Unravelling interior evolution of terrestrial planets using Machine Learning

Siddhant Agarwal, ESO AIA2019

PhD Supervisors

Dr. Nicola Tosi
Prof. Dr. Doris Breuer
Dr. Pan Kessel
Dr. Grégoire Montavon
Prof. Dr. Klaus-Robert Müller

HEIBRIDS
Helmholtz Einstein International
Berlin Research School in Data Science

Agenda

- **Introduction** to mantle convection and the inverse problem
- **Data** used for inversion
- **Results** using Mixture Density Networks
- **Next steps** using this approach
- Acknowledgements
- References

HEIBRIDS
Helmholtz Einstein International
Berlin Research School in Data Science

DLR

TU berlin

Introduction

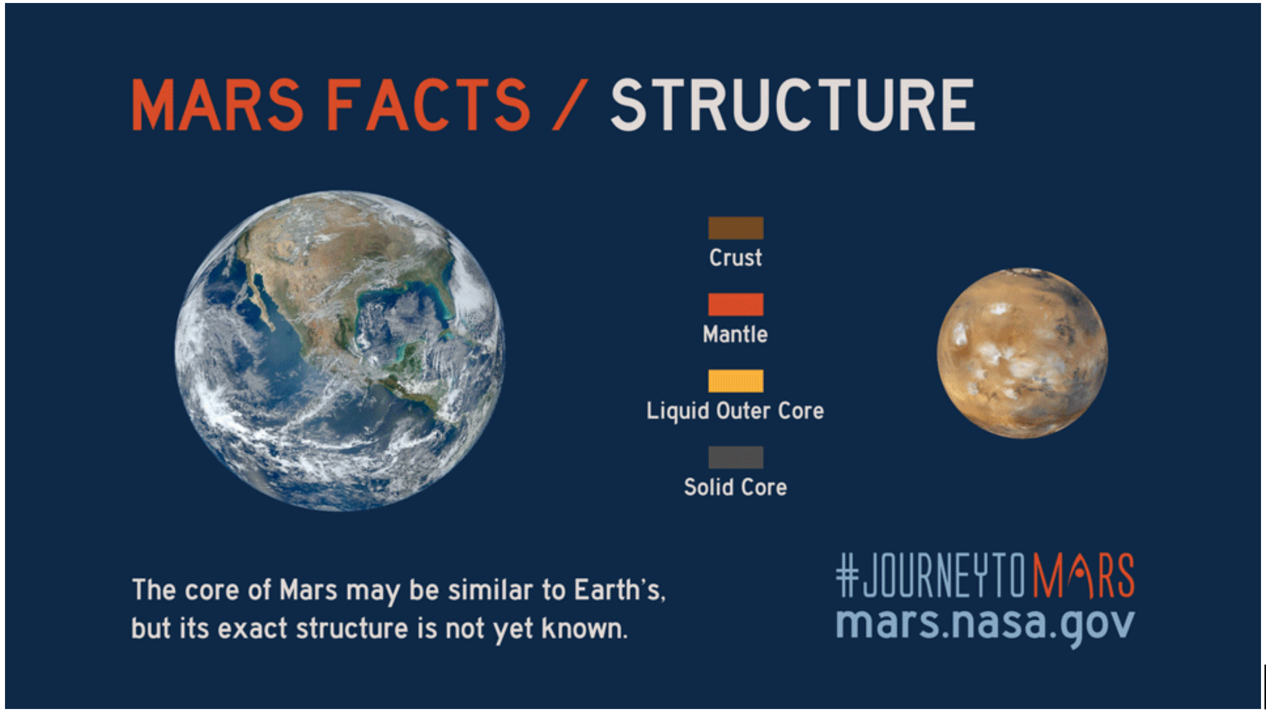
HEIBRIDS
Helmholtz Einstein International
Berlin Research School in Data Science

DLR

TU berlin

Introduction

We are interested in understanding thermal evolution of terrestrial planets like Mars and Earth.

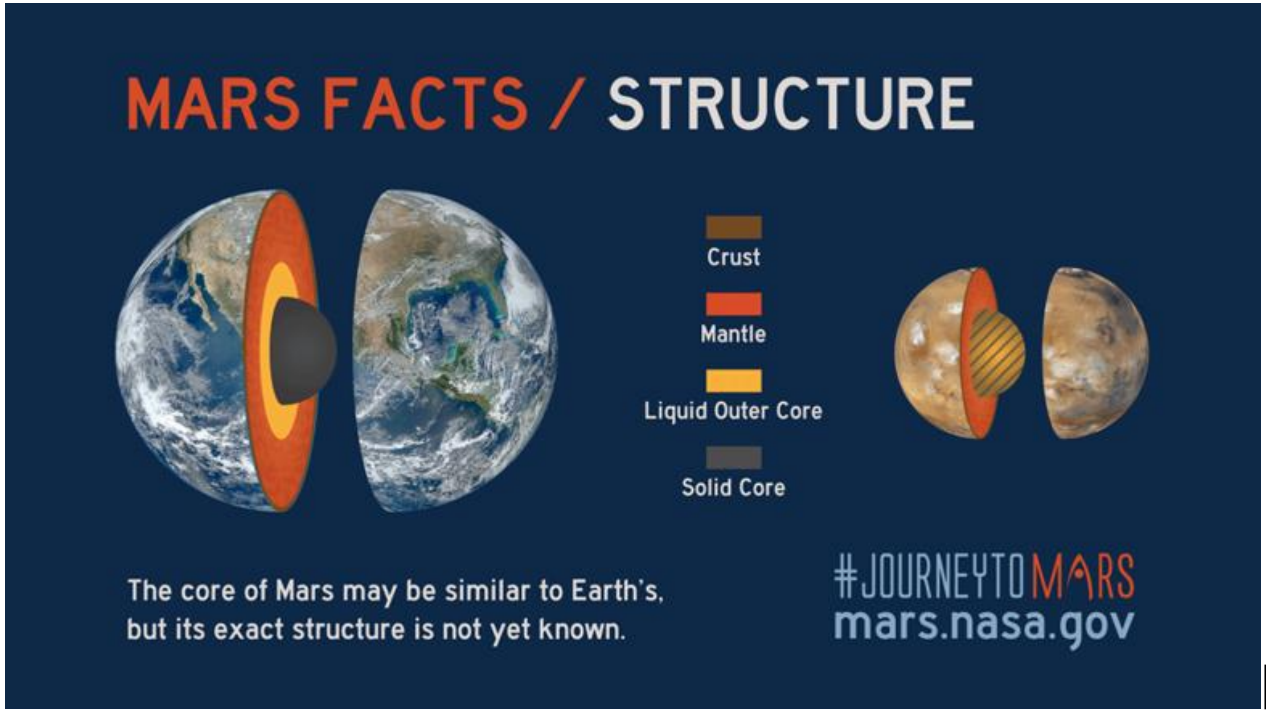


[1]

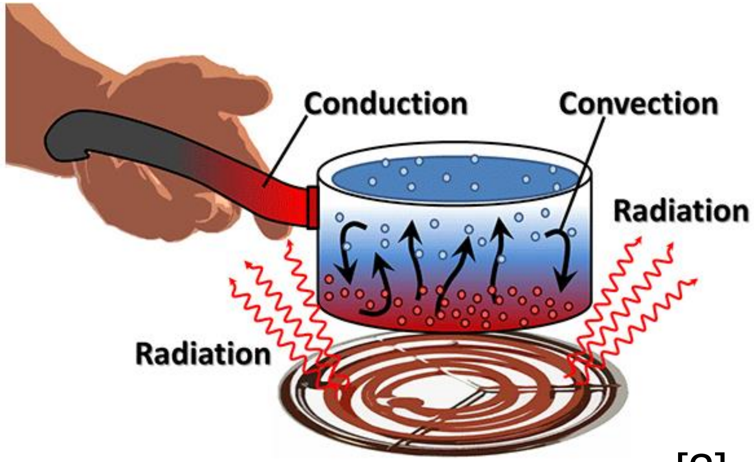
Introduction

We are interested in understanding thermal evolution of terrestrial planets like Mars and Earth.

Mantle convection is an important driver of it.



[1]



[2]

Introduction

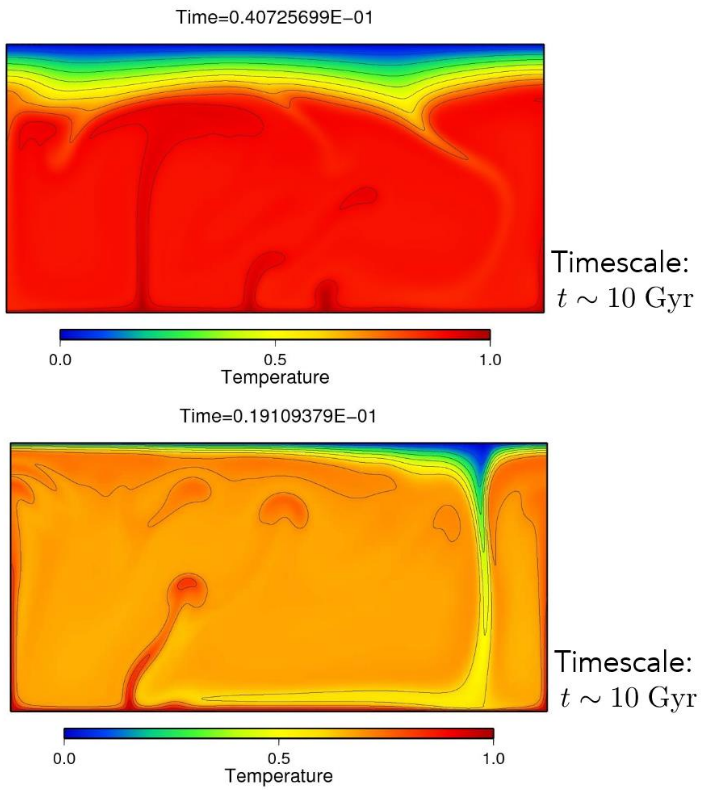
Over geological time scales, rocks behave likes fluids.

[7]

Introduction

Over geological time scales, rocks behave likes fluids.

Hence we use fluid dynamics simulations to study mantle convection.



