

Use of Observing System Simulation Experiments in the United States

Xubin Zeng, Robert Atlas, Ronald J. Birk, Frederick H. Carr, Matthew J. Carrier, Lidia Cucurull, William H. Hooke, Eugenia Kalnay, Raghu Murtugudde, Derek J. Posselt, Joellen L. Russell, Daniel P. Tyndall, Robert A. Weller, and Fuqing Zhang

ABSTRACT: The NOAA Science Advisory Board appointed a task force to prepare a white paper on the use of observing system simulation experiments (OSSEs). Considering the importance and timeliness of this topic and based on this white paper, here we briefly review the use of OSSEs in the United States, discuss their values and limitations, and develop five recommendations for moving forward: national coordination of relevant research efforts, acceleration of OSSE development for Earth system models, consideration of the potential impact on OSSEs of deficiencies in the current data assimilation and prediction system, innovative and new applications of OSSEs, and extension of OSSEs to societal impacts. OSSEs can be complemented by calculations of forecast sensitivity to observations, which simultaneously evaluate the impact of different observation types in a forecast model system.

<https://doi.org/10.1175/BAMS-D-19-0155.1>

Corresponding author: Xubin Zeng, xubin@email.arizona.edu

In final form 28 March 2020

©2020 American Meteorological Society

For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#).

AFFILIATIONS: **Zeng and Russell**—The University of Arizona, Tucson, Arizona; **Atlas and Cucurull**—NOAA/ Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida; **Birk**—The Aerospace Corporation, El Segundo, California; **Carr**—University of Oklahoma, Norman, Oklahoma; **Carrier**—U.S. Naval Research Laboratory, Stennis Space Center, Mississippi; **Hooke**—American Meteorological Society, Washington, D.C.; **Kalnay and Murtugudde**—University of Maryland, College Park, College Park, Maryland; **Posselt**—Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California; **Tyndall**—U.S. Naval Research Laboratory, Monterey, California; **Weller**—Woods Hole Oceanographic Institution, Woods Hole, Massachusetts; **Zhang***—The Pennsylvania State University, University Park, Pennsylvania
* Deceased

An observing system simulation experiment (OSSE) is a modeling experiment used to evaluate the value of a new observing system when actual observational data are not available. An OSSE system includes a nature run,¹ a data assimilation system,² and software to simulate “observations” from the nature run and to add realistic observation errors. OSSEs have been performed to determine whether a new observing system will add value to numerical weather prediction (NWP)³ and analysis; to make design decisions for a new observing system or network; and to investigate the behavior of data assimilation systems and thereby optimally tune these systems in an environment where the “truth” and hence the system’s behavior is known.

OSSEs are essentially an extension of observing system experiments (OSEs) that determine the impact of existing observing systems. Both use data denial experiments, with one assimilating all observations and the other denying the observing system of interest, and the impact is determined by the increase in analysis and forecast errors resulting from the addition of that system. While OSEs use real observations, OSSEs use simulated data from the nature run. They are complemented by forecast sensitivity to observations (FSO) and ensemble FSO (EFSO), which are effective techniques to simultaneously evaluate the impact of different observation types in a forecast model system.

The Weather Research and Forecasting Innovation Act of 2017 mandates the National Oceanic and Atmospheric Administration (NOAA) to undertake OSSEs to quantitatively assess the relative value and benefits of observing capabilities and systems at NOAA. Accordingly the NOAA Science Advisory Board identified the review of OSSE use as one of its key activities, and tasked its Environmental Information Services Working Group to lead this effort in collaboration with the Climate Working Group. The team includes diverse members from the government, academic, and private sectors. Through several iterations, the final white paper was submitted to NOAA in April 2019, and it was subsequently forwarded to the U.S. Congress.

Recognizing the broad importance of OSSEs to the government, academic, and private sectors, and based on the collective expertise and experience of the team, here we present a revised version of the above white paper to briefly review the use of OSSEs in the United States and to develop recommendations with associated rationale for consideration by various organizations. Our goal is to provide an introductory discussion of OSSEs (with no references provided in the text). Our intent

¹ A model simulation with high horizontal resolutions (e.g., a horizontal grid size of a few kilometers in a global model) whose output is assumed to closely represent the true environmental conditions in a statistical sense.

² A system that combines observational data with model output to produce an optimal estimate of the evolving state of the system.

³ A method of weather forecasting that uses mathematical models of the atmosphere and related components of the Earth system to predict the weather based on current weather conditions.

is that readers from different fields (than atmospheric sciences) and those with little or no foreknowledge of OSSEs will enjoy reading the essay and gain some basic knowledge (e.g., what OSSEs are, what organizations are doing them in the United States, their values and limitations, and future directions). Some of these readers may want to know more about OSSEs by reading other regular OSSE publications (starting from the list at the end of this essay) and even consider using OSSEs in their future activities. At the same time, OSSE system developers and users will find the recommendations useful for moving forward.

