

Validation strategy for satellite observations of tropospheric reactive gases

ANDREAS RICHTER*, MARK WEBER*, JOHN P. BURROWS*,
JEAN-CHRISTOPHER LAMBERT[†], ANNE VAN GIJSEL[#]

*Institute of Environmental Physics, University of Bremen

[†]Institut d'Aéronomie Spatiale de Belgique (IASB)

[#]Royal Netherlands Meteorological Institute (KNMI)

richter@iup.physik.uni-bremen.de

Abstract

Satellite observations of tropospheric reactive gases are an integral part of the earth observing system but require continuous validation by independent measurements. For short-lived tropospheric species, the large variability in space and time results in specific challenges for validation, often combined with the scarcity of appropriate validation data. In this paper, the need for validation is discussed, previous work on validation of satellite observations is briefly reviewed, and the challenges and possible approaches for current and future validation networks are evaluated.

I. INTRODUCTION

Over the last two decades, satellite observations of tropospheric composition have become possible using nadir viewing spectrometers operating in the UV, visible, near infrared, and thermal infrared spectral range. Using measurements from instruments such as GOME, SCIAMACHY, OMI, GOME-2, IASI, TES, and MOPITT, global maps of the spatial distribution of many of the most important tropospheric reactive gases including O₃, H₂O, CO, NO₂, SO₂, HCHO, CHOCHO and BrO can be retrieved. These data for the first time provide a global observational view of tropospheric chemistry, and have led to important insights into the relevance of trace gas emission sources, the transport and chemical transformation of reactive species in the atmosphere and the temporal and spatial scales involved. They have also been applied

in studies on air pollution, the assignment of emissions sources and strengths and their changes over time [e.g. Martin, 2008, Wagner et al., 2008, Burrows et al., 2011]. With improving spatial resolution of the sensors and maturity of the retrieval algorithms, applications to chemical weather forecast and routine air quality monitoring will soon become possible.

As any remote sensing observation, satellite data on tropospheric species needs to be validated using independent data with known and documented uncertainties, in order to understand and characterise the capabilities and uncertainties of the satellite measurement, assess data quality, and provide users with quality indicators enabling them to judge the fitness of the data for their purpose. For satellite observations of stratospheric composition, in particular for ozone, validation has been performed on a routine

basis for several decades, and in many cases, both the independent measurements and the methods applied are close to being mature. However, the situation for tropospheric species in general and in particular for the reactive trace gases is not as advanced, and in fact, many of the relevant data products in this field are still poorly validated.

In this manuscript, an attempt is made to discuss the challenges and limitations of validating satellite measurements of reactive tropospheric gases, to review the achievements reached so far, to investigate possibilities to overcome the current limitations, and to formulate recommendations for a validation strategy for the current and future suite of European tropospheric space sensors. The discussion will mainly focus on data products from UV/vis nadir sounders but the concepts are also applicable to tropospheric NIR and TIR data sets

II. THE CHALLENGE

When trying to validate satellite observations of tropospheric reactive gases, a number of challenges become apparent that make this a more difficult task than expected.

The first characteristic of reactive gases is their large variability in space and time. This is the direct result of their reactivity which leads to a short atmospheric lifetime and this, in combination with often strongly localised sources, results in highly varying atmospheric fields. This is true for both the horizontal and the vertical direction. For instance, species such as NO_2 are mainly residing in the boundary layer of polluted regions and will not even be well mixed within this layer. A large spatial variability in combination with atmospheric transport also leads to concentrations rapidly changing in time and as such poses a problem for validation as localised validation measurements are not representative for larger areas and time differences between satellite and validation meas-

urement have to be small to ensure comparability.

A direct result of the large variability is the presence of strong concentration gradients, which complicate validation as the exact position and time of validation measurement has a large impact on the result. In contrast, the satellite measurements usually average over larger areas, smoothing the gradients and not reproducing the variability of the reference measurement.

The short atmospheric lifetime in combination with active photochemistry also leads to diurnal variations of the atmospheric concentrations of some species such as NO_2 , often enhanced by diurnal variations in emissions for example during rush hour or because of the higher probability of thunderstorms and lightning, in the afternoon. Again, this necessitates a good temporal coincidence of satellite and validation measurements.

