Technical Note: Calibration and validation of geophysical observation models

Salama, M.S.; van de Velde, R.J.; van der Woerd, H.J.; Kromkamp, J.C.; Philipppart, C.J.M.; Joseph, A.T.; O'Neill, P.E.; Lang, R.H.; Gish, T.; Werdell, P.J.; Su, Z.

published in
Biogeosciences
2012

DOI (link to publisher)
10.5194/bg-9-2195-2012

document version
Publisher's PDF, also known as Version of record

Biogeosciences, 9, 2195–2201, 2012
www.biogeosciences.net/9/2195/2012/
doi:10.5194/bg-9-2195-2012

Biogeosciences

Technical Note: Calibration and validation of geophysical observation models

M. S. Salama[1,2], R. Van der Velde[1], H. J. van der Woerd[3], J. C. Kromkamp[2], C. J. M. Philippart[4], A. T. Joseph[5], P. E. O'Neill[5], R. H. Lang[6], T. Gish[7], P. J. Werdell[8,9], and Z. Su[1]

[1]Department of Water Resources, ITC, University of Twente, The Netherlands
[2]Royal Netherlands Institute for Sea Research (NIOZ), Yerseke, The Netherlands
[3]Institute for Environmental Studies (IVM), VU University, The Netherlands
[4]Royal Netherlands Institute for Sea Research (NIOZ), The Netherlands
[5]Hydrological Sciences Branch/614.3, Hydrospheric and Biospheric Sciences Laboratory, NASA/GSFC, USA
[6]Dept. of Electrical and Computer Engineering, George Washington University, USA
[7]Hydrology and Remote Sensing Lab, USDA-ARS, USA
[8]NASA Goddard Space Flight Center, Ocean Ecology Branch Code 614.2, USA
[9]Science Systems and Applications, Inc., Lanham, USA

Correspondence to: M. S. Salama (salama@itc.nl)

Received: 3 November 2011 – Published in Biogeosciences Discuss.: 11 January 2012
Revised: 16 May 2012 – Accepted: 21 May 2012 – Published: 18 June 2012

Abstract. We present a method to calibrate and validate observational models that interrelate remotely sensed energy fluxes to geophysical variables of land and water surfaces. Coincident sets of remote sensing observation of visible and microwave radiations and geophysical data are assembled and subdivided into calibration (Cal) and validation (Val) data sets. Each Cal/Val pair is used to derive the coefficients (from the Cal set) and the accuracy (from the Val set) of the observation model. Combining the results from all Cal/Val pairs provides probability distributions of the model coefficients and model errors. The method is generic and demonstrated using comprehensive matchup sets from two very different disciplines: soil moisture and water quality. The results demonstrate that the method provides robust model coefficients and quantitative measure of the model uncertainty. This approach can be adopted for the calibration/validation of satellite products of land and water surfaces, and the resulting uncertainty can be used as input to data assimilation schemes.

1 Introduction

Observation models are widely used for estimating geophysical variables of land and water surfaces from remote sensing data. The simplest form is the empirical linear model, whereby coefficients are derived from regressing measured geophysical variables with observed radiation. In most cases, these empirical models have some physical meaning and are often used because of their simplicity. Examples of land remote sensing applications are available from active/passive microwave remote sensing of soil moisture (e.g. Njoku et al., 2002). Similarly, water quality applications make use of the Lambert-Beer law to model the spectral absorption of light by suspended and dissolved materials as a linear function of their concentrations (D'Sa and Miller, 2005; Robinson, 2004; Salama et al., 2004). Currently, such strategies are proposed for NASA's Soil Moisture Active Passive (SMAP) mission combined radar/radiometer soil moisture product (Entekhabi et al., 2010), the Netherlands' automated monitoring network (IN PLACE: Integrated Network for Production and Loss Assessment in the Coastal Environment), and the NASA Moderate Resolution Imaging Spectroradiometer (MODIS) mission ocean colour products (McClain et al., 2004). This type of model is developed

from comprehensive sets of concurrent remote sensing observations and field measurements, hereafter referred to as matchups. Ideally, the validity of any model is tested against an independent data set. Therefore, the available matchups are often subdivided into independent sets used for derivation of the model coefficients (calibration) and for accuracy assessment (validation). Most studies subdivide matchups into so-called calibration/validation (Cal/Val) sets based on a statistical or regional resemblance (Devlin et al., 2008), but without a clear directive on its effect for model accuracy. This is most likely the case because there has been until now no objective approach for subdividing Cal/Val sets. Many combinations of matchups can be used, specifically when using a large number of points. Each Cal/Val pair has the same probability of occurrence, but provides different results. As such, the selection procedure not only impacts the model's accuracy, but also the accuracy assessment. On the other hand, the selection of Cal/Val pairs can also be thought of as a stochastic sampling from a known probability distribution (e.g. Wang et al., 2005; Salama and Stein, 2009). Such stochastic treatment of matchups within the Cal/Val context has not yet been investigated in the field of Earth observation, but has the advantage of providing a quantitative uncertainty measure for both the model coefficients and derived geophysical variables.

