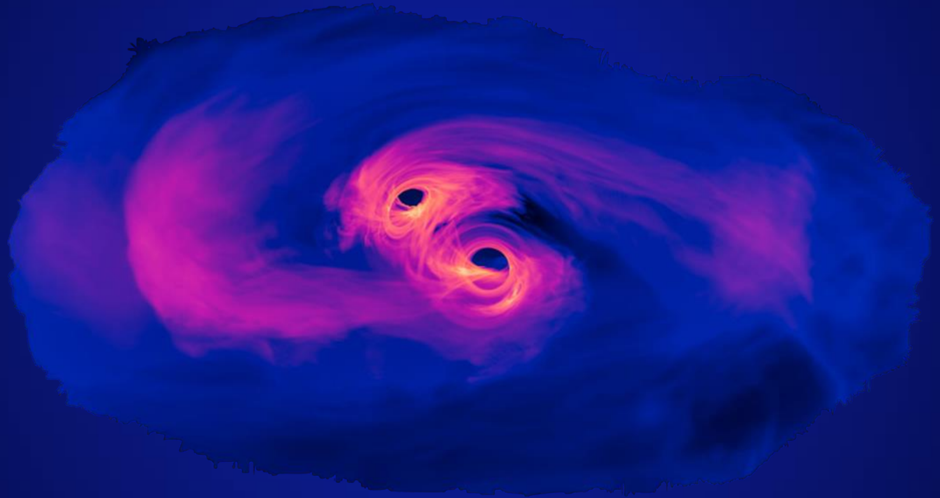
NSF HDR ML Challenge Problems

Josh Peterson on behalf of A3D3

Outline

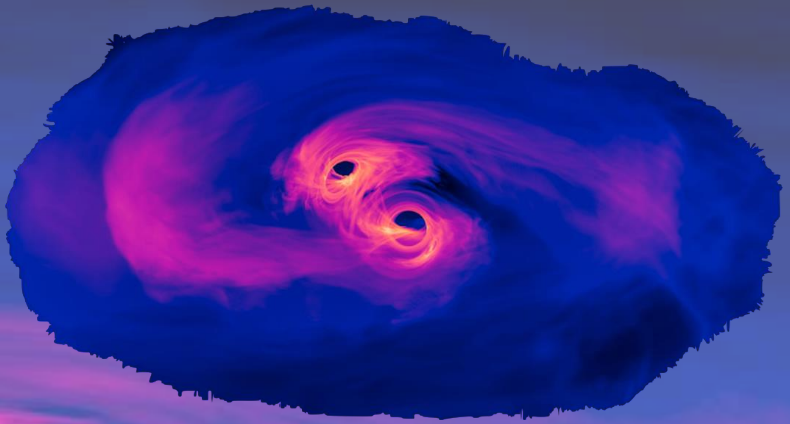
- Gravitational Waves
- Hybrid Butterflies
- Tidal Events

