LA-UR-19-20140
Approved for public release; distribution is unlimited.

Title: Machine Learning

Author(s): Vesselinov, Velimir Valentinov

Intended for: Web

Issued: 2019-01-10

Machine Learning

EES-16 Briefing to ALD-CELS

Velimir V Vesselinov (“monty”)

January 9, 2019

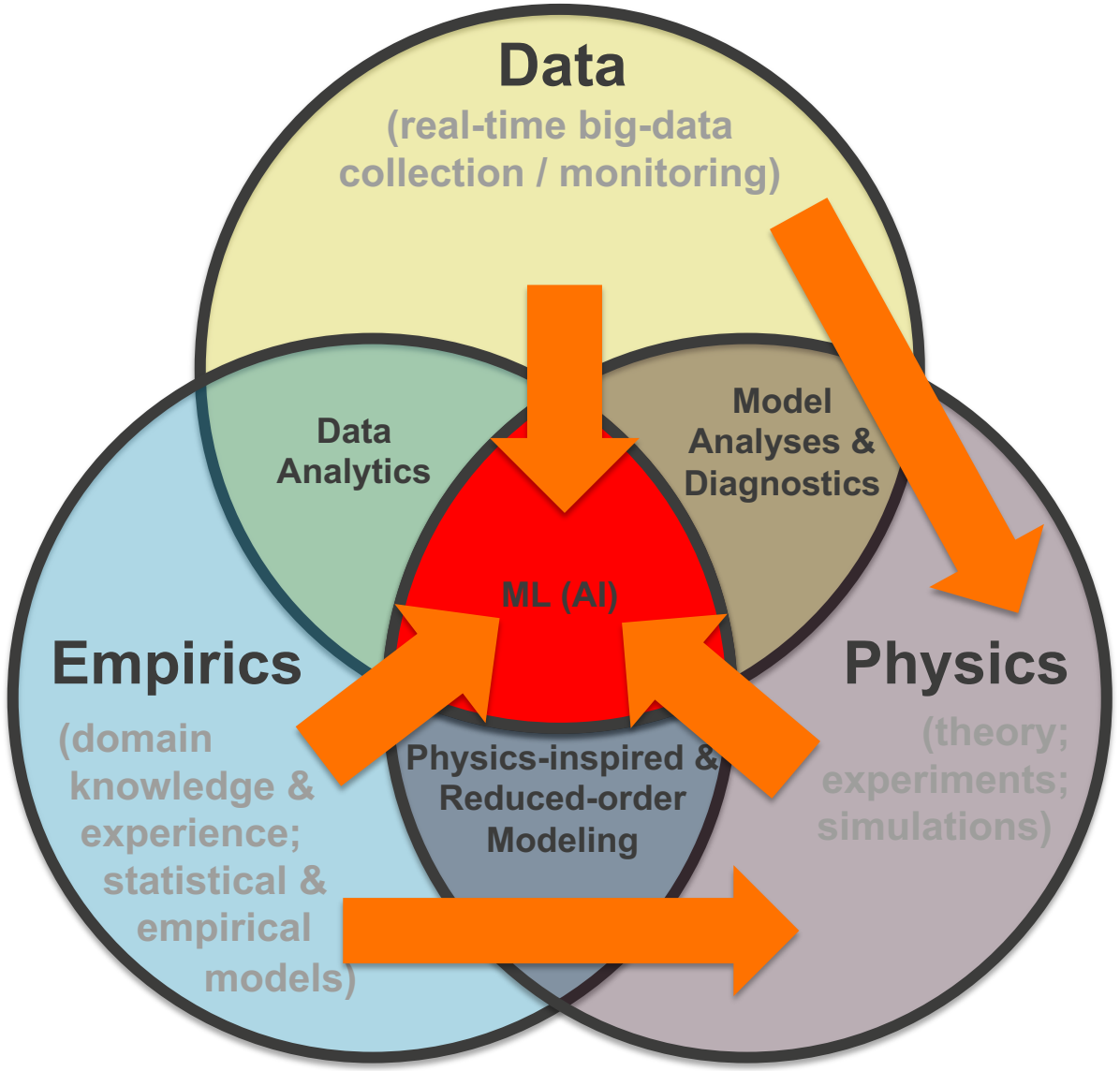
Machine Learning (ML): Overview

(EES-16 & cross-divisional efforts)

