White Paper on

Statistical Post-Processing in NOAA: Key Changes Necessary to make the National Blend of Models and Reforecasting Successful

Thomas M. Hamill[1], Kathryn Gilbert[2], Matthew Peroutka[3], David Myrick[3], Mike Farrar[3], and Hendrik Tolman[4]

[1] *NOAA/OAR Earth System Research Lab, Physical Sciences Division, Boulder, CO*
[2] *NWS/NCEP Weather Prediction Center, College Park, MD*
[3] *NWS/STI Meteorological Development Lab, Silver Spring, MD*
[4] *NWS/NCEP Environmental Modeling Center, College Park, MD*

27 July 2015

Corresponding author: tom.hamill@noaa.gov, (303) 497-3060

Executive Summary

The NWS is making several major changes to the production of its post-processed numerical guidance for weather to weather-climate time scales. Via the "National Blend of Models" project ("*National Blend*" hereafter), the NWS is institutionalizing the production of high-resolution, gridded, statistically post-processed weather guidance over areas of US interest. These will be developed by post-processing and combining (blending) guidance from models and ensemble prediction systems from multiple centers, including several outside the US. When complete, the National Blend will provide guidance for all of the base fields in the NWS National Digital Forecast Database (NDFD) and additional probabilistic information is anticipated. The second major change will be the regular production of reanalyses and reforecasts at the NCEP Environmental Modeling Center (EMC). A third major anticipated change is that the Meteorological Development Lab (MDL) and its partners will develop advanced post-processing methods to exploit the extra multi-center ensemble forecast, reforecast, and high-resolution reanalysis information. Significantly improved numerical guidance is expected, guidance of such high quality that much less manual editing by forecasters will be necessary, freeing their time for more important decision-support roles.

Implementation of the National Blend and reforecasting and ensuring their regular usage through the NWS will require: (a) determining what new data are to be produced; what new reanalyses, reforecasts (at what resolution, frequency, duration, etc.), high-resolution surface reanalyses, additional NDFD elements, and so forth; (b) determining the specifications for an upgraded computational and storage system for reanalysis, reforecasting, and statistical post-processing; (c) identifying what additional WFO-based data storage, communications bandwidth, forecaster workstation capability, and software improvements may be needed if the NDFD is to be augmented with additional probabilistic information; (d) procuring, realigning, and/or maintaining the hardware needed for (a) - (c) above; (e) developing the capacity and then regularly producing global reanalyses and reforecasts, and making this data easily available in a timely fashion for post-processing system development inside and external to NOAA; (f) improving the quality of the high-resolution surface analyses, and generating high-resolution surface reanalyses to match the period of reforecasts; (g) exploring the merging of the North American Ensemble Forecast System with the National Blend; (h) updating and sometimes redesigning the post-processing software to use the greatly expanded training data and to produce a broader range of higher-quality, high-resolution deterministic and probabilistic guidance for the NDFD or other archives.

Funding via NOAA R&D programs will be sought to facilitate these changes.

1. **Introduction**.

Recent weather events such as Hurricane Sandy, the Colorado Front-Range Floods of September 2013, and 26-27 January 2015 Nor'easter have shown that despite significant progress in numerical weather prediction, there still are significant gaps in NWS forecast capabilities. At the same time, there have been rapid developments in computational capacity and in our knowledge of how to improve numerical weather predictions. Consequently, NOAA and the NWS are in the midst of major improvements to the numerical guidance they produce. Via its "Sandy Supplemental" and the Next-Generation Global Prediction System (NGGPS) projects, the NWS will implement a new community, global non-hydrostatic dynamical core for its weather forecast guidance; will upgrade its data assimilation systems to use hybrid 4D-ensemble variational analysis procedures; will improve the fidelity of its model parameterizations; will improve the treatment of initial-condition and model uncertainty in its ensemble prediction systems and extend the system to + 30 days forecast lead; will improve the coupling between state components (atmosphere / ice / ocean / land / aerosol, and so forth); will improve the supporting architecture, including test environments and software optimization; and will improve post-processing. A draft implementation plan for NGGPS covering these upgrades and more is available here.

