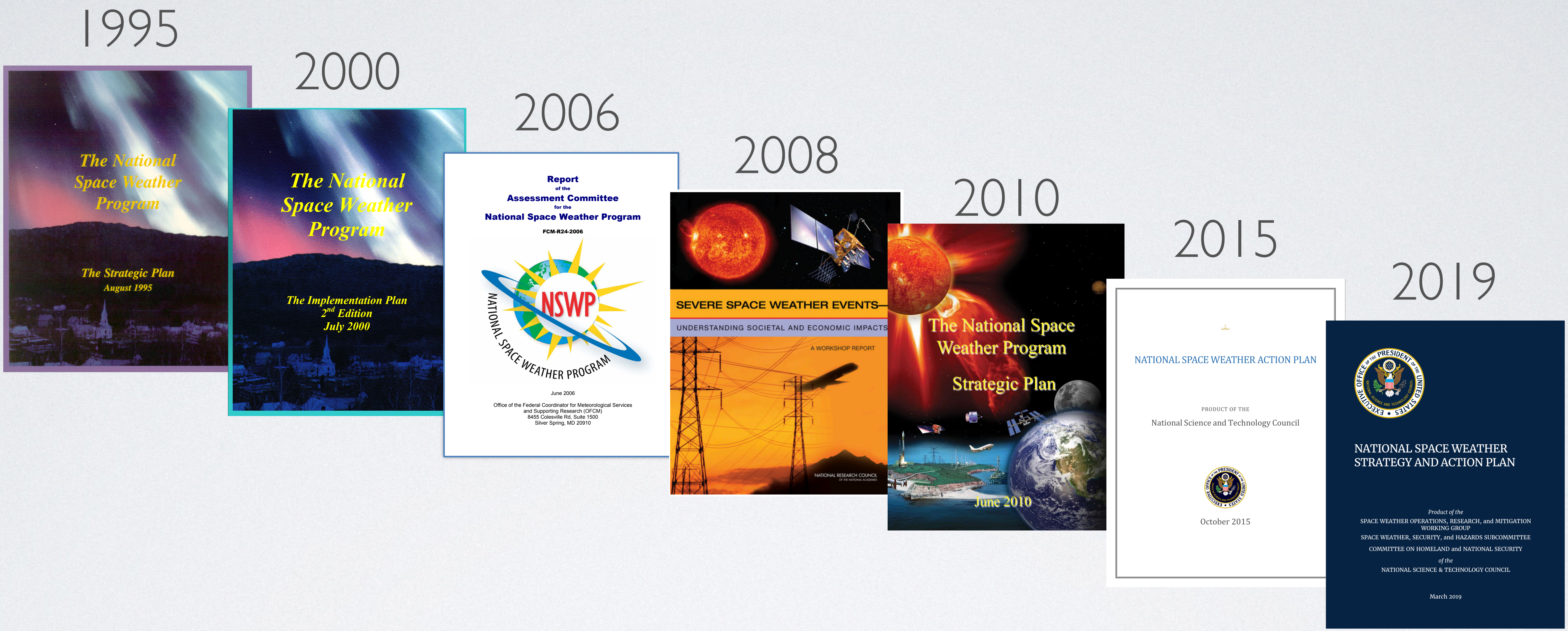
Development of Benchmarks for Extreme Space Weather Events