Thus the second section briefly reviews the use of OSSEs in the United States and the third section discusses the value and limitations of OSSEs. Compared with prior OSSE-related studies, this paper covers both forecasting and other types of OSSEs in the third section, and presents five recommendations on potential actions related to OSSEs in the fourth section. These recommendations are based on the discussions in the second and third sections. A short list of relevant publications with a brief synopsis is also provided at the end of the fourth section for further reading. All acronyms are defined when introduced, but for convenience they are also listed in the appendix.

Brief review of OSSE activities in the United States

Since the 1980s, extensive weather forecast and analysis OSSEs have been developed and conducted, first at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center, and later at the NOAA Atlantic Oceanographic and Meteorological Laboratory in collaboration with operational data assimilation centers, private enterprise, and academic partners. These OSSEs determined correctly the quantitative potential for several proposed satellite observing systems to improve weather analysis and prediction prior to their launch, evaluated trade-offs in orbit configurations, coverage and accuracy for space-based observing systems, and were used in the development of the methodology that led to the first beneficial impacts of satellite surface winds on NWP. The current methodology used for forecast impact OSSEs has been accepted nationally and internationally as the way in which OSSEs should be conducted in order to provide credible results. Today, OSSEs and related capabilities exist at NOAA, NASA, Naval Research Laboratory (NRL), other national centers/laboratories, universities, and the private sector.

Since 2014, OSSEs in NOAA have been performed under NOAA's Quantitative Observing System Assessment Program (QOSAP). QOSAP coordinates the assessment of the impact of current and new observations across NOAA. It uses OSSEs to assess future observing systems, OSSEs to test the value of existing observing systems, and FSO and EFSO as effective techniques to simultaneously evaluate the impact of many different observation types in a forecast model system. QOSAP's primary objective is to improve quantitative and objective assessment capabilities to evaluate operational and future observation system impacts and trade-offs, which can then be used to assess and to prioritize NOAA's observing system architecture. Its main focuses are 1) to increase NOAA's capacity to conduct quantitative observing system assessments, 2) to develop and use appropriate quantitative assessment methodologies, and 3) to inform major decisions on the design and implementation of optimal combinations of observing systems.

Under QOSAP, a state-of-the-art global OSSE system, an advanced hurricane OSSE system, and an internationally recognized first of its kind rigorous ocean OSSE system were developed. For global NWP, an OSSE system was developed for observation impact assessments based on the 7 km NASA Global Earth Observing System, version 5, nature run. QOSAP has begun the initial testing of the global OSSE system using a new 9-km global nature run based on the Integrated Forecasting System at the European Centre for Medium-Range Weather Forecasts. Development of regional OSSE systems for high impact weather and air quality have been initiated, and a 2-km basin-scale ocean nature run has been developed.

Using these systems, a significant number of OSEs and OSSEs in both global and regional (i.e., tropical cyclone) analysis and forecast systems were performed for multiple existing and proposed observing systems and many of these have since been published in the refereed literature. QOSAP met the deadlines to complete OSSEs with Global Navigation Satellite System Radio Occultation and Geostationary Hyper-Spectral Sounder Constellation observations required by the Weather Research and Forecasting Innovation Act of 2017. Additionally, QOSAP conducted OSSEs related to the role of ocean observations in hurricane prediction. Finally, QOSAP began the process to develop the quantitative assessment capability to meet the needs of the NOAA National Ocean Service and National Marine Fisheries Service, and OSSE capabilities for other ocean basins, coastal oceans, and for climate are under development. These capabilities are summarized in Fig. 1.

Besides NOAA, NASA has been conducting OSSEs for decades, most recently by the Global Modeling and Assimilation Office (GMAO). The goal is to determine how much additional information is provided to an analysis of the state of the Earth system by a new set of measurements, relative to the current global observing system. The primary product produced by NASA's modeling and data assimilation infrastructure is the Modern-Era Retrospective Analysis for Research and Applications. GMAO conducts research into how to properly calibrate an OSSE and has produced the aforementioned global 7 km mesoscale-resolving nature run (also used by NOAA). The GMAO OSSE system consists of the NASA Global Earth Observing System, version 5, model and the Gridpoint Statistical Interpolation data assimilation system. Note that this data assimilation method is also used in the NOAA National Centers for Environmental Prediction (NCEP) operational forecast system.

