Correcting orbital drift signal in the time series of AVHRR derived convective cloud fraction using rotated empirical orthogonal function

A. Devasthale1, K.-G. Karlsson1, J. Quaas2,*, and H. Grassl2

1Atmospheric Remote Sensing, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
2Max-Planck Institute for Meteorology, Hamburg, Germany
*now at: Leipziger Institut für Meteorologie, Universität Leipzig, Leipzig, Germany

Correspondence to: A. Devasthale (abhay.devasthale@smhi.se)

Received: 16 May 2011 – Published in Atmos. Meas. Tech. Discuss.: 21 June 2011
Revised: 22 January 2012 – Accepted: 23 January 2012 – Published: 1 February 2012

Abstract. The Advanced Very High Resolution Radiometer (AVHRR) instruments onboard the series of National Oceanic and Atmospheric Administration (NOAA) satellites offer the longest available meteorological data records from space. These satellites have drifted in orbit resulting in shifts in the local time sampling during the life span of the sensors onboard. Depending upon the amplitude of the diurnal cycle of the geophysical parameters derived, orbital drift may cause spurious trends in their time series. We investigate tropical deep convective clouds, which show pronounced diurnal cycle amplitude, to estimate an upper bound of the impact of orbital drift on their time series. We carry out a rotated empirical orthogonal function analysis (REOF) and show that the REOFs are useful in delineating orbital drift signal and, more importantly, in subtracting this signal in the time series of convective cloud amount. These results will help facilitate the derivation of homogenized data series of cloud amount from NOAA satellite sensors and ultimately analyzing trends from them. However, we suggest detailed comparison of various methods and rigorous testing thereof applying final orbital drift corrections.

1 Introduction

Nearly 30 years (1982–present) of data from the Advanced Very High Resolution Radiometers (AVHRRs) onboard the National Oceanic an Atmospheric Administration (NOAA) satellite series constitute the longest continuous meteorological space-based measurements. These long-term measurements at high spatio-temporal resolution and at carefully chosen spectral wavelengths make them extremely valuable for climate monitoring purposes and process-based studies (Rossow and Schiffer, 1999; Karlsson, 2003; Heidinger and Pavolonis, 2009; Schulz et al., 2009). The AVHRR data in principle could be used for investigating variability and trends in essential climate variables at the decadal time scales. Deriving global and regional cloud climatologies is just one example of scientific usefulness of AVHRR data in improving our understanding of the Earth System (Rossow and Schiffer, 1999; Karlsson, 2003; Heidinger and Pavolonis, 2009; Devasthale and Grassl, 2009a, Devasthale and Fueglistaler, 2010). The NOAA satellites on which AVHRR sensors are mounted have however drifted in their orbit during the lifespan (Ignatov et al., 2004). The drifting leads to delay in their time of observation, which subsequently results in inconsistent time-sampling of clouds and other geophysical variables. This observational delay could further lead to artifacts in the time series of clouds (Devasthale and Grassl, 2007). Therefore, it is necessary to address this issue before the long-term AVHRR data can be used for climatological trend analysis.

There have been few efforts in the past to correct for the drift signal in the time series of other geophysical climate variables from AVHRR. For example, Waliser and Zhou (1997) and Lucas et al. (2001) applied rotated empirical functions (REOF) analysis to remove drift and biases arising due to changes in the satellite platforms from the outgoing longwave radiation dataset. Jin and Treadon (2003) corrected the bias in land surface skin temperatures. The correction
of the time series of cloud amount however is not been attempted so far. Since cloud properties are listed as essential climate variables and play a significant role in the Earth’s radiation budget, it is imperative to explore the methodologies to correct for the drift signal in the time series of cloud amount so that the data can be eventually used for trend analysis. In the subsequent sections, we demonstrate the usefulness of REOFs in correcting the drift signal.

In this section, we present a detailed description of the potential influence of orbital drift on the time-series of cloud fraction (Sect. 2.1), the AVHRR data processing carried out for the present study (Sect. 2.2) and a step-by-step description of REOF analysis (Sect. 2.3).