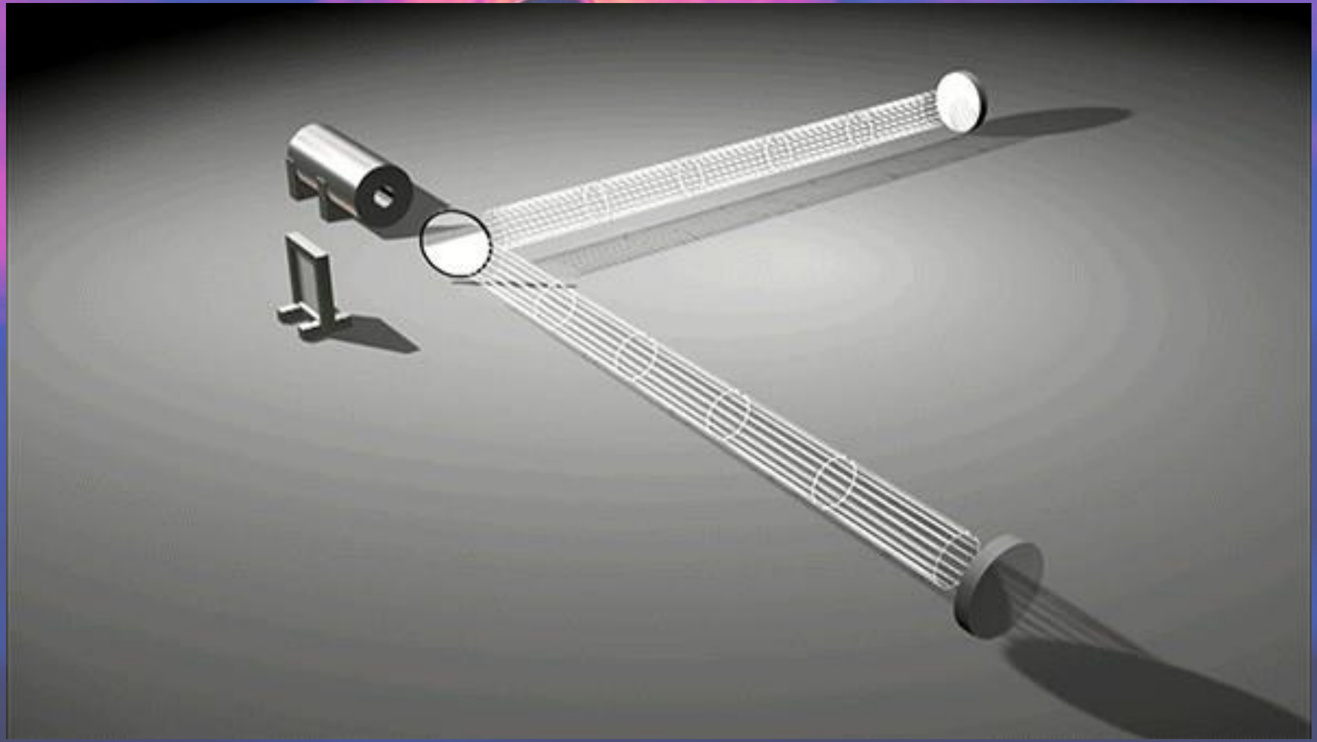


NSF HDR A3D3: Detecting Anomalous Gravitational Wave Signals

Gravitational Waves and Their Detection

Accelerating masses produce deformations in space time that we can detect via interferometry