In contrast to the stratosphere, the spatial distribution of tropospheric species is strongly influenced by the distribution of emission sources such as cities, power plants, biogenic sources, wild fires, sea ice etc., which are very inhomogeneous and often in regions not well accessible. A validation network well representing the variability of atmospheric conditions therefore has to be wide spread, including also difficult to probe regions such as rain forests, cities or the polar sea ice region.

The special characteristics of tropospheric retrievals also have important impacts on the validation needs. In most cases, the sensitivity of the retrieval varies strongly with altitude, usually with lowest values close to the surface where validation measurements are often located. As a result, the retrieval depends critically on a-priori data including the vertical profiles of the species of interest and of temperature, surface reflectance and emissivity, and also the presence of clouds and aerosols. For proper validation, ideally

these input data will also have to be validated in order to be able to decide if any difference observed is linked to the measurement itself or to the ancillary data used.

An additional challenge is the small signal often obtained for tropospheric species, either because their abundances are small or because it is difficult to separate the tropospheric from the stratospheric signals. In many cases, the validation measurements themselves are also not as accurate and precise for these small signals as one would like, adding the uncertainty of the validation data to that of the satellite measurement.

Considering all the above points, an ideal validation measurement for tropospheric species should provide the vertical profile of the species at different times of the day, for all seasons, at good spatial sampling and covering an area typical for a satellite observation. It should also provide a good coverage of all atmospheric situations, have decent observation statistics and sufficient accuracy (bias, precision) and cover also the quantities needed as additional input in the retrievals. Unfortunately, the typical validation measurement falls short in one or even many of these aspects, and in some cases, there exists nearly no independent validation data to compare with.

III. CURRENT VALIDATION WORK

In spite of the difficulties, many studies have been performed validating tropospheric satellite products. These studies have been documented in the literature, for example in the SCIAMACHY book [Gottwald and Bovensmann, 2011], the ACCENT-AT2 book on Remote Sensing of Tropospheric Composition from Space [Burrows et al., 2011] and the AURA validation collection [Schoeberl et al., 2008], but also in many individual articles which are too numerous to be referenced here.

Arguably the best situation exists for tropospheric ozone validation, where data from

the ozone sonde network as well as lidar profiles mainly from the ground but also from aircraft and in-situ aircraft measurements can be applied [e.g. Verstraeten et al., 2013]. While these data provide vertical profiles at relatively high frequency and good systematic error and precision, they still lack coverage in the southern hemisphere and at low latitudes, in spite of specific programs like SHADOZ [Thompson et al., 2007] which in recent years have increased the number of sonde stations in the tropics.

For CO, in-situ aircraft observations are the main source of validation which is assisted by ground-based Fourier Transform Spectrometer (FTS) observations providing column data in cloud free conditions [e.g. Emmons et al., 2009, Kerzenmacher et al., 2012]. Acquired at a limited number of stations set up initially for stratospheric monitoring, these data sets are sparse and usually do not include pollution hot spots or biomass burning areas. Also, regions with challenging retrieval conditions (dark surfaces) are not covered adequately. However, due to the relatively long lifetime of CO, coincidence criteria do not have to be very strict and satellite data can be linked to validation observations by using backward trajectories.

For tropospheric NO₂, different validation approaches have been taken. A large number of surface in-situ measurements of NO₂ are taken within national and local air quality networks. By using assumptions on the vertical distribution of NO₂, these data can be converted to tropospheric columns to be used for validation of satellite measurements [e.g. Ordóñez et al., 2006, Boersma et al., 2009]. While providing good statistics at least locally, such comparisons suffer from the large uncertainty introduced by the conversion from surface mixing ratios to columns, from the lack of accuracy of the instruments employed and from sampling issues when using extremely localised road side measurements. Tropospheric columns

