Application of Data Fusion and Knowledge Extraction in Hydrogeology

Michael Friedel (1,2), Chris Daughney(1), Massimo Buscemma (2,3), Lucas Moriera(4)
(1) GNS Science, Lower Hutt, New Zealand, (2) Mathematics and Statistical Sciences, University of Colorado, USA,
(3) Semeion Institute, Italy, (4) Institute of Education, Science, and Technology, Brazil

H13A-1489

m.friedel@gns.cri.nz

Semeion Centro Ricerche di Scienze della Comunicazione

Abstract: Since the development of personal computers, the modeling of groundwater systems shifted from analytical equations to numerical models. Given the ill-posed nature and non-uniqueness of numerical groundwater models, the use of alternate data fusion and knowledge extraction paradigms is being explored to reduce uncertainty through improvements in the conceptualization and parameterization processes and boundary conditions. This presentation demonstrates the use of data fusion using joint-inverse, artificial adaptive system, and hybrid modeling techniques to assist with these challenges. Examples include using joint-inversion for coupled unsaturated zone and geothermal studies, using artificial adaptive systems in water-quality and groundwater recharge studies including subset selection, using hybrid solutions for remote mapping of surficial aquifers and landscape characteristics, and forecasting climate change.

Goal: Find big-data solutions to hydrogeologic challenges using next-generation computational methods

Motivation: "We're drowning in data and starving for knowledge" *Rutherford D. Rogers*

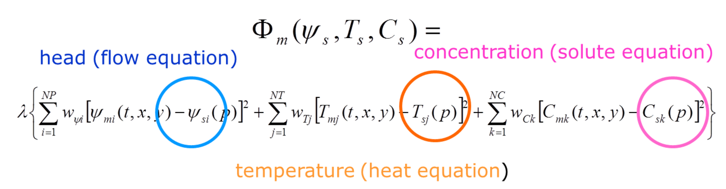
Objective: Use data-fusion to enhance mutual information for improved models

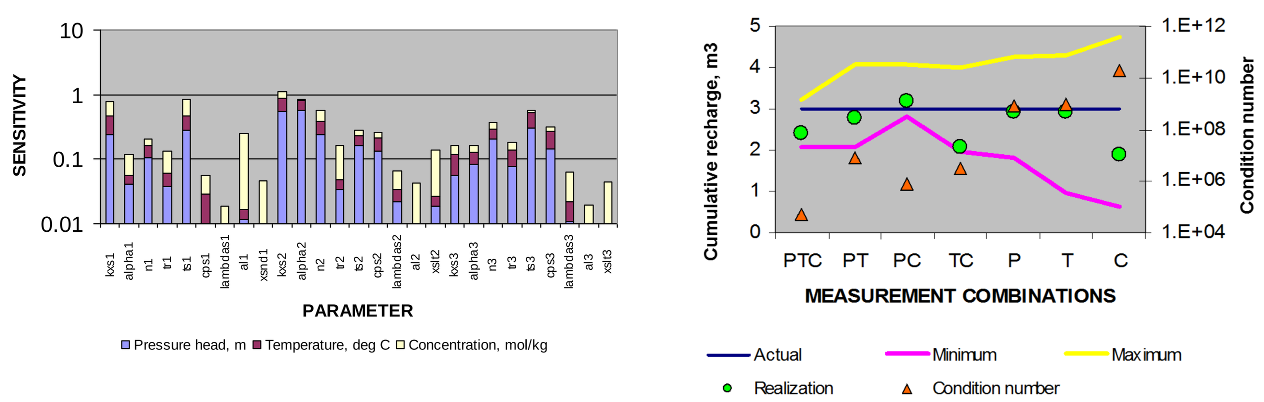
A. JOINT-INVERSION (Top Down)

1. Explicit Inversion of Coupled Partial Differential Equations (Single Model)

Objective: Quantify benefits of crossover effects to reduce recharge uncertainty **Data**: Pressure head, temperature, concentration

