$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The interpretation of probability

Possible interpretations of probability theory (Tarantola, 1988):

1. A purely statistical interpretation: probabilities diescribe the outcome of random processes (in physics, economics, biology, etc.)

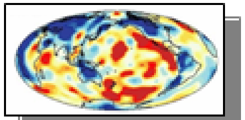
2. Probabilities describe **subjective degree of knowledge** of the true value of a physical parameter. Subjective means that the knowledge gained on a physical system may vary from experiment to experiment.

The key postulate of **probabilistic inverse theory** is (Tarantola 1988):

Let X be a discrete parameter space with a finite number of parameters. The most general way we have for describing any state of information on X is by defining a probability (in general a measure) over X.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Combining states of information



With basic principles from mathematical logic it can be shown that with two propositions f(x) (e.g. two data sets, two experiments, etc.) the combination of the two sources of information (with a logical **and**) comes down to

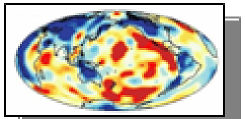
$$\sigma(\mathbf{x}) = \frac{f_1(\mathbf{x})\, f_2(\mathbf{x})}{\mu(\mathbf{x})}$$

This is called the conjunction of states of information (Tarantola and Valette, 1982). Here μ(x) is the non-informative pdf and s(x) will turn out to be the **a posteriori** probability density function.

This equation is the basis for probabilistic inverse problems:

We will proceed to combine information obtained from measurements with information from a physical theory.

Information from physical theories



Solving the forward problem is equivalent to predicting error free values of our data vector d, in the general case

$$d_{cal} = g(m)$$

Examples:

- ground displacements for an earthquake source and a given earth model
- travel times for a regional or global earth model
- polarities and amplitudes for a given source radiation pattern
- magnetic polarities for a given plate tectonic model and field revearsal history
- shaking intensity map for a given earthquake and model
-....

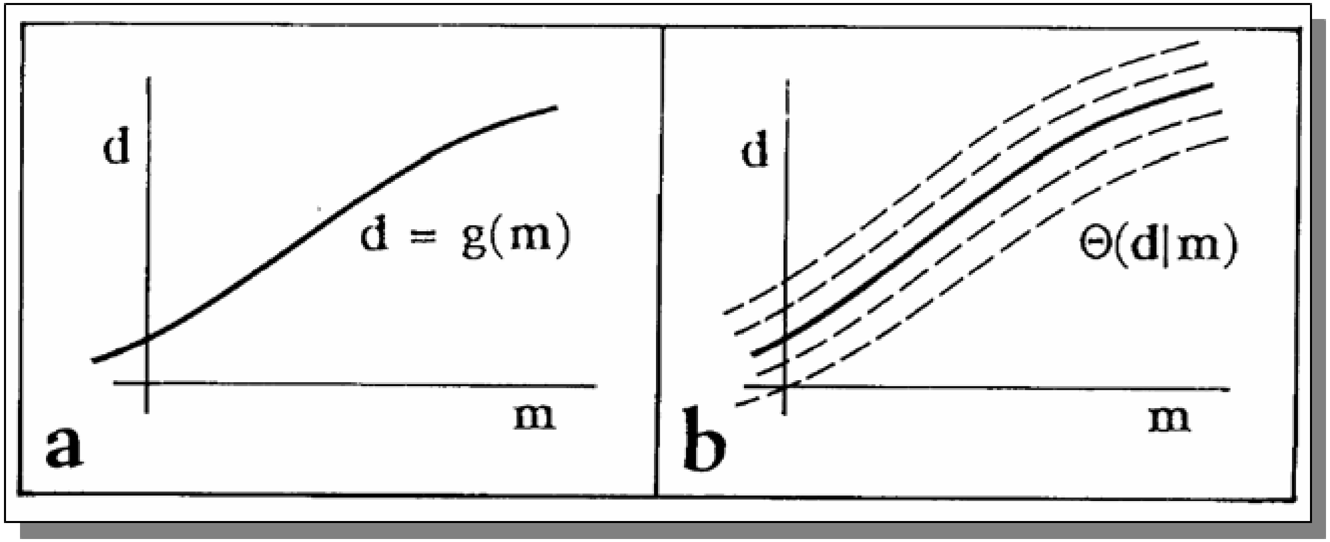
But: Our modeling may contain errors, or may not be the right physical theory, How can we take this into account?

