Meteorologische Zeitschrift, Vol. **16**, No. 5, 525-539 (October 2007) (published online 2007)

Article

Global daily precipitation fields from bias-corrected rain gauge and satellite observations. Part I: Design and Development

MARKUS KOTTEK and FRANZ RUBEL*

Biometeorology Group, University of Veterinary Medicine, Vienna, Austria

(Manuscript received January 16, 2007; in revised form May 7, 2007; accepted May 10, 2007)

Abstract

Global daily precipitation analyses are mainly based on satellite estimates, often calibrated with monthly ground analyses or merged with model predictions. We argue here that an essential improvement of their accuracy is only possible by incorporation of daily ground measurements. In this work we apply geostatistical methods to compile a global precipitation product based on daily rain gauge measurements. The raw ground measurements, disseminated via *Global Telecommunication System* (GTS), are corrected for their systematic measurement errors and interpolated onto a global 1 degree grid. For interpolation ordinary block kriging is applied, with precalculated spatial auto-correlation functions (ACFs). This technique allows to incorporate additional climate information. First, monthly ACFs are calculated from the daily data; second, they are regionalised according to the five main climatic zones of the Köppen-Geiger climate classification. The interpolation error, a by-product of kriging, is used to flag grid points as missing if the error is above a predefined threshold. But for many applications missing values constitute a problem. Due to a combination of the ground analyses with the daily multi-satellite product of the Global Precipitation Climatology Project (GPCP-1DD) not only these missing values are replaced but also the spatial structure of the satellite estimates is considered. As merging method bivariate ordinary co-kriging is applied. The ACFs necessary for the gauge and the satellite fields as well as the corresponding spatial cross-correlation functions (CCFs) are again precalculated for each of the five main climatic zones and for each individual month. As a result two new global daily data sets for the period 1996 up to today will be available on the Internet (www.gmes-geoland.info): A precipitation product over land, analysed from ground measurements; and a global precipitation product merged from this and the GPCP-1DD multi-satellite product.

Zusammenfassung

Globale tägliche Niederschlagsanalysen basieren hauptsächlich auf Satellitenmessungen, oft kalibriert mit Analysen monatlicher Niederschlagsbeobachtungen vom Boden oder kombiniert mit Modellvorhersagen. Eine wesentliche Verbesserung ihrer Genauigkeit ist nur möglich, wenn tägliche Bodenmessungen miteinbezogen werden. In dieser Arbeit wenden wir verschiedene statistische Methoden an um ein globales Niederschlagsprodukt auf Basis von Ombrometermessungen zu erstellen. Zuerst werden die via *Global Telecommunication System* (GTS) empfangenen Niederschlagsbeobachtungen hinsichtlich ihres systematischen Messfehlers korrigiert und auf ein globales 1 Grad Gitter interpoliert. Zur Interpolation wird Ordinary Block Kriging mit vorberechneten räumlichen Autokorrelationsfunktionen (AKFen) verwendet. Diese Technik ermöglicht es, zusätzliche Klimainformation in die Analyse einzubringen. Zum einen werden monatliche AKFen aus den täglichen Daten berechnet; zum anderen werden sie nach den fünf Hauptklimazonen der Köppen-Geiger Klimaklassifikation regionalisiert. Der Interpolationsfehler, ein Nebenprodukt des Krigings, wird verwendet um Gitterpunkte mit Fehlern über einem definierten Grenzwert mit Fehlwerten zu belegen. Für viele Anwendungen stellen Fehlwerte jedoch ein Problem dar. Durch Kombination der Bodenanalysen mit dem täglichen Multi-Satelliten Produkt des Global Precipitation Climatology Project (GPCP-1DD) werden nicht nur diese Fehlwerte abgedeckt, sondern es wird auch die räumliche Struktur der Satellitenniederschläge berücksichtigt. Als Kombinationsmethode wird bivariates Co-Kriging verwendet. Die dafür benötigten AKFen der Bodenanalyse- und Satellitenfelder sowie deren räumliche Kreuzkorrelationsfunktionen (KKFen) werden wieder für jeden individuellen Monat und für die fünf Hauptklimazonen vorberechnet. Als Ergebnis werden zwei neue, globale, tägliche Datensätze für die Periode von 1996 bis zur Gegenwart im Internet (www.gmes-geoland.info) verfügbar sein: Ein Niederschlagsprodukt über Land analysiert aus Bodenmessungen und ein globales Niederschlagsprodukt kombiniert aus ersterem und dem GPCP-1DD Multi-Satelliten Produkt.

1 Introduction

High accuracy global daily precipitation estimates are needed as input for various applications in agricultural sciences (e.g. DE WIT et al., 2005), hydrology (e.g. MARGULIS et al., 2006) and meteorology (e.g. GOTTSCHALCK et al., 2005). Despite the efforts undertaken in the last decade, a (near) real-time global daily precipitation product based not only on satellite estimates but also on operational ground measurements is

*Corresponding author: Franz Rubel, Biometeorology Group, University of Veterinary Medicine Vienna, Veterinärplatz 1, 1210 Vienna, Austria, e-mail: franz.rubel@vu-wien.ac.at

0941-2948/2007/0214 $ 6.75

DOI: 10.1127/0941-2948/2007/0214

still missing today. One can expect, that compared to existing satellite estimates such a precipitation product will have an improved accuracy at least over the land areas. Therefore, within the Core Service bio-geophysical Parameters (CSP) of the European Commission's FP6 Integrated Project GEOLAND[1] an initiative was launched to develop a global daily precipitation product based on bias-corrected rain gauge measurements and existing multi-satellite estimates of precipitation.

Over the past 100 years, substantial progress in global precipitation estimation was made. Based on the first world-wide precipitation maps, compiled at the beginning of the 20th century, it needed about 50 years until the German climatologist Möller published his seasonal precipitation maps for the globe. Further 25 years later, as a milestone, another German climatologist, Jaeger, published his global monthly precipitation maps. Although Jaeger compiled his precipitation maps still by hand, he was the first to provide numerical gridded precipitation values. For that reason his precipitation climatology was widely used for a long time (for a more detailed historical overview as well as references see RUDOLF and RUBEL, 2005). Another 10 years later LEGATES and WILLMOTT (1990) evaluated climatic long-term mean data from world-wide 25 000 rain gauge stations. Using this data set they compiled the first objectively analysed global precipitation climatology based on bias-corrected rain gauge measurements. So far, all global precipitation estimates represent long-term climatic conditions only. With the establishment of the Global Precipitation Climatology Project (GPCP) by the World Climate Research Programme (WCRP) the era of time series of global gridded precipitation estimates began. In 1993 the GPCP delivered the first digital global monthly precipitation maps using gridded gauge data for the land surface and multi-sensor satellite estimates for the oceans. Today, the GPCP-V2 precipitation climatology comprises nearly 30 years (1979-present) of global data fields for individual months by applying a technique to combine various satellite products and gauge data (HUFFMAN et al., 1995; ADLER et al., 2003). These monthly precipitation estimates are calculated operationally on a 2.5° x 2.5° grid in near real-time (about 3 months time delay). Parallel, XIE and ARKIN (1997) compiled the CPC (Climate Prediction Center) Merged Analysis of Precipitation (CMAP) data set also available with monthly resolution on a 2.5° latitude/longitude grid for the period 1979-present. Opposite to the GPCP-V2 product CMAP uses also NCEP/NCAR[2] reanalysis and is therefore not model independent. Similar approaches led to recent climatological data sets with global or regional coverage and with temporal resolutions of up to 3 hours (SHEFFIELD et al., 2006), which have been designed for studies on the terrestrial hydrological cycle rather than for real-time applications.

Today modern computer and network techniques, automatic precipitation gauges as well as new satellites accelerate the progress of improving global precipitation estimates. HUFFMAN et al. (2001) presented the first operational daily precipitation product, the so-called Global Precipitation Climatology Project 1 Degree Daily (GPCP-1DD) product, available for the period 1996-present. As the GPCP-V2 product, the GPCP-1DD is delivered in near real-time and calibrated by gauge data on a monthly basis. Nevertheless, except for various satellite estimates for the tropical belt, a real-time global daily precipitation product based not only on satellite estimates but also on operational ground measurements is still missing today. Main reason is that it is difficult to compile a real-time precipitation product based on rain gauge measurements to have either a pure gauge product over the land areas or a ground truth to calibrate satellite estimates on a daily basis. A daily calibration leads to a significant quality improvement of the satellite estimates at least over the continents as it will be shown in the second part of this paper (RUBEL and KOTTEK, in preparation). With the paper presented here we want to set a first step contributing to the development of a global daily real-time precipitation product based on gauges and satellites. Emphasis lies in the operational analysis of the gauge data by incorporating additional climate information, the application of a gauge-satellite merging algorithm and the verification of this new global daily precipitation product by high resolution continental scale gauge networks over the European Union, East Asia and Australia (part II of this paper, see RUBEL and KOTTEK, in preparation). A verification over North America has been performed in the recently published study from MCPHEE and MARGULIS (2005). In this study the authors demonstrated a relatively high performance of the GPCP-1DD product over the contiguous United States, although verification scores have been calculated only for annual and seasonal aggregated data.

Section 2 of this paper gives an introduction to the ground-based and satellite-based observations used in this study. In section 3 an overview of the statistical structure of precipitation is given. It comprises the calculation of point and block correlation coefficients as well as the fitting of auto-correlation functions (ACFs) and cross-correlation functions (CCFs). The latter are integral parts for the interpolation method introduced in section 4. Generally, a lot of more or less different interpolation methods have been proposed in the literature, especially for the gridding of precipitation. These comprise traditional methods like Thiessen poly-

[1] **G**MES products & services, integrating **EO** monitoring capacities, to support the implementation of European directives and policies related to "**land** cover and vegetation", www.gmes-geoland.info

[2] National Centers for Environmental Prediction / National Center for Atmospheric Research

gons, inverse distance weighting or successive correction methods as well as statistical methods. The latter may be classified in splines (e.g. applied to precipitation by HUTCHINSON, 1995) and optimum interpolation or kriging methods (CRESSIE, 1990). However, from a theoretical point of view, spline and kriging methods are formally equivalent (MATHERON, 1981; DUBRULE, 1983). Here we applied ordinary block kriging and bivariate ordinary co-kriging, which have some potential even when the precipitation data have not been transformed to normal distribution (ATKINSON and LLOYD, 1998; RUBEL and HANTEL, 2001; CLEMENS and BUMKE, 2002). Several authors proposed more sophisticated methods like log-normal kriging (RENDU, 1979; RIVOIRARD, 1990) or disjunctive kriging (YATES et al., 1986), which account for non-Gaussian probability density functions. Such methods have been applied to analyse parameters typically observed by a single-realisation, but rarely applied to multiple realisations of precipitation (a study on artificial precipitation data has been presented by SEO et al., 1999a, 1999b). On a global scale, operational applications of these newer methods are still missing today, because of the problems occurring during back-transformation and the high sensitivity to outliers and extreme values. Moreover, a stable implementation for daily, operational precipitation analyses needs further methodically investigations of these non-linear methods.

