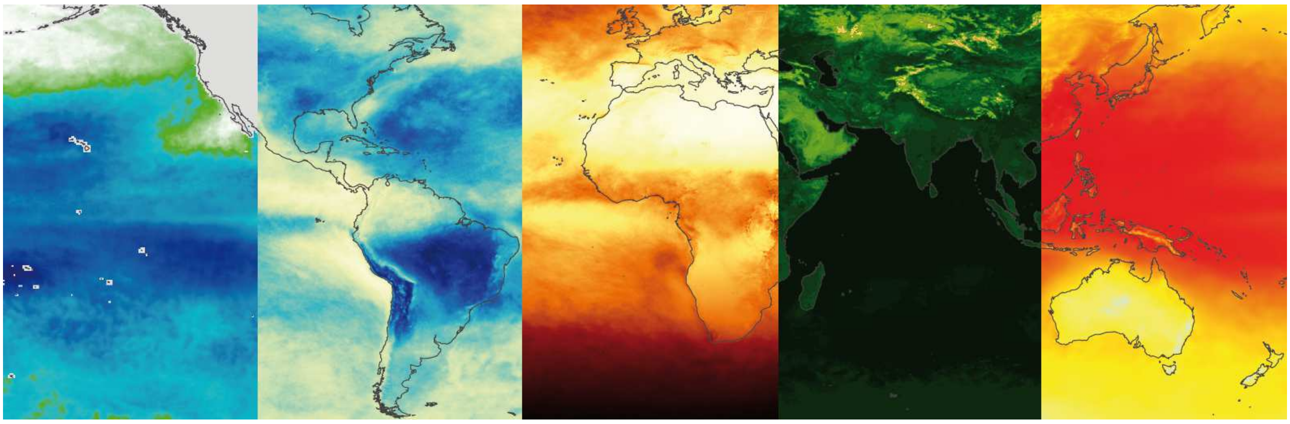

Summary Report from EUMETSAT / CM SAF
Workshop on Applications of Satellite Climate Data Records in Numerical Modeling

15 to 17 November 2016
ECMWF, Reading, UK

Jörg Trentmann, Viju John, Johannes Quaas, Roger Saunders, Steffen Kothe, Christine Träger-Chatterjee, Jörg Schulz

In November 2016 EUMETSAT and the CM SAF organized a three-day workshop on *Applications of Satellite Climate Data Records in Numerical Modeling* hosted by the ECMWF in Reading, UK.

This workshop involved 29 scientists from 10 countries bringing together users of satellite data records in the context of numerical modeling with providers of satellite-based data records. It was organized around three primary topics:

1) Data assimilation and model initialization
2) Process-oriented model evaluation and improvement
3) Model validation, climate trends and attribution studies

Each of these topics was introduced by overview presentations followed by discussion groups to facilitate efficient communication.

The **first day** was devoted to the introduction of the available satellite-based climate data records and a general overview on different aspects of applications of satellite data records in numerical modeling.

Roger Saunders (Met Office) highlighted the substantial increase in the availability of satellite-based climate data records over the past decade, before which only ISCCP and ERBE data were generally used by climate modellers. In the meantime substantial reprocessing efforts have been undertaken by satellite agencies (including EUMETSAT, NOAA, JMA) to generate updated Fundamental Climate Data Records (FCDRs), which are subsequently used to generate climate data records, e.g., in the context of the CM SAF and the ESA-CCI. Roger highlighted some of the requirements of modelers for satellite data:

1) Data records related to model variables
2) Observation simulators
3) Uncertainty information
4) Consistency between satellite data records
5) Commonly used, easy to read data format.

Examples were provided on the use of satellite climate data records documenting the different aspects of established applications of satellite-based climate information in the context of numerical modeling:

1) Assimilation in reanalyses
2) Specification of initial state
3) Model validation
4) Validating model processes
5) Verification of climate models for MIPs (CMIP6)
6) Trends and attribution, models vs obs.

In his summary Roger pointed out the excellent upcoming opportunities of using new satellite-based climate data records in numerical modeling, but also highlighted the importance of the communication and cooperation between the two communities to include the needs of the modelers, incl. the development of satellite simulators.

Jörg Schulz (EUMETSAT) introduced the pathways from satellite observations via the generation of climate data records to applications and decision-making. The role of EUMETSAT in the architecture for climate monitoring from space is mainly in archiving of the original satellite data, and the generation of FCDRs and Thematic Climate Data Records (TCDRs). The EUMETSAT Secretariat provides high quality FCDRs that are used in reanalysis and as input to derive TCDRs from geophysical variables, which are mainly generated in the framework of the EUMETSAT Satellite Application Facilities (e.g, CM SAF and OSI SAF).

Rainer Hollmann (DWD, CM SAF) introduced the Satellite Application Facility on Climate Monitoring (CM SAF) and its current and future product portfolio, which focusses on the energy and water cycle and will contain a global precipitation data record in the near future. He introduced the global (e.g., CLARA, HOAPS) and regional (e.g., SARAH, CLAAS) data records and provided examples of applications of these data records.

