EECS 391
Intro to AI

Probability and Uncertainty

L10: Tue Oct 3

Next Section of Course: Probabilistic Reasoning

- Are we done yet?
- What can we do with search?
- What *can't* we do?
- Is there more to intelligence than problem solving?

Probability and Uncertainty

- fundamental role of uncertainty in reasoning and in AI
  - probability theory can be applied to many of these problems
  - probability theory is the calculus of reasoning with uncertainty
- review of basis probabilistic concepts
  - discrete and continuous probability
  - joint and marginal probability
- the process of probabilistic modeling and reasoning
- simple example

What is the role of probability and inference in AI?

- Many algorithms are designed as if knowledge is perfect, but it rarely is.
- There are almost always things that are unknown, or not precisely known.

- Examples:
  - bus schedule
  - quickest way to the airport
  - sensors
  - speech recognition
  - computer vision
  - joint positions
  - finding an H-bomb

- An agent making optimal decisions must take into account *uncertainty*.

Probability as frequency: *k* out of *n* possibilities

- Suppose we're drawing cards from a standard deck:
  - P(card is the Jack ♥ | standard deck) = 1/52
  - P(card is a ♣ | standard deck) = 13/52 = 1/4
- What's the probability of a drawing a pair in 5-card poker?
  - P(hand contains pair | standard deck) =

$$\frac{\text{\# of hands with pairs}}{\text{total \# of hands}}$$

  - Counting can be tricky (take a course in combinatorics)
  - Other ways to solve the problem?
- General probability of *event* given some conditions:

  P(event | conditions)

Making rational decisions when faced with uncertainty

- *Probability*

  the precise representation of knowledge and uncertainty

- *Probability theory*

  how to optimally update your knowledge based on new information

- *Decision theory: probability theory + utility theory*

  how to use this information to achieve maximum expected utility

- Consider again the bus schedule. What's the utility function?
  - Suppose the schedule says the bus comes at 8:05.
  - Situation A: You have a class at 8:30.
  - Situation B: You have a class at 8:30, and it's cold and raining.
  - Situation C: You have a final exam at 8:30.

Probability of uncountable events

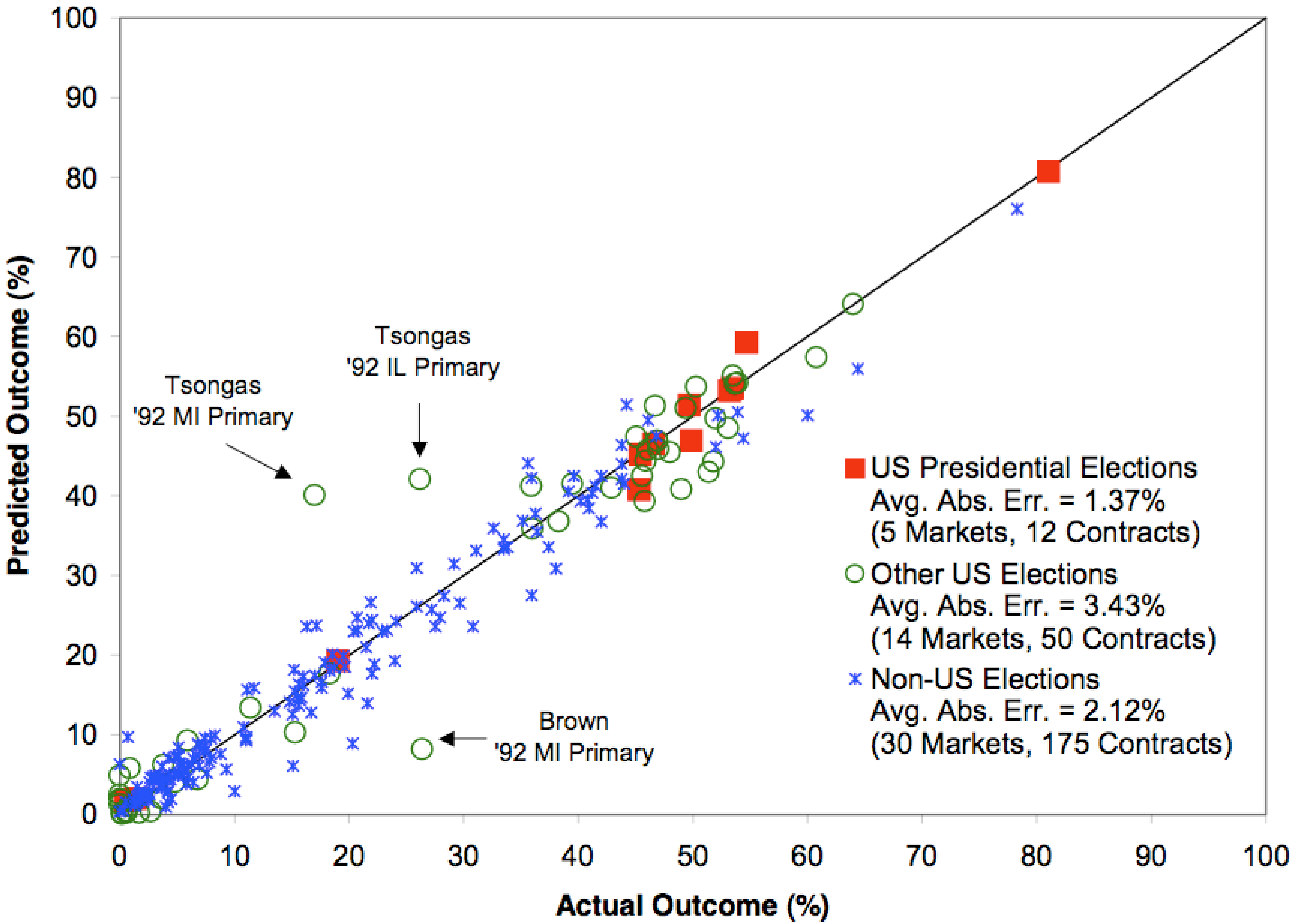
- How do we calculate probability that it will rain tomorrow?
  - Look at historical trends?
  - Assume it generalizes?

- What's the probability that there was life on Mars?
- What was the probability the sea level will rise 1 meter within the century?
- What's the probability that candidate X will win the election?