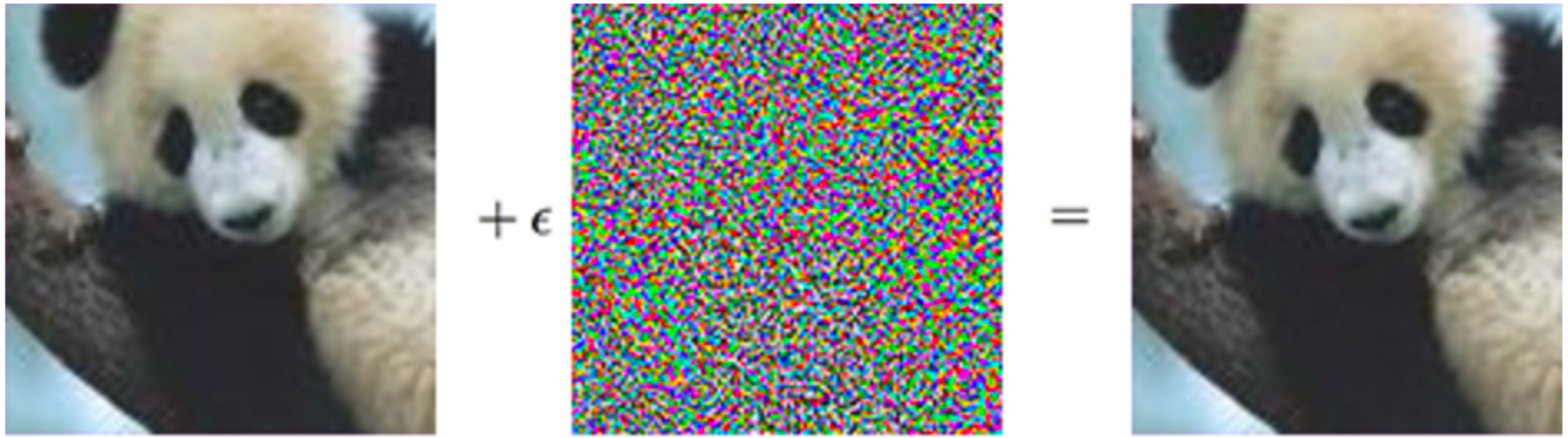
- **Why Machine Learning?**
  - Why ML-Informed Physics?
  - Why Physics-Informed ML?
- **LANL ML strategy related to National & Energy Security**
- **Examples:**
  - ML-Informed Physics for climate modeling and induced seismicity (LDRD-DR)
  - Physics-Informed ML for blind source separation (LDRD-DR)
  - ML forecasting production from tight rock reservoirs (Chevron)

Why Machine Learning (AI)?

- **ML-informed Physics**
  (Physics informed by Data & Empirics)
  - **Feature extraction**
    Knowledge from noise
    Blind Source Seperation
    Anomaly/Signal detection
    (Data → Physics**)**
  - **Knowledge synthesis**
    Reinforcement learning
    (Empirics → Physics**)**
- **[Data/Empirics/Physics]-Informed ML (AI)**
  (bringing together Data, Empirics & Physics; Data Science; Artificial Intelligence)



Adversarial Examples



Goodflellow et al. 2014, Explaining and Harnessing Adversarial Examples

EES-16 involvement in LANL ML efforts

- **Patent:** Alexandrov, B.S., **Vesselinov, V.V.**, Alexandrov, L.B., Stanev, V., Iliev, F.L., Source identification by non-negative matrix factorization combined with semi-supervised clustering, US20180060758A1, https://patents.google.com/patent/US20180060758A1/en
- 3 LRDR-DR projects focused on ML (co-PI's from EES-16)
  - "Tensor Networks: Robust Unsupervised Machine Learning for BigData Analytics" (Alexandrov, **Vesselinov**, Djidjev)
  - "Enabling Predictive Scale-Bridging Simulations through Active Learning" (Germann, **Viswanathan**)
  - "Advancing brittle fracture prediction using dynamic graphs" (Srinivasan, **Viswanathan**)
- DOE funded projects related to Fossil Energy and ML (NRAP, CCSI, FE20-AOP, FE30-AOP)
- 1 CRADA (with Chevron) on ML for oil production (PI: **Vesselinov)**
- ADAPD (Advanced Data Analytics for Proliferation Detection)
- Cross-divisional efforts related to national, global and energy security

