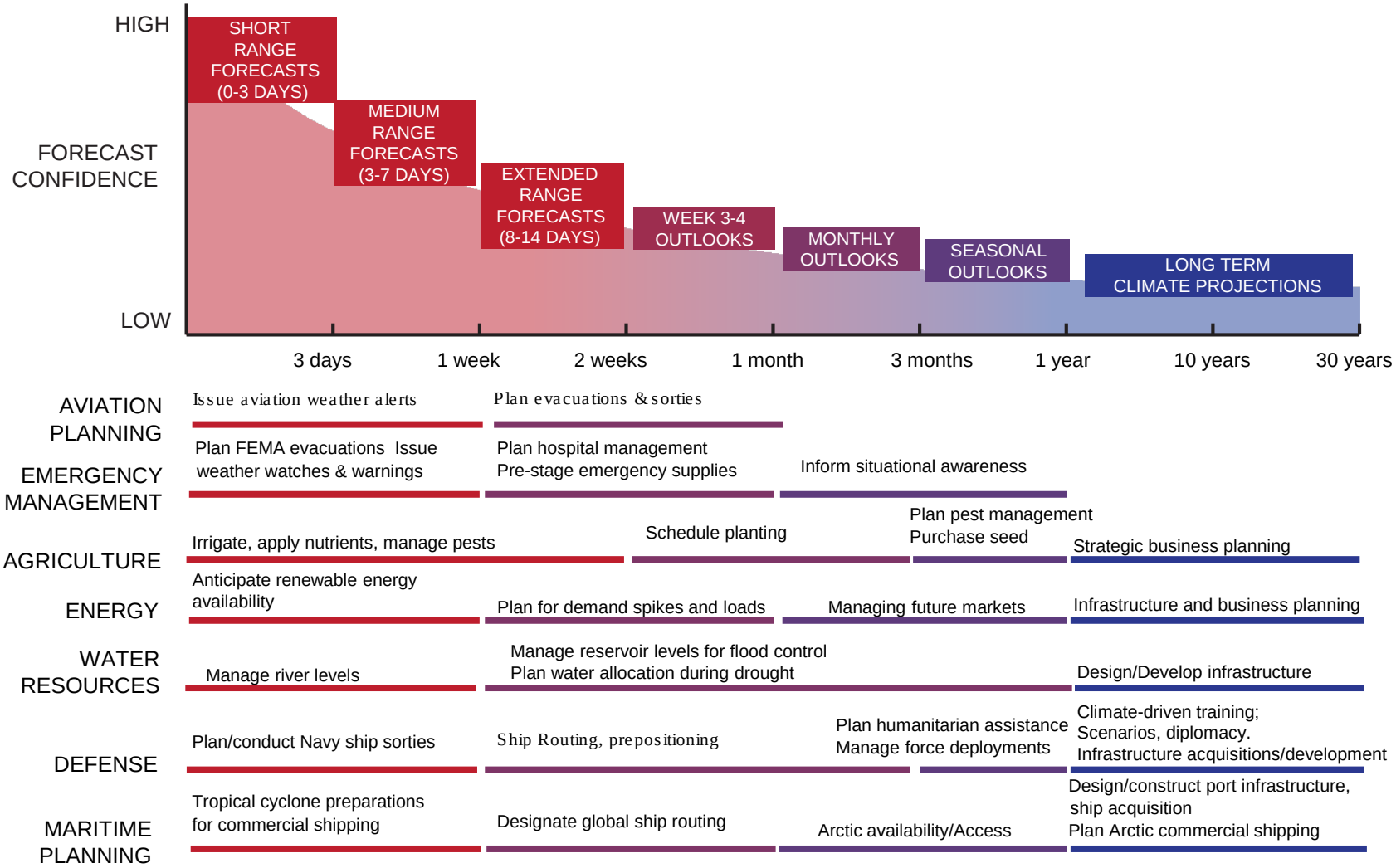
The National Earth System Prediction Capability (National ESPC)

Metrics, Post-processing, and Products for Subseasonal to Seasonal (S2S) Workshop

For DTC UFS Test Plan & Metrics Workshop, 30 July 2018

National ESPC
Earth System Prediction Capability

Forecast Lead Time/Decision Timescale



BAMS 2017: The National Earth System Prediction Capability: Coordinating the Giant
http://journals.ametsoc.org/doi/abs/10.1175/BAMS-D-16-0002.1

