NOAA Technical Memorandum OAR GSD-63
https://doi.org/10.25923/qt1e-2s03

Community Data Assimilation Research and Development

November 2019

Yuanfu Xie
Zoltan Toth

Earth System Research Laboratory
Global Systems Division
Boulder, Colorado
November 2019

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Office of Oceanic and Atmospheric Research

A Technical Memorandum OAR GSD-63

Community Data Assimilation Research and Development

Yuanfu Xie[1]
Zoltan Toth[2]

[1]State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, China Meteorological Administration
[2]Global Systems Division, NOAA/OAR/ESRL

Acknowledgements

The authors appreciate the interest and support they received from members of the NWP community while discussing and developing the concepts presented in this manuscript. Discussions with Sid Boukabara (NOAA), Xin Zhang (JCSDA, formerly with CIRES at GSD), and Timothy Brown (Spire Global, Inc., formerly with CIRES at GSD) are gratefully acknowledged. The presentation of the manuscript benefitted from the expert comments of Nancy Baker (Naval Research Laboratory), Yannick Tremolet (JCSDA, formerly with ECMWF), Mark Iredell (NCEP), Hui Shao (NCAR), and Mark Govett (GSD).

UNITED STATES DEPARTMENT OF COMMERCE

Wilbur Ross
Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Dr. Neil Jacobs
Acting NOAA Administrator

Office of Oceanic and Atmospheric Research

Craig N. McLean Assistant Administrator

Community Data Assimilation Research and Development

9 August 2019

Preface

Except for minor corrections, the paper below is a copy of a manuscript that was submitted for publication in the Bulletin of American Meteorological Society (BAMS) on 3 February 2015. Before submission, the concepts outlined below had been discussed with and embraced by a number of colleagues in the field (Toth et al. 2014). While the manuscript also received an array of positive remarks from the two Reviewers and an Editor for BAMS: “read through… with much interest”; “intriguing essay”; “applaud the ideas”; “interesting and… bold new approach”; “revolutionary framework”; “the ideas proposed could well be adopted”; “I wish you all success”; the paper was rejected, primarily on the ground that it was “premature” and that “widely attended community-wise workshops need to go first”. The events that followed defied such criticism. Under the ample leadership of Dr. Thomas Auligne, Director of the Joint Center for Satellite Data Assimilation (JCSDA), in a short time, the community rallied behind the proposed idea of, and concepts behind a Community Data Assimilation Repository (CDAR) described in the paper below. Just as suggested in the submitted paper, a series of workshops started in 2016 led to the refinement and community-wide adoption of an object-oriented software approach in US data assimilation (DA) efforts, now named Joint Effort for Data assimilation Integration (JEDI, Auligne 2017). Thanks to the tireless efforts of many under the JEDI project, today CDAR is a functioning reality (JEDI 2019). The authors are glad to see that even without a journal publication, their work has such a profound impact on US data assimilation research and future operations.

Abstract

The quality of Numerical Weather Predictions (NWP) critically depends on the choice of initial conditions. In the US alone, a number of major Data Assimilation (DA) systems have been developed to provide estimates of the state of the atmosphere to initialize NWP forecasts. While these schemes contain state-of-the-art algorithms (i.e., the scientific aspect of DA), their computer representation (i.e., software engineering aspect of DA) is based on less advanced, procedure-oriented software design. Each DA system resides in separate software repositories with no chance for software-level interactions between the various systems. Arguably, this poses a major impediment to collaborative research and the efficient transition of research into operations.

To remedy this situation, we suggest the community adopt the use of an Object-Oriented Design (OOD) approach to DA software development. Based on an abstract-level analysis of the common elements and processes (i.e., “objects”) in current and expected DA schemes and with

the introduction of OOD design elements, the various DA schemes can be reformulated to become part of a common, Community Data Assimilation Repository (CDAR). Different strategies (e.g., variational vs. ensemble filter) and techniques (e.g., observational data en- and decoding, quality control) become interchangeable across the currently disjoint DA systems, greatly expanding the technology available for both the research and user communities. Operational centers can subselect objects from CDAR to assemble and optimize the performance of each of their applications, respectively, accelerating the research to operations transition and overall NWP improvements.

1. Introduction

Global Numerical Weather Prediction (NWP) is the backbone of most weather forecasting activities. Even small improvements can have a significant positive impact on a host of other products that depend on the quality of global forecasts. The quality of global forecasts critically depends on how well the state of the atmosphere is captured to initialize the numerical integrations. Over the past 10–15 years, most leading international centers have modernized their data assimilation techniques. Most adapted adjoint-based 4DVAR data assimilation methods: e.g., European Centre for Medium-Range Weather Forecasts (ECMWF, Rabier et al. 2000); United Kingdom Met Office (Rawlins et al. 2007); Japan Meteorological Agency (JMA 2007); Canadian Meteorological Centre (CMC, Gauthier et al. 2007); and the US Navy (Naval Research Laboratory Atmospheric Variational Data Assimilation System – Accelerated Representer - NAVDAS-AR, Xu et al. 2005), while the National Centers for Environmental Prediction (NCEP) introduced a technique where its 3DVAR Gridpoint Statistical Interpolation scheme (GSI, Kleist et al. 2009) is augmented with ensemble-based 3D (later to be expanded to 4D) background error covariance information (“hybrid GSI”, Wang et al., 2013).