This white paper addresses the changes needed to support high-quality post-processing. Why is this component of weather prediction particularly important? Despite the improvements expected in numerical weather prediction in the US from NGGPS and other programs, it's inevitable that the raw numerical guidance will continue to have some systematic errors for the indefinite future. Statistical post-processing is a straightforward way of ameliorating many of the errors in the guidance, extracting the predictable signal, minimizing noise, estimating the state-dependent uncertainty, and correcting biases. NOAA's customers increasingly are requesting high-quality, reliable, unbiased, skillful guidance, deterministic and probabilistic, to improve decision support. Many NWS organizations currently perform post-processing, including regional NWS headquarters, MDL, NCEP (EMC, CPC, NHC, AWC, WPC), and the National Water Center (NWC). Many of them in recent years have found that post-processing with high-quality reforecast data has substantially improved the quality of their products, and hence they advocate for the NWS to support and improve its post-processing capabilities.

The basic concept behind post-processing is simple; use a training data set of past forecasts and observations, ideally from a fixed assimilation/forecast system, to determine the adjustments to the current forecast. Though the concept is simple, the actual post-processing methods can range from relatively simple (for example, surface temperature) to much more complex (such as heavy rainfall, precipitation type, tropical-cyclone intensity, or streamflow prediction).

With much experimentation, NOAA staff have learned some general principles to get the most from post-processing.

a. *Produce reforecasts and high-resolution retrospective analyses*. The post-processing of unusual events such as heavy rainfall, or for long-lead events where skill is small and errors are large, is greatly aided by the availability of a large training data set of *reforecasts.* These are retrospective forecasts generated from the same or very similar model and data assimilation system that is used to generate the real-time forecasts. Improved training data should also include the production of improved, high-quality, high-resolution analyses and reanalyses for model development and validation. The reader is referred to a recent NOAA white paper which provides more rationale for reforecasting and some recommendations on the configuration of the reanalyses and reforecasts. More material on related plans to support an improved infrastructure for probabilistic weather forecast production are also available in Hirschberg et al. (2011).
b. *Develop post-processing techniques appropriate to the variable in question*. Different post-processing methods may be needed depending on the forecast variable; temperature, wind gusts, precipitation amount, precipitation type, and cloud cover are likely to require different algorithms, though they may share some underlying software (I/O routines, minimization procedures, quality control, and so forth). Further, there can be major differences in the skill of post-processing guidance from one method to another; time and attention is needed to determine the best approach for the problem at hand. Some of this technique development is more fundamental in character and will happen more rapidly if OAR works with NWS.
c. *Leverage data from other modeling systems if they are high in quality*. This underlies the case for the National Blend. If other international prediction systems and training data sets are available, significant skill can be added by incorporating their data. In recent years, the US and Canada have found it to their mutual advantage to share weather forecasts; the combination of the two allows them to generate products that are, on average more skillful than either in isolation. This combination has happened through the NAEFS, the North American Ensemble forecast System (Candille 2009). Within the last few years, the US Navy has begun to participate in this data sharing. Following in that line, NWS management has decided to begin development of a broader system, the National Blend, which would include data from multiple models. The National Blend will generate a wide variety of products that can be used as a first guess in the National Digital Forecast Database (NDFD), a NWS high-resolution database that underlies the production of local weather forecasts in the US and associated territories. The NDFD includes elements such as temperature, humidity, wind speed, probability of precipitation, snowfall amount, precipitation type, sky cover, and more. It is envisioned that the National Blend data set will be used heavily by forecasters. Over time, they will decrease their use of station-based Model Output Statistics (MOS; Glahn and Lowry 1972) and NAEFS products. Given the centrality of the National Blend to the NWS future concept of operations, successful technologies that worked well in NAEFS or MOS should be integrated into the National Blend.

The remainder of this white paper discusses what NOAA should do in order to adopt these principles. The changes needed are not trivial. In order to generate and disseminate

reforecasts, the NWS will need sufficient computing and storage and will need to generate high-quality reanalyses for the model initialization (reanalyses serve other purposes, including the diagnosis of climate change and climate variability). Research and development will be needed to build the advanced post-processing techniques for high-impact variables and variables where training data is comparatively sparse, such as for aircraft icing. And there are significant issues with multi-center blending, involving data access and the development of rigorous blending procedures.

The organization of the rest of the white paper is as follows. Section 2 will briefly touch on the assumed organizational responsibilities; these have mostly been worked out by senior NWS and NOAA leadership already. Section 3 provides some background on the underlying scientific and technical challenges that should be considered as plans are made for implementation of these technologies, including issues in reanalysis consistency and the development of advanced surface-based analysis techniques and post-processing techniques. Section 4 then discusses many of the anticipated logistical issues that must be addressed in the coming years, including computation, data storage, dissemination. Section 5 will provide a succinct summary.