In this paper we follow a stochastic approach for selecting Cal/Val sets and demonstrate its use for quantifying uncertainty. The proposed approach combines the bootstrapping method of Efron and Tibshirani (1993) with the Jackknife technique (which leaves out one, or more, observation) and adapts the sample size at each iteration. Bootstrapping and Jackknife methods are usually used to provide the standard error of the derived "plug in" estimates (Efron and Tibshirani, 1993) and have been employed for validating observation models (e.g. Petus et al., 2010; Melin, 2010; Salama and Su, 2010). However, the combination of bootstrapping without replacement with Jackknife sampling and changing the sample size at each iteration is novel and provides not only the accuracy of regressed estimates, but also the underlying probability distribution of regressed estimates and their errors.

The developed method samples from a complete matchup set to populate many sets of Cal/Val pairs. Each pair is used to derive the model coefficients and their associated errors, from which the probability distributions of the calibration and validation result is determined. In this paper the method is demonstrated for two data sets: (i) L-band (1.6 GHz) backscatter (σ°) – soil moisture matchups collected during the 2002 OPE^3 (Optimizing Production Inputs for Economic and Environmental Enhancement) campaign (Joseph et al., 2010a,b), and (ii) matchups of chlorophyll *a* concentrations and derived absorption coefficients obtained from the NASA bio-Optical Marine Algorithm Data (NOMAD, version 2a.) (Werdell and Bailey, 2005).

2 Data sets

2.1 Land application – soil moisture

The 2002 OPE^3 campaign focused on the active and passive microwave remote sensing of soil moisture throughout the corn growth cycle. Part of the field activities consisted of weekly C- (4.75 GHz) and L- (1.6 GHz) band σ° measurements with the NASA/George Washington University (GWU) truck-mounted scatterometer. Further in support of these remote sensing observations, an extensive ground sampling was conducted that included soil moisture. Full details on the data sets collected during the field campaign can be found in Joseph et al. (2010a,b). Here, we only make use of the 75 matchups between the L-band HH polarized σ° observed from a 35° view angle and the measured soil moisture, hereafter referred to as the OPE^3 matchups. The σ° observations are corrected for vegetation effects through application of method described in Joseph et al. (2008), which results in the σ° representative for a bare soil surface. Many studies (e.g. Ulaby et al., 1984; Champaign and Faivre, 1997; Njoku et al., 2002) have demonstrated the following linear relationship between soil moisture and σ° observed under the same land cover conditions:

$$\mathrm{sm} = a\sigma^\circ + b, \tag{1}$$

where sm is the soil moisture content ($\mathrm{m^3\,m^{-3}}$), a is the slope ($\mathrm{m^3\,m^{-3}\,dB^{-1}}$) representing the σ° sensitivity to soil moisture, and b is the offset ($\mathrm{m^3\,m^{-3}}$) accounting for the baseline effects, such as surface roughness, topography, and land cover. Both the σ° sensitivity to soil moisture and the baseline effects depend on the sensing configuration (e.g. wavelength, polarization, view angle) as well as the land surface (e.g. surface roughness, land cover, topography).

2.2 Water application – chlorophyll *a* absorption

The NASA bio-Optical Marine Algorithm Data (NOMAD, version 2a.) set includes measurements of spectral remote-sensing reflectances, spectral marine absorption and backscattering coefficients, and concentrations of water constituents (Werdell and Bailey, 2005). Here, we use only chlorophyll *a* (chl *a*) measurements derived from high performance liquid chromatography (HPLC). The observed radiance spectra and matching HPLC-derived chl *a* concentration consist of 424 matches, hereafter referred to as the NOMAD matchups. The general practice is to derive the absorption coefficients from the observed radiance spectra using semi-analytical inversion models (e.g. Van Der Woerd and Pasterkamp, 2008; Maritorena et al., 2002). Lambert-Beer law is then employed to estimate the absorption per unit mass from derived absorption coefficients and measured concentrations.