Information from physical theories

$\Theta(d,m)$ summarized:

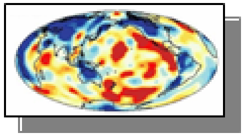
The expected correlations between model and data space can be described using the joint density function $\Theta(d,m)$. When there is an inexact physical theory (which is always the case), then the probability density for data d is given by $\Theta(d|m)\mu(m)$.

This may for example imply putting error bars about the predicted data d=g(m) … graphically



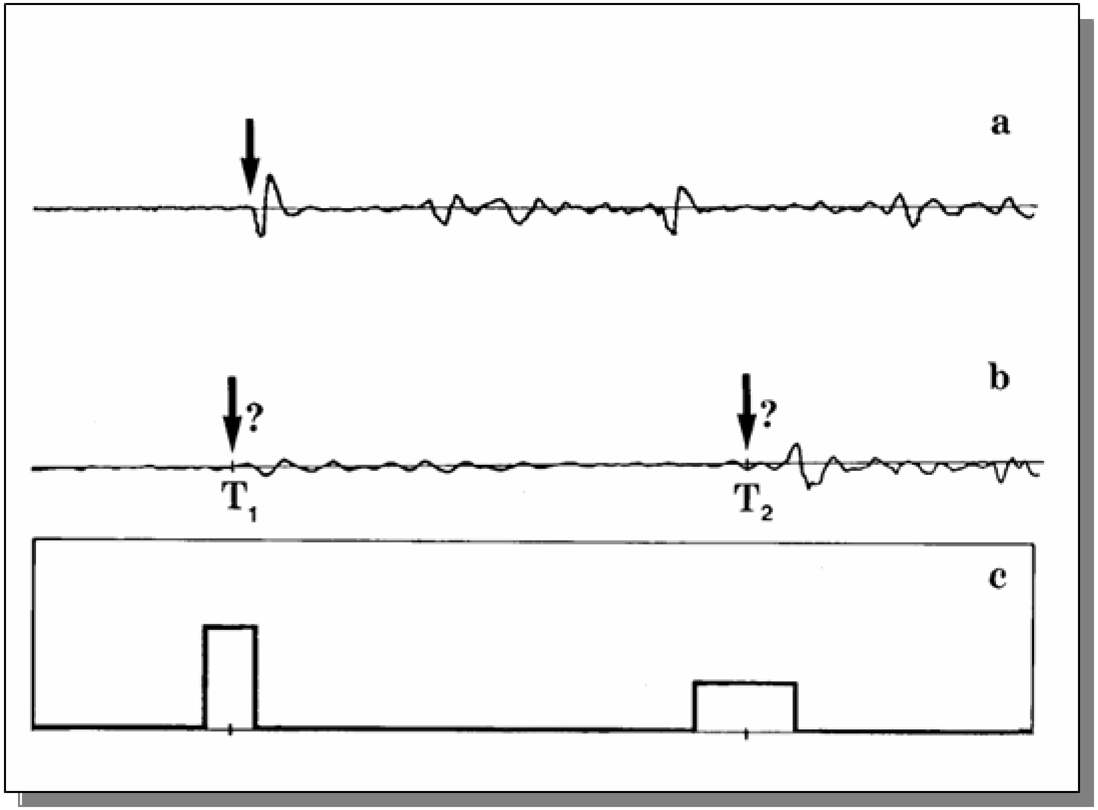
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Information from measurements



An experiment will give us **information on** the true values of observable parameters (but not actually the true values), we will call this pdf $\rho_D(d)$.

