WRITTEN STATEMENT OF

William B. Gail, PhD
Co-founder and Chief Technology Officer, Global Weather Corporation (GWC)
Past-President, American Meteorological Society (AMS)

Before the
Subcommittee on Environment
Committee on Science, Space, and Technology
United States House of Representatives

A Hearing on: ***Advancing Commercial Weather Data: Collaborative Efforts to Improve Forecasts***

20 May 2015

Chairman Bridenstine, Ranking Member Bonamici, and distinguished members of the Subcommittee: It is a privilege for me to be present here today and provide testimony to you. Thank you for your invitation. My name is Bill Gail. I am co-founder and Chief Technology Officer of Global Weather Corporation, a provider of precision weather forecasts to businesses within the energy, media, transportation, and consumer sectors. I am also Past-President of the American Meteorological Society (AMS). I am a member of the newly formed Department of Commerce Data Advisory Council (CDAC) charged with recommending means for expanding the public value of Commerce data such as NOAA's, and was also a member of the National Research Council committee that authored the 2012 study *Weather Services for the Nation: Becoming Second to None*. My academic training is in physics and electrical engineering and I have nearly two decades of experience in the fields of meteorology satellites, weather services, and location-aware software.

Though I'm speaking to you today from my personal perspective, I wear two hats: first as a voice of the weather community through my AMS position, and second as a member of that community building my own startup company. My company has been successful in today's difficult economy precisely because high quality weather information is increasingly needed by businesses across many industries to serve their customers and improve operations.

Let me first commend you for the attention you are giving to the topic of commercial weather data availability, and particularly the controversial role of commercial satellite data. This Subcommittee has raised specific issues about use of commercial satellite data, and now broader questions about NOAA data in general. Though the question of commercial satellite data is itself important, I believe you have been wise to broaden the topic. Addressing these more general issues is critical to NOAA's ongoing success and its future progress.

SUMMARY OF TESTIMONY: A DATA ECOSYSTEM STRATEGY FOR NOAA

In this testimony, I will recommend that NOAA place increased emphasis on the breadth and depth of the data they acquire. This, I believe, is in keeping with both the Subcommittee's objectives and NOAA's goals of enhancing their services. NOAA has long relied on an elaborate NOAA *services ecosystem*, built on partnerships ranging from emergency managers to commercial companies. These partners extend NOAA's data and provide value-added services to end-users. This approach has been highly successful and is the envy of the world.

When it comes to data used by NOAA, however, the ecosystem is much less mature. *The objective of this testimony is to suggest a rationale and approach for creating a NOAA data ecosystem comparable in value to NOAA's highly successful services ecosystem. Such a data ecosystem would promote desirable characteristics of flexibility and robustness, enhancing NOAA's resilience to data loss scenarios and improving its technical performance.*

KEEPING OUR EYE ON THE OBJECTIVE

I want to make clear that *the end goal of all we do is serve the public*, not our institutions. All we do should be measured by that metric. Enhancing the role of the commercial sector is a worthy objective to the extent that it serves this goal. But any change in NOAA data policies or processes must recognize two critical considerations:

- *Weather is global.* Tomorrow's weather here in Washington, DC may have originated last week in patterns over Siberia. We need data from other nations to forecast our weather as much as they need our data to forecast theirs. Historically, this has been accomplished by international agreements enabling free and open sharing of data. This core principle has proven enormously successful. Should we seek changes, such as an increased role for commercial data, this must be viewed in the international context.

- *NOAA is the world's gold standard.* Despite widely discussed weaknesses in some limited areas, NOAA's overall program is still the envy of the world. While we should seek to improve NOAA, we must be very careful not to break what is working well, including the system of satellite data acquisition. During the 1990's modernization, the National Weather Service, as directed by Congress, employed a principle known as "no degradation of services"[1] to guide and monitor all changes made to the system. This was a wise principle then, and should be used to informally guide future changes.

We must keep these considerations in our minds as we proceed through the discussion.

[1] U.S. Congress. 1992. National Oceanic and Atmospheric Administration Authorization Act of 1992. Public Law 102-567, Sections 701-709.