(and coarse vertical profiles) can be measured by Multi-Axis Differential Optical Absorption Spectroscopy (MAX-DOAS) instruments from the ground, and these data are well suited for validation [e.g. Irie et al., 2008]. However, the number of such stations is small and many of them are located in clean air regions. Although DOAS measurements average in the vertical and to a lesser degree also in the horizontal direction, their measurement volume in the troposphere is still much smaller than that of current space instruments. This lack of spatial representativeness and the fairly large uncertainty of individual observations is a problem for validation unless many instruments are distributed over a larger area. The validation of stratospheric columns, that are used to infer the tropospheric column from satellite total column data, can be done using zenith-sky twilight DOAS measurements [e.g. Ionov et al., 2008, Peters et al., 2012]. The spatial variability has been addressed by airborne observations using both in-situ and DOAS remote sensing measurements [e.g. Heue et al., 2005, Martin et al., 2004] but these are very limited in number. Recently, the use of car mounted MAX-DOAS has also been demonstrated for validation [Shaiganfar et al., 2011], providing spatial coverage at low cost. Vertical profile information is currently mainly available from dedicated aircraft flights using in-situ instruments, complemented by MAX-DOAS which mainly resolves the lowest layers.

For other absorbers such as BrO, SO₂, HCHO, CHOCHO and IO, even less validation is available, and it is nearly completely limited to a small number of in-situ observations on the ground and in aircraft and to a few active and passive DOAS and FTS observations [e.g. Vigouroux et al., 2009, Heue et al., 2011, Choi et al., 2012, Dix et al., 2013]. For SO₂, a more extensive network of DOAS instruments has been set-up in the EC project NOVAC [Galle et al., 2009] for monitoring

volcanic emissions, and this could possibly be used to validate satellite data.

IV. STRATEGY FOR THE FUTURE

Any strategy for the future has to address the gaps identified in the current validation activities. The requirements for a good validation strategy are simple – continue acquiring new data, go to the right places, take the right measurements at the right time, accumulate enough data, include validation of ancillary data and facilitate data access. Several actions can and should be taken to move into this direction.

Most importantly, continued operation of existing networks such as the DOAS, MAX-DOAS, lidar, FTS, and ozone sonde networks needs to be secured and maintained. Unfortunately, many of the stations in these networks do not have secure funding and the number of measurements taken is currently declining, further limiting our ability to validate tropospheric satellite observations. As many of the stations are in clean air regions as was appropriate for their original purpose of upper atmospheric observations in the context of the Network for the Detection of Atmospheric Composition Change, NDACC (previously known as Network for the Detection of Stratospheric Change, NDSC) or the Global Atmospheric Watch programme of the World Meteorological Organisation, WMO-GAW, they should be complemented by stations in regions where validation of tropospheric species is needed, for example in pollution hot-spots, biomass burning regions, or areas with large biogenic emissions. Also, there is a lack of stations in the Southern Hemisphere. In some cases, existing networks such as WMO-GAW could be augmented by additional instrumentation such as MAX-DOAS instruments to make them (more) useful for validation while making use of existing infrastructure and experience.

One interesting possibility is the establishment of a small number of end-to-end reference sites (or so-called primary sites), dedicated to the validation of tropospheric data and the intermediate steps of their production. These stations should be equipped to provide high quality measurements of all the quantities needed for validation including the input quantities for the retrievals under different conditions. They should be operated having validation in mind by taking year-round measurements at the right times of day and with a perspective of long-term operation. By strategically placing them on different continents (US, Europe, Asia), such stations could be used to validate both Low Earth Orbit (LEO) satellites and the new generation of geostationary (GEO) observatories to be launched in the coming years. By ensuring that the same LEO satellites are validated by all stations, these validated LEO instruments could serve as transfer standards between the three GEO satellites which do not have overlapping measurements.