Geoffrey Reeves
Los Alamos National Laboratory

Geoff@ReevesResearch.org

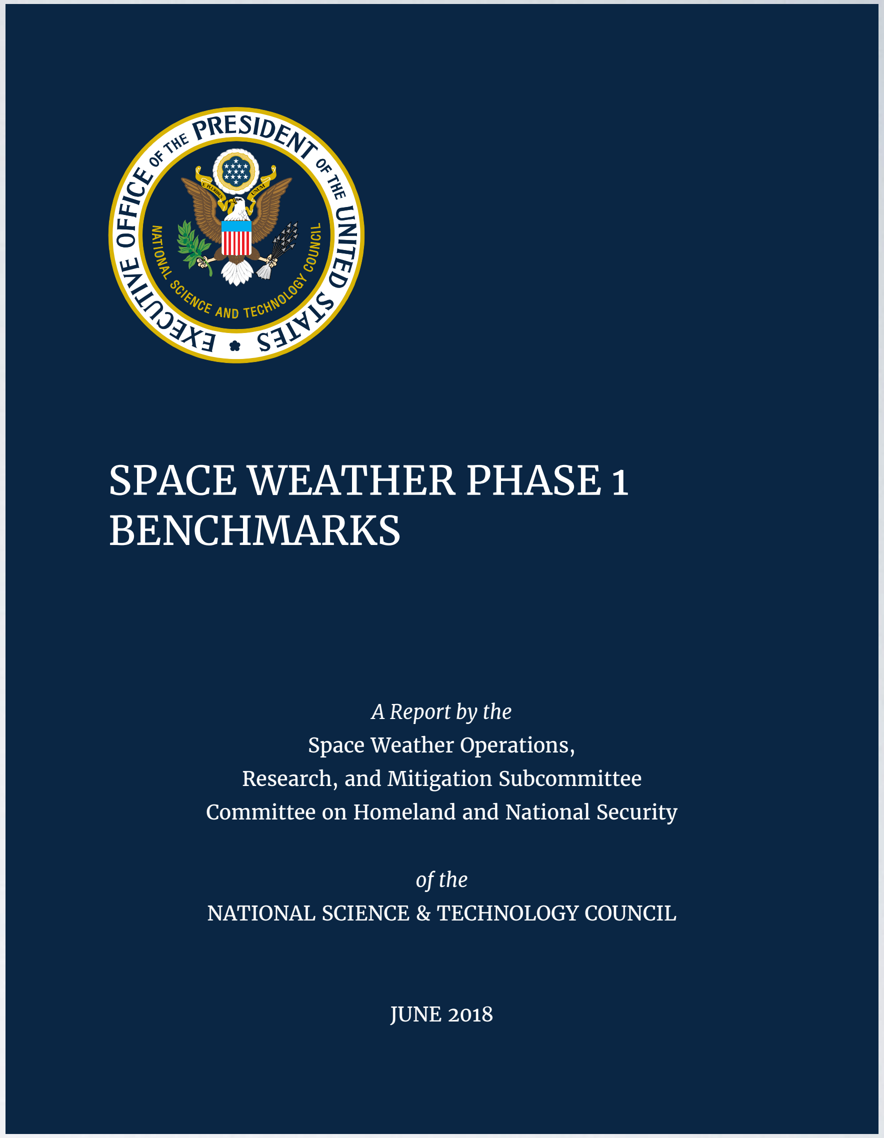
Evolution of the US National Space Weather Program



Resaerch Program → Severe Effects → Actionable Information

Phase 1 Study

- Begun in 2017
- Published in 2018
- Conducted by the Space Weather Operations, Research, and Mitigation (SWORM) subcommittee
- Under the US Department of Homeland Security
- Involved >25 federal departments and agencies


SPACE WEATHER PHASE 1 BENCHMARKS

A Report by the
Space Weather Operations,
Research, and Mitigation Subcommittee
Committee on Homeland and National Security

of the
NATIONAL SCIENCE & TECHNOLOGY COUNCIL

JUNE 2018

Working Toward Phase 2

- Phase 1 was fairly rapid turn-around with little input from the scientific and operator communities
- A 'Next Steps' panel conducted an in depth review of the Phase 1 benchmarks and methodologies
- Also provided recommendations for research and development of improved benchmarks
- Report to be published December 2019

What are benchmarks?

- They are <u>not</u> metrics for model or prediction performance but <u>do</u> help set targets
- The benchmarks specify the **1-in-100 year** and **theoretical maximum** levels of space weather conditions that can affect the critical infrastructure
- They <u>do not</u> evaluate or classify the potential effects of a space weather event on technologies

What is the purpose of benchmarks?

- **Enhance awareness** of threats among critical infrastructure owners and operators
- Provide input for **engineering standards**
- Provide input for **vulnerability & risk assessments**
- Help guide development of **mitigation procedures**
- Establish **thresholds for action**
- Set goals for academic and private sector **research**

Elements of Each Topic Area

- Define the relevant space weather parameters
- Describe and document the methodology (with references)
- Determine 1-in-100 year levels and theoretical maxima

