Proceedings for the 2017 EUMETSAT Meteorological Satellite Conference,
2-6 October 2017, Rome, Italy

CORRECTING CLOUD FRACTIONAL COVER CLIMATOLOGY BY GEOSTATISTICAL MERGING OF SATELLITE AND SYNOPTIC DATA

Jędrzej S. Bojanowski[1], Reto Stöckli[2]
[1] Remote Sensing Centre, Institute of Geodesy and Cartography, Warsaw, Poland
[2] Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland
2

Abstract

Generating a cloud climate data record (CDR) from observations acquired by a set of polar orbiting satellites requires dealing with potential artefacts caused by lack of long-term stability, degradation of satellite sensors, radiometric and geometric calibration problems, and satellite orbital drift. The majority of these aspects have been resolved within the scope of the ESA Cloud_cci project's retrieval algorithms used to produce a new 30y+ (1982-2014) cloud CDR from the AVHRR sensor onboard NOAA satellites (CC4CL-AVHRR). Yet, orbital drift and varying equatorial crossing times of consecutive satellites have not been accounted for. This may inhibit the use of the original dataset for investigating the cloud variability and change over the last 3 decades. We propose and apply a simple statistical method for correcting the original dataset over Europe by a bias removal using quality-checked synoptic observations from 158 sites. The approach builds on kriging with external drift in which ground-based observations are interpolated using satellite data as explanatory variable. Interpolation is applied to monthly means, i.e. not explicitly resolving the diurnal cycle of cloudiness. Averaged over 68 evaluation sites, the corrected (debiased) dataset (MBE=-0.68%, bcRMSE=6.42%) significantly outperforms the original one (MBE=4.05%, bcRMSE=14.90%). The correction also decreases the performance differences among NOAA satellites, and implicitly removes the inhomogeneity in cloud fractional cover (CFC) time series due to changing overpass times. We also show the correction reduces the magnitude of trends in CFC but keeps their sign unchanged.

INTRODUCTION

To be suitable for climate analysis, satellite-derived cloud datasets have to meet the challenging requirements including those for accuracy, precision and decadal stability. The longest cloud fractional cover (CFC) dataset of Cloud_cci project (1982-2014) derived from the afternoon NOAA satellites (Stengel et al., 2017) has been thoroughly evaluated in the Product Validation and Intercomparison Report (Stapelberg et al., 2017).

The Cloud_cci dataset reveals some inhomogeneities in CFC between sensors (e.g. NOAA-7 and NOAA-9) as well as variation of performance within the lifetime of individual sensors. These inaccuracies are related to degradation of satellite sensors, radiometric and geometric calibration problems, and satellite orbital drift. Some of these aspects have been resolved within the scope of Cloud_cci retrieval algorithms (Poulsen et al., 2011; Stengel et al., 2015). However orbital drift and varying equatorial crossing times of consecutive satellites have not been accounted for. This may inhibit the use of the original dataset for investigating the cloud variability and change over the last 3 decades.

Several methods exist for correcting satellite-derived cloud cover time series. Foster and Heidinger (2013) derived corrected CFC daily means from PATMOS-x dataset by fitting a mean diurnal cycle to a sinusoidal function derived over all NOAA's. An average corrected daily value of cloudiness is then interpolated from the fit function using all available ascending, descending, morning, and afternoon satellite overpasses. The application of this correction method is inhibited for the Cloud_cci CC4CL AVHRR-PM dataset since it only utilizes one satellite at a time. Devasthale et al. (2012) succeeded to delineate the signal of orbital drift in the AVHRR-based CFC time series by means of rotated empirical orthogonal function (REOF). However, the method is shown to be sensitive to decisions which EOF modes do reflect unnatural CFC variability. Therefore, it is necessary to rigorously test that the large scale geostatistical features are preserved in the corrected data set. Similar risk of removing from original dataset correct spatio-temporal patterns has been pointed out by Norris and Evan (2015) while detecting (by least squares best-fit) and then removing the CFC variability linked to known artefacts (e.g. orbital drift).

