Data Assimilation Schemes in Colombian Geodynamics - Cooperative Research Plan for 2017 - 2020 Between Universidad EAFIT and TUDelft, With the Help of Universidad de Antioquia and Universidad Nacional de Colombia Sede Medellin

---

Inicio: 1 de Enero de 2017
Final: 30 de Diciembre de 2020

Study of the impact of satellite and ground based data assimilated in meteorological and chemical transport models over natural areas of Colombia

Admission Essay to the PhD Program in Mathematical Engineering

Medellin Air qUality Initiative MAUI

MAUI-RT-04

Entidad Ejecutora	Universidad EAFIT Cra 49 No 7sur - 50 Medellín, Colombia
	Grupo de investigación en modelado matemático – GRIMMAT Grupo reconocido por COLCIENCIAS Categoría A
Responsables	Prof. Olga Lucia Quintero Montoya Prof. Nicolás Pinel Peláez. Investigadores
Entidades Cooperadoras	Department of Applied Mathematics - Tu Delft, Delft The Netherlands TNO
Responsables	Arnold Heemink Arjo Segers

CONTROL DE EDICIÓN Y DISTRIBUCIÓN

Edición	Control*	Nombre y Cargo	Firma	Entidad	Fecha (Día/Mes/Año)
	Creación	Andrés Yarce		Universidad EAFIT	

* Especificar tipo de control: Creación - Revisión - Modificación - Distribución.

Table of contents

1. **PROBLEM STATEMENT**

The term “air pollution” could seem something related to the recent years but, the importance of it for the human being, was recognized far before 18th century when scientist like Rutherford and Lavoisier discovered the chemical composition of atmospheric pollutants (Segers, 2002). With increasing levels of pollution and accelerated population growth, the term becomes increasingly relevant.

Poor air quality (AQ) is a deathly problem global in scale. According to the World Health Organization (WHO), deaths in the world related to pollution are 6,5 million of people each year and, based on data of the Metropolitan area of Medellín, 9,2% of the deaths in the valley are related to the contamination problem.

Human activities generate pollution in the place where are being carried out. Sometimes the pollutants travel to remote areas due to wind and rivers currents and generates problems in the places where are being deposited. Medellín is a source of pollutants that escape from the valley being transported mainly towards to the North and the North-West in different times in the year (Nicolas Pinel, 2017) to different natural areas affecting the equilibrium of the ecosystem generating critical damages. The motivation of this thesis proposal is oriented to understand the chemical atmospheric transport processes of several compounds from urban areas and agriculture production centers to natural zones to initiate the identification of those areas that need more than local conservation effort for the preservation of their ecological functions.

Mathematical models are used to describe the reality between the interaction of diverse variables through certain atmospheric physical processes (Evensen, 2009). Models aims to describe the reality accurately but, generally, exists uncertainty on the simulated outputs of the model and the reality measured.

To reduce the gap of the model output and the exact solution, there is a technique known as Data Assimilation (DA) which is a structured way to combine data from observations with the models with the objective to reduce the uncertainty of the forecast, calibrate the model, suggest new point for take measures and make uncertainty analysis over the model.

Through techniques of analysis of impact is possible to measure the impact of the data that is externally introduced to the model in the assimilation process and also retrieving more data from the available will possible to identify new places to acquire the measurements to incorporate to the model which will give us a broad overview of what is happening with the pollutants are being producing. In accordance with the preceding outlook, the following research questions have been proposed:

- What other kind of satellite or ground based data would help us to have more information to increase the spatial and chemical resolution of the different compounds from the data we have available for assimilate with the LOTOS EUROS model?
- How to measure the impact of the observation of several sensors on the model to have a weighting strategy or sensitivity analysis to improve the forecast using the methods developed by Verlaan (Verlaan & Sumihar, 2016) for Data Assimilation schemes?.
- How to evaluate the effect of new localization strategies developed by Fu et al, 2015 over the observation sensitivity analysis proposed by (Verlaan and Sumihar, 2016) in LOTOS-EUROS model for the Aburrá Valley and areas with protected areas.
- How to optimize the retrieval of data from which it is normally possible to obtain data with an available network of sensors (Dammers et al., 2015)?
- How to determine how and where to obtain data of contamination deposition processes in different ecosystems affected by human impact choosing the places and kind of sensors to install in order to have the greatest representativity for the cases of study we are interested.
- How could be developed a low-cost remote sensing sensor network for monitor large natural ecosystems? In terms of social impact, what will be the problems associated to the disposition of remote sensing networks and what mechanisms could be develop to solve them.
- How to perform further research to formulate the criteria for selection of the redundant data based on the shape of the curve obtained from the observation matrix eigenvalues or another new one? (Krymskaya, 2013).

