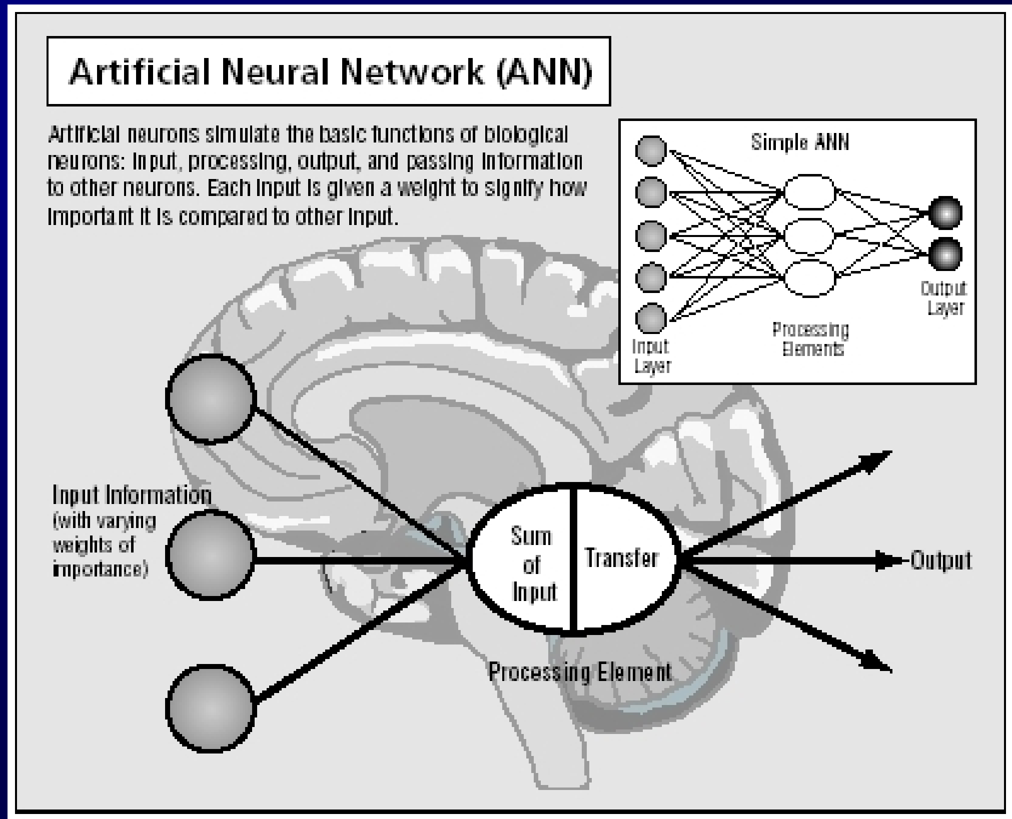
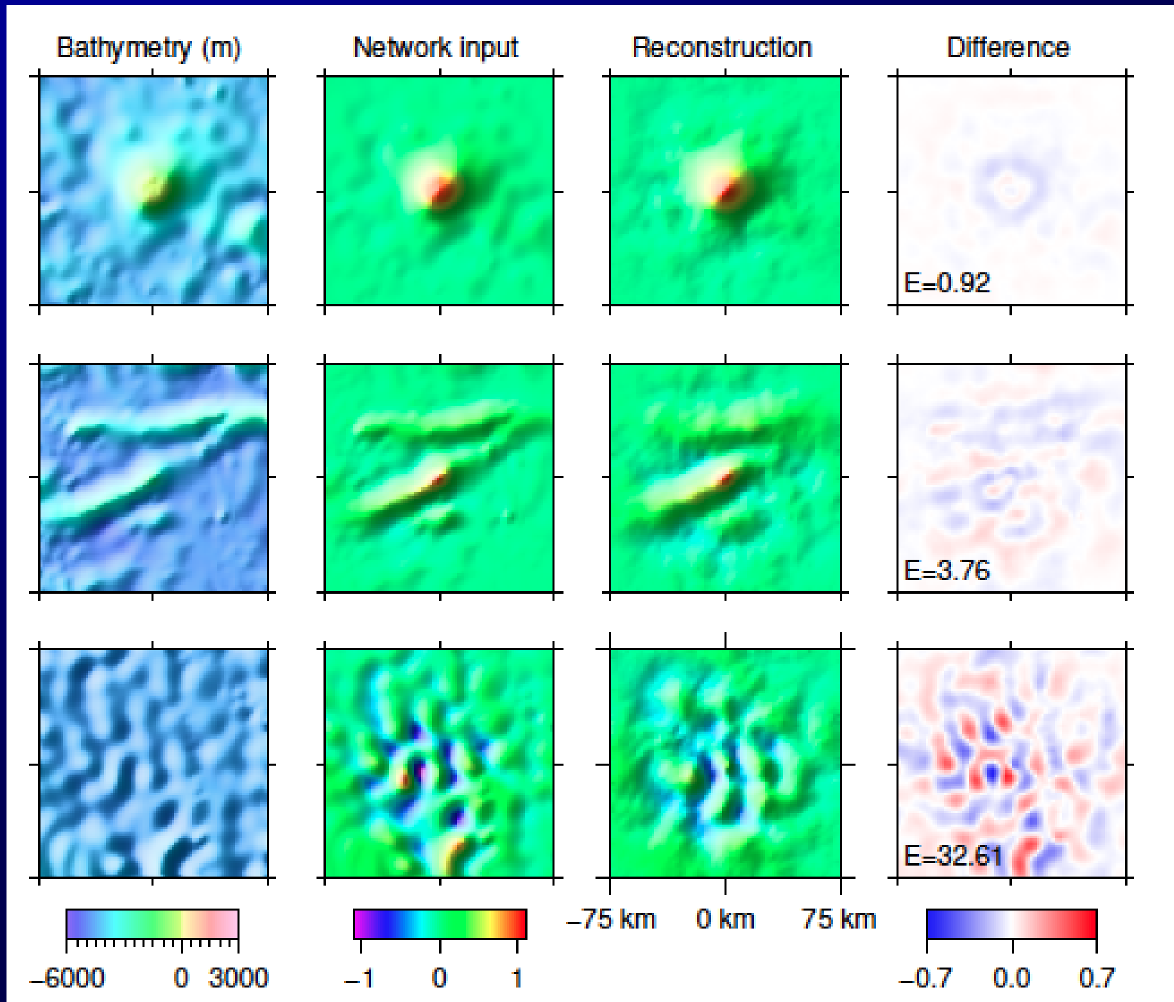
Neural Network Applications in Seismology

Jeannot Trampert
Utrecht University



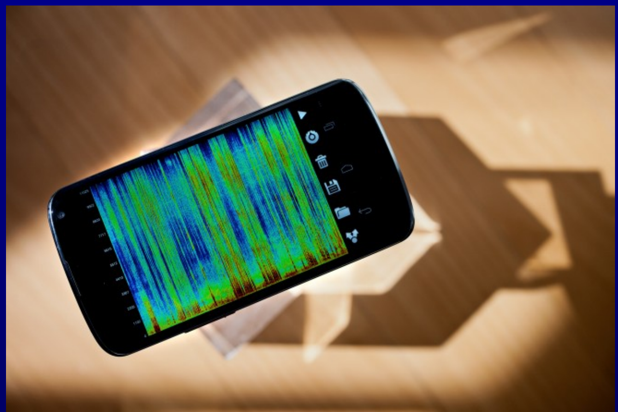
With major contributions
from many excellent PhD students and postdocs

