# ESS2222

Lecture 10 – Machine Learning in Earth Sciences

Hosein Shahnas
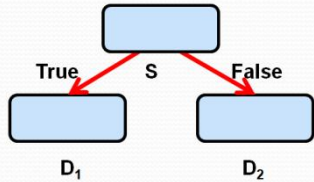
University of Toronto, Department of Earth Sciences,

Outline

- ❑ **Land cover classification**
- ❑ **Geological mapping**
- ❑ **Fast magnitude determination of the seismic events**
- ❑ **Climate change and sea level rise**
- ❑ **Weather forecasting**
- ❑ **Inverse problems in the planetary mantle**
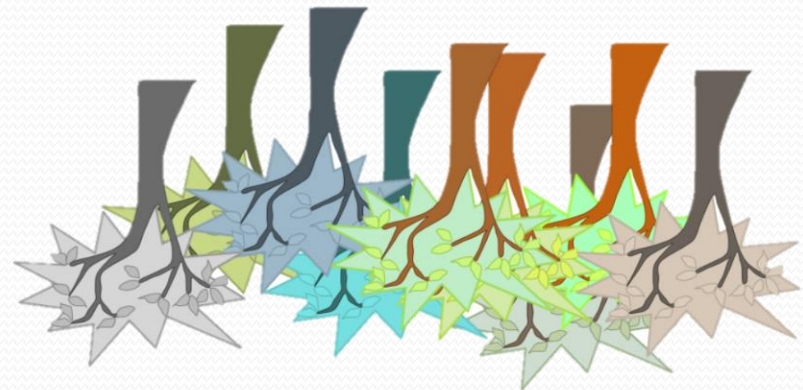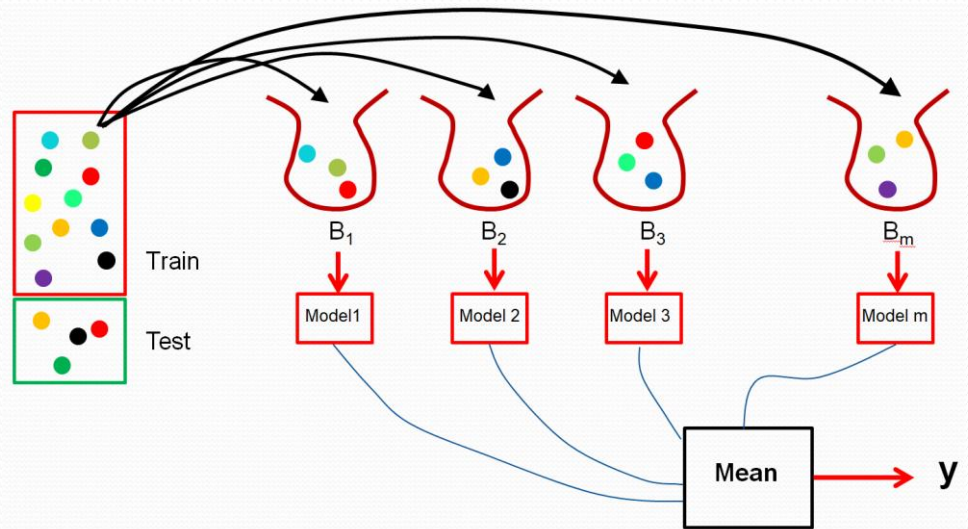- ❑ **The thermal state of the planetary mantle**
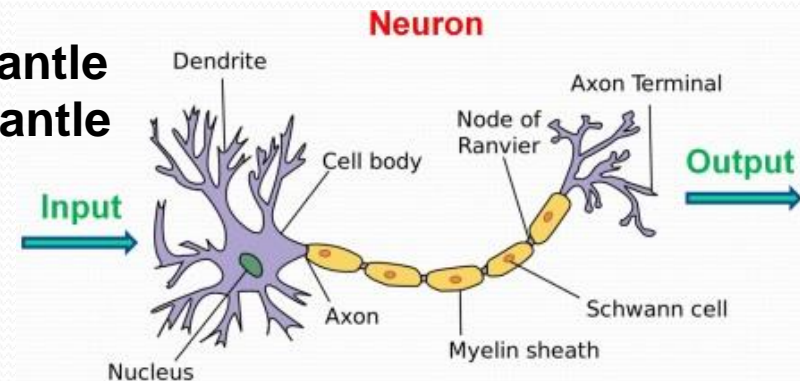
$$gini(D) = 1 - \sum_{c=1}^{C} P^2_c$$
$$\Delta\, gini\,(S) = gini\,(D) - giniS\,(D)$$

Land cover classification and change analysis of the Twin Cities (Minnesota)
The accurate and timely information describing the nature and extent of land resources and changes over time is important, especially in rapidly Growing metropolitan areas.
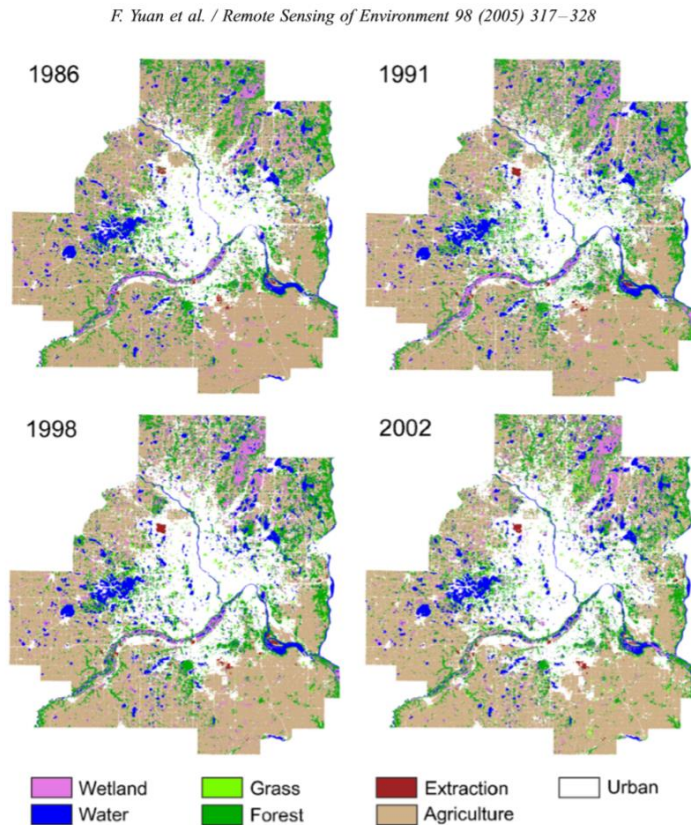

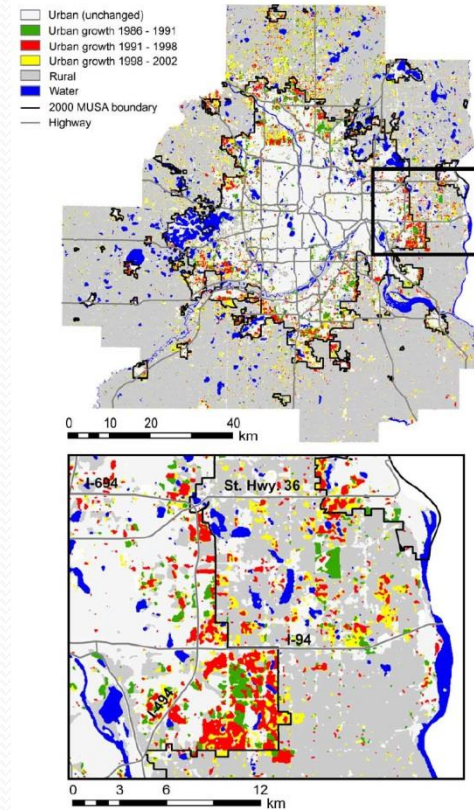Fig. 3. Landsat land cover classifications from 1986 to 2002 for the TCMA.

Fig. 4. Twin Cities Metropolitan Area urban growth from 1986 to 2002 with 2000 MUSA boundary. Rural land cover (agriculture, forest and wetland) that was converted to urban from 1986 to 1991, from 1991 to 1998, and from 1998 to 2002 are highlighted in green, red and yellow, respectively.

4

Geological Mapping Using Machine Learning Algorithms

Geological mapping

a) Remotely sensed spectral imagery
b) Geophysical (magnetic and gravity) data
c) Geodetic (elevation) data
are useful in a number of Earth science applications such as environmental monitoring and mineral exploration.

Remotely sensed imagery has many applications in Earth science applications such as **environmental monitoring** (Munyati, 2000), **land coverage studies** (Yuan et al., 2005), and **mineral exploration** (Hewson et al., 2006; Sabins, 1999).
So improving the techniques used in exploration and lithological identification is important for understanding of regional geology.