For that reason we selected robust ordinary kriging and focus on the main problems of global, daily precipitation analyses. Firstly we perform the regionalisation of the ACFs and CCFs. Typically, in meteorology correlation functions are calculated for different parts of the world as, for example, was demonstrated for the 500 hPa geopotential by BAKER et al. (1987). We propose a new approach, namely the regionalisation according to the well known climate classification after Köppen-Geiger. Secondly the under-sampling of precipitation gauges is considered by applying co-kriging, that is the merging of gauge observations with satellite estimates. The idea behind co-kriging is to benefit from both, the relatively high accuracy of the gauge observations as well as the good spatial coverage of the satellite estimates. The co-kriging method applied follows the theoretical framework of KRAJEWSKI (1987), which provides free parameters to optimise the merging procedure. Resulting precipitation fields and corresponding fields of interpolation errors (error variances) are depicted and discussed in section 5.

2 Data

Several data sets have been used for both, the compilation of the global daily precipitation fields as well as for their verification over selected continental scale regions.

The main input data are those of the synoptic meteorological stations available via the Global Telecommunication System (GTS) network, archived and disseminated by the Global Precipitation Climatology Centre (GPCC) at the German Weather Service. These observations are based on about 7000 synoptic stations worldwide. The precipitation measurements are corrected for systematic measurement errors by applying a statistical correction model developed within the framework of the Baltic Sea Experiment (BALTEX), a European contribution to the Global Energy and Water Cycle Experiment (GEWEX). This bias-correction model, which is based on results from gauge inter-comparison experiments (SEVRUK, 1982; FORLAND et al., 1996; GOODISON et al., 1998) has been described by RUBEL and HANTEL (1999) and was originally applied to global precipitation data sets by UNGERSBÖCK et al. (2001). Its main purpose is to correct for the wind-induced losses, which is the largest error. The correction formulae use observed wind speed and temperature as well as estimated rain intensity. For evaporation and wetting losses, which represent the second largest error of the precipitation measurements, climatological corrections are applied. Further, the correction takes instrument-specific properties into account; these comprise differentiation between unshielded and shielded gauges. Similar statistical bias-correction methods were recently applied by MICHELSON (2004), YE et al. (2004), and XIA (2006).

Indepenent of this synoptic precipitation database, precipitation measurements from about 20000 gauges within the European Union have been collected, 17000 Australian gauges are provided by Australian Bureau of Meteorology and a gridded precipitation data set over East Asia based on 2200 gauges was also available. These data sets are not available in real-time and therefore have exclusively been used for verification purposes as described in part II of this paper.

The next data set used in this study is a gridded Köppen-Geiger climate classification, which has been compiled by KOTTEK et al. (2006). This climate classification is based on recent precipitation fields provided by the GPCC and temperature fields of CRU, the Climatic Research Unit of the University of East Anglia. It has been calculated for the period 1951–2000, provides for each $1^\circ \times 1^\circ$ latitude/longitude grid over the continents one out of 5 main climate classes and was used to regionalise the statistical structure of global precipitation by means of these 5 climate classes.

The last data set is the GPCP-1DD multi-satellite product (HUFFMAN et al., 2001). It is available on the Internet and provides long-record estimates of daily precipitation over the globe on a $1^\circ \times 1^\circ$ latitude/longitude grid as a combination of different satellite estimates. Within the 40° N-S belt data from geostationary satellite's infrared (IR) sensors are used to compute the threshold-matched precipitation index (TMPI) to calculate gridded precipitation estimates at 3-hourly intervals, which are summed to produce the daily value. Estimates outside this belt are computed based on recalibrated Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) data from the polar-orbiting satellites NOAA-12 and NOAA-14. These two satellites with 01:30/13:30 and 07:30/19:30 local time equator crossing times are flying simultaneously and provide global precipitation estimates based on an empirical relationship between rain gauge observations and a function of the cloud-top pressure, fractional cloud cover and relative humidity profile. Beginning with May 2005, Atmospheric Infrared Sounder (AIRS) precipitation estimates have replaced the TOVS estimates at high latitudes because of TOVS instrument termination. The new AIRS data have been adjusted to match the large-scale bias of the TOVS data to maintain homogeneity across the data boundary. Additionally, the GPCP-1DD product is scaled to match the monthly accumulation provided by the GPCP-V2 product for both data regions, which combines satellite and gauge observations at a monthly time scale on a 2.5° × 2.5° grid.

3 Statistical structure of global precipitation

The statistical structure of precipitation is usually described by space and time auto-correlation functions (ACFs). Depending on the specific author equivalent covariance or variogram functions are used (BACCHI and KOTTEGODA, 1995). Precipitation is a complex physical phenomenon; it is not possible to fully describe precipitation processes by just one global ACF. Ideally, precipitation events should be classified by type, e.g. stratiform or convective type precipitation. Each type of precipitation could then be assumed to be a part of an ensemble of homogeneous realisations with defined statistical properties. However, large- and meso-scale precipitation events on the globe contain stratiform and convective components of different intensities and compositions at the same time and it is hardly possible to separate them.

Therefore, in this application the daily precipitation measurements have been used to calculate spatial ACFs for each individual month. This considers higher frequencies of stratiform precipitation during Northern hemisphere winter and higher frequencies of convective type precipitation during summer months as demonstrated by e.g. BERNDTSSON (1987). Additionally, in order to take care of the structure of global precipitation, the spatial ACFs have been regionalised by applying the five main climate zones as defined by the well-known Köppen-Geiger climate classification.

The following procedure describes the empirical estimation of spatial ACFs of precipitation, for each time t, where zero values are generally excluded from the calculation (BACCHI and KOTTEGODA, 1995). Assuming that there exist n instantaneous gauge observations of the precipitation process $Z_{G_t}(\mathbf{u}_i)$, $\mathbf{u}_i \equiv (x_i, y_i)$, $i = 1,\ldots,n$, with multiple realisations $t = 1,\ldots,T$, the auto-correlation R_{GG} (gauge-gauge) is defined as

$$R_{GG}(\mathbf{u}_i,\mathbf{u}_j) = \frac{\sum_{t=1}^{T}[Z_{G_t}(\mathbf{u}_i) - m_G(\mathbf{u}_i)][Z_{G_t}(\mathbf{u}_j) - m_G(\mathbf{u}_j)]}{\sqrt{\sum_{t=1}^{T}[Z_{G_t}(\mathbf{u}_i) - m_G(\mathbf{u}_i)]^2 \sum_{t=1}^{T}[Z_{G_t}(\mathbf{u}_j) - m_G(\mathbf{u}_j)]^2}}, \tag{3.1}$$

where the index G indicates that gauge measurements are used. $m_G(\mathbf{u}_i)$ is the time mean of all realisations of the precipitation process $Z_{G_t}(\mathbf{u}_i)$ at the location $\mathbf{u}_i$.

Assuming further that the precipitation process is homogeneous and isotropic within the model domain (GANDIN, 1993), the correlations depend only on the interstation distance ρ; they are independent of the geographical locations of the stations. Especially on small scales this assumption is rarely satisfied in reality. On large scales or for specific climate zones, however, it leads to acceptable results and provides an opportunity to analyse precipitation fields even over regions with poor gauge coverage.

The result of the empirical analysis of the correlation coefficients is a scatter plot of correlations versus distance. To identify the structure of precipitation a hypothesis about the theoretical ACF has to be made. For the precipitation over the global land areas the non-linear function

$$\widehat{R}_{GG}(\rho) = c_1 \exp\left(-c_2\, \rho^{c_3}\right) \tag{3.2}$$

with the three coefficients c_1, c_2 and c_3 was fitted to the scattergram of the correlation coefficients $R_{GG}(\mathbf{u}_i,\mathbf{u}_j)$. Note that this exponential correlation function has the required property to be positive definite. In order to fit the parameters of the function to the empirically estimated data the coefficients are determined according to the maximum likelihood hypothesis by the use of the least square method. Fig. 1 depicts as an example the northern hemisphere ACFs for January and July as fitted to the corresponding correlation coefficients calculated from the irregularly spaced gauge observations. For practical reasons the scatter plot of the correlation coefficients has been replaced by their mean and standard deviation within pre-defined distance classes (see BACCHI and KOTTEGODA (1995) for a more detailed description on methods to estimate the statistical structure of precipitation). To ensure an adequate sample size, daily precipitation measurements from the entire period 1997 to 2003 have been used.

