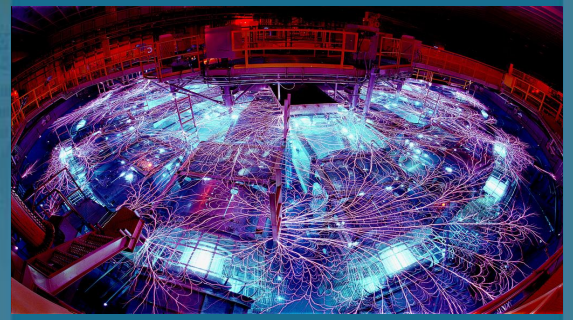
Statistics at Sandia National Laboratories



Presented by: Gabriel Huerta

Sandia Statistical Sciences Department

jghuert@sandia.gov

NISS virtual career fair about Statisticians in National Labs

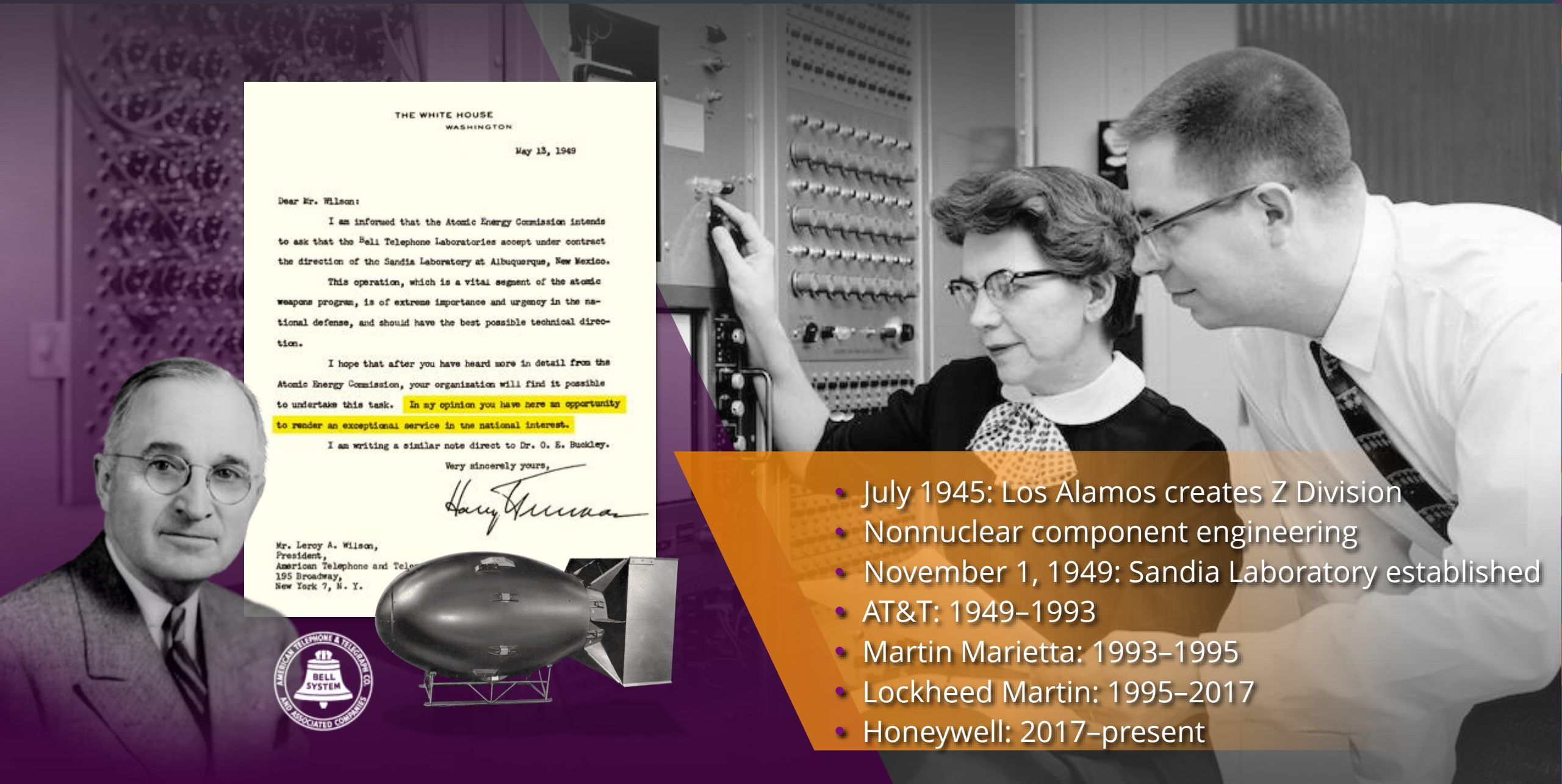
January 19, 2022

Slides 2-6, from SAND 2019-11784 PE

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2022-0568 HR

SANDIA'S HISTORY IS TRACED TO THE MANHATTAN PROJECT



THE WHITE HOUSE
WASHINGTON

May 13, 1949

Dear Mr. Wilson:

I am informed that the Atomic Energy Commission intends to ask that the Bell Telephone Laboratories accept under contract the direction of the Sandia Laboratory at Albuquerque, New Mexico.

