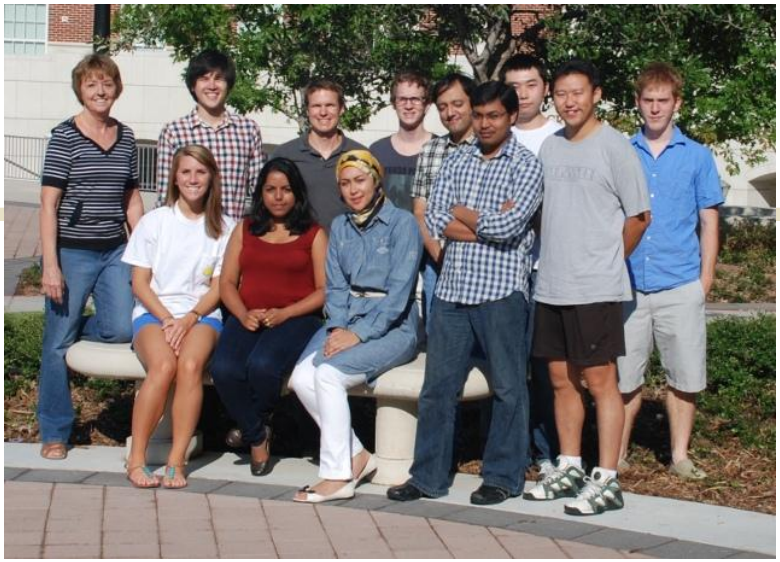
Data Stream Modeling: Hurricane Intensity Prediction

Michael Hahsler & Margaret H Dunham
Computer Science and Engineering

IDA@SMU

Intelligent Data Analysis Lab



Team led by
Margaret H. Dunham
Michael Hahsler

Mission
At IDA@SMU we create novel techniques inspired by knowledge discovery, data mining, machine learning, artificial intelligence and statistical analysis to work with data from various sources.

Current Focus

- Massive data stream modeling: TRACDS™
  - **Hurricane intensity prediction**
  - Effective metagenomic classification for the Human Genome Project
- Recommender systems: R/Apache Mahout

http://www.lyle.smu.edu/IDA

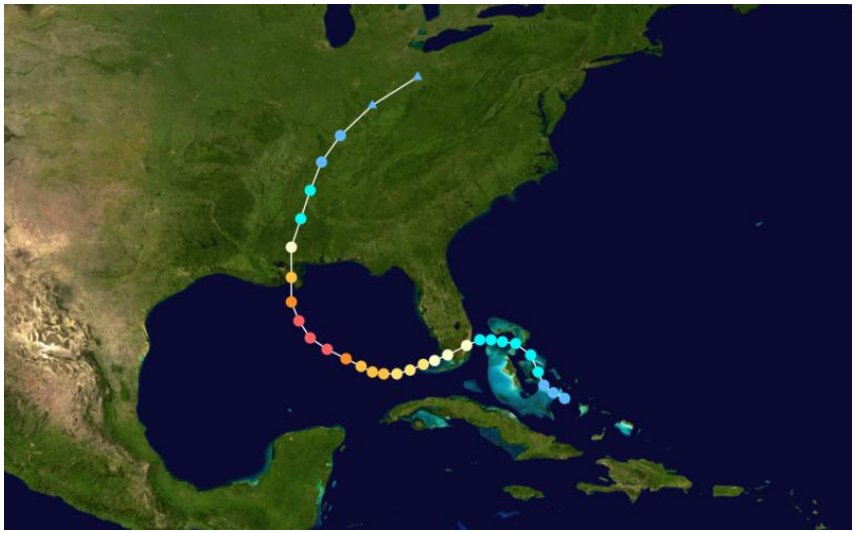
Lower 9th Ward of New Orleans, Louisiana, Feb 27, 2006
Photographer: Mackenzie Schott

Hurricane Intensity Prediction

"Objective: Improve forecast skill to accuracy and confidence levels required for decision-making and risk management"

NOAA's National Weather Service Strategic Plan 2010-2020

- **Intensity = sustained wind speed**
- **Very difficult to predict (rapid intensification)**
- **National Hurricane Center uses**
  - **Dynamical models:** computational intensive and slow
  - **Statistical models:** *Statistical Hurricane Intensity Prediction Scheme (SHIPS)*