[7]

Introduction

Inputs

Viscosity

Initial temperature

Radiogenic elements

⋮

In-house C++ code

Conservation of mass:

$$\frac{\partial \rho}{\partial t} + \nabla . (\rho \vec{u}) = 0$$

Conservation of momentum:

$$\frac{D\rho\vec{u}}{Dt} = -\nabla P + \nabla . \tau + \rho\vec{g}$$

Conservation of energy:

$$\rho c_p \frac{DT}{Dt} - \alpha T \frac{DP}{Dt} = \nabla . (k \nabla T) + \rho H + \phi$$

Outputs

Surface heat flux

Radial contraction

Crustal thickness

⋮

- Mantle convection is governed by several poorly constrained parameters and initial conditions

Introduction

- Mantle convection is governed by several poorly constrained parameters and initial conditions
- In planetary science, the outputs are observable (…sometimes)

Introduction

Inputs

Viscosity

Initial temperature

Radiogenic elements

⋮

[3]

Outputs

Surface heat flux

Radial contraction

Crustal thickness

⋮

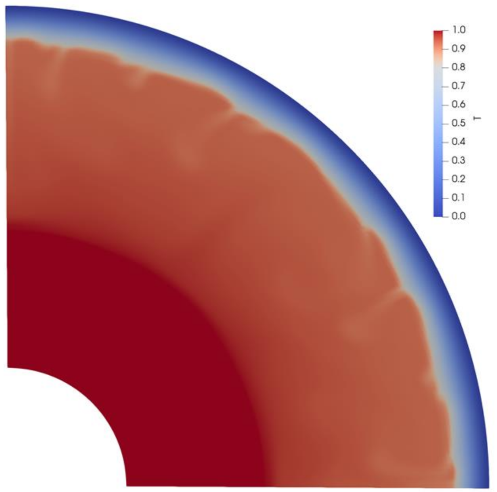
- Mantle convection is governed by several poorly constrained parameters and initial conditions
- In planetary science, the outputs are observable (...sometimes)
- Need Machine Learning for rapid inversion in high-dimensional spaces; Monte Carlo methods are computationally unfeasible

Dataset

Helmholtz Einstein International Berlin Research School in Data Science

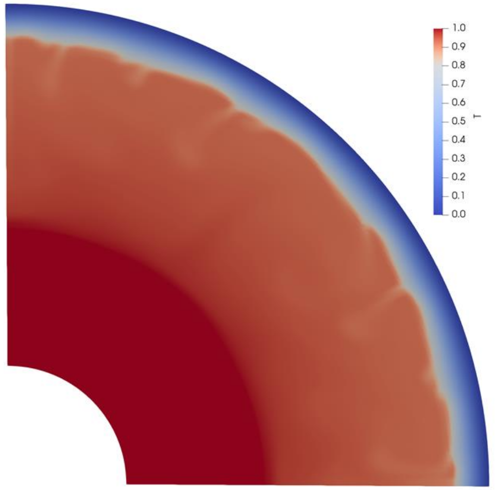
Dataset

- Generated some 3200 2D, quarter-cylinder **evolution** simulations for Mars, with:
  - Compressible convection (Extended-Bousinessq Approximation)
  - Heat production from core and radiogenic elements
  - Temperature and pressure dependent viscosity (Arrhenius)
  - Temperature and pressure dependent thermal conductivity and thermal expansion
  - Solid phase transitions



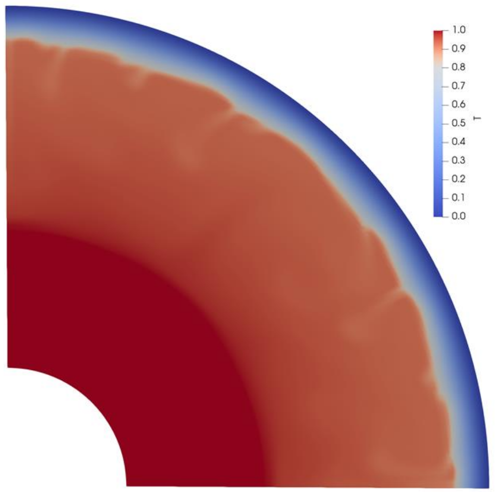
Dataset

- Generated some 3200 2D, quarter-cylinder **evolution** simulations for Mars, with:
  - Compressible convection (Extended-Bousinessq Approximation)
  - Heat production from core and radiogenic elements
  - Temperature and pressure dependent viscosity (Arrhenius)
  - Temperature and pressure dependent thermal conductivity and thermal expansion
  - Solid phase transitions



- ~800 GB of data generated using 60,000 CPU hours; ~800 MB used for ML

Dataset



- Generated some 3200 2D, quarter-cylinder **evolution** simulations for Mars, with:
  - Compressible convection (Extended-Bousinessq Approximation)
  - Heat production from core and radiogenic elements
  - Temperature and pressure dependent viscosity (Arrhenius)
  - Temperature and pressure dependent thermal conductivity and thermal expansion
  - Solid phase transitions

- ~800 GB of data generated using 60,000 CPU hours; ~800 MB used for ML

Parameters

Ra (η_{ref} ϵ $[10^{19}, 10^{22}$ $Pa\ s])$

RaQ (Λ ϵ $[1, 50]$, η_{ref})

$\mathbf{T_{initial}}$ ϵ [1600, 1800 K]

E ϵ [10^5, 5x10^5 J mol^{-1}]

V ϵ [4x10^{-6}, 10x10^{-6} m^3 mol^{-1}]

←

“Observables”

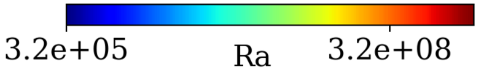
Temperature profile

Radial contraction

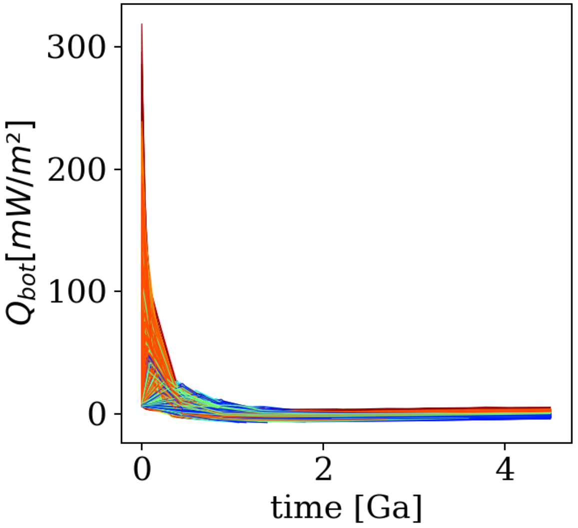
Elastic lithospheric thickness

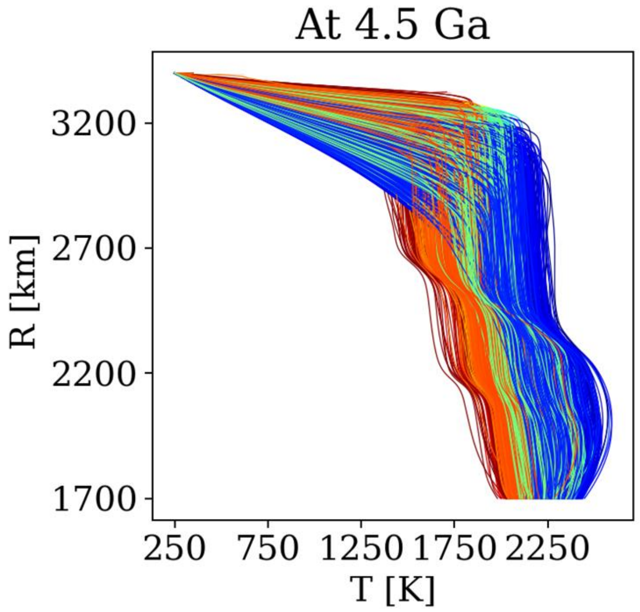
Core heat flux

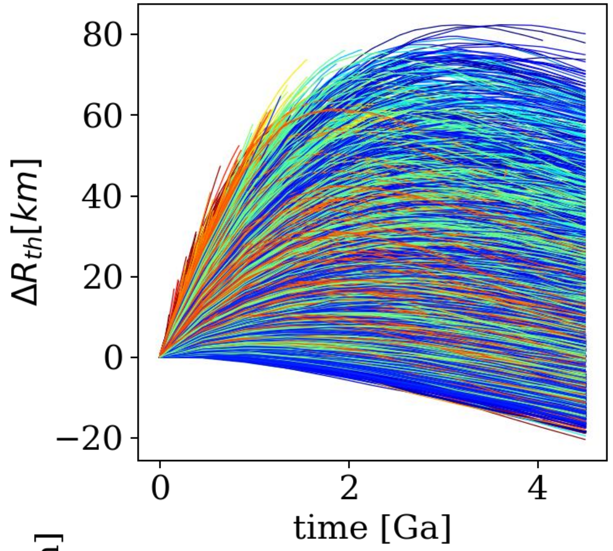
Surface heat flux

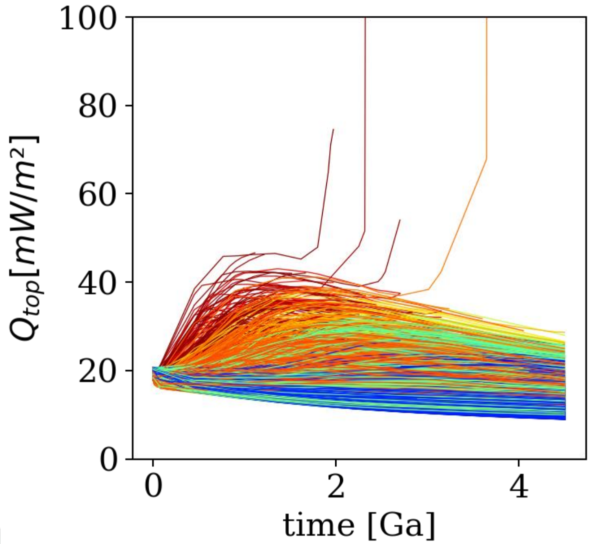


Dataset









"Observables"

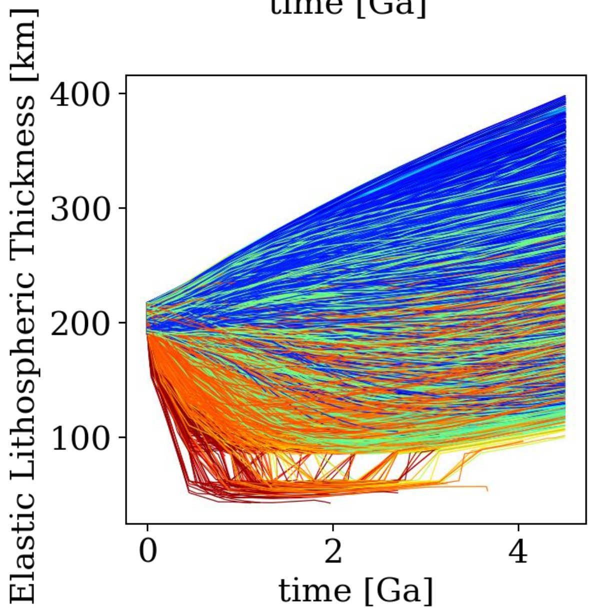
Temperature profile

Radial contraction

Elastic lithospheric thickness

Core heat flux

Surface heat flux



Results

HEIBRIDS
Helmholtz Einstein International
Berlin Research School in Data Science

DLR

TU berlin

Results

- Mixture Density Network (MDN) is promising for inverse problems.

