Springer Earth System Sciences

Diego Fernández-Prieto
Roberto Sabia *Editors*

# Remote Sensing Advances for Earth System Science

## The ESA Changing Earth Science Network: Projects 2011–2013

Springer

# Springer Earth System Sciences

More information about this series at http://www.springer.com/series/10178

Diego Fernández-Prieto · Roberto Sabia
Editors

# Remote Sensing Advances for Earth System Science

## The ESA Changing Earth Science Network: Projects 2011–2013

Springer

*Editors*
Diego Fernández-Prieto
European Space Agency, ESA-ESRIN
Frascati, Rome
Italy

Roberto Sabia
Telespazio-Vega for European Space Agency, ESA-ESTEC
Noordwijk, Zuid-Holland
The Netherlands

ISSN 2197-9596 ISSN 2197-960X (electronic)
Springer Earth System Sciences
ISBN 978-3-319-16951-4 ISBN 978-3-319-16952-1 (eBook)
DOI 10.1007/978-3-319-16952-1

Library of Congress Control Number: 2015950893

Springer Cham Heidelberg New York Dordrecht London


Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

# Preface

## The Changing Earth Science Network Projects 2011–2013

To better understand the various processes and interactions that govern the Earth system and to determine whether the recent human-induced changes could ultimately destabilise its dynamics, both the natural system variability and the consequences of human activities have to be observed and quantified.

In this context, the European Space Agency (ESA) published in 2006 the document "The Changing Earth: New Scientific Challenges for ESA's Living Planet Programme" as the main driver of ESA's new Earth Observation (EO) science strategy. The document outlines 25 major scientific challenges covering all the different aspects of the Earth system, where EO technology and ESA missions may provide a key contribution.

In this framework, and aiming at enhancing the ESA scientific support towards the achievement of "The Challenges", the Agency has launched the "Changing Earth Science Network", an important programmatic component of the new Support to Science Element (STSE) of the Earth Observation Envelope Programme (EOEP). In this preface, the objectives of this initiative are summarised and the list of the projects selected in the second call of the programme is provided. An in-depth overview of such projects will be provided in the following book chapters.

## ESA EO Science Strategy and the Support to Science Element (STSE)

Since their advent, satellite missions have become central in the Earth monitoring and understanding, resulting in significant progresses in a broad range of scientific areas. Although the Earth has undergone significant changes in the past, there is mounting evidence that those occurring during the last 150 years are affecting the various interactions and processes among the different components of the Earth

system. Understanding those changes, their impacts on human lives and how anthropogenic activities affect the Earth system and its climate represent a major scientific endeavour where EO technology is already playing a key role.

In the mid-1990s, ESA set up its Living Planet Programme (LPP) working in close cooperation with the international scientific community to define, develop and operate focused satellite missions addressing some of the key questions at the core of Earth system science.

Moreover, realising the importance of further understanding the Earth and its response to these recent changes, the ESA published "The Changing Earth: New Scientific Challenges for ESA's Living Planet Programme" as the main driver of ESA's new EO science strategy. The document outlines 25 major scientific challenges faced today covering all the different aspects of the Earth system and climate (oceans, atmosphere, cryosphere, land surface, solid Earth), where EO technology and ESA missions may provide a key contribution namely:

The Challenges of the Oceans

1. Quantify the interaction between variability in ocean dynamics, thermohaline circulation, sea level, and climate.
2. Understand the physical and biochemical air/sea interaction processes.
3. Understand the internal waves and the mesoscale in the ocean, its relevance for heat and energy transport, and its influence on primary productivity.
4. Quantify the marine-ecosystem variability, and its natural and anthropogenic physical, biological and geochemical forcing.
5. Understand the land/ocean interactions in terms of natural and anthropogenic forcing.
6. Provide the reliable model- and data-based assessments and predictions of the past, present and future state of the ocean.

The Challenges of the Atmosphere

1. Understand and quantify the natural variability and the human-induced changes in the Earth's climate system.
2. Understand, model and forecast the atmospheric composition and air quality on adequate temporal and spatial scales, using ground-based and satellite data.
3. Better quantify the physical processes determining the life cycle of aerosols and their interaction with clouds.
4. Observe, monitor and understand the chemistry–dynamics coupling of the stratospheric and upper tropospheric circulations, and the apparent changes in these circulations.
5. Contribute to the sustainable development through interdisciplinary research on climate circulation patterns and extreme events.

The Challenges of the Cryosphere

1. Quantify the distribution of sea-ice mass and freshwater equivalent, assess the sensitivity of sea ice to climate change and understand thermodynamic and dynamic feedbacks to the ocean and atmosphere.

2. Quantify the mass balance of grounded ice sheets, ice caps and glaciers; partition their relative contributions to global eustatic sea-level change; and understand their future sensitivity to climate change through dynamic processes.
3. Understand the role of snow and glaciers in influencing the global water cycle and regional water resources, identify the links to the atmosphere and assess likely future trends.
4. Quantify the influence of ice shelves, high-latitude river run-off and land ice melt on global thermohaline circulation, and understand the sensitivity of each of these fresh-water sources to future climate change.
5. Quantify the current changes taking place in permafrost and frozen-ground regimes, understand their feedback to other components of the climate system and evaluate their sensitivity to future climate forcing.

The Challenges of the Land Surface

1. Understand the role of terrestrial ecosystems and their interaction with other components of the Earth system for the exchange of water, carbon and energy, including the quantification of the ecological, atmospheric, chemical and anthropogenic processes that control these biochemical fluxes.
2. Understand the interactions between biological diversity, climate variability and key ecosystem characteristics and processes, such as productivity, structure, nutrient cycling, water redistribution and vulnerability.
3. Understand the pressure caused by anthropogenic dynamics on land surfaces (use of natural resources, and land-use and land-cover change) and their impact on the functioning of terrestrial ecosystems.
4. Understand the effect of land-surface status on the terrestrial carbon cycle and its dynamics by quantifying their control and feedback mechanisms for determining future trends.

The Challenges of the Solid Earth

1. Identification and quantification of physical signatures associated with volcanic and earthquake processes—from terrestrial and space-based observations.
2. Improved knowledge of physical properties and geodynamic processes in the deep interior, and their relationship to Earth-surface changes.
3. Improved understanding of mass transport and mass distribution in the other Earth system components, which will allow the separation of the individual contributions and a clearer picture of the signal due to solid-Earth processes.
4. An extended understanding of core processes based on complementary sources of information and the impact of core processes on Earth system science.
5. The role of magnetic field changes in affecting the distribution of ionised particles in the atmosphere and their possible effects on climate.

To reinforce this strategy, in 2008, it was established the Support to Science Element—STSE (www.esa.int/stse), to provide scientific support for both future and on-going missions, by taking a pro-active role in the formulation of new

mission concepts and products, by offering support to the scientific use of ESA EO multi-mission data and promoting the achieved results.

In this Context, STSE main pillars aim at:

- Developing novel mission concepts in preparation for the next generation of European scientific missions;
- Developing advanced algorithms and innovative products that exploit the increasing ESA multi-mission capacity;
- Reinforcing ESA collaboration with the major international scientific programmes and initiatives in Earth system sciences;
- Support the Next Generation of Earth System European Scientists (The Changing Earth Science Network).

## The Changing Earth Science Network

As one of the main programmatic components of the STSE, ESA launched in 2008 a new initiative—the Changing Earth Science Network—to support young scientists to undertake leading-edge research activities contributing to achieve the 25 scientific challenges of the LPP by maximising the use of ESA data.

The initiative is implemented through a number of research projects proposed and led by early-stage scientists at postdoctoral level for a period of two years. Projects undertake innovative research activities furthering into the most pressing issues of the Earth system, while exploiting ESA missions data with special attention to the ESA data archives and the new Earth Explorer missions.

Specifically, the Initiative Aims at:

- Contributing to the scientific advancement in Member States towards the achievement of the new 25 strategic challenges of the LPP;
- Fostering the use of ESA EO data by the Earth Science community maximising the scientific return (in terms of scientific results and publications) of ESA EO missions;
- Contributing to consolidate a critical mass of young scientists in Europe with a good scientific and operative knowledge of ESA EO missions, assets and programmes;
- Promoting the development of a dynamic research network in ESA Member States addressing key areas of relevance for ESA missions and the ESA science strategy;
- Enhancing interactions, exchanging know-how and allowing cross fertilisation between ESA and Earth science laboratories, research centres and universities.

The first call for proposals, issued in 2008, resulted in the selection of 11 postdoctoral scientists from the Agency's Member States based on the scientific merit of the individual projects. A second call for proposals was issued in early

2010 to be implemented between 2011 and 2013, resulting in a further selection of 10 leading-edge research activities. New calls took place in early 2012 and 2014.

This volume collects some of the results obtained by eight of the second set of projects started in 2011 and completed in 2013. They describe research activities exploiting data coming from several remote sensors on-board a wide suite of ESA and non-ESA satellites. In summary, the projects described in the following provide cutting-edge advanced exploitation of satellite data relevant to a broad range of scientific applications, towards an improved monitoring of the integrated Earth system.

| Acronym | Full project title | Researcher | Institute |
|---|---|---|---|
| CHIMTEA | Chemical Impact of Thunderstorms on Earth's Atmosphere | Enrico Arnone | Istituto di Scienze dell'Atmosfera e del Clima, ISAC-CNR, Bologna, Italy |
| TIBAGS | Tropospheric Iodine Monoxide and Its Coupling to Biospheric and Atmospheric Variables—A Global Satellite Study | Anja Schönhardt | Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany |
| GreenSAR | Greenland and Antarctic Grounding Lines from SAR Data | Noel Gourmelen | School of GeoSciences, University of Edinburgh, UK |
| MESO3D | Sensor Synergies for Studies of Mesoscale and Sub-Mesoscale Ocean Dynamics | Nicolas Rascle | Laboratoire d'Océanographie Spatiale, IFREMER, Plouzané, France |
| SMOSPROC | Study of Ocean Surface Processes and Their Impact on the Retrievals of Salinity from SMOS | Kieran Walesby | National University of Ireland, Galway |
| SMASPARES | SMOS Data Assimilation for Parameter Estimation in Radiative Transfer Models | Carsten Montzka | Institute of Bio- and Geosciences: Agrosphere (IBG-3), Forschungszentrum Jülich, Jülich, Germany |
| PROgRESSIon | PROtotyping the Retrievals of Energy Fluxes and Surface Soil Moisture | George P. Petropoulos | University of Aberystwyth, Aberystwyth, UK |
| GEMMA | Crustal Modelling and Moho Estimation with GOCE Gravity Data | Daniele Sampietro | Politecnico di Milano—Polo Territoriale di Como, Como, Italy |

Diego Fernández-Prieto
Roberto Sabia

# Acknowledgments

The authors of this preface wish to thank all the different institutes and laboratories that have supported the initiative as well as all ESA colleagues acting as internal contact point that have contributed to this with their expertise and dedication.

# Contents

# CHIMTEA—Chemical Impact of Thunderstorms on Earth's Atmosphere

**Enrico Arnone and Bianca Maria Dinelli**

**Abstract** Since their accidental discovery in the 1990s, lightning-related sprites, other transient luminous events (TLEs), and terrestrial gamma-ray flashes have shown us how the impact of thunderstorms extends from the troposphere up to the upper atmosphere and ionosphere. Thunderstorms are a key player for the climate system, in particular through lightning-produced $NO_x$ and troposphere–stratosphere exchange. The CHemical Impact of Thunderstorms on Earth's Atmosphere (CHIMTEA) project focused on TLE-producing thunderstorms and their possible impact on stratospheric $NO_x$ and ozone. The distribution and seasonal cycle of thunderstorm activity were studied through global lightning data and TLE observations over Europe. Michelson Interferometer for Passive Atmosphere Sounding (MIPAS)/Environmental Satellite (ENVISAT) measurements of $NO_x$, ozone, and other related constituents from the upper troposphere to the mesosphere were analyzed with a 2D tomographic approach to quantify thunderstorm-induced changes and explore how to improve their detectability. The study included observations from Global Ozone Monitoring by Occultation of Stars (GOMOS)/ENVISAT, other satellites, and in situ measurements. The sensitivity of the measurements to sprite-$NO_x$ was investigated through ad hoc radiative transfer simulations quantifying reference thresholds. Global and regional observations showed sprite-$NO_x$ to be at the edge of current detectability, with no detectable impact on ozone. Model simulations were performed including for the first time a sprite-$NO_x$ parameterization in the Whole Atmosphere Community Climate Model (WACCM): it was shown that sprites may contribute significantly to tropical $NO_x$ in the middle mesosphere and reach detectable levels above particularly active thunderstorms. Extension of the adopted strategy to study lightning-$NO_x$ was recommended, whereas the modeling and multi-satellite approach was shown to be suitable in support to the upcoming space missions.

E. Arnone (✉) · B.M. Dinelli
Istituto di Scienze dell'Atmosfera e del Clima ISAC-CNR, Bologna, Italy
e-mail: e.arnone@isac.cnr.it



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_1

# 1 Introduction

Clouds are a core component of the Earth's climate system and one of the key shortages in our understanding of a changing climate (IPCC, 2013, www.ipcc.ch). Thunderclouds, in particular, affect the dynamics and chemistry of the atmosphere through convective transport and electrification processes. Beside their severe action at the surface, they impact the interface between the troposphere and the stratosphere, one of the most delicate regions of coupling between chemistry and climate, both dynamically and through lightning-produced $NO_x$. The detection of upper atmosphere lightning and gamma energy emissions from thunderstorms has further highlighted the relevance of thunderstorms also at higher altitude.

## *1.1 Lightning in the Upper Atmosphere*

Since the 1990s, a whole family of upper atmosphere electrical processes have been discovered to occur above thunderstorms: they are known as transient luminous events (TLEs) in their optical and low energy manifestation, and terrestrial gamma-ray emissions (TGFs) in their high energy component (e.g., [18, 19]). In particular, TLEs are produced by the electrical impact of thunderstorms on the above atmosphere, which causes ionization, dissociation, and excitation of neutral air constituents and the consequent well-recognizable optical emissions. They occur in the stratosphere–mesosphere between the top of thunderclouds and the lower ionosphere, the altitude of occurrence determining both their nature and the impact they exert. Above thunderclouds, the formation of streamers (weakly ionized plasma channels) can occur roughly up to 70 km altitude, above which dielectric relaxation timescales become comparable with that of dissociative attachment, leading to diffuse emissions (e.g., [19]). TLEs such as blue jets and gigantic jets are streamer–leader processes injected from thundercloud tops toward the ionosphere and may be considered the upward equivalent of cloud-to-ground (CG) lightning [16]. Sprites [23] are luminous discharges that initiate at about 70–80 km altitude, extend downward to 40 km as streamers and upward to 90 km altitude as diffuse emission, at times in the form of a halo, and are tens of kilometers wide. Elves are diffuse emission rings manifesting the impact of the electromagnetic pulse propagated from a CG on the lower edge of the ionosphere (e.g. [19]). Depending on the relaxation timescales at the altitude of occurrence, TLEs last a few to a few hundred milliseconds. Their continuous occurrence around the globe makes them however a relevant and as yet not considered component of the atmosphere.

### 1.2 Chemistry of Thunderstorms and Transient Luminous Events

In the upper troposphere, lightning is a dominant source of active nitrogen oxides $NO_x$ (N + NO + $NO_2$), which may consequently be transported into the lower stratosphere with tropospheric pollutants by the slow meridional circulation or by deep convection (e.g., [22]). This source contributes to the main production of $NO_x$ by oxidation of tropospheric $N_2O$, and therefore to catalytic cycles affecting ozone (e.g., [9]). Despite several lightning-$NO_x$ estimates, observation of lightning $NO_x$, and in particular of its transport into the stratosphere, remains a challenge for current satellites: this is due to a limited sensitivity to upper troposphere $NO_x$ species both for limb sounders (efficient in the stratosphere) and nadir sounders (strongly weighted by the tropospheric component), and the relatively small size of individual events. Further limitations arise from the reconversion of lightning-$NO_x$ into $HNO_3$ with consequent loss through scavenging, from high dynamical variability of the upper troposphere–lower stratosphere region, and from the large variability among thunderstorm events [8, 22].

In analogy to tropospheric lightning, TLEs were predicted to impact the atmospheric chemistry. Enell et al. [13], Sentman et al. [24], Gordillo-Vazquez [14], and Winkler and Notholt [25] estimated sprite-induced $NO_x$ enhancements within sprite streamers between a few to a few hundreds of percent, with negligible ozone changes. Hiraki et al. [15] estimated orders of magnitude increases in $NO_x$ and significant ozone changes. Two observational studies investigated sprite-induced perturbations with satellite measurements. Arnone et al. [5–7] found a possible sprite-induced $NO_2$ perturbation of 10 % at 52 km altitude and of tens of percent at 60 km altitude in coincidence of active thunderstorms, and no evident sprite global impact. Rodger et al. [21] concluded that TLEs occurring below 70 km altitude exert no significant impact on the neutral chemistry at a global scale. An impact at local scale but negligible at global scale was also suggested by the first laboratory experiments conducted by Peterson et al. [20], although they were criticized for the poor similarity to real TLE conditions.

This paper presents an overview of the CHIMTEA project, which was developed using several datasets (Sect. 2), to improve our understanding of lightning processes in the upper atmosphere (Sect. 3), and investigate the current sensitivity of observations (Sect. 4) and models (Sect. 5) to describe their chemical impact. Concluding remarks and recommendations from the project are given (Sect. 6).

## 2 Instruments and Data

We used lightning data from several sources: World Wide Lightning Location Network (WWLLN), lightning imaging sensor (LIS), and the VLF/LF lightning detection network (LINET). WWLLN is an experimental global lightning very

low-frequency (VLF) network (http://www.wwlln.com) with detection efficiency of about 10 % for CGs and 5 % for intracloud (IC) lightning. Climatological lightning data from the optical transient detector (OTD)/LIS satellite instruments [11] was used for the analysis of observations and model simulations (http://thunder.msfc.nasa.gov). LINET data was available at ISAC: it works in the VLF/LF (low-frequency) range with 3D capability and allows for unprecedented low-amplitude detection power, discerning between CG and IC (http://www.pa.op.dlr.de/linet). TLE data were collected with tens of low-light sensitive cameras around Europe by the *EuroSprite* network ([18], http://www.electricstorms.net). Observations are optical images, which define the "truth" for TLE measurements, and additional radio and infrasound data. The first database of European TLEs was created within CHIMTEA ([2], ACP, *under submission*).

Atmospheric limb observations of the stratosphere–mesosphere (6–70 km altitude) were adopted from the MIPAS2D database [12], obtained from MIPAS/ENVISAT mid-infrared spectroscopic measurements retrieved with the geo-fit multitarget retrieval (GMTR) 2D tomographic code [10]. The 2D approach enables to model horizontal atmospheric inhomogeneities. MIPAS2D data quality was discussed in Dinelli et al. [12]. Pressure, temperature, and ozone-related species and $NO_y$ family are covered by MIPAS2D results (see further details in [1, 4]). Data from the stellar occultation spectrometer GOMOS/ENVISAT, from the microwave limb sounder (MLS)/Aura microwave thermal emission limb sounder, from sondes, and from the SMOS satellite were used in addition.

We performed TLE-$NO_x$ simulations on the WACCM [17], which spans from the Earth's surface to the thermosphere and includes relevant natural processes: the adopted v.4 has updated chemistry, solar proton events, improved stratospheric warming, gravity waves by convection and fronts, and parameterized lightning-$NO_x$. The model was run with specified dynamics, which is controlled by nudging temperature and winds below 60 km to specified observational fields, with nudging strength reduced between 50 and 60 km and free running above 60 km. Simulations were run starting on March 2, 2011, with control simulations run for 60 days, perturbed ones until they converged. Meteorological data from European Center for Medium-Range Weather Forecasting (ECMWF) both ERA-Interim (reanalysis) and from operational processing were used both for comparison and scientific analysis.

## 3 TLE-Producing Thunderstorms

A first climatological study of TLEs over Europe was conducted within CHIMTEA coordinating European observations and releasing the first EuroSprite 2009–2012 summary of TLE observations over Europe and the Mediterranean ([2], *under submission*). The occurrence of both summer and winter thunderstorms makes Europe an ideal region for studies of TLEs under seasonally varying conditions

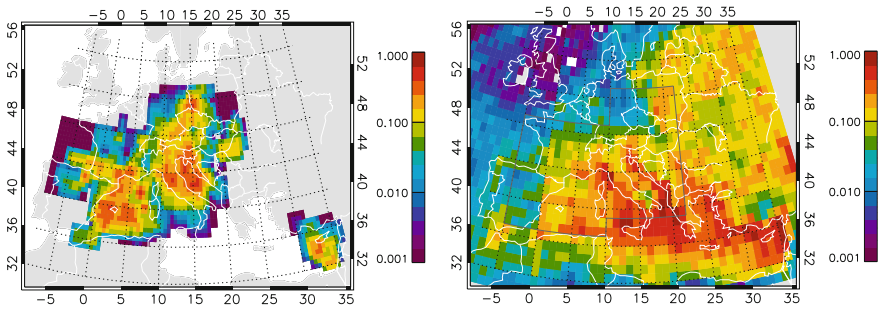


**Fig. 1** *Left* Climatology of observed TLEs (TLEs $10^{-3}$ km$^{-2}$ year$^{-1}$) for 2009–2012. *Right* Climatology of lightning strokes detected by WWLLN (strokes km$^{-2}$ year$^{-1}$) for 2009–2012. The *grey rectangular shapes* delimit the approximate area of TLE observations used in monthly mean calculations of WWLLN data

and a preferential region for the following chemistry part of the study. Over 2009–2012, the number of optical observations of TLEs exceeded 1000 per year. Because of this unprecedented number of observations, it was possible to construct a climatology of over 6000 TLEs observed above 756 thunderstorms. The number of TLEs per thunderstorm was found to follow a power law, with less than 10 TLEs for 567 thunderstorms and up to 147 TLEs above the most prolific one. The vast majority of TLEs were classified as sprites, 294 as elves, 185 halos, 57 upward lightning, 2 blue jets, and 1 gigantic jet. This number of observed sprites largely exceeds that acquired from satellites over equivalent periods of time. A key shortage of the adopted dataset is the inhomogeneity of the observational coverage. The seasonal changes in the climatology were however extracted weighting each season by the total yearly observations. Figures 1 and 2 report the TLE climatology in terms of geographical distribution and seasonal cycle. For comparison, the same diagrams were replicated with a climatology of lightning data from WWLLN. In the

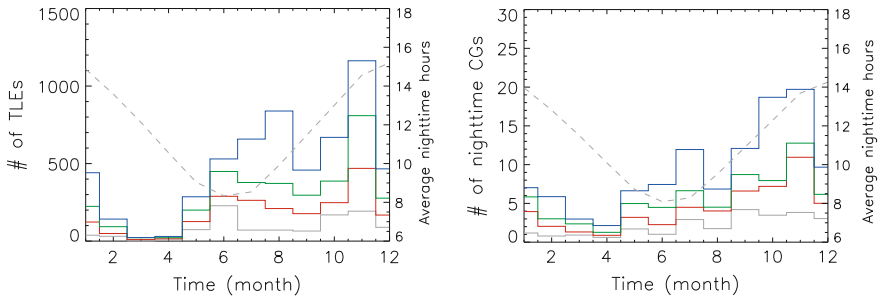


**Fig. 2** Cumulative number of observed TLEs (*left*) and WWLLN CG strokes (*right*) per month including data from the start of the sample (January 2009) to the end of 2009, 2010, 2011 and 2012 (respectively, in *grey*, *red*, *green* and *blue*). The average seasonal cycle of nighttime hours is shown (*dashed grey*). WWLLN data are scaled by $10^4$ and averages calculated over the rectangular shapes shown in Fig. 1

areas covered by the observations, TLE activity in Europe is intense over the Southern regions and Mediterranean Sea. The largest number of TLEs per month is recorded in November, whereas in March and April TLE activity is almost completely halted. This is due to a shift from continental areas in summer to coastal areas and sea in autumn. The peak in November can be reconciled with the number of observations in summer once the change in length of the night is taken into account (see dashed line in the figures), although the larger +CGs/−CGs ratio in autumn/winter maritime thunderstorms should be considered. Elves occur exclusively over autumn/winter maritime thunderstorms. The overall behavior is well consistent among individual years, therefore making the observed seasonal cycle a robust general feature of TLE activity over Europe. The 756 TLE-producing thunderstorms were then used as case studies for several papers and for chemistry–TLE correlations in the next section.

Lightning data from the WWLLN network were used to trace thunderstorm activity. Due to the inhomogeneity of WWLLN detection efficiency, which depends on the distance from the detector, we focused the analysis of WWLLN data over Europe where the available detectors guarantee a satisfactory coverage allowing to study the distribution and seasonal changes of thunderstorm activity. At global scale, WWLLN data were used for correlation with MIPAS measurements (see Sect. 4), assuming WWLLN lightning detections as proxy of thunderstorm rather than individual lightning strokes. Within the regions covered by the observations, the TLE main geographical distribution tends to mimic the distribution of thunderstorm activity, occurring over thunderstorms with an incredible variety of characteristics, from large summer thunderstorms over the continental areas to tiny winter thunderstorms over the sea. This is evident in the northern part of the TLE climatology where large areas with a weak TLE rate are consistent with a drop in lightning activity, e.g., in France and Germany, and partially in Spain contrasting with the adjacent high activity over the Pyrenees. In contrast, the fading of TLE activity in Southern Italy is due to poor coverage. The poor coverage appears to cause similar low rates over Corsica, North-West Italy, and Hungary, where local cameras tend to observe only thunderstorm at a certain distance from the observation spot. Peak TLE rate exceeds $10^{-3}$ km$^{-2}$ year$^{-1}$ in a few hotspots in the Balearic Islands, Italy, and Poland, whereas it is typically around 0.2–0.3 in large adjacent regions. This rate should be compared to lightning rates around 0.2–0.3 within the same regions, pointing to a factor of 1000 in the observed lightning/TLE ratio.

Correlation studies between lightning/TLE activity and atmospheric and surface parameters were performed in order to characterize the observed distributions also for parameterization in models. Surface temperature data for continent and sea areas were investigated showing no simplified correlation with lightning over Europe or globally could be found. Soil moisture (SM) and dielectric constant (DC) data measured by the SMOS satellite were also investigated. A correlation was performed in order to seek a possible bias in lightning activity (either in its production or detection by ground networks) consistent with the availability of soil moisture and soil conductivity. The latter may lead to the use of satellite data to improve prediction of thunderstorm occurrence, without relying on standard parameters such

as convective available potential energy and vertical wind shear. Comparison of SM and DC maps with those of lightning activity (not shown) highlighted a strong variability in SMOS data and difficulties in identifying clear patterns which could be reconciled among the different parameters. Averages over the European continent showed a seasonal evolution of SMOS 2011 data for SM and DC imaginary component with a minimum in summer, and maximum in winter, with no reproduction of the seasonal cycle observed in lightning activity. On the contrary, the DC real component averaged over Europe and the Mediterranean Sea showed a seasonal cycle similar to that of lightning activity. However, once the data was limited to the region of EuroSprite coverage, the temporal behavior changed significantly and the possible correlation was largely lost. This lack of consistency showed the complex interplay of different sources of soil moisture, which integrates precipitation from all cloud types and can therefore bury information on convection within its variability. If a correlation exists, this should be further studied adopting a model approach which incorporates all relevant parameters to describe lightning production.

## 4 Measuring TLE- and Lightning-$NO_x$ from Space

Comparison of MIPAS2D results with MLS, GOMOS, and ECMWF was performed on key regions and seasons. In particular, GOMOS and MIPAS2D $NO_2$ showed a fair agreement in the low to mid-latitude stratosphere, supporting the use of MIPAS2D data to study the impact of TLEs in the range of 50–0.2 hPa, with random errors on $NO_2$ up to 0.7 ppbv at and below 50 km altitude, and around 1–1.3 ppbv at 60 km altitude. A MIPAS2D–GOMOS analysis was performed on the variability of NOy in 2002–2012 [1]. A correlation with solar cycle and geomagnetic activity was found at high-latitude stratopause and lower stratosphere, down to mid-latitude. MIPAS2D data under perturbed conditions were also investigated in the extreme 2011 Arctic vortex [4]. At the low to mid-latitude stratopause, $NO_2$ could be clearly matched to its $N_2O$ source, tracing the occasional changes due to dynamical year-to-year variability (e.g., during quasi-biennial phases): therefore, a dominant influence of TLE-$NO_2$ can be excluded and only secondary contributions should be expected.

We performed a series of radiative transfer simulations mimicking TLE-$NO_2$. The unperturbed atmosphere was from ECMWF and MIPAS data with 1.125° horizontal and 1 km vertical steps, smoothed and with cloud interferences. Sprite perturbations were applied at various latitudes, at 45–90 km altitude, and over a few hundred kilometers. Figure 3 shows retrieved versus reference $NO_2$ for perturbations of +1 to +10 ppbv. The simulations at 60 km altitude show 1–2 ppbv biases and oscillations induced by the retrieval. This uncertainty drops to below 1 ppbv when the sprite perturbation occurs over high background. At 52 and 47 km altitude, retrievals pick also at the +1 ppbv perturbation, with a bias of less than 0.5 ppbv. A low altitude TLE scenario was analyzed with perturbations at 15–45 km altitude. The results

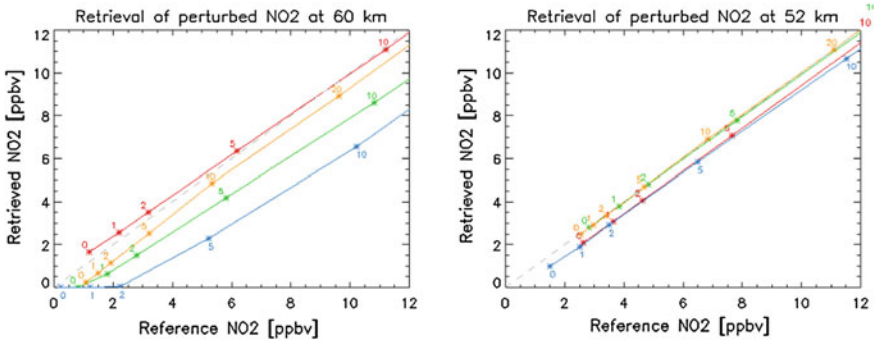


**Fig. 3** Retrieved versus reference $NO_2$ for simulations having perturbed $NO_2$ at 45–90 km altitude. Results are shown at the location of the three sprite-perturbations (colors are latitude 32° S, 9° S, 34° N, the latter having two perturbed grid points) and at altitude 60 (*left*) and 52 (*right*) km, limited to the low magnitudes (see labels)

show a sensitivity to changes of 0.5 ppbv at 36–42 km, and 1 ppbv at 21–33 km altitude. The 18 km altitude still shows a close-to-linear response to the magnitude of the perturbation, although with larger uncertainties.

We also performed simulations with $NO_2$ perturbed at 10–22 km altitude mimicking the effects of lightning or thunderstorm-driven transport. Our simulations show that at 15 km altitude the large amount of $NO_2$ retrieved in some orbits is due to artificial oscillations induced by the retrieval. At 18 km altitude, oscillations are typically of the order of 1 ppbv, but can reach 2 ppbv, and drop to 0.5–1.0 ppbv at 21 km, values below which real perturbations cannot be identified. The retrieval greatly improves at low altitude when the validity of the adopted microwindows is artificially extended down to 9 km. In this case, the insensitive level is shifted to the lowermost 9 km altitude, and the other low altitude levels have sensitivity to at least large amounts of $NO_2$ (10 ppbv) with sufficient accuracy. When further performing simulations removing clouds, the improvement was substantial in terms of retrieved lightning-perturbed $NO_2$ and reduced oscillations. Results of the above simulations showed sensitivity down to 1 ppbv at 60 km, of about 0.5 ppbv at 52 km, and 1–0.5 ppbv at 18–21 km altitudes. Given the presence of background oscillations, this sensitivity can be fully exploited only by averaging a large number of observations in order to reduce the background variability. However, the simulations showed the response of MIPAS $NO_2$ measurements to small changes is largely linear once this minimal amounts are reached, so that if enough $NO_2$ is already present, MIPAS has sensitivity to smaller perturbations.

## *4.1 Thunderstorm Signatures in MIPAS2D $NO_2$*

We sought sprite perturbations above thunderstorms correlating MIPAS2D nighttime $NO_2$ with WWLLN lightning, extending [6, 7]. This approach allowed us to

have a search at global scale both for TLE effects and other thunderstorm impacts above the upper troposphere. The coincidence window was chosen to be 60 min prior to MIPAS measurement, with horizontal dimensions of 500 km along latitude and 60 km along longitude, and a threshold of 10 WWLLN counts. Results by Arnone et al. [6, 7] over the period August–December 2003 showed a perturbation peak in the thunderstorm-$NO_2$ of about 1 ppbv (10 %) at 52 km altitude. An extension of this analysis up to April 2004 confirmed a significant change over latitudes 0–20 North (2–5 % at 52 km), with no significant change in ozone or other species. The analysis was extended to the complete dataset of MIPAS mission using MIPAS2D GRD 2.3. Comparison of WWLLN-$NO_2$ and background $NO_2$ at various altitude showed no significant perturbation of the distributions (e.g., with anomalous peaks) or a change to the mean or median of the distribution. No significant changes were found in other species. We conclude that there is no robust evidence of thunderstorm-induced effects on the adopted MIPAS2D dataset, and therefore perturbations may be below 0.05 ppbv at 52 km altitude. The adopted statistical method based on global detections of lightning is likely missing a large fraction of the MIPAS coincidences. Moreover, the adopted GRD retrieval method is partly smearing $NO_2$ perturbations, and most importantly changing the location of the retrieved profile as compared to the original observation. The analysis will need to be repeated with a version of MIPAS2D suited for studying localized changes.

An analysis was performed on chemistry measurements above individual thunderstorms with documented TLE occurrence by the EuroSprite network (see Sect. 3). MIPAS profiles from 2009 to the end of 2011 were inspected in order to find measurements that sounded the atmosphere above the TLE-active thunderstorms. Despite the large number of TLE-producing thunderstorms observed, no robust coincidence was found with MIPAS sounding, with only 24 close matches (i.e., with spatial distances below 100 km and perfect time match). None of the thunderstorm-$NO_2$ showed a significant change as compared to the background $NO_2$. Because of the lack of a perfect match, this search cannot however be conclusive in constraining the $NO_2$ production by individual TLE-producing thunderstorms.

A further global correlation study was performed with TGF-producing thunderstorms to identify chemical changes induced by high energy emissions or by their low energy tail. The analysis was performed on 11,040 Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI) detections in 2002–2011. The coincidence window was ±3 h and ±300 km, taking into account uncertainty on the geolocation of the TGF footprint and evolution of the thunderstorm. MIPAS $NO_2$ data at 20–60 km altitude in coincidence with TGF activity showed no consistent enhancement as compared to background $NO_2$, therefore pointing to a lack of major regional impact of TGFs on the chemistry of the stratosphere above thunderstorms. A few individual cases of high $NO_2$ were associated to retrieval of poor data. Small $NO_2$ changes, or highly localized changes which were missed by MIPAS may not be excluded.

## 5 Climate-Chemistry Sensitivity to TLE-$NO_x$

Simulations were performed including for the first time a sprite–$NO_x$ parameterization in the WACCM [3]. A simplified parameterization of sprite-$NO_x$ was calculated using vertical profiles from Enell et al.'s [13] ion-neutral chemistry model. Climatological fields of lightning activity from the LIS/OTD [11] were used as a proxy of sprite activity with a rate of 1 sprite every 1,000 lightning flashes (leading to 2–3 sprites/minute globally). The simulations started at magnitude 10 and 100 times the reference sprite-$NO_x$ climatology to produce a sprite-$NO_x$ dominated scenario, followed by lower magnitude perturbations, both constant in time and with a switch-off after 20 days.

Investigation of the climate-chemistry sensitivity of the atmosphere to sprite-like perturbations was performed at varying magnitude, time, and space constraints, allowing the quantification of a threshold for reaching significant changes. Results for the 10x case study are shown in Fig. 4 (with red indicating peak values of 2 ppbv, or about 400 % increase). Horizontal plumes of enhanced $NO_x$ originating above the Americas, Central Africa, and the Maritime Continent buildup and expand to cover much of the tropics. Over the 40 days of the simulation, the enhanced $NO_x$ fills regions of very low background $NO_x$. Tongues of downward transported $NO_x$ can interfere with the sprite-$NO_x$ perturbation at Northern latitudes. Peak values of up to a factor 4 (a factor 20 in the x100 case) enhancement are reached within a couple of weeks above Africa and up to a factor 3 (a factor 10 in the x100 case) above the Maritime Continent, at 66–80 km altitude, down to tens of percent at 60 km. Results for the simulations with the reference 1x perturbation at

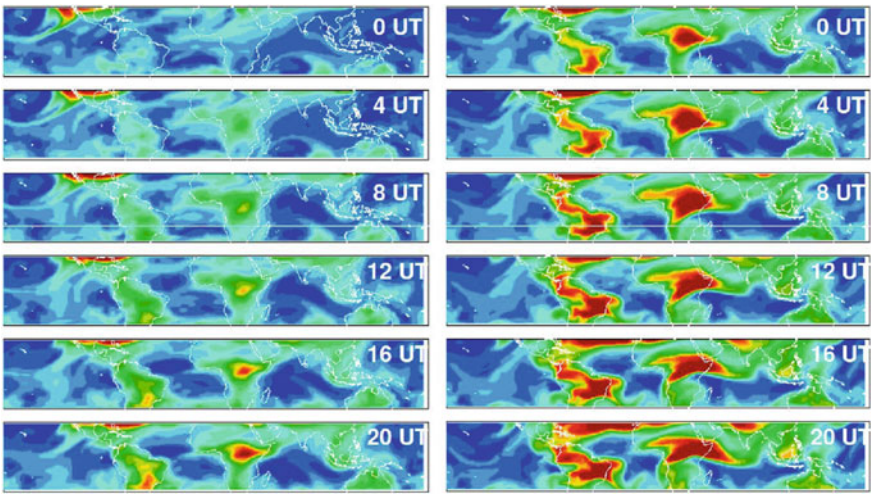


**Fig. 4** WACCM response to sprite-$NO_x$ (x10) at 70 km altitude during day 1 (*left column*) and day 2 (*right column*) of the perturbations, at 4 h time intervals

0.05 hPa (about 70 km) are shown in Fig. 5 as difference between the perturbed (dashed lines) and control simulations. The difference to the control run is significant, and saturates within the tropics at about +0.2 ppbv of $NO_x$ more than the control simulation. Note that the variability of the tropical averaged $NO_x$ is around ±0.02 ppbv. The saturation of the sprite-$NO_x$ source is reached in about 15 days. Oscillations in the timeseries are due to air masses moving in and out of the averaging box region, and $NO_x$ responses building up differently depending on the specific background conditions. The enhancement is also seen in the zonal mean values at the tropics (black line), with a smoother steady increase. A test whereby the sprite-$NO_x$ source is ramped down after 20 days (bold lines) shows that the atmosphere relaxes back to background conditions in about 15 days, i.e., similarly as for the buildup time.

The results show that, without consideration of the competition of transport and sinks, previous attempts to estimate the relevance of sprite-$NO_x$ to the atmosphere at regional or global level are unreliable since the atmosphere needs to reach equilibrium with the sprite source. Our simulations point to a 0.2 ppbv (about 30 %) contribution to Tropical $NO_x$ at 70 km altitude by sprites assuming [13] maximum case (our reference scenario), down to 0.02 ppbv buried within background variability assuming the larger of their typical case. This is above the variability we found for Tropical $NO_x$ at 70 km altitude (about 0.01 ppbv). Below these values the sprite contribution to Tropical $NO_x$ becomes irrelevant. The relaxation timescales for Tropical $NO_x$ at 70 km altitude are of the order of 15 days, so that global effects of sprite-$NO_x$ will be integrated over such time periods even though they occur impulsively.

We also used model results to interpret available observational studies. Considering periods of maximum buildup under slow transport, we can expect peaks of tens of percent change in $NO_x$ at 60–85 km altitude. At this magnitude,

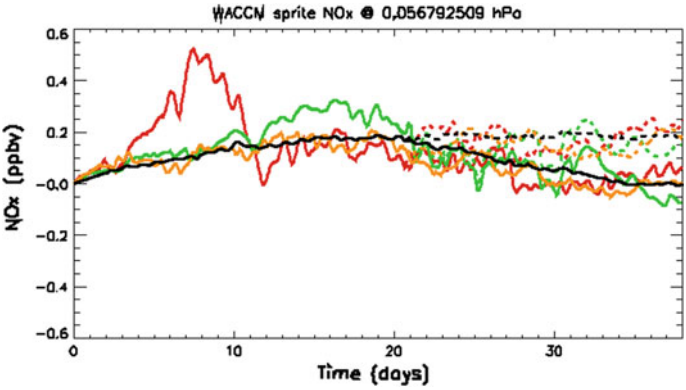


**Fig. 5** WACCM response to sprite-$NO_x$ perturbations in the case 1x (*bold*) and in the case 1x with rampdown after 20 days (*dashed*). Results are shown as difference between the perturbed and reference simulation above target regions South America (*yellow*), Central Africa (*red*), South East Asia (*green*) and Tropics (*black*)

sprite-$NO_x$ have detectable levels under favorable conditions. This confirms what is suggested by the analysis of satellite observations of $NO_x$ by Arnone et al. [6, 7], Rodger et al. [21], and Arnone et al. [5], with significant impact at local scale and negligible at global scale. At this magnitude of perturbation, any impact on the atmosphere would be negligible (see, e.g., [7]). Ad hoc calculations to resemble the method applied by Rodger et al. [21] showed their use of partial column should be limited to 65–80 km altitude in order to have sensitivity to sprite-$NO_x$.

At lower altitudes, it is unlikely that a global effect can be reached by low altitude TLEs such as blue jets because of the much larger number density. However, local effects under favorable transport conditions may not be excluded. Future work will consider the impact of the complete TLE family. On the other hand, the results of our study at sprite altitude are valid for any other kind of thunderstorm-induced $NO_x$ perturbation over the same altitude range, which includes other TLEs, TGFs, or other thunderstorm-induced processes in the lower mesosphere.

## 6 Conclusions and Future Lines

CHIMTEA activities have focused on TLE-producing thunderstorms and their impact on stratospheric $NO_x$ and ozone. The global distribution of thunderstorm activity was traced using global lightning data. Coordination and analysis of an unprecedented number of *EuroSprite* observations produced the first climatology of TLEs over Europe and several case studies. The coordinated European observations were adopted as ground support by the Japanese JEM-GLIMS mission launched in 2012 and will become a core element in ESA 2016 ASIM mission.

The sensitivity of MIPAS/ENVISAT observations to stratospheric variability was compared to extreme conditions occurring under solar proton events and winter polar vortex. The sensitivity to thunderstorm-perturbed $NO_2$ at various altitude was investigated with radiative transfer simulations finding minimum thresholds of 1 ppbv at 60 km altitude, 0.5 ppbv at 52 km and in the range 18–21 km. These thresholds make it largely unfeasible to use MIPAS for robust lightning-$NO_x$ detection. Despite the large MIPAS2D dataset and a significant number of thunderstorms with detected TLE or TGF activity, no robust evidence for thunderstorm-perturbed $NO_x$ was found. Global and regional observations showed sprite-$NO_x$ to be at the edge of current detectability, with no detectable impact on ozone. Strategies for improving sprite-$NO_x$ detectability include the use of refinements in the adopted retrieval and satellite measurements having higher spatial resolution.

Simulations were performed including for the first time a sprite-$NO_x$ parameterization in a global circulation model. Investigation of the climate-chemistry sensitivity to sprite-like perturbations was performed at varying magnitude, time, and space constraints, allowing the quantification of a threshold for reaching significant changes. On the basis of the current knowledge of sprite chemistry, it was

shown that sprites may be able to contribute significantly to tropical $NO_x$ in the middle mesosphere, with a spread component up to 0.2 ppbv (30 %) at 70 km altitude, and may reach detectable levels above particularly active thunderstorms. TLE perturbations are therefore likely to have an atmospheric impact, although at or below the edge of current observational capabilities. A key issue is being able to find a perfect match between TLE activity and chemistry measurements, which has to date been missed. Extension of the adopted strategies to study lightning-$NO_x$ is recommended, whereas the modeling and multi-satellite approach was shown to be suitable in support to the upcoming space missions.

**Acknowledgements** E. Arnone is pleased to acknowledge collaborations: MIPAS2D with Dinelli, Castelli, ISAC-CNR, Papandrea, Carlotti, Univ. Bologna, Italy; WACCM with Smith (NCAR, CO, US), Kero, Enell (SGO, Finland); GOMOS with Hauchecorne, LATMOS, France; MLS by NASA. Thank ECMWF for their data. Lightning data provided and discussed with: LINET (Dietrich, Mugnai, Formenton, ISAC), EUCLID (via EuroSprite), WWLLN (Kero, SGO), LIS (NASA). MIPAS2D-lightning correlation by Kero (SGO). Cloud observations from Meteosat, with Sao Sabbas (INPE, Brazil), Soula (Obs. Midi-Pyrénées, France), Laviola (ISAC). SMOS data from Montzka and Rötzer (Forschungszentrum Jülich, Germany). TLE data from EuroSprite (http://www.electricstorms.net) with Neubert, Chanrion (DTU Space, Denmark), van der Velde (Univ. Politècnica de Catalunya, Spain), Soula (Obs. Midi Pyrénées, France), Bor (Geodetic Geophys. Res. Inst., Sopron, Hungary), Haldoupis (Univ. Crete, Greece). TLE data over Italy from the Italian Meteor and TLE Network (IMTN). TLE camera at ISAC observatory on Mt. Cimone with Prevedelli (Univ. Bologna) and support by Bonasoni and group (ISAC) and IMTN (Italy). RHESSI data with Gjesteland (Univ. Bergen, Norway). Support by ESA through funding, feedback by ESA contact Fehr, CESN coordination by Sabia and Fernandez, access to ESA data. Support by ISAC, through funding, hosting, and access to facilities and expertise.

## References

1. Arnone E, Hauchecorne A (2012) Stratosphere NOy species measured by MIPAS and GOMOS onboard ENVISAT during 2002–2010: Influence of plasma processes onto the observed distribution and variability. Space Sci Rev 168:315–332
2. Arnone E et al (2015) On the distribution and seasonal cycle of transient luminous events over Europe and the Mediterranean Sea. To be submitted to ACP
3. Arnone E et al (2014) WACCM climate chemistry sensitivity to sprite perturbations. J Geophys Res 119(11):6958–6970
4. Arnone E et al (2012) Extreme ozone depletion in the 2010-2011 Arctic winter stratosphere as observed by MIPAS/ENVISAT using a 2-D tomographic approach. Atmos Chem Phys 12 (19):9149–9165
5. Arnone E et al (2009) Seeking sprite-induced signatures in remotely sensed middle atmosphere $NO_2$: latitude and time variations. Plasma Sources Sci Technol 18:1–10
6. Arnone E et al (2008) An estimate of the impact of transient luminous events on the atmospheric temperature. Advances in Geosc. 13(13):37–43
7. Arnone E et al (2008) Seeking sprite-induced signatures in remotely sensed middle atmosphere $NO_2$. Geophys Res Lett 35:L05807
8. Beirle S et al (2010) Direct satellite observation of lightning-produced $NO_x$. Atmos Chem Phys 10:10965–10986. doi:10.5194/acp-10-10965-2010
9. Brasseur GP, Solomon S (2005) Aeronomy of the middle atmosphere, 3rd edn. Springer, The Netherlands

10. Carlotti M et al (2006) GMTR: two-dimensional geo-fit multitarget retrieval model for MIPAS/ENVISAT observations. Appl Opt 45:716–727
11. Christian HJ, Blakeslee RJ, Boccippio DJ, Boeck WL, Buechler DE, Driscoll KT, Goodman SJ, Hall JM, Koshak WJ, Mach DM, Stewart MF (2003) Global frequency and distribution of lightning as observed from space by the optical transient detector. J Geophys Res 108(D1):4005. doi:10.1029/2002JD002347
12. Dinelli BM et al (2010) The MIPAS2D database of MIPAS measurements retrieved with a multi-target 2-dimensional tomographic approach. Atmos Meas Tech 3:355
13. Enell C-F et al (2008) Parameterisation of the chemical effect of sprites in the middle atmosphere. Ann Geophys 26:13–27
14. Gordillo-Vazquez F (2008) Air plasma kinetics under the influence of sprites. J Phys D Appl Phys 41(234):016
15. Hiraki Y et al (2008) Chemistry of sprite discharges through ion-neutral reactions Atmos. Chem Phys 8:3919–3928
16. Krehbiel PR, Riousset JA, Pasko VP, Thomas RJ, Rison W, Stanley MA, Edens HE (2008) Upward electrical discharges from thunderstorms. Nature. doi:10.1038/ngeo162
17. Marsh DR et al (2013) Climate change from 1850 to 2005 simulated in CESM1(WACCM). J Clim 26:7372–7391
18. Neubert T et al (2008) Recent results from studies of electrical discharges in the mesosphere. Surv Geophys 29:71
19. Pasko VP (2010) Recent advances in theory of transient luminous events. J Geophys Res (Space Physics) 115:00–35
20. Peterson H et al (2009) $NO_x$ production in laboratory discharges simulating blue jets and red sprites. J Geophys Res 114: A00E07
21. Rodger CJ, Seppala A, Clilverd MA (2008) Significance of transient luminous events to neutral chemistry: experimental measurements. Geophys Res Lett 35(AQ8):7803
22. Schumann U, Huntrieser H (2007) The global lightning-induced nitrogen oxides source. Atmos Chem Phys 7(14):3823–3907
23. Sentman DD et al (1995) Preliminary results from the Sprites94 aircraft campaign: 1 Red sprites. Geophys Res Lett 22:1205
24. Sentman DD et al (2008) Plasma chemistry of sprite streamers. J Geophys Res 113(11):112
25. Winkler H, Notholt J (2014) The chemistry of daytime sprite streamers—a model study. Atmos Chem Phys 14(7):3545–3556

# TIBAGS: Tropospheric Iodine Monoxide and Its Coupling to Biospheric and Atmospheric Variables—a Global Satellite Study

**Anja Schönhardt, Andreas Richter and John P. Burrows**

**Abstract** In the framework of the TIBAGS project, spatial and temporal variations of iodine monoxide, IO, in the Earth's atmosphere were analysed, and relations between IO and further variables of the biosphere and atmosphere were investigated. The abundances and variations of IO are not well known on a global scale, partly because IO amounts are comparably low. However, due to strong reactivity, also small amounts of IO may have a substantial impact on tropospheric composition. In the present study, satellite data from the SCIAMACHY (Scanning Imaging Absorption spectrometer for Atmospheric CHartographY) sensor on board the ENVISAT satellite is used and a more global view on the subject is obtained. IO amounts are retrieved from measurements of scattered sunlight by using an absorption spectroscopy technique. Two consistent IO data sets are retrieved, one based on near real-time data (2004–2011) and one based on reprocessed consolidated data (2003–2010). Largest amounts of IO are found in the Polar Regions of Antarctica, for example in the Weddell Sea area in spring time. In addition, enhanced IO amounts are detected above some but not all biologically active ocean areas which show high Chlorophyll-a (Chl-a) signals. Correlations between IO and diatom distributions are in some areas stronger than between IO and Chl-a in general, indicating the importance of the specific phytoplankton species present in the ocean water.