2 The orbital drift, AVHRR data processing and REOF methodology

2.1 The nature of the orbital drift and how it may impact the time-series of cloud fraction

The host NOAA satellites that carry AVHRR instruments are sun-synchronous satellites, meaning that they should have fixed equator crossing times (or fixed local time of observation at a given location). In order to fulfill this condition, it is required that these satellites exactly maintain their orbit altitude, which in practice is usually done by tightly monitoring their orbital path and by applying periodic corrections if a deviation occurs. If this is not done properly, the satellite will start drifting from its initial orbit (as happened in case of many of the old NOAA satellites up to NOAA-17). This drifting results in a continuous change in the equator crossing time. The rate of drift is different for each satellite and it may not be constant over time. This is clearly evident in Fig. 1, which shows the monthly mean time of observation (in hours UTC) over the study area in question for JJAS of 1982 to 2006, for the different NOAA satellites in orbit during this period. It is to be noted that the full potential of NOAA satellites for climate monitoring applications – in addition to the weather prediction purposes – was not foreseen then, and therefore the condition on sun-synchronicity was not rigorously met.

The conceptual schematic of how drifting may influence the time-series of cloud fraction is shown in Fig. 2. It shows an idealised diurnal cycle of convection over land and presents three scenarios wherein this diurnal cycle is sampled along its ascending or descending branches. In the first scenario, when the diurnal cycle is sampled along its ascending branch due to a continuous delay in the time of observation due to orbital drift, we may observe the spurious increase in convective cloud fraction during the life span of a sensor. If the diurnal cycle is, on the other hand, sampled along its descending branch, this may lead to spurious decrease in convective cloud fraction as depicted in the second scenario. However, it may happen that the change in the time of observation is around the peak of the diurnal cycle. In such case, the time series may not show any discernible trend or may be small enough to be masked by the real trend. Notice that the magnitude and sign of the spurious trend will eventually depend upon the scenario in question, which in turn depends upon the type of cloud studied, amplitude of the diurnal cycle, the rate of drift, season and geographical position. Therefore, using the slope between two points along the diurnal cycle to correct the time-series in retrieved cloudiness is not practical. As we will show later, the REOF analysis offers a promising way to achieve this goal because of its intrinsic ability to be insensitive to these dependencies.

Fig. 1. Monthly mean time of observation (in UTC) averaged over the entire study area for the ascending passes (i.e. daytime data only) for different NOAA satellites for the JJAS months.

Fig. 2. Conceptual schematic of the potential impact of inconsistent time sampling due to orbital drift and an idealised diurnal cycle of convection on the time-series of convective cloud fraction (refer text for the description).
2.2 AVHRR data processing

We used level 1b AVHRR Global Area Coverage (GAC) data with a nominal spatial resolution of 5 km by 3 km for the analysis. The AVHRR is a five-channel instrument with two channels in the solar spectrum (0.58–0.68 µm and 0.725–1.1 µm), two in the thermal infrared spectrum (10.3–11.3 µm and 11.5–12.5 µm). The third channel falls partly in the solar and in the thermal infrared spectrum (3.55–3.93 µm). The solar and thermal channels of the AVHRRs are intercalibrated as explained in the work of Devasthale and Grassl (2009a,b). The brightness temperatures derived from the thermal channels are used in cloud detection and cloud typing and therefore they are validated in order to ensure consistency of the data across different NOAA platforms (Devasthale and Grassl, 2009b). The data from AVHRR sensors onboard NOAA-7, -9, -11, -14 and -16 satellites are analysed. We analysed only those clouds with channel 4 (11.0 µm) brightness temperatures less than 220 K, which over the Indian summer monsoon area are most likely deep convective clouds and associated optically-thick anvil cirrus clouds. They have strong diurnal cycle amplitude, and previous studies have shown spurious trends in their time-series (Devasthale and Grassl, 2007).

2.3 The REOF analysis

EOF statistical analyses are being used in the atmospheric sciences since the late 1960s. The paper by Hannachi et al. (2007) presents a comprehensive review of EOFs and its variants commonly used in the climate data analysis. The EOF analysis is similar to principal component analysis in many ways except for the advantage that it allows investigations of patterns in both, time and space dimensions, simultaneously. The use of REOFs analysis, on the other hand, offers a few distinct advantages mentioned below over EOFs, which are especially relevant for the present study.