In general, NASA uses Earth observations for two primary purposes: 1) accurate characterization of Earth's atmosphere, oceans, cryosphere, and land surface, and 2) scientific discovery of the processes that govern the evolution of the Earth system, and the linkages among the components of the system. Forecast OSSEs are appropriate for 1, but they are limited in their ability to measure the effectiveness of an observing system for answering specific science

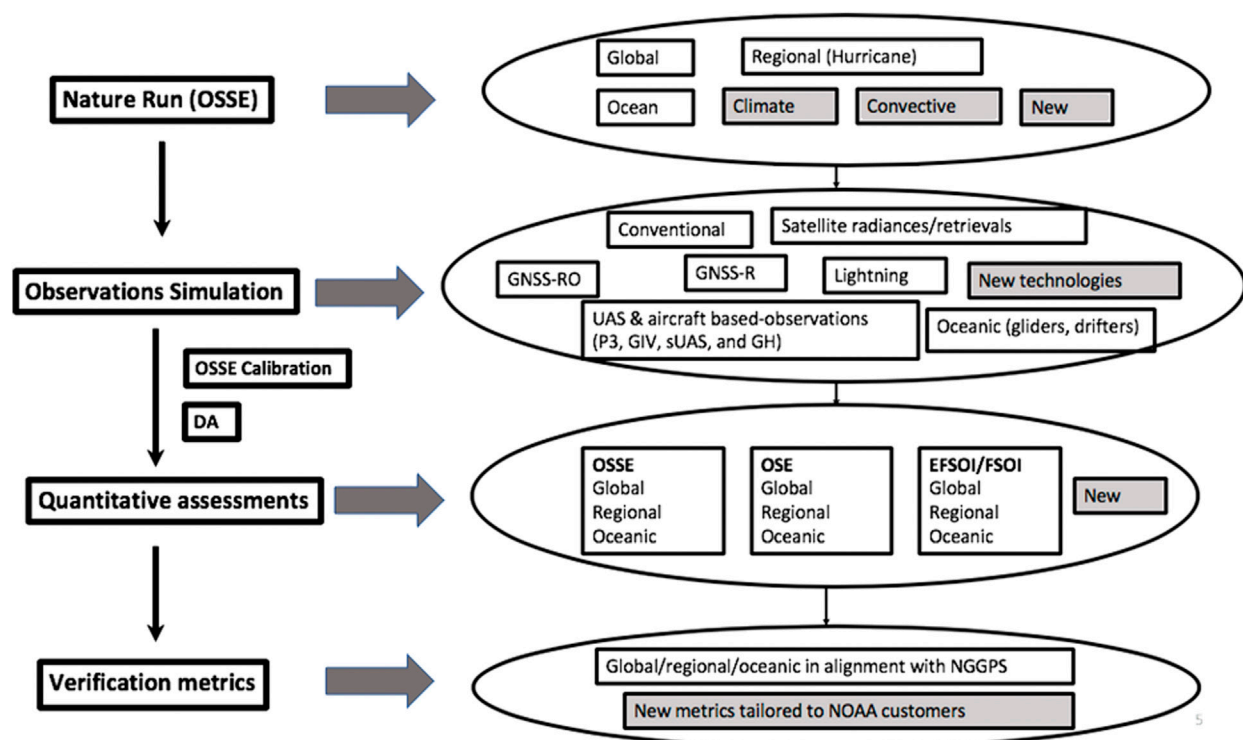


Fig. 1. Diagram showing the different components of the OSSE system at NOAA Atlantic Oceanographic and Meteorological Laboratory. Gray shading indicates a capability to be developed.

questions 2. In addition, there is a significant amount of preparatory research that must be conducted if a forecast OSSE is to provide an accurate measure of the geophysical capability of a new measurement. As such, NASA has broadened the definition of an OSSE to include two additional sets of activities, both of which are also important to the success of a forecast OSSE. A *sampling OSSE* is used to determine whether a candidate measurement system has sufficient temporal and spatial sampling to address a specific science question. A *retrieval OSSE* quantifies the degree to which prospective measurements provide information on a geophysical quantity of interest. In addition to assessing measurement sufficiency, the outcome of a retrieval OSSE can be used to specify uncertainties in a forecast OSSE. Schematics depicting retrieval and sampling OSSEs are shown in Fig. 2.

NRL conducts OSSEs to help meet the mission needs of the U.S. Navy, which requires meteorological and oceanographic information to characterize the battlespace environment to support global-, regional-, and tactical (or local)-scale operations on time scales ranging from minutes to weeks. Because the battlespace is often data sparse, investments in new observation types addressing insufficiently sampled properties are critical. Estimated impacts calculated by OSSEs on Navy NWP forecasts, ocean forecasts, and tactical decision aids (e.g., for integration of environmental information into battle group command decision making) help justify investments in new observing systems. However, running a traditional OSSE can be costly in both personnel and computational resources due to the generation of the nature run and the simulation of both new and existing observations from the global observing

