USING REMOTE SENSING AND GEOSPATIAL DATA SETS TO DELINEATE GROUNDWATER DEPENDENT ECOSYSTEMS IN THE UNITED STATES

PRESENTER

ISABEL C. PEREZ HOYOS

PhD Candidate

ADVISORS

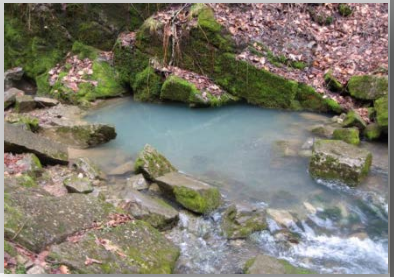
NIR Y. KRAKAUER

REZA KHANBILVARDI

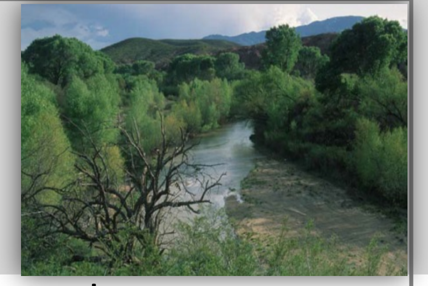
11th Annual NOAA/NESDIS CoRP Science Symposium.

September 16 & 17, 2015. University of Maryland, College Park, MD

WHAT ARE GROUNDWATER DEPENDENT ECOSYSTEMS (GDEs)?


Spring


Fen


Phreatophytes

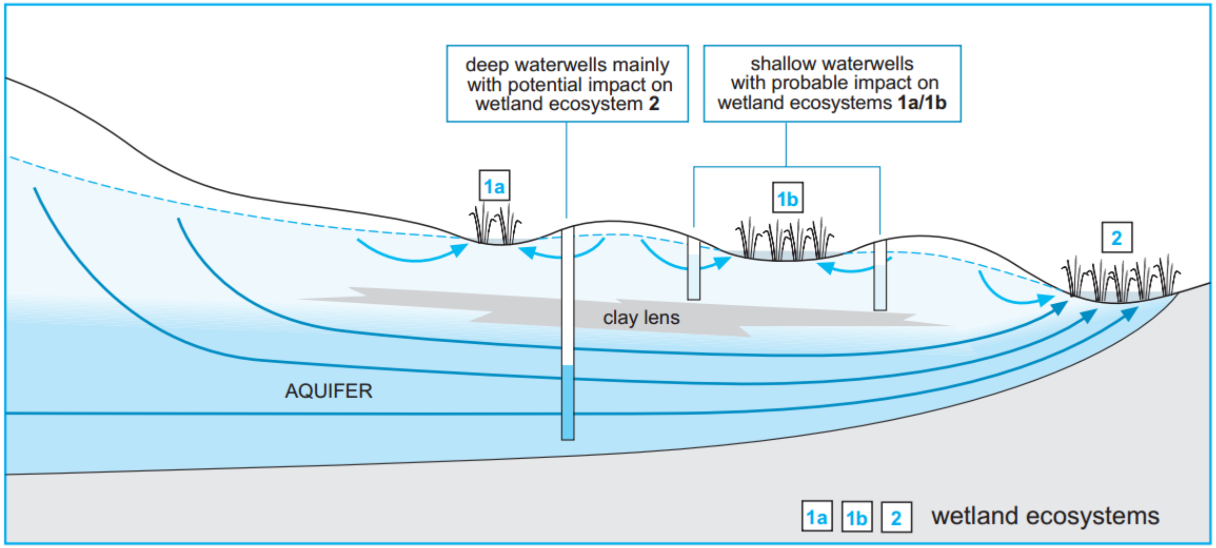
Community of plants, animals, and microorganisms that rely partially or completely on the availability of groundwater to maintain its structure and function

IMPORTANCE OF GDEs

- PROVISIONING SERVICES
  - Food
  - Water
  - Raw Materials
- REGULATING SERVICES
  - Air Quality Regulation
  - Climate Regulation
  - Moderation of extreme events
  - Erosion Prevention
- HABITAT SERVICES
  - Maintenance of biodiversity
- CULTURAL AND AMENITY SERVICES
  - Aesthetic
  - Recreation and Tourism

www.freedrinkingwater.com

MOTIVATION



Foster et al., 2006. World Bank.

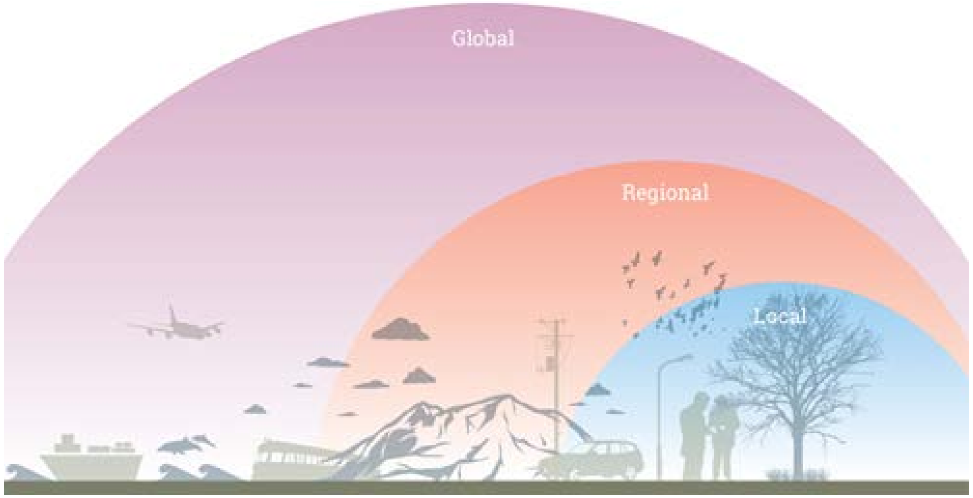
Conservation of GDEs requires:

Strategies that allow for the use of groundwater in a way that is compatible with the persistence of these ecosystems

1. Identification of location and extent of existing GDEs

2. Characterization of groundwater reliance
3. Characterization of ecological response to change

PROBLEM 1



National studies

Significant efforts to map the location and extent of GDEs at a national scale have been undertaken in only two countries

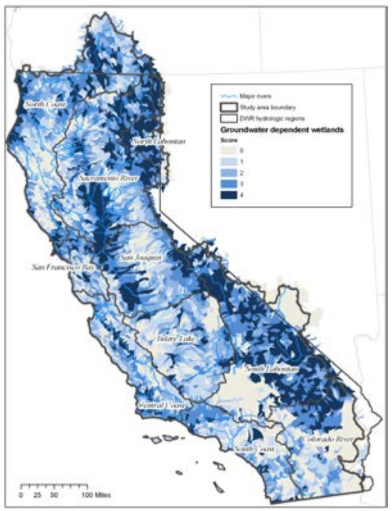
SOUTH AFRICA

- *Colvin et al., (2002)*
- *Probability of occurrence of terrestrial GDEs*
- *Two indicators: Groundwater levels and duration of the moisture growing season*