Example: Uncertainties of a travel time reading



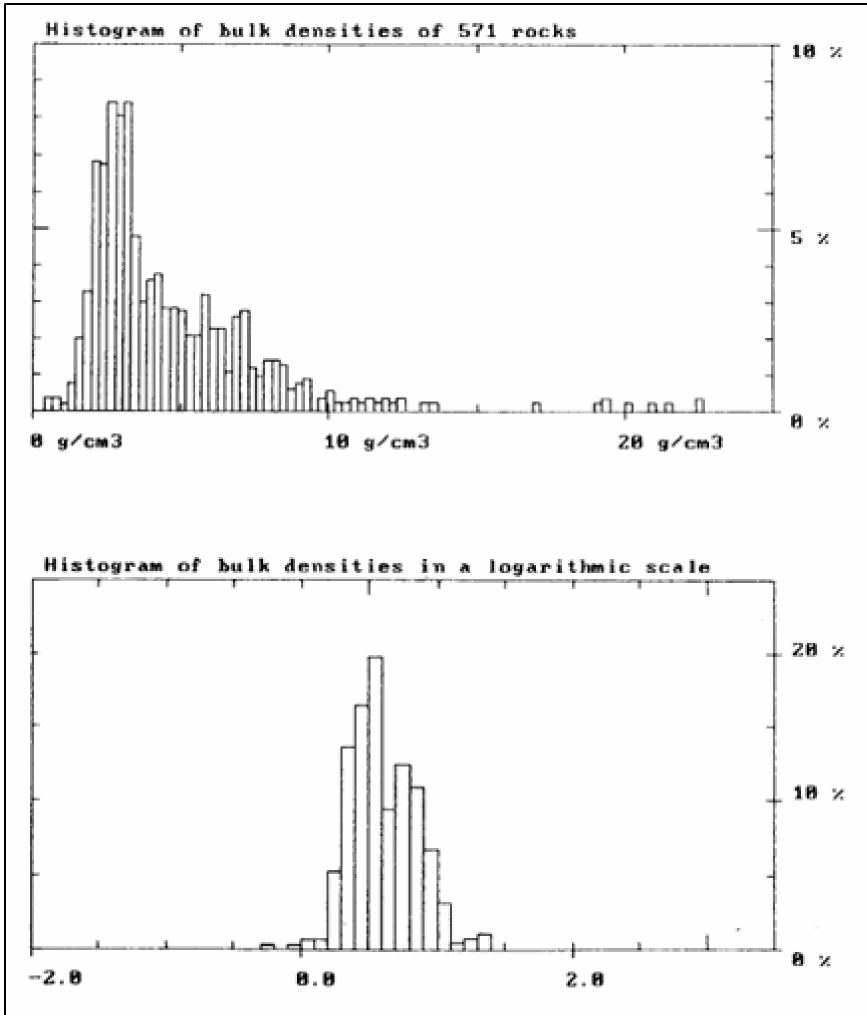
Good data

Noisy data

Uncertainty

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

A priori information on model parameters

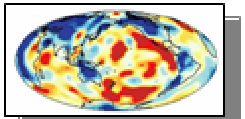


All the information obtained independently of the measurements on the model space is called **a priori information**. We describe this information using the pdf $\rho_M(m)$.

Example: We have no prior information $\rho_M(m)=\mu(m)$, where $\mu(m)$ is the non-informative prior.

Example: We are looking for a density model in the Earth (remember the treasure hunt). From sampling many many rocks we know what densities to expect in the Earth:

<- it looks like **lognormal** distributions are a Good way of describing some physical parameters



