Assessment of Satellite-Derived Sea Surface Salinity in the Indian Ocean

Smitha Ratheesh, Bhasha Mankad, Sujit Basu, Raj Kumar, and Rashmi Sharma

***Abstract*—The study has been motivated by the desire to assess the performance of sea surface salinity (SSS) from the Soil Moisture and Ocean Salinity (SMOS) satellite launched by the European Space Agency. Daily Level 3 product on a $0.25^\circ \times 0.25^\circ$ grid for the year 2010 has been used for this assessment in the Indian Ocean. Various data sets, like the *in situ* data sets available from the Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) buoys and Argo floats and also the data sets from modular ocean model version 3 simulations, have been utilized for this purpose. Comparison made at two buoy locations suggests good quality of SMOS SSS with root-mean-square error of the order of 0.36 and 0.34 psu. The triple collocation method, which explicitly takes into account the error characteristics of the SMOS, Argo, and model data sets, has been used for further validation of the SMOS data. Since the Indian Ocean exhibits characteristically different patterns of SSS in its different subregions, the study area has been divided into different such subregions. The SMOS-derived SSS appears to be of very good quality in the equatorial Indian Ocean and southern Indian Ocean, while the data are of poorer quality in the Arabian Sea and the Bay of Bengal possibly because of the errors in SSS retrieval due to the land contamination and strong winds.**

***Index Terms*—Argo floats, functional relationship (FR), modular ocean model (MOM) version 3 (MOM3) ocean model, Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) buoys, Soil Moisture and Ocean Salinity (SMOS) salinity, triple collocation.**

I. Introduction

OCEAN salinity, along with temperature, controls the oceanic thermohaline circulation. Its knowledge is also essential for studying the variations of the mixed layer depth [1]. As far as the Indian Ocean is concerned, ocean salinity plays a vital role in ocean-atmosphere coupling, an important manifestation of which is the Indian Ocean dipole [2], [3]. Moreover, in the Indian Ocean, the steric part of the sea level rise associated with climate change is crucially dependent on the SSS variability [4]. The Global Ocean Data Assimilation Experiment specifies an accuracy requirement of 0.1 psu on a spatial resolution of $2^\circ \times 2^\circ$ and a temporal resolution of ten days for large-scale circulation studies [5]. Recognizing the importance of salinity in the sea in general and sea surface salinity (SSS) in particular, two satellite missions have been proposed for measuring this parameter from space. The first of these missions is the Soil Moisture and Ocean Salinity (SMOS) satellite of the European Space Agency (ESA), and the second one is the Aquarius satellite of the National Aeronautics and Space Administration. SMOS was launched in November 2009 and is the first satellite to measure SSS from space [5], [6].The repeat cycle of the satellite is 149 days, with 18 days as "almost repeat." Three days is the period necessary for a full Earth coverage.

In this letter, preliminary measurements of SSS from SMOS in the Indian Ocean are investigated. Our study closely follows a similar previous study where SSS from SMOS was investigated in the Atlantic [7]. However, in that paper, SMOS data of only one month, namely, September 2010, were investigated, whereas we have used data for the full 2010. Another major distinction of our study is the use of triple collocation functional relationship (FR) method, unlike the simple linear regression (LR) used in the mentioned study. In LR, it is tacitly assumed that the dependent variable (e.g., Argo-measured SSS or numerical model-generated SSS), against which the measured data (SMOS-measured SSS) are being compared, represents the "physical truth," without any error. This is, of course, far from reality. On the other hand, the triple collocation FR method [8], [9] takes into account the different error characteristics of all the data sets in a very efficient manner.

II. Data and Methodology

ESA reprocessed SMOS-derived L1A data for the year 2010 based on up-to-date algorithm. Taking these data as input, the Centre Aval de Traitement des Donnees SMOS (CATDS)–CATDS Expertise Center- Ocean Salinity has produced one year of composite Level 3 SSS data. These so-called "CATDS/CEC-OS SMOS Level 3 SSS research products" are available as monthly, ten-day, and daily composites. The Level 3 product version is V01. We have used the daily composites available as ten-day running averages. These are gridded at $0.25^\circ \times 0.25^\circ$ resolution (http://www.catds.fr).