While networks of stations provide good statistical and long-term validation, campaign based validation is essential for more detailed analysis. These campaigns should take place in regions close to relevant observations which are not yet validated, for example regions with large spatial gradients, pollution transport, biogenic emissions, shipping emissions, biomass burning, lightning, and bromine explosions. Sometimes, experiments of opportunity such as emission reductions for Olympic Games [e.g. Mijling et al., 2009] or changes in legislation affecting pollutant emissions [e.g. Kim et al., 2006] can also be performed. Often, such campaigns can combine different aims such as addressing a science question (e.g. TRANSBROM campaign [Krüger and Quack, 2013]), an instrument intercomparison (e.g. CINDI [Piters et al., 2012]) or a multi-platform experiment such as DISCOVER-AQ ([\[aq.larc.nasa.gov/\]\(http://aq.larc.nasa.gov/\)\) with validation, while other campaigns are fully dedicated to validation \(e.g. SCIAVALUE \[Fix et al., 2005\]\). However, it is essential to ensure that the needs for validation are addressed from the planning to the execution of measurements and finally the analysis of data. In this context, campaigns already planned by other groups can be used for validation purposes by adding instrumentation for example for column measurements or profiling and ensuring that validation aspects are taken into account when designing the measurement programme.](http://discover-</p></div><div data-bbox=)

In some cases, new developments with potential applications for validation can be supported. Examples are the recent construction of NO₂ sondes [Sluis et al., 2010] to be used in a similar way as O₃ sondes that can fill an important gap in the atmospheric observation system from ground. The Pandora systems of small and flexible remote sensing instruments [Herman et al., 2009] could be used to extend MAX-DOAS networks. Recent developments of highly sensitive Cavity Enhanced or Cavity Ring Down Spectroscopy (CRDS) instrumentation promise better detection limits for a number of species, and smaller and cheaper lidar and FTS instruments could facilitate larger numbers of observations albeit at reduced accuracy. A more radical approach could integrate a large number of small and cheap sensors, usually based on solid-state detectors for crowd measurements with low precision but excellent sampling statistics.

A large potential for validation measurements lies in the use of existing platforms such as cars, trains, ships or commercial aircrafts. By mounting small and automated instruments to these vessels, good spatial coverage and statistics can be achieved without having to cover the large costs of vehicle operation. The usability of such platforms for validation measurements is limited by the constraints imposed by the platform opera-

tors which are often not in line with validation requirements, but as an additional data source, such measurements can provide an important contribution.

In addition to existing platforms, unconventional platforms can become very useful for validation measurements, for example ultra-light aircraft, unmanned aircraft, zeppelins, tethered balloons or buoys in the ocean and sea ice. All these platforms have the potential to extend the range of validation measurements either vertically or to regions not usually accessible by other observations.

One approach to validation which is not linked to independent measurements is the use of chemical data assimilation systems to assess the consistency of a satellite data set. As the data assimilation system implements the chemical and dynamical processes in the atmosphere, it can be used to detect internal biases, spatial offsets and temporal changes in the assimilated data sets. For example, degradation in instrument performance leading to bias in a data product can be picked-up in the quality control of a data assimilation system. This has been successfully used in the past for CO observations from MOPITT and IASI in the MACC assimilation system, and can potentially be extended to other species. In addition, the use of chemical transport models in general and data assimilation systems in particular can increase the number of co-locations usable for validation as the model effectively interpolates in time and space [e.g. Klonecki et al., 2012]. Such techniques work best for long-lived species which are better constrained by transport and chemistry in the model, limiting the applicability of data assimilation for species such as NO₂ or BrO. Finally, comparisons with model data should always be used with care, and can only complement, not replace validation based on real observational data.

As important as collecting new data is facilitating access to existing data. This includes

data sets assembled for validation purposes as well as data from campaigns and observational networks. Ideally, all these data sets should be available for validation through a unified portal, providing a uniform interface and data protocol as well as consistent metadata and formats. Such activities are currently pursued in the context of several initiatives and programs including GEOMS (Generic Earth Observation Metadata Standard, see <http://avdc.gsfc.nasa.gov/GEOMS>), NDACC (<http://www.ndacc.org>), NORS (Demonstration Network Of ground-based Remote Sensing Observations in support of the Copernicus Atmospheric Service, see <http://nors.aeronomie.be>), the ESA and NASA validation data centres (EVDC (<http://nadir.nilu.no/calval>) and AVDC (<http://avdc.gsfc.nasa.gov>)), as well as the CEOS Cal/Val Portal (Committee on Earth Observation Satellites, see <http://calvalportal.ceos.org/>).

