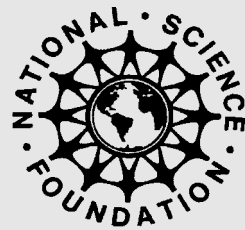


**Citizen Science
Toolkit Conference**

June 20 - 23, 2007

yeah, but...are the data any good?

Sandra Henderson
UCAR Citizen Science Programs Manager



CORNELL LAB OF ORNITHOLOGY

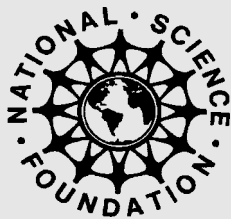
607.254.BIRD telephone
www.birds.cornell.edu

159 Sapsucker Woods Road
Ithaca, New York 14850

This presentation took place at the Citizen Science Toolkit Conference at the Cornell Lab of Ornithology in Ithaca, New York on June 20-23, 2007.

Note that this document did not originate as a formal paper. Rather, it combines an oral presentation with accompanying PowerPoint slides and reflects the more informal, idiosyncratic nature of a delivery prepared specifically for this live event.

Documentation of the conference is meant to serve as a resource for those who attended and for others in the field. It does not necessarily reflect the views of the Cornell Lab of Ornithology or individual symposium participants.



This documentation is supported by the **National Science Foundation** under Grant ESI-0610363.

Any opinions, findings, and conclusions or recommendations expressed in this documentation are those of the authors and do not necessarily reflect the views of the National Science Foundation.

The following is one of three focus point presentations delivered as part of the session titled "Citizen Science Project Design" on day one of the Citizen Science Toolkit Conference

For complete documentation of conference proceedings and to learn more about citizen science and the Citizen Science Toolkit, or to join the ongoing citizen science community, go to:

<http://www.citizenscience.org>

Yeah, but...are the data any good?

Sandra Henderson,
UCAR Citizen Science
Program Manager

Overview

After listening last night and today to the introductions, the talks, and the discussions, I would bet that I'm not the only one in this room who has ever been asked a question of this nature: "Yeah, yeah, yeah, sounds good, but...is the data any good?" Has anyone here ever had to deal with that? How do you answer it?

Personally, I have decided not to spend a tremendous amount of my time trying to convince the nay-sayers or convert the doubting Thomases to these types of programs. The data is very useful to answer the scientific questions we're posing, but also to inform us as citizen science project administrators, managers, and leaders to develop better projects. So I think, yes, the data is very good.

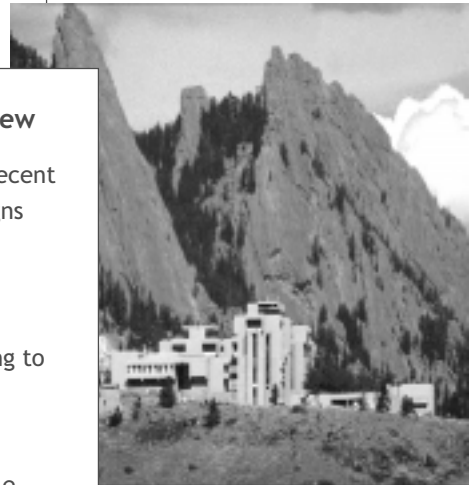
Here is a brief presentation overview. I want to talk about two of our recent citizen science campaigns, GLOBE at Night and Project BudBurst, review some of the concerns we found relating to data quality, give you a lesson learned on data verification that we've put into play this year, and then talk about an upcoming event.

Presentation Overview

- Brief overview of two recent citizen science campaigns
 - GLOBE at Night
 - Project BudBurst
- Review concerns relating to data quality
- Lessons Learned - Data verification - an example
- Coming this fall - Great World Wide Star Count

UCAR
Education and Outreach
Serving UCAR, NCAR, and UOP

Yeah:
How do you
know the data collected
by professionals is any
good?
- Rick Bonney



UCAR and NCAR Citizen Science Campaigns

About Citizen Science at UCAR and NCAR

At UCAR and NCAR, the goal of our citizen science efforts might be a little different than those of a lot of the projects we've heard about so far, and different than a lot of the monitoring and volunteer network campaigns. Our primary goal is to increase public awareness of an environmental issue through data collection. We want to increase public awareness of whatever issue it is we are looking at.