National ESPC
Earth System Prediction Capability

The Weather Research and Forecasting Innovation Act of 2017 Title II Sec. 201

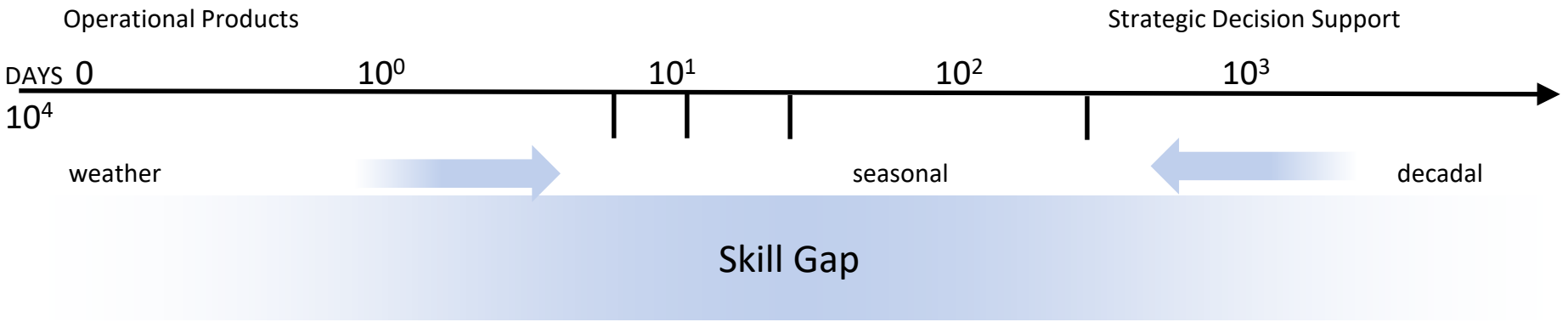
(e) Cooperation.--The Under Secretary shall build upon existing forecasting and assessment programs and partnerships, including--

(1) by designating research and monitoring activities related to subseasonal and seasonal forecasts as a priority in one or more solicitations of the Cooperative Institutes of the Office of Oceanic and Atmospheric Research;

(2) by contributing to the interagency Earth System Prediction Capability; and

(3) by consulting with the Secretary of Defense and the Secretary of Homeland Security to determine the highest priority subseasonal and seasonal forecast needs to enhance national security.

Bridging the Gap



- To extend weather skill past traditional weather scales:
  - Fully coupled air-ocean-land-ice modeling systems needed
  - Multi-model ensembles
  - Improved data assimilation techniques, particularly for ocean-land-ice
- To improve climate model skill at subseasonal scales:
  - Data assimilation, reanalysis/reforecast
  - Process representation

Needed across time scales:

- Improved HPC utilization, i.e. advanced architectures
- Common model architectures
- Multi-model ensemble management
- Uncertainty depiction; metrics suited for longer time scales
- Product creation

- Research agencies: work within mission expertise to improve skill
- Operational agencies: exploit the research for skill improvements
- *Need strategic-level coordination for continuity and efficiency*

Issues with using synoptic verification methods for extended range prediction

- Synoptic verification is focused on grid point by grid point verification and emphasizes the atmosphere
  - Grid point verification penalizes small time and spatial errors, obscuring useful predictive value at longer time scales
  - S2S skill is dependent on aggregating over space and time much as in mesoscale verification where predictive limits are small
  - Lorenz (synoptic) atmospheric time scales are less than two weeks. Predictive time scales in other domains (ocean, ice, land) are much longer. Multi-domain verification is required and may yield predictive value
- Forecast skill at longer ranges is often determined by a model's ability to accurately represent large scale phenomena
  - Verification of specific phenomena - ENSO, QBO, etc., is critical to model development and influences predictive skill at S2S time scales
  - Conditional verification looking at responses to specific phenomena and location of phenomena can extend prediction limits (i.e. MJO phase and CONUS forecast skill)

NAS S2S Report Recommendations

- **Aggregating observations into features or indices** provides added S2S predictability. Feature-based or object-oriented verification, especially ensemble feature-based verification, should be pursued for S2S to support Earth system model development and forecast calibration and validation.
- Two-step verification **correlating a feature, index, or object to a user-valued event** shows promise for extracting useful signal at the limits of predictability.
- Increasing the skill of S2S forecasts through improved and expanded representation of the physical system and expanding the utility of S2S forecasts will require collaboration among **operational, research, and stakeholde**r communities to develop **common S2S forecast skill and verification metrics**, as well as **process-oriented diagnostics** that target S2S processes and phenomena.
- **Retrospective forecasts using the current version** of the forecast system and up-to-date reanalyses are important for advancing calibration and validation efforts of ensemble prediction.

National Academies of Sciences, Engineering, and Medicine. 2016. *Next Generation Earth System Prediction: Strategies for Subseasonal to Seasonal Forecasts*. Washington, DC: The National Academies Press. https://doi.org/10.17226/21873.