Finally, an essential part of a validation strategy is also to ensure the conservation of know-how. As space-borne missions and validation projects come and go, the team of validation scientists is constantly changing, and there is a component of periodic reinvention of the wheel as new generations of scientists encounter the challenges of satellite data validation. It is, therefore, important to provide consistent guidelines to validation groups including information on validation aims, reporting strategies, documentation, fitness for purpose considerations as well as display and interpretation of validation results. In addition, common language and approaches based on metrology should be enforced in validation for the treatment of vertical and horizontal resolution, the treatment of time mismatches, the correct use of appropriate terminology and standards, the differentiation between type A and type B errors, and error reporting in general.

V. CONCLUSIONS

Validation of satellite observations of tropospheric reactive trace gases is a challenging task, both because of the intrinsic variability in the atmospheric fields to be observed and because of the peculiarities of remote sensing of the troposphere. It is an on-going activity that needs continuous support for long-term measurements, campaigns, data analysis, and development of new capabilities.

For tropospheric species in particular, it is crucial for validation measurements to have good and adaptive spatial coverage, coverage of the right quantities, and a combination of long-term observations and dedicated campaigns for specific atmospheric events.

A multi-tiered approach is needed to improve on the current situation where there are by far not enough validation means available for tropospheric species. This can be only overcome by substantial investment in money and time to establish a more troposphere oriented and more comprehensive validations system. Otherwise, we will not have the infrastructure, people, and data needed for tropospheric validation of the many current and upcoming European missions such as GOME-2, IASI, Sentinel-5P, Sentinel-4, and Sentinel-5.

VI. ACKNOWLEDGEMENTS

Parts of this work were performed in the framework of the ESA funded SCIOLOV, Multi-TASTE and VALID projects, and as well as the Belgian BELSPO/ProDEX A3C project and the NL-SCIAvisie project funded by the Dutch government.

REFERENCES

[Boersma et al., 2009] Boersma, K. F., D. J. Jacob, M. Trainic, Y. Rudich, I. DeSmedt, R. Dirksen, and H. J. Eskes (2009). Validation of urban NO₂ concentrations and their diurnal and seasonal variations observed from the

SCIAMACHY and OMI sensors using in situ surface measurements in Israeli cities, *Atmos. Chem. Phys.*, 9, 3867-3879.

[Burrows et al., 2011] Burrows, J. P., Borrell, P., Platt, U. (Eds) (2011). *The Remote Sensing of Tropospheric Composition from Space*. Springer, Berlin, Heidelberg, doi:10.1007/978-3-642-14791-3.

[Choi et al., 2012] Choi, S., Wang, Y., Salawitch, R. J., Canty, T., Joiner, J., Zeng, T., Kurosu, T. P., Chance, K., Richter, A., Huey, L. G., Liao, J., Neuman, J. A., Nowak, J. B., Dibb, J. E., Weinheimer, A. J., Diskin, G., Ryerson, T. B., da Silva, A., Curry, J., Kinnison, D., Tilmes, S., and Levelt, P. F (2012). Analysis of satellite-derived Arctic tropospheric BrO columns in conjunction with aircraft measurements during ARCTAS and ARCPAC, *Atmos. Chem. Phys.*, 12, 1255-1285, doi:10.5194/acp-12-1255-2012.

[Dix et al., 2013] Dix, B., Baidar, S., Bresch, J. F., Hall, S.R., Schmidt, K. S., Wang, S., Volkamer, R. (2013). Detection of iodine monoxide in the tropical free troposphere, *Proc. Natl. Acad. Sci.* 110, 2035-2040, doi:10.1073/pnas.1212386110.

[Emmons et al., 2009] Emmons, L. K., Edwards, D. P., Deeter, M. N., Gille, J. C., Campos, T., Nédélec, P., Novelli, P., and Sachse, G (2009). Measurements of Pollution In The Troposphere (MOPITT) validation through 2006, *Atmos. Chem. Phys.*, 9, 1795-1803, doi:10.5194/acp-9-1795-2009.