This operation, which is a vital segment of the atomic weapons program, is of extreme importance and urgency in the national defense, and should have the best possible technical direction.

I hope that after you have heard more in detail from the Atomic Energy Commission, your organization will find it possible to undertake this task. In my opinion you have here an opportunity to render an exceptional service in the national interest.

I am writing a similar note direct to Dr. O. E. Buckley.

Very sincerely yours,

Harry Truman

Mr. Leroy A. Wilson,
President,
American Telephone and Tele
195 Broadway,
New York 7, N. Y.

- July 1945: Los Alamos creates Z Division
- Nonnuclear component engineering
- November 1, 1949: Sandia Laboratory established
- AT&T: 1949–1993
- Martin Marietta: 1993–1995
- Lockheed Martin: 1995–2017
- Honeywell: 2017–present

WE HAVE FACILITIES ACROSS THE NATION

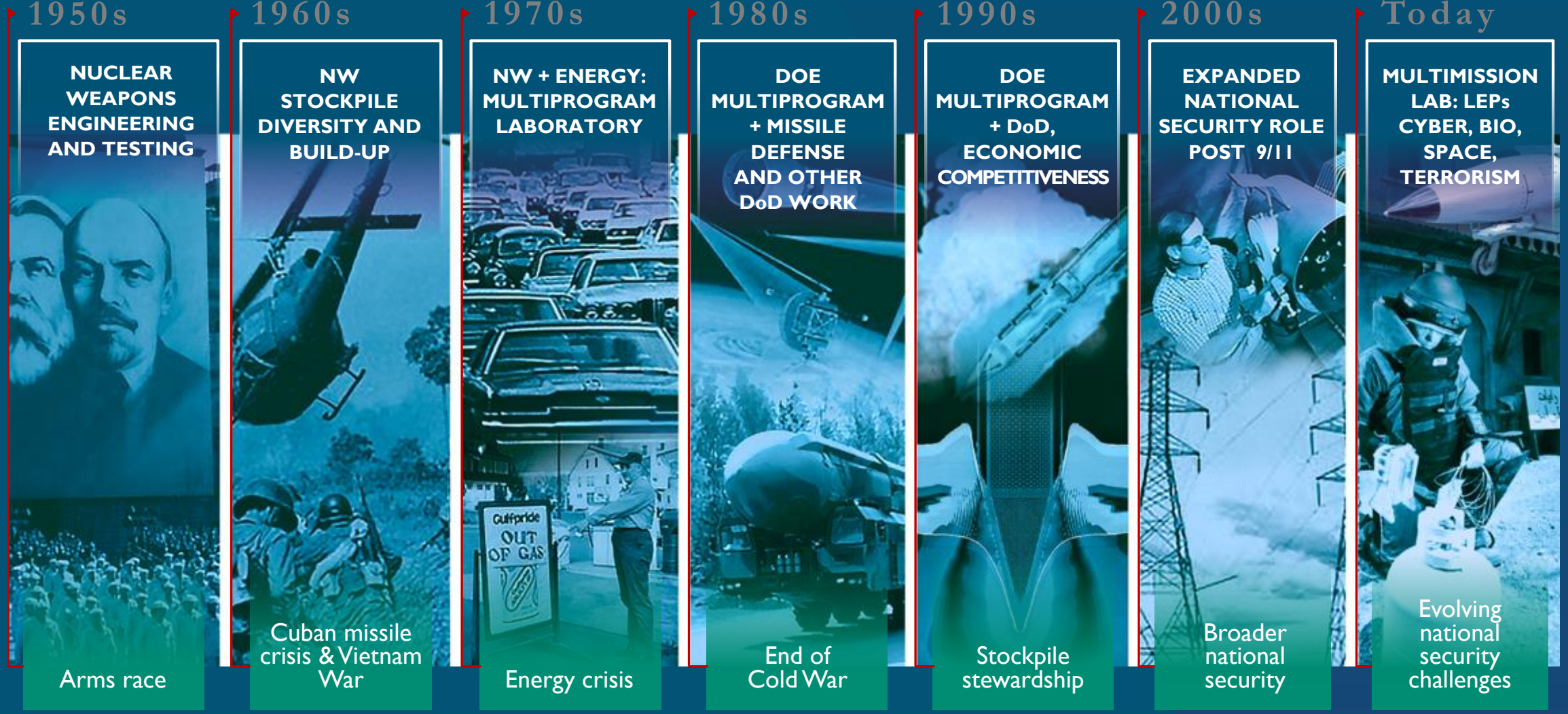
Activity locations

- Kauai, Hawaii
- Waste Isolation Pilot Plant, Carlsbad, New Mexico
- Pantex Plant, Amarillo, Texas
- Tonopah, Nevada