Jean-Noël Thépaut (ECMWF) introduced the Copernicus Climate Change Service (C3S) with its four components, including the Climate Data Store (CDS). The CDS will provide freely available climate data records to support the services by the C3S and will also include satellite-based climate data mainly from available sources that will be linked to the CDS.

On the morning of the **second day** the use of satellite climate data records in data assimilation and model initialization was discussed.

Hans Hersbach (ECMWF) presented an overview on the reanalysis activities at ECMWF. ERA-5 is the global reanalysis currently being generated, which includes multiple reprocessed satellite-based data records that have not been exploited in previous reanalysis, including For example SSMI/I brightness temperature from CM SAF as well as ASCAT-A scatterometer information and Meteosat atmospheric motion vectors (AMVs) from EUMETSAT. The good performance of the variational bias correction was highlighted, but emphasised the value of homogeneous and well-calibrated data records. Data from AMSU-A (Channel 14), SSU (Channel 3), scatterometer and GPS Radio occultation are used as socalled ‘anchoring data sets’ and are not bias-corrected during the assimilation. Historical satellite data from the early satellite periods (i.e., before the 1990s) would have the largest impact on the performance of the reanalysis due to the reduced availability of alternative data. ERA-5 will provide uncertainty information by an ensemble approach.

Shinya Kobayashi (JMA) presented the reanalysis activities at JMA; the latest reanalysis, JRA-55, is considerable improved compared to the previous reanalysis (JRA-25), but is impacted by the fundamental problem of changes in the observing system, a common issue for many reanalysis systems. Shinya pointed out the need for “anchoring observations”, which constrain the model biases and called for data rescue activities of early satellite data and the generation of FCDRs through reprocessing activities. The upcoming JMA reanalysis (JRA-3Q) will include improved satellite observations from such reprocessing activities.

Drew Peterson (Met Office) presented the initialization of the ocean and sea ice components in the Met Office Seasonal Forecast System. Among other data, satellite retrievals of the sea surface temperature (SST) and the sea ice concentration are assimilated to initialize the seasonal forecast system. Drew pointed out the different technical properties of the satellite data with the SST data being assimilated in satellite projection and the OSI-SAF sea ice concentration being considered as gridded data. Huge improvements in seasonal forecasts are expected from information on sea ice thickness that could be assimilated. Hindcasts are used to support the forecast; long (minimum of 15 years) and homogeneous data records are essential for seasonal prediction.

Gianpaolo Balsamo (ECMWF) presented the land surface components of the ECMWF modeling system and their initialization through the Land Data Assimilation System (LDAS), which is currently operationally used at ECMWF and in the ERA-5 reanalysis. The reprocessed data from the ASCAT-A instrument to be used in the ERA-5 reanalysis has resulted in improved performance of the surface soil moisture compared to the NRT ASCAT-A data. Long-term satellite-based soil moisture information has been used to investigate the impact of model parameters and model settings, e.g., increased vertical soil resolution.

On the **afternoon** of the **second day** the use of satellite climate data records in process-oriented model evaluation and model improvement was discussed

Andreas Fink (KIT) presented work on the evaluation of cloud processes in West Africa in climate models using the CM SAF CLAAS, CLARA, and SARAH climate data records. Using cloud observations and surface radiation measurements he indicated some problems in the satellite data records under the particular meteorological situation in West Africa during summer (regional long-lasting low stratus coverage), which currently limit the applicability of satellite data in the evaluation of climate models, which also show substantial biases.

Nicole van Lipzig (KU Leuven) presented results from three studies from Europe, Africa, and Greenland on the long-term evaluation of regional models using satellite data, e.g., the CM SAF CLAAS data record. In particular the improved performance of cloud resolving model simulations and their ability to correctly describe the diurnal evolution of thunderstorms in Lake Victoria was documented; this allows the thermodynamic and dynamic effects to be distinguished. The temporally and spatially high-resolution satellite data allows further investigation of remaining systematic errors in the model simulations. The importance of stable satellite data records with high resolution (in space and time) was emphasized.

Johannes Quaas (University of Leipzig) presented an overview about a selection of studies on constraining the aerosol-cloud-climate forcing. The need for appropriate information from satellite, such as cloud particle number concentration coincident with aerosol information (including CCN), and the value of satellite simulators was emphasized. Both highly-resolved multi-parameter data records as well as consistent long time series are required.

In the **morning** of the **final day** the use of satellite climate data records in operational validation and model performance, including climate trends and attribution studies, was discussed.