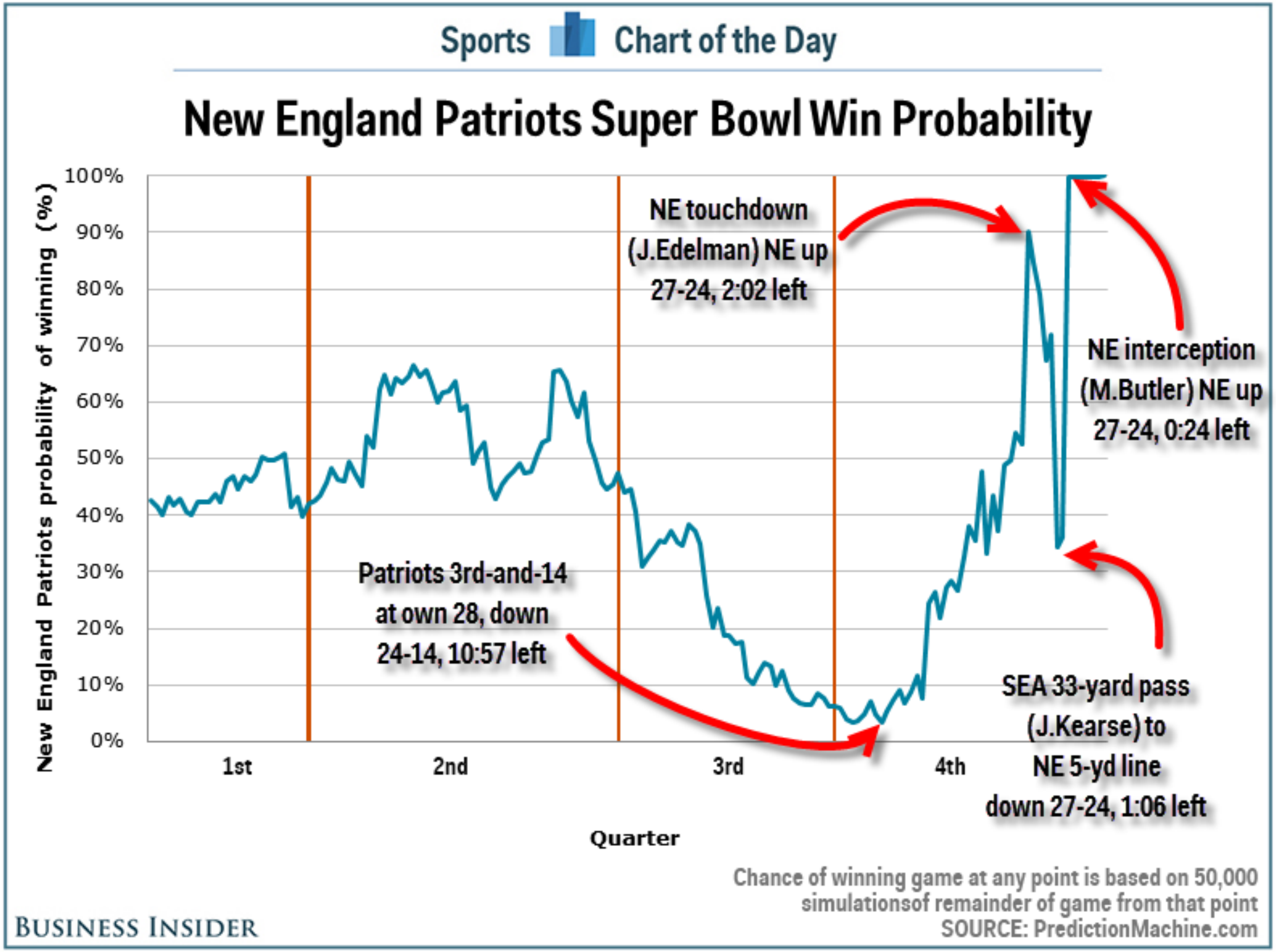
The Iowa Electronic Markets: placing probabilities on single events

- http://tippie.biz.uiowa.edu/iem/
- "The Iowa Electronic Markets are real-money futures markets in which contract payoffs depend on economic and political events such as elections."
- Typical bet: predict vote share of candidate X - "a vote share market"

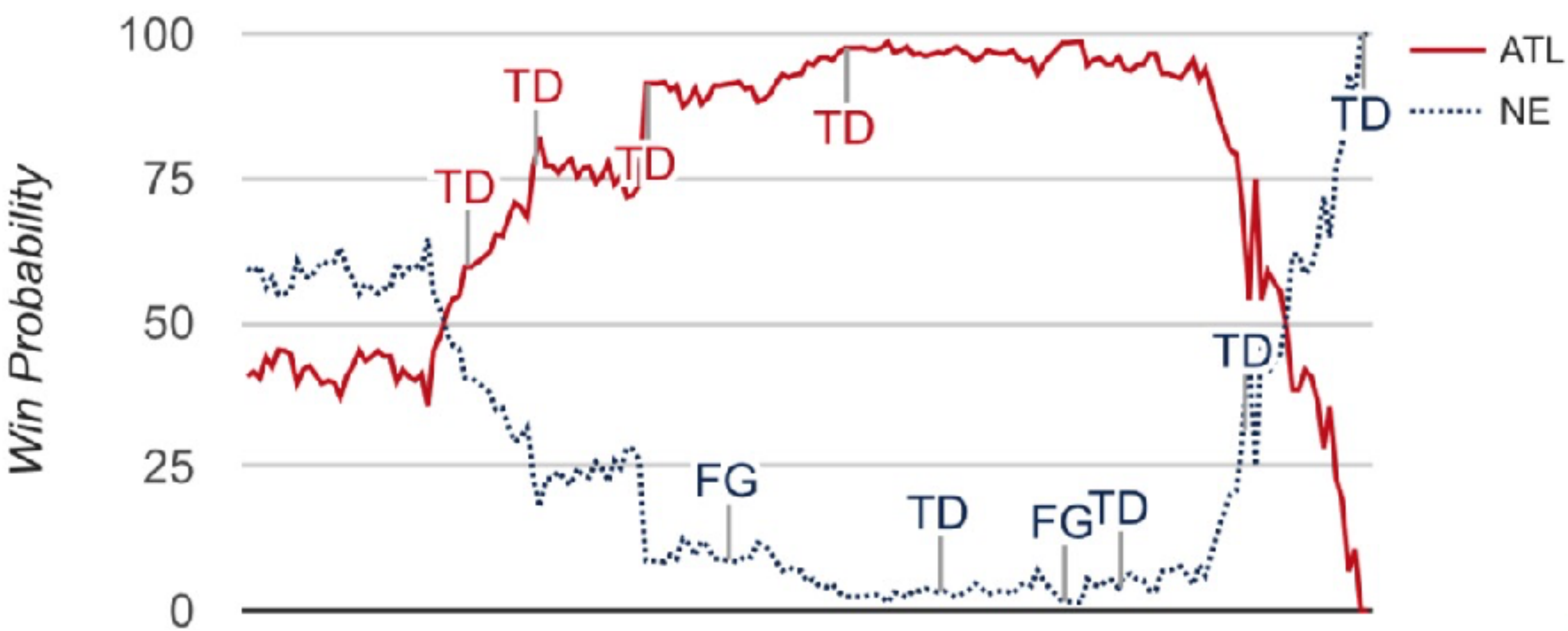
Political futures market predicted vs actual outcomes