Our citizen science projects are very extensive. We are looking at national or even global levels of participation, and they are open to all. There is no prior training, we never meet the people, and we don't have train-the-trainer opportunities. We're looking at how can we do this using existing networks in an entirely Internet-based campaign.

UCAR/NCAR Citizen Science Campaigns

- Primary goal is to increase public awareness of an environmental issue through data collection
- Extensive - national or global participation
- Open to all - prior training or registration not required
- Entirely Internet-based
- Collaborative - draw upon expertise of diverse institutions

Get Out and Observe
the Night Sky!
March 22 - 31, 2006



Engage students worldwide in observing the nighttime sky

Encourage citizen and family science with a hands-on learning activity outside of the classroom

Gather light pollution data from an international perspective

They are also collaborative. We draw upon the expertise of many. We draw upon our expertise as a collaborative university community, a national center in Boulder, and on existing networks that we have. With GLOBE at Night, for example, which is a light pollution astronomy campaign, we also partner with institutions such as the National Astronomy Observatory in Tuscon. In a project like Project BudBurst, we collaborate with the Chicago Botanic Gardens as well as botanists and phenologists at the University of Montana, U.C. Berkeley, and the University of Arizona. We are very much looking at the collaborative aspects of what the community in general has that we can use.

GLOBE at Night

I want to talk to you about our very first foray into this idea of taking on the challenge: Let's do something that will involve the entire world in a citizen science campaign. We started this on January 3, 2006, and when I say "started," I don't mean opening up the Web site or getting the word out, I mean sitting down and planning for an event that went live on March 22. This was about eleven weeks, and in that eleven weeks we developed the protocol, we developed the Web site, we developed the downloadable materials, the marketing, bringing in partners, all aspects. And we did this on a shoestring, there was no budget to do this. We just wanted to see if we could have this kind of outreach and impact through this Internet environment.

Our goals are very straightforward and we have them listed everywhere. In one of the breakout sessions this morning we used the term, "being transparent." This is what we hope to accomplish: First, engage students worldwide in observing the nighttime sky. This is a light pollution study. We wanted to encourage citizen and family science outside of a formal environment. And finally, we wanted to gather light pollution data from an international perspective. It's not an accident that they are listed in that order, looking at the data being actually the last thing.

GLOBE at Night Web Site

- Simple Protocol - observe the constellation Orion and determine the limiting magnitude
- Star Magnitude Charts (by latitude bands)
- Activity packets (English and Spanish)
 - Teachers
 - Parents
- Interactives and games
- Data Entry
- Mapping (ESRI provided GIS)

If you went to the GLOBE at Night Web site, you would find everything you need to participate. There is a very simple protocol. You simply go out between seven and nine during the time period, which was March 22 to 31 for the 2006

event. We have limiting magnitude charts that you download, and you go out and look at Orion. Orion was chosen because at that time of the year it is visible in almost all latitude bands in both hemispheres. There are also downloadable activity packets, and I'd like to take an aside for just a moment.

The packets were in both English and Spanish. As a national center, we are very committed to providing educational materials and opportunities that will reach the changing demographics of the American population. We wanted children to be involved, and we knew that there would be young kids, children who would be out with their parents. We wanted to reach out to the Spanish-speaking population by having these materials available in Spanish. Univision, the Hispanic television network, picked this up and was able to promote it to the Denver affiliate because of the Spanish materials. I encourage you as we develop citizen science programs to be thinking how you can reach all Americans.

We also have interactive games on the Web site to further the interactive content, simple data entry, and ESRI worked with us on providing the mapping.

Then we had to get the word out and as I said, we hit the ground running. We went through our press releases, we did listserves and mass mailings, Web sites, print media—you name it. We were knocking on doors, encouraging everyone to help us get the word out. We were very excited about it. Did it work? Did we get the word out?