COMING SOON: WEATHER FOR EVERY NEED, TIME, AND PLACE

Those of us involved in delivering weather services are confronted daily with the demand for new and different weather use cases. Part of this demand is driven by the rapid adoption of mobile phones, through which businesses and consumers can consume weather in new ways. Part of it is the ease of customizing and delivering weather information to meet these needs, given progress in web services and such things as cloud computing. The broader information world, epitomized by Silicon Valley startups and venture capital backing, gets this growing market demand. I know. In my role leading a startup weather company, I talk to many of these new-technology companies and hear about their needs. Who, a decade ago, would have imagined high-performance sports clothing that anticipates and adjusts to the weather?

The emerging information economy is the context for this change. The number of people throughout the world with access to quality weather information will increase by an order of magnitude in the next decade. Smart phones make that possible. Even those of us with such access already will find we are using weather information perhaps two orders of magnitude more often, as it becomes embedded in apps and smart devices in ways we may never even notice. The information required will have to be highly customized, matched to each user's needs, and delivered when and where they need it. We will no longer produce one forecast for the entire U.S., but instead one for each individual, and perhaps several for each business.

We are undergoing a revolution in weather usage, driven by the context of the world around us. One consequence is that NOAA requires significant advances in the breadth and type of data sources available to them.

THE LOOMING INFORMATION COLLISION

What does this mean for NOAA's data sources? Picture *a train headed down the tracks*. This train represents all of the present data sources – from satellites to balloons – through which NOAA[2] monitors weather and drives our forecast models. The train is now decades old, having been refurbished many times to keep it operating well.

Now imagine a *second train* is rapidly catching the first, travelling on a parallel track that was only recently laid. It represents an emerging breed of data producers, epitomized by the Googles and Microsofts of the world, as well as innovative providers more closely aligned with the weather field. Included are weather observations from automobiles, mobile phones, social networks, and a myriad of other sources never before available. It also reflects the rapidly growing volume of more traditional weather observations from non-NOAA sources such as

[2] Accompanied by many of us in other academic, public, and private organizations, all part of what is referred to as the weather, water, and climate *enterprise*.

mesonets and aircraft, which are still not being effectively integrated into the NOAA data architecture.

The future is not hard to foresee. Like it or not, these parallel tracks cannot remain separate for long. Indeed, we might picture a junction at which the trains will arrive in the near future. The trains can either collide, or with a bit of effort on the part of NOAA, hitch together. Hitching trains in no way means the "new data sources" necessarily replace the old. It does mean the old and new have to coexist, and ideally strengthen each other, in ways they currently don't.

NOAA IN THIS NEW INFORMATION WORLD

This situation is not unique to NOAA. It is something being faced by the weather community internationally, and certainly by many communities beyond ours. The particular issue that has been discussed extensively within this Subcommittee, NOAA's potential use of commercial satellite data, is but one element of this much larger data source transformation. It is a disruptive transformation, not one readily understood and accommodated.

But picture a world in which NOAA does not effectively hitch trains. A future Congress will be holding hearings such as this asking why substantial sources of information about the weather are not used at all in NOAA's weather models. Perhaps some large company will take on this task themselves, combining their data with NOAA's own free and open data to produce forecasts far more accurate than NOAA can. These commercial forecasts would perhaps not themselves be available free and open to the public.

NOAA's new information world will be characterized by data sources far more numerous and diverse than today. Some are very similar to its present sources, derived from government-owned sensors and systems built to NOAA specifications. At the other end of the spectrum is entirely ad hoc data, such as from Twitter. It may come and go in hard-to-anticipate ways, yet is still very valuable. Diversity of data sources is not entirely unfamiliar to NOAA, which already relies heavily on volunteer observer networks, for example. Over the past decade, NOAA has also greatly improved its ability to assimilate data from NASA's scientific satellites.

Why are these new data sources so important to NOAA? In our ongoing efforts to improve forecast skill, new data sources are the raw material we can't do without. They play many different roles:

- *Expanding data sources for NWP[3] assimilation.* NWP models are only as good as the data they assimilate. At some point, additional computing power cannot advance NWP

[3] Numerical Weather Prediction (NWP) models are the workhorses of weather forecasting. They replicate the present state and evolution of the atmosphere at regional and global scales. Human forecasters use them as

forecast skill without additional data at finer spatial and temporal scales. New data is essential, though it must be matched to progress in computing power, model resolution, and underlying model physics.

- *Validating NWP performance.* Assessing performance of NWP models at improved space and time resolutions requires increasingly fine-scale data. This is unlikely to be obtained entirely through traditional processes, such as building more NOAA-quality weather stations. Non-traditional data sources can contribute substantially.