Modeling: Coupled set of PDEs; multi-criteria objective function



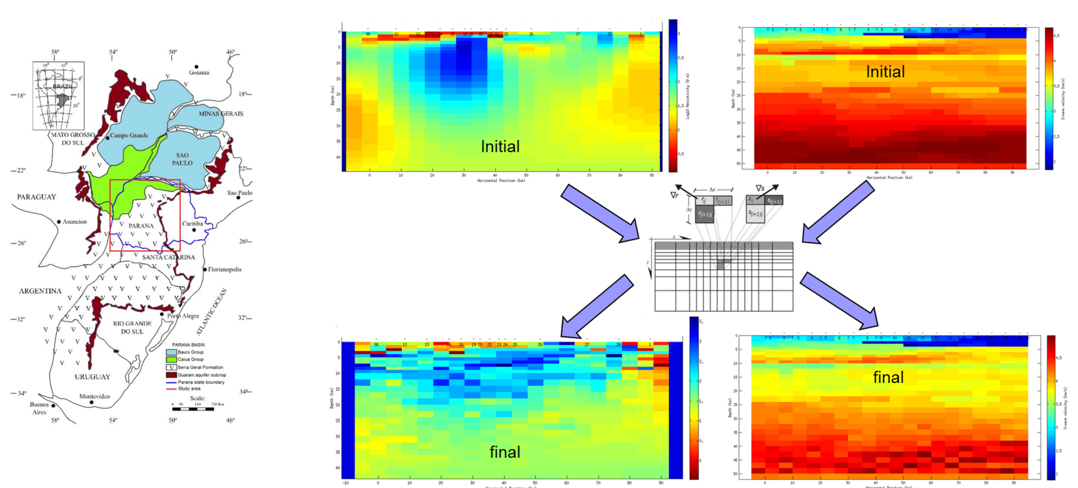


2. Implicit Inversion of Disparate Models

Objective: Quantify improvements in groundwater basin stratigraphy

Data: Receiver function, surface wave dispersion, electric/magnetic fields

Modeling: Cross-gradient constraint, multi-criteria objective function, seismic and magnetotelluric models