WMO S2S Metrics Recommendations

WMO S2S verification recommendations:

- Development of user-relevant metrics, thresholds, onset, etc.
- Identify relevant variables (e.g. rainfall phases) and procedures, beyond standard average events, and phase space methods (e.g. for MJO).
- Implement an S2S framework for evaluating real time and retrospective forecast skill.
- Conditional verification (e.g. verification conditioned by ENSO, MJO or other state).
- Appropriate measures for extremes and event discrimination.
  - Spatial methods. Can we define objects such as watersheds, El Nino location and intensity as an object?
- Account for sampling uncertainty.

Book in preparation: *The Gap between Weather and Climate Forecasting: Sub-seasonal to Seasonal Prediction*; Chapter on "Forecast Verification for S2S Time Scales" (Coelho, Brown, Wilson, Mittermaier, and Casati) (source: Brown)

Metrics, Post-processing, and Products for Subseasonal to Seasonal (S2S) Workshop

NCWCP, 28 Feb - 2 Mar 2018

- 2.5-day workshop (final half day canx due to weather)
- 106 representatives of several federal/state agencies, laboratories, universities, private corporations
- Highlight:
  - User needs, agency capabilities and products, gaps between
  - Potential operational/technological solutions to address gaps
  - Metrics for both developers & users. Reliability important at S2S scales
- Workshop report finished, available on workshop webpage: https://cpaess.ucar.edu/meetings/2018/metrics-and-postprocessing-workshop
- Workshop write-up submitted to BAMS for publication
- Action items/discussion appears in NOAA draft report to Congress: *Subseasonal and Seasonal Forecasting Innovations: Plans for the 21st Century*

Metrics, Post-processing, and Products for Subseasonal to Seasonal (S2S) Workshop

General Recommendations

1. *Build a strong S2S community* including researchers, operational forecasters, social scientists and users similar to the weather and climate communities.
   a. Need appropriate training for users and forecasters
   b. Community software and data libraries for users and developers
   c. Research community access to operational models
   d. Coordinating group to advocate for resources for science and operations to include significant investment in data storage hardware, software, and high performance computing (HPC).

2. *Carefully design an operational configuration that robustly meets user needs and is also accessible to the research community,* to enhance S2S prediction skill. Determine:
   a. Model diversity, ensemble size and resolution
   b. Ensemble generation methodology and initialization frequency
   c. Reforecast characteristics
   d. Required parameterization/process resolution
   e. User targeted post processing

Metrics, Post-processing, and Products for Subseasonal to Seasonal (S2S) Workshop

General Recommendations

3. *Develop and fund observations to address S2S*
   a. Under-observed regions such as the Arctic and oceans
   b. New observation types to address critical parameters such as ice and deep soil moisture.
   c. Enhance coupled data assimilation to better use existing observations.

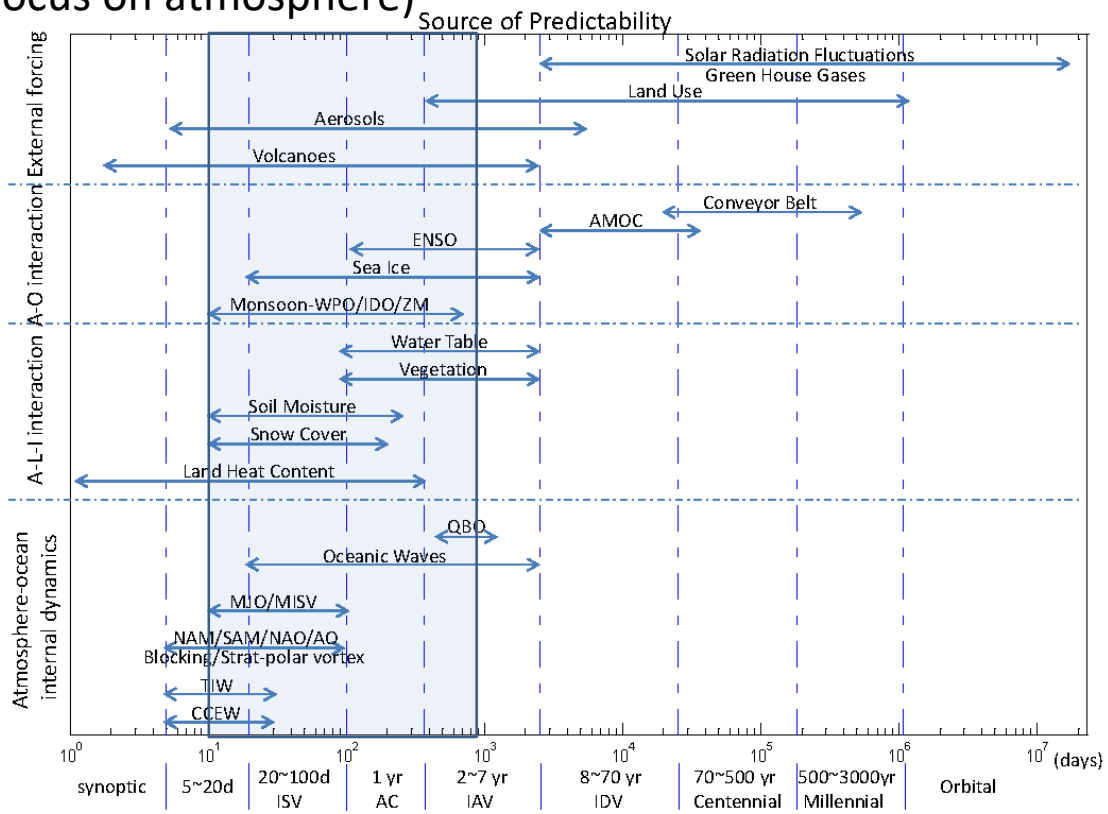
4. *Focus verification and validation more on specific user needs*:
   a. Probabilistic prediction for user-specific time and space scales
   b. Object-oriented verification incorporating both temporal and spatial variations, user value or return-on-investment,
   c. More user engagement in development of tailored products

