

Optimizing Observing Systems Using ASPEN:

An Analysis Tool to Assess the Benefit and Cost Effectiveness of Observations to Earth System Applications

Sid-Ahmed Boukabara and Ross N. Hoffman

ABSTRACT: The Advanced Systems Performance Evaluation tool for NOAA (ASPEN) is developed to help support designing and evaluating existing and planned observing systems in terms of comparative assessment, trade-offs analysis, and design optimization studies. ASPEN is a dynamic tool that rapidly assesses the benefit and cost effectiveness of environmental data obtained from any set of observing systems, whether ground-based or space-based, whether an individual sensor or a collection of sensors. The ASPEN assessed cost effectiveness accounts for the level of ability to measure the environment, the cost(s) associated with acquiring these measurements, and the degree of usefulness of these measurements to users and applications. It computes both the use benefit, measured as a requirements-satisfaction metric, and the cost effectiveness (equal to the benefit-to-cost ratio). ASPEN provides a uniform interface to compare the performance of different observing systems and to capture the requirements and priorities of applications. This interface describes the environment in terms of geophysical observables and their attributes. A prototype implementation of ASPEN is described and demonstrated in this study to assess the benefits of several observing systems for a range of applications. ASPEN could be extended to other types of studies, such as assessing the cost effectiveness of commercial data to applications in all the NOAA mission service areas, and ultimately to societal application areas, and thereby become a valuable addition to the observing systems assessment toolbox.

KEYWORDS: Remote sensing; Satellite observations; Sensitivity studies; Decision support

<https://doi.org/10.1175/BAMS-D-22-0004.1>

Corresponding author: Sid-Ahmed Boukabara, sid.boukabara@noaa.gov

In final form 7 July 2022

©2022 American Meteorological Society

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

AFFILIATIONS: Boukabara—NOAA/NESDIS/Office of Systems Architecture and Advanced Planning, Silver Spring, Maryland; Hoffman—NOAA/NESDIS/Center for Satellite Applications and Research (STAR), and Cooperative Institute for Satellite Earth System Studies, University of Maryland, College Park, College Park, Maryland

The Earth-observing satellite constellation (EOSC), which includes all space-based components of the Global Observing System, currently includes platforms operating in both geostationary Earth orbits (GEO) and low-Earth orbits (LEO) at various inclinations. A number of these sensors are radiometers that measure the microwave, infrared, ultraviolet, and visible spectra. Other instruments include radio occultation sensors, scatterometers, lightning mappers, radars, altimeters, and space weather sensors. Space and solar sensors operate in these LEO and GEO orbits, sometimes hitching a ride on EOSC satellites as well as being stationed at the much more distant, gravitationally stable, Earth–Sun Lagrange points. Together, these observing systems provide for a global environmental monitoring system that covers the atmosphere, biosphere, hydrosphere, oceans, cryosphere, near-space environment, and the Sun (Boukabara et al. 2021). Products are derived on a variety of temporal and geographic scales that are used for a wide range of short- and medium-term warnings, long-term forecasting, and other climate, space weather, ocean, terrestrial weather, and water cycle applications that are critical to the Earth system user community and society in general.

Space agencies around the globe have well-established plans to deploy and exploit Earth-observing satellites for the next two decades (e.g., Simmons et al. 2016). In parallel to these plans, and because of the lengthy processes that often require 15–20 years to design, build, and deploy new Earth-observing satellite systems, some space agencies, including NOAA, have also started formulating the space architecture of the 2030s and beyond. For example, the Geostationary Extended Observations (GeoXO) program¹—with an initial planned launch in 2032—has the responsibility for the evolution of the geostationary component of the NOAA space program in the next-generation architecture. Approaches taken to optimize the architecture range from requirements analysis (e.g., Anthes et al. 2019) to detailed simulation experiments (e.g., Boukabara et al. 2016b).

¹ <https://www.nesdis.noaa.gov/next-generation/geostationary-extended-observations-geoxo>

While, in the past, the evolution of the EOSC was gradual and technology driven, there are currently additional driving factors that add significant complexity to planning and designing the next-generation EOSC. Besides the rapid advances in sensor technology in the recent past, these factors include 1) the expected revolution of medium to very small size satellites that will provide slightly degraded, similar, or even better performance for a fraction of the cost of traditional platforms; 2) the large array of environmental applications, all vying for even more and better environmental data; 3) the diversification of observing methods and technologies that have varying error characteristics and resolutions; 4) the emergence of the private sector as a viable source of environmental data; and 5) the increasing number of new space-faring nations with ambitious space programs that have significant potential to enhance the EOSC. It should be noted that the effective use of very small satellites, including CubeSats, requires overcoming some challenges to data quality (calibration, geolocation, stability) to take advantage of the very positive qualities of these platforms (rapid sensor refresh, low cost, use of open source technology, fast temporal refresh due to multiple numbers, resilience).

Missions, such as the Temporal Experiment for Storms and Tropical Systems (TEMPEST) and the Time-Resolved Observations of Precipitation Structure and Storm Intensity with a Constellation of Smallsats (TROPICS) (Blackwell et al. 2019), should allow us to learn how to best exploit these technologies. Comparisons of such systems to legacy sensors using the methodology described in this paper can account for data quality and geographic and temporal coverage attributes as well as costs, but other considerations, such as risk, launch cadence, development cycle time, and reliability of the sensor manufacturer must also be considered.

The purpose of this paper is to present the Advanced Systems Performance Evaluation tool for NOAA (ASPEN) as an answer to the pressing need to assess and optimize the planning and design of an observing system (including the processing and distribution of observations) given the considerations outlined above. In the sidebar “ASPEN Design Challenges,” we discuss the three critical challenges that the design of a new assessment tool would have to overcome, namely, 1) the need for a solution-agnostic metric when optimizing observing systems, 2) the need for a comprehensive and inclusive assessment using an Earth system approach, and 3) the need for granularity in such assessments in order to capture the influence of the technical details of the observing systems characteristics.

ASPEN is designed to measure the *degree to which one or more observing systems are able to satisfy the needs of one or more environmental applications*. In essence, ASPEN is a performance/gap analysis tool that measures how much an observing system fulfills the (prioritized) requirements ranges of the applications. The higher this degree of satisfaction, the more the observing system is considered beneficial. ASPEN relies on a “universal” representation of the Earth system as a whole and recognizes that 1) an observing system will measure a portion of that Earth system environment and 2) an application expects (i.e., requires and prioritizes) a certain level of knowledge of the same Earth system environment to properly operate. It is the comparison between the application requirements for knowledge of the Earth system environment and the actual observing system performance to deliver that knowledge—both rigorously captured in the observing system performance table and the application requirement ranges table—that allows ASPEN to associate application-relative scores to the observing systems being assessed. The combination of the prioritized scores gives the ASPEN benefit metric. For an observing system, performance can be either planned (i.e., during the design phase) or realized (i.e., during the calibration/validation phase). In some cases the realized performance exceeds the planned performance, and in some cases the reverse holds. Care must be taken when interpreting a comparison of a proposed observing system to an existing one depending on whether the existing observing system performance is planned or realized.

ASPEN is designed to be solution agnostic, comparing observing systems’ geophysical performance to the geophysical requirements of applications and weighing the resulting normalized scores by appropriate priorities. ASPEN is therefore heavily dependent on its input information. Specifically, ASPEN consider two types of inputs: 1) observational requirements and priorities coming from applications and users of the observations, and 2) observational performance coming from observing systems (whether sensors, networks, constellations, or other combinations). These inputs—requirements and associated priorities, as well as the observing systems performance—are captured in tables and expressed in terms of geophysical variables (needed by applications and/or provided by observing systems) and their corresponding attributes (such as spatial coverage, temporal refresh, and uncertainty).

Once the performance of any observing system is put in the right format, ASPEN treats these the exact same way, rigorously. The observing system could be satellite-based, ground-based, or airborne, or even from citizen scientist networks. All that matters is the observing system performance (in the correct format and units). This is demonstrated for radiosondes and several satellite sensors in this paper. Thus, although we show only a few simple example calculations

ASPEN design challenges

The need for a solution-agnostic metric when optimizing the observing systems network

Achieving a fair comparison between different observing systems solutions requires truly solution-agnostic metrics. This will allow a fair comparison and assessment of the complementarity of different observing approaches that provide, for example, temperature sounding (with a certain precision, resolution, spatial coverage, and temporal refresh) from one or more microwave, infrared, and radio occultation sensors. In ASPEN, the observing system benefit assessment is done by comparing the observing system performance defined in terms of geophysical variables and their attributes against the application requirement ranges. This opens the door to the democratization of observing systems assessment, allowing potentially novel, multiuse, and cost-effective solutions.

The need for a comprehensive and inclusive assessment using an Earth system approach

National hydrological and meteorological services require research and operational components to observe the Earth system from a variety of observing systems; to collect, archive, and disseminate these observations and use them in numerous models; to characterize and predict the Earth system and its environmental components; and to use these analyses and forecasts to provide a variety of services to individuals, businesses, and government agencies. Such complex missions demand a comprehensive, Earth system-oriented, assessment approach for planning next-generation observing system architectures. This approach should consider all user needs and all technology options. Also in this process, programmatic and strategic metrics such as system costs, launch cadence, technology availability and technology maturity, and overall risk and reliability must be compared to the agency desired postures for these factors.

ASPEN is designed to provide critical support to such a process. Once the inputs have been prepared, ASPEN is dynamic and nearly instantaneous to execute, and allows side-by-side comparisons of the benefit and cost effectiveness of multiple design options of observing systems to a comprehensive set of applications, representative of the meteorology, oceanography, land/hydrology, and space weather environmental domains. ASPEN therefore offers a path to a second democratization, this time by opening the assessment to include all uses of the observations.

