

Forecast evaluation: No longer just Root-Mean-Squared Error

by Barbara Brown*

There is no doubt that weather forecasting has experienced rapid changes over the past 20 years, with the advent of ensemble models to make operational predictions, the acceptance of probabilistic forecasts by many segments of the meteorological and user communities, and the development of high-resolution prediction capabilities that can create images resembling the output of remote sensing systems like satellite and radar.

During the same period, the processes for determining whether these new forecasting capabilities produce accurate, reliable and useful information have also experienced a renaissance. While these processes have traditionally been called "verification", it now might be more appropriate to call the evolved processes "evaluation" or "assessment." Although "evaluation" is often considered the process of measuring the value of a forecast, the measurement of forecast quality is an important component of the process of estimating forecast value. And while "usefulness" is not often considered an outcome of the verification process, new approaches for verification clearly can provide insights into this important attribute. Today's

forecast evaluation approaches have the ability to provide much more information about forecast performance than has traditionally been possible, and as they continue to evolve, verification methods will have an even greater role in facilitating the optimal use and improvement of weather forecasts.

To understand this evolution of verification approaches and its implications, it is important to consider where the science of verification got its start. Many of us are familiar with the story of Sergeant John Finley and his evaluation of tornado forecasts in the late 1800s. Specifically, Finley determined that his forecasts had excellent accuracy (to be precise, 96.6% of them were correct). However, it was quickly pointed out by Finley's meteorological colleagues that the accuracy rate would have been increased to 98.2% if Finley had never forecasted a tornado!

The importance of this event (in addition to being somewhat humorous) is that it led a number of scientists and mathematicians to think about what *is* the appropriate way to evaluate forecasts? It led to the development of a number of the scores that are still used today to measure forecast performance; to development of the idea that the forecast evaluation process is by nature *comparative*; and to ideas about measuring the value of forecasts (see Murphy 1996 for an excellent overview of the "Finley affair" and its repercussions).

(continued on page 12)



Lightning strikes southern Kansas in June (Photo by Rachel Butterworth)

In This Issue

Flash flood prediction tools	2
Cold advisory for newborn livestock	3
Do improved forecasts save lives	4
Engaging End Users	6
Fire Dogs book review	8
Conferences & Opportunities	9
Jobs & Opportunities	11
About Us & Contact Us	16

Developing flash flood prediction tools for the Front Range David J. Gochis*, David N. Yates** and Wei Yu***

Each year flash floods cause of millions of dollars in property and economic losses and, at times, claim dozens of lives. Currently, there are relatively limited capabilities in predicting flash flood events in terms of lead time, spatial accuracy and event severity. While a few flash flood events are the direct result of improper or failed engineering design (e.g. dam failures), nearly all events possess a weather component that drives a severe hydrological response. Reliable prediction of flash flood events presently resides as a 'nowcasting' challenge whose lead times extend from a few minutes out to about 1-2 hours. Under potential flooding circumstances. Numerical Weather Prediction (NWP) is most often used as a 'guidance' product to help delineate Range during the summers of favorable regions of strong and, often, slow-moving or re-training rainfall events.

Obviously, land surface hydrology plays an important role in flash flood forecasting. Land surface hydrology can directly impact the atmosphere through control of surface energy balance partitioning of incoming solar radiation into sensible and latent heat components and, in turn, planetary boundary layer growth and convective initiation. Land surface hydrology also impacts runoff production through control of soil infiltration capacity, which can often be very limited under certain conditions.

The flash flood prediction system recently developed at the National Center for Atmospheric Research

(NCAR) aims to improve characterization of terrestrial hydrology during inter-storm periods via hydrologic data assimilation and seeks to advance our understanding of how the quality of quantitative precipitation estimates (QPE) and quantitative precipitation forecasts (QPF) from radar, nowcasting and NWP systems, respectively, will impact flood predictions. We coupled a state-of-the-art land surface model-a hydrologically-enhanced variant of the community 'Noah' land surface model called 'Noahdistributed'-to the NCAR High **Resolution Land Data Assimilation** System (HRLDAS) and the Weather Research and Forecasting (WRF) model and deployed the system over the Colorado Front 2008 and 2009. The combined

hydrological model and land data assimilation system is used to update or 'spin-up' land surface and channel flow conditions prior to particular forecast cycle.

In the NCAR system, QPEs are derived in real-time using the Level-II data from a mosaicking, or merging, of three National Weather Service (NWS) NEXRAD radars operating in the Colorado Front Range region—Chevenne, Wyo., Denver, Colo., and Pueblo, Colo. These QPE products have been calibrated and independently validated against the Denver Urban Drainage and Flood Control District ALERT rain gage network, as well as research radars operating within the forecast domain. From these instantaneous radar fields, short-term (~30

(continued on page 11)





Simulated Streamflow Cherry Creek @ Steele -Aug. 8, 2008

Cold Advisory for Newborn Livestock: The Stars Aligned

by Tanja Fransen*

Several years ago, I was informally introduced to the idea of integrating meteorology and social science while on a small commuter plane with three other people during one of my first business trips after moving to Montana. One was a well-known local rancher who just happened to also be the president of the National Cattlemen's Beef Association. He emphasized to me the enormous importance of weather forecasts to cattle operations and the extent of weather impacts on the industry.

A few years later, I attended the first Weather and Society*Integrated Studies (WAS*IS) workshop in Boulder, Colo., and fortuitous events began transpiring in my professional life and building upon one another until the stars all aligned, and the Cold Advisory for Newborn Livestock (CANL) came to life.

When I attended the WAS*IS workshop, Hurricane Katrina had just occurred, and I think many of us wanted to save the world from weather hazards. It quickly dawned on us that this goal was outside of our sphere of influence, and smaller, more manageable projects were born. My group included Jason Samenow, a climate analyst for the U.S. Environmental Protection Agency (EPA), who was working with Dr. Larry Kalkstein on a team that was developing the Heat Health Guidebook. I guickly learned how important Dr. Kalkstein was in the world of biometeorology after our week in

Boulder, listening to discussions mentioning his name.

The WAS*IS project we developed had two components for assessing vulnerability: Heat, which was Jason's portion of the project and cold, which I worked on. We met again a few months later, and presented our results to our WAS*IS group, but I was still at a loss as to what to do with the information I had gathered and how to apply it.