The chl *a* absorption coefficients at the blue band ($\lambda_0 =$ 440 nm) are derived from the observed radiances using the

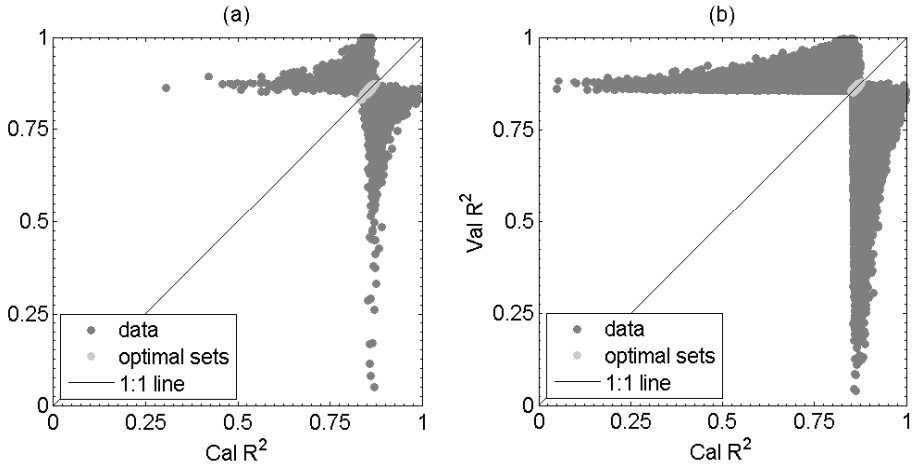

Fig. 1. Determination coefficient, R^2, between measured and observed values of **(a)** soil moisture and **(b)** chl *a* absorption coefficient. The solid line is the 1 : 1 reference line. Light-grey coloured points represent the optimal Cal/Val pairs.

cross entropy method as reported in Salama and Shen (2010). Following the Lambert-Beer law, the absorption coefficient of chl *a* is described as a linear function of the concentration (D'Sa and Miller, 2005, Eq. 10):

$$a_{\text{chl }a}(\lambda_0) = a^*_{\text{chl}\,a}(\lambda_0) C_{\text{chl }a} + \delta(\lambda_0) \quad (2)$$

where $a_{\text{chl}\,a}(\lambda_0)$ is the absorption coefficient of chl *a* (m^{-1}) at the wavelength λ_0 (nm); $a^*_{\text{chl}\,a}(\lambda_0)$ is the specific absorption coefficient that describes the absorption per unit weight ($\text{m}^2\,\text{mg}^{-1}$); $C_{\text{chl}\,a}$ is the concentration ($\text{mg}\,\text{m}^{-3}$); $\delta(\lambda_0)$ is an offset related to sensor noise, retrieval error of $a_{\text{chl}\,a}(\lambda_0)$, ($\text{m}^{-1}$) and the ratio of accessory pigments that are produced in different conditions of growth (nutrients and irradiance), e.g. "xanthine" that acts as sun protection.

The two unknowns in Eq. (2), $a^*_{\text{chl}\,a}(\lambda_0)$ and $\delta(\lambda_0)$, are estimated from regressing $a_{\text{chl}\,a}(\lambda_0)$ versus $C_{\text{chl}\,a}$ using linear-regression model. In practice, Eq. (2) could deviate from linearity depending on the packaging effect, cell sizes, physiology and species composition of the phytoplankton community (Bricaud et al., 1995). For example, the effect of packaging on the variability of $a^*_{\text{chl}\,a}(\lambda_0)$ is smaller in open oligotrophic oceans than in upwelling regions or coastal areas where larger phytoplankton cells are abundant. Hence, the deviation of Eq. (2) from linearity can then be understood based on the water body investigated. The linearity of Eq. (2) for the used data sets is justified in Sect. 4, Fig. 1b.

3 Method

The method randomly subdivides the data into many sets (or Jackknife samples) of Cal/Val pairs. The Cal set is used to derive the coefficients of the observation model, whereas the Val set is employed to check the accuracy of the model. The results are probability distributions of model coefficients and their prediction uncertainties.

The Cal/Val sets are derived from the n available matchups following two rules: (i) both Cal and Val sets must contain at least 7 samples ($k_{\min} = 7$), and (ii) each sample is used once, either for calibration or for validation (i.e. sampling without replacement). The minimum sample size, ($k_{\min} = 7$), is selected according to the method of Cohen et al. (2003), to achieve about 35 % error in the derived slope at 95 % of confidence. This value, 35 %, corresponds to the desired level of accuracy for satellite-derived Chl *a* products (McClain et al., 2006; Bailey and Werdell, 2006).