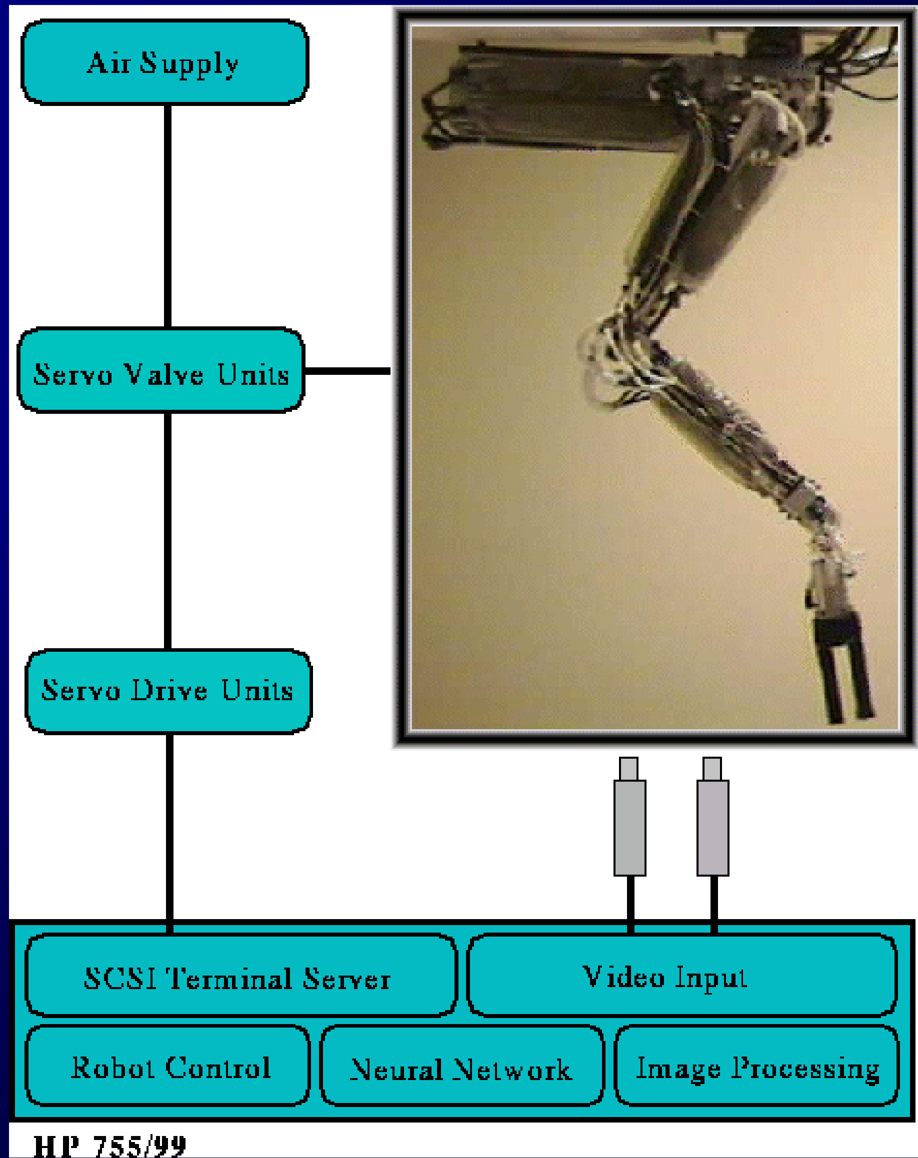
Neural networks are used for pattern recognition and classification



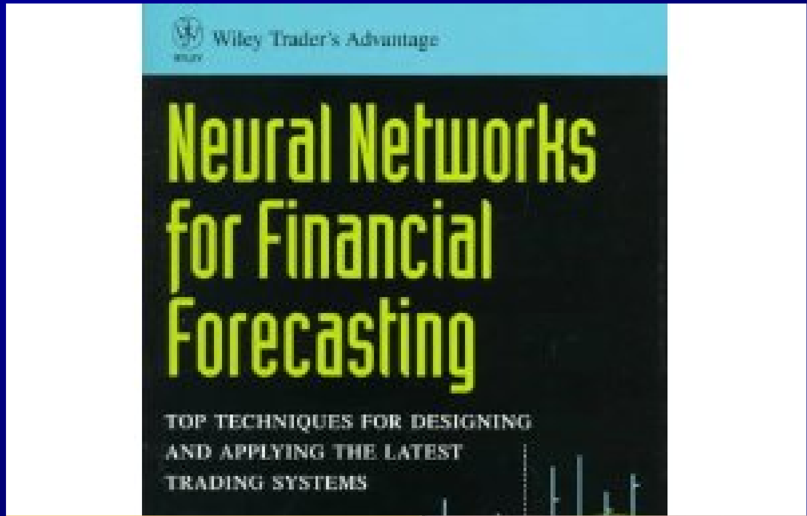
Valentine et al. (2013)

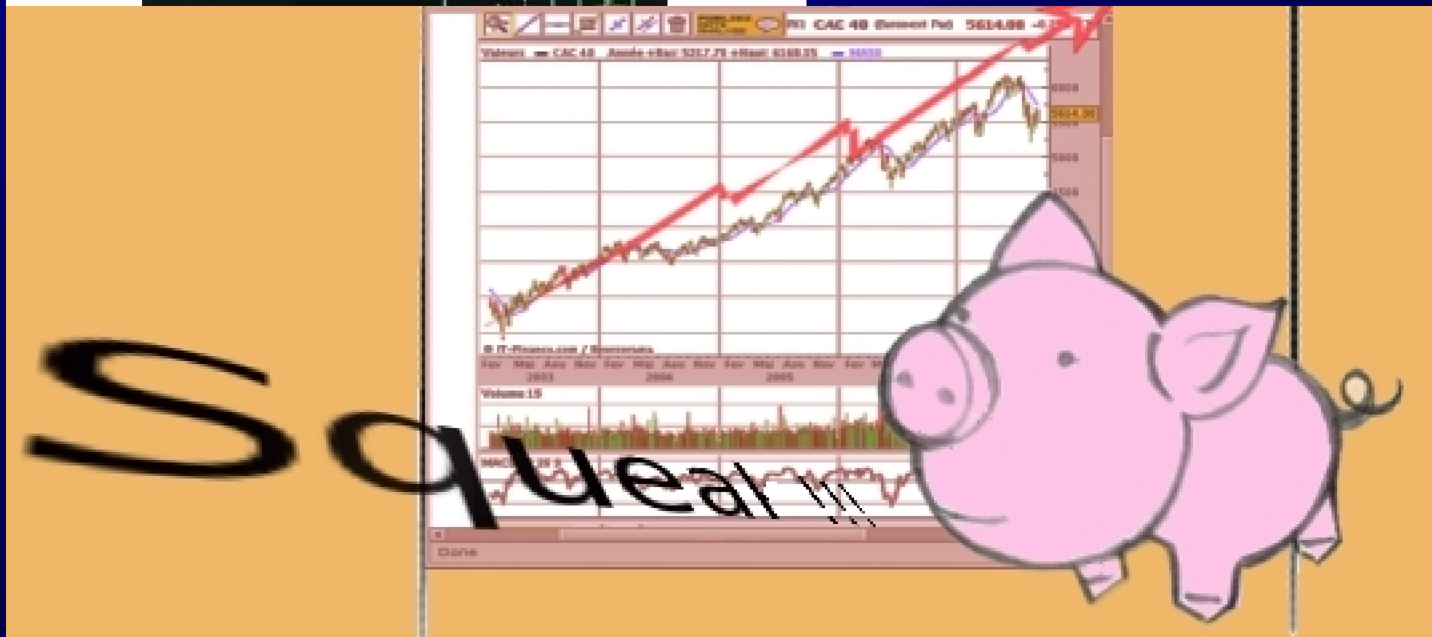
Neural networks are used for control systems