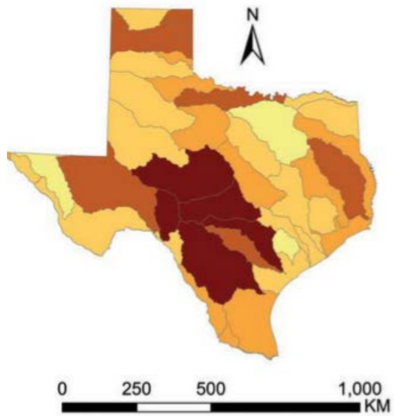
AUSTRALIA

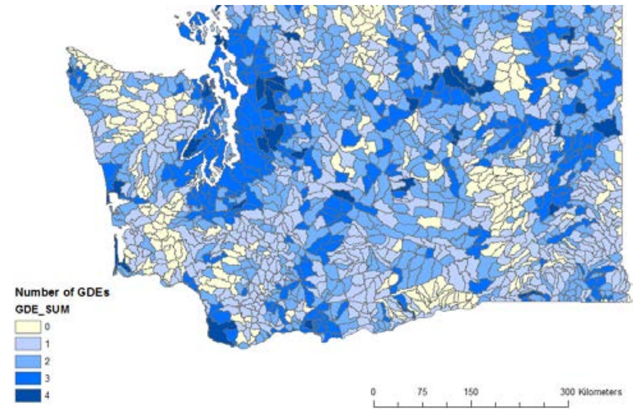
- *Operational GDE atlas (SKM, 2012)*
- *Previously identified GDEs, available literature, geospatial layers and remote sensing data*

PROBLEM 2



Hydrologic Unit Code-12; mean size = 9,570 ha; Cell Size Approx. 10 km





Developing a Ground Water Sustainability Strategy

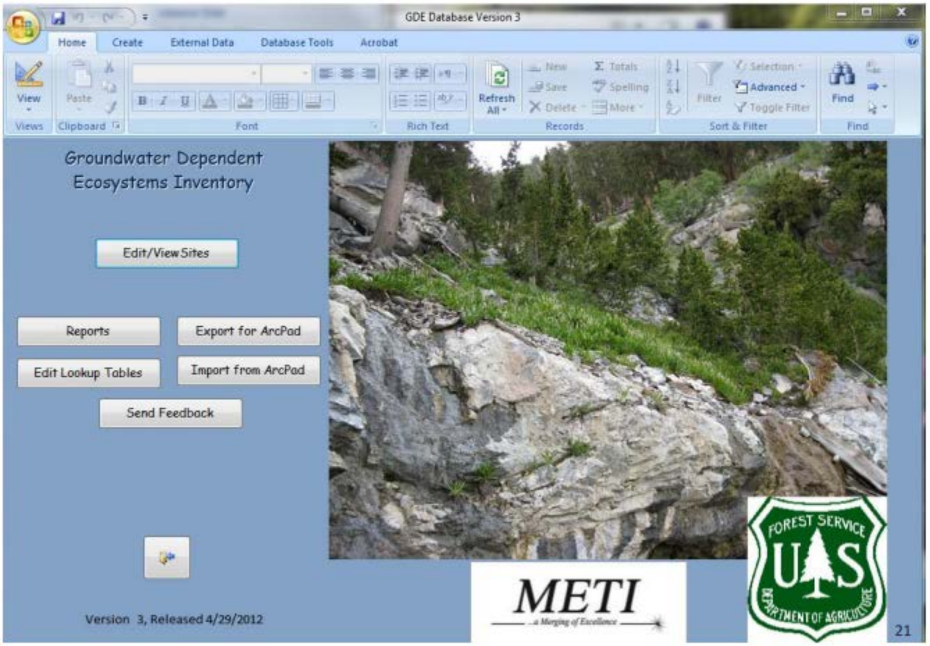
Determining which method or combination of methods to employ in a particular situation to promote a sustainable ground water supply generally should:

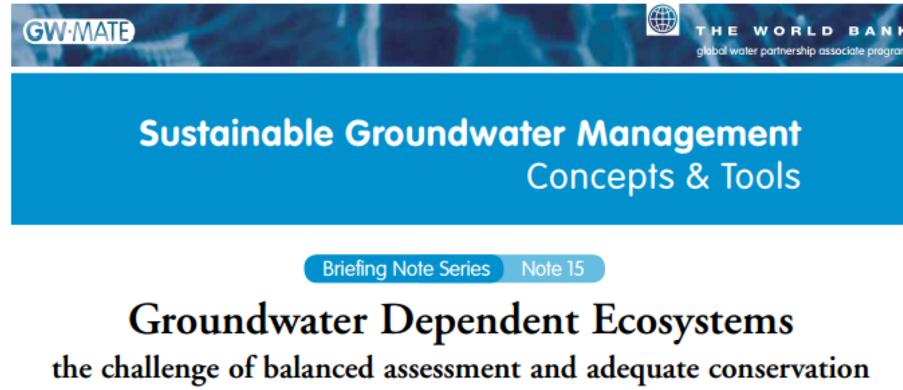
- Be made at a local level, whether that is a state, some government subunit, or an aquifer or ground water basin level. Local decision making provides the necessary flexibility to tailor the strategies to the specific situation. Ground water resource and climatic variability makes a one-size-fits-all approach unworkable. Local ground water management plans can incorporate site-specific information and input from all potentially affected parties. Implementation tools, such as land use planning or conservation measures, are also available at the local level.