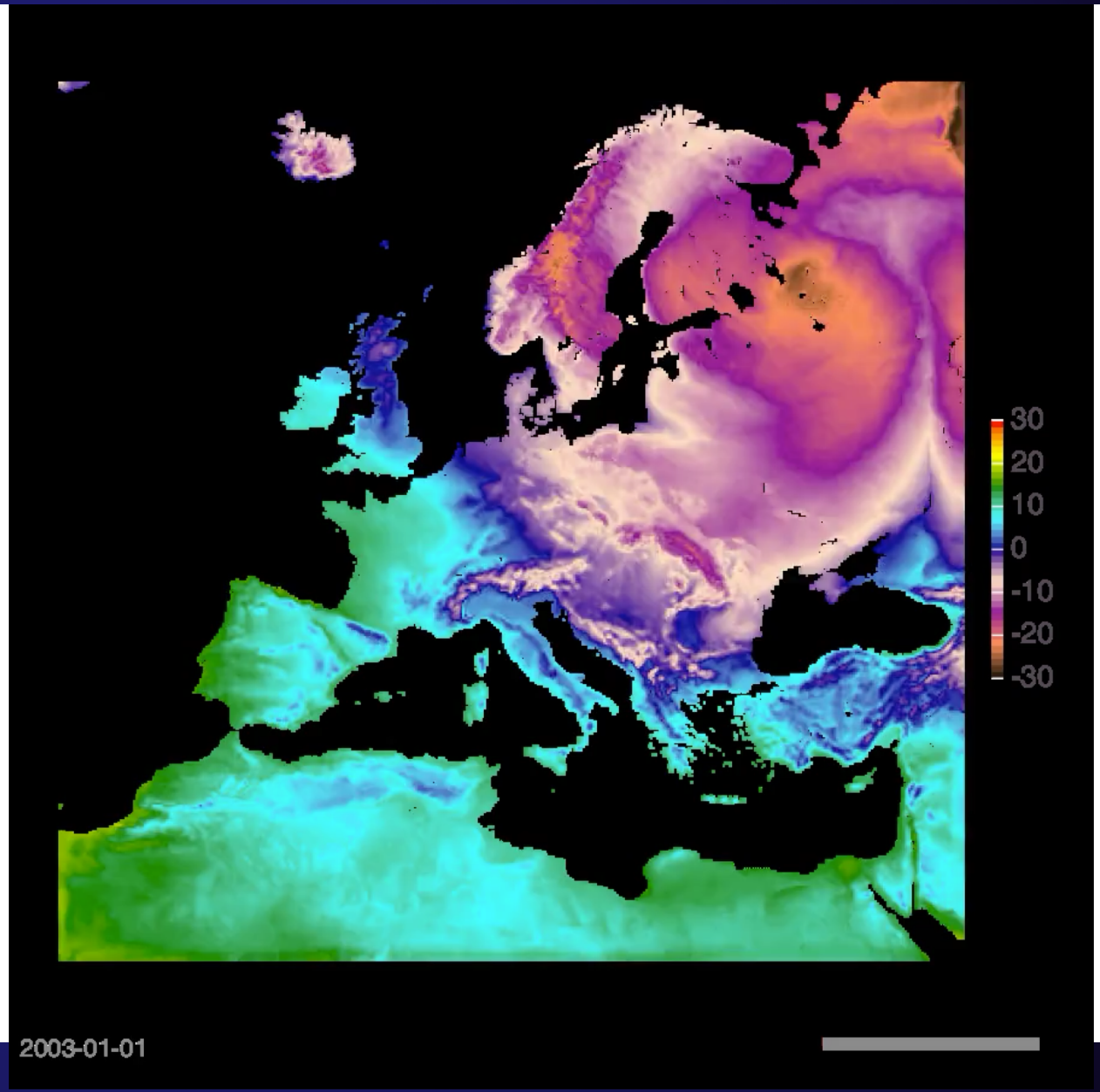
LANL ML strategy related to National & Energy Security

- **Exploit Synergies in subsurface portfolios for Energy Security & National Security**:
  - Rapid prediction of fluid/gas flow in complex, uncertain and dynamic subsurface systems
  - Extraction of more relevant information from the subsurface
- **Machine Learning and Fossil Energy**
  - New DOE Initiative**: double recovery in 10 years from unconventional reservoirs** (currently, recovery is about 5-10%)
  - Multi National Lab collaboration
  - Employ Physics-Informed ML and ML-informed Physics analyses