The ACFs depicted in Fig. 1 demonstrate typical winter and summer conditions of statistical precipitation structure for the northern hemisphere separated by climatic zones. ACFs for the other months are lying within the range of the summer and winter ACFs. For considerations on the southern hemisphere, where the data density is too low to calculate accurate ACFs, simply the summer and winter months were switched. This implies

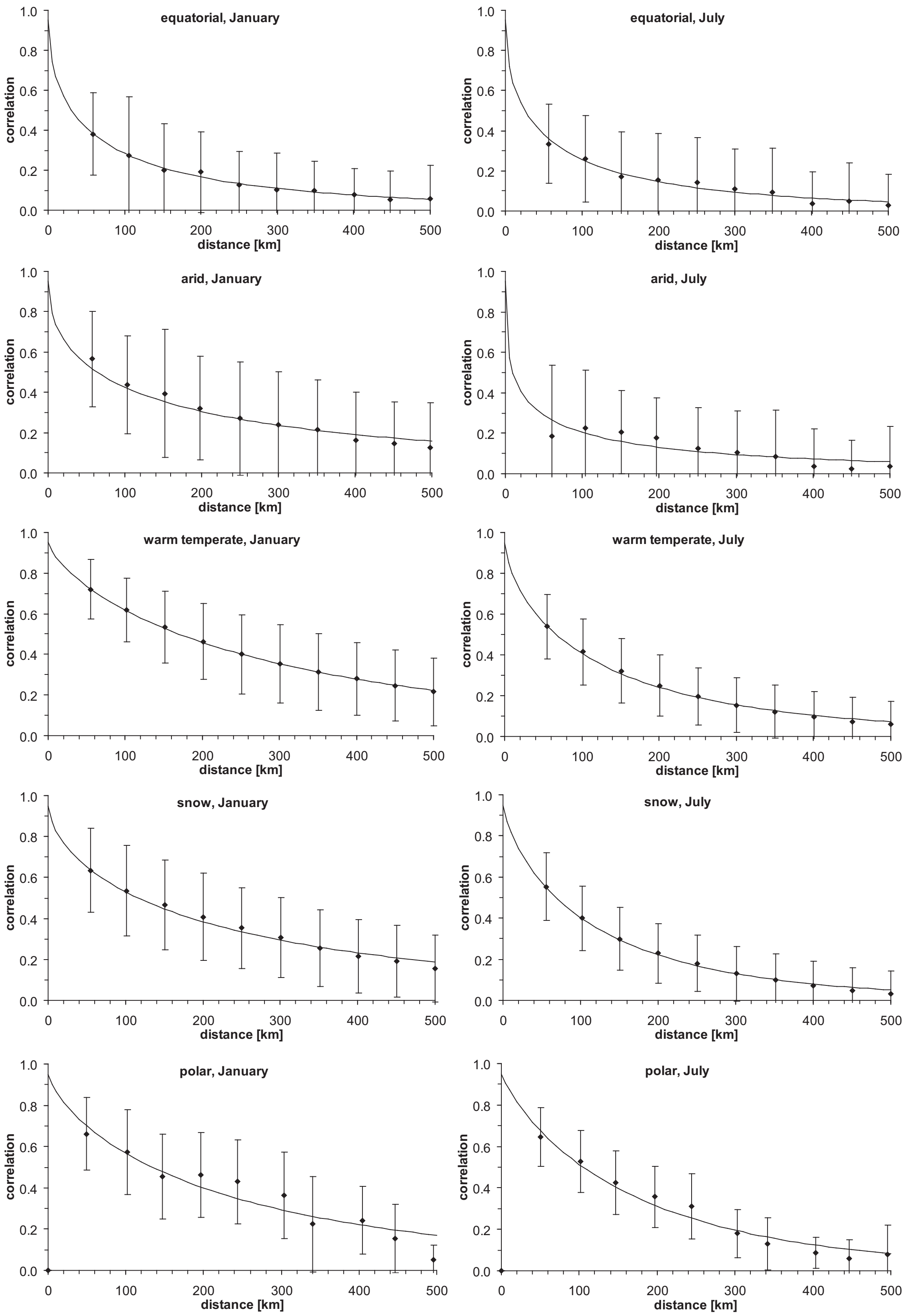

Figure 1: Correlation coefficients estimated from multiple realisations (dots) of synoptic rain gauge measurements with standard deviations (vertical bars) and fitted auto-correlation functions $\widehat{R}_{GG}(\rho)$ (lines) for the five main climatic zones of the northern hemisphere for January (left) and July (right). Temporal resolution 1 day, period 1997–2003.

the assumption that the ACFs depend only on the climate zones. From the curve of the correlation function one can infer the type of precipitation. A strong decrease of the correlation function is characteristic for convective type precipitation. A weak decrease is typical for higher amounts of stratiform precipitation. For the equa-

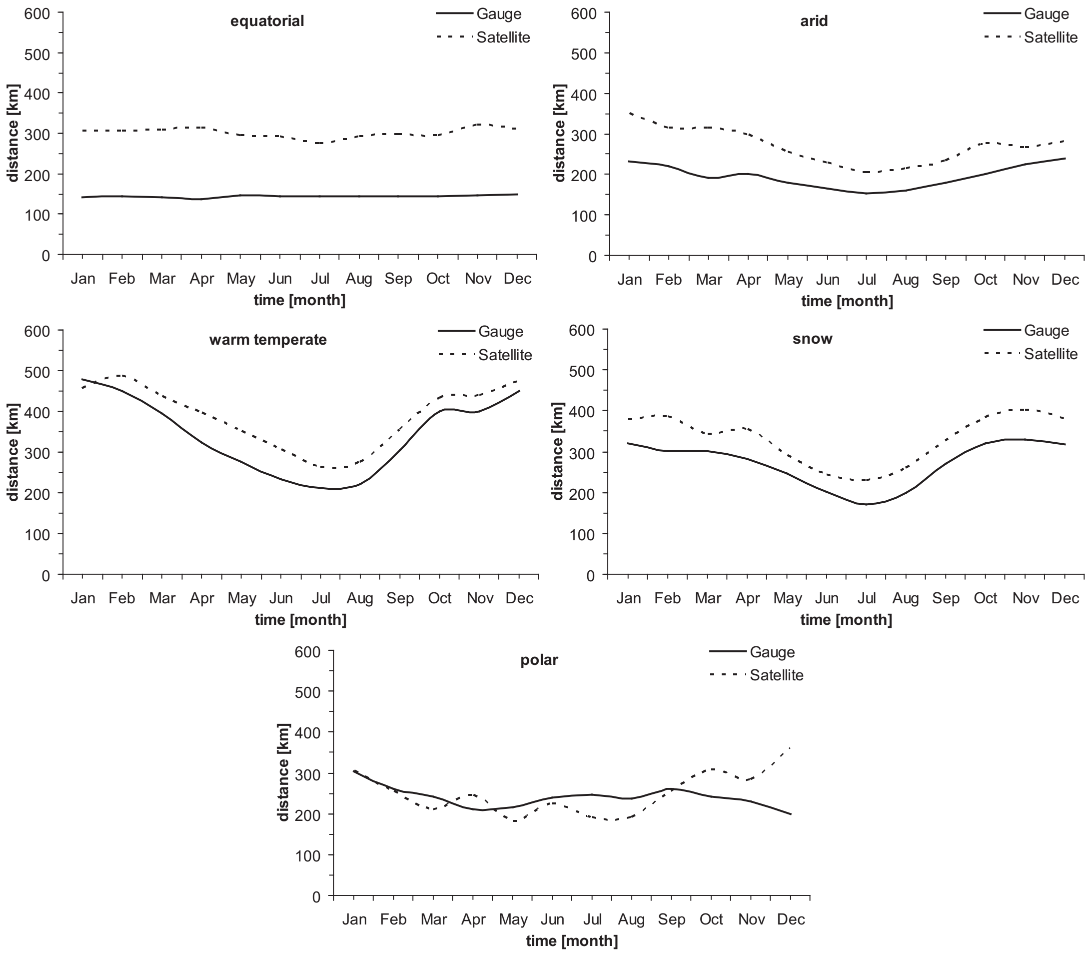

Figure 2: Annual cycle of the decorrelation distances from block auto-correlation functions of precipitation fields analysed from gauges (solid line) and multi-satellite estimates (dashed line), separated for the five main climatic zones of the northern hemisphere. Period 1997–2003.

torial zone both correlation curves for January and July, respectively, show a strong decrease with distance and almost no difference in both curves can be seen. This is consistent with the fact that rainfall in the tropical zone is basically of convective type the whole year round. The warm temperate climate shows a weak decrease in northern winter, typical for stratiform precipitation, contrary to a much stronger decrease in summer, characterised by convective type precipitation. A similar annual cycle is depicted for the snow climate. These two climate zones are characterised by the highest coverage of synoptic rain gauges on the globe, resulting in smooth correlation functions. The opposite picture is shown for the polar zones. Data availability is very rare as it is for the equatorial zone too. The shape of the ACFs for the arid zone shows similar behaviour as for the warm temperate and snow climate ACFs, but slopes much faster. A complete compilation of the parameters describing the 60 ACFs calculated from the synoptic precipitation gauges is given in Tab. 1.

Until now it was assumed that the ACFs were determined from *in-situ* observations; in the following these ACFs will be referred to as *point ACFs*. For the co-kriging method described in the next section also the statistical structure of the precipitation fields analysed from the gauge data and estimated from the satellites are required. Both fields are available on a $1^\circ \times 1^\circ$ grid. The ACFs determined from these gridded data, in the following referred to as *block ACFs*, are characterised by the fact that the correlations are higher than for comparable point ACFs. This is a consequence of the spatial aggregation, which shows a similar effect as the temporal aggregation (RUBEL, 1996b), e.g. annually accumulated precipitation measurements are higher correlated than hourly values. The block correlation coefficients are again calculated according to eq. (3.1), where the index S indicates for satellite. Furthermore, for the co-kriging method also the statistical correlation between both precipitation fields is used. It is described by block cross-correlation functions (block CCFs) fitted to the cross-correlation coefficients R_{GS} (gauge-satellite),

$$R_{GS}(\mathbf{u}_i,\mathbf{u}_j) = \frac{\sum_{t=1}^{T}[Z_{G_t}(\mathbf{u}_i)-m_G(\mathbf{u}_i)][Z_{S_t}(\mathbf{u}_j)-m_S(\mathbf{u}_j)]}{\sqrt{\sum_{t=1}^{T}[Z_{G_t}(\mathbf{u}_i)-m_G(\mathbf{u}_i)]^2\sum_{t=1}^{T}[Z_{S_t}(\mathbf{u}_j)-m_S(\mathbf{u}_j)]^2}}, \tag{3.3}$$

where $m_G(\mathbf{u}_i)$ is the mean of the realisations of precipitation observed by gauges $Z_{G_t}(\mathbf{u}_i)$ at the block $\mathbf{u}_i$ and $m_S(\mathbf{u}_j)$ is the mean of precipitation observed by satellites $Z_{S_t}(\mathbf{u}_j)$ at the block $\mathbf{u}_j$. Again, the block ACFs ($\widehat{R}_{GG}$ and $\widehat{R}_{SS}$) and block CCFs ($\widehat{R}_{GS}$) are described by functions of the type given in eq. (3.2) and determined separately for each month and main climatic zone. The corresponding parameters describing these functional relationships are summarised in Tab. 2.

Altogether 180 block correlation functions have been calculated, which may be characterised by the so-called decorrelation distance ρ_d. It is commonly defined as the distance at which the correlation decreases below $1/e \sim 0.37$ and follows from eq. (3.2) with

$$\rho_d = \left(\frac{\log(c_1)+1}{c_2}\right)^{\frac{1}{c_3}} . \qquad (3.4)$$

Fig. 2 depicts the annual cycle of these decorrelation distances for each of the climate zones as calculated from gauge block ACFs (solid line) and satellite block ACFs (dashed line), respectively. Except for the polar regions, the decorrelation distances calculated from satellite fields are generally higher than those calculated from gauge fields. Apparently, the satellite fields are much smoother than the fields analysed from the gauge data, resulting in higher spatial correlations. Further, the equatorial and polar climates show basically no annual cycle in its precipitation, whereas expectedly the arid, warm temperate and snow climates have a more or less distinctive annual cycle of summer (convective) and winter (stratiform) precipitation.