The need for granularity in an assessment approach

Assessing observing systems accurately and capturing the fine details of their differences requires that the assessment approach be sensitive to a degree of granularity that will allow the distinction of performance differences due to differences in the antennas sizes, or noise levels, or scanning geometry, etc. ASPEN was developed to simultaneously account for all the technical factors above by projecting observing system characteristics into geophysical space performance (as described in the “ASPEN methodology” section) and assessing the geophysical performance against set requirements, also in geophysical space. It provides numerical metrics of the benefits of existing and proposed observing systems—benefits to a single application, a group of similar applications, or a collection of applications, representative of all mission applications. In addition to the benefit due to technical factors, ASPEN also quantifies cost effectiveness as the ratio of benefit to cost. Because the assessment is sensitive to very fine details of the observing systems (including precision/accuracy, spatial and vertical resolution, temporal refresh, and spatial coverage), it can be used for many purposes (see Table 6).

here for illustration, ASPEN is capable of comparing the relative benefits of satellite sensors, ground-based remote sensing systems, and in situ networks of sensors.

The benefits computed by ASPEN could be the benefit to a single application, to groups of applications, or to all the NOAA mission service areas, therefore making it possible to account for the entire Earth system when performing these benefits assessments. In future studies, ASPEN is intended to be used to assess the complementarity of space-based and ground-based observing systems. In the conclusions we describe both the limitations and potential of ASPEN, including potential enhancements to extend ASPEN for trade studies, applications to all the NOAA mission service areas, and ultimately to societal application areas.

The ASPEN concept and metrics

ASPEN methodology. ASPEN provides rapid computations of relative benefits (hereafter benefits), based on the scoring (or normalization) and prioritization (or weighting) described below, of observing systems' performance by the application requirement ranges and priorities. To accomplish this, the computational flow in ASPEN follows Fig. 1. The observing systems performance and the application requirement ranges and priorities are all input tables (second column of Fig. 1) derived from expert elicitation or from calculations conducted for maturity reviews of existing sensors, simulation studies, or sensitivity experiments (first column of Fig. 1). This approach allows us to compare benefits across many possible observing systems, including, for example, multiple permutations of constellations of Earth-observing satellites. Given costs (or relative costs) of the observing systems that are compared, ASPEN also calculates cost effectiveness as the ratio of benefit to cost. ASPEN calculates these metrics (benefit and cost effectiveness) over a wide range of granularity. As described in detail in what follows, benefits at one level of granularity are combined (rolled-up) to the next (less granular) level by means of weighted averages. (Please note that in this discussion, unless otherwise stated, mentions of observations and applications refer to geophysical or ecological observations or applications.)

The ASPEN concept is underpinned by a description of the Earth system environment that is independent of both observing systems and applications. This solution-agnostic and independent frame of reference is critical in achieving a fair assessment and in allowing a single description of application requirements to be compared against many different observing systems. The ASPEN "universal" interface represents the Earth environment components in terms of geophysical variables and attributes of those variables. As such, the ASPEN interface is independent of how sensors or applications work in practice. For example, even though numerical weather prediction (NWP) applications may assimilate radiances directly, for ASPEN,

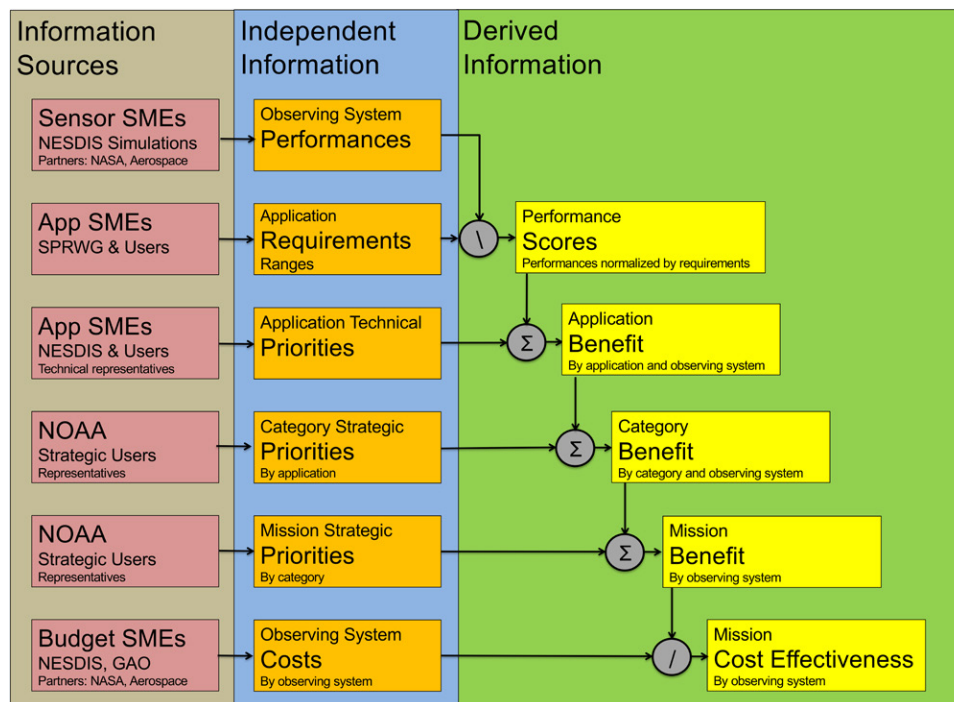


Fig. 1. The flow of information in ASPEN is illustrated, showing from left to right the sources of the independent information (SMEs indicate subject matter experts), the independent information (the ASPEN inputs), and the derived information (the ASPEN outputs). Operators in gray circles include normalization (\backslash), weighted average (Σ), and division ($/$). NESDIS is the National Environmental Satellite, Data, and Information Service. GAO is the Government Accountability Office. Refer to the text for a full description.

the NWP requirements are specified in terms of temperature and humidity since the radiances measured (and assimilated) are tuned to provide temperature and humidity information. Similarly, the performance of a microwave sounder, for example, should also be specified in terms of temperature and humidity since its radiometric measurements (of radiances) are designed to contain temperature and humidity information. Thus, sensor performance in terms of radiances must be converted into geophysical space using a science-based remote sensing approach, such as applying a retrieval algorithm (e.g., Boukabara et al. 2011; Maddy and Boukabara 2021) to a large database of atmospheric profiles. Ideally, the same or similar retrieval algorithm is used for similar sensors so that the relative performances are accurately quantified. Thus, the ASPEN approach assumes the radiances or the brightness temperatures and related geophysical variables (either provided by the observing systems or required by the applications) are projections of each other through the retrieval algorithm or its associated forward problem. However, it should be noted that the retrieval algorithm and its associated forward problem are nonlinear and sometimes require additional a priori information.

The primary ASPEN calculation is to determine the benefit to an application of an observing system. Note that observing system here really refers to observations and/or products emanating from them. In this way, ASPEN is applicable to assessing and prioritizing data and products that are acquired commercially and/or to assessing and prioritizing products generated by federal programs. This is in addition to its main purpose of assessing and optimizing observing systems designs. We will refer to the benefit to an application of an observing system simply as the application benefit. For this calculation, the observing system performance (OSP) in terms of various attributes for different variables is tabulated to capture how well the geophysical variables are observed (first row of Fig. 1) and application requirements ranges (ARR) are formulated as tables of ranges of minimally to maximally useful attributes for each required variable (second row of Fig. 1). Examples of these tables will be given in the “ASPEN interfaces” section.

Given tables of OSP and ARR, ASPEN scores (i.e., normalized) the observing system performances in terms of variables and attributes by the application requirement ranges (second row of Fig. 1). Considering a single geophysical variable and a single attribute, if the observing system performance is x , and the application requirement range is given as a triplet of the form $[x_{\min}, x_{\text{mid}}, x_{\max}]$, where x_{\min} is the minimally useful value, x_{mid} is an intermediate value (not used in the ASPEN prototype), and x_{\max} is the maximally useful value. Then the performance score y is given by the normalization

$$y = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})}. \quad (1)$$

The result y in (1) is truncated to the range $[0, 1]$. This truncation is required because ASPEN measures how well an observing system satisfies the requirements of an application, and this benefit must be bounded between zero and one. In terms of ASPEN supporting the design of observing systems to satisfy the users/applications requirements, the observing system should be neither under- nor overengineered. With the $[0, 1]$ truncation applied, there is no bonus given to an observing system that exceeds the maximum performance levels required by the applications. One important attribute that cannot be usefully normalized by (1) is geographic coverage, for which the performance score is the fraction of the application geographic coverage that is overlapped by the sensor geographic coverage. The geographic regions used in the ASPEN prototype and their overlaps are given in Table 1. This table lists just a few of ASPEN’s large collection of regions, which range from “Polar Regions” to the “U.S. Economic Exclusion Zone” and which could be extended.

In designing sensors and observing systems, we have to not only be aware of the requirements for different observables and their attributes, but also of their relative importance. This helps drive the sensor design in terms of characteristics like antenna size, number of

Table 1. Geographic regions and overlaps used in the ASPEN prototype. The regions in the rows overlap the regions in the columns by the percentages given in the table. CONUS is the continental United States. Meso refers to the moveable GOES mesoscale scan sector.

Region	Area (1,000 km) ²	Global (%)	Tropics (%)	CONUS (%)	Meso (%)	Full disk (%)
Global	511	100	100	100	100	100
Tropics	254	49.72	100	37.76	100	52.91
CONUS	25	4.95	3.76	100	100	11.75
Meso	1	0.20	0.39	3.95	100	0.47
Full disk	215	42.07	44.76	99.81	100	100

channels, and noise levels. In ASPEN, we therefore have to account for the fact that for any given application, some variables are more important than others and for a given variable, some attributes are more important than others. These relative importances are termed application-dependent technical priorities (ATP). Then the application benefit (from an observing system) B is determined (third row of Fig. 1) as

$$B = \sum_i w_i y_i. \quad (2)$$

In (2) the sum is over all variables and attributes and the weights w are the priorities rescaled so that the sum of the w is one. For attributes denoted as fundamental (defined as attributes for which if the requirements are not minimally met, all other attributes do not matter), if the performance score is zero then all other performance scores for that variable are also set to zero during the application benefit calculation. In ASPEN currently only the geographic coverage attribute is considered fundamental since, if no observations fall in the required region, it does not matter how accurate or timely those observations are. (Since all the overlaps in the ASPEN prototype are greater than zero, fundamental attribute processing is not active in the sample calculations shown in this paper.)