2. JUSTIFICATION

This project is elaborated due to the necessity to identify the ecosystems that are being affected by atmospheric pollution transported from far away places to protected natural zones of Colombia. To quantify the deposition of some of these compounds, specifically nitrogenous pollutants and ozone is something that has not been done in Colombia and is of special priority more now that the bio economy is owned as a pillar of economic growth in the country being interesting to have information about the state of ecosystems.

In Colombia, there is not yet a complete model that allows us to make accurate decisions to guide aspects like urban growth, agricultural activities, measure the human development impact in several ways to take early warning decisions. In South America, one of the first times that a regional ensemble data assimilation and forecasting system was carried out was by Dillon et. Al. (Dillon et al., 2015), however, It was focused over the south region of the continent. It is important to develop in Colombia, research with Chemical transport models, that can be supported with the regional available observations such satellite information or sensor networks to improve their simulated output to produce simulated real time concentration and forecast.

The model to be implemented will be the LOTOS-EUROS model, a chemistry transport model developed by TuDelft University, RIVM (http://www.rivm.nl/) and TNO (https://www.tno.nl/en/) in Holland. Another justification for this research is that promotes the technology transfer between countries.

Another of the interest is to measure the impact of the data we are using to assimilate with the models to find out how god they are and how we can retrieve more data for it. We have the necessity of having more data from the different pollutants of interest for the research compounds on the areas we want to have well performed simulated daily concentrations and forecast in order to assimilate the different models to made possible to study the sensitivity observation impact and also to calibrate it. This technique of impact of the data assimilated for the model will suggest the areas that need more measures. Also, is important to develop techniques to take advantage of the data we already have, improving the measures that sensors are taking not only thinking on installing more sensor to optimize the data available.

3. THEORETICAL FRAMEWORK

3.1 Air pollutants

Air pollutants are substances that, present in sufficient concentration, tend to interfere with the human comfort and cause environmental damage. The main pollutants under attention for AQ are divided in primary and secondary. Primary are pollutants that pass into environment in the form they are produced for example Nitrogen Oxides (NO_X =NO+NO_2), Sulfur Dioxide (SO_2), Particulate Matter (PM), Volatile Organic Compound (VOC) etc., and Secondary are pollutants that develop as a result of interaction of primary pollutants and enviromental constituents like for example Ozone (O_3), Nitrogen Dioxide (NO_2), Nitrate (NO_3) and Ammonia (NH_3).

One example of the nitrogenous pollutants that we are interested is the Ammonia. It is emitted from different sources like traffic, fertilizer applications and animal farm production, The major sources for atmospheric NH_3 are agricultural activities and livestock farming, followed by biomass burning (including forest fires) and to a lesser extent fossil fuel combustion (Krupa, 2003). Its role in acidification, eutrophication and its impact in ecosystem and water quality is well documented (Erisman & Schaap, 2004). Modelling the behavior of ammonia from measurements only from the Planet Boundary Layer (PBL) is challenging because capturing the temporal and spatial variability from such sparse ground-based measurements alone is difficult because ammonia gas phase is very reactive with a tropospheric lifetime no more than a few hours (Clarisse et al., 2010).

Another compound of interested is the Ozone. The Ozone depositions affect the vegetation and ecosystems causing damage during the growing seasons. Plants exposed to sufficient ozone can reduce their photosynthetic rate, slowing down its growth and increasing the plant risk of disease and the probability of damage from insects and other problems (“EPA United States Environmental Protection Agency,” n.d.). The ecosystems experience a negative impact from the effects of ozone on individual plants for example the loss of biodiversity, changes to the specific assortment of plants present in a forest, changes to water and nutrient cycles (“EPA United States Environmental Protection Agency,” n.d.).