Benchmarks for Induced Geo-Electric Fields	
Environmental parameter	Intense magnetic storms may induce geo-electric fields of sufficient strength to drive quasi-direct currents in electric power grids, sometimes causing blackouts and damaging transformers.
Methodology for determining benchmarks	Benchmarking for induced geo-electric field amplitudes used two geophysical quantities: the surface impedance relationship between geomagnetic variation and the induced geo-electric field, as well as a measure of geomagnetic activity at Earth's surface. Surface impedance values are obtained by magnetotelluric surveys, which have been completed for about half of the continental United States. Surface geomagnetic activity is routinely measured at magnetic observatories and variometer stations, and geomagnetic variations during a once-per-century event are estimated by a statistical analysis.
1-in-100-year benchmarks	The median once-per-century geo-electric exceedance amplitude among surveyed sites (see Figure 1) is 0.26 volts per kilometer (V/km), with amplitudes exceeding 14 V/km in Minnesota. One standard-deviation error, the result of statistical variance in the geomagnetic data, is estimated to be about 30 percent, which is small compared to the site-to-site differences. The full benchmark of once-per-century geo-electric amplitudes across the United States, where data is available, is displayed in Figure 1.
Theoretical maximum benchmarks	Not feasible to compute benchmarks. Higher frequency amplitudes cannot be reasonably estimated from the observatory data, and while lower frequency harmonics generally yield smaller geo-electric amplitudes, additional investigation would help inform this issue.

Includes deeper textual explanations

Benchmarks for Induced Geo-electric Fields

1. Space Weather Action Plan 1.1.1

Action 1.1.1 of the Space Weather Action Plan states: "The Department of the Interior (DOI), the Department of Commerce (DOC), and the National Aeronautics and Space Administration (NASA), in coordination with the Department of Homeland Security (DHS), the Department of Energy (DOE), and the National Science Foundation (NSF), will: (1) assess the feasibility of establishing functional benchmarks [for induced geo-electric fields] using currently available storm data sets, existing models, and published literature; and (2) use the existing body of work to produce benchmarks [for induced geo-electric fields] for specific regions of the United States."

2. Induced Geo-electric Fields

Geo-electric fields are induced in Earth's electrically conducting interior by time-dependent geomagnetic field variation. During intense magnetic storms, induced geo-electric fields can drive quasi-direct currents of electricity of sufficient strength to interfere with operation of the power grid, sometimes causing blackouts and damaging transformers. Geomagnetic disturbances have affected power grids in the past. For example, in March 1989, an intense magnetic storm caused the collapse of the entire Hydro-Quebec power grid in Canada. More recently, in October 2003, a magnetic storm caused disturbances in power grids in Scotland and Sweden. According to some scenarios, the future occurrence of an extremely intense magnetic storm could result in widespread and possibly cascading failures if the power grid is not sufficiently resilient to the effects of space weather. Even for brief periods of time, loss of power can prove disruptive for communities.

3. Methodology for Establishing Benchmarks for Induced Geo-electric Fields

This task focused on the development of a formal statistical product in terms of maps of geo-electric hazard. For practical evaluation of geo-electric hazards, estimates of two geophysical quantities are needed: (1) the surface impedance relationship between geomagnetic variation and the induced geo-electric field and (2) a measure of geomagnetic activity realized at Earth's surface.

Surface impedance is a function of the three-dimensional conductivity structure of the solid Earth and ocean. It is usually expressed in the Fourier-transformed frequency domain as a tensor. Impedance can differ greatly from one geographic location to another; it is not readily estimated from geological and tectonic models. Impedance is measured, however, during magnetotelluric surveys, such as the one sponsored by the NSF's EarthScope program,[4] which has, so far, been completed for about half of the contiguous United States.

Surface geomagnetic activity is measured at magnetic observatories, such as those operated within the INTERMAGNET consortium,[5] or at variometer stations, such as those of the ULTIMA consortium.[6] For purposes of hazard assessment, analysis of magnetometer time series can be focused on either the time-autocorrelated waveform nature of the data, or it can be focused on statistical analysis of

---

4 A. Schultz et al. "USArray TA Magnetotelluric Transfer Functions: REU60, 2006–2018," doi:10.17611/DP/11455918. Retrieved from the IRIS database August 16, 2017.

5 J. J. Love and A. Chulliat, "An International Network of Magnetic Observatories," *EOS, Transactions, American Geophysical Union* 94, no. 42 (2013): 373–384, doi:10.1002/2013EO42

6 K. Yumoto et al., "ULTIMA of Ground-Based Magnetometer Arrays for Monitoring Magnetospheric and Ionospheric Perturbations on a Global Scale," presented at 2012 Fall Meeting, AGU, San Francisco, California.