NATIONAL GROUNDWATER ASSOCIATION
www.ngwa.org

NOAA-CREST

POTENTIAL USERS





The Nature Conservancy

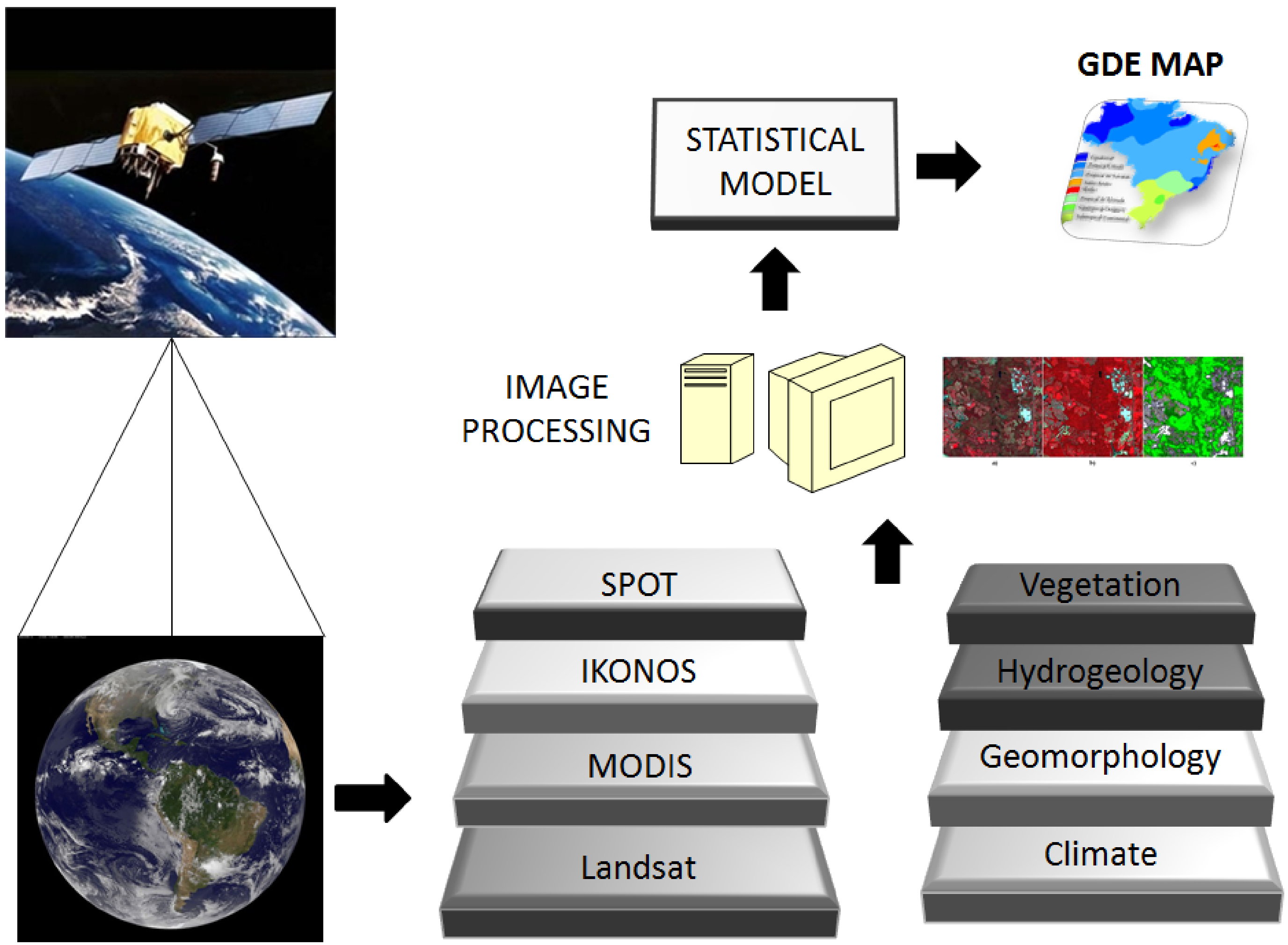
Protecting nature. Preserving life.®

Field Survey Overview

- Personnel with skills in botany, soils, hydrology, geology
  - Level I: 2-3 people
  - Level II: 3-5 people
- Survey time:
  - Level I: less than 2 hours/site
  - Level II: 3-6 hours/site

The Nature Conservancy in Oregon is implementing a comprehensive program to better understand the links between groundwater resources and biodiversity, and to develop and test actions that help ensure conservation of groundwater-dependent ecosystems and species.