ML-informed Physics: Climate modeling

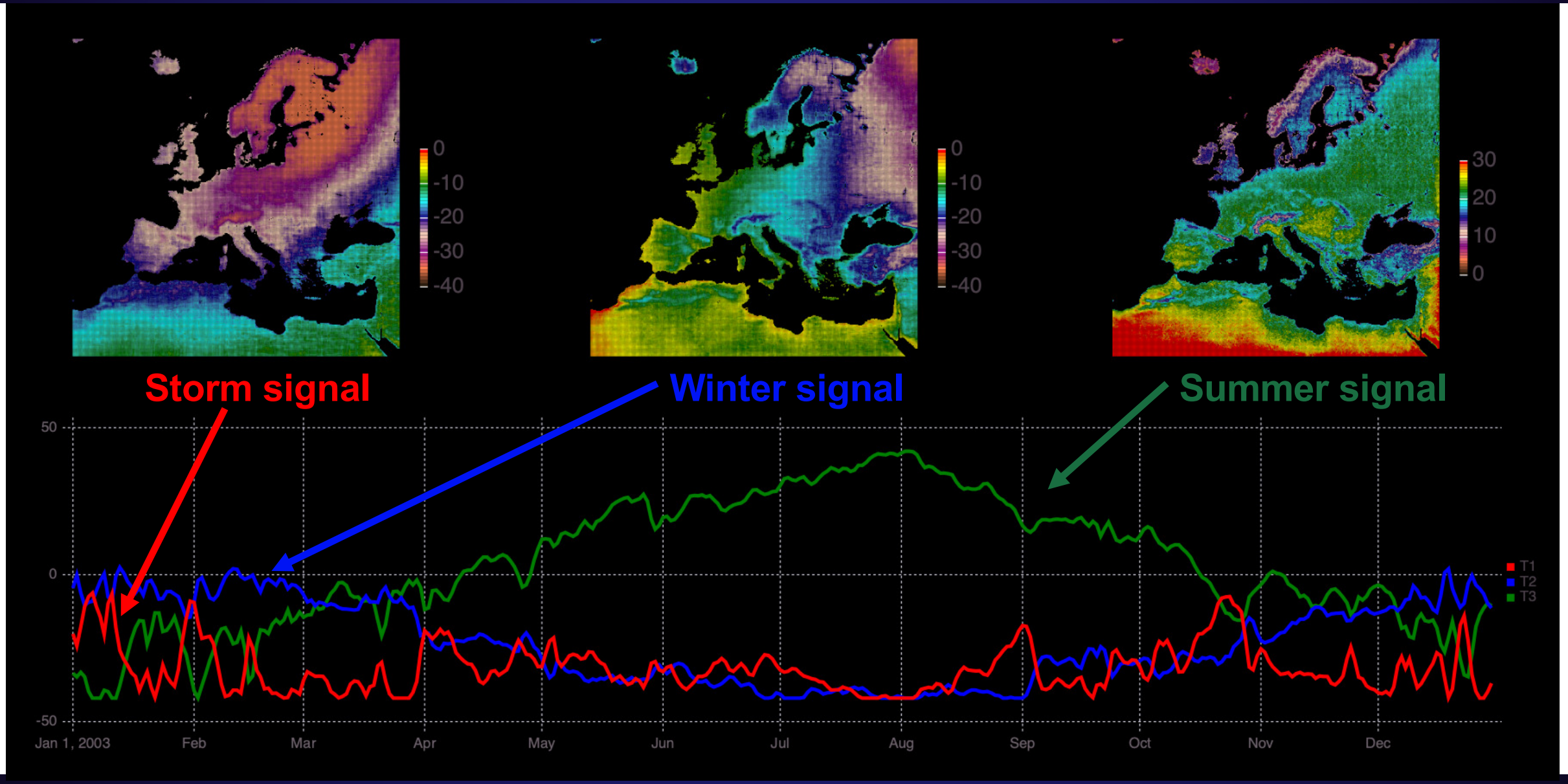
Vesselinov—LDRD-DR

- Collaborative work with Helmholtz Institute, Germany
- Model simulations of fluctuations in the air temperature [^{O}C]
- ML applied to extract features representing governing processes