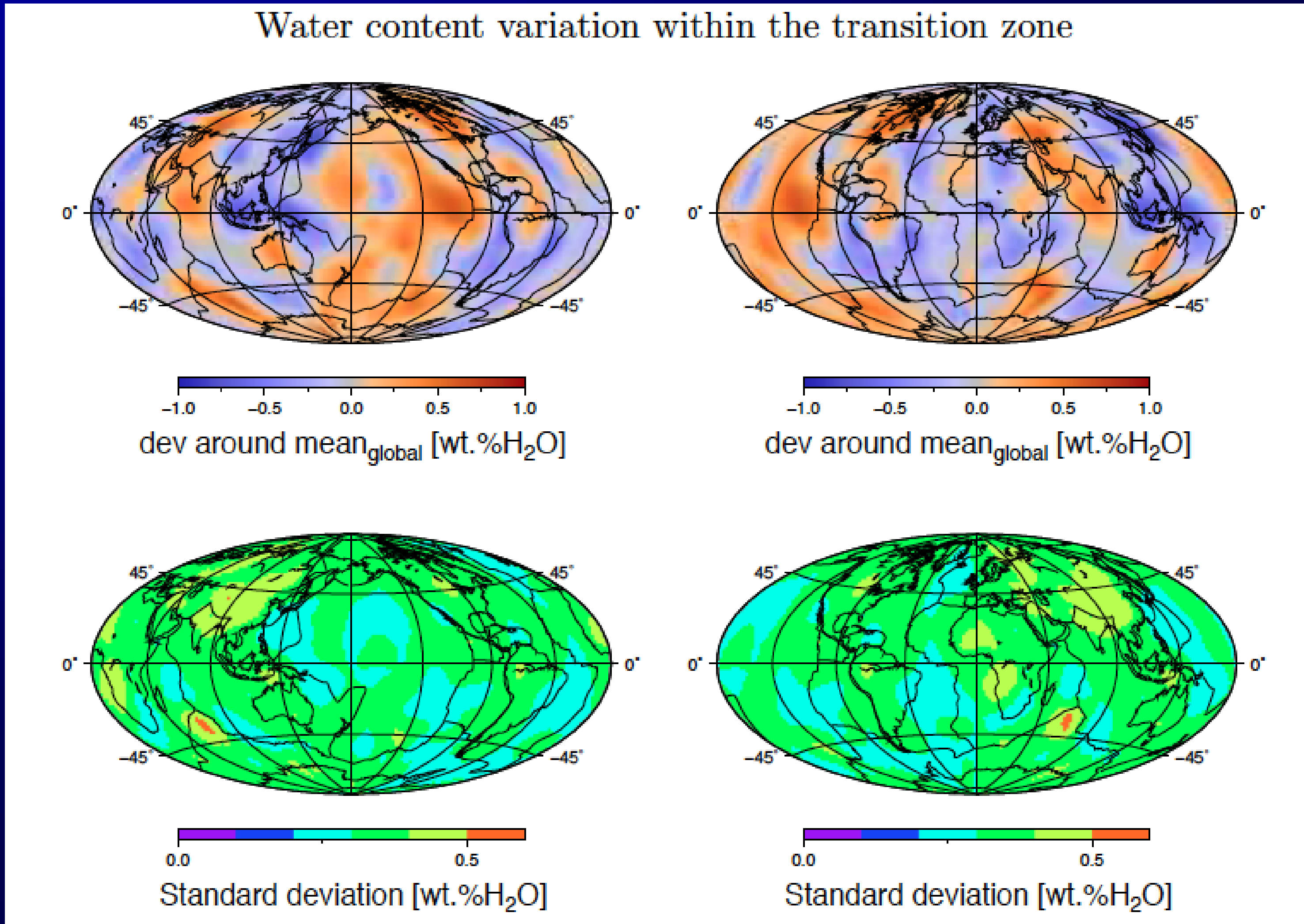


Neural networks are used for forecasting



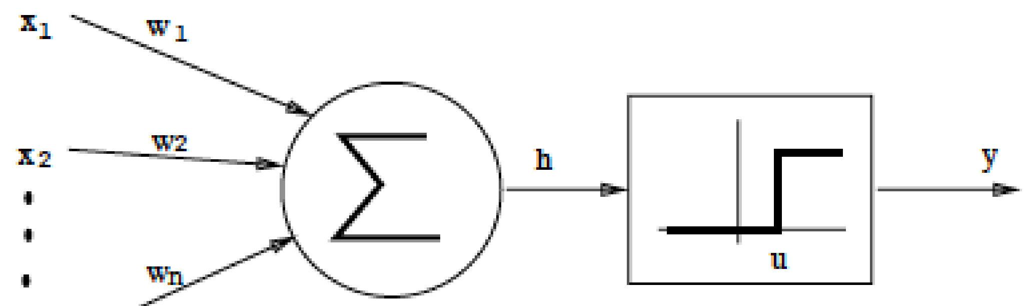


Neural networks are used for function approximation and optimization



Meier et al. (2009)

A single neuron



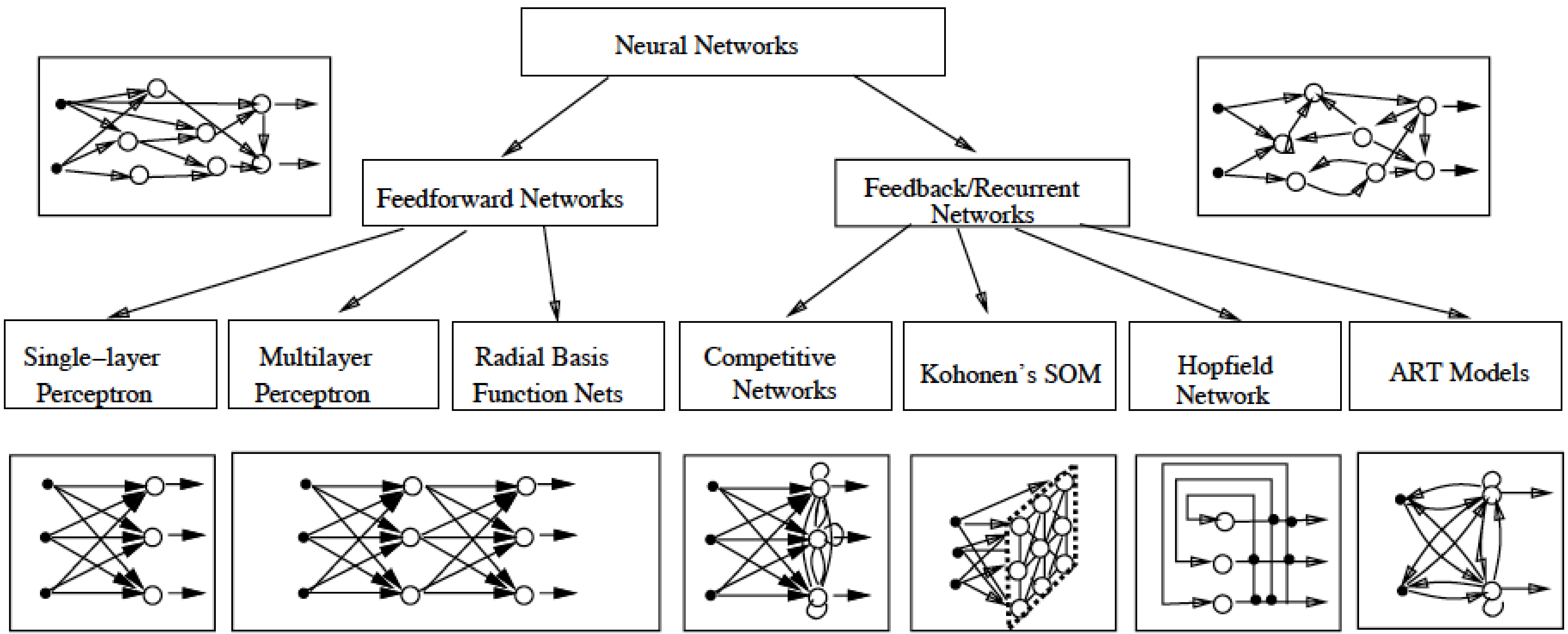
$$y = \theta\left(\sum_{j=1}^{n} w_j x_j - u\right)$$

Figure 3: McCulloch-Pitts model of a neuron. (1943)