Several months after I finished the WAS*IS project, our National Weather Service (NWS) office in Glasgow, Mont. was contacted by Dr. Kalkstein. Dr. Kalkstein was interested in expanding his work to look at cold weather impacts and one of his heat/health partners within NWS had suggested our office as a good starting point. Dr. Kalkstein's work had generally been focused on heatrelated mortality, and he was instrumental in the development of heat/health warning systems across the globe. Now that he was hoping to expand into the realm of cold weather, I was starting to see potential connections with my WAS*IS project. We applied for, and received, a grant through the Cooperative Program for Operational Meteorology, Education and Training (COMET) Partnership Program to look at the feasibility of doing a cold-weather warning system in a rural area.

As the project developed, we found that it was very difficult

to see any identifiable signal in mortality data for humans during arctic outbreaks. The human population in northeastern Montana (approximately 50,000 people), and across the northern Great Plains, is simply too small to yield statistically significant relationships between mortality in any season and the synoptic-scale weather pattern. We researched the main impacts from cold weather in the region and found that livestock losses had a huge impact in Montana.

Thus, the project focus shifted from looking at the feasibility of developing a cold warning system for humans to that of developing such a system for livestock. It is much easier to identify relationships between livestock mortality and weather because the animals are generally exposed to the environment at all times and because there is a much larger cattle population than human population in the area. There are approximately 2.4 million head of cattle in the state of Montana, but the human population has yet to reach one million (USDA-NASS 2008, U.S. Census Bureau 2007). In 2005, weather-related calf losses resulted in a loss of approximately \$6.3 million in Montana alone (USDA-NASS 2006). It turns out that cold greatly affects cattleand cattle producers-across the country; nationally, approximately 95,000 calves die each year due to cold stress (Azzam et al. 1993), resulting in an estimated \$38

(continued on page 14)

From the Director: Do improved forecasts save lives? by Jeff Lazo*

This month I would like to raise the question: What can or should the weather community do to justify increased interest and investment in improving weather forecasting and warnings? Having a good answer to this question would seem to be important to the weather enterprise. However, I won't answer the question but simply raise it.

The mission of NOAA's National Weather Service (NWS) is to "provide weather, hydrologic, and climate forecasts and warnings for the United States ... for the protection of life and property and the enhancement of the national economy" (http:// www.weather.gov/mission.php). Similarly, the function of the "weather enterprise" as defined by the Weather Coalition is to "... develop and distribute weather products and services to the user community in order to protect life, reduce risk to property, and enhance economic competitiveness (http://www.weathercoalition.org/ purpose introduction.html).

It would seem, then, that a critical aspect of improving weather forecasting and warning should focus on reducing fatalities and injuries. Indeed, deaths and injuries from severe weather make headlines somewhere in the United States virtually every month, and numerous forecasters say they are deeply and personally affected when there is loss of life due to weather. Table 1 shows the average annual number of weather-related fatalities during the 1999-2008 tenyear period (http://www.weather. Table 1: Average Annual Weather Related Fatalities in the United States

(1999-200	5)
Heat	162.2
Hurricane*	117.3
Flood	65.9
Tomado	62.9
Wind	43.3
Lightning	42.4
Winter	33.4
Cold	21.2
Other	80.8
Total all weather	629.4
Damages (\$M)	\$23.149

* The 9 year average for hurricanes without the 2005 season that included Hurricane Katrina is 17.4 per year.

gov/os/hazstats.shtml). The loss of 630 people per year seems unacceptable given the advanced state of forecasting and warnings.

The data from which Table 1 was derived are from the NWS Office of Climate, Water, and Weather Services. For lightning, tornadoes, floods and hurricanes, this data is reported from 1940 through 2008, providing a nice time series of data. Using this data, I did a very simple ordinary least squares (OLS) regression on the number of fatalities per 100,000 people in the total U.S. population to see if there were obvious trends in fatalities over time. Results are reported in Table 2. The coefficient shows how much the number of fatalities per one million people has changed per year in the 1940-2008 time period. I should caveat this by saying that I have some concerns about the reliability and validity of the data in the U.S. Natural Hazard Statistics, which are largely based on data from the National Climatic Data Center's Storm Data reports. Emily Laidlaw and I are looking

at this in our work on Storm Data and will report more on this in the future.

Looking at Table 1 first, you can note that for all of the hazards the fatality rates have been decreasing (i.e., the coefficients are all negative), meaning that over time fewer people per one hundred thousand are dying from these weather hazards. For floods and hurricanes this was not statistically significant at usual levels of significance, although the 15.3% significance on floods does suggest that fatalities due to floods have been decreasing.

Does this mean that better forecasts and warnings have saved lives? Not necessarily. A myriad of complex social phenomena may have led to these reductions, including better building codes, more education on safe behavior, and fewer people working out of doors. Making the case that better forecasts have saved lives is not that easy, and arguments relating reduced fatalities with improved

Table 2: OLS Regression of Per 100,000 Fatality Rate (n = 69)

(· · · · · · · · · · · · · · · · · · ·			
Hazard	Coefficients	P-value	
Lightning	-0.33	0.000	
Tornado	-0.17	0.000	
Flood	-0.04	0.153	
Hurricane	-0.01	0.721	
Hurricane (without 2005)	-0.05	0.028	

weather forecasts should be examined carefully. The analysis presented in Table 2 is very simplified, and more sophisticated work can better sort through the details of the relationship between weather forecasts and injuries and fatalities. For instance, Kevin Simons and Daniel Sutter (2008) have looked at the impact of longer lead times of tornado warnings on injuries and fatalities.

When I take the 2005 hurricane season out of the analysis (e.g., Hurricane Katrina and the other hurricanes of 2005), I do detect a statistically significant reduction in hurricane fatalities over time. Does removing the 2005 data from the hurricane data mean that better hurricane forecasts and warnings have saved lives? Again, not necessarily. What happened in 2005 serves as a stark reminder that, even with really good forecasts and warnings, complex social systems, vulnerabilities and constraints will turn extreme weather into disasters. As disasters are largely social constructs (http:// understandingkatrina.ssrc.org/ Dynes Rodriguez/), it often will be more productive to understand and address social issues than it will be to improve forecasts (a common theme in this newsletter).

Does the data in Table 1 show that better forecasts and warnings can save even more lives? Not necessarily! There may simply be a point at which no matter how good the forecasts and warnings are some people are just going to be in the wrong place at the wrong time. And then there will always be those people who head to the shore to ride the waves during a hurricane. That is fundamentally not a forecasting problem.