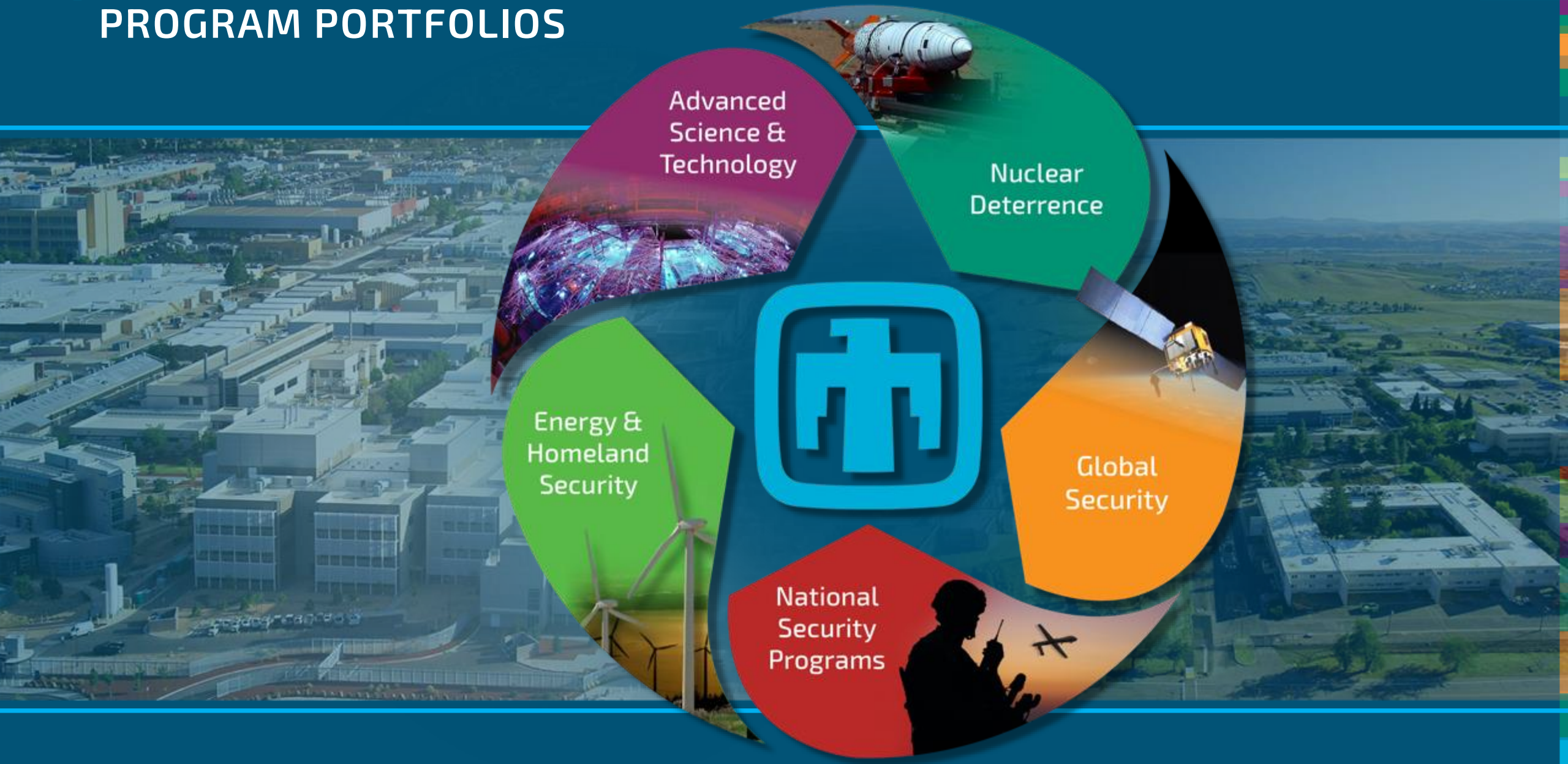
Main sites

- Albuquerque, New Mexico
- Livermore, California

OUR MULTIMISSION ROLE HAS EXPANDED OVER THE DECADES



SANDIA HAS FIVE MAJOR PROGRAM PORTFOLIOS



Nuclear Deterrence

- Sandia is responsible for the non-nuclear components of US stockpile.
  - Each weapon must work immediately if authorized by the President.
  - Each weapon must never go off if not authorized.
  - Weapons remain for decades in complex conditions and survive harsh environments.
- We want to say that we are X% confident that Y% of components will meet a requirement.
- We can't solely rely on experts state of knowledge.
- Many sources of uncertainty that cannot be straightforwardly quantified.
- Provide statistical modeling for weapons lifecycle decision support.