The individual application-specific benefits for an observing system are combined in a weighted average as in (2) using strategic priorities into benefits for entire categories of applications or for all categories of applications (fourth and fifth rows of Fig. 1). These strategic priorities are termed the category-specific priorities and mission priorities, respectively, and these summary benefits are termed the category benefits and the mission benefits, respectively. Users of ASPEN therefore have the ability to assess the relative value of observing systems at the application level, at the *category of applications* level or at the entire mission level.

ASPEN estimates the cost effectiveness of an observing system to an application, category of applications, or all applications by dividing the corresponding benefit by the observing system cost. This is shown only for the mission cost effectiveness in Fig. 1 (sixth row), but the same calculation is also done to determine the application cost effectiveness and the category cost effectiveness.

Based on the above calculations, ASPEN generates a number of outputs, including 1) side-by-side stratified and combined assessments of the benefit and cost effectiveness of various observing systems candidates, 2) intermediate result matrices to provide the traceability and interpretability linkages between the ASPEN metrics (or scores) and the different inputs (next subsection), and 3) efficient frontier plots displaying the benefits from a set of observing systems being assessed, against their costs, to allow us to better select and optimize the most highly cost-effective candidates among them.

Sensitivity, traceability, and interpretability of ASPEN results. ASPEN provides various ways to aid the decision-making process when designing an observing system and in general

to interpret the ASPEN results. These analyses are based on the transparency that is built into the ASPEN design and include 1) simple sensitivity studies, 2) the ability to trace the resulting scores (i.e., particular values of metrics) back to the various inputs, and 3) interpretability analyses, which link ASPEN results directly to observing system design decisions. These tools apply to all ASPEN metrics, including benefit metrics that measure the degree of satisfaction of the application requirements accounting for the (application, category, and mission) priorities and cost-effectiveness metrics that are simply the benefit metrics divided by observing system costs. We briefly describe the three analysis methods in the following paragraphs.

SENSITIVITY ANALYSIS. ASPEN offers the opportunity to easily perform “what if” exercises. It is easy to modify one or more ASPEN inputs (the geophysical characteristics of a specific observing system, as well as the requirements ranges and prioritizations imposed by the applications), redo the calculations, and then examine the sensitivity of any of the ASPEN results to those inputs. This allows the analyst to identify the most influential parameters and to quantify their impact on the ASPEN scores. This facilitates the decision-making process to concentrate efforts and resources in the most optimal way.

TRACEABILITY ANALYSIS. All of ASPEN scores are easily traceable to the ASPEN inputs because all ASPEN intermediate results are readily available and easy to query. The key intermediate results are the performance scores calculated from Eq. (1) and the performance scores multiplied by appropriate weights. This follows since all the ASPEN metrics can be put in the form of a weighted sum of the performance scores. All the weights can be specified in terms of the (application, category, and mission) priorities and the observing system costs. It is only a matter of bookkeeping to determine the weights for each performance score for any metric. All these intermediate results are clearly related to the ASPEN inputs, so we can either track how an input propagates forward through ASPEN or trace back how an output is determined by ASPEN. This allows the analyst 1) to understand the logic of why the scores are what they are and 2) to justify or explain any potential decision made based on ASPEN and link it to actual predetermined requirements and prioritization as well as to observing systems performance.

INTERPRETABILITY ANALYSIS. This analysis allows us to link ASPEN scores to decision-making. It answers the question: Given the scores obtained by ASPEN, how could these influence the design of an observing system? The individual terms in the weighted sums that add up to an ASPEN metric are the most granular contributions to that metric. We can sum these contributions for any subset of interest to determine a less granular contribution. For example, if we sum over all attributes for a geophysical variable, we obtain the contribution to the metric due to that variable (for a given application and sensor). After doing this for each variable, we can create plots showing the contributions to the metric due to the different variables. (An example is shown later in Fig. 7.) If we first divide the contributions by the metric itself, the result is the fractional contribution to the metric, which may likewise be summed over any subset contributing to that metric. Fractional benefit contributions and fractional cost-effectiveness contributions are identical since the divisions by cost cancels out in calculating the fractional cost-effectiveness contributions. At the top level, the ASPEN metrics inform us about the degree to which the observing systems being assessed meet application requirements. The interpretability analysis, by ranking the internal contributions to the metrics from different variables and attributes for different applications, allows us to easily point to which variables and attributes one should invest in to further increase the observing system performance to meet an even higher percentage of the requirements, thereby helping us optimize the design of these systems.

ASPEN interfaces

In ASPEN, the Earth system environmental variables are divided into the following Earth system domains: (i) atmosphere, (ii) biosphere, (iii) ocean, (iv) hydrosphere, (v) cryosphere, and (vi) space. Since an individual sensor might contribute information about several Earth system domains and an application might require information from several Earth system domains, this division of variables into environmental domains is primarily a convenience. Note that the environmental domains listed here are separate from the application categories defined later in this section. For each environmental domain, the ASPEN tool has several (dozens of) variables to represent that domain. For simplicity, we chose to use five key geophysical variables per domain for the purpose of this paper. These are identified with asterisks in Table 2.

The attributes of the geophysical variables are related to their geographic, vertical, and temporal coverages; their horizontal and vertical resolutions; their precision/accuracy; and their availability (e.g., timeliness). The choice of attributes must be equally valid for descriptions of the environment derived from observing systems, as well as for descriptions used in models and other applications. The principal attributes are discussed by Boukabara et al. (2021). The attributes included in ASPEN are defined in Table 3. Observational uncertainty is taken to be a composite standard deviation of the error over the appropriate subsets or dimensions such as over the vertical domain, over clear and cloudy conditions, and over different surface backgrounds. Vertical resolution is taken to be the number of independent pieces of information (the degrees of freedom or d.o.f.) in the vertical at one horizontal location. The attribute “images” indicates whether the observations are taken or used as imagery. Note that imagery will generally have faster refresh and higher horizontal resolution than other data sources.

ASPEN is driven by data elicited from subject matter experts (SMEs) and strategic users, from simulations, and, in the case of strategic priorities, from agency leadership (see Fig. 1 for a breakdown). As part of its development, ASPEN centrally captures information from existing sets of performances and requirements. The performance and requirements sources include those from the Space Platform Requirements Working Group (SPRWG) (SPRWG 2018; Anthes et al. 2019; Maier et al. 2021), the Consolidated Observing User Requirements List (COURL) (Murray et al. 2008), the WMO Observing Systems Capability Analysis and Review (OSCAR) database (WMO 2019), the GeoXO Requirements Working Group (XORWG), and the NOAA Observing System Integrated Analysis (NOSIA) (Helms et al. 2016; St. Germain 2018). In some cases, these performances and requirements are the subject of regular updates and refinements.

In particular, over a 2-yr period beginning in 2015, the SPRWG panel of subject-matter experts carried out an analysis and prioritization of different space-based observations supporting NOAA’s operational services in the areas of weather, oceans, and space weather. NOAA leadership used the SPRWG analysis of space-based observational priorities in different mission service areas, among other inputs, to inform the multi-attribute utility theory (MAUT)-based value model and the NOAA Satellite Observing Systems Architecture (NSOSA) study. The goal of the NSOSA study was to develop candidate satellite architectures for the era beginning in approximately 2030. The SPRWG analysis included a prioritized list of observational objectives together with the quantitative attributes of each objective. These results helped inform NOAA’s assessment of many potential architectures for its future observing system within the NSOSA study. This has led, for example, to some high-level recommendations regarding the NOAA future architecture, including recommendations for a disaggregated approach for the LEO architecture and the exploration of smallsats to possibly support that approach.

In a sense, the SPRWG goals and studies anticipated and motivated the development of the ASPEN tool. One of the goals of the ASPEN development is to have a generic enough and

Table 2. Selected geophysical variables describing the Earth environment. An asterisk indicates the variable is included in the sample calculations presented below. Only selected variables are listed here.

	Geophysical variable	Symbol	Units
Atmosphere	Aerosol concentration	η_A	m^{-3}
	* Air temperature: Profiles	$T(z)$	K
	Carbon dioxide/ CO_2	AC_CO2	ppm
	Incoming shortwave radiation: Surface	ISR	W m^{-2}
	* Relative humidity	RH	%
	* Total lightning	TL	%
	* Wind speed profile: Eastward	u	m s^{-1}
	* Wind speed profile: Northward	v	m s^{-1}
Biosphere	* Fire radiative power	FRP	MW km^{-2}
	* Fire size and location	Fire_size	km
	* Flood standing water: Extent	FI	%
	Land surface albedo	LSA	—
	Land surface temperature	LST	K
	* Normalized difference vegetation index	NDVI	—
	* Soil moisture: Surface wetness	W	$\text{m}^3 \text{m}^{-3}$
	Surface type	ST	—
Cryosphere	Ice surface temperature	IST	K
	* Sea ice age	SIA	years
	* Sea ice concentration	SIC	%
	* Sea ice motion, local	SIM_local	m s^{-1}
	* Snow cover	SC	%
	Snow depth	SD	m
	Snow grain size	SGS	mm
	* Snow water equivalent	SWE	cm
Hydrosphere	Cloud-base height	CBH	km
	* Cloud cover	CC	Fraction
	Cloud drop size (at cloud top)	CPS	μm
	* Cloud liquid water path	LWC	g m^{-2}
	* Cloud-top temperature	CTT	K
	Precipitation rate/snowfall rate	SFR	mm h^{-1}
	* Rain rate	RR	mm h^{-1}
	* Total precipitable water	TPW	mm
Ocean	Bathymetry	B	m
	* Global sea surface wind direction	SSWD	$^\circ$
	* Global sea surface wind speed	SSWS	m s^{-1}
	Ocean color: Chlorophyll-a concentration	Chl	mg m^{-3}
	Salinity	S	PSU
	* Sea surface height	η	cm
	* Sea surface temperature	SST	K
	* Wave height	h	m
Space	Geomagnetic field: MEO	mag_MEO	nT
	Interplanetary solar wind	part_Low	$(\text{cm}^2 \text{s sr keV})^{-1}$
	Multispectral auroral imaging	aurora_image	W m^{-2}
	Solar flux: EUV	sun_flux_EUV	W m^{-2}
	Solar flux: X-ray irradiance	sun_flux_Xray	W m^{-2}
	Solar imagery: Corona	sun_corona	W m^{-2}
	Solar imagery: Magnetogram	Sol_Im_Mag	nm
	Solar imagery: Whitelight	Sol_Im_White	W m^{-2}

Table 3. Attributes of a description of the Earth environment. An asterisk indicates the attribute is included in the sample calculations presented below.