- *Countering NWP latency.* A significant NWP weakness is the latency between data acquisition and forecast release time, resulting from the computational time of the model. Techniques such as post-processing[4], which rely heavily on observations, can be used to counter this latency by adjusting NWP output in near real-time to more closely match current observations.

- *Improving the initial analysis.* The initial analysis is the estimate of current conditions throughout the atmosphere, and it is the starting point for all numerical weather predictions. We are presently limited in our efforts to improve this by lack of observations. Improving accuracy and completeness of the analysis field is a primary driver for improving forecast skill at regional and global scales. A greater spatial density of observations, such as temperature and pressure, would improve the initial analysis.

- *Improving mesoscale severe weather forecasts.* Severe weather at regional and local scales, such as tornado formation and coastal storms, can be strongly impacted by highly localized phenomena. A greater spatial density of observations, such as surface temperature and pressure from mobile phones or vehicles, can improve forecasts for these events. Increased spatial and temporal density for upper atmosphere measurements is highly desired, though less addressable as a byproduct of consumer technologies.

- *Improving underlying climatology.* Climatology models – estimates of the normal spatial variability of weather conditions – are used for downscaling forecasts. Since forecasts are generally made using grid cell sizes larger than the variability of weather activity, downscaling based on climatology is used to estimate what is happening at finer scales. Finer scale observations would improve climatology models.

"guidance" to create a final forecast, but their increasing accuracy means they are used more and more as the accepted forecast without further human modification, especially beyond the one-day forecast.

[4] One example is the technique known as Model Output Statistics (MOS), which has been used for many years to improve forecast accuracy of critical variables such as temperature by about 20% at sensor locations (such as at airports). The National Weather Service employs MOS-based models.

- *Improving application-specific forecasts.* Increasingly, end-users demand forecasts specific to their needs, not simply a generic weather forecast. Wind energy suppliers, for example, need to know wind speed at the 80-100 meter height of wind turbines. This can be substantially different than the wind speed at ground level. Commercial vehicles need to know the exact weather conditions along every road segment, not an average. Knowing that the average condition of a road is dry does not help when there are icy patches.

Society's demands for these high-fidelity forecast improvements are growing. The financial benefits of addressing such improvements are increasingly clear. I prefer to picture NOAA as the leader in this data integration, not a follower. It is an inevitable future. If not NOAA, then it will be someone else. The lessons being learned from today's commercial satellite discussions, particularly regarding radio occultation data, are central to starting this process. We should not look back at why these discussions have been contentious, but look forward at what is possible for NOAA, and for the nation we all serve, by exploring the bigger picture.

CASE STUDIES

Though this vision addresses the future, the issues involved with its implementation are very much present today. Following are four examples of data sources that illustrate both the challenges and the opportunities faced by NOAA today:

- *Satellite temperature and water vapor sounding data.* Satellite sounding data is among the highest value data for improving forecast quality. This Subcommittee has already specifically addressed the potential for NOAA to purchase commercial sounding observations. Such observations have been proposed using both geostationary optical imagers and radio occultation satellites. H.R. 1561 includes provisions for pursuing these interests. However, a more developed and ready program, COSMIC-2, potentially provides a near-term source for valuable radio occultation data. This program, funded largely by Taiwan in cooperation with NOAA and the U.S. Air Force, builds upon the highly cost effective COSMIC program that is already providing data used by NOAA and major weather centers throughout the world. A number of key lessons have emerged from this discussion already:

    - *Cost/benefit/risk of new data and systems*. Evaluating the cost, benefit, and risk of proposed systems (in contrast to doing so for existing data sources) is a critical element of building a robust ecosystem, particularly if NOAA resources contribute to development of the source. Often, seemingly small distinctions such as calibration quality can have major impact on the value of an approach. Such assessments need to be done fully and carefully prior to acquisition decisions.

- *Roles of government and commercial data.* My opinion is that COSMIC-2 is a sound program, of good financial value to NOAA, and should be completed without risk from commercial substitutions. Yet robustness of data sources is valuable to NOAA. To the extent that proposed commercial sources provide redundancy or augment COSMIC-2 with additional data, and to the extent they promote emergence of robust commercial data, they are quite valuable and should be promoted. At minimum, a system capable of replacing COSMIC-2 data after its lifetime must be planned, and commercial sources present an excellent option.