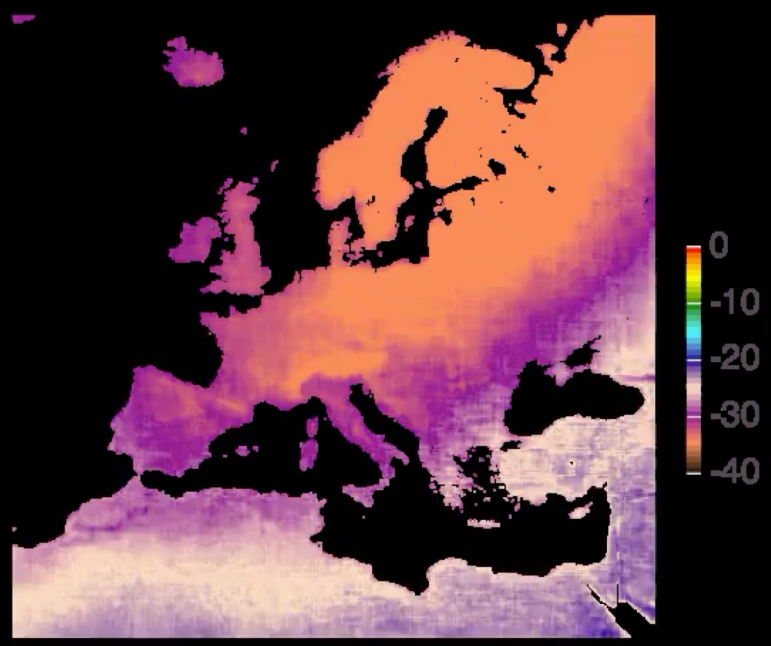
ML-informed Physics: Climate modeling

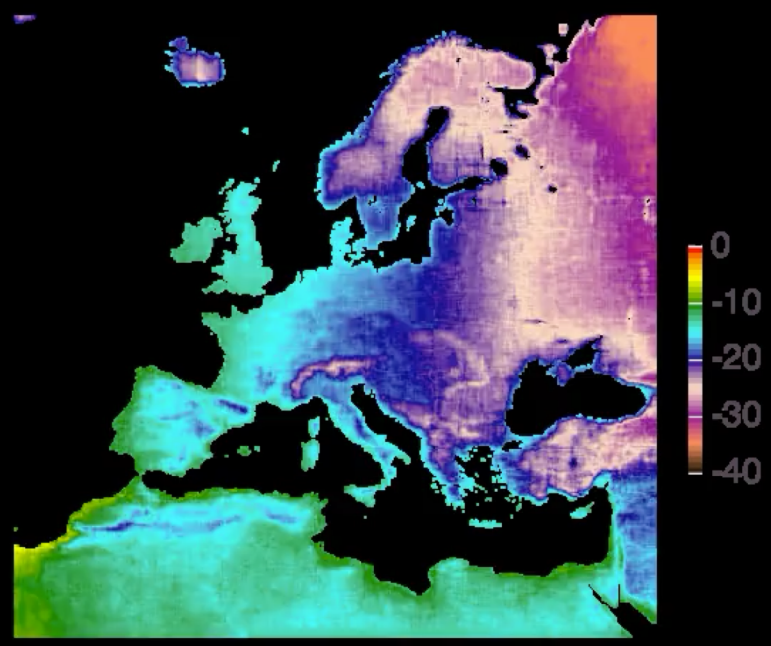
Vesselinov —LDRD-DR

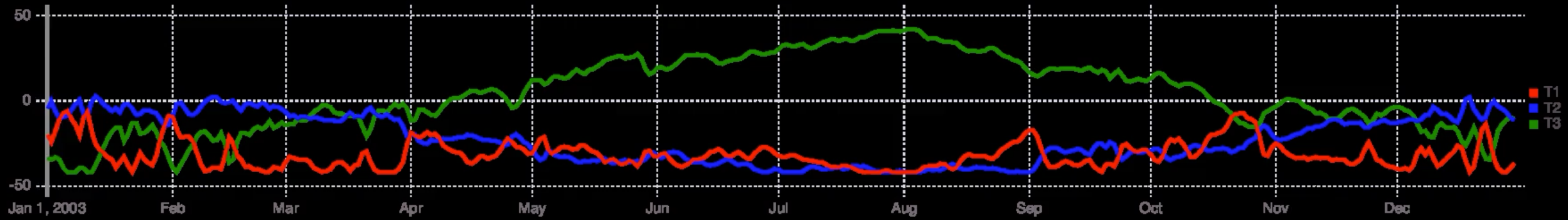


ML-informed Physics: Climate modeling

