



Challenges we face around data



Data Challenges

- Spatio-Temporal Data
- Heterogeneous Data
- Missing Data
- Noisy Data

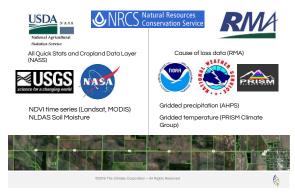


Learning Challenges

- Latent Features
- · Curse of Dimensionality
- Multi-task Learning



Open and easily accessible data are key



What are the needs now and in the future?

Data Repository

■ Data is spread out across agencies - no comprehensive library

- Documentation and Metadata

 Formats or documentation is inconsistent across websites

 Greater adoption of open, self documenting data standards
 (THREDDS, Climate Data Guide).

SLAs and Alerting

• Published SLAs and alerts on changes or operations of production datasets

What are the needs now and in the future?

Accessing Data

- When we are ingesting new datasets, bandwidth is by far the biggest bottleneck. NEXRAD on AWS
- $\bullet \hspace{0.1in}$ Latency and instability of some OPeNDAP servers we've used is so bad that sometimes it is actually quicker to download whole dataset

Actual Data

- Uncertainty around observations should be provided- we think in a
- probabilistic sense.

 Surface wind: . Presently wind is predominantly measured at 10m.