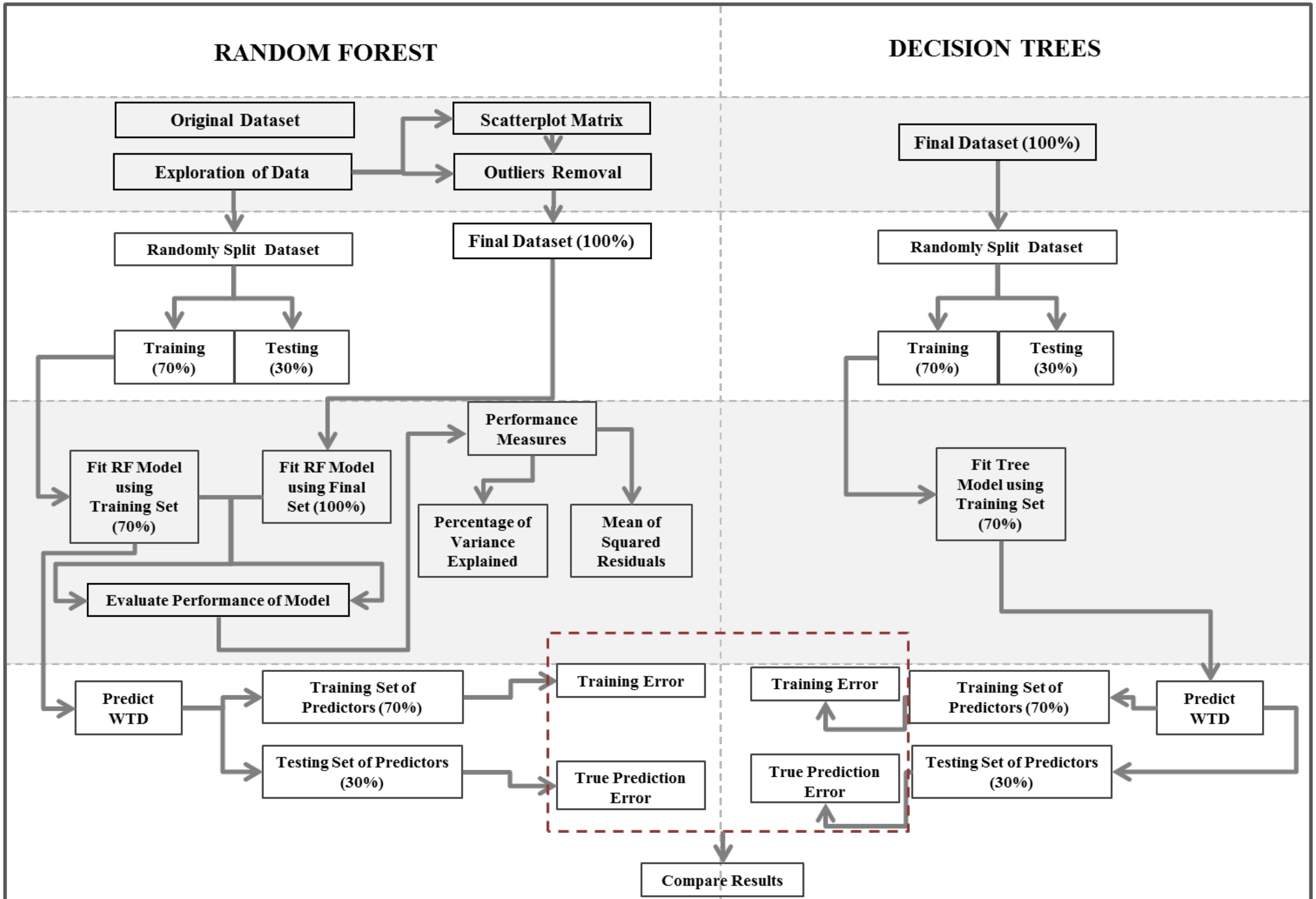
METHODOLOGY



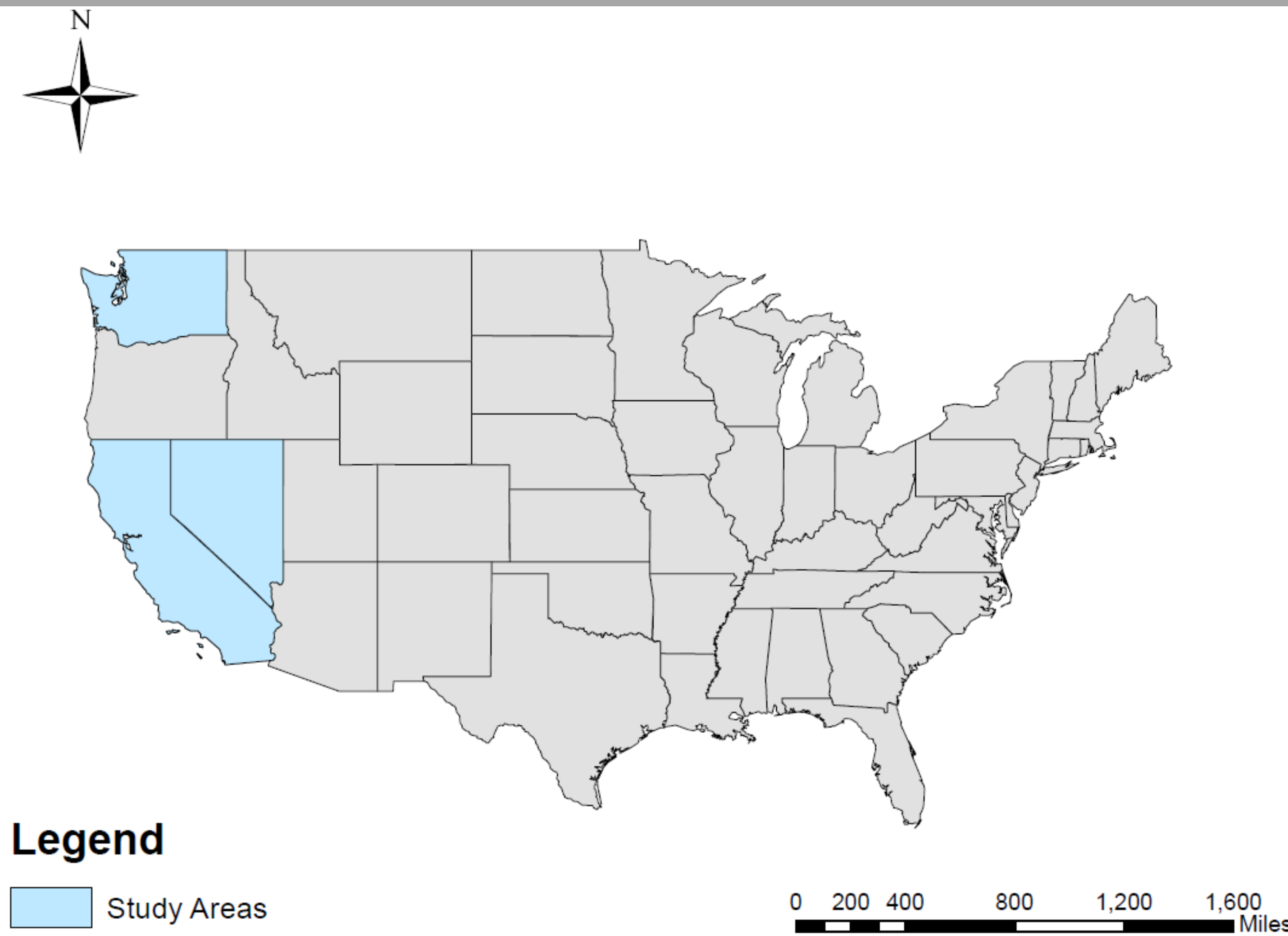
METHODOLOGY

- Ecosystems will generally use resources in proportion to their availability
- The ability of GDEs to use groundwater is related to their root system
- Most of the species in the world develop their root biomass in the first meters of soil
- The greater the aridity, the higher the dependence on the groundwater resource

MODEL DEVELOPMENT AND TESTING



STUDY AREAS

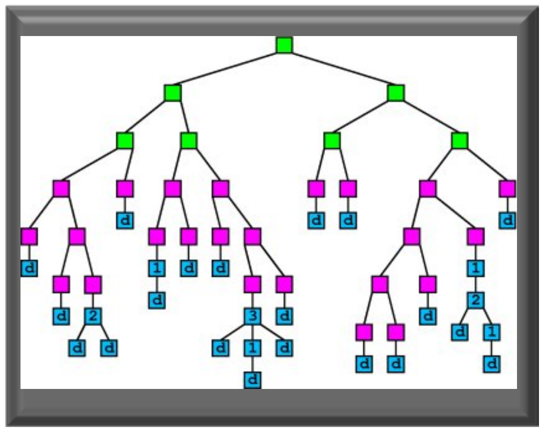


RANDOM FOREST FOR MAPPING GDEs

- PREDICTION

Predictor Variables

Known Response Variable



Ability to predict the response when it is unkown

CATEGORICAL – Classification

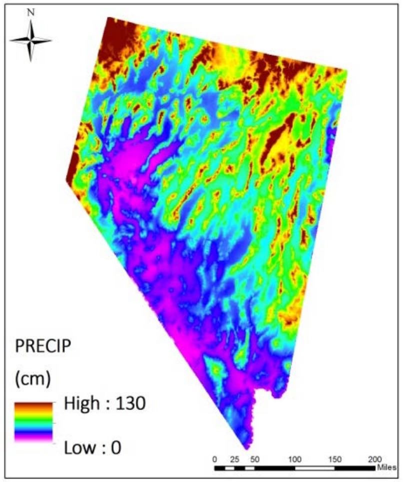
CONTINUOUS – Regression

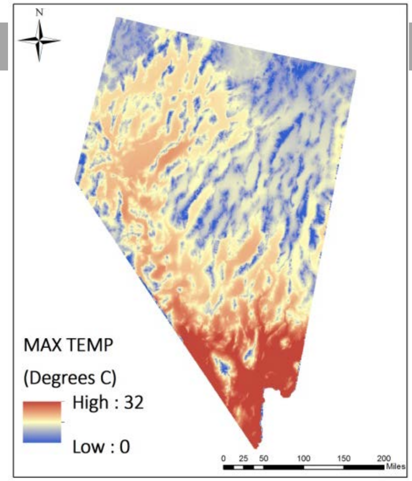
- Large collection of decorrelated decision trees
- Each tree is grown with a random subset of predictors – RANDOM
- A large number of trees are grown (500 to 2000) - FOREST

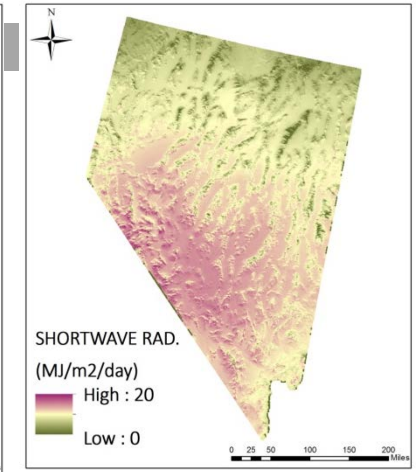
•Modeling the distribution of GDEs at national to global scales is challenging