3.2 Atmospheric Chemistry Transport Models

Atmospheric chemistry models simulate concentration, fluxes, production/loss chemical reactions and deposition of several components. These models receive as inputs data of land use, emission inventories, meteorology, orography and boundary conditions and, with this, solve numerical equation in discretized domains to know the concentration of some species of interest in that coordinate. Depending of the area in which work they can be classified in global, continental, regional and local models and, also if the model is solved for a stationary parcel or for a parcel that “travel” with the flux of the atmosphere, it is classified in Eulerian and Lagrangian respectively.

We are working with LOTOS EUROS model which is a 3D regional Eulerian chemistry transport model which simulates the air pollution in the lower troposphere over Europe and is used for AQ forecasts. The model is designed for intermediate complexity, to favor short computation times, the vertical top is limited up to 3.5 Km and the maximum resolution possible is 9 km x 7 km.

Meteorological grid models are used in conjunction with chemical interaction models to provide gridded output of chemical species and pollutants. Meteorological grid models use mathematical formulations to simulate atmospheric processes such as the change of temperature in time and the change of winds (“EPA United States Environmental Protection Agency,” n.d.) that works as input for the chemical transport models.

For the LOTOS EUROS the main prognostic equation is the continuity equation that describes the change in time of the concentration of the components as a result of the processes of transport (Advection and Diffusion), chemistry, dry and wet deposition and emissions

The equation is given by:

$$\frac{\partial C}{\partial t} + U\frac{\partial C}{\partial x} + V\frac{\partial C}{\partial y} + W\frac{\partial C}{\partial z} =$$

$$\frac{\partial}{\partial x}\left(K_h\frac{\partial C}{\partial x}\right) + \frac{\partial}{\partial y}\left(K_h\frac{\partial C}{\partial y}\right) + \frac{\partial}{\partial z}\left(K_z\frac{\partial C}{\partial z}\right) + E + R + Q - D - W \qquad (1)$$

with C the concentration of a pollutant, U, V and W being the large-scale wind components in west-east, north south and vertical direction respectively. K_h and K_z are the horizontal and vertical turbulent diffusion coefficients. E represents the entrainment and detrainment due to variations in layer height, R gives the amount of material produced or destroyed as a result of chemistry, Q is the constribution by emissions, D and W are loss terms due to processes of dry and wet deposition respectively (Manders-Groot et al., 2016). In the paper of Yanlong et al. (Jia et al., 2016) the wet and dry deposition is explained for the case of the global inorganic nitrogen, the decrease of concentration due to the deposition is

$$\frac{\partial C}{\partial t} = -\Lambda C \quad (2)$$

where the term Λ correspond to a sweep coefficient of the speed of the mass transference of a pollutant from the air to the raindrops. The contribution to the flow of wet deposition is:

$$\Delta D = C_0(1 - e^{-\Lambda t})\Delta z \quad (3)$$

With C_0 the initial concentration and z the height of the cell in the resolution grid

It is common to find ensembled models, although each of these models can perform very well on particular days in particular areas, the ensemble approach aims at providing, on average, forecasts and analyses of better quality than any of them individually. At this time, we are working on a way to integrate the LOTOS-EUROS model with the Weather Research and Forecasting (WRF) model for the region of Colombia. The WRF is a next-generation mesoscale numerical weather prediction system designed for both, atmospheric research and operational forecast needs. This kind of ensemble was not yet reported for the proposed zone of interest to take advantage of the best of each model to complement it.

One example of assembled models is the case of the MACC projects. The MACC is an ensemble that integrates the next models: CHIMERE: (Eulerian Chemistry transport model), the EMEP (European Monitoring and Evaluation Program), the EURAD-IM (European Air Pollution Dispersion Inverse Model), the LOTOS-EUROS (Long Term Ozone Simulation), the MATCH (Multiscale Atmospheric Transport and Chemistry) and the MOCAGE SILAM (Model of Atmospheric Chemistry At large Scale, System for Integrated Modelling of Atmospheric Compositions.

for a case study of Ozone and Ammonia deposition will show the ability of the ensemble in forecast regional ozone pollution events prediction using the models which we are working to understand the effect of this deposition in nature protected areas in Colombia.