A neural network

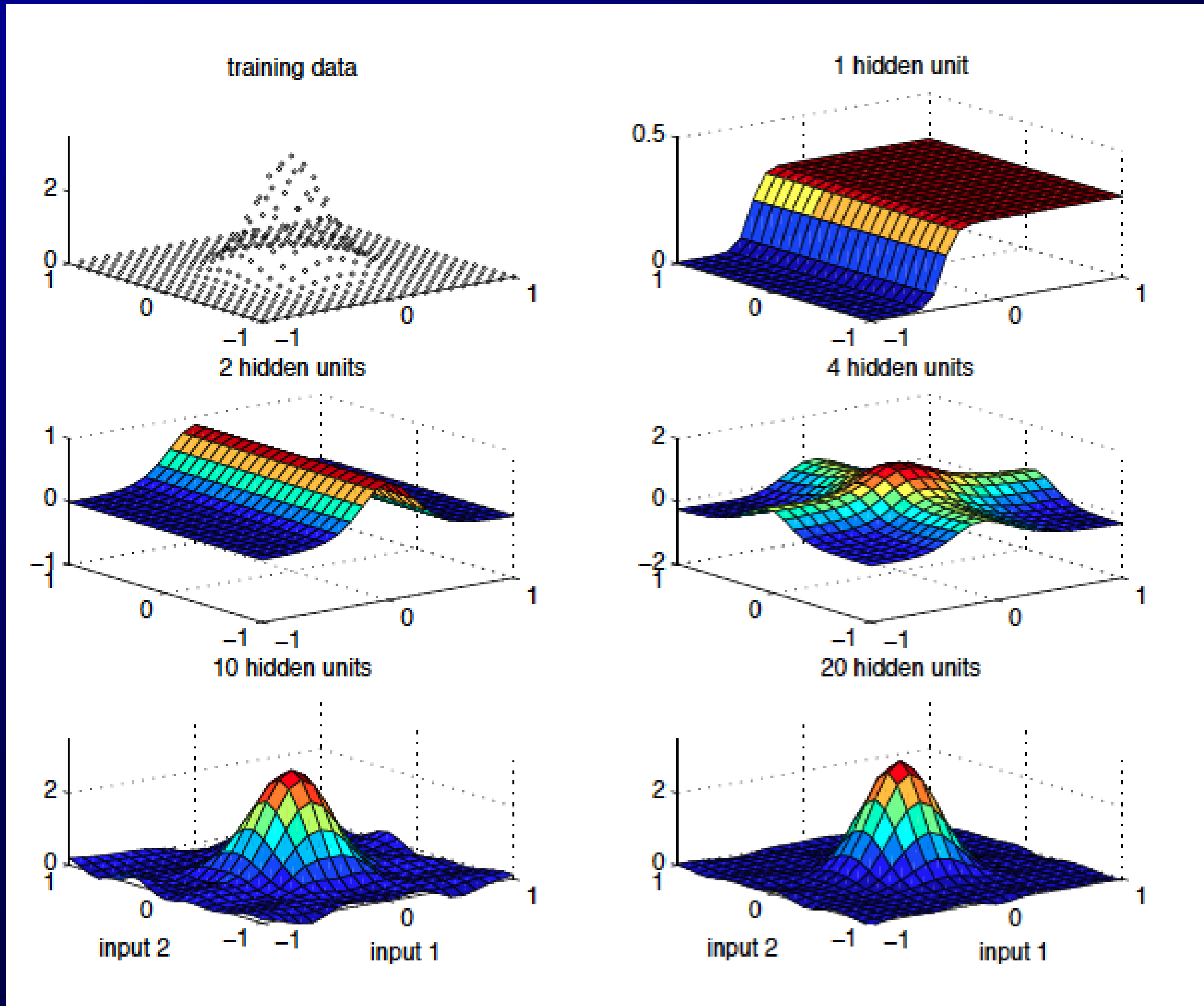
- Architecture
- Activity rule (choose)
- Learning rule (determine w)

Architecture

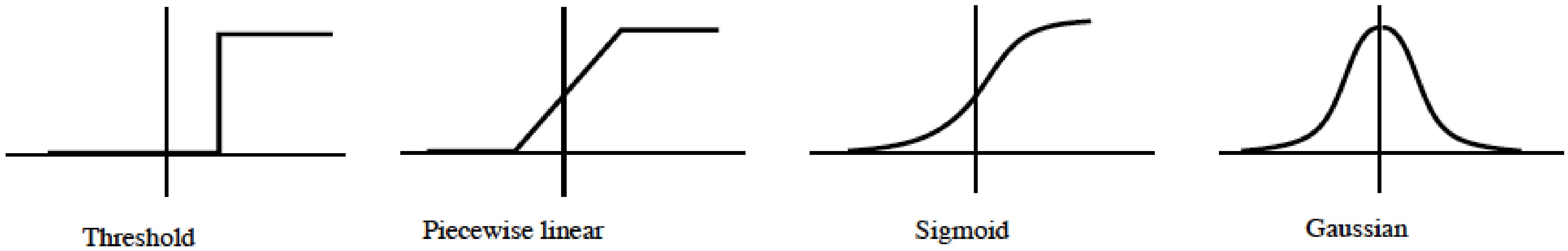


Jain et al. (1996)

Architecture: a Multi Layer Perceptron (MLP) can approximate an arbitrary continuous function