The validation data set consists of the data available from the Research Moored Array for African–Asian–Australian Monsoon Analysis and Prediction (RAMA) buoys and Argo floats. The RAMA buoys were deployed for the improved description, understanding, and prediction of east African, Asian, and Australian monsoon systems. The data have been obtained from the site www.pmel.noaa.gov/tao. The data were averaged in the same way as the SMOS data (as running ten-day averages computed daily) to facilitate proper comparison. Argo is an array of profiling floats reporting temperature and salinity from 2000 m to surface (4–5 m), and the Argo SSS fields were obtained from the official Argo site (http://www.argo.net). Although Argo does not measure salinity exactly at the sea surface, it is possible to use the shallowest Argo measurements as proxy for SSS. This is not a source of large error in regions with a

Manuscript received April 19, 2012; revised June 1, 2012 and June 14, 2012; accepted June 25, 2012. Date of publication August 13, 2012; date of current version November 24, 2012.

The authors are with the Atmospheric and Oceanic Sciences Group, Space Applications Centre, Ahmedabad 380015, India (e-mail: smitha@sac.isro.gov.in; bhasha_mankad@yahoo.com; rumi_jhim@yahoo.com; rksharma@sac.isro.gov.in; rashmi@sac.isro.gov.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2012.2207943

well-mixed surface layer [7]. This condition is generally fulfilled in the Indian Ocean. In addition to these *in situ* data sets, the SSS obtained from the global ocean general circulation model (OGCM) modular ocean model (MOM) version 3 (MOM3) has also been used. This model is being run in-house. The model domain is 180° W – 180° E and 80° S – 80° N with a variable horizontal resolution of $0.5° \times 0.5°$ in the Indian Ocean (30° S – 30° N; 40–110° E) going to a maximum of 2° in the rest of the domain [10]. There are 38 levels in the vertical with 20 levels in the upper 150 m. The model is forced with surface forcings obtained from National Center for Environmental Prediction/National Center for Atmospheric Research Reanalysis products. No satellite data have been assimilated in the ocean model. Also, the model SSS field has not been relaxed to climatology. For the purpose of comparison, the SMOS SSS and model-derived SSS have been colocated with Argo SSS. The nearest SMOS SSS and model-derived SSS corresponding to a specific Argo location (for a particular day) are assumed to be colocated with that particular Argo-derived SSS.

The study area extending from 30° E to 110° E and 30° S to 25° N has been further divided into four subregions, namely, the Bay of Bengal (BOB) spanning (80° E–110° E, 5° N–25° N), the Arabian Sea (AS) spanning (30° E–80° E, 5°N–25° N), the equatorial Indian Ocean (EIO) spanning (30° E–110° E, 5° S – 5° N), and the southern Indian Ocean (SIO) spanning (30° E–110° E, 30° S – 5° S). This is because, in the earlier studies, it has been found that these subregions are more or less homogeneous as far as SSS variability is concerned.

We now briefly recapitulate the essential features of the FR method. Assume that we have three sets of N collocated observations (x_i, y_i, z_i), $i = 1, 2, 3, \ldots, N$, corresponding to a set of deterministic true values T_i. The measurements are generally contaminated by errors $(e_{\mathrm{xi}}, e_{\mathrm{yi}}, e_{\mathrm{zi}})$, which are assumed to be zero-mean random variables. More precisely, after dropping the subscripts, we can write

$$\begin{aligned} x &= X + e_{\mathrm{x}} = T + e_{\mathrm{x}} \\ y &= Y + e_{\mathrm{y}} = a_1 + b_1 T + e_{\mathrm{y}} \\ z &= Z + e_{\mathrm{z}} = a_2 + b_2 T + e_{\mathrm{z}}. \end{aligned} \tag{1}$$

In our case, it is assumed that x represents Argo SSS, while y and z are SMOS-derived and model-derived SSS, respectively. It has been tacitly assumed that Argo provides the most accurate observation of the true SSS. In other words, Argo is closest to the truth, while the other two observations are linear functions of truth, contaminated by zero-mean random errors. Whether the assumption is true, or not, would be known only after estimating the unknown error variances. Another assumption used in FR is that the errors associated with different systems are uncorrelated. In this letter, this assumption can be safely applied as all the three systems, namely, ARGO, model, and SMOS, provide data, which are mutually independent. Simple algebraic manipulation [8] leads to the following estimates of the error variances:

$$\begin{aligned} \langle e_{\mathrm{x}}^2 \rangle &= \langle x^* 2 \rangle - \langle x^* y^* \rangle \langle x^* z^* \rangle / \langle y^* z^* \rangle \\ \langle e_{\mathrm{y}}^2 \rangle &= \langle y^* 2 \rangle - \langle x^* y^* \rangle \langle y^* z^* \rangle / \langle x^* z^* \rangle \\ \langle e_{\mathrm{z}}^2 \rangle &= \langle z^* 2 \rangle - \langle x^* z^* \rangle \langle y^* z^* \rangle / \langle x^* y^* \rangle. \end{aligned} \tag{2}$$