Yes! We were picked up by the *New York Times*, we were the subject of a *Scientific American* podcast, and had coverage from numerous print and media outlets. We felt that we really had been able to get the word out.

NY Times Science Observatory
Scientific American Podcast
Earth Science Picture of the Day
Geography Network front page
Many local papers
Local TV broadcasts

- ### Getting the Word Out
- UCAR and NOAO press releases
 - Listserve and mass mailings
 - Web sites
 - Print media
 - Local broadcast media

For the event itself, we had over 18,000 people participate; over 4,500 data points were entered; all fifty states were represented; we

The Event

- Over 18,000 people participated
- From 96 countries on six continents
- From all 50 US states
- 4591 observations reported
- Averaging 9100 Web site hits a day
- Over 760 people on mailing list



The image shows a world map with numerous small black dots representing observation locations. The map is labeled with 'Pacific Ocean', 'Atlantic Ocean', and 'Indian Ocean'. A legend in the bottom right corner of the map image lists 'Limiting Magnitude' categories: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30.

had people from ninety-six countries, all six continents except Antarctica; and we established a mailing list and a community of individuals interested in light pollution.

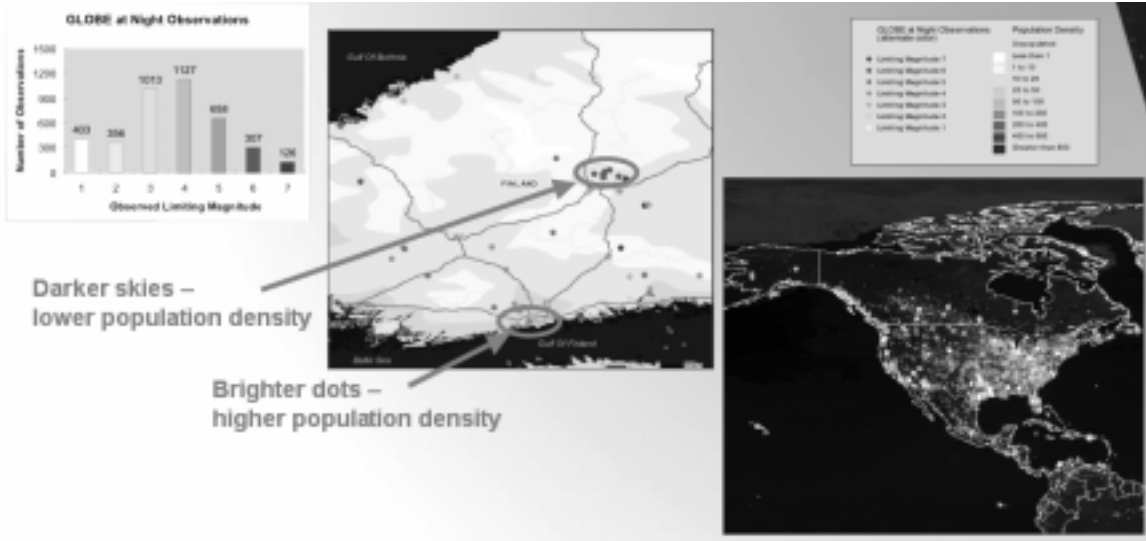
Additional Information from embedded assessments:

- | | |
|--|--|
| <p>Of people reporting data</p> <ul style="list-style-type: none"> • 399 under 12 years old • 949 between 12-14 years old • 871 between 15-18 years old | <p>Teacher feedback</p> <ul style="list-style-type: none"> • 71% Announced in class • 63% Handed out to students (3530 copies) • 24% Recruited science/other school clubs • 31% Shared with other teachers • 17% Contacted local media to publicize |
|--|--|

We had a little bit of embedded feedback information in this campaign so we could learn a little about the people who were participating. If you remember, our first goal was engaging students. Half of the people reporting data were under age eighteen, which told us that this data entry form was pretty straightforward and that kids could participate and enter the data.

What do the GLOBE at Night data tell us? Look at the left graph with a limiting magnitude of one to a limiting magnitude of seven. If you have a limiting magnitude of one, you could be on the strip in Las Vegas or maybe Manhattan. With a limiting magnitude of seven, you could be in a wilderness area in Montana.