Activity rule

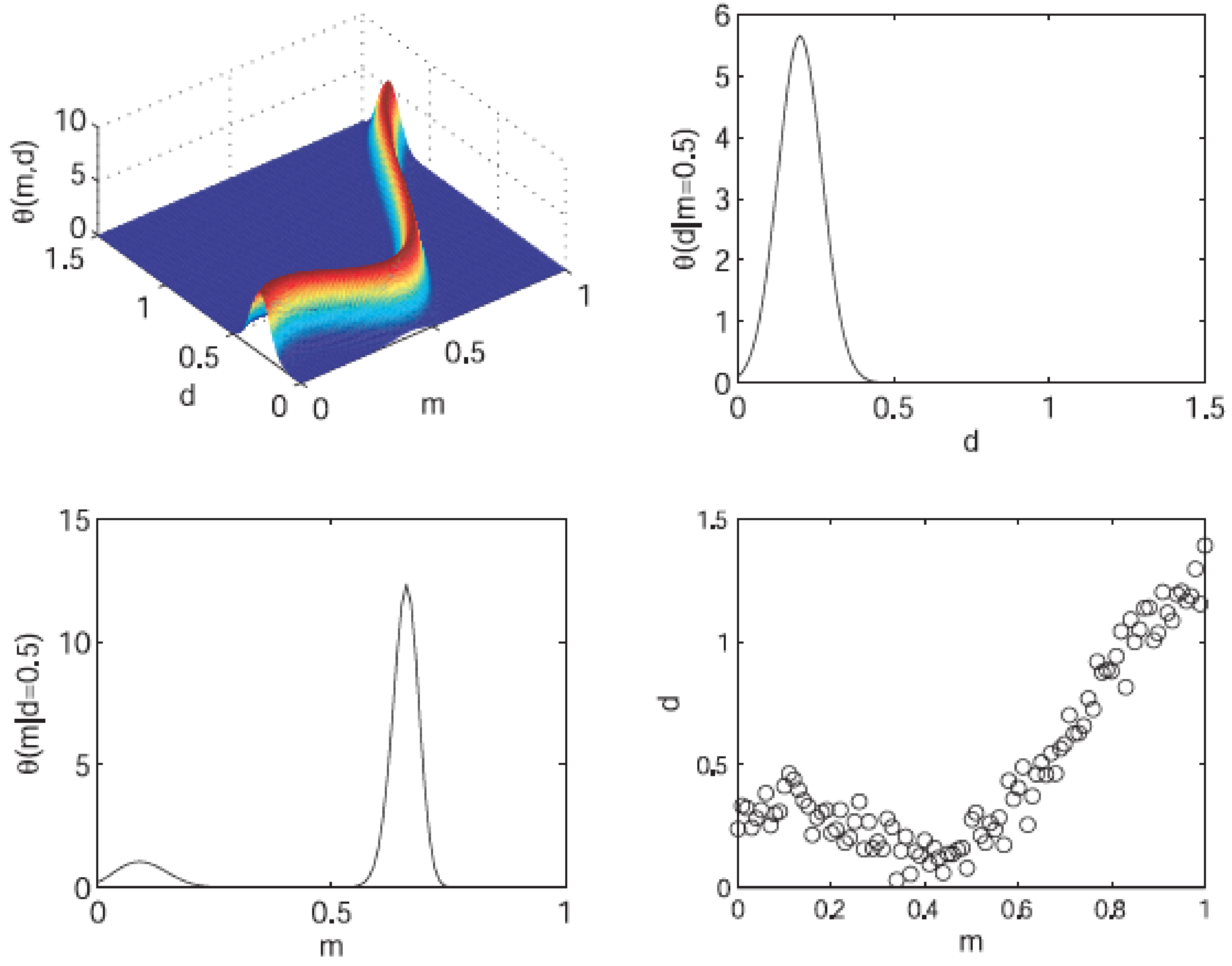


We often use the tanh sigmoid

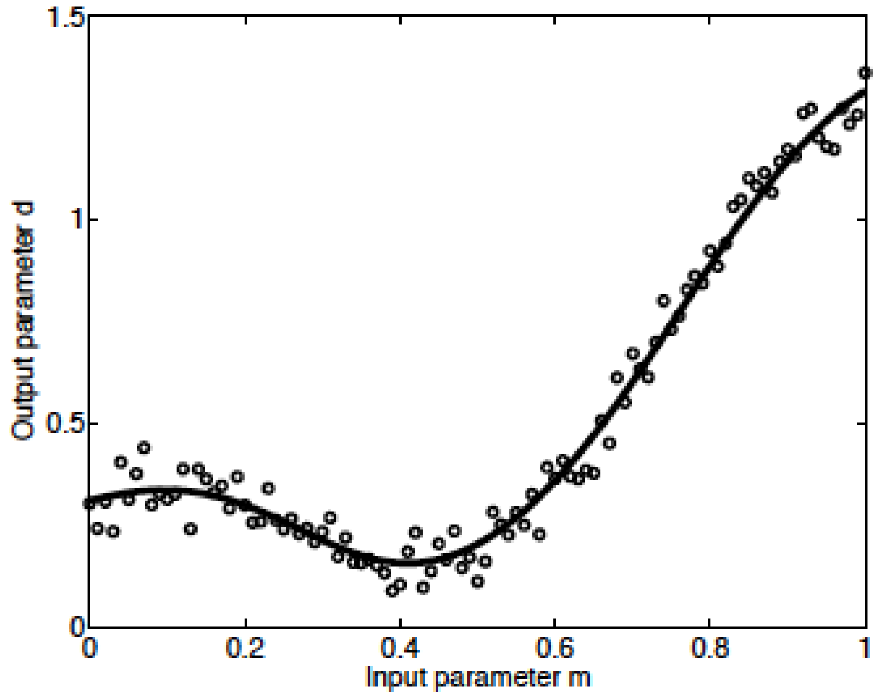
Learning rule

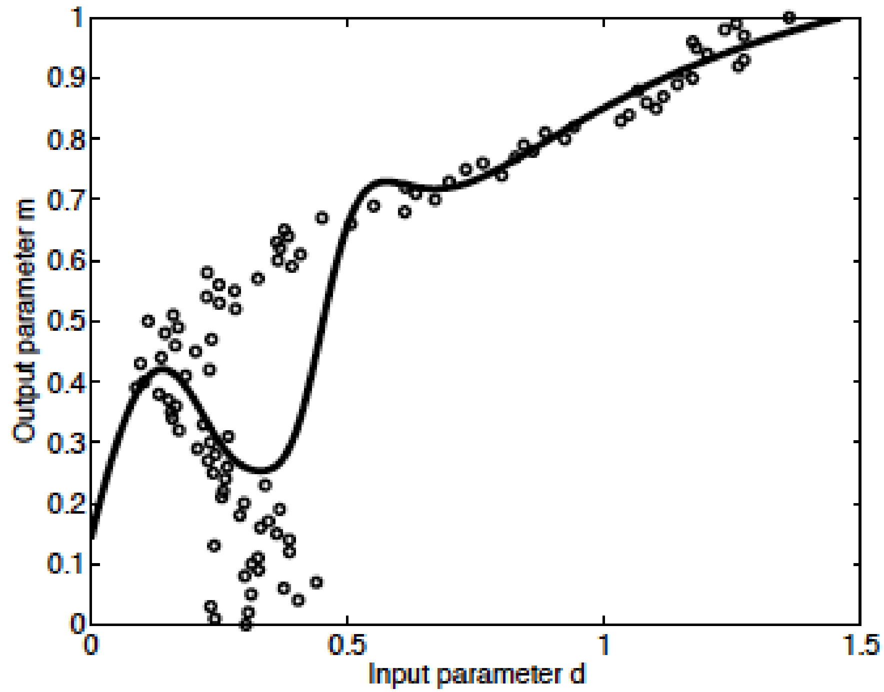
Paradigm	Learning Rule	Architecture	Learning Algorithm	Task
Supervised	Error-correction	Single- or Multi-layer Perceptron	Perceptron learning algorithms Backpropagation ADALINE & MADALINE	pattern classification function approximation prediction, control
	Boltzmann	Recurrent	Boltzmann learning algorithm	pattern classification
	Hebbian	Multi-layer Feedforward	Linear discriminant analysis	data analysis pattern classification
	Competitive	Competitive	Learning vector quantization	within-class categorization data compression
		ART network	ARTMAP	pattern classification within-class categorization
Unsupervised	Error-correction	Multi-layer Feedforward	Sammon's projection	data analysis
	Hebbian	Feedforward or Competitive	Principal component analysis	data analysis data compression
		Hopfield Net	Associative memory learning	associative memory
	Competitive	Competitive	Vector quantization	categorization data compression
		Kohonen SOM	Kohonen's SOM	categorization data analysis
		ART networks	ART1, ART2	categorization
Hybrid	Error-correction & Competitive	RBF network	RBF learning algorithm	pattern classification function approximation prediction, control

Example

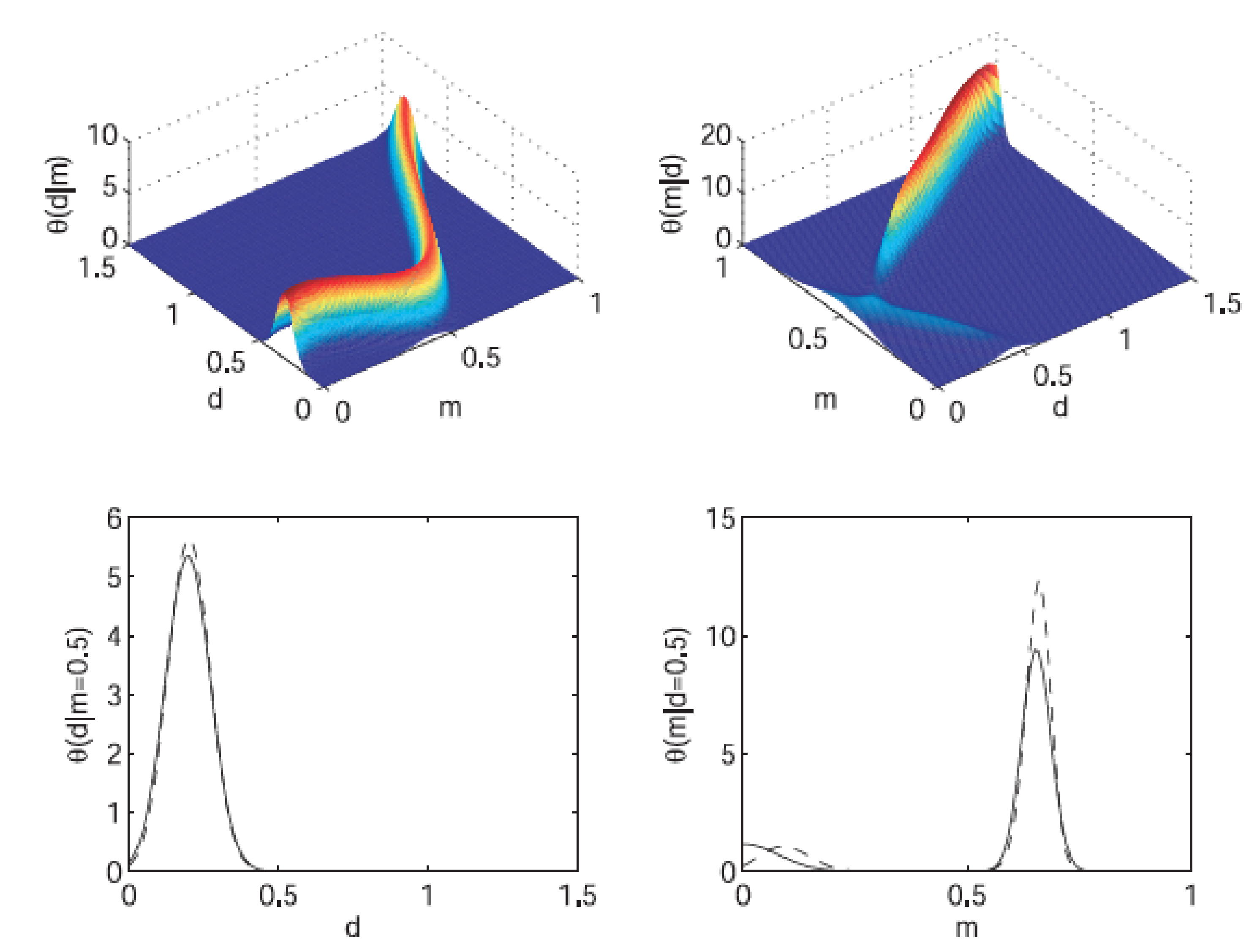

θ(m,d)
10
5
0
1.5
1
0.5
d
0
0
m
0.5
1
θ(d|m=0.5)
6
5
4
3
2
1
0
0.5
1
1.5
d
θ(m|d=0.5)
15
10
5
0
0
0.5
1
m
d
1.5
1
0.5
0
0
0.2
0.4
0.6
0.8
1
m

