SOURCE DETECTION OF ATMOSPHERIC RELEASES USING SYMBOLIC MACHINE LEARNING CLASSIFICATION AND REMOTE SENSING

Mark C. Bowman, Guido Cervone, Dr. Pascale Franzese

George Mason University, Fairfax, VA 22030

ABSTRACT

This paper introduces the National Polar-orbiting Operational Environmental Satellite System (NPOESS) and its use for the identification of the source of atmospheric pollutants. NPOESS is the next generation satellite program, and can be used for the source detection of atmospheric pollutants. The iterative methodology proposed herein uses a combination of ground measurements, atmospheric models, machine learning and remote sensing to identify the characteristics of an unknown atmospheric emission.

Index Terms— Machine Learning, NPOESS, Aerosol, AQ

1. INTRODUCTION

When an airborne toxic contaminant is released into the atmosphere it is rapidly transported by the wind and dispersed by atmospheric turbulence. Contaminant clouds can travel thousands of kilometers within a few days, and spread over areas measuring thousands of square kilometers. A large population can be affected with serious and long term consequences depending on the nature of the hazardous material released. Potential atmospheric hazards include toxic industrial chemical spills, forest fires, intentional or accidental releases of chemical and biological agents, nuclear power plants accidents, and radiological material. Risk assessment of contamination from a known source can be computed by performing multiple forward numerical simulations for different meteorological conditions, and by analyzing the simulated contaminant clouds with clustering and classification algorithms to identify the areas with highest risk [3]. However, often the source is unknown, and must be identified from limited anomalous concentration measurements observed on the ground. There are currently no established methodologies for the satisfactory solution for detecting the sources of atmospheric releases, and there is a great degree of uncertainty towards the effectiveness and applicability of existing techniques.

Several approaches have been developed to identify the source characteristics of an atmospheric emission using ground sensors, numerical models and machine learning algorithms Cervone et al. [3]; Haupt [7]; Haupt et al. [8]; Allen et al. [1]; Delle Monache et al. [6]. All these methods use an iterative process based on machine learning search algorithms to find the characteristics of unknown sources, then performed multiple forward simulations from tentative source locations and use the comparison of simulated concentration with sensor measurements to implement an iterative process that converged to the real source. The strength of the approach lies in the use of proven methodologies that are domain independent, and have been shown to scale well with the increase of the problem complexity.

All these methods are based on the comparison of concentrations from forward transport and dispersion simulations with concentrations from ground sensors. However, in most cases concentrations from ground sensors are not available. Using data from the NPOESS it may be possible to apply the aforementioned source detection methodologies to a greater variety of emission problems, due to the availability of global information of atmospheric contaminants.

2. NPOESS

The National Polar-orbiting Operational Environmental Satellite System (NPOESS) is the next generation of environmental satellites to study the Earth and its climate. NPOESS is a constellation of low orbiting polar satellites capable of achieving near global coverage and is supplemented with a next generation Geosynchronous Orbiting Environmental Satellite (GOES) capability. NPOESS' primary mission is the observation of Earth's land, ocean, and atmosphere, including the troposphere, stratosphere and the near space boundary.

This system, due to come online in approximately 2013, will replace the United States' current constellation of environmental and weather satellites and revolutionize the way remote sensing is used in natural hazard monitoring and response efforts. NPOESS will allow simultaneous or near-simultaneous collection of atmospheric phenomena by multiple sensors. This capability will provide increased understanding of atmospheric phenomena by reducing time between collections of data from multiple sensors and by providing a common platform for data distribution and analysis. The new sensor types flown in NPOESS will also increase collection capabilities. As data from NPOESS becomes available, the environmental monitoring community will be inundated with data which must quickly be assessed and placed into appropriate computer models for analysis. The shear volume expected from this System will quickly overwhelm the majority of conventional processes for analyzing data and will require new applications for the timely reporting of atmospheric phenomena.

The NPOESS' capabilities will significantly increase understanding and observation of atmospheric phenomena by reducing time between collections of data from multiple sensors and providing a common platform for data distribution and analysis. The new sensor types flown in the NPOESS will increase both collection capabilities and data volume, and the transmission advances of the SafetyNet system will significantly decrease data latency. The utility of the Visible Infrared Imaging Radiometer Suite (VIIRS) and the Advanced Technology Microwave Sounder (ATMS) which is co-located with the Cross Track – Infrared Sounder (CrIS) make them ideal candidates for inclusion in a discussion of NPOESS data analysis using machine learning.

