

Computational Research and High Performance Computing at CUNY

Paul Muzio

Director, CUNY High Performance Computing Center

paul.muzio@csi.cuny.edu

(718) 982-3413

- The City University of New York
- Technology Trends
 - Hardware
 - Pervasiveness of computing
 - Need for new algorithms to support emerging applications
- Hurricane Sandy
 - Forecasting, hindcasting, and resilience planning
 - User-facing applications
- Conclusion

- **Senior Colleges**

- (1847) City College
- (1870) Hunter College
- (1919) Baruch College
- (1930) Brooklyn College
- (1937) Queens College
- (1946) New York City College of Technology
- (1955) College of Staten Island
- (1964) John Jay College of Criminal Justice
- (1966) York College
- (1968) Lehman College
- (1970) Medgar Evers College

- **Graduate and Professional Schools**

- (1961) CUNY Graduate Center
- (1973) Sophie Davis School of Biomedical Education
- (1983) School of Law
- (2005) William E. Macaulay Honors College
- (2006) Graduate School of Journalism
- (2006) School of Professional Studies
- (2008) School of Public Health



- **Community Colleges**

- (1957) Bronx Community College
- (1958) Queensborough Community College
- (1963) Borough of Manhattan Community College
- (1963) Kingsborough Community College
- (1968) LaGuardia Community College
- (1970) Hostos Community College

- Largest urban university in the United States.
- Third largest university system in the United States.
 - Faculty
 - 7,331 full-time
 - 3,500 adjunct
 - 269,000 degree program students.
 - 230,000 adult, continuing and professional education students.
- Alumni
 - 12 Nobel prize winners
 - 2 Fields Medal winners
 - Dr. Jonas Salk, Polio Vaccine
 - Robert Kahn, Co-developer of TCP/IP
 - Andy Grove, co-founder and former CEO, Intel Corp.

- Located at the College of Staten Island
- Dedicated 1 Gbit per second link to CUNYnet and NYSERnet
 - Upgradeable
- 4500 square feet (computer room and offices)



- “Supercomputing” has become inexpensive/processor hour

1980



- Discrete components
- \$9 million dollars
- 160 million calculations/second
- 8 megabytes of memory
- 11,000 pounds
- 100,000 watts

2013



- Integrated components
- \$6,000
- 160,000 million calculations/second
- 64 gigabytes of memory (or more)
- 20 pounds
- 1,100 watts

- **Cost to store data has plummeted**
- **Amount of data collected and stored is increasing astronomically**

1978-1983



- 1.4 billion characters of data
- \$600,000
- \$429 per million characters

2013



- 2,000 billion characters of data
- \$150
- \$0.000075 per million characters

Inexpensive Communications Bandwidth

- TCP/IP
- Fiber optics
- Cellular technology

Inexpensive and powerful SENSORS and PDAs



The New Computing Ecosystem



- UBIQUITOUS and JOINED computing environment **S** accessible from everywhere
- Proliferation of sensors gathering **DATA** everywhere on almost everything
- Inexpensive storage allows for **DATA** retention
- An ecology that allows exploration of **UNSTRUCTURED SEMANTIC DATA**

© 2007 CNET Networks, Inc.

Who Am I/Who are you???



- We are constantly broadcasting who we are
- This data is being captured and stored
- **Organizations/people want to learn who we are and predict, from this data, what we will do**

- New application areas – LARGEST HPC GROWTH AREA (my opinion)
 - Humanities
 - Social sciences
 - Personal/organizational/societal relationships
 - Medicine
- Need for new algorithms
 - Semantic data
 - Graphs
 - Dynamic graphs
- New architectures
 - Very-large shared memory
 - Very large number of threads
 - New processor types??