The CCFs calculated from gauge and satellite precipitation fields demonstrate that the correlations between these two precipitation fields are generally much lower than those of the ACFs discussed above. As depicted in Tab. 2, the coefficients c_1, describing the correlations at zero distance, are in the range of 0.21–0.48. Therefore, the concept of decorrelation distances holds for the CCFs of warm temperate (C) and snow (D) climate. For all other climate zones – equatorial (A), arid (B) and polar (E) – the cross-correlations are below 1/e. These low cross-correlations may be explained by a general dissimilarity of the gauge and satellite fields in latter climate zones, caused by both, the undersampling of the gauge data as well as by the relatively low variance of the satellite fields compared to the gauge data, respectively. By reason of reproducibility the coefficients c_1, c_2 and c_3 of the applied ACFs and CCFs of type $\widehat{R}(\rho) = c_1 \exp(-c_2 \rho^{c_3})$ as well as the corresponding decorrelation distances are given in Tab. 1 and Tab. 2.

4 Geostatistical interpolation

The irregular spaced, bias-corrected, synoptic rain gauge measurements are interpolated onto a $1^\circ \times 1^\circ$ latitude/longitude grid by the use of the well known method

Table 1: Point auto-correlation functions (ACFs) of the bias-corrected synoptic rain gauges for the five main climates on the northern hemisphere. The ACFs are of type $\widehat{R}(\rho) = c_1 \exp(-c_2 \rho^{c_3})$ with coefficients c_1, c_2 and c_3, characterised by the decorrelation distance ρ_d in km.

	ACFs gauge				
	month	c_1	c_2	c_3	ρ_d
equatorial (A)	Jan	0.95	0.1029	0.534	64
	Feb	0.95	0.0605	0.632	78
	Mar	0.95	0.0539	0.644	86
	Apr	0.95	0.1292	0.503	53
	May	0.95	0.0388	0.711	89
	Jun	0.95	0.0984	0.530	72
	Jul	0.95	0.1191	0.520	54
	Aug	0.95	0.0699	0.602	76
	Sep	0.95	0.1716	0.445	47
	Oct	0.95	0.1017	0.531	67
	Nov	0.95	0.0929	0.534	78
	Dec	0.95	0.0749	0.559	94
arid (B)	Jan	0.95	0.0808	0.499	139
	Feb	0.95	0.0215	0.750	156
	Mar	0.95	0.0373	0.671	124
	Apr	0.95	0.0626	0.553	136
	May	0.95	0.0671	0.561	112
	Jun	0.95	0.0657	0.594	89
	Jul	0.95	0.2782	0.371	27
	Aug	0.95	0.0809	0.562	80
	Sep	0.95	0.0870	0.537	86
	Oct	0.95	0.0757	0.527	121
	Nov	0.95	0.0381	0.646	146
	Dec	0.95	0.0314	0.659	176
warm temperate (C)	Jan	0.95	0.0134	0.753	285
	Feb	0.95	0.0130	0.771	261
	Mar	0.95	0.0085	0.863	237
	Apr	0.95	0.0167	0.787	170
	May	0.95	0.0288	0.710	137
	Jun	0.95	0.0350	0.688	121
	Jul	0.95	0.0355	0.690	117
	Aug	0.95	0.0358	0.680	124
	Sep	0.95	0.0143	0.807	181
	Oct	0.95	0.0093	0.841	244
	Nov	0.95	0.0064	0.926	222
	Dec	0.95	0.0061	0.902	268
snow (D)	Jan	0.95	0.0312	0.636	215
	Feb	0.95	0.0498	0.568	180
	Mar	0.95	0.0227	0.703	201
	Apr	0.95	0.0193	0.749	182
	May	0.95	0.0201	0.757	162
	Jun	0.95	0.0322	0.696	129
	Jul	0.95	0.0256	0.763	114
	Aug	0.95	0.0294	0.715	129
	Sep	0.95	0.0113	0.842	192
	Oct	0.95	0.0110	0.819	232
	Nov	0.95	0.0165	0.753	217
	Dec	0.95	0.0250	0.680	210
polar (E)	Jan	0.95	0.0164	0.749	226
	Feb	0.95	0.0568	0.551	165
	Mar	0.95	0.0280	0.694	160
	Apr	0.95	0.0177	0.826	124
	May	0.95	0.0197	0.776	148
	Jun	0.95	0.0175	0.789	157
	Jul	0.95	0.0118	0.859	166
	Aug	0.95	0.0285	0.689	162
	Sep	0.95	0.0105	0.865	182
	Oct	0.95	0.0224	0.734	164
	Nov	0.95	0.0286	0.693	157
	Dec	0.95	0.0665	0.545	131

Table 2: Block auto-correlation functions (ACFs) of the gridded gauge data (left), the satellite data (middle) and cross-correlation functions (CCFs) between gridded gauge and satellite estimates (right) for the five main climates on the northern hemisphere. The ACFs and CCFs are of type $\widehat{R}(\rho) = c_1 \exp(-c_2\,\rho^{c_3})$ with coefficients c_1, c_2 and c_3, characterised by the decorrelation distance ρ_d in km.

		ACFs gauge				ACFs satellite				CCFs gauge/satellite			
	month	c_1	c_2	c_3	ρ_d	c_1	c_2	c_3	ρ_d	c_1	c_2	c_3	ρ_d
equatorial (A)	Jan	0.95	0.0012	1.353	142	0.95	0.0020	1.075	308	0.25	5.8E-4	1.250	–
	Feb	0.95	0.0011	1.362	143	0.95	0.0020	1.079	308	0.26	9.4E-4	1.167	–
	Mar	0.95	0.0019	1.259	141	0.95	0.0021	1.062	311	0.27	7.2E-4	1.201	–
	Apr	0.95	0.0011	1.370	138	0.95	0.0022	1.054	315	0.26	1.3E-3	1.096	–
	May	0.95	0.0017	1.268	145	0.95	0.0018	1.099	294	0.25	9.6E-4	1.166	–
	Jun	0.95	0.0022	1.219	144	0.95	0.0022	1.065	293	0.23	4.7E-4	1.285	–
	Jul	0.95	0.0015	1.300	145	0.95	0.0028	1.034	276	0.24	4.1E-4	1.323	–
	Aug	0.95	0.0015	1.300	145	0.95	0.0027	1.032	293	0.24	9.5E-4	1.161	–
	Sep	0.95	0.0013	1.331	143	0.95	0.0025	1.042	297	0.26	6.8E-4	1.216	–
	Oct	0.95	0.0015	1.301	144	0.95	0.0024	1.049	295	0.28	5.8E-4	1.248	–
	Nov	0.95	0.0012	1.341	145	0.95	0.0020	1.070	322	0.29	4.9E-4	1.277	–
	Dec	0.95	0.0025	1.186	149	0.95	0.0024	1.040	312	0.28	7.6E-4	1.196	–
arid (B)	Jan	0.95	0.0013	1.204	233	0.95	0.0076	0.824	351	0.25	1.0E-4	1.443	–
	Feb	0.95	0.0010	1.277	221	0.95	0.0086	0.817	315	0.21	1.8E-5	1.769	–
	Mar	0.95	0.0014	1.234	191	0.95	0.0051	0.908	316	0.24	9.4E-6	1.893	–
	Apr	0.95	0.0026	1.111	201	0.95	0.0046	0.936	298	0.26	1.4E-4	1.450	–
	May	0.95	0.0007	1.379	178	0.95	0.0055	0.931	256	0.26	1.3E-5	1.896	–
	Jun	0.95	0.0005	1.463	166	0.95	0.0056	0.947	228	0.25	1.8E-5	1.867	–
	Jul	0.95	0.0002	1.689	153	0.95	0.0046	1.002	205	0.25	9.3E-6	2.000	–
	Aug	0.95	0.0005	1.478	160	0.95	0.0053	0.966	216	0.28	1.3E-4	1.527	–
	Sep	0.95	0.0008	1.361	179	0.95	0.0042	0.992	233	0.29	1.5E-4	1.499	–
	Oct	0.95	0.0022	1.145	201	0.95	0.0039	0.977	278	0.33	1.2E-4	1.468	–
	Nov	0.95	0.0009	1.285	224	0.95	0.0055	0.922	268	0.24	1.1E-5	1.854	–
	Dec	0.95	0.0008	1.287	239	0.95	0.0066	0.882	281	0.21	6.7E-6	1.928	–
warm temperate (C)	Jan	0.95	0.0001	1.503	479	0.95	0.0019	1.018	458	0.42	3.8E-6	1.908	244
	Feb	0.95	0.0001	1.505	450	0.95	0.0016	1.027	489	0.45	2.6E-6	1.993	284
	Mar	0.95	0.0002	1.459	395	0.95	0.0017	1.040	438	0.40	2.5E-6	2.031	160
	Apr	0.95	0.0004	1.339	323	0.95	0.0016	1.064	397	0.41	5.8E-6	1.936	163
	May	0.95	0.0009	1.244	276	0.95	0.0018	1.066	353	0.38	1.3E-5	1.832	72
	Jun	0.95	0.0011	1.239	232	0.95	0.0020	1.080	307	0.40	1.7E-5	1.827	97
	Jul	0.95	0.0011	1.259	211	0.95	0.0020	1.108	265	0.38	1.1E-5	1.921	67
	Aug	0.95	0.0012	1.237	221	0.95	0.0017	1.125	277	0.40	9.4E-6	1.930	109
	Sep	0.95	0.0004	1.371	304	0.95	0.0015	1.103	355	0.46	1.1E-5	1.850	213
	Oct	0.95	0.0001	1.469	400	0.95	0.0011	1.111	434	0.48	5.0E-6	1.922	287
	Nov	0.95	0.0002	1.407	399	0.95	0.0014	1.072	441	0.45	2.8E-6	2.006	260
	Dec	0.95	0.0001	1.479	449	0.95	0.0015	1.049	476	0.42	1.9E-6	2.032	236
snow (D)	Jan	0.95	0.0003	1.423	320	0.95	0.0107	0.755	380	0.37	2.9E-5	1.639	–
	Feb	0.95	0.0005	1.320	302	0.95	0.0092	0.778	386	0.37	2.5E-5	1.666	28
	Mar	0.95	0.0003	1.424	302	0.95	0.0088	0.801	345	0.38	2.8E-5	1.674	60
	Apr	0.95	0.0003	1.410	283	0.95	0.0057	0.870	355	0.41	1.7E-5	1.764	133
	May	0.95	0.0004	1.406	246	0.95	0.0061	0.889	292	0.39	2.4E-5	1.759	95
	Jun	0.95	0.0009	1.303	203	0.95	0.0067	0.900	244	0.38	2.6E-5	1.771	66
	Jul	0.95	0.0011	1.316	171	0.95	0.0058	0.938	229	0.36	3.6E-5	1.743	–
	Aug	0.95	0.0009	1.312	199	0.95	0.0042	0.972	261	0.38	2.5E-5	1.780	68
	Sep	0.95	0.0003	1.430	271	0.95	0.0039	0.947	326	0.43	2.7E-5	1.717	156
	Oct	0.95	0.0002	1.491	320	0.95	0.0052	0.876	385	0.45	2.6E-5	1.686	196
	Nov	0.95	0.0002	1.437	330	0.95	0.0075	0.806	403	0.42	2.2E-5	1.687	174
	Dec	0.95	0.0003	1.413	318	0.95	0.0089	0.785	382	0.37	1.4E-5	1.758	53
polar (E)	Jan	0.95	1.0E-6	2.401	304	0.95	0.0349	0.577	305	0.25	9.7E-6	1.903	–
	Feb	0.95	8.5E-7	2.504	260	0.95	0.0401	0.570	257	0.24	1.8E-5	1.804	–
	Mar	0.95	8.6E-7	2.534	243	0.95	0.0531	0.538	212	0.21	2.8E-3	0.943	–
	Apr	0.95	8.6E-7	2.599	211	0.95	0.0388	0.580	247	0.22	7.9E-6	1.862	–
	May	0.95	1.1E-6	2.549	215	0.95	0.0623	0.523	182	0.25	1.5E-5	1.872	–
	Jun	0.95	9.9E-7	2.512	240	0.95	0.0334	0.618	224	0.35	4.6E-6	2.055	–
	Jul	0.95	1.5E-6	2.420	246	0.95	0.0261	0.683	193	0.32	3.7E-6	2.095	–
	Aug	0.95	1.8E-6	2.403	238	0.95	0.0251	0.691	192	0.33	3.6E-5	1.726	–
	Sep	0.95	1.6E-6	2.390	260	0.95	0.0192	0.703	256	0.33	4.2E-5	1.657	–
	Oct	0.95	1.2E-6	2.479	243	0.95	0.0268	0.623	307	0.30	2.8E-4	1.365	–
	Nov	0.95	1.8E-6	2.422	230	0.95	0.0251	0.643	284	0.31	4.0E-6	2.008	–
	Dec	0.95	1.4E-6	2.534	200	0.95	0.0208	0.649	361	0.30	4.5E-6	1.964	–