Attribute	Symbol	Units	Definition
Images	I	logical	This parameter is represented by imagery
* Geographic coverage	D	—	Geographic region observed
* Horizontal density	N_A	$(100 \text{ km})^{-2}$	Number of observations within swath per 100 km square region.
* Horizontal resolution	δ_x	km	GIFOV or ground-projected instantaneous field of view
* Temporal refresh	T_R	h	Time between observations at a location
Vertical extent bottom	E_b	km	Bottom of vertical region observed
Vertical extent top	E_t	km	Top of vertical region observed
* Vertical resolution	N_v	d.o.f	Independent pieces of information in one GIFOV
* Error standard deviation	s.d.	units	Estimated standard deviation of the errors
Validity range low	V_l	units	Low value that can be observed
Validity range high	V_h	units	High value that can be observed
Robustness	N_s	—	Number of sources making this observation
Continuity	T_c	years	Time for which the observations can be intercalibrated
* Data latency	T_L	h	Time from "image taken" to full relay of data to a ground station

solution-agnostic approach so we can fairly optimize the overall constellation, but with a repeatable, flexible tool where we can easily test “what if” scenarios. The observing systems assessed by ASPEN could be of any type—ground-based or space-based, passive or active, microwave, infrared, or ultraviolet, in geostationary or low-Earth orbit, from sensors acquired and deployed or from commercially acquired datasets—as long as we can develop an appropriate observing system performance table.

For the purpose of illustration, only a restricted set of observing systems, mostly space-based sensors, are included in the examples presented. The sensors used were chosen to be representative and are listed in Table 4. Costs must be associated with each sensor and could include launch costs, ground system costs, and other costs in addition to the contract costs, and the interpretation of the cost effectiveness determined by ASPEN should account for which costs were included in the calculation. The observing system costs in Table 4 are the annualized costs for a single sensor based on public information along with some reasonable assumptions (described in the figure caption). These are not definitive, but are reasonable and used here only to demonstrate the calculation of ASPEN cost-effectiveness metrics. They should not be taken at face value as the true costs of these sensors.

As described in the section “The ASPEN concept and metrics,” the ASPEN interface documents the performance of an observing system to provide information about different geophysical variables (Table 2) in terms of a set of observation attributes (Table 3) in an observing system performance table. This table has entries describing the performance of the observing system for all relevant geophysical variable (rows in the table) and attributes (columns in the table). Figure 2a displays the Advanced Temperature and Moisture Sounder (ATMS) performance table that is used in the sample calculations. In Fig. 2 only the symbols are used to label the rows and columns. Please refer to Tables 2 and 3 for the definitions of the symbols used here and in the figures that follow. For example, in this figure, row 1 for temperature and row 2 for relative humidity indicate that ATMS retrievals of these quantities have error standard deviations of 2 K and 20%, respectively, and 8 and 5 degrees of freedom in the vertical, respectively.

The confidence in the observing system performance table entries, e.g., for the observation error standard deviation or degrees of freedom in the vertical, is also collected by ASPEN in

Table 4. Sample sensors. To determine the observing system costs (million U.S. dollars), the total program costs [\$18.8 billion for the Joint Polar Satellite System (JPSS), <https://www.jpss.noaa.gov/faq.html>, and \$10.8 billion for GOES-R, <https://www.goes-r.gov/resources/faqs.html>] are simply divided (i) by the number of satellites (5 for JPSS and 4 for GOES-R), (ii) by 5 assuming each satellite is active for 5 years (although in fact the GOES-R operational design lifetime is 10 years), and (iii) proportionally by the sensor contract costs (assessed from public sources) and for the JPSS sequence—*Sumoi NPP* through *JPSS-4*—reported by <https://www.jpss.noaa.gov/news.html>). Costs for radiosondes are for twice daily launches of a global network of 800 radiosondes (locations) available in near-real time (<https://www.ncei.noaa.gov/products/weather-balloon/integrated-global-radiosonde-archive>) at \$200 per launch (<https://www.weather.gov/media/key/Weather-Balloons.pdf>). These costs are very rough estimates and are only meant to be used for illustration purposes.

	Sensor	Sensor type	Program	Orbit	Cost
*	ABI	Imager	GOES-R	GEO	314
	EXIS	EUV X-ray sensor			32
*	GLM	Lighting mapper			92
	MAG	Magnetometer			7
	SEISS	In situ particle flux sensor			30
	SUVI	Solar UV imager			65
*	ATMS	MW	JPSS	LEO	111
*	CrIS	IR			199
	OMPS	UV			120
*	VIIRS	Imager			322
*	Raob	Radiosondes	Global		117
	Raob	Radiosondes	NWS		13

the form of uncertainty and then used in Monte Carlo calculations to map the uncertainty in the ASPEN inputs to the uncertainty in the ASPEN outputs. This, in theory, should give us a sense of the reliability of the ASPEN output. This applies to uncertainties of inputs both for observing systems and for applications. The uncertainties can come from various information sources, including from the confidence of individual SMEs, from the dispersion of the estimates from multiple SMEs, and from the uncertainties derived from sensor simulation studies. Since ASPEN is computationally efficient, the size of the Monte Carlo ensemble used is large enough to provide robust uncertainty estimates for all of the ASPEN metrics. It is, however, recognized that capturing representative uncertainties and their correlations for all ASPEN inputs is quite challenging and remains only partially met.

As mentioned previously, each application has different requirements and priorities for what variables are needed and for each variable and each attribute a range of useful values. For this purpose, ASPEN collects application requirement ranges and application priorities in tables that are similar in shape to the observing system performance tables.

The applications are grouped into five application categories—Earth system, meteorology, oceanography (including fisheries and coastal applications), land (including hydrology and ecology), and space weather. For each application category, for simplicity in this study only a few representative application areas are listed in Table 5. There are three types of applications: 1) models, such as atmospheric global and regional prediction models; 2) forecast analyst support systems that aid the operational forecaster to visualize and analyze environmental data; and 3) mission service areas, which are major topical/application areas corresponding to a core agency function, each of which is supported by multiple models and forecast analyst support systems. ASPEN is applicable to all three types of applications so long as accurate and consistent application requirement ranges and application priorities (e.g., Figs. 2b and 3a) are provided.

The list of ASPEN applications and application categories will be gradually expanded in the future. The current list is sufficient to estimate useful NOAA mission benefits because

	a							b						
	D	N_A	delta_x	T_R	N_V	s.d.	T_L	D	N_A	delta_x	T_R	N_V	s.d.	T_L
T(z)	Global	16	25	12	8	2	1.5	Global	[0.25,1,4]	[50,10,2]	[24,12,6]	[3,8,15]	[2.5,2,1.5]	[12,4,0]
RH	Global	16	25	12	5	20	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	[0,3,8]	[24,16,8]	[12,4,0]
TL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
u	NA	NA	NA	NA	NA	NA	NA	Global	[0.25,1,4]	[50,10,2]	[24,12,6]	[3,8,15]	[8,5,2]	[12,4,0]
v	NA	NA	NA	NA	NA	NA	NA	Global	[0.25,1,4]	[50,10,2]	[24,12,6]	[3,8,15]	[8,5,2]	[12,4,0]
FRP	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
fire_size	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
FI	Global	16	25	12	1	0.2	1.5	NA	NA	NA	NA	NA	NA	NA
NDVI	NA	NA	NA	NA	NA	NA	NA	Global	[1,4,16]	[50,10,2]	[24,12,6]	NA	[25,15,1]	[12,4,0]
W	Global	16	25	12	1	0.2	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	[0,3,8]	[2,1,0.5]	[12,4,0]
SIA	Global	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SIC	Global	16	25	12	1	16	1.5	Global	[0.25,1,4]	[50,10,2]	[24,12,6]	NA	[25,15,10]	[36,12,0]
SIM_Local	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SC	Global	16	25	12	1	20	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	NA	[25,10,5]	[12,4,0]
SWE	Global	16	25	12	1	3.5	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	NA	[8,4,2]	[36,12,0]
CC	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
LWC	Global	16	25	12	1	4	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	[0,3,8]	[10,5,2]	[12,4,0]
CTT	Global	16	25	12	1	2	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	[0,3,8]	[10,5,1]	[12,4,0]
RR	Global	16	25	12	1	0.7	1.5	Global	[1,4,16]	[50,10,2]	[24,12,6]	NA	[2,1,0.5]	[12,4,0]
TPW	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OSWD	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
OSWS	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
eta	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
SST	Global	16	25	12	1	3.2	1.5	Global	[0.25,1,4]	[50,10,2]	[24,12,6]	NA	[4,2,1]	[36,12,0]
h	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

ATMS Performance

Global NWP Requirement Ranges

Fig. 2. ASPEN (a) ATMS performance table and (b) global NWP requirement range table used in the sample calculations. When an observable (or attribute) is not relevant, NA (i.e., not available) is entered and ASPEN will ignore this cell during computations. In the figure, the row labels are the variable symbols from Table 2 for those variables marked by an asterisk in Table 2, and the column labels are the attribute symbols from Table 3 for those attributes marked by an asterisk in Table 3.

the prototype list includes the mission service areas which indirectly include all relevant NOAA applications. Figure 2b shows a typical application requirements range table that is used in the sample calculations. For example, the third column (horizontal resolution) shows that for all variables needed by global NWP the minimally useful resolution is 50 km and the maximally useful resolution is 2 km.

