



Exploring Machine Learning to Correct Satellite-Derived Sea Surface Temperatures

Stéphane Saux-Picart, Pierre Tandeo, Emmanuelle Autret, Blandine Gausset

► To cite this version:

Stéphane Saux-Picart, Pierre Tandeo, Emmanuelle Autret, Blandine Gausset. Exploring Machine Learning to Correct Satellite-Derived Sea Surface Temperatures. *Remote Sensing*, 2018, 10 (2), pp.224 -. 10.3390/rs10020224 . hal-01759072

HAL Id: hal-01759072

<https://hal.science/hal-01759072>

Submitted on 5 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Exploring Machine Learning to Correct Satellite-Derived Sea Surface Temperatures

Stéphane Saux Picart ^{1,*}, Pierre Tandeo ², Emmanuelle Autret ³ and Blandine Gausset ¹ 

¹ Météo-France/Centre de Météorologie Spatiale, Avenue de Lorraine, B.P. 50747, 22307 Lannion CEDEX, France; blandine.gausset@gmail.com

² IMT Atlantique, Lab-STICC, UBL, 29238 Brest, France; pierre.tandeo@imt-atlantique.fr

³ Ifremer, Laboratoire d'Océanographie Physique et Spatiale, ZI Pointe du Diable CS 10070, 29280 Plouzané, France; emmanuelle.autret@ifremer.fr

* Correspondence: stephane.sauxpicart@meteo.fr; Tel.: +33-296-05-67-07

Received: 20 November 2017; Accepted: 27 January 2018; Published: 1 February 2018

Abstract: Machine learning techniques are attractive tools to establish statistical models with a high degree of non linearity. They require a large amount of data to be trained and are therefore particularly suited to analysing remote sensing data. This work is an attempt at using advanced statistical methods of machine learning to predict the bias between Sea Surface Temperature (SST) derived from infrared remote sensing and ground “truth” from drifting buoy measurements. A large dataset of collocation between satellite SST and in situ SST is explored. Four regression models are used: Simple multi-linear regression, Least Square Shrinkage and Selection Operator (LASSO), Generalised Additive Model (GAM) and random forest. In the case of geostationary satellites for which a large number of collocations is available, results show that the random forest model is the best model to predict the systematic errors and it is computationally fast, making it a good candidate for operational processing. It is able to explain nearly 31% of the total variance of the bias (in comparison to about 24% for the multi-linear regression model).

Keywords: machine learning; systematic error; sea surface temperature; random forest

1. Introduction

Characterising the error associated with data, from observations or model outputs, is essential for correct use and analysis. It is, however, often a very complex problem requiring many assumptions. When large amounts of data are available, data-driven methods provide a convenient way to work around that complexity, especially for remote sensing data. Recently in this domain, authors proposed to use relevant machine learning methods (see, e.g., [1] or [2]). These methods are based on statistical models and are able to automatically solve lots of regression and/or classification problems. Here, we focus on a regression problem and we test various classical methods using linear and nonlinear assumptions.

Sea Surface Temperature (SST) is a variable which has been estimated over a long time period from infrared radiometers' acquisitions on board satellites. It is used in many domains such as meteorological, climatic or ecosystem studies. For many of these applications it is important to know the accuracy of the data being used. This has been recognised internationally by the community of satellite SST data producers and users and a formal recommendation has been put forward by the Group for High Resolution SST (GHRSST) in the GHRSST Data Specification version 2.0 [3] to include Sensor-Specific Error Statistics (SSES) in distributed products.

The error is the difference between the measured SST and the true SST [4]. Systematic errors in satellite estimation of SST may have various origins [5]. One of the sources is the retrieval algorithm itself. Other sources include the calibration of the sensor which may not be accurate and contamination by sea ice, undetected clouds or atmospheric Saharan dust [6]. These effects result in inaccuracies in retrieved SST.

There is no consensus on how to derive SSES, and each data producer uses their own methodology. Methodologies are often based on regression of the error against a set of explanatory variables using a large dataset of collocations between drifting buoys and satellite retrievals. One approach, based on look-up tables (LUT) established from comparisons with in situ observations, provides discrete values of SSES. This is the case for the SSES hypercube of bias and standard deviation [7] used for observations from the Moderate Resolution Imaging Spectroradiometer (MODIS). Castro et al. [8] also proposed, for many infrared and microwave SST products, bias corrections from LUT representing bias and standard deviation dependencies on retrieval conditions such as wind speed, water vapor, view angle and SST. In Petrenko and Ignatov [9] and in Petrenko et al. [10], the SSES method is based on the segmentation of the SST domain and local regression coefficients are applied for each segment. In order to avoid possible discontinuities and noise introduced by these methods, other definitions of continuous SSES are proposed by Tandeo et al. [11] and Xu et al. [12]. A different approach is to model and propagate the uncertainties independently of in situ data [13], as used by the SST Climate Change Initiative [14].

The objective of this study is to make a first evaluation of the potential of advanced statistical methods of machine learning to model and predict the bias between satellite derived SST products and drifting buoy measurements, considered to be ground truth.

This study is based on a large data set of collocations between satellite and in situ measurements, described in Section 2. The methods are explained in Section 3 and interpretation of the results are presented in Section 4. Discussion and conclusion are in Section 5.