Contents lists available at ScienceDirect
Physics of the Earth and Planetary Interiors
journal homepage: www.elsevier.com/locate/pepi

Using pattern recognition to infer parameters governing mantle convection

Suzanne Atkins [a,*], Andrew P. Valentine [a], Paul J. Tackley [b], Jeannot Trampert [a]
[a] Department of Earth Sciences, Utrecht University, P.O. Box 80.115, 3508 TC Utrecht, The Netherlands
[b] Institute of Geophysics, ETH Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland

[4]

Geophysical Journal International
Geophys. J. Int. (2013) **195,** 408–422 doi: 10.1093/gji/ggt220
Advance Access publication 2013 June 26

Bayesian inference of Earth's radial seismic structure from body-wave traveltimes using neural networks

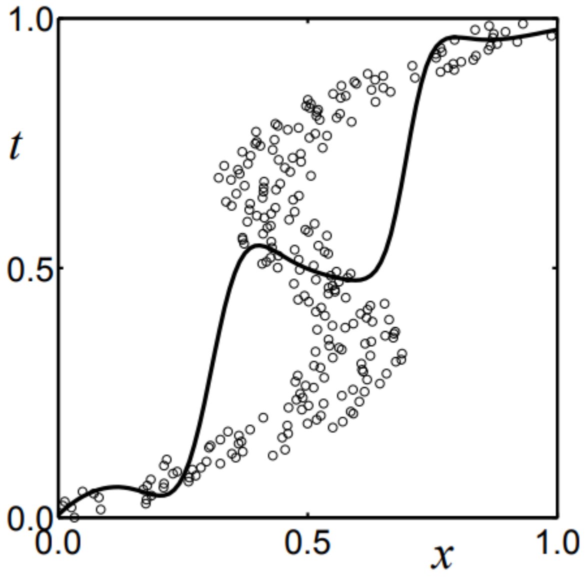
Ralph W. L. de Wit, Andrew P. Valentine and Jeannot Trampert
Department of Earth Sciences, Utrecht University, Budapestlaan 4, 3584 CD, Utrecht, the Netherlands. E-mail: r.w.l.dewit@uu.nl

[5]

Accepted 2013 May 29. Received 2013 May 15; in original form 2013 February 14

Results

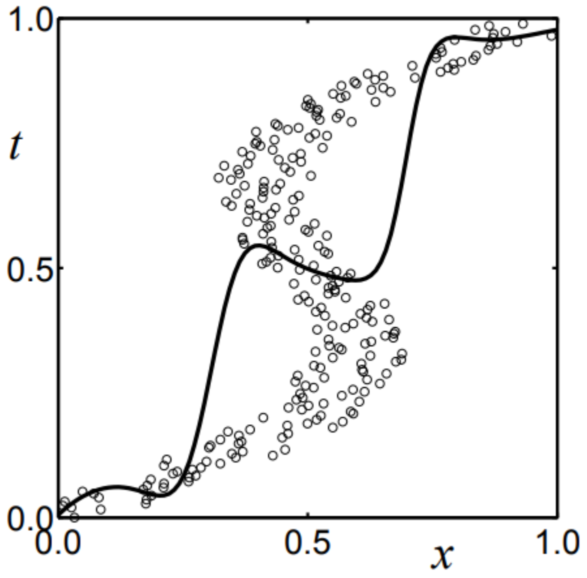
- Mixture Density Network (MDN) is promising for inverse problems.
- Based on the algorithm by Bishop [6]

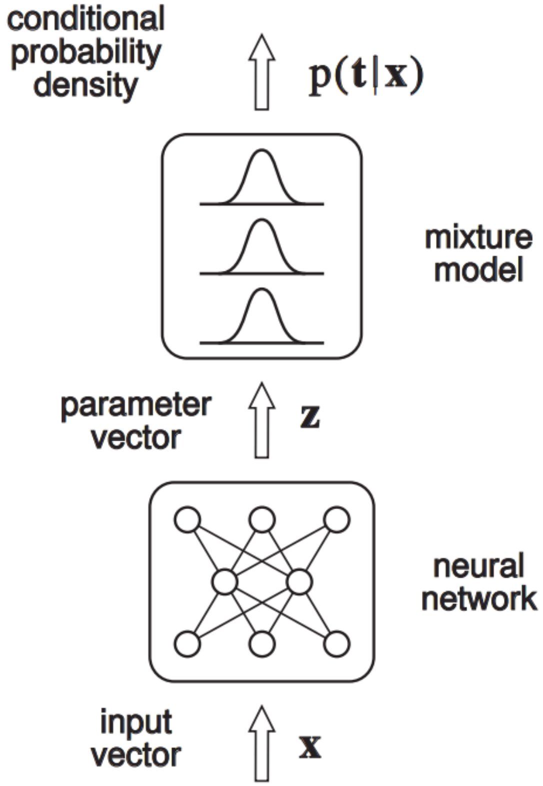


$$E^S = \frac{1}{2}\sum_{q=1}^{n}\sum_{k=1}^{c}\left[f_k(x^q; w) - t_k^q\right]^2$$

Results

- Mixture Density Network (MDN) is promising for inverse problems.
- Based on the algorithm by Bishop [6]



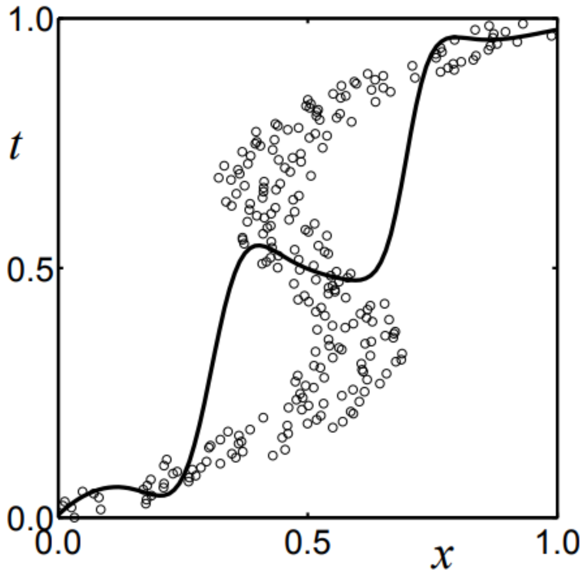


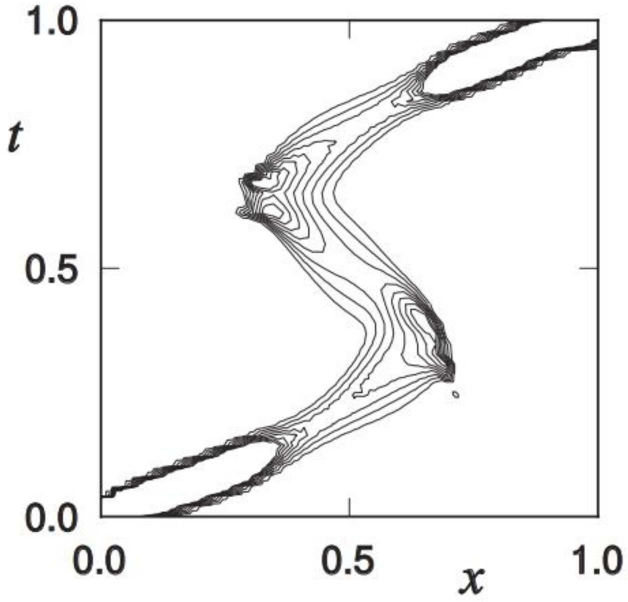
$$E^S = \frac{1}{2}\sum_{q=1}^{n}\sum_{k=1}^{c}\left[f_k(x^q;w) - t_k^q\right]^2$$

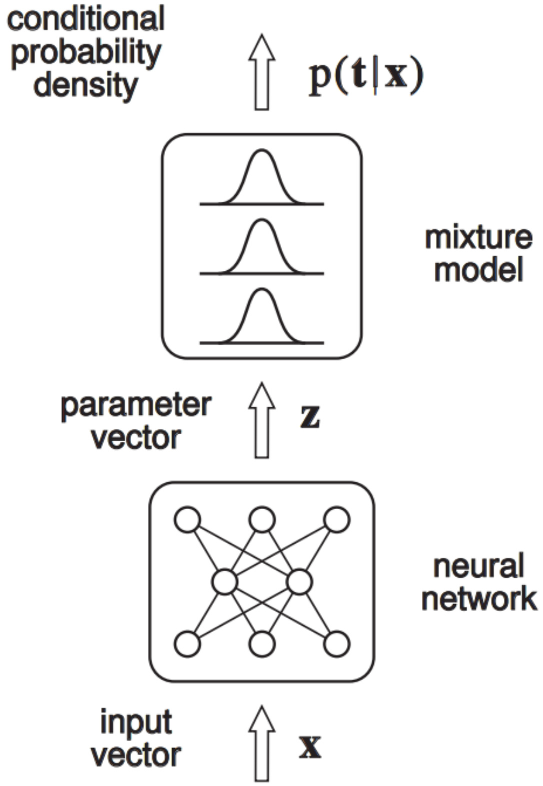
$$E^S = -\ln\left\{\sum_{i=1}^{m}\alpha_i(x^q)\emptyset_i(t^q|x^q)\right\}$$

Results

- Mixture Density Network (MDN) is promising for inverse problems.
- Based on the algorithm by Bishop [6]



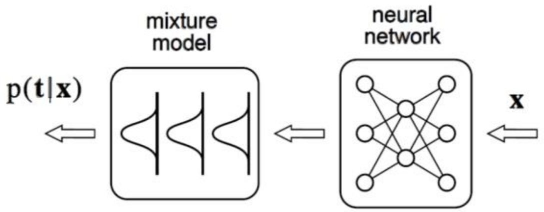


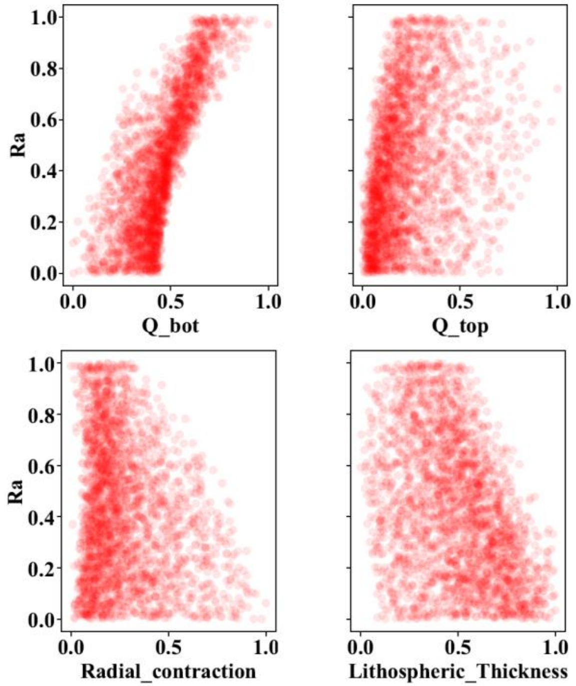


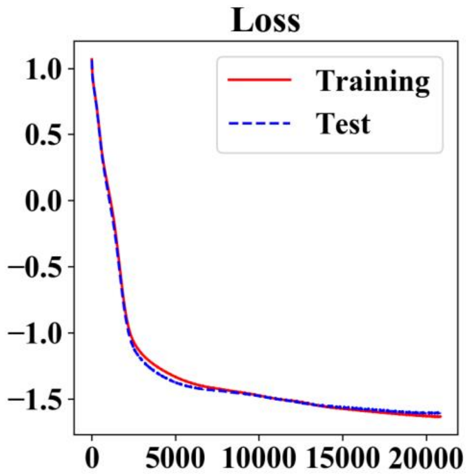
$$E^S = \frac{1}{2}\sum_{q=1}^{n}\sum_{k=1}^{c}\left[f_k(x^q; w) - t_k^q\right]^2$$