5. *Improve public/private policy, especially the federal funding model; empower a federal coordinating authority for the S2S enterprise*.
   a. If S2S is to be a community-based endeavor, improving communication across agencies is critical: *"Each agency appropriation being done separately does not lend itself toward community-wide objectives."*

The National ESPC Executive Steering Group and the Interagency Weather Research Coordination Committee (IWRCC) both operate within the Office of the Federal Coordinator for Meteorology organizational structure. Both groups are addressing Weather Act tasking; the IWRCC works interagency research in S2S, while ESPC coordinates operational advances. Establishing an effective linkage between these two groups may advance both research and operations in S2S and enable a more effective coordination mechanism.

National ESPC Workshop
Metrics Recommendations

Include/emphasize **metrics appropriate to other domains** such as ocean/land/ice/bio (don't just focus on atmosphere)



National Research Council. 2010. *Assessment of Intraseasonal to Interannual Climate Prediction and Predictability*. Washington, DC: The National Academies Press. https://doi.org/10.17226/12878.

National ESPC Workshop
Metrics Recommendations

Scorecard metrics for developers and users, agreed-upon by operational and research communities:

- Match/overlap metrics with those used in synoptic forecasting, on the short end
- Include user-relevant metrics and thresholds, e.g. degree days, onset/duration/strength etc.
  - use object-oriented methods similar to precipitation in mesoscale meteorology to track features
  - apply to other variables and larger regions
- Connect process-related metrics important to longer-scale modes of variability to user needs, **e.g.**

blocking	flood/drought
ENSO	watershed precipitation, heat/cold waves
upper ocean heat content	tropical development, fisheries productivity, precipitation (atm rivers)
Arctic ice thickness/coverage/formation/breakup	community support, infrastructure, extended prediction

- Reliability of predicting extreme events (detection vs. false alarm)

More National ESPC Workshop Metrics Recommendations

For ensembles:

- Sophisticated users require clear **measures of reliability**
  - Need to add conditional reliability (region, index phases, etc.)
- **Ensemble distribution/spread**
  - Need to translate to user **decision tools** that better describe ensemble space
- Identification of extremes, onset of events, length of events, etc.

Consider a JEDI-like **organization** for post-processing, verification and validation, in a multi-agency approach.

- Significant challenge in meeting needs of multiple users: no one agency should have primacy
- Need a spectrum of metrics choices to transparently balance competing needs

Make tools and relevant data easily available to the community, to assist broad participation.

Website

http://earthsystemprediction.gov

BACKUPS

What is National ESPC?

- An integrated National Capability meeting the U.S. Federal need for Earth System Prediction by coordinating common science and technology to accelerate the capability of operational products and services
  - For the protection of life and property in the US
  - For the economic development, aviation, maritime, shipping, agriculture of the US
  - National defense and homeland security (worldwide)
  - Strategic decision making
- Includes:
  - Near term, medium range and extended range weather (< 90 days)
  - Seasonal and inter-annual climate (90 Days+)
  - Sub-decadal to decadal
  - Leveraging research and development efforts and operational capabilities
- Works within mission of each agency to further national goals

Effort is broadly consistent with WMO's S2S Prediction Plan and various national reports.

Strong need Identified for Inter-Agency Coordination