Example: The MLP fails if the function is multi-valued





Example: The Mixture Density Network extension works in the multi-valued case



We have now all the tools to solve a Bayesian inference problem using a MDN

$$\sigma(m|d) = k\rho(m)L(m)$$

$$\sigma(m|d) = \sum_{j=1}^{K} \alpha_j(d,w)\phi_j(m|d,w)$$

$$\phi_j = (m|d,w) = \frac{k}{\eta_j(d,w)} \exp\left[-\frac{\left\| m - \mu_j(d,w) \right\|^2}{2\eta_j(d,w)^2} \right]$$

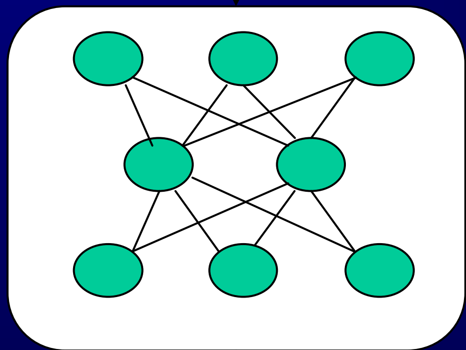
$_j$, $_j$ and $_j$ are outputs from standard MLP

3) Forward propagating a new datum through the trained network (i.e. solving the inverse problem)

Network Input:
Observed or synthetic data for training

$\mathbf{d}^{obs}$

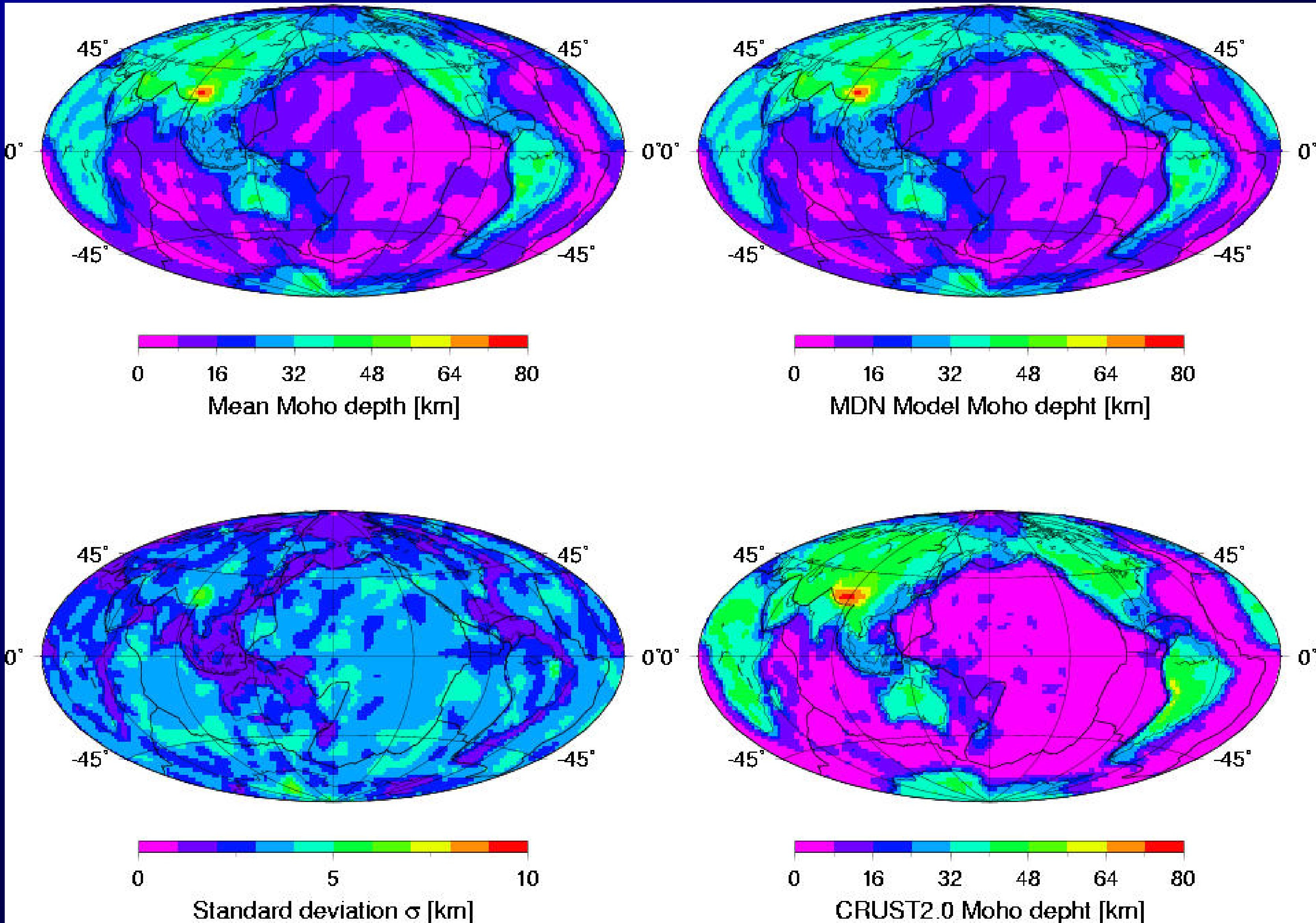
Neural Network



Network Output:
Conditional probability density

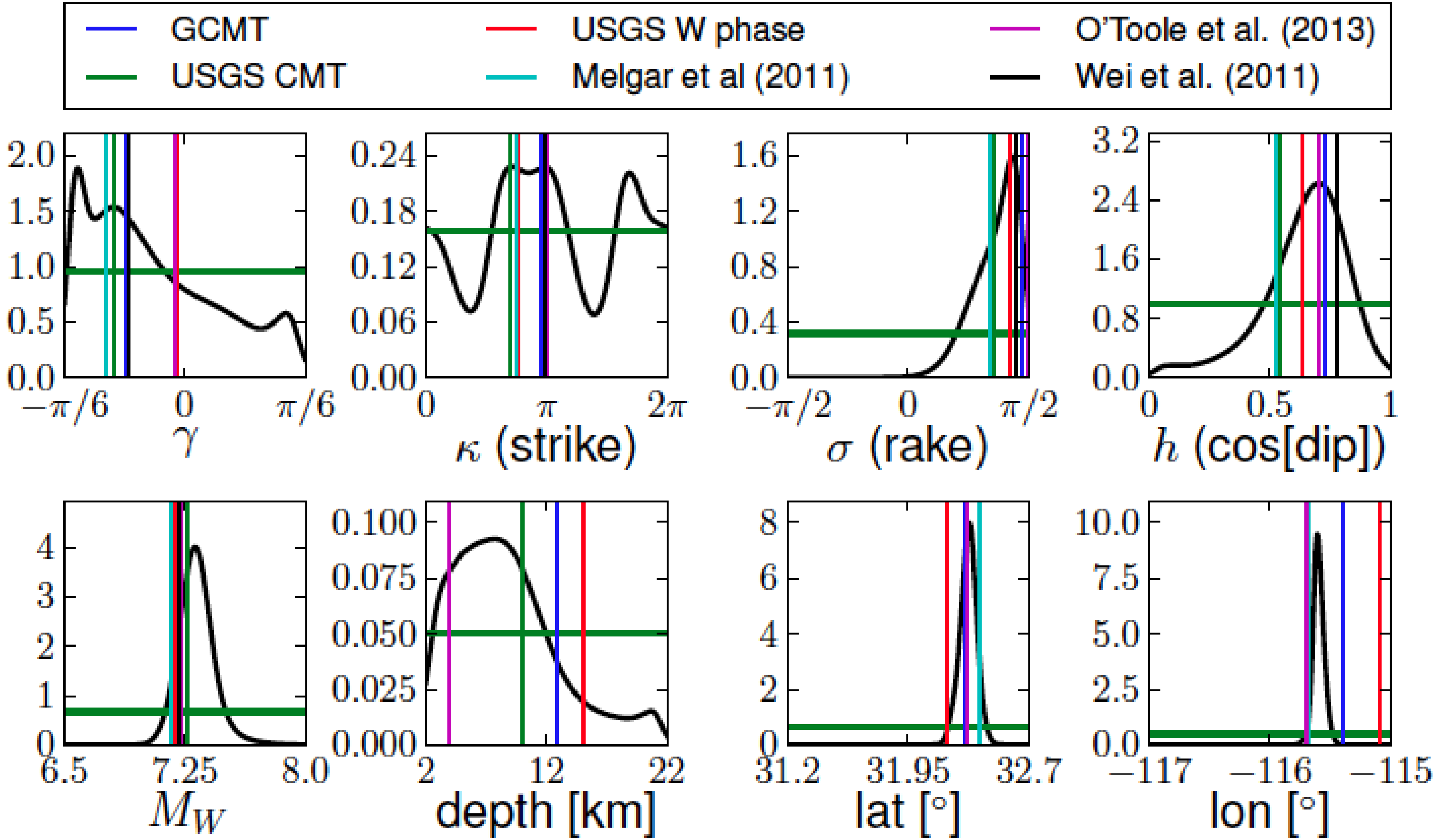
$\mathbf{p}(m|\mathbf{d}^{obs},\mathbf{w}^*)$