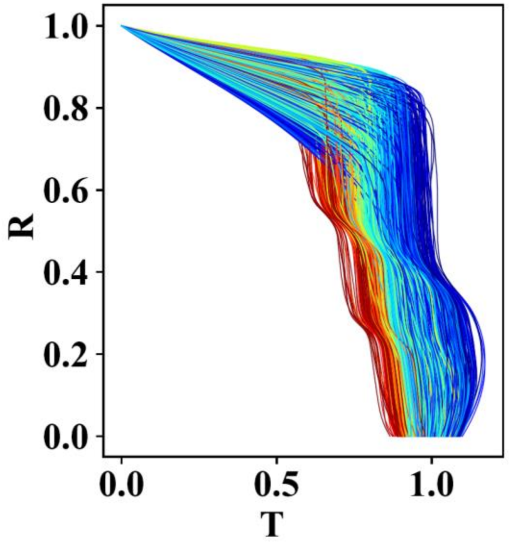
$$E^S = -\ln\left\{\sum_{i=1}^{m}\alpha_i(x^q)\emptyset_i(t^q|x^q)\right\}$$

Results

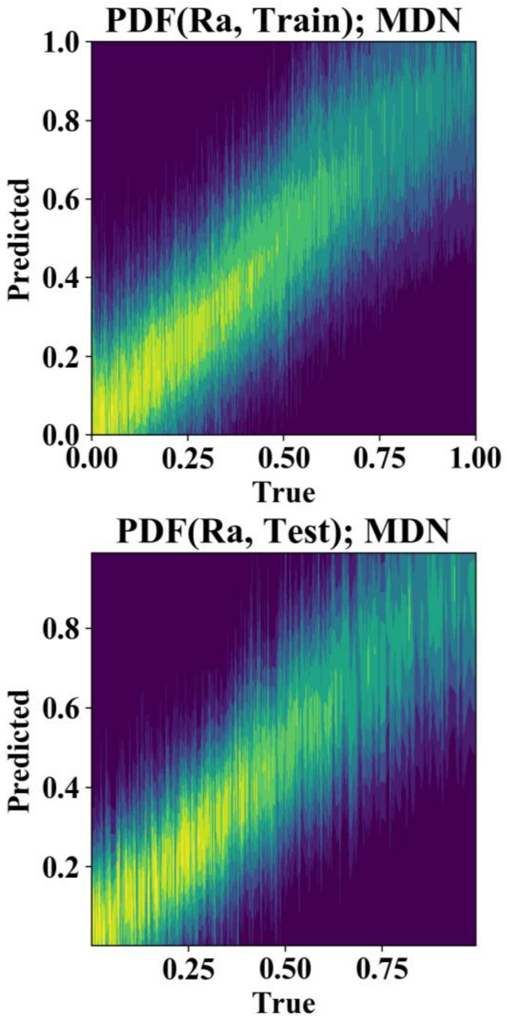

mixture model
neural network
p(t|x)
x

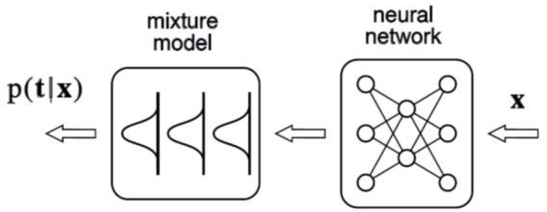

Ra
Q_bot
Q_top
Radial_contraction
Lithospheric_Thickness
0.0
0.5
1.0

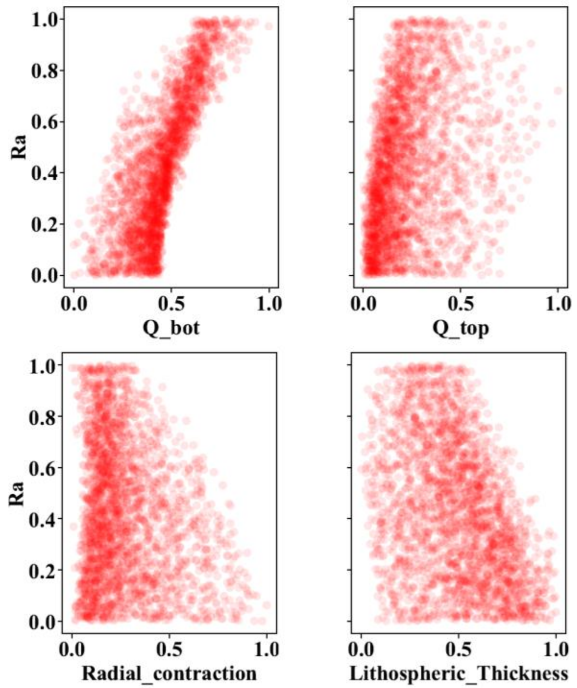

Loss
Training
Test
1.0
0.5
0.0
−0.5
−1.0
−1.5
0
5000
10000
15000
20000

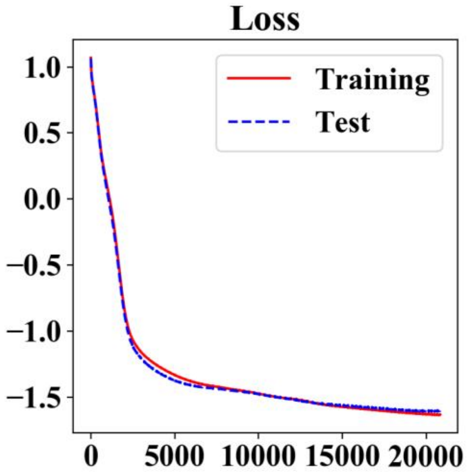

R
T
0.0
0.2
0.4
0.6
0.8
1.0
0.5

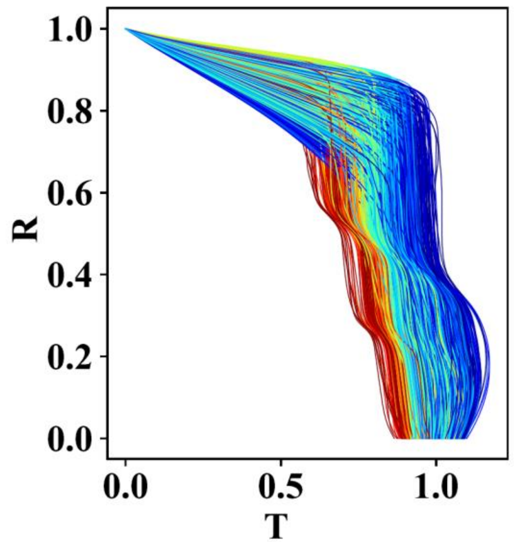
Results


PDF(Ra, Train); MDN
Predicted
True
PDF(Ra, Test); MDN
Predicted
True

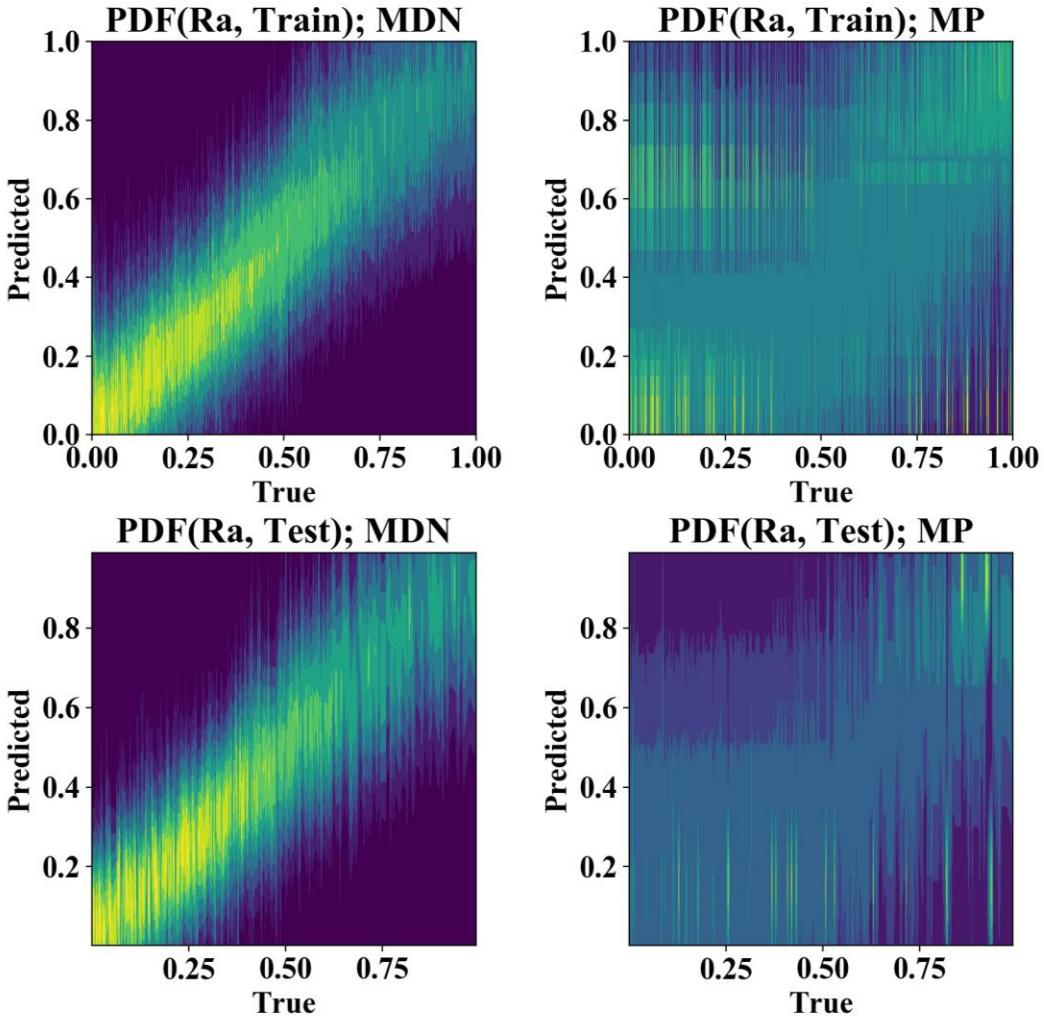

p(t|x)
mixture model
neural network
x

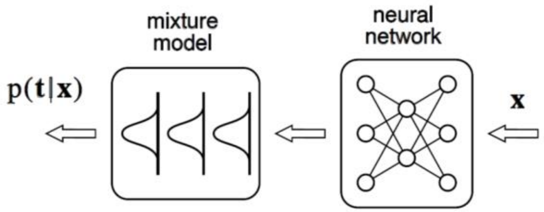

Ra
Q_bot
Q_top
Radial_contraction
Lithospheric_Thickness

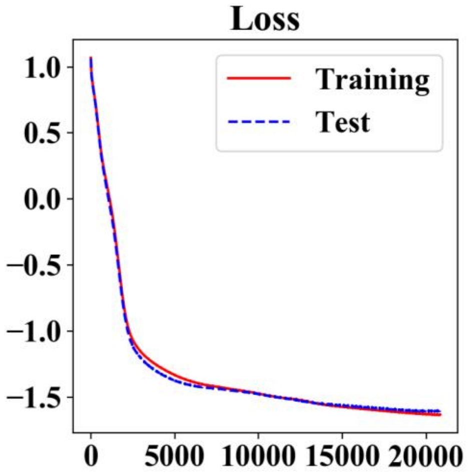

Loss
Training
Test