Global security – Nuclear forensics

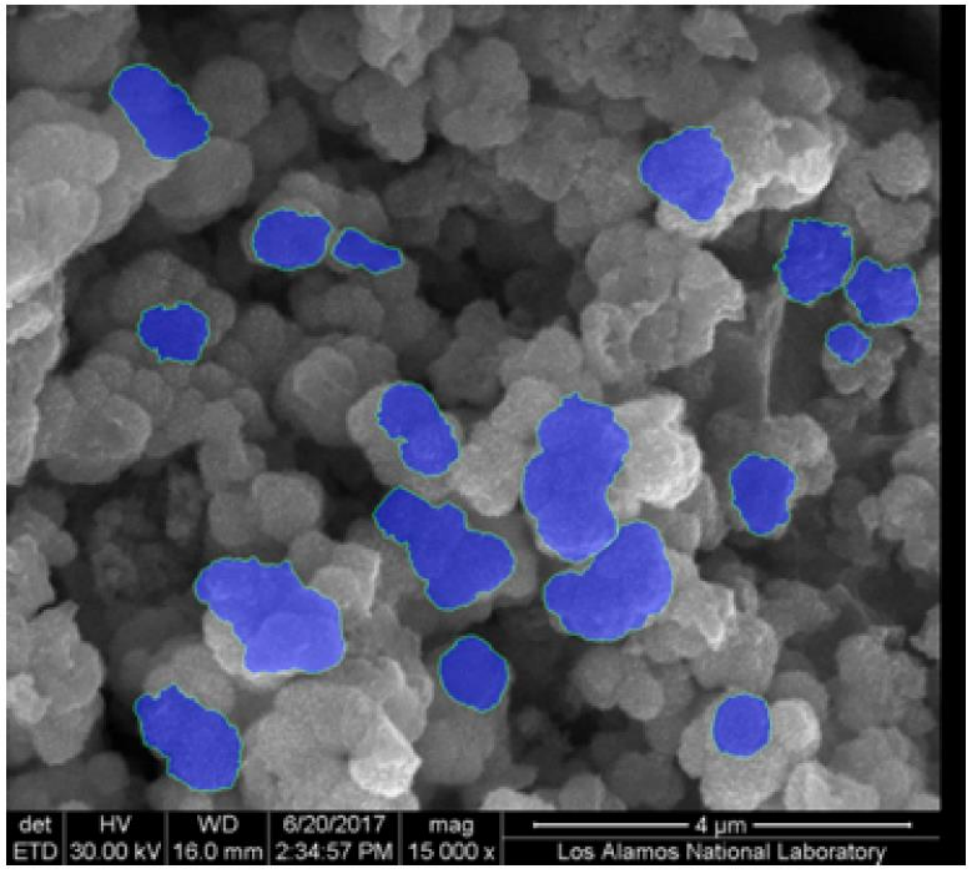


Figure 1. SEM image from MAMA software. The blue shapes are the particles from which measurements such as vector area and pixel area are calculated.

Since the 1960's, Sandia has been involved in treaty verification

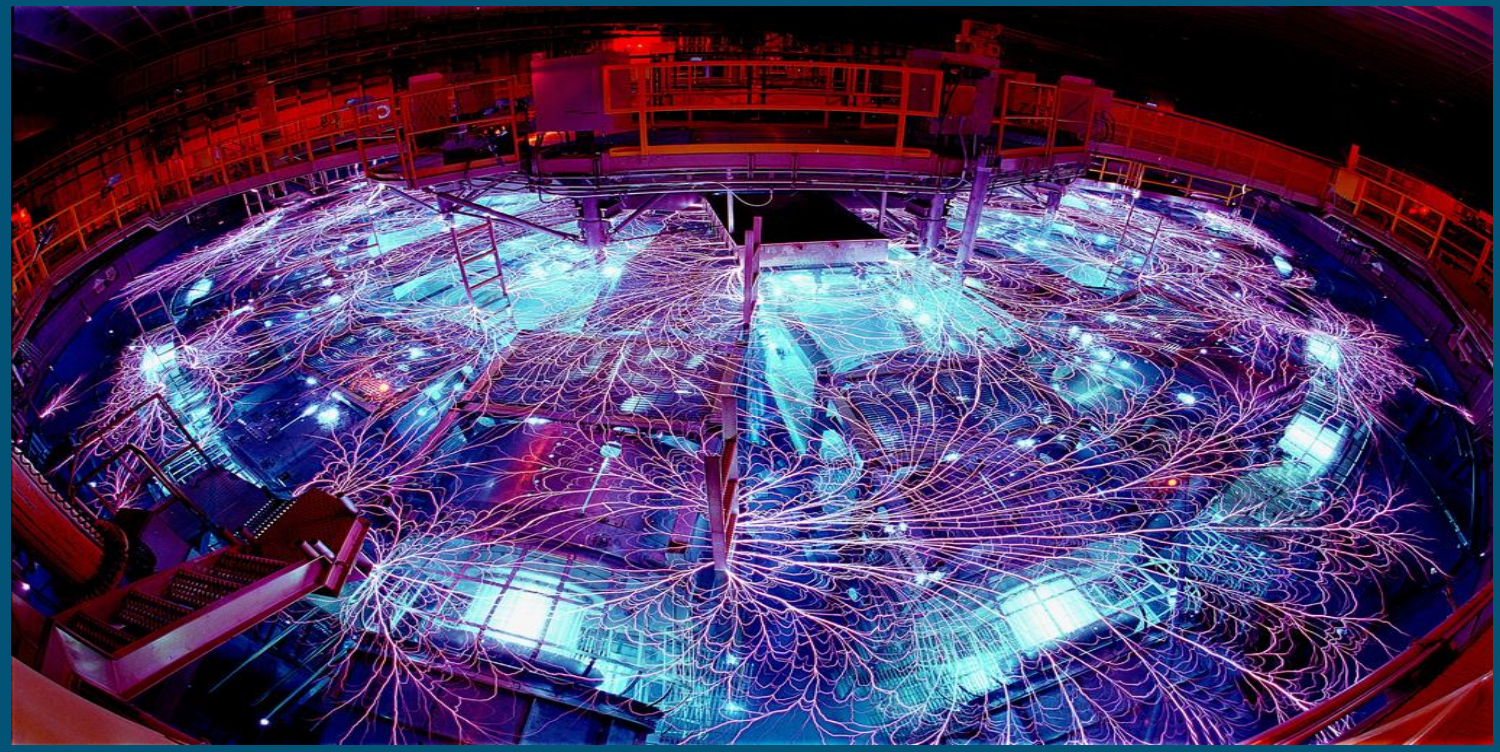
- Sensors are deployed on Earth, in orbit, and seismically to detect foreign nuclear activities
- Sandia statisticians play an important role in developing the detection and classification algorithms for the processing of sensing signatures.