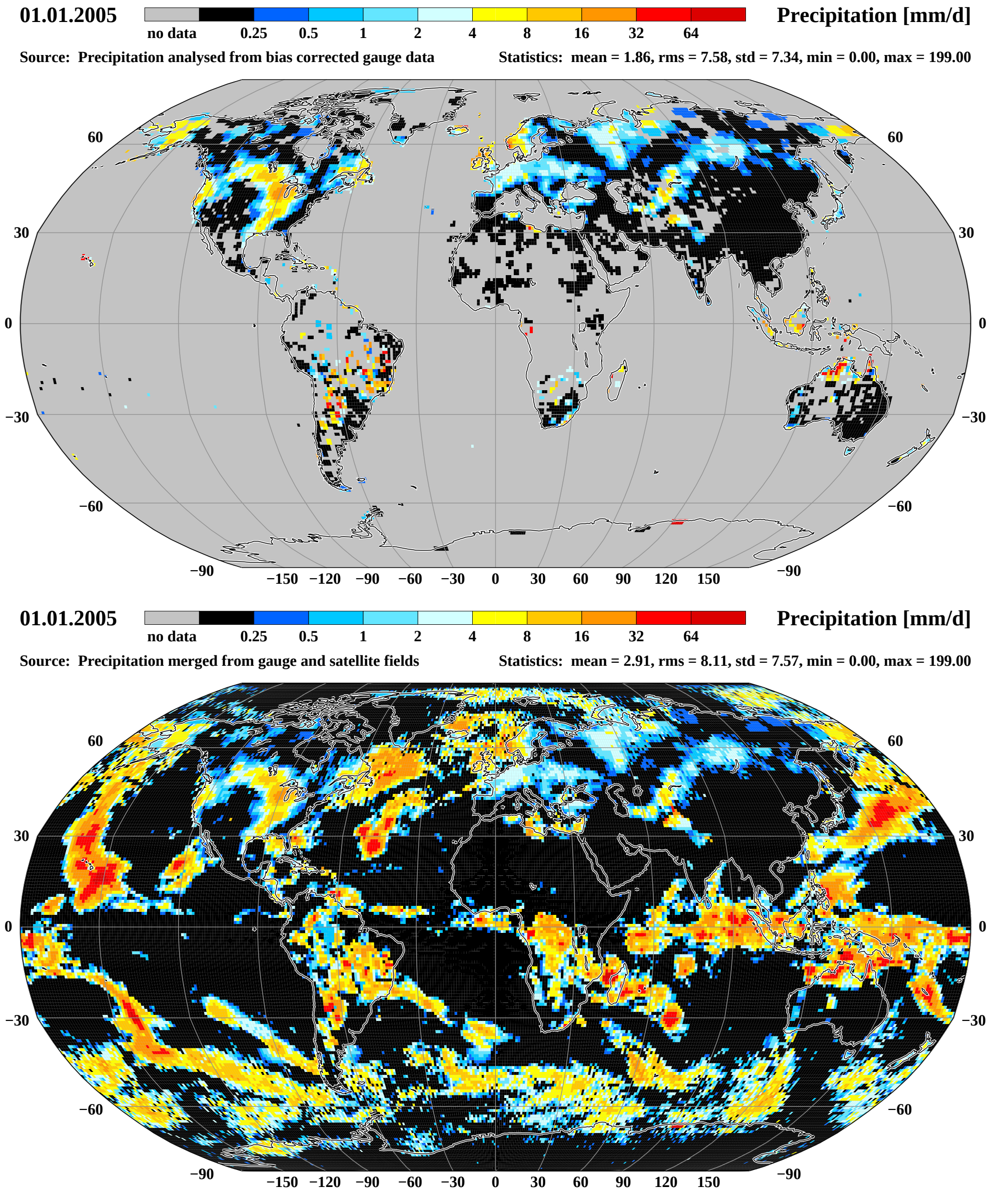

Figure 3: Daily precipitation field analysed from bias corrected synoptic rain gauge measurements (top) and corresponding precipitation field merged from the latter and the GPCP-1DD multi-satellite field (bottom). Date January 1, 2005. Units mm/day.

called *optimal averaging with normalised weights* (GANDIN, 1993) or *ordinary block kriging* (CRESSIE, 1993). Thus, a global precipitation product analysed from ground measurements is produced on a regular grid over the continents. To fill the data gaps over the oceans and over regions, where the synoptic station density is low, this product is combined with the GPCP-1DD multi-satellite product by applying *field merging* or *ordinary co-kriging*. Although both methods are discussed in detail in the geostatistical literature, applications to precipitation are sporadic. Kriging has been applied e.g. by RUBEL (1996a) and RUBEL and HANTEL (2001) to analyse daily precipitation fields over the Baltic Sea drainage basin. Co-kriging has been studied e.g. by KRAJEWSKI (1987) and SEO et al. (1990a), who introduced some numerical experiments with artificial data to merge rain gauge and radar measurements. Therefore, in this paper only a brief statement on ordinary block kriging and co-kriging methods is given while focus is on the application to global precipitation data.

Ordinary block kriging estimates an unknown areal precipitation $Z_A(\mathbf{u}_0)$, defined by

$$Z_A(\mathbf{u}_0) \quad = \quad \frac{1}{A}\int_A Z(\mathbf{u})\,dA, \tag{4.1}$$

from the linear combination

$$\widehat{Z}_A(\mathbf{u}_0) \quad = \quad \sum_{i=1}^{n} \lambda_{G_i} Z_G(\mathbf{u}_i), \tag{4.2}$$