Beyond the GSI used operationally at NCEP of the National Oceanographic and Atmospheric Administration (NOAA) and being adopted for operations at the Air Force Weather Agency (AFWA), there are a number of other well-established data assimilation systems in the US, including the Weather Research Forecasting (WRF) Data Assimilation scheme (WRFDA, Barker et al. 2012) and its 4DVAR version (Zhang et al. 2014) developed at the National Center for Atmospheric Research (NCAR) and used by AFWA; and the Local Analysis and Prediction System (LAPS, Hiemstra et al 2006); and Space and Time Multiscale Analysis System (STMAS, Xie et al. 2011) developed by the Global Systems Division of NOAA, that are used operationally and also by large parts of the US weather community. In addition, there is a significant number of other data assimilation systems and initiatives supported, e.g., by the Global Modeling and Assimilation Office (GMAO, Norris and Silva 2007) of the National Aeronautics and Space Administration (NASA); NOAA’s Atlantic Oceanographic and Meteorological Laboratory (AOML, Aksoy et al. 2013); the National Severe Storm Laboratory (NSSL, Gao et al 2012); NCAR (Data Assimilation Research Testbed – DART, and other schemes, Anderson et al. 2009);

and meteorological departments at various academic institutes such as the University of Maryland and Texas A&M University (TAMU, Szunyogh et al., 2008); Pennsylvania State University (PSU, Zhang et al. 2013); the University of Washington; and State University of New York (SUNY, Torn and Hakim, 2008).

The output of DA schemes is used, not only to initialize NWP forecasts, but also to aid human forecasters in their situational awareness, nowcasting, and very short-range forecast activities. As each of the major systems has been developed and configured to serve different needs, there are good reasons for the existence of multiple data assimilation configurations. Unfortunately, there is a lack of an infrastructure allowing the various groups to interact in their development work. Each major system (e.g., GSI, NAVDAS, WRFDA, LAPS) has its own software repository. No attempt, however, has been made to connect most data assimilation efforts on the software development level. We see this as a major impediment to faster progress on the national level in data assimilation and global NWP. Techniques tested and implemented in one system cannot be easily and directly compared with alternative techniques used in the other systems. Progress in each of the systems therefore often appears piecemeal and suboptimal.

The lack of a common platform or software infrastructure hinders not only DA development but also negatively affects the user community. Since the existing systems or their components are not interchangeable and the adoption of a system often requires significant efforts, users must decide, upfront, which system they want to use. Since each system has its stronger and weaker application areas, the users end up in a situation where the system of their choice will likely perform suboptimally, outside of their main area of interest.

To remedy the current situation, we need to accelerate US data assimilation developments, empower the user community, and regain US leadership in global NWP forecasting. We propose the establishment of a Community Data Assimilation Repository (CDAR) to aid in these efforts. In the rest of this article, we will outline a possible approach, inviting the DA developer and user communities to engage in a discussion that we feel is critical to the continued success of US NWP efforts.

2. Data Assimilation Systems

Each of the data assimilation systems developed by the community contains dozens of theoretically well-separated components (or techniques) designed to fulfill specific needs such as quality control of observational data, interpolation of observational or model forecast data, the conversion between model and observed variables, the minimization of a chosen metric (e.g., the cost function in variational schemes), etc. Various DA systems differ in their choice for each of these functions and the way the functions are connected to form a DA system.

Today, DA systems that are used in or considered for NWP operations fall into two groups; variational approaches and ensemble filter approaches (Kalnay 2003). The international community has not reached consensus as to the method(s) that may offer the best performance in terms of quality, computational efficiency, and other desirable qualities, especially as the requirements for higher spatial resolution analyses are expected to expand in the coming decade. Adjoint-based, variational data assimilation methods, however, have led to consistently superior performance when compared with operational or research output from alternative methods used at other centers (for historical and recent results see, e.g., WMO 2019).

To create a DA system, one must choose a number of techniques to serve specific functionalities and arrange them in a scientifically-sensible manner. This is what we call the scientific aspect of DA systems that is subject to intense research. The arrangement of selected techniques then needs to be transformed into computer software using selected software architecture design principles and coding and scripting languages. A good computer architecture design is critical for the future development and use of DA systems. The present paper does not address the scientific aspect of DA systems, but rather focuses on their software engineering aspect.

3. Community Data Assimilation Repository (CDAR)

DA systems, like other components of NWP technology, evolve under various constraints such as the availability of human and computational resources and time and management pressure. Many elements of these systems are not ideal - this is true for both their scientific and software aspects. Since the software aspects of the various major DA systems have been developed in isolation from their peers, their scientific components are not interchangeable. The maintenance and further development of some of these complex systems also pose special challenges.