2. **Assumptions on organizational responsibilities.**

The analysis to follow below takes the following as given: (a) EMC will become the eventual regular provider of reanalyses and reforecasts, per Hendrik Tolman's vision presented as EMC Director at the 2014 NCEP Production Suite Review[1]. (b) EMC will, over time, transfer its post-processing activities such as NAEFS to other partners within the NWS, allowing it to concentrate resources on the production of raw analysis and forecast guidance (per Hendrik Tolman's personal communication[2]). (c) Accordingly, MDL will continue to grow its center of expertise for NWS post-processing and is the most likely candidate for taking over the functionality of NAEFS, possibly to be merged with the National Blend in the future. (d) OAR scientists such as those at ESRL/PSD will continue their development of advanced post-processing techniques and will work collaboratively with NWS partners to port these technologies for operational use. It is anticipated that projects like NGGPS and USWRP will provide funding to support these efforts. (e) Given the volume of data produced with reanalyses and reforecasts, all options need to be explored to identify the best method to serve them to the broader community in a timely fashion at a reasonable cost. This could include serving the data up from EMC or NCO, and/or cloud storage and dissemination with external partners (as being explored through NOAA/CIO's Big-Data initiative).

The last assumption above is worth further discussion. Currently forecast data archival is performed by the National Centers of Environmental Information (NCEI), with the climate

[1] This commitment is based on the estimate of reforecast configurations from the Hamill et al. white paper. If much larger reforecasts prove necessary, then this commitment will be reconsidered.
[2] With the caveat that this fits in a integral NOAA strategy.

component being the former NCDC. Their archive is designed to preserve the data with no data loss for a very long time and to have near 100% data availability. There are large costs to achieve high reliability, and it is cumbersome to transfer large amounts of data rapidly from EMC to NCEI. This problem will become geometrically worse as EMC assumes the role as a regular generator of reanalyses and reforecasts. For some data sets such as the most recent forecast guidance, NCEP must also have very high reliability in its ability to provide data quickly. However, for other data sets such as its anticipated newly generated reanalyses and reforecasts, as long as there is a tape backup, the failure of one disk or an outage for a period of hours to days is comparatively inconsequential. That may delay the technique development of post-processing, but such development occurs over months and is not as time critical. The new data that NCEP or an outside vendor hosts would be readily available to its NOAA partners such as MDL and ESRL, as well as to outside customers from other government agencies and the commercial sector of the weather enterprise.

Regardless of what organizational arrangements are made to support the end goals, the scientific, technical, and logistical challenges discussed below will still remain.

3. **Underlying scientific and technical challenges**.

Reanalysis/reforecasting, advanced post-processing technique development, and the National Blend together have the potential to dramatically improve the weather forecast guidance. Each will have both scientific and technical challenges.

a. *Reanalysis/reforecasting*. For optimal performance, reforecasts require reanalyses with consistent statistics (that is, little change in bias and accuracy) for their initialization. The recently generated reanalysis for the NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010) involved many tens of person years of effort. However, because of the substantial changes in observation quality and amount during the reanalysis period, the statistical characteristics of the CFSR were later found to change with time. For example, sea-surface temperature biases in the tropical Pacific changed around 1999, at the advent of microwave-radiance assimilation (Kumar et al. 2012). Because of such challenges, NOAA's Climate Program Office recently funded a task force of researchers to develop methods for minimizing analysis error changes during periods with changes to observation-network characteristics. While such methods can be leveraged during the production of the next-generation reanalyses, it is prudent to expect that there will still be scientific challenges in addition to the practical challenges of obtaining sufficient computational resources and managing the input and output data. Should regular reanalyses not be available for reforecast initialization, then older reanalyses must be used. Ideally, the reanalysis initial conditions would be adjusted in some manner so that they have more similar bias characteristics to the current operational analysis; the exact procedures for performing such an adjustment would be an area requiring further research.