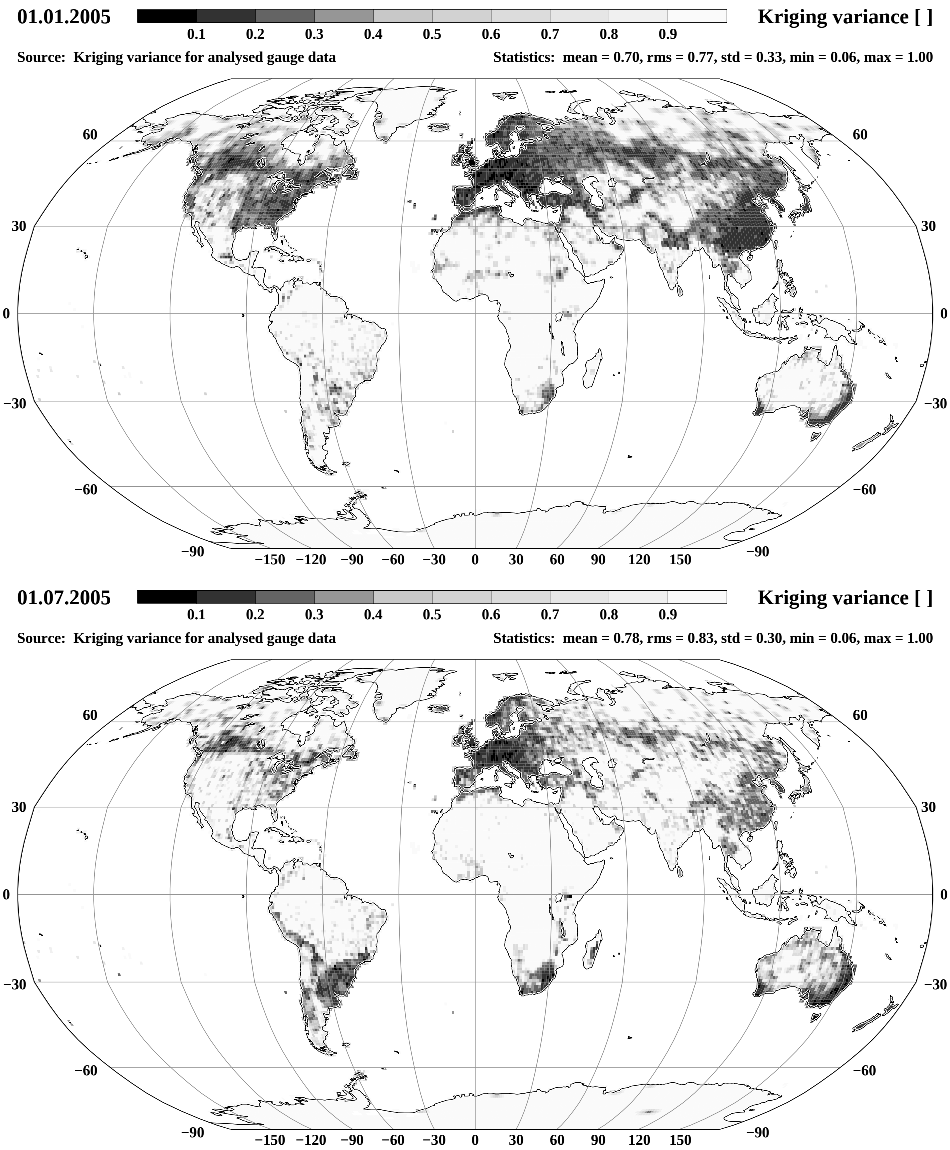

Figure 4: Daily kriging variance (interpolation error normalised with respect to the variance) for January 1, 2005 (top) and July 1, 2005 (bottom).

where $\widehat{Z}_A(\mathbf{u}_0)$ is the precipitation estimate of a grid area A represented by its central point $\mathbf{u}_0 \equiv (x_0, y_0)$, the values $Z_G(\mathbf{u}_i)$ are point precipitation measures from gauges at the locations $\mathbf{u}_i \equiv (x_i, y_i), i = 1, \ldots, n$, and λ_{G_i} are the appropriate weighting factors which have to be estimated.

The statistically optimal determination of the weighting factors follows from the linear system of equations written here in form of the empirically estimated spatial auto-correlation function $\widehat{R}$,

$$\begin{aligned}
&\sum_{j=1}^{n} \lambda_{G_j} \widehat{R}_{GG}(\mathbf{u}_i, \mathbf{u}_j) + \mu = \widehat{R}_{0G}(\mathbf{u}_0, \mathbf{u}_i) \\
&for \quad i = 1, \ldots, n \\
&\sum_{i=1}^{n} \lambda_{G_i} = 1,
\end{aligned} \tag{4.3}$$

where $\widehat{R}_{0G}$ are the correlations between the unknown areal precipitation value and the rain gauge values observed at the location of the stations, $\widehat{R}_{GG}$ are the correlations between the rain gauge observations and μ is the Lagrangian multiplier. Additionally, for each precipitation field the kriging variance (also called estimation variance, error variance or normalised interpolation error),

$$\sigma^2_{GG}(\mathbf{u}_0) \quad = \quad \widehat{R}_{00}(\mathbf{u}_0, \mathbf{u}_0) - \sum_{i=1}^{n} \lambda_{G_i} \widehat{R}_{0G}(\mathbf{u}_0, \mathbf{u}_i) - \mu, \tag{4.4}$$

may be calculated. The kriging variance is a function of the size of the grid area, the station density and the statistical structure of the precipitation field. Eqs. (4.3) and (4.4) are valid for point kriging, i.e. the estimation of an unknown precipitation value $\widehat{Z}(\mathbf{u}_0)$ at the point $\mathbf{u}_0$.

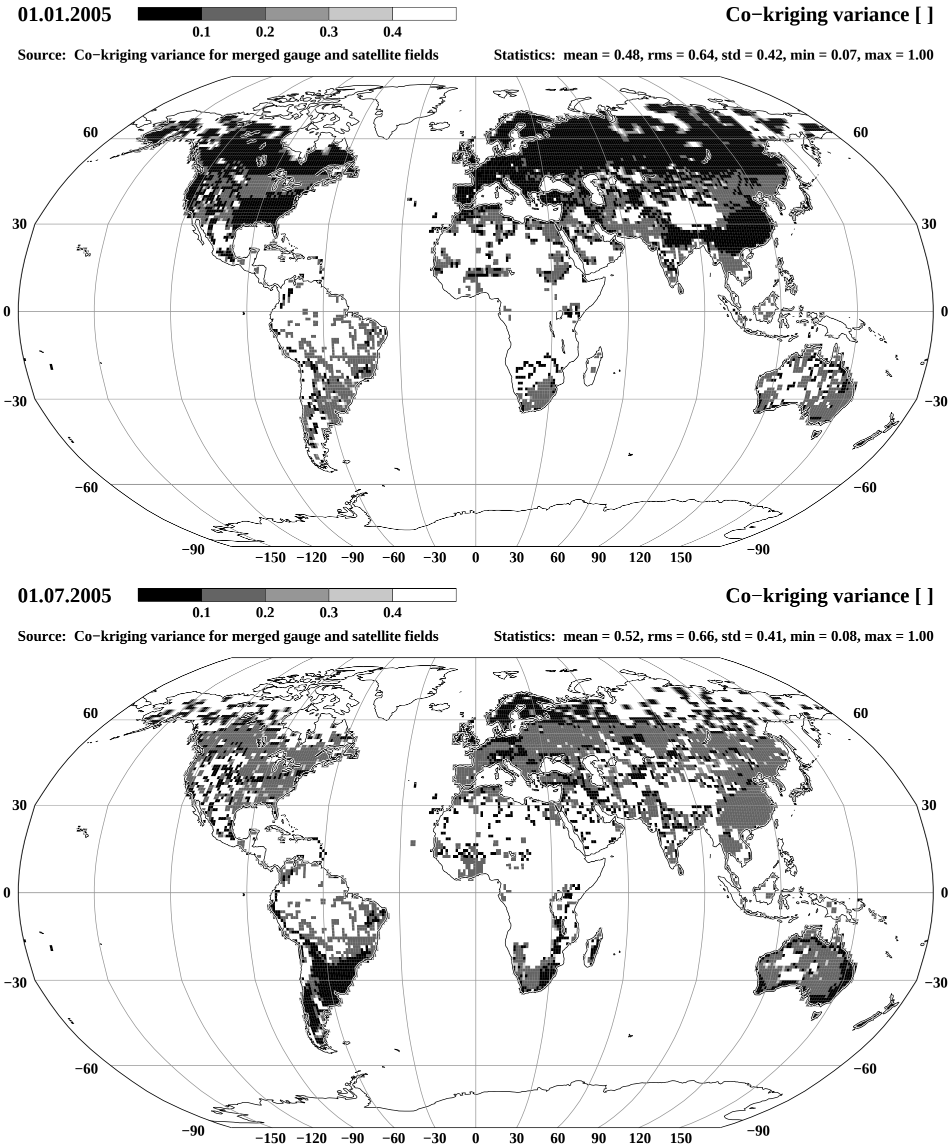

Figure 5: Daily co-kriging variance (merging error normalised with respect to the variance) for January 1, 2005 (top) and July 1, 2005 (bottom). Regions where the co-kriging system was applied are marked by grey and black areas, whereas over white areas only one data set, mostly the satellite data, have been taken directly.

Here we want to estimate an unknown areal precipitation $\widehat{Z}_A(\mathbf{u}_0)$ for a grid area (block) with the central point $\mathbf{u}_0$. Block kriging yields this by integrating the correlations $\widehat{R}_{0G}$ and $\widehat{R}_{00}$ over the grid area A, see e.g. GANDIN (1993),

$$\widehat{R}_{0G}(\mathbf{u}_0, \mathbf{u}_i) = \frac{1}{A}\int_A \widehat{R}_{0G}(\mathbf{u}, \mathbf{u}_i)\, dA \tag{4.5}$$

$$\widehat{R}_{00}(\mathbf{u}_0, \mathbf{u}_0) = \frac{1}{A^2}\int_A\int_{A'} \widehat{R}_{00}(\mathbf{u}, \mathbf{u}')\, dA\, dA'. \tag{4.6}$$

As the correlations of the unknown (true) precipitation values are not known, they are usually approximated by the correlations calculated from the data ($\widehat{R}_{0G} \simeq \widehat{R}_{00} \simeq \widehat{R}_{GG}$). Thus, the optimal estimation of the weights λ_G depends only on the empirically estimated spatial autocorrelation function $\widehat{R}_{GG}(\rho)$. For the numerical integrations of the ACFs in eqs. (4.5) and (4.6) the grid area is approximated by a 5-point operator (RUBEL, 1996a).

A well known problem of kriging applications are negative weights, which may result in negative precipitation values. Although negative weights are seldom, the correction procedure proposed by JOURNEL and RAO (1996) was implemented to avoid them. The correction removes the sample with the largest negative weight; i.e. after solving the kriging system a positive constant equal to the modulus of the largest negative weight (if any) is added to all weights.