So, given all this, does the possibility of reducing weatherrelated fatalities provide a strong argument for increased interest and investment in improving weather forecasting and warning? Again, not necessarily! Based on an idea that Rick Anthes presented last summer at a National Research Council (NRC) meeting at Wood's Hole, Table 3 shows the fatalities rates per 100,000 for the 15 leading causes of death in 2006 in the United States according to the Center for Disease Control—and the total weather related fatality rate derived from Table 1. While some of the 15 leading causes listed in Table 3 may include some component of weather (e.g., heat-related stress leading to heart failure or weatherrelated automobile accidents), decision makers allocating resources to reduce fatalities based on this information probably wouldn't be too concerned with weather forecasting and warnings.

So—returning to my question: What can or should the weather community do to justify increased interest and investment in improving weather forecasting and warnings? I presented the data in the tables here simply to raise the question of whether arguing that improved weather forecasts will save lives is the strongest argument in support of improved weather forecasting. I'll let you find the answer.

*Jeff Lazo (lazo@ucar.edu) is the director of NCAR's SIP.

Reference

Simmons, K.M. and D. Sutter. 2008. Tornado Warnings, Lead Times, and Tornado Casualties: An Empirical Investigation. *Weather and Forecasting*. **23**(2):246–258.

Table 3: Death per 100,000 for the 15 leading causes of death in 2006: United States and 10-yr average for weather

Cause of death	Per 100,000
All causes	810.4
Diseases of heart	211.0
Malignant neoplasms	187.0
Cerebrovascular diseases	45.8
Chronic lower respiratory diseases	41.6
Accidents (unintentional injuries)	40.6
Diabetes mellitus	24.2
Alzheimer's disease	24.2
Influenza and pneumonia	18.8
Nephritis, nephrotic syndrome and nephrosis	15.1
Septicemia	11.4
Intentional self-harm (suicide)	11.1
Chronic liver disease and cirrhosis	9.2
Essential hypertension and hypertensive renal disease	8.0
Parkinson's disease	6.5
Assault (homicide)	6.2
All weather hazards (1999-2008 average - with 2005 hurricanes) Source: <u>http://www.cdc.gov/nchs/data_access/Vitalstatsonline.htm</u> , <u>http://www.weather.gov/os/hazstats.shtml</u> , U.S. Census Bureau.	0.000978

Strategies for Engaging End Users to Optimize Weather Product Utility

by Bill Mahoney*

The primary objective of the weather enterprise is to provide weather hazard alerts, advisories, and prediction information to support decision making to save lives and property. We know from experience and research results that many users of weather information misinterpret the information or have difficulty extracting relevant information, as products are not generally tailored by the providers for individual decision makers.

Given the millions of end users and billions of forecasts that are obtained each year (Lazo 2009), it is not practical to tailor the information for each individual end user, so most weather products are generalized. The commercial weather services add value to the forecasting process by tailoring products to meet the needs of their clients. Tailoring products improves the likelihood that products will be properly interpreted, resulting in benefits to end users.

There has been a lot of discussion lately about the use and value of weather information and how the weather community should focus on improving the way we provide products and communicate risk and uncertainty. Congressional staffers have stated in AMS-sponsored meetings that more should be done to improve the delivery of weather products and services, as hundreds of weather products are available online without much guidance about how to interpret or use the information. It is wellrecognized that a perfect weather advisory or forecast may be useless if misused. We really don't have a good grasp of the overall utility of our weather products because we don't have a good measure of how the information is being used or misused. I suspect that a significant fraction of the weather information available today is not well understood by the public at large and, hence, is not as useful as we would like to think it is because its full potential is not being realized.

I believe that a renewed focus on improving product delivery of the current suite of weather products would quickly generate benefits that would likely exceed the incremental improvements in the product forecasting skills that are being developed on an ongoing basis. Millions of dollars are spent each year on research and development aimed at improving weather forecast skills by a few percent. How much is spent per year to improve the utility of existing weather products by studying how products are being used by individuals and various business sectors? I believe the return on investment from optimizing the current suite of weather products would be large if the weather community did more to infuse social science research. methods, and capabilities into the weather product development production line.

Now with the above 'soap-box-'

oriented discussion as background, I am going to discuss some basic lessons that I have learned from 25 years of engaging end users during the user needs assessment process associated with developing user-oriented weather decision support systems. The goal of this article is to share some wisdom on this topic so that others can be more productive in their quest to understand more fully what decision makers want or need with respect to weather products and services. I have found that following these recommendations improves the likelihood that long term relationships can be established between the weather information provider and end user, resulting in products that have a high level of utility.

Decision Support Systems: Let me begin by defining a decision support system (DSS). Many people think that a weatheroriented DSS is a computer workstation that provides usertailored weather information to end users. While this is an example of a sophisticated DSS, a DSS can actually include newspaper weather pages, radio and television weather broadcasts, text messages, and web content. People get their weather information from various sources and each of them support their decision-making process. When we think about optimizing weather information, we need to look at all weather information sources. In this article, I am mainly focusing on a discussion of best

practices for engaging end users when the envisioned DSS will allow user-tailoring.

Assessing End User Needs:

Although many users of weather information believe they know what they want, I have found that in the vast majority of cases, they really don't know what is available or even possible to provide. Obviously, more sophisticated and experienced users of weather information (e.g., pilots, energy traders) tend to be more proactive in searching for weather content and more aware of product capabilities and limitations; however, the vast majority of weather information users are not well-versed in our field, have limited exposure to weather capabilities, and are often not even sure where to obtain information beyond a radio or television broadcast. I have been surprised many times by the number of people who are in a business that is critically sensitive to weather (e.g., transportation, agriculture, construction, commercial trucking, etc.) and yet have not sought out weather information beyond what they passively receive from the media.

You can almost be certain that an end user will know his or her job better than you, so it is important not to come across as a "pointy-headed" scholar who "knows" the person's job or decision process better than they do. When engaging end users to obtain information about their needs for weather information, it is imperative to develop a relationship of trust with them. You really need to "get inside their head" and environment and work through some scenarios that will help you appreciate their daily duties. If at all possible, spend time shadowing them on the job. It may take a while for them to be comfortable with a "stranger", but once they get used to your presence, they will open up and provide valuable insight into their daily operations and decision processes.