High-quality, high-resolution data sets such as surface temperature, wind, precipitation type and amount data sets are needed in the statistical post-processing procedures. Post-processed guidance is generated through a process of training against such analyses, and hence the ultimate quality of the product is largely dependent on the quality of these supporting data sets. For surface temperature, dew point, and winds, currently the high-resolution training data is provided by the RTMA (Real-Time Mesoscale Analysis; Pondeca et al. 2011) / URMA (Un-Restricted Mesoscale Analysis, which has a later data cutoff time). These methods currently use the High-Resolution Rapid Refresh (HRRR; Smith et al. 2008) to provide the background forecast. A common assumption in such assimilation methods is that the background forecast is unbiased. If it is not, then either the biases must be corrected, or the analysis will inherit the background forecast's bias, especially in data-sparse regions. Methods for improving the background forecast so that they provide unbiased guidance are perhaps the most challenging of several improvements that are needed within the RTMA/URMA system. This may be an area where further collaboration between OAR and NWS would be helpful.

An additional current challenge is that there are several areas of US interest, such as for Alaska and Guam, where there are not currently any high-resolution analysis or reanalysis data sets available, due to the sparsity of observation data. In such situations, either much different data sets may need to be used, such as satellite-derived precipitation estimates, or the guidance may need to be presented with minimal statistical post-processing.

b. *Advanced post-processing technique development*. The National Blend will require NOAA to depart substantially from their past procedures for data generation and statistical post-processing. The existing MDL post-processing procedures leverage ensembles in a limited way and are more oriented around station-based rather than grid-based post-processing. The post-processing methods for the National Blend must be able to extract maximum information from the shorter training sets of past forecasts and observations that will be available initially. However, when reforecasts become more readily available, the post-processing methods must be updateable with improved algorithms to appropriately use this richer data for improved forecast guidance. The development of advanced techniques in OAR and their transition to the NWS will need to be supported through programs like NGGPS and USWRP.

c. *Leveraging multi-center forecast data*. Beyond the logistical challenges of basing the post-processing on data sets from multiple center's prediction systems, discussed in the next section, there are significant scientific challenges to their synthesis. Different center's prediction systems may come with training-data sets of different lengths, and hence post-processing methods are needed that can both leverage the rich information that may be provided by some center's reforecasts and yet can still produce acceptable post-processed guidance when much less training data is available. The optimal

methods for combining the forecast data is also deserving of further research. One could imagine a post-processing method that would perform the correction of each center's systematic errors together with the downscaling and the combination of the data. Alternatively, these might be broken into two or three distinct steps. Ideally, the synthesis of multi-center post-processed guidance would be consistent with Bayesian principles, modeling the dependence of the analyzed state based on multiple models and their error relationships (Hodyss et al. 2015).

4. **Implementation issues.**

This white paper's abstract provided a list of several major implementation steps necessary to make the National Blend and reforecasting operational in the near future. Not every issue that each organization must wrestle with is identified; we highlight here only major ones that have some novelty to them with respect to current institutional roles and practices. Again, these issues are: (a) determining what new data sets are to be produced; what new reanalyses, reforecasts (at what resolution, frequency, duration, etc.), high-resolution surface reanalyses, additional NDFD elements, and so forth; (b) determining the specifications for an upgraded computational and storage system for reanalysis, reforecasting, and statistical post-processing; (c) identifying what additional Weather Forecast Office (WFO) - based data storage, communications bandwidth, forecaster workstation capability, and software improvements may be needed if the NDFD is to be augmented with additional probabilistic information. Again, we note possible synergies with NOAA/CIO's Big-Data initiative; (d) procuring and maintaining the hardware needed in (a) - (c) above; (e) developing the capacity for the regular production of global reanalyses and reforecasts, and making these data easily available in a timely fashion for post-processing system development inside and external to NOAA; (f) improving the quality of the high-resolution surface analyses, and generating high-resolution surface reanalyses to match the period of reforecasts; (g) changing the existing North American Ensemble Forecast System substantially; pending agreement with Canada, international sharing of raw guidance will continue, but product generation in the US will be switched to the National Blend; (h) updating and sometimes redesigning the post-processing software to use the greatly expanded training data and to produce a broader range of higher-quality, high-resolution deterministic and probabilistic guidance for the NDFD or other archives. We now discuss these issues in turn.

a. *Determine what new data sets are to be produced*.

Qualitatively, we have a general knowledge of what new data sets EMC should produce. These include global reanalyses, reforecasts, and high-resolution surface analyses and reanalyses. However, details of the specific configurations still need to be determined.