As the volume of data increase:
1) manual interpretation cannot maintain the pace with the amount of incoming data and
2) manual image interpretation is generally subjective and can be inconsistent among interpreters

Machine learning techniques can be employed in geological mapping and interpretation (Harvey & Fotopoulos, 2016) as a rapid approach of geological mapping in contrast to conventional field expedition techniques.

Geological Mapping Using Machine Learning Algorithms
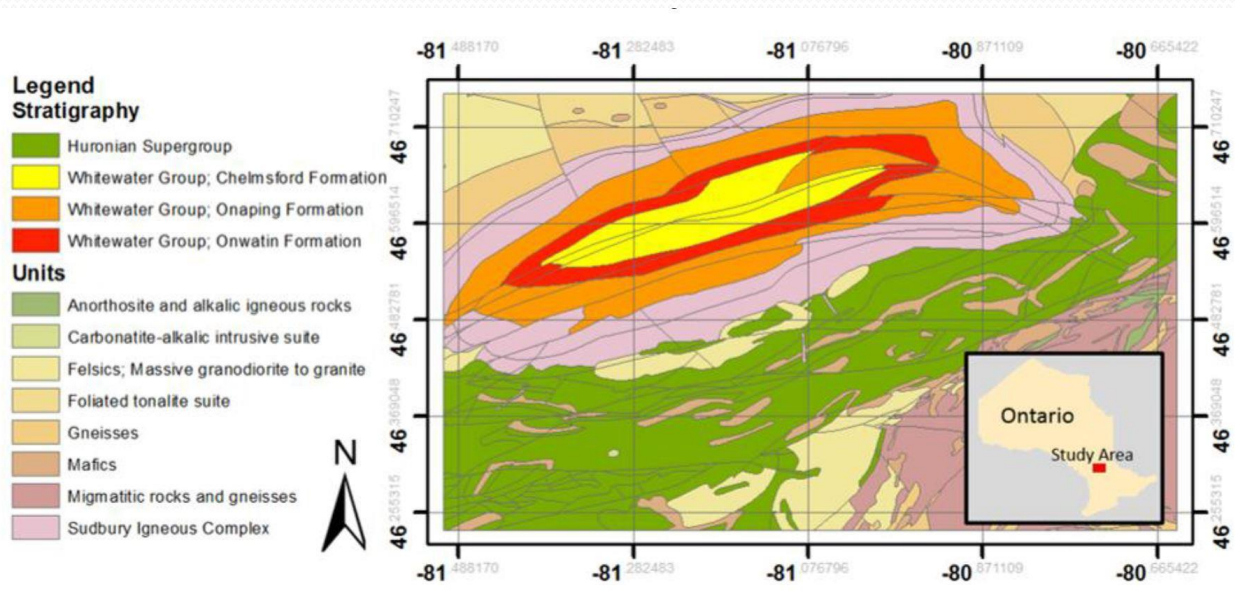
Training data

Figure 1. Map showing major stratigraphy groups and other major units in the Sudbury region (Ontario Geological Survey, 2011).

Features

Feature	Source and Filename	Units	Original Resolution
Landsat 4-5 TM Bands 1-7 October 2011	USGS LT50190282011278EDC00	Spectral Response 16-bit data	30 m × 30 m
Digital Elevation Model	USGS; SRTM n46_w081_1arc_v3 n46_w081_1arc_v3	metres	30 m × 30 m
Total Magnetic Intensity	OGS; MNDM ONMAGONL from GDS1036	nanoTelsa	200 m × 200 m
Bouguer Gravity Anomaly	OGS; MNDM ONGRAVTY1	milliGal	1000 m × 1000 m
Bedrock Geology	OGS Geopoly from MRD126-REV1	Discrete Geological Units	Resampled to study area density

Table 1. Summary of data, features for classification and validation, and class label inputs. Includes source, units, and original resolution.

Harvey and Fotopoulos (2016)

Geological Mapping Using Machine Learning Algorithms

All the inputs features:

Total magnetic intensity
Elevation
Gravity　　　　→　　are used to create a digital signature for each rock-type
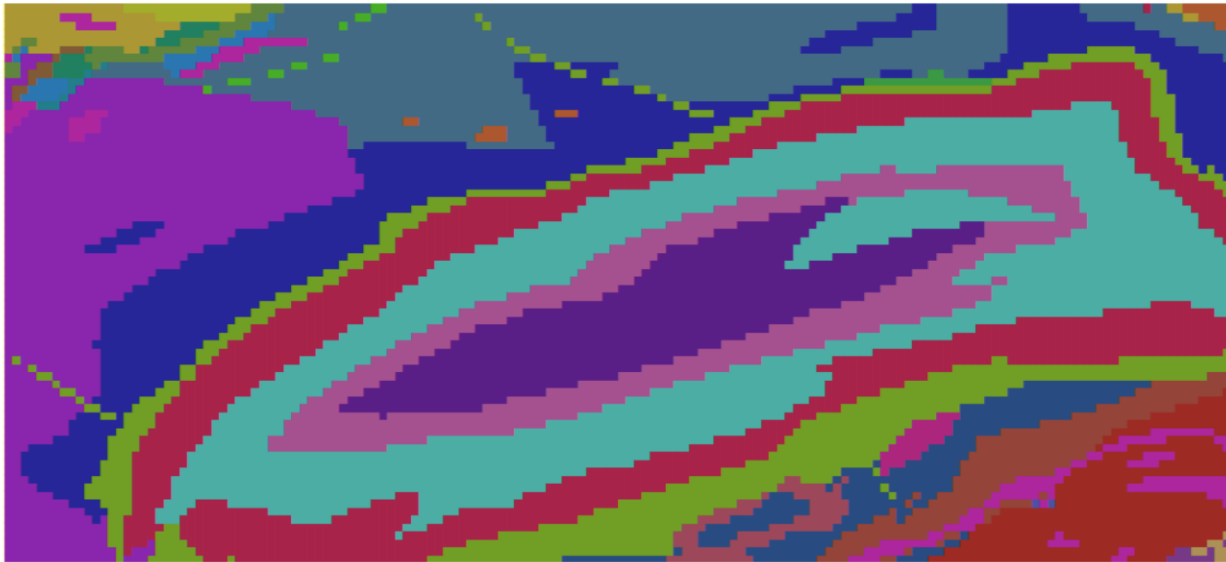Spectral images

Figure 2. Rocktype map of the Sudbury Basin and surrounding area. Refer to Table 3 for legend, rocktype descriptions, and proportions within the study area (Ontario Geological Survey, 2011).

Geological Mapping Using Machine Learning Algorithms

Class labels

Legend	% Cover	Rocktype Description
	0.11	Amphibolite, gabbro, diorite, mafic gneisses
	0.24	Basaltic and andesitic flows, tuffs and breccias, chert, iron formation, minor metasedimentary and intrusive rocks
	7.07	Carbonaceous slate
	0.08	Commonly layered biotite gneisses and migmatites; locally includes quartzofeldspathic gneisses, ortho- and paragneisses
	0.44	Conglomerate, sandstone, siltstone, argillite
	0.22	diorite, quartz diorite, minor tonalite, monzonite, granodiorite, syenite and hypabyssal equivalents
	0.25	Gabbro, anorthosite, ultramafic rocks
	0.82	Granite, alkali granite, granodiorite, quartz feldspar porphyry; minor related volcanic rocks (1.5 to 1.6 Ga)
	13.54	Granophyre
	18.53	Lapilli tuff, breccia, felsic flows and intrusions, minor carbonate and cherty
	2.72	Mafic, intermediate and felsic metavolcanic rocks, intercalated metasedimentary rocks and epiclastic rocks
	10.80	Massive to foliated granodiorite to granite
	0.33	Murray Granite 2388 Ma, Creighton Granite 2333 Ma: granite
	1.64	Nipissing mafic sills (2219 Ma): mafic sills, mafic dikes and related granophyre
	0.14	Norite, gabbro, granophyre
	7.79	Norite-gabbro, quartz norite, sublayer and offset rocks
	0.24	Quartz sandstone, minor conglomerate, siltstone
	3.50	Quartz-feldspar sandstone, argillite and conglomerate
	0.38	Quartz-feldspare sandstone, sandstone with minor siltstone, calcareous siltstone and conglomerate
	0.85	Rhyolitic, rhyodacitic, dacitic and andesitic flows, tuffs and breccias, chert iron formation, minor metaseds and intrusive rocks
	0.09	Sandstone, siltstone, conglomerate, limestone, dolostone
	0.13	Siltstone, argillite, sandstone, conglomerate
	0.05	Siltstone, argillite, wacke, minor sandstone
	2.33	Siltstone, wacke, argillite
	10.70	Tonalite to granodiorite-foliated to gneissic-with minor supracrustal inclusions
	10.40	Tonalite to granodiorite-foliated to massive
	6.67	Wacke, minor siltstone