[Fix et al., 2005] Fix, A., Ehret, G., Flentje, H., Poberaj, G., Gottwald, M., Finkenzeller, H., Bremer, H., Bruns, M., Burrows, J. P., Kleinböhl, A., Küllmann, H., Kuttippurath, J., Richter, A., Wang, P., Heue, K.-P., Platt, U., Pundt, I., Wagner, T. (2005). SCIAMACHY validation by aircraft remote sensing: design, execution, and first measurement results of the SCIA-VALUE mission, *Atmos. Chem. Phys.*, 5, 1273-1289.

[Galle et al., 2009] Galle, B., M. Johansson, C. Rivera, Y. Zhang, M. Kihlman, C. Kern, T. Lehmann, U. Platt, S. Arellano and S. Hi-

- dalgo (2009). NOVAC – A global network for volcanic gas monitoring, network layout and instrument description. *J. Geophys. Res.*, 115, D05304, doi:10.1029/2009JD011823.
- [Gottwald and Bovensmann, 2011] Gottwald, M. and Bovensmann, H. (Eds.) (2011). *SCIAMACHY – Exploring the Changing Earth's Atmosphere*, Springer, Dordrecht, Heidelberg, London, New York, doi:10.1007/978-90-481-9896-2.
- [Herman et al., 2009] Herman, J., A. Cede, E. Spinei, G. Mount, M. Tzortziou, and N. Abuhassan (2009). NO₂ column amounts from ground-based Pandora and MFDOAS spectrometers using the direct-sun DOAS technique: Intercomparisons and application to OMI validation, *J. Geophys. Res.*, 114, D13307, doi:10.1029/2009JD011848.
- [Heue et al., 2005] Heue, K.-H., A. Richter, T. Wagner, M. Bruns, J. P. Burrows, C. v. Friedeburg, W. D. Lee, U. Platt, I. Pundt, P. Wang (2005). Validation of SCIAMACHY tropospheric NO₂-columns with AMAXDOAS measurements, *Atmos. Chem. Phys.*, 5, 1039-1051.
- [Heue et al., 2011] Heue, K.-P., Brenninkmeijer, C. A. M., Baker, A. K., Rauthe-Schöch, A., Walter, D., Wagner, T., Hörmann, C., Sihler, H., Dix, B., Frieß, U., Platt, U., Martinsson, B. G., van Velthoven, P. F. J., Zahn, A., and Ebinghaus, R. (2011). SO₂ and BrO observation in the plume of the Eyjafjallajökull volcano 2010: CARIBIC and GOME-2 retrievals, *Atmos. Chem. Phys.*, 11, 2973-2989, doi:10.5194/acp-11-2973-2011.
- [Ionov et al., 2008] Ionov, D. V., Y. M. Timofeyev, V. P. Sinyakov, V. K. Semenov, F. Goutail, J.-P. Pommereau, E. J. Bucsela, E. A. Celarier, and M. Kroon (2008). Ground-based validation of EOS-Aura OMI NO₂ vertical column data in the midlatitude mountain ranges of Tien Shan (Kyrgyzstan) and Alps (France), *J. Geophys. Res.*, 113, D15S08, doi:10.1029/2007JD008659.
- [Irie et al., 2008] Irie, H., Kanaya, Y., Akiyama, H., Tanimoto, H., Wang, Z., Gleason, J. F., and Bucsela, E. J. (2008). Validation of OMI tropospheric NO₂ column data using MAX-DOAS measurements deep inside the North China Plain in June 2006: Mount Tai Experiment 2006, *Atmos. Chem. Phys.*, 8, 6577-6586, doi:10.5194/acp-8-6577-2008.
- [Kerzenmacher et al., 2012] Kerzenmacher, T., Dils, B., Kumps, N., Blumenstock, T., Clerbaux, C., Coheur, P.-F., Demoulin, P., García, O., George, M., Griffith, D. W. T., Hase, F., Hadji-Lazaro, J., Hurtmans, D., Jones, N., Mahieu, E., Notholt, J., Paton-Walsh, C., Raffalski, U., Ridder, T., Schneider, M., Servais, C., and De Mazière, M. (2012). Validation of IASI FORLI carbon monoxide retrievals using FTIR data from NDACC, *Atmos. Meas. Tech.*, 5, 2751-2761, doi:10.5194/amt-5-2751-2012.
- [Kim et al., 2006] Kim, S.-W., A. Heckel, S. A. McKeen, G. J. Frost, E.-Y. Hsie, M. K. Trainer, A. Richter, J. P. Burrows, S. E. Peckham, and G. A. Grell (2006). Satellite observed U.S. power plant NO_x emission reductions and their impact on air quality, *Geophys. Res. Lett.*, 33, L22812, doi:10.1029/2006GL027749.
- [Klonecki et al., 2012] Klonecki, A., Pommier, M., Clerbaux, C., Ancellet, G., Cammas, J.-P., Coheur, P.-F., Cozic, A., Diskin, G. S., Hadji-Lazaro, J., Hauglustaine, D. A., Hurtmans, D., Khatatov, B., Lamarque, J.-F., Law, K. S., Nedelec, P., Paris, J.-D., Podolske, J. R., Prunet, P., Schlager, H., Szopa, S., and Turquety, S. (2012). Assimilation of IASI satellite CO fields into a global chemistry transport model for validation against aircraft measurements, *Atmos. Chem. Phys.*, 12, 4493-4512, doi:10.5194/acp-12-4493-2012.
- [Krüger and Quack, 2013] Krüger, K. and Quack, B. (2013). Introduction to special issue: the TransBrom Sonne expedition in the tropical West Pacific, *Atmos. Chem. Phys.*, 13, 9439-9446, doi: 10.5194/acp-13-9439-2013.
- [Mijling et al., 2009] Mijling, B., R. J. van der A, K. F. Boersma, M. Van Roozendaal, I. De Smedt, and H. M. Kelder (2009). Reductions