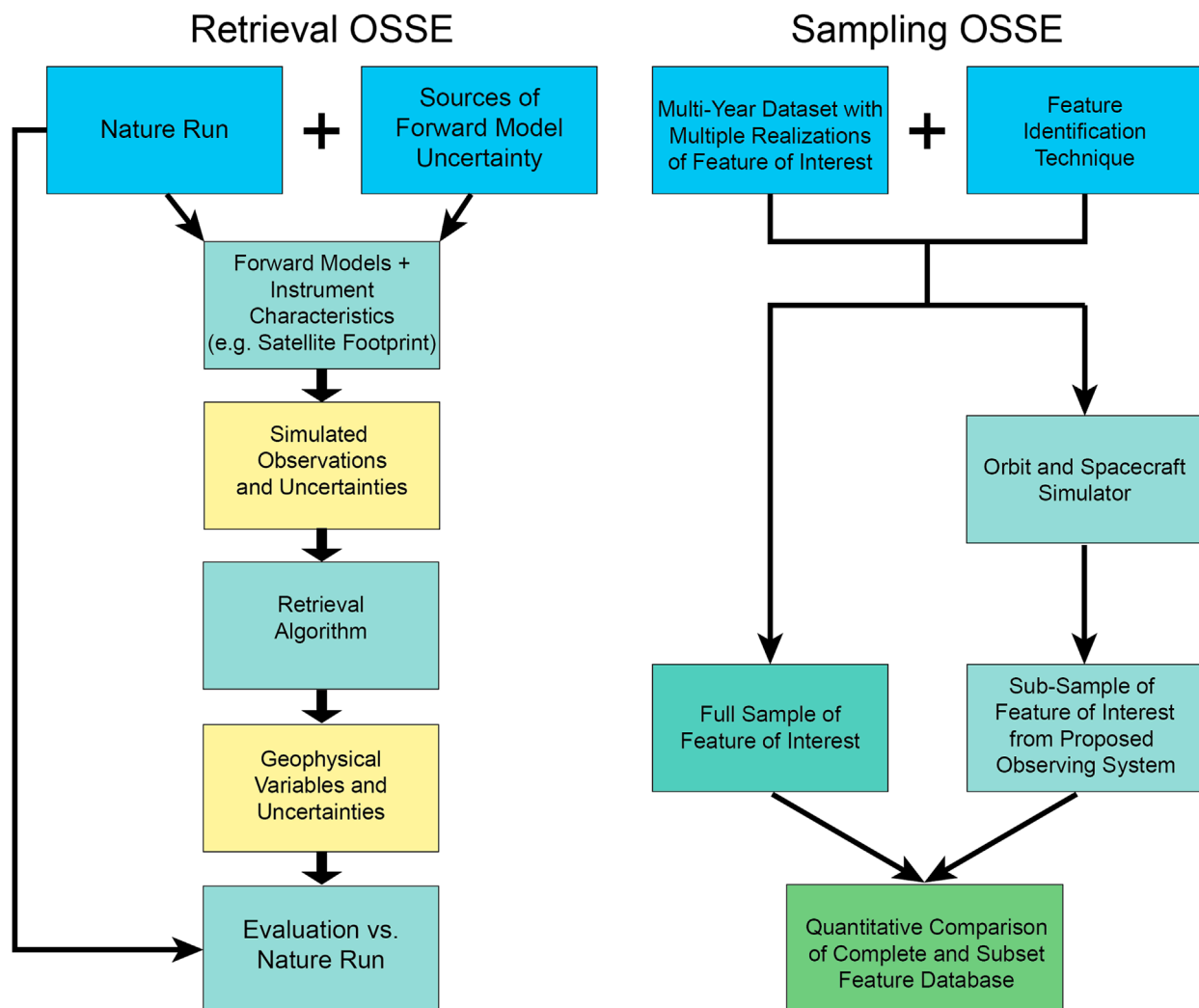


Fig. 2. Schematics depicting retrieval and sampling OSSEs as used in NASA.

system. Instead of running traditional OSSEs to estimate observation impacts, NRL has run several variants of the methodology (e.g., the historical OSSE; Fig. 3) in recent years to derive similar statistics for

- the Coupled Ocean–Atmosphere Mesoscale Prediction System using the NCEP Global Forecast System analysis fields (to replace the nature run),
- the Navy Coastal Ocean Model (by simulating observations from a nature run or using the model data from a different year—but the same month and day), and
- the Navy Global Environmental Model to study impacts of potential observations (e.g., stratospheric ozone) on middle atmosphere prediction.

Among different types of OSSEs, the global forecasting OSSE is mature with extensive peer-reviewed publications, while other types of OSSEs (e.g., the sampling or historical OSSE) are relatively new and need further development. When forecasting OSSEs are performed for specific cases, they are usually called quick OSSEs. While forecasting or quick OSSEs use a nature run to simulate all observations (including observations currently part of the global observing system), historical OSSEs use an alternate model to simulate only the new observations whose impacts are to be estimated and allow for the examination of actual weather events in the historical record (Fig. 3). While historical OSSEs may be useful for understanding specific cases (like quick OSSEs), their use in decision-making for observing system design is cautioned.

The private sector recognizes the value of developing, evolving, and applying OSSEs to inform decisions on investments in observing system capabilities. Assessments are conducted to inform plans and designs for commercial sector observing systems, including making the case to investors for the value of the remote sensing instruments and data streams. Both private sector developers and national centers/laboratories assess the value of alternatives through OSSEs to inform decisions on design alternatives for government systems.