Table 3. Legend and rock type descriptions for Figure 2. Includes % of how much of the study area each rock type covers. Adapted from Ontario Geological Survey (2011).
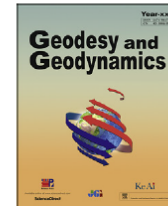
Fast Magnitude Determination

Contents lists available at ScienceDirect

Geodesy and Geodynamics

Fast magnitude determination using a single seismological station record implementing machine learning techniques

Luis H. Ochoa[*], Luis F. Niño, Carlos A. Vargas

Universidad Nacional de Colombia, Carrera 45 N° 26-85 Edificio Manuel Ancizar - Of: 330, Bogotá, Colombia

ARTICLE INFO

ABSTRACT

In this work a Support Vector Machine Regression (SVMR) algorithm is used to calculate local magnitude (Ml) using only five seconds of signal after the P wave onset of one three component seismic station. This algorithm was trained with 863 records of historical earthquakes, where the input regression parameters were an exponential function of the waveform envelope estimated by least squares and the maximum value of the observed waveform for each component in a single station. Ten-fold cross validation was applied for a normalized polynomial kernel obtaining the mean absolute error for different exponents and complexity parameters. The local magnitude (Ml) could be estimated with 0.19 units of mean absolute error. The proposed algorithm is easy to implement in hardware and may be used directly after the field seismological sensor to generate fast decisions at seismological control centers, increasing the possibility of having an effective reaction.

Early Earthquake Warning System

Early warning systems employ **dense seismological networks** **to localize and determine the magnitude of the earthquake using at** **least 3 stations****.**

L.H. Ochoa et al. / Geodesy and Geodynamics xxx (2017) 1–8

Fig. 1. Seismological Colombian network. Modified from Servicio Geológico Colombiano (SGC).

42 stations transmitting in real time

The problem with this method:
The density of stations in some high seismic risk areas is not enough to make such localization calculations fast.

An alternative solution:
The seismological records of previous events recorded at one single station can be used to localize and **estimate** **the magnitude of the event.**

10

Early Earthquake Warning System

Implementation of an early earthquake warning system for the city of Bogota, Colombia

Seismic early warning systems (SEWS) **emit an alert**, few seconds after the event initiates, from few seconds to a few tens of seconds before the stronger shaking movement arrives.

Fig. 2. Area of study. Bogota–Colombia and its surroundings.

The main task:

Estimation of **magnitude** and **source location** of an earthquake in a short period of time accurately

863 records of historical earthquakes are used in training of a Support Vector Machine Regression (SVMR)–model to calculate (predict) local magnitude (Ml) using only **five seconds** of signal after the P wave onset of one three component seismic station.

SVMR–model
10-fold cross validation
Polynomial kernel

Early Earthquake Warning System

Historical data:

The seismic catalogue with 2164 seismic events, selected between Jan. 1st,1998 and Oct. 27th, 2008, located at less than 120 km from the seismic station.
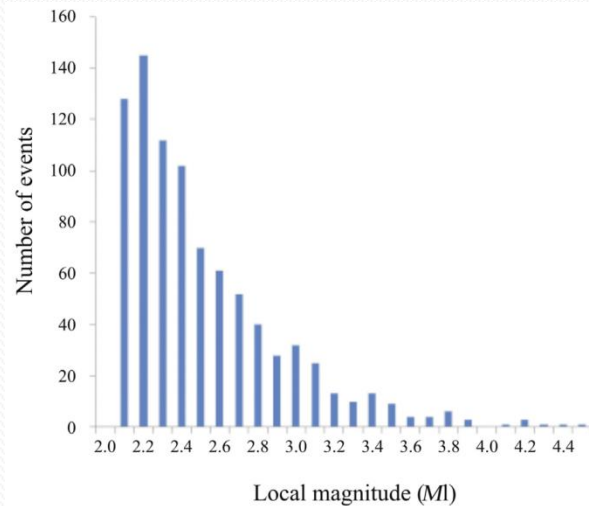

Fig. 3. Local magnitude statistical distribution.
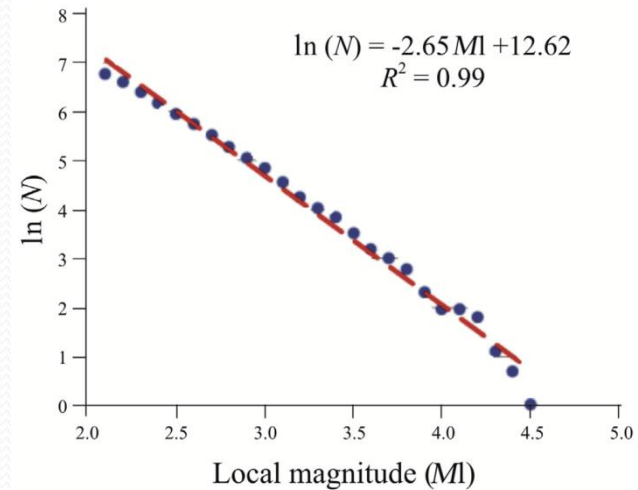

$$\ln (N) = -2.65\, Ml + 12.62$$
$$R^2 = 0.99$$

Fig. 4. Gutenberg–Richter relation for selected seismic events.

The parameters for Gutenberg -Richter relation change from one seismic region to another which should be taken into account in the training machine leaning models for different regions.

Early Earthquake Warning System

SVM-Regression Model

Tainting data:

863 historical data from seismic events (2164 events filtered to exclude Ml < 2) and anomalous values

Noise filtering:

High-pass filter with a cut-off frequency of 0.075 Hz
low-pass filter with a cut-off frequency of 150 Hz.
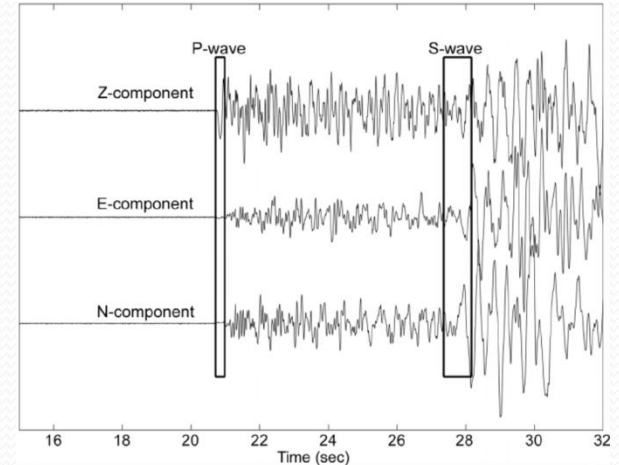
Three-component raw waveforms recorded directly at seismic station

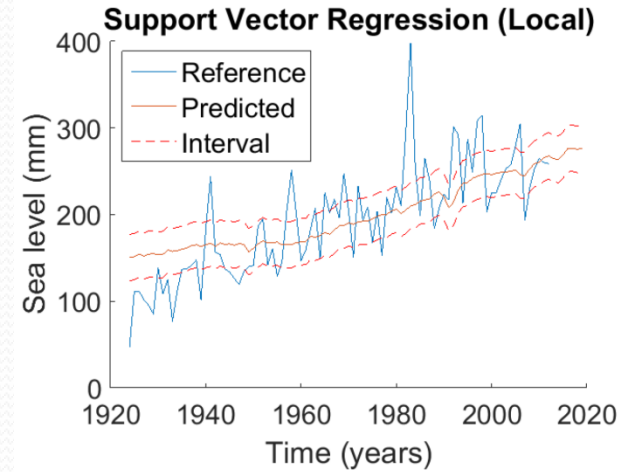# Climate Change and Sea Level Rise

Features
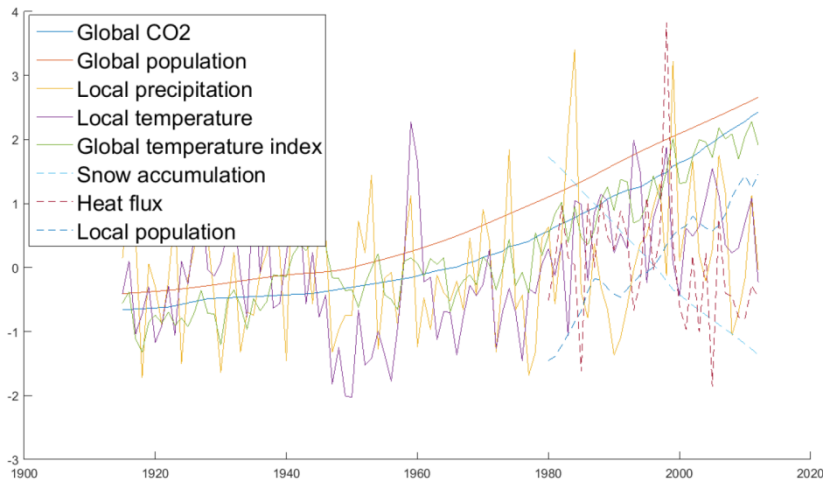
(Alahmadi & Kolmas, 2015)