## 1 Background Information

The focus of the TIBAGS project is set on the trace gas iodine monoxide in the atmosphere. Iodine compounds are relevant for tropospheric composition for several reasons. Iodine radicals react with ozone, whereby iodine monoxide, IO, is formed and tropospheric ozone is destroyed. This also affects levels and lifetimes of

A. Schönhardt (✉) · A. Richter · J.P. Burrows
Institute of Environmental Physics, University of Bremen, Otto-Hahn-AHee 1,
28359 Bremen, Germany
e-mail: schoenhardt@iup.physik.uni-bremen.de



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_2

other tropospheric species. In addition, IO may lead to the formation of fine atmospheric particles which influence the radiation budget. IO is hence an indicator of active iodine photochemistry, and iodine chemistry is considered to be relevant for understanding tropospheric composition, especially in the marine boundary layer and Polar Regions (e.g. [6, 22], and references therein). Precursors of IO include molecular iodine, $I_2$, as well as halocarbons. These may be emitted, e.g. by algae and phytoplankton ([15], and references therein), or via inorganic pathways from the ocean [7].

Several field studies have investigated these precursors as well as atmospheric levels of IO. However, all field studies are necessarily restricted in time and space, and the spatial and temporal distributions of atmospheric iodine are only partly known. Satellite measurements are an additional and valuable source of information as they yield near global observations over time scales of many years.

In the framework of the TIBAGS project, abundances of IO are retrieved from space on a nearly global scale. The retrieval of IO from the Scanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY) has initially been demonstrated by Schönhardt et al. [24]. Based on near real-time data, now the time series covers the years 2004–2011. Comparisons between the amounts and distributions of IO with selected parameters of the atmosphere and biosphere shall improve our understanding of source regions and links to other processes. The studied parameters include atmospheric trace gases such as bromine monoxide, BrO, and the short-lived organic compounds formaldehyde, HCHO, and glyoxal, CHOCHO, as well as compounds dissolved in the ocean waters such as Chlorophyll-a and individual phytoplankton species.

Focus areas of the TIBAGS project are the Polar Regions, especially Antarctica, as well as the world's ocean areas. An overview of IO above Antarctica is shown in Fig. 1. The Southern Hemispheric map depicts the average IO vertical column

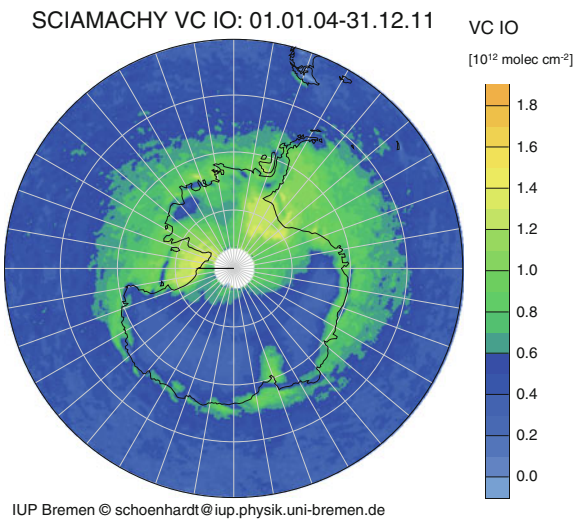


**Fig. 1** Average IO vertical column amounts for eight years (2004–2011) above the Antarctic region

amount for the eight years from 2004 to 2011. IO amounts are enhanced above the sea ice region, above the shelf ice, along the coast lines, as well as above parts of the continent. Details on the IO retrieval and the resulting observations are described in the following sections.

## 2 The Satellite Sensor

Column amounts of IO may be retrieved by using absorption spectroscopy. The well-established and widely used method of differential optical absorption spectroscopy (DOAS) [16, 17] was applied in the present study to detect IO amounts in the radiances recorded by the SCIAMACHY sensor on board the European Space Agency's (ESA's) Environmental Satellite ENVISAT. SCIAMACHY is a spectrometer measuring in the ultraviolet (UV), visible and infrared (IR) wavelength regions, in three different geometries: nadir, limb and occultation [3, 5, 11, 12]. In the present study, nadir measurements in the visible were used.

## 3 Data Analysis

The DOAS retrieval initially yields slant column amounts of IO, which describe the amount of IO integrated along the slant light path. The slant columns are the result of a least-squares optimization routine based on the Lambert–Beer law. An actual radiance measurement $I$ is compared to a background measurement $I_0$, which in the present case is an Earthshine radiance chosen from a background region. The difference between the two measurements is caused by several atmospheric effects, including absorption, scattering and reflection, of the electromagnetic radiation. Absorption by trace gases in the atmosphere is one important contribution. In the DOAS method, only those spectral effects, which quickly vary with wavelength (high-pass filter), are further analysed. Low-frequency effects are effectively filtered out by the subtraction of a polynomial, a quadratic polynomial in the present case.

The applied DOAS retrieval for IO uses the wavelength window between 416 and 430 nm. The IO absorption cross-section measured by Gómez Martín et al. [10] is applied to identify the IO absorption bands in the measurements. Additional trace gases taken into account are $O_3$ and $NO_2$. In addition, the Ring effect is taken into account, an effect that is caused by inelastic scattering on molecules in the Earth's atmosphere and leads to an infilling of absorption lines, especially the solar Fraunhofer lines. The Ring effect is calculated separately by radiative transfer (RTF) calculations, and an effective Ring spectrum is fitted as a pseudo-absorber cross-section in the same manner as the trace gases. A linear intensity offset is also taken into account.

## 4 Considerations on the Air Mass Factor

As the slant column is intrinsically dependent on the respective light paths which the radiation has taken through the absorber layer(s), this light path needs to be estimated when the slant columns are to be converted into the more comprehensible values of vertical columns. The vertical column is the amount of the absorber per ground area integrated vertically through the atmosphere and usually given in molecules per $cm^2$ (molec/$cm^2$). The light path taken by the radiation going from the sun through the atmosphere and into the satellite sensor is computed by RTF calculations.

In the present study, the RTF code SCIATRAN is applied [20] to calculate the so-called air mass factor (AMF) which is the light path length through the absorber layers relative to a single vertical transmission. Hence, the vertical column $VC_i$ of trace gas $i$ is given by the ratio of the slant column $SC_i$ to the AMF $a$.

$$VC_i = \frac{SC_i}{a(\lambda, p)}$$

The calculated AMF $a$ depends on the wavelength $\lambda$ and on a parameter set $p$ including, e.g. the surface reflectance, solar zenith angle and the absorber profile. The parameters used in the RTF calculation need to be adapted to the respective measurement scenario.

By considering the variation of AMF with the solar zenith angle, SZA, the dependency of the slant column value on the SZA is effectively eliminated in the vertical column value. For satellite geometry in nadir observation, typical variations of AMF with changing SZA are shown in Fig. 2.

The left graph is computed for a Polar scenario with snow/ice cover and corresponding albedo of 0.9, and SZA variation from 30 to 84°. The standard IO

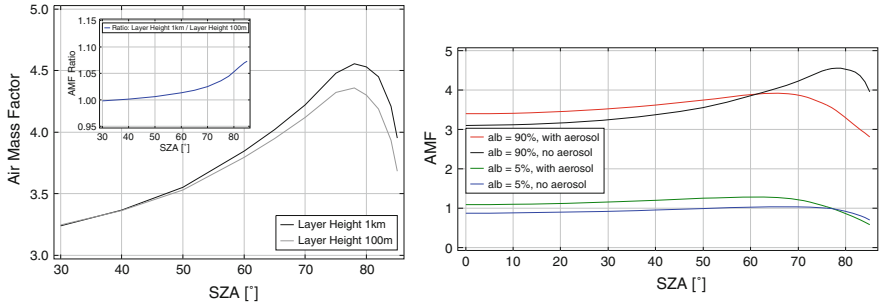


**Fig. 2** *Left* AMF for IO at 90 % albedo calculated for different mixing layer heights, for 1 km shown in *black* and for 100 m shown in *grey*. The *inset* shows the ratio of the two curves, maximum deviation at large SZA does not exceed about 7 %. *Right* AMF for different scenarios, testing the influence of aerosols on the IO AMF for a bright (90 %) albedo scene (*red* and *black* curves) and a dark (5 %) albedo case (*green* and *blue* curves)

retrieval only uses measurements up to 84° SZA, as the signal-to-noise ratio (SNR) strongly decreases for larger angles. The left figure compares the AMF for two different profiles, both box profiles of constant mixing ratios, one going from the ground up to 1 km (black), the other from ground to 100 m (grey). The actual IO profile is not well known. Recent observations report on IO in the free troposphere [8, 18], but IO is mostly found in the boundary layer (e.g., [6, 21]).

For the results in Fig. 2, a pure Rayleigh atmosphere is considered, i.e. without scattering on aerosols. For Antarctica, this is a valid first assumption. The largest difference of 7 % between the two settings is found for large SZA. As the influence of the mixing layer height on the AMF is comparatively small for this high albedo scene, the choice of the layer height does not influence the derived IO vertical column amount much. The results from the black curve with 1 km mixing layer height are typically used for the calculation of the vertical IO columns in Polar studies such as the Antarctic map in Fig. 1.

Aerosols are frequently present in other scenes and may strongly influence the AMF as the light path is influenced by additional scattering processes on the aerosol particles. Whether the AMF increases or decreases with aerosol load depends on the relative vertical position of the aerosol and absorber layers as well as on the aerosol type. The right graph of Fig. 2 presents the aerosol influence on the AMF value. As an example case for ocean scenarios, the aerosol type considered here is maritime aerosol with a visibility of 10 km. The aerosol is mixed with the IO layer and is partly situated above the IO. The calculated AMF results are again plotted versus SZA, where the cases for 90 % albedo are shown in red and black (with and without aerosols) for comparison, and the cases for 5 % albedo typical for water surfaces are shown in green and blue. For small SZA, aerosols enhance the sensitivity towards IO detection, i.e. lead to an increase of the AMF. For the bright case of 90 % albedo, the enhancement lies around 10 %, while for the dark scene, aerosols enhance the sensitivity by around 25 %. At larger SZA (>62° for 90 % albedo, and >77° for 5 % albedo), aerosols lead to a decrease of the AMF. Only for the 90 % case at low sun (large SZA) a difference larger than 25 % is found with lower sensitivity in the presence of aerosols. For smaller SZA, therefore, not considering aerosol influence on the AMF may lead to a limited overestimation of IO vertical columns while for larger SZA, some underestimation may occur.

In the details, therefore, the IO amount is dependent on the aerosol load, but for moderate aerosol amounts, the influence does not dominate over regional or temporal changes. The IO overview maps presented in the next sections are computed for aerosol-free scenarios.

## 5 Detection Limit and Averaging

Considering slant column amounts, the IO detection limit for a single SCIAMACHY measurement lies at $7 \times 10^{12}$ molec/cm$^2$. The vertical column detection limit depends on the AMF. Above snow and ice, the detection limit lies

around $1.7 \times 10^{12}$ molec/cm$^2$. As the IO amounts are often not much larger than the detection limit, temporal and/or spatial averaging is necessary to improve SNR and thus data quality. In addition, absolute IO amounts need to be treated with caution. IO maps are typically generated as averages over time spans of several months. This way, the statistical error on the measurements is reduced. In order to resolve smaller scale temporal variations, single calendar months of subsequent years are averaged. Using this strategy, a time series of IO maps through different seasons may be generated (cf. Sect. 6).

## 6 Observations Above Antarctica

Using SCIAMACHY satellite data, both IO and BrO are retrieved for many years and compared for the same time periods above the Southern Hemisphere. Also based on satellite observations, the regions of IO enhancement are compared to the sea ice cover in the respective time periods.

### *6.1 Spatial and Temporal Variations of IO Vertical Columns*

The temporal averaging period in order to obtain a SNR of sufficient quality for the IO vertical column product depends on several aspects such as the time of year, the location on the Earth and the surface conditions. Usually, a suitable averaging period for IO data is a few months.

In order to resolve temporal variations in the IO amounts not only on a seasonal basis but on a smaller time scale, IO columns of single calendar months are averaged over subsequent years. In this way, monthly variations which reoccur every year can be resolved. This procedure is suitable owing to the fact that many features in the spatial pattern of IO are repeated annually. As a result, monthly maps, each averaged over eight years from 2004 to 2011, are produced. They are used for comparison with BrO columns and other parameters. The temporal evolution of IO enhancements above the Antarctic regions is thus observed. In Schönhardt et al. [23] maps for the time period 2004–2009 have been published. Figure 3 shows the time series of IO from October (Antarctic spring time) through summer to March (Antarctic autumn).

For direct comparison, the same time series is plotted for BrO vertical columns in Fig. 5 in Sect. 6.2 below.

IO amounts in September are rather scattered due to low light levels resulting in lower signal at the satellite. Some locally enhanced IO amounts along the coast west of the Antarctic Peninsula are detected and some scattered amounts above the ice shelves. In October, enhanced IO vertical columns of up to about $1.6 \times 10^{12}$ molec/cm$^2$ are spread over wide parts of the shelf ice areas, especially

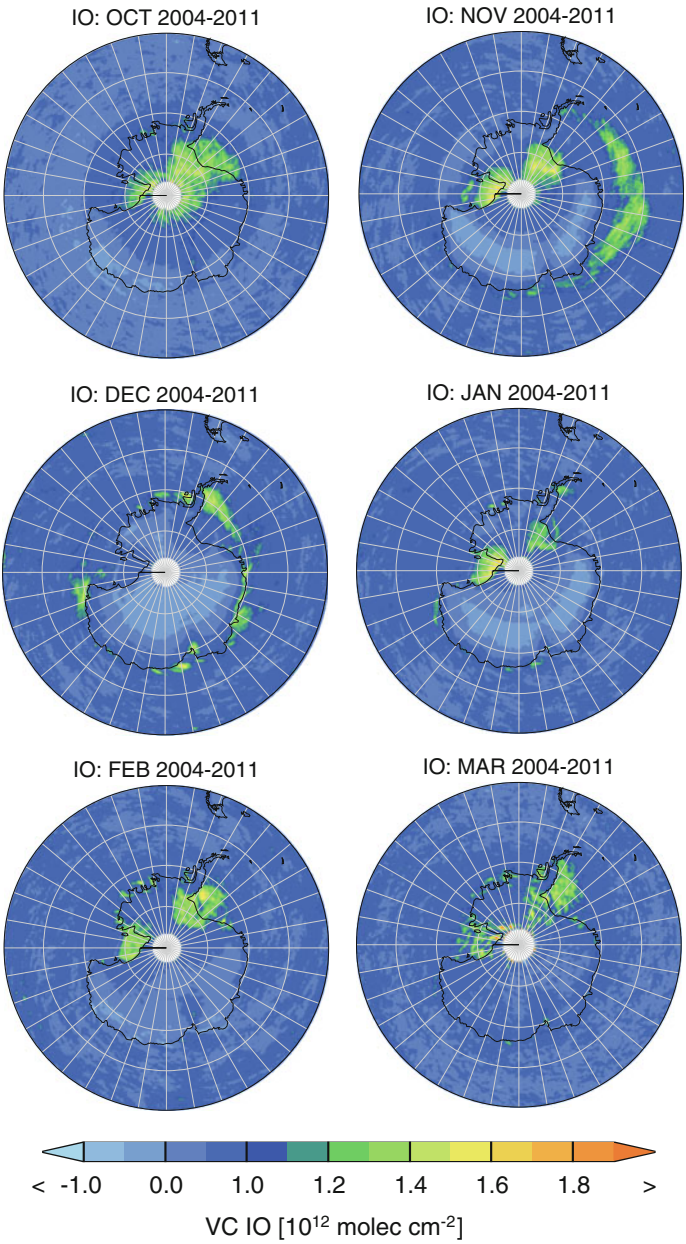


**Fig. 3** Monthly averages of IO vertical columns above the Antarctic continent for eight years (2004–2011). The AMF applied here assumes a ground reflectance of 90 % suitable for clean snow and ice

the Weddell Sea and the Ross Sea areas, in coastal locations and the continent. Between October and November then a distinct feature of IO enhancement evolves at some distance of the coast, and has fully developed into a circular region of IO amounts above the sea ice in the November average. The circular enhancement retreats somewhat in December and moves closer to higher latitudes, and IO is still visible along the coasts on some sea ice patches. Later in autumn (January/February), IO is mostly found on the shelf ice areas. In parts, the areas of IO occurrence overlap and agree with the sea ice cover.

The main spatial features are repeated from year to year. This is demonstrated by Fig. 4 showing Antarctic monthly means of IO observations for one year (top row), six years (middle row, the time period used in Schönhardt et al. [23]) and the full eight years of the IO near real-time product (bottom row). A longer averaging period reduces noise effects and emphasizes the main features of IO enhancement. For the comparison in Fig. 4, the calendar months November (left column), December (middle column) and January (right column) have been chosen, as the IO spatial pattern changes noticeably during this time of year with enhanced IO amounts above the ring-shaped sea ice around Antarctica in November, and reducing amounts and spatial extent towards January.

Clearly, local values may be larger when using shorter averaging periods. It is remarkable, however, that spatial patterns and IO amounts are conserved rather clearly when averaging over several years. Some persistence in the seasonal variation of IO is the reason for this behaviour.

## *6.2 Comparison of IO and BrO Distributions*

Bromine monoxide, BrO, is a molecule that is in principle similar to IO. However, the atmospheric relevance as well as sources and sinks may be quite different. In the Polar Regions, BrO is regularly generated by a mechanism called the bromine explosion (cf., e.g. [25], and references therein). During these events that begin shortly after Polar sunrise in early spring time, atmospheric BrO increases rapidly and is present above large areas. The release mechanism is an inorganic process. The relevant and necessary conditions for the bromine explosion to take place are a matter of continuing research.

In order to investigate the relations between IO and BrO distributions, satellite observations of the two species are compared. BrO vertical columns are retrieved from SCIAMACHY observations in the UV wavelength range [19, 30]. Taking into account the stratospheric amounts of BrO, a stratospheric AMF has been applied [19]. More details on AMF considerations for BrO observations can be found in studies by Begoin et al. [2] and Theys et al. [27]. The BrO data were averaged for the same periods and the same region as for IO and are shown in Fig. 5.

The BrO distributions show some clear differences towards the IO observations, but also some general similarities. One similarity lies in the fact that both, IO and

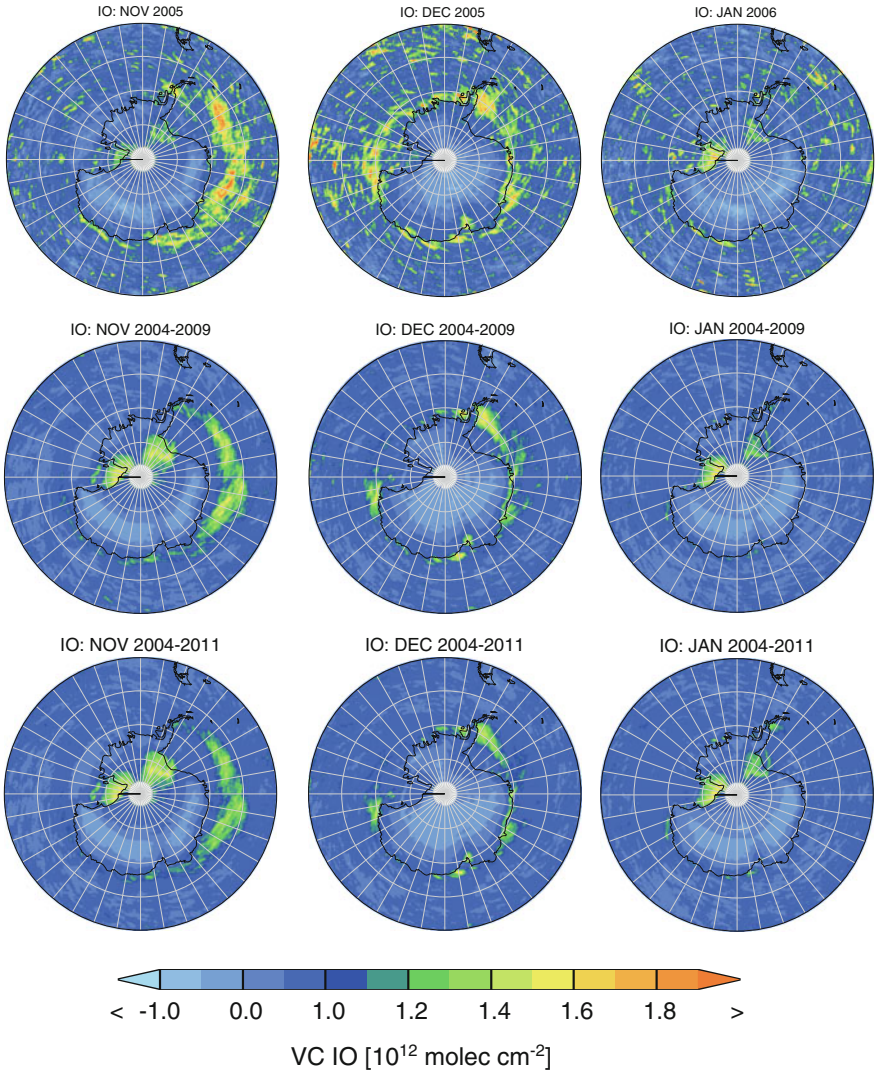


**Fig. 4** Comparison of IO vertical columns over Antarctica for an averaging period of one year (*top row*), six years (*middle row*, as used in Schönhardt et al. [23]) and eight years (*bottom row*). Main patterns and regional enhancements of IO reappear from year to year

BrO amounts, are enhanced above the Antarctic region in spring time. The details of the spatial and temporal distributions however, are quite different. BrO is mostly present above the sea ice around the continent, and enhancement starts with Polar sunrise in August (not shown), and is fully developed in September. Through the summer months, BrO amounts decrease and mostly vanish in autumn, except for

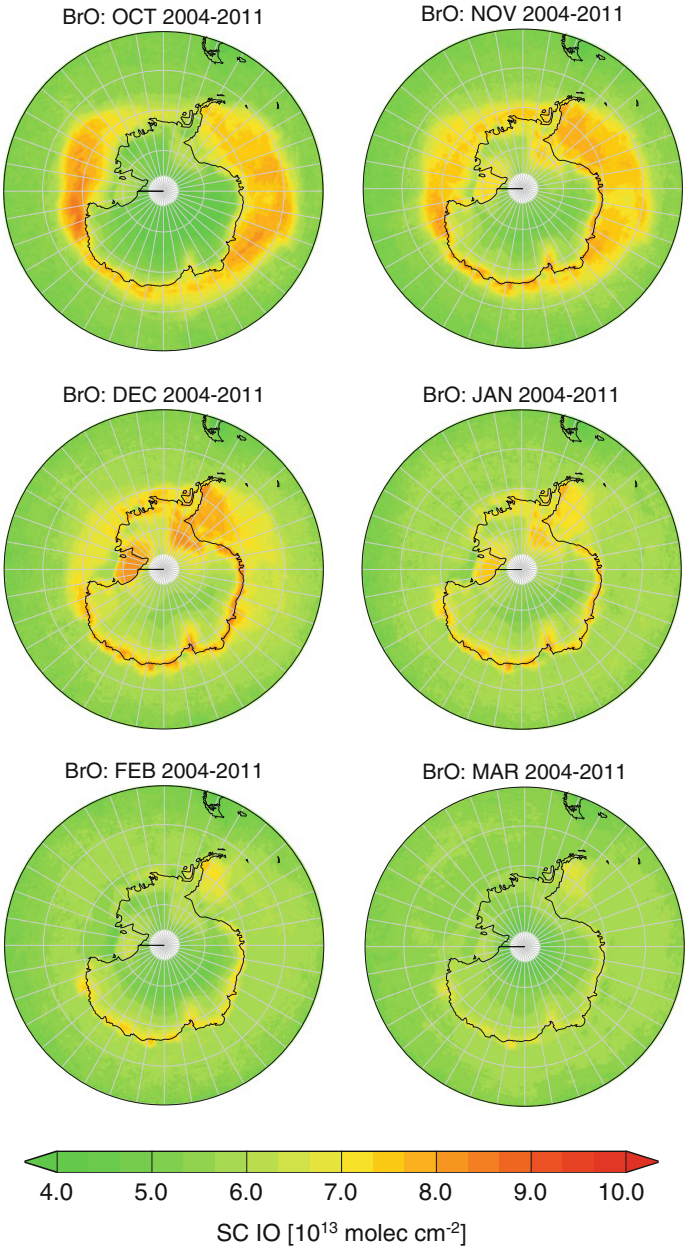


**Fig. 5** BrO vertical column averages above the Antarctic region for the same averaging periods as for IO in Fig. 3

enhancements along the coast lines and above the ice shelves (there especially in December and January).

Resulting from the differences between the IO and BrO spatial and temporal distributions, it can be concluded that at least some individual release pathways exist for iodine and bromine species in the South Polar Region. Although both, IO and BrO, occur in Antarctic Spring time, IO is present above the sea ice for a comparably shorter time period concentrated more towards later spring, while BrO is already present on the sea ice prominently from early spring onwards.

## 6.3 *Relation of the Halogen Oxides to Sea Ice Cover*

Both halogen oxides, IO and BrO, show enhancements above the sea ice covered area around the Antarctic continent. Comparisons with sea ice cover data have been performed using ice concentration data from the AMSR-E instrument [13, 26]. AMSR-E is the Advanced Microwave Scanning Radiometer for EOS, a passive microwave radiometer on board the NASA AQUA satellite belonging to the Earth Observing System (EOS). The ice concentration data is provided by ZMAW (Centre for Marine and Atmospheric Sciences) in Hamburg, Germany, and can be downloaded from the Integrated Climate Data Center, KlimaCampus, at the University of Hamburg (ICDC, http://icdc.zmaw.de/seaiceconcentration_asi_amsre.html?&L=1). The ice concentration is defined as the percentage of the representative AMSR-E satellite pixel covered by sea ice.

Figure 6 (right) shows the monthly ice concentration for November averaged over six years from 2004 to 2009, together with the IO map (left) for the same time period. The close spatial correlation of sea ice area with BrO enhancements

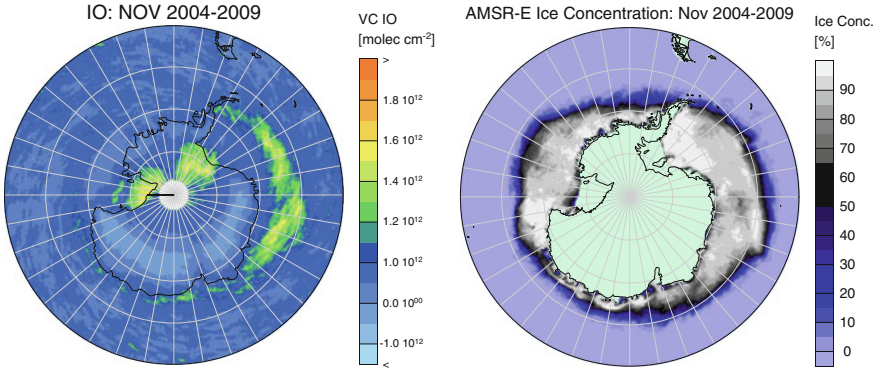


**Fig. 6** IO vertical columns (*left*) from SCIAMACHY observations and ice concentration (*right*) derived from AMSR-E data in November above the Southern Polar Region averaged over six years from 2004 to 2009. The ice map is based on daily data provided by the Integrated Climate Data Center, Hamburg (http://icdc.zmaw.de)

(cf. Fig. 5) is apparent. In addition, the curved region of enhanced IO in November is located above still present sea ice. In November, the density of the sea ice is reduced in some areas, as visible in Fig. 6 by darker grey patches within the sea ice area. The ice sheets start to break up and retreat in spring time. When the sea ice becomes more porous, and more open leads and polynyas develop, the water gets into contact with the atmosphere above. As iodine compounds are presumably emitted by biological species such as phytoplankton or ice algae, which prefer the habitat underneath the sea ice sheets, iodine input to the atmosphere may be facilitated as soon as the ice sheets break up. Convection above open water areas further supports the insertion of gaseous species, as the water is warm compared to the spring time Antarctic boundary layer air. These considerations form a possible explanation for the temporal behaviour of the observed IO occurrence in late spring. As further evidence of a connection to a biological source, Chlorophyll-a data have been consulted. This is discussed in Sect. 7.

In addition to providing a habitat for biological species, the sea ice cover also changes the radiation conditions in the respective areas. This has two consequences. On one hand, the stronger light reflection improves the visibility of IO above ice covered regions. On the other hand, the photochemical situation is different above and next to the ice. While the first aspect enhances the observed IO amounts above sea ice, the second aspect can influence in both directions, as iodine precursors as well as the IO amount itself are altered by changing light conditions. In any case, the Antarctic ice region is an especially interesting area for iodine research.

In a study by Atkinson et al. [1], the Weddell Sea area has been selected as focus area for iodine measurements in and above the sea ice and the ocean. The field study includes measurements in the water, ice and atmosphere, and the results emphasize that the Weddell Sea area is rich in iodine chemistry. Atmospheric iodine chemistry, however, is not yet well enough understood to make full atmospheric modelling possible. Not all observed data may be reproduced by model calculations. Some iodine source terms might still be unknown.

## 7 Relations Between IO and Biospheric Parameters

Previous studies have demonstrated that iodine compounds are emitted by biological species such as microalgae and macroalgae. Several chemical compounds are thereby emitted by various organisms in different speciations and different amounts [9, 28]. Satellite IO observations are compared to Chlorophyll-a (Chl–a) concentrations in the oceans, an indicator of active biology. For this comparison, IO data from a reprocessed data set for 2003–2010 is used. The data set is fully consistent with the near real-time data set from 2004 to 2011, but includes a more complete time series during the initial year 2003. For the investigation of Chl-a concentrations, the ESA merged GlobColour product is used. This product is based on the three instruments SeaWiFS (Sea-viewing Wide Field-of-view Sensor), MERIS (Medium Resolution Imaging Spectrometer) and MODIS (Moderate

Resolution Imaging Spectroradiometer). Information and data are available from the website http://www.globcolour.info.

## 7.1 Comparison Between IO and Chlorophyll-a Above Antarctica

The oceans surrounding the Antarctic continent are a region of strong biological activity. Following the detection of IO within the sea ice zone above the biology-rich waters, the first interesting comparison is that of IO and Chl-a around Antarctica. Long-term averages over many years are compiled and presented in Fig. 7 for IO (left) and Chl-a concentrations (right). IO vertical columns are computed here with an AMF that assumes a surface albedo of 5 %, appropriate for oceans in the visible spectral range. A spatial mask is applied to the IO data, showing the results only for regions, where Chl-a data is also available. The Chl-a data can only be derived from above-satellite ground pixels which are entirely free of ice, otherwise the signal from the ice covered part of the field-of-view would dominate the measured signal. This way, also the IO data is plotted only above regions which are at least for some part of the year free of ice. As the low albedo is applied, IO amounts in this map differ from the ones above. Some overestimation of the absolute IO amount may occur when applying the low albedo scenario.

Two main features can be derived from Fig. 7. First of all, the Antarctic proves to be an area rich in biological productivity, demonstrated by fairly large Chl-a concentrations all around the Antarctic continent with local maxima in the Weddell Sea, the Ross Sea and just off the coast over long distances as well as the area

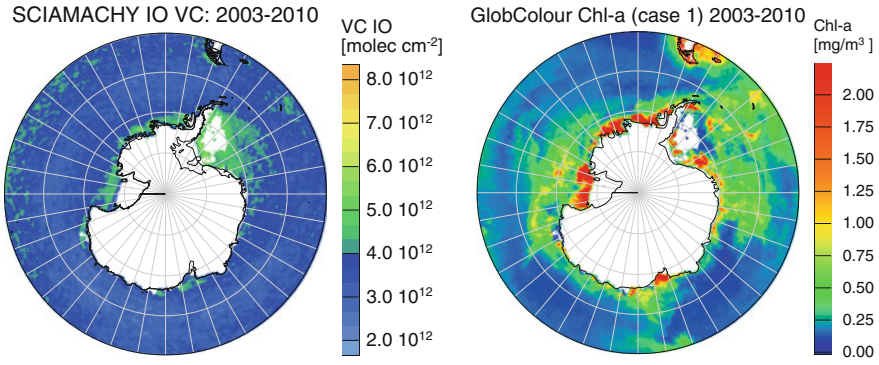


**Fig. 7** Antarctic maps of IO vertical columns in the atmosphere (*left*) and Chlorophyll-a concentrations in the ocean (*right*). For the IO data, conversion to vertical columns assumes an albedo of 5 % suitable for observations above water bodies, and data is masked by the area, where Chl–a data is available, i.e. data is only shown above regions which are free of ice at least in parts of the year

offshore of the Amery ice shelf. The second striking observation is the spatial overlap of enhanced values of IO with many of these Chl-a rich locations. Especially the Weddell Sea and the waters of the Ross Sea offshore of the Ross ice shelf show both, enhanced IO and enhanced Chl-a, in similar spatial patterns. Naturally, a correlation does not represent causality. However, from this comparison, a relation between biological production and the release of iodine precursors seems probable.

## 7.2 *Comparison Between IO and Chlorophyll-a Above Ocean Areas*

In contrast to the Antarctic analysis, spatial correlation between IO and Chl-a is not a general feature elsewhere. Other regions on the globe do not show such a clear spatial relation. Especially, regions with strong biological productivity can be found where no significant IO signal is detected.

For biological productivity, the ocean upwelling regions are important. Here, colder deep water masses rise to the surface and are often rich in nutrients and organically produced gaseous compounds may enter the atmosphere. Figure 8 compares atmospheric IO amounts from SCIAMACHY with oceanic Chl-a

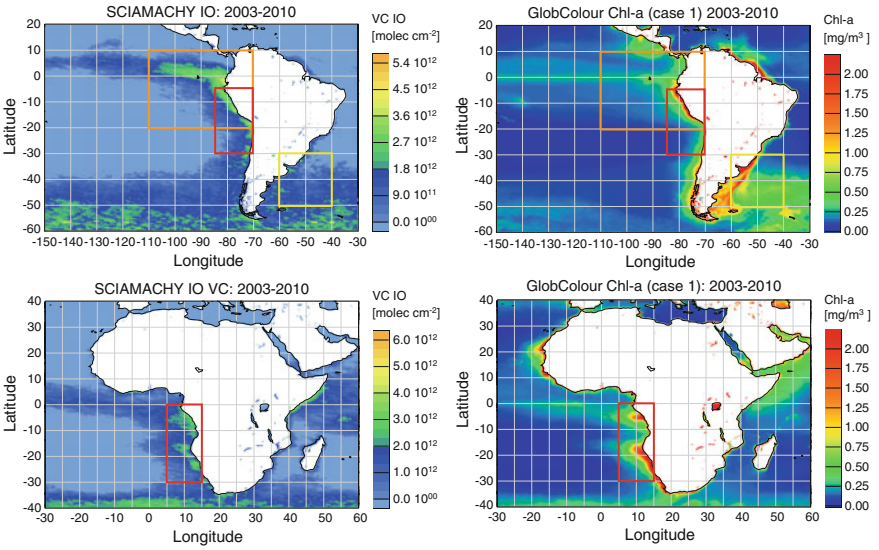


**Fig. 8** Comparison of eight year averages (2003–2010) of atmospheric IO (*left*) and oceanic Chl–a concentrations (*right*). The areas marked by coloured boxes are used for the computation of spatial correlation coefficients

concentrations from the GlobColour project, both for an eight year average from 2003 to 2010. The maps in the top show the Eastern Pacific and Southwest Atlantic around South America, and the bottom maps focus on the waters around Africa. Field studies have also reported on elevated IO amounts above the Eastern Pacific [8, 14, 29]. Along the West coast of South America, a large upwelling region is situated where the Humboldt current from Antarctica surfaces. As seen in the enhanced Chl-a data, off the coasts of Peru and Chile, biological activity is present. Enhanced Chl-a amounts are also seen further into the Pacific towards the Galápagos Islands.

Above the Peru and Northern Chile upwelling area, Chl-a enhancements coincide with enhanced IO abundances, which also spread further into the ocean towards the Galápagos Islands. Around the South tip of Chile and Argentina, where Chl-a is large especially on the East coast, IO amounts are smaller and the relation between the two compounds is less prominent. North of Brazil, around the Amazon estuary, the Chl-a amounts are large, and IO amounts are also detected, however, much less strongly as compared to the Eastern Pacific. For the coloured boxes in Fig. 8 (top), example correlation coefficients for the spatial correlations are determined to be $r = 0.29$ and 0.27 for the yellow and orange areas, respectively, and somewhat larger at $r = 0.43$ for the red box, each with an uncertainty around 0.02. These coefficients are not overly large, but significantly positive, and the red box close to the coast reveals the strongest correlation of these three areas.

Around the African continent the diverse relation can be seen even better. Figure 8 (bottom) shows that Chl-a is enhanced especially above the Mauritanian upwelling region (off the African Northwest coast) and above a two parted area off the African Southwest coast, partly coinciding with the Benguela current. In the Southwest, the spatial patterns of IO and Chl-a are very similar, while in contrast, no significantly enhanced IO is found above the Mauritanian upwelling in the North. For the Southwest region marked by a red box, the spatial correlation lies at $r = 0.50$.

The Chl-a and IO spatial patterns are also similar, e.g. towards the open ocean at the Southern limit of the displayed map as well as at the West coast of Somalia. Clear differences appear along the coasts of the Arabian Peninsula where no enhanced IO is detected above areas where Chl-a is present.

All in all, an ambiguous picture is received from the analysis of Fig. 8. Enhancements in IO and Chl-a concentration coincide for some areas, while other locations do not exhibit a spatial relation between the two variables, i.e. Chl-a concentrations are high while no large IO is found. Possibly, the phytoplankton types as well as surface and atmospheric conditions or other than direct biological source pathways play a role for the emissions of iodine compounds.

## 7.3 *Comparison Between IO, Chlorophyll-a and Diatoms for Southeast Asia*

Following from the ambiguous picture discussed in the previous section, investigations of individual phytoplankton types are of interest. Using the PhytoDOAS method [4], a few different phytoplankton species may be distinguished. From the absorption spectrum characteristic for each different organism, the PhytoDOAS method retrieves an equivalent Chlorophyll-a concentration indicative of the amount of phytoplankton present in the light absorbing upper ocean layers. In the context of iodine release, the distributions of diatoms are specifically interesting and have therefore been investigated for some selected regions. Diatom maps display the equivalent Chl-a amount, which should be proportional to the amount of diatom organisms in the upper ocean layers. The absolute diatom amount still depends somewhat on the unknown vertical phytoplankton profile.

Figure 9 shows the comparison of IO abundances (top), total Chl-a concentration (bottom left) and Chl-a from diatoms (bottom right) in the oceans of Southeast Asia. This region is of specific importance for the exchange between the stratosphere and the troposphere because of strong convective transport. Therefore, this area is the

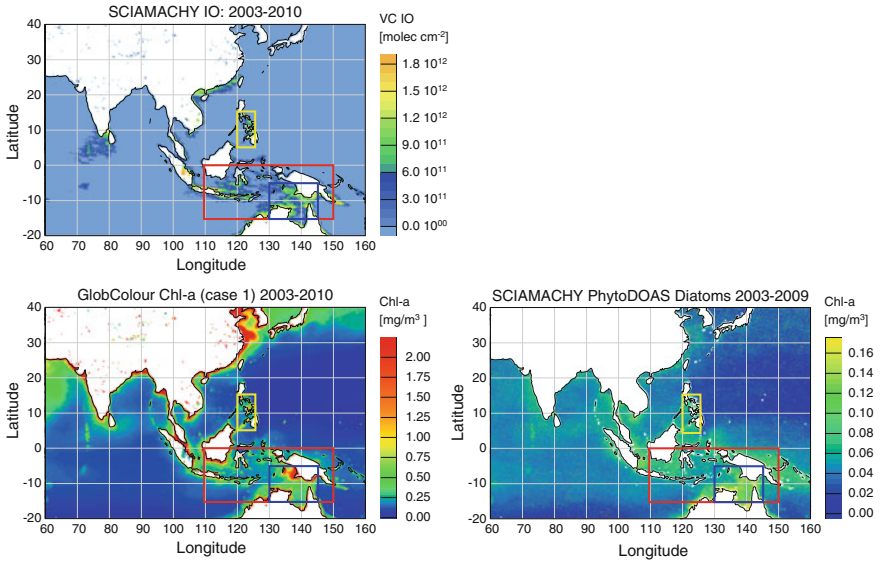


**Fig. 9** Comparison between long-term averages of atmospheric IO (*top*) with total Chl-a (*bottom left*) and Chl-a from diatoms (*bottom right*) in the oceans of Southeast Asia. Marked areas are used for example calculations of spatial correlation coefficients. In several locations, a spatial relation between IO and diatoms, as well as between IO and total Chl-a is observed, while other areas such as the East China Sea are rich in Chl-a with no diatoms and no IO. Diatom data are courtesy of Astrid Bracher, AWI Bremerhaven and Tilman Dinter, University of Bremen, in the framework of HGF project Phytooptics and EU project SHIVA

central research area of the SHIVA project (Stratospheric Ozone—Halogen Impacts in a Varying Atmosphere). Data is averaged over the years 2003–2010, for diatoms from 2003 to 2009 as data of the year 2010 was not available.

Several locations show enhanced IO, where Chl-a and diatoms are present in the water, especially at some coast lines, e.g. of Southern China, and between some islands, such as between the Philippines, between New Guinea and Australia, and further west in the sea gate between Sri Lanka and India. In most of these locations, the Chl-a content in the water is large, and diatoms are detected, except for the coast of Southern China, with large Chl-a but little diatom detections.

On the other hand, there is strong Chl-a occurrence in the Yellow Sea and East China Sea, but no IO is detected there and diatoms are not prominent there either.

Further to the qualitative comparison of the regional maps, some spatial correlation coefficients have been computed. For example, the areas between the islands as marked by three coloured boxes in Fig. 9 have been analysed. For IO and Chl-a, the correlation coefficients are $r = 0.33$, 0.31 and 0.38 for the yellow, red and blue areas, respectively. For IO and diatoms, the correlation coefficients are larger with $r = 0.48$, 0.47 and 0.62, respectively.

These results and observations are a further indication that there is no general one-to-one relationship between iodine compounds and all Chl-a producing biological species. If iodine species are biogenically produced, some differentiation amongst the emitting phytoplankton species is taking place which is strong enough that it becomes noticeable in the satellite measurements. In the selected regions, IO abundances correlate much better with the diatom distributions than with total Chl-a patterns, while the enhanced IO is not accompanied by strong diatom occurrence at the coast of South China. Consequently, although diatoms show a closer relation with IO in several locations than Chl-a, also diatoms are no safe indication for iodine emissions. Although the picture remains to some extent ambiguous, the spatial overview from satellite reveals many interesting regions with a spatial link between IO and the underlying biological situation.

## 8 Summary

The TIBAGS project has provided the opportunity to continue research on atmospheric iodine measured from space. Long-term data sets of IO observations from the SCIAMACHY instrument have been retrieved and investigated in order to increase our knowledge on spatial and temporal distributions and variations of atmospheric IO abundances. Largest amounts of IO have been detected in the Antarctic, and in the long-term data sets re-occurring IO maxima in the same regions each year have been found. Comparisons between IO and BrO above the South Polar Region show that besides the mutual appearance in Antarctic Spring, the spatial distribution as well as the temporal evolution differ in their details. Separate release pathways are most probably the reason for the differences.

In ocean areas around the South American and African continent, Chlorophyll-a concentrations in the ocean waters coincide with enhanced IO in several places but not in all. The spatial correlation between IO and Chl-a in some places suggests the importance of biological activity, while the missing one-to-one relationship indicates that other influencing factors are relevant for the release of iodine from the ocean. In the ocean areas of South-East Asia, correlations between IO and Chl-a as well as between IO and diatoms, a specific phytoplankton species known to emit iodine precursor substances, have been investigated. Spatial correlation coefficients between IO and diatom abundances are larger than between IO and total Chl-a, supporting the laboratory result that iodine release is to some extent dependent on the prevalent phytoplankton species. Closer investigations in ocean areas are needed to further improve our understanding on iodine release. Efforts should include field studies on local air composition and measurements of gaseous compounds as well as phytoplankton species in the ocean waters in addition to continuing long-term satellite observations.

**Acknowledgments** The TIBAGS project has been financially supported by ESA within the CESN framework. Further financial support was received from the State and University of Bremen, the German Aerospace Center DLR, and the European Union. SCIAMACHY data are provided by ESA and DLR. Support by Vladimir Rozanov on the application of SCIATRAN is gratefully acknowledged. Sea ice concentration data from AMSR-E observations are available at ICDC, Integrated Climate Data Center, ZMAW, in Hamburg, http://icdc.zmaw.de/seaiceconcentration asi amsre.html?&L = 1. Chl-a data are provided through the ESA GlobColour Project: ACRI & the GlobColour Team, funded by ESA with data from ESA, NASA and GeoEye, MERIS/MODIS/SeaWiFS merged product, information and data are available at http://www.globcolour.info. Diatom data are courtesy of Astrid Bracher, Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, and Tilman Dinter, University of Bremen, in the framework of HGF project Phytooptics (VH-NG-300) and EU project SHIVA (226224-FP7-ENV.2008.1.1.2.1).