What do the GLOBE at Night Data Tell Us?



“The GLOBE at Night data shows brighter skies in areas with more people. By submitting our observations to GLOBE at Night, you made an important contribution to science. You have provided valuable information that would have otherwise been impossible to obtain!”

The curve is what we would expect. When the data was analyzed, it wasn't a surprise. Again, darker skies have lower population densities, brighter dots would show higher population densities.

What were some of the data quality challenges? As I said, we did this entirely Internet-based, so we did not have the opportunity for feedback to see if people understood the protocol. I like the way Karen Oberhauser put it in her presentation: "We thought it was easy and straightforward." We didn't have that opportunity for feedback.

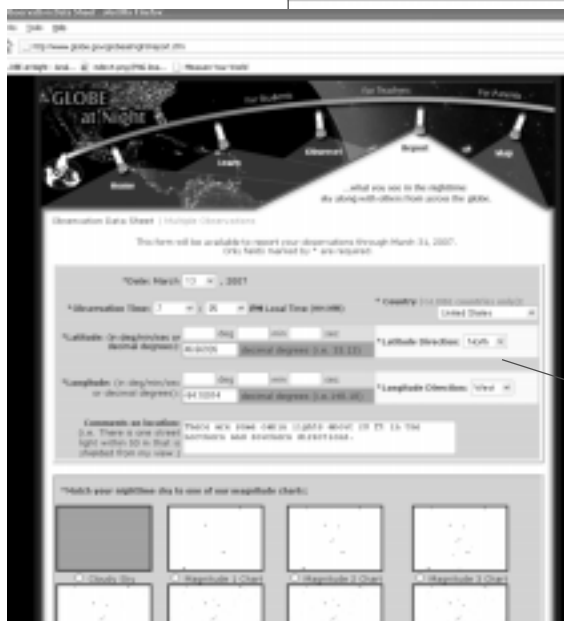
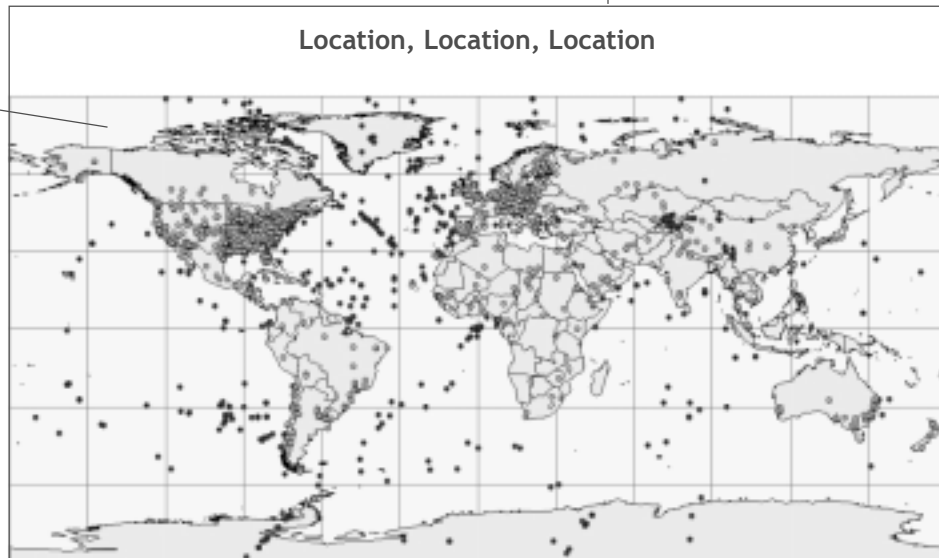
Another major challenge was location, how to get latitude and longitude. This was international and remember, we didn't have people register and there weren't specified study sites. GPS units certainly helped if you had access to them, but there are large parts of the world in which those are just not available. We were able to develop different methods using Maporama and Google Earth to show people how to get the latitude and longitude.

And finally, there was the issue of reporting data. I know you probably have all had experience with this. We know people were out there doing the activity because we get this kind of feedback, but they don't enter the data.