Application requirements are based on SME inputs and have an element of technically informed subjectivity for current applications. ASPEN can also be used to estimate benefits of observing systems to planned upgrades of current applications as well as entirely new applications, provided corresponding upgraded or new application requirement ranges and priority tables are created. In these cases, the SME inputs will be even more subjective. However, for current applications and possibly for applications upgrades currently being tested and

evaluated, the degree of subjectivity could be alleviated in part by using observing system simulation experiments (OSSEs), observing system experiments (OSEs), and other techniques such as forecast sensitivity observation impact (FSOI) (e.g., Boukabara et al. 2016a,b; Joo et al. 2013; Eyre 2021) in order to assess the true application sensitivity and thereby refine the ASPEN application inputs. (See the sidebar “ASPEN validation” for an example of comparing ASPEN to FSOI results.) Subjective elements will always remain, however, for the requirements and priorities of applications that do not exist in a form that allows objective testing. In this way ASPEN can help to plan observing systems consistent with future upgrades and future new applications while accounting for uncertainty.

As an example of the application priorities, Fig. 3a displays the global NWP priority table. The priorities do not have to sum to 1, but do in Fig. 3a. In any case, during the calculations the priorities are rescaled into weights that do sum to 1. Note that the global NWP requirement ranges table of Fig. 2b and the global NWP priority table of Fig. 3a are conformable. The highest global NWP priorities are given to temperature and then wind components and relative humidity and to geographic coverage, error standard deviation, and vertical resolution. Figure 4 compares the application priority tables for the 10 applications considered in our sample calculations. The global NWP priorities from Fig. 3a are repeated just left of the center of Fig. 4. Subtle differences are apparent when comparing the meteorology applications, but the importance of key variables for certain applications stand out, including normalized difference vegetation index (NDVI) for land surface, sea ice cover for sea ice, and wave height for waves.

For the purpose of sensor design and trade studies, within an ASPEN application category (e.g., within all meteorology applications), not all applications are equal. That is, some applications are foundational, such as global NWP, and others are secondary and/or downstream and dependent on the foundational application(s). Similarly, the different application categories all support the NOAA mission but to different degrees. The category and mission priorities,

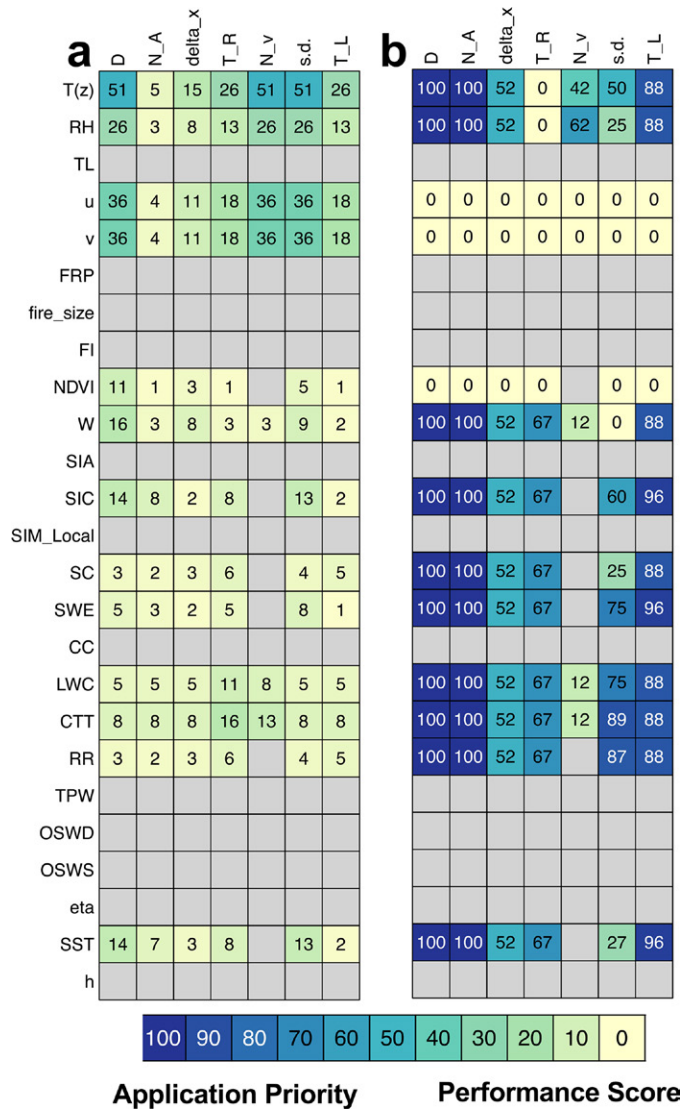


Fig. 3. The (a) global NWP priority table used in the sample calculations and (b) the calculated performance scores by variable and attribute of the ATMS performances scored by the global NWP requirements ranges. For display, the values are multiplied by 1,000 in (a) and by 100 in (b). See the color scale at the bottom. Gray indicates NA. Row and column labels are as in Fig. 2. In (a), since the original values sum to 1, the sum of all the values in the figure sum to 1,000. Note that the sum of the product of the two panels is the application benefit giving the total benefit of ATMS for global NWP.

ASPEN validation

In Fig. SB1, ASPEN benefits to global NWP are compared to FSOI metrics (Langland and Baker 2004) from four global NWP centers—NOAA’s Environmental Model Center (labeled EMC), NASA’s Global Modeling and Assimilation Office (GMAO), the Naval Research Laboratory (NRL), and the United Kingdom’s Meteorological Office (Met Office). There is quite a lot of variability among the FSOIs from center to center. Note that since radiosonde ranks first in all cases, the figure panels are effectively normalized by the radiosonde score. Each of the other three sensors ranks second in at least one case, but CrIS ranks second for NRL, Met Office, and ASPEN. Figure SB1 serves as a preliminary validation of ASPEN, showing that ASPEN’s relative benefits for global NWP fit in with an ensemble of FSOI results from different global NWP centers. Other activities aiming at validating ASPEN results are the subject of ongoing and future work.



Fig. SB1. ASPEN benefits to global NWP compared to FSOI metrics from four global NWP centers labeled EMC, GMAO, NRL, and Met Office. For consistency with the ASPEN benefits, the absolute value of the FSOI metrics (J kg^{-1}) are plotted. ASPEN benefits are from the ASPEN implementation running 7 Sep 2021. FSOI metrics are for 0000 UTC cycles in January 2015 from the Joint Center for Satellite Data Assimilation Impact of Observing Systems FSOI Intercomparison Exercise (ios.jcsda.org).

given in Table 5, which are used in the sample calculations, were chosen subjectively for demonstration purposes.

ASPEN prototype results

The examples presented in this section use the full ASPEN functionality but restricted to the 25-variable, 7-attribute, 6-sensor, 10-application subset that are selected (indicated by asterisks) in Tables 2–5, respectively.

A sample table of performance scores is presented in Fig. 3b. Note that the weights (i.e., rescaled priorities) and the performance scores are matrices of the same shape (see Fig. 3) as the observing system performance table. There is one scaled performance matrix for each pair of observing system and application. In the figure a value of zero (yellow cell)

Table 5. ASPEN prototype applications and application categories. An asterisk indicates the application is included in the sample calculations presented below. Types are model, forecast analyst support (FAS), and mission service area (MSA). The columns labeled CSP and MSP contain category and mission strategic priorities, respectively. Note that the category and mission priorities given here are for illustration purposes only and for prototyping ASPEN. Actual weights to be used for real evaluation of observing systems are the subject of an ongoing vetting process.

Category	Application	Type	CSP	MSP
Earth system	Assessments of Climate Changes and Its Impacts (CLI-ACC)	MSA	0.5	0.5
	Climate Mitigation and Adaptation Strategies (CLI-CMA)	MSA	0.8	
	Climate Prediction and Projections (CLI-CPP)	MSA	1	
	Climate Science and Improved Understanding (CLI-SIU)	MSA	0.5	
	Science and Services and Stewardship Advanced (WRN-SSS)	MSA	0.3	
	* Unified Forecast System (UFS)	Model	1	
Meteorology	Aviation Weather and Volcanic Ash (WRN-AWX)	MSA	0.3	1
	* Global NWP	Model	1	
	* Hurricane Prediction	Model	0.7	
	Hurricane/Tropical Storms (WRN-HUR)	MSA	0.7	
	Marine Weather and Coastal Events (WRN-MWX)	MSA	0.7	
	Public Weather (WRN-PWX)	MSA	0.5	
	Regional NWP, Nowcasting/Short-Term Forecasting	FAS	0.4	
	* Severe Storms, Monitoring and Forecasting	FAS	0.5	
	Severe Weather (WRN-SEV)	MSA	0.5	
	Winter Weather (WRN-WWX)	MSA	0.4	
Land/hydrology	* Fire Monitoring	FAS	0.5	0.4
	Fire Weather (WRN-FWX)	MSA	0.5	
	* Flood Forecasting	FAS	1	
	Integrated Water and Prediction Information (WRN-IWPI)	MSA	1	
	* Land Surface Assimilation and Modeling	Model	0.5	
Oceanography	Coastal Water Quality (RC-CWQ)	MSA	0.8	0.6
	Ecosystem Monitoring Assessment and Forecast (HO-ECO)	MSA	0.6	
	Fisheries Monitoring Assessment and Forecast (HO-FMA)	MSA	1.0	
	Habitat Monitoring and Assessment (HO-HAB)	MSA	0.6	
	Healthy Ocean Science Services and Stewardship Advances	MSA	0.8	
	Marine Transportation (RC-MTS)	MSA	0.8	
	Planning and Management (RC-PAM)	MSA	0.4	
	Protected Species Monitoring (HO-PSM)	MSA	0.5	
	Resilient Coasts Science Services and Stewardship Advances	MSA	0.6	
	Resilience to Coastal Hazards and Climate Change (RC-RCC)	MSA	0.8	
	* Global Ocean Forecasting	Model	0.8	
	* Sea Ice Monitoring	FAS	0.8	
	* Wave Forecasting	Model	0.8	

indicates there is an unmet requirement, while an NA (blank gray cell) indicates there is no requirement. Here we see that ATMS satisfies or nearly satisfies the needs of global NWP for geographic coverage, data density, and latency for those variables observed by ATMS, but only partially for other attributes, notably vertical resolution and error standard deviation.