2. Data

The Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the Geostationary satellite Meteosat Second Generation (MSG) operates in the thermal infrared channels, enabling SST retrieval. Hourly SST products are computed operationally at the Centre de Météorologie Spatiale (CMS) in the framework of the EUMETSAT Ocean and Sea Ice (OSI SAF) project. The basis of SST retrieval is a split-window algorithm using the 10.8 and 12 μm channels. A complete description of the retrieval methodology can be found in Le Borgne et al. [15].

As part of the operational processing of MSG data, SST products are created as well as a Match-up DataSet (MDS) by collocating satellite information and in situ measurements (drifting buoys) collected from the Global Telecommunication System with a 5 days delay to ensure sufficient coverage. A satellite match-up is searched for within ± 1 hour and information is extracted for a 5×5 pixel box around the position of the measurements. The MDS includes satellite SST but also includes some other variables used in the SST processing such as Numerical Weather Prediction (NWP) model output (e.g., total atmospheric content of water vapour) and level 1 SEVIRI brightness temperature in the channels of interest for SST retrieval.

Thereafter, only night-time data are used to minimize the difference between skin SST retrieved by infrared sensors like SEVIRI and bulk SST measured by drifting buoys. The quality level (QL) is a confidence indicator designed to help users filter out data that are not sufficiently good for their application. As per recommendations formulated in the Product User Manual [16] only data with a higher QL are considered (3, 4 or 5) because for $QL < 3$ the SST retrieved is likely to be contaminated by undetected clouds.

The MDS used in this work covers two years (August 2013–July 2015) and contains 485,600 match-ups. The period from August 2013 to July 2014 is used as the learning sample (with 249,318 match-ups) to train the statistical models and the other period from August 2014 to July 2015 is used as test sample (with 236,282 match-ups) to validate statistical models.

Infrared sensors are sensitive to the skin temperature of the sea, whereas drifting buoy measurements are taken at a depth of between 20 to 30 cm which can lead to significant differences [17]. However, since this study focuses on night-time data only, it is expected that these differences are small and therefore we make the assumption that the drifting buoy measurements constitute the sea truth. The accuracy of SEVIRI SST, ΔSST , is therefore defined as:

$$\Delta\text{SST} = \text{SST}_{\text{sat}} - \text{SST}_{\text{buoys}} \quad (1)$$

where SST_{sat} is the SST estimation given in SEVIRI products and $\text{SST}_{\text{buoys}}$ is the temperature measured by drifting buoys.

Figure 1a shows the spatial distribution of ΔSST averaged over the training dataset for $5 \times 5^\circ$ boxes and Figure 1b represents the number of data points in each box. A strong negative bias is noticeable in the intertropical zone, primarily due to the high atmospheric water vapour content in this region [18], and secondarily due to the presence of Saharan dust in the atmosphere [6]. On the other hand positive biases are observed in the southern hemisphere and around the Mediterranean Sea due to the drier atmosphere. Note that due to both cloud coverage and geographical distribution of drifting buoys, the spatial distribution of the match-ups is not homogeneous at all, as shown by Figure 1b.

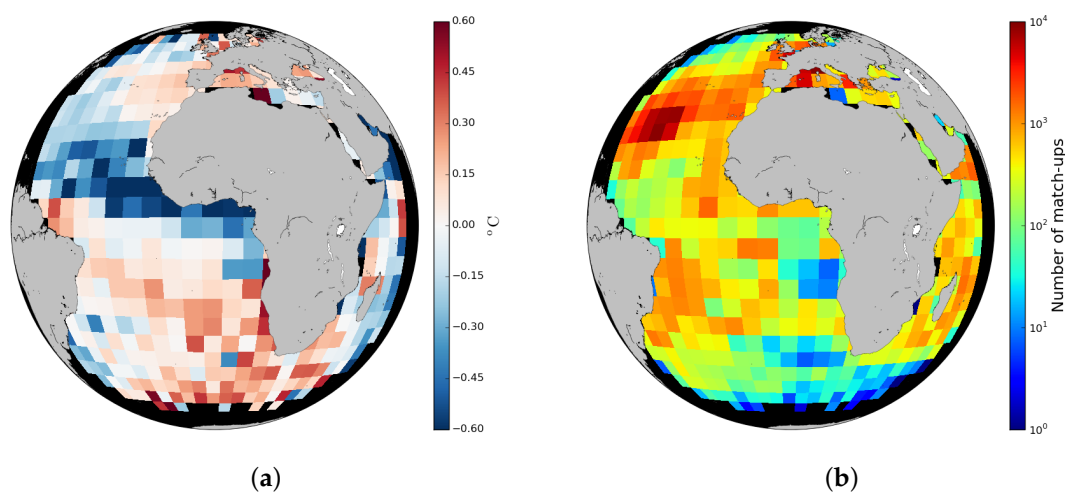


Figure 1. Training dataset (August 2013–July 2014, 249 318 match-ups): (a) mean difference between satellite and drifting buoys (ΔSST); (b) number of match-ups.