NCEP/EMC has indicated a willingness, pending identification of resources, to generate three tiers of reanalyses and reforecasts: (i) a shorter one (2-3 years) for the Global Forecast System (GFS) on a yearly upgrade schedule; (ii) ~20 years for the GEFS on a biennial

schedule; and (iii) ~1979-current for the CFS every four years. What are now three separate models (GFS, GEFS, CFS) will tentatively morph into three applications of a Unified Coupled Global Model, all ensemble based with their own reanalysis and reforecasts. Ideally, the reanalyses will be generated using a fully modern version of the forecast and assimilation system, at the same resolution and with the same assimilation system and model parameterization suite used operationally. It is possible this could be prohibitively expensive; in particular, the modern data assimilation suite may require O(100+) high-resolution members for estimation of covariances, and the cycling of 100+ high-resolution members may be prohibitively expensive. Hence, an important initial challenge will be to determine how much computational resources will realistically be available (step b, below) for reanalysis and reforecasting and to set the global reanalysis/reforecast configurations accordingly. Similarly, the computational expense of the high-resolution surface analyses should be considered. Given that the National Blend is requiring products to be generated not only for the CONUS, but also for Alaska, Hawaii, Guam, Puerto Rico, and 10-m winds over much of the globe (to drive ocean wave models), local, high-resolution analyses and reanalyses of winds, temperatures, precipitation amounts, and so forth may be needed for these additional domains.

Suggested configurations of reforecasts (e.g., the number of days between reforecasts, the number of members, and so forth) that minimize computational expense while providing sufficient value were discussed in the recent white paper (Hamill et al. 2014). Since the distribution of this white paper, several interested parties, notably hydrologists, have asserted that a more expansive reforecast may be needed, a data set that captures more of the extreme events than would a once-every-fifth or seventh-day reforecast. Perhaps some finer temporal granularity may be possible during periods where there were events of great hydrologic significance, or for forecasts of short leads, rather than finer granularity through the entire period of the reforecasts. Such methods for minimizing computational expense should be discussed periodically with partners.

b. *Determine the specifications for an upgraded computational and storage system for reanalysis, reforecasting, and statistical post-processing*.

This step will presumably be done in consultation with step (a) above; unrealistic requirements for computational resources in (a) may mean no reanalysis/reforecast at all. It should be noted that in the FY16 President's budget, $1.76M was allocated (see p. 86) to keep the Weather and Climate Operational Supercomputing System (WCOSS) system as a computational platform after its functional life span, in large part for the computation of reanalyses and reforecasts (personal communication, Steve Lord). The size of this machine may thus be useful in determining reanalysis / reforecast configurations.

Estimates will also be needed for data storage. For reference, the ~ 30-year, daily, 11-member, 2012-era GEFS reforecast had 28 of its fields stored at the native resolution (~ ½ degree) and 99 fields at 1-degree resolution. Data were stored for every 3 h to +72 h and every 6 h thereafter. Storage of this data set in grib2 format required 150-200 TB. Storage of the full

model states (on tape) was O(1 PB). With multiple higher-resolution reanalyses, reforecasts, and high-resolution surface analyses all required, a ballpark estimate is that ~ 2 PB of dedicated online storage would be desirable, and a deeper tape storage of ~ 10 PB. Of course, a more thorough analysis should be conducted, which may show creative ways of minimizing storage. Also, note that the disk storage need not be as on-time reliable as with real-time operational data; in case of disk failure, data should be set up to be readily recovered from tape, causing only a temporary delay of availability. Note that a key element of a sustainable modeling and reforecast enterprise is storing resulting data directly at its intended distribution point. More than four years after completing the CFSRR, NCEP is still working on transitioning the entire data set to NCEI. For future versions of reforecasts, this should be avoided by having an integrated data distribution strategy implemented before the data are generated.

It is possible that cloud storage of the data may be a less expensive option; exploration of such alternatives is encouraged. A partnership with cloud providers might be possible, whereby they host the data and make it free of charge in NOAA and with a nominal expense for outside-user access, which is an option being explored under NOAA CIO's Big-Data initiative.

c. *Identify what infrastructure improvements will be needed were the NDFD augmented with additional probabilistic information*.

Many national reports and plans (e.g., NRC's Fair Weather (NRC 2003), NRC's Completing the Forecast (NRC 2006), and the AMS Ad-Hoc Committee on Uncertainty in Forecasts (Hirschberg et al. 2011)) have recommended that the NWS fully embrace probabilistic forecasting to provide improved decision support. These recommendations informed the NWS Weather Ready Nation strategic plan (NWS 2011). This probabilistic data would be most visible and readily available to many NWS customers if it were available through the NDFD.