2.1 VIIRS

The Visible Infrared Imaging Radiometer Suite (VIIRS) is perhaps the most effective tool within the suite of sensors on the NPOESS constellation. VIIRS is a planned follow-on to the very successful MODIS sensor flown on the AQUA and TERRA spacecrafts. It will be available on three of the NPOESS satellites and will provide an unprecedented three looks daily over the Earth. VIIRS products will provide a spatial reference map for comparison of the atmospheric environmental sensors co-located upon the NPOESS satellites for the immediate production of advanced data products comparing multiple layers of information.

VIIRS, built by Raytheon Inc. is a heritage replacement system. It provides the radiometric accuracy of the Advanced Very High Resolution Radiometer (AVHRR) with the spatial accuracy of the Operational Linescan System (OLS) of 0.65 km. VIIRS also provides an enhancement of spatial, spectral and radiometric resolution over the heritage sensors of MODIS and AVHRR. VIIRS has visible and infrared imagery ranging capabilities from 0.38km at nadir for the fine resolution "I" bands to 0.78 km at nadir for the multispectral bands and features 21 spectral bands which will provide information on landforms, sea temperatures, atmospheric gases and cloud formations.

2.2 ATMS and CrIS

The use of passive microwave remote sensing in monitoring hazards is perhaps one of the most promising aspects of the planned NPOESS constellation. Microwave sensors receive and measure emissions from the Earth's surface and its atmosphere. These emissions are then characterized by the imaging system in order to better define the current conditions of the Earth.

The NPOESS constellation will provide passive microwave scanning via the Advanced Technology Microwave Sounder (ATMS) which is co-located with the Cross Track – Infrared Sounder (CrIS) on two NPOESS vehicles. Together these two sensors are the Cross Track Infrared and Microwave Sounding Sensors (CrIMSS). Another NPOESS vehicle will include a microwave imaging sounder (MIS) which is being developed as a low cost alternative the ATMS/CrIS.

ATMS is a 22 channel sensor that collects frequencies between 23.8 and 183 GHz. The CrIS will produce operational atmospheric temperature, moisture and pressure profiles using high vertical resolution. It operates on 3 infrared bands:

- *Long wave Infrared*- LWIR images between 650-1095 cm^{-1}
- Midwave Infrared- MIDWIR images 1210-1750 cm^{-1}
- *Shortwave Infrared*- SWIR images 2155-2550 cm^{-1}

2.3 SafetyNet

SafetyNet is the data collection system specially designed for NPOESS. Coinciding with the anticipated 2013 launch of the initial NPOESS satellite, SafetyNet, developed by Raytheon, Inc., will come online. Fifteen data reception centers located in ten countries are the backbone of the SafteyNet system. These 15 centers provide nearly continuous collection of NPOESS sensor data and are linked by dedicated fiber optic cable to the National Data Exploitation (NDE) centers.

These NDEs will stand to gain the most benefit from the use of appropriate machine learning algorithms. Currently, data from U.S. Polar Orbiting Environmental Satellite (POES) is downloaded in bursts to a ground station and funneled to the appropriate agency for processing. This is a time-consuming, inefficient process measured in hours. The NPOESS' SafetyNet system will distribute five times as much environmental data in 30 minutes- approximately ¼ of the time required by the POES. This significant decrease in data latency necessitates the use of machine learning to process the data collected by NPOESS.

One area of research which stands to gain from NPOESS' increases in data collection and decreases in data latency is source detection of atmospheric releases. Currently, there are no established methodologies for determining the source of atmospheric releases from space, and there is a great degree of uncertainty towards the effectiveness and applicability of existing techniques. At present, the nature and source of atmospheric hazard releases must be inferred from the anomalous levels of contaminant concentration measured by sensors on the ground.

3. SEARCH BASED ON MACHINE LEARNING

Iterative machine learning algorithms based on machine learning classifiers, such as genetic algorithms, Bayesian inference, Monte Carlos simulations, Markov Chain and symbolic rule induction have been developed to identify the characteristic of atmospheric pollutant sources. All of these methods are based on an iterative approach which tries to minimize the error between a set of ground measurements and numerical transport and dispersion simulations.

These methods start by generating a set of candidate sources at random within the boundaries of the search space, and generate transport and dispersion simulations from these sources. Each candidate solution is evaluated according to an error function, which quantifies how well the measured concentrations match the simulated values.