In this work, we use ten variables to model the bias in satellite-derived SST. These are listed in Table 1. Atmospheric water vapour is the primary cause of error in SST retrieval algorithms and therefore the integrated water vapour from the ECMWF model is one of the most important variables to model the bias between satellite and in situ SST. The differences between the 3.9 and the 8.7 μm channels and between the 10.8 and the 12.0 μm channels provides information on the presence of atmospheric Saharan dust [6] which can strongly affect the quality of the retrieval. These differences are averaged over 5×5 pixel boxes (of the MDS) in order to smooth out the effect of radiometric noise. Despite the

fact that this study focuses only on night-time data, it is important to include the wind speed and solar zenith angle to take into account possible residual diurnal warming effects at the beginning of the night. The number of valid (clear sky) pixels in the 5×5 pixel boxes is also used as it provides information on the level of cloudiness around the central pixel. Additionally, the standard deviation of SST in the boxes is also informative of the spatial variability (presence of thermal fronts) which may be a source of error in satellite to in situ comparison. Finally the SST is also used as a model input variable in order to account for SST dependent error in the algorithm. Since the algorithm is calibrated on a global scale, it may indeed show weaknesses for retrieving extreme temperatures (low or high).

Note that all this information is available without delay. This would allow an online statistical model to produce SST error in real time so long as the model is already trained.

Table 1. Description of the Match-up DataSet (MDS) ancillary variables.

Name	Description
Latitude	Latitude of in situ measurements
Wind speed	Near surface wind speed (ECMWF)
Solar zenith angle	Angle between zenith and sun position
Satellite zenith angle	Angle between zenith and satellite position
Integrated water vapour	Integrated water vapour in the atmosphere (ECMWF)
IR_039 – IR_087 averaged in box	Difference between channel $3.9 \mu\text{m}$ and $8.7 \mu\text{m}$ averaged in 5×5 pixels box
IR_108 – IR_120 averaged in box	Difference between channel $10.8 \mu\text{m}$ and $12.0 \mu\text{m}$ averaged in 5×5 pixels box
Number of valid pixels	Number of valid retrievals (quality level 3, 4 or 5) in 5×5 pixels box
SST STD	Standard deviation of SST in 5×5 pixels box
SST	SST retrieved from SEVIRI

3. Methods

The goal of this study is to estimate ΔSST defined in Equation (1) in order to operationally adjust SST_{sat} measurements. This will be achieved by modelling the impact of simultaneous covariates presented in Table 1 and denoted as $\{X_i, i = 1, \dots, p\}$. The relationship between ΔSST and covariates can be either nonlinear as for the latitude between 20°N and 60°N (Figure 2a) or linear with a negative slope (Figure 2b) for the integrated water vapour. Note that nonlinear interactions of covariates can also be detected (not shown). In this paper, we compare 4 regression models classically used in machine learning. They are described below.

- The first model used is a simple linear regression expressed as

$$\Delta\text{SST} = \alpha_0 + \sum_{i=1}^p \alpha_i X_i + \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,j} X_i X_j \quad (2)$$

where the α parameters correspond to the intercept, the linear and quadratic effects of covariates, and interactions between covariates.

- The second model, LASSO (Least Absolute Shrinkage and Selection Operator, see Tibshirani et al. [19]), is similar to the first one, except with a sparsity constraint on the α parameters. Thus, it is a subversion of Equation (2), where some of the α values are null. The LASSO model is based on a numerical optimization to find the alpha parameters that minimize the following expression:

$$\min_{\alpha} \frac{1}{2N} \sum_{i=1}^N \left| \Delta\text{SST} - \alpha_0 + \sum_{i=1}^p \alpha_i X_i + \sum_{i=1}^p \sum_{j=1}^p \alpha_{i,j} X_i X_j \right|^2 + \lambda \sum_{i=1}^p |\alpha_i| \quad (3)$$

where N corresponds to the number of training samples used to learn the model, $P = 1 + p + p^2$ is the total number of alpha parameters and λ is estimated by cross validation.

- The third model, GAM (Generalized Additive Model, see Hastie et al. [20]), uses nonlinear functions to model the impact of the covariates such as

$$\Delta SST = \alpha_0 + \sum_{i=1}^p f_i(X_i) + \sum_{i=1}^p \sum_{j=1}^p f_{i,j}(X_i X_j) \quad (4)$$

where functions f_i and $f_{i,j}$ are adjusted using local linear regressions, as in Figure 2a,b.

- The last model, random forest (see Breiman et al. [21]), applies N random samplings with replacement such as

$$\Delta SST = \frac{1}{N} \sum_{i=1}^N t_i(X_1, \dots, X_p) \quad (5)$$

where t_i are the different regression trees (see Breiman et al. [22]). A tree is based on simple decision criteria on the X covariates such as: if $X < \text{threshold}$ then $\Delta SST = \text{value1}$ else $\Delta SST = \text{value2}$. The threshold value is learned from the data, maximizing the difference between value1 and value2. Then, we recursively split the dataset in 2 leaves at each node of the tree. In this paper, we use trees with a maximum of 1000 nodes and the forest is based on 100 trees.