of NO₂ detected from space during the 2008 Beijing Olympic Games, *Geophys. Res. Lett.*, 36, L13801, doi:10.1029/2009GL038943.

[Martin, 2008] Martin, R.V. (2008). Satellite remote sensing of surface air quality, *Atmos. Environ.*, 42, 7823-7843.

[Martin et al., 2004] Martin, R.V., D.D. Parrish, T.B. Ryerson, D.K. Nicks Jr., K. Chance, T.P. Kurosu, A. Fried, B.P. Wert, D.J. Jacob, and E. D. Sturges (2004). Evaluation of GOME satellite measurements of tropospheric NO₂ and HCHO using regional data from aircraft campaigns in the southeastern United States, *J. Geophys. Res.*, 109, D24307, doi:10.1029/2004JD004869.

[Ordóñez et al., 2006] Ordóñez, C., A. Richter, M. Steinbacher, C. Zellweger, H. Nüß, J. P. Burrows, and A. S. H. Prévôt (2006). Comparison of 7 years of satellite-borne and ground-based tropospheric NO₂ measurements around Milan, Italy, *J. Geophys. Res.*, 111, D05310, doi:10.1029/2005JD006305.

[Peters et al., 2012] Peters, E., Wittrock, F., Großmann, K., Frieß, U., Richter, A., and Burrows, J. P. (2012). Formaldehyde and nitrogen dioxide over the remote western Pacific Ocean: SCIAMACHY and GOME-2 validation using ship-based MAX-DOAS observations, *Atmos. Chem. Phys.*, 12, 11179-11197, doi:10.5194/acp-12-11179-2012.