## References

1. Atkinson HM, Huang R-J, Chance R, Roscoe HK, Hughes C, Davison B, Schönhardt A, Mahajan AS, Saiz-Lopez A, Hoffmann T, Liss PS (2012) Iodine emissions from the sea ice of the Weddell Sea. Atmos Chem Phys 12:11229–11244. doi:10.5194/acp-12-11229-2012
2. Begoin M, Richter A, Weber M, Kaleschke L, Tian-Kunze X, Stohl A, Theys N, Burrows JP (2010) Satellite observations of long range transport of a large BrO plume in the Arctic. Atmos Chem Phys 10:6515–6526. doi:10.5194/acp-10-6515-2010
3. Bovensmann H, Burrows JP, Buchwitz M, Frerick J, Noël S, Rozanov VV, Chance KV, Goede APH (1999) SCIAMACHY: mission objectives and measurement modes. J Atmos Sci 56:127–150
4. Bracher A, Vountas M, Dinter T, Burrows JP, Röttgers R, Peeken I (2009) Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. Biogeosciences 6:751–764. doi:10.5194/bg-6-751-2009
5. Burrows JP, Hölzle E, Goede APH, Visser H, Fricke W (1995) SCIAMACHY—scanning imaging absorption spectrometer for atmospheric chartography. Acta Astronaut 35:445–451
6. Carpenter LJ (2003) Iodine in the marine boundary layer. Chem Rev 103:4953–4962

7. Carpenter LJ, MacDonald SM, Shaw MD, Kumar R, Saunders RW, Parthipan R, Wilson J, Plane JMC (2013) Atmospheric iodine levels influenced by sea surface emissions of inorganic iodine. Nat Geosci 6:108–111. doi:10.1038/ngeo1687
8. Dix B, Baidar S, Bresch JF, Hall SR, Schmidt KS, Wang S, Volkamer R (2013) Detection of iodine monoxide in the tropical free troposphere. PNAS 110(6):2035–2040
9. Giese B, Laturnus F, Adams FC, Wiencke C (1999) Release of volatile iodinated C1–C4 hydrocarbons by marine macroalgae from various climate zones. Environ Sci Technol 33:2432–2439
10. Gómez Martín JC, Spietz P, Burrows JP (2007) Kinetic and mechanistic studies of the $I_2/O_3$ photochemistry. J Phys Chem A 111:306. doi:10.1021/jp061186c
11. Gottwald M, Bovensmann, H, Lichtenberg G, Noël S, von Bargen A, Slijkhuis S, Piters A, Hoogeveen R, von Savigny C, Buchwitz M, Kokhanovsky A, Richter A, Rozanov A, Holzer-Popp T, Bramstedt K, Lambert J-C, Skupin J, Wittrock F, Schrijver H, Burrows JP (eds) (2006) SCIAMACHY—monitoring the changing Earth's atmosphere, DLR, Institut für Methodik der Fernerkundung (IMF). Available online at http://atmos.caf.dlr.de/projects/scops/sciamachybook/sciamachybookdlr.html. Last access on 20 July 2012
12. Gottwald M, Bovensmann H (eds) (2011) SCIAMACHY—exploring the changing Earth's atmosphere. Springer, Heidelberg. doi:10.1007/978-90-481-9896-2. ISBN 978-90-481-9895-5
13. Kaleschke L, Lüpkes C, Vihma T, Haarpaintner J, Bochert A, Hartmann J, Heygster G (2001) SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis. Can J Remote Sens 27:526–537
14. Mahajan AS, Gómez Martín JC, Hay TD, Royer SJ, Yvon-Lewis S, Liu Y, Hu L, Prados-Roman C, Ordóñez C, Plane JMC, Saiz-Lopez A (2012) Latitudinal distribution of reactive iodine in the Eastern Pacific and its link to open ocean sources. Atmos Chem Phys 12:11609–11617. doi:10.5194/acp-12-11609-2012
15. Pedersén M, Collén J, Abrahamsson K, Ekdahl A (1996) Production of halocarbons by seaweeds: an oxidative stress reaction? Sci Marina 60:257–263
16. Platt U, Perner D (1980) Direct measurements of atmospheric $CH_2O$, $HNO_2$, $O_3$, $NO_2$, $SO_2$ by differential optical absorption in the near UV. J Geophys Res 85(C12):7453–7458
17. Platt U, Stutz J (2008) Differential optical absorption spectroscopy—principles and applications, Springer, Berlin. ISBN 978-3-540-21193-8
18. Puentedura O, Gil M, Saiz-Lopez A, Hay T, Navarro-Comas M, Gómez-Pelaez A, Cuevas E, Iglesias J, Gomez L (2012) Iodine monoxide in the north subtropical free troposphere. Atmos Chem Phys 12:4909–4921. doi:10.5194/acp-12-4909-2012
19. Richter A, Wittrock F, Eisinger M, Burrows JP (1998) GOME observations of tropospheric BrO in Northern Hemispheric spring and summer 1997. Geophys Res Lett 25:2683–2686
20. Rozanov VV, Rozanov AV, Kokhanovsky AA, Burrows JP (2014) Radiative transfer through terrestrial atmosphere and ocean: software package SCIATRAN. J Quant Spectrosc Radiat Transfer 133:13–71
21. Saiz-Lopez A, Mahajan AS, Salmon RA, Bauguitte SJ-B, Jones AE, Roscoe HK, Plane JMC (2007) Boundary layer halogens in coastal Antarctica. Science 317:348–351. doi:10.1126/science.1141408
22. Saiz-Lopez A, Plane JMC, Baker AR, Carpenter LJ, von Glasow R, Gómez Martín JC, McFiggans G, Saunders RW (2012) Atmospheric chemistry of iodine. Chem Rev 112:1773–1804. doi:10.1021/cr200029u
23. Schönhardt A, Begoin M, Richter A, Wittrock F, Kaleschke L (2012) Gómez Martín, J. C., and Burrows, J. P.: Simultaneous satellite observations of IO and BrO over Antarctica. Atmos Chem Phys 12:6565–6580. doi:10.5194/acp-12-6565-2012
24. Schönhardt A, Richter A, Wittrock F, Kirk H, Oetjen H, Roscoe HK, Burrows JP (2008) Observations of iodine monoxide columns from satellite. Atmos Chem Phys 8:637–653. doi:10.5194/acp-8-637-2008
25. Simpson WR, von Glasow R, Riedel K, Anderson P, Ariya P, Bottenheim J, Burrows J, Carpenter LJ, Frieß U, Goodsite ME, Heard D, Hutterli M, Jacobi H-W, Kaleschke L, Neff B, Plane J, Platt U, Richter A, Roscoe H, Sander R, Shepson P, Sodeau J, Steffen A, Wagner T,

Wolff E (2007) Halogens and their role in polar boundary-layer ozone depletion. Atmos Chem Phys 7:4375–4418. doi:10.5194/acp-7-4375-2007
26. Spreen G, Kaleschke L, Heygster G (2008) Sea ice remotesensing using AMSR-E 89 GHz channels. J Geophys Res 113:C02S03. doi:10.1029/2005JC00338
27. Theys N, Van Roozendael M, Hendrick F, Yang X, De Smedt I, Richter A, Begoin M, Errera Q, Johnston PV, Kreher K, De Mazière M (2011) Global observations of tropospheric BrO columns using GOME-2 satellite data. Atmos Chem Phys 11:1791–1811. doi:10.5194/acp-11-1791-2011
28. Tokarczyk R, Moore RM (1994) Production of volatile organohalogens by phytoplankton cultures. Geophys Res Lett 21(4):285–288
29. Volkamer R, Coburn S, Dix B, Sinreich R (2010) The eastern Pacific Ocean is a source for short lived trace gases: glyoxal and iodine oxide. Clivar Exchanges 53(2):30–33
30. Wagner T, Platt U (1998) Satellite mapping of enhanced BrO concentration in the troposphere. Nature 395:486–490

# GreenSAR—Greenland and Antarctic Grounding Lines from SAR Data

**Noel Gourmelen, JeongWon Park and Andrew Shepherd**

**Abstract** The Greenland and Antarctic Ice Sheet (GIS and AIS) are currently losing mass. The GreenSAR project aims at improving the knowledge of essential climate variables by exploiting Earth Observation (EO) datasets from past, present and future European Space Agency (ESA) satellite missions. The project is built around the measure of grounding line migration at the ice–ocean interface. Here, we show that the grounding line of the Pine Island Glacier (PIG) has retreated continuously for the past 20 years due to sustained thinning. The Petermann Gletscher glacier grounding line shows little change in the past 20 years despite the two recent large calving events reducing the area of the floating tongue by 40%.

## 1 Introduction

The Greenland and Antarctic Ice Sheet (GIS and AIS) are currently losing mass [30]. The GreenSAR project aims at improving the knowledge of essential climate variables by exploiting Earth Observation (EO) datasets from past, present and future European Space Agency (ESA) satellite missions. The project is built around the measure of grounding line migration at the ice–ocean interface. The results presented here focus on the Pine Island Glacier (PIG) and Petermann Gletscher glacier, two key regions of the Antarctic and Greenland Ice Sheets. The PIG is the most dynamic ice covered region of the AIS, showing the fastest retreat and largest mass loss of all AIS regions; here we seek to map the grounding line retreat since it was last precisely measured in 2000. The Petermann Gletscher is the largest outlet

N. Gourmelen (✉)
School of GeoSciences, University of Edinburgh, Edinburgh, UK
e-mail: noel.gourmelen@ed.ac.uk

J. Park
Yonsei University, Seoul, South Korea

A. Shepherd
CPOM, University of Leeds, Leeds, UK



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_3

glacier of northern GIS, showing very little change over the last 20 years [21]; using novel EO data we seek to revisit the current state of stability of the Petermann Gletscher outlet glacier.

## 2 Retreat of the Grounding Line of the Pine Island Glacier, West Antarctic Ice Sheet

### *2.1 Introduction*

The West Antarctic Ice Sheet (WAIS) contains enough ice to raise eustatic sea level by over 3 m [2], and the Amundsen Sea sector of the WAIS is susceptible to accelerated retreat due to the presence of a bedrock topography that lies well below sea level and deepens inland and the absence of substantial floating ice shelf barriers [17]. Satellite observations show that glaciers draining this sector are retreating [22], thinning [28], accelerating [13] and losing mass [25]. Observations of ice shelf thinning [29] in the face of increased glacier discharge and a numerical simulation of glacier response to external forcing [19] suggest that the surrounding ocean is the source of this imbalance. The PIG is a major tributary of the WAIS Amundsen Sea sector. Satellite observations of relative tidal motion show that the PIG hinge-line retreated by up to 25 km between 1992 and 2009 [14, 22]. This retreat corresponds to a reduction in ice thickness of around 90 m at the glacier terminus considering the recent geometry [33], a value that is consistent with direct observations of thinning acquired by satellite altimetry during the same period [35].

The potential sea level contribution due to ice mass losses from the Amundsen Sea sector over the twenty-first century is a source of considerable uncertainty [16]. A hypothetical scaling of glacier discharge rates indicates a potential twenty-first century sea level contribution in the range of 4–15 cm from the PIG alone [20]. However, an extrapolation of the recent PIG volume trend acceleration [35] provides a much smaller estimated contribution of around 2 cm by the year 2100—a value that is consistent with the 1.8 cm likely estimate of a basin-scale model of the glacier response to ocean forcing [14]. While changes occurring in the vicinity of the PIG grounding line are expected to cause thinning inland for decades to come [13, 19], the presence of an extended region of lightly grounded ice at the glacier terminus [4] has promoted rapid retreat over recent years. Based on the geometry farther inland, and even allowing for an increase in ocean melting, the grounding line has been expected to stabilise soon and for several decades [14]. Here, we use satellite radar interferometry and satellite radar altimetry to analyse the rate of hinge-line retreat and thinning to establish whether the glacier has reached this anticipated state of relative stability.

## 2.2 Methods

We use interferometric synthetic aperture radar (InSAR) data acquired by the European Remote Sensing (ERS-1 and ERS-2) satellite systems to measure changes in the position of the PIG hinge-line over the period 1992–2011 [6, 7, 22]. The InSAR dataset consists of synthetic aperture radar (SAR) images recorded in 1992, 1996, 1999, 2000 and 2011 during dedicated periods when the satellites orbited in short-repeat cycles. The temporal spans (baselines) of the InSAR data vary between 6 days (1992), 3 days (2011) and 1 day (1996, 1999 and 2000). In order to locate the glacier hinge-line (the limit of tidal flexure) we subtract consecutive interferograms, corrected for the effects of topography. This procedure eliminates the signal of ice flow that is common to each individual interferogram, and reveals the relative surface motion due mainly to tidal flexure of the floating glacier tongue. We then mapped the locus of the glacier hinge-line at different time periods (Fig. 1) by minimising the departure between the observed tidal flexure and that of a model elastic beam. Based on the degree of data misfit to this model, and the tidal phase gradient in the vicinity of the hinge-line, we estimate the uncertainty in hinge-line position to be 0.13, 0.35, 0.51 and 0.29 km in 1992, 1996, 1999–2000 and 2011, respectively. The origin of these uncertainties lies in high-frequency noise of the InSAR measurement; we do not observe a systematic long wavelength misfit that could indicate a departure of the observed flexure from the elastic beam model assumption that is commonly used to locate hinge-line positions [23, 32].

Since 1992, the PIG hinge-line has retreated by as much as 28.4 km along the central section of flow (Fig. 1). However, the degree of hinge-line migration has varied in space and time, and so we computed the average rate of hinge-line retreat along different regions of the grounding zone (Table 1). In the first instance, we calculated the rate of hinge-line retreat along three adjacent 5 km wide sections spanning northern (N5), central (C5) and southern (S5) portions of the glacier trunk (Fig. 1), to provide a detailed picture of how the migration has occurred over time. We also calculated the rate of hinge-line retreat along a 10 km wide central section (C10, Fig. 1) of the glacier to facilitate a comparison with satellite altimeter observations acquired over a similar area [28]. The average rate of hinge-line retreat is computed as the reduction in grounded area between successive hinge-line positions, divided by the section width perpendicular to the direction of ice flow. Considering these sections, retreat varies from 9.5 to 27 km. On average, an area of $14.2 \pm 1.3$ km$^2$ has become ungrounded each year (Fig. 2a).

Changes in the position of the PIG hinge-line correspond to changes in the ice thickness due to hydrostatic equilibrium [22, 31]. We used this relationship to calculate the rate of ice thinning associated with hinge-line retreat from the InSAR dataset. The change in ice thickness over an intervening time period $\dot{h}$ (positive for thickening) is given by

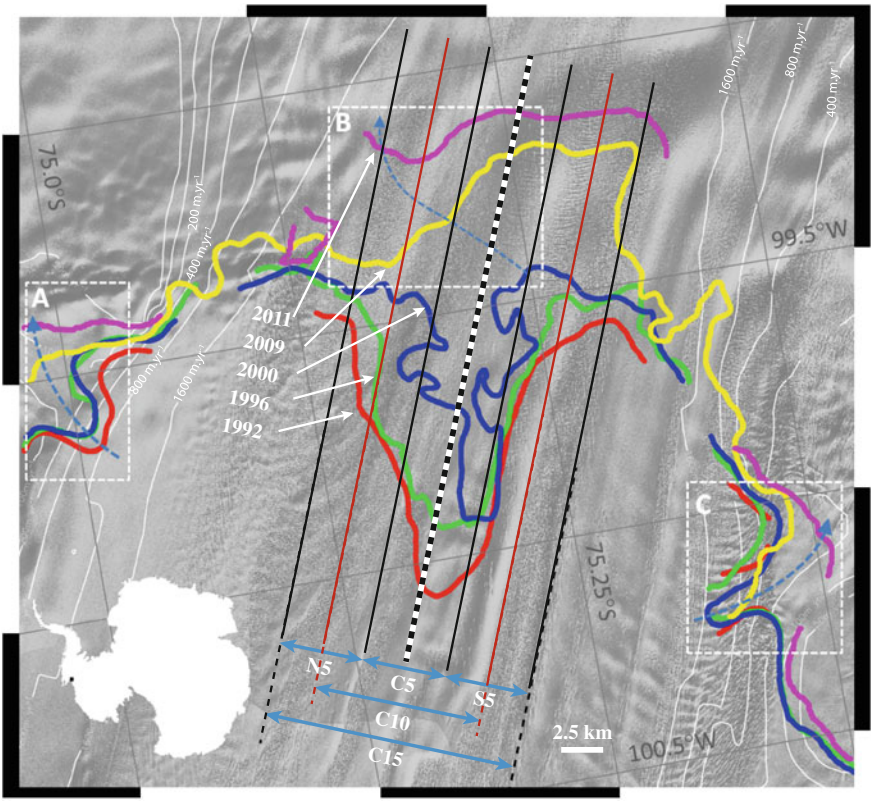


**Fig. 1** Hinge-line positions of the Pine Island Glacier (*coloured lines*) superimposed on a Landsat image (*greyscale*). Successive hinge-line positions are marked with coloured lines (*red* = 1992, *green* = 1994, *blue* = 2000, *yellow* = 2009, *magenta* = 2011). *White* contours show ice velocity (m year$^{-1}$). Average rates of hinge-line retreat (Table 1; Fig. 2) are calculated within the northern, central and southern 5 km sections (N5, C5 and S5, respectively), and within the central 10 and 15 km sections (C10 and C15, respectively). *Dotted line* marks the profile along which the satellite radar altimeter data are acquired (Fig. 3). The greatest hinge-line retreat has occurred towards the centre of the fast flowing glacier. *Boxes A, B* and *C* highlight other regions of notable hinge-line retreat (*blue dotted arrows*). Rapid hinge-line retreat has occurred beyond the margins of the glacier in the vicinity of *boxes A* and *C* in recent years, despite little apparent change in the glacier surface geometry, suggesting increased ocean melting has occurred in these regions. Rapid hinge-line retreat in the vicinity of *box B* has occurred, despite the apparent position of stability in 2009, due to a combination of factors; a narrow channel of lightly grounded ice has favoured retreat in a northwest direction in tandem with an evolution of the ice surface slope, which has also promoted retreat

$$\dot{h} = \left(\frac{\rho_w}{\rho_i}\right)\dot{z} - \left[\alpha - \beta\left(1 - \frac{\rho_w}{\rho_i}\right)\right]\dot{x} \tag{1}$$

**Table 1** Retreat rate of the PIG hinge line along parallel sections of the glacier breadth (see Fig. 1)

| | Northern 5 km, N5 (km year$^{-1}$) | Central 5 km C5 (km year$^{-1}$) | Southern 5 km S5 (km year$^{-1}$) | Central 10 km C10 (km year$^{-1}$) | Central 15 km C15 (km year$^{-1}$) |
|---|---|---|---|---|---|
| 1992–1996 | 0.48 + 0.24 | 0.75 + 0.17 | 0.64 + 0.27 | 0.67 + 0.22 | 0.63 + 0.23 |
| 1996–2000 | 1.21 + 0.23 | 1.00 + 0.21 | 0.54 + 0.26 | 1.36 + 0.23 | 0.92 + 0.23 |
| 2000–2009[a] | 0.35 + 0.13 | 1.37 + 0.15 | 0.80 + 0.16 | 1.21 + 0.15 | 0.84 + 0.15 |
| 2009[a]–2011 | 2.81 + 0.72 | 1.23 + 0.78 | 1.32 + 0.81 | 1.53 + 0.87 | 1.79 + 0.87 |
| 2000–2011 | 0.79 + 0.09 | 1.33 + 0.12 | 0.89 + 0.13 | 1.26 + 0.09 | 1.00 + 0.09 |

[a]Using the hinge-line located in 2009 with coarse resolution speckle-tracked range offsets [16]

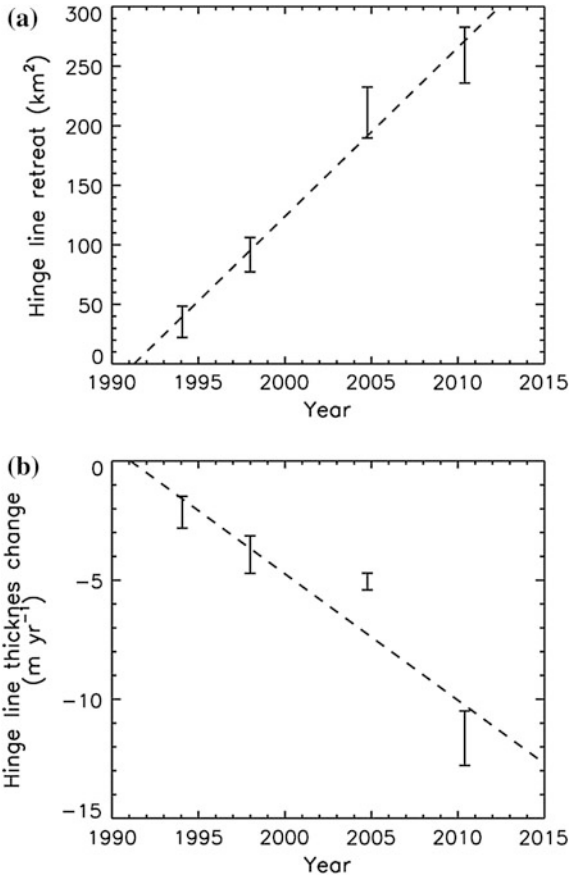


**Fig. 2** **a** Retreat and **b** ice thickness change at the hinge-line of the Pine Island Glacier. Retreat area is computed within the central 15 km of the glacier (see Fig. 1) to capture the entire signal. Thickness changes are computed within the central 10 km for ease of comparison with satellite altimeter data (Fig. 3)

where $\dot{z}$ is the change in ocean tide, $\dot{x}$ (positive for retreat) is the change in hinge line position, $\rho_w$ is the density of sea water = 1027.5 kg $m^{-3}$ [8], $\rho_i$ is the density of ice = 900 kg $m^{-3}$ [10], and $\alpha$ and $\beta$ are the slopes of the ice surface and base, respectively. The local balance between surface and basal slopes governs the rate of hinge-line migration for a given change in ice thickness. Repeat observations of ice surface elevation (e.g. [35]) suggest that the surface slope in the vicinity of the PIG hinge-line has altered considerably over time.

To account for this variation, we estimate $\alpha$ averaged along the location and between the dates of successive hinge-line positions using a sequence of glacier surface slopes generated from elevation models derived from satellite and airborne data [1, 5, 15]. We estimate $\beta$ averaged along the location of each hinge-line position using a model of the bedrock elevation [33]. In the absence of direct observations, we use a model of Antarctic Ocean tide (an updated version of the regional inverse model described by Padman et al. [18]) to assess the fluctuation in grounding line position due to tidal displacement. However, because tide models tend to perform less well when predicting tidal phase [27], we estimate the uncertainty associated with $\dot{z}$ using the maximum modelled tidal range in the vicinity of the PIG (1.8 m) following a conservative approach [22], and by taking account of variability in the basal slope [3]. Using these assumptions, we determined the rates of ice thickness change in the vicinity of the PIG grounding line over the survey period (Fig. 2b) associated with the observed hinge-line retreat.

We obtained additional estimates of ice thickness changes inland of the PIG hinge-line from repeat-pass satellite radar altimeter data. For this exercise, we processed a continuous record of data acquired by the ERS-1, ERS-2 and ENVISAT satellite radar altimeters between 1994 and 2010. Time-series of surface elevation change were developed at crossing points of the satellites' ground tracks falling within the PIG drainage basin during 35-day orbit repeat mission phases, using the method of dual cycle cross-overs [36]. Elevation measurements were corrected for the lag of the leading edge tracker, surface scattering variation, dry atmospheric mass, water vapour, the ionosphere, solid Earth tide and ocean loading tide [28, 34, 35]. To cross-calibrate the observations recorded by successive satellites, we corrected for the differences between the average elevation changes occurring during periods of mission overlap. At each crossing point we formed a time-series of elevation change, and these data were then averaged into 10 km by 10 km grid cells. For each grid cell we computed the average rate of elevation change during discrete time periods (e.g. Fig. 3), and we estimated the error associated with each time-series according to the variance of the data.

## 2.3 Results

Although rates of hinge-line retreat have been high near to the centre of the PIG, they diminish rapidly towards the glacier margins. Within the central 15 km of the glacier (C15, Fig. 1; Table 1), the average rate of hinge-line retreat has remained broadly

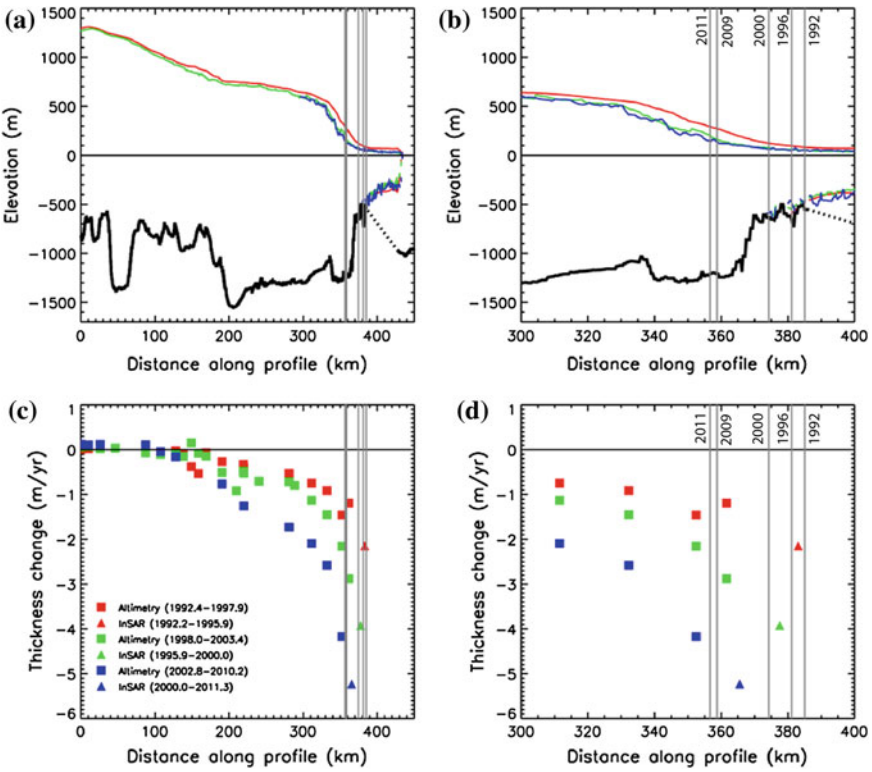


**Fig. 3** **a**, **b** Geometry and **c**, **d** rate of ice thickness change along a profile of the Pine Island Glacier (see Fig. 1) during three time intervals. Thickness changes are determined using either satellite altimetry (*squares*) or satellite InSAR (*triangles*). *Vertical grey lines* chart migrating glacier hinge line position which, since 2009, has been situated in a location where the bedrock has shoaled

constant between 1992 and 2011 at $0.95 \pm 0.09$ km year$^{-1}$. However, there have also been considerable temporal variations in the hinge-line retreat rate over the course of our survey, with episodes of rapid and asymmetric migration. During the periods 1992–2000, 2000–2009 and 2009–2011, for example, hinge-line retreat was concentrated within the central, southern and northern sections of the PIG, respectively. The peak rate of hinge-line retreat ($2.8 \pm 0.7$ km year$^{-1}$) occurred within the northernmost 5 km wide section of the PIG during the most recent period captured by our InSAR data (2009–2011). These irregular patterns can be explained in part by the limiting influence of the glacier geometry. For a constant rate of ice thinning, for example, the degree of hinge-line retreat is dependent on the surface and bedrock slopes which vary in space and, in the case of the ice surface, over time.

The manner in which the PIG geometry develops upstream of the hinge-line tends to impede retreat, because the bedrock and surface slopes shoal and steepen inland, respectively (Fig. 3a). In consequence, the near-constant rate of hinge-line retreat during our survey must reflect an accelerating rate of ice thinning (Fig. 2b).

Within the central 10 km of the PIG, the rate of ice thinning associated with hinge-line retreat has increased, progressively, from 2.1 ± 0.7 m year$^{-1}$ between 1992 and 1996 to 11.6 ± 1.1 m year$^{-1}$ between 2009 and 2011. According to the InSAR data, the average rate of ice thinning at the PIG hinge-line has accelerated by 0.53 ± 0.15 m year$^{-2}$ over the 19-year survey period.

We compared rates of hinge-line ice thinning determined from the InSAR dataset with those determined farther inland using satellite radar altimetry. According to the altimeter dataset, ice thinning is observed to start around 200–250 km inland and increase sharply towards the PIG hinge-line (Fig. 3). To facilitate a more detailed comparison, we estimated rates of ice thinning during successive time periods when the respective datasets were coincident (1992–1998, 1998–2003, and 2000–2010). These periods were selected to optimise the degree of temporal overlap, the duration of the altimeter data and the provenance of the altimeter data in relation to the three satellite platforms. Rates of ice thinning determined from InSAR along the central 10 km of the PIG match closely those determined from altimetry (Fig. 3), illustrating that the independent techniques are complementary. Along a stream-wise profile of the PIG, temporally averaged thinning rates are greater at the hinge-line (as determined by InSAR) by a proportion that is consistent with the rate of increase farther inland (as determined by altimetry). InSAR thus provides observation beyond the locus of the altimeter data of the migration of the calving margin, where flux variability can be estimated with better accuracy.

## *2.4 Discussion*

The rate of hinge-line retreat we have recorded may be compared with that determined during the overlapping periods of previous surveys [14, 22, 24]. For example, average retreat rates of 0.64 and 0.55 km year$^{-1}$ have been estimated across a 14 km wide centre profile [22] during the periods 1992–1996 and 1996–2000, respectively, as compared to our estimated retreat rates of 0.63 ± 0.23 and 0.92 ± 0.23 km year$^{-1}$ across a partially overlapping 15 km wide profile. The hinge-line position we have recorded in the year 2000 is, however, considerably less regular than that of a previous survey [24], showing greater retreat to the south of the glacier centre-line. When compared to the 2009 hinge-line position located using technique of radar speckle tracking [14], which has coarse resolution, the 2011 location is between 2 and 5 km farther inland—more than double the average rate of retreat since 1992 (Table 1). The hinge-line positions are irregularly shaped, and retreat within the central portion of the glacier has been asymmetrical at all times. Rapid migration tends to occur along relatively confined sections of the hinge-line. Prior to 2009, for example, steep surface slopes to the north of the PIG hinge-line presented an obstacle to retreat in that sector (Fig. 1). Since then, the hinge-line has retreated across a submarine bedrock ridge (see Fig. 3b), and the principal trajectory of retreat has been northwards and inland, eroding a grounded promontory that was present in 2009.

Accelerated thinning of ice at the grounding line of the PIG has occurred in the face of well-documented increases in the rate of glacier discharge [25], indicating that ocean-driven melting has substantially exceeded the additional ice inflow. This conclusion is consistent with the findings of a survey of changes in oceanographic conditions at the glacier terminus over the period 1994–2009 [9], which reveals strengthening ocean circulation within an enlarged sub-ice-shelf cavity leading to faster ice melting. The emergence of an enlarged ocean cavity has resulted, in part, from retreat of the glacier across a submarine bank upon which it was formerly grounded [11], allowing relatively warm ($\sim$4 °C above freezing) seawater to access the glacier grounding line. The geometry of the glacier in the vicinity of 2011 hinge-line does not favour retreat, and further retreat is at odds with simulations of the PIG evolution under conditions of increased ocean melting [14]. However, similar conditions were present at other times during our survey, and more recent simulations of the PIG evolution that utilise an adaptive (finer) grid suggest that the grounding line may retreat farther inland if ocean melting persists (S. Cornford, personal communication). It is also possible that the increased ocean melting has

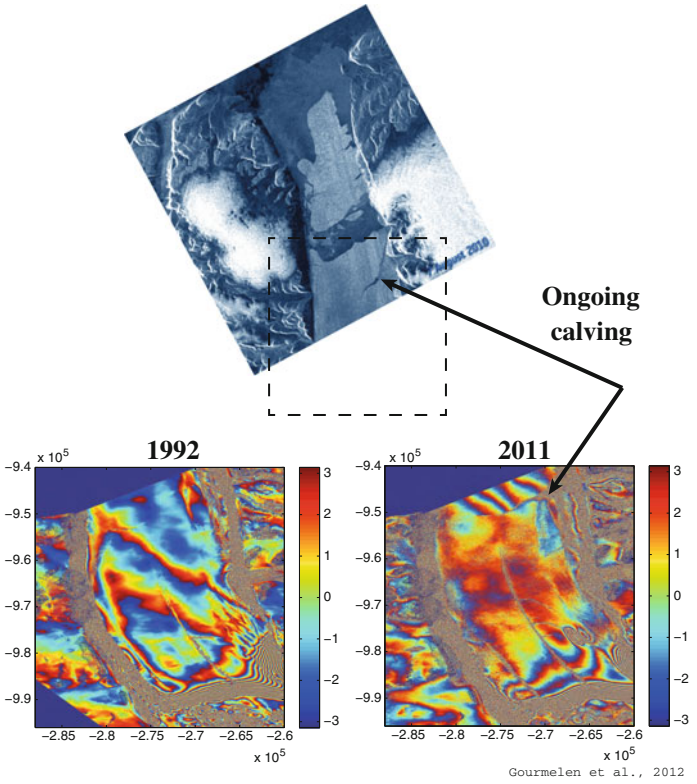


**Fig. 4** Initial stage of the 2012 calving event on the floating tongue of Petermann glacier, seen in the double difference interferograms performed in the years 1992 (*lower left*) and 2011 (*lower right*)

exceeded that imposed during existing simulations. Similar geometrical obstacles to retreat were, for example, present at other times during our survey, and retreat has progressed at other times presumably as a result of changes in both the rate of ocean melting and the ice surface slope (a dynamical response to reduced grounding [13]). Although the hinge-line has retreated since 2009, analysis of glacier velocities reveals that the glacier speed has yet to increase (I. Joughin, personal communication).

## 3 Grounding Line Retreat of Petermann Gletscher, Greenland Ice Sheet

Surface displacement maps over the Petermann glacier are calculated using datasets from the European Remote Sensing, ERS 1&2 Synthetic Aperture Radar, SAR, satellites [6, 7, 12, 26]. The SAR dataset consists of radar images collected in 1992 and 2011 during specific 3-day repeat phase, the so-called "ice phase" carried out by the European Space Agency. In order to locate the grounding line, we subtract two consecutive topography-corrected interferograms (i.e. double difference approach—Figs. 2 and 4); this process eliminates the ice flow signal common to the two interferograms, only remaining the difference of surface movement related to

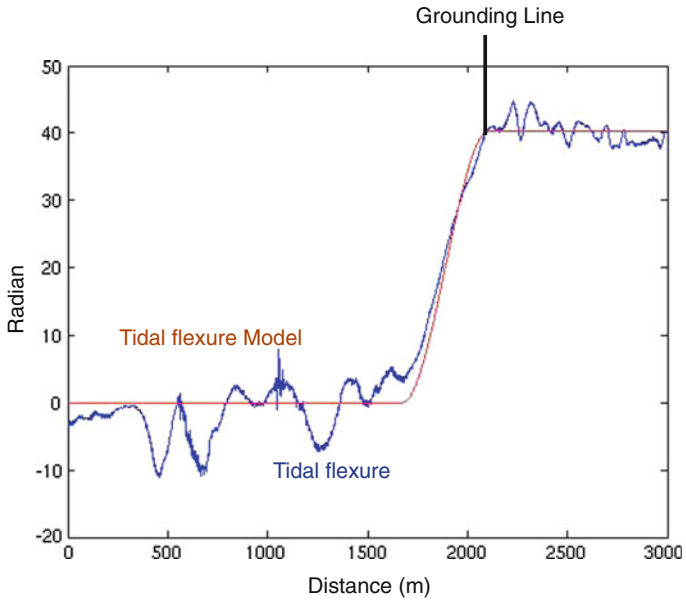


**Fig. 5** Profile over the double difference interferograms showing tidal flexure. The grounding line is localised at the onset of flexure between grounded (*right portion of profile*) and floating (*left portion of profile*) sections of the glacier using a best-fitting elastic beam model

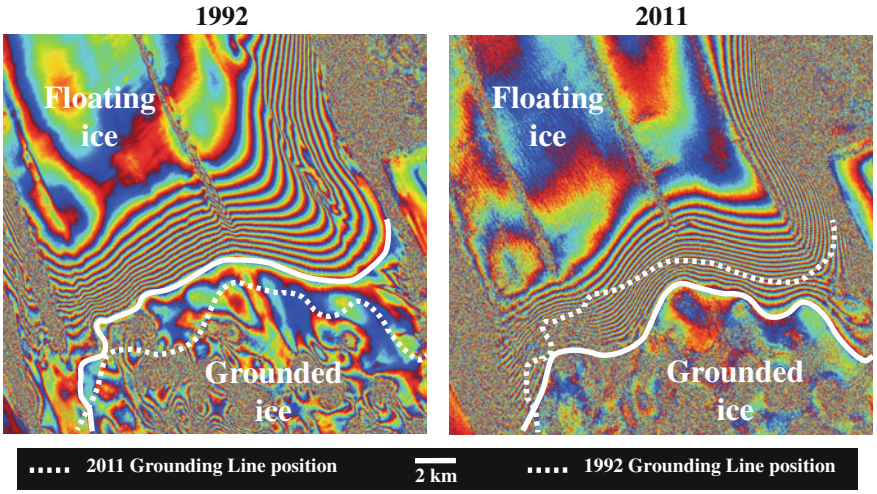


**Fig. 6** Double difference interferogram for the years 1992 (*left*) and 2011 (*right*) showing tidal fringes and grounding line migration of 2 km

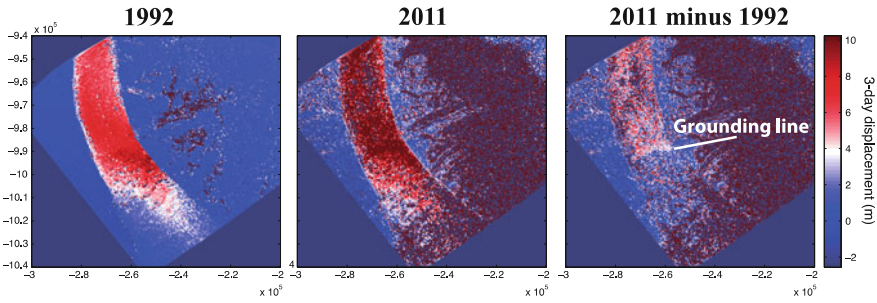


**Fig. 7** Displacement of the Petermann glacier obtained from speckle tracking applied to SAR dataset acquired in 1992 (*left*), 2011 (*right*), and displacement difference showing displacement gradient at the location of the grounding line

variation in tidal flexure of the floating tongue (Fig. 2) [22]. We note that a similar approach based on speckle tracking gives similar results although with a degraded spatial resolution (Fig. 5). The grounding line position is then mapped by minimising the residuals between the observed flexure and an elastic beam model of flexure (Fig. 4) [22].

Here we use the detailed surface topography and bedmap datasets obtained by the University of Kansas (The Center for Remote Sensing of Ice Sheets, CReSIS) to compute detailed thinning rates across the glacier's grounding line (Fig. 6). Our estimated retreat rate errors are a combination of the data—beam model residuals, the strength of the tidal signal in the double difference interferograms, and of the

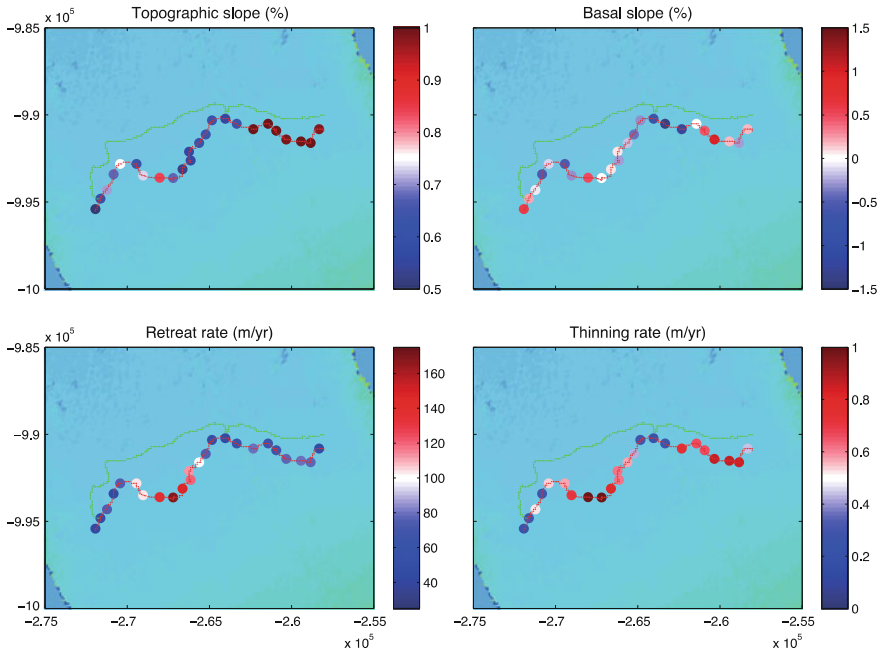


**Fig. 8** Petermann glacier in the region surrounding the grounding line (*thin green* (1992) and *red* (2011) *line*). The *coloured dots* represent spatial variation of surface slope (*top left*), basal slope (*top right*), grounding line retreat (*bottom left*) and thinning rate (*bottom right*)

tidal range. Mean retreat of the grounding line between 1992 and 2011 is 2 km, a mean rate of 100 m year$^{-1}$. This reveals a constant grounding line retreat since 1992 as the rate of retreat for the 1992–2011 period is similar to the rate calculated for the period 1992–1996 [22] (Figs. 7 and 8).

## 4 Conclusions

The GreenSAR project has performed the analysis of the 2011 ERS2 ice phase, the last ERS2 acquisitions before decommissioning of the satellite, over the Greenland and Antarctic Ice Sheets. We have shown:

At Pine Island Glacier:

- The grounding line of the PIG is retreating at a constant rate of 0.95 km year$^{-1}$ since 1992
- The melting at the grounding line of the PIG is accelerating since 1992 at a rate of 0.53 m year$^{-2}$
- Based on these observations, the lower limits of sea level projection might be too conservative

At Petermann Gletscher:

- The grounding line retreat has been constant since 1992 at a mean rate of 100 m $year^{-1}$
- Thinning at the grounding line has been constant since 1992 at a mean rate of 0.8 m $year^{-1}$
- This constant thinning is at odds with observation of surface height stability from IceSat measurements

This study confirms the dramatic changes affecting the Amundsen Sea sector of the Western Antarctic Ice Sheet with no signs of slowing down.

**Acknowledgments** I thank Diego Fernandez, Yves-Louis Desnos, Marcus Engdhal, Roberto Sabia from ESA for their support. We acknowledge the use of data and/or data products from CReSIS generated with support from NSF Grant ANT-0424589 and NASA Grant NNX10AT68G. DEMs used are the ASTER GDEM (http://gdem.ersdac.jspacesystems.or.jp) and GIMP DEM (http://bprc.osu.edu/GDG/gimpdem.php).

## References

1. Bamber JL, Bindschadler RA (1997) An improved elevation dataset for climate and ice-sheet modelling: validation with satellite imagery. Ann Glaciol 25:439–444
2. Bamber JL, Riva REM, Vermeersen BLA, LeBrocq AM (2009) Reassessment of the potential sea-level rise from a collapse of the West Antarctic Ice Sheet. Science 324:901–903. doi:10.1126/science.1169335
3. Bindschadler R, Choi H, Wichlacz A, Bingham R, Bohlander J, Brunt K, Corr H, Drews R, Fricker H, Hall M, Hindmarsh R, Kohler J, Padman L, Rack W, Rotschky G, Urbini S, Vornberger P, Young N (2011) Getting around Antarctica: new high-resolution mappings of the grounded and freely-floating boundaries of the Antarctic Ice Sheet created for the international polar year. Cryosphere Discuss 5:183–227. doi:10.5194/tcd-5-183-2011
4. Corr HFJ, Doake CSM, Jenkins A, Vaughan DG (2001) Investigations of an "ice plain" in the mouth of Pine Island Glacier, Antarctica. J Glaciol 47(156):51
5. DiMarzio J, Brenner A, Schutz R, Shuman CA, Zwally HJ (2007) GLAS/ICESat 1 km laser altimetry digital elevation model of Greenland. National Snow and Ice Data Center, Boulder, Colorado, USA (Digital media)
6. Gabriel AK, Goldstein RM, Zebker HA (1989) Mapping small elevation changes over large areas—differential radar interferometry. J Geophys Res-Solid Earth Planets 94:9183–9191
7. Goldstein RM, Engelhardt H, Kamb B, Frolich RM (1993) Satellite radar interferometry for monitoring ice-sheet motion—Application to an Antarctic ice stream. Science 262:1525–1530
8. Jacobs SS, Hellmer HH, Jenkins A (1996) Antarctic ice sheet melting in the Southeast Pacific. Geophys Res Lett 23(9):957
9. Jacobs SS, Jenkins A, Giulivi CF, Dutrieux P (2011) Stronger ocean circulation and increased melting under Pine Island Glacier ice shelf. Nat Geosci 4:519. doi:10.1038/ngeo1188
10. Jenkins A, Vaughan DG, Jacobs SS, Hellmer HH, Keys JR (1997) Glaciological and oceanographic evidence of high melt rates beneath the Pine Island Glacier, West Antarctica. J Glaciol 43(143):114
11. Jenkins A, Dutrieux P, Jacobs SS, McPhail SD, Perrett JR, Webb AT, White D (2010) Observations beneath Pine Island Glacier in West Antarctica and implications for its retreat. Nat Geosci 3(7):468

12. Joughin IR, Winebrenner DP, Fahnestock MA (1995) Observation of ice sheet motion in Greenland using Satellite Radar Interferometry. Geophys Res Lett 22:571–574
13. Joughin I, Rignot E, Rosanova CE, Lucchitta BK, Bohlander J (2003) Timing of recent accelerations of Pine Island Glacier, Antarctica. Geophys Res Lett 30(13):1706
14. Joughin I, Smith BE, Holland DM (2010) Sensitivity of 21st century sea level to ocean induced thinning of Pine Island Glacier, Antarctica. Geophys Res Lett 32(L20502). doi:10.1029/2010GL044819
15. Krabill WB (2011) IceBridge ATM L2 ice elevation, slope, and roughness. National Snow and Ice Data Center, Boulder, Colorado USA (Digital media)
16. Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Nada A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C (2007) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds) Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, Cambridge
17. Mercer JH (1978) West Antarctic Ice sheet and $CO_2$ greenhouse effect: a threat of disaster. Nature 271:321–325
18. Padman L, Fricker HA, Coleman R, Howard S, Erofeeva L (2002) A new tide model for the Antarctic ice shelves and seas. Ann Glaciol 34:247
19. Payne AJ, Vieli A, Shepherd A, Wingham DJ, Rignot E (2004) Recent dramatic thinning of largest West Antarctic ice stream triggered by oceans. Geophys Res Lett 31(23):L23401
20. Pfeffer WT, Harper JT, O'Neel S (2008) Kinematic constraints on glacier contributions to 21st-century sea-level rise. Science 321(5894):1340–1343
21. Pritchard HD, Arthern RJ, Vaughan DG, Edwards LA (2009) Extensive dynamic thinning on the margins of the Greenland and Antarctic ice sheets. Nature 461:971–975
22. Rignot E (1998) Hinge-line migration of Petermann Gletscher, north Greenland, detected using satellite-radar interferometry. J Glaciol 44:469–476
23. Rignot EJ (1998) Fast recession of a West Antarctic glacier. Science 281:549–551
24. Rignot E (2002) Ice-shelf changes in Pine Island Bay, Antarctica, 1947-2000. J Glaciol 48 (161):247
25. Rignot E (2008) Changes in West Antarctic ice stream dynamics observed with ALOS PALSAR data. Geophys Res Lett 35(12)
26. Rignot E, Jezek EC, Sohn HG (1995) Ice flow dynamics of the Greenland Ice-Sheet from Sar interferometry. Geophys Res Lett 22:575–578
27. Shepherd A, Peacock NR (2003) Ice shelf tidal motion derived from ERS altimetry. J Geophys Res 108(C6):3198
28. Shepherd A, Wingham DJ, Mansley JAD, Corr HFJ (2001) Inland thinning of Pine Island Glacier, West Antarctica. Science 291(5505):862
29. Shepherd A, Wingham D, Rignot E (2004) Warm ocean is eroding West Antarctic Ice Sheet. Geophys Res Lett 31(23)
30. Shepherd A, Ivins ER, Geruo A, Barletta VR, Bentley MJ, Bettadpur S, Briggs KH, Bromwich DH, Forsberg R, Galin N, Horwath M, Jacobs S, Joughin I, King MA, Lenaerts JT, Li J, Ligtenberg SR, Luckman A, Luthcke SB, McMillan M, Meister R, Milne G, Mouginot J, Muir A, Nicolas JP, Paden J, Payne AJ, Pritchard H, Rignot E, Rott H, Sørensen LS, Scambos TA, Scheuchl B, Schrama EJ, Smith B, Sundal AV, van Angelen JH, van de Berg WJ, van den Broeke MR, Vaughan DG, Velicogna I, Wahr J, Whitehouse PL, Wingham DJ, Yi D, Young D, Zwally HJ (2012) A reconciled estimate of ice-sheet mass balance. Science 338:1183–1189. doi:10.1126/science.1228102
31. Thomas RH, Bentley CR (1978) A model for Holocene retreat of the West Antarctic Ice Sheet. Quat Res 10:150
32. Vaughan DG (1995) Tidal flexure at ice shelf margins. J Geophys Res 100:6213–6224

33. Vaughan DG, Corr HFJ, Ferraccioli F, Frearson N, O'Hare A, Mach D, Holt JW, Blankenship DD, Morse DL, Young DA (2006) New boundary conditions for the West Antarctic Ice sheet: subglacial topography beneath Pine Island Glacier. Geophys Res Lett 33(9)
34. Wingham DJ, Ridout AJ, Scharroo R, Arthern RJ, Shum CK (1998) Antarctic elevation change from 1992 to 1996. Science 282(5388):456–458
35. Wingham DJ, Wallis DW, Shepherd A (2009) Spatial and temporal evolution of Pine Island Glacier thinning, 1995–2006. Geophys Res, Lett 36
36. Zwally HJ, Brenner AC, Major JA, Bindschadler RA, Marsh JG (1989) Growth of Greenland ice-sheet—Measurement. Science 246(4937):1587

# Sea Surface Roughness Manifestations Around Ocean Fronts

**Nicolas Rascle, Bertrand Chapron, Frédéric Nouguier and Alexis Mouche**

**Abstract** Ocean fronts are often visible at the ocean surface as well-marked choppy rough water or, contrarily, as anomalously smooth sea surface. As such, high-resolution satellite images—e.g., obtained by synthetic aperture radars (SARs) or by radiometers viewing areas in and around the sun glitter—at times can provide clear observations of mesoscale and submesoscale oceanic fronts. These observations have thus a powerful potential to monitor the upper ocean dynamics, by providing essential information on oceanic fronts. In that perspective, we review recent advances in the qualitative and quantitative interpretation of satellite surface roughness anomalies.