In the next step the areal precipitation measurements calculated with eq. (4.2) are merged with precipitation estimates from satellite data available on the same grid (blocks). Standardised ordinary co-kriging (ISAAKS and

SRIVASTAVA, 1989) is used to calculate the merged precipitation $\widehat{Z}_A(\mathbf{u}_0)$ from a linear combination of gauge and satellite blocks,

$$\widehat{Z}_A(\mathbf{u}_0) = \sum_{i=1}^{n} \lambda_{G_i} Z_G(\mathbf{u}_i) + \sum_{j=1}^{m} \lambda_{S_j} (Z_S(\mathbf{u}_j) - \widehat{m}_S + \widehat{m}_G). \quad (4.7)$$

This approach uses only one nonbias condition, namely $\sum \lambda_G + \sum \lambda_S = 1$, which requires the adjustment of the satellite estimates in eq. 4.7 to make sure that their mean is equal to the mean of the gauge values. The means of the gauge values $\widehat{m}_G$ and the satellite values $\widehat{m}_S$ are computed simply as arithmetic averages. Eq. 4.7 was implemented by combining up to $n = 9$ neighbouring blocks of gauge precipitation $Z_G(\mathbf{u}_i)$ linearly with appropriate weights λ_{G_i} and up to $m = 9$ blocks of satellite estimates $Z_S(\mathbf{u}_j)$ with appropriate weights λ_{S_j} as demonstrated by KRAJEWSKI (1987)[3]. The system is laid out to nine precipitation blocks of each field, but not all of the nine blocks have to be covered with precipitation values, i.e. data gaps can occur over the oceans and deserts in the rain gauge fields or over regions with instrument malfunctions in the satellite fields. Then less than nine (but at minimum one) blocks of precipitation of the considered field serve as input to the linear system. Having no available block in one field, the precipitation value from the other field at the considered block is taken directly. The weights λ_{G_i} and λ_{S_j} can be determined from the linear system of equations

$$\begin{aligned}
& \sum_{j=1}^{n} \lambda_{G_j} \widehat{R}_{GG}(\mathbf{u}_i,\mathbf{u}_j) + \sum_{j=1}^{m} \lambda_{S_j} \widehat{R}_{GS}(\mathbf{u}_i,\mathbf{u}_j) \\
+ \quad & \mu = \widehat{R}_{0G}(\mathbf{u}_0,\mathbf{u}_i); \quad i = 1,\ldots,n \\
& \sum_{i=1}^{n} \lambda_{G_i} \widehat{R}_{GS}(\mathbf{u}_i,\mathbf{u}_j) + \sum_{i=1}^{m} \lambda_{S_i} \widehat{R}_{SS}(\mathbf{u}_i,\mathbf{u}_j) \\
+ \quad & \mu = \widehat{R}_{0S}(\mathbf{u}_0,\mathbf{u}_j); \quad j = 1,\ldots,m \\
& \sum_{i=1}^{n} \lambda_{G_i} + \sum_{j=1}^{m} \lambda_{S_j} = 1,
\end{aligned} \quad (4.8)$$

where $\widehat{R}_{0G}$ and $\widehat{R}_{0S}$ are now (unknown) correlation coefficients between the true block process and the gauge or satellite blocks, respectively. $\widehat{R}_{GG}$ are the correlation coefficients of the gauge blocks, $\widehat{R}_{SS}$ the correlations of the satellite blocks and $\widehat{R}_{GS}$ the cross-correlations between gauge and satellite blocks. μ is the Lagrangian multiplier. Finally, the co-kriging variance may be calculated as

$$\begin{aligned}
\sigma^2_{GS}(\mathbf{u}_0) = \widehat{R}_{00}(\mathbf{u}_0,\mathbf{u}_0) \quad & - \quad \sum_{i=1}^{n} \lambda_{G_i} \widehat{R}_{0G}(\mathbf{u}_0,\mathbf{u}_i) \quad (4.9) \\
& - \quad \sum_{j=1}^{m} \lambda_{S_j} \widehat{R}_{0S}(\mathbf{u}_0,\mathbf{u}_j) - \mu.
\end{aligned}$$

Again, the unknown correlations are approximated by correlations calculated from the data ($\widehat{R}_{0G} \simeq \widehat{R}_{00} \simeq \widehat{R}_{GG}$ and $\widehat{R}_{0S} \simeq \widehat{R}_{GS}$) and the method of JOURNEL and RAO (1996) is applied to avoid negative weights.

5 Results

Results of this study comprise a new approach for the regionalisation of ACFs and CCFs of daily global precipitation by using of the Köppen-Geiger climate classification, the application of these ACFs and CCFs to compile improved precipitation products from gauge and satellite data as well as the resulting precipitation products following from this procedure.

The first new precipitation product introduced here consists of global daily precipitation fields on a $1^\circ \times 1^\circ$ latitude/longitude grid as calculated from bias corrected synoptic rain gauges. These precipitation fields have been compiled by applying the ordinary block kriging method described above and the point ACFs listed in Tab. 1. It covers the period Oct. 1996 to present. Fig. 3, top, depicts as an example the precipitation field for January 1, 2005, in units of mm/day. Grey colours indicate that no observations are available or the station density is too low to perform a precipitation analysis. This is the case over the oceans as well as over major parts of the land regions within the 30° N-S belt. Over the mid-latitudes however, the station density is high enough to analyse daily precipitation with a spatial resolution of 1°. Combining this first precipitation product with the GPCP-1DD satellite data leads to the second product, the merged precipitation product, as e.g. depicted for January 1, 2005, in Fig. 3, bottom. It has been calculated by applying ordinary co-kriging with the block ACFs and CCFs listed in Tab. 2. Note that the calculated ACFs and CCFs are very uncertain over the sparse gauged polar regions. Therefore, the co-kriging system was not applied over the polar regions, where the merged product has been determined directly from either the satellite estimates or, if available, from the gauge values. Again, it covers the period Oct. 1996 to present.

Additionally, the kriging variance, a kind of interpolation error within the interval 0 (no error) and 1 (maximum error = variance), is given for all precipitation fields and depicted as example for January 1 and July 1, 2005 in Fig. 4. It demonstrates typical winter and summer conditions in error characteristics. Generally, lower errors occur in winter time due to mainly stratiform precipitation corresponding to higher correlations and wider decorrelation distances (northern hemisphere

[3]Note, that KRAJEWSKI (1987) applied traditional ordinary 5-point co-kriging with the two nonbias conditions $\sum \lambda_G = 1$ and $\sum \lambda_S = 0$, in which the second condition tends to limit severely the influence of the secondary (remote sensed) variable.

in Fig. 4, top, and southern hemisphere in Fig. 4, bottom). The opposite picture is shown in summer. White areas in the figures indicate maximum errors identical to missing values. The kriging variance enables the user to decide whether or not a grid values interpolation error is low enough to be used in a certain application. The corresponding co-kriging variance is shown in Fig. 5. It depicts the location of grid areas at which a merged precipitation of different accuracy (black and grey areas) was estimated by co-kriging and at which only one data set was available (white areas).

At least over land areas, the precipitation products presented here offer the possibility to continue the time series of global daily precipitation estimates, where the GPCP-1DD product failed due to operational problems with single satellites. Fig. 6 depicts the precipitation time series for the period 1996-present. As shown in the figure, the global mean of precipitation varies between 2.5 and 2.7 mm/day (long-term mean 2.61 mm/day, after ADLER et al., 2003), the global mean over the oceans between 2.7 and 3.0 mm/day (long-term mean 2.84 mm/day). The global mean of the land areas varies within the range of 1.8 and 2.5 mm/day (long-term mean 2.09 mm/day) whereas differences between the GPCP-1DD and the merged product are marginal. Thus, local improvements do not influence the global values.

6 Discussion

The gauge reporting times which are often not consistent with the accumulation period of the satellite product are an important point not addressed till now. Due to the fact that the synoptic network grew historically and was read off manually in former times the reporting times are slanted toward the local time of the day. For example, following a study undertaken by the GPCC for January 23, 1999, a total of 4 209 synoptic stations were reporting precipitation. Thereof 1 057 stations reported at 00:00 UTC, located in Asia, on Pacific islands as well as in Mexico. Further 150 stations located in Australia have reporting times at 23:00 UTC, thus 1 hour earlier. In Europe, Africa and Canada 2 388 stations report 6 hours later namely at 06:00 UTC. The reporting time of the 568 synoptic stations in the U.S. is 12:00 UTC. The rest of 46 stations has other reporting times. From this it follows that only a bit more than 25 % of the gauge reporting times coincide with the satellite estimates. Contrary the satellite precipitations in the outer tropics are merely based on 4 instantaneous measurements per day from NOAA satellites with 01:30/13:30 and 07:30/19:30 local time equator crossing times (see section 2). Thus starting with 00:00 UTC the accumulation time is an extrapolated one. Depending on the geographical location the first measurement to determine the daily accumulated precipitation can occur hours after 00:00 UTC. Hence the temporal allocation of the satellite estimates as well as the ground measurements is a general problem. As will be shown in the second part of this paper (RUBEL and KOTTEK, in preparation) the sampling error of the satellite precipitations exceeds the reporting time induced error of the gauges by far, whereas this study is restricted to reporting time errors up to 6 hours (these are 85 % of all gauges except for the U.S. and South America). With increasing automation of the gauge measurements one can expect also increasing global coverage of 6 hourly gauge reports making it possible to adjust the accumulation period of the gauges to those of the satellite estimates.

Another important point is related to the applied interpolation method. Appearing from eq. 4.2 and eq. 4.7 the kriging and the co-kriging method, respectively, demonstrate a statistically optimal procedure to calculate a weighted mean. Note that in either case a weighted mean implies a certain smoothing, what generally is not desirable. This becomes apparent especially at the co-kriging with the used 9-point operator. The intensity of the smoothing is objectively determined by the spatial ACFs and CCFs. Practically, the smoothing is lower as supposed due to the fact that the weights rapidly decrease with increasing distance from the central block. On the other hand, if several of the 9 blocks are not covered with precipitation values the remaining blocks are weighted to a greater extent and the full potential of the used interpolation method arrives to bear. From a comparison of global daily means and variances of the satellite, gauge and merged precipitation products it appears that the daily means differ up to 20 % (on average 8 %). The variance of the gauge product is on average about 11 % higher than the variance of the satellite product (see also the discussion in section 3). Therefore, the merged precipitation field should have a mean adjusted to the gauge product and additionally a higher variance compared to the mere satellite product. Contrariwise one has to accept that the merged field will show a slightly lower variance (about 3 %) compared to the gauge field. Thus, the idea behind co-kriging is that the full spatial coverage of the GPCP-1DD precipitation product is combined with the higher accuracy of the gauge analyses.