In (2), $\langle \rangle$ denotes statistical average and x^*, y^*, z^* denote the variables after the removal of the corresponding averages. In a similar manner, the coefficients a_1, a_2 and b_1, b_2 can also be

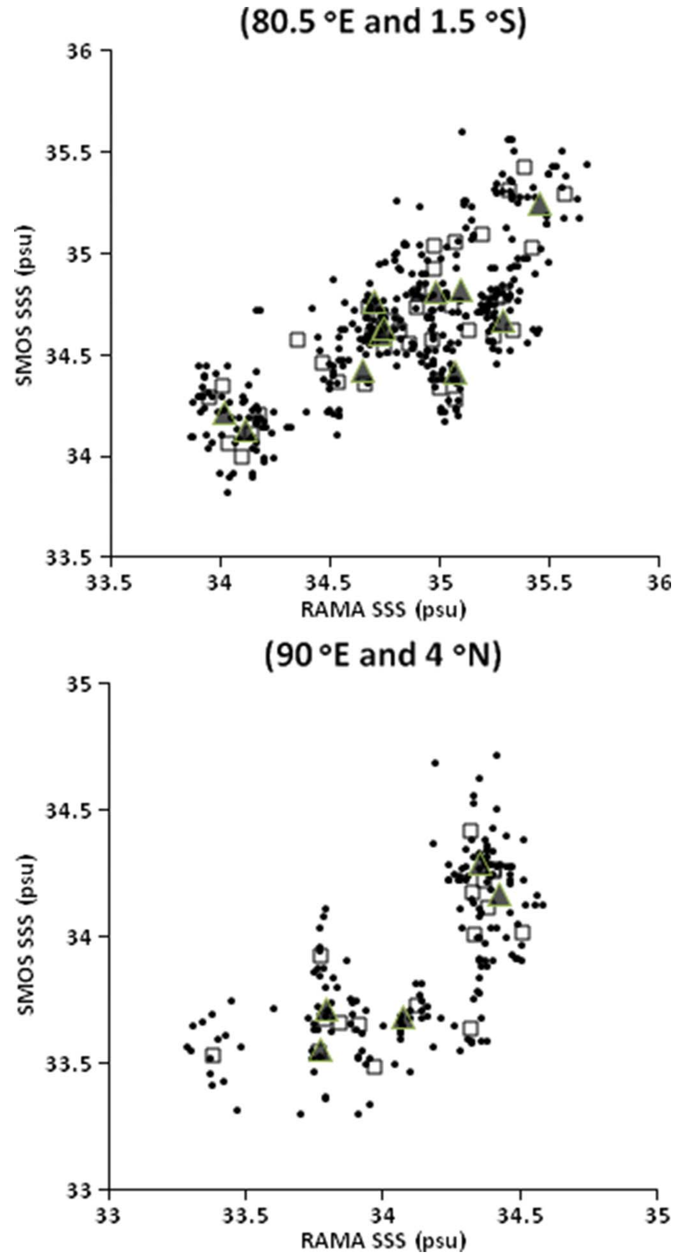

Fig. 1. Scatter plots between daily SMOS SSS and SSS measured by RAMA buoys in the EIO (top panel) 1.5° S and 80.5° E (bottom panel) 4° N and 90° E. Overlaid on each plot are the ten-day and monthly averaged comparisons also. The solid circles are for daily comparison, the squares are for ten-day averaged comparison, and the triangles are for monthly averaged comparison.

easily obtained, and the corresponding expressions are being omitted for the sake of economy.