- *Challenges around open data policies.* Open data policies may limit the business case options for commercial sounding providers and thus preclude such data from even coming available. This illustrates a clear conflict between two desirable goals: a) maintain open data for the benefit of the entire community, and b) promote the emergence of commercial data sources that can benefit the weather community and the public. The resolution of this conflict is not simple, with strong disagreement even about whether such data is defined as falling under the domain of data that should be free and open or not. We need thoughtful evolution of currently policies to resolve this, which should be accomplished within the context of the World Meteorological Organization (WMO).

- *Opportunity versus threat.* Commercial proposals such as those that have been put forth in this area reflect substantial initiative and risk-taking on behalf of the companies and their investors. Such initiative is often the seed of breakthroughs that take our field to new levels. New ideas always bring challenges that must be worked through to determine if the ideas are worthy or not. NOAA should embrace this business innovation process and the new ideas it produces. Issues associated with proposed opportunities should be worked out through open community dialogue with the proposers, along the lines recommended in the 2003 National Research Council report *Fair Weather: Effective Partnership in Weather and Climate Services*. All parties are best served when proposers offer credible plans that can be properly evaluated and seek community dialogue as the primary means for promoting their initiatives, though the need to protect competitive information should be respected.

- *Aircraft flight observations data.* Aircraft flight data is also among the highest value data for improving forecast quality. NOAA currently purchases data from U.S. carrier long haul flights (known as AMDAR[5]) through a commercial aggregator. The data is

[5] AMDAR is the acronym for Airborne Meteorological DAta Reporting. TAMDAR is the acronym for Tropospheric Airborne Meteorological DAta Reporting.

redistributed largely free and open to other international meteorological agencies but not to other parties. NOAA receives equivalent data from Europe and other partners. Aviation weather data from regional flights (known as TAMDAR), which provide more detail on the lower atmosphere, are also commercially available. Unlike AMDAR, they are sold commercially to meteorological agencies separately on a non-open basis. NOAA is not presently purchasing this data, while other agencies are. As presently configured, many airlines do not participate. For those that do, the partner relationships are fragile and there is risk of losing these data sources.

- *Surface observations data.* NOAA has an established network of more than 900 surface observations stations known as the Automated Surface Observation System (ASOS), deployed largely during the 1990's. Other nations have equivalent systems. Data are freely available for all users, with some technological limitations. Over recent years, both commercial and academic institutions have deployed additional networks[6]. Today, there are perhaps 50,000 surface observation stations of varying quality throughout the world, some with open data policies and some not. Yet NOAA uses few of these observations today.

- *Vehicle observations data.* Vehicle observations are an example of a non-traditional data source that could be used by NOAA and other meteorological agencies. With millions of vehicles travelling at any given time even within the U.S., the number of potential observation locations is orders of magnitude larger than traditional ground observation stations. There is some research suggesting vehicle data, through its fine spatial density, can improve near-term severe weather forecasts such as for tornadoes.

Each of these case studies provides different lessons for a NOAA data ecosystem. Understanding the lessons is essential to building a robust ecosystem.

EXTENDING NOAA'S DATA ECOSYSTEM

The U.S. weather enterprise has been built upon the concept of NOAA providing data, from observations to model output, free and open for use by others. The result has been an enormous ecosystem of value-added providers and researchers who themselves further the public benefit. Today, it is estimated that more than 90% of weather information reaching the public passes through this value-added process, most of which is commercial. Other nations and world regions

---

[6] A 2008 National Research Council study titled *Observing Weather and Climate From the Ground Up: A Nationwide Network of Networks* discussed the emergence of networks of ground-based observing stations, established by businesses, state and local governments, and even individuals. Some are research networks, some commercial, and some consumer. The report recommended, and the community has struggled to implement, aggregating these individual networks through a national-scale system that makes the data readily accessible.

approach this differently, with less reliance on value-added partners. As a result, the U.S. has the most vibrant and productive weather enterprise in the world.

This is a *services ecosystem strategy*. NOAA amplifies its resources for providing weather services through the resources of companies such as The Weather Company, AccuWeather, Earth Networks, and my own company Global Weather Corporation. Indeed, these companies are only the tip of the iceberg, with many more behind them. If NOAA were to attempt providing the services of the ecosystem as a whole, its budget would need to be many times its current value. Some NOAA relationships within this ecosystem are contractual and formal, but most are informal, driven by NOAA itself or mediated through organizations such as the American Meteorological Society. It works. The 2012 National Research Council study *Weather Services for the Nation: Becoming Second to None* emphasized the importance of leveraging this ecosystem to serve the nation.

