An Objectively Optimized Earth Observing System

David J. Lary, Oleg Aulov, Andrew Rickert
UMBC/JCET and NASA/GSFC
Code 610.3, Greenbelt, MD 20771, USA
Email: David.J.Lary@nasa.gov

Abstract—This paper describes one vision for future earth observing systems. New in this vision is the desire for symbiotic communication to dynamically guide an earth observation system. An earth observation system which is not just a single satellite acting on its own but a constellation of satellites, and sub-orbital platforms such as unmanned aerial vehicles, and ground observations interacting with computer systems used for modeling, data analysis and dynamic observation guidance. An autonomous Objectively Optimized Observation Direction System (OOODS) is of great utility for earth observation. In particular, to have a fleet of smart assets that can be reconfigured based on the changing needs of science and technology. The OOODS integrates a modeling and assimilation system within the sensor web allowing the autonomous scheduling of the chosen assets and the autonomous provision of analyses to users. The OOODS operates on generic principles that could easily be used in configurations other than the specific examples described here. Metrics of what we do not know (state vector uncertainty) are used to define what we need to measure and the required mode, time and location of the observations, i.e. to define in real time the observing system targets. Metrics of how important it is to know this information (information content) are used to assign a priority to each observation. The metrics are passed in real time to the sensor web observation scheduler to implement the observation plan for the next observing cycle. The same system could also be used to reduce the cost and development time in an Observation Sensitivity Simulation Experiment (OSSE) mode for the optimum development of the next generation of space and ground-based observing systems.

TABLE OF CONTENTS

1. INTRODUCTION

2004 was the fortieth anniversary of the NASA Nimbus program. The Nimbus satellites, first launched in 1964, carried a number of instruments: microwave radiometers, atmospheric sounders, ozone mappers, the Coastal Zone Color Scanner (CZCS), infrared radiometers, etc. Nimbus-7, the last in the series, provided significant global data on sea-ice coverage, atmospheric temperature, atmospheric chemistry (i.e. ozone distribution), the Earth's radiation budget, and sea-surface temperature.

What will the earth observing systems of the future look like? Autonomy is likely to be a key feature.

2. AUTOMATION ON MANY LEVELS

It is very likely that the observing systems of the future will increasingly involve the integration of models and data assimilation systems. Automation will be of great value in both the direction of observations and for many parts of the associated software systems, especially if the observing and analysis systems are to be dynamically reconfigurable.

Autonomous Observation Direction Systems

A prime area for the application of automation is in the autonomous direction of observations. However, to include autonomy, objective metrics to direct the system are required. The modeling and assimilation system engineering diagram for our OOODS system is shown in Figure 1. It is desirable if generic classes of metrics could be used so that the approach could be easily applied to many different areas of earth observation. For example, an autonomous Objectively Optimized Observation Direction System (OOODS) could use two specific metrics to perform the optimization for a given sensor web capability. Firstly, metrics of what we do not know (state vector uncertainty) can be used to define what we need to measure and the required mode (i.e. global survey, rapid scan, step-and-stare or zoom in), and the time and location of the observations, i.e. to define in real time the observing system targets. The state vector uncertainty is provided by

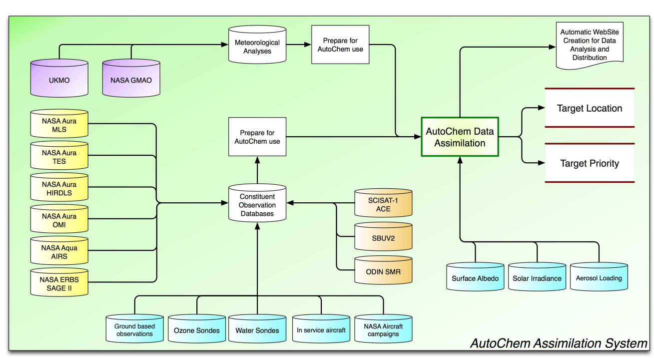

Figure 1. The modeling and assimilation system engineering diagram for OOODS.