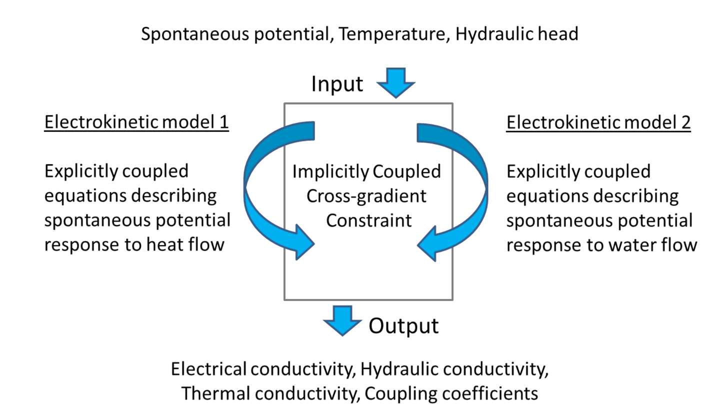


3. Implicit Inversion of Explicitly Coupled Models

Goal: Understand interaction of geothermal and groundwater systems

Data: spontaneous potential (SP), temperature, hydraulic head

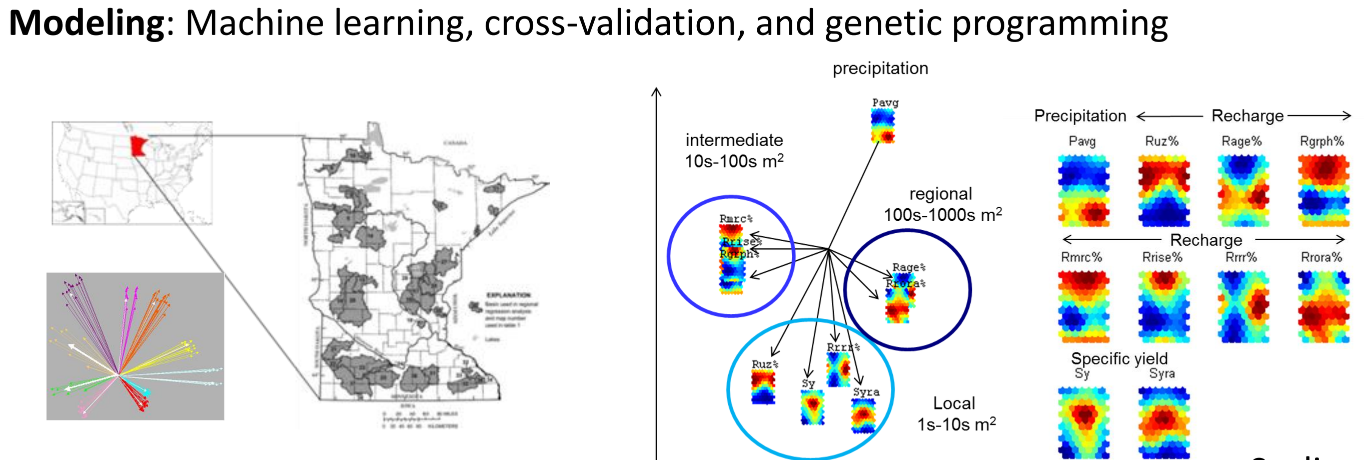
Modeling: Cross-gradient constraint, multi-criteria objective, electrokinetic models

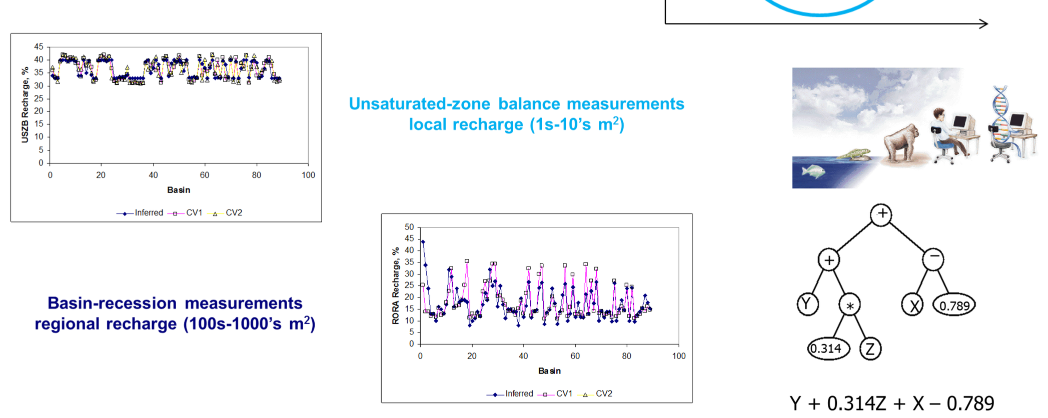


B. ARTIFICIAL ADAPTIVE SYSTEMS (Bottom Up)

1. Intelligent Scaling of Groundwater Recharge

Goal Develop equations for estimating groundwater recharge from scale-dependent measurements
Data: Local (1s-10s m^2): unsaturated-zone water balance; Intermediate (10s-100s m^2): water-table fluctuations, age dating; Regional (100s-1000s m^2): streamflow recession
Modeling: Machine learning, cross-validation, and genetic programming