The number of Cal/Val pairs is computed as $nr = n - 2k_{\min} + 1$. Now, for each $i = [k_{\min}, n - k_{\min}]$, the method forms a Cal/Val pair by increasing/decreasing the number of data points in the sets (forming the Jackknife sample). The number of possible combinations, npc_i, for the i-th Cal/Val pair is

$$\text{npc}_i = \binom{n}{k_i} = \frac{n!}{k_i!\,(n-k_i)!} \quad (3)$$

where n is the total number of data points, k_i is the number of samples in the Cal or Val set during the i-th iteration. For data sets with $n > 20$ (holds for both the OPE[3] and NOMAD matchups), the number of possible combinations (npc_i) is large (e.g. 1.9848 E9 for 75 over 7 in OPE[3]), and therefore npc_i is reduced to the number of used combinations, nuc_i, by bootstrapping $\text{nuc}_i = 10 \log \text{npc}_i$ combinations from npc_i. In principle, each combination nuc_i has the same probability of occurrence; therefore, the uniform distribution is used to select nuc_i unique combinations from npc_i (bootstrap method of Efron and Tibshirani, 1993). Each formed Cal/Val set is used for the calibration and the subsequent validation of the empirical model. The validation is always performed using type-II model (Bevington and Robinson, 2003), while the calibration depends on the model, e.g. for linear model we use the type-I regression. The accuracy of the empirical model is assessed using two statistical measures: (i) the

mean absolute error between derived and measured values (MAE) and (ii) the determination coefficient (R^2). The algorithm produces three probability distributions (PDs): two for the calibration coefficients, PD_c, and one for the accuracy measure, PD_v. The above method is implemented in the following model, called GeoCalVal:

1. Take k_i samples for the Cal set and $n-k_i$ for the Val set;
2. Compute npc_i from Eq. 3 and $nuc_i = 10 \log npc_i$;
3. Use the uniform distribution to generate nuc_i unique combinations of Cal sets and their complements for Val sets;
4. Compute model coefficients from the calibration set and store them in PD_c;
5. Use the new model coefficient to estimate the geophysical variables from the Val set;
6. Compute the uncertainty of step 5 and store them in PD_v;
7. Increase k_i by one and repeat steps 1 to 7.

4 Results and discussions

4.1 Optimal Cal/Val pairs

The determination coefficients, R^2, of the Cal set are plotted against those of the Val set in Fig. 1 for all possible combinations. The data point position with respect to the x-axis is an indication for the ability of the model to fit the matchups of the Cal set, whereas its position with respect to the y-axis represents the model's performance in deriving the geophysical variables, here soil moisture and Chl *a* absorption coefficient ($a_{\mathrm{chl}\,a}(\lambda_0)$). Obviously, both the Cal and Val R^2 depend on the number of data points, reaching their maxima when all data points are included, which suggests for the Cal sets that the used observation models in Eqs. (1) and (2) are indeed linear.

Both OPE^3 and NOMAD matchups produce a narrow region of Cal/Val pairs, for which the calibration R^2 is similar to validation R^2, about 0.75–0.85 (light-grey coloured data points in Fig. 1). In other words, within these Cal/Val subdivisions the model validity and the accuracy assessment are balanced. This region defines the optimal setups for subdividing matchups into Cal/Val sets. The underlying mechanisms of the data points in this region are investigated further. We found that the optimal Cal/Val sets are obtained when the arithmetic mean, μ, and dispersion, σ, of each set are equal to those of the original data set. As such, the optimal Cal/Val pair satisfies the following condition:

$$\begin{aligned} &\mu_{\mathrm{cal}} = \mu_{\mathrm{val}} = \mu_{\mathrm{data}}, \\ \mathrm{AND} \quad &\\ &\sigma_{\mathrm{cal}} = \sigma_{\mathrm{val}} = \sigma_{\mathrm{data}}, \end{aligned} \tag{4}$$

Table 1. Estimated parameters of the best fit *t*-location-scale distribution to model coefficients and MAE uncertainties. The degree of fit is expressed in standard error.

OPE^3 matchups			
	μ	σ	ν
slope [$m^3\,m^{-3}\,dB^{-1}$]	34.0870	1.3729	2.7342
standard error	0.0160	0.0186	0.0882
intercept [$m^3\,m^{-3}$]	−26.5426	0.2137	2.7668
standard error	0.0025	0.0029	0.0887
MAE [$m^3\,m^{-3}$]	0.0244	0.0025	3.3380
standard error	3E−05	3E−05	0.1128
NOMAD matchups			
	μ	σ	ν
slope [$m^2\,mg^{-1}$]	0.0304	0.0013	1.9722
standard error	2.6E−6	3E−6	0.0085
intercept [m^{-1}]	0.0195642	0.00128274	2.3505
standard error	2.6E−06	3E−06	0.0114
MAE [$mg\,m^{-3}$]	0.6043	0.0534	3.4957
standard error	0.0001	0.0001	0.02012

where the subscripts "cal", "val" and "data" are for the calibration, validation and original data sets, respectively.