More recently, the Department of Commerce has formed a Commerce Data Advisory Council (CDAC) to advise the Secretary of Commerce on ways that improved access to all Commerce data can benefit the nation. One possibility is facilitating an ecosystem of organizations that access, organize, share, and add value to Commerce data, similar to the NOAA model. We know this makes sense, as illustrated through the tremendous value already provided by companies such as Zillow using Commerce data.

We need to extend this model of an ecosystem, used so successfully by NOAA on the downstream services side, to the upstream data side through a *data ecosystem strategy*. That exists only in rudimentary form today. Our goal should be to extend NOAA's existing data acquisition model, not undermine it. We need to move rapidly, but cautiously, to succeed.

What is in the data ecosystem? Data potentially available to NOAA through a data ecosystem may be divided into five classes. These are my informal categories, selected to illustrate the issues involved in this testimony. They should not be considered definitive, particular in their characterization of data quality.

- *Class I - NOAA-quality data.* This data is specified by NOAA, and generally acquired through systems built to NOAA requirements. It meets operational expectations for reliability, availability, and continuity. Examples include the GOES and JPSS satellite systems, the ASOS ground-based sensors, and NOAA ocean buoys. In addition, equivalent data may be obtained from international partners such as weather agencies in Europe and Japan.

- *Class II - Research-quality data.* This data is specified and acquired by a research organization such as NASA. It is often of a quality that satisfies NOAA needs for assimilation into NWP systems. It may not fully meet operational expectations of availability and continuity. Examples of highest-quality data include NASA research satellites such as Aqua[7]. Other examples, sometimes of lesser yet still acceptable quality, include academic mesonets of ground-based weather stations.

- *Class III - Commercial-quality data.* This data is specified and acquired by a commercial organization. The data generally meets quality standards that satisfy the data's commercial purpose and are enforced by the acquiring organization. In some cases, such as with lightning data used in commercial weather forecasting, this purpose may be closely aligned with NOAA's mission. In other cases, such as vehicle temperature sensors, the purpose may be different and the quality standards may be lower. Some data, as from sensors on commercial aircraft, can be extremely high quality and reliable.

- *Class IV - Consumer-quality data.* This data is ancillary to consumer use of devices such as mobile phones. It generally has no quality standards and is not originally intended for weather application uses. However, there is an enormous and rapidly growing volume of such data. An example is pressure data from mobile phones.

- *Class V - Ad-hoc data.* This data is ancillary to a variety of business and consumer uses and is often poorly characterized. It may take a wide variety of forms. An example is Twitter data on storm impacts that can help guide rapid assessment of affected areas.

As used today by NOAA, this ecosystem is in its infancy. NOAA has traditionally focused its efforts on Class I data. Over the last decade, it has made increasing use of Class II data through access to NASA research satellites. The Joint Center for Satellite Data Assimilation (JCSDA), funded and operated jointly by NOAA and NASA, has played a key role in accomplishing this. It has been one of the success stories in expanding NOAA's data ecosystem. But there is still much research-quality data, such as from mesonets, which is not effectively used. Commercial data such as that from aircraft is not adequately employed. Existing NOAA systems, such as radar and surface observations stations, are aging. And Classes III-V data is used in only very limited ways.

Building a more robust NOAA data ecosystem is certainly not a simple endeavor. But, as with services, the benefits can be substantial. This is not an initiative that needs to be planned fully before implementation can begin. A robust data ecosystem can be built incrementally, through gradual changes to the present system. Key to success is setting long-term goals and strategies to start down the path of making this change. This was indeed the path taken with weather services during the 1990's and accelerated by the National Research Council report *Fair Weather: Effective Partnerships in Weather and Climate Services* in 2003. Many of the principles and

---

[7] Assimilation of NASA satellite data into NOAA forecast models has been a data ecosystem success story through a joint effort known as the Joint Center for Satellite Data Assimilation (JCSDA). Its efforts have ensured that NASA research data is sufficiently characterized to enhance forecast model performance prior to operational use.

guidelines established within this report for weather services may be applied to the development of a data ecosystem.