IEEEAC paper # 1007

the integrated data assimilation system. Secondly, metrics of how important it is to make these observations (information content) [?] are used to assign a priority to each observation. The information content is also provided by the integrated assimilation system. The calculation of both of these metrics is described below.

These two metrics are then passed in real time to the smart sensor web observation scheduler that is aware of each assets observing capabilities to implement the observation plan for the next observing cycle. The sensor web will typically involve a suite of orbital, sub-orbital, aerial and ground based assets. The optimum observation schedule information will depend on the asset. For a satellite instrument it will typically include, pointing information, viewing mode, and micro window selection for every part of the next observation cycle. For unmanned aerial vehicles or aircraft missions it would be an optimal flight plan (i.e. time and route). For balloon launches it would be optimal launch time and location.

As an aside, it is worth noting that the same system could also be used to reduce the cost and development time in an Observation Sensitivity Simulation Experiment (OSSE) mode for the optimum development of the next generation of space and ground-based observing systems.

Parallelization— The shear scale of the task in creating an autonomous observation direction system has led to the extensive use of parallelization via the Message Passing Interface (MPI). MPI is a computer software standard that allows many computers to communicate with one another. It is used in computer clusters. In our case, we have used the free MPICH2 high performance and widely portable implementation of the MPI-2 standard on a cluster of Mac OS X machines. MPI was used for both the creation of the massively parallel modeling and assimilation system and the data queries. The large data volumes involved have led us to use a set of automatically synchronized databases which are queried using massively parallel queries.

Data Biases— The system we have outlined will typically be fusing data from many sources. In such a situation biases are ubiquitous. When combining observations from many sensors over a long time period biases will always be an issue. If they are not dealt with they can hinder us addressing the scientific issues the measurements were taken to address. Two companion studies have shown how this issue can be elegantly dealt with using neural networks [?], [?]. [?] was chosen as a NASA Aura mission science highlight.

A complimentary study— to ours [?] describes an adaptive cyberinfrastructure for real-time multiscale weather forecasting. Currently, scientists generate today's forecasts on a fixed time schedule. However, [?] point out that new radar technologies and improved model physics are enabling on- demand forecasts in response to current weather events. These forecasts ingest regional atmospheric data in real time and can consume large computational resources in real time as well. Two highly complementary projects are developing a hardware and software framework to enable real- time multiscale forecasting. Collaborative Adaptive Sensing of the Atmosphere and Linked Environments for Atmospheric Discovery are stand-alone systems that offer distinct benefits to their respective user communities, but when used together, promise a paradigm shift in atmospheric science research.

Automatic Code Generation

If the observing system is to be dynamically reconfigurable it is of great use if a high level of automatic code generation is used in the creation of the model and assimilation system that will be providing the objective measures used by the OOODS just described above.

An example of a fully automated code generation and documentation system that provides this information for atmospheric chemistry is NASA's AutoChem automatic code generation (e.g. `www.AutoChem.info`). AutoChem is an automatic code generator and documentor for atmospheric chemistry modeling and assimilation [?], [?]. Given a set of reaction databases and a user supplied list of required species it will automatically select the reactions involving those constituents. It then constructs the ordinary differential equation (ODE) time derivatives, symbolically differentiates the time derivatives to give the Jacobian, and symbolically differentiates the Jacobian to give the Hessian and the adjoint. It also documents the whole process in a set of LaTeX and PDF files. In addition, a huge number of observations of many different constituents from a host of platforms are available from this site in an atmospheric chemistry observational database.

AutoChem typically creates in less than a second the modeling and assimilation system that would take approximately a man year to write by hand. Once the model and assimilation system has been run AutoChem also automatically creates a cross linked web site for analysis and data mining (e.g. `www.CDACentral.info`). The automatic creation of web sites for data mining of the analyses greatly facilitates the scientific analysis needed to understand and answer major scientific questions, and can be used by policy makers to establish sound policy decisions, thus increasing the accessibility and utility of Earth Science data.