Research Objective:
Identify neural
correlates of individual
differences in behavior

- Neuroscience
- Computer science
(graph application,
i.e., mapping social
behavior to
brain/neuron
structure)

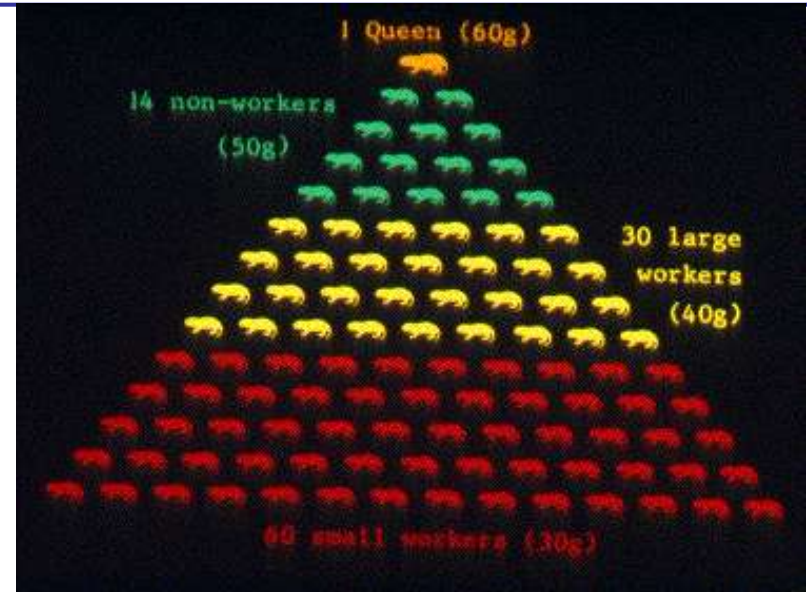


Courtesy: Dr. D. McCloskey, College of Staten Island

Naked African Mole Rat

- Not a mole, nor a rat, but the only cold-blooded mammal