Nuclear Forensics.

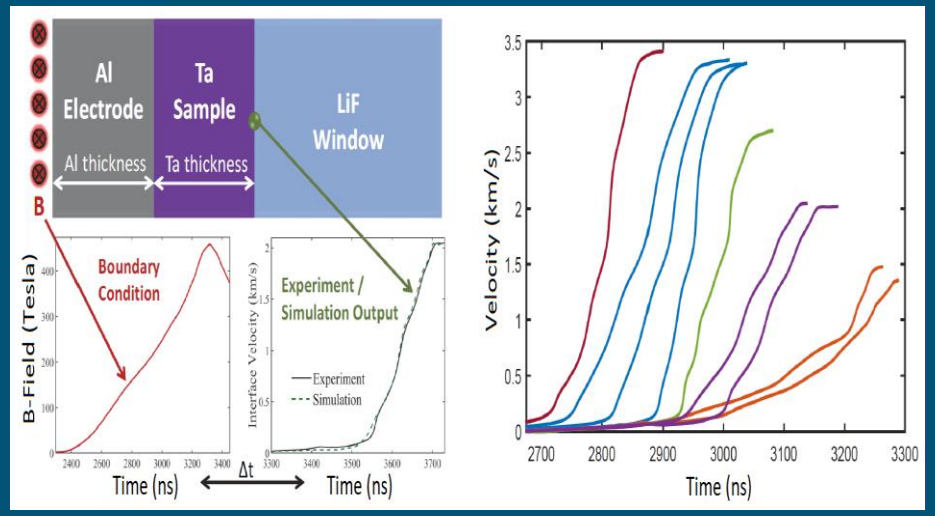
- Suppose a piece of interdicted material is found. We would like to know where it came from.
- Sandia statisticians are developing inverse prediction techniques for chemometrics, based on measured physical, chemical, and scanning electron microscope (SEM) features of materials, to detect signature characteristics associated with the pedigree of the material based on a statistically designed experiment.
- Sample SEM images of particles. What conditions created it?

Advanced Science and Technology – Z- machine

- Sandia's Z machine is Earth's most powerful pulsed-power facility and X-ray generator. Z compresses energy in time and space to achieve extreme power and intensity found nowhere else on Earth. The Z accelerator is an integral part of Sandia's Pulsed Power Program.
- Many different applications across the laboratory.