Win Probability for Super Bowl XLIX: Patriots vs Seahawks

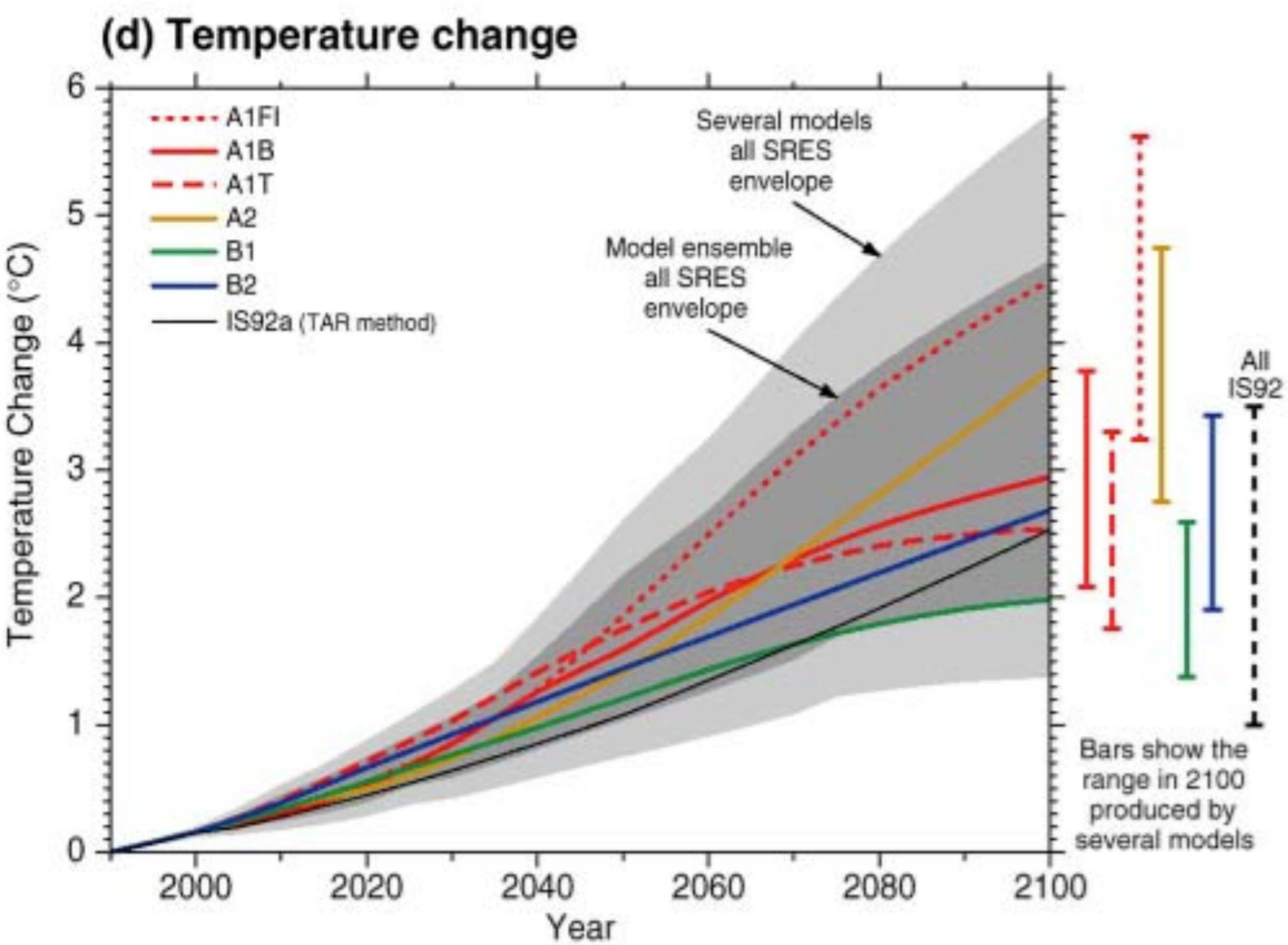


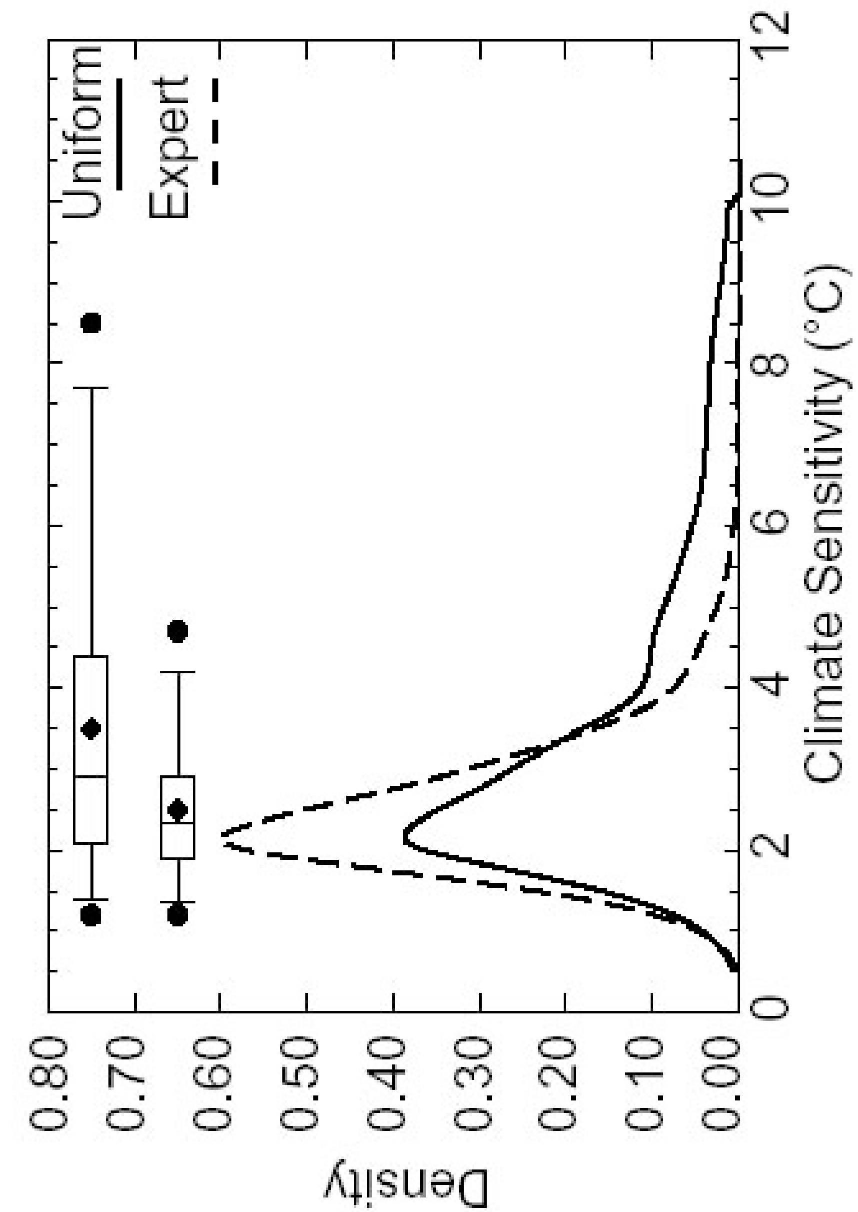
Win Probability for Super Bowl LI (Falcons vs Patriots)