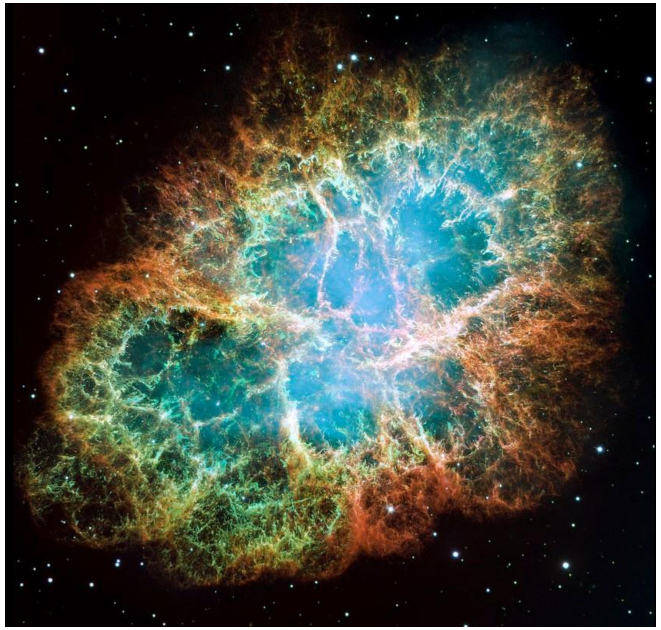


A SIGNAL WILL APPEAR IN AT LEAST TWO INTERFEROMETERS, WITH THE TIME DELAY BECAUSE OF THE DISTANCE BETWEEN THE DETECTORS

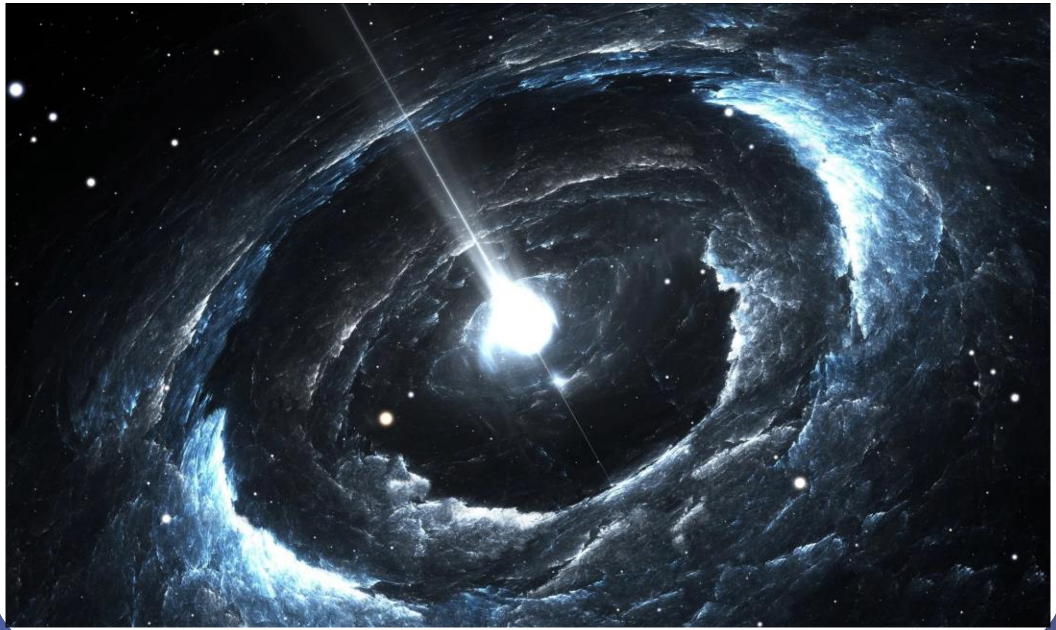
GWAK Anomalous Gravitational Wave Sources

Known "Unknowns" possible signal sources that are poorly modelled and therefore cannot be easily detected using the match filtering pipeline

Core-collapse supernova (CCSN)



Neutron Star Glitches



UNKNOWN "UNKNOWNS" NEW, UNEXPECTED GW SOURCES

WE REFER TO THEM AS ANOMALOUS AND AIM TO DEVELOP A SEMI-SUPERVISED APPROACH WHICH WOULD LET US TO DISCOVER ANOMALOUS SIGNALS WITHOUT EXPLICIT MODELLING

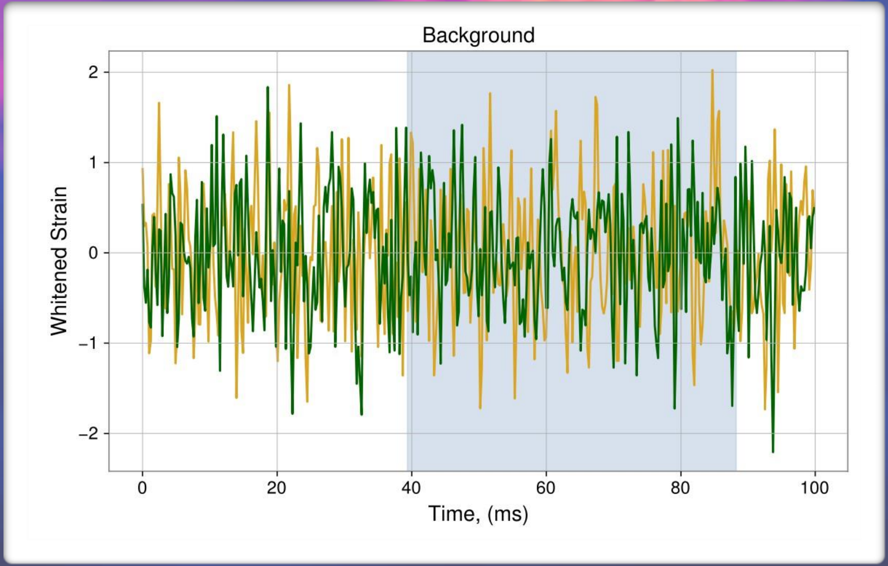
?

Background dataset

Sampling rate is 4096 Hz, meaning there are 4096 data points recorded every second

The data is divided into segments of 50 milliseconds each, which contains 200 data points (50 milliseconds * 4096 samples/second = 200 samples)

The dimension of the input data is (N, 200, 2), where N represents the number of data segments. The last dimension of 2 corresponds to the data streams from the two LIGO interferometers in Hanford, Washington, and Livingston, Louisiana