To make advances in DA techniques quickly accessible to the wider research and user communities and thus accelerate the advances in data assimilation research and their transition to operations, we propose the creation of a Community Data Assimilation Repository (CDAR, see Fig. 1). CDAR can be a centralized or more likely distributed repository where participating agencies, groups, and

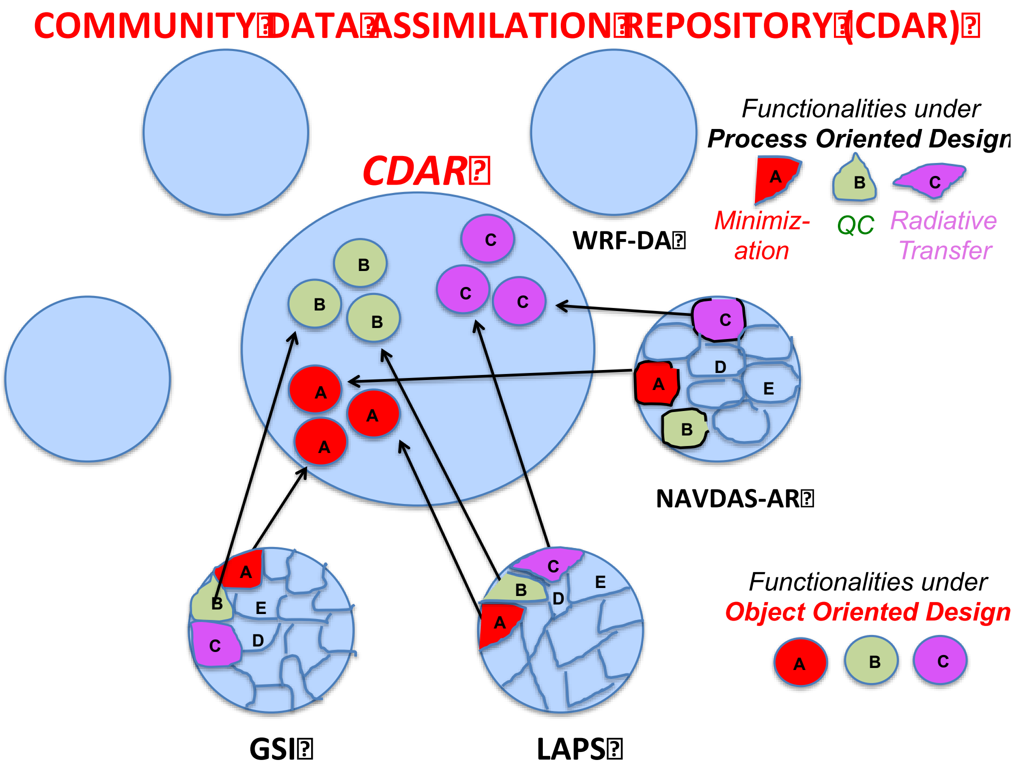

Figure 1. Schematic for a Community Data Assimilation Repository (CDAR, oval in center). Current data assimilation systems (discs around CDAR) are comprised of hard-to-separate or inseparable algorithms that are therefore not interchangeable across the systems. In contrast, CDAR contains a collection of algorithms for each well-defined data assimilation functionality from which users can optimally configure schemes for any specific user application.

researchers share their existing and newly-developed DA techniques. For CDAR to work, participants must define and adhere to specific software protocols (discussed in more detail in the next section) that:

- Support the implementation and use of current and anticipated DA techniques. The CDAR protocols must be able to support both variational and ensemble filter type DA approaches;
- Facilitate the easy exchange of DA techniques among participants. Scientists can continue their usual, off-line development and commit to CDAR as their work matures;
- Allow users to configure DA techniques according to their needs. Such plug and play flexibility can make the designation of DA schemes such as GSI, WRFDA, LAPS, etc. obsolete as users will be able to choose components to mix and match various techniques from the CDAR repository according to their needs and will;
- Accelerate research to operations transition. Operational centers will have a much larger pool of candidate techniques developed and tested by the community to choose from;

- Anticipate advances in computer science and software engineering. A future-oriented approach will allow subsequent updates of or additions to the CDAR protocols, serving the evolving needs of the community.

4. Technical Approach

Though the concept of CDAR may be appealing, is there a viable technical approach to make it a reality? Below we describe a possible approach that can be assessed as to its technical soundness and ease of implementation.

4.1. Object-Oriented Software Design

Traditional, procedure-oriented software design (Schach 2011) has focused on satisfying specific computational requirements identified in relation to an end-to-end computational procedure. Such an approach may offer software that is well suited for a specific application but not amenable for easy modifications under changing requirements. Object-Oriented Design (Gamma 1995, Rouson 2011) on the other hand, starts with a conceptual analysis of a problem, followed by the design of modular and reusable objects (such as a data structure along with functions or other sets of calculations). Such an abstract approach at understanding the underlying problem and its possible solution is not bound by specifics of its most apparent use case or scenario, allowing for an easy expansion or generalization of the software if and when desired.

Since data assimilation is a complex and relatively fast developing field with frequent upgrades to the observing systems, analysis methodologies, and forecast models used, the implementation of related software should preferably be organized according to higher level and more general considerations instead of low-level details. This allows both the developers and users to easily modify implementation details in a particular area of the software without affecting the functioning of the overall system. For example, in a well-designed repository, one should be able to switch from the use of a variational to ensemble filter type scheme without any changes to data ingest and post-processing procedures. Or the details of specific observational instruments (such as radiosondes or mesonets) should be handled by particular routines that interface with the observational data files and remain hidden from the rest of the software. Where to draw the line demarcating details hidden from the bulk of the software is perhaps the most critical step in the design of an OOD software repository.