What were some of the data quality challenges we found in the 2006 G@N Internet-based campaign?

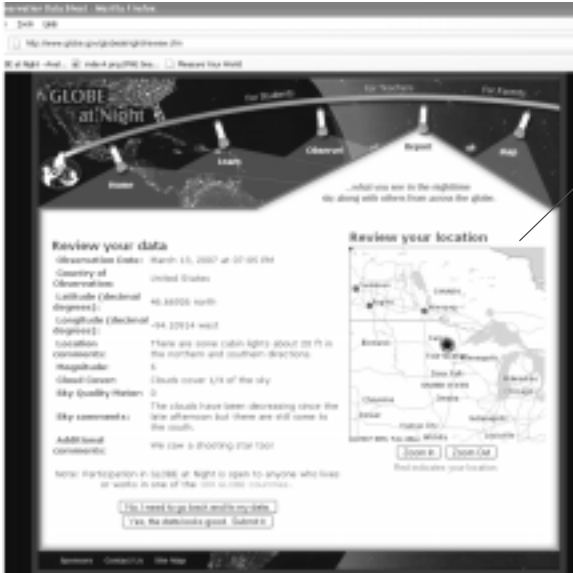
- Training/instruction in the protocol
- Location - how to identify (and verify) longitude and latitude - changes in 2007 event
- Reporting data - we know lots of folks participate, but do not enter/report data

What I'd like to do in the interest of time is focus on location. The gray dots represent data that we were able to use. The black dots represent data that we felt were just erroneous—say out in the middle of the ocean.



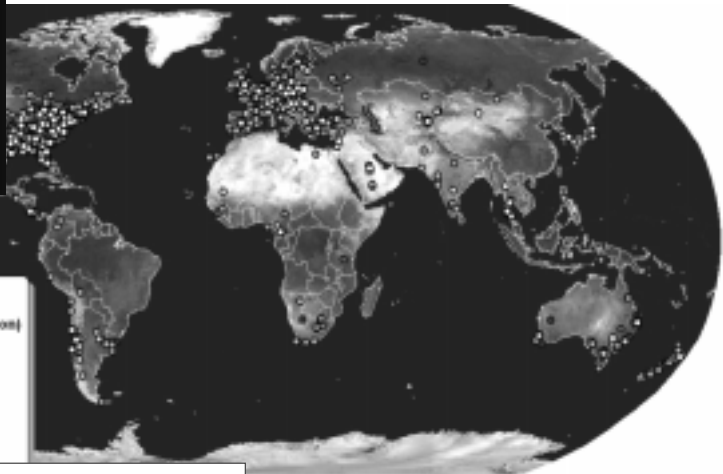
2006 Campaign: 4590 submitted, 3990 accepted, 600 rejected (13.07%)

This data told us we had to make a change in doing the latitude-longitude data entry. This is just a screen capture. You enter your latitude and longitude as well as comments on your location. What you



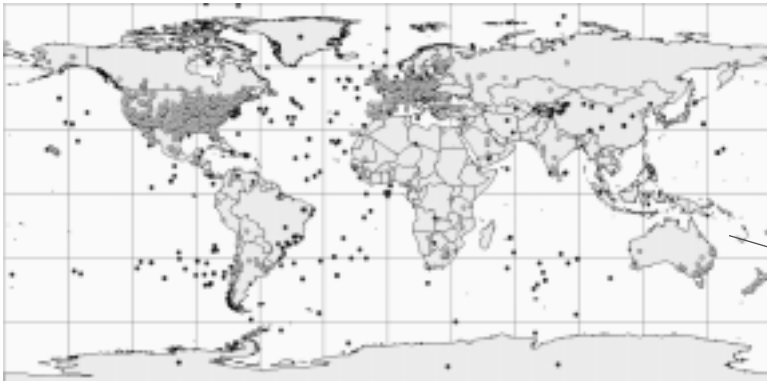
get then immediately is verification. ESRI worked with us on this. If you put in this latitude and longitude data you get a verification screen and can see—yeah, I’m probably pretty close to Fargo—and you get that immediate feedback.