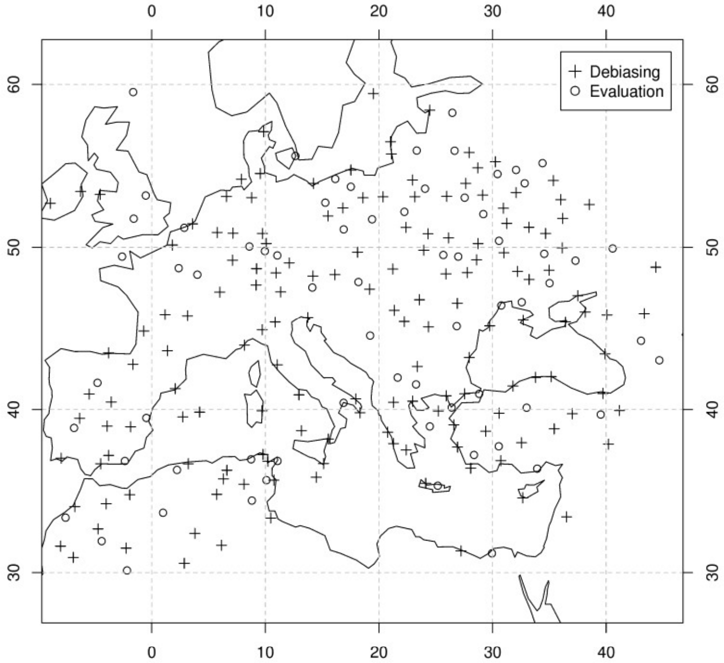

Figure 1: **SYNOP sites used for debiasing and evaluation of CC4CL AVHRR-PM CFC monthly means.**

In this study, we propose a simple statistical method for correcting CFC by debiasing the CC4CL AVHRR-PM CFC data using synoptic observations. The method is based on interpolation of these observations using satellite data as an explanatory variable. In this context, the proposed method relies on a strong assumption that synoptic observations are accurate and homogenous. Moreover, we assume that CFC errors caused by the artefacts such as orbital drift and transition between satellites are related to the occurrence and intensity of cloudiness diurnal cycle. Hence, we assume that these errors are spatially correlated (i.e. can be interpolated). Interpolation can be applied to monthly means, i.e. not explicitly resolving the diurnal cycle of cloudiness.

METHODS

Synoptic observations from the European Centre for Medium-Range Weather Forecasts (ECMWF) archive were used for debiasing and evaluation. The archive initially contains data for over 6000 globally distributed sites. From these we selected sites for a geographic range of 30N to 60N and 10W to 45E. In order to ensure the collection of long-term homogenous data series, we selected stations where observations were continuously performed in 1982–2014, at least every 6 hours with a maximum break of 30 days. For each site we used cloud amount observed with the highest temporal frequency (up to 1 hour) that was reported for the whole 33-year period. Thus the frequency of observations could vary between sites, but remained stable in time for each site. Instantaneous observations were transformed from the okta scale to cloud fractional cover, and aggregated to monthly means. Further, we excluded sites for which the Standard Normal Homogeneity Test (SNHT, Alexandersson, 1986, Khaliq and Ouarda, 2007) detected any inhomogeneity in a time series of cloud amount monthly anomalies.

Since SYNOP stations are unevenly distributed in geographic space, a regular grid was created. For each 2x2 degree geographic grid cell the SYNOP sites with the most valid observations were used. In case two or more sites were present within one grid cell, one site was used as training and one as validation site. This selection procedure yielded a total of 158 SYNOP sites used for training, and 68 sites used for validation (Figure 1).

Satellite data consisted of 33 years of CFC monthly means derived from NOAA-AVHRR afternoon satellites (CC4CL AVHRR L3C). These included data from consecutive afternoon NOAA satellites: NOAA-7, NOAA-9, NOAA-11, NOAA-12 (morning satellite due to a gap in data from afternoon satellites), NOAA-14, NOAA-16, NOAA-18 and NOAA-19.