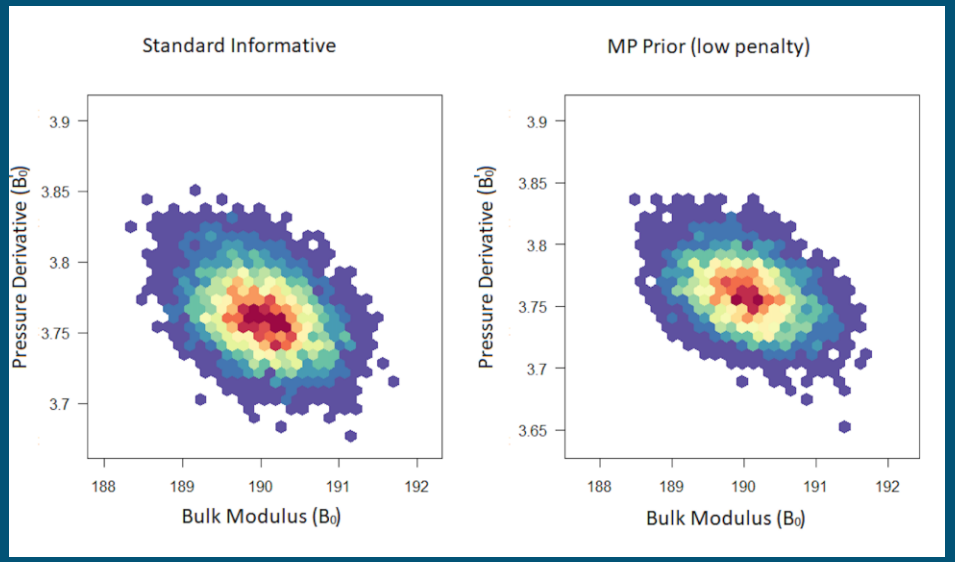


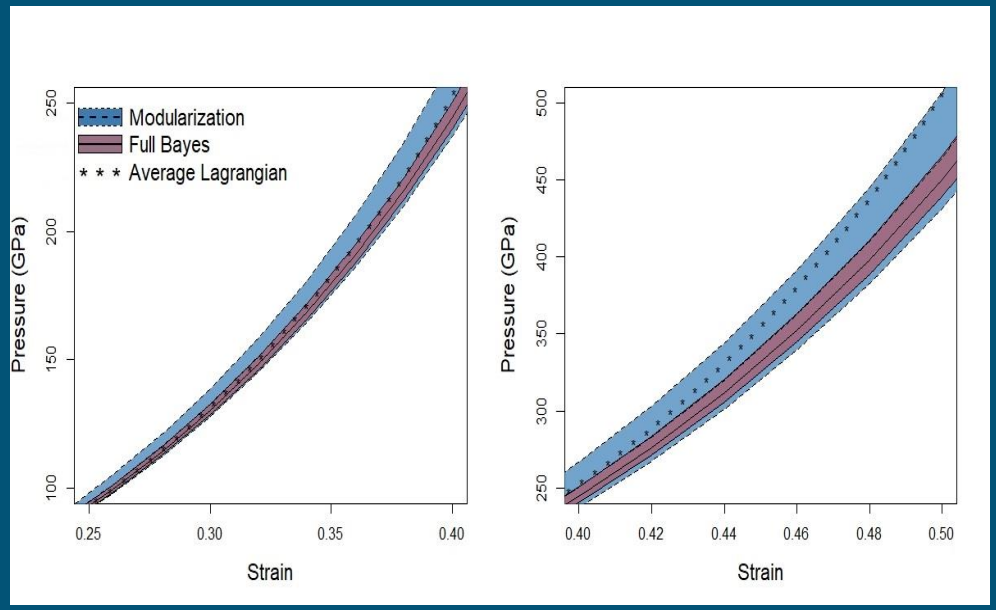
Analysis with the Z Machine



Material sciences

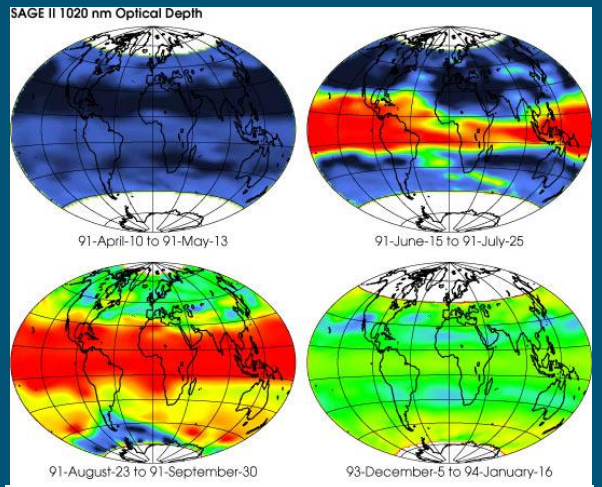
- Statisticians had been involved in the development of ***uncertainty quantification*** methods to predict important physical characteristics in material sciences by coupling experimental data and computer simulations.





Statistical research for detecting climate pathways

Data Fusion



https://earthobservatory.nasa.gov/images/1510/global-effects-of-mount-pinatubo

Pulling together the multi-resolution data from varying sources to create a near-global picture of relevant processes.

Spatial Statistics Methods

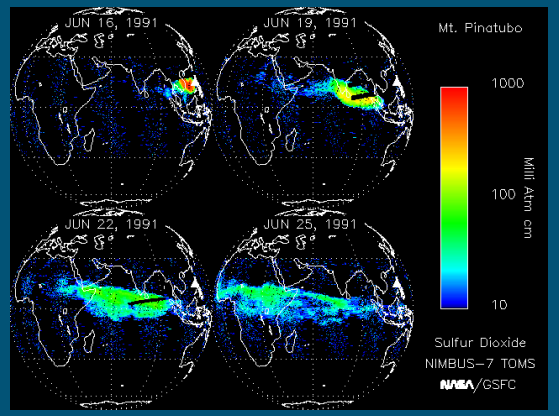
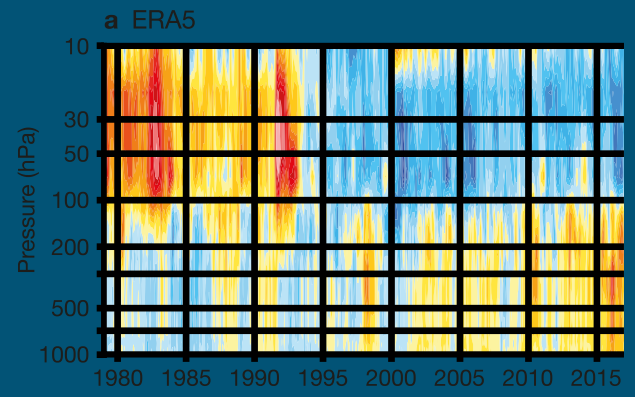


Figure: Sulfur Dioxide from TOMS on NOAA Nimbus-7 immediately after eruption

Quantifying the spatio-temporal evolution of aerosols and climate impacts with interpretable approaches.

Deep Learning



https://www.ecmwf.int/en/newsletter/159/meteorology/global-reanalysis-goodbye-era-interim-hello-era5

Deep learning will add computational efficiently and account for complex dynamics that lead to long-term climate impacts.

Change Point Methods

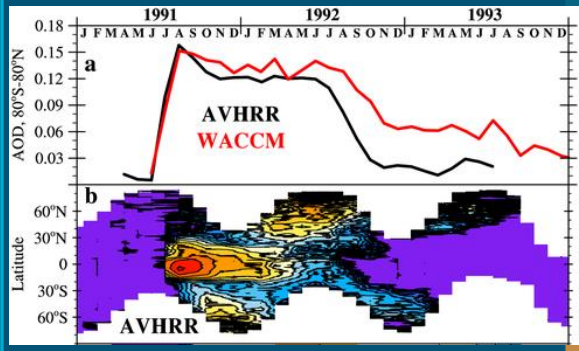


Figure: Stratospheric Aerosol Optical Depth from AVHRR instrument (Mills et al., 2016).