- Lives over 30 years.
- Navigates using magnetic fields.
- Lives under extreme hypoxia.
- Males and females are barely distinguishable
- Communicate through chirps
- Have individual recognition.
- Is cancer resistant.
- Does not feel pain

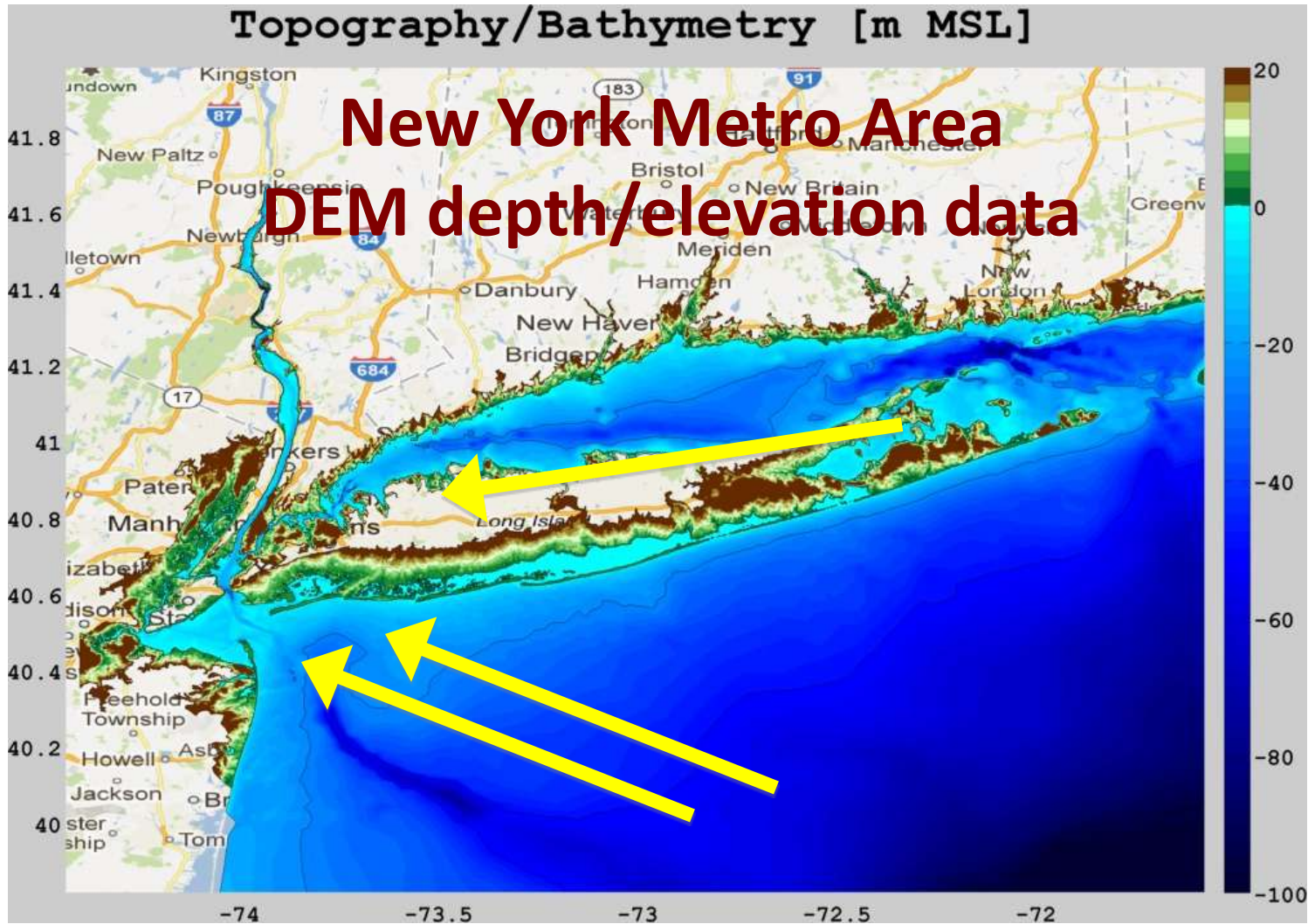
Courtesy: Dr. D. McCloskey, College of Staten Island