•Conventional statistical approaches not enough

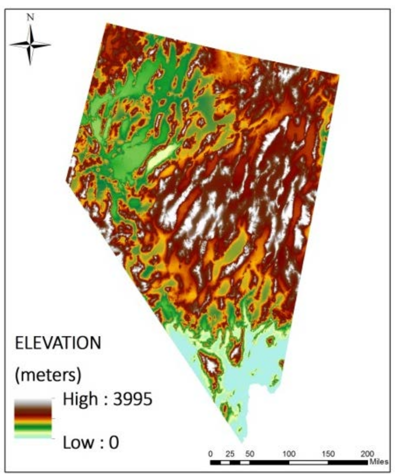
PREDICTOR VARIABLES

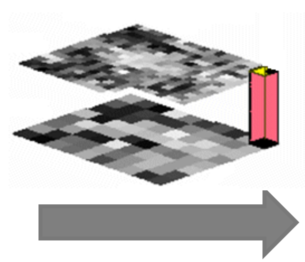




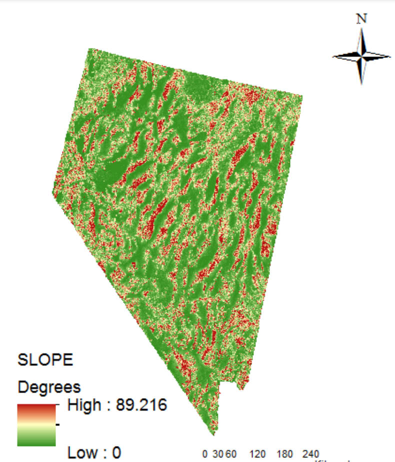


DAYMET 30 Year Average (1981-2010) of Daily Values (Thornton et al., 2014)



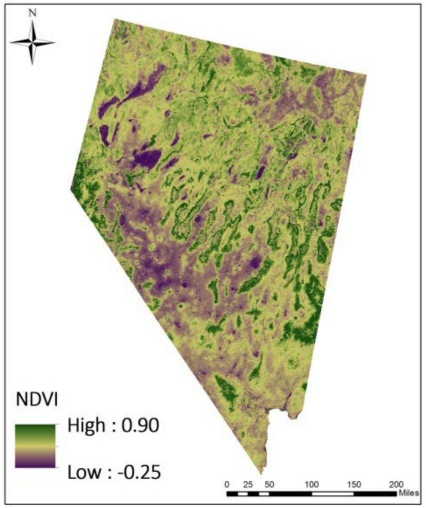


Spatial Analyst Tools

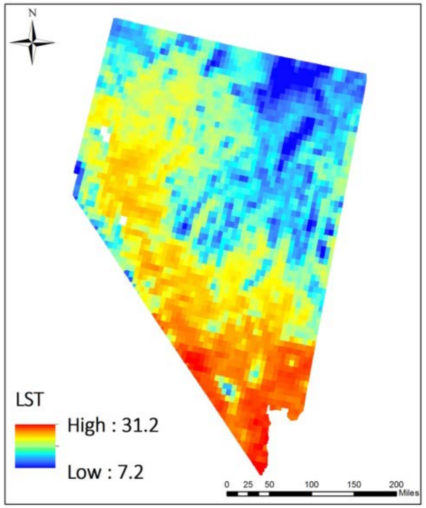


Digital Elevation Model at 30 m spatial resolution (USGS, 1999)