Change Point Detection methods can identify the underlying fundamental shifts in climate processes due to significant events

Spatio-temporal statistics

Dynamic spatio-temporal models are intuitive and interpretable ways to represent complex, dependent, physical processes that change across space and time simultaneously.

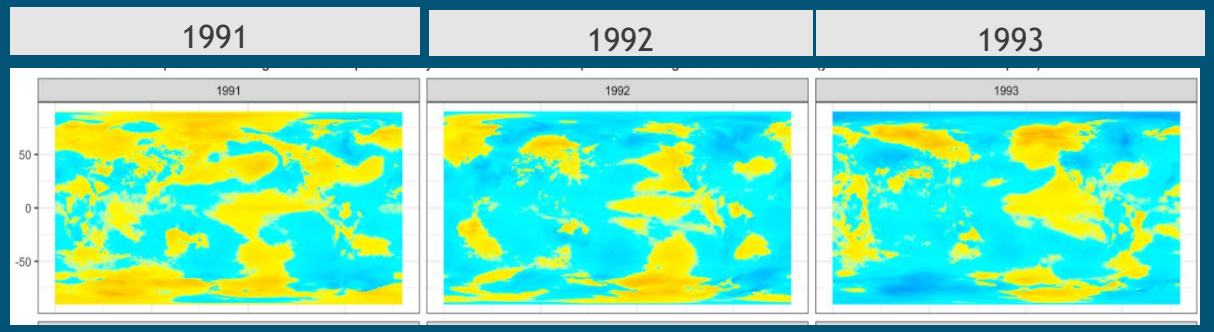


Figure: Yearly spatial temperature anomalies from ERA-5 reanalysis in1992 and 1993 anomalies show clear effects from the 1991 eruption.

- Nearest Neighbor Gaussian Process (NNGP) priors can make dynamic models for large data more efficient.

Datta et al. (2016)

$$\mathbf{y}_t(\mathbf{s}) = \mathbf{X}_t(\mathbf{s})'\boldsymbol{\beta}_t + \mathbf{u}_t(s) + \boldsymbol{\epsilon}_t(\mathbf{s}),\ \boldsymbol{\epsilon}_t(\mathbf{s}) \overset{iid}{\sim} N(0, D)$$

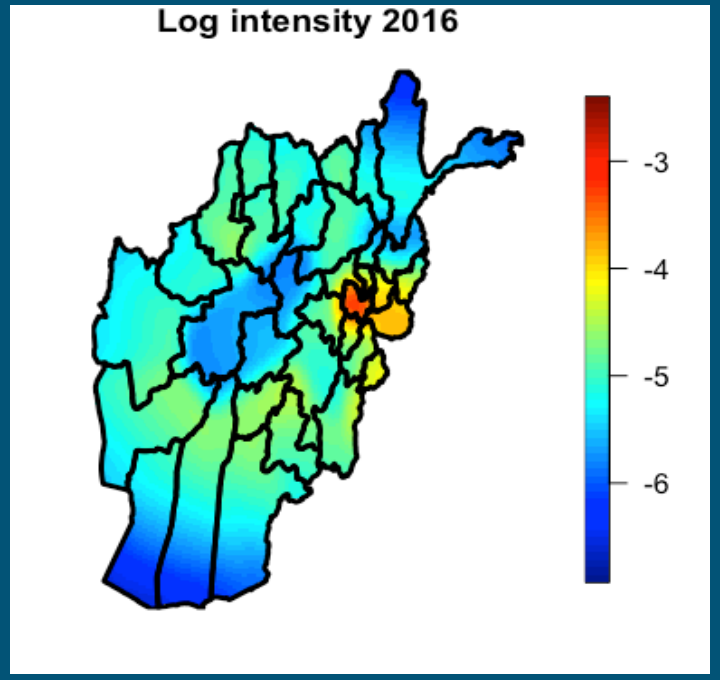
$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t,\ \boldsymbol{\eta}_t \overset{iid}{\sim} N(0, \boldsymbol{\Sigma}_\eta),\ \boldsymbol{\beta}_0 \sim N(\mathbf{m}_0, \boldsymbol{\Sigma}_0)$$