ADDRESSING CHALLENGES, LEVERAGING OPPORTUNITIES

Expanding this data ecosystem, as NOAA has done so well with services, makes sense. Yet there are always obstacles. NOAA's present observational system has faced a number of recent challenges. Most are well known and extensively documented[8]. They include:

- *Growing cost and procurement delays.* There is broad awareness of recent challenges with acquisition cost and schedule for NOAA's satellite systems. This has been a motivation for Congressional hearings seeking ways to improve the process. It is appropriate to question these acquisitions, and to seek improvements. Yet there is still enormous potential for addressing these issues from within traditional approaches to such data acquisition. Alternatives such as block buys, fixed price procurements, requirements simplification, and oversight streamlining present attractive options.

- *Aging technology.* As noted in the National Research Council *Weather Services for the Nation: Becoming Second to None* report, much of NOAA's present technology, such as radars and surface observation stations, was designed during the 1980's and built during the 1990's. Some have been upgraded, but the basic technology is now over two decades old. This limits its uses and potential improvements.

- *Limited flexibility.* The paradigm for NOAA's present observing systems is largely that of design-to-requirements. While this ensures that NOAA obtains the data it needs to fulfill its mission, it does limit flexibility and alternatives as well as access to new data sources that are unforeseen by NOAA. In some cases, this lack of flexibility translates into reduced robustness. We have seen this recently with the threat of a polar satellite gap and the challenging search for viable alternative data sources.

- *Restricted scope.* Among the biggest opportunities for new data capabilities is Classes III-V data. Partly due to its design-to-requirements paradigm, NOAA presently has no consistent means to access these advances. The growth is being driven by a variety of largely commercial trends, from use of big data to consumer adoption of mobile phones. Many of the uses of this data will be entirely new to NOAA, so substantial effort is

---

[8] The National Research Council completed two complementary reports in 2011 and 2012 regarding the National Weather Service. The first was a retrospective review titled *The National Weather Service Modernization and Associated Restructuring*, assessing lessons from that critical program. The second was a view to the future for the National Weather Service titled *Weather Services for the Nation: Becoming Second to None.*

required to leverage the data. One major constraint is that there is limited availability of Class III-V data for the atmosphere except at the Earth surface.

Evolving the existing NOAA data system would ideally resolve current challenges and create new opportunities. Establishing a truly comprehensive data ecosystem will require changing many of the traditional NOAA views of its data sources. This in no way means the new data necessarily replaces the old. Indeed, NOAA has many data needs that are unlikely to be met without data systems being built specially to their specifications. Yet the challenges of evolving today's NOAA data system into a more comprehensive data ecosystem should not be underestimated. A sample of the important issues includes:

- *Replacing government data sources with commercial.* Substituting commercial data for NOAA's Class I data is very challenging in practice. Class I data meets requirements developed by NOAA. Since those requirements are typically quite specialized, it will be rare for commercial data to exist that has not been specifically designed to meet NOAA's requirements. Without other markets for the data, NOAA effectively bears the whole cost of the data, whether provided as commercial data or as a system purchase built by commercial contractors. Yet a commercial capability, should one arise with a viable business model for at least some portion of Class I data, could bring value in terms of cost reduction. As noted below, NOAA would also need assurance of the existence of multiple sources. The replacement of Class I data through commercial sources is an area with large potential rewards, but also significant challenges.

- *Augmenting government data sources with commercial.* Expanding data sources to augment NOAA data, either to increase robustness or add new data types, is a wise strategy. This is particularly true for data sources that are already commercial products of Class III-V, available whether NOAA buys them or not. It may also be true for proposed Class I-II NOAA-centric projects, such as has been discussed for commercial sounding data, if demonstrated by rigorous cost/benefit analysis.

- *Assessing value of new or alternative data sources.* In most cases, it is not a simple matter to understand just how much new or alternate data sources can enhance NOAA's mission. Performing system-wide tradeoffs of data value are essential to any comprehensive data ecosystem. Traditionally, we trade off only one type of satellite data against another, or data only within one class rather than across classes to simplify the analysis. Such things as OSE/OSSE studies need to trade off a broad set of options, including satellite versus ground-based sources. NOAA should have available a flexible set of tradeoff tools, appropriate to the cost and risk of the new data being considered.

- *Technical capacity to use new data sources.* New data aggregation, analysis, assimilation, and statistics techniques will be needed to deal with new data sources. That is a broad technical challenge for our academic community. For example, we know temperature observations from vehicles are of poor quality. With millions of them at any given instant, however, can we extract high quality data? This new way of working with data – instilling data quality after the observation is made, rather than designing it into the sensor – is among the research progress we will need.

