Remote Sensing of Environment 164 (2015) 334–343

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Evaluation of different covariance models for the operational interpolation of high resolution satellite Sea Surface Temperature data over the Mediterranean Sea

B. Buongiorno Nardelli [a,b,*], A. Pisano [a], C. Tronconi [a], R. Santoleri [a]

[a] *CNR—Istituto di Scienze dell'Atmosfera e del Clima-Gruppo Oceanografia da Satellite-Roma, Roma, Italy*
[b] *CNR—Istituto per l'Ambiente Marino Costiero, Napoli, Italy*

ARTICLE INFO

Article history:
Received 7 November 2014
Received in revised form 10 April 2015
Accepted 21 April 2015
Available online 8 May 2015

Keywords:
Sea surface temperature
Satellite observations
Optimal interpolation
Covariance model
Mediterranean Sea

ABSTRACT

Daily gap-free Sea Surface Temperature (SST) fields at high resolution are required by several operational users working on monitoring and forecasting the status of the marine environment. Existing instruments cannot provide such fields, and careful interpolation of available data is required to provide gap-free SST estimates. In the framework of the MyOcean projects (now turning into the European Copernicus Marine Service), several satellite (and/or satellite and in situ) interpolated products are distributed in real time, and continuous research and development activities are carried out to improve their quality. In this paper, we describe the work done to improve the High Resolution (1/16°) interpolated product covering the Mediterranean Sea, that is obtained by combining all available satellite infrared images. Three Optimal Interpolation schemes based on different space–time covariance models and background fields are derived, compared and validated versus in situ drifting buoy measurements (leading to a minimum standard deviation error (STDE) of 0.46 °K and mean bias error (MBE) of −0.15 °K), as well as by adopting a holdout validation approach with artificial clouds (with STDE ranging between 0.22 °K and 0.55 °K, and MBE always between 0.03 °K and 0.05 °K, depending on the cloud configuration considered). Almost negligible differences are found between the three products, revealing only a slight improvement when using spatially varying covariance parameters and a daily climatology as background, which is less risky in case of prolonged cloudy conditions, though attaining the same performance with small/rapid clouds. This scheme has thus been implemented in the Mediterranean SST operational chain since April 2014.

1. Background

Sea Surface Temperature (SST) is a fundamental variable for many scientific and operational applications (see http://www.ghrsst.org). It is primarily required by the meteorological and marine operational forecasting systems to constrain their numerical prediction models (e.g. Chelton & Wentz, 2005; Dobricic et al., 2007), but also by public institutions and private companies working, for example, on marine environment and security managing, fisheries, tourism, marine transportation, offshore exploration and extraction. Most of these users need a high resolution (HR) gap-free estimate of the SST, known as level 4 (L4) data (as examples of environmental analyses requiring gap-free SST see Volpe, Buongiorno Nardelli, Cipollini, Santoleri, & Robinson, 2012; Liu et al., 2013; Bonanno et al., 2015). In situ instruments, however, only provide sparse point observations, and even satellite sensors are affected by both coverage and physical constraints. More specifically, space-borne infrared sensors cannot 'see' through clouds, while microwave measurements are contaminated by several factors (such as rain, land, ice etc.), so that satellite images are also always characterized by data voids (Robinson, 2004). Interpolation of available data is thus a crucial step to provide an accurate L4 SST field.

Within the GMES (Global Monitoring for Environment and Security)/Copernicus MyOcean projects, funded by the European Commission, a number of global and regional SST near-real-time (NRT) daily L4 products based on the combination of available satellite (and/or satellite and in situ) measurements have been developed and are currently distributed through a single user interface (http://www.myocean.eu/web/24-catalogue.php). All these products are obtained by applying a statistical technique known in literature as Optimal Interpolation (OI).

First introduced by Gandin (1965) and by Bretherton, Davis, and Fandry (1976) to the oceanographic community, classical optimal interpolation (OI) method can be applied in several different ways, depending on a number of assumptions, from the definition of the state vector to be reconstructed, which can require very different covariance models and calibration techniques, to the numerical algorithms used. Through OI, SST L4 data are obtained as a linear combination of the observations (namely, of the SST anomalies with respect to a background field, also

---

[*] Corresponding author at: CNR-IAMC Calata Porta di Massa, 80133, Napoli, Italy.

http://dx.doi.org/10.1016/j.rse.2015.04.025

called *first guess* field), weighted directly with their correlation to the interpolation point and inversely with their cross-correlation and error. OI thus requires that the background error covariance and observation error covariance are known, and combines these error covariances and observations to provide the minimum variance solution (i.e. the solution that minimizes the mean square error of the analysis).

Though, theoretically, the background covariance matrix should be estimated directly from all available observations, the large amount of satellite SST data makes this approach impractical for NRT applications, and most of the SST L4 operational systems are thus based on predetermined parametric (analytical) covariance functions estimated from historical observations or from numerical model output (e.g. Donlon et al., 2012; Martin et al., 2012; Reynolds & Chelton, 2010; Roberts-Jones, Fiedler, & Martin, 2012, and references therein). Covariance functions depending only on spatial distance are also generally preferred. These OI schemes take the previous day analysis as the background field and daily increments with respect to this background are used as input data, eventually relaxing to daily/weekly climatologies in case of prolonged lack of observations.