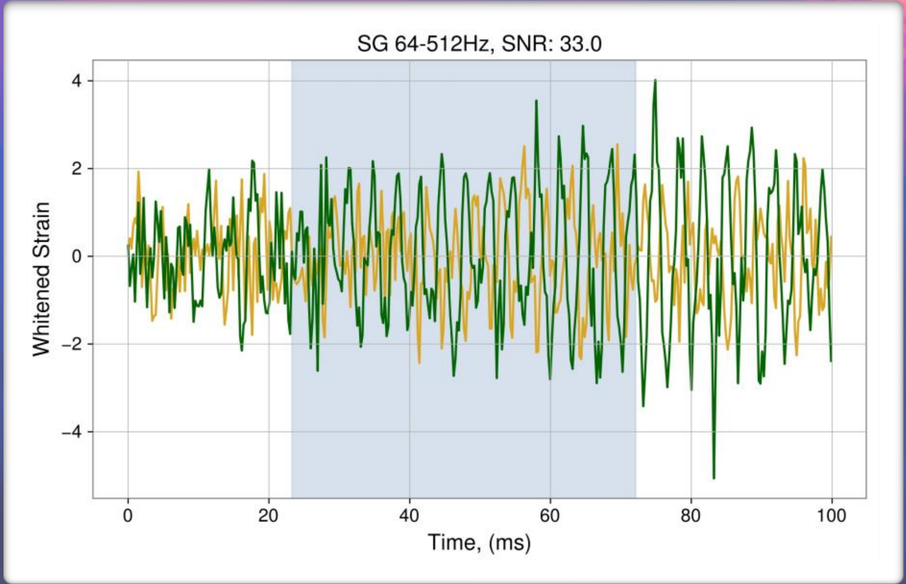


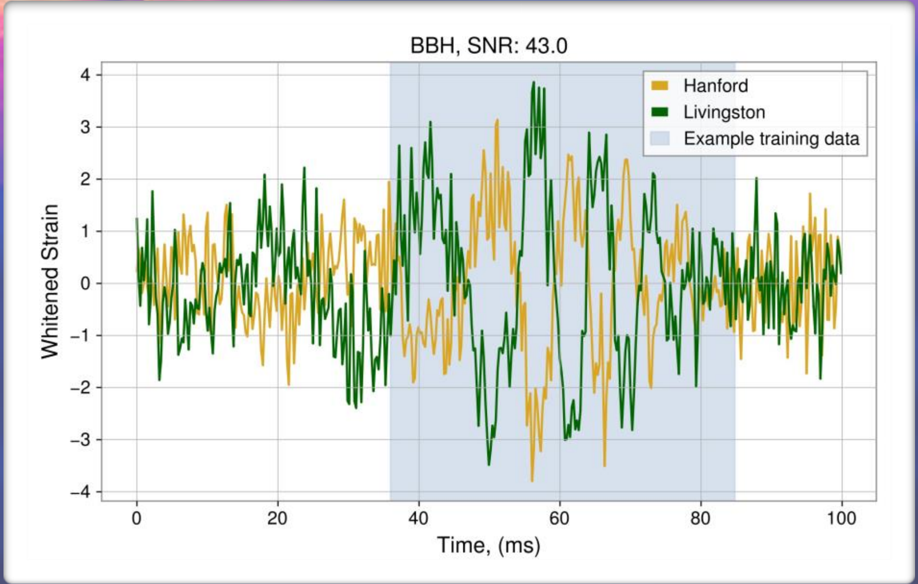
SIGNAL DATASETS

SAMPLING RATE IS 4096 HZ, MEANING THERE ARE 4096 DATA POINTS RECORDED EVERY SECOND

THE DATA IS DIVIDED INTO SEGMENTS OF 50 MILLISECONDS EACH, WHICH CONTAINS 200 DATA POINTS (50 MILLISECONDS * 4096 SAMPLES/SECOND = 200 SAMPLES)

THE DIMENSION OF THE INPUT DATA IS (N, 200, 2), WHERE N REPRESENTS THE NUMBER OF DATA SEGMENTS. THE LAST DIMENSION OF 2 CORRESPONDS TO THE DATA STREAMS FROM THE TWO LIGO INTERFEROMETERS IN HANFORD, WASHINGTON, AND LIVINGSTON, LOUISIANA