Equation (4) and the equal R^2 for both Cal and Val sets (light-grey coloured points in Fig. 1) are the criteria that should be used to determine the sample size of the optimal Cal/Val set.

4.2 The underlying distribution

Figure 2 shows the derived probability distributions (PDs) of model coefficients, PD_c, and the associated uncertainties, PD_v, for the two matchup sets, OPE^3 and NOMAD. The resulting PDs of model coefficients have high kurtosis (acute peak around the mean) values and flat tails, i.e. more prone to outliers. Different values of $k_{\min}$ were tested (not shown), and the results show that all derived PDs from both data sets (OPE^3 and NOMAD matchups) can be described by the *t*-location-scale probability distribution (the black lines in Fig. 2) of the form (Evans et al., 1993)

$$f = \frac{\Gamma(0.5\nu + 0.5)}{\sigma\sqrt{\nu\pi}\,\Gamma(0.5\nu)} \left[1 + \nu^{-1}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-(0.5\nu+0.5)} \tag{5}$$

where μ, σ and ν are the mean, standard deviation and shape factor (or the degree of freedom), respectively. The gamma function Γ is equivalent to the factorial function $n!$ extended to non-integral arguments. The distribution in Eq. (5) means that the standard variates of the data points follow the Student *t* distribution. The function in Eq. (5) is fitted to the distributions of derived model coefficients and MAEs by varying the parameters μ, σ and ν, which are listed in Table 1 with their standard errors.

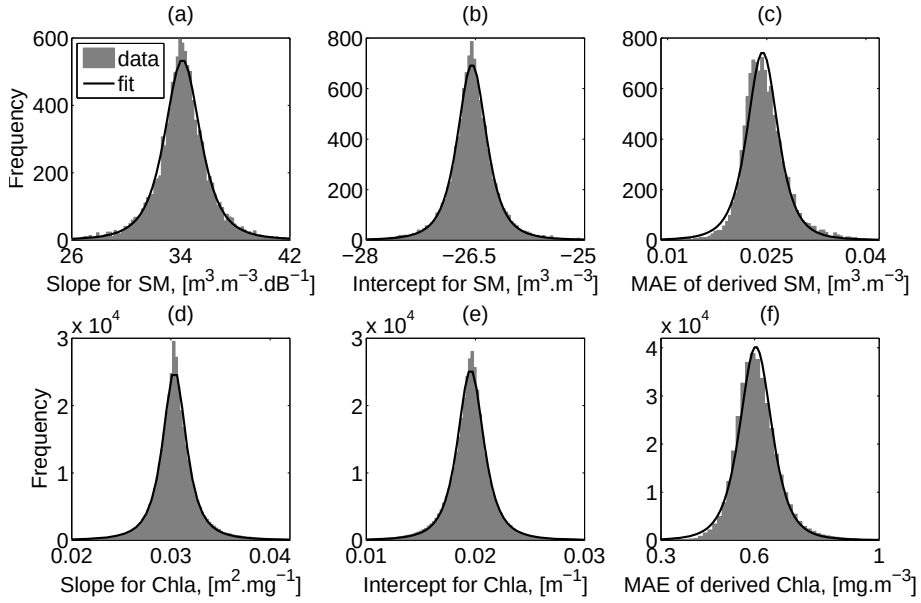

Fig. 2. Derived probability distributions of model coefficients **(a, b, d, e)** and associated uncertainties **(c, f)** for the OPE^3 data (upper panels) and NOMAD matchups (lower panels). The solid lines are the fits by Eq. (5) with coefficients given in Table 1.

The reason for having flat tails in the PDs of Fig. 2 is due to the fact that the accuracy of model coefficients depends on the size of the Cal set. In other words, for a large Cal set we expect to have higher accuracy as most data points are used; however, this makes them also sensitive to outliers in the Val set, because most of the data points have been used to create the Cal set. For a linear observation model, the *t*-probability density function should, thus, be employed to describe the distributions PD_c and PD_v, regardless of the original distribution of geophysical measurements or remote sensing observations. For example, the NOMAD matchups set has a log-normal distribution, while OPE^3 is close to uniform distribution (not shown here) for measurements, residuals and observations. Yet, the distribution of derived coefficients follows, for both data sets, Eq. (5). This is basically a confirmation of previous statistical studies, for example Singh (1988) showed that the normality distribution is not always a valid assumption for linear models, and the *t*-distribution is broader and therefore better suited. In this regard, having the result of our sampling scheme reproducing the *t*-distribution is another validation of the correctness of the GeoCalVal. The proposed method reveals the shape of the underlying probability distribution without any a priori assumption on its parameters (e.g. degree of freedom). For non-linear models there is no straightforward theoretical approximation of the expected probability distribution. If we would follow the theory, we would have no means to justify our assumption on the underlying probability distribution and its parameters. The only objective approach is by evaluating all possible combination sets as is proposed through the GeoCalVal method.