1. The REOF analysis reduces dimensionality of the data set even more effectively than original EOF analysis. This has been shown in the works by Waliser and Zhou (1997) and Lucas et al. (2001) in the present study. This allows relatively easy identification and delineation of orbital drift signal in the data.

2. The REOFs are independent of the domain size. This entails that although we successfully applied this analysis over the Indian monsoon region, it can be spatially extended and/or applied to other domains.

3. It is relatively easy to interpret physically. For example, by reducing dimensionality and aggregating drift signal only in few modes, one can compare spatial patterns of modes that are or are not, respectively, affected by the drift signal in a more physical way.

The VARIMAX rotation used in the present analysis is briefly explained below (based on the works of Hannachi et al., 2007; Kaiser, 1958 and Richman, 1986). The rotation is done by seeking a rotation matrix \( R \) to construct the REOFs \( B \) from:

\[
B = U_m \cdot R \quad (1)
\]

where, \( U_m \) is a \( p \times m \) matrix with \( p \) being the spatial points, and \( m \) is the number of leading EOF loadings (that are to be rotated). In case of orthogonal VARIMAX rotation \( R = R \).

The criterion for choosing the rotation matrix \( R \) is expressed as the maximization problem.

\[
\text{max } f \left( U_m, R \right) \quad (2)
\]

In particular, the VARIMAX rotation maximizes a simplicity criterion based on the following:

\[
\text{max } \left( f(B) = \sum_{k=1}^{m} \left[ p \sum_{j=1}^{p} b_{j,k}^2 - \left( \sum_{j=1}^{p} b_{j,k} \right)^2 \right] \right) \quad (3)
\]

where \( b_{i,j} \) are the elements of the REOFs matrix \( B \) in Eq. (1).

The practical methodology to correct for orbital drift signal is conceptually based on the original works of Waliser and Zhou (1997) and Lucas et al. (2001). The corrected data set is obtained in six well-defined steps described below.

a. First, the time series’ (at each 1° × 1° grid point) of monthly mean convective cloud anomalies are prepared by subtracting the time series means. Note that the convective clouds are defined here as pixels with channel 4 brightness temperatures less than 220 K.

b. In the next step, an EOF analysis is performed and the first twenty modes are retained because together they explain more than 99 % of variance in the data set.

c. The EOF loadings, defined as the time series of the first seven principal components (or EOF modes), are rotated using the VARIMAX rotation mentioned above (Kaiser, 1958; Richman, 1986). Only the first seven modes were rotated as the other modes showed extremely weak correlation with equator crossing time (< ±0.02; see Fig. 4 and explanations below).

d. The time-series of seven modes are inspected visually and the modes that contain an orbital drift signal, i.e. resembling the increasing/decreasing trend during the life span of the sensor and then showing sudden jumps at the change of satellite, are identified. In a more comprehensive study this procedure can easily be performed automatically by correlating loadings with equator crossing times and then choosing the modes with the highest correlation for further analysis (e.g. see Fig. 4).

e. Synthetic loadings were then computed for these contaminated modes by fitting a linear regression between the EOF loadings and the local time of observation.
The climatological distribution of deep convective clouds (with AVHRR channel 4 brightness temperatures < 220 K) over the study area for the summer (JJAS) monsoon season. The occurrence frequencies are shown in percent.

f. These synthetic loadings were removed from the anomaly dataset. This yields a new dataset with the orbital drift signal removed.