Path of Hurricane Katrina (2005)
Color shows intensity

Katrina Facts:

- Category 5 - 175 mph
- Damage: >$100 billion
- Fatalities: >1,800

PIIH – Prediction Intensity Interval Model for Hurricanes

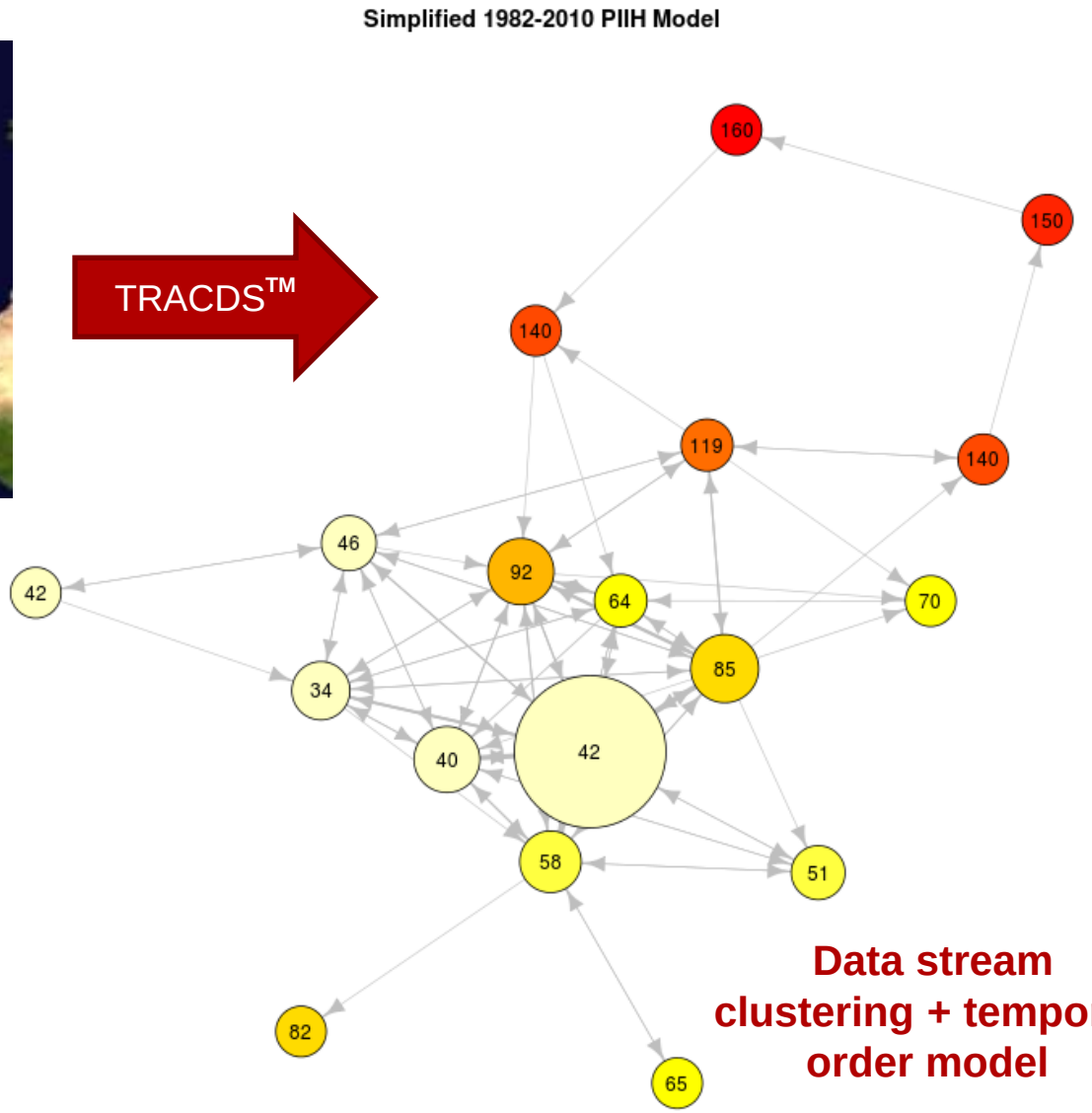
Historic hurricane data

Features

- Current wind speed
- Various temperatures
- Time of the year
- Direction of movement
- GOES Satellite Data (IR)