G@N 2007 – 8,491 Observations



- Limiting Magnitude**
- m=1 (most light pollution)
 - m=2
 - m=3
 - m=4
 - m=5
 - m=6

Location, Location, Location



2007 Campaign: 9371 submitted, 8491 accepted, 880 rejected (9.39%)

In the 2007 event, not only did we have increased numbers, we had more data that was actually usable.

The gray again is data we were able to use and the black represents data we weren’t able to use. This time we rejected less than we did the first time in terms of it being useful data.

Changes in Latitude, Changes in Longitude
(with apologies to Jimmy Buffett)

| Maximum Change | Latitude | Longitude |
|----------------|--------------------|---------------------|
| ≥90° | 6 (0.07%) | 297 (3.50%) |
| ≥60–89.9° | 111 (1.31%) | 145 (1.71%) |
| ≥30–59.9° | 241 (2.84%) | 86 (1.01%) |
| ≥1–29.9° | 123 (1.45%) | 108 (1.27%) |
| <1° | 265 (3.12%) | 263 (3.10%) |
| Total | 746 (8.78%) | 899 (10.59%) |

2007 n=8, 491

When we looked at the numbers we were able to identify where and how 1,000 people made changes in longitude and latitude. We felt this was a good thing, and we probably had a lot more data that was usable, but I find that the top one is particularly interesting.

If you have a degree change of greater than or equal to ninety degrees, you’re basically talking a

hemisphere. They were way off. In this case, and in many cases, a negative sign might not have been put in. That was something we felt we were able to learn. We required participants to verify their location, and it seems to have helped their rejection rate a little bit. And eleven percent of the people did choose to make changes based on that.

So, was the G@N data any good?

1. When compared, results are consistent with satellite data
2. 2007 event allowed for better lat/long verification
3. We never intended to topple any paradigms in astronomy

So was the data any good? It was very consistent compared to the satellite data. The 2007 event allowed for this better lat/long verification. And please keep in mind, never in those goals did we expect to topple any paradigms in the field of astronomy. This is not what we were going after. So we weren't surprised that the data showed what we expected.

Using the wonders of PowerPoint, over lunch I was able to throw

another slide in. One thing that Sam Droege said during his presentation that I really think is key in any of these events, and I know other speakers have said this too, is feedback. You have to do something with the data. If you just let it go into the black data hole, that's just wrong. It's just not the right thing to do. Within six weeks after the event, it is always our goal to have a report out. We did do that in a simple and straightforward manner. We sent it out to the listserve, publicized it throughout the Web site, and anyone who participated could go in and see it—even the people who were out floating, by the way. We left their data there, it's just we didn't use it in the analysis of light and dark. They could drill down, using GIS tools, and go in and see their data point and compare it with others. So Sam, thank you for saying that first thing this morning. I hope we hear that all week long: Get the feedback to the folks you want to engage again.



Project BudBurst

I want to get a little more down to earth, get away from astronomy, and talk briefly about Project BudBurst. This was the outcome of a citizen science meeting last fall of the National Phenology Network in Milwaukee. I and a number of other people were part of the citizen science working group. I know

What did we learn?

1. Requiring participants to verify their location seems to have reduced the rejection rate by about 4%, and
2. Given the opportunity, about 11% of participants adjusted or corrected their initial latitude and or longitude.

What Sam said this morning¹

If we want continued participation and data input, as CS program leaders and managers, we MUST provide feedback to participants in a timely manner!

¹ See Sam Droege's presentation in the document, "Just Because You Paid Them Doesn't Mean Their Data Are Any Better."

Project BudBurst

- Outcome of CS working group at the October 2006 National Phenology Network meeting
- Based on the successful G@N model
- Secured small initial funding from BLM through CBG
- Pilot launched spring 2007

most of you have had this experience. You go to meetings and they want you to be in working groups and develop white papers, and so on. We had been talking in the citizen science group and said, "Let's do something. Let's do something national, a simple thing to capture the imagination." We wrote a couple of pages, and decided to go forward with it.