- Is one of only 2 eusocial mammalian species (both African Mole rats)
- Each colony has a queen and breeding male
- Offspring will likely never reproduce (less than 5% chance)

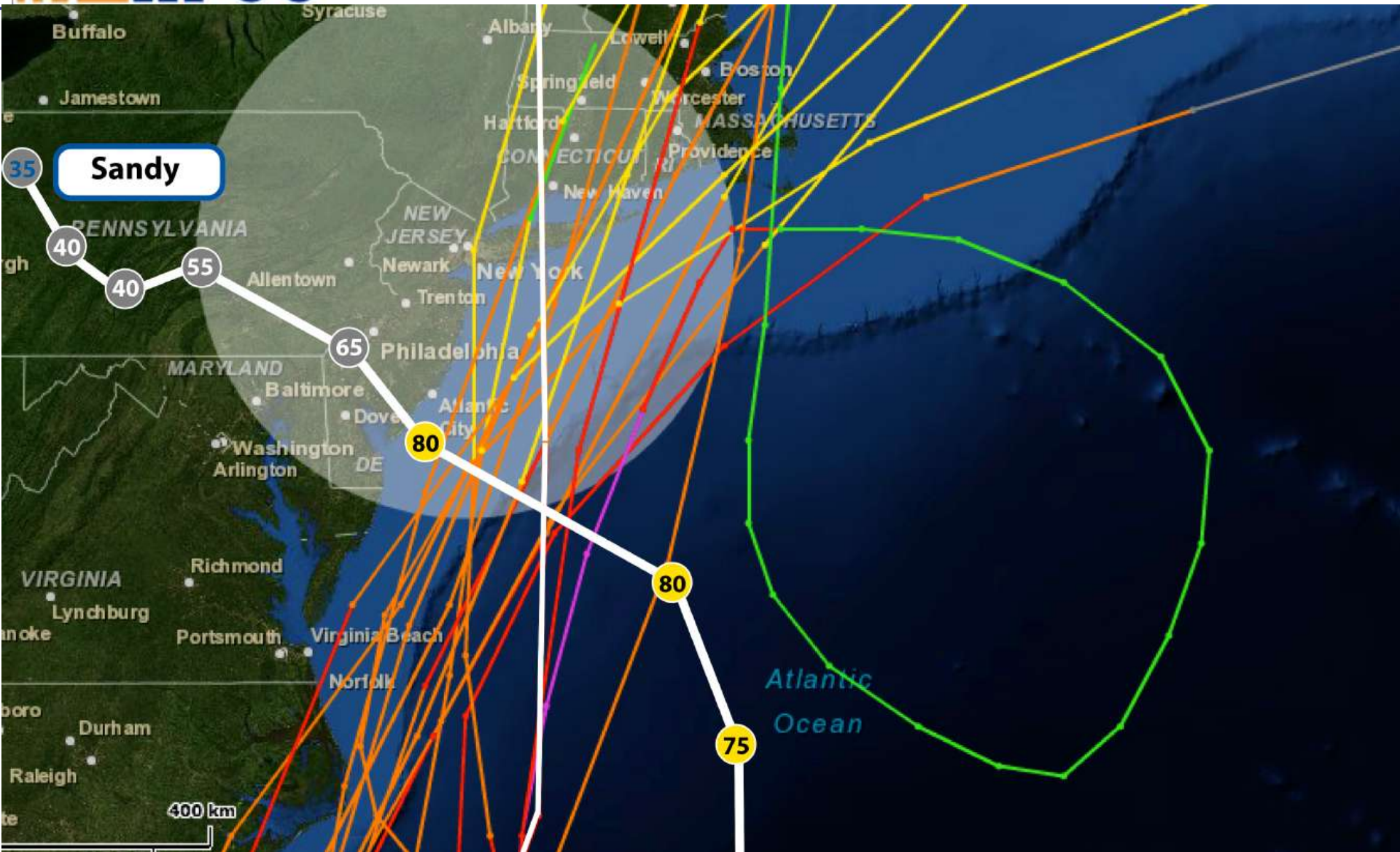
- Tropical storm-force winds spanned an area of ocean 940 miles in diameter (29 October 2012)
- region of ocean covered by 12-foot high seas spanned an area of 1560 miles in diameter
- Reached New Jersey shore at high tide and astronomical Full Moon
 - But would have been worse if it arrived at first high tide of the day or at the New Moon
- 3.4 meters (MLLW) predicted surge (tip of Manhattan)
- Actual was 4.3 meters