Figure 2: Prediction of sea level rise in San Francisco.
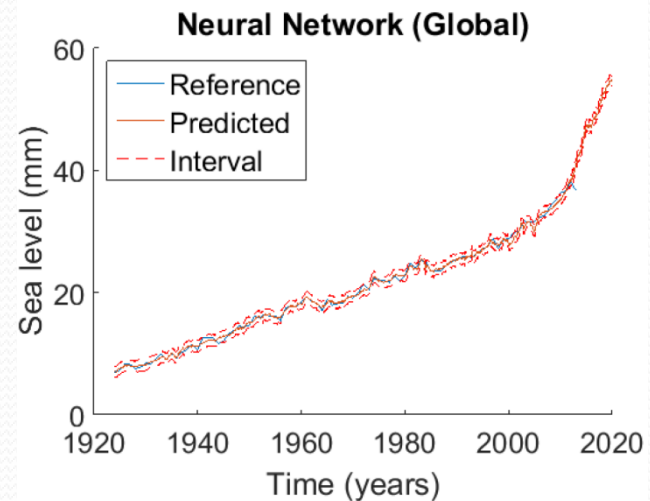
Figure 3: Prediction of global sea level rise.

Unsupervised Learning

Table I. Original air-pollution Database

Date	CO_2	RPM	SO_2	NO_X
1/1/2009	85	183	12	95
2/1/2009	95	289	14	125
3/1/2009	112	221	10	101
4/1/2009	114	191	11	97
5/1/2009	100	175	11	101
6/1/2009	78	149	7	93
……….	…….	…….	……	……..
1/2/2009	120	197	10	105
2/2/2009	115	151	10	85
3/2/2009	98	154	8	96
4/2/2009	90	195	8	93
……….	…….	…….	……	……..

Table VI. Weather forecasting from September, 2009 to June 2010

| Date	New data inserted into	Weather Category
1/9/2009	Cluster2	Hot, dry and smogy
2/9/2009	Cluster3	dusty, fly ash, smogy, fog, Mist
3/9/2009	Cluster2	Hot, dry and smogy
4/9/2009	Cluster2	Hot, dry and smogy
………...	……………..	………………
28/9/2009	Cluster3	dusty, fly ash, smogy, fog, Mist
29/9/2009	Cluster3	dusty, fly ash, smogy, fog, Mist

Chakraborty et al., 2014

RPM: Respirable particulate matter
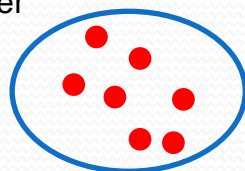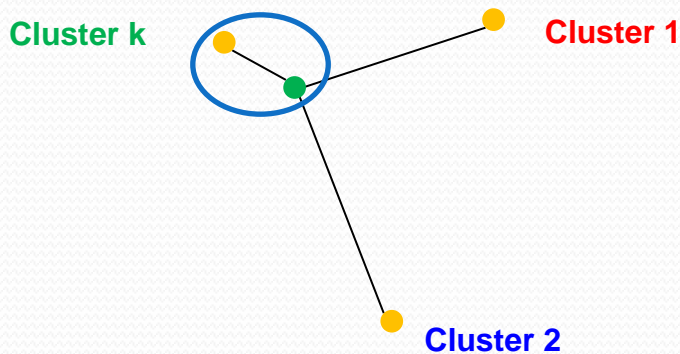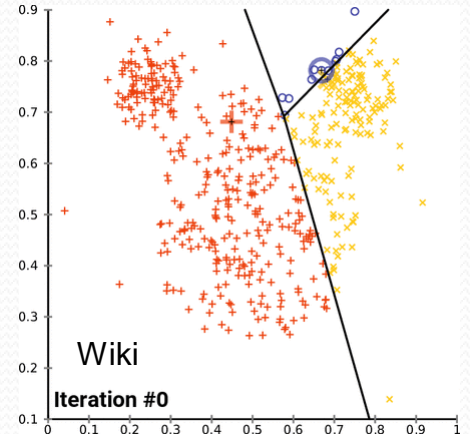
Cluster 1
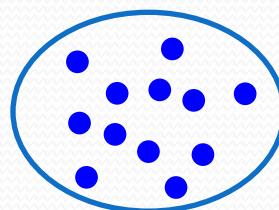
Cluster 2
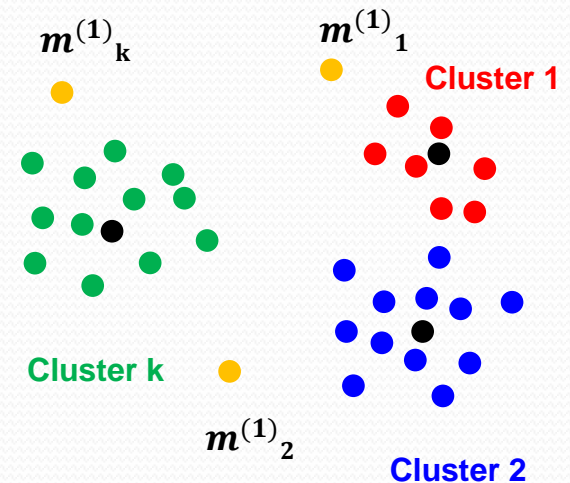
Cluster 3

15
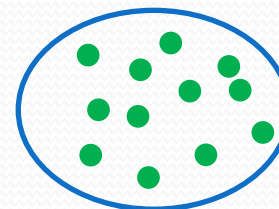
K-Means Clustering

Algorithm:

1) Assign **random means** (centroids) for K-cluster (orange points): $m^{(1)}_1, m^{(1)}_2, \ldots \quad \ldots m^{(1)}_k$

2) Assign each data point to the cluster whose mean has the **least** squared Euclidean distance (the "nearest" mean)

3) Calculate the **new mean** (centroid) for each cluster

$$m^{(t+1)}_i = \frac{\sum_{j=1}^{n_i} x_j}{n_i} \qquad 1 \le i \le k$$

4) Iterate until the centroids **do not change significantly.**

Wiki

Iteration #0

Cluster k

Cluster 1

Cluster 2

● **Initial centroids (Start)**

$m^{(1)}_k$

$m^{(1)}_1$

Cluster 1

Cluster k

$m^{(1)}_2$

Cluster 2

● **Final centroids (End)**

16

Inverse Problem in Geodynamics

PUBLICATIONS

JGR

## Journal of Geophysical Research: Solid Earth

Inverse Problems in Geodynamics Using Machine Learning Algorithms

M. H. Shahnas[1,2] iD, D. A. Yuen[2] iD, and R. N. Pysklywec[1] iD

[1]Department of Earth Sciences, University of Toronto, Toronto, Ontario, Canada, [2]Department of Earth Sciences and Minnesota Supercomputing Institute, University of Minnesota, Twin Cities, Minneapolis, MN, USA

Key Points:
- We show that machine learning techniques can be successfully employed in geodynamics problems
- Machine learning algorithms can be used in solving of inverse problems in geodynamics

Correspondence to:
M. H. Shahnas,
shahnas@umn.edu

Abstract During the past few decades numerical studies have been widely employed to explore the style of circulation and mixing in the mantle of Earth and other planets. However, in geodynamical studies there are many properties from mineral physics, geochemistry, and petrology in these numerical models. Machine learning, as a computational statistic-related technique and a subfield of artificial intelligence, has rapidly emerged recently in many fields of sciences and engineering. We focus here on the application of supervised machine learning (SML) algorithms in predictions of mantle flow processes. Specifically, we emphasize on estimating mantle properties by employing machine learning techniques in solving an inverse problem. Using snapshots of numerical convection models as training samples, we enable machine learning models to determine the magnitude of the spin transition-induced density anomalies that can cause flow stagnation at midmantle depths. Employing support vector machine algorithms, we show that SML techniques can successfully predict the magnitude of mantle density anomalies and can also be used in characterizing mantle flow patterns. The technique can be extended to more complex geodynamic problems in mantle dynamics by employing deep learning algorithms for putting constraints on properties such as viscosity, elastic parameters, and the nature of thermal and chemical anomalies.