AutoChem is being used in the validation and analysis of results from the NASA Aura platform (e.g. `aura.gsfc.nasa.gov`).

Machine Learning

The whole approach described depends in large part on the integration of a data assimilation system. When considering data assimilation of atmospheric chemistry, one of the computationally most expensive tasks is the time integration of a large and stiff set of ordinary differential equations (ODEs). However, very similar sets of ODEs are solved at adjacent grid points and on successive days, so similar calculations are repeated many thousands of times. This is the type of application that benefits from adaptive, error monitored, machine-learning technology. Our ODE solver already employs adaptive time stepping with error monitoring, if this is extended to an adaptive use of machine learning then there are literally massive potential savings in computational expense. A prototype code has been developed that we would like to extend here for use within the ODE solver. Early work seems promising that such an approach would work [**?**], [**?**]. A success in this area would mean a dramatic reduction in the computational cost of assimilation and hence of the entire dynamic data retrieval control system.

Other Areas of Automation

Automatic parallelization will greatly facilitate the implementation and automatic adaption of the system for different problems and its possible use on a variety of hardware. Automatic documentation of both software and data products facilitate both code maintenance, and the production and quality monitoring of self-consistent analyses. The use of automatic compression can minimize both the required cost of storage and dissemination, and the required time for electronic product transfer/download.

3. Relevancy Scenarios

We consider two relevancy scenarios, one for immediate application, and the other for future systems currently being designed. However, before considering these scenarios it is worth noting that GOES-R and all planned geostationary platforms of other agencies such as Eumetsat and NASDA have an optional rapid scan mode. This enables the assets to scan a limited region (e.g. of a 1000 km x 1000 km) every minute if required. Knowing when best to use this rapid scan mode will be an issue for all these platforms. The methodology described here could help autonomously answer this question.

Current Scenario

A practical issue that faces the ongoing long-term NASA Aura validation effort is deciding the optimum validation balloon launch times. The OOODS described here can ingest the suite of observations made by NASA Aura and other platforms and produce assimilated constituent analyses. The state vector uncertainty of the analyses will then be used to define target regions of large uncertainty. The relative priority of the different target regions will then be determined using the information content fields derived from the assimilated analyses. Then by considering the Aura overpasses in the next 24 hours the best launch times and locations will be determined. It will then automatically send a set of emails to the balloon launch teams at these sites giving optimum launch times. It could also provide optimal flight plans for any UAV and aircraft missions involved in the validation.

Future Scenario

The requirements for the next generation of earth observing system for air quality are currently being discussed by NASA and NOAA. Key issues for this observing system will be what are the spatial scales on which observations are required, what are the most important constituents to observe and how does this change spatially and temporally, what are the optimum observation times for each constituent, and when should instrument zoom in, step and stare, rapid scan or global survey modes be used. The OOODS described here will be of great utility in autonomously addressing all of these issues. In this scenario, there is a daily OOODS cycle. As in the scenario above, the OOODS will ingest the full suite of relevant sensor web observations made by NASA and other platforms observing constituents, aerosols, surface reflectivity and cloud properties. These will be used to produce assimilated constituent analyses. The state vector uncertainty of the analyses will then be used to define target regions of large uncertainty. The relative priority of the different target regions will then be determined using the information content fields derived from the assimilated analyses. The metrics are then passed in real time to the system observation scheduler. The scheduler will then be able to do the following tasks. Upload to the satellite instruments involved their observing mode, pointing information, and (if required) micro window selections for the next 24 hours. Dispatch the flight plans to any unmanned aerial vehicles involved. Send emails to sonde and balloon launch teams giving optimum launch times.

The OOODS components and simulator just described would also be of use in the context of Observation Sensitivity Simulation Experiments (OSSE). A NASA OSSE capability is currently being developed by the NASA Research and Analysis program to determine the optimum configuration of the next generation of space and ground-based observing systems.