III. Results

Two stations from the EIO were selected for the comparison of SMOS salinity data with RAMA buoy salinity data. The result is shown in Fig. 1. The match between the two data sets is reasonably good. The coefficient of correlation (R) between SMOS SSS and buoy SSS is 0.74 for the buoy located at 1.5° S and 80.5° E. R improves to 0.78 and 0.84 for the ten-day averaged and monthly averaged comparisons, respectively. The root-mean-square error (rmse) has been found to be 0.36, 0.33,

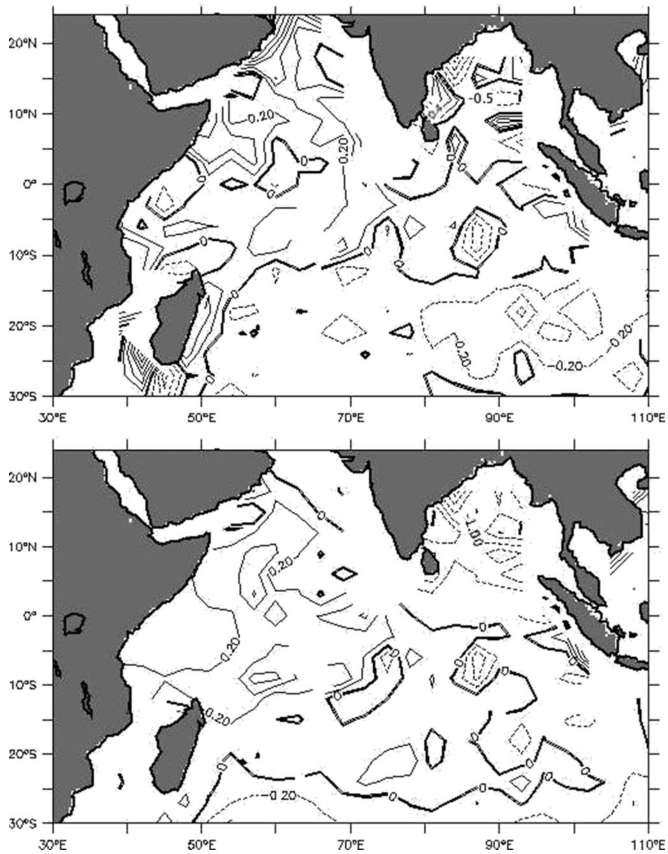

Fig. 2. Annually averaged maps of $3° \times 3°$ gridded differences between (top panel) Argo and SMOS SSS and (bottom panel) Argo and model SSS.

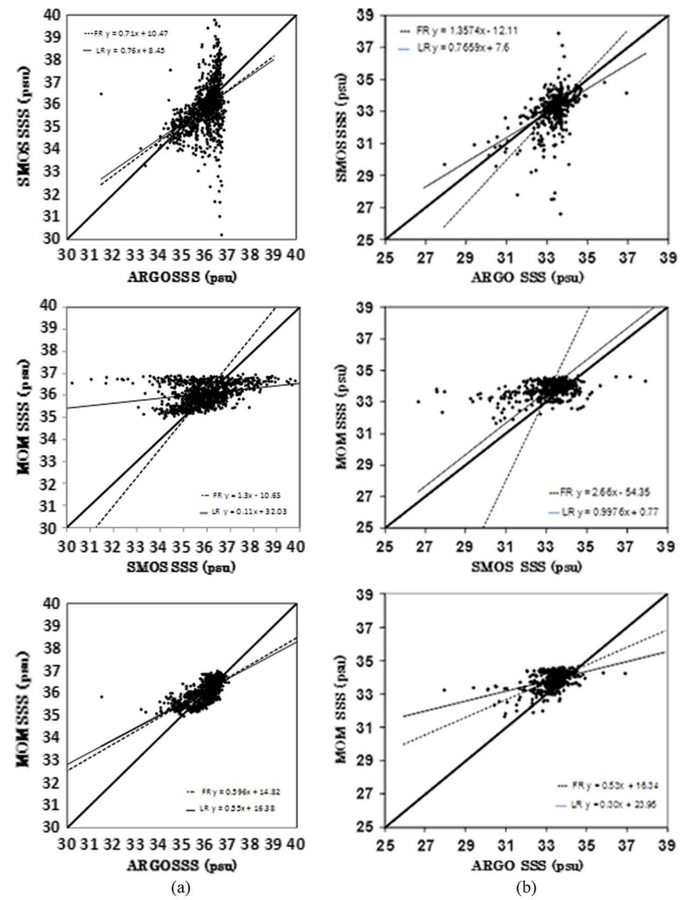

Fig. 3. (a) Scatter plots of daily SSS from SMOS, Argo, and model simulation using triple collocation method for the EIO. The solid line is for the LR fit, and the dotted line is for the FR fit. (Bold solid) One-to-one line is also drawn in the figure. (b) Same as in (a) except that it is for the AS.

and 0.28 psu for the daily, ten-day, and monthly comparisons, respectively. For the buoy located at 4° N and 90° E, the R values are 0.58, 0.72, and 0.98 while the rmses are 0.34, 0.29, and 0.21 psu, respectively.