	Corrected		**Uncorrected**	
NOAA	**MBE**	**bcRMSE**	**MBE**	**bcRMSE**
AVHRR-PM	-0.68	6.42	4.05	14.90
7	-0.70	6.57	5.28	14.66
9	-0.73	6.55	4.25	14.90
11	-0.91	6.99	3.04	15.17
12	-0.24	6.97	9.16	15.41
14	-0.79	6.12	5.77	14.65
16	-0.80	6.03	3.68	14.08
18	-0.21	6.27	2.82	15.25
19	-0.53	6.32	3.25	14.52

***Table 1:* Mean bias error (MBE) and bias-corrected root mean square error (bcRMSE) of CC4CL-AVHRR-PM uncorrected and corrected (debiased) mean monthly cloud fraction as compared to synoptic observations aggregated for NOAA missions.**

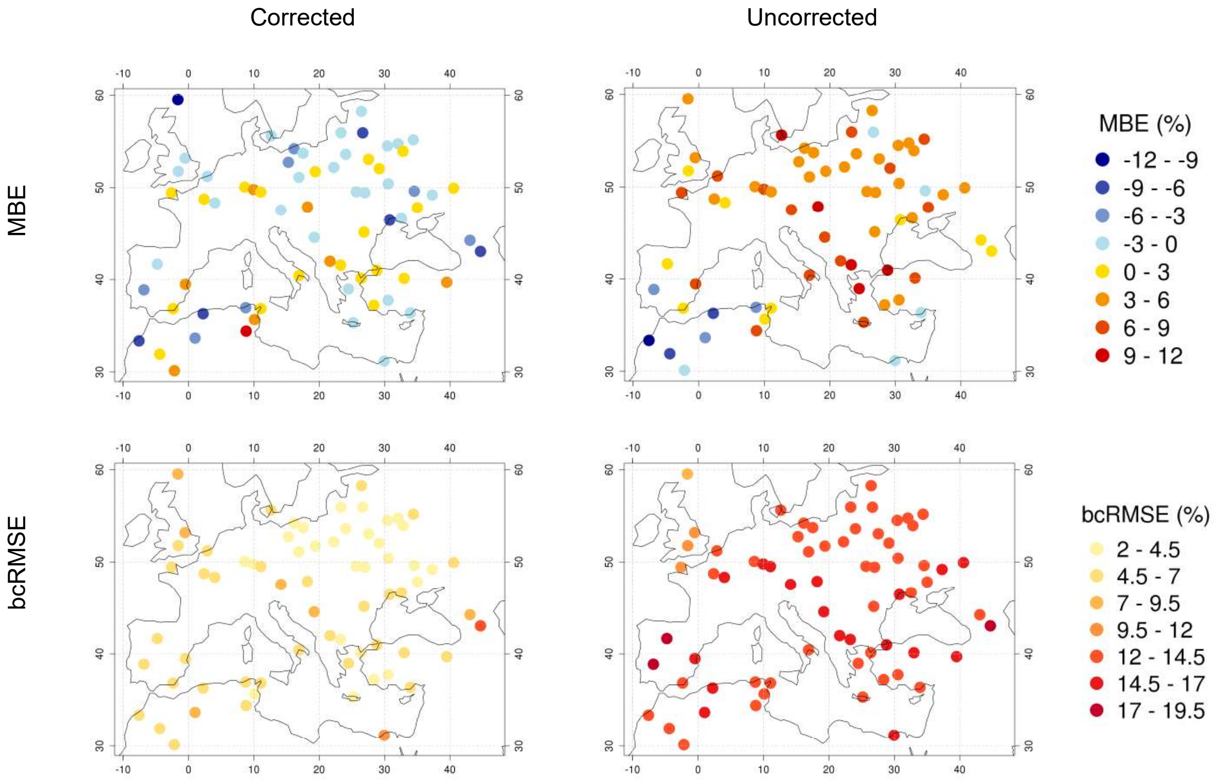

***Figure 2:* Mean bias error (MBE) and bias-corrected root mean square error (bcRMSE) of CC4CL-AVHRR-PM uncorrected and corrected (debiased) mean monthly cloud fraction as compared to synoptic observations.**

Debiasing of satellite-derived CFC employed interpolation of synoptic observations using satellite data as an explanatory variable. For each month (i.e. 1982-01, 1982-02, etc.) kriging with external drift was applied to SYNOP-based monthly means (at 158 sites). The variogram was fit automatically to regression residuals and kriging performed using the R package ‘automap’ (Hiemstra et al., 2009). Based on explanatory analysis, the spherical variogram model was chosen and imposed for each fitting. The native spatial resolution of 0.05 degree was preserved during interpolation.