Suppose that in addition to deterministic forecasts of particular variables, the values associated with the 10th, 25th, 50th, 75th, and 90th percentiles of the forecast distribution were also conveyed via the NDFD (other methods of conveying probabilistic information are also possible, such as probabilities for categories, or parameters of probability distributions). This would thus result in the need for storage and dissemination of ~5x more NDFD data. Additionally, one might envision other products, such as probabilistic accumulations of precipitation over longer time periods (e.g., one day, three days, storm totals), and such probabilistic products could not be derived after the fact from probabilistic accumulations over shorter periods. Hence yet more data would need to be stored and transmitted if such variables were deemed needed in the NDFD.

What are the consequences of disseminating this value-added probabilistic information via the NDFD rather than via alternatives? Scientifically, the generation of such probabilistic data through the National Blend project is possible. Many of the algorithms used in the prototypes are probabilistic and could provide such information readily; others will take some but not an undue amount of research and development to convert.

Still, there are many issues associated with augmenting the NDFD with probabilistic data based on the National Blend. For example, for probabilistic information to be incorporated into the National Blend, the probabilistic data may need to be sent to WFOs. The forecasters presumably would then need to be able to affect the probabilities, either explicitly (editing the probabilities) or implicitly (e.g., editing the deterministic forecast value, which is used to shift probabilities). Were explicit grid editing of probabilities desired, this would potentially add to the workload of forecasters, conflicting with the vision outlined in Weather-Ready Nation, where forecasters are envisioned to perform less grid editing and instead provide more decision-support services. Should "implicit" grid editing be used, such as adjusting probabilities after the modification of a deterministic forecast, some R&D may be necessary in order to make sure that the adjustment procedure is a valid one.

Before committing to communicating the additional information via the NDFD and allowing explicit forecaster editing of the probabilities, the NWS will need to determine the hardware, software, and communications requirements to permit AWIPS-II consoles to do this. A current major limitation is local storage, which may need to be addressed through increased local storage or cloud storage. Another limitation is that existing AWIPS-II will need additional software to facilitate modification of the probabilistic information in ways that make it consistent with deterministic forecast modifications.

The full NDFD solution as scoped above is not the only way in which the NWS could convey the probabilistic information from the National Blend. Additional probabilistic fields could be disseminated through the NDFD but could be considered exempt from local grid editing; presumably then they would only need to be stored centrally. Another possibility is that such data would be computed in the National Blend but stored only in the NDGD (National Digital Gridded Database). Yet another possibility is that probabilistic information is conveyed by other standalone web applications, e.g., WPC's product here. In such circumstance, the NWS should strive to make the underlying data available not only via web interface but also accessible as downloadable files, such that more sophisticated users can leverage this information in their decision-support algorithms.

d. *Procure and maintain the hardware*.

Once the specifications for a system architecture to support the National Blend and reforecasting/reanalysis are identified, the NWS will need to procure the associated hardware and budget for its continued operation and maintenance.

e. *Develop the capacity to regularly produce and archive global reanalyses and reforecasts*.

Reanalysis will require a dedicated staff at EMC and partners to examine the quality of the reanalysis guidance and make algorithmic adjustments as necessary. For example, ECMWF has a staff of ~10 people dedicated to their suite of reanalysis products; as such, a

similar number at EMC should be expected. This staff is recommended to leverage the work of the recent NOAA/OAR/CPO Reanalysis Task Force.

Reforecasting as a procedure is relatively straightforward, provided that initial conditions are available from an ensemble-based reanalysis, as they will be for operational forecasts at EMC. EMC staff are encouraged to consult with staff from ESRL/PSD with regards to how to set up a reforecast archive and dissemination system that is convenient for the user community; some of this infrastructure could be recycled.

f. *Improve the quality of the high-resolution surface analyses, and generate high-resolution surface reanalyses to match the period of reforecasts*.

High-quality surface data assimilation will require attention to many details, including the quality control of the observational data, the modeling of background-error covariances, which may be significantly anisotropic and flow dependent near the surface, and the removal of bias from the first guess forecasts (near-surface first-guess fields are often contaminated by significant bias). It is presumed here that EMC will continue to upgrade its surface data assimilation to improve upon existing methodologies, potentially as part of a regular global full-atmospheric reanalysis.