## 1 Introduction

Gradients of surface currents at oceanic fronts are often qualitatively reported as choppy water with anomalously steep and breaking waves or, contrarily, as anomalously smooth sea surface. As wide swath images of sea surface roughness are now routinely obtained using active or passive sensors—such as microwave radars including synthetic aperture radar (SAR) [5] or optical radiometers viewing areas in and around the sun glitter [3]—a wide range of oceanic phenomena are often reported to manifest at scales of about 1–30 km, including internal waves, mesoscale and submesoscale features such as fronts, filaments, and spiraling eddies [1, 10]. An example of surface roughness in the sun glitter is shown in Fig. 1.

N. Rascle (✉) · B. Chapron · A. Mouche
Laboratoire d'Océanographie Spatiale, Institut Francais de Recherche pour l'Exploitation de la Mer (IFREMER), ZI Pointe du Diable, 29280 Plouzané, France
e-mail: nicolas.rascle@ifremer.fr

F. Nouguier
Université de Toulon, CNRS/INSU, IRD, Mediterranean Institute of Oceanography (MIO), UM 110, 83957 La Garde, France



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_4

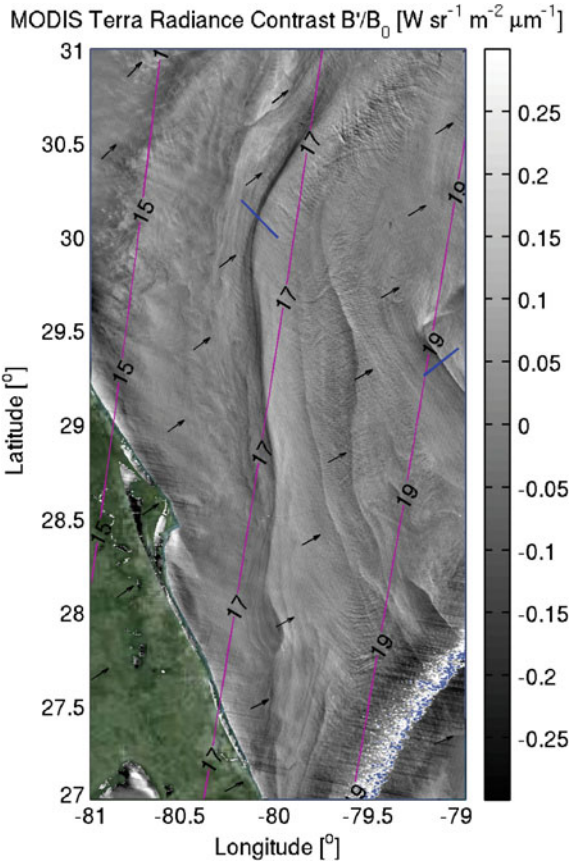


**Fig. 1** Radiance contrasts $B'/B_0$ around the sun glitter in the Gulf Stream region on April 1st, 2010 by MODIS Terra. Contours and *arrows* show the zenith and azimuth angles of the reflective facets. The coast of Florida is shown in *green* and clouds in *white*. More details on the image and on the two transects shown in *blue* are given in Sect. 7

Surface roughness contrasts can be essentially related to three mechanisms. (I) The presence of surfactants, possibly accumulated in zones of surface current convergence, that can dampen the short gravity waves [9]. (II) The modification of the sea surface temperature that can alter the atmospheric boundary layer and thus modify the wind, which in turns modifies the short waves [4]. (III) A surface current, and more specifically its spatial gradient, that can directly refract the waves [14].

Hereafter, we focus on mechanism (III), which can be confidently isolated from the other mechanisms. Mechanism (I) due to convergence processes and surfactants is likely limited to very low wind speed [21]. Mechanism (II) involves modification of the atmospheric boundary layer which seemingly occurs at larger spatial scales than the current refraction [17]. Moreover, the effects of wind modification (II) and current refraction (III) have been recently shown to be separable using SAR measurements at different polarizations [18].

Based on the conservation of wave action, Phillips [22] studied the modulation of an individual wave over a varying current and later extended to modulations of a complete spectrum of wind waves [8].

Surface roughness variations in sun glitter images are controlled by specularly reflected sunlight [6], and related to changes of wave slope distribution, mostly related to variations of mean square slope (mss) [20]. For radar images, the backscatter intensity is less directly affected by specular reflection, and other geometrical properties of the surface must be more precisely taken into account [17]. Details on how to estimate mss from glitter brightness are presented in [20]. For simplicity, we assume an unequivocal relationship between surface roughness and mss, which is an adequate assumption at least for optical observations.

In the present paper, we wish to address and review the following questions: What is the surface roughness signature of oceanic fronts? In particular, what are the signs of roughness anomalies? Does their observability depend on the wind direction, on the sensor look angles? Does it vary whether the sensor is sensitive to short or long waves? How does the propagation of waves modify the observation of the front spatial extent?

Following the framework proposed by Kudryavtsev et al. [17], we consider the conservation of wave action in a relaxation approximation [2], Sect. 2. The wave refraction by current is given in Sect. 3 and, neglecting wave propagation, the mss anomalies are related to divergence of surface currents and strain in the wind direction, Sect. 4, as previously reported [24]. In Sect. 5, effects of wave propagation are included. Kudryavtsev et al. [17] proposed that a dominant mechanism of surface roughness modulation by currents is the generation of short waves by longer breaking ones. This effect is discussed in Sect. 6. With a set of three simultaneous sun glitter images of the Gulf Stream, obtained at different azimuth angles, interpretation is made following the proposed framework, Sect. 7, and conclusions are given in Sect. 8.

## 2 Wave–Current Interactions and Surface Roughness

### *2.1 Surface Waves in Currents*

In the presence of current, the conservation of wave action $N(\mathbf{x}, \mathbf{k}, t)$, related to the spectral displacement level at a fixed wavenumber, reads [22]

$$\frac{\partial N}{\partial t} + \left(c_{gi} + u_i\right)\frac{\partial N}{\partial x_i} = k_j \frac{\partial u_j}{\partial x_i}\frac{\partial N}{\partial k_i} + S_{in} + S_{nl} - S_{ds}, \tag{1}$$

where $\mathbf{c_g}$ is the group velocity, $\mathbf{u}$ is the current, $\mathbf{k}$ is the wavenumber, and $S_{xx}$ are the action sources which include input from the wind, wave–wave non-linear interactions, and dissipation. Repeated indices $i, j = 1\ldots2$ indicate summation over horizontal components.

## 2.2 Relaxation Time

Following a relaxation approach [2, 12, 15], the action is supposed to experience small disturbance $N'$ with respect to a background value $N_0$, $N(\mathbf{x},\mathbf{k},t) = N_0(\mathbf{k}) + N'(\mathbf{x},\mathbf{k},t)$, with $N' \ll N_0$, leading to:

$$\frac{\partial N'}{\partial t} + \left(c_{gi} + u_i\right)\frac{\partial N'}{\partial x_i} = k_j \frac{\partial u_j}{\partial x_i}\frac{\partial N_0}{\partial k_i} - \frac{N'}{\tau_c}, \tag{2}$$

where the source terms act to restore equilibrium and are linearized with a relaxation time $\tau_c(\mathbf{k})$ as

$$\frac{1}{\tau_c} = -\frac{\partial}{\partial N}(S_{in} + S_{nl} - S_{ds}). \tag{3}$$

If all action source components are function of the action level, with the wind input written as $S_{in} = \beta N$, with $\beta$ the wind input factor, the relaxation time will take the form $\tau_c = (n\beta)^{-1}$, with $n(k)$ related to the energy dissipation [26].

The relaxation length is defined from the relaxation time as $l_c = c_g \tau_c$. In the downwind direction, the wind input factor is approximately $\beta \propto \omega(u_*/c)^2$ [23], with $u_*$ the wind friction velocity, $\omega$ the intrinsic frequency, and $c$ the phase speed, to give $l_c \propto gk^{-2}u_*^{-2}$.

Figure 2 shows the relaxation length calculated using the model of Kudryavtsev et al. [17]. More details on the calculation, especially in the crosswind and upwind directions, are given in their paper. We note that, at a fixed wind speed, this length scale decreases as $k^{-2}$, and shorter waves adjust quickly back to an equilibrium state. We also note that Fig. 2 uses a wind speed of 5 m s$^{-1}$ but $l_c$ at a given wavenumber $k$ decreases with the square of the wind speed ($l_c \propto u_*^{-2}$).

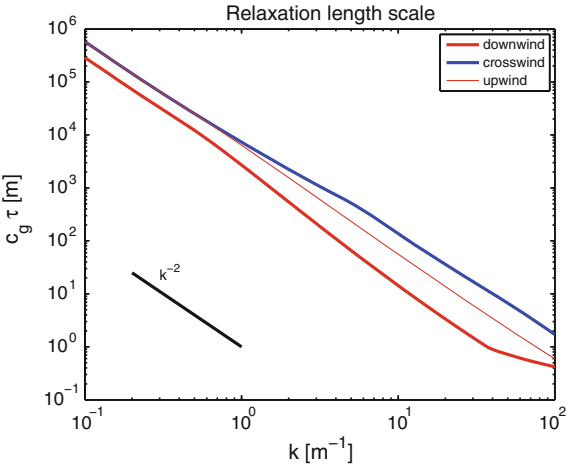


**Fig. 2** Relaxation length scale $l_c = c_g\tau_c$ for a wind speed of 5 m s$^{-1}$, as function of the wavenumber $k$. The different *curves* correspond to upwind propagating waves ($\phi = \phi_w + 180°$), crosswind waves ($\phi = \phi_w + 90°$), and downwind waves ($\phi = \phi_w$)

### 2.3 Changes of Mean Square Slope

From the wave action anomaly $N'$, we can calculate the scale distributed changes of roughness properties.

In this paper, we will make the assumption that those properties, which may differ with sensor geometries and characteristics, can all be related to the wave mean square slopes (mss) in the two main axes which are the upwind and the crosswind directions.

The mss anomalies in the upwind and crosswind directions are calculated from the wave action anomaly by

$$\begin{aligned} \mathrm{mss}'_u(\mathbf{x}) &= \int_k \int_\phi \omega^{-1} k N' k^2 \cos^2(\phi - \phi_\mathrm{w}) \mathrm{d}k k \mathrm{d}\phi, \\ \mathrm{mss}'_c(\mathbf{x}) &= \int_k \int_\phi \omega^{-1} k N' k^2 \sin^2(\phi - \phi_\mathrm{w}) \mathrm{d}k k \mathrm{d}\phi, \end{aligned} \tag{4}$$

where $\phi_\mathrm{w}$ is the wind direction related to the $x$-axis.

As the mean square slope quantities change, the overall density of breaking elements will necessarily be modulated to provide contrast in high-resolution radar and/or glitter images.

Additionally, a sensor is sensitive to the distribution of slopes in the azimuthal direction $\varphi_r$ which reflects the signal. That reflective direction is the sensor look direction for a mono-static configuration (e.g., radar), whereas for a bi-static configuration as in the sun glint case, it is the direction of the specular facets between the sun and the radiometer. In this paper we make the assumption that the sensor is sensitive to the anomaly of the mss projected in the reflective direction $\varphi_r$, which writes

$$\mathrm{mss}'_{\varphi_r} = \cos^2 \varphi_r \mathrm{mss}'_u + \sin^2 \varphi_r \mathrm{mss}'_c. \tag{5}$$

## 3 Current Refraction

### 3.1 Canonical Current Gradients

The action source term due to current variation is

$$\begin{aligned} R_u(x,k) &= k_j \frac{\partial u_j}{\partial x_i} \frac{\partial N_0}{\partial k_i} \\ &= [\, k_x \quad k_y \,] \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial N_0}{\partial k_x} \\ \frac{\partial N_0}{\partial k_y} \end{bmatrix} \end{aligned} \tag{6}$$

where notations $x$, $y$, and $u$, $v$ are equivalent to notations $x_1$, $x_2$ and $u_1$, $u_2$. This term can be interpreted as an advection of action in the phase space, at a velocity equal to the product of the first two tensors on the left-hand side.

This term writes in polar coordinates

$$\begin{aligned} R_u(\mathbf{x},k,\phi) = N_0 \Bigg[ & \frac{\partial u}{\partial x}\left(\cos^2(\phi)m_k - \cos(\phi)\sin(\phi)m_\phi\right) \\ & + \frac{\partial u}{\partial y}\left(\cos(\phi)\sin(\phi)m_k + \cos^2(\phi)m_\phi\right) \\ & + \frac{\partial v}{\partial x}\left(\cos(\phi)\sin(\phi)m_k - \sin^2(\phi)m_\phi\right) \\ & + \frac{\partial v}{\partial y}\left(\sin^2(\phi)m_k + \cos(\phi)\sin(\phi)m_\phi\right)\Bigg] \end{aligned} \tag{7}$$

where $k$ and $\phi$ are the wavenumber magnitude and direction, and where we define $m_k = \partial \ln N_0 / \partial \ln k$ and $m_\phi = \partial \ln N_0 / \partial \phi$.

## 3.2 Centered Current Gradients

The current deformation tensor can be separated in divergence $D$, vorticity $V$, strain $S_t$, and shear $S_h$,

$$\begin{aligned} D &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}, \quad S_t = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}, \\ V &= \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}, \quad S_h = \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}, \end{aligned} \tag{8}$$

$$\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} D+S_t & -V+S_h \\ V+S_h & D-S_t \end{bmatrix}. \tag{9}$$

Using that form of the current deformation tensor, the refraction source term can be rewritten as

$$\begin{aligned} R_u(\mathbf{x},k,\phi) = \frac{N_0}{2}\big[ & Dm_k - Vm_\phi \\ & + S_t\big(\cos(2\phi)m_k - \sin(2\phi)m_\phi\big) \\ & + S_h\big(\sin(2\phi)m_k + \cos(2\phi)m_\phi\big)\big]. \end{aligned} \tag{10}$$

With $\phi_{\mathrm{w}}$ the wind angle from the $x$-axis, the strain $S_t^w$ and shear $S_h^w$ in the wind direction express from the strain $S_{\mathrm{t}}$ and shear $S_{\mathrm{h}}$ in the $x$-direction as

$$\begin{bmatrix} S_{\mathrm{t}}^{\mathrm{w}} \\ S_{\mathrm{h}}^{\mathrm{w}} \end{bmatrix} = \begin{bmatrix} \cos 2\phi_{\mathrm{w}} & \sin 2\phi_{\mathrm{w}} \\ -\sin 2\phi_{\mathrm{w}} & \cos 2\phi_{\mathrm{w}} \end{bmatrix} \begin{bmatrix} S_{\mathrm{t}} \\ S_{\mathrm{h}} \end{bmatrix}. \tag{11}$$

The divergence $D$ and vorticity $V$ do not depend on any chosen direction.

## 4 First Approximation: Without Propagation

In this section we assume that the typical length scale $L$ of the current is much larger than the relaxation length scale $l_c = \tau_c c_g$ of the waves impacting the surface roughness. In that case the advection term on the left-hand side of (1) can be ignored. In a steady state, the action anomaly $N'$ locally responds as

$$N'(\mathbf{x}, \mathbf{k}) = \tau_c R_u. \tag{12}$$

### *4.1 Wind Wave Spectral Symmetry*

We will now assume that the angular spread of the wave spectrum is symmetrical about the wind direction, which is a good approximation for the short wind waves [7]. We can set the $x$-axis in the wind direction ($\phi_{\mathrm{w}} = 0$), and $N_0(\mathbf{x}, k, \phi)$ is then an even function of $\phi$. The same property is transferred to the relaxation time $\tau_c(k, \phi)$. Note that $m_k N_0$ is even, whereas $m_\phi N_0$ is odd. The integrals in (4) over the wave direction $\phi$ of any odd function will then cancel out.

Now expanding the four terms of (10), the second ($V$) and the fourth ($S_{\mathrm{h}}$) terms create an action anomaly $N'(\mathbf{x}, k, \phi)$ which is an odd function of $\phi$, i.e., the action anomaly of waves traveling in the direction $\phi$ is exactly opposite to that of waves traveling in the direction $-\phi$. As a key consequence, vorticity $V$ and shear $S_{\mathrm{h}}$ do not create mss anomaly in (4). Only divergence $D$ and strain $S_{\mathrm{t}}$ could create roughness changes.

If the wind is not in the $x$-direction (i.e., if $\phi_{\mathrm{w}} \neq 0$), the previous conclusion holds in a frame of reference aligned with the wind. That is, only divergence $D$ and strain in the wind direction $S_{\mathrm{t}}^{\mathrm{w}}$ modulate mss, whereas vorticity $V$ and shear in the wind direction $S_{\mathrm{h}}^{\mathrm{w}}$ have no impact.

### *4.2 Divergence, Strain, and Polarization Index*

Out of four types of current deformation, only two—divergence $D$ and strain in the wind direction $S_t$—create mss anomalies. Accordingly, the current tensor in (9) can be reduced to its diagonal terms.

The mss anomaly can thus be rewritten from (10) and (4) in the form

$$\mathrm{mss}'_u = [D + b_u S_t] \times a_u, \tag{13}$$

where the parameters $a_u$ and $b_u$ only depend on the properties of the wave field.

The amplitude parameter $a_u$ is defined as

$$a_u = \int_k \int_\phi \omega^{-1} \tau_c k^3 \frac{N_0}{2} m_k \cos^2 \phi \, \mathrm{d}k k \mathrm{d}\phi, \tag{14}$$

and the so-called polarization parameter $b_u$ is defined as

$$b_u = \frac{\int_k \int_\phi \omega^{-1} \tau_c k^3 N_0 \left(\cos(2\phi) m_k - \sin(2\phi) m_\phi\right) \cos^2 \phi \, \mathrm{d}k k \mathrm{d}\phi}{\int_k \int_\phi \omega^{-1} \tau_c k^3 N_0 m_k \cos^2 \phi \, \mathrm{d}k k \mathrm{d}\phi}. \tag{15}$$

Similarly the parameters $a_c$ and $b_c$ can be defined for the crosswind component of the mss, replacing $\cos^2 \phi$ by $\sin^2 \phi$ in the above.

The amplitude parameters $a_u$ and $a_c$ determine the mss response to a current divergence. They are negative for all waves shorter than the spectral peak, because $m_k < 0$. This means that divergent (convergent) currents will always appear as a negative (positive) mss anomaly on surface roughness images.

The parameters $b_u$ and $b_c$ determine the mss response to a current strain. They also depend on the angular distribution of the wave spectrum. As found, $b_u$ is positive, i.e., for an upwind looking instrument (i.e., sensitive to the upwind mss anomaly), a positive (negative) strain will appear as a negative (positive) anomaly, similar to a divergence (convergence) effect. Interestingly, $b_c$ can be negative, i.e., a current strain can induce a crosswind mss anomaly with an opposite sign to that of the upwind mss anomaly.

More details can be found in [24].

## 5 One-Dimensional Cases

Considering ocean frontal structures with small spatial scale, typically less than 10 km, but uniform along one direction, for instance along the $y$-axis, two types of current deformation are of interest, $\partial u/\partial x$ and $\partial v/\partial x$.

### *5.1 Case 1: Acrossfront Current 'divergence'*

We first consider case 1 where the current is normal to a front (Fig. 3) of spatial length scale $L$. As a numerical example, the current is specified as

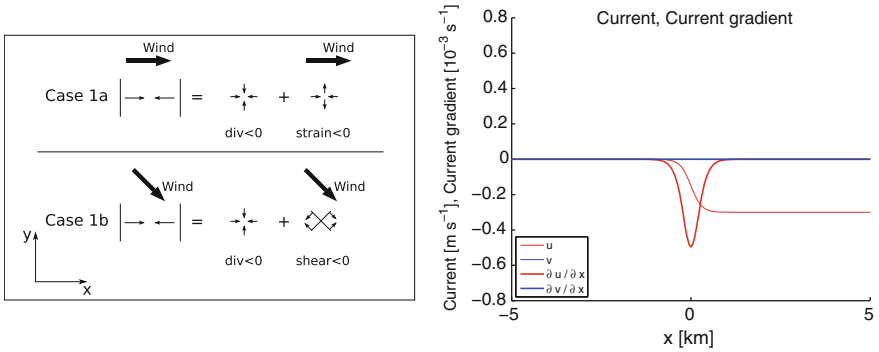


**Fig. 3** Overview of the case 1, i.e., oceanic front with acrossfront current. *Left* Sketch of the current and of the current deformation along the wind direction. *Right* Current and current gradient as function of the $x$-coordinate

$$u(x) = u_0 \frac{1 + \tanh(x/L)}{2}, \quad v = 0, \tag{16}$$

with $u_0 = -0.3$ m s$^{-1}$ and $L = 300$ m. The wind is set to 5 m s$^{-1}$ westerly ($\phi_w = 0$, case 1a) or northwesterly ($\phi_w = -45°$, case 1b). Such acrossfront current is the sum of a convergence and a negative strain in the $x$-direction (case 1a), which is also the sum of a convergence and a negative shear in the northwest direction (case 1b).

The mss anomaly for case 1a is highlighted in Fig. 4. The mss is separated into mss supported by long waves ($k < 0.6$ m$^{-1}$), short waves ($k > 30$ m$^{-1}$), and intermediate waves.

In Fig. 4a, the wave propagation is neglected. The mss anomalies are positive for all waves, as expected from a convergent current. Only waves longer than the spectral peak (for which $m_k > 0$ and thus $a_u > 0$ and $a_c > 0$, see Sect. 4.2) produce a small negative anomaly, visible on the crosswind mss of long waves.

Once the propagation is included (Fig. 4b), the mss anomaly is now shifted toward the downwind side. This is particularly marked for long waves ($k < 0.6$ m$^{-1}$) which have typical relaxation length $l_c$ larger than 10 km. Intermediate waves have relaxation length of the order of 1 km. Short waves ($k > 30$ m$^{-1}$) have a relaxation length scale $l_c$ shorter than 10 m. Anomalies follow the local current gradient, but are small due to very short relaxation time [see, e.g. (12)].

In Fig. 4c, we switch on the generation of short waves by wave breaking. This will be discussed in Sect. 6.

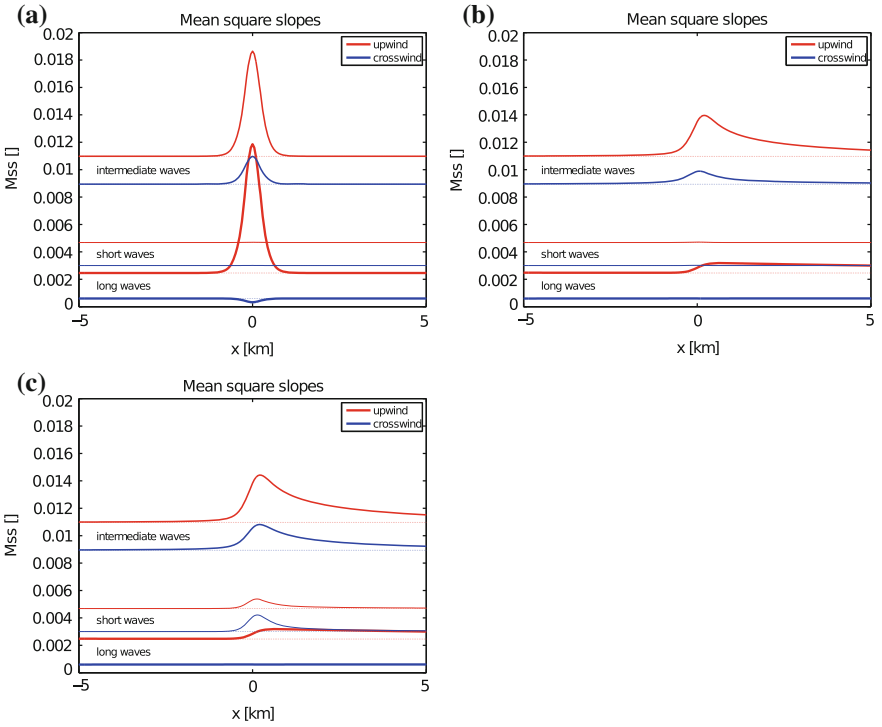


**Fig. 4** Mss response to case 1a (acrossfront convergent current with acrossfront wind). The mss is separated into mss supported by long waves ($k < 0.6$ m$^{-1}$), short waves ($k > 30$ m$^{-1}$), and intermediate waves. **a** Without propagation and without wave breaking. **b** With propagation but without wave breaking. **c** With propagation and wave breaking. Note that without propagation in **a**, the magnitude of the current has been reduced by half in order to remain in the range of validity of our assumption $N' \ll N_0$

## 5.2 *Case 2: Alongfront Current 'shear'*

We now consider case 2 where the current is along the front (Fig. 5). For the numerical example, the current is specified as

$$u = 0, \quad v(x) = u_0 \frac{1 + \tanh(x/L)}{2}, \tag{17}$$

with again $u_0 = -0.3\,\mathrm{m\,s^{-1}}$ and $L = 300\,\mathrm{m}$. The wind is still set to 5 m s$^{-1}$ and we investigate three directions, acrossfront (westerly, $\phi_w = 0$, case 2a), oblique (northwesterly, $\phi_w = -45°$, case 2b), and alongfront (northerly, $\phi_w = -90°$, case 2c). Such alongfront current is the sum of a positive vorticity and a positive shear in the $x$-direction (case 2a), which is also the sum of a positive vorticity and a negative strain in the northwest direction (case 2b) or the sum of a positive vorticity and a negative shear in the $y$-direction (case 2c).

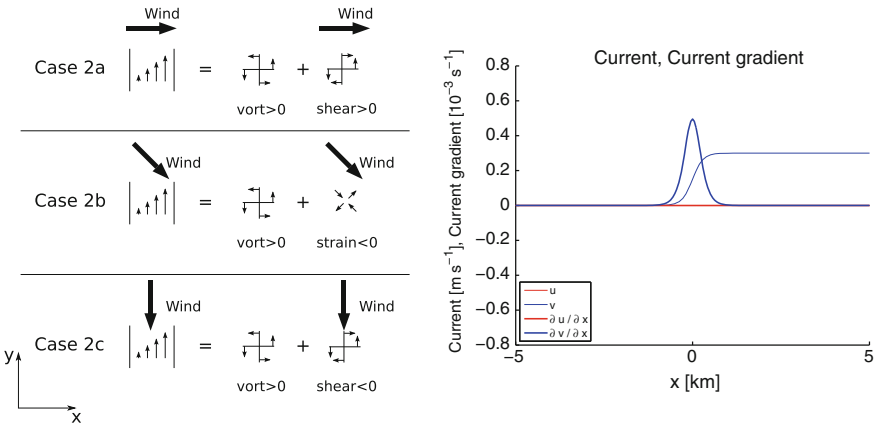


**Fig. 5** Overview of the case 2, i.e., oceanic front with alongfront current. *Left* Sketch of the current and of the current deformation along the wind direction. *Right* Current and current gradient as function of the *x*-coordinate

### 5.2.1 Without Propagation

We first neglect the wave propagation. When the wind is normal to the front (case 2a) or along the front (case 2c), there is no divergence or strain in the wind direction. As discussed in Sect. 4, the action anomaly $N'(k, \phi)$ is anti-symmetrical about the wind direction (Fig. 6a, c). No mss anomaly (not shown) is thus expected once integrated over the direction $\phi$.

On the other hand, when the wind is oblique to the front (case 2b), the strain in the wind direction creates an action anomaly which is not anti-symmetrical about the wind direction (Fig. 6b). A mss anomaly is thus expected once integrated over the direction $\phi$.

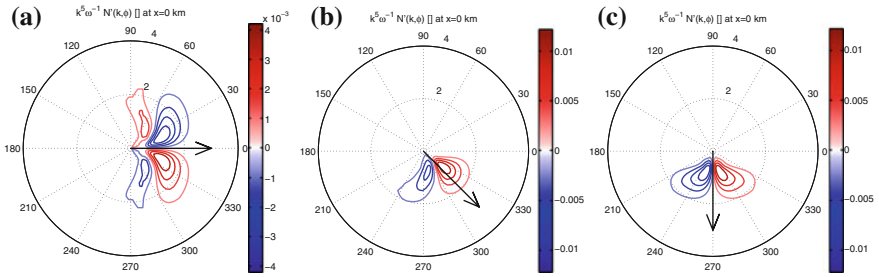


**Fig. 6** Spectrum of the saturation anomaly $k^5\omega^{-1}N'(k, \phi)$ at the center of the front ($x = 0$). The radial coordinate scale is $\log(k) + 2$. The *black arrow* shows the wind direction for **a** Case 2a, **b** Case 2b, and **c** Case 2c. Note that propagation and wave breaking are not included in those calculations

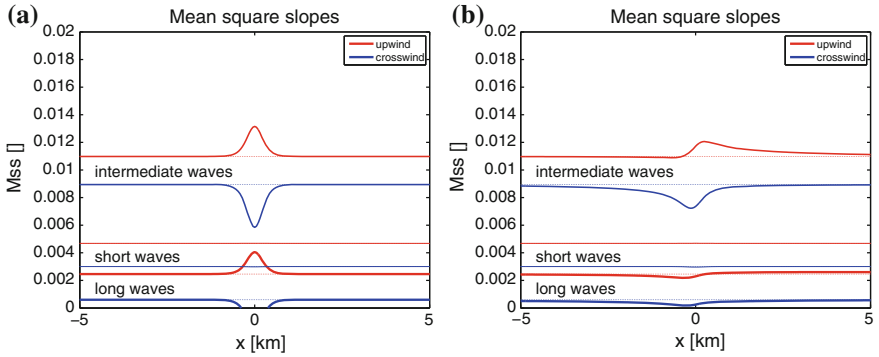


**Fig. 7** Mss response to case 2b (alongfront current with oblique wind). The mss is separated into mss supported by long waves ($k < 0.6$ m$^{-1}$), short waves ($k > 30$ m$^{-1}$), and intermediate waves. **a** Without propagation and without wave breaking. **b** With propagation but without wave breaking. **c** With propagation and wave breaking. Note that without propagation in **a**, the magnitude of the current has been reduced by half in order to remain in the range of validity of our assumption $N' \ll N_0$

The mss anomaly for case 2b is shown in Fig. 7a. The mss anomaly is due to the negative current strain in the wind direction, as sketched in Fig. 5. As a result, waves traveling in the upwind direction are compressed and create a positive mss anomaly, whereas waves traveling in the crosswind direction are elongated and create a negative mss anomaly. Short waves produce a negligible mss anomaly due to their short relaxation time.

### 5.2.2 With Propagation

In Fig. 7b, we switch on the wave propagation. For intermediate waves, the upwind mss anomaly is shifted downwind, as in the case of a convergent current. However, the anomalies in the crosswind direction and in both directions for the long waves are more difficult to explain. For that purpose, we shall get back to case 2a and case 2c.

In case 2a of a wind normal to the front, the action anomaly is symmetrical about the $x$-axis (see Fig. 6a). As a consequence, the geometry being uniform along the $y$-direction, the advection by the group velocity $c_g$ cannot create any action anomaly. On the contrary in case 2c of wind along the front, the group velocity can advect the negative (positive) action anomaly outside of the front toward the $x < 0$ ($x > 0$) (see Fig. 8). This effect of propagation is dominant in case 2c because without propagation, there is no strain in the wind direction and thus no mss anomaly.

In case 2b, this effect of propagation is of second order. It only slightly modifies the mss anomalies compared to the expected anomaly due to strain in the wind direction (compare Fig. 7a, b).

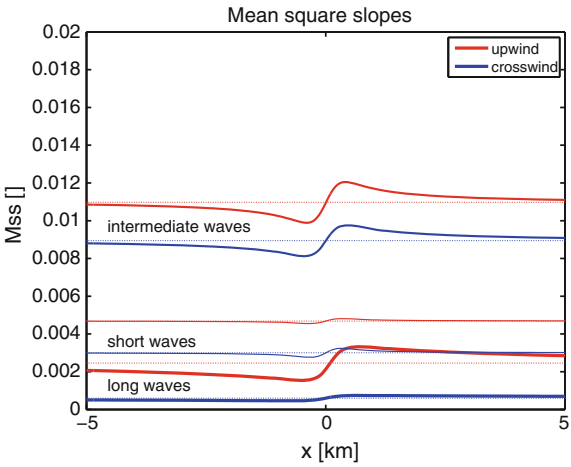


**Fig. 8** Mss response to case 2c (alongfront current with alongfront wind). The mss is separated into mss supported by long waves ($k < 0.6$ m$^{-1}$), short waves ($k > 30$ m$^{-1}$), and intermediate waves. Propagation and wave breaking are included

## 6 Generation of Short Waves by Wave Breaking

To take into account the plausible mechanism of generation of short waves by longer breaking waves, Kudryavtsev et al. [17] added a term in the action conservation (1), to specify an isotropic source of action

$$Q_{\mathrm{wb}}(k) = \omega^2 k^{-5} I_{\mathrm{wb}}, \tag{18}$$

proportional to the dimensionless rate of longer breaking waves

$$I_{\mathrm{wb}}(k) \propto \omega^{-1} \int_{\phi} \int_{k<k_m} \omega^{-1} \beta k^3 N \, \mathrm{d}k k \mathrm{d}\phi, \tag{19}$$

where $k_m = k/10$ sets the shortest (but long) breaking waves which can produce mechanical energy at wavenumber $k$.

This term plays a crucial role in the definition of the background wave spectrum (without current) of short gravity waves. This role is certainly very modest for downwind propagating waves, but is dominant for crosswind propagating waves. More specifically, as crosswind propagating waves do not experience strong direct wind input, the background spectral level is specified from the balance between dissipation and generation by wave breaking (see Fig. 1 of [17]). Also, upwind propagating waves are in equilibrium between generation by wave breaking and negative energy input from the wind.

This wave-breaking term also plays a crucial role in the response of short gravity waves to current because the modulation of the wave-breaking source dominates the modulation of action due to current refraction (see Fig. 7 of [17]).

The effect of wave breaking has been calculated for the idealized one-dimensional fronts. The result for case 1a is shown in Fig. 4c. Long and

intermediate waves experience most of the modulation due to current refraction (Fig. 4b). Those waves, when the energy source from their breaking is included, in turn create large mss anomalies of the short waves (Fig. 4c). Although the breaking energy source (19) is isotropic, its impact is larger on the crosswind mss than on the upwind mss. The reason is that, as stated above, the wave-breaking source is the dominant source of crosswind propagating waves, whereas downwind propagating waves are closely related to wind input, and thus have shorter relaxation times (see e.g. Fig. 2).

Similarly, the effect of wave breaking for case 2b is shown in Fig. 7c. We note that the wind input factor of Plant [23] is $\beta(k,\phi) \propto (u_*/c)^2 |\cos(\phi - \phi_w)| \cos(\phi - \phi_w)$. Therefore, the term $I_{wb}$ is mainly related to downwind propagating waves. As a result both the upwind and crosswind mss anomalies of short waves follow the upwind mss anomaly of intermediate waves, which is positive in case 2b, and they do not follow the crosswind anomaly of intermediate waves, which is negative in that case.

## 7 Observations

Case 1 of divergent/convergent front has been discussed many times with images of internal waves [1, 3]. That particular case does not show any change of surface roughness anomalies with the sensor look angle, because upwind and crosswind mss anomalies have similar signs (Fig. 4). Here, we wish to illustrate from multiple look angles case 2 of non-divergent front with current shear.

We use a set of three optical images around the sun glitter area in the Gulf Stream region obtained within a few hours interval on April 1st, 2010. Those images are shown in Figs. 1, 9, and 10 and were obtained from (a) the 900 nm radiance measured by MEdium Resolution Imaging Spectrometer (MERIS) [13] on board *Envisat*, at 15:42 UTC, (b) and (c) the 850 nm radiance measured by MODerate resolution Imaging Spectroradiometer (MODIS) [25] on board *Terra* at 16:30 UTC, and on board *Aqua* at 18:05 UTC, respectively. The three images have a horizontal resolution of about 250 m. The observed radiance $B = B_0 + B'$ is separated into a slowly varying background $B_0$ and a local anomaly $B'$ by horizontally filtering at a scale $L$ = 25 km.

There is a transfer function between the radiance contrast $B'/B_0$ and the mss contrast $\mathrm{mss}'/\mathrm{mss}_0$ [20]. The details of this transfer function will be described elsewhere. In this paper for the sake of simplicity, it will be ignored and we limit our discussion to the radiance contrast.

Sun glitter observations operate in a bi-static configuration and are sensitive to the amount of reflective surface facets between the sun and the radiometer. The zenith and azimuth angles of those reflective facets are shown in Figs. 9 and 10. As assumed, the radiance contrast in that azimuthal direction is related to the contrast of the mss (of all waves, i.e., long, intermediate, and short waves) projected along the same direction.

A transect in the Gulf Stream internal front (the northern one in blue on Fig. 1) is detailed in Fig. 9. The wind was blowing approximately from the northwest, as evidenced by the presence of wind streaks [11]. If we make the hypothesis that the Gulf Stream current was non-divergent and flowing northward, then the wind and current have a configuration similar to that of case 2b. The radiance contrast from MERIS, which is looking upwind, is positive. The contrast from MODIS Terra, which is looking crosswind, is negative. Those are in good qualitative agreement with the upwind and crosswind mss contrasts expected for case 2b (see Fig. 7c for the intermediate waves, as those waves carry most of the mss and thus dominate the total mss anomaly). We also note that the spatial shift between the negative minimum of MODIS Terra and the positive maximum of MERIS is also in qualitative agreement with the model calculations for case 2b. Finally, MODIS Aqua, which is looking with a slanting angle, does not show any significant contrast, which is also well in agreement with the calculations.

A second transect in another front of the Gulf Stream (the southern one in blue in Fig. 1) is detailed in Fig. 10. The wind is also northwesterly there and thus almost aligned with the front. If we again make the hypothesis that the current is non-divergent and with a current shear as in case 2, the wind and current have then

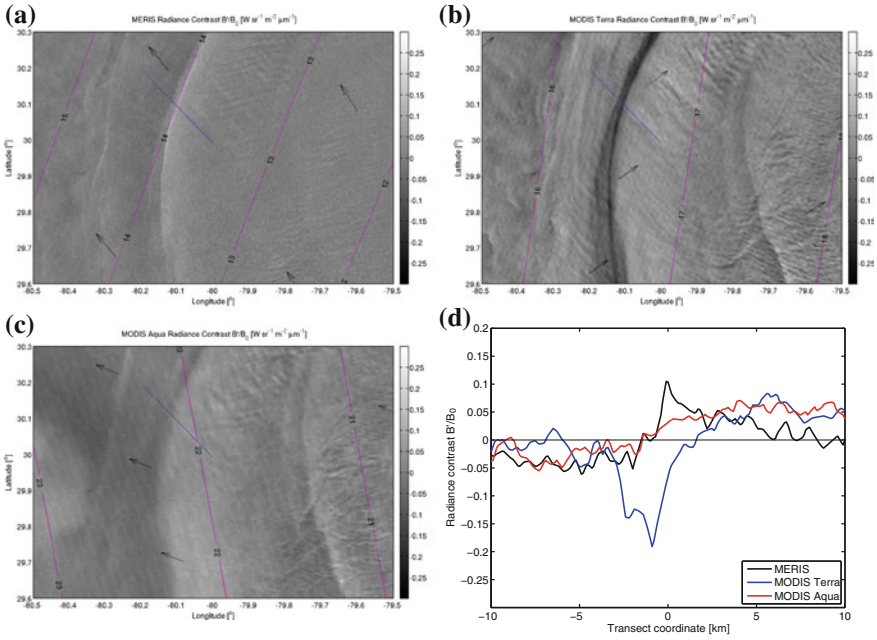


**Fig. 9** Radiance contrasts $B'/B_0$ around the sun glitter in the Gulf Stream region on April 1st, 2010, by **a** MERIS, **b** MODIS Terra, and **c** MODIS Aqua. Contours and *arrows* show the zenith and azimuth angles of the reflective facets. **d** Details of the transect shown in *blue* on the images

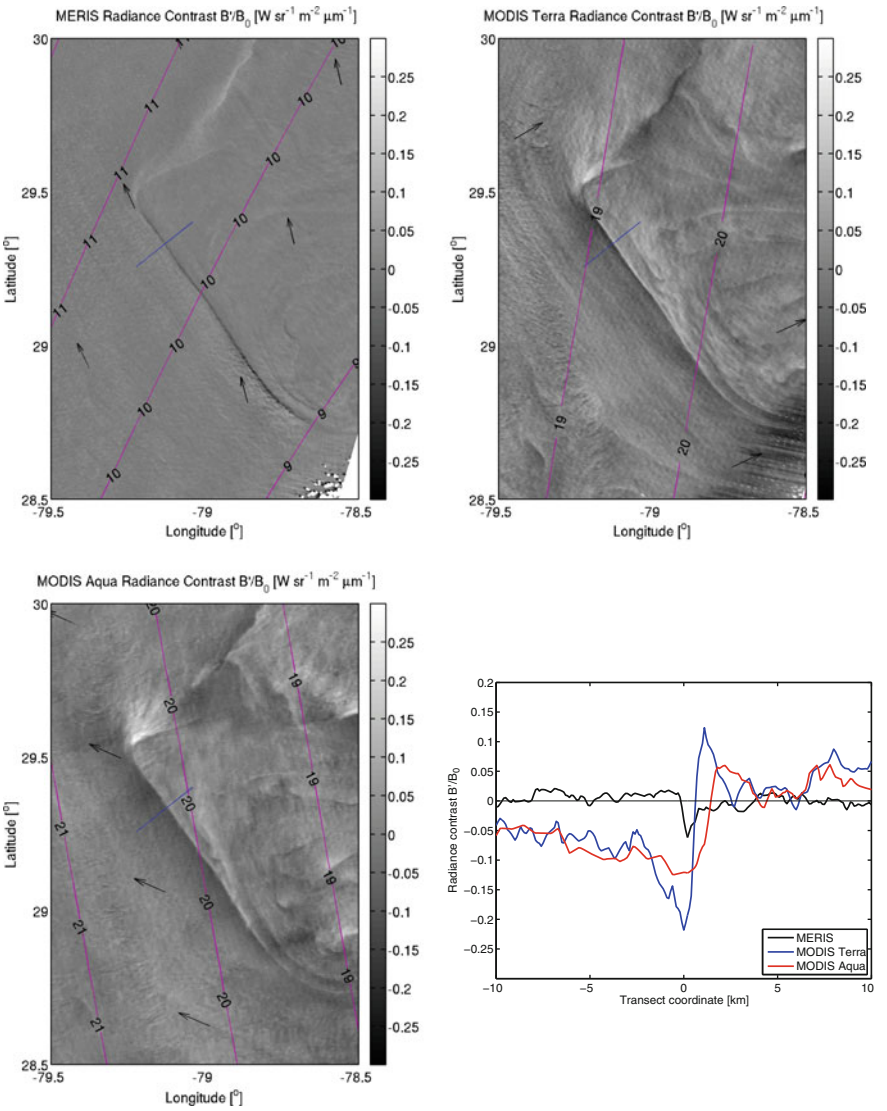


**Fig. 10** Same as Fig. 9 but at a different location in the Gulf Stream region

a configuration similar to that of case 2c. The radiance contrast from MODIS Aqua and from MODIS Terra both exhibits a dipole pattern, with a negative anomaly westside of the front and a positive anomaly eastside. Again, this is in good agreement with the mss contrasts expected for case 2c (Fig. 8).

# 8 Discussion

These observations showed an encouraging qualitative agreement between the signs of the surface roughness anomalies and the expected sign of the current gradient around oceanic fronts. More refined quantification of the current gradients from surface roughness observations can then be attempted and will be subject to further dedicated investigations.

But, in this discussion, we wish to focus on another important information contained in surface roughness variations, related to the spatial scale of the oceanic fronts. As discussed in Sect. 2.2, waves of different scales have very different relaxation length scales. As different sensors might not be sensitive to similar wave scales, it is important to more precisely identify which waves are mostly modified by the current gradients and which waves are supporting most of the mss variations. We could then infer some relaxation length scales from this analysis.

## *8.1 Waves with Dominant Action Contrast*

### 8.1.1 Response to Oscillatory Currents

Solutions of (2) in the Fourier space have been studied by Kudryavtsev et al. [19] (see their Eq. 1).

$$\frac{\hat{N}'}{N_0} = \left(\frac{1}{\tau_c} + ic_g K\right)^{-1} \frac{\partial \hat{u}_j}{\partial x_i} m_k^{ij}, \tag{20}$$

where we defined $m_k^{ij} = k_j/N_0 \partial N_0/\partial k_i$ and any quantity $z(\mathbf{x})$ is defined as a sum of Fourier spectral components as $z(\mathbf{x}) = \int \hat{z}(\mathbf{K}) \exp(i\mathbf{K}\cdot\mathbf{x})\mathbf{dK}$, with $\mathbf{K}$ the wavenumber.

- For small waves, $1/\tau_c \gg \partial u/\partial x$ and the action anomaly becomes weak, i.e., $N' \ll N_0$. The physical reason is that short waves are tightly coupled to local wind forcing ($\tau_c$ is short) and thus cannot have a large anomaly from their background equilibrium.
- For large waves, $c_g K \gg \partial u/\partial x$ and the action anomaly also becomes weak. The physical reason is that long waves propagate too fast and therefore they cannot respond to currents which oscillate with a short spatial length scale.
- It is thus waves of intermediate size which have the strongest response to oscillatory currents. Kudryavtsev et al. [19] found that the response is maximum for waves with $k \propto u_*^{-1}(gK)^{1/2}$, which indeed increases with the current wavenumber $K$.

### 8.1.2 Response to Isolated Current Front

Philips [22] studied solutions of (2) for an isolated current front similar to that of case 1a. He found that the action contrast $N'/N_0$ due to the front is an increasing function of $\tau_c \partial u/\partial x$ (see his Eq. 4).

- Contrary to oscillatory currents, the action contrast increases for the longest waves (see his Eq. 7). The reason is that the relaxation length scale increases for the longest waves, and therefore the response to a given current gradient increases for the longest waves.
- The propagation does not introduce an intermediate wave scale for an isolated front. This is a difference from the case of long waves traveling over a rapidly oscillating current, which as stated in the previous paragraph do not respond because they travel too quickly over the oscillating current variations. In other words, an isolated current front of length scale $L$ involves different Fourier components $K$ with $0 \leq K \leq 2\pi/L$, and therefore there is no intermediate wavenumber $k$ scaling with $L^{-1/2}$ as in the previous paragraph.
- The action anomaly increases for longer waves. This is valid up to the waves at the spectral peak, where $m_k$ cancels out and so does the action anomaly. There is therefore a scale limit due to the wave properties around the spectral peak.

The action contrast $N'/N_0$ at the center of the front for case 1a is shown in Fig. 11. The contrast is shown as function of wavenumber $k$ and for different directions. The action contrast is larger for long waves (around the spectral peak $k \simeq 0.3\,\mathrm{m}^{-1}$) and for intermediate waves ($k \simeq 1$).

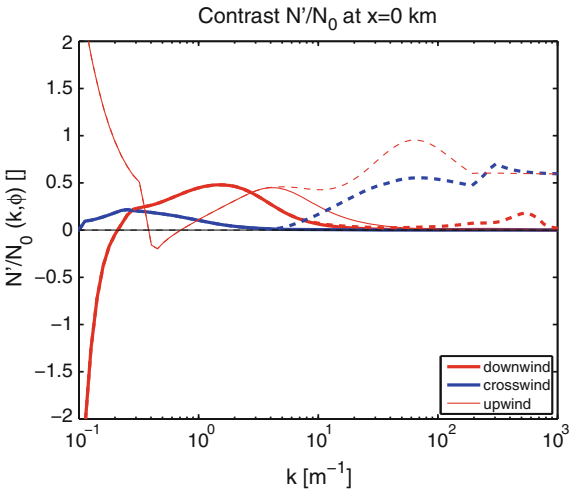


**Fig. 11** Action contrast $N'/N_0$ at the center of the front for the case 1a. The contrast is shown as function of wavenumber $k$. The different *curves* correspond to upwind propagating waves ($\phi = \phi_w + 180°$), crosswind waves ($\phi = \phi_w + 90°$), and downwind waves ($\phi = \phi_w$). Dashed curves include the effect of wave breaking

When the effect of wave breaking is included, short waves are generated and thus exhibit large action contrasts (Fig. 11, dashed curves), but those contrasts are not directly related to wave refraction by the current.

## 8.2 Waves with Dominant mss Contrast

To highlight the contribution of the different wavenumbers $k$ to the mss anomaly, we rewrite (4) as

$$\begin{aligned} \mathrm{mss}'_u(\mathbf{x}) &= \int_k \left[ \int_\phi \omega^{-1} k N' k^2 \cos^2(\phi - \phi_w) k \mathrm{d}\phi \right] \mathrm{d}k, \\ \mathrm{mss}'_c(\mathbf{x}) &= \int_k \left[ \int_\phi \omega^{-1} k N' k^2 \sin^2(\phi - \phi_w) k \mathrm{d}\phi \right] \mathrm{d}k, \end{aligned} \tag{21}$$

where the terms in brackets are shown in Fig. 12a. The larger mss anomaly occurs for waves around $k \simeq 1\,\mathrm{m}^{-1}$. Mss anomaly of shorter waves is small but sums up to an important contribution, as shown in Fig. 12b using a logarithmic distribution. Whereas the contribution to the background mss is approximately equally supported by waves longer and shorter than $k \simeq 6\,\mathrm{m}^{-1}$ (Fig. 12c), the mss anomaly is mainly supported by waves longer than $k \simeq 6\,\mathrm{m}^{-1}$, i.e., long and intermediate waves (see Fig. 12b, solid curves).