The solution to the inverse problem

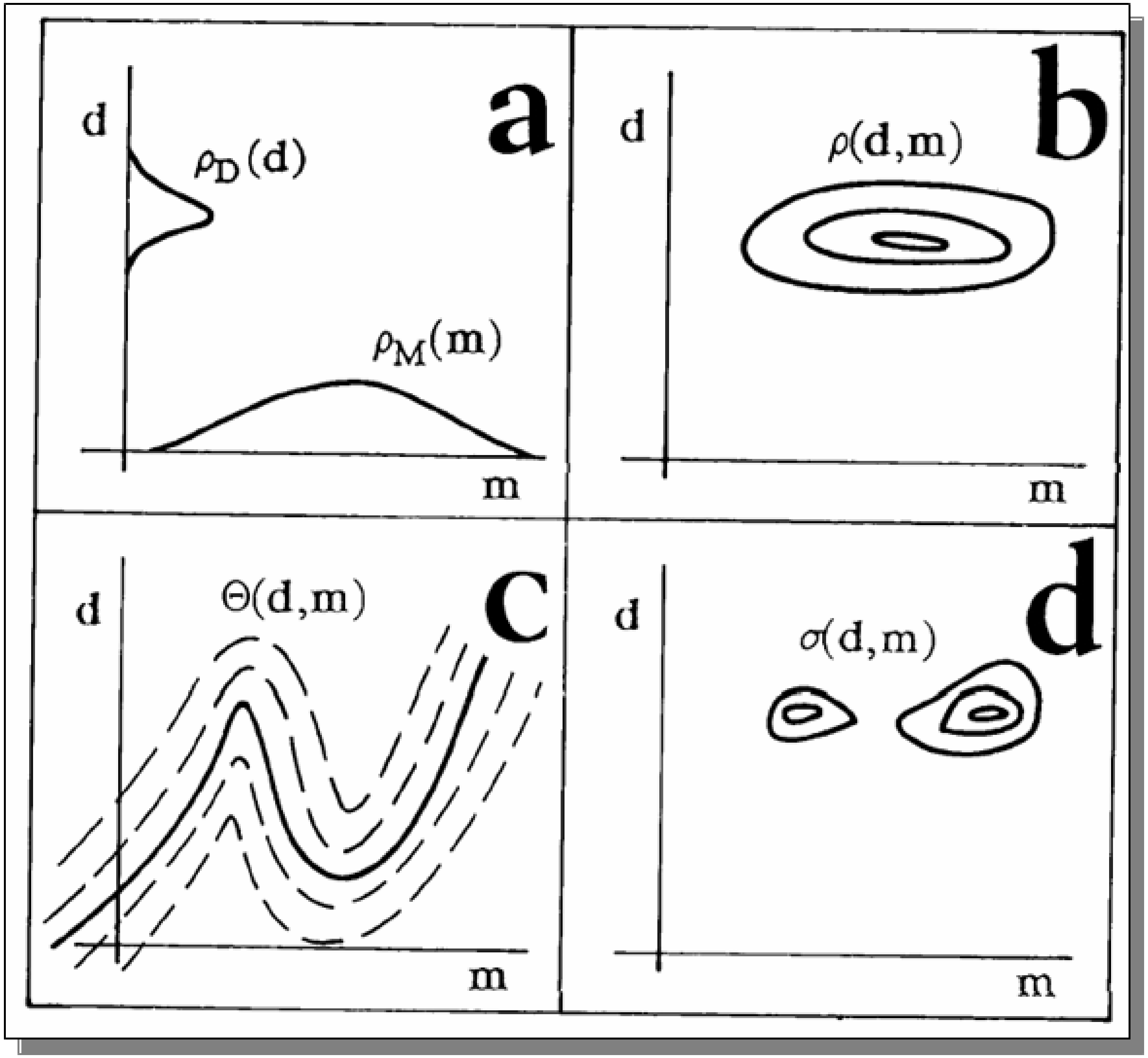
The information obtained **a priori** which we described with ρ(d,m) is now combined with information from a physical theory which we decribe with Θ(d,m). Following the ideas of conjunction of states of information, we define the **a posteriori probability density funtion** as **the** solution to an inverse problem

$$\sigma(d,m) = \frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Let's try and look at this graphically …