4. Conclusion

A vision for future earth observing systems has been described where there is symbiotic communication to dynamically guide an earth observation system. Where the earth observing system is a constellation of satellites, and sub-orbital platforms such as unmanned aerial vehicles, and ground observations interacting with computer systems used for modeling, data analysis and dynamic observation guidance. The earth observing system includes an autonomous Objectively Optimized Observation Direction System that use metrics of what we do not know (state vector uncertainty) to define what we need to measure, and metrics of how important it is to know this information (information content) to assign a priority to each observation. The metrics are passed in real time to the sensor web observation scheduler to implement the observation plan for the next observing cycle. The same system automatically creates cross-linked web sites for data mining and analysis.

The same system could also be used to reduce the cost and development time in an Observation Sensitivity Simulation Experiment (OSSE) mode for the optimum development of the next generation of space and ground-based observing systems.

More details can also be found in the invited Royal Society Vision article [**?**].

5. Acknowledgements

The author thanks NASA ESTO for research support of grant number AIST-05-0035.

References

[1] B. Khattatov, J. Gille, L. Lyjak, G. Brasseur, V. Dvortsov, A. Roche, and J. Waters, "Assimilation of photochemically active species and a case analysis of UARS data," *J. Geophys. Res. (Atmos.)*, vol. 104, no. D15, pp. 18 715–18 737, 1999.

[2] D. J. Lary, D. W. Waugh, A. R. Douglass, R. S. Stolarski, P. A. Newman, and H. Mussa, "Variations in stratospheric inorganic chlorine between 1991 and 2006," *Geophys. Res. Lett.*, vol. 34, no. 21, NOV 13 2007.

[3] D. J. Lary and O. Aulov, "Space-based measurements of hcl: Intercomparison and historical context," *JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES*, vol. 113, no. D15, APR 18 2008.

[4] B. Plale, D. Gannon, J. Brotzge, K. Droegemeier, J. Kurose, D. McLaughlin, R. Wilhelmson, S. Graves, M. Ramamurthy, R. D. Clark, S. Yalda, D. A. Reed, E. Joseph, and V. Chandrasekar, "Casa and lead: Adaptive cyberinfrastructure for real-time multiscale weather forecasting," *Computer*, vol. 39, no. 11, pp. 56–+, 2006.

[5] M. Fisher and D. Lary, "Lagrangian 4-dimensional variational data assimilation of chemical-species," *Q. J. R. Meteorol. Soc.*, vol. 121, no. 527 Part A, pp. 1681–1704, 1995.

[6] D. J. Lary, B. Khattatov, and H. Y. Mussa, "Chemical data assimilation: A case study of solar occultation data from the ATLAS 1 mission of the atmospheric trace molecule spectroscopy experiment (ATMOS)," *J. Geophys. Res. (Atmos.)*, vol. 108, no. D15, 2003.

[7] D. J. Lary, M. D. Muller, and H. Y. Mussa, "Using neural networks to describe tracer correlations," *Atmospheric Chemistry and Physics*, vol. 4, pp. 143–146, 2004.

[8] D. J. Lary and H. Y. Mussa, "Using an extended kalman filter learning algorithm for feed-forward neural networks to describe tracer correlations," *Atmospheric Chemistry and Physics Discussions*, vol. 4, pp. 3653–3667, 2004.

[9] D. J. Lary and A. Koratkar, Eds., *Data Assimilation and Objectively Optimized Earth Observation, Chapter 16*, ser. Advances in Earth Science: From Earthquakes to Global Warming, Royal Society Series on Advances in Science. Imperial College Press, 2007.

David Lary *is a full research professor at NASA Goddard Space Flight Centre, Greenbelt, MD. His research interests include objectively optimized observing systems, atmospheric chemical data assimilation, and neural networks. He is the author of 50 articles in the areas indicated above. Dr. Lary obtained a First Class double honors degree in Physics and Chemistry from Kings College in London in 1987 and a Ph.D. in Computer Modeling of Atmospheric Chemistry from Cambridge University in 1991.*