Iron Spin Transition in the Lower Mantle
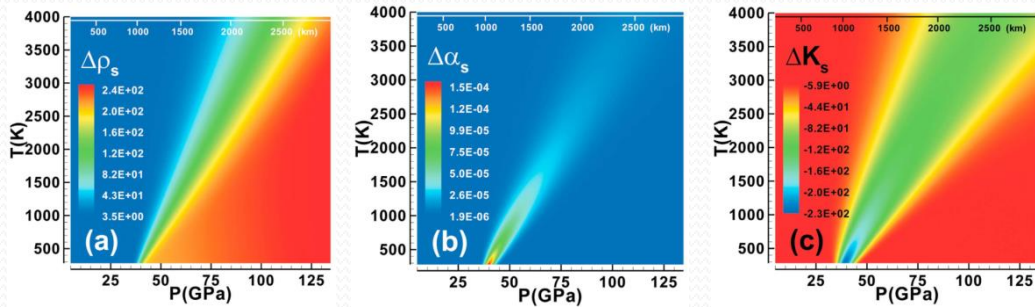
(a) density in kg/m^3, (b) thermal expansivity in 1/K, (c) bulk modulus in GPa, in **Fp** [Wu et al., 2009, Shahnas et al., 2011]
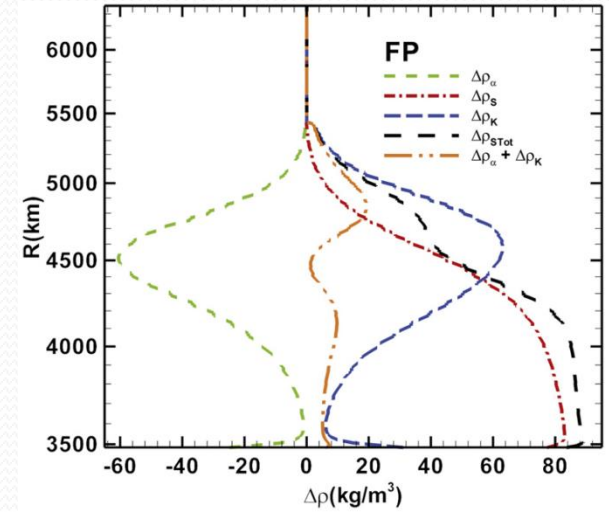
$$V(n) = nV_{LS}(P,T) + (1-n)V_{HS}(P,T)$$

$$\alpha(n)V(n) = nV_{LS}\alpha_{LS} + (1-n)V_{HS}\alpha_{HS} + (V_{LS} - V_{HS})\frac{\partial n}{\partial T}\bigg|_P$$
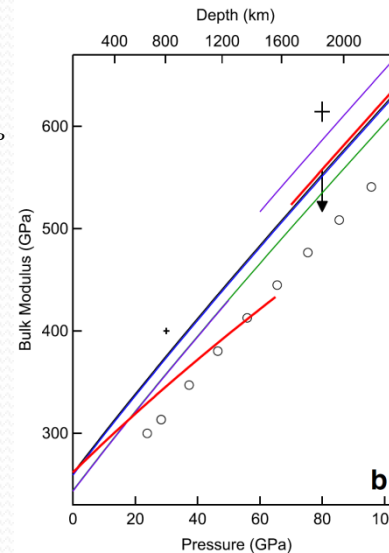
$$\frac{V(n)}{K_T(n)} = n\frac{V_{LS}}{K_{LS}^T} + (1-n)\frac{V_{HS}}{K_{HS}^T} - (V_{LS} - V_{HS})\frac{\partial n}{\partial P}\bigg|_T$$

$$\Delta\rho_{STot} = \Delta\rho_s + \Delta\rho_\alpha + \Delta\rho_K$$

$$\Delta\rho_\alpha = -\rho_0\Delta\alpha_s(T - T_0)$$

$$\Delta\rho_K = \rho_0\left[\left(\frac{1}{K_T + \Delta K_s} - \frac{1}{K_T}\right)(p - p_0)\right]$$

$$\rho = \rho_r\left[1 - \alpha(T - T_r)\right] + \Delta\rho_i(\Gamma_i - \Gamma_{ri}) + \frac{1}{K_T}(p - p_r) + \Delta\rho_s$$


The effect of spin transition on the bulk modulus of **Pv**. Fe^{3+} Al: Catalli et a. (2011),Fe^{2+} : (Lundin et al. (2008), Al: Yagi et al. (2004), Fe^{3+} : Catalli et al. (2010), Mg-Pv: Lundin et al. (2008), PREM: Dziewonski and Anderson (1981) [Catalli et al, 2011].

Mid-mantle Seismic Heterogeneities

Mid-mantle stagnations are prevalent globally in seismic tomographic inversions, but previous explanations for their existence are not satisfactory.

Iron spin transition in the lower mantle minerals can significantly influence the thermoelastic properties of the mantle material.

Numerical experiments explore how the electronic spin transition in iron modifies the mantle flow, and in particular the fate of sinking cold slabs and rising plumes [Shahnas et al. JGR, 2011; Shahnas et al., G^3, 2016; Shahnas et al. GJI, 2017; Shahnas et al. EPSL, 2017; Li et al., 2018].
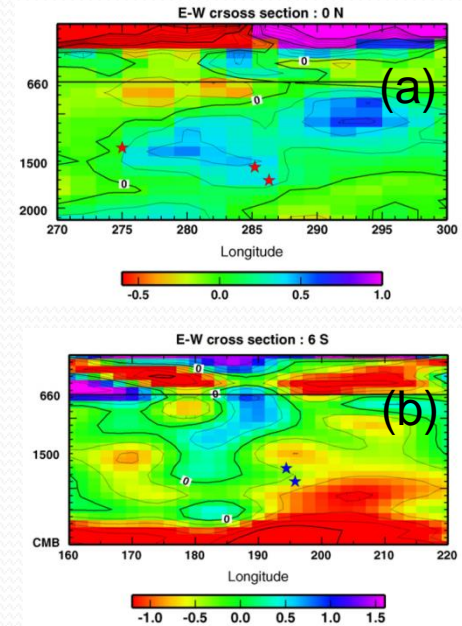
(a) Mid-mantle slabs in the South America region, (b) Pacific superplume [Kaneshima and Helffrich, 2010]
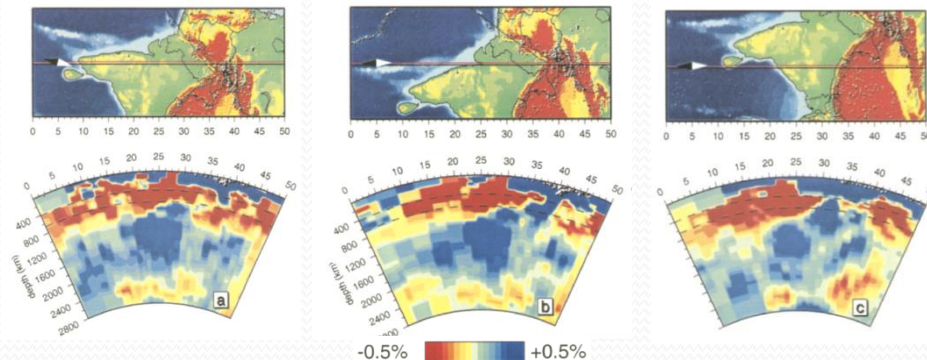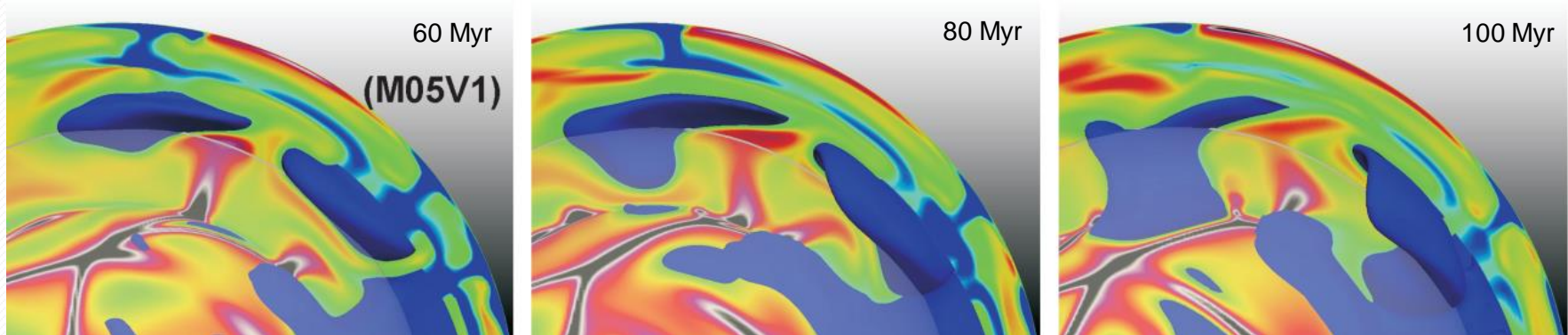