- *Extending open data principles.* Any ecosystem strategy such as this leads naturally to a reassessment of the principle of free and open data, which if adhered to literally may preclude access to important data sets. In general, U.S. promotion of this principle, led by NOAA (and now more broadly by the Department of Commerce), is sound and should be applauded. It has been a foundational principle for the growth of all weather services within the U.S. Given the global nature of weather data, and the corresponding importance of data we use from other nations, every effort should be made to support this. But inevitably, there are some data sources that will not be made available to NOAA (and to the global weather community) under such open data conditions. Such data can contribute to NOAA's core goal of enhancing public welfare, including safety. By rigid adherence to the open data principle, such data – and the benefit to public safety that comes with it – may not be made available to NOAA or other international weather agencies. This presents a dilemma; in such cases, the open data principle may not serve the public good.

- *International agreements.* NOAA is committed to international policy agreements through the World Meteorological Organization (WMO). Their Resolution 40, in particular, states "members shall provide on a free and unrestricted basis data and products which are necessary for the provision of services in support of the protection of life and property and the well-being of all nations . . ." Details of this principle, including guidance as to what should fall within the domain of "free and open", are included in the resolution. The principle of free and open data remains sound, but WMO's implementation was developed in an era of vastly different needs from today. It is appropriate to refine the original resolution, and there is some indication that WMO is receptive to doing so. In particular, it needs to evolve from viewing commercial data as a risk to traditional weather service data to being a complement. Free and open data is not an end-goal of its own, but rather a means to best serve the public. When it begins precluding access to data that can help NOAA (and international partners) keep the public safe, it introduces issues of its own that need to be resolved.

- *Resources.* Finally, there is the inescapable resource challenge. Finding the resources to accomplish this may, in today's budgetary environment, be the biggest challenge. We know that NOAA's return on investment to the nation is enormous. Our economy suffers from a nearly $1 trillion economic inefficiency resulting from our sensitivity to weather and climate[9]. Farmers experience this through drought, and energy suppliers through unseasonal weather. Reducing this inefficiency, through improved weather information, is a rare lever we have for driving economic growth. It is a worthy use of resources.

SEEKING INNOVATIVE PARTNERSHIPS

Our two trains will not hitch properly if we rely only on traditional public-private partnering mechanisms such as data buys. These mechanisms reflect the old information world from decades back, not the new. The new information world is characterized by business models, like freemium and shareware, that were unheard of when the data buy paradigm was first developed. The commercial information sector is innovating all sorts of new business models. It may be that none fit the need for NOAA's data acquisition, but the proliferation of new business models should itself be a lesson that new approaches can be found with focused effort.

Other U.S. government agencies have explored very innovative public-private partnerships. For example, the intelligence community has used direct investment in technology companies through the widely discussed In-Q-Tel non-profit venture capital firm to seed innovation. Key technologies that benefit us all, such as Google Maps, have emerged.

Could this be done within NOAA? The In-Q-Tel model may not be directly applicable, but it does illustrate the potential for new approaches. To illustrate the possibilities, consider a commercial satellite system that produces foundational data for NOAA along with additional data to be sold commercially. The additional data might allow post-processing of any NOAA forecast model using the foundational data to produce more accurate specialized results. This is not a perfect mechanism. But it does illustrate that new partnership ideas, with the potential to bridge the issue of open and proprietary data, are possible.

NOAA has a long and successful history of data buys. These include radar imagery for ice monitoring, ocean color data for identifying such things as algal blooms, lightning data, and more. NOAA claims that it has adequate procurement tools to accomplish data buys. Data buys will remain an important element of the data ecosystem, but innovative new mechanisms will be needed as well.

We may think of data buys as falling within one of two categories. In the first category are what we might call *project data buys*. They involve data specified by NOAA, systems designed specifically around NOAA's needs, and limited markets for the data outside NOAA – in essence, a data project. Benefits of this type of data buy, as compared to NOAA procuring the system

[9] Jeffrey K. Lazo, Megan Lawson, Peter H. Larsen, Donald M. Waldman, U.S. Economic Sensitivity to Weather Variability, *Bulletin of the American Meteorological Society*, June 2011.

itself, may exist but are limited. For example, if NOAA funding were not available, this data source would cease to exist. It is thus not an independently robust data source, and NOAA will likely pay the full cost of the data. The second category is *product data buys*. They involve data products for which there is a market separate from NOAA. The data source exists independent of NOAA, and can be considered independently robust if there are multiple suppliers (so failure of any one company does not jeopardize the source). As one of many buyers, NOAA would not be paying the full cost of the data. This distinction is critical for NOAA when considering options.