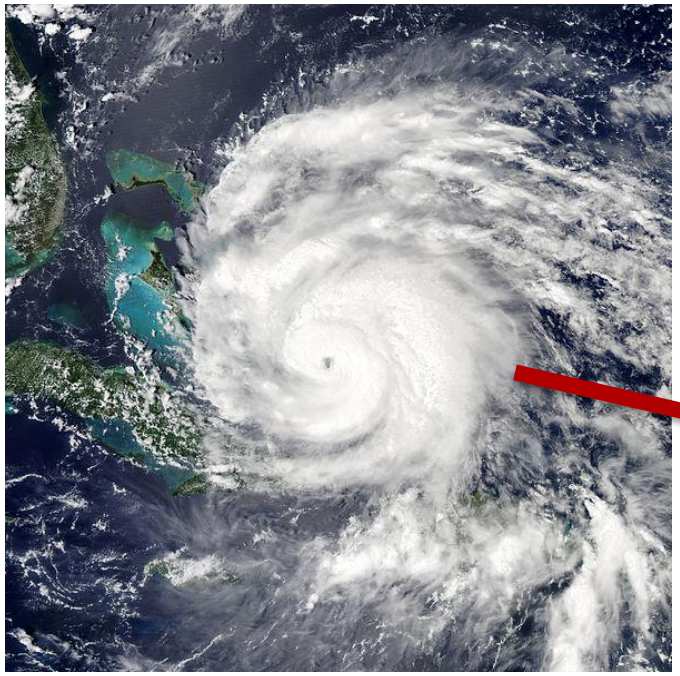
Currently 23 features from the Statistical Hurricane Intensity Prediction Scheme (SHIPS)