3.3 Data Assimilation

Data Assimilation is a methodology used to produce a regular, physically consistent, four-dimensional representation of the state of the atmosphere from an heterogeneous array of in-situ and remote instruments which sample imperfectly and irregularly in space and time (Daley, 1991). The ultimate target of a modeler is to have the residuals of the state of the system compared with the observational networks of sensors as small as possible. Under the name of data assimilation a variety of methods exist which all try to reach this target (Segers, 2002). This term refers to the fact that all methods try to merge model forecast and measurements using the benefit of both sources of information.

In different regions, different data assimilation projects are being developed. One example is the Air Quality Modelling International Initiative (AQMEII) (Solazzo et al., 2013), an operational evaluation of 12 regional-scale chemical transport models which was developed to model North America's and Europe's latitudes. Another one is the Monitoring Atmospheric Composition and Climate: Interim Implementation (MACC II) ensemble (Marécal et al., 2015).

There exist two classes of data assimilation techniques: Variational and Linear Filters methods

The Variational data assimilation methods are based on minimization of cost function using proper orthogonal decomposition (POD) adjoint method (Altaf, El Gharamti, Heemink, & Hoteit, 2013). Objective function is typically the sum of squared differences between the data and the corresponding model values. From this perspective exist basically two methods, the 3DVar and the 4DVar. 3DVar method uses a static, flow-independent, climatological background error covariance (BEC) that is often spatially homogeneous and anisotropic. The 4DVar method allows the fitting of the model forecast trajectory to observations distributed over a period of time so as to provide more accurate model state estimations that are also more consistent with the prediction model (Liu, Xue, Liu, & Xue, 2016).

The Ensemble Kalman filter is a linear adaptive filter, the state analyzed is a linear combination of the forecast state and the data elements where the analyzed state is adapted proportional to the residue. With a Kalman Filter is possible to make estimation to predict some response value and be seen as an extension of the Optimum Interpolator (OI) (Segers, 2002), accounting for the evolution of errors from previous time.

An important difference between 4DVar and Kalman filtering is the form of the final result. 4DVar provides an assimilated result in the form of piece-wise model evaluation, with discontinuities at the evaluation interval; the Kalman filters provides the result in terms of mean and covariance (Segers, 2002).

The LOTOS-EUROS model is equipped with a data assimilation package with the ensemble Kalman filter technique and uses a library developed by the company Deltares which is named http://openda.org/. Kalman and variationally methods are suitable to be used in online forecast applications, for offline applications such a parameter estimation, the variation approach is often favored due to its clear insight into how parameters are optimized, by comparison of model forecast based on certain parameter values measured. Compared to the 4DVar method the Kalman Filter is in general quite simple to implement since only the forward model is in use (Verlaan & Sumihar, 2016).

3.4 Study of the impact of data in a model

In air quality application, the equation system that describes the evolution of trace gas concentration for several species in time is discretized according to:

$$x(k+1) = Mx(k) + w(k) \qquad (4)$$

where $x(.)$ is n-dimensional vector representing the state of the system at a given time where the elements are gas-phase concentrations, the state space operator M describes the time evolution from the time k to $k+1$ of the state vector. Unknown disturbances are represented by the vector $w(.)$ Is constructed to allow for model errors. This random term $w(.)$ is assumed to have a normal distribution $N(0, Q)$ with zero mean and Q the covariance matrix.