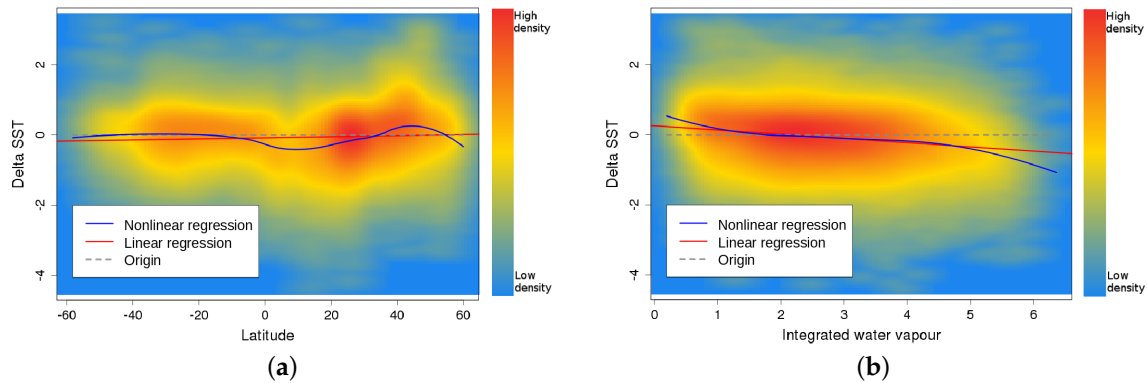


Figure 2. Density scatterplot of ΔSST as function of (a) latitude and (b) integrated water vapour fitted using linear and nonlinear (lowess) regressions.

Hereinafter, the statistical estimates of ΔSST proposed by the presented 4 models will be denoted by $\widehat{\Delta SST}$. In order to select the best model we use the adjusted R^2 (denoted as R_{adj}^2) which is negatively impacted by the number of parameters in the model and the Root Mean Square Error (RMSE). The R_{adj}^2 is given by the following equation:

$$R_{adj}^2 = 1 - \frac{\sum^N ||\Delta SST - \widehat{\Delta SST}||^2}{\sum^N ||\Delta SST - \overline{\Delta SST}||^2} \frac{N-1}{N-P-1} \quad (6)$$

where $\widehat{\Delta SST}$ corresponds to the estimations given by one of the 4 models presented above and $\overline{\Delta SST}$ the mean value computed with the N training samples. The use of the adjusted R_{adj}^2 enables a model to be found with a good fit and a low number of parameters to avoid over-fitting.

4. Results

Using the training dataset (249,318 match-ups) the four regression models presented above are determined. These models are then applied to the test sample (236,282 match-ups) to predict the bias in

the satellite estimate of SST (Equation (1)). The performance of each model is assessed using R_{adj}^2 and RMSE. Results are presented in Table 2.

Table 2. Results of R_{adj}^2 and RMSE for the different models.

	Linear Regression	LASSO	GAM	Random Forest
R_{adj}^2	24.65%	24.43%	28.44%	30.96%
RMSE	0.576 K	0.576 K	0.562 K	0.554 K

These results show that the link between Δ SST and X covariates is clearly nonlinear. Indeed, R_{adj}^2 of GAM and random forest are 4 and 6% better than linear regression (30.96% and 28.44% against 24.65%). LASSO results are very similar to linear regression. The RMSE of nonlinear models are also improved in comparison to linear models.

We denote corrected SST_{corr} as the satellite SST to which the predicted bias has been removed: $SST_{corr} = SST_{sat} - \Delta\widehat{SST}$, and we compare the corrected SST to in situ SST. Figure 3 illustrate the zonal performances of the four models on the test sample (August 2014 to July 2015) in comparison with uncorrected SST difference (in black line). All four models are able to reduce the strong negative bias between 5°S and 35°N. Linear regression and LASSO models both over estimate the bias around 5°S and amplify it North of 50°N. On the contrary GAM and random forest give better results consistently. At high latitude (where fewer match-ups are available) the random forest estimates the bias more accurately than the GAM model.

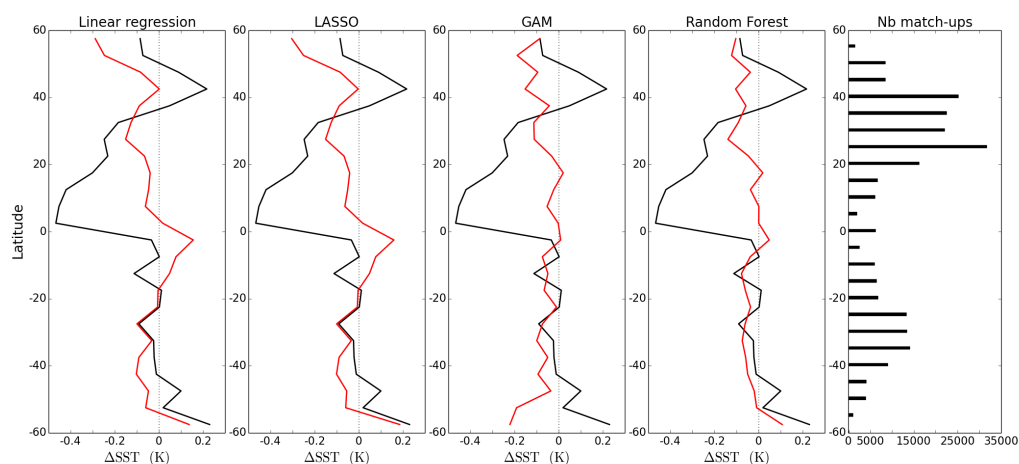


Figure 3. Δ SST = $SST_{sat} - SST_{buoys}$ (black line) and $\Delta\widehat{SST} = SST_{corr} - SST_{buoys}$ (red line) averaged per latitude bins of 5° for linear regression, Least Square Shrinkage and Selection Operator (LASSO), Generalised Additive Model (GAM) and random forest models (from left to right) for the test sample. Far right plot shows the number of match-ups for each latitude bin.