Different methods employ different techniques to continue the search for the best solution. In genetic algorithms, the best solutions are chosen and modified according to pseudo-random operators inspired by biological evolution [7,8]. In Bayesian inference, the probabilities are updated according to the result of the error functions [6]. In non-Darwinian evolution, the best solutions are modified according to a reasoning process which determines the attribute combinations that are discriminates with respect to the worst performing.

The advantage of such methodologies consists in the fact that although they were developed for ground measurements, they can easily be extended to the use of remotely sensed data. In fact, at each error function evaluation, the result from the atmospheric transport and dispersion simulation can be compared against the remotely sensed data, instead of the ground measurements. New error functions will have to be investigated to determine how to best compare simulations with observations, as the remote sensing data might be noisy, contain missing data, or have an incorrect temporal coverage.

4. CONCLUSIONS

The increases in data collection capability and transmission speed inherent in the NPOESS will necessitate new methodologies for data processing if the collected data is to be exploited to any degree of its potential. Traditional human-reliant analytical processes are too time consuming and labor intensive to handle the increase in data presented by the NPOESS. Machine learning algorithms are the only way the volume of data in the near future will be exploitable.

The next step in this process is to take actual data from VIIRS and ATMS/CrISS and integrate the different search methodologies to identify the best candidate. Then compare the results to those from data collected by ground sensors and determine if it is possible to use spaceborne collection to determine sources for aerosol source detection.

5. REFERENCES

[1] Allen, C. T., Young, G. S., Haupt, S. E., 2007. “Improving pollutant source characterization by better estimating wind direction with a genetic algorithm.” Atmos. Environ. 41 (11), 2283–2289.

[2] Bowman, M., Cervone G., 2009 “The Next Generation of Remote Sensing For Natural Hazard and Environmental Monitoring: National Polar-orbiting Operational Environmental Satellite System (NPOESS).” Proceedings of the International Symposium on the Remote Sensing of the Environment, Stresa, Italy 1 – 5 May 2009.

[3] Cervone, G., Franzese, P., Keesee, A. P. K., 2009, AQ Learning, An Application for the Source Detection of Atmospheric Emissions, WIREs, Computational Statistics, to appear

[4] Cervone, G., Franzese, P., Ezber, Y., Boybeyi, Z., 2008. “Risk assessment of atmospheric emissions using machine learning.” Natural Hazards and Earth System Science 8, 991–1000.

[5] Clinton, President William J., May 5, 1994. Presidential Decision Directive/NSTC-2 “Convergence of U.S.-Polar-Orbiting Operation Environmental Satellite Systems”, Washington DC. http://www.ipo.noaa.gov/About/NSTC-2.html

[6] Delle Monache, L., Lundquistand, J., Kosovi´c, B., Johannesson, G., Dyer, K., Aines, R., Chow, F., Belles, R., Hanley, W., Larsen, S., Loosmore, G., Nitao, J., Sugiyama, G., Vogt, P., 2008. “Bayesian inference and markov chain monte carlo sampling to reconstruct a contaminant source on a continental scale.” J. Appl. Meteor. Climatol.

47, 2600–2613.

[7] Haupt, S. E., Dec. 2005. “A demonstration of coupled receptor/ dispersion modeling with a genetic algorithm.” Atmos. Environ. 39 (37), 7181–7189.

[8] Haupt, S. E., Young, G. S., Allen, C. T., August 2007. “A genetic algorithm method to assimilate sensor data for a toxic contaminant release.” Journal of Computers 2 (6), 85–93.

[9] S. Mango, "The Roles of NPOESS in the Next Generations: Potential Contributions to GEOSS Societal Benefit Areas," NPOESS Integrated Program Office. Fairfax, VA, briefing, February 20, 2008.

[10] NASA, 2008. "Fiscal Year 2009 Budget Estimates United States National Aeronautics and Space Administration,” Washington DC. http:// www.nasa.gov

[11] “Ensuring the Climate Record from the NPOESS and GOES-R Spacecraft: Elements of a Strategy to Recover Measurement Capabilities Lost in Program Restructuring.” National Research Council of the National Academies. Washington DC., p.p. 1-4, 2008.

[12] United States Congressional Record. “Acquisition Decision Memorandum, NPOESS System Review 5 June 2006,” United States General Printing Office. Washington DC, 2006.