4.3 Effect of sample size

Fixing the number of sampling points will result in PD with lower kurtosis, i.e. the PD will be less peaked. That means adapting the sample size will increase the accuracy of the derived parameters (slope and intercept in this case), as the dispersion will also be reduced. The importance of adapting the size of the sample is related to the common practice in calibration and validation of Earth observation products. Here, we search for the optimal division (thus, sample size) of the Cal/Val sets, such that the Cal set produces EO-model coefficients that enable generating EO products (estimated from the Val set) with an accuracy satisfying the mission requirements. Hence, one of the statistical questions addressed within our manuscript is what are the criteria to define the optimal sample size needed for calibrating observation models so that it produces EO products within the designed mission accuracy and within the accuracy of the calibration itself? For example, we can condition the iterative scheme in GeoCalVal to stop when the criteria defined in Sect. 4.1 are met (these are, Eq. 4 and coloured points in Fig. 1). This will however be at the cost of losing information on the probability distributions of regression coefficients, their errors and the shape of the underlying distributions. This information can only be derived if we change the sample size and study its effect on the accuracy of calibration and validation (as shown in Fig. 1). Thus, GeoCalVal makes use of the proposed sampling scheme, because only through this approach we can, in an objective manner, identify the probability distributions of coefficients and associated uncertainties of observation

models through optimal divisions of the data set into Cal and Val pairs.

4.4 Application of the GeoCalVal model

The detailed knowledge on the PDs of uncertainties and uncertainty sources embedded within the remote sensed geophysical variable (shown in Fig. 2) can be used as input for data assimilation schemes (Reichle, 2008). On the other hand, these PDs can also be employed to derive the probability distribution of uncertainty within the remote sensing observations itself, i.e. one PD per observation. The relationship between measurements and observations is described by a model of the form, $\mathbf{Y} = f(\Phi, \mathbf{X})$, in which Φ is the set of n model coefficients, $\Phi = [\phi_1, \phi_2 ... \phi_n]$, $\mathbf{X}$ is the set of m geophysical measurements (with $m > n$) and $\mathbf{Y}$ is the corresponding remote sensing observations. Assuming that the fluctuations in the measured quantity, $\mathbf{X}$, and derived model coefficients, Φ, are uncorrelated, we approximate the second moment using the truncated Taylor series expansion:

$$\sigma_{\mathrm{y}}^2 = w_{\mathrm{x}}^2 \sigma_{\mathrm{x}}^2 + \sum_{i=1}^{n} w_{\phi_i}^2 \sigma_{\phi_i}^2 \quad (6)$$

where w is the partial derivative of $\mathbf{Y}$ with respect to the measurements $\mathbf{X}$ and each model coefficient, ϕ_i. The terms, σ^2, are the corresponding variances. For the linear model $Y = a \times X + b$, the uncertainty in Eq. (6) becomes $\sigma_{\mathrm{y}}^2 = a^2 \sigma_{\mathrm{x}}^2 + x^2 \sigma_{\mathrm{a}}^2 + \sigma_{\mathrm{b}}^2$. The coefficient a and the uncertainties terms σ_{a}^2 and σ_{b}^2 are quantified from the derived probability distributions of model coefficients, $\mathrm{PD_c}$. Measurement uncertainty, σ_{x}^2, is either assumed (e.g. NOMAD matchups) or estimated from available measurements (e.g. OPE^3). In the NOMAD data set, the concentrations of Chl *a* were estimated using high-performance liquid chromatography (HPLC) method. Many studies (Claustre et al., 2004; Hooker et al., 2005) found that the error in HPLC estimation of Chl *a*, on average, varies between 7 % and 25 %. On the other hand, each observation site in the OPE^3 data set contains 21 soil moisture measurements. The standard deviation of these measurements, per observation, can be used as a proxy for σ_{x}. Estimated values of σ_{x}, a, σ_{a} and σ_{b} form the inputs to Eq. (6) to produce the PD, quantifying the uncertainty of each remote sensing observation. This results in a PD of uncertainty per data point that has $\sum_i^{nr} \mathrm{nuc}_i$ number of samples, i.e. number of all used combinations. It should, however, be noted that this uncertainty should not be confused with observation errors associated with remote sensing retrievals, which included also other components (e.g. model goodness-of-fit and inversion uncertainty).