In the following development we focus on the random forest model because it gives better results and is faster to apply (once trained) than the GAM procedure, making it a better choice for operational applications.

The global mean and standard deviation of Δ SST for the test dataset (August 2014 to July 2015) before correction are equal to -0.082 and 0.664 K respectively, and after correction they are reduced to -0.071 and 0.547 K respectively. This improvement of the standard deviation can be visualized using

Figure 4 which represents the probability density functions of ΔSST and $\widehat{\Delta\text{SST}}$: when the predicted bias is applied the distribution is narrower and more Gaussian.

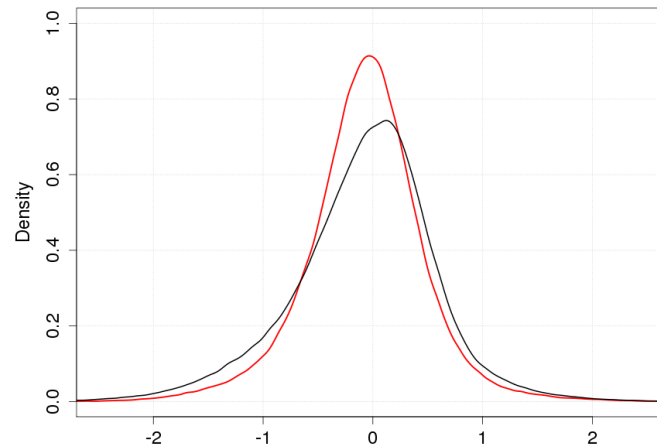


Figure 4. Histograms of $\Delta\text{SST} = \text{SST}_{\text{sat}} - \text{SST}_{\text{buoys}}$ (black line) and $\widehat{\Delta\text{SST}} = \text{SST}_{\text{corr}} - \text{SST}_{\text{buoys}}$ with correction by random forest model on test sample (red line).

Figure 5 shows the geographical distribution of the mean and standard deviation of the difference between SST_{sat} and $\text{SST}_{\text{buoys}}$ (Figure 5a,b respectively) and between SST_{corr} and $\text{SST}_{\text{buoys}}$ (Figure 5c,d respectively). Comparison of Figure 5a,c illustrate the overall reduction in the bias after subtracting the predicted bias to the satellite SST. The standard deviation is also reduced (Figure 5b,d) but high values are still observed in a number of regions. For instance, around the coast of West Africa, in the Mediterranean Sea and in the Red Sea: these regions are subject to atmospheric mineral dust events occurring only during a few months every year. High standard deviation is also visible in the Gulf Stream region and south of South Africa where strong SST gradients are observed.

Here we focus on a case study: the random forest model is applied to a satellite scene (30 April 2015 at 12 a.m., see Figure 6). This scene is composed of 792 408 clear-sky pixels ($QL > 2$). This scene was chosen because it corresponds to a large event of Saharan dust visible on Figure 6c which represents the Saharan Dust Index (SDI, a dimensionless quantity correlated to the concentration of mineral dust particles, [6]). SDI above 0.1 indicates an amount of mineral dust particles in the atmosphere that significantly affect SST retrieval. Figure 6a,b show the integrated water vapour content and the wind speed at 10m from ECMWF NWP respectively. The predicted error from the random forest model is shown in Figure 6d.