The wave refraction by the current thus involves long and intermediate waves. Again, once the wave-breaking effect is included, steep short waves are generated, and those waves support an important part of the mss anomaly (Fig. 12b, dashed

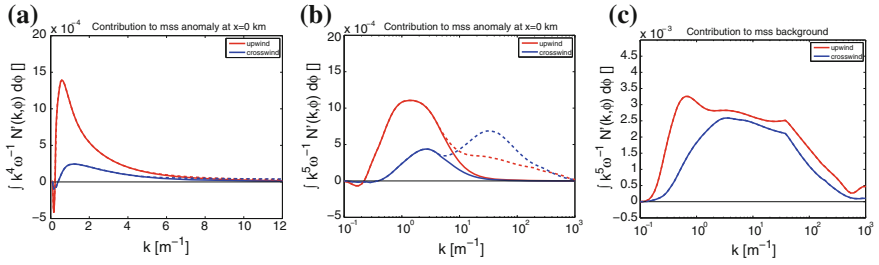


**Fig. 12** **a** Mss anomaly at the center of the front ($x = 0$ km) in case 1a. The contribution of linear wavenumber intervals d$k$ to the anomaly is shown. **b** Contribution of logarithmic intervals d ln $k$, i.e., the linear contribution is multiplied by $k$ and plotted with a log scale, so that the area under each *curve* still gives the overall mss anomaly. **c** Contributions of logarithmic intervals d ln $k$ to the background mss. The *dashed curves* include the impact of wave breaking

curves), especially in the crosswind direction. But these anomalies are not related to direct refraction of short waves by the current.

## 8.3 Relaxation and Current Gradient Length Scale

The mss anomaly due to current gradient is mainly generated by the refraction of long and intermediate waves. Those waves, in particular those traveling in the downwind direction, in turn break and create short waves in all directions.

The relevant relaxation length scale of the mss anomaly is thus the one associated with long and intermediate waves, even though the short waves in fine contribute to the mss anomaly.

This is apparent in the calculations of case 1a (Fig. 4). Short waves experience no direct modulation due to current refraction (Fig. 4b). They exhibit a modulation once the wave-breaking term is included (Fig. 4c), and in such case they follow a relaxation rate similar to that of the intermediate waves (of the order of 1 km for a 5 m $s^{-1}$ wind speed), which is much longer than their own relaxation rate (of the order of 10 m).

On the contrary, long waves ($k < 0.6$ $m^{-1}$) have relaxation length $l_c$ greater than 10 km.

We thus concluded that sensors sensitive to all waves (e.g., optical sensors at low incidence angles) will detect oceanic frontal structures with a relaxation rate of the order of a few kilometers because they include long and intermediate wave relaxation. On the other hand, sensors sensitive to short waves (e.g., SARs at high incidence angles in certain conditions, optical sensors away from the specular domain) will detect structures with a finer resolution, of the order of the relaxation rate of the breaking intermediate waves.

## 8.4 Numerical Illustration

This relaxation length scale is illustrated with a forward numerical simulation.

We use surface current calculated from numerical simulation of mesoscale and submesoscale oceanic turbulence in a beta-plane channel performed with a primitive equation ocean model. The simulation is forced by an unstable westerly zonal flow. The turbulence is intensified at the surface where the submesoscale dynamics is active and the spectrum of horizontal kinetic energy is shallow (slope in $k^{-2}$). The averaged horizontal kinetic energy is of about 10 cm $s^{-1}$ at the surface. The horizontal resolution of the model is of 2 km and the vertical resolution is of about 3 m near the surface. More details can be found in [16].

Using these surface currents, we perform a forward simulation of the mss variations using the wave model of [17], which includes propagation and wave

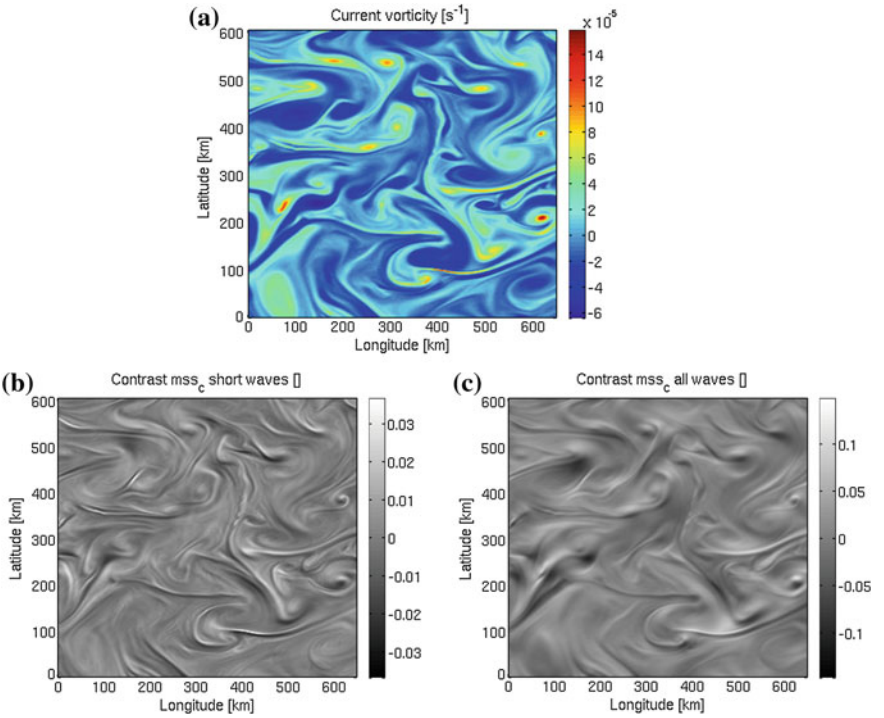


**Fig. 13** Crosswind mss contrast $\mathrm{mss}_c'/\mathrm{mss}_c$ from a forward numerical simulation. The wind is set to westerly 5 m $s^{-1}$ and the waves are supposed fully developed. **a** Surface current vorticiy. **b** Mss contrast of short waves only ($k > 30$ $m^{-1}$). **c** Mss contrast of all waves

breaking. The horizontal resolution of the wave model is also 2 km. The wind is westerly 5 $ms^{-1}$ and background waves are supposedly fully developed.

Figure 13 shows the crosswind mss contrasts $\mathrm{mss}_c'/\mathrm{mss}_c$ of short waves ($k > 30$ $m^{-1}$) only and of all waves. As discussed above, the contrast of short waves is due to the modulation of intermediate waves and thus exhibits a finer resolution than the contrast of all waves.

## 9 Conclusion

Owing to wave refraction by surface currents, ocean fronts are routinely observed on images of sea surface roughness. In this paper, we discuss a simplified and efficient framework to help interpret the observations. As proposed, the conservation of action is linearized with a relaxation time approach and the overall surface roughness contrast is reduced to wave mss anomalies.

Two types of fronts are discussed: a front with acrossfront current divergence and a front with alongfront current shear. Neglecting the wave propagation, these fronts will appear through their divergence and through their strain in the wind direction.

The impact of wave propagation is highlighted, as well as the impact of wave breaking and the associated generation of short waves. Wave propagation appears to reduce and shift the roughness anomalies in the propagation direction. Yet, it does not qualitatively change the roughness anomalies, except in the specific case of alongfront current shear with alongfront wind, where dipole anomalies appear. The impact of wave breaking links the short waves to the longer breaking waves. Short waves then become sensitive to the current gradient along the wind direction instead of the current gradient along their own propagation direction. Furthermore, now bound to longer wave modulations, short waves then exhibit a relaxation length scale two orders of magnitude longer than their own relaxation scale.

The present analysis framework can thus lead to practical analysis strategies, and can help in combining different observations to more efficiently uncover the upper ocean dynamics at finer spatial scales. As suggested, an azimuthal diversity (upwind and crosswind) would likely help to distinguish and interpret the surface roughness contrast manifestations. As well and already reported for radar signals, the use of polarization sensitivity can also quite directly lead to better discriminate and interpret contrasts associated to breaking waves. Following this analysis, future investigations will be directed toward improving our ability to more unambiguously retrieve surface deformation field from combined high-resolution optical and radar images.

**Acknowledgments** N. Rascle was supported by the ESA (European Space Agency) through the STSE MESO3D project and by the "Laboratoire d'Excellence" LabexMER (ANR-10-LABX-19), co-funded by a grant from the French government under the program "Investissements d'Avenir". We also acknowledge the financial support of the ANR (French Agence Nationale pour la Recherche) through the REDHOTS project and of the ESA through the GlobCurrent project. We thank V. Kudryavtsev, RSHU (Russian State Hydrometeorological University), St. Petersburg for the fruitful discussions and for providing the numerical code at the core of the present discussion. We also thank F. Collard, OceanDataLab, for motivating the analysis and the interpretation of numerous surface roughness observations. We thank S. Delwart at ESA-ESRIN for providing the Level-1 MERIS data and we used the website http://lance-modis.eosdis.nasa.gov to download the Level-1 MODIS data.

## References

1. Alpers W (1985) Theory of radar imaging of internal waves. Nature 314(6008):245–247
2. Alpers W, Hennings I (1984) A theory of the imaging mechanism of underwater bottom topography by real and synthetic aperture radar. J Geophys Res Oceans (1978–2012) 89 (C6):10529–10546
3. Apel JR, Byrne HM, Proni JR, Charnell RL (1975) Observations of oceanic internal and surface waves from the Earth Resources Technology Satellite. J Geophys Res 80(6):865–881

4. Beal R, Kudryavtsev V, Thompson D, Grodsky S, Tilley D, Dulov V, Graber H (1997) The influence of the marine atmospheric boundary layer on ERS 1 synthetic aperture radar imagery of the Gulf Stream. J Geophys Res Oceans (1978–2012) 102(C3):5799–5814
5. Beal RC, DeLeonibus PS, Katz I (1981) Spaceborne synthetic aperture radar for oceanography, vol 7. Johns Hopkins University Press
6. Cox C, Munk W (1954) Measurement of the roughness of the sea surface from photographs of the sun's glitter. J Opt Soc Am 44(11):838–850
7. Donelan MA, Hamilton J, Hui WH (1985) Directional spectra of wind-generated waves. Phil Trans Roy Soc London A 315:509–562
8. Dulov V, Kudryavtsev V (1990) Imagery of the inhomogeneities of currents on the ocean surface state. Soviet J Phys Oceanogr **1**(5):325–336 (1990). doi:10.1007/BF02196830. url http://dx.doi.org/10.1007/BF02196830
9. Espedal HA, Johannessen OM, Johannessen JA, Dano E, Lyzenga D, Knulst J (1998) COASTWATCH'95: ERS 1/2 SAR detection of natural film on the ocean surface. J Geophys Res Oceans (1978–2012) 103(C11):24969–24982
10. Fu LL, Holt B (1983) Some examples of detection of oceanic mesoscale eddies by the SEASAT synthetic-aperture radar. J Geophys Res Oceans (1978–2012) 88(C3):1844–1852
11. Gerling T (1986) Structure of the surface wind field from the Seasat SAR. J Geophys Res Oceans (1978–2012) 91(C2):2308–2320
12. Hughes B (1978) The effect of internal waves on surface wind waves 2. Theoretical analysis. J Geophys Res 83(C1):455–465
13. Huot J, Rast M, Delwart S, Bezy J, Levrini G, Tait H (2001) The optical imaging instruments and their applications: AATSR and MERIS. ESA Bull 106(June)
14. Johannessen JA, Shuchman RA, Digranes G, Lyzenga D, Wackerman C, Johannessen OM, Vachon P (1996) Coastal ocean fronts and eddies imaged with ERS 1 synthetic aperture radar. J Geophys Res Oceans (1978–2012) 101(C3):6651–6667
15. Keller W, Wright J (1975) Microwave scattering and the straining of wind-generated waves. Radio Sci 10(2):139–147
16. Klein P, Hua B, Lapeyre G, Capet X, Le Gentil S, Sasaki H (2008) Upper ocean turbulence from high-resolution 3D simulations. J Phys Oceanogr 38(8):1748–1763
17. Kudryavtsev V, Akimov D, Johannessen J, Chapron B (2005) On radar imaging of current features: 1. Model and comparison with observations. J Geophys Res Oceans 110(C7):C07016
18. Kudryavtsev V, Kozlov I, Chapron B, Johannessen J (2014) Quad-polarization SAR features of ocean currents. J Geophys Res Oceans 119(9):6046–6065
19. Kudryavtsev V, Myasoedov A, Chapron B, Johannessen JA, Collard F (2012) Imaging mesoscale upper ocean dynamics using synthetic aperture radar and optical data. J Geophys Res Oceans (1978–2012) 117(C4)
20. Kudryavtsev V, Myasoedov A, Chapron B, Johannessen JA, Collard F (2012) Joint sun-glitter and radar imagery of surface slicks. Remote Sens Environ 120:123–132
21. McWilliams J, Colas F, Molemaker M (2009) Cold filamentary intensification and oceanic surface convergence lines. Geophys Res Lett 36(18)
22. Phillips OM (1984) On the response of short ocean wave components at a fixed wavenumber to ocean current variations. J Phys Oceanogr 14:1425–1433
23. Plant WJ (1982) A relationship between wind stress and wave slope. J Geophys Res 87:1961–1967
24. Rascle N, Chapron B, Ponte A, Ardhuin F, Klein P (2014) Surface roughness imaging of currents shows divergence and strain in the wind direction. J Phys Oceanogr 44(8):2153–2163
25. Salomonson VV, Barnes W, Maymon PW, Montgomery HE, Ostrow H (1989) MODIS: advanced facility instrument for studies of the Earth as a system. IEEE Trans Geosci Remote Sens 27(2):145–153
26. Yurovskaya M, Dulov V, Chapron B, Kudryavtsev V (2013) Directional short wind wave spectra derived from the sea surface photography. J Geophys Res Oceans 118(9):4380–4394

# The Impact of Near-Surface Salinity Structure on SMOS Retrievals

**K.T. Walesby and B. Ward**

**Abstract** The European Space Agency recently launched their Soil Moisture and Ocean Salinity (SMOS) mission, providing measurements of sea surface salinity on a global scale for the first time. However, SMOS is only able to sense the upper 1 cm of the ocean, and there are questions as to how representative this point measurement is of the upper several metres of the water column. Here we present results from investigations into near-surface salinity structure. These observations were made using a novel, upwardly rising, microstructure profiler and the data came from the tropical North Atlantic Ocean. Analysis is presented whereby the in situ data was used to quantify the strength of near-surface salinity gradients in this region. A comparison was also conducted between the in situ observations and the co-located SMOS Level 3 data. This showed that the difference between these two datasets was normally larger than the salinity gradients seen in the upper 5 m of the ocean. Effectively, this implies that near-surface salinity gradients cannot explain the discrepancy between the two datasets. Future research is required to repeat this analysis using higher temporal resolution Level 2 data.

## 1 Introduction

One of the most robust predictions of the climate models is for an intensification in the global water cycle [15]. This change is linked to the fact that a warmer atmosphere is able to contain more moisture within it. Interestingly, however, land-based measurements have thus far been rather inconclusive in detecting this expected trend [5, 9]. It must be remembered that the bulk of the global hydrological cycle takes place on or above the ocean. For example, 86 % of the

K.T. Walesby (✉) · B. Ward
National University of Ireland, Galway, Ireland
e-mail: kieran.walesby@nuigalway.ie

B. Ward
e-mail: bward@nuigalway.ie



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_5

evaporation and 78 % of the precipitation on this planet occur here [15]. Thus, although much of the scientific interest in this intensification of the hydrological cycle understandably concerns changes over land, a balanced perspective of this issue requires consideration of the ocean too. Recent research has provided evidence that the global water cycle over the ocean may, indeed, be intensifying [6].

Naturally, making in situ observations over the ocean presents considerable technical and logistical challenges. However, a possible way forward uses changes in sea surface salinity as a proxy for precipitation changes [14]. The launch of the European Space Agency's Soil Moisture and Ocean Salinity (SMOS) mission permitted the accurate determination of sea surface salinity on a global scale [13]. In areas where there is an excess of precipitation over evaporation, there will be a corresponding freshening of the surface waters. In contrast, an excess of evaporation over precipitation will result in an increase in sea surface salinity. In other words, this approach seeks to use the ocean itself as a rain gauge [14].

This satellite-based approach has one significant uncertainty. Sea surface salinity is not solely determined by the balance of evaporation and precipitation, as was presented in the above picture. Instead, surface salinity can be redistributed horizontally by advection, and vertically by a variety of turbulent mixing processes. The effects of these must be properly understood, under a range of conditions, if the above approach is to be used.

Since vertical mixing processes in the upper oceanic boundary layer, and their associated impact on sea surface salinity (SSS) retrievals by SMOS, are the key focus of this study, a review of these processes is provided in Sect. 2. The different datasets used in the analysis presented here are described in Sect. 3, with particular emphasis on the novel instrumentation used. Finally, Sect. 4 reports the results found by this study and discusses their implications for SMOS.

## 2 A Review of Upper Ocean Mixing Processes

Since mixing in the upper ocean is central to this investigation, there follows in this section a short introduction to important aspects of some of these processes. Figure 1 shows a stylised schematic of the turbulent processes found in the upper ocean. Several of the most relevant processes to the work presented here are described in more detail in the following subsections. Thorpe provides a fuller review of mixing processes found in the upper ocean [18].

### *2.1 Surface Waves*

That the state of the ocean surface is influenced by winds in the atmosphere is a matter of common experience. The effects of the wind on the sea do not cease at the air–water interface, but instead can be seen through much of the upper oceanic

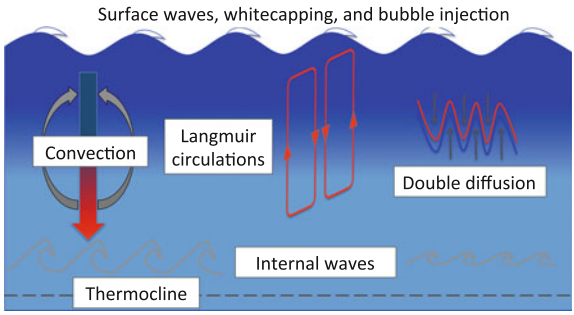


**Fig. 1** A stylised schematic indicating the major mixing processes in the upper oceanic boundary layer. The position of the thermocline—a sharp gradient in temperature which inhibits mixing with the oceanic interior—is indicated

boundary layer (OBL). Wind-generated waves represent one of the most important sources of turbulent kinetic energy in the upper OBL. Atmospheric winds blowing over an initially flat sea surface form a kind of two-fluid shear instability. The winds follow the surface of the sea, with the flow pattern resulting in differential pressures on the windward and leeward sides of a wave. This wave energy can then be dissipated by wave breaking.

Large breaking waves are associated with whitecaps and foam, and can thus be detected visually. Wave breaking also occurs on smaller scales, however, and this does not have a visual signature. Microscale wave breaking has been shown to be a fundamental physical mechanism contributing to gas transfer across the air–sea interface [21]. These small-scale breakers can also have a significant impact on the momentum and energy budgets. Video-based techniques have been used to study breaking waves. Such studies have reported whitecaps accounted for 11–17 % of the momentum flux, with small-scale breakers being responsible for the remainder [7, 11]. Determining the prevalence of small-scale breaking remains a research priority.

One of the effects of wave breaking is to inject momentum into the OBL [1]. Generally, however, the effects of wave breaking are only felt in the upper few metres of the ocean. Despite this comprising only a relatively small portion of the upper OBL, they play a crucial role in determining the evolution of this layer, and by extension, the wider ocean, weather and climate [10]. An important consequence of wave breaking is that turbulence production in this surface region is dominated by the turbulent kinetic energy flux. In contrast, the equivalent region of the atmospheric boundary layer (ABL) is dominated by shear production [17].

## *2.2 Internal Waves*

Waves can also exist below the surface and are known as “internal waves”. Fluctuations in wind velocity can result in oscillations at near-inertial frequencies.

Close to the coast, tidal forcing can also be important. These oscillations are found throughout most of the upper ocean, but are thought to be especially important sources of mixing at the thermocline since momentum rapidly diffuses through the mixed layer and is focused here. This increased concentration of shear increases the probability that some of these internal waves will break down to form turbulence. An example of this was observed using a microstructure profiler in the Labrador Sea [19]. Wain and her colleagues measured elevated levels of turbulence at the crest of one of these waves, associated with a breaking event. This indicates that such waves may be an important source of turbulence in this region.

The most common, and well-known, mechanism by which internal waves dissipate is by the Kelvin–Helmholtz shear instability. This occurs as a consequence of the effect that the superposition of random internal waves has on the local Richardson number (Ri). Ri gives an estimate of the likelihood of turbulence occurring in a particular flow. Physically, it represents the ratio of the buoyancy to shear terms in the turbulence kinetic energy budget. Shear associated with the passing of an internal wave is thus in competition with the stabilising effect of stratification in the OBL. When this balance leads to the local Ri dropping below a critical value (generally 0.25), the onset of turbulence can occur [12]. The Kelvin–Helmholtz shear instability comes about once Ri becomes subcritical. Then the inertial shear forms distinct vortices or "billows". When viewed in cross-section, these appear like surface water waves breaking on the seashore. Eventually, these overturn and are dissipated as turbulent mixing. Some of these turbulent motions themselves go on to generate further high-frequency waves.

## *2.3 Convective Mixing*

Convection is another important turbulence production process within the oceanic boundary layer. It acts to redistribute fluid parcels vertically. Fluid parcels at the surface lose heat to the atmosphere and eventually sink, moving cool water downwards. The depth to which a fluid parcel sinks is determined by the local stratification. The motion of these convective plumes acts to generate small-scale turbulence. Typically, this surface cooling will occur at night. Observations of this have been reported previously [3, 4]. It is not, however, solely a nocturnal phenomenon. Surface density can increase sufficiently to trigger convection due to the effects of synoptic weather features (e.g. cold air outbreaks), and there is also usually a strong seasonal cycle [8]. In the ABL, convection is governed by temperature differences with height. Within the ocean, water density is controlled by both temperature and salinity. This means that convection here can also be triggered by an excess of evaporation over precipitation.

During the day, heating caused by incoming solar radiation reduces the density of the water close to the surface. This acts to stabilise the water column and suppress convection. Similarly, when precipitation exceeds evaporation the

near-surface water becomes freshened, reducing its density, and again convection is suppressed. The most significant effect of convective circulations is to mix properties such as heat and bubbles which have travelled across the air–sea interface from close to the surface to throughout the mixed layer. There is a corresponding upward flow within a convective cell which returns water from depth to the surface.

# 3 Methodology

## *3.1 ASIP*

Investigation of near-surface salinity gradients requires an instrument capable of profiling turbulence within the upper oceanic boundary layer with very high resolution. The Air–Sea Interaction Profiler (ASIP) is an autonomous, upwardly rising, microstructure profiler which is extremely well suited to this task. It was developed by the National University of Ireland, Galway and is the main experimental tool which will be used here [20].

ASIP allows the measurement of small-scale fluctuations in temperature, conductivity (used, together with temperature, to calculate salinity), velocity shear (from which dissipation of turbulent kinetic energy—$\varepsilon$—can be estimated) and various biological variables. $\varepsilon$ provides a measure of how turbulence levels vary with depth through the water column. When ASIP reaches the surface at the end of each profile, its onboard global positioning system (GPS) transmits the coordinates of its position by satellite.

ASIP carries a mixture of fast and slow frequency instruments for temperature and conductivity. There are two fast-response temperature thermistors and one for the fast-response conductivity sensor (all by Sea-Bird Electronics, and modified by Rockland Scientific Instruments). In order to measure temperature and conductivity at almost the same point, and thereby determine salinity with as much precision as possible, one of these thermistors was located approximately 1 mm from the conductivity sensor (Fig. 2).

The disadvantage of the fast-response microstructure sensors is that they require continuous calibration. The slow-response instruments on ASIP consist of a CTD (Conductivity–Temperature–Depth) and NBOSI (Neil Brown Ocean Sensors, Inc.). Although the slower response of these sensors—particularly the CTD—means they are not able to capture turbulent variability with the same resolution as the microstructure instruments, they do have more temporally stable calibrations. This means that the microstructure instruments could be re-calibrated using the CTD after each ASIP profile.

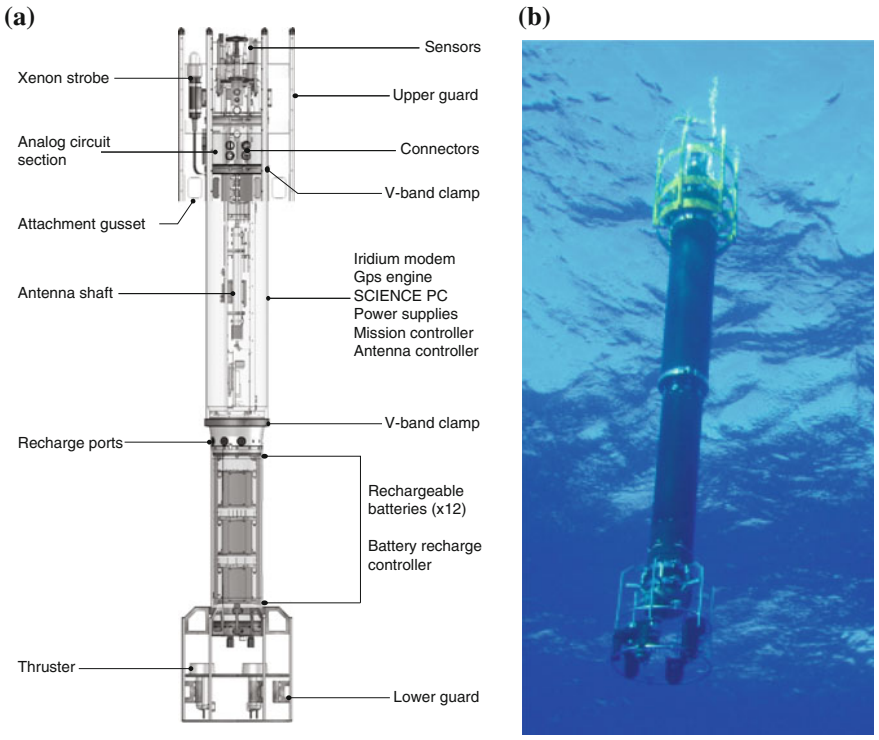


**Fig. 2** **a** A schematic of ASIP outlining several key features. **b** ASIP floating at the surface, prior to beginning a profile in the tropical North Atlantic

## *3.2 SPURS Cruises*

The field-work element of this project consisted of examining data from a cruise in the tropical North Atlantic Ocean. This cruise formed part of the larger project, Salinity Processes in the Upper ocean Regional Study (SPURS), and was termed Strasse (August–September 2012). The aim of this expedition was to bring together satellite measurements of sea surface salinity with high-quality oceanographic measurements, thereby allowing either side of the air–sea interface to be intensively probed. It also offered the opportunity to validate the satellite measurements against in situ data.

SMOS uses L-band frequency radiometry to estimate salinity in a surface layer of the order of 1 cm thickness. Since most in situ measurements occur over a depth range from 1 cm to a few decimetres, the challenge here was to understand the gradients in salinity which exist in this region. ASIP was deployed several times during Strasse, yielding hundreds of profiles in total.

### 3.3 SMOS Data

For this project, SMOS Level 3 products were used. These were produced by the SMOS Barcelona Expert Centre and are available online [16]. These data were binned into 3-day averages, located on a 0.25° × 0.25° resolution, global grid.

## 4 Results and Discussion

### 4.1 ASIP Dataset from the Strasse Cruise

In this section, some general analyses from ASIP data collected during the Strasse cruise are presented. Conclusions of particular relevance to SMOS retrievals of SSS are highlighted. Figure 3 displays contour plots of temperature, salinity and $\log(\varepsilon)$ over the course of all seven ASIP deployments during Strasse. These plots provide an overview of the conditions which prevailed during Strasse, and the high-quality measurements which ASIP was able to make, right up to the air–sea interface.

In the top panel of Fig. 3, the diurnal variation in temperature can be seen, with warming occurring during the daytime, and surface cooling at night which results in buoyancy loss and convective mixing. The salinity data shows the very salty waters which were found in the Strasse region (Fig. 3). The salinity variations seen in this time series are a consequence of a combination of diurnal effects, and ASIP crossing boundaries between different water masses. This issue will be revisited later in this section. The bottom panel in Fig. 3 shows a contour plot of $\log(\varepsilon)$. Again, this variable displays some diurnal dependence. During daytime, most of the turbulence is often confined to a relatively shallow layer close to the surface. During the night, the turbulence is able to penetrate to a much greater depth (>30 m), although this picture also varies somewhat with conditions.

In an attempt to remove the water mass variation effect and isolate any possible diurnal cycle in salinity, all of the data from the Strasse cruise shown in Fig. 3 were binned by time in order to generate a composite diurnal cycle. The results of this analysis are displayed in Fig. 4.

The diurnal cycle, which was seen for temperature in Fig. 3, is confirmed by the composite data. During the daytime, incoming solar radiation warms the upper several metres of the water column and stratifies this water. Under nocturnal conditions, the water column becomes well mixed.

In contrast, the composite salinity data still does not show a strong diurnal cycle. In addition, several abrupt changes in salinity are apparent over the course of this composite day. The likely explanation for this is that any diurnal cycle which may be present is small, and is masked by the variations in salinity due to changes in water mass.

Also plotted in Fig. 4 is a black line which represents at which depth the local salinity anomaly (with respect to the surface) first exceeded 0.05 PSU. The depth at

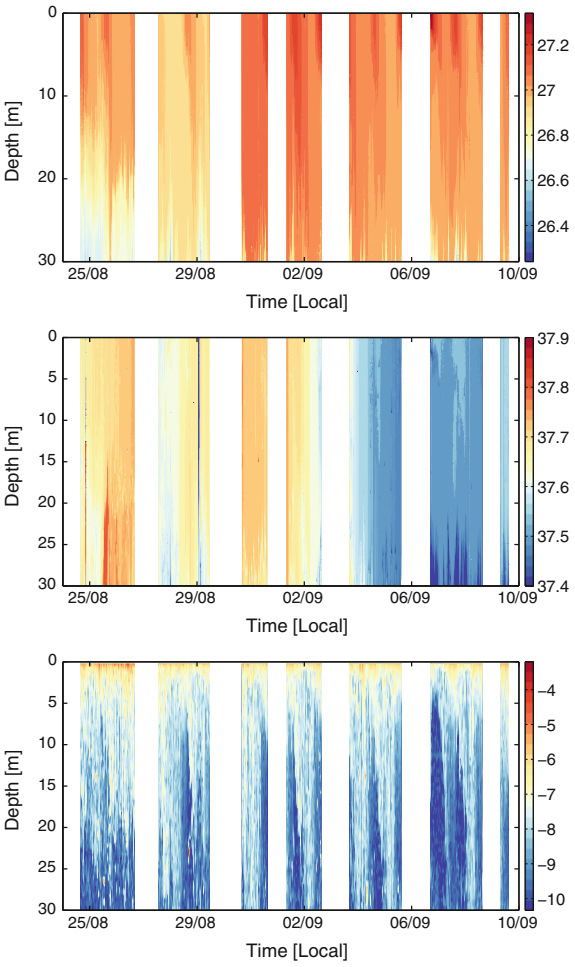


**Fig. 3** Contour plots of (from *top*) temperature, salinity and $\varepsilon$ from all ASIP deployments during Strasse

which this threshold was exceeded was calculated individually for each profile, before a mean was calculated for each hour of the composite day. This procedure largely removes the effect of water mass variations. This provides a simple measure of the strength of near-surface salinity gradients—with clear consequences for the representativeness of SMOS retrievals of SSS. When these gradients are strong, this threshold is exceeded at a very shallow depth. Under more homogeneous conditions, this threshold is only crossed at greater depths.

Interestingly, this metric does appear to show a diurnal cycle which is readily understood. Under daytime conditions, solar heating leads to both near-surface stratification and enhanced evaporation. This generates increased salinity close to the surface, and the $\Delta S_{0.05\mathrm{PSU}}$ threshold is within 2 m of the surface. In contrast, the well-mixed layer which is found during the night—once the surface waters have lost sufficient buoyancy by cooling—means that this threshold is now exceeded at

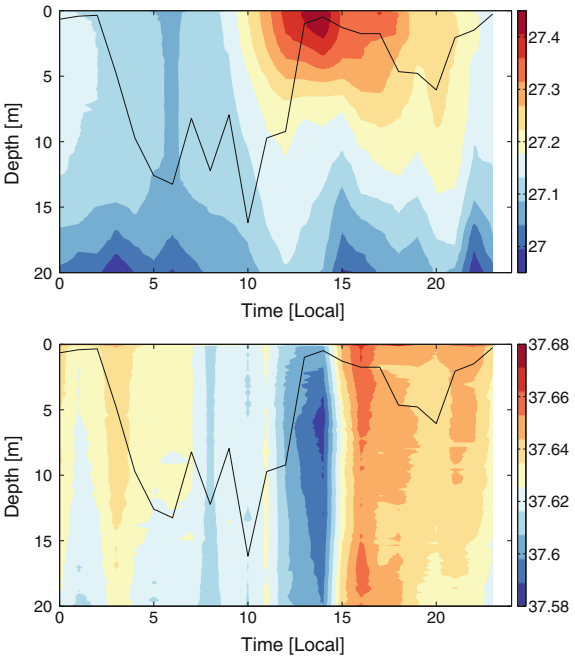


**Fig. 4** Composite diurnal cycles in temperature (*upper panel*) and salinity (*lower panel*) as observed by ASIP during the Strasse cruise. The *black line* indicates the depth at which the salinity anomaly (with respect to the surface value) first exceeded 0.05 PSU

depths in excess of 5 m. This analysis indicates that SMOS measurements of SSS are least representative of the mean salinity in the upper oceanic boundary layer during the daytime, and most representative during the latter portion of the night.

This conceptual view is neatly illustrated by the ASIP data collected during Deployment 7. Although salinity variations for most of the deployments shown in Fig. 3 were driven by water mass changes, a small diurnal cycle appeared to be present in Deployment 7. Contour plots for temperature, salinity, density and $\varepsilon$ are all shown in Fig. 5. The diurnal cycle in heating over the upper few metres of the ocean is apparent here. There is also a small cycle in salinity too ($\sim$0.05 PSU), with salinity close to the surface increasing during the daytime. This variation slightly lags behind the temperature cycle. This is expected, since both the surface stratification and enhanced evaporation are instigated by the daytime solar heating, and it is these processes which result in the near-surface salinity increase.

During the daytime, temperature and salinity changes have competing effects on the near-surface density. Figure 5 also shows that the density-reducing effect of solar heating dominates, and the near-surface water becomes stably stratified during the day, isolating this layer from the underlying water and allowing evaporation to increase the salinity at the near-surface. After sunset, surface cooling leads to the surface density increasing until the onset of convective mixing. The signature of this mixing can be seen in the dissipation data, as the turbulence penetrates to depths of 30 m during the latter portion of the night, prior to restratification.

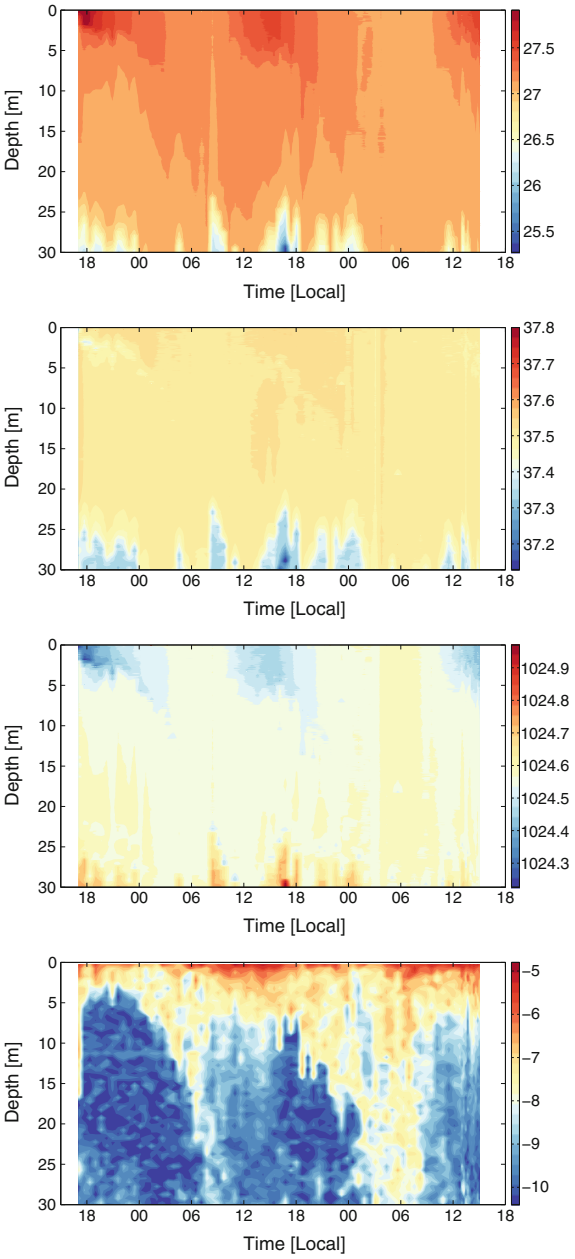


**Fig. 5** Contour plots of (from *top*) temperature, salinity, density and $\log(\varepsilon)$ for a single ASIP deployment during Strasse. This deployment shows evidence of a diurnal cycle in salinity, a consequence of the stratified near-surface layer which is formed by daytime solar heating

## 4.2 Comparison Between ASIP and SMOS Data

In this section, the ASIP data from the Strasse cruise was compared against the appropriate SMOS Level 3 data. As was described in Sect. 3.3, this satellite data was provided in 3-day averaged bins at 0.25° resolution. This clearly created a mismatch in temporal resolution with the ASIP data—which profiled every 5–10 min—that should be borne in mind in the analysis which follows.

For an initial comparison between ASIP and SMOS, SSS maps of the area surrounding ASIP were produced for each ASIP deployment (Fig. 6). The ASIP measurement of near-surface salinity was based on the mean of the upper 0.5 m of each profile. Salinity, as measured by the thermosalinograph (TSG) on board the ship, is also plotted. Figure 6 provides an estimate of the variability in surface salinity in the region ASIP was deployed.

It can be seen that this particular ASIP deployment took place in particularly salty and uniform water, although to the east and west the surface salinity became more heterogeneous. A direct comparison of the SMOS data immediately around ASIP reveals some considerable differences of up to 0.3 PSU. The TSG data tends to support the ASIP measurements, and provides a measure of the spatial variations in near-surface salinity which were not picked up by SMOS. This can partly, however, be explained by the 3-day averaging the SMOS data undergoes.

In order to investigate this further, and provide a more quantitative analysis of the SMOS–ASIP difference, a method had to be developed to compare SMOS Level 3 data with ASIP measurements. The approach used here involved averaging all the ASIP data for each deployment (generally approximately 36 h in duration), yielding a single salinity measurement for each (with an associated standard deviation). The ASIP profiles were sampled to provide measurements of salinity at depths of 1 cm and 1 m.

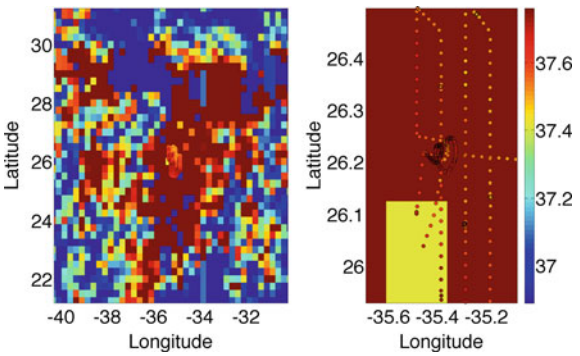


**Fig. 6** Comparison between salinity as measured by SMOS, ASIP (average of upper 0.5 m of each profile) and the shipboard TSG (nominal 5 m depth) for a single ASIP deployment. The *left-hand panel* shows the large-scale satellite picture. The *right-hand panel* shows the area immediately surrounding ASIP, with ASIP/TSG data plotted in *filled squares/circles*

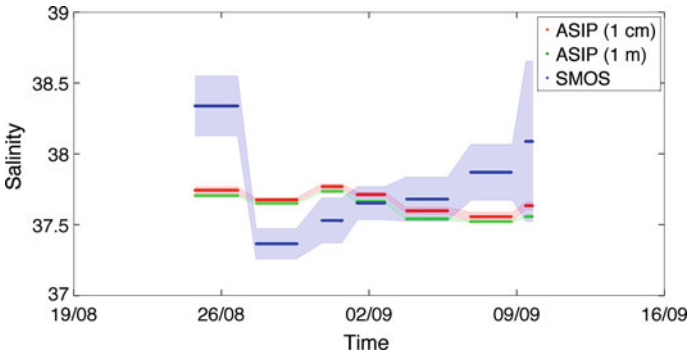


**Fig. 7** Time series of salinity for all deployments during the Strasse cruise. SMOS data is shown in *blue*, and ASIP data from depths of 1 cm/1 m are plotted in *red/green*. The *circular markers* represent the mean, and the *shaded area* one standard deviation either side of this

The SMOS data is provided in grid boxes 0.25 × 0.25° wide, and clearly ASIP did not always remain within any single one of these boxes for the entire deployment. For this reason, a small acceptance window was defined for each ASIP deployment, which extended 0.25° either side of the maximum/minimum ASIP latitude/longitude for that deployment. These SMOS data were then averaged to provide a mean and standard deviation which could be compared with ASIP.

The results of this comparison are presented in Fig. 7. Interestingly, both sets of ASIP measurements lie beyond one standard deviation of the SMOS data for most ASIP deployments. Indeed, the ASIP–SMOS discrepancy was frequently an order of magnitude greater than the difference between the two ASIP datasets. However, this discrepancy was also close to the accuracy of 0.2 PSU which SMOS aspires to (when averaged over Global Ocean Data Assimilation Experiment (GODAE) scales, see [2]).

In one sense, this is an encouraging result: it implies that, at least in this evaporative region of the ocean, near-surface salinity gradients are relatively small and these cannot explain the ASIP–SMOS difference. This is further supported by Fig. 8, which shows the difference between salinity anomaly (with respect to the surface) measured by ASIP at depths of 1 and 5 m, for a single, representative ASIP deployment. It is, however, very important that the reasons for the large difference between the ASIP and SMOS time series be investigated further. This would certainly require analysis of SMOS Level 2 data, with careful co-location of satellite and ASIP time series.

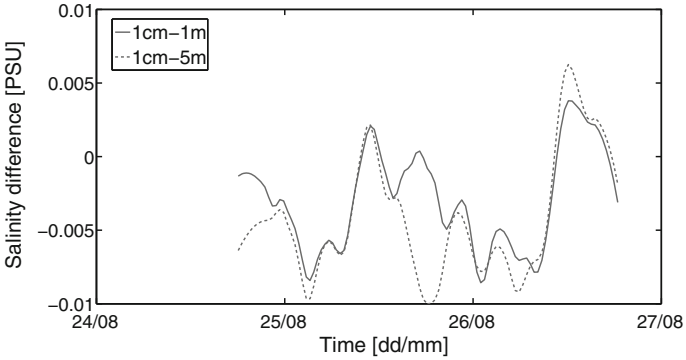


**Fig. 8** Salinity anomaly (with respect to the surface), as measured by ASIP during a single, representative deployment during Strasse. The *solid/dashed lines* present the anomaly at depths of 1 and 5 m. Comparison with the results in Fig. 7 indicates that near-surface salinity gradients are unable to explain the differences seen between ASIP and the SMOS Level 3 salinity data

## 5 Summary and Future Work

This report has summarised investigations into how mixing processes in the upper oceanic boundary layer impact on near-surface salinity gradients. This issue was investigated using a novel, upwardly rising microstructure profiler, known as the Air–Sea Interaction Profiler (ASIP), and SMOS Level 3 data. An ASIP dataset from the tropical North Atlantic Ocean was analysed and presented here. This cruise formed part of the larger initiative, Salinity Processes in the Upper ocean Regional Study (SPURS), which took place in the tropical North Atlantic Ocean.

Contour plots of temperature, salinity and dissipation rate of turbulent kinetic energy ($\varepsilon$) were presented for the entirety of this cruise. Unusually, these data were acquired right up to the surface—a consequence of the novel instrumentation and platform used in this study. A clear diurnal cycle in temperature and dissipation rate was observed. The variations seen in near-surface salinity were less clear, and were apparently caused by a combination of diurnal forcing and water mass changes (as ASIP drifted with the currents and waves). All of the data from the seven ASIP deployments during Strasse were combined to yield a composite diurnal cycle. This analysis showed that near-surface gradients in salinity were, on average, larger during daytime than at night. This was explained by the enhanced evaporation rates which were seen during the daytime, and data from one deployment demonstrated this particularly clearly. Comparisons between ASIP measurements of salinity and SMOS Level 3 data were also carried out. These showed that ASIP–SMOS differences were significantly larger than the salinity gradients observed over the upper 5 m of the ocean. Effectively, this implies that near-surface salinity gradients cannot explain the discrepancy between these data.

**Acknowledgements** KTW acknowledges a fellowship jointly funded by the European Space Agency and the Irish Research Council.

## References

1. Banner ML, Peirson WL (1998) Tangential stress beneath wind-driven air-water interfaces. J Fluid Mech 364:115–145
2. Boutin J, Martin N, Reverdin G, Yin X, Gaillard F (2013) Sea surface freshening inferred from SMOS and ARGO salinity: impact of rain. Ocean Sci 9:183–192
3. Brainerd KE, Gregg MC (1993) Diurnal restratification and turbulence in the oceanic surface mixed layer: 1. Observations. J Geophys Res 98:22645–22656
4. Brubaker JM (1987) Similarity structure in the convective boundary layer of a lake. Nature 330:742–745
5. Dai A, Qian T, Trenberth KE, Milliman JD (2009) Changes in continental freshwater discharge from 1948 to 2004. J Clim 22:2773–2792
6. Durack PJ, Wijffels SE (2010) Fifty-year trends in global ocean salinities and their relationship to broad-scale warming. J Clim 23:4342–4362
7. Gemmrich JR, Banner ML, Garrett C (2008) Spectrally resolved energy dissipation rate and momentum flux of breaking waves. J Phys Ocean 38:1296–1312
8. Houpert L, Testor P, de Madron XD, Estournel C, D'Ortenzio F (2013) Seasonal cycle of oceanic mixed layer and upper-ocean heat fluxes in the Mediterranean Sea from in-situ observations. In: EGU General Assembly Conference Abstracts 15, pp 11209
9. Huntington TG (2006) Evidence for intensification of the global water cycle: review and synthesis. J Hydrol 319:83–95
10. Kagan BA (2006) Ocean atmosphere interaction and climate modeling (vol 11). Cambridge University Press, Cambridge
11. Melville WK, Matusov P (2002) Distribution of breaking waves at the ocean surface. Nature 417:58–63
12. Miles JW (1961) On the stability of heterogeneous shear flows. J Fluid Mech 10:496–508
13. Reul N, Fournier S, Boutin J, Hernandez O, Maes C, Chapron B, Alory G, Quilfen Y, Tenerelli J, Morisset S (2013) Sea surface salinity observations from space with the SMOS satellite: a new means to monitor the marine branch of the water cycle. Surv Geophys 1–42
14. Schmitt RW, Montgomery ET (2000) Salinity: a missing piece in the climate puzzle. Backscatter
15. Schmitt RW (2008) Salinity and the global water cycle. Oceanography 21:12
16. SMOS Barcelona Expert Centre (2014) Ocean reprocessed dataset. Available via SMOS-BEC. http://cp34-bec.cmima.csic.es/. Cited March 2014
17. Stull RB (1987) An introduction to boundary layer meteorology. Kluwer, Boston
18. Thorpe SA (2005) The turbulent ocean. Cambridge University Press, Cambridge
19. Wain D, Ward B, Lilly J, Callaghan A (2013) Breaking internal waves on the pycnocline of the Labrador Sea. In: EGU General Assembly Conference Abstracts 15, pp 10039
20. Ward B, Fristedt T, Callaghan AH, Sutherland G, Sanchez X, Vialard J (2014) The air-sea interaction profiler (ASIP): an autonomous upwardly-rising profiler for microstructure measurements in the upper ocean. J Atmos Oceanic Technol (in press)
21. Zappa CJ, Asher WE, Jessup AT (2001) Microscale wave breaking and air-water gas transfer. J Geophys Res: Oceans 106:9385–9391

# SMASPARES–SMOS Data Assimilation for Parameter Estimation in Radiative Transfer Models

**Carsten Montzka, Cho Miltin Mboh and Kathrina Rötzer**

**Abstract** The validation of the SMOS Level 2 soil moisture product in the Rur and Erft catchments, Germany, showed that two main directions for enhancement should be followed: (i) improving radio frequency interference (RFI) mitigation strategies, and (ii) improving the parameterization of the radiative transfer model (RTM). Therefore, in this chapter two methods are developed to investigate the characteristics of RTM parameters, with a strong focus on soil surface roughness and vegetation opacity. One approach uses a dual state-parameter estimation technique in a data assimilation environment to select adequate parameters. It is a one-dimensional synthetic experiment neglecting spatial pattern of soil moisture as well as parameters. The spatial scale comes into play by investigating the feasibility of parameter estimation from synthetic disaggregated SMOS brightness temperature time series for the Rur and Erft catchments. A new partial grid search approach to parameter estimation (PAGSAPE) is developed in order to reduce computational cost compared to ensemble methods, which is important for future global applications.