To assess the accuracy of debiased satellite CFC, we used the mean bias error, i.e. mean difference between debiased CFC and reference SYNOP data. To express the precision, the bias-corrected root mean squared error (bcRMSE) was used. Finally, for a trend analysis we used linear trends derived using Theil-Sen estimates (Theil, 1950) and their significance was estimated with the Mann-Kendall test (Kendall, 1938; Mann, 1945) and adjusted using the method of Benjamini and Hochberg (1995).

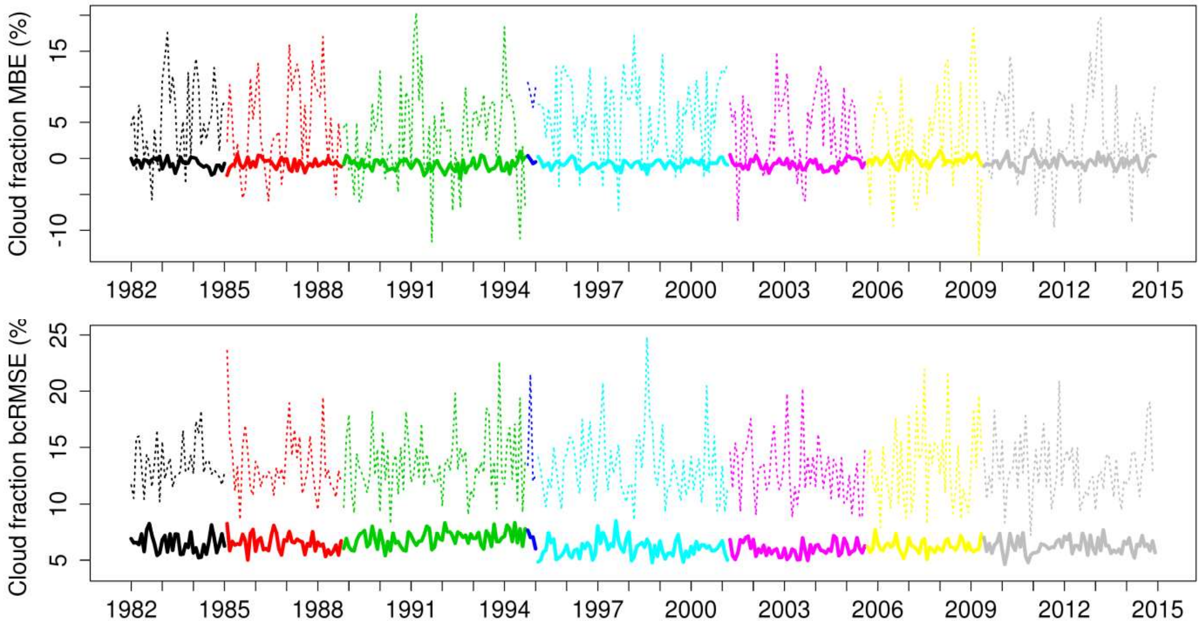

Figure 3: **Time series of mean bias error (upper panel) and bias-corrected root mean square error (lower panel) of CC4CL-AVHRR-PM as compared to synoptic observations. Dashed line reveals performance of uncorrected data, thick solid line of corrected (debiased) data. Colours represent consecutive satellite missions: NOAA-7 (black), NOAA-9 (red), NOAA-11 (green), NOAA-12 (dark blue), NOAA-14 (light blue), NOAA-16 (purple), NOAA-18 (yellow) and NOAA-19 (grey).**