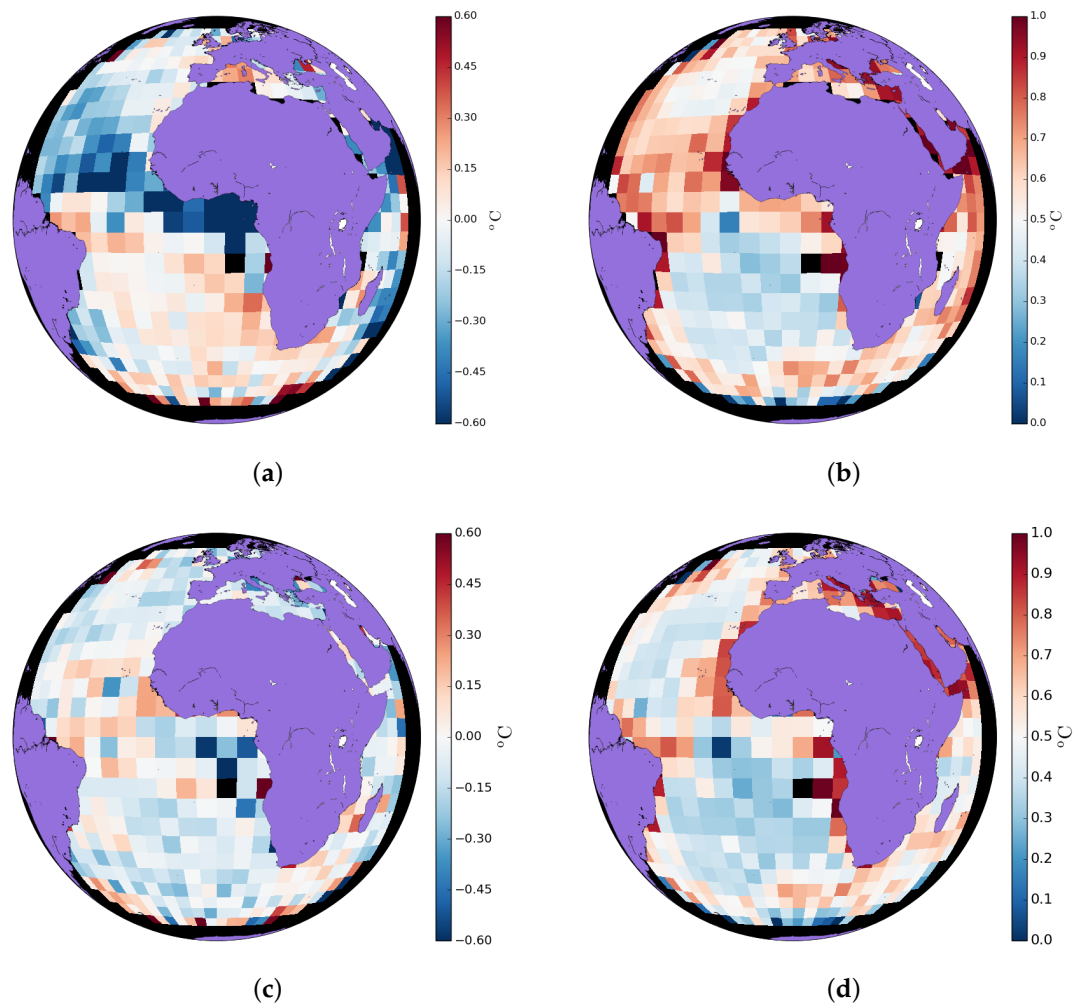


Figure 5. (a) Mean and (b) standard deviation of $SST_{sat} - SST_{buoys}$ on test sample; (c) Mean and (d) standard deviation of $SST_{corr} - SST_{buoys}$ with correction by random forest model on test sample.

Large scale features in the predicted error can be visually correlated to atmospheric features. The predicted error is largely negative where the atmosphere is humid (integrated water vapour above 4.5 g cm^{-2}) in combination with high a SDI (of the order of 0.3 or above): this is the case around the Southern coast of West Africa. On the other hand, positive errors occur when the atmosphere is drier than average (integrated water vapour below 3 g cm^{-2}) in combination with low SDI values (below 0.2): this is the case off the coast of Brazil where a south eastward thong of drier atmosphere leads to positive predicted error. It is interesting to note that where there is a combination of dry atmosphere and higher than normal SDI, the predicted error is often positive: this is the case in the Mediterranean Sea and off the coast of Namibia. There is no visible correlation between predicted error and wind speed, which is not altogether surprising since at 12 am residual diurnal warming would be minimal (and probably only observed in the western part of the domain).