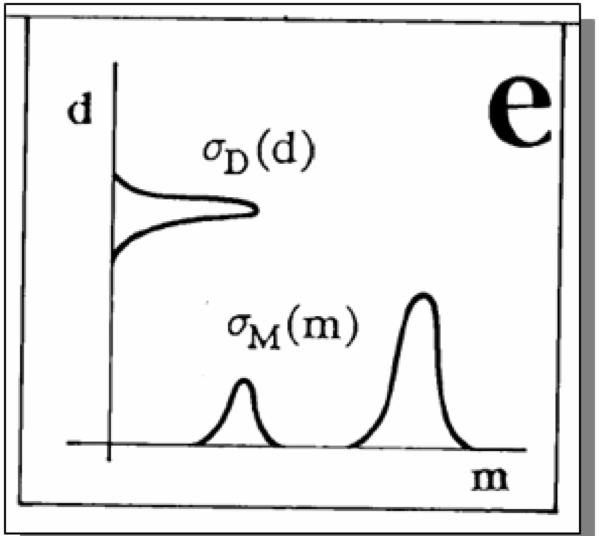
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The solution to the inverse problem



The (only) goal of this lecture is to understand these figures!

The rest is details …

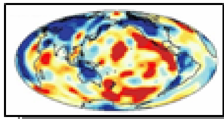


$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

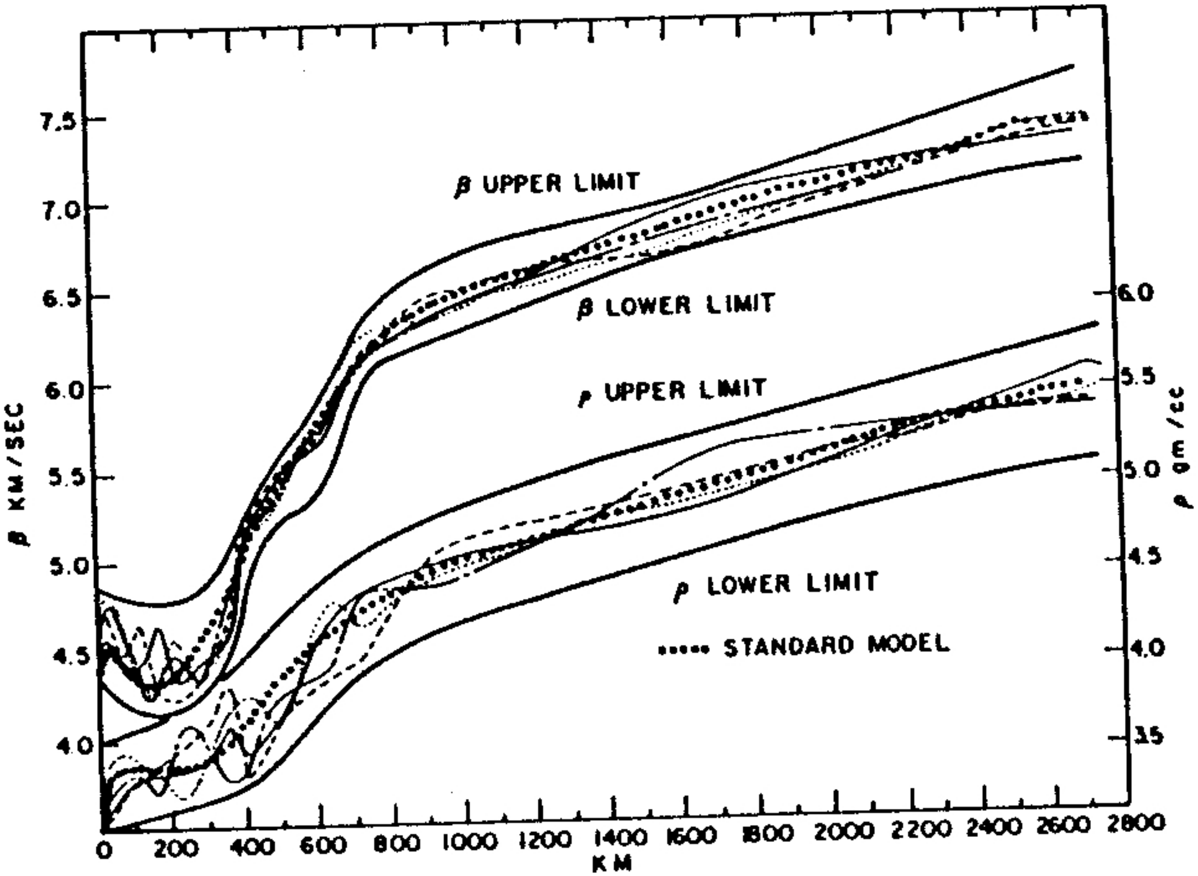