The primary results of ASPEN are the application benefits to the different applications due to the different observing systems. The benefits may be displayed to show the importance of each sensor (Fig. 5a). For example, while ATMS provides a benefit of 0.43 for global NWP and 0.41 for the Unified Forecast System (UFS), its benefit for the other applications

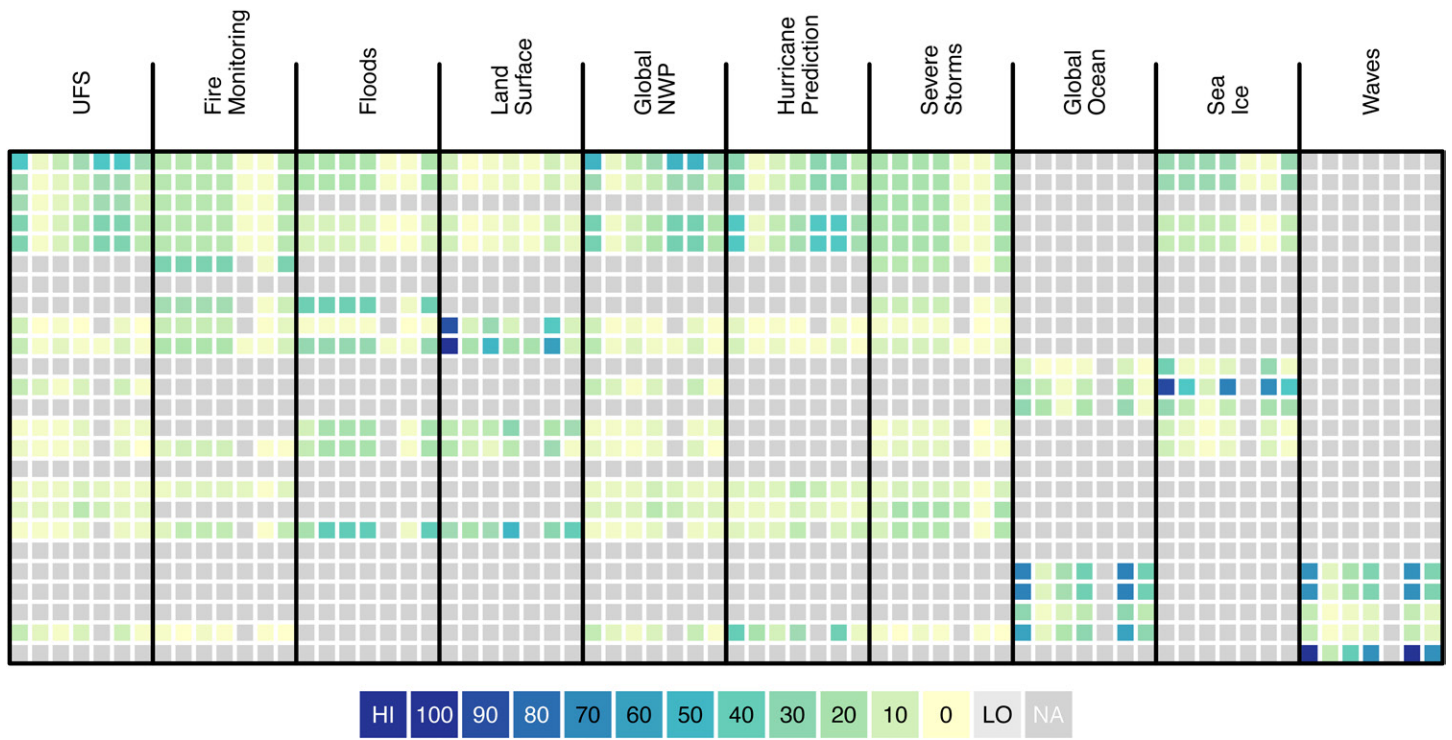


Fig. 4. Application priorities for the 10 applications used in the sample calculations. In each section of the figure, the row and column definitions are as in Figs. 2 and 3.

is much less due in part to the increased requirements on the part of the other applications for better temporal coverage for smaller-scale phenomena. In contrast, the benefits due to ABI are fairly similar from application to application (except for waves). ABI has the greatest benefit, and GLM provides the least benefit. By dividing by the observing system costs, ASPEN obtains the application-relative cost effectiveness of the observing systems, which

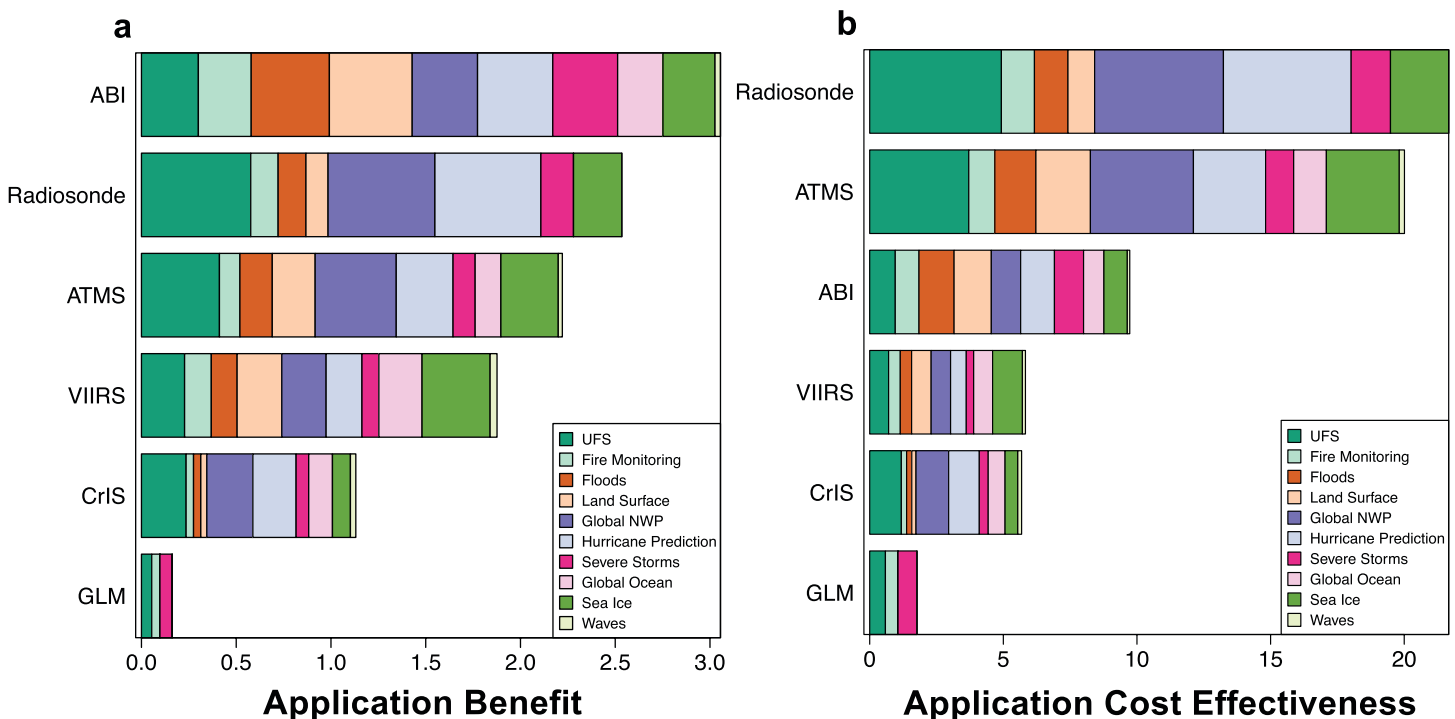


Fig. 5. The application (a) benefit and (b) cost effectiveness (benefit per billion U.S. dollars) summed over applications and sorted by the importance of each sensor. The different colors (legend) show the contribution of each application.

can also be displayed in terms of importance (Fig. 5b). This shows that after radiosondes, ATMS has the greatest cost effectiveness, ABI the next greatest cost effectiveness, and GLM the least cost effectiveness, while VIIRS and CrIS have approximately equal cost effectiveness. It is important to note that as GLM is a relatively new sensor, increased investment in the use of GLM will increase the benefits it provides and likely its ranking among other sensors.

The application benefits can be rolled up by categories (applying the category priorities) and then for the mission (applying the mission priorities) to obtain the category and mission benefits. Dividing by costs gives the mission-relative cost effectiveness (Fig. 6). Again, ATMS is the most cost effective of the satellite sensors. Of course, cost effectiveness (Figs. 5b and 6) depends strongly on the costs provided to ASPEN (see Table 4 and the discussion in the “ASPEN interfaces” section).