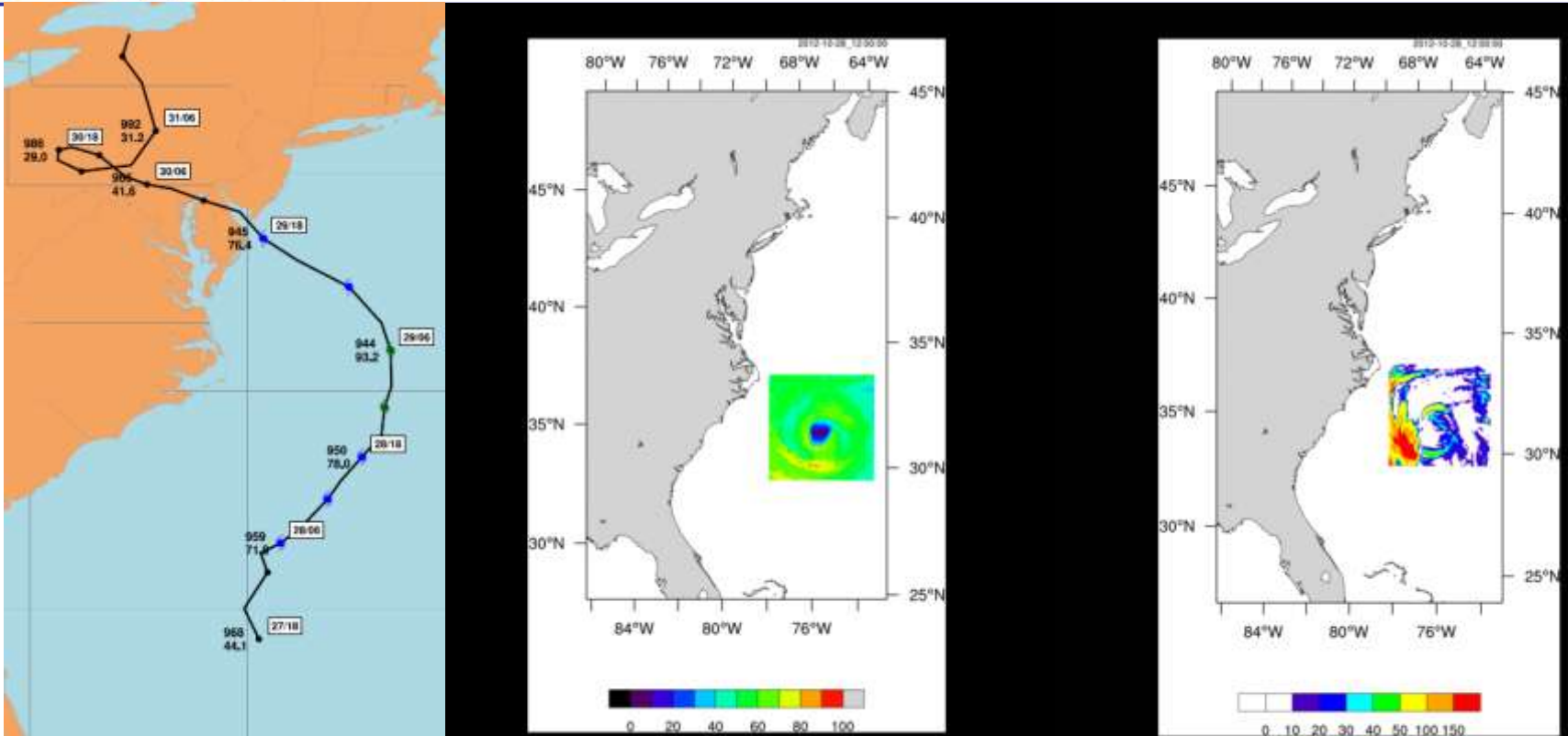
Hurricane Sandy



Courtesy: Dr. A. Benimoff, College of Staten Island

Hurricane Sandy

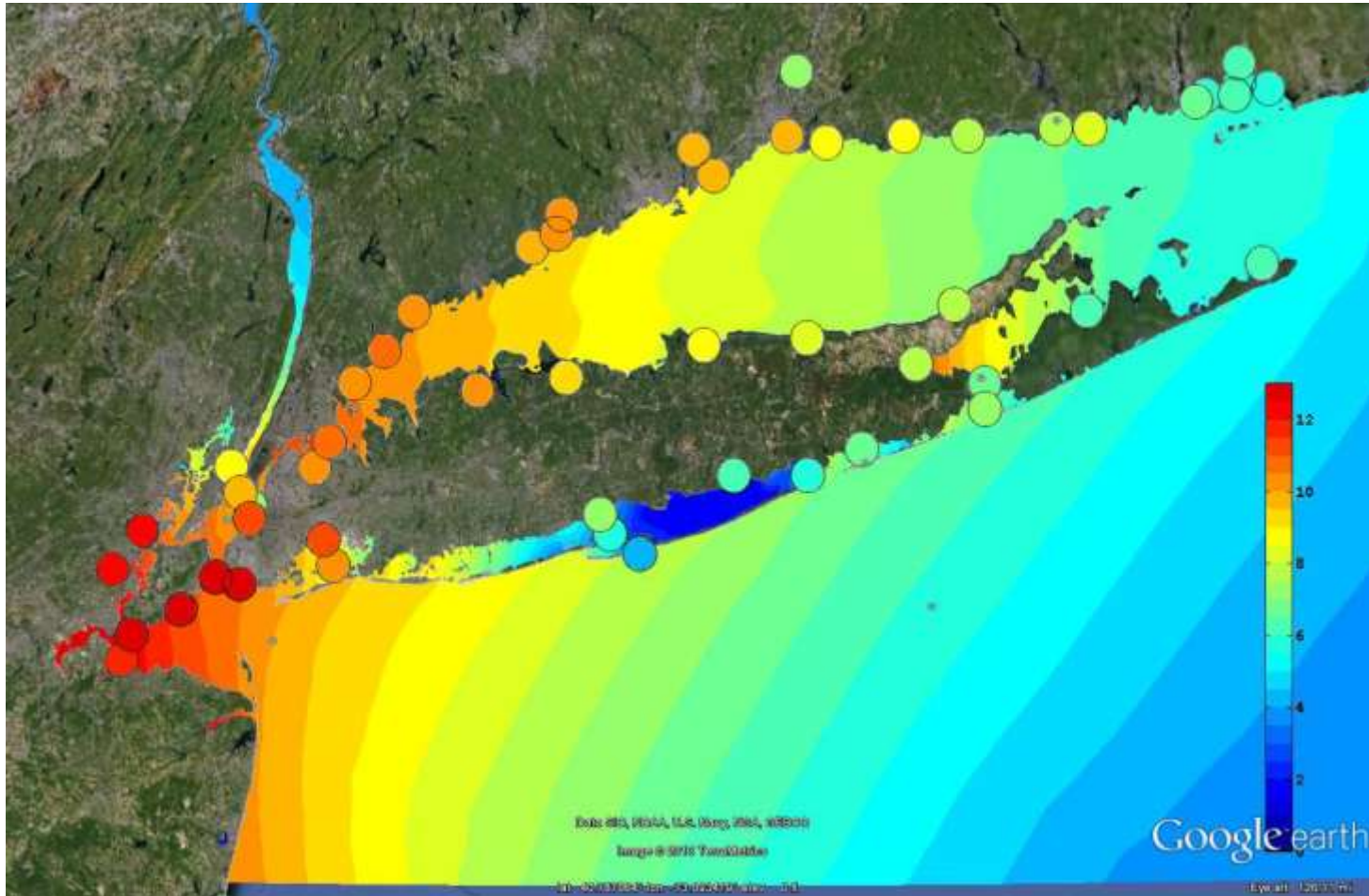




- Hurricane Sandy WRF hindcast simulation Track of eye (left)
 Wind speed, MPH (center)
 6-hr accumulated precipitation, mm (right)
 24 hr spin-up, 72 hour forecasting using NAM input

Courtesy: Drs. M. Arend, J. Gonzalez, CCNY

Surge Model vs Measured Data



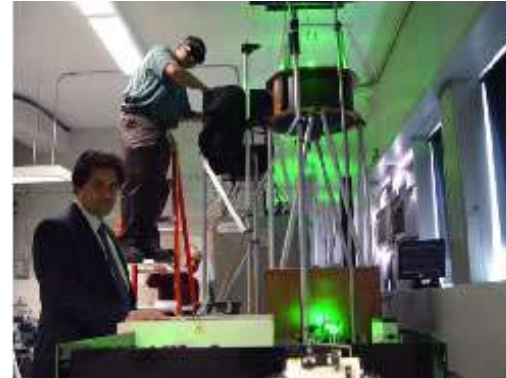
ADICRC output (background) compared to USGS measured elevation points (colored circles). Courtesy: College of Staten Island

- **Modeling and simulation**
 - Surge prediction was underestimated partially due to the MET models
 - Some of the ocean models were likely too frictional, slowing the water.
 - Science of wind-to-ocean momentum transfer under hurricane strength winds is still young
 - Need improved models to address the effectiveness of soft (marsh lands, dunes, oyster beds) and hard protective systems
 - Interaction of high winds and urban structures
 - Re-work Evacuation Zone Maps
 - ADCIRC is the storm surge model used by FEMA and Corps of Engineers
 - College of Staten Island and CUNY collaborates with Stevens Institute and Renaissance Computing Institute and University of North Carolina

- **CREST Satellite Receiving Station**
- **MODIS Aqua and Terra Satellites**
- **Level 0**
- **Level 1B**
 - HDF product
 - Cloud Mask
- **Level 2 using algorithms listed below**
 - Aerosol, cloud properties, cloud profiles, corrected reflectances, fire, ocean color, sea surface temperature, sea ice, snow cover, vegetation