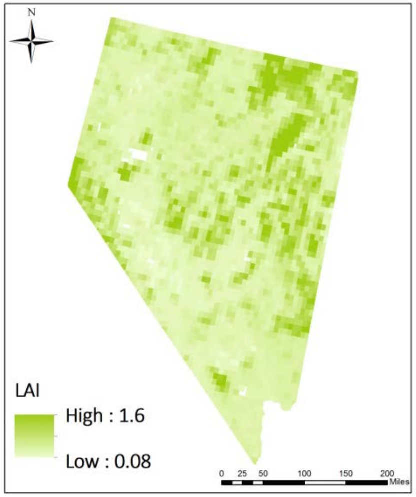
REMOTE SENSING VARIABLES



TERRA/MODIS
MOD13A2
Resolution: 1 km
Monthly
Feb 2000 – Dec 2012



TERRA/MODIS
MOD11A2
Resolution: 1 km
Monthly
Feb 2000 – Dec 2012

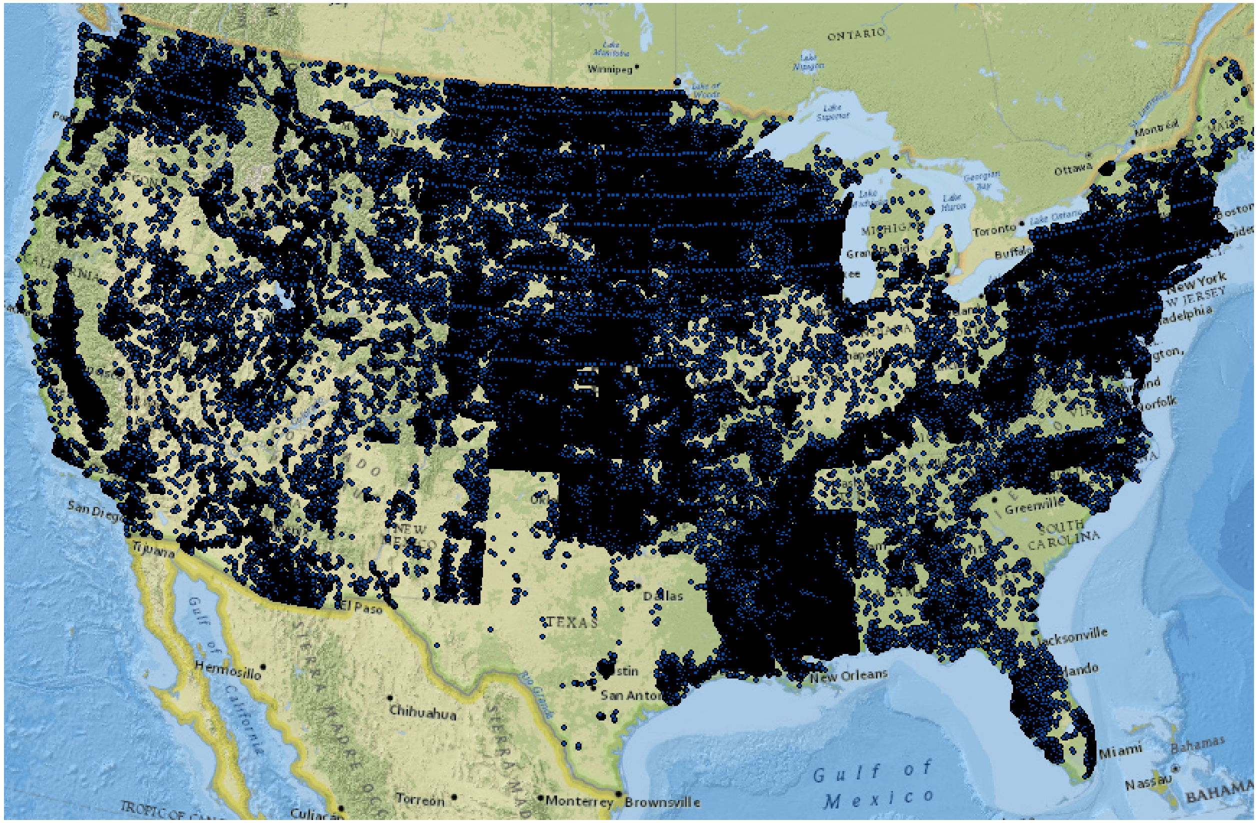


TERRA/MODIS
MOD15A2
Resolution: 1 km
Monthly
Feb 2000 – Dec 2012

RESPONSE VARIABLE

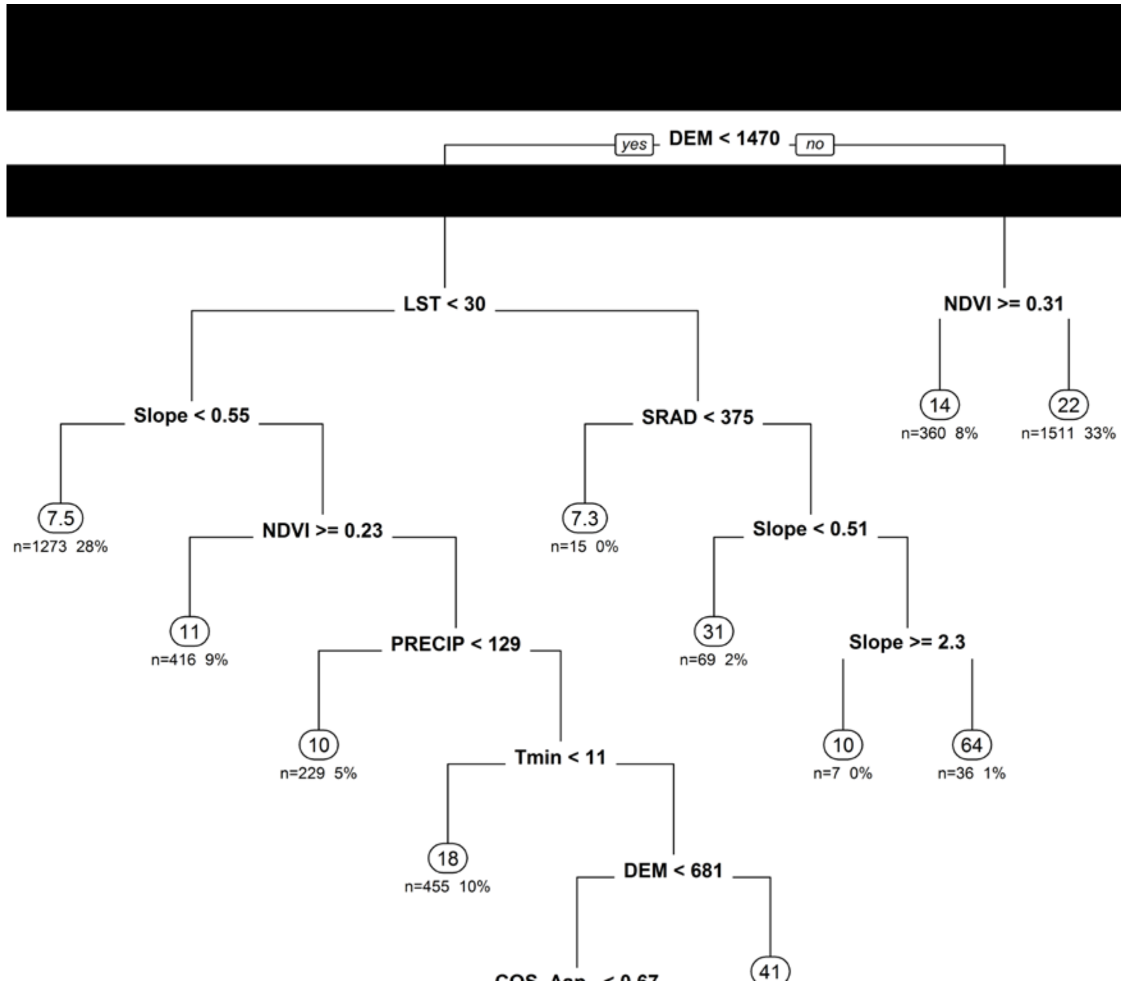
Water table depth observations were compiled from USGS archives (1927-2010) for more than 550,000 sites