More often than not, my discussions with end users begin with the users stating that their current level of weather information is "good enough" or that weather does not really affect their job. After some additional exploring (which can take hours or days), you learn that weather actually has a significant impact on their daily decision process and that they don't know what additional information could be provided to support their decisions. I like to call this process "exploring the art-ofthe-possible". What often starts out as an awkward initial dialogue can end up with an end user who has transitioned from being a skeptic to someone who becomes quite excited about the possibility of having more relevant information available to them

Assessing the Culture: A major pitfall to avoid is coming into a work environment advertising that you have a solution before you have even begun to explore user issues, concerns and needs. Don't assume you know what problem you are trying to solve. I remember a time when I was working with airline cargo pilots exploring what weather products they wanted in the cockpit. The consensus was that they wanted to see satellite cloud images, temperatures, and precipitation data. After several rounds of dialogue I asked why they needed all these data, and they responded that they wanted to assess the likelihood of in-flight icing at their flight level. They didn't just ask for a route-specific in-flight icing product because they assumed that it did not exist or was not possible to generate. So, it is important to spend time asking probing questions because the answers may not come in a straightforward manner.

It is also important to understand the culture of the organization you are working with. In some workplaces, a computer-based DSS may be seen as threatening as it could be perceived as a system that will eventually be used to replace someone's job. Thus, the role of the system should be articulated. An assessment should be performed to determine if a DSS could be used to help someone perform mundane tasks, such as data analysis, or higherlevel analytical tasks required to support more advanced decisions. Will the weather DSS be used by lower skilled workers, supervisors, middle managers, senior managers, or executives? Knowing the job category(s) that will utilize the DSS will allow intelligent judgments to be made about the design and provide insight into the human factors issues

Technical Infrastructure: It probably goes without saying that an evaluation of the technical capabilities of the target organization is required so that a solution is not selected that cannot be implemented or supported. For

Review of The Fire Dogs of Climate Change: An Inspirational Call to Action*

by Ilan Kelman**

"In the Chinese Year of the Fire Dog (February 2006-2007), I opened my heart a little to the Earth" (p. 10) writes Sally Andrew as the first sentence in The Fire Dogs of Climate Change: An Inspirational Call to Action. Her discoveries send her into despair at the wreck that humanity is making of the planet, followed by elation at those around the world who select different pathwayspathways that reduce and undo the damage. Her focus is contemporary climate change, one of humanity's many environmental and human catastrophes.

Fire Dogs is 142 pages plus a greenhouse gas emissions flowchart. The book is balanced between Andrew's personal journeys and facts about climate change, covering the science and actions necessary to tackle the problem. These chapters are interspersed, continually keeping the reader engaged with different writing styles, a plethora of ideas, and varying perspectives of the same challenge.

Andrew's exploration to understand and constructively deal with the horrors that climate change brings is based in her previous antiapartheid activism and personal battle with a debilitating illness. The personal anecdotes are highly emotive, conveying the meaning of learning about the mess that humanity has created of Earth and what can and should be done to change. *Fire Dogs* becomes an impressive collection of inspiring examples on action and on technology. That clever separation ensures that the reader does not fall into the common trap of assuming that technology can and will be a saviour. Much more is needed. Andrew's educational examples, including a "Educator's Guide," further reinforce the need for education, behavioural change, and attitude change, rather than investing all expectations in technology.

While some might criticize the heavy focus on referencing Web sites, the focus appeals to those who most need to change to deal with climate change and other sustainability concerns—those who are affluent enough to have Internet access. Consequently, Andrew is astute in admitting and emphasising the root causes of climate change.

It's not simply greenhouse gases, but that "Our Western Capitalist approach to development has failed. We can't continue with 'growth' and 'development' that is driven by profit, rather than the needs of this planet and the life (including people) that lives on it." (p. 44). We cannot solve problems through similar solutions that created them. For instance, Andrew makes a well-argued plea against nuclear power, providing common sense for a debate often obscured by the factoids and halftruths of the pro-nuclear lobby.

Similarly, Andrew asks pointed, poignant questions that are often sidestepped: "There is no choice about whether to switch over to renewable energy. The question is when: now or later?" (p. 52). Highlighting fundamentals leaves room for direct arguments on solutions that are often trampled in rhetoric, such as Andrew's common sense and accurate description of how sustainable energy solutions create jobs-not harm the economy as is usually claimed without evidence by the fossil fuel lobby.

The major approaches presented by *Fire* Dogs are all given a healthy dose of realism. Examples are the succinct descriptions of the advantages and disadvantages of biofuels and carbon capture and storage. Andrew raises needed questions about today's use of wind energy and then explains the importance of alternative approaches involving wind energy. As such, the book does not paralyze with fear or hopelessness. It grasps the problem, personalizes it, and pragmatically explains what can and cannot—as well as what should and should not-be done.

Some minor inconsistencies emerge, such as on page 74, where Andrew indicates support for the Sea Shepherd Conservation Society, which is "opposed in principle to all whaling by any people". On the same page, she implies that she has solidarity with indigenous lifestyles and peoples. Many indigenous groups prefer sustainable whaling to flying in more tinned food. Nonetheless, Andrew writes and acts from the heart, while remaining refreshingly hyperbole-free.

An example is Andrew mentioning but not emulating Bill McKibben's campaign, which favours style over substance. McKibben effectively uses social media to mobilize crowds, particularly youth, to advocate for 350 ppm or less of CO₂ in the atmosphere. He has had an impressive impact and is worthy of notice. Yet McKibben does not address deeper issues beyond the superficial number. Examples are: (i) how his goal could be achieved without causing more harm and (ii) whether ppm CO_2 in the atmosphere is the best metric-or, more philosophically, whether a quantitative metric is an appropriate target.

In contrast, Andrew displays both style and substance. One chapter asks "How Can <u>You</u> Make a Difference?" and then provides practical answers. *Fire Dogs* thus achieves its aim of "An Inspirational Call to Action" but goes further. A call to action, yes, but also providing what you need to make your own blueprint for acting.

**The Fire Dogs of Climate Change: An Inspirational Call to Action* by Sally Andrew is available from Findhorn Press, 2009. For more information, please visit http://sallyandrew. findhornpress.com.

**Ilan Kelman is a senior research fellow at the Center for International Climate and Environmental Research - Oslo (CICERO) in Norway. To contact Ilan, please visit http://www. ilankelman.org/contact.html.