4.2. Design Elements and Principles

In OOD, software programs are considered as interactions among objects containing data and procedures to serve particular functions. The definition of objects and the corresponding interfaces at the right (not too specific and not too general) level of abstraction is at the core of good OOD software development practices.

Some of the design principles used in OOD software development, along with some examples on how these can benefit and support collaborative DA research and development are introduced below.

Encapsulation and information hiding

According to these principles, the components of OOD software must be made self-contained to the extent possible. Only information necessary for proper interactions with other objects is revealed to the developers of other objects and the users.

For an example in meteorological data assimilation, consider the ingest of radiosonde observations (RAOB). The developers of new minimization algorithms in variational data assimilation do not need to know how RAOB observations are retrieved, whether the observational data are at mandatory or other levels, or how data quality control is done. Future changes or improvements in the ingest of RAOB will not affect developers of other parts of the software either, as long as data interfaces stay the same. Details on RAOB observation arrays or on reading those data are therefore handled and contained (or encapsulated) in the appropriate data ingest objects or modules and kept hidden from other objects.

A potentially large benefit from following the principle of information sharing parsimony for community software development is that experts can work on a particular object or objects without unnecessary interference with changes made by other community members to other objects.

Inheritance

According to this principle, objects or classes are connected through their interfaces without knowing about their implementation details. A new object belonging to a so-called subobject or subclass can inherit certain properties of its parent class while possessing some specific properties in terms of data and function elements. Despite such unique details, subclasses will not affect the way their parent classes interact with other classes.

Data assimilation systems deal with data from a variety of inhomogeneous observing systems such as RAOB, profilers, or radars. Most DA systems follow the traditional, procedure-oriented design and require users to create entirely new data structures and functions for each new observation type to be added to the system. OOD, on the other hand, recognizes that the various observational datasets share some common features like observation values, error estimates, and the location and time of observations. In OOD, a single observation object can be defined with features common across all types of observations. For observation types with unique characteristics, a subclass of observations can be created. The subclass for radar radial winds, for example, will inherit all desired data structures and functionalities from its parent class (observations), such as observation values, locations, and errors. In addition, as a subclass, it can be extended with features needed only for that type of observation, such as the nyquist-unfolding function (Zhang and Wang 2006) for the processing of radial wind data.

Benefits of the inheritance principle include savings associated with the elimination of redundancy in the development of common data structures and functions that can be inherited from a parent class and the ease with which new subclasses can be configured with the addition of special functionalities to parent class properties.

Polymorphism

Via the use of common interfaces, polymorphism in OOD supports access to different objects and subobjects (see https://en.wikipedia.org/wiki/Polymorphism). Polymorphism renders software more generally applicable and versatile by minimizing and localizing changes required for modified or expanded applications, thus making references to different functions more transparent and easier.

Object-oriented design is inherently developer friendly as it provides independence and flexibility in software development and is highly suitable for community development of future data assimilation systems. The resulting software is easy to reconfigure as changes to components of the software can be developed independently from other parts of the software. The added cost in terms of the upfront, abstract-level analysis of the problems may be compensated by the ease at which OOD software can be reconfigured later on. Today, OOD is followed as standard practice in the computer industry and is widely used at businesses around the world (Schach 2011).

4.3. Technical implementation considerations

The development of an OOD-based Community Data Assimilation Repository requires the use of a programming language with object-oriented capability, as well as a careful object-oriented analysis of current and expected data assimilation systems.

Programming languages

In their ambitious effort to transform most of their NWP system, including numerical modeling and data assimilation into an Object-Oriented Prediction System (OOPS, Trémolet, 2013), ECMWF considered the use of C++, and Python programming languages that support object-oriented design and implementation. OOPS is using C++ for the OOP layer and Fortran for the inner model layer. Even though C++ and Python are commonly used in other communities, such as the computer industry, their use is not so widespread in the meteorological community. Fortunately, Fortran 2003 and 2008 extend their predecessor's capabilities in handling inheritance, polymorphism, and encapsulation. Therefore, it may be possible to design and implement CDAR using a single language (Fortran) that is widely used in the meteorological community.

OOD Analysis of DA systems

Data assimilation systems can be partitioned into three major components: data ingest, the analysis core, and post-processing. An OOD DA repository could follow this structure, with appropriate and unified interfaces connecting the main components (see Fig. 2). The interface between the data ingest and the analysis core components will be the processed observations from various sources and the background field(s). The interface between the analysis core and post-processing will be the array of analysis grid(s) themselves. The interfaces and other technical details can be formally defined through open discussions following a protocol agreed upon by the data assimilation community.

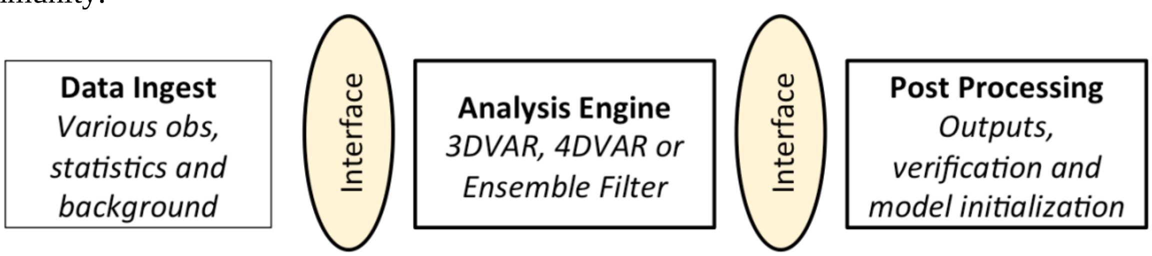

Figure 2. The main components of CDAR based on an object-oriented analysis of DA systems.