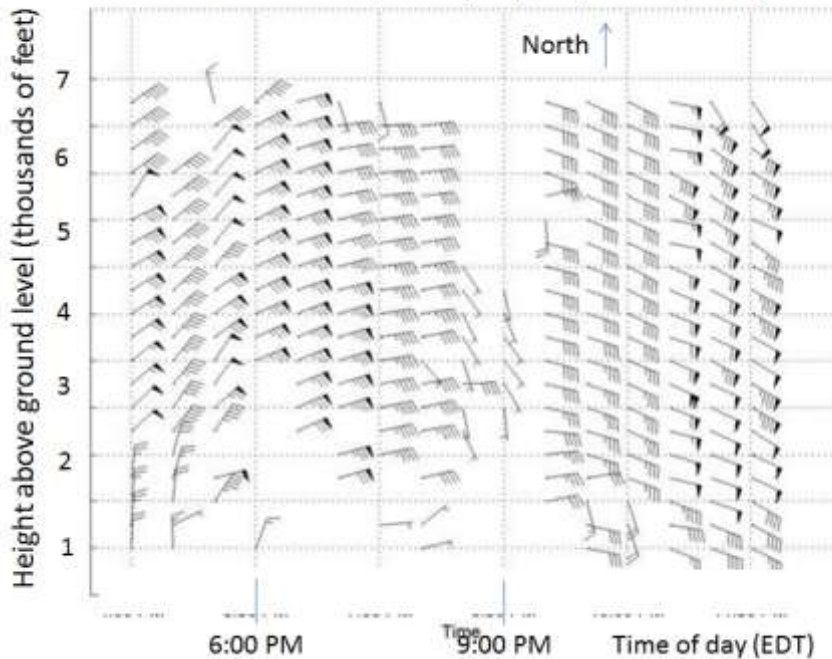


- **Lidar Transmitters**
- **Lidar Receiver Stations**

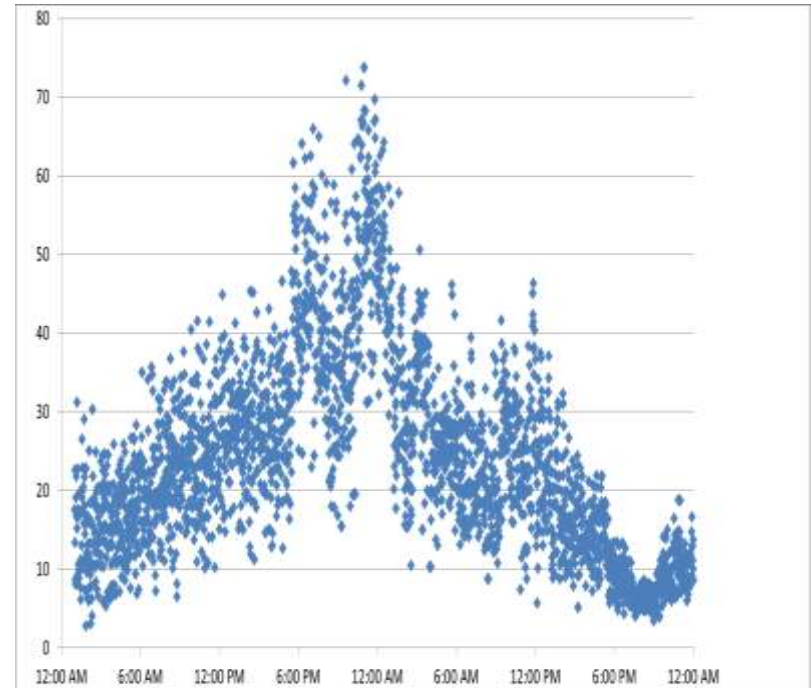


- **NYMetNet**
 - Approximately 400 ground stations across New York City collection air quality data

Hurricane Sandy Wind Barbs Measured by NYCMetNet
Radar Wind Profiler, Jersey City, NJ (Oct. 29, 2012)



- Peak recorded sustained wind (1/2 hr average)
- Speed: 125 mi/hr
- Time and date: 15:30 PM Oct 29
- Height: 1.65 km above ground level