Contribute to WSW

See an article you'd like to respond to? Want to share your views on a societal impacts topic and request feedback? Have program highlights or research updates to share? *Weather and Society Watch* is continually accepting a wide variety of articles, photographs, and book reviews. Not sure if you should contribute something? Have more questions? Please contact Emily Laidlaw at laidlaw@ucar.edu.

Conferences & Opportunities

2010 Summer Colloquium: Forecast Verification in the Atmospheric Sciences and Beyond

Host: NCAR Advanced Study Program (ASP) Date: June 6–18, 2010 Location: Boulder, Colorado For More Information: Please visit: http://www.asp.ucar.edu/colloquium/summer_colloquia.php.

This colloquium is designed for graduate students who have completed at least one year of studies in meteorology and climatology or computer science, economics or statistics with an interest in weather and climate. The Colloquium will address the many facets of they dynamic and rapidly evolving field of forecast verification and evaluation. ASP will fund travel and local expenses for about 25 student participants. Invited lecturers will discuss verification from the perspective of the meteorologist and climatologist, the statistician, the scientist and the decision maker. Speakers from the fields of meteorology, statistics, economics and health sciences will share a wide range of perspectives. For more information, please visit http://www.asp.ucar.edu/colloquium/summer_colloquia.php.

Invitation to AMS Joint Session on Environmental Security: National Security Implications of Global Climate Change

Date: January 19, 2010 (During AMS Annual Meeting in Atlanta, Ga.) Location: Atlanta, Georgia

For More Information: Please visit http://www.sip.ucar.edu/news/news.jsp.

The American Meteorological Society (AMS) is hosting a two-part joint session on "Environmental Security: National Security Implications of Global Climate Change", to be held in conjunction with the 90th AMS Annual Meeting at the Georgia World Congress Center in Atlanta, Ga. These special sessions will be held on Tuesday, January 19, 2010, 11:00 a.m.-noon, and 1:30-3:00 p.m., respectively, as part of the Fifth Symposium on Policy and Socio-Economic Research, and jointly with the Second AMS Conference on International Cooperation in the Earth System Sciences and Services and the First Environment and Health Symposium. For more information, please visit http://www.sip.ucar.edu/news/news.jsp.

End Users (continued from page 7)

example, many of us assume that most organizations today have good network bandwidth. Well, that is not always a good assumption. Many users of weather information spend a significant amount of time in the field where access to computer networks is limited. In this case, information may need to be refined so that it can be provided remotely via cell phone, PDA, or paging system. I have seen some excellent capabilities fall short when end users lose patience waiting for information updates to arrive over slow networks. Make sure the proposed solution matches the information technology capabilities of the host organization.

Technology Champion: When considering introducing new technology into an organization it's important to understand who will champion the technology. If the technology is being promoted by management, then one can expect some push back or skepticism from staff. There can be significant challenges if you are brought into an organization without a significant amount of support from the real users of the technology. More time will be required to develop relationships and get buyin from the staff if staff believes the new system is being pushed on them. A strong recommendation in this scenario is to ask the management representative to identify a few staff members that may be receptive to the new technology and have them work closely with the providers before the system is introduced to the broader group. If the actual users of the system are the champions,

then you can rely on them to make the case for introducing it into the workplace. Management tends to be receptive to new capabilities if they perceive that it will lead to enhanced productivity. In both scenarios, it is critical to work closely with the real end users through the user needs assessment and system prototyping processes.

Application Categories: When scoping out a DSS for a particular end user or organization, it is important to understand its envisioned application category. Will it be used for strategic planning (condition prediction), tactical planning (alert functionality), operations management (productivity), incident management (problem notification), risk management, and/or as a real-time or off-line training tool? Each of these uses are somewhat unique (although there is overlap) and will dictate, to some degree, the amount and accuracy that will be required. For example, a DSS that focuses on tactical alerts will typically require more accuracy than products that provide more general predictive information necessary to support longer term strategic planning. A thorough discussion with the end users is required to provide a clear understanding of the system's uses, and the information should be used to design the system and set performance requirements or goals.

Setting Appropriate Expectations: This is a big deal! It is absolutely critical that appropriate expectations be set early in the user needs exploration process. During the process of exploring

the art-of-the-possible it is fine to think outside the box, but the discussion should not drift to the point where the end user thinks that by working with you they will soon have access to the perfect weather information system. I have seen end users get really excited about new capabilities only to be disappointed later when the new weather products did not perform as they had anticipated. It is better to be cautiously optimistic about performance than overshoot and have end users focus on the lack of capability when they actually have significantly improved their capability over the baseline.

So, in summary, the best way to engage end users is to be a good listener, patiently probe their decision space to glean information about their daily process, understand their organizational culture, set appropriate expectations, and never oversell a system's capability. After all, Mother Nature has a way of getting in the way when we least expect it.

*Bill Mahoney (mahoney@ucar. edu) is the director of the Weather Systems & Assessment Program in the Research Applications Laboratory (RAL) at the National Center for Atmospheric Research (NCAR).

References

Lazo, J.K., R. E. Morss, and J. L. Demuth, 2009: 300 Billion Served: Sources, Perceptions, Uses, and Values of Weather Forecasts. *Bulletin of the American Meteorological Society*.**90**, 785– 798.

Flash flood (continued from page 2)

min) extrapolations are made using the NCAR Thunderstorm Identification, Analysis and Tracking (TITAN) System to provide rapid nowcasts of ongoing convective activity. Additionally, OPFs out to 36-hour lead times from the NCAR WRF model are used to produce operational NWP forecasts of rainfall that are then fed into the Noah-distributed hydrological prediction system.