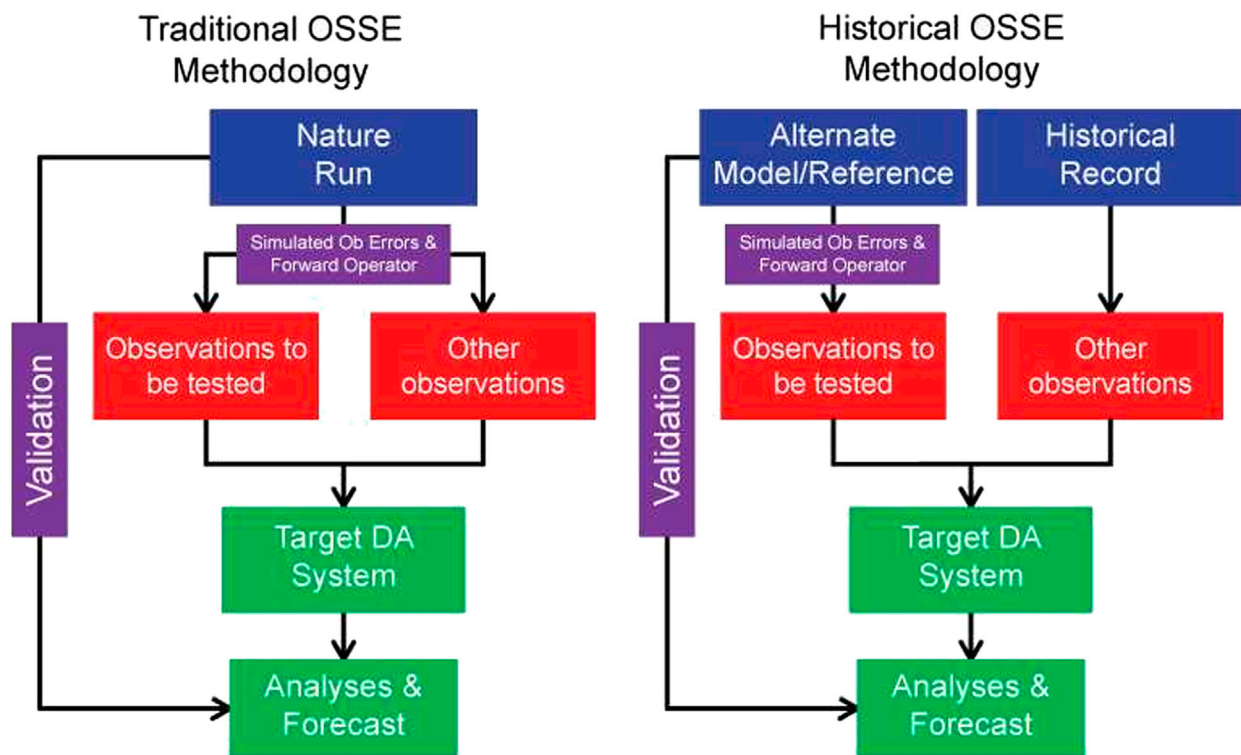


Fig. 3. Schematic depiction of (left) the traditional OSSE methodology (which, when applied to case studies, is known as the quick OSSE methodology) and (right) the historical OSSE methodology currently being explored by NRL.

University efforts are diverse, involving regional and global OSSEs and related activities, including quick OSSE case studies. Two examples (one on regional OSSEs, and the other on a tool that can be used in conjunction with OSSEs) are given below, and additional discussions are included in the short list of references at the end of the “Recommendations on potential actions related to OSSEs” section. A comprehensive summary of such efforts, including discussion of advantages, disadvantages, and special difficulties in performing robust regional OSSEs, is beyond the scope of this paper.

Besides OSSEs, operational or research data assimilation systems (e.g., at NCEP) can provide real-time assessments of the sensitivity of the final analysis to the individual observations, or more commonly collections of observations (e.g., all microwave radiances), used in the analysis. This is known as the FSO method. This was first done with adjoint data assimilations systems—because of the use of the tangent linear model, this approach is limited to short-range forecast impacts (e.g., less than 1–2 days). FSOs can also be conducted using ensemble Kalman filter data assimilation systems. These EFSO methods use ensemble perturbations and future observations to evaluate sensitivity to observations. For instance, at the University of Maryland, EFSO research has leveraged the fact that FSO and EFSO sensitivity can determine whether each observation is beneficial or detrimental (e.g., during the 6-h forecasts). Furthermore, University of Maryland has used EFSO to develop a fully flow-dependent quality control, called Proactive Quality Control (PQC). This is able to identify and then delete, for example, the 10% most detrimental observations, resulting in large forecast improvements. A collection of detrimental observations can also facilitate improving the observation and quality control algorithms. Combining OSSEs with FSO or EFSO provides more information about each observing system, and, as an example, OSSE + EFSO/PQC is found to be much more effective and useful than OSSEs alone in University of Maryland studies.