Vesselinov —LDRD-DR



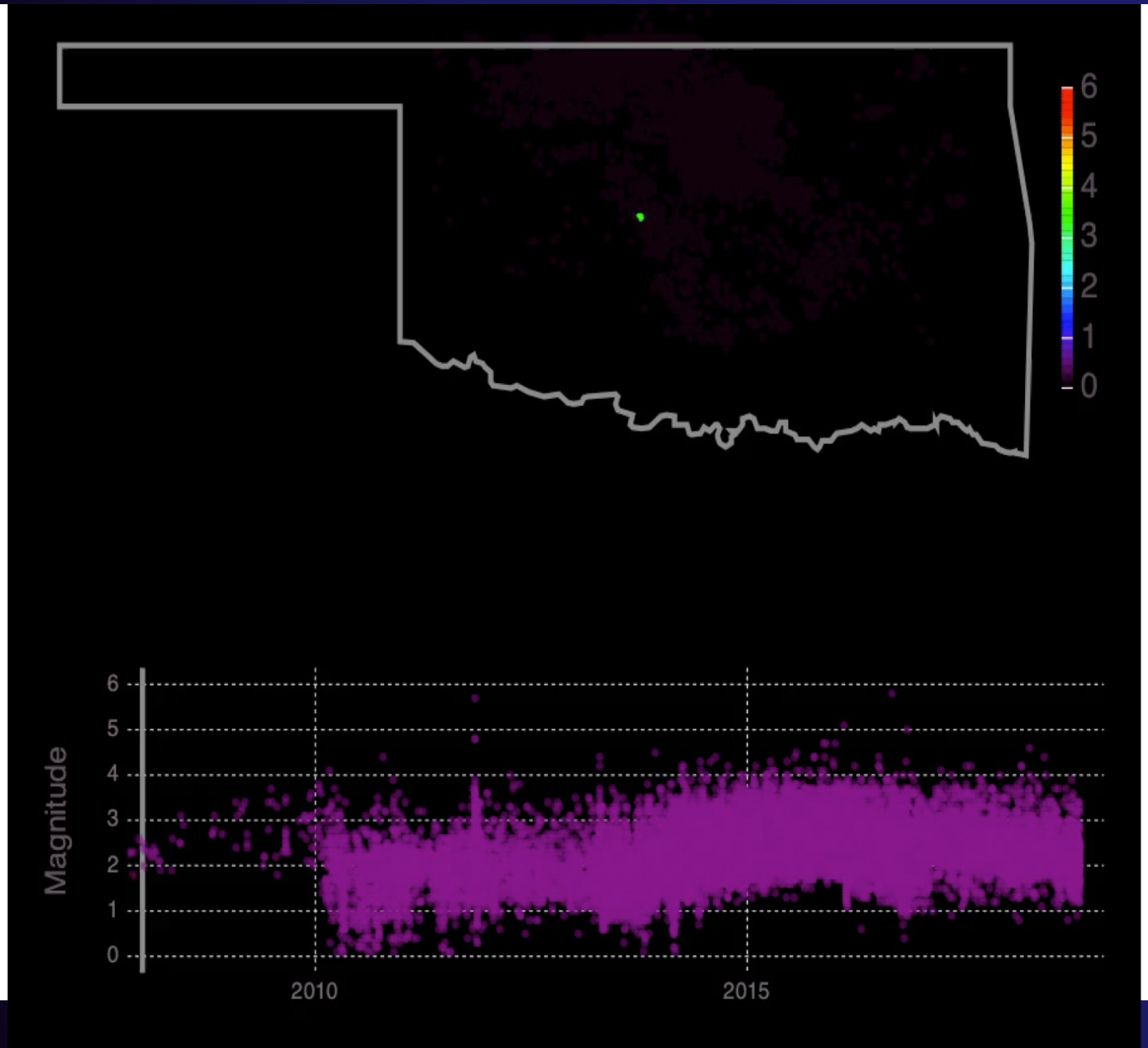


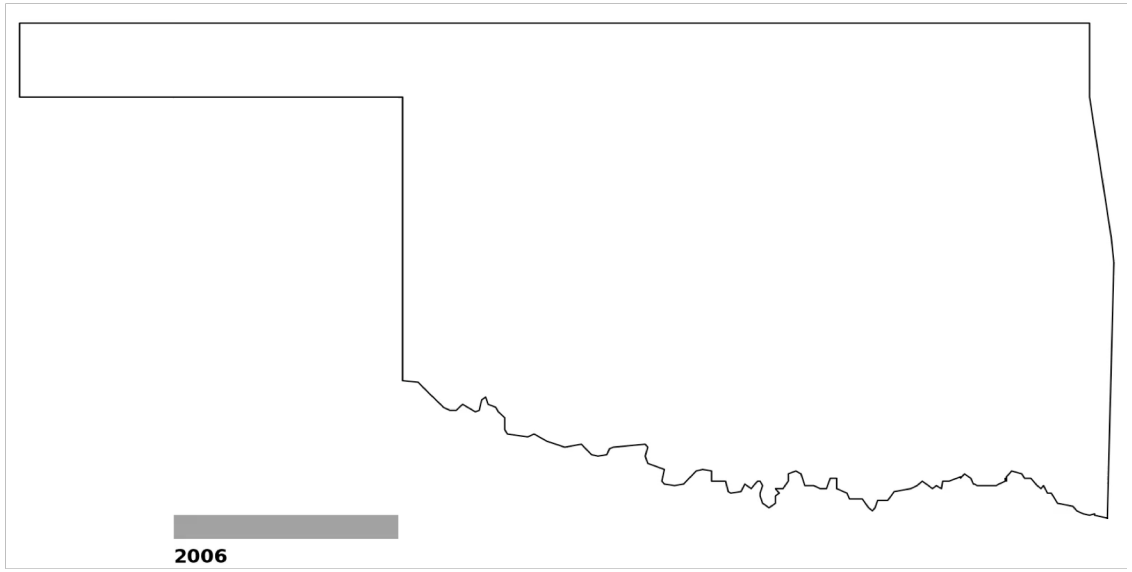


ML-informed Physics: Oklahoma Seismicity

Vesselinov, O'Malley—LDRD-DR

Los Alamos NATIONAL LABORATORY EST. 1943