Tomographic results displayed in three vertical sections (a-c) along the lines indicated in the maps, showing the P-wave velocity anomalies under India and adjacent areas. [Van der Voo, 1999].
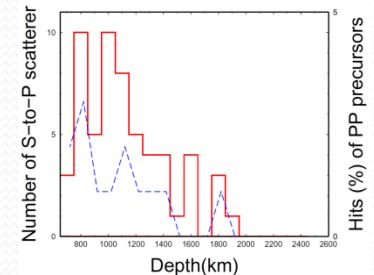
Red line: The histogram of the number of mid-lower mantle S-to-P scatterers in the western Pacific as a function of depth (Kaneshima,2016)

19

Flow Stagnation

[Shahnas et al., EPSL, 2017]
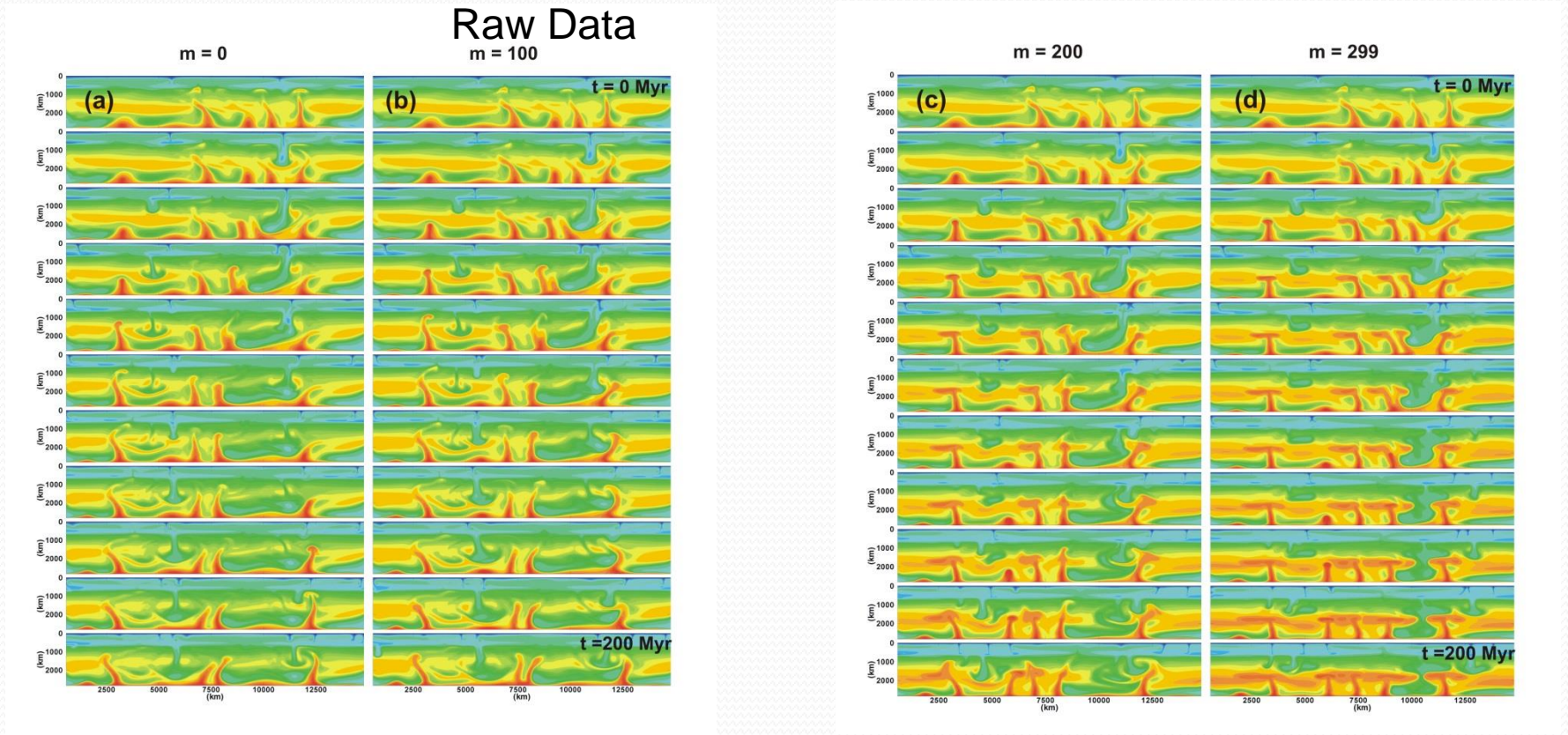
Stagnation in the Lower Mantle

Responsible Mechanisms:
1 – Viscosity
2 – Composition
3 – Phase Transition
4 – Dissociation Transition
5- Other Material Property Changes

Tackling the Problem
a) Inverse Problem Approach
b) Machine Learning Approach

Testing and Training Samples

Raw Data

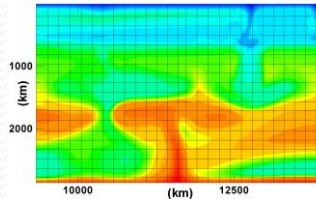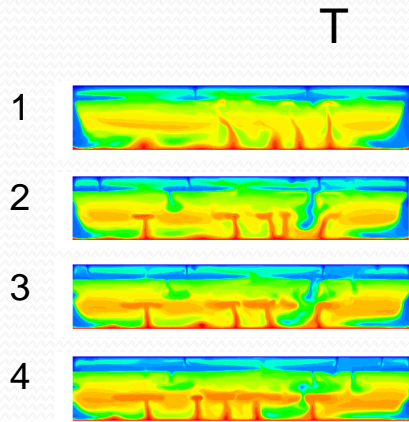$$\Delta \rho_{STot(m)} = \Delta \rho_{sFp} + \Delta \rho_{\alpha Fp} + \Delta \rho_{KFp} + m\beta \Delta \rho_{KPv}$$

$$\left| \Delta \rho_{KPv} \right| = \left| \Delta \rho_{KFp} \right|, \qquad \beta = 0.01, \qquad 0 \le m \le 299.$$

[Wu and Wentzcovitch, 2009; Catalli, 2010; Shahnas et al., 2011, 2017]

n samples (images)
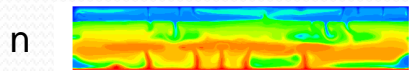
2D: m x m pixels

T

1

2

3

4
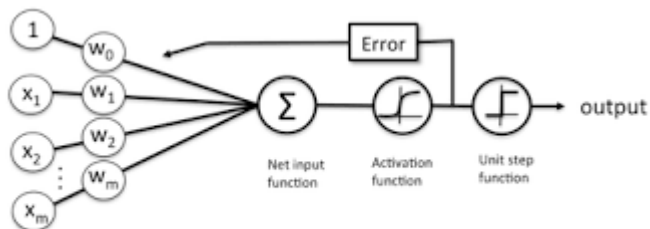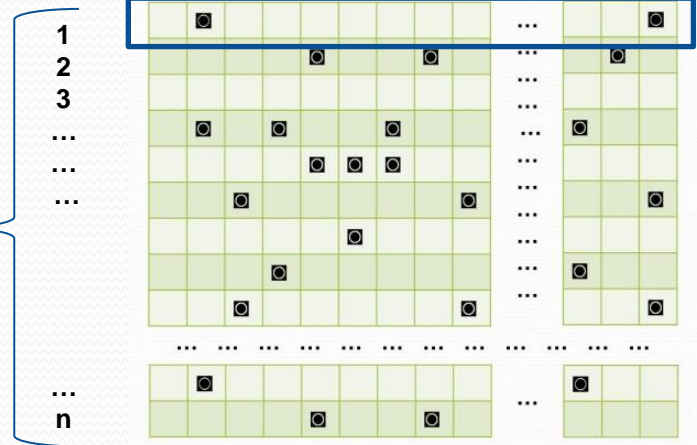
n

2D to 1D array

1D: m^2

...

m^2

1
2
3
...
...
...