Values and limitations of OSSEs

As discussed in the “Brief review of OSSE activities in the United States” section, the values of OSSEs have been demonstrated at NOAA, NASA, and NRL in evaluating the impact of new observing systems on operational forecasts when actual observational data are not available. OSSEs are also valuable for testing new data assimilation methodologies and for testing observation targeting strategies. They have been further enhanced by

- combining OSSEs with EFSO (or FSO): OSSEs can see the forecast impact of one particular (proposed) observing system, but they cannot determine the individual impact of all the other (simulated) observing systems, which EFSO (or FSO) can do;
- using different approaches to replace the nature run—as any given nature run may have specific deficiencies (e.g., possibly using historical data in case studies);
- using a variety of OSSEs (for forecasting, sampling, and retrievals)—as forecast OSSEs may not be able to address the impact of observations in answering specific science questions; and
- using OSEs for current observation systems (e.g., for calibration of OSSEs).

While most OSSEs are done for global satellite systems, they can be used to assess new observing systems for regional scales as well. For instance, on the storm scale, OSSEs have been conducted at the University of Oklahoma to assess the value of using dual-polarization radar data in NWP, as well as to assess different scanning strategies and network configurations. On the conterminous U.S. scale, OSSEs can be used to assess the value of increasing the density of vertical profiling systems, as recommended by the National Academies “Network of Networks” Report. Besides OSSEs, several National Academies reports have recommended

that new observing systems be deployed in regional testbeds (including urban testbeds) for evaluation using OSEs, before investing in a nationwide system.

Regarding the limitations of OSSEs, their reliability and effectiveness depend critically on the data assimilation methodology and the forecast models. In particular, the relative impacts of different observation systems may depend critically on the data assimilation system. For example, the NCEP operational data assimilation system in 2018 demonstrated no quantifiable impacts of cloud/rain-affected satellite radiances in the operational model performance while the more advanced European Centre data assimilation and prediction system in 2018 put such radiances as the most impactful source of observations.

The reliability and effectiveness of OSSEs also depend on the estimated error characteristics of a new observing system. As the true error characteristics would be unknown prior to launch, retrieval and sampling OSSEs can help specify realistic errors to a certain degree. This issue can also be addressed by running OSSEs for the best and worst case scenarios, with the former usually taken to be the expected errors prior to launch and the latter more challenging to define (as it depends on “unknown unknowns”).

The use of extreme events (e.g., a major hurricane event), including their nature runs, as truth for OSSEs may have utility for understanding specific cases (also called quick OSSEs). However, we strongly caution against their use in decision-making for observing system design, as the skill scores for (especially global) operational NWP models are judged by a large number of cases or seasons, and by many metrics, not by a single event. In general, individual events in the real world or in nature runs can have case- and flow-dependent predictability, which will have significant impacts on the assessment of certain observing systems. A continuous long nature run, on the other hand, is likely to drift away from a truthful representation of Earth’s natural processes due to errors in the model physics and dynamics and their interactions, or in the boundary conditions.

While OSSEs for the global atmosphere are mature, further development of OSSEs for the ocean and other Earth system components is needed. In particular, despite its major role in the Earth system, the ocean is sparsely observed. The lack of observations stems in part from the technical challenges of sustaining observations of the ocean and from the cost of maintaining observing arrays and networks across the ocean basins. In this context, OSSEs, especially combined with EFSO, are a valuable tool to inform those that fund ocean observations about the impact of specific observations on model fields. Additional ocean OSE efforts are also valuable. For instance, the large, cooperative EU project AtlantOS will focus on a forward design for basin-scale in situ observations, with a quantitative focus informed by OSE/OSSE work.

Because ocean models are deficient in capturing important modes and variability present (partly due to a lack of observations), the conclusions from ocean OSEs and OSSEs about the impacts of observing system elements need to be questioned and considered with care. For instance, recent community efforts indicate that tropical Pacific OSE/OSSE studies are expensive (usually) and often inconclusive, in large part due to the large systematic errors in models and dependence on parameterization assumptions. Therefore, multiple lines of evidence are encouraged to support expected impacts (e.g., for the current and proposed future Tropical Pacific Observing System).

Recommendations on potential actions related to OSSEs

NOAA is mandated to perform OSSEs by the Weather Research and Forecasting Innovation Act of 2017, Section 107. Other organizations (e.g., NASA and NRL) also perform OSSEs regularly. Indeed, OSSEs have been successfully used in major decision-making in the past in the United States. For instance, there was a proposed data buy where NASA and NOAA were each required to spend \$150 million to buy a particular type of data. A joint NOAA–NASA

OSSE was performed to determine the data requirements for this observing system. It was determined that the minimum requirements to ensure a beneficial impact on weather prediction for this observing system could not be met, and it was not in the nation's best interest to procure those data.

