison of CMIP3 and CMIP5 climate model ensembles for temperature and precipitation. *Climate Dynamics*, **43**, 1–28, https://doi.org/10.1007/s00382-014-2070-3.

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

Levizzani, V., and E. Cattani, 2019: Satellite remote sensing of precipitation and the terrestrial water cycle in a changing climate. *Remote Sensing*, **11**, 2301, https://doi.org/10.3390/rs11192301.

Levizzani, V., C. Kidd, D.B. Kirschbaum, C.D. Kummerow, K. Nakamura and F.J. Turk (eds.), 2020a: *Satellite Precipitation Measurement*. Vol. 1, Springer Nature, Cham, *Advances Global Change Research*, **67**, 450 pp, https://doi.org/10.1007/978-3-030-24568-9.

Levizzani, V., C. Kidd, D.B. Kirschbaum, C.D. Kummerow, K. Nakamura and F.J. Turk (eds.), 2020b: *Satellite Precipitation Measurement*. Vol. 2, Springer Nature, Cham, *Advances in Global Change Research*, **69**, 712 pp, https://doi.org/10.1007/978-3-030-35798-6.

Li, R.C.Y., W. Zhou, C.M. Shun and T.C. Lee, 2017: Change in destructiveness of landfalling tropical cyclones over China in recent decades. *Journal of Climate*, **30**, 3367–3379, https://doi.org/10.1175/JCLI-D-16-0258.1.

Li, X., W.-K. Tao, T. Matsui, C. Liu and H. Masunaga, 2010: Improving a spectral bin microphysics scheme using TRMM satellite observations. *Quarterly Journal of the Royal Meteorological Society*, **136**, 382–399, https://doi.org/10.1002/qj.569.

Ling, J., and C. Zhang, 2011: Diabatic heating profiles in recent global reanalyses. *Journal of Climate*, **26**, 3307-3325, https://doi.org/10.1175/JCLI-D-12-00384.1.

Machado, L.A.T., M.A.F. Silva Dias, C. Morales, G. Fisch, D. Vila, R. Albrecht, S.J. Goodman, A.J.P. Calheiros, T. Biscaro, C. Kummerow, J. Cohen, D. Fitzjarrald, E.L. Nascimento, M.S.

Sakamoto, C. Cunningham, J.-P. Chaboureau, W.A. Petersen, D.K. Adams, L. Baldini, C.F. Angelis, L.F. Sapucci , P. Salio, H.M.J. Barbosa, E. Landulfo, R.A.F. Souza, R.J. Blakeslee, J. Bailey, S. Freitas, W.F.A. Lima and A. Tokay, 2014: The Chuva Project: How does convection vary across Brazil? *Bulletin of the American Meteorological Society*, 95, 1365–1380, https://doi.org/10.1175/BAMS-D-13-00084.1.

Matiu, M., M. Petitta, C. Notarnicola and M. Zebisch, 2020: Evaluating snow in EURO-CORDEX regional climate models with observations for the European Alps: Biases and their relationship to orography, temperature, and precipitation mismatches. *Atmosphere*, **11**, 46, https://doi.org/10.3390/atmos11010046.

Neale, R.B., J.H. Richter and M. Jochum, 2008: The impact of convection on ENSO: From a delayed oscillator to a series of events. *Journal of Climate*, **21**, 5904-5924, https://doi.org/10.1175/2008JCLI2244.1.

Nikulin, G., C. Jones, F. Giorgi, G. Asrar, M. Büchner, R. Cerezo-Mota, O. Bøssing Christensen, M. Déqué, J. Fernandez, A. Hänsler, E. van Meijgaard, P. Samuelsson, M. Bamba Sylla and L. Sushama, 2012: Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *Journal of Climate*, **25**, 6057–6078, https://doi.org/10.1175/JCLI-D-11-00375.1.