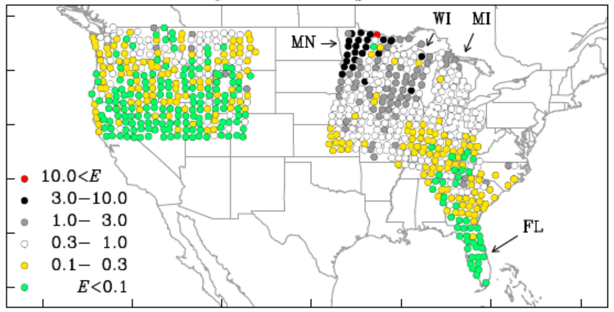
characteristic features identified in the data. These two approaches are orthogonal, but knowing the results of both is useful. This report takes a statistical approach for benchmarking induced geo-electric field amplitudes that are unlikely to occur more than once in 100 years.

To use the measured impedances and to perform a statistical analysis of observatory data, Love et al. focused on sinusoidal variation over a finite window of time.[7] Analysis of geomagnetic variation is limited on the high-frequency end of the spectrum by the one-minute sampling rate of the historical magnetic observatory data. For specificity, the amplitudes of geomagnetic activity Fourier waveforms having period of 240 seconds and persisting over a duration of 600 seconds were estimated from approximately 30 years of observatory data. This was done for both north-south (p_x) and east-west (p_y) magnetic vector components. These amplitudes were then extrapolated using a simple statistical model to once-per-hundred-year values. The frequency domain multiplication of a Fourier magnetic field amplitude with an impedance tensor gives a geo-electric amplitude.

4. Benchmarks

For the one-in-100-year benchmark, detailed results are discussed in Love et al.[8] A map of once-per-century geo-electric exceedance amplitudes (E_e^x) for p_x is shown in Figure 1. Depending on location, once-per-century geo-electric exceedance amplitudes can exceed 1 volt per kilometer (V/km) in many places across the northern Midwest United States and some places in the Eastern United States. Among the surveyed sites, the median geo-electric amplitude is 0.26 V/km, but because of the combination of geographic differences in geomagnetic activity and Earth-surface impedance, geo-electric amplitudes differ by over two orders of magnitude. At some sites in Minnesota, for example, once-per-century amplitudes exceed 3.00 V/km. Across other areas, such as in Florida, these amplitudes are less than 0.1 V/km. In northern Minnesota, once-per-century amplitudes exceed 14.00 V/km, while just over 100 kilometers away, amplitudes are only 0.08 V/km. One standard-deviation error, the result of statistical variance in the geomagnetic data, is estimated to be about 30 percent, which is small compared to the differences.

At some sites in the northern Midwest United States, once-per-century geo-electric amplitudes exceed 2 V/km, which is the level inferred to have been realized in Quebec during the March 1989 storm. As a point of reference only, amplitudes in some regions of northern Minnesota exceed the once-per-century baseline amplitude of 8 V/km (without latitude corrections) used by the North American Electric Reliability Corporation (NERC) in its benchmark study using synthetic Earth impedances.[9]



---

7 J. J. Love et al., "Geoelectric Hazard Maps for the Continental United States," *Geophysical Research Letters*, 43, no. 18 (2016.): 9415–9424, doi:10.1002/2016GL070469

8 Ibid.

9 NERC, "Benchmark Geomagnetic Disturbance Event Description" (2014): 1–26.

Example: Atmospheric Expansion Benchmarks

- Some benchmarks include multiple scenarios
- Benchmarks are quantitive with uncertainties
- Identifies areas where benchmarks are not currently possible

1-in-100-year benchmarks	Cause of Upper Atmosphere Expansion	Altitude (km)	Benchmark (percent neutral density increase)[m]	Associated Uncertainty
	Solar Extreme Ultraviolet and Far Ultraviolet Radiation	250	50%	± 30%
		400	100%	± 30%
		850	200%	± 30%
	Solar EUV Radiation Enhancement during Solar Flares	400	75%	factor of 2
	Coronal Mass Ejections Driving Geomagnetic Storms	400	400%	± 100%
Theoretical maximum benchmarks	Solar Extreme Ultraviolet and Far Ultraviolet Radiation	250	100%	factor of 2
		400	160%	factor of 2
		850	300%	factor of 2
	Solar EUV Radiation Enhancement during Solar Flares	400	135%	factor of 2
	Coronal Mass Ejections Driving Geomagnetic Storms	400	Not feasible to compute benchmarks	± 100%