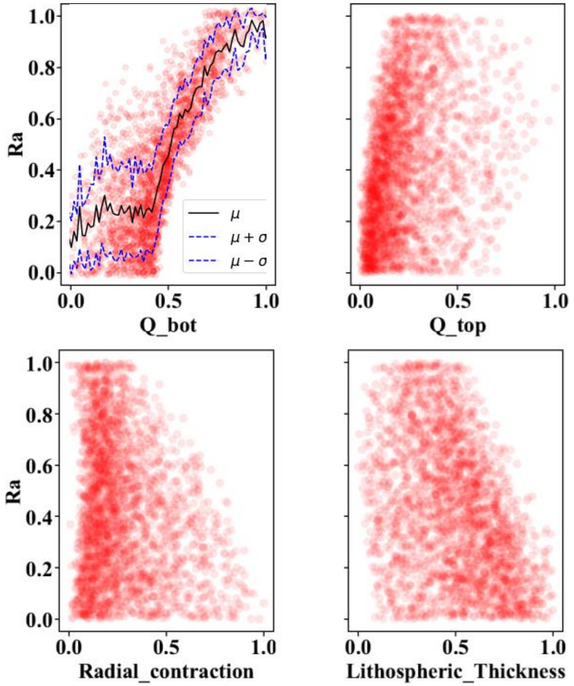

R
T

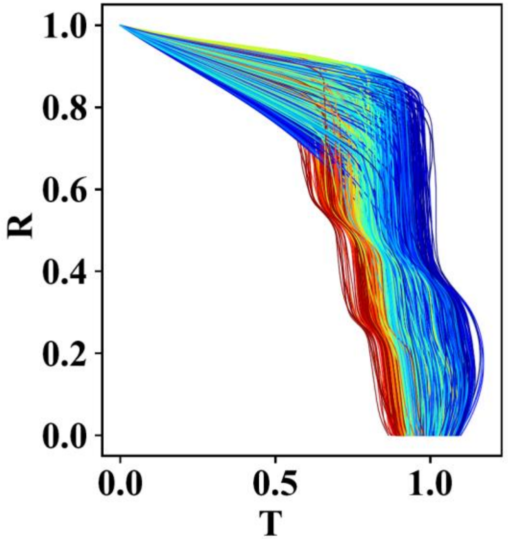
Results


PDF(Ra, Train); MDN
PDF(Ra, Train); MP
PDF(Ra, Test); MDN
PDF(Ra, Test); MP
Predicted
True
0.00 0.25 0.50 0.75 1.00
0.0 0.2 0.4 0.6 0.8 1.0

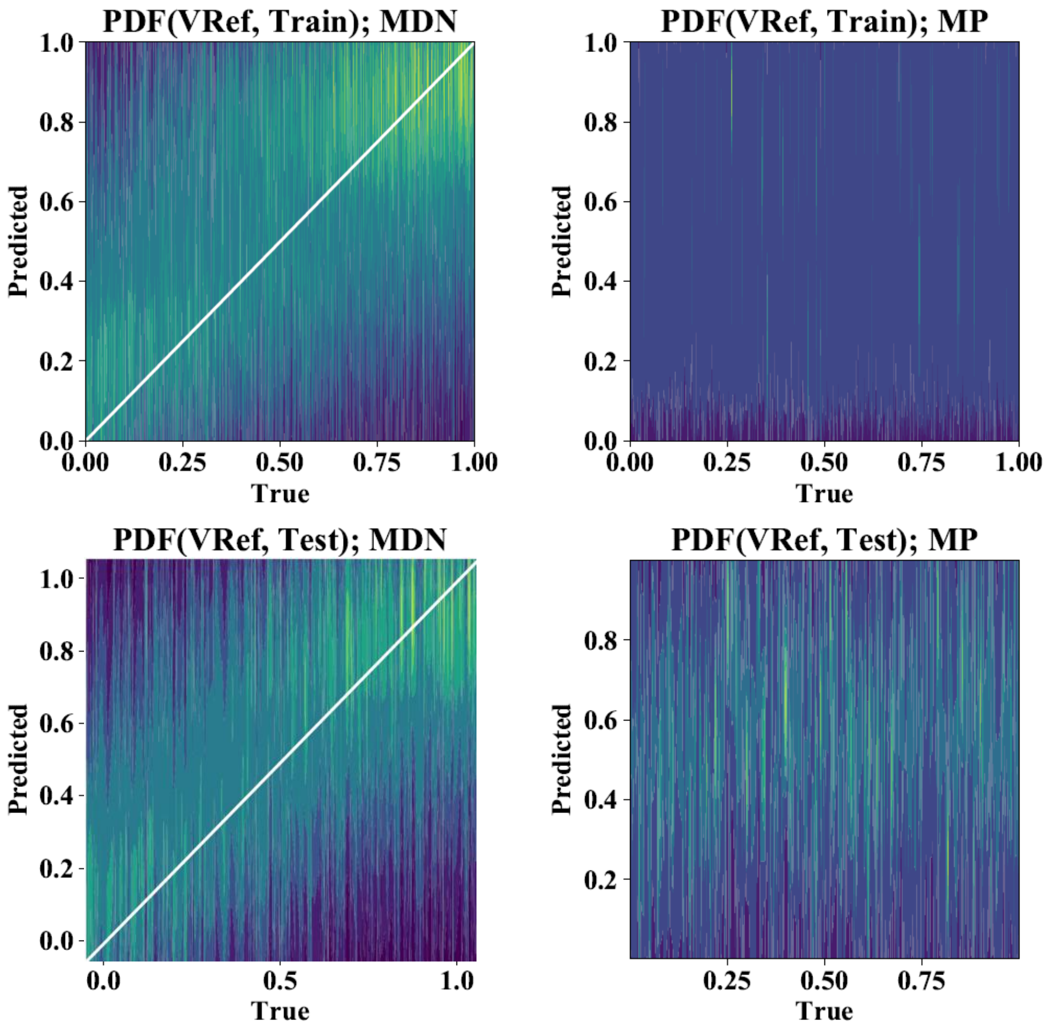

p(t|x)
mixture model
neural network
x

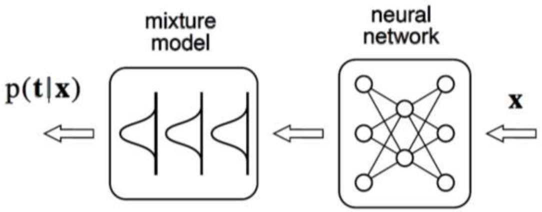

Loss
Training
Test
1.0 0.5 0.0 −0.5 −1.0 −1.5
0 5000 10000 15000 20000

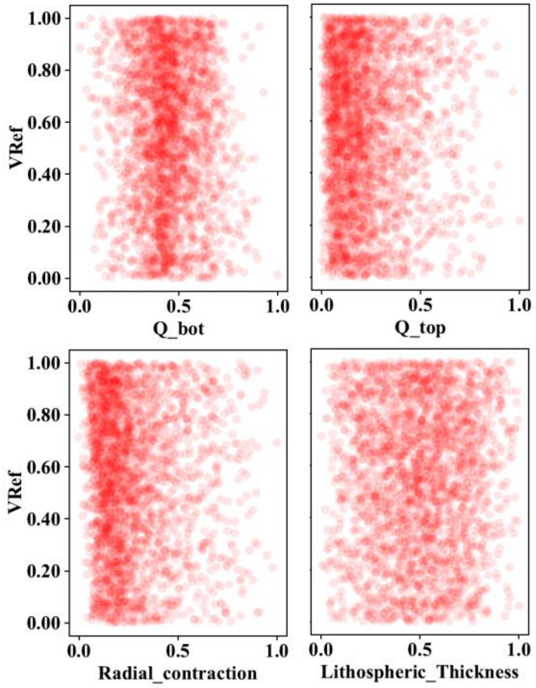

Ra
Q_bot
Q_top
Radial_contraction
Lithospheric_Thickness
μ
$\mu+\sigma$
$\mu-\sigma$
0.0 0.5 1.0

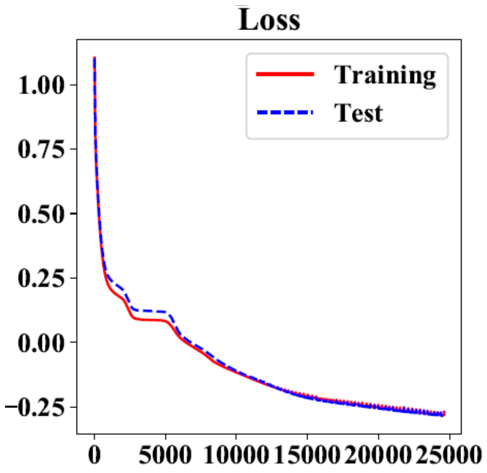

R
T
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.5 1.0

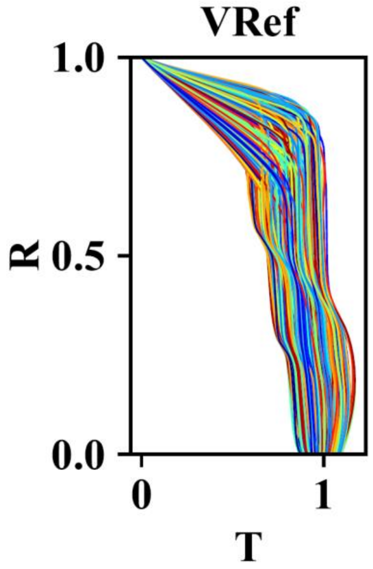
Results


PDF(VRef, Train); MDN
PDF(VRef, Train); MP
PDF(VRef, Test); MDN
PDF(VRef, Test); MP
Predicted
True
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.25
0.50
0.75
1.00
0.5