Probabilistic assessment of dangerous climate change

from Mastrandrea and Schneider (2004)



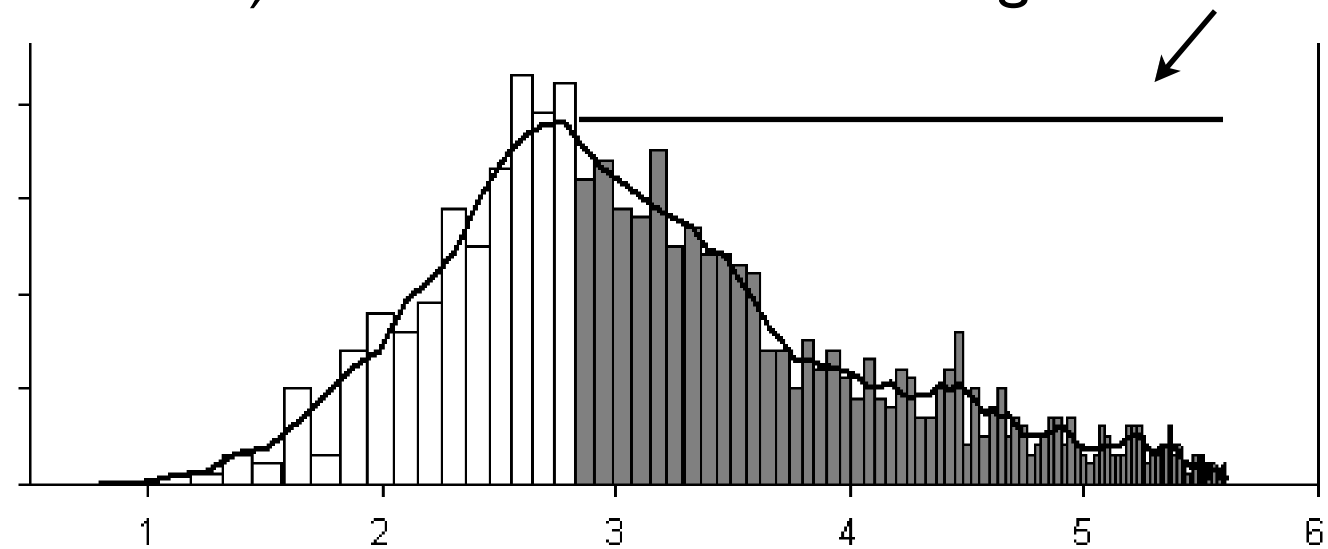


from Forrest et al (2001)

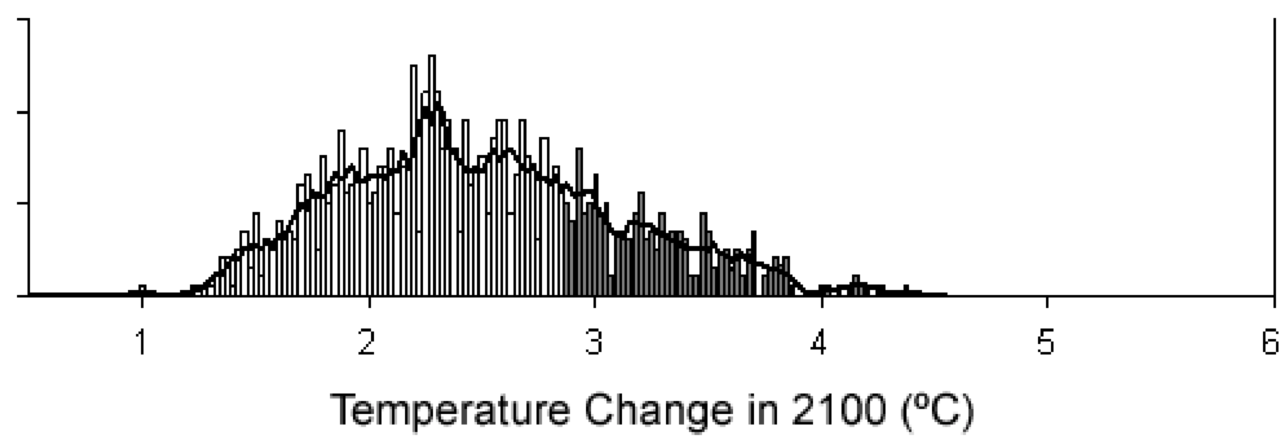
Factoring in Risk Using Decision Theory

P(“DAI” = 55.8%)