January-November 2019
Next Steps Space Weather Benchmark Panel

Geoff Reeves, Chair
Thomas J. Colvin, Executive Secretary
Jericho Locke, Executive Secretary

Induced Geo-Electric Fields
Pete Riley, Leader
Jeff Love
Antti Pulkkinen
Adam Schultz
Emanuel Bernabeu
Alan Thomson

Ionizing Radiation
Christina Cohen, Leader
Joe Giacalone
Therese Moretto Jorgensen
Juan Rodriguez
Tim Guild
Delores Knipp

Ionospheric Disturbances
Susan Skone, Leader
Anthea Coster
Keith Groves
Jonathan Makela
Ethan Miller
Roger Varney

Solar Radio Bursts
Dale Gary, Leader
Tim Bastian
Gregory Fleishman
Stephen White
Angelos Vourlidas
Jade Morton
Jasmina Magdalenic

Upper Atmosphere Expansion
David Jackson, Leader
Sean Brunsma
Yue Deng
Eric Sutton
Tzu-Wei Fang
John Emmert

Next Steps Benchmark Panel

We Answered the Following Questions

* Are the current benchmark <u>quantities</u> well-aligned with the objectives and use cases stated in the Phase 1 Document?
* Are the benchmark <u>values</u> reasonable and up-to-date based on current understanding? (data, models, and gaps)
* Is the <u>methodology</u> used to derive the benchmark values up-to-date, rigorous, and compelling?

And Made Two Types of Recommendations

* Recommendations (with priorities) for updates that could be done now or in the near term
* Recommendations (with priorities) for longer-term studies, data collection, or research that would improve the benchmark values, reduce their uncertainties, or improve their usability

We Adopted the Following Guidelines

* The benchmarks should be ‘technology agnostic’. We focused on physical quantities not on the effects on particular systems.
* The benchmarks should have utility for current preparedness & planning but should also apply to future systems that have yet to be designed
* We focused on developing benchmarks with utility to the security of US infrastructure but recognize that more international collaboration and consensus on future space weather benchmarks is needed
* Not every important quantity can be a benchmark. Any additional benchmarks need to add value without sacrificing ease of use

General Conclusions

* The Phase I benchmark panel did an amazing job in a short amount of time

* In general the Phase I benchmark quantities are: well-aligned with the objectives and use cases but we provide recommendations for other quantities that could enhance the value and/or utility of the benchmarks for end users

* The benchmark values are mostly reasonable and up-to-date but we recommend some updates, identify some gaps, and make recommendations for refining and improving the benchmark values

* Some of the methodology used were up-to-date, rigorous, and compelling but we identified some gaps and made recommendation for other methodologies that should be considered in development of improved benchmark values.

Each Focus Area Group Developed a Detailed Analysis, Identifying Gaps and Making Recommendations Specific to That Focus Area

In Addition, We Identified Some
Cross-Cutting Issues and Recommendations

Cross-Cutting Issues

* We recommend that, in addition to 1-in-100 year and worst case, developing 1-in-N year benchmarks would add value, confidence and utility
* Benchmarks would benefit from a dedicated data collection plan prioritizing both data continuity and new data sources
* Capturing duration along with intensity of events would enhance their usability and value

We Made Near and Long-Term Research Recommendations

* The committee recognizes that improving the space weather benchmarks represents a new direction for the research community and for research funding agencies
* The goals are aimed more at quantification and less at basic physical understanding than past activities (but not mutually exclusive)
* Future benchmarks could benefit from greater application of techniques for extreme value statistical analysis and uncertainty quantification (UQ)
* Benchmark-focused research will require non-traditional research investments including: cleaning data sets to remove artifacts; cross-calibration of heterogeneous data sets; making data sets more publicly available, etc.
* Benchmark-focused research will require development of models aimed specifically at long-term analysis and/or prediction of extremes

Specific Research Recommendations

- The panel recommends that research funding agencies, such as NASA and NSF, implement new research programs that directly address the unique applied research demands of improving space weather benchmarks.
- The panel also suggests that research funding agencies also consider how research priorities in modeling might more effectively advance physical models with the goal of understanding long-term and extreme space weather conditions.
- We also note that the goals of SWORM would benefit from a review, similar to this one, but led by stakeholders, planners, and the user community