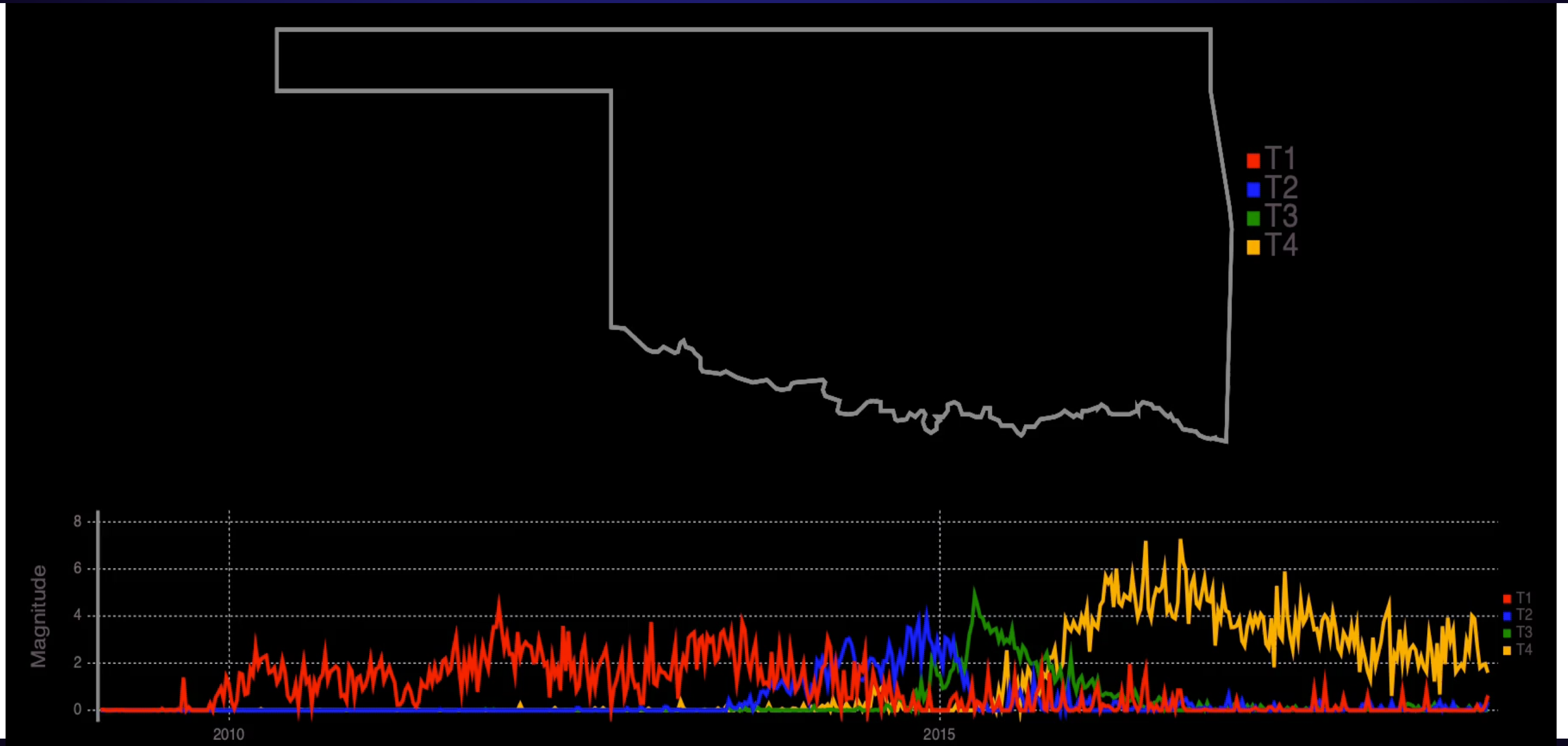
What are processes affecting seismic events?

- 32k+ events over a period of increased injections of waste water from hydraulic fracturing
- Can unsupervised machine learning extract any features?