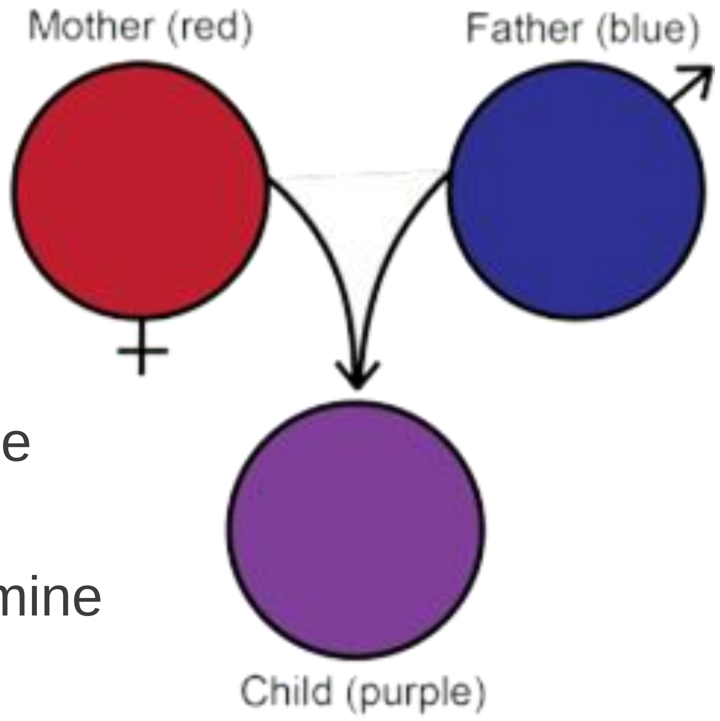




Anomaly Detection: Hybrid Butterflies

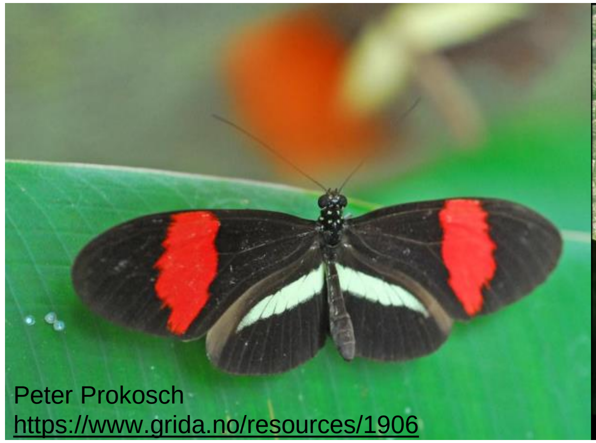
Hybrid Detection

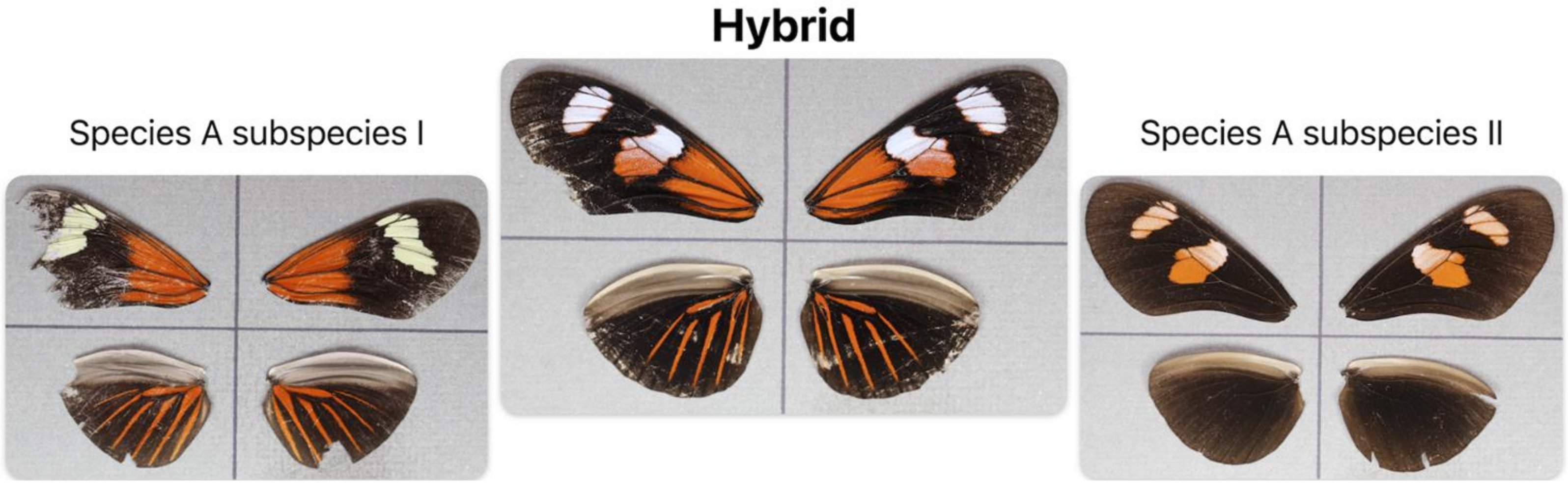
- Researchers have sought a means to detect hybrids since the creation of the field of taxonomy.
- Detecting hybrids would give taxonomists the ability to determine what constitutes a true *species* or *subspecies.*
- The question is ***how***?
    - *How* do we recognize a hybrid?
    - What does a hybrid look like?