## 1 Introduction

The objectives of European Space Agency's (ESA's) Soil Moisture and Ocean Salinity (SMOS) mission over land are to provide soil moisture observations for weather forecasting, climate monitoring, and investigating the global freshwater cycle [1, 2]. SMOS has been providing two-dimensional brightness temperature images of the Earth surface at a frequency of 1.4 GHz (L-band) since it was successfully launched in November 2009 [2]. The SMOS level-2 processor contains a radiative transfer model (RTM) that generates soil moisture from SMOS brightness temperatures (**Tb**). The necessary parameters are either obtained through

C. Montzka (✉) · C.M. Mboh · K. Rötzer
Institute of Bio- and Geosciences: Agrosphere (IBG-3),
Forschungszentrum Jülich, 52425 Jülich, Germany
e-mail: c.montzka@fzjuelich.de



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_6

auxiliary data products, e.g., forecast and analyses fields from The European Centre for Medium-Range Weather Forecasts (ECMWF), or from the peer-reviewed literature [3]. Prior to the launch, the optimization of L-MEB (L-band Microwave Emission of the Biosphere) model for different surfaces has been addressed by numerous studies, most of which were based on the analysis of ground-based and airborne L-band data, e.g. [4–12]. During the lifetime of the mission, the processor has been continuously improved based on extensive validation work.

However, the accuracy of the retrieval can be limited through uncertainties in the (1) observations, (2) microwave emission model, (3) prescribed state variables, and (4) parameter fields. These errors can be systematic and static, e.g., a wrong soil classification at a given location, systematic and temporarily varying, e.g., a seasonal bias in the soil temperature data, or random, e.g., instrument noise. These uncertainties in the observations, the forward model, and the auxiliary data will manifest themselves in systematic and random errors in the retrieved state variables, i.e., soil moisture and vegetation water content. Random uncertainties are partly considered by the SMOS Level 2 processor in that the accuracies of the observations and the retrieved state variables can be specified. However, systematic uncertainties can hardly be addressed in current retrievals.

In this chapter we first evaluate the accuracy of the SMOS Level 2 soil moisture product in the Rur and Erft catchments, Germany, for the years 2010–2012. We then present two methods to estimate RTM parameters for enhanced soil moisture retrieval from SMOS. The first is based on a particle filter data assimilation procedure, presented for a single pixel, whereas the second is a partial grid search inversion approach with data assimilation capabilities presented with a two-dimensional spatial application for the Rur and Erft catchments. The development of this second method was necessary, because ensemble methods, especially the particle filter, are known to be extremely computationally expensive.

## 2 The Rur and Erft Catchments

The study area consists of catchments of the rivers Rur and Erft and is located at the Belgian–Dutch–German border region near the city of Aachen (Fig. 1). The Rur is a tributary of the river Meuse and the Erft of the river Rhine. With a total area of 4125 $km^2$, the region has a wide range of land use, soil types, and meteorological characteristics [13].

In the northern part the soils mainly evolved from loess deposited by the rivers Rhine and Meuse. The annual precipitation is about 650–850 mm, the land cover is mainly fertile agricultural lands. The southern part of the region covers the bedrock of the Eifel Mountains and is characterized by an annual precipitation of about 850–1300 mm with dominant forest and grassland vegetation [13]. The study area is particularly interesting for this kind of investigation due to the presence of several Terrestrial Environmental Observatories (TERENO) sites (e.g., Selhausen,

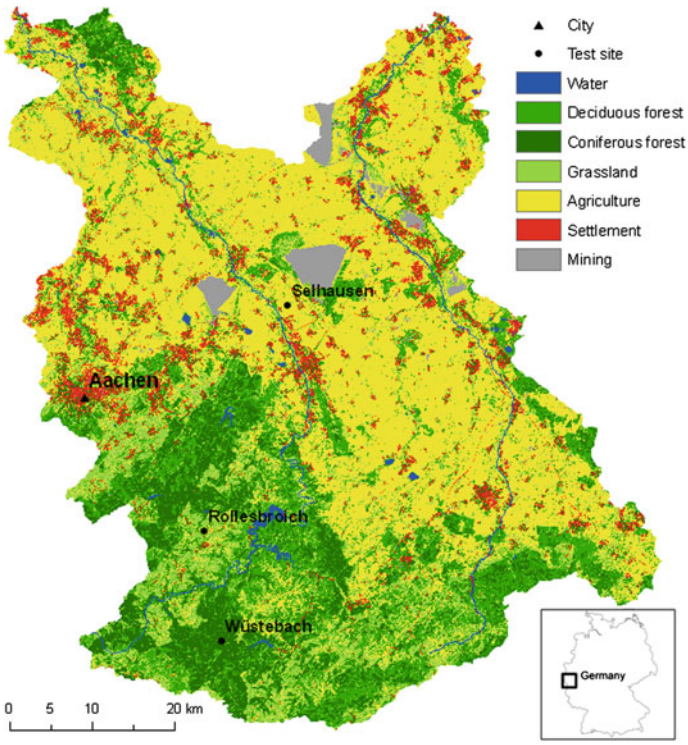


**Fig. 1** The Rur and Erft catchments

Wüstebach, Rollesbroich) which are highly instrumented to provide a wide range of in situ data [14]. In situ soil moisture of these test sites was used for the validation of the soil moisture reference [15, 16].

## 3 SMOS L2 Accuracy in the Rur and Erft Catchments

The SMOS Level 2 soil moisture product, derived with the SMOS Level 2 Processor for Soil Moisture version 5.51 [17], is validated for the Rur and Erft catchments in the west of Germany for the years 2010–2012. As reference soil moisture, output from the hydrological model WaSiM-ETH [18] is used. The grid-based hydrological model calculates soil moisture in hourly time steps and a horizontal spatial resolution of 200 m, which is then averaged over the extents of the pixels of the SMOS product. The soil is discretized in five layers from which the topmost layer of 0–5 cm is used, in order to take into account the penetration depth of the L-band microwave sensor. We gave preference to a modeled reference rather than to in situ data, as then it is possible to investigate a larger area with pixels for

which no in situ data is available. Additionally, scale differences of remotely sensed and in situ soil moisture could bias the results.

Comparison of time series of the SMOS soil moisture product to the modeled reference showed that SMOS is able to catch the trend of the model, but, especially for the pixels in the southern part of the study area a substantial bias gets visible. Correlation analysis exhibited an overall correlation coefficient of 0.28 for the three years. The development of the correlation coefficient over the seasons (Fig. 2a) shows an increase from 2010 to 2011. This is probably a result of ESA's effort to detect sources of radiofrequency interferences (RFI), which led to the switch off of many sources in the beginning of 2011 [19]. For 2012 the correlation decreases again, as a new RFI source was detected in Poland in summer 2012, while the low correlation in spring is probably caused by inaccuracies of the soil moisture reference. For all years the correlation decreases from spring to autumn. The decrease from spring to summer can be explained by the increase of vegetation biomass, which attenuates the microwave signal from the soil. The further decrease of correlation coefficients in autumn could indicate problems with the parameterization of litter or with the flagging of snow and frost, because in theory the decreasing influence of the vegetation canopy should lead to increasing correlation coefficients. Local frost or snow occurrences, which are not present in the coarse-scaled auxiliary data from ECMWF, could be an explanation for that.

The spatial distribution of correlation coefficients in the study area (Fig. 2b) shows values of up to 0.62, with the lowest values of less than 0.30 in the northwestern parts of the study area. Comparison with the distribution of probability of RFI exhibits high similarities: correlation coefficient of the correlation coefficient reference/SMOS and RFI probability of the pixels is −0.82. This demonstrates further a strong influence of RFI on correlation in the study area despite RFI filtering in the data. The SMOS soil moisture product shows a dry bias compared to modeled soil moisture. This bias is more or less constant over the seasons in the three years of the study period (Fig. 2a). A clear distinction is visible between the northern and southern part of the study area (Fig. 2c). The southern part of the study area shows bias up to 0.20 $m^3\ m^{-3}$ and higher, in the northern region they are constantly lower than that with values up to 0.10 $m^3\ m^{-3}$. This distribution was already reported in Montzka et al. [20, 23] for the time period of May and June 2010, which was in the commissioning phase of SMOS. This study makes apparent that the dry bias is a temporally stable phenomenon.

The situation improved with the implementation of Mironov Dielectric Mixing Model [21] in the Level 2 soil moisture processor version 5.51 with an absolute decrease of the bias of 0.04 $m^3\ m^{-3}$, respectively, for all seasons of the years 2010 and 2011, as well as for the whole years, in comparison to soil moisture retrieved in the processor version 5.00 (not shown), which used the Dobson model [22]. However, the remaining bias is still considerable. The temporal consistency and spatial variability of the bias indicate that the reasons are topographic differences or parameterization of different kinds of land use or other spatially distributed input parameters to the retrieval model rather than accuracy levels of the vegetation

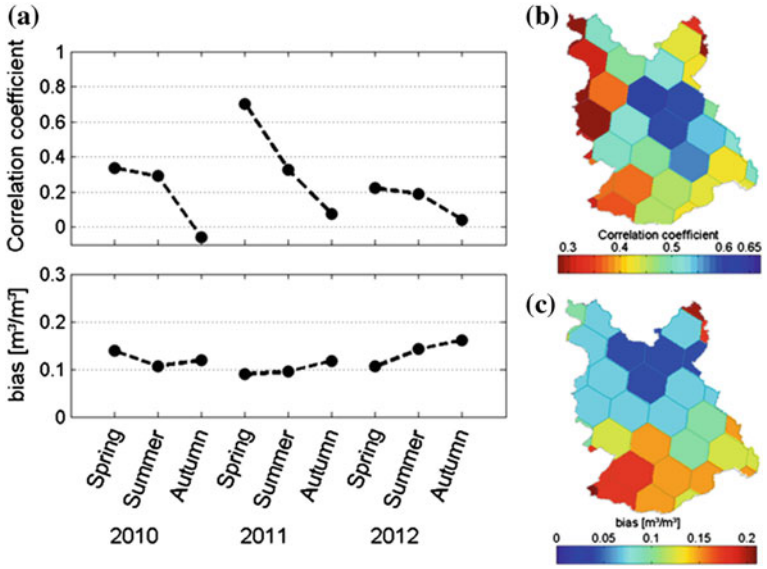


**Fig. 2** Seasonal correlation coefficients and biases of modeled soil moisture and SMOS soil moisture product for all pixels completely located in the study area (**a**), and spatial distribution of correlation coefficients (**b**) and bias (**c**) in the study area

opacity retrieval over the year or the variations of RFI in the region, which also show temporal variations due to the RFI probability given in the SMOS product.

In general the validation of the SMOS soil moisture products in the Rur and Erft catchments through a modeled reference gives promising results, as it is able to reproduce the trend of the modeled reference quite well. A challenge will be to deal with the problems caused by the dry bias, which is temporally stable, but shows considerable spatial differences.

## 4 Radiative Transfer Parameter Estimation in a 1D System

In order to reduce the bias and improve the overall SMOS soil moisture retrieval accuracy, a synthetic one-dimensional data assimilation experiment is performed. An in situ soil moisture and soil temperature reference is used to drive a RTM. By assimilating synthetic SMOS **Tb**, RTM parameters can be adjusted in an operational manner. The aim of this experiment is to retrieve temporally changing vegetation opacity $\tau$ and constant soil surface roughness *hs* in the presence of a systematic difference in the simulated **Tb**. More details about this experiment can be found in Montzka et al. [23].

### 4.1 Experimental Design

The hydrological model HYDRUS-1D was forced with meteorological data from the station in the city of Aachen (50.78°N, 6.09°E, 202 m ASL) for the year 2010 to generate soil moisture and soil temperature time series. The soil was parameterized according to the three horizons luvic cambisol [24, 25] of the TERENO test site Selhausen. The precipitation and air temperature forcing data were perturbed to generate ensembles with 1000 members mimicking the typical forcing uncertainty. This results in uncertain top soil moisture (−2 cm) as well as soil temperature (−2 and −50 cm) input for the RTM. The resulting uncertainty in the modeled soil moisture data is comparable to the envisaged accuracy of 0.04 $m^3$ $m^{-3}$ for the SMOS and SMAP missions.

In general, the soil penetration depth of microwave sensors such as SMOS is a function of dielectric permeability and electrical conductivity which will be mainly influenced by soil moisture, and is therefore not constant in time and space. Nevertheless, Kerr [26] stated that the penetration depth can be assumed to be approximated as the average over the first 5 cm. In order to simulate this condition, the computed soil moisture from the HYDRUS outputs is averaged over the first 5 cm. In order to mimic the typical temporal sampling of a spaceborne instrument operating in a sun-synchronous polar orbit we selected data every three days, in this case simulations for 6:00 a.m. local time. In addition, a random error of 2 K has been added to the **Tb** computations to mimic instrument noise.

The RTM used here is the L-band Microwave Emission of the Biosphere Model (L-MEB) L-MEB proposed by [27]. L-MEB parameters are sampled uniformly in their feasible range. Here the Latin hypercube sampling [28] without any further constraints is used to generate the plausible initial parameter sets. Then model states replicates and the parameters ensemble are incorporated in the L-MEB model and forwarded in time. A Sequential Importance Resampling Particle Filter (SIR-PF) is applied to assimilate synthetic SMOS observations. A persistent overestimation of simulated **Tb** by a factor of 5 % was assumed simulating a typical offset as reported in the previous section. The prior "bias" uncertainty range is ±20 %, i.e., $\mathbf{Tb}_{biased} = \mathbf{Tb}_{true} \pm 0.2 * \mathbf{Tb}_{true}$. After calculating the weight vector for the model states, resampling is applied not only for the states but also for the parameters. The parameters include temporal variable τ (two growth cycles with harvest), temporal stable *hs*, and the systematic difference term. We adopted the procedure by Moradkhani et al. [29], who recommended a minor parameter particle perturbation after each assimilation step to avoid sample impoverishment. Here we used a minor perturbation value of 2 % of the prior parameter range.

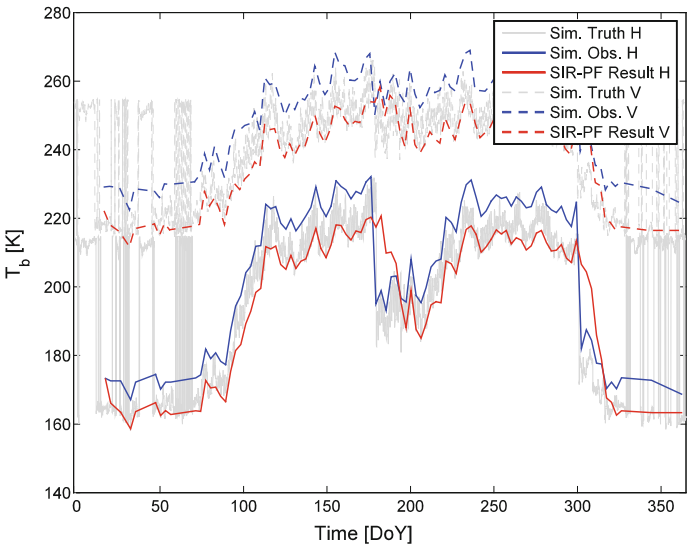


**Fig. 3** Simulated "true" **Tb**, simulated **Tb** observations in ascending node (morning overpass) with +5 % bias and particle filtering result

## *4.2 Results of the 1D Experiment*

Figure 3 shows the "true" **Tb** and the 5 % "bias" of simulated L-band **Tb**. The latter are placed at the top border of the diurnal Tb cycle, which is not plausible for SMOS ascending node data for the simulated region. Nevertheless, the SIR-PF result accounts for the systematic difference in simulated **Tb** observations and represents colder day times which occur in the early mornings.

Moreover, the two growth cycles were visible by relatively higher **Tb**, which are rapidly reduced during harvest. In general, the SIR-PF was able to correctly follow the vegetation growth, but after harvest and the following period a systematic discrepancy occurred in terms of a time delay of assimilated **Tb** compared to the "simulated truth." This behavior could be explained by the uncertainty evolution of $\tau$ and *hs* (Fig. 4). While $\tau$ was able to follow the dynamic growth of the vegetation, harvest and corresponding large change of opacity were too abrupt to be estimated in time. One reason for this temporal mismatch is that the assimilation was performed only every 72 h. The second reason is that after several resamplings during the stable growth season the uncertainty of $\tau$ was strongly reduced with all particles having a value of $\tau \sim 0.3$ and without particles with $\tau \sim 0$. The minor parameter perturbation, which was implemented to account for sample impoverishment, helped to change $\tau$ to a correct level. It took approximately 10 days until the correct $\tau$ was retrieved. This type of delay was expected; an immediate response would only be possible if the model predictions were very uncertain while the measurements were characterized by very small uncertainty. To a certain extent, *hs* as well

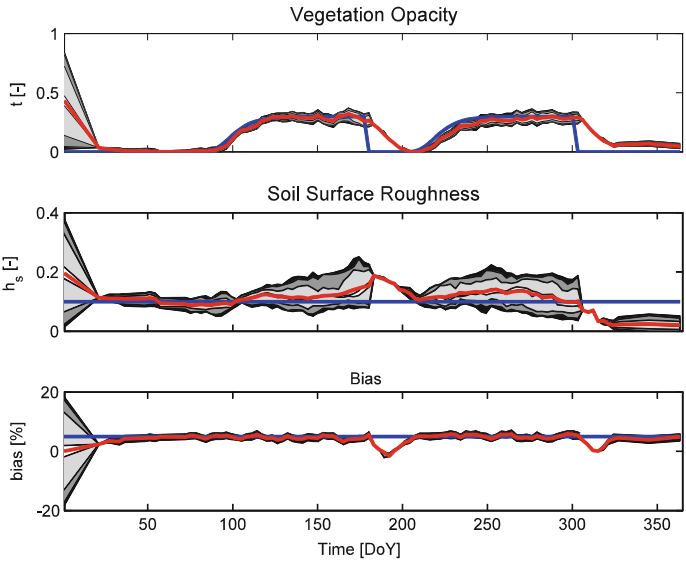


**Fig. 4** Uncertainty evolution of τ [−] (*top*), roughness parameter *hs* [−] (*middle*), and bias [%] (*bottom*). *Shaded areas* correspond to 95, 90, 68, and 10 percentile confidence intervals; *red line* shows the average and the *blue line* the simulated true value

as the bias estimation compensated for inadequately estimated τ during harvest. In light of this reason we excluded the frozen soil conditions from the analysis. To a certain extent, this problem can be solved by assimilation of multi-incidence angle **Tb** observations in real SMOS data.

## 5 Radiative Transfer Parameter Estimation in a 2D System

Also in a 2D system accurate estimates of the vegetation opacity and soil surface roughness at the catchment scale are very important for the successful retrieval of surface soil moisture from SMOS **Tb**. These parameters can be estimated from SMOS **Tb** by inverse modeling or ensemble-based data assimilation systems which compare **Tb** observations with corresponding simulations, subject to different distributions of soil surface roughness and vegetation opacity until a close fit is found. The simulations require a RTM which has to be forced with appropriate meteorological, soil condition, and vegetation data surface data. Forcing data are seldom perfect, they usually contain uncertainties. Just like other satellite products, SMOS **Tb** observations are uncertain. It is therefore necessary to address these sources of uncertainty in order to accurately estimate soil surface roughness and vegetation opacity from SMOS **Tb**. In this work we numerically investigate the

feasibility of estimating soil surface roughness and vegetation opacity in the Rur and Erft catchments in Germany based on SMOS **Tb** simulated with the Community Microwave Emission Modeling (CMEM) [21]. We propose a novel grid search approach to parameter estimation which explicitly accounts for these sources of uncertainty during the inversion for spatial and temporal evolution of soil surface roughness and vegetation opacity in the Rur and Erft catchments.

## *5.1 Model Set up and Numerical Experiment*

CMEM meteorological forcing includes snow depth in water equivalent (m), snow density (kg/m$^3$), soil temperature (K) for three levels, skin temperature (K), air temperature (K), and volumetric soil moisture (m$^3$ m$^{-3}$) for three levels. Soil condition forcing includes surface geopotential (km), sand fraction (%), and clay fraction (%). Vegetation forcing includes fraction and types of low vegetation, fraction and types of high vegetation, leaf area index of low vegetation for each pixel. These meteorological, soil condition, and vegetation input were then used to force CMEM and simulate **Tb** at 1 km resolution with a temporal frequency of 3 days at 6:00 a.m. and a look angle of 40° to mimic the SMOS morning overpass. These input data were generated from a 1 km resolution Community Land Model (CLM [21]) set up for the region. Based on a plausible set of radiative transfer model parameters (RTP, Table 1) for the Rur and Erft catchments, a noise-free input data **Tb** maps in horizontal and vertical polarizations were generated and the reference data set for the simulations was determined by considering an observation error ($\mathbf{Tb}_{p,ref} = \mathbf{Tb}_p + (\mathbf{Tb}_p * \mathbf{Err}_{Tb})$, with $\mathbf{Tb}_p$ being the simulated noise-free L-band **Tb** at polarization *p and* $\mathbf{Err}_{Tb}$ being a vector of randomly sampled fractions between −0.05 and 0.05. To account for forcing uncertainty, 10 of the input fields were also corrupted with relative errors sampled within expected feasible ranges (Fig. 5). An inversion scheme is set up to numerically investigate the feasibility of retrieving the spatial and temporal evolution of the soil surface roughness and vegetation opacity of the Rur and Erft catchments during the period of April–September 2010. Sixty-three **Tb** maps were used and each map contains about 5287 grid cells with an average resolution of 1 km.

Taking the observation and forcing uncertainties into account, in addition to six unknown radiative transfer model parameters ($\mathbf{RTP} = [b1_{grass}, b1_{crops}, b2_{grass}, \tau_{NAD,con}, \tau_{NAD,dec}, \sigma]$ see Table 1), the inverse problem of estimating the spatio-temporal evolution of soil surface roughness and vegetation opacity maps involves over 2 million unknowns. Cognizant of the ill-posedness of the inverse problem, and the challenge involved in tackling such a highly multi-dimensional problem, we propose a novel inversion method with a fairly good flexibility to handle very high multi-dimensionality called partial grid search to parameter estimation (PAGSAPE). In the PAGSAPE method, sensitivity curves for each of the unknowns are generated while other unknowns are held constant and properties of the sensitivity curves so generated are exploited to get plausible estimates of the unknowns.

**Table 1** Reference radiative transfer parameters used in this study

| Variable | Parameter | Value | | Ref. |
|---|---|---|---|---|
| | | Hpol | Vpol | |
| Incidence | $\theta$ | 40° | | |
| Roughness | $Q$ | 0 | | [27] |
| | $N_{C3crops}$ | 0 | −1 | [27] |
| | $N_{C3grass}$ | 1 | 0 | [27] |
| | $N_{dec}$ | 1 | 2 | [5] |
| | $N_{con}$ | 1.75 | 0 | [5] |
| | $\delta$ | 0.3809 | | [27] |
| | $H_{max}$ | 0.3 | | [23] |
| Vegetation | $\omega_{C3grass}$ | 0.05 | 0.05 | [27] |
| | $\omega_{C3crops}$ | 0 | 0 | [27] |
| | $\omega_{dec}$ | 0.07 | 0.07 | [27] |
| | $\omega_{con}$ | 0.08 | 0.08 | [27] |
| | $bw_0$ | 0.41 | | [3] |
| | $w_0$ | 0.34 | | [3] |
| | $ttp_{C3crops}$ | 1 | 2 | [27] |
| | $ttp_{C3grass}$ | 1 | 1 | [27] |
| | $ttp_{dec}$ | 0.49 | 0.46 | [5] |
| | $ttp_{con}$ | 0.8 | 0.8 | [5] |
| | $b1_{C3grass}$ | 0.0375 | | [27] |
| | $b1_{C3crops}$ | 0.05 | | [27] |
| | $b2_{C3grass}$ | 0.05 | | [27] |
| | $b2_{C3crops}$ | 0.05 | | [27] |
| | $\tau_{nad,dec}$ | 0.7 | | [5] |
| | $\tau_{nad,con}$ | 0.69 | | [5] |

$\theta$, $N$, $H_{min}$, and $H_{max}$ are, respectively, the incidence angle, the maximum soil roughness parameter, and the minimum soil roughness parameter, while $b1_{C3grass}$, $b1_{C3crops}$, $b2_{C3grass}$, and $b2_{C3crops}$ are the Wigneron et al. [27] parameters for the low vegetation opacity at nadir and $\tau_{NAD,for}$ is the optical depth for high vegetation ($\tau_{NAD,dec}$ for deciduous forest and $\tau_{NAD,con}$ for coniferous forest) at nadir

The feasible sampling range for each parameter is finely discretized into several units. To generate sensitivity curves for the radiative transfer parameters (RTP), the lower boundaries of the ranges of all other parameters (both RTP and forcing errors) are considered as the first proposal. Each parameter is systematically varied within its feasible range and forward model runs are performed while the rest of the parameters are held constant. Equation 1 is used to compute objective functions values for the variation:

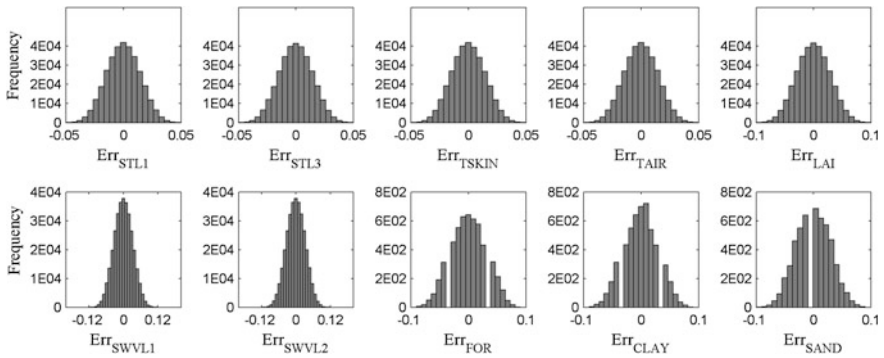


**Fig. 5** Frequency distribution of the forcing data errors (relative). $\mathbf{Err}_{\mathrm{STL1}}$, $\mathbf{Err}_{\mathrm{STL3}}$, $\mathbf{Err}_{\mathrm{Tair}}$, $\mathbf{Err}_{\mathrm{LAI}}$, $\mathbf{Err}_{\mathrm{SWVL1}}$, $\mathbf{Err}_{\mathrm{STVWL3}}$, $\mathbf{Err}_{\mathrm{FOR}}$, $\mathbf{Err}_{\mathrm{CLAY}}$, and $\mathbf{Err}_{\mathrm{SAND}}$ are, respectively, the relative errors on soil temperature level 1 (0–5 cm), soil temperature level 3 (15–30 cm), air temperature low vegetation leaf area index, volumetric soil moisture content level 1, volumetric soil moisture content level3, forest fractions, clay fraction, and sand fraction of the soil

$$\mathrm{OF}_1 = 0.5\sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\left(\mathbf{Tbh}_{\mathbf{all}} - \mathbf{Tbh}^{*}_{\mathbf{all}}(\mathrm{RTP})\right)^2 + \left(\mathbf{Tbv}_{\mathbf{all}} - \mathbf{Tbv}^{*}_{\mathbf{all}}(\mathrm{RTP})\right)^2\right)} \quad (1)$$

$\mathrm{OF}_1$ is the overall root mean square deviation between $N$ reference **Tb** at horizontal and vertical polarizations ($\mathbf{Tbh}_{\mathbf{all}}$, $\mathbf{Tbv}_{\mathbf{all}}$) and their corresponding simulations $\left(\mathbf{Tbh}^{*}_{\mathbf{all}}, \mathbf{Tbv}^{*}_{\mathbf{all}}\right)$. To generate the "pseudo sensitivity curves" for forcing data errors, all points in the matrix of a considered forcing data field are simultaneously and systematically varied within its feasible sampling range while the radiative transfer parameters and other forcing data fields are held constant. Equation 2 is used to estimate the vertical and horizontal observation errors ($\mathbf{Err}_{\mathrm{Tbh}}$, $\mathbf{Err}_{\mathrm{Tbv}}$) and the corresponding forcing error ($\mathbf{Err}_d$) for each input field. The sensitivity curves are derived with respect to each point in the considered forcing data field under perturbation. This results in as many sensitivity curves as the number of observation points. For static forcing fields, Eq. 2 is re-arranged and a pixel-wise minimum is determined from the sensitivity curves of each point of the catchment over time.

$$\mathrm{OF}_{2p} = \left|\left(\mathbf{Tbp}_{\mathbf{all}}\left(\mathbf{Err}_{\mathbf{Tbp}}\right) - \mathbf{Tbp}^{*}_{\mathbf{all}}(\mathbf{Err}_d)\right)\right| \quad (2)$$

If the feasible sampling range of the error on each observation point is discretized into N units and the feasible sampling range of the error on each forcing data point is discretized into M units, the estimation of the forcing error of each point and the corresponding observation error of that point involves M forward runs of CMEM and (2M × 2N) objective function evaluations in each inversion loop. From the sensitivity curve of each parameter, another estimate $X_{\mathrm{new}}$ of the

parameter $X_{\min}$ obtained from the minimum of the sensitivity curve is possible by applying the Newton's method of optimization expressed as:

$$X_{\text{new}} = X_{\min} - \frac{\text{OF}_1(X_{\min})}{\text{OF}_2(X_{\min})} \tag{3}$$

where $\text{OF}_1(X_{\min})$ and $\text{OF}_2(X_{\min})$ are, respectively, the first and second derivatives of the sensitivity curve evaluated at its minimum $X_{\min}$. The derivatives $\text{OF}_1(X_{\min})$ and $\text{OF}_2(X_{\min})$ are obtained from the corresponding sensitivity curve using a finite difference approximation about $X_{\min}$. In this study we are however not interested in parameter estimates from the Newton's method. We instead use the method to reduce the feasible sampling range of each parameter so as to enable faster convergence to a plausible set of parameters.

By computing the first and second derivatives of the whole sensitivity curve with respect to each point on the sensitivity curve, points along the sensitivity curve with a positive second derivative fall in a region, which we will refer to in this work as the Newton confidence zone (NCZ). The point of the NCZ with the smallest ε (where ε is the absolute difference between the point and its Newton–Raphson estimate) value $X_{\text{est}}$ is considered trustworthy and is compared against $X_{\min}$, the estimate based on the minimum of the sensitivity curve. If the $X_{\min}$ differs from $X_{\text{est}}$ by more than $2\Delta X$, where $\Delta X$ is the discretization step of the parameter under perturbation, $X_{\min}$ is rejected in favor of $X_{\text{est}}$ otherwise, $X_{\min}$ is accepted and the NCZ is also considered trustworthy. When $X_{\min}$ is rejected in favor of $X_{\text{est}}$, the whole feasible sampling range of the parameter is adopted as the NCZ. Parameter estimates determined using this procedure are used as proposals for the next inversion loop and a new set of sensitivity curves are generated again, while new parameter estimates are determined again. This inversion procedure continues until the parameter estimates do not change significantly over at least five inversion loops. To speed up convergence the inversion is constrained within the narrowest NCZ of each parameter as PAGSAPE proceeds to further loops.

## *5.2 Results of the 2D Experiment*

Figure 6 shows the simulated soil surface roughness and vegetation opacity for the Rur and Erft catchments for chosen days of the year while Fig. 7 shows the post-inversion root mean square deviation for both parameters. Spatio-temporal changes are noticeable especially for soil surface roughness (Fig. 6) and the inversion results in a general improvement of predictions for soil surface roughness and vegetation opacity in the north of the catchment. Such improvements are not very remarkable in the south of the catchment. The lack of remarkable improvements in the south of the catchment is due to dominant forest vegetation which greatly attenuates SMOS **Tb** signals. Accurate estimation of soil surface roughness and vegetation opacity also depends on accuracy of the forcing data. Figure 4 shows

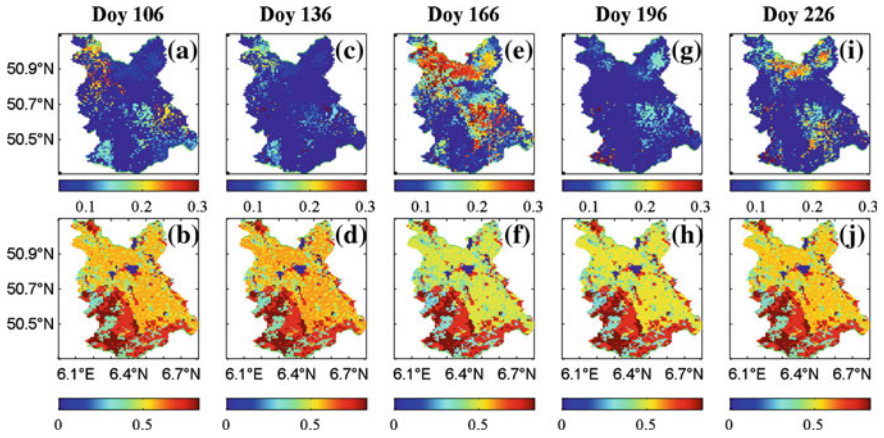


**Fig. 6** CMEM-simulated soil surface roughness (**a**, **c**, **e**, **g**, and **i**) and vegetation opacity (**b**, **d**, **f**, **h**, and **j**) in the Rur and Erft catchments for selected days of the year (Doy) in 2010

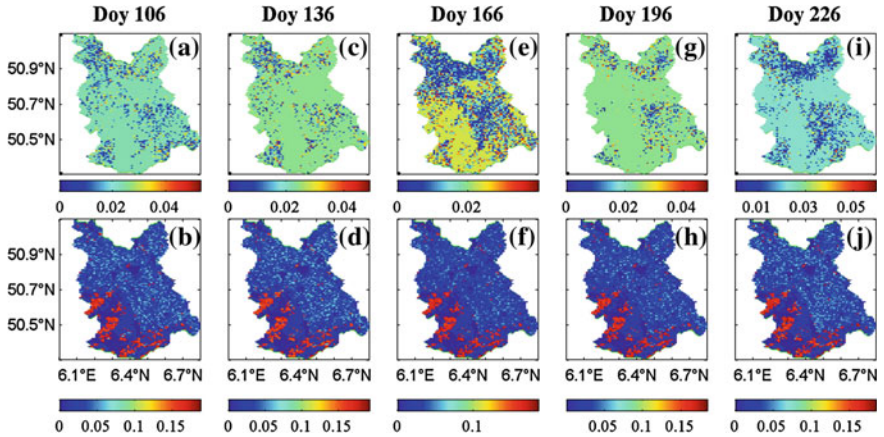


**Fig. 7** Post-inversion root mean square deviation of residuals for soil surface roughness (*top panels*) and vegetation opacity (*bottom panels*) in the Rur and Erft catchments

that the improvement of soil surface roughness and vegetation opacity estimates in the north of the catchment is a response from the improvement of soil moisture and temperature profiles in that part of the catchment. Compared to ensemble-based data assimilations systems which can be used to handle similar problems, the PAGSAPE is relatively more flexible and less computationally demanding. By fixing all other parameters during the estimation of each parameter, the PAGSAPE method avoids the curse of dimensionality and takes full advantage of the system sensitivity to drive the search forward and escape from convergence issues emanating due to parameter correlations.

## 6 Conclusion and Outlook

In the Rur and Erft catchments, the correlation of the SMOS L2 product to a modeled soil moisture reference is relatively larger in spring, but decreasing in summer and further decreasing in autumn for the years 2010–2012. Here, a more detailed analysis of the vegetation parameterization is suggested. A dry bias as compared to the modeled soil moisture reference is relative stable and varies between 0.15 and 0.19 $m^3\ m^{-3}$. These validation results indicate that there are two aspects influencing the soil moisture product accuracy: (i) As an external source RFI has a large impact on the accuracy level and needs to be adequately filtered, and (ii) the parameterization of the RTMs needs to be improved. The latter issue was addressed in this study by two parameter estimation approaches.

One suggestion presented in this chapter to enhance the accuracy of the SMOS Level 2 soil moisture product is by means of data assimilation with dual state-parameter retrieval. A parameter retrieval system within a data assimilation framework for the L-band Microwave Emission of the Biosphere (L-MEB) model is evaluated by a synthetic experiment. We analyzed the ability of the system to estimate temporally constant as well as variable parameters, namely soil surface roughness and vegetation opacity, in the presence of a systematic difference of +5 % between synthetic true Tb and synthetic Tb observations. The analysis shows a good performance to simultaneously estimate vegetation opacity, soil surface roughness as well as this "bias". Further enhancements may be obtained by constraining the "bias" estimation. Moreover, it was shown by Montzka et al. [23] that real SMOS Tb observations in ascending node can be assimilated to L-MEB to retrieve vegetation opacity, soil surface roughness, and the systematic difference.

Another suggestion to enhance the SMOS Level 2 soil moisture product is by means of the PAGSAPE approach. This numerical case study of the Rur and Erft catchments shows that it is feasible to obtain plausible estimates of the spatial and temporal evolution of soil surface roughness and vegetation opacity from SMOS **Tb** downscaled to 1 km. However, good estimates are mostly possible where the vegetation is not too dense. Other studies [30] in the region indicate that radio frequency interference is another factor which may hamper the retrieval of good estimates of soil surface roughness and vegetation opacity from SMOS **Tb**. This work also underscores the difficulty to expect in any attempt to estimate soil surface moisture from **Tb** under dense vegetation based on **Tb**. The PAGSAPE developed for this study can be applied to the parameterization of other earth systems. Due to its high flexibility to handle multi-dimensionality and relatively good rapidity, the next step forward is to apply the method for the joint estimation of soil surface roughness, vegetation opacity, and soil moisture at the catchment scale. This study was based on L-band SMOS **Tb** simulations for a look angle of 40°. Another aspect warranting investigation will be to explore the information content of more angles in the feasible view range of about 20°–40°.

**Acknowledgments** This study was supported by the German Ministry of Economics and Technology through the German Aerospace Center (50EE1040) and by the European Space Agency (Support to Science Element (STSE) Program: SMASPARES). In situ soil moisture for the Rur and Erft catchments were made available by the Terrestrial Environmental Observatories (TERENO) initiative.

## References

1. Kerr YH, Waldteufel P, Wigneron JP, Martinuzzi JM, Font J, Berger M (2001) Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission. IEEE Trans Geosci Remote Sens 39(8):1729–1735
2. Mecklenburg S, Drusch M, Kerr YH, Font J, Martin-Neira M, Delwart S, Buenadicha G, Reul N, Daganzo-Eusebio E, Oliva R, Crapolicchio R (2012) ESA's soil moisture and ocean salinity mission (SMOS). IEEE Trans Geosci Remote Sens 50(5):1354–1366
3. Kerr Y, Waldteufel P, Richaume P, Davenport I, Ferrazzoli P, Wigneron J-P (2010) SMOS level 2 processor soil moisture ATBD. Toulouse SO-TN-ESL-SM-GS-0001, 24/10/2010, 2010
4. de Rosnay P, Calvet JC, Kerr Y, Wigneron JP, Lemaitre F, Escorihuela MJ, Sabater JM, Saleh K, Barrie JL, Bouhours G, Coret L, Cherel G, Dedieu G, Durbe R, Fntz NED, Froissard F, Hoedjes J, Kruszewski A, Lavenu F, Suquia D, Waldteufel P (2006) SMOSREX: a long term field campaign experiment for soil moisture and land surface processes remote sensing. Remote Sens Environ 102(3–4):377–389
5. Grant JP, Wigneron JP, Van de Griend AA, Kruszewski A, Sobjaerg SS, Skou N (2007) A field experiment on microwave forest radiometry: L-band signal behaviour for varying conditions of surface wetness. Remote Sens Environ 109(1):10–19
6. Guglielmetti M, Schwank M, Matzler C, Oberdorster C, Vanderborght J, Fluhler H (2008) FOSMEX: forest soil moisture experiments with microwave radiometry. IEEE Trans Geosci Remote Sens 46(3):727–735
7. Juglea S, Kerr Y, Mialon A, Wigneron JP, Lopez-Baeza E, Cano A, Albitar A, Millan-Scheiding C, Antolin MC, Delwart S (2010) Modelling soil moisture at SMOS scale by use of a SVAT model over the Valencia anchor station. Hydrol Earth Syst Sci 14(5):831–846
8. Panciera R, Walker JP, Kalma JD, Kim EJ, Hacker JM, Merlin O, Berger M, Skou N (2008) The NAFE'05/CoSMOS data set: toward SMOS soil moisture retrieval, downscaling, and assimilation. IEEE Trans Geosci Remote Sens 46(3):736–745
9. Saleh K, Kerr YH, Richaume P, Escorihuela MJ, Panciera R, Delwart S, Boulet G, Maisongrande P, Walker JP, Wursteisen P, Wigneron JP (2009) Soil moisture retrievals at L-band using a two-step inversion approach (COSMOS/NAFE'05 Experiment). Remote Sens Environ 113(6):1304–1312
10. Wigneron JP, Waldteufel P, Chanzy A, Calvet JC, Kerr Y (2000) Two-dimensional microwave interferometer retrieval capabilities over land surfaces (SMOS mission). Remote Sens Environ 73(3):270–282
11. Merlin O, Walker JP, Panciera R, Escorihuela MJ, Jackson TJ (2009) Assessing the SMOS Soil Moisture Retrieval Parameters With High-Resolution NAFE'06 Data. IEEE Geosci Remote Sens Lett 6(4):635–639
12. Hasan S, Montzka C, Rüdiger C, Ali M, Bogena H, Vereecken H (2014) Soil moisture retrieval from airborne L-band passive microwave using high resolution multispectral data. ISPRS J Photogrammetry Remote Sens 91:59–71

13. Montzka C, Bogena HR, Weihermüller L, Jonard F, Bouzinac C, Kainulainen J, Balling JE, Loew A, Dall'Amico JT, Rouhe E, Vanderborght J, Vereecken H (2013) Brightness temperature and soil moisture validation at different scales during the SMOS validation campaign in the Rur and Erft catchments, Germany. IEEE Trans Geosci Remote Sens 51 (3):1728–1743. doi:10.1109/TGRS.2012.2206031
14. Zacharias S, Bogena H, Samaniego L, Mauder M, Fuss R, Putz T, Frenzel M, Schwank M, Baessler C, Butterbach-Bahl K, Bens O, Borg E, Brauer A, Dietrich P, Hajnsek I, Helle G, Kiese R, Kunstmann H, Klotz S, Munch JC, Papen H, Priesack E, Schmid HP, Steinbrecher R, Rosenbaum U, Teutsch G, Vereecken H (2011) A network of terrestrial environmental observatories in Germany. Vadose Zone J 10(3):955–973. doi:10.2136/Vzj2010.0139
15. Bogena H, Kunkel R, Puetz T, Vereecken H, Kruger E, Zacharias S, Dietrich P, Wollschlager U, Kunstmann H, Papen H, Schmid HP, Munch JC, Priesack E, Schwank M, Bens O, Brauer A, Borg E, Hajnsek I (2012) TERENO—long-term monitoring network for terrestrial environmental research. Hydrol Wasserbewirts 56(3):138–143
16. Bogena HR, Herbst M, Huisman JA, Rosenbaum U, Weuthen A, Vereecken H (2010) Potential of wireless sensor networks for measuring soil water content variability. Vadose Zone J 9:1–12
17. Al Bitar A, Leroux D, Kerr YH, Merlin O, Richaume P, Sahoo A, Wood EF (2012) Evaluation of SMOS soil moisture products over continental U.S. using the SCAN/SNOTEL network. IEEE Trans Geosci Remote Sens 50(5):1572–1586. doi:10.1109/TGRS.2012.2186581
18. Jasper K, Calanca P, Fuhrer J (2006) Changes in summertime soil water patterns in complex terrain due to climatic change. J Hydrol 327(3–4):550–563. doi:10.1016/j.jhydrol.2005.11.061
19. Oliva R, Daganzo-Eusebio E, Kerr YH, Mecklenburg S, Nieto S, Richaume P, Gruhier C (2012) SMOS radio frequency interference scenario: status and actions taken to improve the RFI environment in the 1400–1427 Mhz band. IEEE Trans Geosci Remote Sens 50(5):1427–1439
20. Aksoy M, Johnson JT (2013) A study of SMOS RFI over North America. IEEE Geosci Remote Sens Lett 10(3):515–519. doi:10.1109/LGRS.2012.2211993
21. Mironov VL, Dobson MC, Kaupp VH, Komarov SA, Kleshchenko VN (2004) Generalized refractive mixing dielectric model for moist soils. IEEE Trans Geosci Remote Sens 42(4):773–785. doi:10.1109/Tgrs.2003.823288
22. Dobson MC, Ulaby FT, Hallikainen MT, Elrayes MA (1985) Microwave dielectric behavior of wet soil. 2. Dielectric mixing models. IEEE Trans Geosci Remote Sens 23(1):35–46. doi:10.1109/Tgrs.1985.289498
23. Montzka C, Grant JP, Moradkhani H, Franssen HJH, Weihermuller L, Drusch M, Vereecken H (2013) Estimation of radiative transfer parameters from L-band passive microwave brightness temperatures using advanced data assimilation. Vadose Zone J 12(3). doi:10.2136/Vzj2012.0040
24. Bauer J, Weihermuller L, Huisman JA, Herbst M, Graf A, Sequaris JM, Vereecken H (2012) Inverse determination of heterotrophic soil respiration response to temperature and water content under field conditions. Biogeochemistry 108(1–3):119–134
25. Weihermüller L, Huisman JA, Lambot S, Herbst M, Vereecken H (2007) Mapping the spatial variation of soil water content at the field scale with different ground penetrating radar techniques. J Hydrol 340(3–4):205–216
26. Kerr YH (2007) Soil moisture from space: where are we? Hydrogeol J 15(1):117–120
27. Wigneron JP, Kerr Y, Waldteufel P, Saleh K, Escorihuela MJ, Richaume P, Ferrazzoli P, de Rosnay P, Gurney R, Calvet JC, Grant JP, Guglielmetti M, Hornbuckle B, Matzler C, Pellarin T, Schwank M (2007) L-band microwave emission of the biosphere (L-MEB) model: description and calibration against experimental data sets over crop fields. Remote Sens Environ 107(4):639–655
28. Iman RL, Helton JC, Campbell JE (1981) An approach to sensitivity analysis of computer-models.1. Introduction, input variable selection and preliminary variable assessment. J Qual Technol 13(3):174–183

29. Moradkhani H, Hsu KL, Gupta H, Sorooshian S (2005) Uncertainty assessment of hydrologic model states and parameters: sequential data assimilation using the particle filter. Water Resour Res 41(5):W05012. doi:10.1029/2004WR003604
30. Rötzer K, Montzka C, Bogena H, Wagner W, Kerr YH, Kidd R, Vereecken H (2014) Catchment scale validation of SMOS and ASCAT soil moisture products using hydrological modeling and temporal stability analysis. J Hydrol 519, 934–946. doi:10.1016/j.jhydrol.2014.07.065

# PROgRESSIon—Investigating the Prototyping of Operational Estimation of Energy Fluxes and Soil Moisture Content Using a Variant of the "Triangle" Inversion Methodology

**George P. Petropoulos and Gareth Ireland**

**Abstract** Accurately estimating the spatio-temporal distribution of energy, mass and momentum at the surface–atmosphere interface can help develop a better understanding of the complex interactions of the Earth system. By linking deterministic land surface process model, such as SimSphere, to the spatialised information provided by Earth observation (EO) data, a more powerful synergistic avenue can be developed to take advantage of the temporal and spatial benefits of both modelling and EO-based approaches. The "triangle" utilises the distribution of land surface temperature (LST) and vegetation index (VI) formed by a satellite-derived scatterplot, linked with SimSphere under a full range of vegetation cover and soil moisture, to derive spatial estimates of energy fluxes and soil moisture content (SMC). To this end, the objective of this study was to implement the "triangle" technique using Advanced Along-Track Scanning Radiometer (AATSR) satellite data products to derive and subsequently validate spatially explicit maps of land surface heat fluxes and SM for different ecosystems in Europe. The "triangle"-derived estimations of soil moisture exhibited a minor overestimation of the in-situ observations, and an average error of 0.097 vol $vol^{-1}$. In overall, results were comparable to those of previous validation studies of the "triangle" implementation. Results for the LE and *H* fluxes were within the accuracy range of 50 $Wm^{-2}$, with root mean square difference (RMSD) of 41.15 $Wm^{-2}$ and 44.37 $Wm^{-2}$, respectively. Furthermore, there was a good agreement between the "triangle"-derived and in-situ observed instantaneous LE and *H* fluxes, exhibited by high *R* values (0.88 and 0.69, respectively). Our study is one of the few studies

G.P. Petropoulos (✉) · G. Ireland
Department of Geography and Earth Sciences, University of Aberystwyth, Aberystwyth, Ireland
e-mail: george.petropoulos@aber.ac.uk



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_7

validating the "triangle" over different ecosystems in Europe. It is a significant step forward in supporting the operational development of this method using remote sensing data in deriving key land surface parameters on a global scale.