Moho depth inversion using phase velocity maps



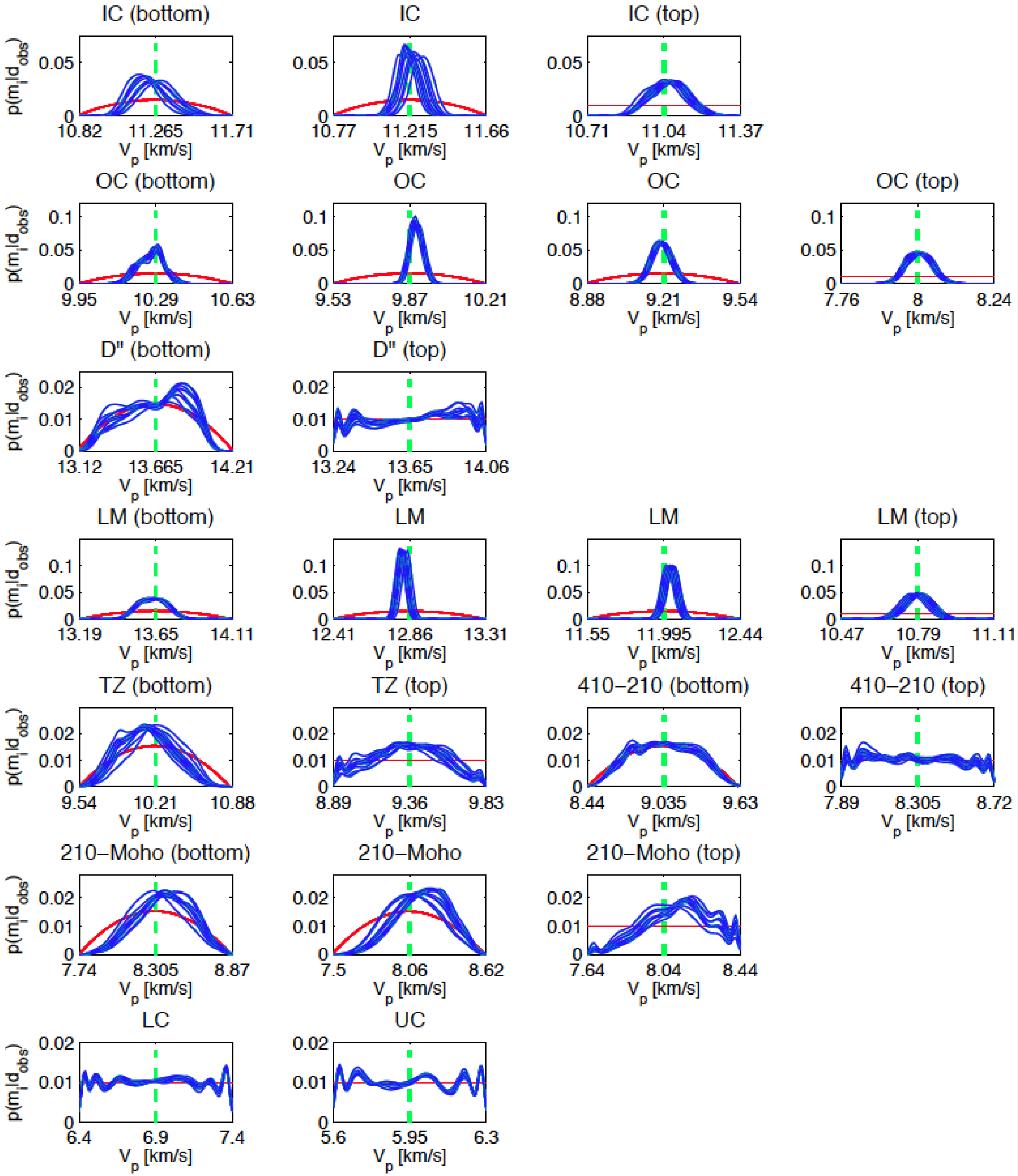
Meier et al. (2007)

CMT inversion using static displacements
2010 El-Mayor Cucapah event



Kaeufl et al. (2013)