A successful NOAA ecosystem must, in the long run, be more than a list of data buys. An ecosystem is not necessarily a set of contractual relationships, but often simply working relationships and group interactions. This is what NOAA has learned so well from the services side. For example, the National Weather Service (NWS) has implemented a program called Weather Ready Nation that has developed thousands of informal partnerships already to amplify NWS efforts.

In many ways, NOAA already has a strong start on building an ecosystem. The Joint Center for Satellite Data Assimilation (JCSDA) is an excellent example of NOAA's willingness to expand the data ecosystem from operational to research satellites, as well as innovation in establishing new institutions and processes that make it possible. This paradigm should now be extended to the broader data community.

A PATH FORWARD

With expanding societal needs, NOAA will be required to grow capabilities at a rate that likely exceeds its resources for acquiring data. The best solution is to leverage data investments being made outside NOAA, in the commercial and academic communities. While NOAA will long need NOAA-specified data similar to today's GOES and JPSS systems, building an ecosystem of data suppliers – drawing from all five data classes and calling upon innovative new techniques – is a wise strategy to keep pace with the technological advances going on around NOAA.

I believe the concept of a NOAA data ecosystem, comparable in importance to NOAA's successful services ecosystem, is worthy of the substantial attention it would need for implementation. It will require guidance, support, and resources from Congress. It will motivate enhanced collaboration with NOAA's international partners to do similarly. And it will involve close collaboration with the community as a whole. To accomplish this, I would like to suggest the following:

1. NOAA, with support from Congress, should establish and build upon the concept of a *data ecosystem*, equivalent to what it has done successfully for services, to enhance its

operations. This will enable NOAA to better leverage the results of the information revolution going on throughout the commercial world.

2. NOAA should lead the international community in following this model. NOAA's efforts should be pursued within the context and goals of its international collaborations, including the WMO 40 data policies. NOAA should lead efforts to extend WMO 40 to recognize the context of new data sources. NOAA's ability to function within the context of global meteorology requires us to respect international definitions and guidance such as that for open data.

3. General legislative guidance on broadening the data ecosystem is valuable, but decisions on which particular data source options should be pursued are best left to NOAA.

4. NOAA, and Congress, should seek external guidance, such as through the National Research Council, regarding approaches and challenging issues (such as updates to open data principles) of this initiative.

5. As needed, a data ecosystem can be implemented in small steps toward the long-term goal of a vibrant data ecosystem. Near-term opportunities, such as the emergence of commercial options for satellite sounding data, should be used as examples to address and resolve issues, rather than deferred while NOAA establishes architectures or plans.

6. NOAA, and its data ecosystem organizations, should be informally guided in all efforts by the principle established during the Modernization of "no degradation of services", as well as the overarching goal of serving the public.

Weather legislation isn't considered within Congress often. In deliberating the evolution of data sources used by NOAA, I urge you to take a decade-scale view. The legislation you pass needs to stay relevant despite the enormous advances expected within information technology over that timescale. In this context, providing NOAA with the resources needed to develop a true data ecosystem will pay off to the nation many times over.

BIOGRAPHY

William B. Gail is co-founder and Chief Technology Officer of Global Weather Corporation, a provider of precision forecasts for weather-sensitive business sectors, and is the current Past-President of the American Meteorological Society. He was previously a Director in the Startup Business Group at Microsoft, Vice President of mapping products at Vexcel Corporation, and Director of Earth science programs at Ball Aerospace. Dr. Gail received his undergraduate degree in Physics and his PhD in Electrical Engineering from Stanford University, where his research focused on physics of the Earth's magnetosphere. During this period, he spent a year as cosmic ray field scientist at South Pole Station.

Dr. Gail is a Fellow of the American Meteorological Society and a lifetime Associate of the US National Academy of Science's research council. He is currently a member of their Board on Atmospheric Sciences and Climate, and has participated on many prior Academy committees, including the 2012 review of the National Weather Service and the 2007 Earth Sciences Decadal Survey. He is a member of the US Commerce Data Advisory Council and serves or has served on a variety of other editorial, corporate, and organizational boards. His book *Climate Conundrums: What the Climate Debate Reveals About Us* was published in 2014.