Duane Waliser (JPL/NASA) provided an overview on the Obs4MIPs initiative, which identifies, documents, and disseminates observations for climate model evaluation (e.g., CMIP) and serves as a direct connection between data providers and the model evaluation community. He highlighted the strategy of Obs4MIPs, i.e., to format the observations in a consistent way to the model output and to host the data side-by-side with the model output on the Earth System Grid (ESG). There are currently about 125 data records provided by Obs4MIPs, many of those based on satellite observations; it is expected that in the near future more than 200 data records will be accessible. The increasing number of available data records (and model simulations) enhances the need to provide guided information (e.g., maturity) on the data records and to easily search and access the data records.

Richard Allan (University of Reading) presented examples of applications of satellite climate data records in the assessment of decadal variability and its representation in climate models. The importance of homogeneous data records and of the consistency between independent data records to allow their application for climate process studies was emphasised (e.g., the trends in cloud cover in relation to the trend in the surface energy flux). Inhomogeneities in modern reanalysis systems were highlighted and these were found to be particularly acute for some variables (e.g. precipitation) and when involving complex observing systems. It was argued that climate simulations driven by the observed SST and sea ice fields and realistic radiative forcing may be considered the most basic and homogeneous reanalysis system. It was concluded that the combined use of satellite data and appropriate model simulations will lead to a better characterisation of deficiencies in both data records and advances in scientific understanding.

Thomas Haiden (ECMWF) presented the use of satellite data in the operational validation of the ECMWF IFS forecasts. It was highlighted that satellite climate data should be temporally homogeneous to enable attribution of changes in the performance measures to changes in the quality of the forecast. The usefulness of solar radiation (both the reflected at the top-of-the-atmosphere and the incoming at the surface) to infer the cloud forecast skill was emphasized; both data records are used to derive the deterministic and the ensemble forecast skill. The value of satellite-based cloud products for model evaluation will be assessed in the near future.

Torben Koenigk (SMHI) presented work on the representation of Arctic Sea Ice in global climate models in comparison with satellite climate data records from the CM SAF (CLARA SAL) and OSI SAF (Sea Ice Concentration). The substantial spread in sea ice concentration and sea ice albedo in the CMIP5 model simulations was highlighted; the comparison to the satellite data documents the inability of most models to reproduce the sea ice distribution and its surface albedo as well as their decline over the past decades. Satellite data records of the sea ice thickness and melt the pond fraction (both are sensitive parameters in the model simulations) would help to further investigate and improve the model representations of Arctic sea ice.

Discussion Groups

Detailed feedback by the participants was provided during the two parallel discussion group sessions on each of the main topics, which is summarized here.

In general, **satellite-based climate data records** have been proven useful and important in the field of numerical modeling. While FCDRs are mainly used in data assimilation, TCDRs are widely used in many applications in the context of numerical modeling (incl. data assimilation). All levels of temporal and spatial aggregation of the satellite data (from Level1 to Level3) are used. Long-term climate data records (> 20 years) are requested, also at all levels of aggregation: Level 1 (FCDR) via Level 2 (TCDR) on satellite sensor resolution to Level 3 (TCDR), e.g., as monthly averages. To reduce the data amount for the data transfer and on the side of the user it was suggested to provide 'hosted processing'-platforms to aggregate the highly resolved information of Level 2 data according to the user needs. The close cooperation and interaction with the data providers for technical and scientific support is seen as very important, including for data that is accessed from dataset collections like Obs4MIPs.

The **temporal stability** of the long term satellite climate data records is of relevance in all applications. However, the absolute requirements on stability differ for each application, e.g., data assimilation vs operational verification vs decadal variability and trend estimation. It was suggested to provide specific data records with high stability, but accepting reduced accuracy. In general, the stability of data records should be reported with the data.

The provision of **uncertainty** information with the satellite data records has been discussed intensively. In general, uncertainty information will be used by the users and should be provided with the data.
In data assimilation, uncertainty is defined as the known bias and the standard deviation; in addition data assimilation requires information on the error covariance, which could also be provided with the satellite data record.
In the context of model evaluation (mainly using TCDRs) the uncertainty information is not as clearly defined and used; it was suggested to consider applying an ensemble retrieval approach to estimate uncertainty and to validate the TCDR with reference observations to assess the uncertainty. Also a quality flag could be used as a measure of uncertainty. It is envisioned to have uncertainty information available for every grid point of the satellite data record.

The **consistency** of the multi-parameter satellite climate data records has also been addressed as an important quantity of satellite climate data particularly for assimilaton, but also for process studies; however, it has been pointed out that consistency must not be achieved at the expense of independence, for example through a common retrieval algorithm that enforces consistency among parameters. For model evaluation and process studies consistent and independent data are required.

Satellite simulators are seen as valuable tools to enable consistent comparison between model and satellite data in 'observation space'; it was highlighted that simulators with different complexities are available and the application of simulators should not be the only approach. A specific type of satellite simulator (i.e., forward model and its adjoint) is required for data assimilation applications.

The rescuing of **historical satellite data**, i.e., prior to the current satellite era, i.e., in the 1970s, would be very beneficial, in particular for reanalysis.