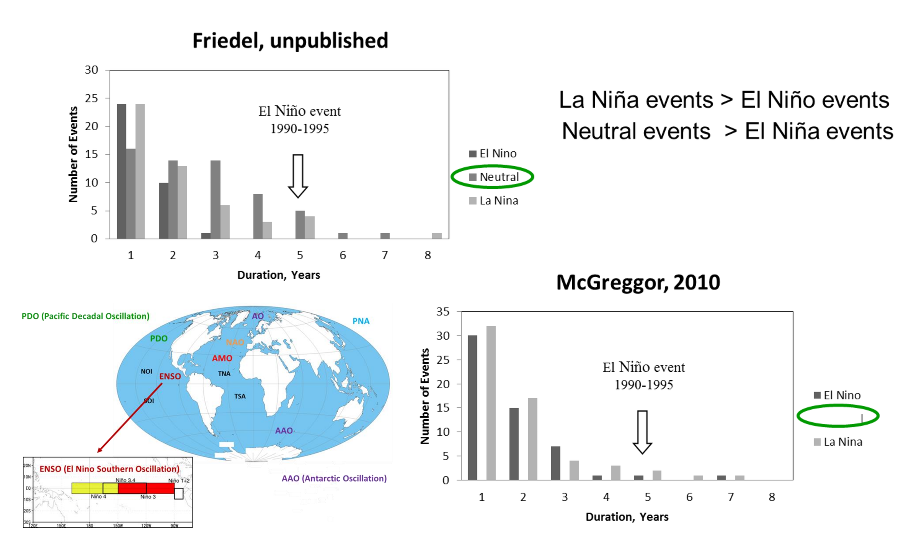


Scaling Equations

- Local recharge:
  - Local: unsaturated-zone water balance
  - Ruz = f(Rrise, Rrora), others …
- Regional recharge
  - Intermediate: water-table fluctuations, age dating
  - Rrise = f(Ruz, Rrora) , others …
- Basin recharge
  - Streamflow recession-curve-displacement
  - Rrora = f(Ruz, Rrise) , others …

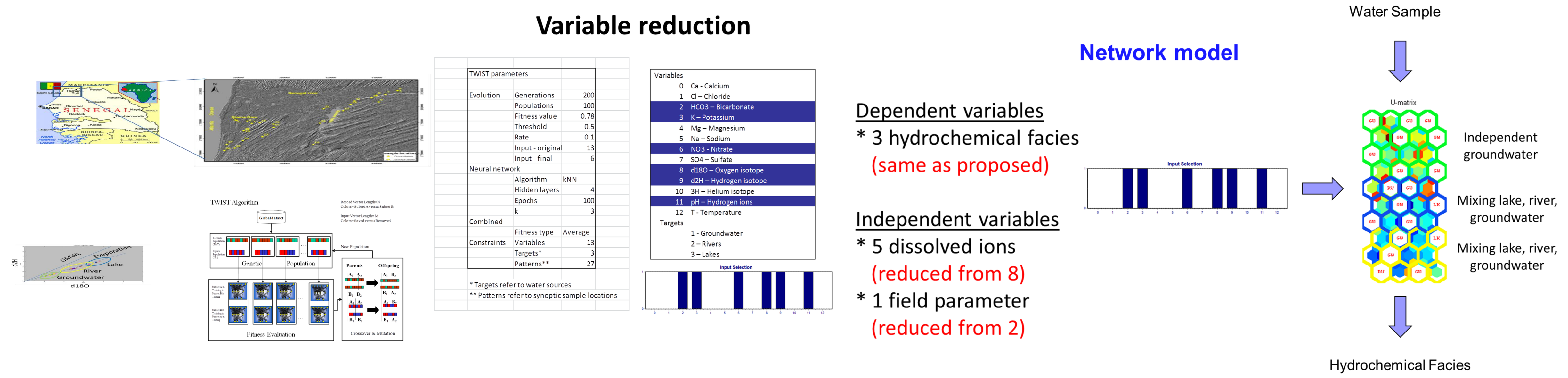
2. Climate Change: 1650-1977, 0-90N

Objective Quantify persistence of El Nino Southern Oscillation **Data**: Modern and paleoclimate temperatures **Modeling**: Machine learning, cross-validation