p(t|x)
mixture model
neural network
x

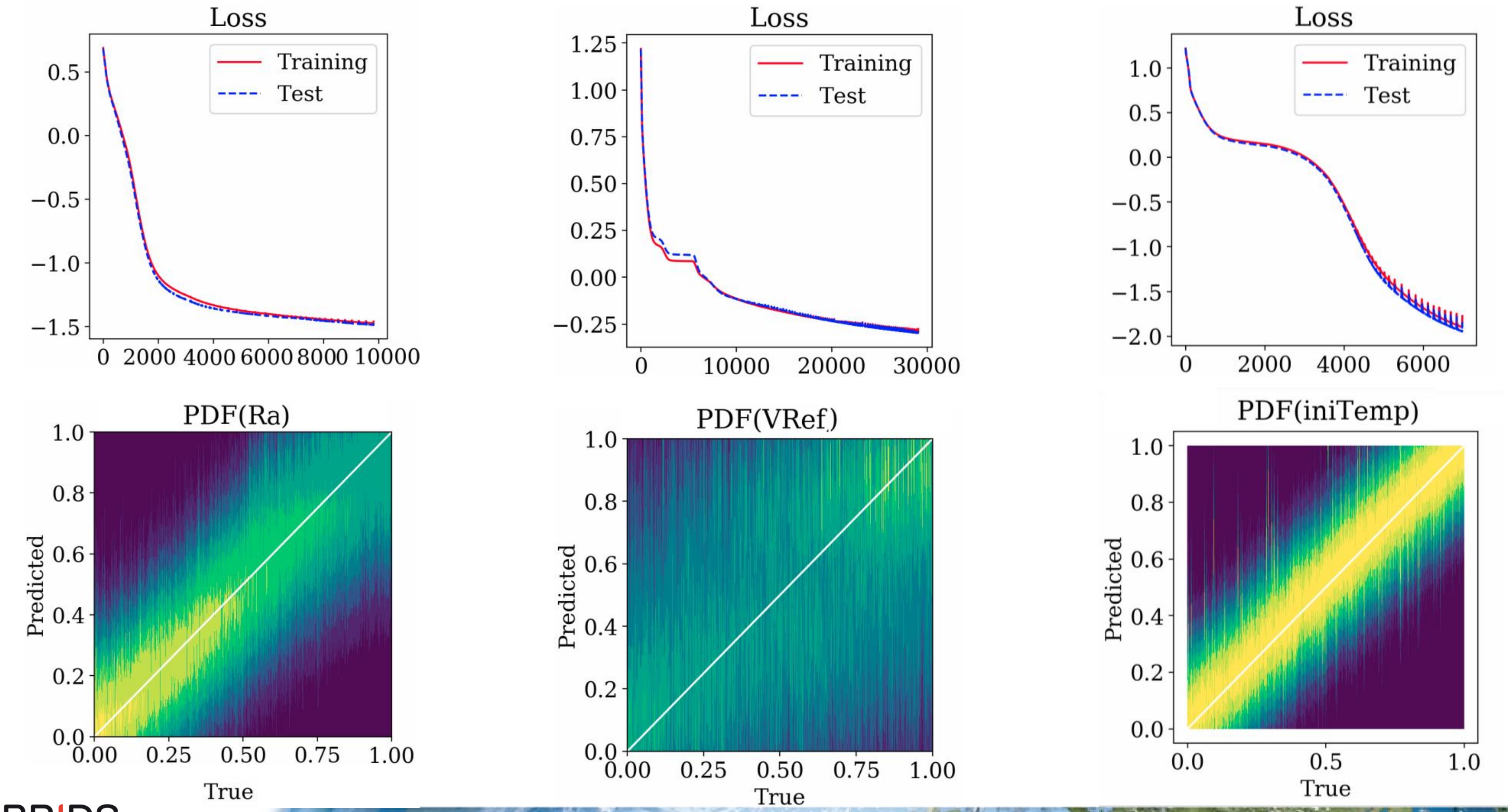

VRef
Q_bot
Q_top
Radial_contraction
Lithospheric_Thickness
0.00
0.20
0.40
0.60
0.80
1.00
0.0
0.5
1.0


Loss
Training
Test
1.00
0.75
0.50
0.25
0.00
−0.25
0
5000
10000
15000
20000
25000


VRef
R
T
1.0
0.5
0.0
0
1

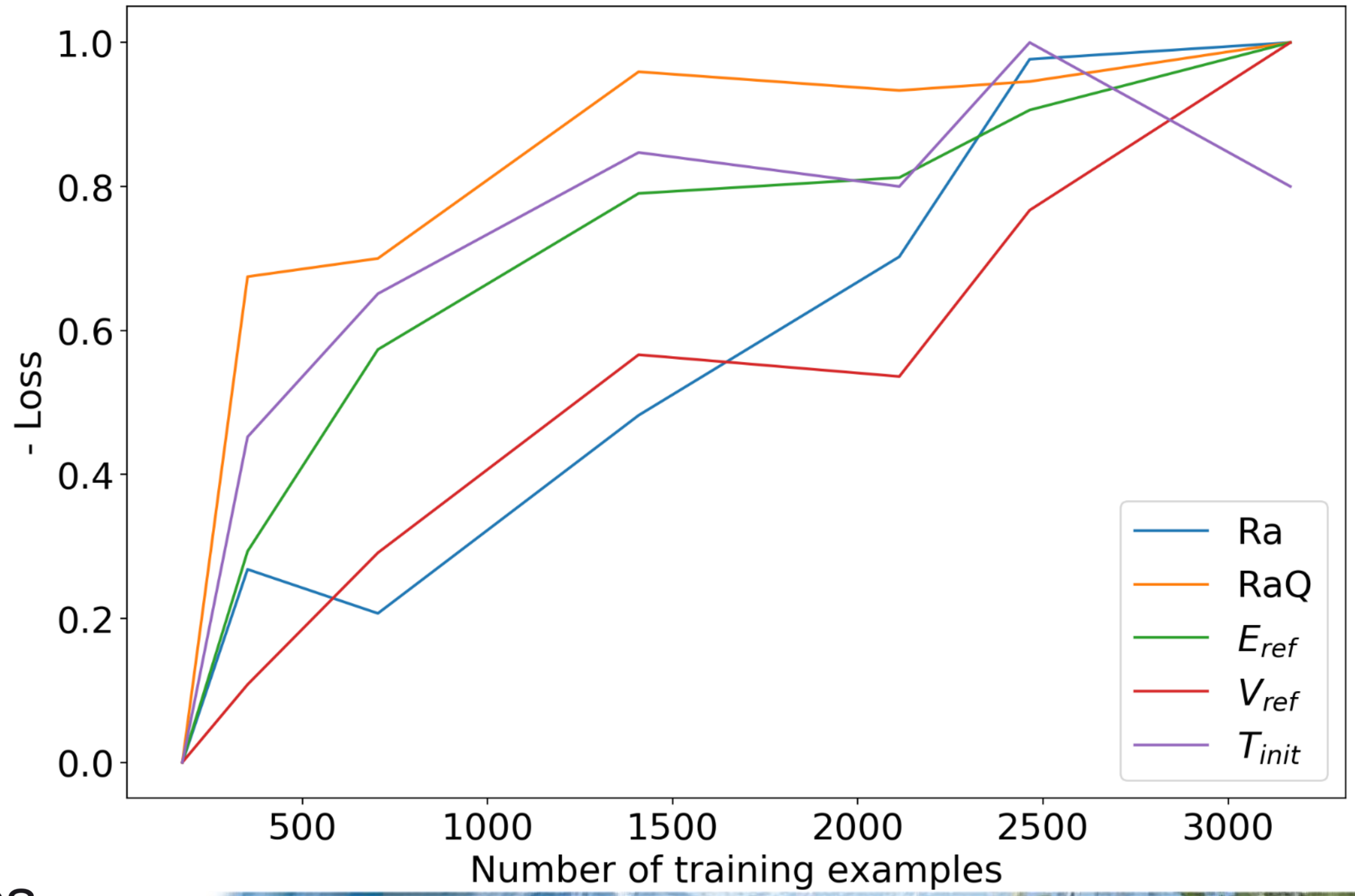
Results

Using loss as a measure of ‘constrainability’

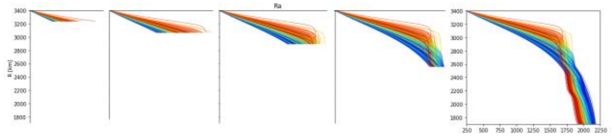


Results

Asymptotic behavior of loss shows number of simulations is sufficient for this set of parameters and observables



Results



Parameters / Observables (colour scale: -(Mean Loss), 0.0–1.5)

Parameters	Q_{bot}	Q_{top}	$R_{contract}$	L_{thick}	$T_{prof;\,90,\,10}$	$T_{prof;\,80,\,20}$	$T_{prof;\,70,\,30}$	$T_{prof;\,60,\,40}$	$T_{prof;\,50,\,50}$	$T_{prof;\,0,\,100}$
Ra	0.66	0.14	0.06	0.16	0.49	0.99	1.21	1.31	1.35	1.43
RaQ	0.61	0.33	0.14	0.34	0.57	1.01	1.30	1.36	1.39	1.44
ERef	0.01	0.09	0.01	0.09	0.14	0.22	0.36	0.36	0.38	0.63
VRef	-0.07	-0.07	-0.11	-0.09	-0.11	-0.07	-0.03	-0.02	-0.02	0.08
T_{init}	-0.06	-0.05	-0.02	-0.02	-0.03	-0.02	-0.00	0.01	0.06	0.07

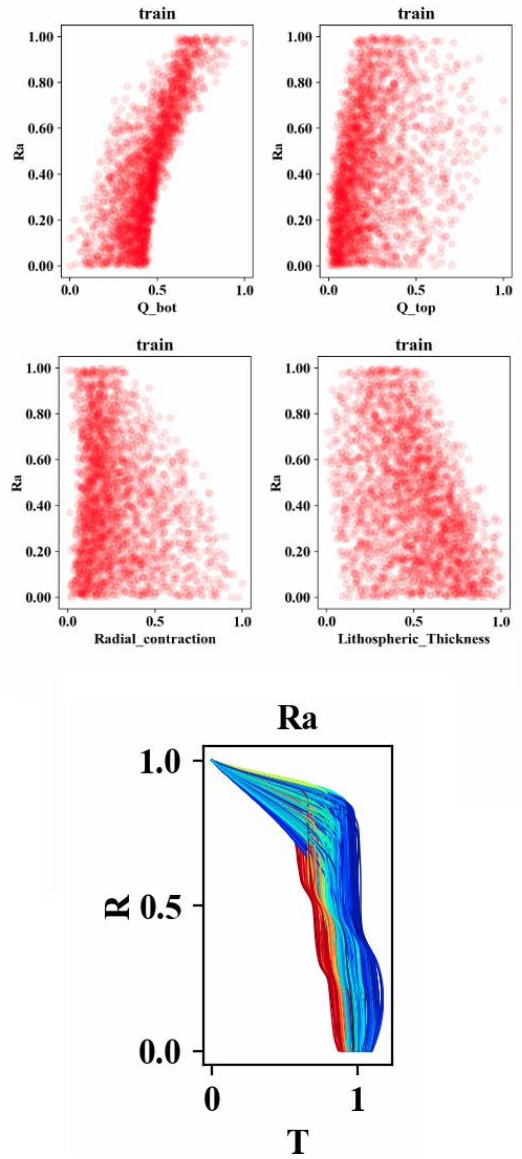
Parameters	Q_{bot}, Q_{top}	$Q_{bot}, R_{contract}$	Q_{bot}, L_{thick}	$Q_{top}, R_{contract}$	Q_{top}, L_{thick}	$R_{contract}, L_{thick}$	$Q_{bot}, Q_{top}, R_{contract}$	$Q_{bot}, Q_{top}, L_{thick}$	$Q_{top}, R_{contract}, L_{thick}$	$Q_{bot}, Q_{top}, R_{contract}, L_{thick}$	$Q_{bot}, Q_{top}, R_{contract}, L_{thick}, T_{prof;\,0,\,100}$
Ra	1.11	1.02	1.12	0.59	0.26	0.60	1.15	1.11	0.65	1.21	1.59
RaQ	1.19	1.05	1.19	0.71	0.44	0.73	1.25	1.21	0.79	1.33	1.60
ERef	0.18	0.11	0.18	0.15	0.09	0.16	0.24	0.35	0.15	0.39	0.69
VRef	-0.06	-0.08	-0.07	-0.07	-0.10	-0.07	-0.07	-0.09	-0.09	-0.07	0.29
T_{init}	-0.04	-0.03	-0.01	0.34	-0.02	0.27	0.94	-0.02	0.42	0.95	1.92