7 Conclusions

Main goal of this work was to improve the quality of existing global daily satellite estimates of precipitation over land areas, where precipitation gauge observations are available. Therefore, the most widely used GPCP-1DD satellite product was merged with gauge analyses by applying kriging and co-kriging. This merged product shows a significant improvement in terms of verification scores, compared to the original multi-satellite product. For example, the verification over the entire European Union results in a rank-order correlation coefficient which increases from 0.49 (GPCP-1DD) to 0.86 (merged product). The true skill score (TSS) increases from 0.36 (GPCP-1DD) to 0.67 (merged product). Detailed verification results over Europe, East Asia

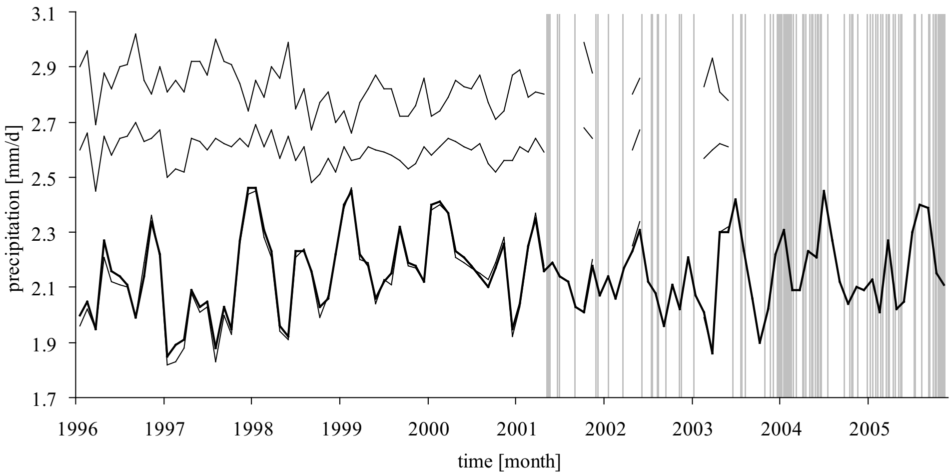

Figure 6: Time series of monthly averaged global precipitation from the GPCP-1DD product for the oceans (upper curve), total (middle curve) and land areas (lower thin curve) compared to the merged product over land (lower fat curve). Grey bars indicate days with TOVS instrument malfunctions. Monthly averaged precipitations have been calculated only for months without TOVS instrument malfunctions. Period Oct. 1996–Oct. 2005.

and Australia will be presented in part II of this paper (RUBEl and KOTTEK, in preparation).

The merged precipitation product presented here is distributed by the Core Service Bio-Geophysical Parameters of the European Union FP6 Integrated Project GEOLAND (www.gmes-geoland.info).

Acknowledgements

This work was funded by the FP6 Integrated Project GEOLAND (SIP3-CT-2003-502871). The synoptic data are provided by the Global Precipitation Climatology Centre (GPCC), operated at the German Weather Service. The GPCP-1DD multi-satellite product is provided by the Global Precipitation Climatology Project (GPCP), www1.ncdc.noaa.gov/pub/data/gpcp/1dd.

References

ADLER, R. F., G. J. HUFFMAN, A. CHANG, R. FERRARO, P.-P. XIE, J. JANOWIAK, B. RUDOLF, U. SCHNEIDER, S. CURTIS, D. BOLVIN, A. GRUBER, J. SUSSKIND, P. ARKIN, E. NELKIN, 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). – J. Hydrometeor. **4**, 1147–1167.

ATKINSON, P. M., C. D. LLOYD, 1998: Mapping precipitation in Switzerland with ordinary and indicator kriging. – J. Geogr. Inf. Decis. Anal. **2**, 65–76.

BACCHI, B., N. T. KOTTEGODA, 1995: Identification and calibration of spatial correlation patterns of rainfall. – J. Hydrol. **165**, 311–348.

BAKER, W., S. BLOOM, J. WOOLLEN, M. NESTLER, E. BRIN, T. SCHLATTER, G. BRANSTATOR, 1987: Experiments with a three-dimensional statistical objective analysis scheme usign FGGE data. – Mon. Wea. Rev. **115**, 272–296.

BERNDTSSON, R., 1987: On the use of cross-correlation analysis in studies of patterns of rainfall variability. – J. Hydrol. **93**, 113–134.

CLEMENS, M., K. BUMKE, 2002: Precipitation fields over the Baltic Sea derived from ship rain gauge measurements on merchant ships. Boreal Env. Res. **7**, 425–436.

CRESSIE, N. A. C., 1990: The origins of kriging. – Math. Geol. **22**, 239–252.

—, 1993: Statistics for Spatial Data. – John Wiley & Sons, New York, 900 pp.

DE WIT, A. J. W., H. L. BOOGAARD, C. A. VAN DIEPEN, 2005: Spatial resolution of precipitation and radiation: The effect on regional crop yield forecasts. – Agric. Forest Meteor. **135**, 156–168.

DUBRULE, O., 1983: Two methods with different objectives: Splines and kriging. – Math. Geol. **15**, 245–257.

FØRLAND, E. J., P. ALLERUP, B. DAHLSTRÖM, E. ELOMAA, T. JONSSON, H. MADSEN, J. PERÄLÄ, P. RISSANEN, H. VEDIN, F. VEJEN, 1996: Manual for Operational Correction of Nordic Precipitation Data. – Report Nr. 24/96, Norwegian Meteorological Institute, Oslo, 66 pp.

GANDIN, L. S., 1993: Optimal Averaging of Meteorological Fields. – Office Note 397, National Meteorological Center. (Available on request from Franz Rubel), 68 pp.

GOODISON, B. E., P. Y. T. LOUIE, D. YANG, 1998: WMO solid precipitation measurement intercomparison, final report. WMO/TD-No. 872, World Meteorol. Organ., Geneva, 212 pp.

GOTTSCHALCK, J., J. MENG, M. RODELL, P. HOUSER, 2005: Analysis of multiple precipitation products and preliminary assessment of their impact on global land data assimilation system land surface states. – J. Hydrometeor. **6**, 573–598.

HUFFMAN, G. J., R. F. ADLER, B. RUDOLF, U. SCHNEIDER, P. R. KEEHN, 1995: Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. – J. Climate **8**, 1285–1295.

HUFFMAN, G. J., R. F. ADLER, M. M. MORRISSEY, S. CURTIS, R. JOYCE, B. MCGAVOCK, J. SUSSKIND, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. J. Hydrometeor. **2**, 36–

50.
HUTCHINSON, M. F., 1995: Interpolating mean rainfall using thin plate smoothing splines. – Int. J. Geogr. Inf. Syst. **9**, 385–403.
ISAAKS, E. H., R. M. SRIVASTAVA, 1989: Applied Geostatistics. – Oxford Univ. Press, New York, 561 pp.
JOURNEL, A. G., S. E. RAO, 1996: Deriving conditional distributions from ordinary kriging. – Stanford Center for Reservoir Forecasting (Report No. **9**), Stanford, 25 pp.
KOTTEK, M., J. GRIESER, C. BECK, B. RUDOLF, F. RUBEL, 2006: World map of the Köppen-Geiger climate classification updated. Meteorol. Z. **15**, 259–263.
KRAJEWSKI, W. F., 1987: Cokriging radar-rainfall and rain gage data. *J. Geophys. Res.*, **92**, 9571–9580.
LEGATES, D. R., C. J. WILLMOTT, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation. – Int. J. Climatol. **10**, 111–127.
MARGULIS, S. A., D. ENTEKHABI, D. MCLAUGHLIN, 2006: Spatiotemporal disaggregation of remotely sensed precipitation for ensemble hydrologic modeling and data assimilation. – J. Hydrometeor. **7**, 511–533.
MATHERON, G., 1981: – In: Merriam, D. F. (Ed.), Down-to-Earth Statistics: Solutions Looking for Geological Problems, Ch. Splines and Kriging: Their Formal Equivalence, pp. 77–95. Syracuse University Geological Contributions, Syracuse.
MCPHEE, J., S. A. MARGULIS, 2005: Validation and error characterization of the GPCP-1DD precipitation product over the contiguous United States. – J. Hydrometeor. **6**, 441–459.
MICHELSON, D. B., 2004: Systematic correction of precipitation gauge observations using analyzed meteorological variables. – J. Hydrol. **290**, 161–177.
RENDU, J. M., 1979: Normal and lognormal estimation. – Math. Geol. **11**, 407–422.
RIVOIRARD, J., 1990: A review of lognormal estimators for in situ reserves. – Math. Geol. **22**, 213–221.
RUBEL, F., 1996a: PIDCAP – Quick Look Precipitation Atlas. – Österreichische Beitr. Meteor. Geophys. **15**, 97 pp.
—, 1996b: Scale dependent statistical precipitation analysis.– Proc. Int. Conf. on Water Resources & Environmental Research: Towards the 21st Century, 29-31 Oct. 1996, Kyoto, Japan, 317–324, http://i115srv.vu-wien.ac.at/staff/rub/pub/Kyoto.pdf.
RUBEL, F., M. HANTEL, 1999: Correction of daily rain gauge measurements in the baltic sea drainage basin. – Nord. Hydrol. **30**, 191–208.
—, —, 2001: Baltex 1/6-degree daily precipitation climatology 1996–1998. – Meteor. Atmos. Phys. **77**, 155–166.
RUBEL, F., M. KOTTEK, in preparation: Global daily precipitation fields from bias-corrected rain gauge analyses and satellite observations. Part II: Verification results. – Meteorol. Z.
RUDOLF, B., F. RUBEL, 2005: Global Precipitation. – In: HANTEL, M. (Ed.): Observed Global Climate. New Series on Landolt-Börnstein, Numerical Data and Functional Relationships, Springer, Berlin, 11.1–11.24.
SEO, D.-J., W. F. KRAJEWSKI, D. S. BOWLES, 1990a: Stochastic interpolation of rainfall data from rain gages and radar using cokriging, 1. Design of experiments. – Water Resour. Res. **26**, 469–477.
—, —, —, 1990b: Stochastic interpolation of rainfall data from rain gages and radar using cokriging, 2. Results. – Water Resour. Res. **26**, 915–924.
SEVRUK, B., 1982: Method of correction for systematic error in point precipitation measurement for operational use. – WMO-589, Geneva, World Meteorological Organization, 91 pp.
SHEFFIELD, J., G. GOTETI, E. F.WOOD, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. – J. Climate **19**, 3088–3111.
UNGERSBÖCK, M., F. RUBEL, T. FUCHS, B. RUDOLF, 2001: Bias correction of global daily rain gauge measurements. – Phys. Chem. Earth (B) **26**, 411–414.
XIA, Y., 2006: Optimization and uncertainty estimates of WMO regression models for the systematic bias adjustment of NLDAS precipitation in the United States. – J. Geophys. Res. **111**, D08102, DOI: 10.1029/2005JD006188.
XIE, P., P. A. ARKIN, 1997: Global Precipitation: A 17-year monthly analysis based on gauge observations, satellite estimations and numerical model outputs. – Bull. Amer. Meteor. Soc. **78**, 2536–2558.
YATES, S. R., A. W. WARRICK, D. E. MYERS, 1986: Disjunctive kriging 1. overview of estimation and conditional probability. – Water Resour. Res. **22**, 615–621.
YE, B., D. YANG, Y. DING, T. HAN, T. KOIKE, 2004: – A bias-corrected precipitation climatology for China. – J. Hydrometeor. **5**, 1147–1160.