The suite of meteorological and hydrological products, collectively referred to as the NCAR Flash Flood Prediction System (FFPS), has been run in real-time during the response. Plans are underway to summers of 2008 and 2009 as part of a hydrometeorological forecast demonstration project. The summer of 2008 turned out to be anomalously dry in the Colorado Front Range region, resulting in one of the driest summers on record through early August. Conversely, the early summer of 2009 was one of the wettest summers on record, although much of the severe thunderstorm activity that did occur moved rapidly across the landscape and didn't pose significant flash flood hazards. ucar.edu) is a scientist for the

Nevertheless, flash flood events of different magnitudes and return periods did occur during both summers along the Colorado Front Range, which afforded the opportunity to both evaluate and improve the forecast system. One such event occurred on August 8, 2008, and the attached figure shows the accumulated rainfall occurring over a 6-hour period on the evening of August 8. The main event occurred over a small region in the southern part of the

Denver metropolitan area within the Cherry Creek and Harvard Gulch watersheds. Both observed and model-simulated hydrological responses from this storm event are shown in the flood hydrographs where the multiple model runs (dark blue, light blue, and green lines) indicate sensitivity in modeled runoff when the model's infiltration scaling parameter is varied over a plausible range of values. As with most hydrological models, runoff response is sensitive to the representation of infiltration. but the ensemble of simulations does bracket the observed operate the FFPS again during the primary convective season of 2010, in July and August, and additional retrospective flood events are also being evaluated.

*David J. Gochis (gochis@) ucar.edu) is a scientist for the Hydrometeorology Applications Program (HAP) within the **Research Applications Laboratory** (RAL) at NCAR.

**David N. Yates (yates@) Hydrometeorology Applications Program within the Research Applications Laboratory (RAL) at NCAR.

***Wei Yu (weiyu@ucar.edu) is a software engineer for the Hydrometeorology Applications Program within the Research Applications Laboratory (RAL) at NCAR.

Jobs & **Opportunities**

Post-doctoral Research Position: Socioeconomics of Weather and Communication. Use. and Value of Weather Forecasts, National Center for Atmospheric Research

The National Center for Atmospheric Research (NCAR) is seeking a Postgraduate (post-doctoral) Scientist to conduct research on the socioeconomic aspects of weather and the communication, use, interpretation, and value of weather forecasts and warnings. The scientist will join a multidisciplinary team of researchers in NCAR's Mesoscale and Microscale Division, Integrated Science Program, and Collaborative Program on the Societal Impacts and Economic Benefits of Weather Information.

The position will include work on projects related to everyday weather, hazardous weather such as floods and hurricanes, and weather forecast uncertainty information, applying knowledge and methodologies from the scientist's field of expertise. As a member of a growing interdisciplinary program, the postgraduate scientist will also have a unique opportunity to influence an expanding area of research at the interface of science and society.

To apply, please submit a curriculum vitae, 1-2 page cover letter and/or statement of interest, and abstract of Ph.D. dissertation or recent publication at: https://hostedjobs. openhire.com/epostings/submit. cfm?version=1&company id=15947. (click on "Scientific" under "Current Job Openings", then on job #10037).

Initial consideration will be given to applications received prior to February 7, 2010. For more information, contact Dr. Rebecca Morss (morss@ucar. edu).

Evaluation (continued from page 1) Unfortunately, in the 1970s and 1980s many of the scores that were commonly used were the same scores that had been developed in the 1890s, soon after the Finley episode - that is, almost all of the scores in common usage had been defined almost 100 years before. Although new methods had been developed to handle new types of forecasts (e.g., the Brier score for probability forecasts, the anomaly correlation for model fields), typical verification processes that were in use involved the computation of relatively simple metrics: most verification was done for administrative purposes (e.g., to track overall performance over time) and relied on one or two basic scores (e.g., Root-Mean-Squared Error – RMSE; Critical Success Index – CSI)). Individual errors or distributions of errors were rarely considered, and sampling variability and biases that could impact interpretation of the verification results were consistently ignored. Verification was essentially regarded as a necessary but relatively uninteresting aspect of the forecasting process. Measures were selected and applied uniformly, essentially without consideration of the forecast user. Thus, these approaches typically did not provide comprehensive forecast evaluations.

Things began to change as Allan Murphy and others became interested in decompositions of scores that revealed what the scores could *really* tell us about things like calibration, resolution, and correlation; and the new verification researchers investigated new ways of interpreting

verification measures from a statistical perspective, leading to diagnostic evaluations of forecasts, and informative displays (e.g., "conditional quantile" plots) that began to address the uncertainty associated with verification information. One very important message from Murphy's work is the concept that the quality of a forecasting system cannot be summarized using a single number. In fact, a variety of metrics are required to represent the variety of attributes of forecast quality. These developments set the stage for user-relevant verification approaches that have been developed over the last decade, and which are beginning to be applied in operational as well as forecast development settings.

In considering the need for userrelevant verification approaches, one only has to understand that different users may have very different ideas regarding the kinds of errors that are important. For a farmer or reservoir manager, the timing and quantity of precipitation are important. For an airline strategic flight planner, the northsouth location of convective storms in the central U.S. is important, but the east-west location may be of less interest. For a utility company, large errors in temperature forecasts, or the timing of large wind speed increases may be important. The variety of weather forecast users ranges from forecast developers to operational forecasting offices to sophisticated users, and could include the general public if appropriate information were available.

To meet the needs of this large

mix of potential users of forecast evaluation information requires a broad variety of scores, summary measures, and approaches for examining errors. An example of a relatively simple new approach that can help meet these requirements is the use of box plots and other graphical methods to examine distributions of errors. Distribution approaches allow users to understand the frequencies of different types of errors - not just the mean error, which is not likely to provide a satisfactory level of information to actually meet the needs of most forecast users or developers (see figure on p. 13 for example).

Spatial verification approaches represent another major development over the last decade that can provide user-specific as well as overall performance information. Spatial methods were developed partially in response to the fact that traditional (gridto-grid comparison) methods could not measure the apparent improvement in forecasts associated with higher-resolution forecasts, and partly to provide more meaningful information about forecast performance. For example, some of these methods can provide information about forecast performance as a function of spatial scale, or measure the spatial displacement between forecasted and observed storm systems. Because these methods are relatively new, their capabilities have not yet been fully explored or exploited for forecast evaluation and improvement, especially in operational settings. Other efforts have also focused on the development of methods to

evaluate forecasts of extreme or rare events. These methods are of particular importance as populations become more vulnerable to extreme events (e.g., flooding, winds) and the warning capabilities of weather services become a more critical component of emergency services.

A final important development is the increased availability and use of methods to evaluate the statistical uncertainty in verification measures. Application of these approaches is critical, for example, for determining whether a model "improvement" has actually led to better forecasts, or if an apparent improvement may just be an artifact of the particular (typically small) sample used for the verification study. Still too commonly, researchers and model developers point to minor changes in verification statistics as indicative of an improvement in performance, when closer examination of the statistical sample would not support that conclusion. Use of statistical hypothesis tests and confidence intervals can help avoid these kinds of errors and lead to much stronger conclusions about improved performance.