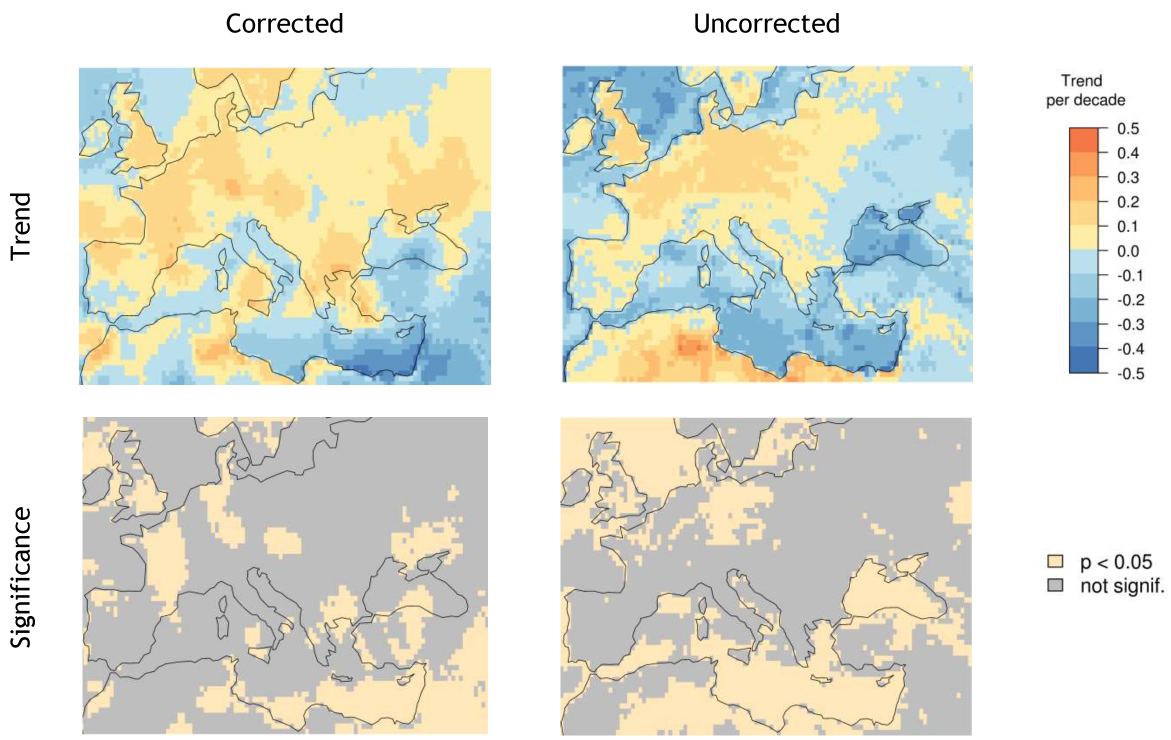

Figure 4: **Map of CC4CL-AVHRR-PM Theil-Sen monotonic trend (upper panels) and its statistical significance according to the Mann-Kendall test adjusted using Benjamini-Hochberg method (lower panels) based on the cloud fraction monthly standardized anomalies in 1982-2014.**

RESULTS

Averaged over 68 evaluation sites (Table 1), the corrected (debiased) dataset (MBE=-0.68%, bcRMSE=6.42%) significantly outperforms the original one (MBE=4.05%, bcRMSE=14.90%). Figure 2 reveals that the performance differs among sites. For the vast majority of them (11 out of 68), the original dataset overestimates the reference SYNOP CFC, while after correction 28 sites overestimate and 40 underestimate the reference. There are 16 sites for which the correction method increased the absolute bias, most of them located at the edges of the interpolation area. Concurrently, the bcRMSE decreased for all sites during correction.

The performance of the debiased dataset (Table 1) is stable among NOAA missions (absolute bias <1%, and bcRMSE within 6-7%). Compared to the uncorrected data, the correction also decreases the performance differences among NOAA satellites. Figure 3 displays concurrently lower mean errors and lower variability. The errors are also more stable in time. These facts reveal that debiasing can implicitly remove the inhomogeneity in CFC time series due to changing overpass times.

Trends in CFC monthly anomalies are comparable in original and corrected dataset (Figure 4). The magnitude of trends is mostly smaller, but the correction keeps their signs unchanged (positive trends are less positive, negative trends are less negative). More prominent differences can be observed over water. It has to be noted however that CFC there was interpolated from inland sites. Therefore, the sharp edges between trends on land and water are obscured in debiased data. Nevertheless, the similarity of trends over land before and after the correction encourages concluding that observed signal is related to natural changes in the last 30 years. Performance of both CFC datasets over water should be further investigated prior any conclusions about potentially observed trends.