3 Results of the REOF analysis

Figure 3 shows the climatological distribution of deep convective clouds over the Indian subcontinent. Most of the convection is organized along the two main branches of the monsoonal winds; the one approaching the subcontinent at the western coast over the Arabian Sea and the other approaching it from the Bay of Bengal. The climatological distribution of convective clouds, therefore, shows a high frequency over these regions (Devasthale and Grassl, 2009a). The Bay of Bengal and the central and northeast parts of India along the Himalayan mountain ranges even support the deepest convection penetrating into the tropical tropopause layer (Devasthale and Fueglistaler, 2010). The diurnal cycle of such clouds has a large amplitude. In the present study, the EOF analysis is done on this regional time-series of clouds with AVHRR channel 4 brightness temperatures below 220 K as mentioned in Sect. 2. The analysis steps a-c described in Sect. 2.3 are followed. The REOFs analysis disentangles orbital drift signals better than unrotated EOFs. This is demonstrated in Fig. 4, which shows the correlation of the first seven EOF loadings with the local time of observation using the unrotated and rotated EOFs. In the unrotated case (Fig. 4, left panel), five of the seven modes are contaminated by the orbital drift signal, while in the latter case (Fig. 4, right panel) a strong signal is seen in only two modes (modes 1 and 3). The spatial pattern of REOFs for the first seven modes is shown in Fig. 5. The variance explained by the first seven modes is 19.0 %, 13.5 %, 6.98 %, 5.16 %, 4.08 %, 3.67 % and 3.05 % respectively. Mode 2 shows the pattern of variability that we are actually looking for in the data. Its spatial pattern resembles closely the climatological spatial distribution of deep convective clouds and has low correlation with the time of observation. The loadings of Modes 1 and 3 are regressed against the time of observation. The linear relationship between the loadings and the local time of observation is used to compute new synthetic loadings. They are shown in Fig. 6 in red color. These synthetic loadings are constructed to represent the contribution from the orbital drift signal to the temporal evolution of convective and thick anvil cloud fraction in the AVHRR time series. The corrected dataset is computed by removing these synthetic loadings from the original dataset. Figure 7 shows the spatial distribution of
Fig. 6. EOF loadings for modes 1 and 3 in rotated case (black line) and corresponding synthetic loadings (in red line) computed by linearly regressing the loadings with the time of observation.

Fig. 7. Spatial pattern of correlation of time series of deep convective cloud amount at each grid point with the time of observation for the original AVHRR dataset (left panel) and after drift signal removal (right panel). The correlations are reduced considerably in the latter case.

Fig. 8. Frequency distribution of optically thick convective cloud fraction (BT < 220 K) for JJAS 2006 for original and corrected AVHRR data compared with MODIS over the study area.

Table 1. Statistical comparison of uncorrected and corrected AVHRR data with the MODIS cloud product for JJAS 2006 over the study area.

<table>
<thead>
<tr>
<th></th>
<th>Observed AVHRR/N16</th>
<th>Corrected AVHRR/N16</th>
<th>Modis/AQUA</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>max</td>
<td>0.3622</td>
<td>0.3469</td>
<td>0.3342</td>
</tr>
<tr>
<td>mean</td>
<td>0.05441</td>
<td>0.04566</td>
<td>0.04357</td>
</tr>
<tr>
<td>median</td>
<td>0.04629</td>
<td>0.03109</td>
<td>0.03299</td>
</tr>
<tr>
<td>std dev</td>
<td>0.04834</td>
<td>0.04654</td>
<td>0.04448</td>
</tr>
</tbody>
</table>