http://waterdata.usgs.gov/nwis



REGRESSION TREES

- Regression Tree for Nevada

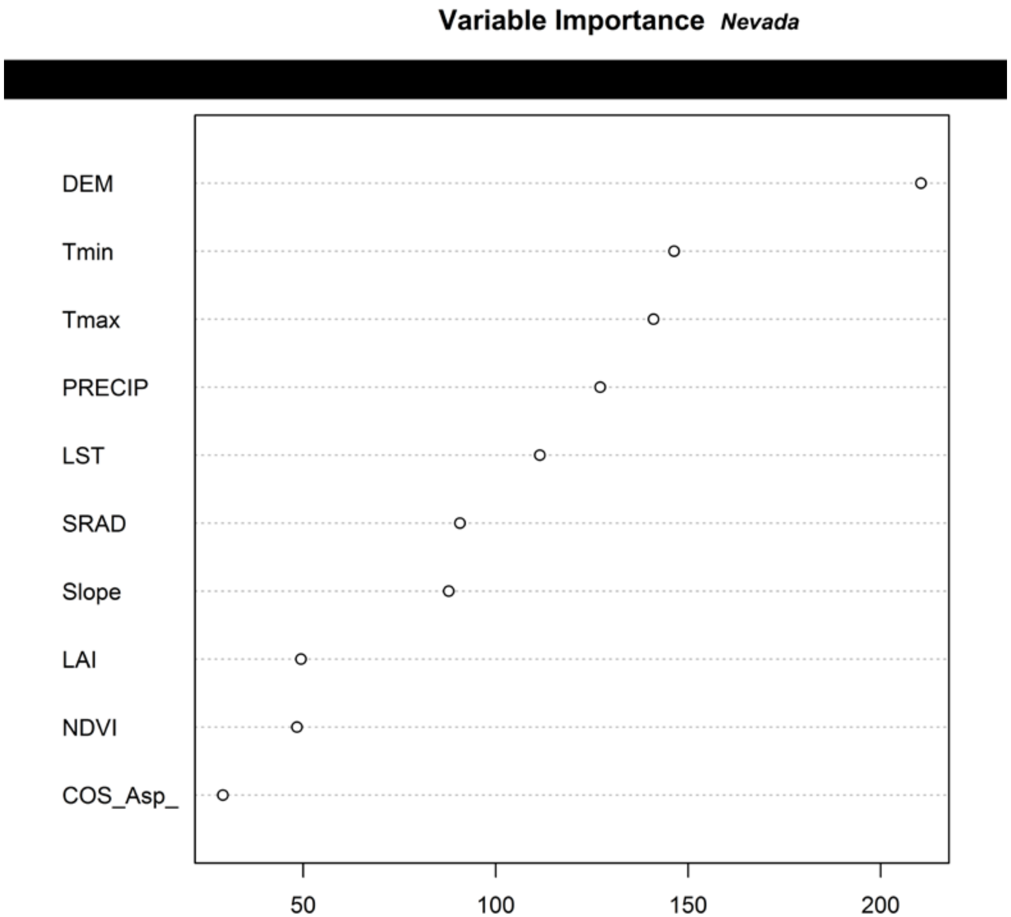


RANDOM FOREST RESULTS TRUE PREDICTION ERROR

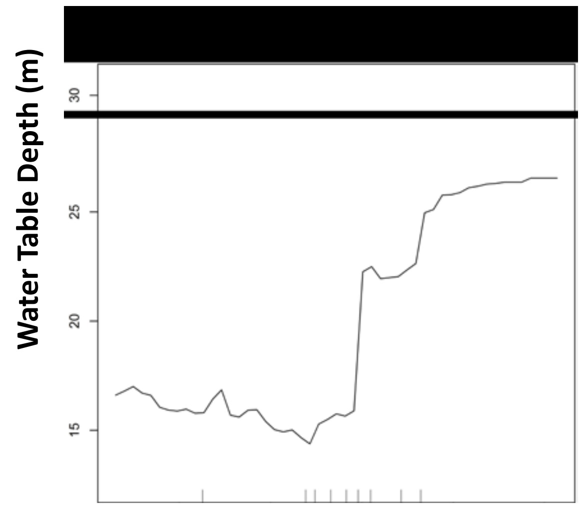
Performance Metric	Regression Procedure		
	Random Forest	Regression Trees	RF (No RS)
Nevada (n= 1940)			
Mean Absolute Error (MAE)	7.520	10.839	7.630
Mean Squared Error (MSE)	141.490	221.597	144.948
Root Mean Squared Error (RMSE)	11.895	14.886	12.039
Normalized Root Mean Squared Error (NRMSE)	0.118	0.147	0.119
Pearson's r	0.711	0.467	0.702
Kendall's Tau	0.555	0.330	0.549
Spearman's Rho	0.737	0.444	0.733
R-squared (as squared Pearson's r)	0.505	0.218	0.493
R-squared (as explained variance/total variance)	0.413	0.235	0.405
Nash-Sutcliffe Efficiency (NSE)	0.499	0.218	0.474
California (n= 6296)			
Mean Absolute Error (MAE)	5.301	9.441	5.426
Mean Squared Error (MSE)	73.279	164.560	75.801
Root Mean Squared Error (RMSE)	8.560	12.828	8.706
Normalized Root Mean Squared Error (NRMSE)	0.089	0.133	0.090
Pearson's r	0.806	0.454	0.799
Kendall's Tau	0.624	0.322	0.617
Spearman's Rho	0.801	0.440	0.794
R-squared (as squared Pearson's r)	0.650	0.207	0.638
R-squared (as explained variance/total variance)	0.559	0.224	0.542
Nash-Sutcliffe Efficiency (NSE)	0.647	0.206	0.632

RANDOM FOREST RESULTS VARIABLES ANALYSIS

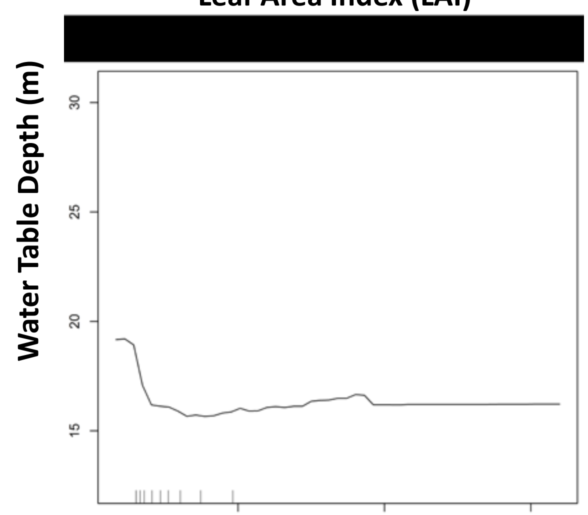
NOAA CREST



Elevation(m)



Leaf Area Index (LAI)