To understand and interpret the results obtained, various interpretability plots can be examined. For example, Fig. 7 shows that for ATMS contributing to the mission, the fractional contribution is greatest for temperature, then relative humidity, then sea surface temperature. Figure 7 also shows the breakdown by application. For example, sea surface temperature is very important for UFS, global NWP, global ocean, and hurricane prediction, but less so for the other applications, some of which are smaller scale and/or are focused on land areas. Also, as expected, the greatest contribution of ATMS to global NWP (and UFS) is due to temperature profile information. The temperature contribution is approximately twice that of humidity for global NWP (and UFS), but approximately equal for hurricane prediction.

Summary and concluding remarks

ASPEN is designed to be a science-based and yet efficient tool for comparative assessment of observation systems and thereby to support the decision process leading to the design, selection, and ultimately deployment of new space-based or ground-based assets or to the selection and acquisition of commercial data either as complement or as a baseline component of the Global

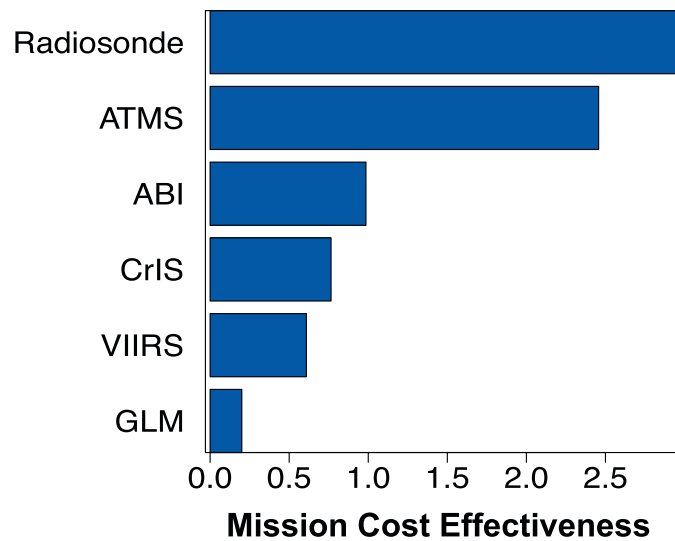


Fig. 6. Mission-relative cost effectiveness (benefit per billion U.S. dollars) sorted by the importance of each observing system.

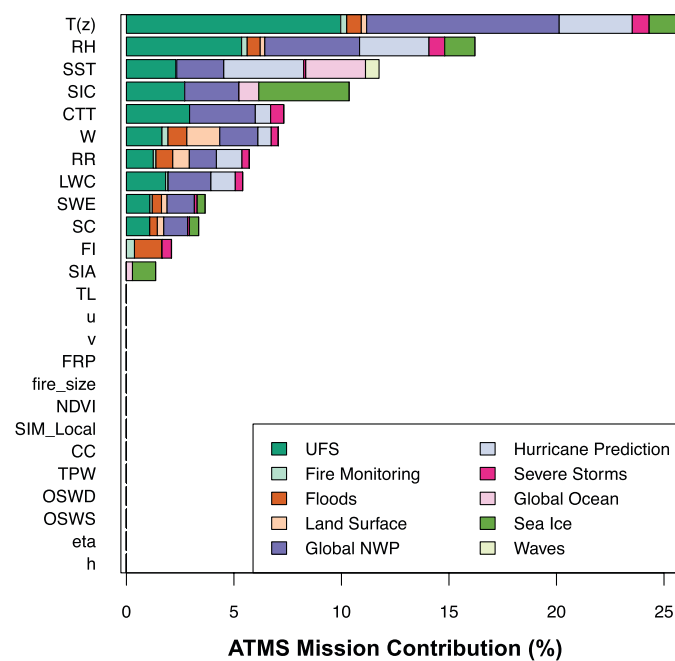


Fig. 7. Sums of the fractional contributions of ATMS to the mission benefit sorted by the importance of each variable. The sums are over attributes and applications. The different colors show the contribution for each application.

Table 6. Applicability of ASPEN.

Use	Description
Observing system optimization	ASPEN allows the direct optimization of the attributes of the geophysical variables observed by an observing system, a sensor or a constellation, accounting for diverse applications in the ASPEN application categories—meteorology, oceanography, etc.—in terms of benefit or cost effectiveness. This is useful when designing a system that caters to multiple users and/or applications with varying priorities.
Sensor optimization	By mapping sensor specifications to the attributes of the geophysical variables observed, ASPEN can support sensor design, e.g., in trade space assessments, which often must balance benefit or cost effectiveness vs design risk, such as risk associated with technological maturity or launch cadence.
Payload optimization	To determine the best set of sensors to host on a spacecraft or to include in a mission, ASPEN can be applied to all possible permutations of the potential spacecraft or mission sensor bundles.
Network optimization	To determine the best configuration of satellites and sensors provided by one or a set of partners, ASPEN can assess the benefit and cost effectiveness of the different candidate sets of contributed system resources.
Complementarity study	The relative benefits and cost effectiveness of the ground- and space-based observing systems determined by ASPEN can be used to optimize the overall Global Observing System, accounting for a wide range of applications. A similar approach could be used to evaluate a potential Arctic Weather Satellite against the existing or planned suite of polar LEO sensors.
Gap analysis	ASPEN can be used to determine how well the overall observing system is meeting the requirements of the applications and users by considering the change in benefit and cost effectiveness when adding or removing one component of the system or when evolving or designing a new configuration.
Commercial data cost-effectiveness assessment	ASPEN can determine the relative cost effectiveness of observations from alternative providers or systems knowing only the attributes of the geophysical variables provided by the different sources.
Data exchange assessment	ASPEN can be applied to an existing or proposed data exchange as part of an international partnership.
Forward-looking assessment	The ASPEN application requirements and priorities tables can account for projected evolution of applications, albeit with increased uncertainty. Using such tables allows ASPEN to be used in all the above modes of applicability, extending an assessment based on current existing requirements and priorities to account for future requirements and priorities.

Observing System. The applicability of ASPEN is summarized in Table 6. These are all areas where ASPEN has the *potential* to support decision-making.

Key characteristics of ASPEN include the following:

- 1) ASPEN is an interactive, web-based tool that is nearly instantaneous and dynamic. Since ASPEN inputs are contained in tables, it is a simple matter to update these based on changes in SME knowledge or on updates to an analytical approach (e.g., to compute observing systems performance) or on reevaluation of decision-maker priorities.
- 2) ASPEN is comprehensive: although we focused on NWP in our examples, ASPEN is designed to account for multiple applications, environmental domains, users, and mission service areas. All that is required to apply ASPEN to these is a table of requirements and associated priorities in the ASPEN expected format.
- 3) ASPEN provides comparative (side by side) assessments of multiple solutions, which is critical when choices have to be made due to budget or other constraints.
- 4) ASPEN provides a performance and gap analysis at the detail level appropriate for design optimization (resolution, uncertainty, temporal refresh, which have sensor design impacts on antenna, channel choices, noise levels, etc.) but also at summary levels to assess at the level of satisfaction of requirements for an individual application and/or overall mission level
- 5) ASPEN is dynamically adaptable and is flexible enough to perform “what if” scenarios.
- 6) ASPEN provides traceability and transparency—all resulting metrics are traceable to the applications requirements and priorities and to the observing systems performances, thus providing a high degree of transparency connecting all inputs and results in support of decision-making.
- 7) ASPEN allows for uncertainty estimation based on collecting uncertainty estimates for all inputs (as discussed in the “ASPEN interfaces” section). It is expandable to assessing

new sensors and constellations and applies to all environmental applications as long as we can reasonably quantify their observational requirements.

- 8) ASPEN is meant to be an agency-wide tool since individual observing systems contribute to many applications and products. ASPEN accounts for both technical and strategic priorities. ASPEN is designed to roll up the benefit and cost-effectiveness metrics of a sensor for individual applications across categories of applications or across the full suite of applications.

Assumptions and limitations of ASPEN. ASPEN makes a number of critical assumptions, and users of ASPEN should be careful to consider how these affect the application of ASPEN to their problem. These assumptions and caveats include the following:

- The ASPEN benefit is a measure of how well the observing system meets the application or user requirements. But the ASPEN benefit metrics are not necessarily synonymous with prediction skill. Specifically, for forecast applications, ASPEN does not assess the impact on forecasts in terms of anomaly correlation coefficients, RMSE scores, or other usual measures of forecast skill. However, based on experience with forecast impact tests (OSEs and OSSEs), the more an observing system scores at the high end of the requirements ranges, especially for high-priority measurements, the more likely the forecast impact of this observing system will be positive.
- Since the observing system performances are overall values for any given variable and attribute, ASPEN results lack the detailed granularity and nuance that is sometimes necessary. In reality performance can vary with region, height, synoptic conditions, and other factors.
- The ASPEN normalization is linear. ASPEN currently scores performance linearly within the requirement range. However, ASPEN does collect three requirement values. In the future this will allow specifying nonlinear response functions that would be shaped similar to a logistics curve.
- ASPEN inputs are in geophysical space. Whether observations are actually used in geophysical or radiance space does not impact ASPEN since the objective is to compare performance versus requirements. However, limitations in the transformation from radiance to geophysical space, including the impact of a priori information in that process, should be considered, especially when assigning a sensor vertical resolution and error standard deviation.
- The definition and determination of costs depends upon the use case. Costs are an input to ASPEN and there are many nuances that might be included when specifying these costs. Costs definitions range from sensor costs alone to the total cost including sensor, system, and data exploitation costs. Within any application of ASPEN, a consistent cost approach should be applied across the observing systems being assessed.
- ASPEN has no mechanism to directly account for risk and mitigation strategies. A first-order approach would be to adjust costs for these factors.
- The reliability and applicability of ASPEN outputs depend obviously on the quality and consistency of the ASPEN inputs in the context of the user's purpose. In particular, sensor performance is sometimes based on remote sensing science and sometimes on SMEs, but estimates from different sources may not be consistent. In general, even though SMEs are expert, specifying the ASPEN inputs will involve some degree of subjectivity and variability.