To go to the next step – for verification information to truly lead to improved use of forecasts – will require a greater application of the new methods and approaches in operational as well as research settings. In addition, it will require collaborations between meteorologists, verification experts, social scientists and users to develop forecast evaluation methods that truly represent the needs and interests of specific users (Morss et al. 2008). Among other things, the involvement of social scientists would help identify the types of verification information that are relevant for particular user groups, as well as appropriate communication approaches for specific forecast users and decision makers.

In conclusion, the science of verification has come a long way over the last two decades. As new methods and approaches are gradually integrated into forecasting and forecast development processes, new ideas and approaches will continue to be developed and assimilated and will further improve verification, forecasting and decision-making processes. Fortunately, the verification community is vibrant and has grown significantly over the last 10 years, and it is ready to respond to new needs and ideas.[1] It is likely (and my hope) that the next 10 years will be characterized by further interactions between verification scientists, users, and social scientists, which will enhance the development of userrelevant verification concepts and methods.

*Barbara Brown (bgb@ucar.edu) is director of the Joint Numerical Testbed program in the Research Applications Laboratory at the National Center for Atmospheric Research. She also is a member and past chair of WMO's Joint Working Group on Forecast Verification.

References

Finley, J. P., 1884: Tornado predictions. *Amer. Meteor. J.*, **1**, 85–88.

Morss, R.E., J.K. Lazo, B.G. Brown, H.E. Brooks, P.T. Ganderton, and B.N. Mills, 2008: Societal and economic research and applications for weather forecasts: Priorities for the North American THORPEX program. *Bulletin of the American Meteorological Society*, **89**, 335-346.

Murphy, A.H., 1996: The Finley Affair: A signal event in the history of forecast verification. *Weather and Forecasting*, **11**, 3-20.

Footnote

[1] For more information see <u>http://www.</u> cawcr.gov.au/projects/verification/ and <u>http://www.wmo.int/pages/prog/arep/</u> wwrp/new/Forecast_Verification.html.





Livestock (continued from page 3)

million loss to producers (Dietz et al. 2003). It was obvious that mitigating this loss would be in keeping with the NWS's mission "for the protection of life and property and the enhancement of the national economy."

With this information, Dr. Kalkstein brought in Dr. Katrina Frank, who has a background in animal science and bioclimatology, to assist with the project. This collaboration was essential to the successful completion of the project because Dr. Frank's background in livestock production, specifically biothermal responses, gave us an understanding of the biology and production systems of animals that was lacking among the previous team members. As I got to know Katrina better through our monthly conference calls, I encouraged her to apply for the annual summer WAS*IS workshop, both because of her potential contributions to the workshop and to help bring more WAS*IS collaboration into the project. More stars were aligning as she was selected to attend the workshop in 2007 and better understand the connection from research to operations to end users.

For this project to accomplish the goal of developing a warning system for livestock that would be utilized by producers in the area, it was necessary that operational products and their impacts be connected with the needs of the products' users. It was useless for us, as product developers, to create a product without first consulting the end users to ensure that the proposed product would meet their needs. In order to determine what those needs were, we incorporated some of the methods learned in WAS*IS about qualitative and quantitative survey methods and asked a few of the livestock producers in northeastern Montana what they needed from us. Only once we had gathered information about what product(s) would be useful to the stakeholders could we proceed with development of an operational advisory system.

We received input from ranchers in northeastern Montana who were weather spotters, cooperative weather observers, or known by the NWS to have a significant interest in weather. One rancher even sat down with two of us to discuss our questions in detail. He provided further explanations on cattle, cattle production, and actions taken during inclement weather. He was very pleased that we were undertaking a project that he felt would benefit his ranching operations.

Overall, the producers identified newborn calves as the animals of most concern to them in harsh weather conditions. They are "concerned with calf losses during calving season" (producer with 500 head of cattle, Valley Co., Mont.) because "calves are our saleable product, so no calves, no sales, no income" (producer with 300 head of cattle, Prairie Co., Mont.).

An unintended, but quite fortuitous, result of reaching out to the ranching community was that we were able to strengthen our relationships with cattle producers who were happy to see us showing an interest in their livelihoods and working on a system that would help them. By incorporating the ranchers' input from the start, we knew that we were working on something that would be economically beneficial to an end user.

We developed a preliminary database of several weather events that resulted in calf losses based on the information provided by the ranchers. We added to the database by reviewing events entered into the NWS Storm Data database that mentioned livestock losses. This brought us a total of eight significant events to review. We then looked further into the climate data during those events, exploring variables such as:

- maximum and minimum temperatures both during and prior to each event
- average winds and gusts
- measured snowfall/ precipitation
- type of airmass
- length of the event

Having these data allowed us to see the range of weather events that had caused losses, or were deemed "significant" by the producers.

Given that our database did not yield enough events that had resulted in calf losses to develop a statistically significant weather/ mortality relationship, we had to incorporate the producers' and Dr. Frank's knowledge of how calves respond to cold and do a literature review to generate a decision tree for advisory issuance. Through this process, we identified a 'newborn' calf as one that was less than 24-hours old because these calves are least able to regulate their body temperature (Sanko et al. 1991).

In keeping with the spirit of WAS*IS to involve the end users as much as possible, Dr. Frank and I organized a user's workshop the summer before the system was to debut, and she was able to fly to Montana to attend and to meet some of the key users who were providing us feedback. The goal was to show the livestock community what we had developed and to get their feedback on the draft decision tree before it was implemented. We also got feedback on what they expected as a final outcome, how they wanted to access the data, what the final format should look like and how often they needed the information. This workshop was held twice, once during the afternoon and a second time during the morning, to allow as many attendees to come as possible. We were able to tweak the system based on their feedback and develop the documentation needed for NWS to run the system experimentally.