The state of observations $y(.)$ in (5) is defined by the observation operator H that maps state variables x to observations y and have an uncertainty $v(k)$ assumed to be a white Gaussian error with zero mean and covariance denoted by R. These variables $v(k)$ and $w(k)$ are what is known as noise in Kalman filtering literature and are considered unknown a priori

$$y(k) = Hx(k) + v(k) \qquad (5)$$

To study the impact of observations, an analysis step is added to (4) and (5). We denote the estimate for the state x at time k based upon observations until time l as $\hat{x}(k|l)$. Now a linear analysis update can be written as

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K[y(k) - H\hat{x}(k|k-1)] \qquad (6)$$

with the notation of the corresponding forecast changing to

$$\hat{x}(k+1|k) = M\hat{x}(k|k) \qquad (7)$$

where the left term denotes the estimate $\hat{x}$ for time $k+1$ based on observation up to and including time k. In the observation sensitivity experiments, one would like to study the impact of various set of observations on the accuracy of subsequent forecast. (Verlaan & Sumihar, 2016). A measure to study the impact based on observations of the form (Todling, 2012):

$$J(k,l,m) = \big(y(k+m) - H\hat{x}(k+m|l)\big)' R^{-1}\big(y(k+m) - H\hat{x}(k+m|l)\big) \qquad (8)$$

with k being the time to start the forecast, l is the time of the assimilated observation, and m is the forecast lead-time considered for validation, $\hat{x}$ denotes the estimate for the state x at some time based upon observation until time l. For measure the impact of the observations at the most recent analysis update to the analysis update, (6) is extended to

$$\hat{x}(k|k) = \hat{x}(k|k-1) + sK[y(k) - H\hat{x}(k|k-1)] \qquad (9)$$

with $0 \leq s \leq 1$. When $s = 0$, observations $y(k)$ are ignored and with $s = 1$ the observation are fully included. It could be understood as a tuning to regulate the amount of the observation taken into account. The corresponding cost function from (8) becomes

$$J_s(k,l,m) = \big(y(k+m) - H\hat{x_s}(k+m|l)\big)'R^{-1}\big(y(k+m) - H\hat{x}_s(k+m|l)\big) \qquad (10)$$

The impact of observations at time k can be written as $\triangle J(k,m) = J_1(k,m) - J_0(k,m)$, and this gradient is commonly approximated with a trapezoid rule:

$$\triangle J(k,m) = \int_0^1 \frac{dJ_s(k,m)}{ds} ds \approx 1/2 \left[\frac{dJ_s(k,m)}{dx}|_{s=0} + \frac{dJ_s(k,m)}{dx}|_{s=1}\right] \qquad (11)$$

A common approach to compute $\triangle J(k,m)$ is with an adjoint model (Daescu D., 2009; Langland RH., 2004) derived by noting that the functional $J_s(k,l,m)$ is a concatenation of three steps: Analysis, forecast, and evaluation of cost at forecast time. This calculus depending on what order of accurate of the trapezoid rule and with higher order it requires high computation. To find another expression to calculate it, Verlan (Verlaan & Sumihar, 2016) propose the estimates of observation sensitivity based on the Ensemble Kalman filter (EnKF) (Evensen, 2009; Gerrit., Jan van Leeuwen, & Evensen, 1998) computing the forecast error covariance by integrating an ensemble of randomly perturbated initial analysis states in time with random perturbation added to the forcing. It is possible to define a way to understand how the model is working in comparison with the real values with next expression that will be one of the focus for impact analysis

$$\Delta J(k,m) = [\big(y(k+m) - H\hat{x}(k+m|k)\big) + y(k+m) - H\hat{x}(k+m|k-1))]'R(k+m)^{-1}$$
$$D(k+m|k-1)D(k|k-1)'D(k|k-1)D(k|k-1)'D(k|k-1)' + R(k))^{-1}$$
$$(y(k) - H\hat{x}(k|k-1) \qquad (12)$$

where the forecast tangent can be approximated with an ensemble forecast being $D(k|l)$ the square root of the covariance, i.e.,

$$[D(k|l)]_{:,i} = \left(\frac{1}{\sqrt{q-1}}\right)\left(H\xi_i(k|l) - \left(\frac{1}{q}\right)\sum_{j=1}^{q} H\xi_j(k|l)\right) \qquad (13)$$

where $\xi_i(k|l)$ is an ensemble of state vectors generated with the random perturbation added to the force.