Data ingest

As the ingestion of various data types has some common elements as well as elements that are unique to specific observation types, an OOD approach can provide much needed support for community system development. The inhomogeneity of observation types and characteristics can be well organized under an OOD architecture proposed for CDAR. Fig. 3 illustrates the abstraction of the data ingest problem, fully exploiting the inheritance principle of OOD.

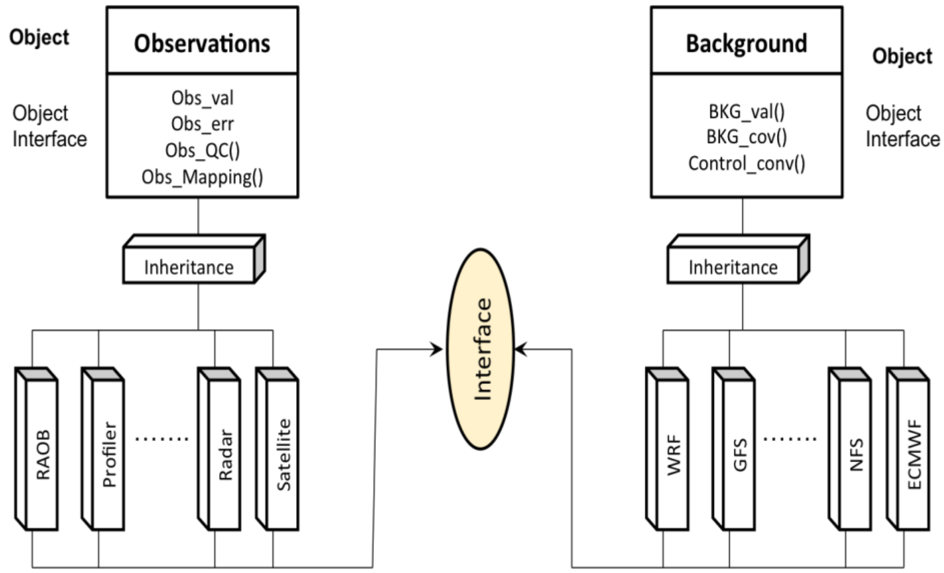

Figure 3. Object-oriented design of the CDAR data ingest component.

Analysis Core

As pointed out in section 2, the large number of data assimilation systems used by the community fall into one of two categories. CDAR will allow both variational and ensemble filter approaches to share commonly used objects. Such examples include the objects of various observations, or the object evaluating the discrepancy between observations and the analysis (see Fig. 4). Some other objects, on the other hand, will be used only by variational schemes such as those related to the background and other constraints (see items connected to the rhomboids in Fig. 4).

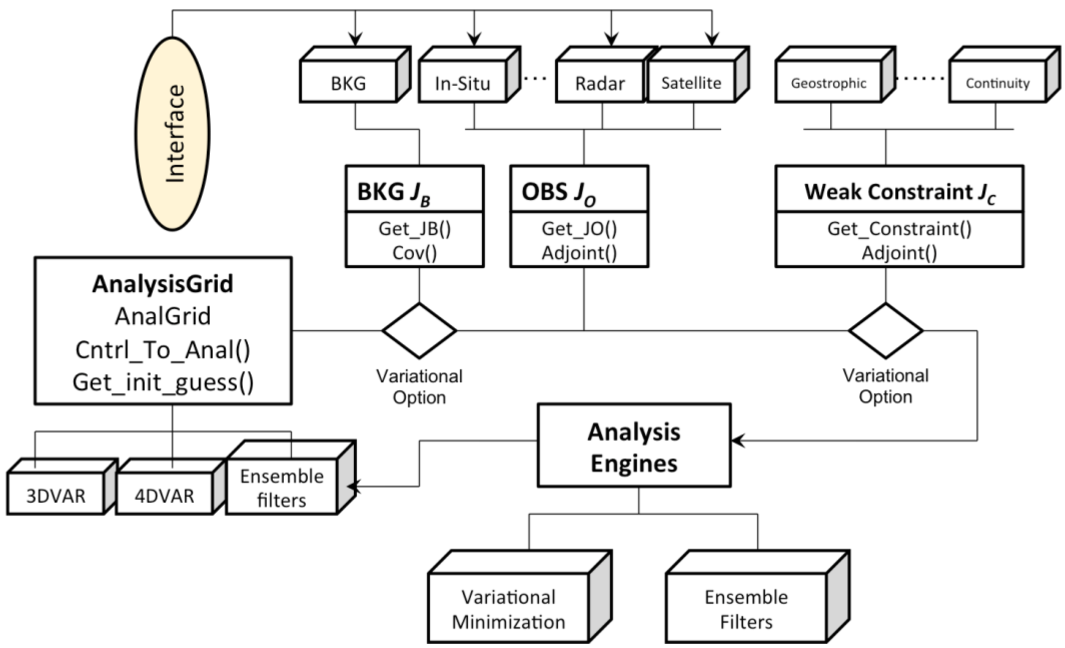

Figure 4. Object-oriented design for the analysis core with two choices (variational or ensemble filter) for the analysis engine. Note that many supporting objects (e.g., background term, observations, constraints) are shared by the different analysis engines in the proposed CDAR architecture.