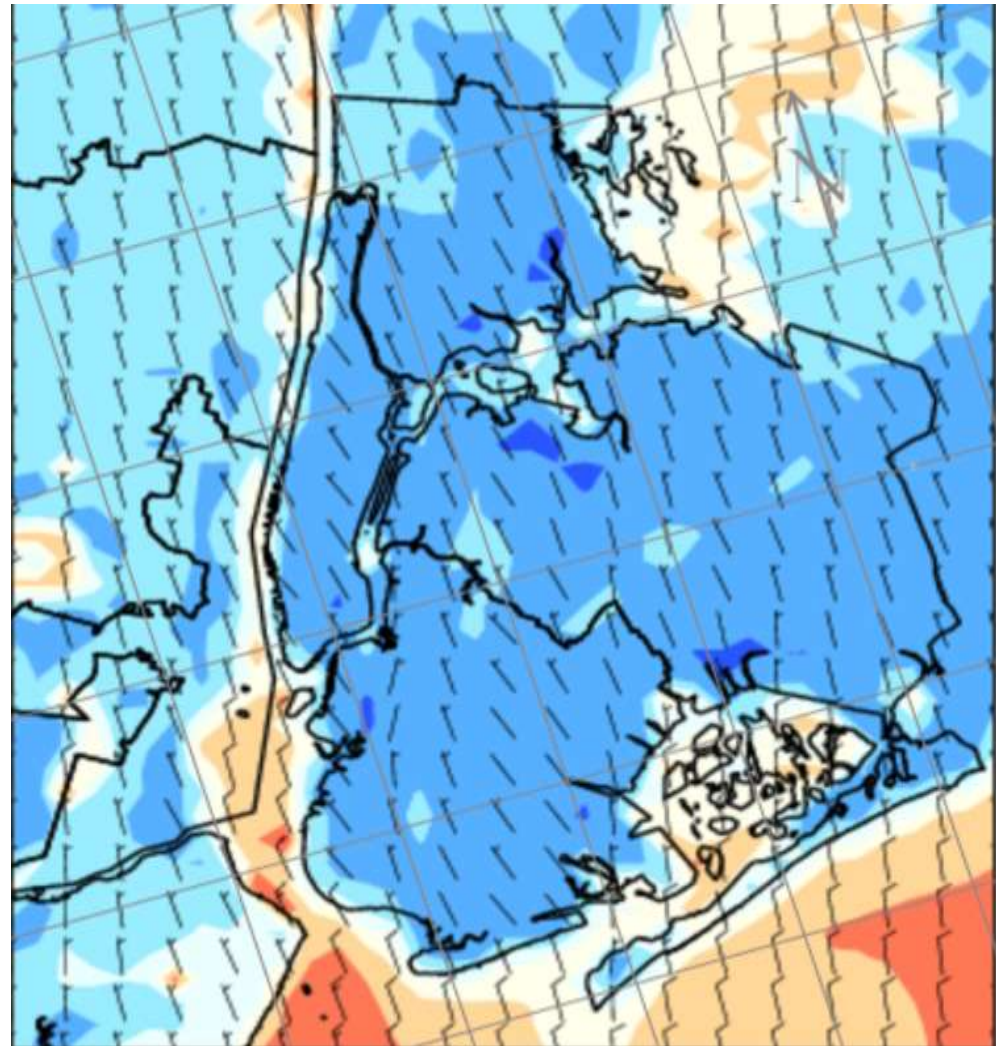


- 1-min average
- Midtown Manhattan Building
- Sonic anemometer

Courtesy: Drs. M. Arend, J. Gonzalez, CCNY

- uWRF
 - Includes effects of building height, heating, and reflectivity
 - 333 meter resolution
- Automated feeds
 - Raw data from CCNY to HPCC
 - Predictions fed back to CCNY
- 2-day predictions posted to NYCMetNet

<http://air.cuny.cuny.edu/ws/wrfn/thindex.wrfmetnet.php>



Courtesy: Drs. M. Arend, J. Gonzalez, CCNY

- Office of Data Analytics
- Develop a template for customer-facing programs to facilitate real-time intelligence, efficient data collection/sharing
 - City agency data, electricity, fuel, telecommunications
 - Apps like “Gas Buddy”
 - Up-to-date geosupport system at granular building level
 - Police Domain Awareness System



- Inventory of vulnerable/homebound people
 - Mobilize critical personnel
 - Dialysis and critical prescriptions
- Upgrade telecommunications systems
 - Backup cell-tower facilities
 - Standardized/recognizable messaging systems

- Traditional HPC (PDEs) is continuing to evolve and is allowing us to develop more accurate physical models
 - Remains a major research/educational area
- The computational ecosystem is enabling us to address problems in research domains where HPC was not previously exploited
 - Social sciences, humanities, business analytics
- Requires
 - New algorithms and approaches
 - Curriculum updates (students are often ahead of the faculty)