ML-informed Physics: Oklahoma Seismicity

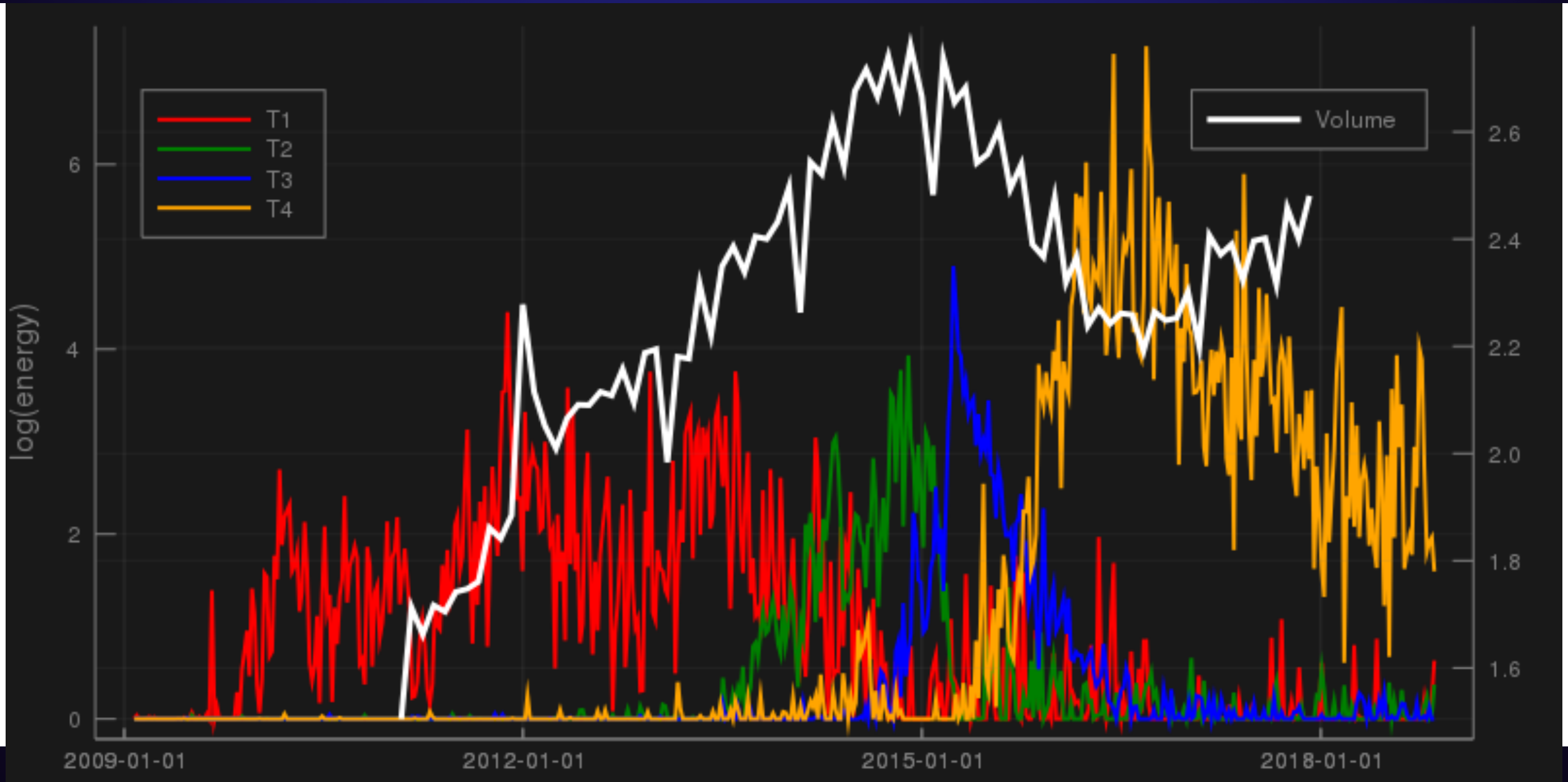
Vesselinov, O'Malley —LDRD-DR

Los Alamos NATIONAL LABORATORY EST. 1943



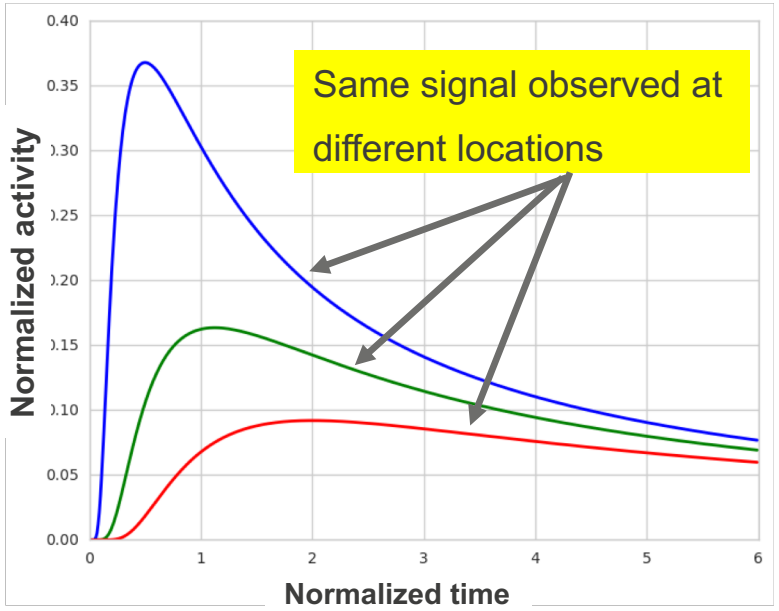
ML-informed Physics: Oklahoma Seismicity
Vesselinov, O'Malley —LDRD-DR

Los Alamos NATIONAL LABORATORY EST. 1943



Physics-informed ML: Feature extraction with Green's functions (Vesselinov —LDRD-DR)

- Datasets frequently are impacted by various physical processes
- Analyses are improved when information about these processes is included in the ML methodology (for example, diffusion and Doppler shifts)
- We have developed novel ML methods that allow for physics-informed ML



Green's function

$$G(\boldsymbol{x},t) = \frac{1}{4\pi\sqrt{D_x D_y}t} e^{-\frac{(x-u_x t)^2}{4D_x t}} e^{-\frac{y^2}{4D_y t}}$$

- Iliev, F.L., Stanev, V.G., **Vesselinov, V.V.**, Alexandrov, B.S., Nonnegative Matrix Factorization for identification of unknown number of sources emitting delayed signals *PLoS ONE*, 10.1371/journal.pone.0193974. 2018.
- Stanev, V.G., Iliev, F.L., Hansen, S.K., **Vesselinov, V.V.**, Alexandrov, B.S., Identification of the release sources in advection-diffusion system by machine learning combined with Green function inverse method, *Applied Mathematical Modelling*, 10.1016/j.apm.2018.03.006, 2018.

LANL leads the ML dialog national and internationally through workshops, conferences, symposia, etc.

Steering Committee

- Paul Johnson (LANL), Greg Beroza (Stanford), Eric Bylaska (PNNL), Maarten de Hoop (Rice), **Hari Viswanathan** (LANL)

Organizing Committee

- Kipton Barros (LANL), Harsha Bhat (ENS, Paris), George Guthrie (LANL), Aric Hagberg (LANL), Claudia Hulbert (ENS, Paris), **Jeffrey Hyman** (LANL), **Satish Karra** (LANL), Youzuo Lin (LANL), **Maruti Mudunuru** (LANL), **Daniel O'Malley** (LANL), Zhigang Peng (Georgia Tech), Pietro Poli (U. Grenoble). Chris Ren (LANL), Bertrand Rouet-LeDuc (LANL), Gowri Srinivasan (LANL), James Theiler (LANL), Daniel Trugman (LANL), Brendt Wohlberg (LANL)