Parametric approaches have the advantage of reducing the amount of data to be taken in input, as covariance functions clearly decay at increasing distance and observations that are found far from the interpolation point can be safely excluded from the analysis, which makes all these algorithms theoretically sub-optimal, but computationally efficient. However, approximated OI approaches can also be easily extended to higher-dimensional Euclidean spaces, including also temporal decorrelations (e.g. Le Traon, Nadal, & Ducet, 1998; Marullo et al., 2014), or even to higher dimensional spaces (e.g. Buongiorno Nardelli, 2012). More complex space–time models, even if requiring additional computational time, directly take into account temporal correlations instead of limiting to a spatial approach or hybrid spatial/climatological relaxation approach, which implies the discretional definition of a temporal e-folding scale. In fact, in ‘truly’ optimal space–time interpolation, one would consider a temporal sequence of the SST field as the state vector, and several realizations of these space–time data should be used to estimate the covariance. In the approximated space–time covariance models, this turns up into the need to define a mixed space–time covariance function and to estimate corresponding parameters. Simplified space–time models, however, allow simpler estimations by assuming the covariance function can be approximated as the product of space and time covariance functions, separately.

The Mediterranean Sea and Black Sea L4 SST products developed by the Consiglio Nazionale delle Ricerche-Istituto di Scienze dell'Atmosfera e del Clima-Gruppo di Oceanografia da Satellite (CNR-ISAC-GOS) within MyOcean projects, as well as in previous projects/research activities, have been obtained directly through a space–time OI approach (Buongiorno Nardelli, Colella, Santoleri, Guarracino, & Kholod, 2010; Buongiorno Nardelli, Tronconi, Pisano, & Santoleri, 2013; Buongiorno Nardelli et al., 2003; Marullo, Buongiorno Nardelli, Guarracino, & Santoleri, 2007; Santoleri, Marullo, & Böhm, 1991).

The most recent algorithm used for the Mediterranean (MED) HR L4 product (at 1/16° resolution) has been fully described and validated in Buongiorno Nardelli et al. (2013), and was operational within the MyOcean system until April 2014 (corresponding to the release of an upgraded MyOcean service, hereafter referred to as MyOcean V4). This algorithm was basically the same used by Marullo et al. (2007), briefly recalled hereafter. The background correlation function was assumed to depend separately from time and space lag:

$$C = C(\Delta r, \Delta t) = C(\Delta r) \cdot C(\Delta t) \qquad (1)$$

and was defined as the product of two negative exponential functions with fixed and isotropic decorrelation space and time scales. These scales were estimated by a least-square fit of observed lagged-correlations in space and time, but were taken uniform all over the interpolation domain.

Since the MyOcean V4 release, a new covariance model has been adopted, relaxing this last assumption, i.e. using non-uniform space and time decorrelation scales, and also adopting a different functional dependence, based on the tests described hereafter. In particular, different background fields (and corresponding covariances) have also been tested, trying to improve the accuracy of the analysed field. This paper thus describes the new models tested and provides the CAL/VAL of the current (V4) operational MyOcean HR L4 SST data over the Mediterranean Sea.

2. Data

2.1. Satellite observations

All satellite data used for the SST L4 processing at CNR are provided by the Group for High-Resolution Sea Surface Temperature (GHRSST) Global and Regional Data Assembly Centres (GDAC, RDAC; see www.ghrsst.org for more details). The sensors (and platforms) ingested are: the Advanced Along Track Scanning Radiometer (AATSR, installed on the European ENVIronment SATellite, ENVISAT), MODerate resolution Imaging Spectroradiometer (MODIS, on both Aqua and Terra satellites), Advanced Very High Resolution Radiometer (AVHRR, on METeorological Operational satellite, METOP, and National Oceanic and Atmospheric Administration (NOAA) satellites), and Spinning Enhanced Visible and Infrared Imager (SEVIRI, installed on Meteosat Second Generation, MSG). The data considered consist of one year (2011) of the GHRSST Level 2P (L2P), that contain single sensor SST observations on the native grid/swath, geo-location data, error estimates (SSES, Single Sensor Error Statistics, i.e. bias error and standard deviation error), land and ice flags, as well as additional auxiliary fields for each pixel, referred to as dynamic flags. The information contained in the L2P files is used for a pre-selection of the data that are ingested in the processing chain, mapping super-collated (namely merged, multi-sensor) observations on the interpolation grid (these data will be referred to as L3S in the following). In particular, to avoid both diurnal warming effects and residual cloud contamination, we keep only the night-time observations that are flagged with quality level 5 (best observations) in the original data file. L3S data are then obtained by removing from each L2P image a large scale sensor bias estimated through an iterative procedure that adjusts all images to a reference sensor (AATSR or METOP) and keeps only one value per pixel, as detailed in Buongiorno Nardelli et al. (2013).

2.2. In situ measurements and matchup data

GHRSST group on satellite SST validation (STVAL) indicates drifting buoy measurements as the baseline for the calibration and validation of all SST products (https://www.ghrsst.org/ghrsst/tags-and-wgs/stval-wg/sses-common-principles/). The L4 validation activities described in the following sections are thus based on the compilation of a match-up database between satellite SST L4 and surface drifting buoys measurements distributed by the MyOcean In Situ Thematic Assembly Centre (In Situ-TAC), covering the year 2011. Standard quality flags are used in the CNR processing chain to exclude all suspect measurements, keeping only the highest quality data (for more details see also the MyOcean In Situ TAC Product User Manual, http://catalogue.myocean.eu.org/static/resources/myocean/pum/MYO2-INS-PUM-013-V1.2.pdf). Furthermore, specific quality control (QC) procedures have been set up and applied to drifter measurements before building our matchup dataset, as described in Buongiorno Nardelli et al. (2013). The matchup is restricted to night-time data (between 9 p.m. and 6 a.m. local time), keeping only one value per buoy for each date, in order to get fully independent measurements (autocorrelated information would be given by sequential measurements from the same drifter within the time period considered, see also Bayley and Hammersley, 1946). In fact, the single SST value per buoy/day used here is the closest (in time) to the position time (the drifting buoys collect SST data at