Pathak, R., S. Sahany, S.K. Mishra and S.K. Dash, 2019: Precipitation biases in CMIP5 models over the South Asian region. *Scientific Reports*, **9**, 9589, https://doi.org/10.1038/s41598-019-45907-4.

Prein, A.F., and A. Gobiet, 2017: Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *International Journal of Climatology*, **37**, 305–327, https://doi.org/10.1002/joc.4706.

Prein, A.F., and A.J. Heymsfield, 2020: Increased melting level height impacts surface precipitation phase and intensity. *Nature Climate Change*, https://doi.org10.1038/s41558-020-0825-x.

Rysman, J.-F., S. Berthou, C. Claud, P. Drobinski, J.-P. Chaboureau and J. Delanoë, 2018: Potential of microwave observations for the evaluation of rainfall and convection in a regional climate model in the frame of HyMeX and MED-CORDEX. *Climate Dynamics*, **51**, 837–855, https://doi.org/10.1007/s00382-016-3203-7.

Simpson, J., R.F. Adler and R.F. North, 1988: A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. *Bulletin of the American Meteorological Society*, **69**, 278–295, https://doi.org/10.1175/1520-0477(1988)069<0278:APTRMM>2.0.CO;2.

Simpson, J., C. Kummerow, W.-K. Tao and R.F. Adler, 1996: On the Tropical Rainfall Measuring Mission (TRMM). *Meteorology and Atmospheric Physics*, **60**, 19–36, https://doi.org/10.1007/BF01029783.

Smith, L.A., 2002: What might we learn from climate forecasts? *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 2478–2492, https://doi.org/10.1073/pnas.012580599.

Stephens, G.L., and C.D. Kummerow, 2007: The remote sensing of clouds and precipitation from space: A review. *Journal of the Atmospheric Sciences*, **64**, 3742–3765, https://doi.org/10.1175/2006JAS2375.1.

Stephens, G.L., T. L'Ecuyer, R. Forbes, A. Gettlemen, J.-C. Golaz, A. Bodas-Salcedo, K. Suzuki, P. Gabriel and J. Haynes, 2010: Dreary state of precipitation in global models. *Journal of Geophysical Research: Atmospheres*, **115**, D24211, https://doi.org/10.1029/2010JD014532.

Sun, Y., S. Solomon, A. Dai and R.W. Portmann, 2006: How often does it rain? *Journal of Climate*, **19**, 916–934, https://doi.org/10.1175/JCLI3672.1.

Tang, G., M.P. Clark, S.M. Papalexiou, Z. Ma and Y. Hong, 2020: Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sensing of Environment*, **240**, 111696, https://doi.org/10.1016/j.rse.2020.111697.

Tao, W.-K., S. Lang, W. Olson, R. Meneghini, Y. Yang, J. Simpson, C. Kummerow, E. Smith and J. Halverson, 2001: Retrieved vertical profiles of latent heat release using TRMM rainfall products for February 1998. *The Journal of Applied Meteorology and Climatology*, **40**, 957–982, https://doi.org/10.1175/1520-0450(2001)040%3C0957:RVPOLH%3E2.0.CO;2.

Tao, W.-K., E. Smith, R. Adler, Z. Haddad, A. Hou, T. Iguchi, R. Kakar, T.N. Krishnamurti, C. Kummerow, S. Lang, R. Meneghini, N. Nakamura, T. Nakazawa, K. Okamoto, W. Olson, S. Satoh, S. Shige, J. Simpson, Y. Takayabu, G. Tripoli and S. Yang, 2006: Retrieval of latent heating from TRMM measurements. *Bulletin of the American Meteorological Society*, **87**, 1555–1572, https://doi.org/10.1175/BAMS-87-11-1555.

Tao, W.-K., D. Wu, S. Lang, J. Chern, A. Fridlind, C. Peters-Lidard and T. Matsui, 2016a: High-resolution model simulations of MC3E, deep convective-precipitation systems: Comparisons between Goddard microphysics schemes and observations. *Journal of Geophysical Research: Atmospheres*, **121**, 1278–1306, https://doi.org/10.1002/2015JD023986.