Observables

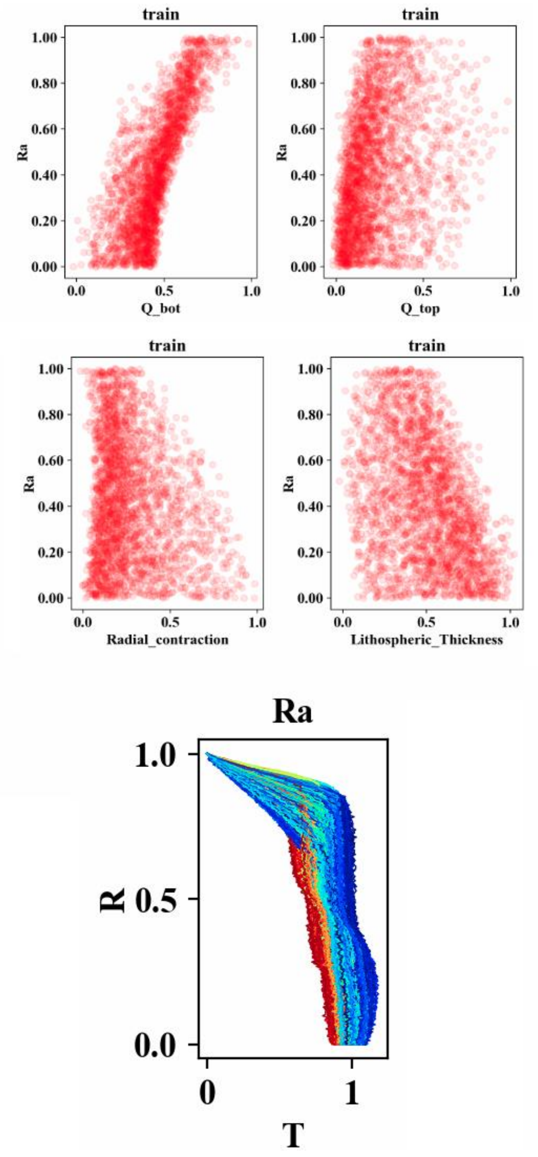
-(Mean Loss)

Results

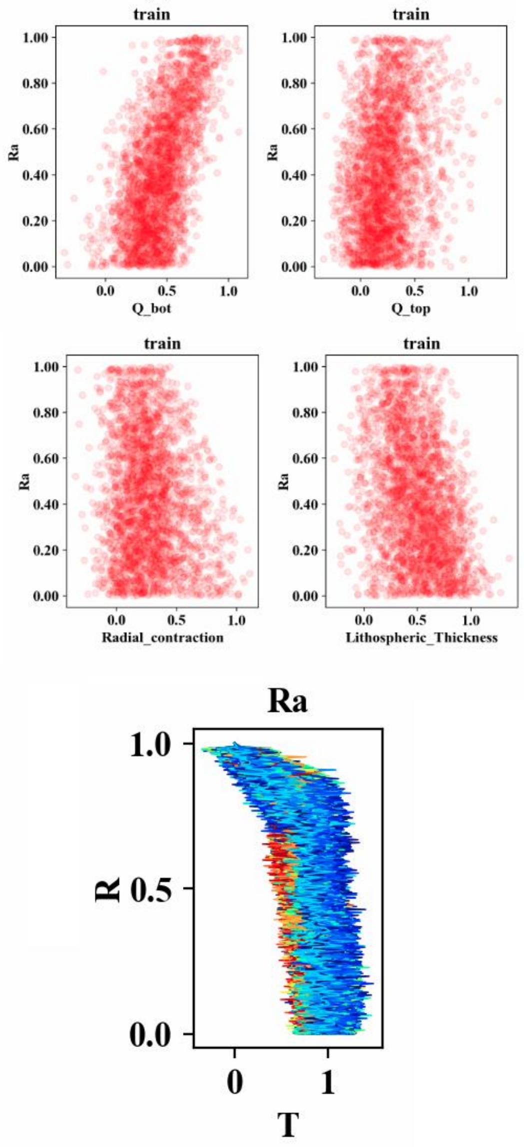
σ = 0% noise



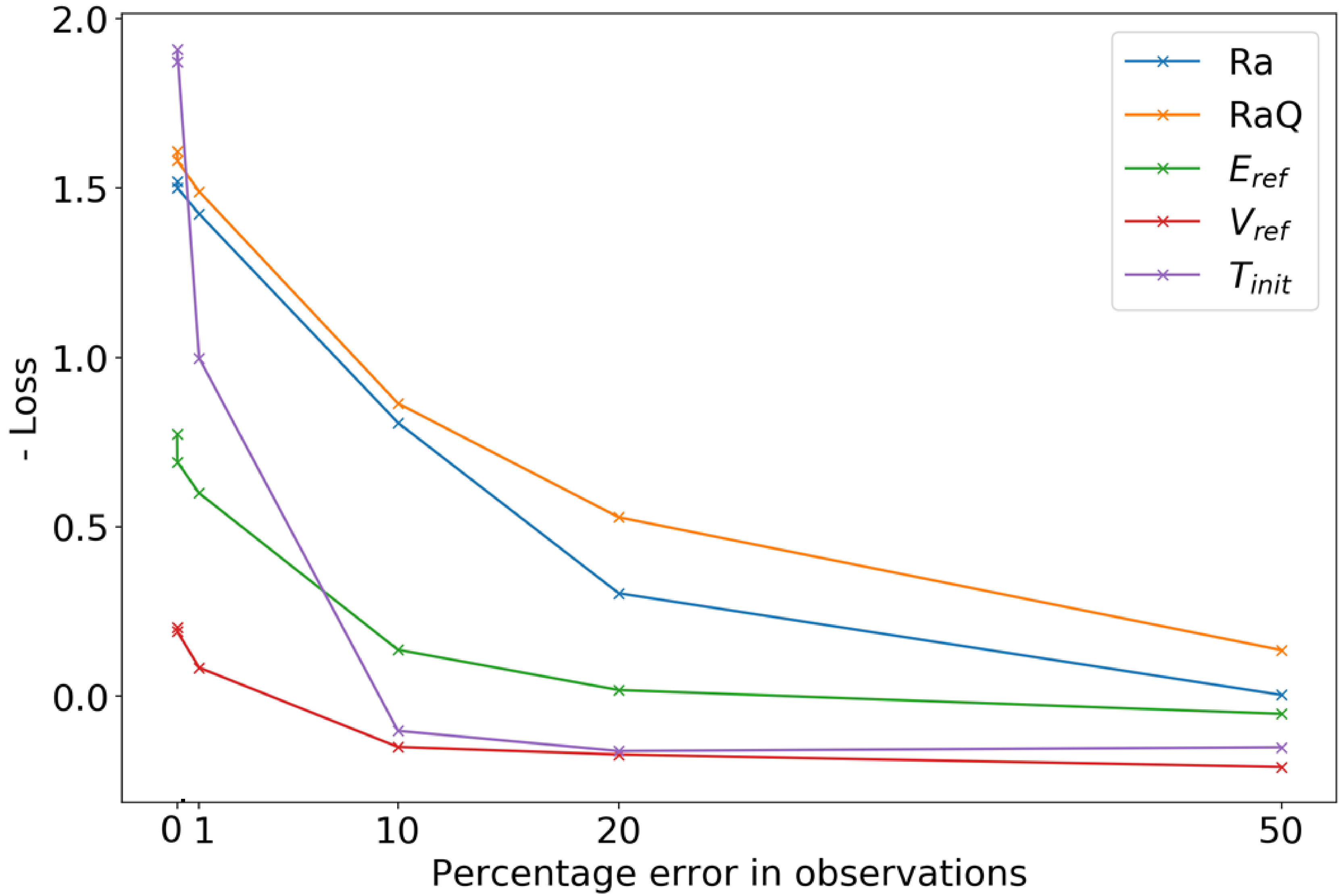
σ = 1% noise



σ = 10% noise



Results



2.0
1.5
1.0
0.5
0.0
- Loss
0 1
10
20
50
Percentage error in observations
Ra
RaQ
E_{ref}
V_{ref}
T_{init}

Future steps

Helmholtz Einstein International Berlin Research School in Data Science

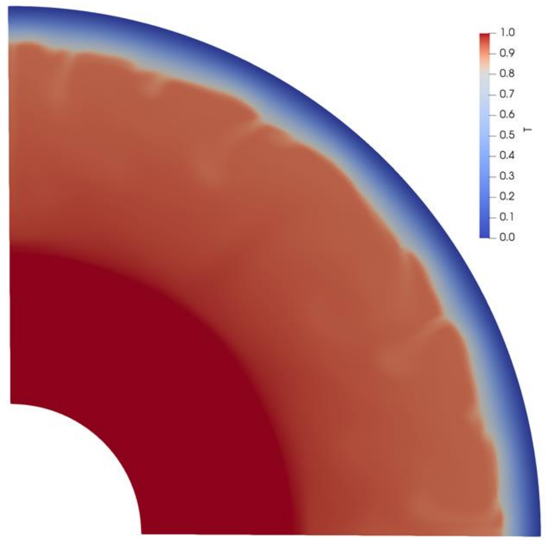
Future steps

- Generate a new dataset with more parameters and observables
    - Increased degeneracy and uncertainty expected

Parameters		Observables
Ra (η_{ref} $\in [10^{19}, 10^{22}$ Pa s])	←	Temperature profile
RaQ ($\Lambda \in [1, 50], \eta_{ref}$)		Radial contraction
$\mathbf{T_{initial}} \in [1600, 1800$ K]		Elastic lithospheric thickness
$\mathbf{E} \in [10^5, 5\text{x}10^5$ J mol^{-1}]		Core heat flux
$\mathbf{V} \in [4\text{x}10^{-6}, 10\text{x}10^{-6}$ m^3 mol^{-1}]		Surface heat flux
$\mathbf{\Delta\ T_{cmb}} \in [0, 300$ K]		Crustal thickness
$\mathbf{K_{regolith}}, \mathbf{D_{regolith}} \in$ [[const], [0, 5km]]		Duration of volcanism