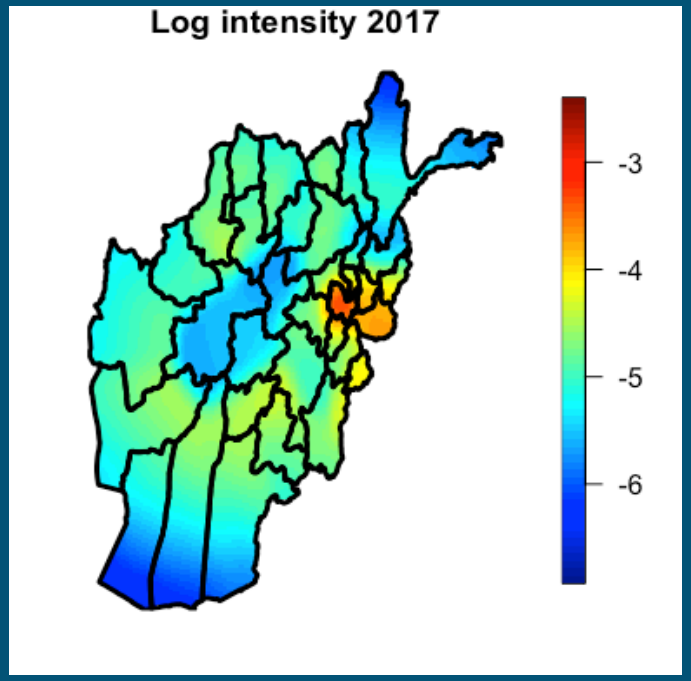
$$\mathbf{u}_t(\mathbf{s}) = \mathbf{u}_{t-1}(\mathbf{s}) + \mathbf{w}_t(\mathbf{s}),\ \mathbf{w}_t(\mathbf{s}) \overset{ind}{\sim} NNGP(\mathbf{0}, \tilde{\mathbf{C}}(\cdot,\cdot\,|\,\boldsymbol{\theta}_t))\ .$$

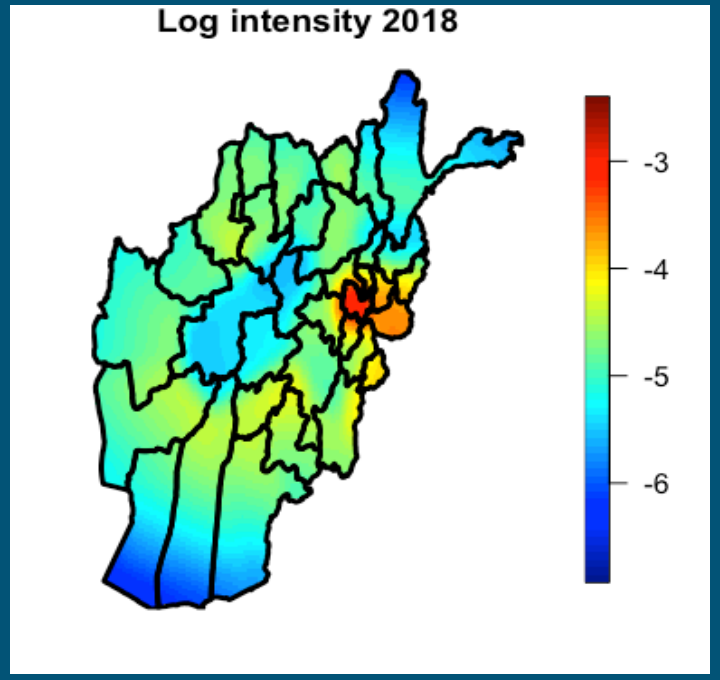
- Baseline trends (e.g. seasonal, ENSO effects) can be captured in $X_t(s)'\beta_t$.
- Assuming $X_t(s)'\beta_t$ captures our general space-time trends well, $\epsilon_t(s)$ and $w_t(s)$ capture the measurement error and spatial variability, respectively
  - $\epsilon_t(s)$ will also capture temporal trends outside of $X_t(s)'\beta_t$
- θ_t allows for a temporally varying spatial distribution

National security and intelligence science

- Statisticians are involved in developing space-time marked point processes to study *terrorism* events globally.
- Plots of fitted log-intensities 2016-2018. for Afghanistan. Covariates: *Pop. Density, terrain, voting, no. of languages.*
- Extreme spatiotemporal intensities (for attacks that produce at least 20 casualties), integrated over year.







University collaborations and teaching

- Some of the universities Sandia statisticians have ongoing research collaborations with:
  - University of Illinois Urbana-Champaign
  - Ohio State University
  - Brigham Young University
  - Florida State University
  - University of Washington
  - University of New Mexico

- **Sandia supports many students from universities all across the USA through internships (summer and year round).**

- Teaching activities within Sandia
  - *Data Analysis Techniques, Quantification of Margins and Uncertainties, Intro to R, Intro to Measurement Uncertainty*, others

Areas of expertise and career path

Some of the areas of emphasis of the Department are

- Experimental design for physical and computer of experiments.
- V&V/UQ studies (validation and verification/uncertainty quantification)
- Bayesian analysis
- Functional data analysis
- Statistical signal processing and time series analysis
- Spatial and spatio-temporal data analysis
- Computational statistics
- Statistical machine learning
- Measurement error, repeatability, and reproducibility plans and analysis
- Data visualization

Staff has varied backgrounds (MsC-Ph.D) and career paths. *(student/postdoc → staff; staff -> management; academia --> staff)*

https://sandia.jobs/. Postings for student internships 680151, 680149

https://www.sandia.gov/careers/career-possibilities/students-and-postdocs/internships-co-ops/

Summary

- Statisticians at Sandia support a variety of critical mission areas,
  - Nuclear Deterrence and indirect support of nuclear deterrence
  - Global Security
  - National Security
  - Advanced Science and Technology
  - University Research Collaborations, Internships, and Teaching

Come visit us at www.sandia.gov and http://www.sandia.gov/statistics/

Acknowledgement: Special thanks to my manager Adele Doser and co-workers Lekha Patel, Lyndsay Shand and Adah Zhang who provided material for this presentation.