Based on the findings in the “Brief review of OSSE activities in the United States” section and discussion in the “Values and limitations of OSSEs” section, here are our recommendations on potential actions related to OSSEs in the United States:

Recommendation 1: OSSE, OSE, FSO, EFSO, and PQC research efforts should be coordinated nationally (e.g., sharing of software tools) to avoid duplication of effort (e.g., via the QOSAP program). Each method has its pros and cons, and all should be used to assess the relative benefit of different observing systems. Besides full-scale OSSE experiments, simple experiments can be very powerful (e.g., for sampling strategies and data value evaluation).

Recommendation 2: The OSSE development for Earth system models (e.g., for sea ice prediction) needs to be accelerated. Furthermore, global nature runs based on Earth system models (at 5 km grid spacing, preferably 3 km, and possibly 1 km) should be developed as the basis for a variety of OSSEs for exploring observation impacts over many different regional domains across the globe. This may require access to high-performance computers or partnerships among agencies. In parallel, studies and other observing campaigns should move forward to resolve processes, develop parameterizations of unresolved processes, and provide the basis for improving the models used in OSSEs.

Recommendation 3: Data assimilation and prediction systems will continue to improve. OSSEs are used to evaluate the observational networks of the future, sometimes decades from now. Therefore, the choice of observations and investment decisions based on OSSEs (and EFSOs) need to consider the potential impact of deficiencies in the current data assimilation and prediction system (e.g., by using the most advanced data assimilation method among different centers).

Recommendation 4: Besides existing OSSE activities in the United States, OSSEs should be used to perform the following:

- Assess the value of partnership in satellite remote sensing with foreign agencies (e.g., India) and the private sector (e.g., purchasing data from privately launched satellites).
- Assist the exploration of strategies for the most effective and efficient approach for sea ice prediction (including observations, models, and data assimilation). For instance, should NOAA request ice-breakers? How many?
- Compare the value of satellite deployment strategies (e.g., a small number of large satellites in geostationary orbit versus a large number of small and cube satellites in polar orbits) for weather, climate, and Earth system prediction.
- Conduct a data need analysis; for instance, what are the greatest new observational needs at NOAA? What combination of old and new systems will work best?

Recommendation 5: OSSEs have been primarily used to evaluate the impacts of observing systems and/or observation denial on forecast performance based on physical parameters, while treating all forecast locations, times, and circumstances as equal. The approach should be extended to assess societal impacts on lives and property. In other words, there are national priorities (e.g., saving lives) where monetary impacts are not the primary consideration, and then there are priorities constrained by financial resources. This could be a possible additional avenue of research. OSSEs based on an Earth system model with social systems and the built environment included would enable us to assess impacts of propagating the physical Earth system information through the social systems.

As an example, while OSSEs provide quantitative analyses of future observing system impacts for a specific model, the effects on products that rely on that model can only be estimated qualitatively. For instance, the NOAA National Environmental Satellite, Data, and Information Service Technology, Planning and Integration for Observation division has developed a qualitative assessment tool for supporting investment decisions, called the NOAA Observing System Integrated Analysis, version 2, also known as NOAA's Value Tree. This Value Tree is based on the survey of subject matter experts across all NOAA organizations to gauge the impacts of Earth observation investments on NOAA's key products and services. Therefore the aforementioned OSSE, OSE, FSO, EFSO, and PQC tools should be used in concert with the current integrated analysis system to determine NOAA's future observing needs.

As mentioned earlier, OSSEs have great power for cost-effectively (when compared to the cost of large observing systems) and rapidly exploring the impact of the relative contributions made to NWP by a wide range of observing technologies—and indeed providing insights into a number of observing configurations that might be prohibitively expensive and time consuming to develop by any other means. On the other hand, attention should not be confined to OSSEs to the exclusion of other research and development activities, such as

- actual deployment and use of observing technologies in pilot programs and demonstration projects,
- complementing advances in NWP per se with corresponding improvements in public risk communication and the use of new technologies such as data analytics and artificial intelligence,
- basic social science research toward similar ends, and
- research and development in valuing weather information.

The opportunities—and the public stakes (with respect to health and safety and building resilience to hazards; development of renewable natural resources; and protecting the environment and ecosystems)—are so high and so urgent as to demand a national pursuit of all these diverse research and development and technology transfer paths in parallel, rather than in sequence or in isolation. More attention to OSSEs and development of their potential is needed, but in a manner balanced by additional attention to other opportunities across the board.

Perhaps the greatest benefit of research and development on OSSEs goes beyond the guidance they can provide by themselves with respect to any particular observing system development and deployment decision. Instead it is about the enriched perspective they provide regarding strategic approaches to investment in Earth observations, science, and services in support of the national agenda. There is an analogy to the famous statement “Individual plans are worthless, but planning is vital” as quoted by Eisenhower and others.