TRACDS™

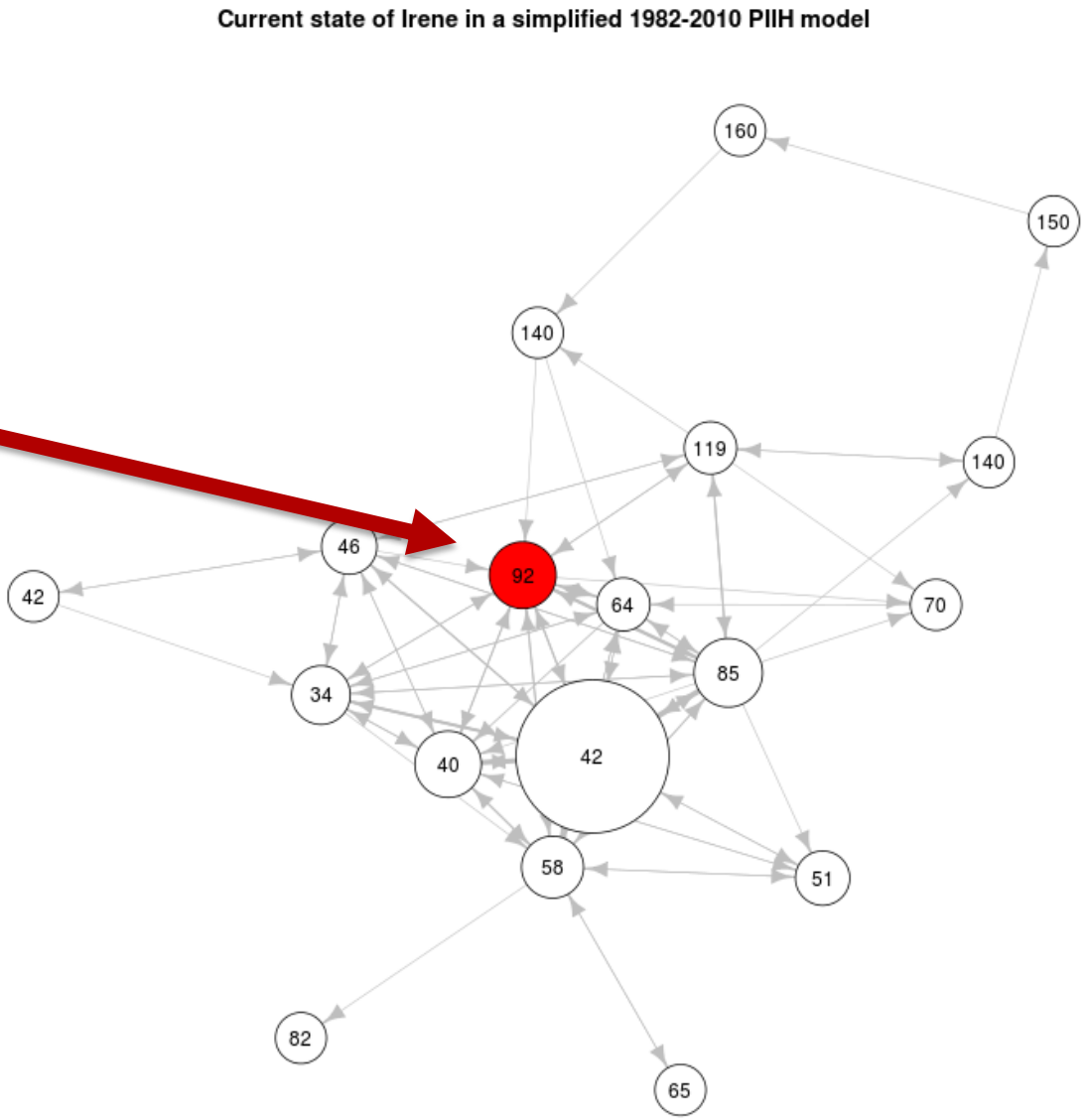


Data stream clustering + temporal order model

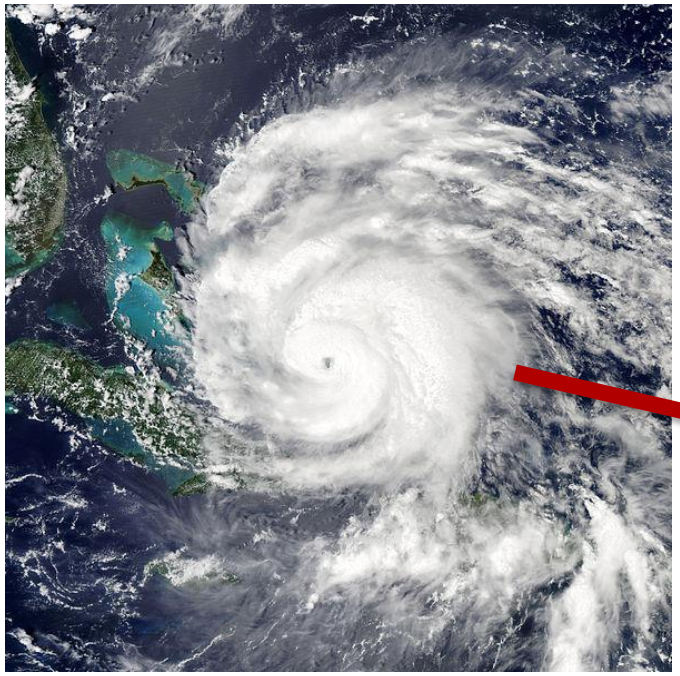
Prediction using PIIH – Irene (2011)