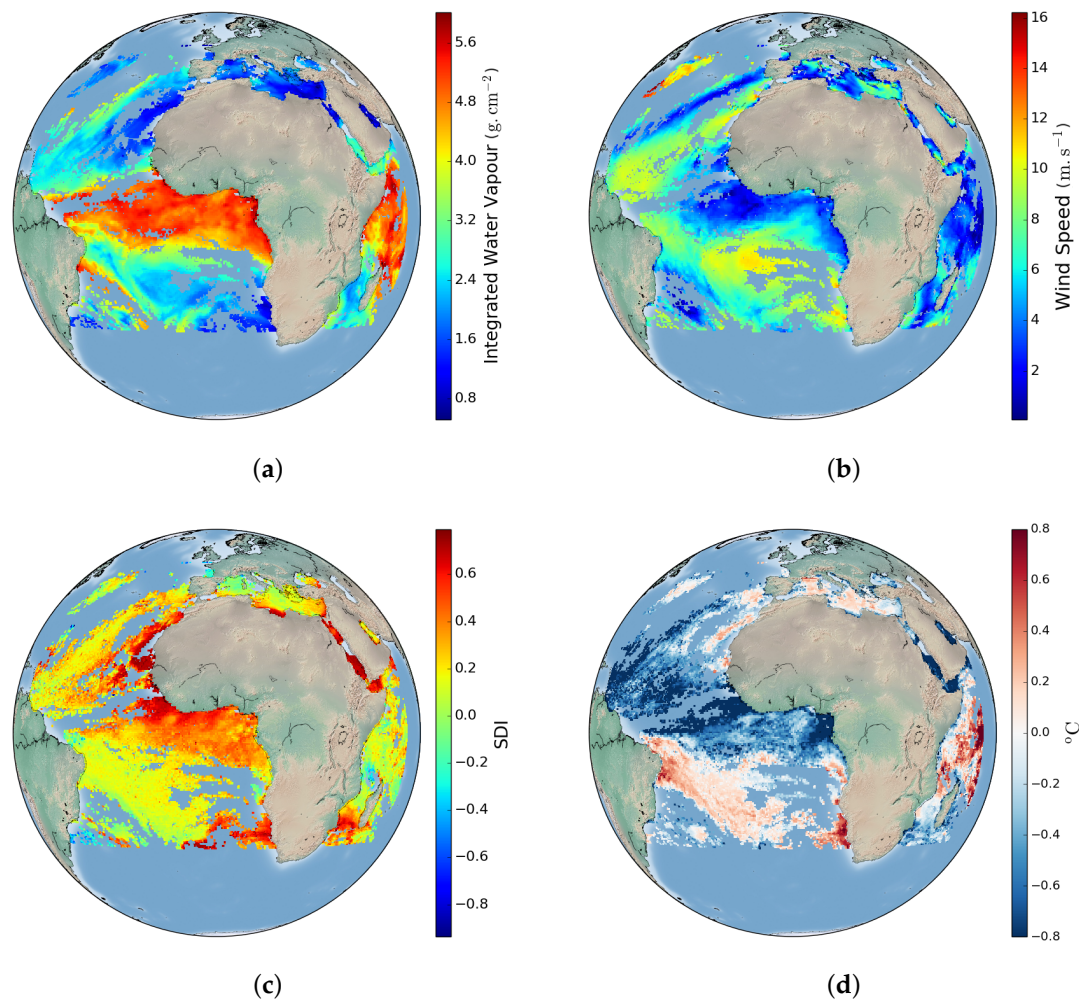


Figure 6. (a) Integrated water vapour; (b) Wind speed; (c) Saharan Dust Index; (d) Δ SST predicted by random forest model (30 April 2015 at 12 a.m.)

5. Discussion and Conclusions

The Match-up DataSet (MDS) of Météo-France Centre de Météorologie Spatiale (CMS) which collocates satellite and in situ Sea Surface Temperature (SST) measurements has been used to define statistical models of SST bias (Δ SST predicted by a model) for the Spinning Enhanced Visible and Infrared Imager (SEVIRI) sensor on-board Meteosat Second Generation (MSG) geostationary satellite. Linear regression, LASSO, GAM and the random forest model were used to fit Δ SST using the information of ten covariates.

It was shown that the nonlinear models (GAM and random forest) perform better in predicting the bias of satellite SST retrieval than linear models. They clearly manage to reduce the zonal biases associated with high water vapour content.

The random forest model was preferred over the GAM because of its slightly better results but mostly because it is quicker to run once the model has been trained making it a better choice for operational application. The random forest has then been studied further and applied on a study case (one 15-min of acquisition of the SEVIRI instrument).

Ocean and Sea Ice Satellite Application Facility (OSI SAF) operational processing of SST performed at CMS uses a very basic principle to derive SST bias: a MDS is used to compute the bias per quality level. The bias is then attributed to each pixel according to its respective quality level. This method provides discrete fields of bias which are to be avoided according the Group for High Resolution SST (GHRSST) Data Specification version 2.0 [3]. Above all this method provides a R_{adj}^2 equals to 16.74 %, so a statistical model like random forest, with a R_{adj}^2 equals to 30.96 %, is twice as efficient in capturing the variance of the error and would therefore be a valuable update to the processing chain.

The main limitation of the use of statistical models to predict the error in SST retrieval is scarceness of the in situ data. Despite the large number of match-ups some areas or phenomena are not sampled sufficiently to have a significant impact on the model. This can be seen when comparing Figure 5a,c. The model is well able to estimate the bias in the satellite retrieval at large spatial scales, which is largely due to the inability of the SST algorithm to cope with a varying atmosphere. However the high standard deviations in localised areas on Figure 5d may suggest that the model does not fully capture the spatial variability of the error.

Currently, the methodology has been proven successful in predicting bias in SST derived from a geostationary satellite which provides a large number of collocations with in situ data due to high temporal resolution of acquisitions (15 min). Although no test has been done on the minimum size of MDS required to train the random forest model, it is anticipated that for polar orbiting satellites a longer period would be needed. This is certainly a limitation associated with data-driven methodologies.

The large amount of data required to train the random forest model properly, means that an accurate model cannot be built to estimate the SST error in the first few month of the life of a satellite because too few match-ups are available. This is certainly true for polar orbiting satellite which do not provide as many match-ups as geostationary satellite. A temporary model could be built even prior to the launch of a satellite by using radiative transfer simulations of brightness temperature to train the model and updated later when sufficient in situ match-ups become available.

More work could be done to assess the random forest performance during daytime or to determine whether inclusion of simulations of brightness temperature as a covariate would be beneficial. Nevertheless, advanced statistical models such as random forest are promising for evaluation of the systematic error in SST retrieval from space with respect to in situ measurements. It is worth noting that these techniques may be applied in many other remote sensing contexts as long as large match-up datasets are available.

Acknowledgments: The data from the EUMETSAT Satellite Application Facility on Ocean and Sea Ice used in this study are accessible through the SAF's homepage <http://osi-saf.eumetsat.int>. The authors would like to thank Météo-France for funding Blandine Gausset's master thesis. The authors also wish to thank the reviewers for their very constructive comments.