Hybrid Detection: Butterflies

- Consider these two species:
- Hybridization may lead to a variety of resulting patterns.
- There are several [dominant] genes that control color pattern on wings.
  - Ex: red on hindwings is a dominant trait.
- Dominance: hybrids may look like one parent.
- In practice, identifying hybrids requires knowledge of their parent species/subspecies.




Hybrid
Species A subspecies I
Species A subspecies II

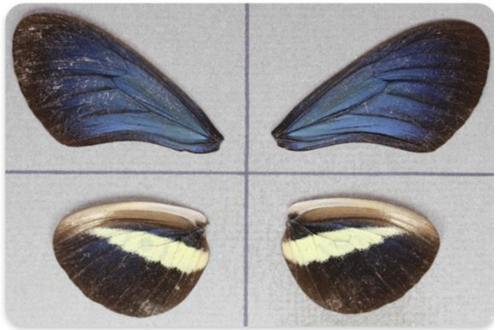
The Challenge: Find the Hybrids

- Among Species A & B, can your algorithm find…
  - Species A signal hybrids?
  - Species A non-signal hybrids?
  - Species B hybrids (mimics of Species A signal hybrids)?

Species A subspecies I

Species A subspecies II

Species A subspecies III



Species A subspecies IV

Species B subspecies II

Species B subspecies I

Detecting anomalous sea level rise events

iHARP
NSF HDR INSTITUTE

AS OUR OCEAN WARMS, SEA LEVEL RISES

We know seas are rising and we know why. The urgent questions are by how much and how quickly.

SEA LEVEL RISE: 1880 - 2017

Sea levels have risen about **8 inches** since the beginning of the 20th century. The ocean is projected to rise by as much as **3 feet or more** by the end of this century.

Earth's climate history shows there have been times when ice sheets rapidly changed and created multiple meters of sea level rise in a century. As Earth's ice sheets continue to change, a key question facing scientists now is: Could human-caused global warming be pushing us toward one of those times?

Making Better Predictions of Sea Level Rise

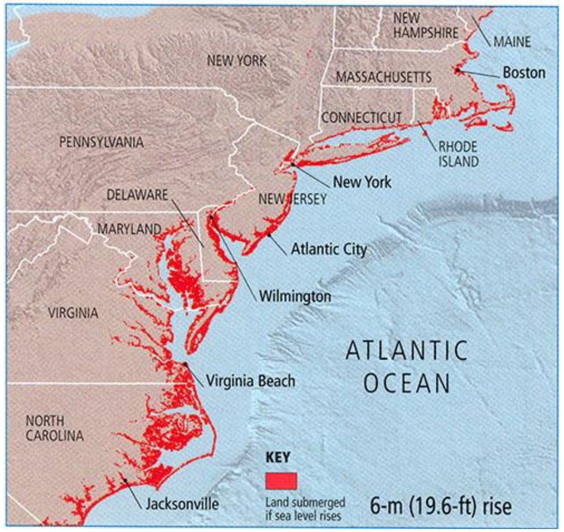
As the ocean rises, the ability to provide even more precise information about coastal sea level rise is crucial

More Damaging Flooding

Sea level rise will create a profound shift in coastal flooding over the next 30 years by causing tide and storm surge heights to increase and reach further inland. By 2050, “moderate” (typically damaging) flooding is expected to occur, on average, more than 10 times as often as it does today, and can be intensified by local factors.