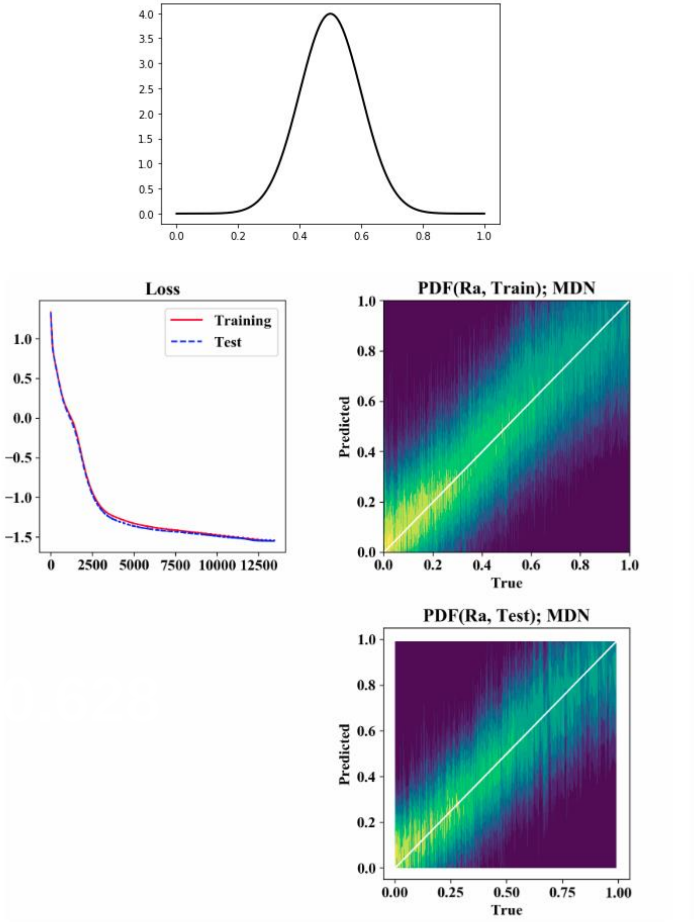
Future steps

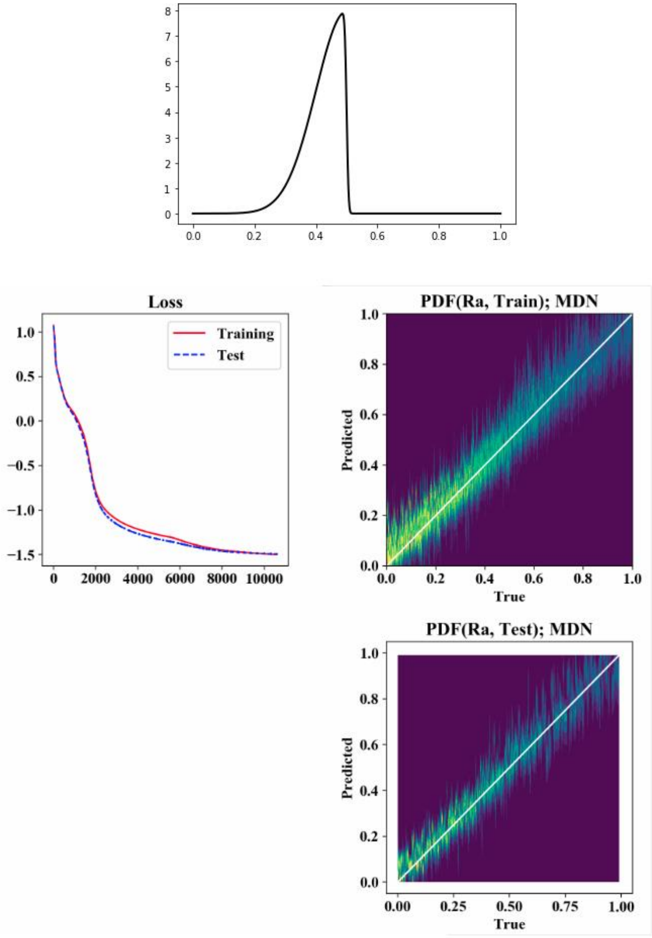
- Generate a new dataset with more parameters and observables
- Investigate higher-dimensional observables



Future steps

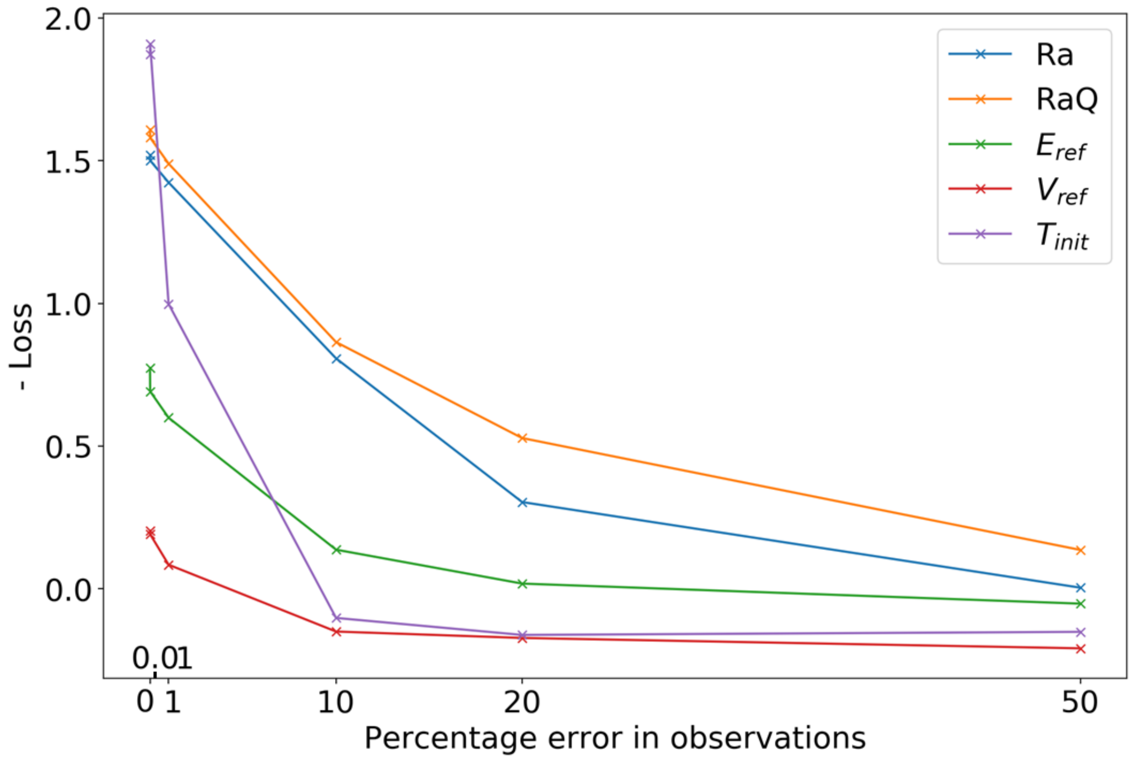
- Generate a new dataset with more parameters and observables
- Investigate higher-dimensional observables
- Explore some algorithmic modifications





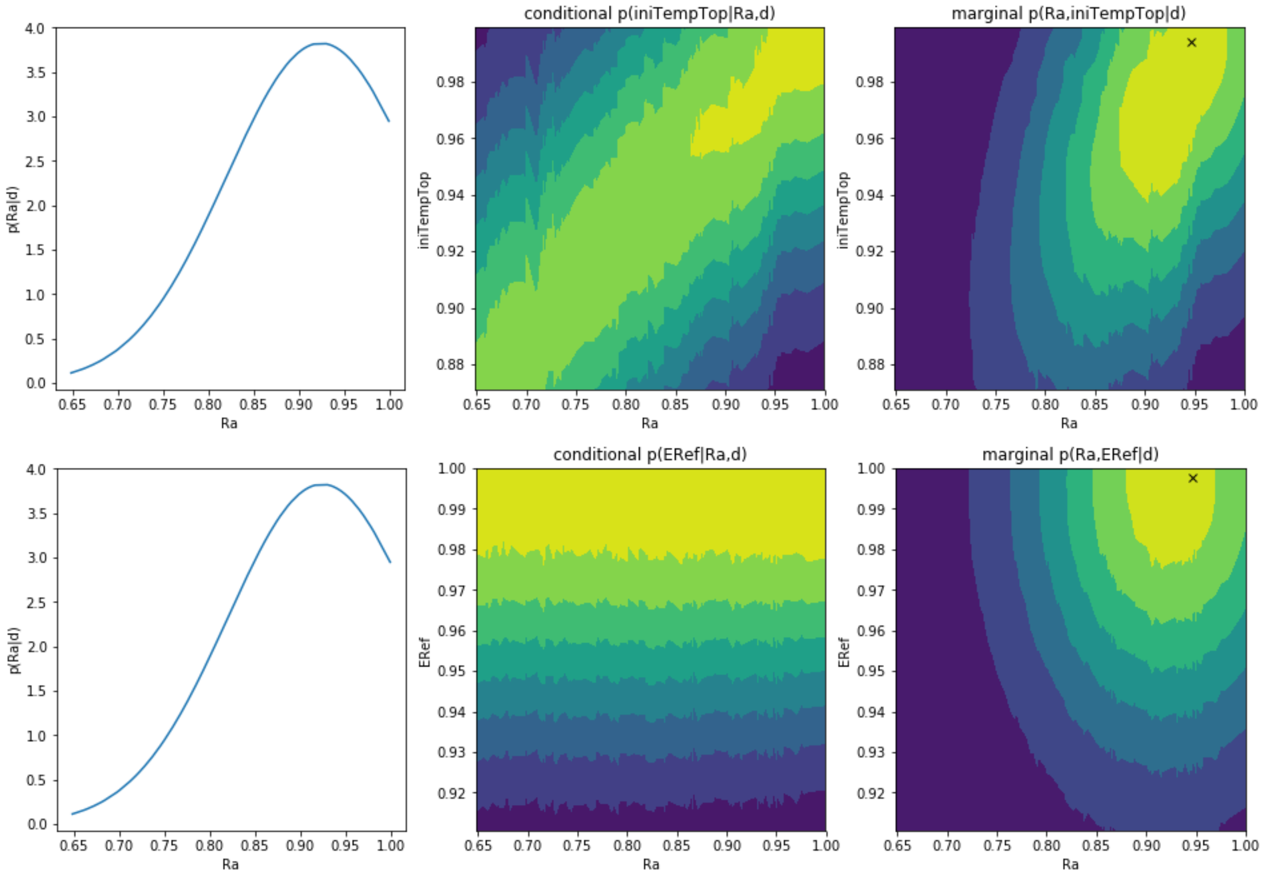
Future steps

- Generate a new dataset with more parameters and observables
- Investigate higher-dimensional observables
- Explore some algorithmic modifications
- Quantify precision requirements for each observable



Future steps

- Generate a new dataset with more parameters and observables
- Investigate higher-dimensional observables
- Explore some algorithmic modifications
- Quantify precision requirements for each observable
- Build a parameter-dimensional marginal PDF [4]



Acknowledgments

*I acknowledge the North-German Supercomputing Alliance (**HLRN**) for providing HPC resources that have contributed to the research results reported in this presentation.*

*I acknowledge the support of the Helmholtz Einstein International Berlin Research School in Data Science (**HEIBRiDS**), the German Aerospace Center (**DLR**) and Technical University of Berlin (**TUB**).*

References

1. “NASA's InSight Mars Lander.” *NASA*, NASA, 26 June 2019, mars.nasa.gov/insight/.

2. Gonzalez, Carlos. “What's the Difference Between Conduction, Convection, and Radiation?” *Machine Design*, 7 Aug. 2017, www.machinedesign.com/whats-difference-between/what-s-difference-between-conduction-convection-and-radiation.

3. Marshall, Rick. “Terminator: Dark Fate: Everything We Know about the New Movie so Far.” *Www.digitaltrends.com*, Digital Trends, 23 May 2019, 12:06PM PST, www.digitaltrends.com/movies/terminator-sequel-trilogy-news-cast/.

4. Atkins, Suzanne, et al. “Using Pattern Recognition to Infer Parameters Governing Mantle Convection.” *Physics of the Earth and Planetary Interiors*, vol. 257, 2016, pp. 171–186., doi:10.1016/j.pepi.2016.05.016.

5. Wit, Ralph W. L. De, et al. “Bayesian Inference of Earth's Radial Seismic Structure from Body-Wave Traveltimes Using Neural Networks.” *Geophysical Journal International*, vol. 195, no. 1, 2013, pp. 408–422., doi:10.1093/gji/ggt220.

6. Bishop, Christopher M. “Mixture Density Networks.” *Neural Computing Research Group*, Feb. 1994, doi:https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/bishop-ncrg-94-004.pdf.

7. Tosi, Nicola, et al. “Interior Dynamics and Evolution of the Terrestrial Planets.” 5th Potsdam - Berlin Colloquium. 2019, Berlin.