...
n

n images, one per line

Schematic of a logistic regression classifier.

output

Error

Σ

1 — w_0

x_1 — w_1

x_2 — w_2

x_m — w_m

Net input function

Activation function

Unit step function

Feature Reduction

Raw Data
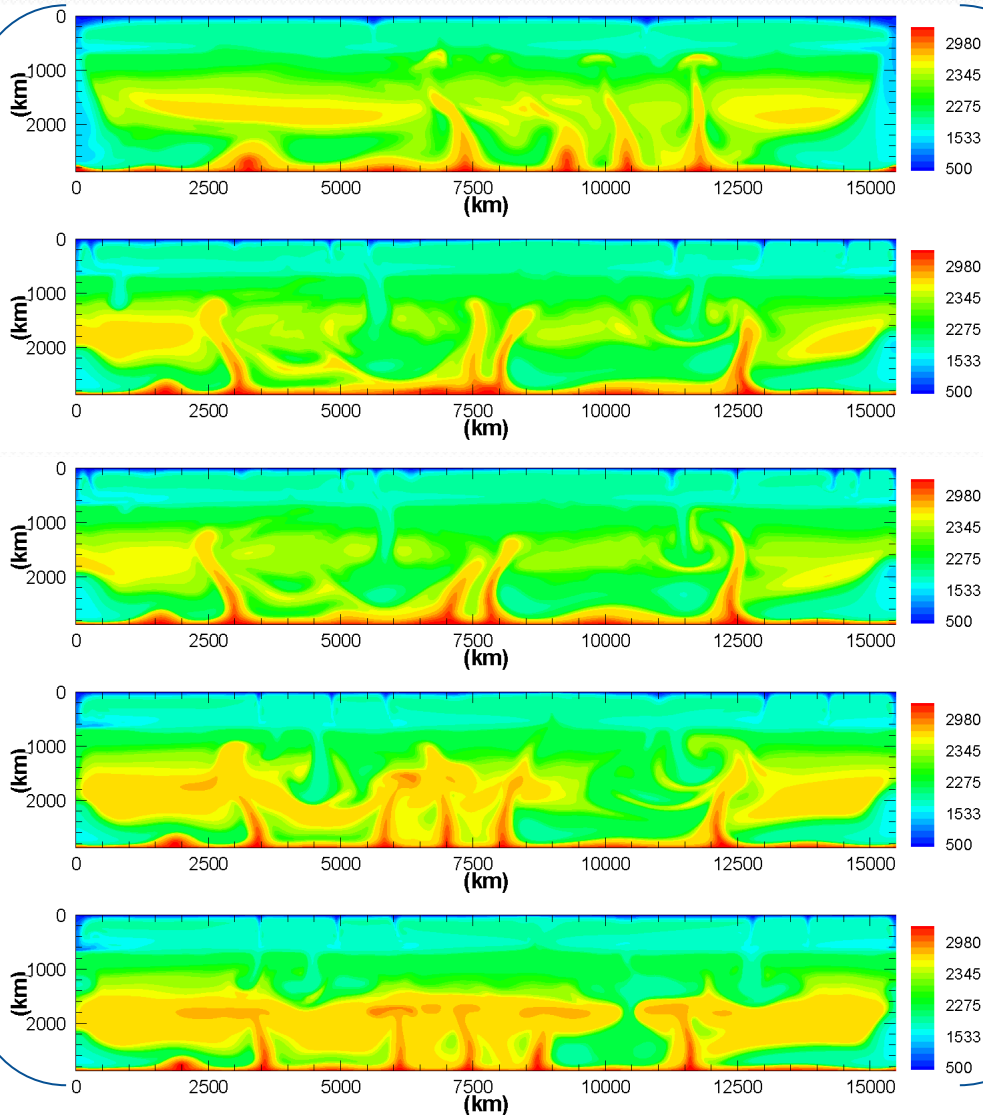
Meaningful Features
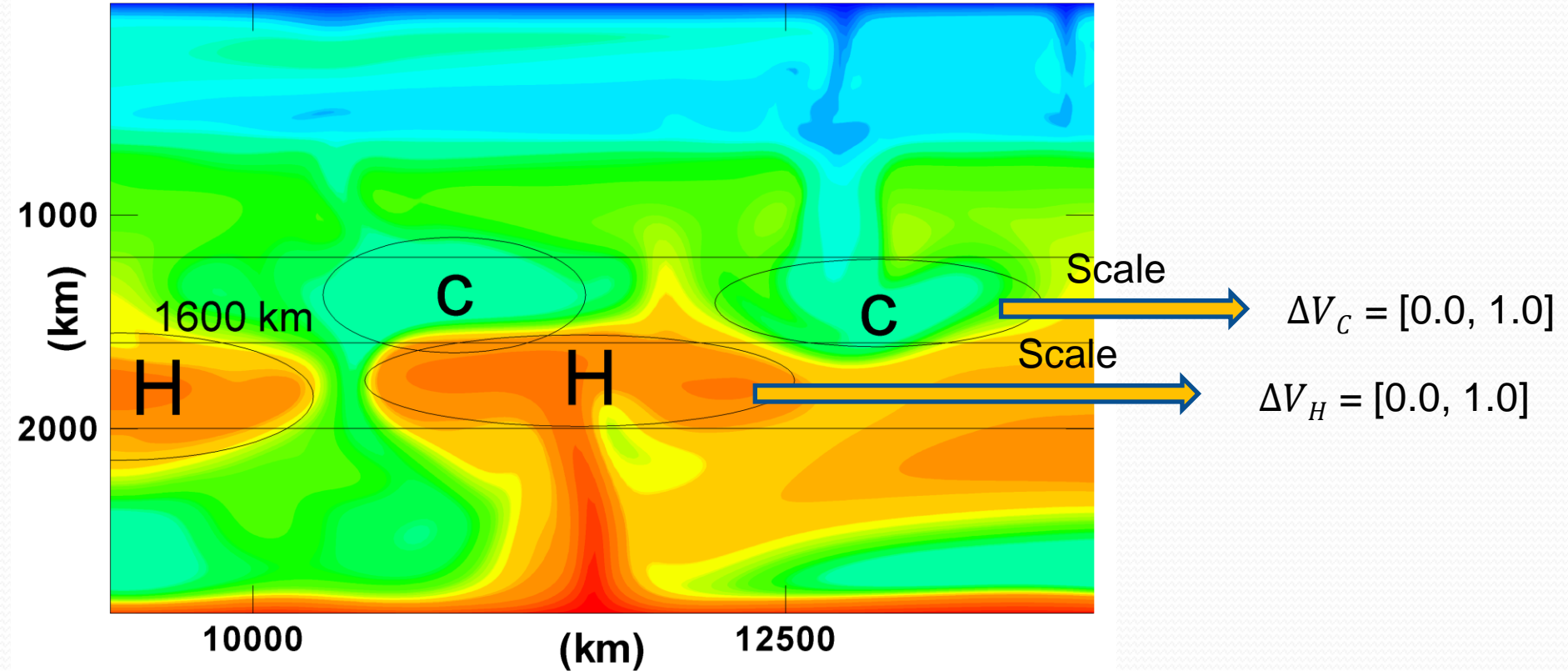
$$T_C < T_{ave}^{up} - 60^o \qquad 1200 \le d_{up} \le 1600 \ km \qquad \text{Cold material}$$

$$T_H > T_{ave}^{dn} + 60^o \qquad 1600 \le d_{dn} \le 2000 \ km \qquad \text{Hot material}$$

24

Class Labels in Inverse Problem

Table1 - Normalized volume fractions (features) representing the degrees of the slab (ΔV_C) and plume (ΔV_H) stagnation.