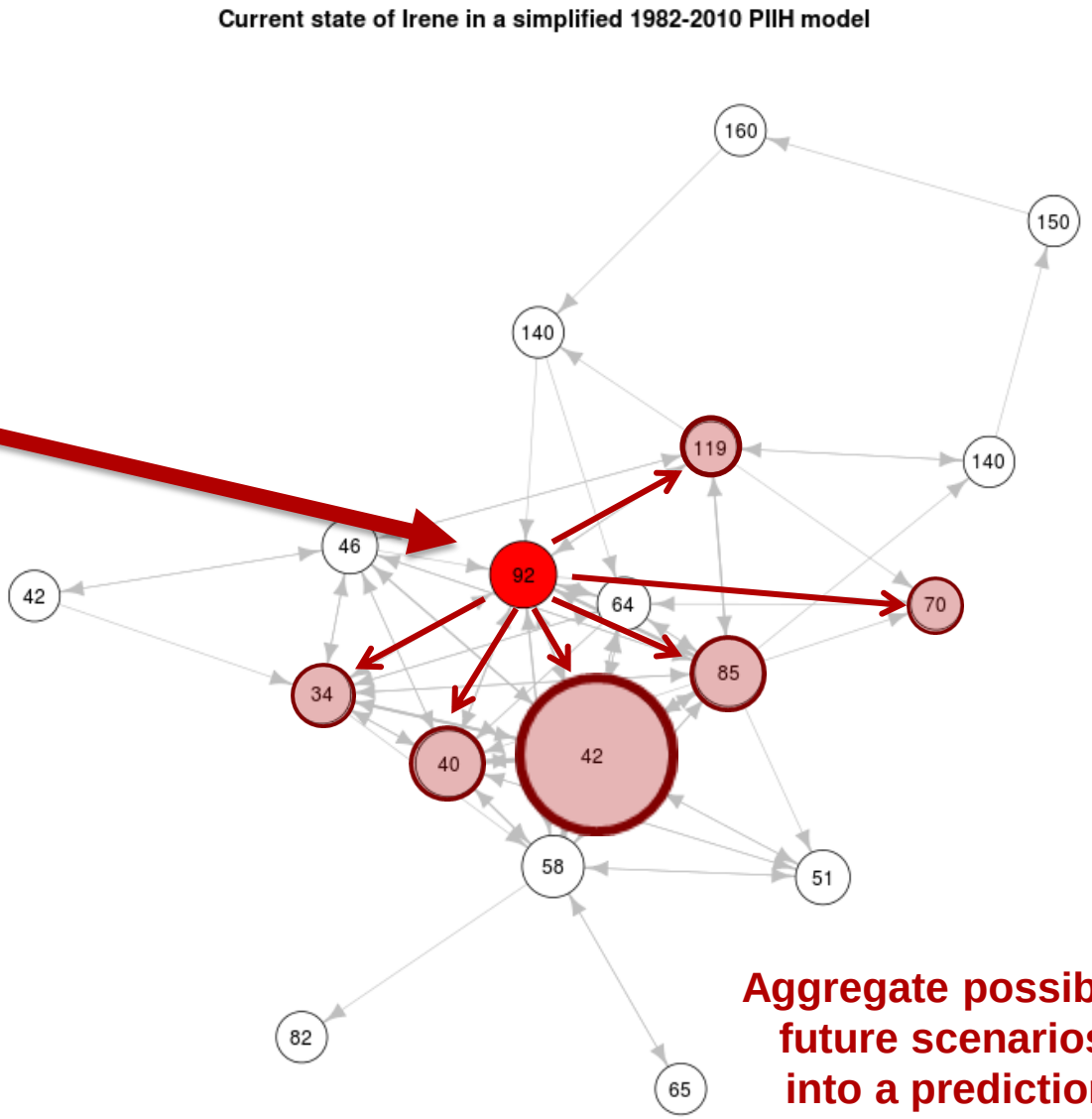
Current features of hurricane



Prediction using PIIH – Irene (2011)