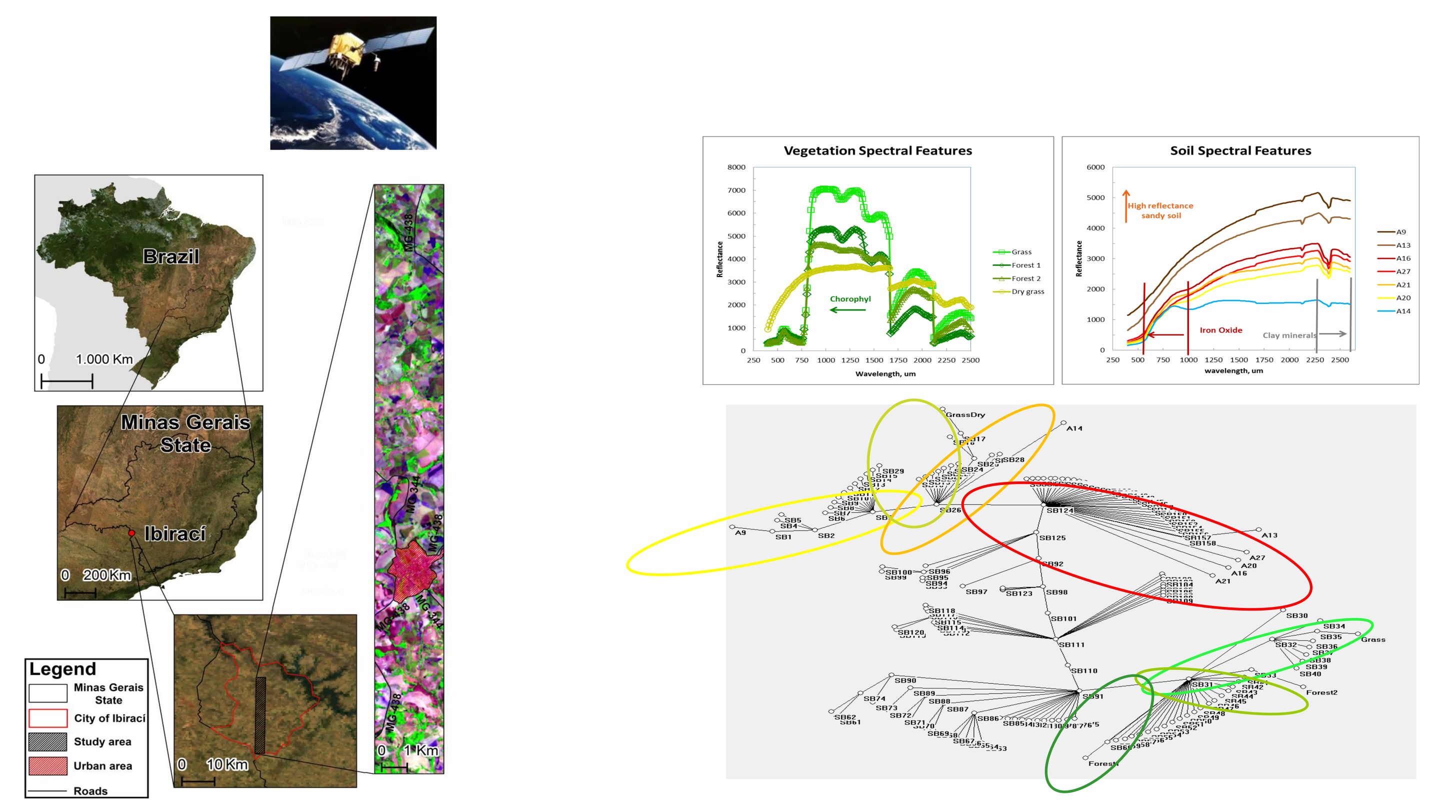
3. Intelligent Input Selection for Water Quality Modeling of River Delta Aquifer