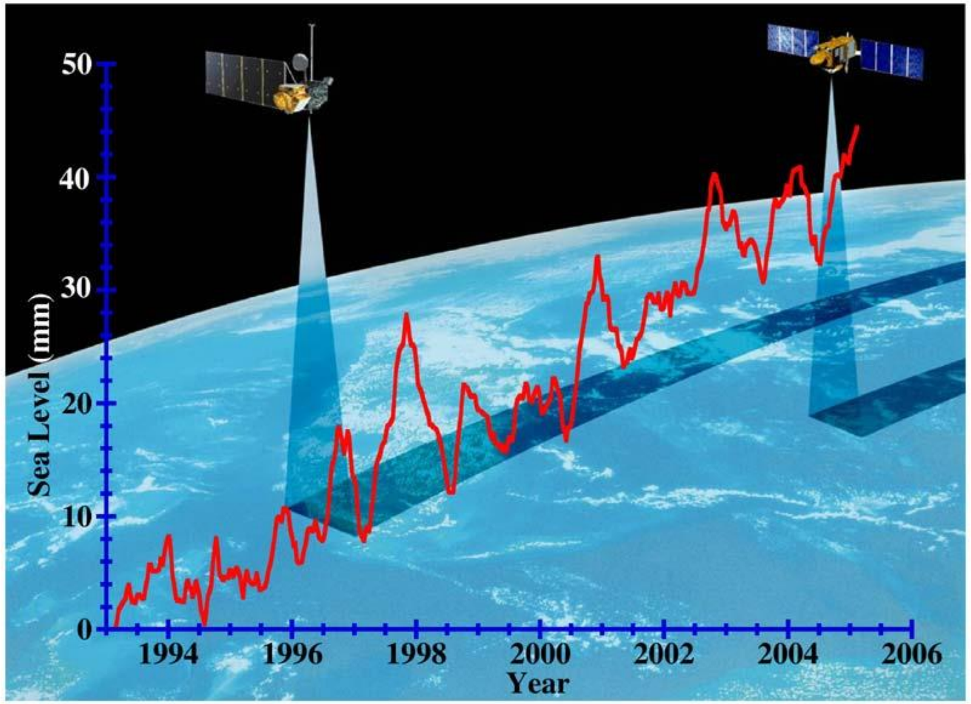
NORTHEAST COASTLINE

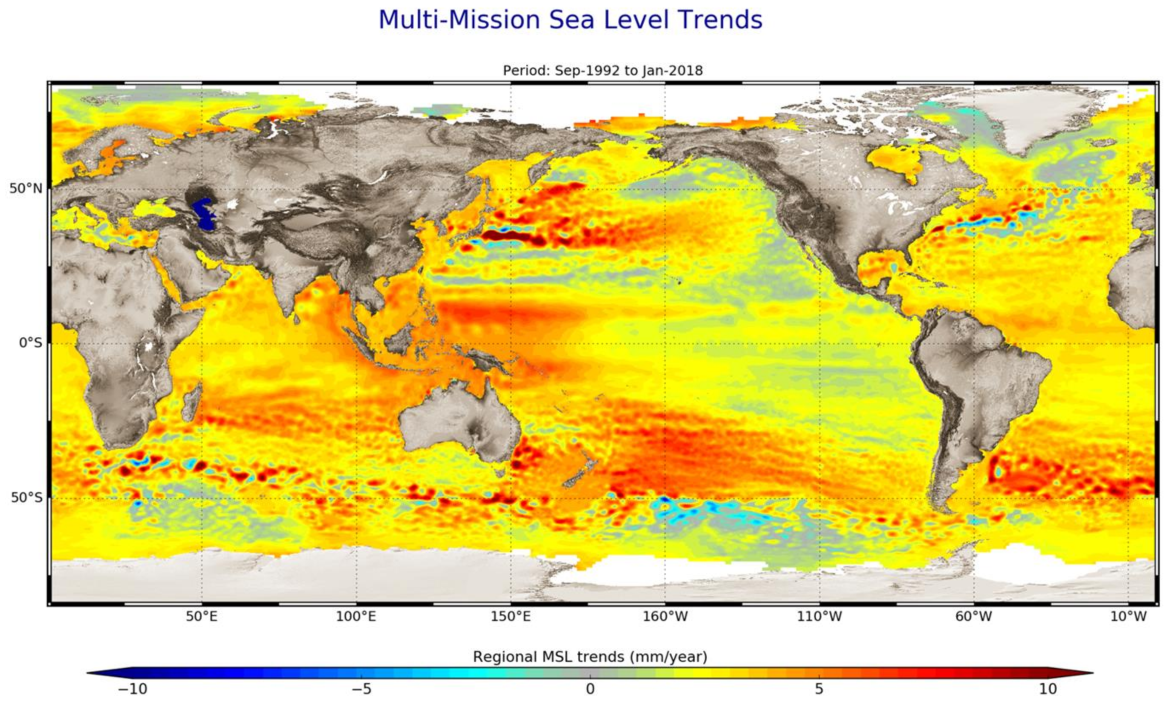
Most of New York City and Boston would be submerged if sea level were to rise by 6 m (19.6 ft).