Post-processing

The post-processing components of data assimilation systems enable the use of output from data assimilation schemes in various functions such as in the initialization of NWP models, in situational awareness and nowcasting activities, in verification of NWP forecasts, and in the evaluation of the data assimilation schemes themselves by comparing their output with observations (Fig. 5). With OOD, data access becomes transparent both from a user application and product point of view.

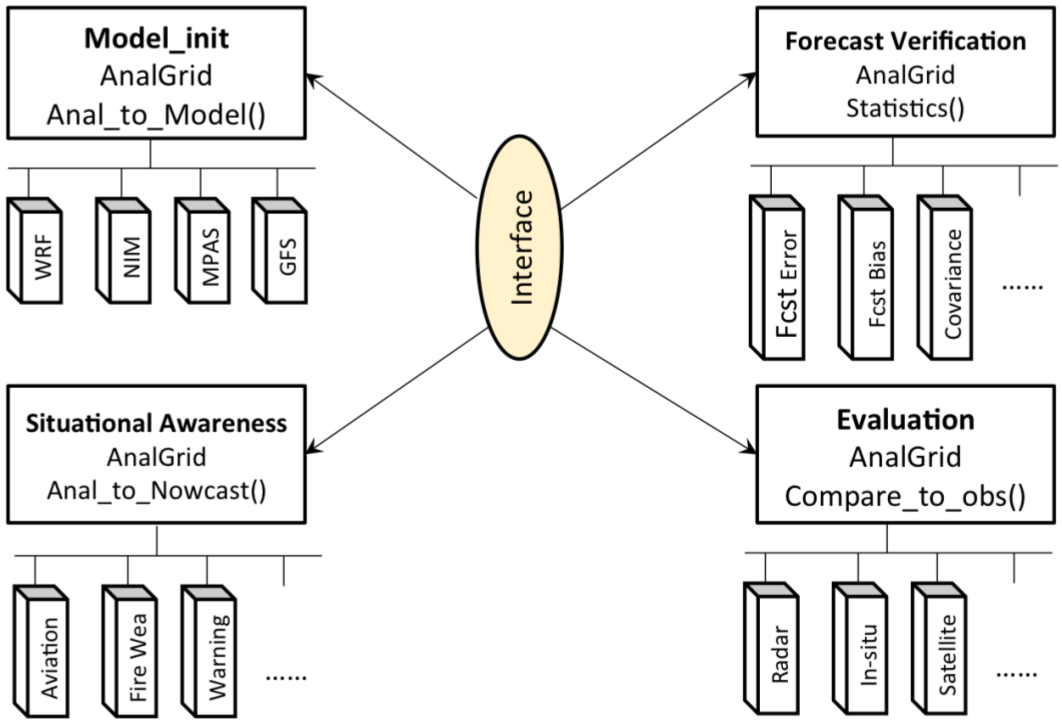

Figure 5. Object-oriented design of the post-processing part of the proposed CDAR architecture.

In summary, an OOD architecture generalized from actual data assimilation systems containing several options for various abstract object types offers a number of benefits. Researchers and users alike can configure and optimize their system with flexibility (e.g., choice of observation types, variational or ensemble filter, numerical forecast initialization or forecast verification schemes) for specific applications at hand. For example, data ingest can handle NWP forecasts from any numerical model using function overloading (polymorphism) to deal with different forecast data formats, while the verification and other post-processing objects will seamlessly connect with output from the analysis core (i.e., the analysis and its error covariance matrix).

5. Potential obstacles

As with any other new initiative, potential pitfalls and obstacles may hinder the development of CDAR and thus need to be carefully considered, including:

- Initial skepticism. This can be addressed by considering experience with related work in other fields and other regions (e.g., the Object-Oriented Prediction System (OOPS) being developed at ECMWF, and with carrying out well-designed pilot project(s);
- Computational performance. There is a well–understood tradeoff between performance (with software optimized for very specific applications) and flexibility and extensibility (with properly developed OOD architecture). The granularity of the objects in OOD

directly affects both its computational efficiency and utility. Objects too small or too specific will hinder parallelization and hence computational efficiency, while objects too large will diminish the potential advantages of OOD. The community must seek a balance between computational efficiency and utility of OOD by selecting an appropriate level of complexity for objects across the various components of CDAR;

- Reliability and security in operations. This can be mitigated by allowing operational centers full control over their system configuration as well as satisfying their security requirements;
- Cost of conversion. Costs and service interruptions can possibly be minimized by a gradual adoption of CDAR. CDAR protocols can first be followed in the development of new techniques while existing codes can be converted as resources permit and/or when their update becomes necessary for any other reason.