After this preliminary comparison, it was imperative to check whether the known general structure of SSS [1] has been reasonably well picked up by the SMOS data. It has been observed that (figure not shown) SMOS could pick up the low salinity in the northern BOB occurring due to the large river discharge coupled with high rainfall amount. Both the high salinity in the AS region and the relatively low salinity in the eastern EIO due to the intrusion of low saline water from the Pacific have also been nicely picked up by SMOS. These features are well captured by MOM3 simulations too. However, in order to see the regional differences among the three data sets, we show (see Fig. 2) the spatial distribution of the differences between gridded annually averaged SSS from Argo and from SMOS (top panel) and between Argo SSS and model SSS (bottom panel). All the three data sets of SSS are gridded to a uniform grid size of $3° \times 3°$. This is simply to ensure that sufficient number of Argo observations is available in a grid. It is clear from Fig. 2 that the differences are not too significant in the SIO. However, the same cannot be said about the northern Indian Ocean. It is evident that SMOS underestimates SSS in the northern AS and overestimates it in the northwest BOB region. On the other hand, the model performs exceedingly well in the entire AS, while systematically overestimating the salinity in the entire BOB. Comparisons of Level 2 SMOS SSS data with the high-resolution hybrid coordinate ocean model simulations of SSS reveal large differences in the Indian Ocean [11]. Overestimation of the SSS by SMOS in the BOB has also been reported [12].

After this gross large-scale assessment of the quality of SMOS data, we proceeded to the triple collocation analysis study. Fig. 3(a) shows the scatter plots between the data sets for the EIO region. The dotted line in this and in Fig. 3(b) represents the FR line. For distinguishing the FR method from the standard LR analysis, we also show the LR line in these figures as a solid line. The 45° line is also shown as a standard benchmark. In the EIO [see Fig. 3(a)], the error variances are the following: For x(Argo), it is 0.05 $(\text{psu})^2$; for y(SMOS), it is 0.11 $(\text{psu})^2$; and for z(MOM3), it is 0.12 $(\text{psu})^2$. The general tendency of the collocated points is to align themselves more along the fitted FR line. Of course, this is not crystal clear from the figure, since the FR and LR lines are close to each other, which is a consequence of the fact that the error variances are of the same order of magnitude [8], with Argo being the most accurate, which is to be expected. In making this inference, we have been mainly guided by the original paper related to FR method [8]. Anyway, the top panel of Fig. 3(a) establishes the fact that the SMOS data are of high quality in the EIO. The reason for the fairly good match could be the low wind variability in this region with minimal land contamination. In Fig. 3(b), we show the result for the AS. Here, large scatters are noted in the panels involving SMOS SSS.

There appears to be unrealistically large variation in SMOS SSS ranging from 30 to 40 psu, while the upper limit in the model and Argo is 37 psu. The reason for this might be either land contamination or inaccurate retrieval under high winds (occurring in the AS during southwest monsoon), or maybe a combination of both. For the sake of records, we report the error variances, which are 0.02 psu^2 (Argo), 1.1 psu^2 (SMOS), and 0.12 psu^2 (MOM3). We have also performed the triple colocation analysis in the other two regions, namely, the BOB and the SIO. The corresponding figures are not shown for the economy of space. Suffice it to say that the SIO results are similar to the results for the EIO, pointing to the high quality of SMOS data in this region, while the BOB results are analogous to that for the AS, pointing to the poorer performance of SMOS in this region. The reason for the mismatch in the BOB is the same as that for the AS.

IV. Conclusion

The objective of this study is to validate the SMOS-derived SSS data for the Indian Ocean. The study has been done using *in situ* as well as model data sets. Comparison with RAMA buoy data has been done for two stations in the Indian Ocean, and a reasonably good match between the two data sets has been observed. The annual structure of SSS observed by SMOS has been found to be in reasonably good agreement with that observed by Argo floats and simulated by an OGCM. A triple collocation analysis has been performed among the Argo, SMOS, and model data sets. This method takes into account the different error characteristics of the various data sets. SMOS data have been found to be of very high quality in the EIO and SIO. The data are of poorer quality in the AS and the BOB, possibly due to the land contamination in the satellite footprint. High winds during monsoon in the AS and the BOB might have also contributed to the error in SSS retrieval by SMOS.