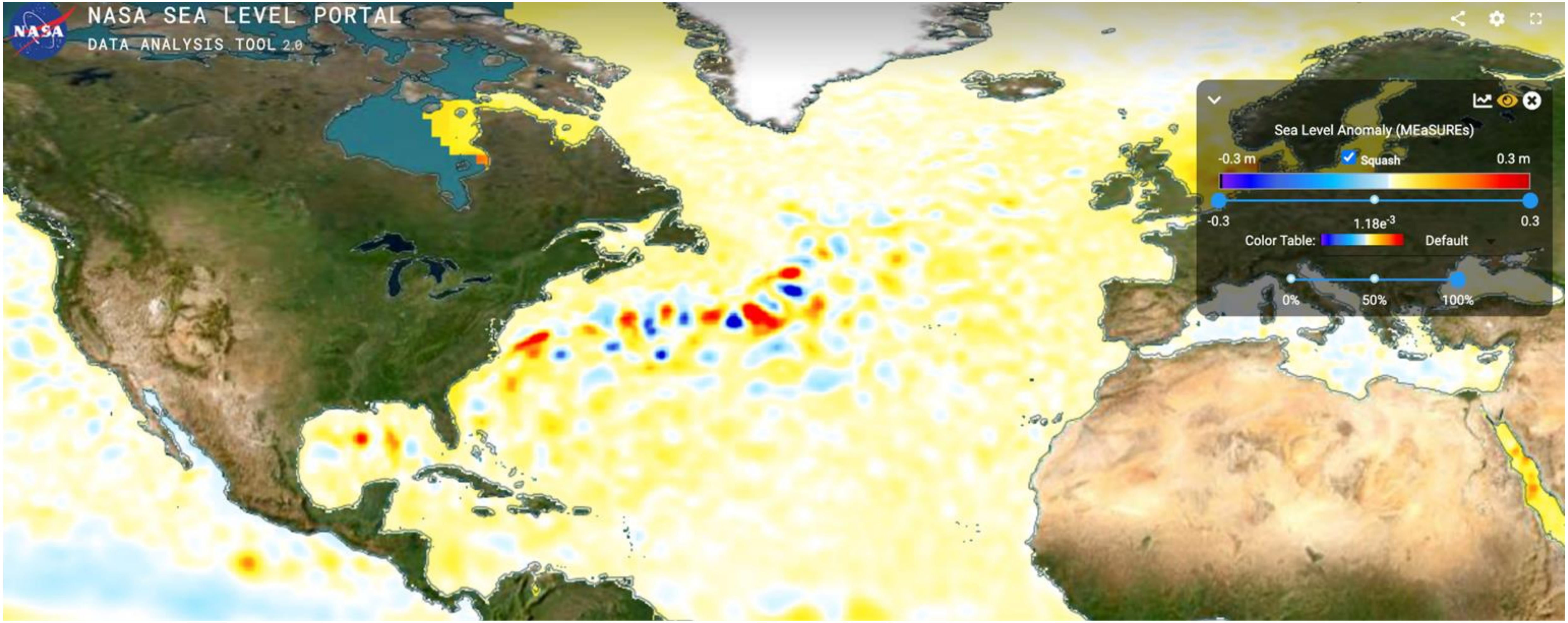
Continual Tracking

Continuously tracking how and why sea level is changing is an important part of informing plans for adaptation. Our ability to monitor and understand the individual factors that contribute to sea level rise allows us to track sea level changes in a way that has never before been possible (e.g., using satellites to track global ocean levels and ice sheet thickness). Ongoing and expanded monitoring will be critical as sea levels continue to rise.





Machine Learning Challenge: Detect anomalous flooding events from satellite sea level maps



Machine Learning Challenge: Detect anomalous flooding events from satellite sea level maps

- We provide daily satellite sea level anomaly data over the North Atlantic for the past 30 years
- We provide dates of anomalous flooding along US East coast stations for the past 30 years
- Challenge is to detect anomalous flooding events along the US East Coast with the maps of sea level over the North Atlantic