If it continues to prove difficult to generate high-quality surface analyses, then perhaps alternative methods for generating them should be explored. See here for an example of a proposed statistically based method.

g. *Evolve the existing North American Ensemble Forecast System, and explore merging functionality into the National Blend*.

In 2003, the NWS entered into an agreement with Environment Canada for the sharing of global ensemble forecast data, including both raw guidance and bias-corrected information. More recently, the agreement was extended to include data from the US Navy global prediction system. Some additional detail is provided in a 2011 briefing. Some rather simple methodologies were used to provide gross bias corrections of the data for users wishing to generate products incorporating this. Given that there is much overlap with the National Blend, it would make sense to merge functionalities where feasible. There may be some practical difficulties in doing so, however. The list of variables that are bias corrected (see above) is more extensive than the list of National Blend variables, so additional post-processing and additional storage would be required for the National Blend, were these variables to be incorporated. There will also be additional personnel expense associated with planning and then making software changes, as well as potentially forming new agreements with Canada, US Navy, and other potential partners. Despite the initial up-front cost, there should be cost savings over the longer run, in that only one system would need to be maintained rather than two.

h. *Update/redesign the post-processing software to use the greatly expanded training data and to produce a broader range of higher-quality, high-resolution deterministic and probabilistic guidance for the NDFD or other archives*.

MDL has tried to exploit the forecast data at hand for statistical post-processing, which in the past has often been quite limited. With the EMC commitment to generate extensive reforecasts, MDL will need to redesign their post-processing software to be able to use much more extensive data sets. Given the need for a major redesign, this may also afford the opportunity to consider different archival formats, different process flows, and so forth. Another benefit of this redesign will be enhanced collaboration among NOAA scientists and their non-NOAA counterparts.

To adapt to all these changes, the NWS will require a staff of highly trained professionals with knowledge of the latest approaches being discussed in the literature, and with modern computer-science skills. MDL is the most logical place for this, given their critical mass with expertise in post-processing. Such a staff ideally would collaborate with scientists engaged in post-processing technique development at other NOAA institutions, including ESRL (e.g., precipitation amount and type in PSD), tropical cyclone track and intensity (at AOML and NHC), tornado warn-on-forecast (at NSSL and SPC), and hydrologic applications (at NWC). This collaboration would help improve the quality of guidance produced by all. This collaboration might involve more formal tasks for MDL, such as developing and maintaining a library of shared post-processing techniques and verification methods.

5. **Conclusions**.

This white paper addresses some scientific and implementation details related to three major anticipated changes to post-processing in NOAA: (a) the regular production of reanalyses, reforecasts, and high-resolution surface analyses/reanalyses; (b) the development of advanced post-processing techniques, and (c) and the generation of new, improved quality gridded guidance from multi-model, multi-center ensembles via the National Blend of Models Project. For post-processing to be a success, NOAA must plan for an orderly process of system development and data archival/dissemination. The authors hope that this document will facilitate NOAA's planning for these substantial changes, and that this document can inform strategic and budget planning in the years to come. There were many other issues related to post-processing that were not discussed here. See the appendix for a short summary of these.

The National Blend and the institutionalization of the regular production of reanalyses and reforecasts are major changes to the NOAA and NWS production of automated guidance. Though these changes will require additional resources, the anticipated impacts are large: the automated, post-processed guidance will in most cases be so skillful and reliable that forecasters will be freed from the task of much of the current manual grid editing. NWS forecasts will be improved in their skill and consistency, and the forecasters in times of

high-impact weather will not be distracted by grid editing and will be able to provide full-time attention to decision support for its customers.

The authors are happy to consult with any NOAA managers that wish to discuss these issues in greater detail.

Acknowledgments:

The authors would like to thank many people for consultations that were useful in producing this document, including Yuejian Zhu (NCEP/EMC), Jeff Whitaker (ESRL/PSD), Gil Compo (CIRES, University of Colorado), Arun Kumar (NCEP/CPC), Steve Lord (formerly of NWS/OST), John Schaake (NWS/OHD, retired), Peter Houtekamer and Normand Gagnon (Canadian Meteorological Centre), and Dave Novak (NCEP/WPC).