Dangerous Climate Change



P(“DAI” = 27.4%
Carbon Tax 2050
= $174/Ton



Knutti and Sedláček, *Nature Climate Change (2013)*

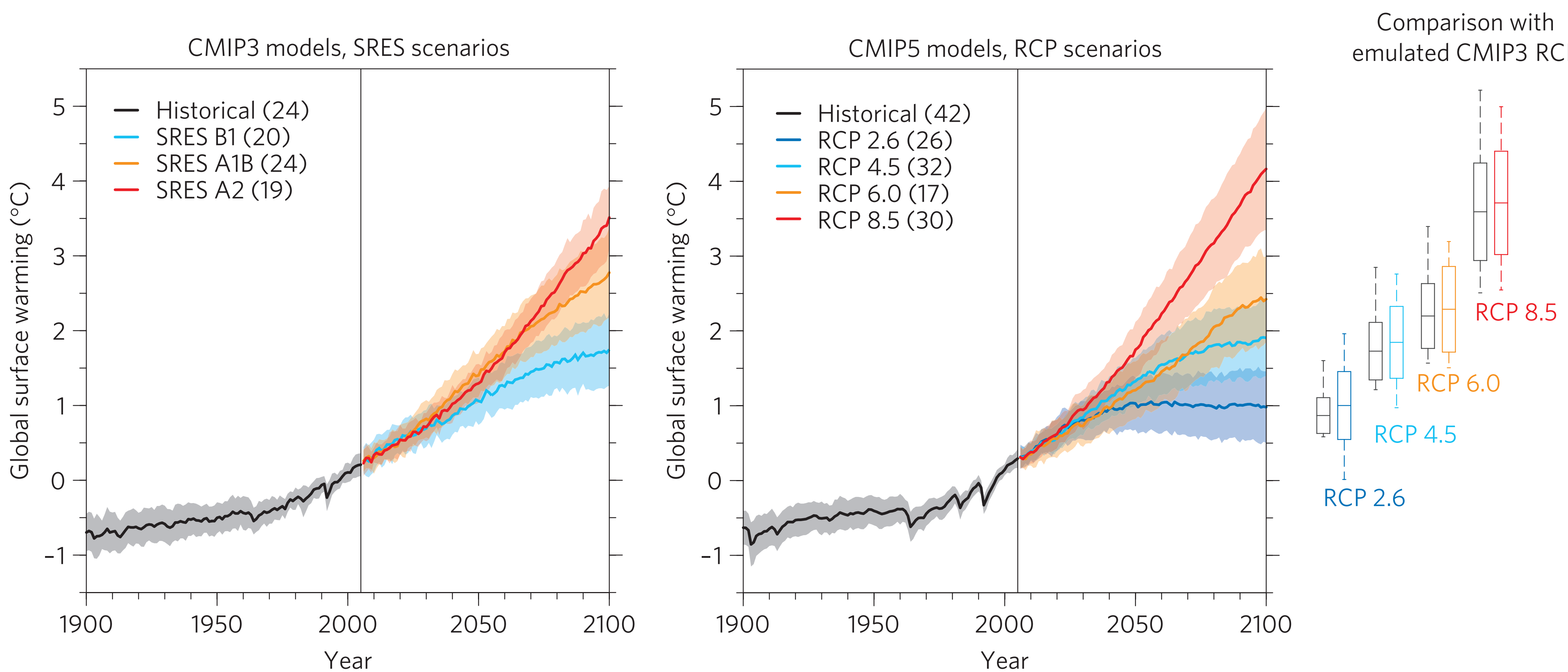
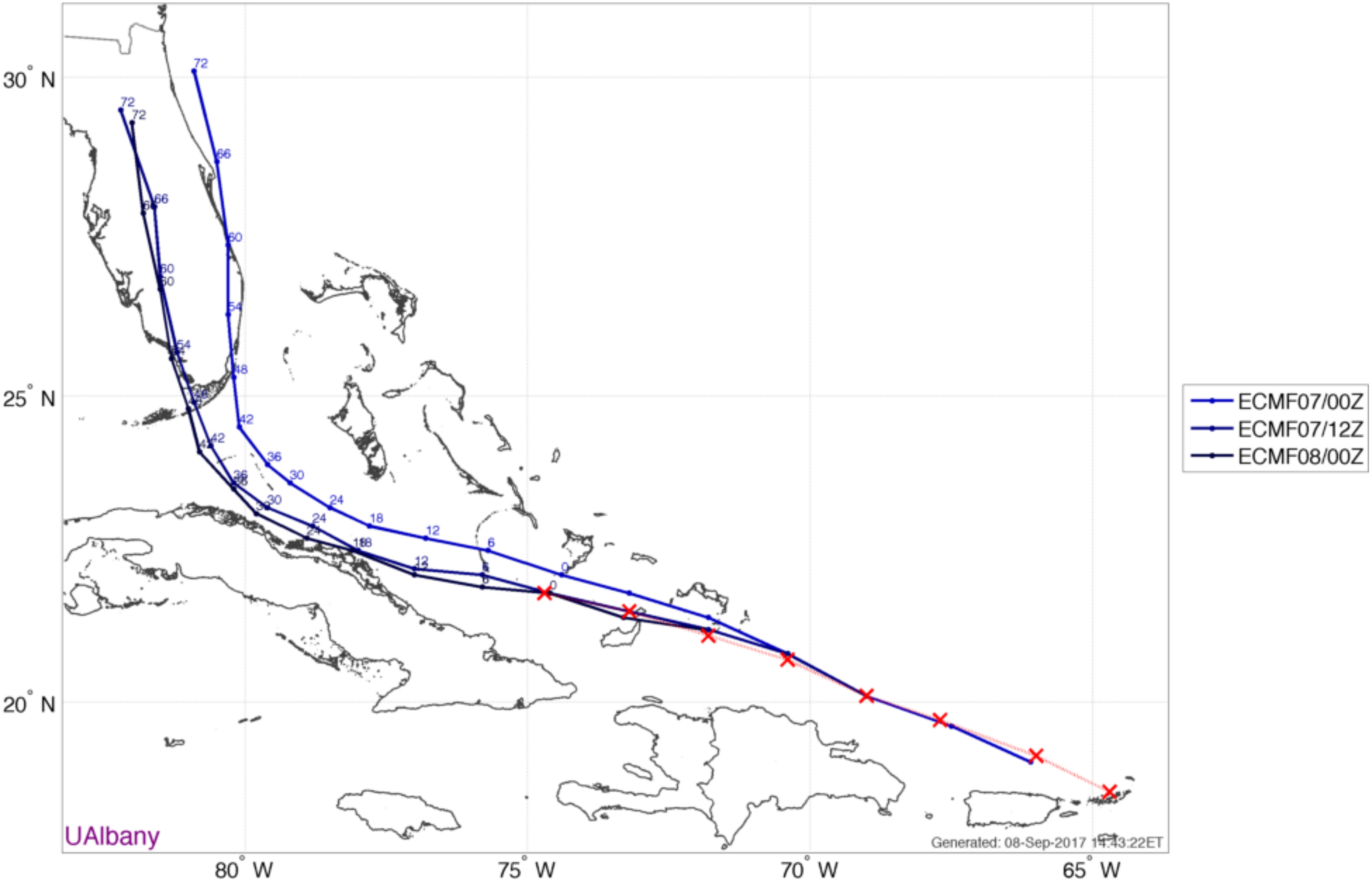
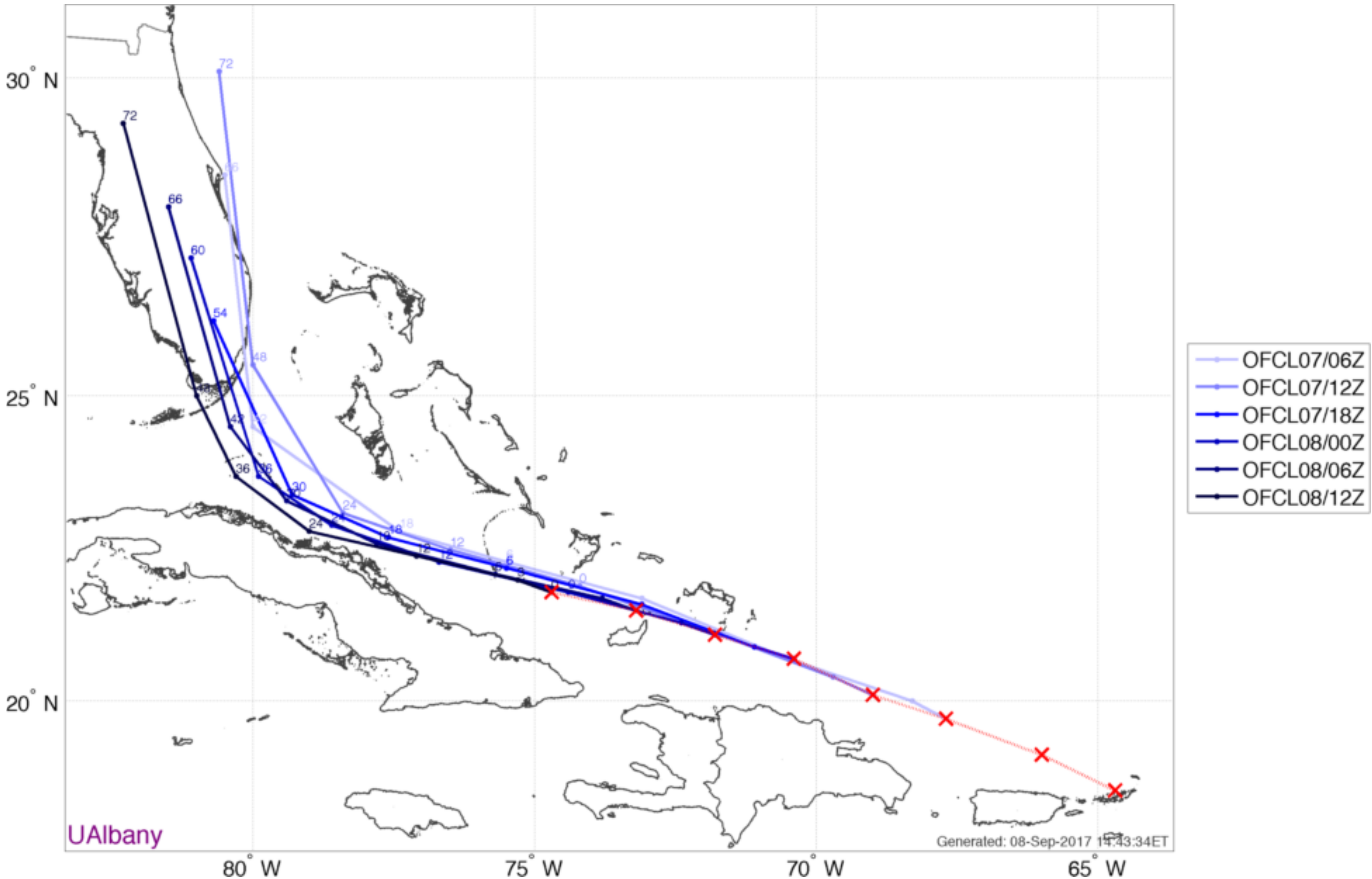