Potential of ASPEN. ASPEN might be used to support many types of decision-making. A primary use of ASPEN might be to provide a gap analysis, at a global scale, to check if and to what extent gaps exist (or will exist in the future) in our ability to measure the Earth system environment. Table 6 highlights examples of some of the potential ASPEN use cases. In these

cases, the benefits and cost effectiveness of several observing systems would be assessed in terms of a suite of representative applications, e.g., all the mission service areas listed in Table 5. In the cases listed in Table 6, the comparisons would be between observing systems or variations of an observing system (observing system optimization), between a series of sensors being considered in a trade study (sensor optimization), between different potential satellite hosting manifests (payload optimization), between different proposed networks of sensors or constellations of satellites (network optimization), between space-based and ground-based observing systems (surface/space complementarity study), between comparisons of perturbations of a given observing system (gap analysis), between different proposed data buys (commercial data cost-effectiveness assessment), and between proposed sets of data offered during data exchange negotiations (data exchange assessment). In the case of forward-looking assessment, this paradigm is slightly modified to assess current and future observing systems in terms of the requirements and priorities of future enhanced and new applications.

Future enhancements and extensions are planned for ASPEN in terms of adding applications, sensors, and constellations; improving capabilities for trade studies, Monte Carlo uncertainty estimates, and sensitivity analysis; improved ease of use; conducting further validation and calibration; and ultimately extending ASPEN to societal application areas. For example, in the current ASPEN prototype the observing system performance and application requirements for variables like atmospheric temperature are defined as averages for the entire column. Future versions of ASPEN will differentiate between atmospheric layers (e.g., by replacing the variable “temperature” with a series of variables including boundary layer temperature, free troposphere temperature, and upper troposphere–lower stratosphere temperature). Similarly, performances are defined as averages over all conditions, and the precision (i.e., error standard deviation) varies considerable for some observing systems for cloudy versus noncloudy, land versus ocean, and day versus night. Future versions of ASPEN will include multiple precision attributes for these different conditions.

In summary, ASPEN is ideal for comparative assessments of multiple observing systems accounting for performance and cost. ASPEN is complementary to other impact tools (including OSE, OSSE, and FSOI approaches, all of which could be used to calibrate ASPEN). The sidebar “ASPEN validation” presents a first-order calibration using a comparison ranking of observing systems by ASPEN and FSOI. As a result of all these characteristics, in addition to the full transparency offered by the ASPEN traceability mechanism, ASPEN can be used to discover and highlight optimal (cost-effective) solutions using a defensible and demonstrable approach to gain wide buy-in from the various stakeholders, including stakeholders from both the user community and the sensor designers. This could be done at an agency level or at the international level to optimize the Global Observing System.

Acknowledgments. We thank our many colleagues for discussions over the last several years that motivated and led up to the design of ASPEN. We thank all those who participated over the last 2–3 years in developing ASPEN and the various inputs for ASPEN, including special thanks to Stacy Bunin, Roberto Montoro, Roderick Cardenas, Monica Coakley, Louis Cantrell, and David Helms. We thank the three anonymous reviewers for very thoughtful and helpful comments that greatly improved this paper. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the author(s) and do not necessarily reflect those of NOAA or the U.S. Department of Commerce.

Data availability statement. The best way to engage in the ASPEN project is through the direct contact with the authors. We will be happy to explore ways to collaborate with members of the science community. The inputs for the sample calculations shown in this paper are available as excel files from the authors.

Appendix: Acronyms

Acronyms used in the text are listed here. Acronyms used only in a table are defined in the table caption. Acronyms of the authors institutions are defined on the title page. Common acronyms (e.g., UTC and RMSE) and proper names (e.g., names of specific institutions and systems such as NOAA and ATMS) are not expanded in the text when first used. ASPEN acronyms are in boldface text below.

ABI	Advanced Baseline Imager
ARR	Application-dependent requirement range
ASPEN	Advanced Systems Performance Evaluation tool for NOAA
ATMS	Advanced Technology Microwave Sounder
ATP	Application-dependent technical priority
COURL	Consolidated Observing User Requirement List
CrIS	Cross-track Infrared Sounder
EMC	Environmental Modeling Center (NOAA/NWS)
EOSC	Earth-observing satellite constellation
FSOI	Forecast sensitivity observation impact
GEO	Geosynchronous equatorial orbit; geostationary Earth orbit
GLM	Geostationary Lightning Mapper
GMAO	Global Modeling and Assimilation Office
GOES	Geostationary Operational Environmental Satellite
GeoXO	Geostationary Extended Observations (satellite system)
LEO	Low-Earth orbit
MAUT	Multi-attribute utility theory
NA	Not available
NASA	National Aeronautics and Space Administration
NDVI	Normalized difference vegetation index
NOAA	National Oceanic and Atmospheric Administration
NOSIA	NOAA Observing System Integrated Analysis
NSOSA	NOAA Satellite Observing Systems Architecture
NRL	Naval Research Laboratory
NWP	Numerical weather prediction
OSP	Observing system performance
OSCAR	Observing Systems Capability Analysis and Review (WMO)
OSE	Observing system experiment
OSSE	Observing system simulation experiment
RMSE	Root-mean-square error
SME	Subject matter expert
SPRWG	Space Platform Requirements Working Group
UFS	Unified Forecast System
UTC	Coordinated universal time
VIIRS	Visible Infrared Imaging Radiometer Suite
WMO	World Meteorological Organization (Geneva)
XORWG	GeoXO Requirements Working Group

References

- Anthes, R.A., and Coauthors, 2019: Developing priority observational requirements from space using multi-attribute utility theory. *Bull. Amer. Meteor. Soc.*, **100**, 1753–1774, <https://doi.org/10.1175/BAMS-D-18-0180.1>.
- Blackwell, W. J., and Coauthors, 2019: Microwave atmospheric sounding CubeSats: From MicroMAS-2 to TROPICS and beyond. *9th Conf. on Transition of Research to Operations*, Phoenix, AZ, Amer. Meteor. Soc., J3.5, <https://ams.confex.com/ams/2019Annual/meetingapp.cgi/Paper/352453>.
- Boukabara, S.-A., and Coauthors, 2011: MiRS: An all-weather 1DVAR satellite data assimilation and retrieval system. *IEEE Trans. Geosci. Remote Sens.*, **49**, 3249–3272, <https://doi.org/10.1109/TGRS.2011.2158438>.
- , K. Garrett, and V. K. Kumar, 2016a: Potential gaps in the satellite observing system coverage: Assessment of impact on NOAA's numerical weather prediction overall skills. *Mon. Wea. Rev.*, **144**, 2547–2563, <https://doi.org/10.1175/MWR-D-16-0013.1>.
- , and Coauthors, 2016b: 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>.
- , J. Eyre, R. A. Anthes, K. Holmlund, K. St. Germain, and R. N. Hoffman, 2021: The Earth-Observing Satellite Constellation: A review from a meteorological perspective of a complex, interconnected global system with extensive applications. *IEEE Geosci. Remote Sens. Mag.*, **9**, 26–42, <https://doi.org/10.1109/MGRS.2021.3070248>.
- Eyre, J. R., 2021: Observation impact metrics in NWP: A theoretical study. Part I: Optimal systems. *Quart. J. Roy. Meteor. Soc.*, **147**, 3180–3200, <https://doi.org/10.1002/qj.4123>.
- Helms, D., M. Austin, L. McCullough, R. C. Reining, A. Pratt, R. Mairs, L. O'Connor, and S. Taijeron, 2016: NOAA Observing System Integrated Analysis (NOSIA-II) methodology report. NOAA Tech. Rep. NESDIS 147, 86 pp., <https://doi.org/10.7289/v52v2d1h>.
- Joo, S., J. Eyre, and R. Marriott, 2013: The impact of MetOp and other satellite data within the Met Office global NWP system using an adjoint-based sensitivity method. *Mon. Wea. Rev.*, **141**, 3331–3342, <https://doi.org/10.1175/MWR-D-12-00232.1>.
- Langland, R. H., and N. L. Baker, 2004: Estimation of observation impact using the NRL atmospheric variational data assimilation adjoint system. *Tellus*, **56A**, 189–201, <https://doi.org/10.3402/tellusa.v56i3.14413>.
- Maddy, E. S. and S. A. Boukabara, 2021: MIIDAPS-AI: An explainable machine-learning algorithm for infrared and microwave remote sensing and data assimilation preprocessing - Application to LEO and GEO sensors. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, **14**, 8566–8576, <https://doi.org/10.1109/JSTARS.2021.3104389>.
- Maier, M. W., and Coauthors, 2021: Architecting the future of weather satellites. *Bull. Amer. Meteor. Soc.*, **102**, E589–E610, <https://doi.org/10.1175/BAMS-D-19-0258.1>.
- Murray, J., D. Helms, and C. Miner, 2008: Sensor performance considerations for aviation weather observations for the NOAA Consolidated Observations Requirements List (CORL CT-AWX). *Proc. SPIE*, **7088**, 708802, <https://doi.org/10.1117/12.795233>.
- Simmons, A., and Coauthors, 2016: Observation and integrated Earth-system science: A roadmap for 2016–2025. *Adv. Space Res.*, **57**, 2037–2103, <https://doi.org/10.1016/j.asr.2016.03.008>.
- SPRWG, 2018: NOAA Space Platform Requirements Working Group (SPRWG) final (cycle 2b) report. NOAA/NESDIS Tech. Rep., 177 pp., <https://nesdis-prod.s3.amazonaws.com/2021-09/SPRWG18.pdf>.
- St. Germain, K., 2018: Overview of the NOAA Satellite Observing Systems Architecture (NSOSA). NOAA Tech. Rep., 14 pp., <https://www.space.commerce.gov/wp-content/uploads/2018-06-NSOSA.pdf>.
- WMO, 2019: Space-based capabilities (OSCAR/Space). World Meteorological Organization, <https://www.wmo-sat.info/oscar/spacecapabilities/>.