$\Delta V_C < 0.11$	$0.11 \leq \Delta V_C < 0.20$	$0.20 \leq \Delta V_C < 0.26$	$\Delta V_C \geq 0.26$
Degree 1 C_1	Degree 2 C_2	Degree 3 C_3	Degree 4 C_4

$\Delta V_H < 0.08$	$0.08 \leq \Delta V_H < 0.16$	$0.16 \leq \Delta V_H < 0.20$	$\Delta V_H \geq 0.20$
Degree 1 H_1	Degree 2 H_2	Degree 3 H_3	Degree 4 H_4

Table 2- a) Class labels based on the amount of the spin dependent density anomaly in Pv

| Class 1	Class 2 …	… Class 14	Class 15
$0 \leq m \leq 20$	$21 \leq m \leq 40\ldots$	$\ldots 261 \leq m \leq 280$	$281 \leq m \leq 299$

b) Stagnation class labels based on the stagnation features C_i and H_i, $1 \leq i \leq 4$

H_1-C_1	Class 1	H_1-C_2	Class 2	H_1-C_3	Class 3	H_1-C_4	Class 4
H_2-C_1	Class 5	H_2-C_2	Class 6	H_2-C_3	Class 7	H_2-C_4	Class 8
H_3-C_1	Class 9	H_3-C_2	Class 10	H_3-C_3	Class 11	H_3-C_4	Class 12
H_4-C_1	Class 13	H_4-C_2	Class 14	H_4-C_3	Class 15	H_4-C_4	Class 16

The Thermal State of the Planetary Mantle

$$\theta = \frac{g(f)}{1+f^{-\frac{3}{2}}} + C(f, Ra_B) \frac{\left((1+f+f^2)H/3\right)^{\alpha}}{Ra_B^{\beta}}$$

$$Q = \alpha Ra_B^{\beta} C^{\gamma}$$
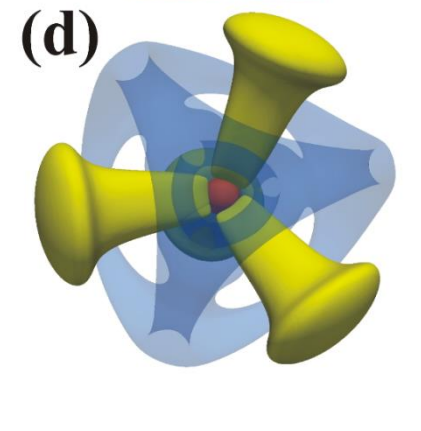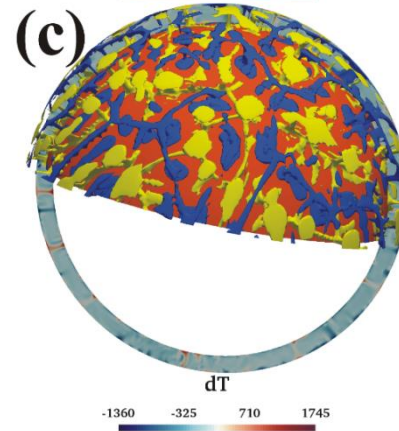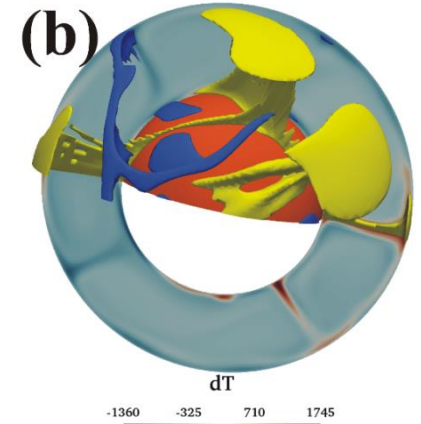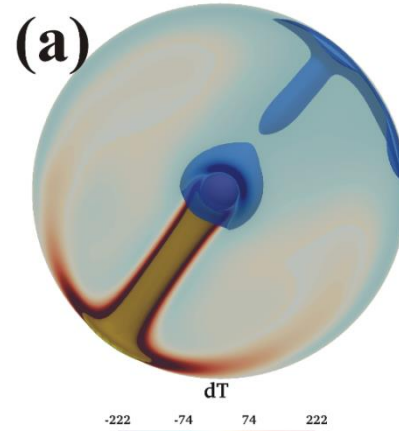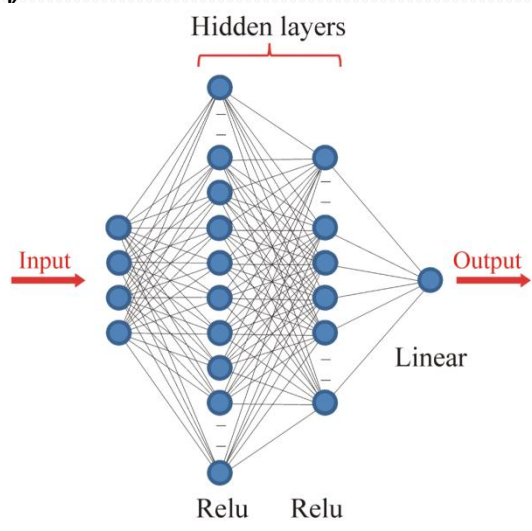
Shahnas & Pysklywec, 2019
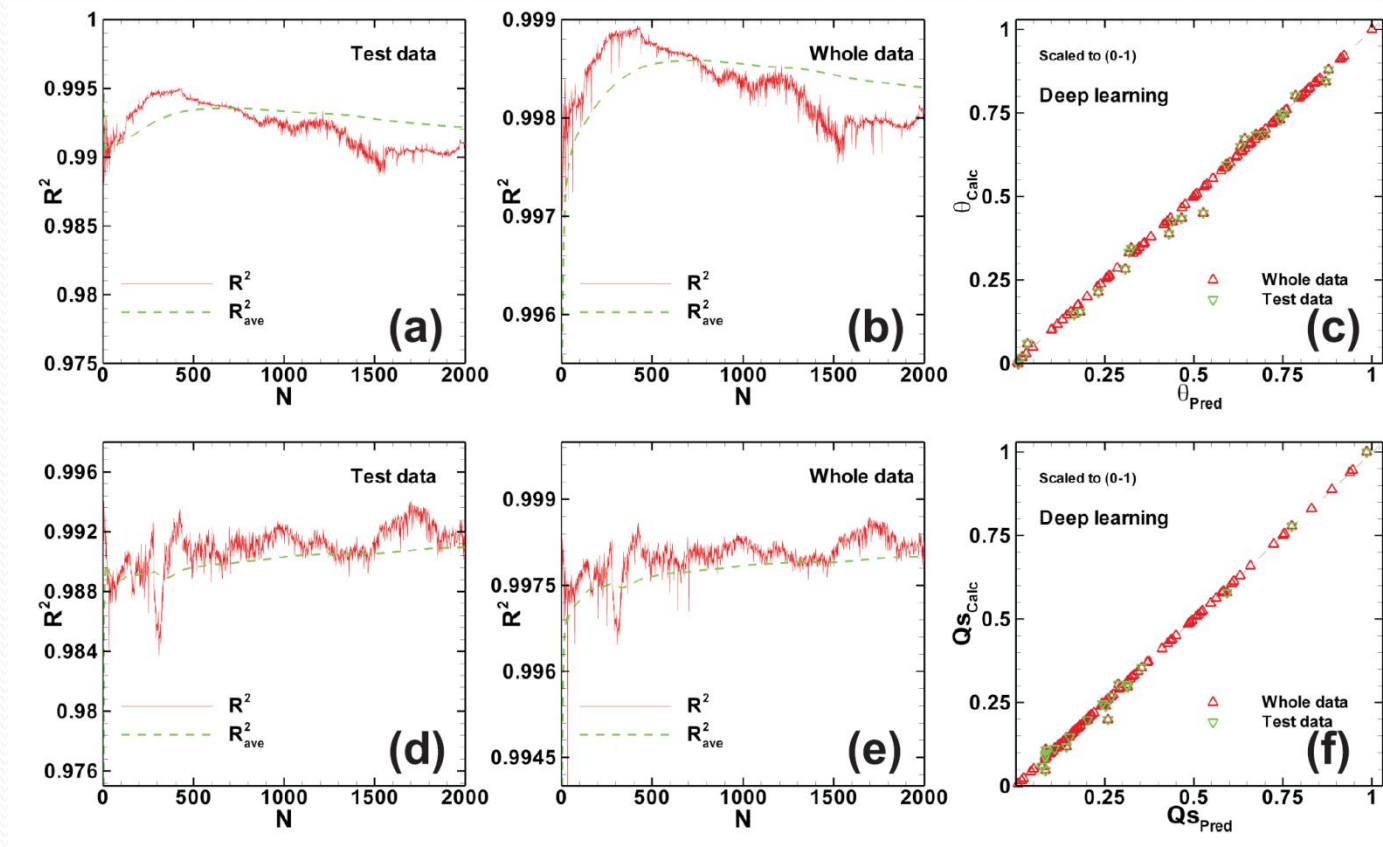
The Thermal State of the Planetary Mantle

Figure 7- R^2- and $\overline{R^2}$-scores for a) Test model, b) Whole data, for deep learning model specified in the text, for predicting mantle mean temperatures. c) Non-dimensional calculated mean temperatures (θ_{Calc}) versus the predicted values (θ_{Pred}). R^2- and $\overline{R^2}$-scores for d) Test model, e) Whole data, for deep learning model specified in the text, for predicting surface mean heat fluxes. f) Non-dimensional calculated surface mean heat fluxes (Qs_{Calc}) versus the predicted values (Qs_{Pred}). The non-dimensional mean heat fluxes are scaled to 0-1.