Figure 1 | Global temperature change and uncertainty. Global temperature change (mean and one standard deviation as shading) relative to 1986–2005 for the SRES scenarios run by CMIP3 and the RCP scenarios run by CMIP5. The number of models is given in brackets. The box plots (mean, one standard deviation, and minimum to maximum range) are given for 2080–2099 for CMIP5 (colours) and for the MAGICC model calibrated to 19 CMIP3 models (black), both running the RCP scenarios.

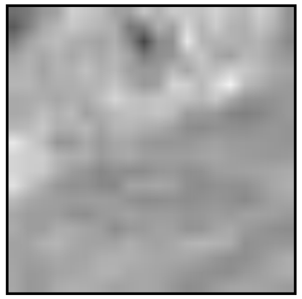
Hurricane Forecasting: European model

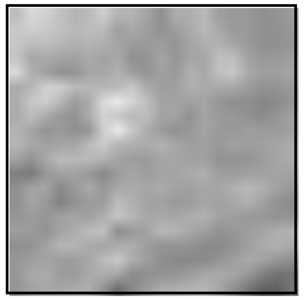


Hurricane Forecasting: National Hurricane Center

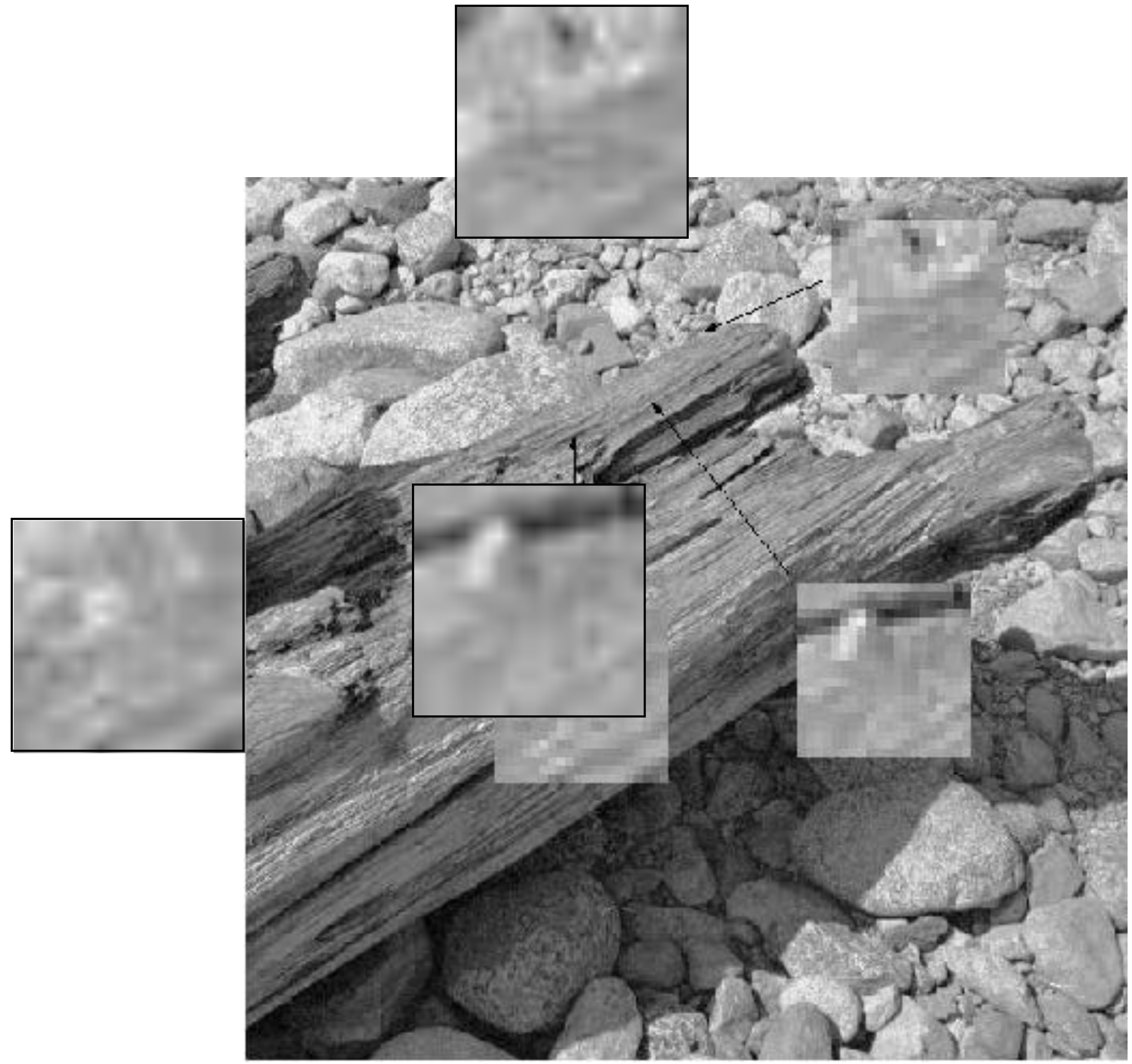


Uncertainty in vision: What are these?

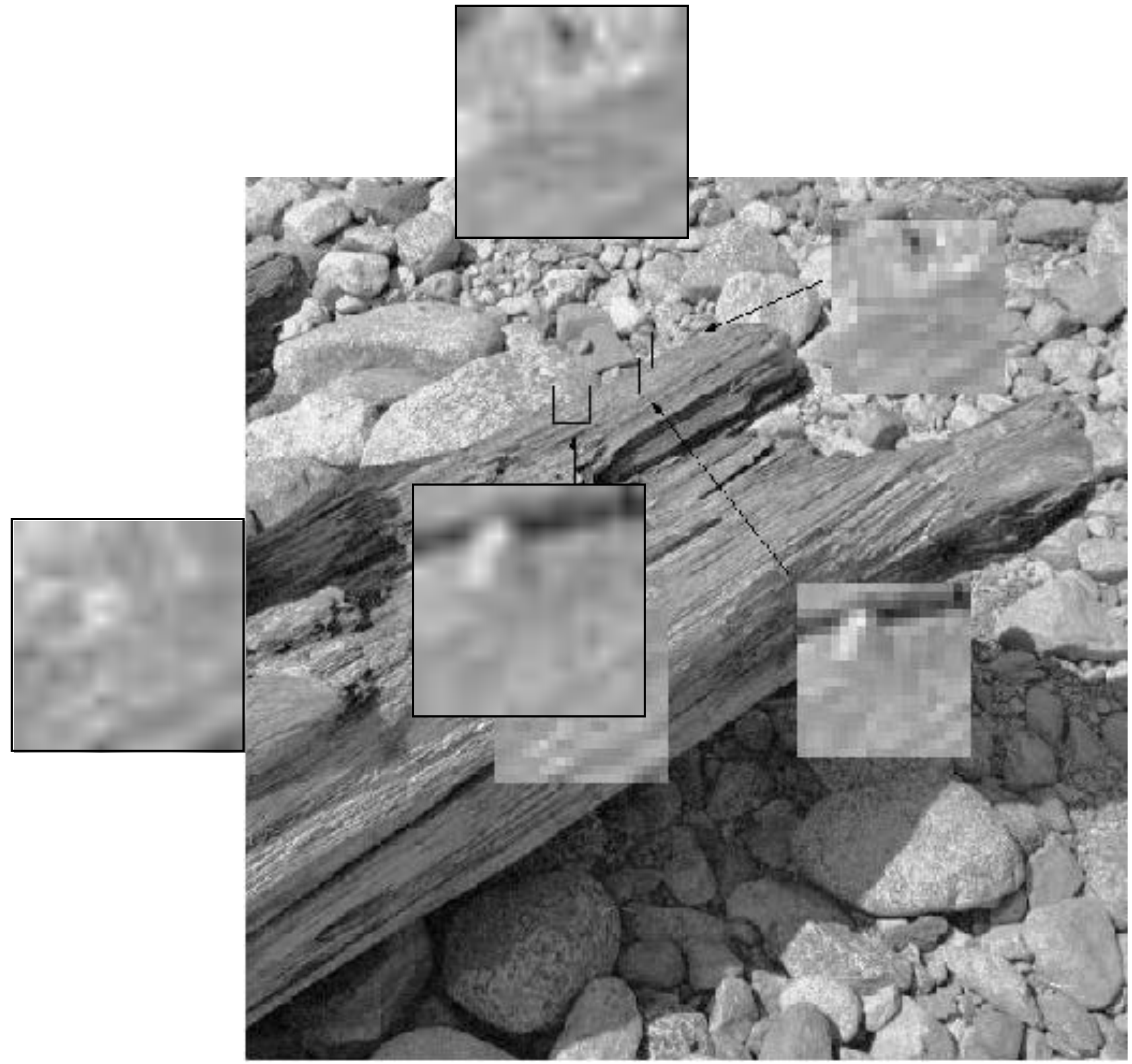




Uncertainty in vision

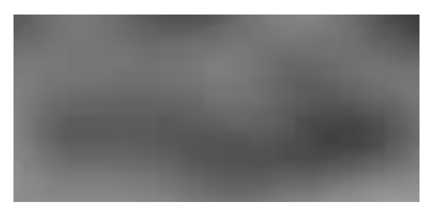


Edges are not as obvious they seem



An example from Antonio Torralba

What's this?



We constantly use other information to resolve uncertainty

Image interpretation is heavily context dependent

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 10 20

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5	10	20

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 **10** 20

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 10 **20**

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 10 20

"*Fill that canteen with fresh spring water*"

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 10 **20**

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 **10** 20

This phenomenon is even more prevalent in speech perception

- It is very difficult to recognize phonemes from naturally spoken speech when they are presented in isolation.
- All modern speech recognition systems rely heavily on context (as do we).
- HMMs model this contextual dependence explicitly.
- This allows the recognition of words, even if there is a great deal of uncertainty in each of the individual parts.

dB SNR

5 10 20

The process of probabilistic inference

1. *define* model of problem
2. *derive* posterior distributions and estimators
3. *estimate* parameters from data
4. *evaluate* model accuracy

John Craven and the missing H-Bomb

- In Jan. 1966, used Bayesian probability and subjective odds to locate H-bomb missing in the Mediterranean ocean.

John Craven and the missing H-Bomb

- Jan 1966 John Craven, deep ocean expert, got call: "I'm missing a H-bomb!"
- B-52 collided with refueling tanker at 30,000 feet off coast of Spain. 3 of 7 crew died in crash.
- And they lost 4 Hydrogen Bombs.
- Three recovered almost immediately, fourth was lost in Mediterranean.
- Height of the Cold War.
- Pres. Johnson knew Soviets would be looking for it, but most thought it would never be found.
- How to find it?