5 Conclusions

In this paper we present the GeoCalVal model for an objective selection of calibration/validation data sets to assess the performance of observation model for geophysical variables. GeoCalVal combines two traditional re-sampling methods (bootstrapping and Jackknife) and adapts the sample size at each iteration. This combination of bootstrapping with Jackknife sampling and changing the sample size at each test iteration is novel and provides not only the accuracy of regressed estimates and associated errors but also their underlaying distributions. The GeoCalVal tests all probable combinations of Cal/Val setup and considers the effect of changing the sample size on the accuracy of regressed estimates. The end results are probability distributions of model coefficients (calibration) and uncertainties in the estimates (validation).

GeoCalVal is applied to two matchups sets, which shows that

- GeoCalVal provides an optimal setup for subdividing matchups into Cal/Val sets;
- the coefficients and associated uncertainties of linear observation models follow the *t*-location scale distribution, i.e. the distribution of their standard variate follows the Student *t* distribution;
- the derived PDs provide complete information on the variations of model coefficients, their uncertainties and the accuracy of observations, which can be employed in time series analyses and data assimilation schemes;
- the optimal Cal/Val sets are obtained when the arithmetic mean and dispersion of the Cal/Val sets are equal to those of the original data set;
- the presented method is applicable to any data set and can be adjusted to any observation model regardless of the application area, e.g. water quality or surface hydrology.

Acknowledgements. The authors would like to thank NASA Ocean Biology Processing Group and individual data contributors for maintaining and updating the SeaBASS database. This work was supported in part by the Earth and Life Sciences (ALW) division of the Netherlands Organization for Scientific Research (NWO), in the framework of the *National Programme Sea and Coast Research – Changing Carrying Capacity*, the IN PLACE project.

Edited by: S. W. A. Naqvi

References

Bailey, S. and Werdell, P. J.: A multi-sensor approach for the on-orbit validation of ocean color satellite data products, Remote Sens. Environ., 102, 12–23, 2006.

Bevington, P. and Robinson, K.: Data Reduction and Error Analysis for the Physical Sciences, 3rd edn., McGraw-Hill, New York, 2003.

Bricaud, A., Babin, M., Morel, A., and Claustre, H.: Variability in the Chlorophyll-specific absorption coefficients of naturnal phytoplankton: analysis and parameterization, J. Geophys. Res., 100, 13321–13332, 1995.

Champaign, I. and Faivre, R.: Sensitivity of the radar signal to soil moisture: variation with incidence angle, frequency, and polarization, IEEE T. Geosci. Remote, 35, 781–783, 1997.

Claustre, H., Hooker, S., Van Heukelem, L., Berthon, J., Barlow, R., Ras, J., Sessions, H., Targa, C., Thomas, C. S., Van der Linde, D., and Marty, J.: An intercomparison of HPLC phytoplankton pigment methods using in situ samples: application to remote sensing and database activities, Mar. Chem., 85, 41–61, 2004.

Cohen, J., Cohen, P., West, S. G., and Aiken, L. S.: Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 3rd edn., Lawrence Earlbaum Associates, Mahwah, NJ, 2003.

Devlin, M., Barry, J., Mills, D., Gowenc, R., Foden, J., Sivyerb, D., and Tett, P.: Relationships between suspended particulate material, light attenuation and Secchi depth in UK marine waters, Estuar. Coast. Shelf S., 79, 429–439, 2008.

D'Sa, E. and Miller, R.: Bio-optical properties of coastalwaters, in: Remote Sensing of Coastal Aquatic Environments, edited by: Miller, R., del Castillo, C., and McKee, B., Springer, Dordrecht, The Netherlands, 129–155, 2005.

Efron, B. and Tibshirani, R.: An introduction to the Bootstrap, 57, Monographs on Statistice and Applied Probability, Chapman & Hall/CRC, 1993.

Entekhabi, D., Njoku, E., O'Neill, P., Kellogg, K., Crow, W., Edelstein, W., Entin, J., Goodman, S., Jackson, T., Johnson, J., Kimball, J., Piepmeier, J., Koster, R., Martin, N., McDonald, K., Moghaddam, M., Moran, S., Reichle, R., Shi, J., Spencer, M., Thurman, S., Tsang, L., and Van Zyl, J.: The Soil Moisture Active Passive (SMAP) mission, Proc. IEEE, 98, 704–716, 2010.

Evans, M., Hastings, N., and Peacock, B.: Statistical Distributions, 2nd Edn., Wiley, New York, 1993.