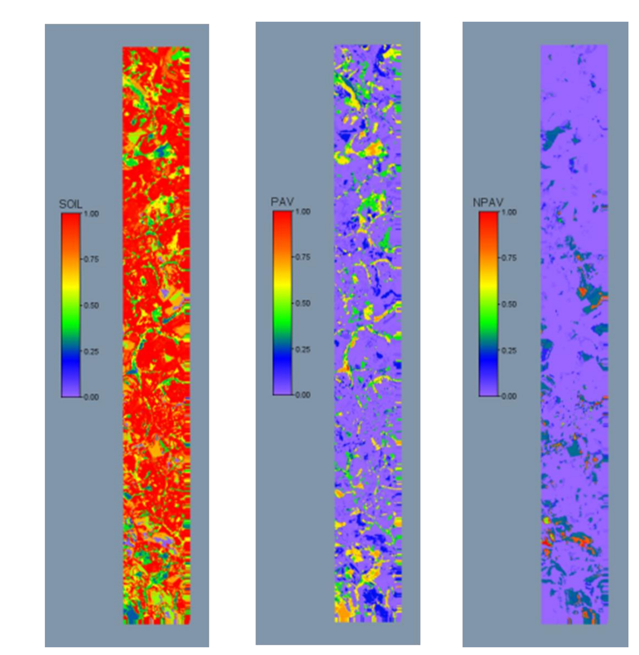
Objective: Quantify improvements to classifier of hydrochemical facies fiollowing variable reduction **Data**: Dissolved ions, field parameters, isotopes **Modeling**: machine-learning, genetic algorithm-supervised artificial neural network, statistics



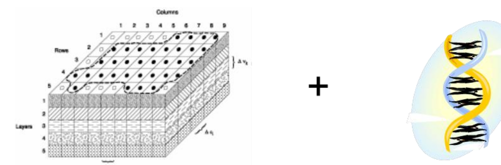
4. Remote Classification of Soil and Vegetation Components from Satellite Hyperspectral Data

Objective: Test training with independent spectral libraries (7 soil and 5 vegetation components)
Data: 200 hyperspectral reflectance bands **Modeling**: Machine-learning, boosting (ensamble)



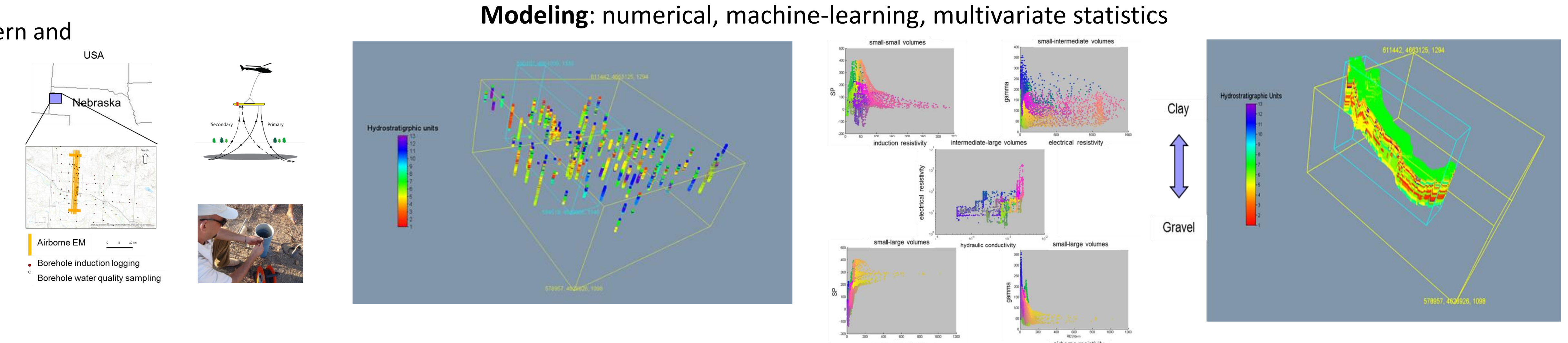


C. HYBRID MODELING (Combined)



1. Estimating Hydrostratigraphy from Hydrogeophysical Data

Objective: Inform groundwater model construction **Data**: Borehole hydrogeologic and geophysical, airborne electromagnetic
Modeling: numerical, machine-learning, multivariate statistics