Tao, W.-K., Y.N. Takayabu, S. Lang, S. Shige, W. Olson, A. Hou, G. Skofronick-Jackson, X. Jiang, C. Zhang, W. Lau, T. Krishnamurti, D. Waliser, M. Grecu, P.E. Ciesielski, R.H. Johnson, R. Houze, R. Kakar, K. Nakamura, S. Braun, S. Hagos, R. Oki and A. Bhardwaj, 2016b: TRMM latent heating retrieval: Applications and comparisons with field campaigns and large-scale analyses. *Meteorological Monographs*, **56**, 2.1–2.34, https://doi.org/10.1175/AMSMONOGRAPHS-D-15-0013.1.

Tapiador, F.J., 2010: A joint estimate of the precipitation climate signal in Europe using eight regional models and five observational datasets. *Journal of Climate*, **23(7)**, 1719–1738, https://doi.org/10.1175/2009JCLI2956.1.

Tapiador, F.J., A. Navarro, A. Jiménez, R. Moreno and E. García-Ortega, 2018: Discrepancies with satellite observations in the spatial structure of global precipitation as derived from global climate models. *Quarterly Journal of the Royal Meteorological Society*, **144**, 419–435, https://doi.org/10.1002/qj.3289.

Tapiador, F.J., A. Navarro, V. Levizzani, E. García-Ortega, G.J. Huffman, C. Kidd, P.A. Kucera, C.D. Kummerow, H. Masunaga, W.A. Petersen, R. Roca, J.-L. Sánchez, W.-K. Tao and F.J. Turk, 2017: Global precipitation measurements for validating climate models. *Atmospheric Research*, **197**, 1–20, https://doi.org/10.1016/j.atmosres.2017.06.021.

Tapiador, F.J., R. Roca, A. Del Genio, B. Dewitt, W. Petersen and F. Zhang, 2019: Is precipitation a good metric for model performance? *Bulletin of the American Meteorological Society*, **100**(2), 223–233, https://doi.org/10.1175/BAMS-D-17-0218.1.

Trenberth, K.E., A. Dai, R.M. Rasmussen and D.B. Parsons, 2003: The changing character of precipitation. *Bulletin of the American Meteorological Society*, **84**, 1205–1218, https://doi.org/10.1175/BAMS-84-9-1205.

vanZanten, M.C., and B. Stevens, 2005: Observations of the structure of heavily precipitating marine stratocumulus. *Journal of the Atmospheric Sciences*, **62**, 4327– 4342, https://doi.org/10.1175/JAS3611.1.

Voosen, P., 2016: Climate scientists open up their black boxes to scrutiny. *Science*, **354**, 401–402, https://doi.org/10.1126/science.354.6311.401.

Washington, R., R. James, H. Pearce, W.M. Poka, and W. Moufouma-Okia, 2013: Congo Basin rainfall climatology: Can we believe the climate models? *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**, 20120296, https://doi.org/10.1098/rstb.2012.0296.

Wu, M., G. Nikulin, E. Kjellström, D. Belušić, C. Jones and D. Lindstedt, 2020: The impact of regional climate model formulation and resolution on simulated precipitation in Africa. *Earth System Dynamics*, **11**, 377–394, https://doi.org/10.5194/esd-11-377-2020.

Yang, M., G.J. Zhang and D.-Z. Sun, 2018: Precipitation and moisture in four leading CMIP5 models: Biases across large-scale circulation regimes and their attribution to dynamic and thermodynamic factors. *Journal of Climate*, **31**, 5089–5106, https://doi.org/10.1175/JCLI-D-17-0718.1.

Yosef, Y., E. Aguilar and P. Alpert, 2020: Is it possible to fit extreme climate change indices together seamlessly in the era of accelerated warming? *International Journal of Climatology*, https://doi.org/10.1002/joc.6740.