4. OBJECTIVES

4.1 General Objectives

Identify the areas in Colombia that need more than local conservation efforts using atmospheric chemical transport areas and methodologies combined with data assimilation techniques supported in observation sensitivity analysis impact of the satellite, ground sensors and different other sources of data available

4.2 Specific Objectives

- To implement diverse methodologies of Data Assimilation schemes for the LOTOS EUROS model over Colombia.
- To evaluate the effect of new localization strategies developed by Fu et. Al. over the observation sensitivity propose by (Verlaan and Sumihar, 2016).
- To apply techniques to retrieve more information of the observations that experience uncertainty due to the systematic error in order to optimize the available sources of data.
- To analyze and quantify the impact of the selection of the data subset to be used in a Data Assimilation scheme over the Chemistry Transport Models over Colombia.
- To propose different techniques to determine the most appropriate zones to make proper comparison between the capabilities of prediction of the model and the measures available.

5. METHODOLOGY

Different methodologies such as weather and atmospheric chemistry composition modelling and re-analysis techniques are used to understand the transport of contamination (Marécal et al., 2015). These models are powered by data that come from different satellite platforms (e.g. Infrared Atmospheric Sounding Interferometer (IASI) instrument on board the MetOp polar orbiting satellite, Tropospheric Emission Spectrometer TES (Van Damme et al., 2014) Ozone Monitoring Instrument (Jia et al., 2016)), multispectral or infrared spectrometer on ground based sensors (e.g. Fourier Transform Infrared FTIR Observations (Dammers et al., 2015)), UAV using Synthetic Aperture Radar (SAR) (Baxter & Bush, 2014), high altitude balloons (Basart et al., 2016), rocketsondes (Mount, 2010), and ground based platforms (e.g. LIDAR, Specific chemical deposition sensors (Fredriksson, Galle, Nyström, & Svanberg, 1979)).

Commercial aircraft routes have been involved in monitoring programs to get measurements in situ from instrumentation on board to feed several models. One example of in situ sensing from airplanes is the program MOZAIC which means *Measurements of Ozone, water vapor, carbon monoxide and nitrogen oxides by Airbus In-service aircraft* (Solazzo et al., 2013) started in 1993 as joint effort of european scientists, aircraft manufactures and airlines. Other references of works to register atmospheric data during regular long distance flights are: the CARIBIC program (Civil aircraft for global measurements of trace gases and aerosols in the tropopause region) (Ans & Einrich, 1999) , the JAL project (Foundation Japan Airlines) (Hidekazu Matsueda; Hisayuki Y. INOUE, 1996) and the NOXAR (*Measurement of Nitrogen Oxides and Ozone along air routes*) (Wernli, 2001). In Medellín in 2016, the professor Miguel Ángel Gordillo, from Universidad de Granada (Spain), flew over the city valley with a plane equipped with an aethalometer, instrument that detects particles of soot or black carbon in the air (Nacional, 2016).

For the detection of this components from space are often used sun synchronous polar orbits satellites to cover most of the area of the planet. Using this kind of platforms, with specific payloads dedicated to the mapping with multispectral or hyperspectral instruments, is possible to identify the reflectivity emission from the atmospheric compound of particular components of interest (Van Damme et al., 2014).

For ground measure, we have the SIATA project (https://siata.gov.co) that operates a network of sensors that monitor, among many other things, meteorological conditions and air quality parameters in the Aburrá Valley. SIATA currently has different network of sensors dedicated to specific purposes. There are: pluviometry, level, meteorological, real time and infrared camera, disdrometers, accelerograph, soil monitoring, electric field monitoring, pyranometer networks and remote sensors capabilities like the hydro meteorological radar, a radiometer, a ceilometer network and a vertical wind profiler. In the next table is possible to see a consolidated report of the sensing capabilities that SIATA already has in operation. Nowadays the total number of stations that SIATA has in the valley of different networks is 183 and outside 12.

The initial data of interest from SIATA data we are interested to make the analysis, are from the air quality network sensors. The sensor that SIATA use for measure the pollutants related to air quality is the PQ 200 monitor stations. This is a certified EPA instrument that have the possibility to measure PM 10, PM 2.5, PM coarse and PM 1.0 with a flow range of 10-20 LPM.