6. Conclusions

This paper reviewed some new tools for enhanced community involvement in the development and transition of new data assimilation methodologies into US NWP operations. Unlike GSI, WRFDA, or LAPS, the proposed Community Data Assimilation Repository (CDAR) is not a data assimilation system. Rather, it is a protocol that if adopted by the NWP community, will allow developers to place their DA techniques into a common repository. Such an arrangement will allow developers to easily interchange DA techniques and will reduce overall maintenance costs. Moreover, it will enable users to freely configure their DA systems drawing from a wide array of techniques and sources that best suit their needs given their applications. This is analogous to how other community software systems such as WaveWatch lll (Tolman 2009) and the Hybrid Coordinate Ocean Model (HYCOM, Chassignet et al. 2006) are configured differently at various operational centers. Depending on their specific needs and requirements, WaveWatch contributing partners such as the MetOffice in the UK or the US Navy use different components but all drawn from the same software repository.

CDAR's object-oriented design features, not available in today's fragmented DA software development environment, will empower the interagency testbeds (the Joint Center for Satellite Data Assimilation (JCSDA), and the Developmental Testbed Center (DTC)), and will potentially lead to closer collaboration across agencies. In turn, we may observe a significant acceleration in transitioning new techniques from the academic research community into public and private sector operations.

Like most bold initiatives, CDAR will have its skeptics and potential obstacles. With this article we invite members of the data assimilation research and user communities to take a careful look at the concept of CDAR and critique and modify the approach suggested here. We believe an open dialogue will lead to a comprehensive assessment of the options available to the community, and eventually to the selection of the most promising way forward.

The data assimilation community can learn from the collaborative experience accumulated through the development and use of the Earth System Modeling Framework (ESMF, Collins et al. 2005) by the US NWP modeling community. ESMF provides tools and protocols, some of which are object oriented, that attempt to enhance ease of use, portability, interoperability, and reuse of various climate and weather numerical modeling components. CDAR can benefit from some of the ESMF tools such as its regridding algorithms, and potentially from some of its protocols and user experiences. Other tools and object-oriented design protocols will have to be developed for DA applications (e.g., for handling inhomogeneous types of observations).

Looking ahead, if the community chooses to engage in this effort, the potential of CDAR can be explored via the following steps:

- Community discussions about CDAR, possibly facilitated by NOAA's Virtual Laboratory (VLab, Various 2015), using virtual and in-person meetings and other venues;
- Consensus decision on the principles and key design elements of CDAR;
- Participants commit to carry out all new DA development work in compliance with accepted CDAR guidelines;
- With time, as various elements of the participants' DA systems are continuously revised, most DA techniques become CDAR compliant;
- Participating developers interchange their techniques for testing and use in CDAR;
- Users configure their DA systems from a wide variety of techniques and sources to best suite their specific applications.

The authors are looking forward to fruitful community-wide discussions and collaboration facilitated by the proposed object-oriented software design on a scale not seen before; first in data assimilation, then possibly in other areas such as forecast verification and statistical post-processing.

References:

Aksoy, Altuğ, Sim D. Aberson, Tomislava Vukicevic, Kathryn J. Sellwood, Sylvie Lorsolo, Xuejin Zhang, 2013: Assimilation of High-Resolution Tropical Cyclone Observations with an Ensemble Kalman Filter Using NOAA/AOML/HRD's HEDAS: Evaluation of the 2008–11 Vortex-Scale Analyses. *Mon. Wea. Rev.*, **141**, 1842–1865.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009: The Data Assimilation Research Testbed: A Community Facility. *Bull. Amer. Meteor. Soc.*, **90**, 1283–1296.

Auligne, Thomas, 2017: Joint Effort for Data assimilation Integration. Presentation at a community data assimilation workshop, taken on 8 July 2019 from: https://www.da.ucar.edu/sites/default/files/Auligne_JCSDA_WS.pdf

Barker, D., and Coauthors, 2012: The Weather Research and Forecasting Model's Community Variational/Ensemble Data Assimilation System: WRFDA. *Bull. Amer. Meteor. Soc.*, **93**, 831–843.

Chassignet, E.P., H.E. Hurlburt, O.M. Smedstad, G.R. Halliwell, P.J. Hogan, A.J. Wallcraft, and R. Bleck, 2006: Ocean prediction with the Hybrid Coordinate Ocean Model (HYCOM). In: Ocean Weather Forecasting: An Integrated View of Oceanography, Chassignet, E.P., and J. Verron (Eds.), Springer, 413-426.

Collins, N., G. Theurich, C. DeLuca, M. Suarez, A. Trayanov, V. Balaji, P. Li, W. Yang, C. Hill, and A. da Silva (2005): Design and Implementation of Components in the Earth System Modeling Framework. *International Journal of High Performance Computing Applications*, Volume 19, Number 3, pp. 341-350.

Gamma, E., R. Helm, R. Johnson, and J. Vlissides, 1995: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

Gao, J., and D. Stensrud, 2012: Assimilation of reflectivity data in a convective-scale, cycled 3DVAR framework with hydrometeor classification. J. Atmos. Sci., 69, 1054–1065.

Gauthier, P., M. Tanguay, S. Laroche, S. Pellerin and J. Morneau, 2007. Extension of 3D-Var to 4D-Var: Implementation of 4D-Var at the Meteorological Service of Canada. Mon. Wea. Rev. 135:2339-2354.