CONCLUSIONS

Our study demonstrates that geostatistical merging of satellite-based cloud fractional cover with synoptic observations (called debiasing) significantly improves the accuracy and precision of the former. ESA Cloud_cci's AVHRR-derived dataset (1982-2014) was merged using kriging with external drift with cloud observations at 158 sites over Europe. The correction also decreases the performance differences among NOAA satellites, and implicitly removes the inhomogeneity in cloud fractional cover time series due to changing overpass times. We performed the merging for Europe, however the same approach can easily be transfered to other regions, for which ground-based cloud observations are available.

REFERENCES

Alexandersson, H., (1986) A homogeneity test applied to precipitation data. J. Climatol. 6, 661–675.

Benjamini, Y., and Hochberg, Y., (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57, 289–300.

Devasthale, A., Karlsson, K., Quaas, J., & Grassl, H., (2012) Correcting orbital drift signal in the time series of AVHRR derived convective cloud fraction using rotated empirical orthogonal function. Atmospheric Measurement Techniques, 5, 267-273.

Foster, M. J., Heidinger, A., (2013) PATMOS-x: Results from a Diurnally Corrected 30-yr Satellite Cloud Climatology. J. Clim., 26, 414–425.

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, Ch.J.W., Heuvelink, G.B.M., (2009) Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network, Computers & Geosciences 35, 1711-1721.

Kendall, M.G., (1938) A New Measure of Rank Correlation. Biometrika 30, 81–93.

Khaliq, M.N., Ouarda, T.B.M.J., (2007) On the critical values of the standard normal homogeneity test (SNHT). Int. J. Climatol. 27, 681–687.

Mann, H.B., (1945) Nonparametric Tests Against Trend. Econometrica 13, 245–259.

Norris, J. and A. Evan, (2015) Empirical Removal of Artifacts from the ISCCP and PATMOS-x Satellite Cloud Records. J. Atmos. Oceanic Technol., 32, 691–702.

Poulsen, C. A., Watts, P. D., Thomas, G. E., Sayer, A. M., Siddans, R., Grainger, R. G., Lawrence, B. N., Campmany, E., Dean, S. M., Arnold, C., (2011) Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR. Atmospheric Meas. Tech. Discuss. 4, 2389–2431.
Stapelberg, S., Stengel, M., Karlsson, K.-G., Meirink, J, F., Bojanowski, J.S., and Hollmann, R. (2017) ESA Cloud_cci Product Validation and Intercomparison Report (PVIR), issue 4, rev. 1, 19 april 2017 edn., available from http://www.esa-cloud-cci.org/?q=documentation.
Stengel, M., Stapelberg, S., Sus, O., Schlundt, C., Poulsen, C., Thomas, G., Christensen, M., Carbajal Henken, C., Preusker, R., Fischer, J., Devasthale, A., Willén, U., Karlsson, K.-G., McGarragh, G. R., Proud, S., Povey, A. C., Grainger, D. G., Meirink, J. F., Feofilov, A., Bennartz, R., Bojanowski, J., and Hollmann, R. (2017) Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud_cci project, Earth Syst. Sci. Data, in press.
Stengel, M., Mieruch, S., Jerg, M., Karlsson, K.-G., Scheirer, R., Maddux, B., Meirink, J. F., Poulsen, C., Siddans, R., Walther, A., Hollmann, R., (2015) The Clouds Climate Change Initiative: Assessment of state-of-the-art cloud property retrieval schemes applied to AVHRR heritage measurements. Remote Sens. Environ. 162, 363–379.
Theil, H., (1950) A rank-invariant method of linear and polynomial regression analysis. Parts 1–3. Ned Akad Wetensch 53, 386–392, 521–525, 1397–1412.

ACKNOWLEDGEMENTS

This work was supported by the European Space Agency (ESA) through the Cloud_cci project (contract No.: 4000109870/13/INB) and by the National Science Centre, Poland under the POLONEZ grant No 2015/19/P/ST10/03990 that received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665778. The work was carried out at the Federal Office of Meteorology and Climatology MeteoSwiss and the Institute of Geodesy and Cartography. Computing was performed at the Swiss National Supercomputing Centre (CSCS).