Other approach to stablish the criteria to ask for more data or run the model for obtain simulated concentrations in areas of all Colombia territory is to characterize in terms of diversity and available data some characteristics like the land use distribution, the availability of Ideam sensing capabilities and the population density. Once we characterize those, we can stablish the best areas from make comparisons and assimilation with the models, to proceed with the impacts estimates.

6. EXPECTED RESULTS

- Formation as a PhD in mathematical engineering
- At least one (1) scientific paper in an indexed journal
- 1 conference paper
- Four (4) report working papers
- Capabilities to estimate the optimal position to install on ground sensors that maximize the perform of the assimilated models over the region of Aburrá valley and natural zones of interest

7. REFERENCES

Altaf, M. U., El Gharamti, M., Heemink, A. W., & Hoteit, I. (2013). A reduced adjoint approach to variational data assimilation. *Computer Methods in Applied Mechanics and Engineering*, *254*, 1–13. https://doi.org/10.1016/j.cma.2012.10.003

Ans, W. H., & Einrich, G. H. (1999). CARIBIC — Civil Aircraft for Global Measurement of Trace Gases and Aerosols in the Tropopause Region, 1373–1383. https://doi.org/10.1175/1520-0426(1999)016<1373:Ccafgm>2.0.Co;2

Basart, S., Benedictow, A., Blechschmidt, A. M., Chabrillat, S., Clark, H., Christophe, Y., … Zerefos, C. (2016). Observations characterization and validation methods document. *Report of the Copernicus Atmosphere Monitoring Service, Validation Subproject (CAMS-84)* ., (March). Retrieved from https://atmosphere.copernicus.eu/sites/default/files/repository/CAMS84_2015SC1_D.84.8.1-2016Q1_201603.pdf

Baxter, R. A., & Bush, D. H. (2014). Use of Small Unmanned Aerial Vehicles for Air Quality and Meteorological Measurements. *T&B Systems, Inc, Valencia California*, 19.

Clarisse, L., Shephard, M. W., Dentener, F., Hurtmans, D., Cady-Pereira, K., Karagulian, F., … Coheur, P. F. (2010). Satellite monitoring of ammonia: A case study of the San Joaquin Valley. *Journal of Geophysical Research-Atmospheres*, *115*. https://doi.org/10.1029/2009jd013291

Daescu D., T. R. (2009). Adjoint estimation of the variation in model functional output due to the assimilation of data. *Mon. Weather. Rev.*, *137*, 1705–1716.

Daley, R. (1991). *Atmospheric Data Analysis* (Cambridge).

Dammers, E., Vigouroux, C., Palm, M., Mahieu, E., Warneke, T., Smale, D., … Erisman, J. W. (2015). Retrieval of ammonia from ground-based FTIR solar spectra. *Atmospheric Chemistry and Physics*, *15*(22), 12789–12803. https://doi.org/10.5194/acp-15-12789-2015

Dillon, M. E., Skabar, Y. G., Ruiz, J., Kalnay, E., Collini, E. a., Echevarría, P., … Kunii, M. (2015). Application of the WRF-LETKF Data Assimilation System over Southern South America: Sensitivity to model physics. *Weather and Forecasting*, 151111130547009. https://doi.org/10.1175/WAF-D-14-00157.1

EPA United States Environmental Protection Agency. (n.d.). Retrieved from https://www3.epa.gov/scram001/metmodel.htm

Erisman, J. W., & Schaap, M. (2004). The need for ammonia abatement with respect to secondary PM reductions in Europe. *Environmental Pollution*, *129*(1), 159–163. https://doi.org/10.1016/j.envpol.2003.08.042

Evensen, G. (2009). *Data assimilation: The Ensemble Kalman filter*. Bergen, Norway. https://doi.org/10.1007/978-3-642-03711-5

Fredriksson, K., Galle, B., Nyström, K., & Svanberg, S. (1979). Lidar system applied in atmospheric pollution monitoring. *Applied Optics*, *18*(17), 2998–3003. https://doi.org/10.1364/AO.18.002998

Gerrit., B., Jan van Leeuwen, P., & Evensen, G. (1998). Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, *126*(6), 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2

Hidekazu Matsueda; Hisayuki Y. INOUE. (1996). Measurement of atmospheric CO2 and CH4 using a commercial airliner from 1993 to 1994. *Atmospheric Environment*, *30*, 1647–1655.