Measurement Category	Measurement Type		Hydrostratigraphic Unit 1	2	3	4	5	6	7	8	9	10	11	12	13
Hydrogeologic	Lithology	Gravel (%)	100	95	88	57	0	5	0	0	0	1	9	3	7
		Sand (%)	0	3	12	43	99	90	9	7	0	1	1	1	14
		Silt (%)	0	0	0	0	0	0	82	0	0	0	1	0	1
		Clay (%)	0	0	0	0	0	1	2	92	80	3	0	2	3
		Mudstone (%)	0	0	0	0	0	0	7	1	0	90	79	1	0
		Siltstone (%)	0	2	0	0	1	4	0	0	20	5	2	93	0
		Brule (%)	0	0	0	0	0	0	0	0	0	0	8	0	75
	Hydraulic	Ks (m/s)	0.00178	0.00121	0.00109	0.00145	0.00121	0.00167	0.00024	0.00066	0.00051	0.00083	0.00057	0.00090	0.00059
	Water-quality	SC (mS/cm)	495.7	1081.9	1189.0	1206.3	995.3	666.7	968.4	716.6	631.7	1011.0	1015.2	1041.5	766.3
Geophysics	Borehole	RESind (ohm-m	46.8	38.5	148.7	102.9	41.1	54.3	11.8	13.1	5.5	9.8	11.5	17.1	20.7
		RES16N (ohm-r	86.6	91.8	98.4	313.7	129.8	67.8	227.5	21.6	12.3	16.4	58.1	56.2	33.5
		RES64N (ohm-r	156.7	163.1	151.1	758.4	278.8	104.6	766.8	30.7	16.3	24.4	150.5	136.7	41.7
		Gamma (cps)	65.2	83.2	113.3	99.5	76.5	120.4	102.8	125.0	165.5	114.8	116.8	142.1	125.7
		SP (mv)	52.7	30.3	38.3	110.6	246.2	218.4	57.6	72.5	15.9	59.1	-2.0	100.4	56.4

Summary: The data-fusion paradigms presented are useful for extracting knowledge for solutions to current challenges in hydrogeology.

References

Buscema, M., Breda, M., Lodwick, W., 2013, Training with input selection and testing (TWIST) algorithm, Journal of Intelligent Learning Systems, 5, 29-38
Buscema, M., Petritoli, R., Pieri, G., Sacco, P.L., 2010, AutoContracxtive Maps, Technical Paper 32, Semeion Institute, Rome, Italy.
Friedel, M.J., *in review*, Application of machine learning for estimating and scaling hydrostratigraphic units from hydrogeophysical data, Journal of Hydrology.
Friedel, M.J., Vicente, E., Iwashita, F., Koga-Vicente, A., Franceschini, M.H.D, *in review,* Toward real-time classifying of soil and vegetation components using hyperion imagery, Remote Sensing Journal.
Friedel, M.J., Esfahani, A., Iwashita, F., 2015, Toward real-time 3D mapping of surficial aquifers using a hybrid modeling approach, Hydrogeology Journal. (DOI) 10.1007/s10040-015-1318-2
Esfahani, A.A., Friedel, M.J., 2014, Forecasting conditional climate-change using a hybrid approach, Environmental Modelling & Software, 52, 83-97
Friedel, M.J., Iwashtia, F., 2013, Hybrid modeling of spatial continuity for applications to environmental inverse problems, Environmental Modelling & Software, 43, 60-79.
Friedel, M.J., 2012, Data-driven modeling of surface temperature anomaly and solar activity trends, Environmental Modelling & Software, 37, 217-232