Author Contributions: The idea of this study was initially designed by Pierre Tandeo, Emmanuelle Autret and Stéphane Saux Picart. Blandine Gausset did most of the work during her master thesis at Météo-France and Pierre Tandeo performed some extra analysis. The paper was mostly written by Stéphane Saux Picart and Blandine Gausset with help of the the author co-authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine learning in geosciences and remote sensing. *Geosci. Front.* **2016**, *7*, 3–10.
2. Zhang, L.; Zhang, L.; Du, B. Deep learning for remote sensing data: A technical tutorial on the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 22–40.

3. GHRSSST Science Team. The Recommended GHRSSST Data Specification (GDS) 2.0. 2011. Available online: <https://www.ghrsst.org/about-ghrsst/governance-documents/> (accessed on 29 January 2018).
4. Merchant, C.; Paul, F.; Popp, T.; Ablain, M.; Bontemps, S.; Defourny, P.; Hollmann, R.; Lavergne, T.; Laeng, A.; de Leeuw, G.; et al. Uncertainty information in climate data records from Earth observation. *Earth Syst. Sci. Data* **2017**, *9*, 511–527.
5. Merchant, C.J.; Horrocks, L.A.; Eyre, J.R.; O’Carroll, A.G. Retrievals of sea surface temperature from infrared imagery: origin and form of systematic errors. *Q. J. R. Meteorol. Soc.* **2006**, *132*, 1205–1223.
6. Merchant, C.J.; Embury, O.; Le Borgne, P.; Bellec, B. Saharan dust in nighttime thermal imagery: Detection and reduction of related biases in retrieved sea surface temperature. *Remote Sens. Environ.* **2006**, *104*, 15–30.
7. Evans, R.; Kilpatrick, K. The MODIS hypercube. In Proceedings of the 8th International GHRSSST Science Team meeting, Melbourne, Australia, 14–17 May 2007. Available online: <https://www.ghrsst.org/meetings/8th-international-ghrsst-science-team-meeting-ghrsst-viii/> (accessed on 29 January 2018).
8. Castro, S.L.; Wick, G.A.; Jackson, D.L.; Emery, W.J. Error characterization of infrared and microwave satellite sea surface temperature products for merging and analysis. *J. Geophys. Res. Ocean.* **2008**, *113*, doi:10.1029/2006JC003829.
9. Petreko, B.; Ignatov, A. SSES in ACSPO. In Proceedings of the GHRSSST XV Science Team Meeting, Cape Town, South Africa, 2–6 June 2014. Available online: <https://www.ghrsst.org/meetings/15th-international-ghrsst-science-team-meeting-ghrsst-xv/> (accessed on 29 January 2018).
10. Petrenko, B.; Ignatov, A.; Kihai, Y.; Dash, P. Sensor-Specific Error Statistics for SST in the Advanced Clear-Sky Processor for Oceans. *J. Atmos. Ocean. Technol.* **2016**, *33*, 345–359.
11. Tandeo, P.; Autret, E.; Piolle, J.; Tournadre, J.; Ailliot, P. A Multivariate Regression Approach to Adjust AATSR Sea Surface Temperature to In Situ Measurements. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 8–12.
12. Xu, F.; Ignatov, A.; Liang, X. Towards continuous error characterization of sea surface temperature in the advanced clear-sky processor for oceans. In Proceedings of the 89th AMS Annual Meeting, 16th Conference on Satellite Meteorology and Oceanography, Phoenix, AZ, USA, 10–16 January 2009; pp. 11–15.
13. Bulgin, C.; Embury, O.; Corlett, G.K.; Merchant, C. Independent uncertainty estimates for coefficient based sea surface temperature retrieval from the Along-Track Scanning Radiometer instruments. *Remote Sens. Environ.* **2016**, *178*, 213–222.
14. SST-CCI. SST CCI Uncertainty Characterisation Report v2; SST-CCI-UCR-UOE-002. 2013. Available online: www.esa-sst-cci.org (accessed on 29 January 2018).
15. Le Borgne, P.; Roquet, H.; Merchant, C. Estimation of Sea Surface Temperature from the Spinning Enhanced Visible and Infrared Imager, improved using numerical weather prediction. *Remote Sens. Environ.* **2011**, *115*, 55–65.
16. OSI SAF. *Geostationary Sea Surface Temperature Product User Manual*; Technical Report; EUMETSAT: Berlin, Germany, 2011.
17. Donlon, C.J.; Minnett, P.J.; Gentemann, C.; Nightingale, T.J.; Barton, I.J.; Ward, B.; Murray, M.J. Toward Improved and Validation of Satellite and Sea Surface and Skin Temperature and Measurements and for Climate and Research. *J. Clim.* **2002**, *15*, 353–359.
18. Marsouin, A.; Le Borgne, P.; Legendre, G.; Péré, S.; Roquet, H. Six years of OSI-SAF METOP-A AVHRR sea surface temperature. *Remote Sens. Environ.* **2015**, *159*, 288–306.
19. Tibshirani, R.J. Regression Shrinkage and Selection via the lasso. *J. R. Stat. Soc.* **1996**, *58*, 267–288.
20. Hastie, T.J.; Tibshirani, R.J. *Generalized Additive Models*; Chapman & Hall/CRC: London, UK, 1990.
21. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32.
22. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth: Belmont, CA, USA, 1984.