It needs to be emphasized that the discussions and recommendations in this essay do not represent the positions or perspectives of NOAA nor any other organizations. Also note that this essay provides an introductory discussion of OSSEs, based on the collective expertise and experience of the authors, rather than a comprehensive review of all relevant publications and practices. For readers who desire more information on a particular OSSE-related topic, a short list of relevant publications with a brief synopsis is provided here for further reading:

- OSSEs for global NWP applications (most prior studies are in this category):
 - Atlas (1997) provided an overview of OSE and OSSE methodology for global NWP studies in the 1980s and 1990s, and evaluated the relative utility of the principal atmospheric observing systems and the potential for new observing systems.

- Boukabara et al. (2016) documented the Community Global OSSE Package for short- to medium-range global NWP applications. This system is designed to evolve, both to improve its realism and to keep pace with the advance of operational systems.
- Hoffman and Atlas (2016) discussed how OSSE systems need to evolve as operational forecast and data assimilation systems evolve and with the expected development of new observing systems. It also provided a detailed checklist to guide the design of an OSSE system and OSSE experiments.
- OSSEs for regional and local weather:
 - Xue et al. (2006) represents one of the first OSSE studies in storm-scale OSSEs by evaluating the impact of data from radar networks on thunderstorm analysis and forecasting.
 - Atlas et al. (2015) summarized early applications of global OSSEs to hurricane track forecasting and new experiments using both global and regional models.
- Oceanic OSSEs:
 - Halliwell et al. (2014) developed a new ocean OSSE system in the open Gulf of Mexico that has followed, for the first time, the complete set of design strategies and rigorous validation techniques developed for the atmosphere.
 - Kamenkovich et al. (2017) represents one of the recent OSSE studies in this area by assessing the number of profiling floats in the Southern Ocean that is needed for the reconstruction of biogeochemical variables.
- FSO/EFSO:
 - Hotta et al. (2017) discussed EFSO that enables the quantification of how much each observation has improved or degraded the forecast. In particular, an EFSO-based fully flow-dependent quality control scheme was developed.

Acknowledgments. We thank Brad Colman, John Snow, Susan Avery, the editor (Robert Fovell), and three anonymous reviewers for helpful and insightful input that has substantially improved the readability. We dedicate this work to the memory of coauthor Fuqing Zhang.

Appendix: Acronyms

EFSO	Ensemble forecast sensitivity to observations
FSO	Forecast sensitivity to observations
GMAO	Global Modeling and Assimilation Office
NASA	National Aeronautics and Space Administration
NCEP	National Centers for Environmental Prediction
NOAA	National Oceanic and Atmospheric Administration
NRL	Naval Research Laboratory
NWP	Numerical weather prediction
OSE	Observing system experiment
OSSE	Observing system simulation experiment
PQC	Proactive Quality Control
QOSAP	Quantitative Observing System Assessment Program

References

- Atlas, R., 1997: Atmospheric observations and experiments to assess their usefulness in data assimilation. *J. Meteor. Soc. Japan*, **75**, 111–130, https://doi.org/10.2151/jmsj1965.75.1B_111.
- , L. Bucci, B. Annane, R. Hoffman, and S. Murillo, 2015: Observing system simulation experiments to assess the potential impact of new observing systems on hurricane forecasting. *Mar. Technol. Soc. J.*, **49**, 140–148, <https://doi.org/10.4031/MTSJ.49.6.3>.
- Boukabara, S. A., and Coauthors, 2016: Community Global Observing System Simulation Experiment (OSSE) Package (CGOP): Description and usage. *J. Atmos. Oceanic Technol.*, **33**, 1759–1777, <https://doi.org/10.1175/JTECH-D-16-0012.1>.
- Halliwell, G. R., A. Srinivasan, H. Yang, D. Willey, M. Le Henaff, V. Kourafalou, and R. Atlas, 2014: Rigorous evaluation of a fraternal twin ocean OSSE system for the open Gulf of Mexico. *J. Atmos. Oceanic Technol.*, **31**, 105–130, <https://doi.org/10.1175/JTECH-D-13-00011.1>.
- Hoffman, R. N., and R. Atlas, 2016: Future observing system simulation experiments. *Bull. Amer. Meteor. Soc.*, **97**, 1601–1616, <https://doi.org/10.1175/BAMS-D-15-00200.1>.
- Hotta, D., T.-C. Chen, E. Kalnay, Y. Ota, and T. Miyoshi, 2017: Proactive QC: A fully flow-dependent quality control scheme based on EFSO. *Mon. Wea. Rev.*, **145**, 3331–3354, <https://doi.org/10.1175/MWR-D-16-0290.1>.
- Kamenkovich, I., A. Haza, A. R. Gray, C. O. Dufour, and Z. Garraffo, 2017: Observing system simulation experiments for an array of autonomous biogeochemical profiling floats in the Southern Ocean. *J. Geophys. Res. Oceans*, **122**, 7595–7611, <https://doi.org/10.1002/2017JC012819>.
- Xue, M., M. Tong, and K. K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Oceanic Technol.*, **23**, 46–66, <https://doi.org/10.1175/JTECH1835.1>.