Hooker, S. B., Van Heukelem, L., Thomas, C. S., Clauster, H., Ras, J., Barlow, R., Sessions, H., Shlüter, L., Perl, J., Trees, C., V., S., Head, E., Clementson, L., Fishwick, J., Llewellyn, C., and Aiken, J.: The second SeaWiFS HPLC analysis Round-Robin experiment (SeaHARRE-2), Tech. Rep. 212787, Goddard Space Flight Center, NASA, Greenbelt, Maryland 20771, 2005.

Joseph, A. T., Van der Velde, R., O'Neill, P. E., Lang, R. H., and Gish, T.: Soil moisture retrieval during a corn growth cycle using L-band (1.6 GHz) radar observations, IEEE T. Geosci. Remote, 46, 2365–237, 2008.

Joseph, A., van der Velde, R., O'Neill, P. E., Choudhury, B. J., Lang, R. H., Kim, E. J., and Gish, T.: L-band brightness temperature observations over a corn canopy during the entire growth cycle, Sensors, 10, 6980–7001, 2010a.

Joseph, A., van der Velde, R., O'Neill, P. E., and Lang, R. H.: Effects of corn on C- and L-band radar backscatter: a correction method for soil moisture retrieval, Remote Sens. Environ., 114, 2417–2430, 2010b.

Maritorena, S., Siegel, D., and Peterson, A.: Optimization of a semianalytical ocean color model for global-scale applications, Appl. Optics, 41, 2705–2714, 2002.

McClain, C., Feldman, G., and Hooker, S. B.: An overview of the SeaWiFS project and strategies for producing a climate research quality global ocean bio-optical time-series, Deep-Sea Res. Pt. II, 51, 5–42, 2004.

McClain, C. R., Feldman, G. C., Hooker, S. B., and Bontempi, P.: Satellite data for ocean biology, biogeochemistry, and climate research, EOS Transactions, 87, 337–343, 2006.

Mélin, F.: Global distribution of the random uncertainty associated with satellite-derived chla, IEEE Geoscience and Remote Sensing Letters, 7, 220–224, 2010.

Njoku, E., Wilson, W., Yueh, S., Dinardo, S., Li, F., Jackson, T., Lakshmi, V., and Bolten, J.: Observations of soil moisture using a passive and active low-frequency microwave airborne sensor during SGP99, IEEE T. Geosci. Remote, 40, 2659–2673, 2002.

Petusa, C., Chustb, G., Gohinc, F., Doxarand, D., Froidefonda, J. M., and Sagarminagab, Y.: Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery, Cont. Shelf Res., 30, 379–392, 2010.

Reichle, R. H.: Data assimilation methods in the Earth sciences, Adv. Water Resour., 31, 1411–1418, 2008.

Robinson, I.: Measuring The Ocean From Space: The Principle And Methods Of Satellite Oceanography, Chapter 6: Ocean color remote sensing, Springer Praxis Books, Berlin, Heidelberg, New York, 2004.

Salama, M. S. and Shen, F.: Stochastic inversion of ocean color data using the cross-entropy method, Opt. Express, 18, 479–499, 2010.

Salama, M. S. and Stein, A.: Error decomposition and estimation of inherent optical properties, Appl. Optics, 48, 4947–4962, 2009.

Salama, M. S. and Su, Z.: Bayesian model for matching the radiometric measurements of aerospace and field ocean color sensors, Sensors, 10, 7561–7575, 2010.

Salama, M. S., Monbaliu, J., and Coppin, P.: Atmospheric correction of Advanced Very High Resolution Radiometer imagery, Int. J. Remote Sens., 25, 1349–1355, 2004.

Singh, R. S.: Estimation of error variance in linear regression models with errors having multivariate student-t distribution with unknown degrees of freedom, Economic Letters, 27, 47–53, 1988.

Ulaby, F. T., Allen, C. T., Eger III, G., and Kanemasu, E.: Relating the microwave backscattering coefficient to leaf area index, Remote Sens. Environ., 14, 113–133, 1984.

Van Der Woerd, H. and Pasterkamp, R.: Hydropt: a fast and flexible method to retrieve chlorophyll-*a* from multispectral satellite observations of optically complex coastal waters, Remote Sens. Environ., 112, 1795–1807, 2008.

Wang, P., Boss, E., and Roesler, C.: Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color, Appl. Optics, 44, 4074–4084, 2005.

Werdell, J. and Bailey, S.: An improved in-situ bio-optical data set for ocean color algorithm development and satellite data product validation, Remote Sens. Environ., 98, 122–140, 2005.