Having demonstrated that the main purpose of identifying and removing the orbital drift signal is achieved, it is important to verify whether the information to be analysed is not removed at the same time. Thus, the remaining question is whether the natural variability in the corrected dataset is preserved. The data from MODerate resolution Imaging Spectroradiometer (MODIS) onboard NASA’s Aqua satellite are used to answer this question. In contrast to the NOAA satellites, the Aqua satellite orbit is stable in its time of observation. A comparison is possible since Aqua satellite on which the MODIS instrument is mounted also observes the study region in the afternoon orbit similar to that of NOAA-16 satellite (which has drifted in orbit). The REOF analysis is performed on AVHRR/NOAA-16 data (JJAS months) for 2001 to 2006 period. The AVHRR and MODIS data (Level 2 Cloud Products, Version 5) from 2006 were used for comparison, as the drift is largest towards later years than the year 2001 (e.g. see Fig. 1). A comparison of the cloud cover statistics is given in Table 1. For all statistical quantities, the corrected AVHRR data is closer to the MODIS than the uncorrected AVHRR data. The histogram of cloud fraction (Fig. 8) confirms that indeed the frequency distribution of corrected AVHRR data compares better with the MODIS reference data than for original data. Waliser and Zhou (1997)
and Lucas et al. (2001) have previously argued that the use of synthetic loadings removes only orbital drift signal from the dataset: our comparison results support their argument. Figure 9 shows the spatial distribution of uncorrected (original) AVHRR convective cloud fraction, corrected AVHRR cloud fraction after removing the drift signal, MODIS cloud fraction and the difference of uncorrected AVHRR and corrected AVHRR convective cloud fractions for the JJAS months of 2006. The difference image shows large spatial variability. A careful investigation shows that the correction reduces the diagnosed convective cloud fraction over areas where the diurnal cycle is sampled along its ascending branch during the life span of the sensor, while the cloud fraction is slightly increased over areas where the cycle is sampled along its descending branch. This is physically consistent. For example, Fig. 10 shows the diurnal cycle of convection derived, for the same time period (JJAS 2006), over two areas (marked as rectangles in Fig. 9) using Meteosat Visible and Infrared Imager (MVIRI) onboard the Meteosat-5 geostationary satellite providing images at every 30 min. This figure demonstrates that, over the first area over the northeast India where the diurnal cycle is sampled along its ascending branch, the cloud fraction is reduced in the corrected AVHRR data, while the opposite is true for the area over the Bay of Bengal. These two regions are in fact good realizations of the first two scenarios described in the conceptual schematic presented in Fig. 2.

4 Conclusions and discussions

We demonstrate that REOF analysis efficiently delineates the orbital drift signal in the time series of convective cloud fraction. These results have special significance in the context of climate monitoring from space, since NOAA satellite sensors (AVHRR, HIRS, TOVS) offer the longest continuous data records of essential climate variables from space. However, the two key issues of accurate intercalibration of AVHRR sensors and removal of orbital drift signal need to be addressed before these data can be used for long-term trend analysis of cloud properties. While the scientific community is currently arriving at a consensus on the accurate intercalibration methodologies, the results presented in this study try to address the other important issue of orbital drift removal. It has to be noted here that we present a case study where the orbital drift signal is extreme, since the diurnal cycle of deep convective clouds has very large amplitude. This provides an upper bound of the possible orbital drift impact. However, the drift signal in the total cloud fraction could be weak (due to weak diurnal cycle amplitude of total cloud fraction resulting from the compensating effects of phase lags of maximum in the diurnal cycles of different cloud types) or even absent (depending on cloud type, latitude and season). This explains the fact that in our case study the first and the third modes of EOFs were contaminated by a drift signal (dominating the variability in the dataset), while in the study by Waliser and Zhou (1997) and Lucas et al. (2001) where they investigated the outgoing longwave radiation time series, the affected modes were 4th and 3rd (relatively weak
contamination) respectively. In a cautionary note, we would like to mention that there is no unique way to determine how many modes are to be rotated. For example, the Preisendorfer N rule significance test as used by Lucas et al. (2001) can be used or, as done in the present study, some threshold on correlation of data time-series with equator crossing time can also be effectively used. Although we used the most commonly used rotation (VARIMAX), there are few other rotation methods available which could be more suitable. Additionally, a detailed comparison with other statistical tools (e.g. Hilbert Huang Transform) and methods (e.g. using diurnal cycle slope) is also needed to examine the relative benefits of other methods. It is also necessary to rigorously test that the large scale statistical features are preserved in the corrected data set. All of these issues need consideration, and will be investigated in future, before we apply final corrections to the climatological time series. Nonetheless, the REOFs analysis certainly is a promising approach.

Acknowledgements. Authors would like to thank two anonymous referees for their constructive remarks that lead to considerable improvement of the original manuscript. The authors are grateful to NOAA/CLASS for providing access to the level 1b AVHRR data, and NASA DAAC for providing MODIS cloud products. This study was funded by the EUMETSAT’s Satellite Application Facility on Climate Monitoring (CM-SAF).

Edited by: A. Kokhanovsky

References