## 1 Introduction

Earth's atmosphere, hydrosphere and biosphere are under a constant exchange of mass, energy and momentum via flux of sensible ($H$) and latent heat (LE) [1]. Sensible heat has a strong influence on the turbulent nature of the near-surface atmosphere by changing molecular movement through heat transfer, whereas latent heat is directly linked to the global water and carbon cycle, with more than half of the land surface evapotranspiration occurring as plant-controlled stomatal transpiration [2]. Quantifying the spatio-temporal dynamics of these fluxes can help develop an understanding of the complex processes and feedback mechanisms that interplay within the Earth system, and are of significant practical value in a large number of regional and global scale applications [3]. Soil moisture is the main control on the partitioning of the available energy at the Earth's surface into sensible and latent heat exchange through evaporation and transpiration processes [4]. Thermal inertia, temperature and shortwave albedo are also affected by the spatial distribution of soil moisture, where accurate information on its distribution can be important to understand evaporation, infiltration and runoff processes [5]. Frequent soil moisture observations at different spatial scales are thus of crucial importance to many environmental and bio-geophysical applications, such as flood forecasting [6, 7], meteorology [8, 9], agricultural applications [10, 11], and global and regional circulation climate models [12, 13], amongst others.

In view of the importance of information on the spatial and temporal dynamics of these mass and energy fluxes, a number of deterministic mathematical models representing the complex relationships of the Earth system have recently emerged in the environmental or geotechnical modelling fields (amongst others, Gamerith et al. [14], Heydari et al. [15], Miro et al. [16], Xu et al. [17]). One such category of deterministic approaches, namely Soil–Vegetation–Atmosphere Transfer (SVAT) models, has been designed to simulate the interaction between plant canopy processes and the environment [18]. These are essentially mathematical models with the ability to simulate, at fine temporal and spatial resolutions, the multifarious and complex interactions of energy and mass transfers through the 1-D vertical column at the soil–vegetation–atmosphere interface [19–21]. They require a large set of input parameters and initial state variables that are spatially and temporally distributed, where in general, model complexity and the number of parameters increase simultaneously [22]. SimSphere is a SVAT model, first developed by Carlson and Boland

in 1978 and is currently globally distributed from the Aberystwyth University.[1] However, SVAT models, being one-dimensional vertical representations of the land surface interactions, are rarely capable of producing distributed estimates of soil moisture without the use of remotely sensed data as inputs [23].

By linking 1D SVAT models as SimSphere to the spatialised information provided by airborne or satellite Earth Observation (EO) data, a more powerful synergistic avenue can be developed to take advantage of the benefits of both modelling and EO-based approaches [24]. To this end, combining remotely sensed surface conditions within a surface temperature ($T_s$)/vegetation index (VI) feature space with a SVAT model, allows for a framework that can assess the spatial variability in mass and energy exchanges to quantify spatially variable fluxes [25].

This method, termed the "triangle", relies on the triangular distribution of land surface temperature (LST) and VI formed by a satellite-derived scatterplot, linked with SimSphere under a full range of vegetation cover and soil moisture [26]. The triangular or trapezoidal shape in a $T_s$/VI feature space is due to the fact that for the given vegetation amount present at the surface, the temperature ranges from a minimum with strongest evaporative cooling to a maximum with weakest cooling, defining the triangle's boundaries [27]. Such a triangle is characterised by four physical bounds: on the right-hand side, the slanting dry edge representing the warmest pixels and on the left-hand side representing the coldest pixels, the nearly vertical wet edge, and the bare soil and full vegetation cover limits. These edges respectively represent two limiting cases of soil moisture and evaporative fraction (EF) for each VI value (i.e. the soil surface radiant temperature limits for the highest and lowest temperatures at a given fractional vegetation cover ($F_r$) (or normalised difference vegetation index (NDVI)) [28, 29]. Variation along the triangle's base represents bare soil ranging from wet to dry (left to right), where for bare soils, at constant irradiance, $T_s$ is primarily determined by soil moisture content (SMC), via evaporative control and thermal properties of the surface [30]. The triangle's apex equates to full vegetation cover [28] (Fig. 1). A number of studies have been concerned with the examination of the main factors driving the shape of the $T_s$/VI scatterplot, an overview of which can be found in recent reviews [31, 32].

One of the ESA-funded project PROgRESSIon aims was to investigate the prototyping of operational estimation of energy fluxes and SMC from the Advanced Along-Track Scanning Radiometer (AATSR) using a variant of the "triangle" approach. In this context, the specific objectives presented herein, were twofold: (a) to derive spatially explicit maps of land surface heat fluxes and SMC over different ecosystems in Europe, from the implementation of AATSR Level 2 full resolution geophysical products as inputs to the "triangle" method and (b) to validate the maps of land surface fluxes and SMC against in-situ measurements acquired from the CarboEurope flux network to assess the accuracy of the "triangle" method in estimating these parameters.

[1] http://www.aber.ac.uk/en/iges/research-groups/earth-observation-laboratory/research/simsphere/.

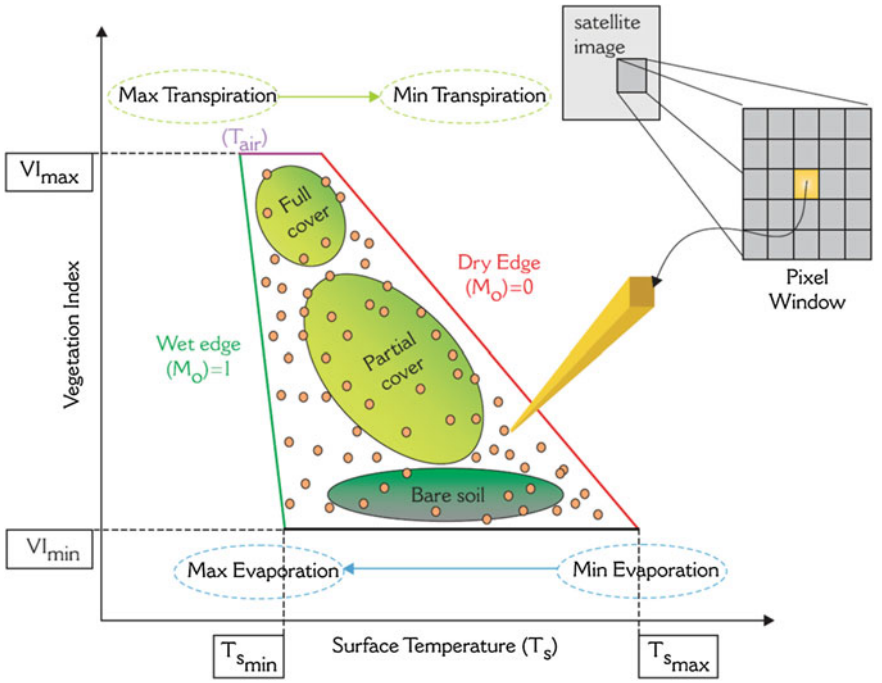


**Fig. 1** Summary of the main physical properties and interpretations of the satellite (or airborne) derived $T_s/F_r$ feature space (adapted from Petropoulos et al. [31])

## 2 Datasets and Study Sites

The development of the various products from AATSR sensor using a variant of the "triangle" required the utilisation of three different groups of data types: satellite, land surface model and ancillary in-situ data.

### *2.1 AATSR Satellite Data*

AATSR is one of the Announcement of Opportunity (AO) instruments on board the European Space Agency (ESA) satellite ENVISAT. This operational space-borne instrument is the most recent in a series of instruments designed primarily to measure sea surface temperature (SST) with an accuracy of 0.3 K. The AATSR provides measurements of reflected and emitted radiation taken at the following wavelengths: 0.55, 0.66, 0.87, 1.6, 3.7, 11 and 12 μm. The satellite is a dual-view,

multichannel, imaging radiometer where the spatial resolution is 1 km at nadir with a 512 km swath, whereas its temporal resolution is 1–3 days [33]. In terms of AATSR satellite data, the ATS_NR_2P product was utilised in developing the retrievals of the energy fluxes and SMC within PROgRESSIon.

This is the AATSR Gridded Sea Surface Temperature Level 2 full resolution geophysical product. This product is a full spatial resolution (approximately 1 km by 1 km) product which contains the values of various geophysical parameters for each pixel, and thus the dataset is switchable dependent on the pixel classification. Parameters include SST over oceans, NDVI over land, cloud top temperature for cloud pixels, and surface brightness temperature/radiance and top of atmosphere (TOA) brightness temperature/radiance for unclassified pixels. It is distributed using 512 by 512 km minimum scenes, where near real-time (NRT) products are available in payload data handling station (PDHS) within 3 h of data take. The SST and other parameters (e.g. NDVI) are systematically derived from the gridded brightness temperatures of the AATSR Level 1B full resolution product (https://earth.esa.int/documents/10174/437508/Vol-07-Aats-4C.pdf).

## *2.2 Land Surface Process Model*

In terms of the land surface process model, the SimSphere model was used, although generally any other model with similar functionalities can be theoretically used. A detailed description of its architectural design and operation as well as an overview of its use so far can be found in Petropoulos et al. [34]. Briefly, SimSphere has been developed to simulate the various physical processes that take place as a function of time in a column that extends from the root zone below the soil surface up to a level higher than the surface vegetation canopy. The model performs simulations over a 24-h cycle, starting from a set of initial conditions given in the early morning (at 05:30 h local time) and simulates the continuous evolving interaction between soil, plant and atmospheric layers. In the present work, the most recent version of SimSphere, available from the Department of Geography and Earth Sciences of Aberystwyth University, UK (see footnote 1), was used to implement the "triangle" with the AATSR data. This is the most recent model version which has been recently enhanced by adding various functionalities described in Petropoulos et al. [35], which allowed a much faster and computationally efficient implementation of the model with the AATSR using the "triangle" method.

### 2.3 Ancillary In-Situ Data

Radiosonde Observations: Available at 06.00 h, used for SVAT parameterisation, and, optionally, in atmospherically correcting any satellite observations if required. Such data were provided from the Forecast Systems Laboratory (FSL)/National Climatic Data Center (NCDC) and University of Wyoming (Department of Atmospheric Sciences) databases[2] and were available at no cost. These data were acquired for all days for which the AATSR data had been acquired.

CarboEurope In-situ Measurements: In-situ measurements of various parameters characterising land–surface interactions required in SimSphere parameterisation and in validating the derived maps. Those include mainly meteorological forcing data (e.g. wind speed/direction, atmospheric pressure, precipitable water) as well as some other data related to vegetation condition (e.g. biome type, leaf area index (LAI), average vegetation height, etc.), soil characteristics of the area (soil type, thermal inertia, soil moisture) and location of the study region (e.g. geographical coordinates, aspect, slope). Such data were obtained from various sites belonging to global in-situ validated operational networks, namely the CarboEurope validated observational network (Table 1). CarboEurope is part of FLUXNET, which is nowadays the largest global network of micro-meteorological flux and ancillary parameter measurement sites [36]. Its role is to coordinate regional measurement networks so that ground observations of an array of parameters can be obtained at a global scale, ensuring site to site comparability, coordinating simultaneous improvements to existing network plans and the operation of a global archive and distribution centre [37–39]. All in-situ data were acquired from the CarboEurope web-site.[3]

## 3 "Triangle" Implementation Using AATSR Products

The "triangle" based on combining the biophysical properties encapsulated in a satellite-derived scatterplot developed between the surface temperature ($T_s$) and VI maps with a SVAT model (herein SimSphere) for deriving spatially distributed estimates of turbulent heat fluxes (instantaneous and daytime average ones) as well as surface soil moisture (SMC). "Triangle" implementation involved the following steps.

Initial $F_r$ and $T_s$ computation: To compute the NDVI, the following equation was used, where NIR and RED denote the near-infrared and the red surface spectral reflectance, respectively (1).

$$\mathrm{NDVI} = \frac{\rho_{\mathrm{NIR}} - \rho_{\mathrm{RED}}}{\rho_{\mathrm{NIR}} + \rho_{\mathrm{RED}}} \tag{1}$$

[2] http://weather.uwyo.edu/upperair/sounding.html.

[3] http://gaia.agraria.unitus.it/.

**Table 1** Location and characteristics of the CarboEurope flux tower sites used for the in-situ eddy covariance observations in this study

| Site name | Site abbreviation | Geographic coordinates (lat/long) | Country | Ecosystem type/land cover | Elevation (m) |
|---|---|---|---|---|---|
| Aguamarga | ES-Agu | 36.8347/−2.2511 | Spain | Annual broadleaf shrub | 195 |
| Amoladeras | ES-Amo | 36.9405/−2.0329 | Spain | Annual broadleaf shrub | 53 |
| Collelongo-Selva Piana | IT-Col | 41.8493/13.5881 | Italy | Mixed deciduous beech forest | 1645 |
| Monte Bondone | IT-MBo | 46.0296/11.0829 | Italy | Grasslands | 1547 |
| Renon/Ritten (Bolzano) | IT-Ren | 46.5878/11.4347 | Italy | Evergreen needleleaf forest | 1794 |
| Lecceto | IT-Lec | 43.3046/11.2706 | Italy | Evergreen needleleaf forest | 269 |
| Nonantola | IT-Non | 44.6898/11.0887 | Italy | Deciduous broadleaf forest | 14 |
| Malga Arpaco | IT-Mal | 46.1167/11.7028 | Italy | Evergreen needleleaf forest | 1730 |
| Bonis | IT-Bon | 39.4778/16.5347 | Italy | Evergreen needleleaf forest | 1170 |
| Negrisia | IT-Neg | 45.7476/12.4467 | Italy | Woody crop | 9 |
| Castellaro | IT-Cas | 45.0700/8.7175 | Italy | Cereal crop | – |
| Espirra | PT-Esp | 38.6394/−8.6018 | Portugal | Hardwood forest | 95 |
| Mitra (Evora) | PT-Mi1 | 38.5406/−8.0001 | Portugal | Evergreen broadleaf forest | 264.2 |
| Mitra IV Tojal | PT-Mi2 | 38.4765/−8.0246 | Portugal | Grassland | 190 |

NDVI values range from −1 to +1, where negative values correspond to an absence of vegetation. $F_r$ was then derived from NDVI following the methods of Gillies and Carlson [40] and Choudhury et al. [41]:

$$F_r = N^{*^2} \tag{2}$$

where $N^*$ is a scaled NDVI defined as:

$$N^* = \frac{\mathrm{NDVI} - \mathrm{NDVI_o}}{\mathrm{NDVI_s} - \mathrm{NDVI_o}} \tag{3}$$

where the subscripts $s$ and $o$, respectively, denote the values for dense vegetation and bare soil. The derivation of $F_r$ allowed us to plot both the SVAT-simulated and the measured surface radiant temperatures from the satellite sensor on the same scale. Finally, following the transformation of NDVI to $F_r$, $T_s$ normalization was performed:

$$T_{\text{scaled}} = \frac{T_{\text{o}} - T_{\text{min}}}{T_{\text{max}} - T_{\text{min}}} \tag{4}$$

where $T_{\text{max}}$ and $T_{\text{min}}$ are the maximum and minimum $T_{\text{s}}$, for wet vegetated pixels and for the dry/bare soil respectively interpolated from the scatterplot bounds ($T_{\text{s}}$ essentially corresponds to the radiometric temperature value of any pixel in the scene). Subsequently, SimSphere model was integrated with the satellite observations of $T_{\text{s}}$ and $F_r$ in order to derive the inversion equations that provided the spatially explicit maps of land surface fluxes and SMC. The process was composed of the following steps:

- Initially, SimSphere was parameterised using the time and geographic location as well as the site-specific atmospheric, biophysical and geophysical characteristics. Parameters were adjusted based on radiosonde observations.
- Following that, model tuning was performed where the SVAT model input parameters were further adjusted and the SVAT model was iterated repetitively until the simulated and observed extreme values of $F_r$ and $T_{\text{s}}$ in the $T_{\text{s}}/F_r$ scatterplot were matched. Specifically, initial model simulations endeavored to align observed $T_{\text{s}}$ with two end points ($\text{NDVI}_{\text{o}}$, $\text{NDVI}_{\text{s}}$) where they intersect the "dry" edge.
- Once SimSphere tuning was completed, simulation time (corresponding to the satellite overpass) was kept the same as the SVAT model ran repeatedly, varying $F_r$ and SMC over all possible values (0–100 % and 0–1, respectively), for all possible theoretical combinations of SMC, $F_r$. The result was a matrix of model outputs for a number of simulated parameters: SMC, $F_r$, $T_{\text{s}}$, LE and $H$.
- Finally, this output matrix was used to derive a series of empirical, non-linear (quadratic) equations, relating $F_r$ and $T_{\text{s}}$ to each of the other variables of interest: $H$, LE, latent heat flux ratio ($\text{LE}/R_n$) and sensible heat flux ratio ($H/R_n$). The set of physically-based relationships between the various surface-atmosphere parameters, as described by the detailed bio-physical descriptions included in SimSphere and inherent in the matrix outputs, are used to derive a series of simple, empirical relations relating each of these parameters to just the locations of $F_r$ and $T_{\text{s}}$ recorded at that location. Since these variables of $F_r$ and $T_{\text{s}}$ are derivable from the satellite data, these empirical equations were then used to derive the required spatially explicit maps of the land surfaces LE and $H$ fluxes as well as of SMC from the satellite products of $F_r$ and $T_{\text{s}}$. The quadratic polynomial equations derived from the SimSphere matrix model outputs have the general form (5) (here shown for the version relating $M_0$ to $F_r$ and $T_{\text{s}}$ and/or $T_{\text{scaled}}$):

$$M_0 = \sum_{p=0}^{3} \sum_{q=0}^{3} a_{pq} \left(T^*_{\text{scaled}}\right)^p (F_r)^q \tag{5}$$

where the coefficients $a_{p,q}$ are derived from non-linear regression between the matrix values of $F_r$, $T_{\text{scaled}}$ and $M_0$ and $p$ and $q$ vary from 0 to 3. Thus, by this method, $F_r$ and $T_{\text{s}}$ (or equally $T_{\text{scaled}}$) recorded at each location are used to

generate output values for $H$ and LE for a range of measured values of $F_r$ and $T_s$. Since these variables of $F_r$ and $T_{scaled}$ are derivable from the satellite data, empirical equations such as this can then be used to obtain the required spatially explicit maps of the LE and $H$ fluxes as well as of SMC from the satellite observations [32].

The "triangle" method was implemented for each AATSR image acquired for the selected test sites using the AATSR Gridded Surface Temperature Level 2 full resolution geophysical product.

## 4 Validation Approach

Direct comparisons between the predicted and the corresponding in-situ measurements, which had been acquired near concurrently, were conducted. Where necessary, in-situ measurements were interpolated to synchronise with the sensor overpass time on each of the selected test sites.

For the quantitative evaluation of the agreement between predicted and observed parameters of both SMC and LE/$H$ fluxes (both instantaneous and daytime-averaged), point-by-point comparisons formed the main validation approach employed. Such point-based comparisons have been the most common approach followed in analogous validation experiments of satellite-derived maps of surface energy fluxes and SMC, including past verification exercises of the "triangle" method" (e.g. [42–45]) as well as of operational products (e.g. [46]).

To quantify the level of agreement between the "triangle"-derived (inverted) measures and the reference estimates (in-situ) observations, a series of appropriate statistical measures was computed, including the mean ($X$), the RMSD, the correlation coefficient ($R$), the bias, the scatter, and the mean absolute difference (MAD) (Table 2).

## 5 Results and Discussion

The main results from the statistical comparisons between the in-situ and the predicted parameters for the case of the 2P AATSR product "triangle" implementation are summarised in Table 3. Associated scatterplots are exhibited in Figs. 2, 3, 4 and 5, whereas examples of product outputs for the case of the AATSR_2P implementation for selected days are shown in Fig. 6.

With regards to the SMC comparisons, as can be observed (Table 3 and Fig. 2), the 2P AATSR SMC estimation shows a minor overestimation of the in-situ observations, exhibiting a bias error of 0.033 vol vol$^{-1}$; however, this amount of overestimation is not significant and the model seems to estimate the in-situ values to a relatively accurate degree. The AATSR "triangle" implementation predictions

**Table 2** Definition of the quantitative measures used to assess the agreement between the "triangle"-derived estimates, and the in-situ observations

| Name | Description | Mathematical definition |
|---|---|---|
| Bias/MBE | Bias (accuracy) or mean bias error | $\text{bias} = \text{MBE} = \frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)$ |
| Scatter/MSD | Scatter (precision) or mean standard deviation | $\text{scatter} = \frac{1}{(N-1)}\sum_{i=1}^{N}\left(P_i - O_i - \overline{(P_i - O_i)}^2\right)$ |
| RMSD | Root mean square difference | $\text{RMSD} = \sqrt{\text{bias}^2 + \text{scatter}^2}$ |
| MAD | Mean absolute difference | $\text{MAD} = N^{-1}\sum_{i=1}^{N}\lvert P_i - O_i\rvert$ |

Subscripts $i = 1 \ldots N$ denotes the individual observations', $P$ denotes the predicted values, and $O$ denotes the "observed" values, in our case those obtained from the selected CarboEurope sites. The horizontal bar in scatter/MSD ratio equation denotes the mean value

**Table 3** Results for the comparison between "triangle"-derived estimates, and the in-situ observations

| AATSR 2P product | SWC | LE | $H$ | LE/$R_n$ | $H/R_n$ |
|---|---|---|---|---|---|
| Mean | 0.148 | 194.19 | 140.26 | 0.467 | 0.301 |
| Bias | 0.033 | 9.39 | −11.15 | 0.046 | −0.002 |
| Scatter | 0.091 | 40.06 | 42.94 | 0.125 | 0.145 |
| RMSD | 0.097 | 41.15 | 44.37 | 0.133 | 0.145 |
| MAD | 0.080 | 34.43 | 36.39 | 0.114 | 0.120 |
| $R$ | 0.435 | 0.88 | 0.69 | 0.707 | 0.611 |
| $N$ | 37 | 22 | 40 | 22 | 35 |

Statistical results—mean (vol vol$^{-1}$/Wm$^{-2}$), bias (vol vol$^{-1}$/Wm$^{-2}$), scatter (vol vol$^{-1}$/Wm$^{-2}$), RMSD (vol vol$^{-1}$/Wm$^{-2}$), MAD (vol vol$^{-1}$/Wm$^{-2}$)

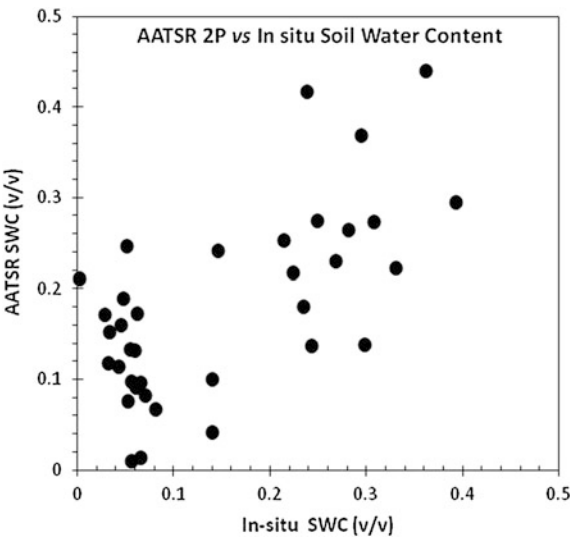


**Fig. 2** Scatterlpot comparison of the "triangle"-derived estimated, and the in-situ observed SMC values

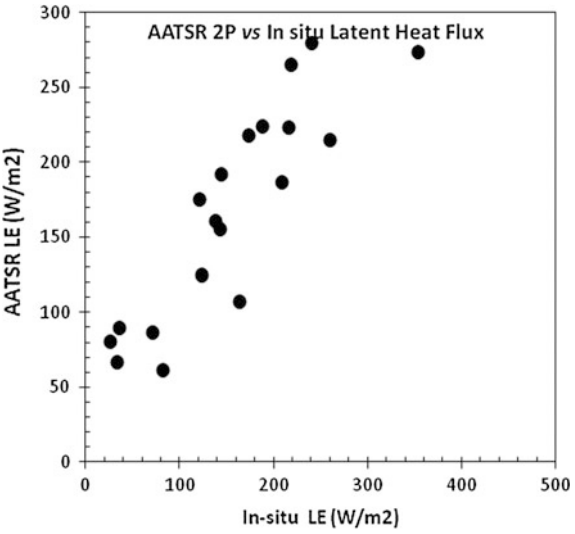


**Fig. 3** Scatterlpot comparison of the "triangle"-derived estimated, and the in-situ observed latent heat flux values

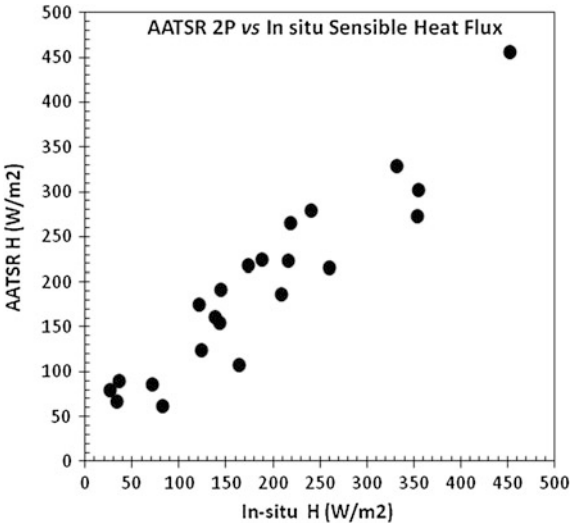


**Fig. 4** Scatterlpot comparison of the "triangle"-derived estimated, and the in-situ observed sensible heat flux values

showed a moderately good agreement with the retrieved in-situ data, ($R$ = 0.660). Error range found was relatively low (RMSD = 0.097 vol vol$^{-1}$), showing a small range of error. MAD results suggested good model prediction (MAD = 0.080 vol vol$^{-1}$), whereas scatter results (scatter = 0.091 vol vol$^{-1}$) displayed slightly higher values for the agreement between both datasets, showing to some extent a moderately unstable estimation of SMC, with minor dispersion or variance from the in-situ measurements.

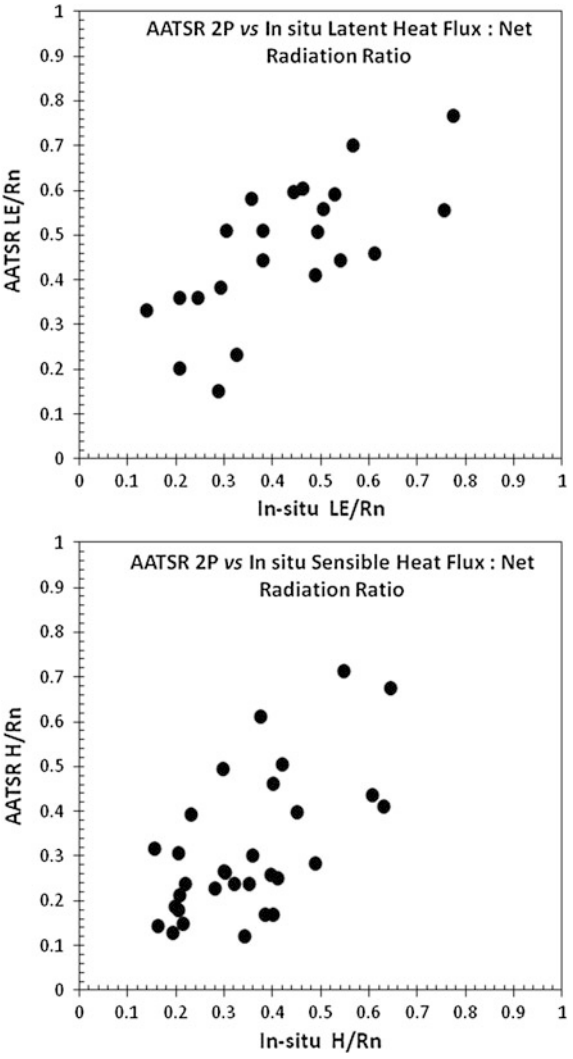


**Fig. 5** Scatterlpot comparison of the "triangle"-derived estimated, and the in-situ observed daytime-averaged latent and sensible heat flux values (daytime averaged fluxes derived from the ratio of instantaneous fluxes to radiation)

The validity of the "triangle" in deriving spatially distributed maps of instantaneous SMC has been previously examined in the literature using different satellite sensors, with varying degree of agreement reported. Results from the current study are comparable, or to some extent of better accuracy, to those reported by Capehart and Carlson (1997) who performed comparisons of SMC derived from the "triangle" method using Advanced Very High Resolution Radiometer (AVHRR) data versus SMC simulated from a soil hydrological model. They reported a low degree of correspondence ($R^2$ from 0.266 to 0.441 and an RMSD varying from 0.150 to 0.190 vol vol$^{-1}$ respectively). However, Capehart and Carlson (1997) found the "triangle"-derived SMC to consistently underestimate the SMC derived from the

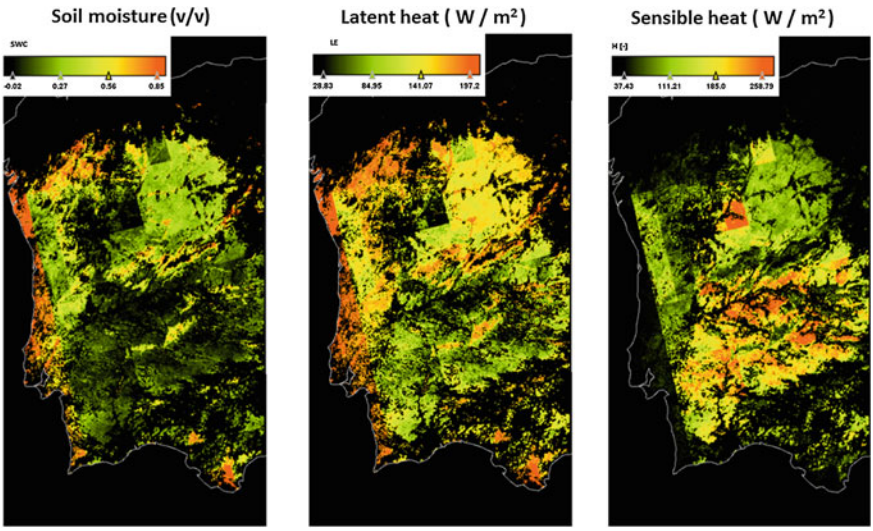


**Fig. 6** Examples of output products of the SimSphere model run for an AATSR 2P image acquired over Spain, 24 July 2011, 10:57 (95 % contrast stretch). Examples of output products (LE daily, *H* daily) of the SimSphere run for an AATSR 2P image acquired over Spain, 24 July 2011, 10:57 (95 % contrast stretch)

hydrological model, adverse to the results of the present study. Authors attributed the low correlations of their comparison to the poor mismatch in both the horizontal and vertical scales of the satellite data and the hydrological model, stating that the satellite-derived Tkin/Trad may be responding to the soil water content in a layer much shallower than the minimum resolution of the hydrological model. The results of our study are also comparable, or better, to those reported in the Gillies et al. [44] verification study of the "triangle" method. In comparisons between SMC measurements and those predicted from the "triangle" using high spatial resolution airborne data (of a spatial resolution similar to LANDSAT TM) they reported $R^2$ and standard errors in the estimation of SMC varying from 0.290 to 0.790 and from 8.73 to 8.25 %, respectively. In another study, Chauhan et al. [43] using a variant of the "triangle" and data from the Special Sensor Microwave Imager (SSM/I) and AVHRR for a site in Southern Great Plains reported a RMSD of less than 0.050 vol $vol^{-1}$ in the estimation of SMC from their proposed algorithm. However, it should be noted that this study results are not directly comparable to those reported by Chauhan et al. [43], primarily because of the following two reasons: (1) they implemented a modified version of the "triangle" (by including in the regression equation the "surface albedo" term), and (2) they were concerned mostly with sites that had no vegetation cover (bare soil). These factors significantly affected the accuracy of their results.

For the case of the instantaneous LE fluxes retrievals from the "triangle" method implementation, validation results obtained indicated the highest *R* value of all

parameters ($R$ = 0.94) (Fig. 3). The bias was relatively high, exhibiting a moderate overestimation of the in-situ data by the model predictions (Bias = 9.39 $Wm^{-2}$). However, error distribution was relatively low for the comparisons of this parameter (RMSD = 41.15 $Wm^{-2}$/MAD = 34.43 $Wm^{-2}$), with low scatter results suggesting a stable model estimation (Scatter = 40.06 $Wm^{-2}$). In terms of the LE fluxes comparisons, the agreement reported here in close correspondence with previous findings of Gillies et al. [44] who reported a mean St. Error of 34.73 $Wm^{-2}$. In this study the validation of LE derived from the NS001 multispectral scanner (30 m spatial resolution) carried on board NASA's C-130 aircraft was assessed against ground observations from FIFE (Sellers et al. [47]) and MONSOON'90 [48] field experiments. In addition, the results are also comparable to those of Brunsell and Gillies [42] who validated the "triangle" using both airborne (TIMS) and satellite (NOAA AVHRR), and field observations from the Southern Great Plains 1997 (SGP) Hydrology experiment (which was conducted in Oklahoma, USA) reporting a RMSD between the different comparison schemes ranging from 18.00 to 90.00 $Wm^{-2}$.

With regards to the instantaneous $H$ flux comparisons, overall, correlations with in-situ data were relatively good, showing a moderate agreement with the in-situ data ($R$ = 0.69), comparable to those reported earlier in the SMC comparisons (Fig. 4). There is a significant bias underestimation of the observed $H$ fluxes (Bias = −11.15 $Wm^{-2}$), showing a considerably larger estimation bias in comparison to the LE fluxes comparison (Bias = 9.39 $Wm^{-2}$). Error range is again relatively good (RMSD = 44.37 $Wm^{-2}$); however results for the LE fluxes comparisons exhibit better performance. In terms of the correspondence of the estimated $H$ fluxes to the in-situ data, results are in comparison to those reported in the "triangle" verification study of Gillies et al. [44] performed using the NS001 multispectral scanner (airborne instrument with spatial resolution similar to that of LANDSAT). They reported an $R^2$ of 0.83 and St. Error of 39.61 $Wm^{-2}$ for comparisons of the $H$ fluxes, and standard errors varying between 25.00 and 55.00 $Wm^{-2}$. Results presented herein were also comparable to those reported by Brunsell [49] and Brunsell and Gillies [42], who verified the "triangle" using airborne (from TIMS) and satellite data (from AVHRR) at SGP, USA, and reported average agreement (between the different comparison schemes) varying from 21.00 to 145.00 $Wm^{-2}$ and average agreement between the high spatial resolution (TIMS) data and the eddy fluxes (for both comparison schemes) varying from 45.00 to 80.00 $Wm^{-2}$.

In terms of the LE/$R_n$ and $H/R_n$ results, representing the daytime latent and sensible heat fluxes respectively, they exhibited lower correlation coefficients to the instantaneous flux comparisons (i.e. LE/$R_n$ = 0.707 and $H/R_n$ = 0.611) (Fig. 5). Estimation of the daytime averages of both parameters was comparable to the instantaneous fluxes comparison results, with the daytime averaged LE overestimating, and the daytime averaged $H$ underestimating the in-situ measurements. However the range of over- and under-estimation was significantly more accurate, 0.046 and −0.002 for LE/$R_n$ and $H/R_n$ respectively. Error range was also a lot higher for the daytime average fluxes in comparison to the instantaneous fluxes,

exhibited by significantly higher RMSD and MAE values (RMSD LE/$R_n$ = 0.133 and $H/R_n$ = 0.145, MAD LE/$R_n$ = 0.114 and $H/R_n$ = 0.120) respective to the required accuracy of operational retrieval. Overall, the results suggest that estimation of instantaneous $H$ and LE fluxes through the implementation of the "triangle" method was closely tied to the performance of the "triangle" in deriving the daytime averaged $H$ and LE fluxes. This is an important observation as it signifies an inherent link between the accurate retrieval of predicted instantaneous fluxes and the daytime average ones by the "triangle" method implementation. This finding is in agreement with similar past verification studies of the "triangle" method (i.e. [44] and Brunsell and Gilles [42]). In addition, generally the "triangle"-predicted LE and LE/$R_n$ had greater agreement with the observations in comparison to the derived $H$ and $H/R_n$ fluxes. This could perhaps be related to the accuracy by which the LE and $H$ are derived in SimSphere, as the same model is used to invert both the $H/R_n$ and LE/$R_n$ fluxes. Although no previous studies had assessed the implementation of the "triangle" method in the estimation of daytime averaged $H$ fluxes, the level of estimation accuracy observed for the comparisons for the $H/R_n$ fluxes parameter was similar to that found in previous LE/$R_n$ comparisons. These results are comparable to those reported by other methods deriving the daytime average LE flux as an expression of the EF (e.g. [50–53]). It is also worthwhile to note that visual inspection of the derived maps done here showed significant spatial patterns coherent between the predicted $H/R_n$ fluxes, and both the predicted LE/$R_n$ and the predicted SMC and $F_r$ parameters, which also substantiates the validity of the $H/R_n$ flux maps inverted from the "triangle" method.

## 6 Conclusions

This study has analysed the effectiveness of implementing the "triangle" method for deriving spatially explicit maps of SMC, LE and $H$ fluxes. To our knowledge, it is one of the few studies to implement the "triangle" method to estimate these fluxes at the meso-scale, assessing the accuracy of this technique over different ecosystems in Europe by comparing the model estimations with validatory in-situ observations. Furthermore, it is the only study to analyse the effectiveness of integrating AATSR data, specifically the AATSR Gridded Surface Temperature Level 2 full resolution geophysical product, with the SimSphere land surface process model, to implement the "triangle" inversion methodology for the retrieval of surface fluxes. Results from validation studies such as this provides strong supportive evidence of the potential value of this method in accurately deriving estimates of land surface fluxes, and are important steps to support the operational development of such models. Future work needs to concentrate on evaluating how implementing the "triangle" method with higher resolution satellite products will affect the accurate estimation of LST and $F_r$ from the sensor data, which in turn will affect the accurate estimation of the "triangle"-derived surface fluxes and SMC. Furthermore, evaluating the effect of satellite product pre-processing levels and atmospheric correction

on the estimation of LST or $F_r$ from the satellite could be another avenue of exploration for future work. The more accurate derivation of these parameters could have a significant effect on the "triangle"-derived estimates of fluxes and SMC. Assessing the use of different validation sites that vary in homogeneity, land cover type, and climatic conditions could provide a more cohesive outlook on the operational applicability of the "triangle" inversion modelling technique for practical use in a wide range of disciplines.

**Acknowledgments** This work was produced in the framework of the PROgRESSIon (Prototyping the Retrievals of Energy Fluxes and Soil Moisture Content) project, funded by the European Space Agency (ESA) Support to Science Element (STSE) under contract STSE-TEBM-EOPG-TN-08-0005. Dr. Petropoulos gratefully acknowledges the financial support provided by the Agency. Dr. Petropoulos also gratefully acknowledges the financial support provided by the European Commission under the Marie Curie Career Re-Integration Grant "TRANSFORM-EO" project which also supported a part of the participation of G. Ireland in the completion of this work.

## References

1. Wang Y, Li X, Tang S (2013) Validation of the SEBS-derived sensible heat for FY3A/VIRR and TERRA/MODIS over an alpine grass region using LAS measurements. Int J Appl Earth Obs Geoinf 23:226–233
2. Petropoulos GP (2008) Retrieval of surface energy fluxes and related land surface parameters at CARBOEUROFLUX test sites derived the SimSphere model and ASTER imagery analysis. PhD thesis, King's College London
3. Ezzahar J, Er-Raki S, Marah H, Khabba S, Amenzou N, Chehbouni G (2012) Coupling soil-vegetationatmosphere-transfer model with energy balance model for estimating energy and water vapor fluxes over an olive grove in a semi-arid region. Glob Meteorol 1(1):e1
4. Vereecken H, Huisman JA, Pachepsky Y, Montzka C, van der Kruk J, Bogena H, Weihermüller L, Herbst M, Martinez G, Vanderborght J (2014) On the spatio-temporal dynamics of soil moisture at the field scale. J Hydrol 516:76–96
5. Sánchez-Ruiz S, Piles M, Sánchez N, Martínez-Fernández J, Vall-llossera M, Camps A (2014) Combining SMOS with visible and near/shortwave/thermal infrared satellite data for high resolution soil moisture estimates. J Hydrol 516:273–283
6. Brocca L, Melone F, Moramarco T, Wagner W, Naeimi V, Bartalis Z, Hasenauer S (2010) Improving runoff prediction through the assimilation of the ASCAT soil moisture product. Hydrol Earth Syst Sci Discuss 7(4):4113–4144
7. Koster RD, Mahanama SP, Livneh B, Lettenmaier DP, Reichle RH (2010) Skill in streamflow forecasts derived from large-scale estimates of soil moisture and snow. Nat Geosci 3(9): 613–616
8. Albergel C, Calvet JC, De Rosnay P, Balsamo G, Wagner W, Hasenauer S, Naeimi V (2010) Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in southwestern France. Hydrol Earth Syst Sci 14(11):2177–2191
9. Drusch M (2007) Initializing numerical weather prediction models with satellite-derived surface soil moisture: data assimilation experiments with ECMWF's integrated forecast system and the TMI soil moisture data set. J Geophys Res 112:D03102
10. Bolten JD, Crow WT, Zhan X, Jackson TJ, Reynolds CA (2010) Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. Sel Top Appl Earth Observations Remote Sens, IEEE J 3(1):57–66

11. Odhiambo JJO, Bomke AA (2007) Cover crop effects on spring soil water content and the implications for cover crop management in south coastal British Columbia. Agric Water Manag 88(1):92–98
12. Loew A, Holmes T, de Jeu R (2009) The European heat wave 2003: early indicators from multisensoral microwave remote sensing? J Geophys Res 114:D05103
13. Miralles DG, Holmes TRH, De Jeu RAM, Gash JH, Meesters AGCA, Dolman AJ (2011) Global land-surface evaporation estimated from satellite-based observations. Hydrol Earth Syst Sci 15(2)
14. Gamerith V, Neumann MB, Muschalla D (2013) Applying global sensitivity analysis to the modelling of flow and water quality in sewers. Water Res 47(13):4600–4611
15. Heydari S, Miranda-Moreno LF, Lord D, Fu L (2014) Bayesian methodology to estimate and update safety performance functions under limited data conditions: a sensitivity analysis. Accid Anal Prev 64:41–51
16. Miro S, Hartmann D, Schanz T (2014) Global sensitivity analysis for subsoil parameter estimation in mechanized tunneling. Comput Geotech 56:80–88
17. Xu E, Zhang H (2013) Spatially-explicit sensitivity analysis for land suitability evaluation. Appl Geography 45:1–9
18. Ridler ME, Sandholt I, Butts M, Lerer S, Mougin E, Timouk F, Kergoat L, Madsen H (2012) Calibrating a soil–vegetation–atmosphere transfer model with remote sensing estimates of surface temperature and soil surface moisture in a semi arid environment. J Hydrol 436:1–12
19. Li Y, Zhou J, Kinzelbach W, Cheng G, Li X, Zhao W (2013) Coupling a SVAT heat and water flow model, a stomatal-photosynthesis model and a crop growth model to simulate energy, water and carbon fluxes in an irrigated maize ecosystem. Agric For Meteorol 176:10–24
20. Olioso A, Chauki H, Courault D, Wigneron JP (1999) Estimation of evapotranspiration and photosynthesis by assimilation of remote sensing data into SVAT models. Remote Sens Environ 68(3):341–356
21. Olioso A, Inoue Y, Ortega-Farias S, Demarty J, Wigneron JP, Braud I, Jacob F, Lecharpentier P, Ottle C, Calvet J-C, Brisson N (2005) Future directions for advanced evapotranspiration modeling: assimilation of remote sensing data into crop simulation models and SVAT models. Irrigat Drain Syst 19(3–4):377–412
22. Demarty J, Ottlé C, Braud I, Olioso A, Frangi JP, Bastidas LA, Gupta HV (2004) Using a multiobjective approach to retrieve information on surface properties used in a SVAT model. J Hydrol 287(1):214–236
23. Moran MS, Peters-Lidard CD, Watts JM, McElroy S (2004) Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. Can J Remote Sens 30(5): 805–826
24. Olioso A (1992) Simulation des 6changes d'6nergie et de masse d'un convert v6gandal, dans le but de relier ia transpiration et la photosynthese anx mesures de reflectance et de temp6rature de surface. Doctorate thesis, pp 1–260
25. Brunsell NA, Ham JM, Owensby CE (2008) Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tall-grass prairie environment. Remote Sens Environ 112(6):2977–2987
26. Wang H, Jia G (2013) Regional estimates of evapotranspiration over Northern China using a remote-sensing-based triangle interpolation method. Adv Atmos Sci 30:1479–1490
27. Shu Y, Stisen S, Jensen KH, Sandholt I (2011) Estimation of regional evapotranspiration over the North China Plain using geostationary satellite data. Int J Appl Earth Obs Geoinf 13(2): 192–206
28. Carlson TN (2007) An overview of the—triangle methodl for estimating surface evapotranspiration and soil moisture from satellite imagery. Sensors 7:1612–1629
29. Tang R, Li ZL, Tang B (2010) An application of the $T_s$–VI triangle method with enhanced edges determination for evapotranspiration estimation from MODIS data in arid and semi-arid regions: implementation and validation. Remote Sens Environ 114(3):540–551

30. Mallick K, Bhattacharya BK, Patel NK (2009) Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI. Agric For Meteorol 149(8):1327–1342
31. Petropoulos GP, Carlson TN, Wooster MJ, Islam S (2009) A review of $T_s$/VI remote sensing based methods for the retrieval of land surface fluxes and soil surface moisture content. Adv Phys Geography 33(2):1–27
32. Petropoulos GP (2013) Remote sensing of land surface turbulent fluxes and soil moisture: state of the art. Taylor & Francis Book. ISBN: 978-1-4665-0578-0, p 506
33. European Space Agency (2007) AATSR product handbook. Available at: https://earth.esa.int/handbooks/aatsr/CNTR.htm
34. Petropoulos GP, Carlson TN, Wooster M (2009) An overview of the use of the Simsphere soil vegetation atmospheric transfer (SVAT) model for the study of land atmosphere interactions. Sensors 9:4286–4308
35. Petropoulos GP, Griffiths HM, Tarantola S (2013) A sensitivity analysis of the Simsphere SVAT model in the context of EO-based operational products development. Environ Model Softw 49:166–179
36. Baldocchi DD, Valentini R, Running SR, Oechel W, Dahlman R (1996) Strategies for measuring and modeling carbon dioxide and waterfluxes over terrestrial ecosystems. Glob Change Biol 2:159–168
37. Gamnitzer U, Karstens U, Kromer B, Neubert RE, Meijer HA, Schroeder H, Levin I (2006) Carbon monoxide: a quantitative tracer for fossil fuel $CO_2$? J Geophys Res 111:D22302
38. Levin I, Karstens U (2007) Quantifying fossil fuel $CO_2$ over Europe. In: Dolman AJ, Freibauer A, Valentini R (eds) Observing the continental scale greenhouse gas balance of Europe. Springer, Heidelberg
39. Levin I, Karstens U (2007) Inferring high-resolution fossil fuel $CO_2$ records at continental sites from combined $14CO_2$ and CO observations. Tellus 59B:245–250
40. Gillies RR, Carlson TN (1995) Thermal remote sensing of surface soil water content with partial vegetation cover for incorporation into climate models. J. Appl. Meteorol. 34(4):745–756
41. Choudhury BJ, Ahmed NU, Idso SB, Reginato RJ, Daughtry CS (1994) Relations between evaporation coefficients and vegetation indices studied by model simulations. Remote sensing of environment 50(1):1–17
42. Brunsell NA, Gillies RR (2003) Scale issues in land-atmosphere interactions: implications for remote sensing of the surface energy balance. Agric For Meteorol 117:203–221
43. Chauhan NS, Miller S, Ardanuy P (2003) Spaceborne soil moisture estimation at high resolution: a microwave–optical/IR synergistic approach. Int J Remote Sens 22:4599–4622
44. Gillies RR, Carlson TN, Cui J, Kustas WP, Humes KS (1997) Verification of the "triangle" method for obtaining surface soil water content and energy fluxes from remote measurements of the normalized difference vegetation index NDVI and surface radiant temperature. Int J Remote Sens 18:3145–3166
45. Jiang L (2000) Estimation of land surface evaporation map over large areas using remote sensing data. PhD Thesis, University of Cincinnati
46. Validation Report Evapotranspiration Products LSA-16 (MET), LSA (DMET) (2010) Doc Number, SAF/LAND/RMI/VR/06. Available at: http://landsaf.meteo.pt/GetDocument.do;jsessionid=9588908831FFC55ECD304FB798539712?id=309. Accessed on 1 Sept 2011
47. Sellers P, Hall FG, Asrar J, Strubel DE, Murthy RE (1992) An overview of the First Intenational Satellite Land Surface Climatology project (ISLSCP) Field Experiment (FIFE). J. Geophys. Res. 97:18345–18371
48. Kustas WP, Goodrich DC (1994) Preface. MONSOON'90 multidisciplinary experiment. Water Resour Res 30(5):1211–1225
49. Brunsell NA (2003) An examination of scale issues involved with remotely sensed data. PhD Thesis, University of Utah
50. Jiang L, Islam S (2003) An intercomparison of regional heat flux estimation using remote sensing data. Int J Remote Sens 24:2221–2236

51. Nishida K, Nemani R, Glassy J, Running S (2003) Development of an evapotranspiration index from aqua/MODIS for monitoring surface moisture status. IEEE Trans Geosci Remote Sens 41(2):493–501
52. Verstraeten WW, Veroustraete F, Feyen J (2005) Estimating evapotranspiration of European forests from NOAA-imagery at satellite overpass time: towards an operational processing chain for integrated optical and thermal sensor data products. Remote Sens Environ 96:256–276
53. Wang K, Li Z, Cribb M (2006) Estimation of evaporative fraction form a combination of day and night land surface temperatures and NDVI: a new method to determine the Priestley-Taylor parameter. Remote Sens Environ 102:293–305

# Crustal Modelling and Moho Estimation with GOCE Gravity Data

**Daniele Sampietro**

**Abstract** GOCE observations are an extremely innovative and useful product for the study of the Earth crust at regional and global scales: on the one hand, they can be considered as a constraint to verify crustal models, on the other hand combining GOCE gravity observations with seismic data and taking into account additional information it is possible to retrieve important information on the Earth crust structure. After one year only of GOCE observations, thanks to the GOCE Exploitation for Moho Modelling and Applications (GEMMA) project, it has been possible to globally estimate the depth of the boundary between the Earth's crust and mantle, usually called Moho, with unprecedented resolution. The knowledge of the Moho is a key topic in Solid Earth sciences: the new GOCE Moho has been used, for instance, as background information to improve our ability to understand and model earthquakes or for the study of the Earth's heat flux and heat production which in turn constitutes a basic knowledge to understand the plate tectonics and the thermal evolution of our planet.