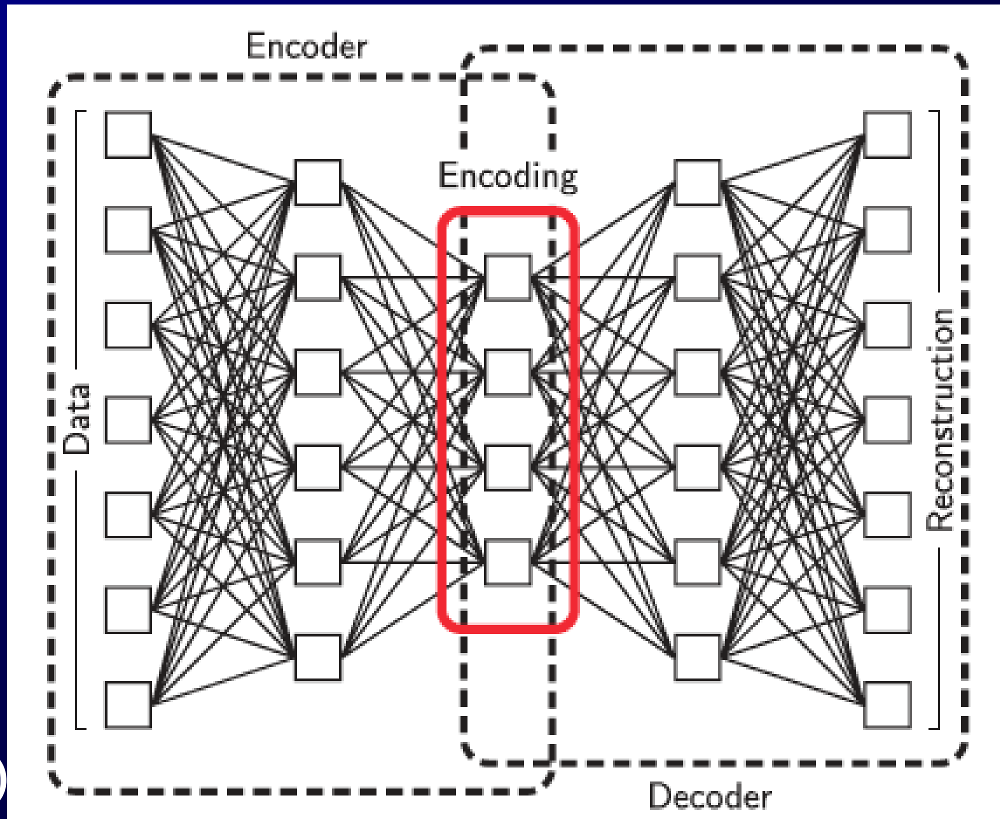
Travel time inversion of EHB data



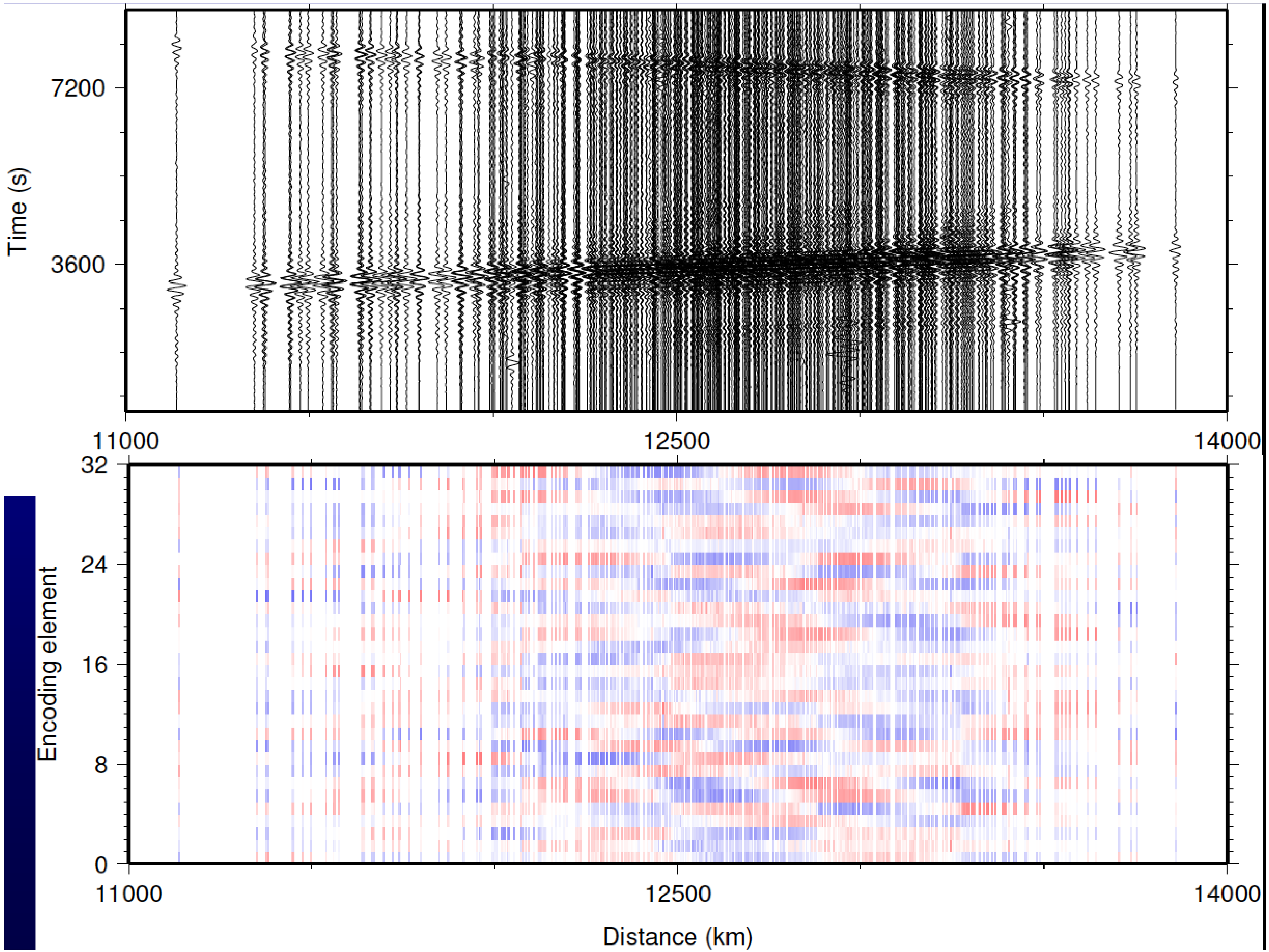
De Wit et al. (2013)

Generalization to waveform inversion

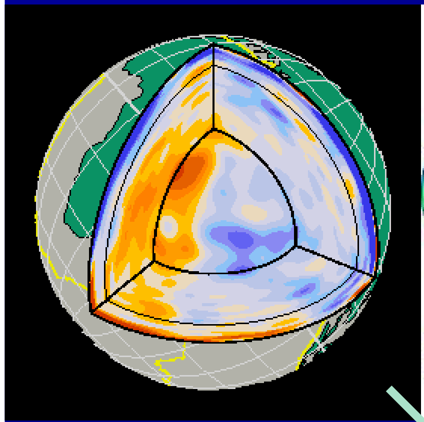
- Trivial extension: 1 input per time step
- Autoencoder

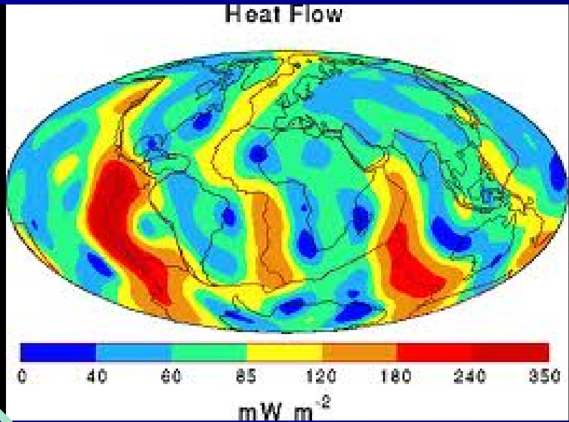


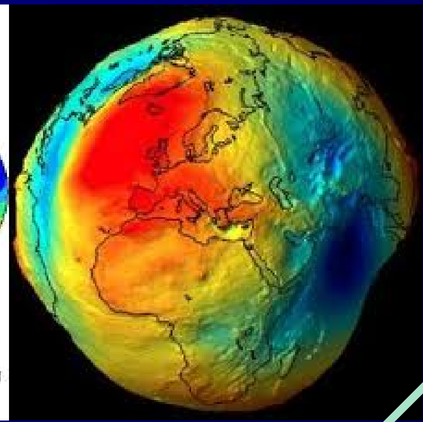
Valentine and Trampert (2012)

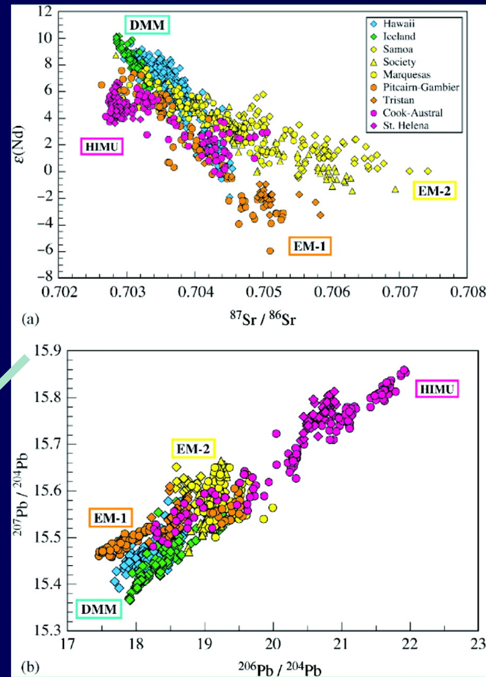

Time (s)
7200
3600
11000
12500
14000
Encoding element
32
24
16
8
0
11000
12500
14000
Distance (km)

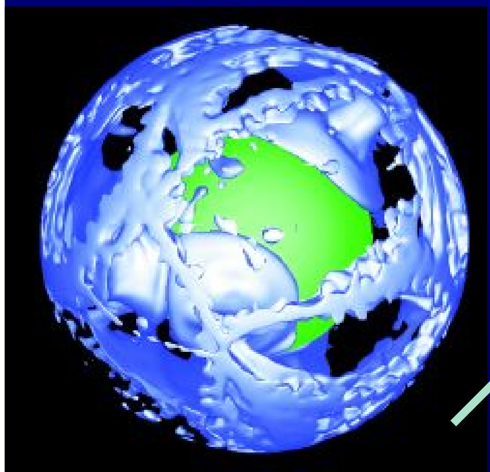
iGEO











Thermo-chemical convection

Physical and chemical model parameters using neural networks