John Craven and the missing H-Bomb

- Small object, somewhere in 100's of square miles of ocean…
- But, Craven knew Bayesian Theory.
- Made map of area
- Asked different groups to place "bets"

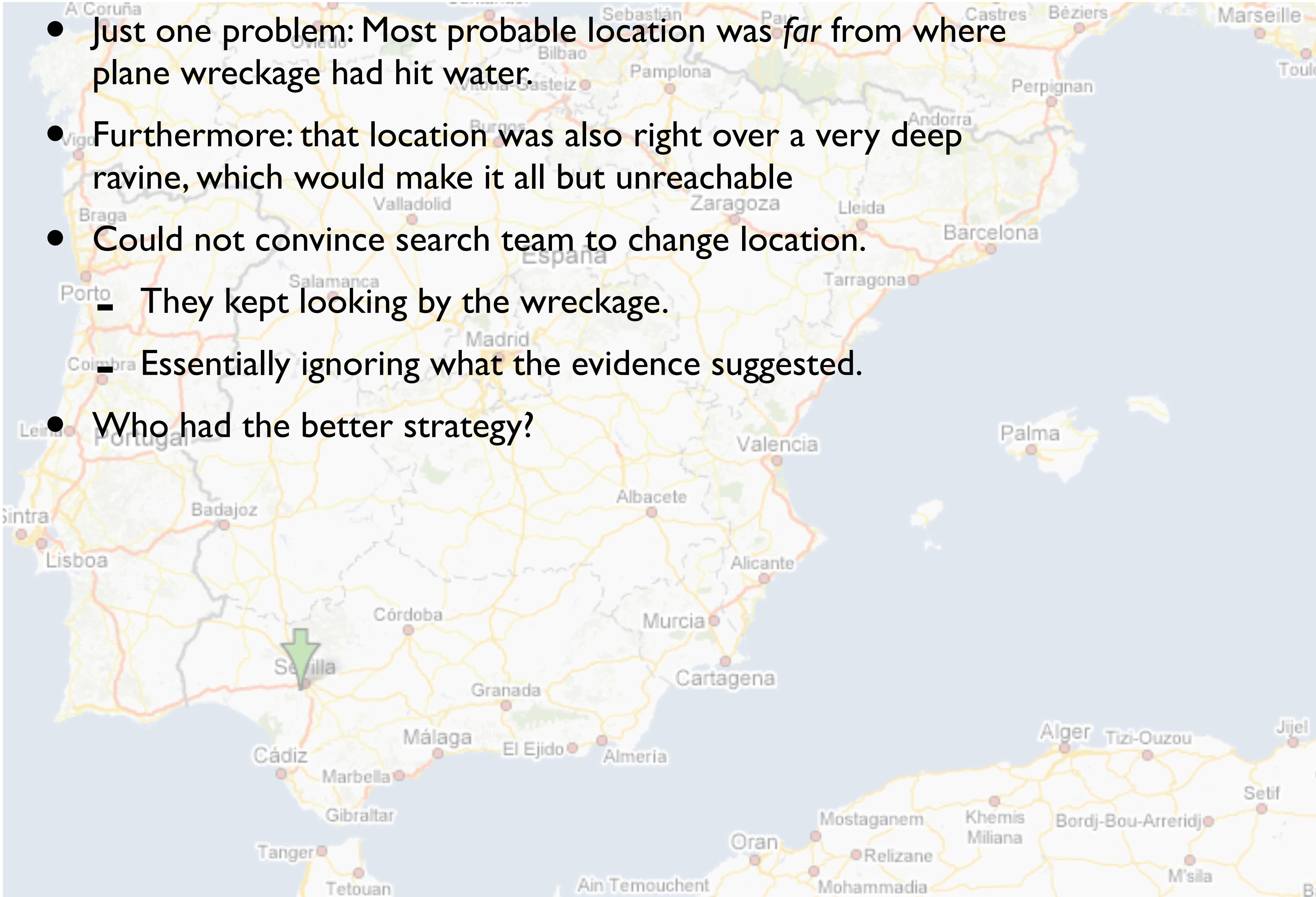
type of collision

prevailing wind direction

0, 1, or 2 parachutes open?

- Combine results into a probability map of location

John Craven and the missing H-Bomb



- Just one problem: Most probable location was *far* from where plane wreckage had hit water.
- Furthermore: that location was also right over a very deep ravine, which would make it all but unreachable
- Could not convince search team to change location.
  - They kept looking by the wreckage.
  - Essentially ignoring what the evidence suggested.
- Who had the better strategy?

John Craven and the missing H-Bomb

- Only one thing changed their minds
- A fisherman, reputed to be the best in the Mediterranean, had seen something go in the water …

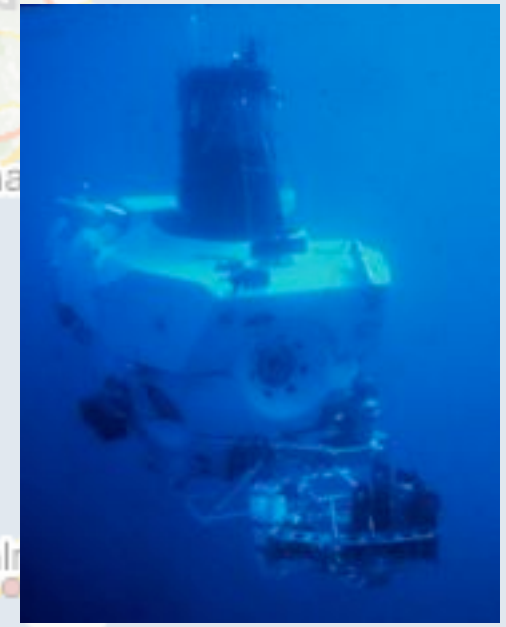
    … right at the highest probability location.

- With no other leads, they went to look there, but needed deep-diving submersibles.

John Craven and the missing H-Bomb

- Only one thing changed their minds
- A fisherman, reputed to be the best in the Mediterranean, had seen something go in the water …

  … right at the highest probability location.
- With no other leads, they went to look there, with deep-diving submersibles.
- Pres. Johnson wanted update
- Craven sent the latest prob. map and Johnson blew up.
- He called in Academics from Cornell and MIT, but they said he had the best plan.
- That same day, the crew of the Alvin, on its 10th dive, spotted a parachute attached to a cylindrical object, 2550 feet down, on a 70 degree slope.
- Right where Craven's probability map said it should be.



The process of probabilistic inference

1. *define* model of problem
2. *derive* posterior distributions and estimators
3. *estimate* parameters from data
4. *evaluate* model accuracy

type of collision

prevailing wind direction

0, 1, or 2 parachutes open?