Hiemstra, C.A., G. E. Liston, R.A. Pielke Sr.,D. L. Birkenheuer, and S. C. Albers, 2006: Comparing Local Analysis and Prediction System (LAPS) assimilations with independent observations. *Wea. Forecasting,* **21**, 1024–1040.

JEDI 2019: Joint Effort for Data assimilation Integration. Taken from the JCSDA website on 8 July 2019: https://jointcenterforsatellitedataassimilation-jedi-docs.readthedocs-hosted.com/en/latest/index.html

JMA, 2007: Outline of the operational numerical weather pre- diction at the Japan Meteorological Agency. Appendix to WMO Technical Progress Report on the Global Data- Processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP), Japan Meteorological Agency, Tokyo, Japan. [Available online at http://www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline-nwp/index.htm.]

Kalnay, E., 2003: Atmospheric modeling, data assimilation and predictability. Cambridge University Press.

Kleist, D. T., David F. Parrish, John C. Derber, Russ Treadon, Wan-Shu Wu, Stephen Lord, 2009: Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting.* 24, pp. 1691-1705.

Norris, Peter M., Arlindo M. da Silva, 2007: Assimilation of Satellite Cloud Data into the GMAO Finite-Volume Data Assimilation System Using a Parameter Estimation Method. Part I: Motivation and Algorithm Description. *J. Atmos. Sci.*, **64**, 3880–3895.

Rabier, F., H. Jarvinen, E. Klinker, J.-F. Mahfouf, and A. Simmons, 2000: The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simpli- fied physics. *Q. J. R. Meteorol. Soc.*, **126,** 1143–1170.

Rawlins, F., S. P. Ballard, K. J. Bovis, A. M. Clayton, D. Li, G. W. Inverarity, A. C. Lorenc and T. J. Payne, 2007: The Met Office global four-dimensional variational data assimilation scheme. *Q. J. R. Meteorol. Soc.* **133**: 347–362.

Rouson, D., J. Xia, and X. Xu, 2011: Scientific Software Design: The Object-Oriented Way. Cambridge University Press, New York, NY, USA.

Schach, S., 2011: Object-Oriented and Classical Software Engineering. McGraw-Hill, New York, NY, USA.

Szunyogh, I., E. J. Kostelich, G. Gyarmati, E. Kalnay, B. R. Hunt, E. Ott, E. Satterfield*, and J. A. Yorke, 2008: A Local Ensemble Transform Kalman Filter data assimilation system for the NCEP global model. *Tellus,* **60A,** 113-130.

Tolman, H. L., 2009: User manual and system documentation of WAVEWATCH III version 3.14. NOAA / NWS / NCEP / MMAB Technical Note **276**, 194 pp.+ Appendices.

Torn, R. D., and G. J. Hakim, 2008: Performance characteristics of a pseudo-operational ensemble Kalman filter. *Mon. Wea. Rev.*, **136,** 3947--3963. DOI: 10.1175/2008MWR2443.1

Toth, Z., Yuanfu Xie, Sid Boukabara, Xin Zhang, and Timothy Brown, 2014: Community Data Assimilation Research and Development. Manuscript, available from the authors.

Tremolet, Y., 2013: Object-Oriented Prediction System. Available at: http://www.data-assimilation.net/Events/Year3/OOPS.pdf

Various authors, 2015: Virtual Laboratory. Available at: http://www.nws.noaa.gov/mdl/vlab/

Wang, X. G., D. F. Parrish, D. Kleist, J. Whitaker, 2013: GSI 3DVar-based ensemble-variational hybrid data assimilation for NCEP global forecast system: single resolution experiments. *Mon. Wea. Rev.*., **141**, 4098–4117.

WMO, 2019: World Meteorological Center – Lead Center for Deterministic Forecast Verification, available at: https://apps.ecmwf.int/wmolcdnv/

Xie, Y. F., S. Koch, J. McGinley, S. Albers, P. Beringer, M. Wolfson, and M. Chan, 2011: A Space-Time Multiscale Analysis System: A Sequential Variational Analysis Approach. *Mon. Wea. Rev.*, **139**, 1224-1240.

Xu, L., T. Rosmond, and R. Daley, 2005: Development of NAVDAS-AR: Formulation and initial tests of the linear problem. Tellus, 57A, 546-559.

Zhang, J. and S. Wang, 2006: An Automated 2D Multipass Doppler Radar Velocity Dealiasing Scheme, J. Atmos. Oceanic Tech. Vol. 23, 1239-1248.

Zhang, F., M. Zhang, and J. Poterjoy, 2013: E3DVar: Coupling an ensemble Kalman filter with three-dimensional variational data assimilation in a limited-area weather prediction model and comparison to E4DVar. *Monthly Weather Review*, **141**, 900-917.

Zhang, X., X. Huang, J. Liu, J. Poterjoy, Y. Weng, F. Zhang, and H. Wang, 2014: Development of an Efficient Regional Four-Dimensional Variational Data Assimilation System for WRF. *J. Atmos. Oceanic Technol.*, **31**, 2777–2794.
doi: http://dx.doi.org/10.1175/JTECH-D-13-00076.1