Acknowledgment

The authors would like to thank the Director of the Space Applications Centre and the Deputy Director of the Earth, Ocean, Atmosphere, Planetary Sciences and Applications Area for the encouragement and the Group Director of the Atmospheric and Oceanic Sciences Group for the motivation toward this study. Soil Moisture and Ocean Salinity data were obtained from the CATDS, operated for the Centre National d'Etudes Spatiales, France, by l'Institut Francais de Recherche Pour l'Exploitation de La Mer (Brest, France). The authors would also like to thank the Tropical Atmosphere Ocean Project Office of National Oceanic and Atmospheric Administration/Pacific Marine Environmental Laboratory for the Argo data set used in the study.

References

[1] R. Sharma, N. Agarwal, I. M. Momin, S. Basu, and V. K. Agarwal, "Simulated sea surface salinity variability in the tropical Indian Ocean," *J. Climate*, vol. 23, no. 24, pp. 6542–6554, Dec. 2010.

[2] N. H. Saji, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, "A dipole mode in the tropical Indian Ocean," *Nature*, vol. 401, no. 6751, pp. 360–363, Sep. 1999.

[3] P. J. Webster, "The role of hydrological processes in ocean-atmosphere interaction," *Rev. Geophys.*, vol. 32, no. 4, pp. 427–476, Nov. 1994.

[4] S. George, R. Sharma, N. Agarwal, S. Basu, and A. Sarkar, "Dynamic height anomaly from Argo profiles and sea level anomaly from satellite altimetry: A comparative study in the Indian Ocean," *Int. J. Remote Sens.*, vol. 32, no. 18, pp. 5105–5113, Sep. 2011.

[5] J. Font, A. Camps, A. Borges, M. Martin-Neira, J. Boutin, N. Reul, Y. H. Kerr, A. Hahne, and S. Mecklenburg, "SMOS: The challenging sea surface salinity measurement from space," *Proc. IEEE*, vol. 98, no. 5, pp. 649–665, May 2010.

[6] Y. H. Kerr, P. Waldteufel, J. P. Wigneron, S. Delwart, F. Cabot, J. Boutin, M. J. Escorihuela, J. Font, N. Reul, C. Gruhier, S. E. Juglea, M. R. Drinkwater, A. Hahne, M. Martin-Neira, and S. Mecklenburg, "The SMOS mission: New tool for monitoring key elements of the global water cycle," *Proc. IEEE*, vol. 98, no. 5, pp. 666–687, May 2010.

[7] C. J. Banks, C. P. Gommenginger, M. A. Srokosz, and H. M. Snaith, "Validating SMOS ocean surface salinity in the Atlantic with Argo and operational ocean model data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1688–1702, May 2012.

[8] S. Caires and A. Sterl, "Validation of ocean wind and wave data using triple collocation," *J. Geophys. Res.*, vol. 108, no. C3, pp. 3098–3114, 2003.

[9] A. Chakraborty, R. Kumar, and A. Stoffelen, "Validation of ocean surface winds from the OCEANSAT-2 scatterometer using triple collocation," *Remote Sens. Lett.*, vol. 4, no. 1, pp. 85–94, Jan. 2013.

[10] N. Agarwal, R. Sharma, S. Basu, A. Sarkar, and V. K. Agarwal, "Evaluation of relative performance of QuikSCAT and NCEP re-analysis winds through simulations by an OGCM," *Deep Sea Res. I., Oceangr. Res. Paper*, vol. 54, no. 8, pp. 1311–1328, 2007.

[11] B. Subrahmanyam, G. Grunseich, and E. S. Nyadjro, "Preliminary SMOS salinity measurements and validation in the Indian Ocean," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 19–27, Jan. 2013.

[12] N. Reul, J. Tenerelli, J. Boutin, B. Chapron, F. Paul, E. Brion, F. Gaillard, and O. Archer, "Overview of the first SMOS sea surface salinity products. Part I: Quality assessment for the second half of 2010," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1636–1647, May 2012.