[Peters et al., 2012] Peters, A. J. M., Boersma, K. F., Kroon, M., Hains, J. C., Van Roozendaal, M., Wittrock, F., Abuhassan, N., Adams, C., Akrami, M., Allaart, M. A. F., Apituley, A., Beirle, S., Bergwerff, J. B., Berkhout, A. J. C., Brunner, D., Cede, A., Chong, J., Clémer, K., Fayt, C., Frieß, U., Gast, L. F. L., Gil-Ojeda, M., Goutail, F., Graves, R., Griesfeller, A., Großmann, K., Hemerijckx, G., Hendrick, F., Henzing, B., Herman, J., Hermans, C., Hoexum, M., van der Hoff, G. R., Irie, H., Johnston, P. V., Kanaya, Y., Kim, Y. J., Klein Baltink, H., Kreher, K., de Leeuw, G., Leigh, R., Merlaud, A., Moerman, M. M., Monks, P. S., Mount, G. H., Navarro-Comas, M., Oetjen, H., Pazmino, A., Perez-Camacho,

M., Peters, E., du Piesanie, A., Pinardi, G., Puentedura, O., Richter, A., Roscoe, H. K., Schönhardt, A., Schwarzenbach, B., Shaiganfar, R., Sluis, W., Spinei, E., Stolk, A. P., Strong, K., Swart, D. P. J., Takashima, H., Vlemmix, T., Vrekoussis, M., Wagner, T., Whyte, C., Wilson, K. M., Yela, M., Yilmaz, S., Zieger, P., and Zhou, Y. (2012). The Cabauw Intercomparison campaign for Nitrogen Dioxide measuring Instruments (CINDI): design, execution, and early results, *Atmos. Meas. Tech.*, 5, 457-485, doi:10.5194/amt-5-457-2012.

[Schoeberl et al., 2008] Schoeberl, M. R., A. R. Douglass, and J. Joiner (2008). Introduction to special section on Aura Validation, *J. Geophys. Res.*, 113, D15S01, doi:10.1029/2007JD009602.

[Shaiganfar et al., 2011] Shaiganfar, R., Beirle, S., Sharma, M., Chauhan, A., Singh, R. P., and Wagner, T. (2011). Estimation of NO_x emissions from Delhi using Car MAX-DOAS observations and comparison with OMI satellite data, *Atmos. Chem. Phys.*, 11, 10871-10887, doi:10.5194/acp-11-10871-2011.

[Sluis et al., 2010] Sluis, W. W., Allaart, M. A. F., PETERS, A. J. M., and Gast, L. F. L.: The development of a nitrogen dioxide sonde, *Atmos. Meas. Tech.*, 3, 1753-1762, doi:10.5194/amt-3-1753-2010, 2010

[Thompson et al., 2007] Thompson, A.M., J.C. Witte, H.G.J. Smit, S.J. Oltmans, B.J. Johnson, V.W.J. H. Kirchhoff, and F.J. Schmidlin (2007). Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2004 tropical ozone climatology: 3. Instrumentation, station-to-station variability, and evaluation with simulated flight profiles, *J. Geophys. Res.*, 112, D03304, doi:10.1029/2005JD007042.

[Verstraeten et al., 2013] Verstraeten, W. W., Boersma, K. F., Zörner, J., Allaart, M. A. F., Bowman, K. W., and Worden, J. R. (2013). Validation of six years of TES tropospheric ozone retrievals with ozonesonde measurements: implications for spatial patterns and

temporal stability in the bias, *Atmos. Meas. Tech.*, 6, 1413-1423, doi:10.5194/amt-6-1413-2013.

[Vigouroux et al., 2009] Vigouroux, C., Hendrick, F., Stavrakou, T., Dils, B., De Smedt, I., Hermans, C., Merlaud, A., Scolas, F., Senten, C., Vanhaelewyn, G., Fally, S., Carleer, M., Metzger, J.-M., Müller, J.-F., Van Roozendael, M., and De Mazière, M. (2009). Ground-based FTIR and MAX-DOAS observations of formaldehyde at Réunion Island and comparisons with satellite and model data, *Atmos. Chem. Phys.*, 9, 9523-9544, doi:10.5194/acp-9-9523-2009.

[Wagner et al., 2008] Wagner, T., S. Beirle, T. Deutschmann, E. Eigemeier, C. Frankenberg, M. Grzegorski, C. Liu, T. Marbach, U. Platt, and M. Penning de Vries (2008). Monitoring of atmospheric trace gases, clouds, aerosols and surface properties from UV/vis/NIR satellite instruments. *J. Opt. A Pure Appl. Opt.*, 10, 104019, doi:10.1088/1464-4258/10/10/104019.