Appendix:

To maintain its clear focus on the two overriding issues of reforecasting and the National Blend, this document did not touch on several topic areas related to post-processing. (1) The issues associated with the development of supporting analysis data sets in areas of US interest such as Alaska and Guam was only touched upon briefly. If post-processed products are to be expanded to all areas where the NWS has responsibilities, then more comprehensive analysis and reanalysis data sets will be needed. The production of these may be particularly challenging given the paucity of data (e.g., precipitation analyses in regions of sparse in-situ data). (2) The post-processing of very short-range forecasts and very long-range forecasts were not discussed here either. Consequential decisions may be made on short-range forecasts, such as evacuation and protection in advance of a tornado. Consequential decisions are also made at very long time scales, such as a farmer's purchase of seed stock for the following planting season based on seasonal climate information. For nowcasts or very short-range forecasts (e.g., 0-6 hours), some judicious combination of nowcasting, extrapolation methods, and numerical weather predictions may provide the best guidance. These methods are being actively explored at research facilities like NSSL, ESRL/GSD, and NCAR. (3) The issues involved with the synthesis of short-range weather forecast guidance (~ 0-2 days) into the National Blend was also not considered here. This guidance commonly is produced by limited-area, high-resolution models, sometimes (as with HRRR) with hourly refreshes of the short-range forecasts, as opposed to the 6- or 12-hourly refreshes of the medium-range guidance considered here. A concept of operations for the inclusion of this data into the National Blend is needed, but this is beyond the scope of the current work. (4) Issues related to the post-processing of seasonal climate forecasts were also not considered. For these time scales, a different class of methods may be required, methods that work in the subspace of the few modes of climate variability where there may be remaining predictable signal, modes such as El Nino/Southern Oscillation, the Madden-Julian Oscillation, the North Atlantic Oscillation, the Pacific-North American pattern, and the Pacific Decadal Oscillation.

References:

Candille, G., 2009: The multiensemble approach: the NAEFS example. *Mon. Wea. Rev.*, **137**, 1655–1665. doi: http://dx.doi.org/10.1175/2008MWR2682.1

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203-1211.

Hodyss, D., E. Satterfield, J. McClay, T. M. Hamill, and M. Scheuerer, 2015: The overweighting of climatology in multi-model Bayesian Model Averaging. *Mon. Wea. Rev.*, in preparation. Available from Daniel.Hodyss@nrlmry.navy.mil.

National Research Council, 2003: *Fair Weather: Effective Partnerships in Weather and Climate Services.* 238 pp. Available at http://tinyurl.com/nl39lbu .

National Research Council, 2006: *Completing the Forecast. Characterizing and Communicating Uncertainty for Better Decisions*. Available at http://tinyurl.com/completing-the-forecast .

Hirschberg, P.A., E. Abrams. A. Bleistein, W. Bua, L. Delle Monache, T. W. Dulong, J. E. Gaynor, B. Glahn, T. M. Hamill, J. A. Hansen, D. C. Hilderbrand, R. N. Hoffman, B. H. Morrow, B. Philips, J. Sokich, N. Stuart, 2011: A weather and climate enterprise strategic implementation plan for generating and communicating forecast uncertainty information. *Bull. Amer. Meteor. Soc.*, **92**, 1651-1666.

National Weather Service, 2011: *Weather-Ready Nation. NOAA's National Weather Service Strategic Plan 2011*. Available at http://www.nws.noaa.gov/com/weatherreadynation/files/strategic_plan.pdf

Hamill, T. M., and others, 2014: *White Paper: A Recommended Reforecast Configuration for the NCEP Global Ensemble Forecast System*. Available at http://www.esrl.noaa.gov/psd/people/tom.hamill/White-paper-reforecast-configuration.pdf .

Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the nonstationarity in the bias of sea-surface temperature forecasts for the NCEP Climate Forecast System (CFS) Version 2. *Mon. Wea. Rev.*, **140**, 3003–3016. doi: http://dx.doi.org/10.1175/MWR-D-11-00335.1

Pondeca, M., and others, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: current status and development. *Wea. Forecasting*, **26**, 593–612. doi: http://dx.doi.org/10.1175/WAF-D-10-05037.1

Saha, S., and others, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057. doi: http://dx.doi.org/10.1175/2010BAMS3001.1

Smith, T. L., S. G. Benjamin, J. M. Brown, S. Weygandt, T. Smirnova, B. Schwartz, 2008: Convection Forecasts from the Hourly Updated, 3-km High Resolution Rapid Refresh (HRRR) Model. Preprints, 24th Conf. on Severe Local Storms, Savannah, GA, Amer. Meteor. Soc. (https://ams.confex.com/ams/pdfpapers/142055.pdf)