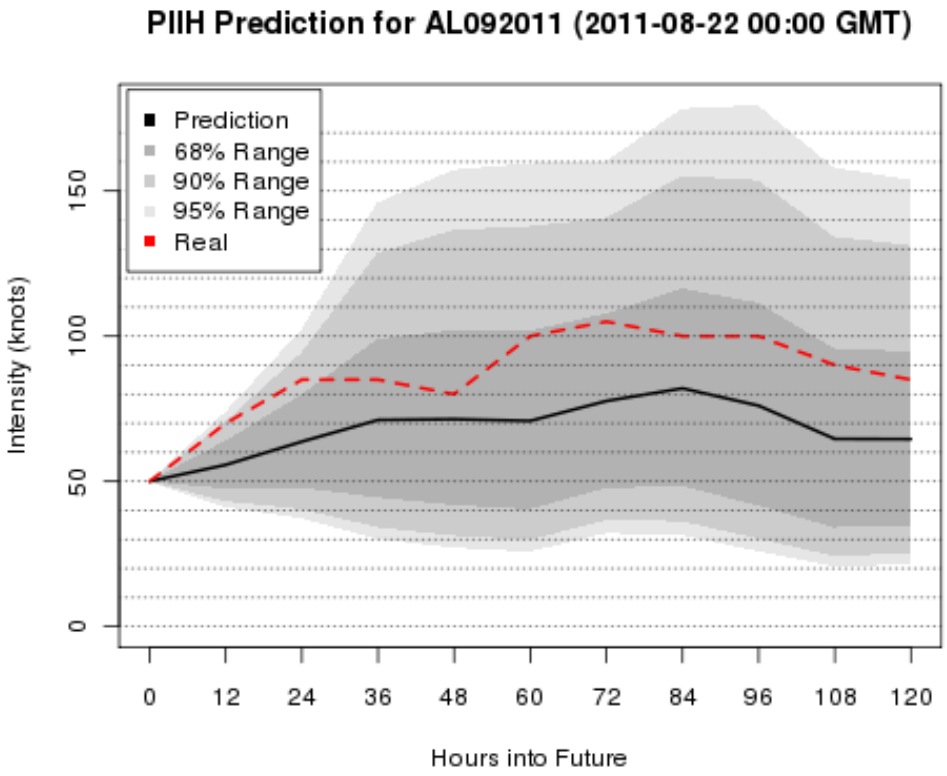


Current features of hurricane



Aggregate possible future scenarios into a prediction

Prediction Results for Irene (2011)



	MAD	MSE
PIIH	***14.28***	***310.79***
SHIFOR 5*	12.64	229.49
LGEM	15.06	411.73
SHIPS	14.80	319.64
D-SHIPS	17.11	500.36

MAD … Mean average deviation
MSE … Mean squared error
* Baseline model

http://IDA.lyle.smu.edu/PIIH/

PIIH: Learning Prediction Intensity Interval model...

IDA@SMU

SMU | Bobby B. Lyle School of Engineering

Navigation
- PIIH home
- IDA home
- TRACDS
- Contact us

Hurricane Info
- NHC (NOAA/NWS)
- RSS feed (NHC)
- 2011 Season (NHC)
- 2011 Season (WP)
- NHC Forecast Models
- Tropical cyclone scales

Data
- Live Isdiag data
- Isdiag archive
- ATCF archive
- ATCF current
- SHIPS+devel. data
- File formats

Development
- PIIH 2010

48-hour formation potential: Low <30% Medium 30-50% High >50%

Image courtesy of NHC/NOAA

Tropical Cyclones Intensity Forecasts for Tropical Cyclone/Depressions using PIIH*

NAME	ATCF ID	PIIH	Comparison	Advisory Archive (NHC)
ARLENE	AL012011	Predictions	Model Accuracy	Advisories and Graphics
BRET	AL022011	Predictions	Model Accuracy	Advisories and Graphics

Internet | Protected Mode: On 100%

Cooperation & Media Coverage

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
UNITED STATES DEPARTMENT OF COMMERCE

James Franklin
Branch Chief, Hurricane Specialist Unit, NHC, NOAA

Mark Demaria
Chief of the NESDIS Regional and Mesoscale Meteorology Branch, CIRA, NOAA

theguardian

Weatherwatch: Can the intensity of a hurricane be predicted?

David Hambling
guardian.co.uk, Wednesday 12 October 2011 18.05 EDT

A larger | smaller

Devastation in New Orleans after Hurricane Katrina in 2005. A new modelling technique may be able to forecast the intensity of an approaching storm. Photograph: Mario Tama/Getty Images

Researchers at the Southern Methodist University in Dallas, Texas are developing a new modelling technique to predict the speed of hurricane winds. Known as the Learning Prediction Intensity Interval model, it is based on data mining using an advanced machine learning process. The computer itself works out the pattern of intensity development from a large pool of raw data, unlike existing methods where humans cherry-pick the most relevant historical data for a regression model to fit the current situation.

Current predictions just give one figure for wind speed, with an error range of about 15mph per day ahead. The new method gives the range of wind speeds (say, 50 to 70 knots), that can be expected with 95% certainty, and the ranges expected with 90% and 68% probability. The researchers are already putting up live predictions for the 2011 hurricanes as they happen from their current model (ida.lyle.smu.edu/PIIH/). Next year they plan to have a fully operational version.

© 2011 Guardian News and Media Limited or its affiliated companies. All rights reserved.

Future Work

1. **Deploy model with NOAA**
   - Add decay model over land
   - Evaluate additional features
   - Predict rapid intensification
   - Interface with NOAA's systems

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
UNITED STATES DEPARTMENT OF COMMERCE

2. **Improve the TRACDS™ model**
   - Data stream clustering
   - Higher-order effects
   - Improve model selection and outlier handling

IBM

Possible collaboration with research team at **Watson Research Center**

Thank you!

http://www.lyle.smu.edu/IDA