One of the most difficult tasks we encountered was finding a name for the system we had developed. The team eventually decided on Cold Advisory for Newborn Livestock (CANL). Some had suggested the Livestock Advisory, but that was very similar to an outdated product that NWS had issued many years ago. It also implied that a watch, warning or advisory would be issued, and that was not the goal of this system. The intention was to aid producers with a decision support system to prepare for hazardous weather. The CANL system was run in an experimental mode from February through May 2009 in northeastern Montana. The graphics were displayed on the NWS Glasgow website, where there was a link that allowed people to provide feedback on the system. We received just eight responses through the online feedback, but we did have ranchers who called us and spoke to us in person. One of the first calls was from an elderly man who had dialup Internet access. He didn't really understand some of the data and functionality of the web site (e.g., we used "RH" instead of "Relative Humidity"; the initial images were too small, but clicking on them enlarged them), and based on my 30-minute discussion with him we were able to go back and make some great improvements to the webpage. Pre-testing the webpage would have been a good idea, one that we won't forget in the future!

We advertised the system in a variety of ways including agricultural newsletters, NWS newsletters, on the NWS Glasgow homepage, through the local media, and by mailing information to known livestock producers in northeastern Montana. Getting the word out to the public is ongoing and a very important aspect of continuing the CANL system.

In the fall of 2009, we held another users' workshop to get feedback after the system had run for a season. Drs. Frank and Kalkstein were able to attend via teleconference. At this workshop, we showed the producers some of the events and forecasts from the previous winter. After reviewing many of the events and non-events, we collected the feedback and modified the decision tree and criteria slightly.

From the onset, we felt that a huge challenge would be expanding the system to other areas of the country. However, during the winter of 2008-2009, the Dakotas and southeastern Montana were hit with many significant winter storms. There were headlines almost daily tallying livestock losses, with economic losses reaching millions of dollars. The NWS offices in Aberdeen, S.D., Billings, Mont., Bismarck, N.D., and Great Falls, Mont., all joined in to be part of the CANL system experimental test period starting January 18, 2010, and running through May 31, 2010.

While in the end, we aren't saving the world as I had hoped after Hurricane Katrina, we are making a difference to the people who are using the CANL system and those who hopefully will use the system in the near future. A chance discussion with a rancher, an incredible workshop, new collaborations, partnerships, friendships, and a lot of hard work helped align those stars to be bright enough to make a difference in our corner of the world. The elderly rancher who had dial-up ended his call with, "I'll be using this day and night during calving season. Thank you!"

*Tanja Fransen (Tanja.Fransen@noaa. gov) is the Warning Coordination Meteorologist for the National Weather Service office in Glasgow, Montana. She is also a proud participant in and champion of the Weather and Society * Integrated Studies (WAS*IS) workshop series. (continued on page 16)

Livestock (continued from page 15)

**Dr. Katrina Frank (kfrank@live. com) is currently in Enterprise, Alabama, and working with Dr . Larry Kalkstein on various research projects for the University of Miami's Synoptic Climatology Laboratory. She is also a strong supporter and participant from the Weather and Society * Integrated Studies Workshop.

Acknowledgements

This project was inspired by Lynn Cornwell (1951-2008) and set in motion by Julie Adolphson and Dr. Larry Kalkstein through a grant from the COMET Partnership Project, (S07-62730). Those who assisted with the project include, Tom Salem, Corey Bogel, Don Simonsen and Bill Martin (NWS Glasgow) and Scott Sheridan (Kent State).

References

Azzam, S. M., J. E. Kinder, M. K. Nielsen, L. A. Werth, K. E. Gregory, L. V. Cundiff and R. M. Koch, 1993: Environmental Effects on Neonatal Mortality of Beef Calves. Journal of Animal Science. 71:282-290.

Dietz, R. E., J. B. Hall, W. D. Whittier, F. Elvinger, and D. E. Eversole, 2003: Effects of feeding supplemental fat to beef cows on cold tolerance in newborn calves. Journal of Animal Science. 81:885-894.

Sanko, R. L., M. J. Guthrie, and R. D. Randel, 1991: Response to environmental temperatures in Brahman calves during the first compared to the second day after birth. Journal of Animal Science. 69:4419-4427.

United States Census Bureau. 2007. National and State Population Estimates. [Accessed 18 March, 2008, from http://www.census.gov/ popest/states/NST-ann-est.html].

United States Department of Agriculture – National Agricultural Statistics Service. 2008. Montana Data – Cattle & Calves, Cattle Inventory – January 1. [Accessed 18 March, 2008, from http://www.nass.usda.gov/ QuickStats/PullData_US.jsp].

United States Department of Agriculture – National Agricultural Statistics Service. 2006. 2005 Montana and United States Cattle Predator Losses. [Accessed 17 March, 2008, from http://www.nass.usda.gov/Statistics_by_ State/Montana/Publications/Press_Releases_ Livestock/catprdls.htm].

About Weather and Society Watch

Weather and Society Watch is published quarterly, in January, April, July and October, by the Societal Impacts Program (SIP) at the National Center for Atmospheric Research (NCAR). The University Corporation for Atmospheric Research (UCAR) operates NCAR with support from the National Science Foundation and other sponsors.

The purpose of *Weather and Society Watch* is to provide a forum for those interested in the societal impacts of weather and weather forecasting to discuss and debate relevant issues, ask questions, and stimulate perspective. The newsletter is intended to serve as a vehicle for building a stronger, more informed societal impacts community.

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of NSF or other sponsors. Contributions to *Weather and Society Watch* are subject to technical editing at the discretion of SIP staff.

Weather and Society Watch is available on the World Wide Web at: http:// www.sip.ucar.edu/news/. Archives of WeatherZine, a previous weather impacts newsletter upon which Weather and Society Watch was modeled, are available on the Web at http://sciencepolicy.colorado.edu/zine/archives/.

About SIP

SIP, initiated in 2004 and funded by NOAA's U.S. Weather Research Program (USWRP) and NCAR, aims to improve the societal gains from weather forecasting. SIP researchers work to infuse social science and economic research, methods and capabilities into the planning, execution and analysis of weather information, applications, and research directions. SIP serves as a focal point for developing and supporting a closer relationship between researchers, operational forecasters, relevant end users, and social scientists concerned with the impacts of weather and weather information on society. Program activities include primary research, outreach and education, and development and support for the weather impacts community.

Contact Us

For additional information or to submit ideas for a news item, please contact:

SIP Director: Jeff Lazo (lazo@ucar.edu) Managing Editor: Emily Laidlaw (laidlaw@ucar.edu) Contributing Copy Editor: Dale Laidlaw

To send mail about Weather and Society Watch, please write to:

Jeff Lazo Societal Impacts Program National Center for Atmospheric Research P.O. Box 3000 Boulder, CO 80307