## 1 Introduction

The boundary between the Earth crust and mantle is usually modelled as a discontinuity surface, called Mohorovičić discontinuity or Moho from the name of the Croatian seismologist who discovered it in 1909 [22]. Even if it is well known that the Moho can locally present very complex features, like duplications, fragmentations, subductions, etc., a sharp separation defined by a single surface well approximates its global behaviour. This surface is generally defined from the seismic point of view as the discontinuity separating rocks having $V_P$ (P-wave) velocities of 6 km/s from those having velocities of about 8 km/s [20]. The com-

D. Sampietro (✉)
GReD s.r.l., c/o Politecnico di Milano - Polo Territoriale di Como, via Valleggio 11, 22100 Como, Italy
e-mail: Daniele.sampietro@g-red.eu



D. Fernández-Prieto and R. Sabia (eds.), *Remote Sensing Advances for Earth System Science*, Springer Earth System Sciences,
DOI 10.1007/978-3-319-16952-1_8

positional differences velocity is usually reflected in a corresponding increase of density, which allows the Moho estimation also from gravity data. Concerning seismic methods different attempts to estimate regional and global crustal models have been made in the past years. At regional scale the problem is well revisited, amongst many others, in a recent work by Carbonell et al. [6], while at global scale the first attempt probably dates back to 1982 with the work of Soller et al. [42] where a compilation of crustal thickness and seismic velocity values, from published seismic refraction and surface wave data, have been used to contour the crustal thickness on a global Van der Grinten projection. After that a noteworthy result was the CRUST5.1 model [23] based on seismic refraction and, later on, its updated versions, i.e. CRUST2.0 [3] and CRUST1.0 [16]. These models describe the crust structure by giving information on thickness and density of a number of global components (e.g. ice, oceans, soft and hard sediments, upper, middle and lower crust) on a grid with a resolution of 5°, 2° and 1°, respectively. Moreover, for each cell of the grid, the crustal type (e.g. oceanic, continental plateaus, rift, orogenetic regions, etc.) and the upper mantle density are also given. The models of the CRUST series are based on seismic refraction data published from 1950, on a detailed compilation of ice and sediment thickness, and on statistical predictions for regions such as most of Africa, South America, Greenland and oceans where no or very few seismic measurements were available. The models are delivered without any error map, the only information available on their accuracy is an approximated value for the observation error of seismic profiles considered as 10 % of the Moho depth itself [7].

The main disadvantage related to this kind of model is their inconsistency with the gravitational field: in fact, comparing the second radial derivative of the gravitational potential of the CRUST2.0 and CRUST1.0 with the actual one observed by GOCE, differences with a standard deviation (std) of about 1000 and 1300 mE are found [33]. Moreover, since they are mainly based on controlled-source seismic surveys, many regions of the world are not properly covered by data. In any case, even where seismic data are available, they have been in general acquired during different campaigns, at different times, with different instruments and elaborated with different processing strategies: the resulting crustal models are therefore extremely inhomogeneous in precision and accuracy. In order to overcome some of these limitations, expensive controlled-source experiments have been complemented (or substituted) by passive seismic studies that use natural seismic sources: in 2007 Meier et al. [19] directly used surface wave data to globally infer the crustal thickness; the resulting Moho, which is delivered with the corresponding error estimate, is limited by the lateral resolution of the input phase and group velocity maps ranging between 500 and 1000 km. A review on the use of passive seismic methods for Moho estimation can be found in the work of Lebedev et al. [18].

As for gravimetric methods different models have been developed, e.g. starting from the one of Oldenburg [25] to Shin et al. [38] and Braitenberg and Ebbing [4] at regional scale or the work of Sünkel [44], Moritz [24] and Sjöberg [40] at global scale. Here two main problems have to be faced: the former is related to the intrinsic indetermination of the inverse gravimetric problem, while the latter is related to the

maximum resolution of Moho models derived from gravity data. As for the first one, it is well known that, in order to guarantee the uniqueness of the solution of the inverse gravimetric problem, in general some very restrictive hypotheses should be introduced [32]. To overcome this limitation the solution should be regularized by introducing some a priori constraints or, as an alternative, one can model and remove from the gravity data the effect of every known subsurface density anomaly, apart from the one between crust and mantle, and then invert the residual field (eventually considering also seismic information) to obtain a new model consistent as much as possible with the available observations. As for the second problem, namely the maximum resolution achievable from gravity data, global solutions have been limited, up to now, to a maximum resolution of about 150 km achievable with data from the Gravity Recovery and Climate Experiment (GRACE) mission [38].

Nowadays the European Space Agency (ESA) Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) mission [8], exploiting for the first time the concept of satellite gradiometry, has observed the Earth gravitational field from space with very high accuracy and spatial resolution allowing for a global estimate of the Moho with unprecedented resolution. Numerical experiments [30] have shown the possibility to estimate the Moho up to a resolution of 90 km with only 1 year of GOCE observations; such a resolution increases to 70/80 km with the fourth release of GOCE model, as can be seen in Fig. 1 where the signal degree variances from CRUST2.0 model are compared with the Moho error degree medians estimated from the second and fourth release of GOCE time-wise models [26]. Of course even better results are expected with the already released fifth generation of GOCE models.

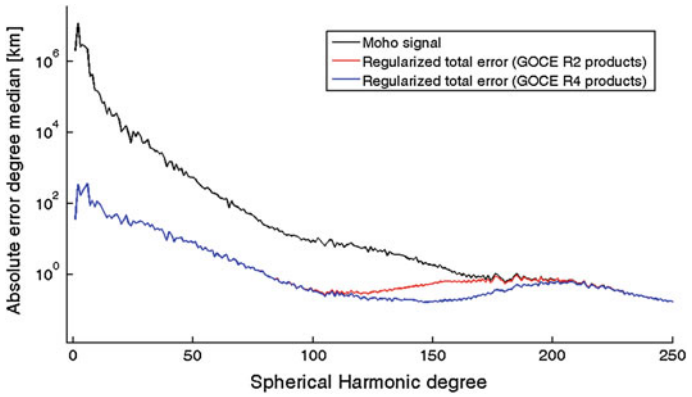


**Fig. 1** Moho signal degree variances from CRUST2.0 model (*black line*) compared with estimated errors from GO_CONS_GCF_2_TIM_R2 (*red line*) and GO_CONS_GCF_2_TIM_R4 (*blue line*). Errors are obtained by combining linearization error with observation error propagated in terms of Moho and regularized at high degrees

In this scenario the GOCE Exploitation for Moho Modelling and Applications (GEMMA) project, founded by ESA through the Support To Science Element (STSE) program, aimed at improving the actual knowledge on the Earth crust by fully exploiting the potentiality of GOCE observations both at regional and global scales. This required to study and implement new methodologies to invert the gravitational field (and in particular, the second radial derivative of the anomalous gravitational potential), to study the propagation of GOCE observation errors in terms of Moho depth and to study and implement ad hoc regional and global terrain correction procedures.

## 2 Inverse Gravimetric Problem

The inverse gravimetric problem consists, in general, in the determination of the internal density distribution $\rho$ of a body $B$ from a functional of its exterior gravitational potential. This problem is based on Newton's law of gravitation: according to this law the gravitational potential $T$ and its derivatives can be seen as functions of the mass density distribution: $T = f(\rho)$. This means that the relation $T = f(\rho)$ has to be inverted in order to recover the unknown mass density distribution. According to Hadamard's criteria [12] the inverse gravimetric problem is known to be an ill-posed problem. In fact, neither the existence nor the uniqueness nor the stability of the solution is in principle guaranteed [35]. The non-uniqueness can be treated, for instance, by considering hypotheses on the shape of the density discontinuity. In principle, one can think that it is better to use a rough geophysical hypothesis and to find a unique solution, rather than accepting a solution that can be very far from reality because it corresponds to a purely mathematical criterion [36]. It has been proved that for a two-layer model and under simple assumptions (known topography and constant density or known depth of compensation) the inverse gravitational problem admits a unique solution [33]. The uniqueness of the solution is also proved when a linear vertical gradient density distribution is considered as unknown. Once the uniqueness is guaranteed, we are entitled to apply to the corresponding inverse problem a regularization method and we know from literature [37] that in this way we can approximate the true solution, dominating the inherent instabilities. A thorny issue to be faced, in order to solve the inverse gravimetric problem, is therefore how to separate the various signals, coming from different geological structures, mixed up into the observed gravimetric data. As a matter of fact this can be achieved only with the help of additional geological information by modelling crustal heterogeneities, as well as unwrapping the contributions of large deep features from those closer to the surface.

Two main methodologies to estimate the Moho depth from GOCE data at different scales have been developed in the GEMMA project and will be discussed in the following sections. The general procedure is however similar in both cases: one

has first to reduce the gravitational observations using the most accurate a priori geological information (possibly gravity independent), thus simplifying the initial problem in a two-layer problem. After that, the residual field can be inverted, by means of linearized equations relating the observed functional with the unknown Moho depth. Possible instability, due to observation errors, is dominated by means of an opportune regularization operator.

## 3 Local Solution

The local solution developed in the framework of the GEMMA project is based on the inversion of regional grids of the gravitational potential and of its second radial derivative, computed at mean GOCE altitude. These two quantities represent the most important observations coming from the GOCE satellite: the gravitational potential can be obtained by means of the energy conservation approach applied to satellite-to-satellite track observations (see, e.g. Visser et al. [47] or Jekeli [14]), while the grid of second radial derivative can be obtained by merging all the observed gravity gradients with the so-called space-wise approach [29].

### *3.1 Local Inversion Algorithm*

We recall here only the main concepts, leaving the interested reader to consult Sampietro [31], Reguzzoni and Sampietro [28], and Sampietro et al. [34] for details.

Considering a spherical coordinate system $(\varphi, \lambda, r)$, with $\varphi$ spherical latitude, $\lambda$ the longitude and $r$ the radius, the gravitational potential $T$ observed at point $P$ (e.g. at mean GOCE altitude) due to the masses between the Moho and the topography can be computed as:

$$T = \int_{\varphi_{\min}}^{\varphi_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{H_Q}^{M_Q} \frac{\rho G r^2 \cos\varphi \, \mathrm{d}r\mathrm{d}\lambda\mathrm{d}\varphi}{\sqrt{r_P^2 + r_Q^2 + 2 r_P r_Q \cos\psi_{PQ}}} \tag{1}$$

where $G$ is the gravitational constant equal to $6.67384 \times 10^{-11}$ $\mathrm{m^3 kg^{-1} s^{-2}}$, $M_Q$ and $H_Q$ are the values of the radial coordinate corresponding, respectively, to the actual unknown Moho depth, and to the actual topography at point $Q$, and $\psi_{PQ}$ is the angular distance between points $P$ and $Q$. The notation used is systematized in Fig. 2. Modelling the Moho as $M_Q = \bar{R} + \delta D$, and linearizing Eq. 1 around $\bar{R}$ in the $r$ direction we get:

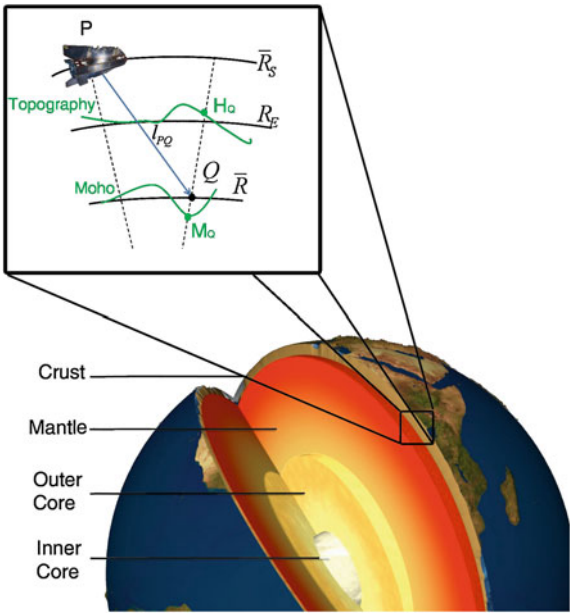


**Fig. 2** Schematic representation of the problem. The observation point $P$ is here represented by the GOCE satellite, while the integration one, $Q$, is running in the Earth crust, i.e. in the masses between the actual topography and the Moho. $\bar{R}_S$ represents the mean GOCE satellite altitude, while $R_E$ is the radius of the average topography and $\bar{R}$ is the radius corresponding to the mean Moho depth. Finally $l_{PQ}$ is the geometrical distance between points $P$ and $Q$, i.e. the denominator of Eq. 1

$$T = \int_{\varphi_{\min}}^{\varphi_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} \int_{H_Q}^{\bar{R}} \frac{\rho G r^2 \cos\varphi \mathrm{d}r\mathrm{d}\lambda\mathrm{d}\varphi}{\sqrt{r_P^2 + r_Q^2 + 2r_P r_Q \cos\psi_{PQ}}} \mathrm{d}r + \int_{\varphi_{\min}}^{\varphi_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\Delta\rho G r^2 \cos\varphi \mathrm{d}\lambda\mathrm{d}\varphi}{\sqrt{r_P^2 + \bar{R}^2 + 2r_P \bar{R} \cos\psi_{PQ}}} \delta D \tag{2}$$

where $\Delta\rho$ is the difference between the crust and the mantle mass density at point $Q$, and $\delta D$ is the difference between the actual Moho depth in $Q$ and $\bar{R}$. Note that the first integral in Eq. 2 does not depend on the actual Moho depth, therefore it can be numerically computed and its effect can be removed from $T$ obtaining the residual field $\delta T$:

$$\delta T = \int_{\varphi_{\min}}^{\varphi_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\Delta\rho G r^2 \cos\varphi \, \mathrm{d}\lambda\mathrm{d}\varphi}{\sqrt{r_P^2 + \bar{R}^2 + 2r_P \bar{R} \cos\psi_{PQ}}} \delta D. \tag{3}$$

Applying the same reasoning also to the second radial derivative of the gravitational potential ($T_{rr}$) we get a similar expression:

$$\delta T_{rr} = \int_{\varphi_{\min}}^{\varphi_{\max}} \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\left(3r_P^2 \cos^2 \psi_{PQ} + 4\bar{R} r_P \cos \psi_{PQ} + 2\bar{R}^2 - r_P^2\right) \mathrm{d}\lambda \mathrm{d}\varphi}{\left(r_P^2 + \bar{R}^2 + 2 r_P \bar{R} \cos \psi_{PQ}\right)^{5/2}} \bar{R} G \Delta\rho \delta D. \quad (4)$$

To estimate the Moho we have to solve the system obtained by inverting Eqs. 3 and 4 degraded with the corresponding observation noise. As for the noise in the two equations it should be stressed that none of the two quantities is a direct observation of the GOCE mission: as reminded above the potential $T$ is derived from GPS tracking data, while the second radial derivatives $T_{rr}$ are obtained by processing the gradiometer observations. In any case the resulting potential is known to have an almost white error, while the second radial derivatives have a time-correlated error with spectral characteristics almost identical to the original observations [21]. Note that $T$ contains the very low frequencies of the gravitational signal while $T_{rr}$ allows to infer the Moho up to a resolution better than 100 km [34].

The optimal solution of the problem can be obtained by means of a least squares collocation algorithm, considering $\delta D$ as a random signal:

$$\delta \hat{D} = C_{y,\delta D}^T \, C_{y,y}^{-1} \, y \quad (5)$$

where

$$C_{y,\delta D} = \begin{bmatrix} C_{T_{rr},\delta D} \\ C_{T,\delta D} \end{bmatrix}, \quad C_{y,y} = \begin{bmatrix} \overline{C}_{T_{rr},T_{rr}} & C_{T_{rr},T} \\ C_{T,T_{rr}} & \overline{C}_{T,T} \end{bmatrix}, \quad y = \begin{bmatrix} \delta T_{rr} \\ \delta T \end{bmatrix}.$$

Here $y$ are the observations, $C_{a,b}$ is the covariance matrix between $a$ and $b$ and $\overline{C}_{a,b}$ the covariance matrix between $a$ and $b$ plus the error covariance matrix $C_{v_a,v_b}$.

Since we will suppose the noise of the two observations to be uncorrelated, the error covariances are present only on the diagonal blocks of the matrix $C_{y,y}$. It can be noticed that all the needed covariance matrices, with the exception of the error covariance matrices, can be computed by propagating $C_{\delta D,\delta D}$ through Eqs. 3 and 4. The solution of Eq. 5 is quite heavy in terms of computational time, because it requires to invert a large matrix, namely $C_{y,y}$, which contains the cross-covariances between satellite observations. However, it should be observed that if we suppose to have gridded observations, and to compute the Moho depth on the same grid and we accept a further approximation in the observations equations, $\delta \hat{D}$ can be efficiently computed in terms of multiple input single output (MISO) Wiener filter in the frequency domain [39]. In fact, by applying the following mapping:

$$\begin{cases} \lambda \leftrightarrow \xi = \bar{R}\lambda \cos\bar{\varphi} \\ \varphi \leftrightarrow \eta = \bar{R}\varphi \end{cases} \tag{6}$$

where $\overline{\varphi}$ is the mean latitude of the considered region, Eqs. 3 and 4 can be rewritten as:

$$\delta T = \int_{\eta_{\min}}^{\eta_{\max}} \int_{\xi_{\min}}^{\xi_{\max}} \frac{\cos\varphi \Delta\rho G \mathrm{d}\xi \mathrm{d}\eta}{\cos\bar{\varphi}\left[(r_P - \bar{R})^2 + \frac{r_P}{\bar{R}}\ell_{PQ}^2\right]^{\frac{1}{2}}} \delta D \tag{7}$$

and

$$\delta T_{rr} = \int_{\eta_{\min}}^{\eta_{\max}} \int_{\xi_{\min}}^{\xi_{\max}} \frac{\left(3\ell_{PQ}^4 + 4\bar{R}\ell_{PQ}^2(2r_P - 3\bar{R}) + 8\bar{R}^2(r_P - \bar{R})^2\right)\cos\varphi G \Delta\rho \mathrm{d}\xi \mathrm{d}\eta}{4\bar{R}^2 \cos\bar{\varphi}\left[(r_P - \bar{R})^2 + \frac{r_P}{\bar{R}}\ell_{PQ}^2\right]^{\frac{5}{2}}} \delta D \tag{8}$$

where $\ell_{PQ} = \sqrt{(\xi_P - \xi_Q)^2 + (\eta_P - \eta_Q)^2}$. Supposing finally to deal with observations at a constant altitude (e.g. mean GOCE altitude), namely $r_P = \text{const.}$, Eqs. 7 and 8 take the characteristics of convolution integrals, i.e. they can be seen as the integral of the product of two functions, one depending only on $\xi_Q$ and $\eta_Q$, and the other depending on the differences $\xi_P - \xi_Q$ and $\eta_P - \eta_Q$:

$$\delta T = \int\int k_T(\xi_P - xi_Q, \eta_P - \eta_Q)\Delta\rho\delta D \cos\varphi \mathrm{d}\xi \mathrm{d}\eta \tag{9}$$

with

$$k_T(\xi_P - xi_Q, \eta_P - \eta_Q) = \frac{G}{\cos\bar{\varphi}\left[(r_P - \bar{R})^2 + \frac{r_P}{\bar{R}}\ell_{PQ}^2\right]^{\frac{1}{2}}} \tag{10}$$

and

$$\delta T_{rr} = \int\int k_{T_{rr}}(\xi_P - xi_Q, \eta_P - \eta_Q)\Delta\rho\delta D \cos\varphi \mathrm{d}\xi \mathrm{d}\eta \tag{11}$$

with:

$$k_{T_{rr}}(\xi_P - xi_Q, \eta_P - \eta_Q) = \frac{3\ell_{PQ}^4 + 4\bar{R}\ell_{PQ}^2(2r_P - 3\bar{R}) + 8\bar{R}^2(r_P - \bar{R})^2}{4\bar{R}^2 \cos\bar{\varphi}\left[(r_P - \bar{R})^2 + \frac{r_P}{\bar{R}}\ell_{PQ}^2\right]^{\frac{5}{2}}}. \tag{12}$$

Applying the convolution theorem the collocation solution can be rewritten in terms of Wiener filter in the frequency domain as:

$$\delta\hat{D} = \mathscr{F}^{-1}\left(\frac{S_{\delta D}\mathscr{F}(k_T)S_{v_{rr}}}{S_{v_{rr}}S_v + \mathscr{F}\left(k_{T_{rr}}^2\right)S_{\delta D}S_v + \mathscr{F}(k_T^2)S_{\delta D}S_{v_{rr}}}\mathscr{F}(\delta T) + \frac{S_{\delta D}\mathscr{F}(k_{T_{rr}})S_v}{S_{v_{rr}}S_v + \mathscr{F}\left(k_{T_{rr}}^2\right)S_{\delta D}S_v + \mathscr{F}(k_T^2)S_{\delta D}S_{v_{rr}}}\mathscr{F}(\delta T_{rr})\right) \tag{13}$$

where $\mathscr{F}(\cdot)$ and $\mathscr{F}^{-1}(\cdot)$ are the Fourier and the inverse Fourier transform operators, respectively.

The approximation introduced by this "almost spherical" mapping has been investigated by computing for regions with increasing size the difference between the collocation solution (that considers the right geometry of the problem) and the "almost spherical" one, that makes use of the Fourier transform [31]. Results show that for relatively small regions (area smaller than 10°) the mean of the difference between the two solutions is practically the same (and smaller than 0.5 km), while, as expected, for bigger regions the "almost spherical" approximation introduces an error in the low frequencies, that however is always smaller than 1.5 km.

## 3.2 Numerical Results

During the GEMMA project a Moho model of the Himalayas and the Tibetan Plateau has been computed [34]. In particular, two grids directly estimated at mean GOCE satellite altitude from eight months of observations (one of the gravitational potential and one of its second radial derivatives) have been used as input. The power spectral densities required to solve Eq. 13, namely $S_{\delta D}$, have been computed by applying the Fourier transform to a theoretical covariance function obtained by modelling as the product of two Gaussian functions in $\xi, \eta$, an empirical anisotropic covariance computed from CRUST2.0 model. As for the error spectra of GOCE gridded observations, they are estimated from Monte Carlo samples [21].

Concerning the data reduction, firstly the effect of sediments in terms of GOCE observables at satellite altitude has been computed and removed from the GOCE grids. The sediment model used is taken from CRUST2.0 and refined with a local dataset [5]. Lateral variations of density contrast (again from CRUST2.0) have been taken into account in the inversion keeping the uniqueness of the solution. The resulting Moho model is shown Fig. 3, its error covariance function, predicted by propagating the Monte Carlo error covariance matrix of GOCE grids to the results, shows a variance of about 2.5 km$^2$ and a correlation length of about 2.5°.

The estimated GOCE-only Moho model has been compared with two available Moho profiles [46, 48] derived from seismic observations. Results seem to be satisfactory since the GOCE Moho well interpolates the seismic one: showing

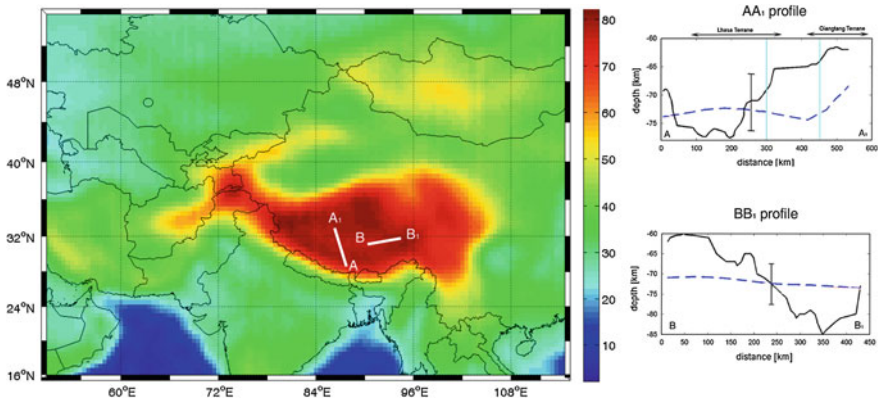


**Fig. 3** Moho estimated inverting GOCE observations. *Black lines* are seismic profiles, *dashed blue lines* are GOCE solutions. Unit (km)

differences of the order of 4.5 km (std). Comparing this value with the predicted error covariance we can conclude that the intrinsic error of GOCE observations should allow, in principle, to recover the Moho with an accuracy of 1.6 km (std) or better; however, the proposed solution propagates the errors in the sediment and in the density crustal models to the estimated Moho depth, thus degrading its accuracy. Moreover in the $AA_1$ profiles, it is clearly visible a region in which the behaviours of gravity and seismic Moho are different. This difference is found to be in correspondence with the collision between Indian and Eurasian plates where doubling as well as fragmentation of the Moho is known to occur. Therefore, the discrepancy can be explained as a consequence of the fact that the two-layer hypothesis is not acceptable in those critical regions. A full model should comprise the subduction of the Indian lithosphere and density variations at lower crustal level across the strike of the Himalayan belt. On the other hand, this numerical result shows that, in principle, discrepancies between GOCE Moho model and seismic profiles can be used to detect the presence of such kind of density anomalies by simply mapping the largest residuals.

## 4 Global Solution

The global solution developed in the framework of the GEMMA project is based on the inversion of global gravity field models (GGM). These models, delivered in terms of coefficients of geopotential spherical harmonics series, are mainly derived from satellite measurements and are becoming nowdays more and more detailed and accurate. The GEMMA solution is tailored on the second release of GOCE space-wise solution, however it can be easily extended to newer releases of GOCE models or to other GGMs.

## 4.1 Global Inversion Algorithm

The global inversion algorithm is based on a linearized expression [41, 43], relating the coefficients of the gravitational potential due to the Moho discontinuity, $\delta T_{nm}$, to the coefficients of a functional $\delta\omega$ defined as the product between the Moho undulation $\delta D$ with respect to a reference spherical Moho of radius $\bar{R}$ and its density contrast $\Delta\rho$. We revise here the main step of the derivation of the solution, while for a more complete treatment of this subject the interested reader should refer to Reguzzoni et al. [30] and Reguzzoni and Sampietro [27].

The anomalous potential generated by the global Moho undulation, see Eq. 1, can be expressed in terms of series of fully normalized spherical harmonics [13] $\bar{Y}_{nm}$ of degree $n$ and order $m$ as:

$$\delta T = -\frac{GR_E^3}{r_P}\sum_n \left(\frac{R_E}{r_P}\right)^n \left(1-\frac{\bar{R}}{R_E}\right)^{n+3}\frac{1}{2n+1} \times \sum_{m=-n}^{n}\bar{Y}_{nm}(P)\int_0^{\pi}\int_0^{2\pi}\frac{\Delta\rho}{n+3}\left[1-\left(1+\frac{\delta D}{R_E-\bar{R}}\right)^{n+3}\right]\bar{Y}_{nm}(Q)\cos\varphi \mathrm{d}\lambda\mathrm{d}\varphi. \tag{14}$$

Linearizing with respect to $\frac{\delta D}{R_E-\bar{R}}$, and approximating $(1\pm x)^n$ with $1\pm nx$ Eq. 14 simplifies in:

$$\delta T = -\frac{3GM}{r_P\rho_E(R_E-\bar{R})}\sum_n \left(\frac{R_E}{r_P}\right)^n \left(1-\frac{\bar{R}}{R_E}\right)^{n+3}\frac{1}{2n+1} \times \sum_{m=-n}^{n}\bar{Y}_{nm}(P)\int_0^{\pi}\int_0^{2\pi}\delta\omega\bar{Y}_{nm}(Q)\cos\varphi \mathrm{d}\lambda\mathrm{d}\varphi. \tag{15}$$

where $\rho_E$ is the average density of the Earth, i.e. 5.496 kg/m$^3$. Note that the double integral in Eq. 15 gives the spherical harmonic coefficients $\delta\omega_{nm}$ of the function $\delta\omega$. Comparing now this quantity with the spherical harmonic expression of the anomalous potential one can finally get the sought relation between $\delta\omega_{nm}$ and $\delta T_{nm}$:

$$\delta\omega_{nm} = \frac{\rho_E(R_E-\bar{R})(2n+1)}{3}\left(1-\frac{R_E}{\bar{R}}\right)^{n+3}\delta T_{nm}. \tag{16}$$

A remark is in order: Eq. 16 allows to compute $\delta\omega_{nm}$, and therefore $\delta\omega$, given the spherical harmonic coefficients of the gravitational potential due to the masses between a constant reference Moho $\bar{R}$ and the actual Moho. In order to compute this quantity the observed gravitational potential should be properly reduced by removing the gravitational effect of all the crust components apart from the one

related to the Moho (e.g. sediments, oceans, topography, etc.), possibly considering gravity-independent models. This is a crucial problem since an erroneous disentangling of the gravitational field will be reflected in an error in the final solution. Moreover, since the results of the inversion should be $\delta D$ and $\Delta\rho$ (and not $\delta\omega$) a further step is required. However, one has to note that while reducing the data a density model for crust and upper mantle is required, thus implicitly defining $\Delta\rho$. Therefore once $\delta\omega$ is computed, the final Moho depth can be found simply as:

$$\delta D = \frac{\delta\omega}{\Delta\rho}. \tag{17}$$

## 4.2 Numerical Results

Two global solutions have been computed during the GEMMA project using the principle described above: GEMMA2012C [30] and GEMMA1.0 [27]. Note that, even if both these solutions are based on Eq. 16, they are completely different in terms of data reduction as well as solution strategy. The GEMMA2012C has been computed by reducing GOCE anomalous gravitational field by the effects of ice sheets, topography, bathymetry (from Etopo1 model [1]) and sediments [3]. The crystalline crust density $\rho$ has been modelled by assembling two layers of CRUST2.0, namely the upper crust layer below continents and the middle crust layer below oceans. After this reduction, the spherical harmonic coefficients of the obtained anomalous potential are computed and then transformed into $\delta\omega_{nm}$ coefficients by Eq. 16. From these estimated coefficients a harmonic synthesis of $\delta\omega$ up to degree 210 is performed on a global grid with resolution $0.5^\circ \times 0.5^\circ$. Assuming a density contrast $\Delta\rho = \rho_m - \rho$, where $\rho_m$ is the upper mantle density (again taken from the CRUST2.0 model), the Moho depth is finally obtained using Eq. 17. This gravimetric-only model (see Fig. 4) shows several inconsistencies: firstly it shows an unrealistic oceanic crust, with an average thickness of about 15 km, secondly many features in the continental crust seem to be smoothed or not even present. In order to mitigate these problems, that are basically due to the rough approximation in the description of the crystalline crust density, the GOCE-only solution has been directly combined with the CRUST2.0 global seismic model. In detail the considered observations are $\delta\omega$ obtained from GOCE data and $\Delta D_{CR2}$ and $\delta\rho_{CR2}$ from the CRUST2.0 model. The solution is found by linearizing the GOCE observation equation around the estimated GOCE-only Moho $\delta D_{\mathrm{GOCE}}$ and its a priori density contrast $\Delta\rho_{\mathrm{GOCE}}$:

$$\begin{cases} \delta\omega = \Delta\rho_{\mathrm{GOCE}}\delta D + \delta D_{\mathrm{GOCE}}\Delta\rho + v_\omega \\ \delta D_{CR2} = \delta D + v_{\delta D_{CR2}} \\ \Delta\rho_{CR2} = \Delta\rho + v_{\Delta\rho_{CR2}} \end{cases} \tag{18}$$

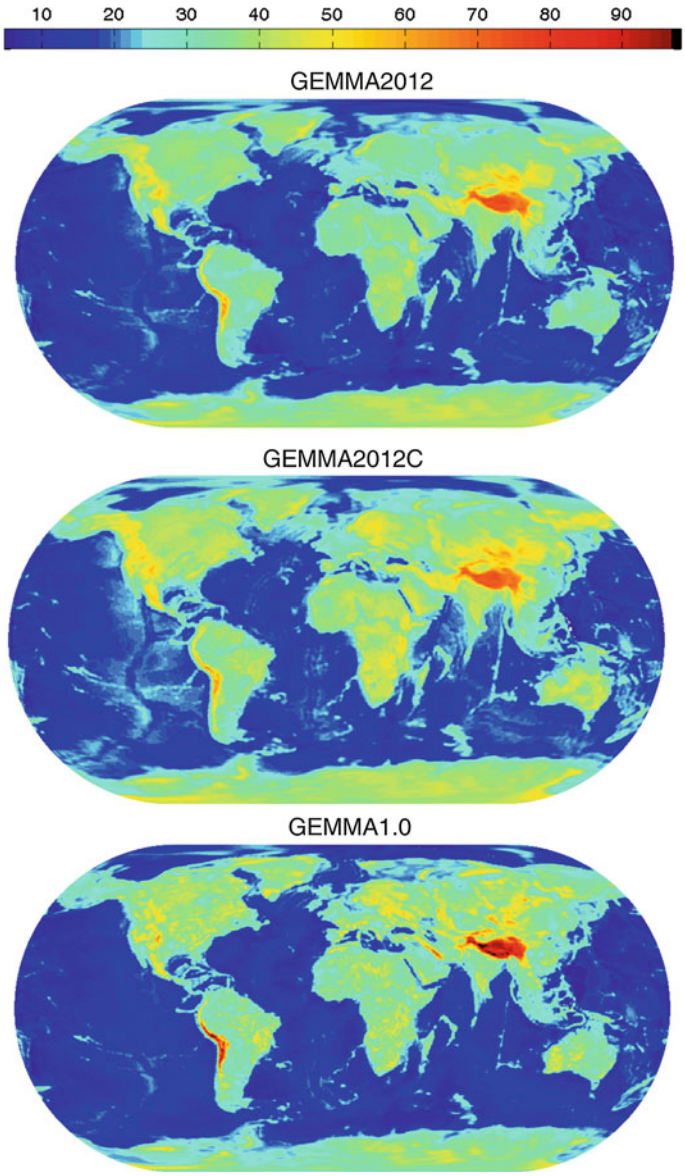


**Fig. 4** GEMMA2012, GEMMA2012C and GEMMA1.0 Moho models. Unit (km)

where $v_{\omega}$, $v_{\delta D_{CR2}}$ and $v_{\Delta\rho_{CR2}}$ are the observation errors, considered independent from each other and from all other grid knot errors. In order to account for the correlation of the Moho undulation in the solution of Eq. 18 the following pseudo-observation equation is added to the system:

$$0 = \delta D + \eta_{\delta D} \tag{19}$$

where $\eta_{\delta D}$ is a random variable modelling the Moho undulation stochastic structure. Basically the above equation states that the searched solution $\delta\hat{D}$ should have the same spatial correlation of $\eta_{\delta D}$. The covariance function of $\eta_{\delta D}$ is modelled by a Gaussian exponential function and is estimated by fitting the empirical covariance of the CRUST2.0 Moho depth [15].

Note that the combination in Eq. 18 allows us to obtain a more realistic Moho (see Fig. 4) to the detriment of consistency with the model itself: in order to keep unchanged the fitting with gravitational data, to a change in $\delta D$ should correspond a change in $\Delta\rho$ in such a way that $\delta\omega = \delta D\Delta\rho$ remains unchanged. At the same time $\Delta\rho$ should be in principle defined in the data reduction step as the difference between $\rho_m$ and $\rho$. Moreover, in order to apply Eq. 17, $\rho$ should be constant in the radial direction too. This is a very limiting hypothesis since it is well known [7] that, due to the increasing pressure and temperature, the crustal density tends to increase with depth.

In order to overcome these limitations an iterative procedure has been developed and applied to GOCE data to product the GEMMA1.0 model. Details on the procedure can be found in Reguzzzoni and Sampietro [27], while in the following, only the main results are reported. Basically the new procedure allows to take into account the dependency of the crust density on the radial direction in such a way that the density model used in the data reduction is coherent with the density contrast in Eq. 17. Moreover, the new algorithm allows us to adjust the a priori density model of the crystalline crust for scale factors, thus overcoming another important problem of gravity derived Moho models. Another important difference between GEMMA2012C and GEMMA1.0 is in the data reduction: in fact in GEMMA1.0, the Earth surface has been divided into a set of 139 geological homogeneous provinces (according to United States Geological Survey (USGS) Geologic Province and Thermo-Tectonic Age Maps [11]) classified into eight main geological types (i.e. shield, basin, platform, extended crust, orogenetic crust, oceanic crust, igneous provinces and mid-oceanic ridges). A specific radial variation of the crustal density, derived from Christensen and Money [7], is thus assumed for each type of province. Finally the last difference between GEMMA2012C and GEMMA1.0 is in the use of seismic information (derived from CRUST2.0 model). In fact in GEMMA2012C a direct combination, according to Eq. 18, between a GOCE only model and the CRUST2.0 one was performed. On the contrary in order to compute GEMMA1.0 seismic derived information is integrated into the inversion procedure. Basically the CRUST2.0 model is used, inside the iterative procedure, to set a scale coefficient for each geological province, that multiplies the crystalline crust density in such a way that the mean Moho depth of each province retrieved from the inversion is “close” to the mean depth obtained by seismic observations. Note that the developed algorithm weakly combines gravity and seismic data since only few parameters (one for each geological province) are estimated from seismic data. An empirical error map (with a global mean

of 3.5 km and a std of 1.6 km) has also been computed for GEMMA1.0 by propagating GOCE observation error and taking into account errors in the upper mantle density variation and in the definition of the shape and densities of the geological provinces.

The GEMMA1.0 model is shown in Fig. 4. Comparing the result with the GEMMA2012C model it can be seen that the mean oceanic Moho depth is now corrected and that in general it gives a more refined crustal structure. Starting from the South America it is interesting to see how the GEMMA model is able to detect all the main known features: the Andean range is well defined and shows remarkable details as the thinning of the crust in the northern Puna and between Ecuador and Peru. Another interesting feature is the presence of a thin crust between the Andean range and the cratonic areas: the presence of this feature, not visible in CRUST2.0, seems to be confirmed also by other seismic models, e.g. [2]. The thickening of the crust in correspondence to the Paranà basin as well as the presence of the Trans-Brazilian lineament, the Chaco and the Oriente basins are also visible in the model. Concerning the African Moho the GEMMA1.0 Moho seems to properly describe (differences with seismic observations smaller than 2 km) some interesting features not present in the other models as, for example the Garoua Rift in Cameroon (seismic observations from [45]), the Afar Depression in Ethiopia (seismic observations from [9]) or the East African Rift (seismic observations from [17]).

GEMMA1.0 also shows a deep and defined orogenetic crust (e.g. in the Himalayas, the Andes, the Rocky Mountains, the Alps and the Urals). This is due partially to the high resolution of the model, e.g. below the Alps and the Urals, and partially to the effect of unmodelled density anomalies. For example, the effect of the subduction of the Nazca plate under the South American Plate or of the collision between the Indian and Eurasian plate in the Himalayas, not modelled in GEMMA1.0, causes an anomalous thickening of the crust in such regions.

## 5 Conclusions

During the GEMMA project GOCE observations have been used in order to retrieve information on the Earth crustal structure. Different algorithms have been developed to estimate the Moho at regional and global scales. These algorithms and the numerical results obtained prove that GOCE observations can be profitably used to study the Earth crust–mantle interface and point out that the most critical step, in all the procedures implemented, is related to the data reduction.

Among the results the GEMMA1.0 model seems to be an improvement of the state of the art of the global crustal knowledge for different aspects: it combines GOCE data with CRUST2.0 seismic information (where judged reliable), it is well consistent with the actual gravity field, thus overcoming one of the main limitations of seismic Moho models. It is derived from a uniform and homogeneously distributed dataset (differently from seismic derived models where large areas are

uncovered by observations). The resulting crustal model has a planimetric resolution of only 0.5° × 0.5°. The whole model, i.e. the top and the bottom of each layer, the density distribution and the corresponding gravitational signal, is freely available through a web processing service at the webpage http://gocedata.como.polimi.it.

## References


1. Amante C, Eakins BW (2009) ETOPO1 1 Arc-minute global relief model: procedures, data sources and analysis. NOAA Technical Memorandum NESDIS NGDC-24, 19 p
2. Assumpção M, Feng M, Tassara A, Julià J (2013) Models of crustal thickness for South America from seismic refraction, receiver functions and surface wave tomography. Tectonophysics 609:82–96
3. Bassin C, Laske G, Masters G (2000) The k. AGU, EOS Trans 81
4. Braitenberg C, Ebbing J (2009) New insights into the basement structure of the West Siberian Basin from forward and inverse modelling of GRACE satellite gravity data. J Geophys Res-Sol Ea 114(B6):B06402. doi:10.1029/2008JB005799
5. Braitenberg C, Wang Y, Fang J, Hsu HY (2003) Spatial variations of flexure parameters over the Tibet-Qinghai Plateau. Earth Planet Sci Lett 205:211–224
6. Carbonell R, Levander A, Kind R (2013) The Mohorovičić discontinuity beneath the continental crust: an overview of seismic constraints. Science 609:353–376
7. Christensen NI, Mooney WD (1995) Seismic velocity structure and composition of the continental crust: a global view. J Geophys Res-Sol Ea 100(B6):9761–9788
8. Drinkwater MR, Floberghagen R, Haagmans R, Muzi D, Popescu A (2003) GOCE: ESA's first earth explorer core mission. Earth gravity field from space? From sensors to earth sciences. Springer, Netherlands, pp 419–432
9. Dugda MT, Nyblade AA, Julia J, Langston CA, Ammon CJ, Simiyu S (2005) Crustal structure in Ethiopia and Kenya from receiver function analysis: implications for rift development in eastern Africa. J Geophys Res-Sol Ea 110(B1). doi:10.1029/2004JB003065
10. Dziewonski AM, Anderson DL (1981) Preliminary reference earth model. Phys Earth Planet In 25(4):297–356
11. Exxon (1995). Tectonic map of the world, 18 sheets, scale 1:10,000,000. Technical Report. Exxon, Houston, Texas
12. Hadamard J (1923) Lectures on Cauchy problem in linear partial differential equations. Oxford University Press, London
13. Heiskanen WA, Moritz H (1967) Physical geodesy. W. H. Freeman, London
14. Jekeli C (1999) The determination of gravitational potential differences from satellite to satellite tracking. Celest Mech Dyn Astr 75:85–101
15. Koch KR (1999) Parameter estimation and hypothesis testing in linear models, 2nd edn. Springer, Berlin
16. Laske G, Masters G, Ma Z, Pasyanos M (2013) Update on CRUST1.0—a 1-degree global model of earth's crust. Geophys Res Abstr 15
17. Last RJ, Nyblade AA, Langston CA, Owens TJ (1997) Crustal structure of the East African Plateau from receiver functions and Rayleigh wave phase velocities. J Geophys Res-Sol Ea 102(B11):24469–24483
18. Lebedev S, Adam JMC, Meier T (2013) Mapping the Moho with seismic surface waves: a review, resolution analysis, and recommended inversion strategies. Tectonophysics 609: 377–394
19. Meier U, Curtis A, Trampert J (2007) Global crustal thickness from neural network inversion of surface wave data. Geophys J Int 169(2):706–722
20. Meissner R (1973) The 'Moho' as a transition zone. Geophys surv 1(2):195–216

21. Migliaccio F, Reguzzoni M, Sansò F, Tselfes N (2008) An error model for the GOCE space-wise solution by Monte Carlo methods. In: Sideris MG (ed) International Association of Geodesy Symposia? Observing our changing earth? vol 133, pp 337–344
22. Mohorovičić A (1992) Earthquake of 8 October 1909. Geofizika 9(1):3–55
23. Mooney WD, Laske G, Masters TG (1998) CRUST 5.1: a global crustal model at 5 × 5. J Geophys Res-Sol Ea 103(B1):727–747
24. Moritz H (1990) The figure of the earth: theoretical geodesy and the earth's Interior. Wichmann
25. Oldenburg DW (1974) The inversion and interpretation of gravity anomalies. Geophysics 39:526–536
26. Pail R, Goiginger H, Mayrhofer R, Schuh WD, Brockmann JM, Krasbutter I, Höck E, Fecher T (2010) GOCE gravity field model derived from orbit and gradiometry data applying the time-wise method. In: Proceedings of the ESA living planet symposium, ESA Publication SP-686, ESA/ESTEC, pp 978–992
27. Reguzzoni M, Samietro D (2014) GEMMA: an earth crustal model based on GOCE satellite data. Int J Appl Earth Obs Geoinf. doi:10.1016/j.jag.2014.04.002
28. Reguzzoni M, Sampietro D (2012) Moho estimation using GOCE data: a numerical simulation. In: Kenyon SC, Pacino MC, Marti U (eds) International Association of geodesy symposia, "geodesy for planet earth" vol 136, pp 205–214
29. Reguzzoni M, Tselfes N (2009) Optimal multi-step collocation: application to the space-wise approach for GOCE data analysis. J Geodesy 83(1):13–29
30. Reguzzoni M, Sampietro D, Sansò F (2013) Global Moho from the combination of the CRUST2.0 model and GOCE data. Geophys J Int 195(1):222–237
31. Sampietro D (2011) GOCE exploitation for Moho modeling and applications. In: Proceedings of the 4th international GOCE user workshop, vol 31, Munich, Germany
32. Sampietro D, Sansò F (2012) Uniqueness theorems for inverse gravimetric problems. In: Sneeuw N, Nóvak P, Crespi M, Sansò F (eds) International Association of Geodesy Symposia, VII Hotine-Marussi symposium on mathematical geodesy? vol 137, pp 111–115
33. Sampietro D, Reguzzoni M, Negretti M (2014) The GEMMA crustal model: first validation and data distribution. In: Proceedings of ESA living planet symposium 2013, ESA SP-722
34. Sampietro D, Reguzzoni M, Braitenberg C (2014) The GOCE estimated Moho beneath the Tibetan Plateau and Himalaya. In: Rizos C, Willis P (eds) International Association of Geodesy Symposia, "Earth on the edge: science for a sustainable planet" vol 139, pp 391–397
35. Sansò F (1980) Internal collocation. Memorie dell'Accademia dei Lincei XVI(1)
36. Sansò F, Barzaghi R, Tscherning CC (1986) Choice of norm for the density distribution of the earth. Geophys J Roy Astr S 87(1):123–141
37. Schock E (1984) On the asymptotic order of accuracy of Tikhonov regularization. J Optim Theory App 44(1):95–104
38. Shin YH, Shum CK, Braitenberg C, Lee SM, Xu H, Choi KS, Baek JH, Park JU (2009) Three dimensional fold structure of the Tibetan Moho from GRACE gravity data. Geophys Res Lett 36(1):L01302. doi:10.1029/2008JB005799
39. Sideris MG (1996) On the use of heterogeneous noisy data in spectral gravity field modeling methods. J Geodesy 70(8):470–479
40. Sjöberg LE (2009) Solving Vening Meisnesz-Moritz inverse problem in isostasy. Geophys J Int 179(3):1527–1536
41. Sjöberg LE, Bagherbandi M (2011) A method of estimating the Moho density contrast with a tentative application of EGM08 and CRUST2.0. Acta Geophys 59(3):502–525
42. Soller DR, Ray RD, Brown RD (1982) A new global crustal thickness map. Tectonics 1(2):125–149
43. Strang van Hees GL (2000) Some elementary relations between mass distributions inside the Earth and the geoid and gravity field. J Geodyn 29:111–123
44. Sünkel H (1985) An isostatic Earth model. Report No. 367, Department of Geodetic Science and Surveying. The Ohio State University, Columbus

45. Tokam APK, Tabod CT, Nyblade AA, Julià J, Wiens DA, Pasyanos ME (2010) Structure of the crust beneath Cameroon, West Africa, from the joint inversion of Rayleigh wave group velocities and receiver functions. Geophys J Int 183(2):1061–1076
46. Tseng TL, Chen WP, Nowack RL (2009) Northward thinning of Tibetan crust revealed by virtual seismic profiles. Geophys Res Lett 36:L24304. doi:10.1029/2009GL040457
47. Visser PNAM, Sneeuw N, Gerlach C (2003) Energy integral method for gravity field determination from satellite orbit coordinates. J Geodesy 77(3–4):207–216
48. Zhang Z, Klemperer SL (2005) West-east variation in crustal thickness in northern Lhasa block, central Tibet, from deep seismic sounding data. J Geophys Res 110:B09403. doi:10.1029/2004JB003139