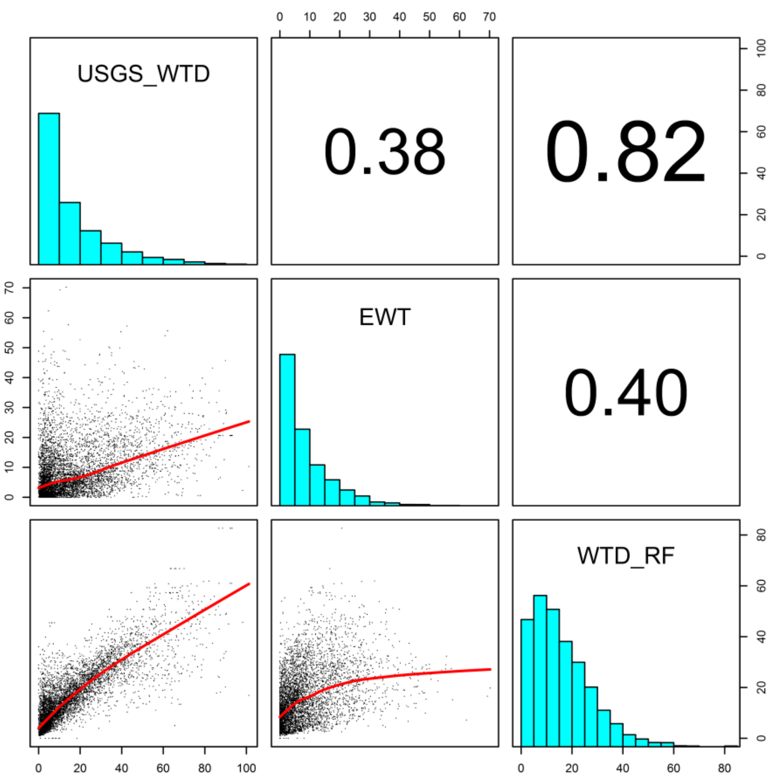
MODEL VALIDATION

USGS WATER TABLE OBSERVATIONS

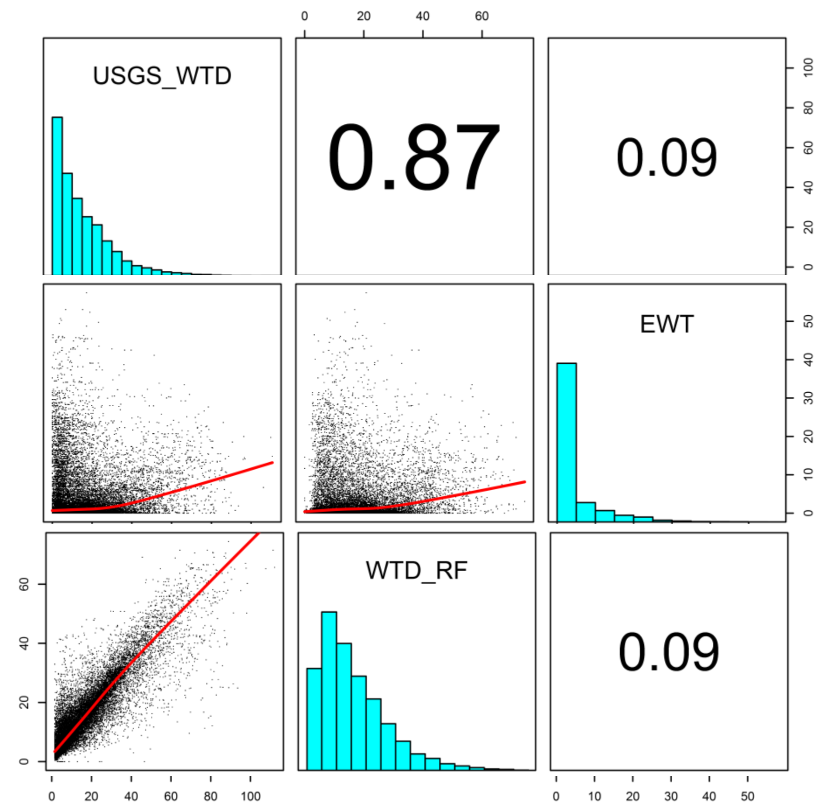
RANDOM FOREST RESULTS

EQUILIBRIUM WATER TABLE DEPTH (Fan et al., 2013)

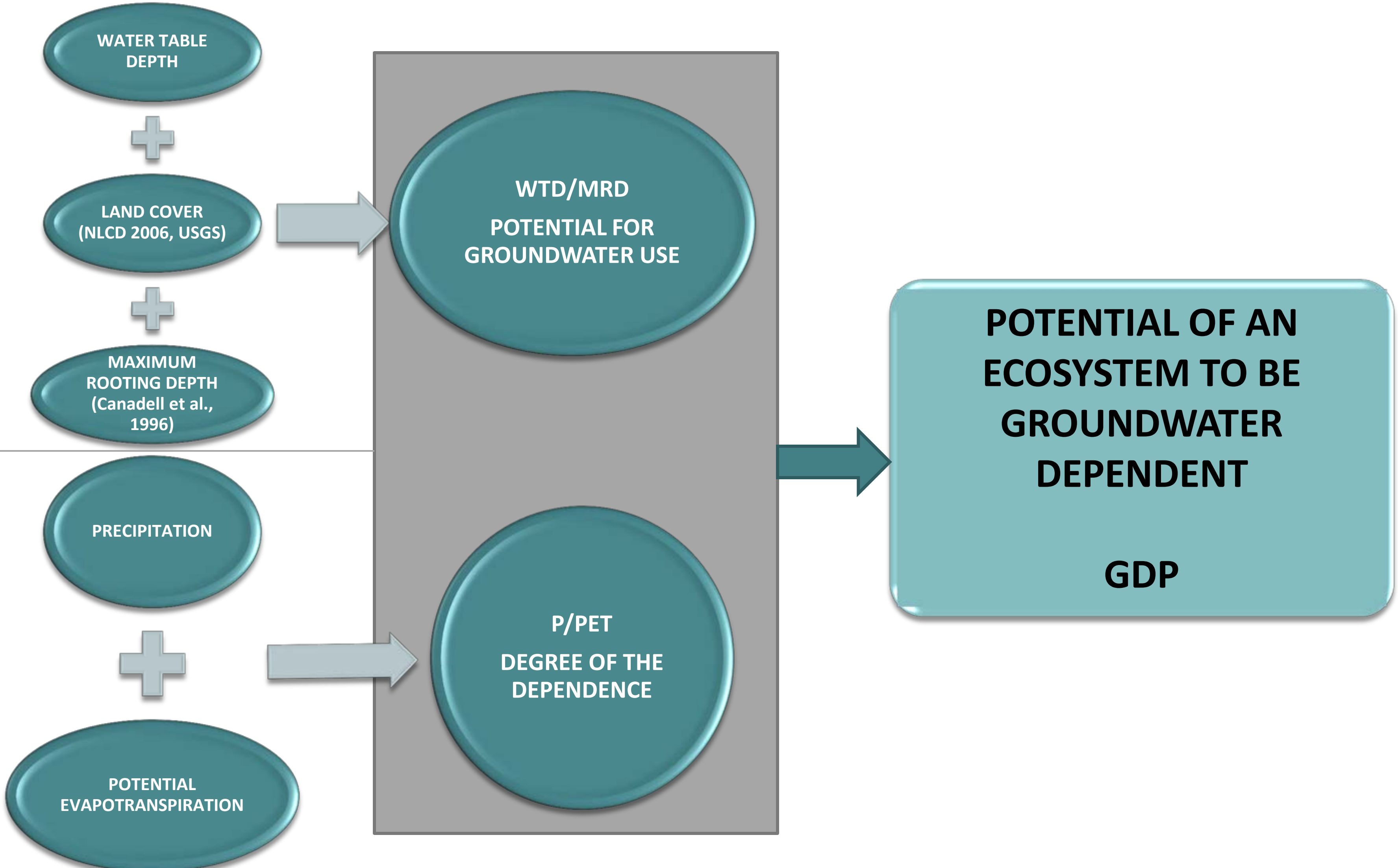
NEVADA (METERS)



CALIFORNIA (METERS)



GROUNDWATER DEPENDENCE POTENTIAL



REFERENCES

Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D. 1996. Maximum rooting depth of vegetation types at the global scale. Oecologia, 108(4), pp.583–595

Fan, Y. et al., 2007. Incorporating water table dynamics in climate modeling: Water table observations and equilibrium water table simulations. Journal of Geophysical Research, 112(D10), p.D10125. Available at: http://doi.wiley.com/10.1029/2006JD008111 [Accessed January 16, 2015].

Fan, Y., H. Li, G. Miguez-Macho, 2013. Global patterns of groundwater table depth, Science, 339 (6122): 940-943, doi:10.1126/science.1229881

Perez Hoyos, I.C., Krakauer, N. & Khanbilvardi, R., 2015. Random forest for identification and characterization of groundwater dependent ecosystems. , 196, pp.89–100. Available at: http://library.witpress.com/viewpaper.asp?pcode=WRM15-008-1

Thornton, P.. et al., 2014. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. Oak Ridge National Laboratory Distributed Active Archive Center. Available at: http://daac.ornl.gov/ [Accessed March 14, 2015]

Trabucco, A., and Zomer, R.J. 2009. Global Potential Evapotranspiration (Global-PET) and Global Aridity Index (Global-Aridity) Geo-Database. CGIAR Consortium for Spatial Information. Available online from the CGIAR-CSI GeoPortal at: http://www.csi.cgiar.org

CONCLUSIONS

- Although regression trees are constructed based on a continuous response variable, they still produce piecewise constant models. Their clear advantage is that they are easy to interpret and their results can provide insights into the nature of the data explored.

- The prediction accuracy of regression trees is reduced in comparison with smoother models such as random forest.

- Random Forest Algorithm has been found to provide superior predictive capability that could be useful in detecting GDEs.

- The poor spatial coverage of field observations could be complemented by geospatial data sets that provide cost-effective ways to monitor continuously large and remote areas.

THANK YOU