Kay Havens, my colleague from the Chicago Botanic Garden, took this two-page white paper we had written, submitted it to BLM, and got \$20,000 to let us get this little seedling thing going. As a result we launched the pilot in the spring of 2007.



Project BudBurst

A National Phenology Network Field Campaign for Citizen Science

April-June 2007

Join us this spring in collecting important climate change data on the timing of leafing and flowering of trees and flowers in your area through Project BudBurst! This national citizen science field campaign targets native tree and flower species across the country

www.budburst.org

Learn why phenology is important

Participate!

Does climate change affect budburst?

Download free materials

Report your observations online

Map results from around the country

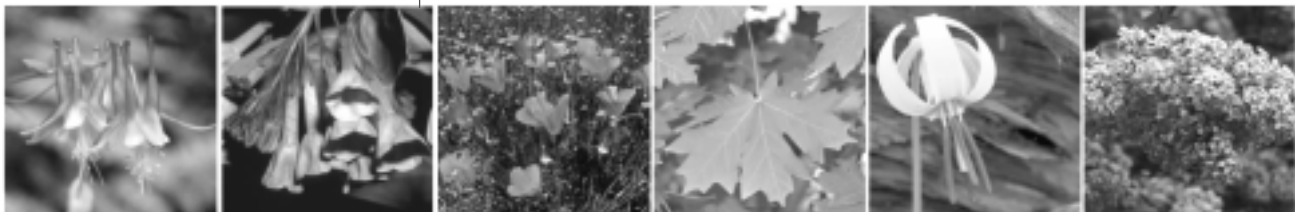
It was a phenology study, and I know that many of you are very familiar with this idea, but the link for us, particularly at NCAR and UCAR where there was a lot of interest in this, was the climate change aspect because that's the basis of so much of our research.

We launched it in 2007 and I know a number of you here have heard about it. We had such positive feedback on this and excitement. We launched it in April of 2007 and it just ended last Friday, so the data hasn't been analyzed; it's still pretty new.

One of the things we did get immediately in April, not surprisingly, was: "Wow, this sounds like

Results of Pilot Campaign

- Ended last Friday (June 15)
- Over 500 data entered; hundreds of e-mails of support and interest
- Numerous organizations, agencies, and institutions to collaborate next year
- 2008 event will be expanded to a full year to capture a more complete pheno picture



a great idea, but I live in South Carolina and everything has long since budded out and leafed out.” So this was truly a pilot, just a proof-of-concept to see if it worked.

We had a wide range of collaborators. When you are trying to do these extensive, large-scale projects, having collaborators and really building on the community is absolutely invaluable.

As I said, it ended last Friday and we did have over 500 data entered and hundreds of e-mails of interest and support. I want to thank Craig Tufts, who is here, for always supporting this and getting the word out for us. We have had numerous organizations and agencies that want to collaborate with us on this, and we are going to expand the 2008 event to start in January, run through December, and be able to capture a lot of phenological events throughout the year.

Great World Side Star Count

Finally, we want to go back to the heavens, and we are stealing this from the birding folks. We are going to develop the Great World Wide Star Count. It starts in October from the first through the fifteenth, so I would invite all of you to get a kid, go outside, and look heavenwards. I bet everyone in this room got their start in their career at some point because you just wanted to be outside and look around. I strongly encourage you to go outside, go onto the Great World Wide Star Count Web site, which is open to everybody, enter data, and let us know what your night skies look like.

“ _____
When you are trying to do these extensive, large-scale projects, having collaborators and really building on the community is absolutely invaluable.
_____ ”



Join thousands of other students, families, and citizen scientists counting stars this fall for the **Great World Wide Star Count!** This international event encourages everyone to go outside, look skyward after dark, count the stars they see in certain constellations, and report what they see online. This Windows After Dark citizen science campaign is designed to raise awareness about light pollution and the night sky as well as encourage learning in astronomy. The Great World Wide Star Count will be held from **October 1st through October 15, 2007.**

For more information visit
[www.windows.ucar.edu/
citizen_science/starcount](http://www.windows.ucar.edu/citizen_science/starcount)
or contact Dennis Ward at dward@ucar.edu