Jia, Y., Yu, G., Gao, Y., He, N., Wang, Q., Jiao, C., & Zuo, Y. (2016). Global inorganic nitrogen dry deposition inferred from ground- and space-based measurements. *Scientific Reports*, *6*, 19810. https://doi.org/10.1038/srep19810

Krupa, S. V. (2003). Effects of atmospheric ammonia (NH3) on terrestrial vegetation: A review. *Environmental Pollution*, *124*(2), 179–221. https://doi.org/10.1016/S0269-7491(02)00434-7

Krymskaya, M. V. (2013). *Quantification of the impact of data in reservoir modeling*. TuDelft.

Langland RH., B. N. (2004). Estimation of observation impact of real observation in regional numerical weather prediction using an ensemble Kalman filter. *Tellus*, *56A*, 189–201.

Liu, C., Xue, M., Liu, C., & Xue, M. (2016). Relationships among Four-Dimensional Hybrid Ensemble–Variational Data Assimilation Algorithms with Full and Approximate Ensemble Covariance Localization. *Monthly Weather Review*, *144*(2), 591–606. https://doi.org/10.1175/MWR-D-15-0203.1

Manders-Groot, A. M. M., Segers, A. J., Jonkers, S., Schaap, M., Timmermans, R., Hendriks, C., … Banzhaf, S. (2016). LOTOS-EUROS v2.0 Reference Guide. TNO 2016 R10. *TNO Innovation for Life*.

Marécal, V., Peuch, V. H., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., … Ung, A. (2015). A regional air quality forecasting system over Europe: The MACC-II daily ensemble production. *Geoscientific Model Development*, *8*(9), 2777–2813. https://doi.org/10.5194/gmd-8-2777-2015

Mount, G. H. (2010). The Measurement of Air Pollution from Space What Pollutants Would You Want to Measure ? *ESRP 285*, (April).

Nacional, A. de N. U. (2016). Desde avión miden partículas contaminantes en el aire de Medellin. Retrieved from http://agenciadenoticias.unal.edu.co/detalle/article/desde-avion-miden-particulas-contaminantes-en-el-aire-de-medellin.html

Nicolas Pinel. (2017). Potential Urban Pollution impacts on protected areas in Colombia through atmospheric teleconnections. In *CMAS Community Modeling and Analysis system*.

Segers, A. (2002). *Data assimilation in atmospheric chemistry models using Kalman filtering*. The Netherlands: Delft University Press.

Solazzo, E., Bianconi, R., Pirovano, G., Moran, M. D., Vautard, R., Hogrefe, C., … Galmarini, S. (2013). Evaluating the capability of regional-scale air quality models to capture the vertical distribution of pollutants. *Geoscientific Model Development*, *6*(3), 791–818. https://doi.org/10.5194/gmd-6-791-2013

Todling, R. (2012). Comparing Two Approaches for Assessing Observation Impact. *Monthly Weather Review*, *141*(Modeling, Global Office, Assimilation), 1484–1505. https://doi.org/10.1175/MWR-D-12-00100.1

Van Damme, M., Clarisse, L., Heald, C. L., Hurtmans, D., Ngadi, Y., Clerbaux, C., … Coheur, P. F. (2014). Global distributions, time series and error characterization of atmospheric ammonia (NH3) from IASI satellite observations. *Atmospheric Chemistry and Physics*, *14*(6), 2905–2922. https://doi.org/10.5194/acp-14-2905-2014

Verlaan, M., & Sumihar, J. (2016). Observation impact analysis methods for storm surge forecasting systems. *Ocean Dynamics*, *66*(2), 221–241. https://doi.org/10.1007/s10236-015-0912-0

Wernli, D. B. J. S. D. J. H. (2001). Nitrogen oxides and ozone in the tropopause region of the northern hemisphere: measurenments from commercial aircraft in 1995/1996 and 1997. *Journal of Geophysical Research-Atmospheres*, *106*, 673–699.