Monte Carlo Methods

$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Monte Carlo Methods



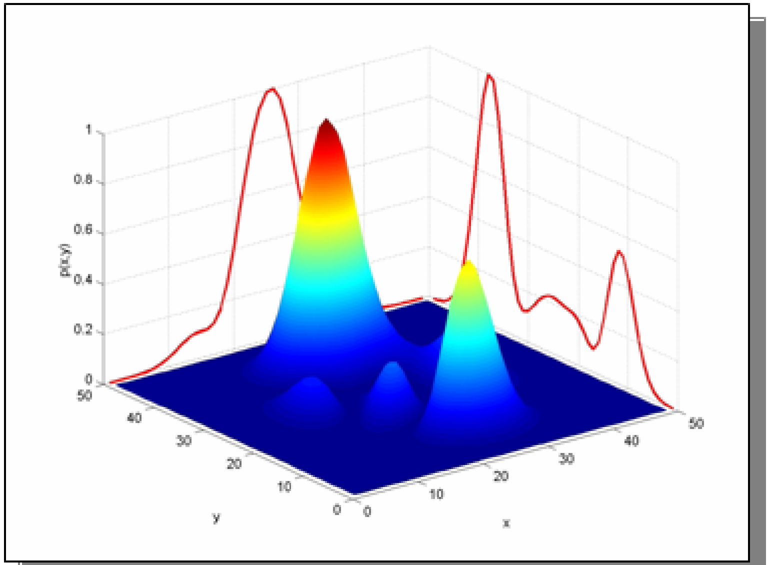
Early applications of Monte Carlo methods to the determination of Earth’s structure (Press 1968).

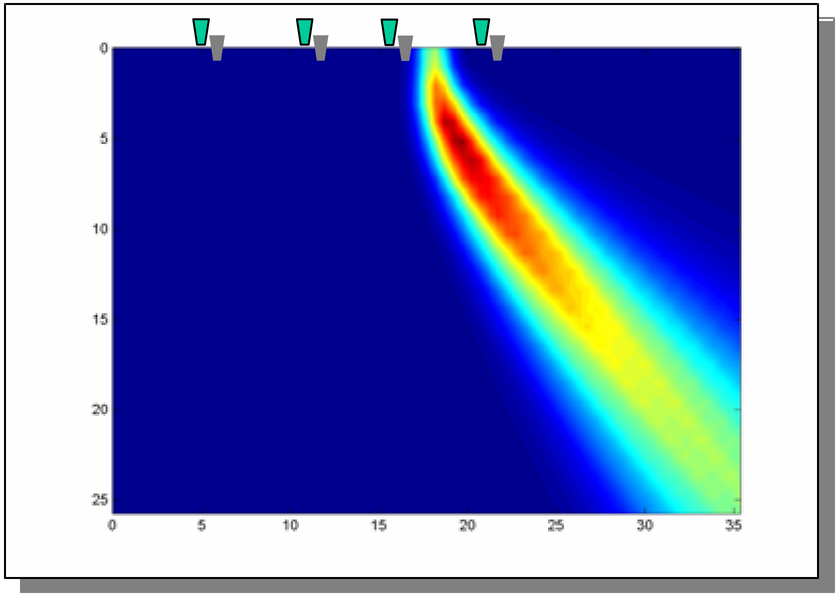


Monte Carlo Methods

The goal:

How can we efficiently **sample** the a posteriori pdf …





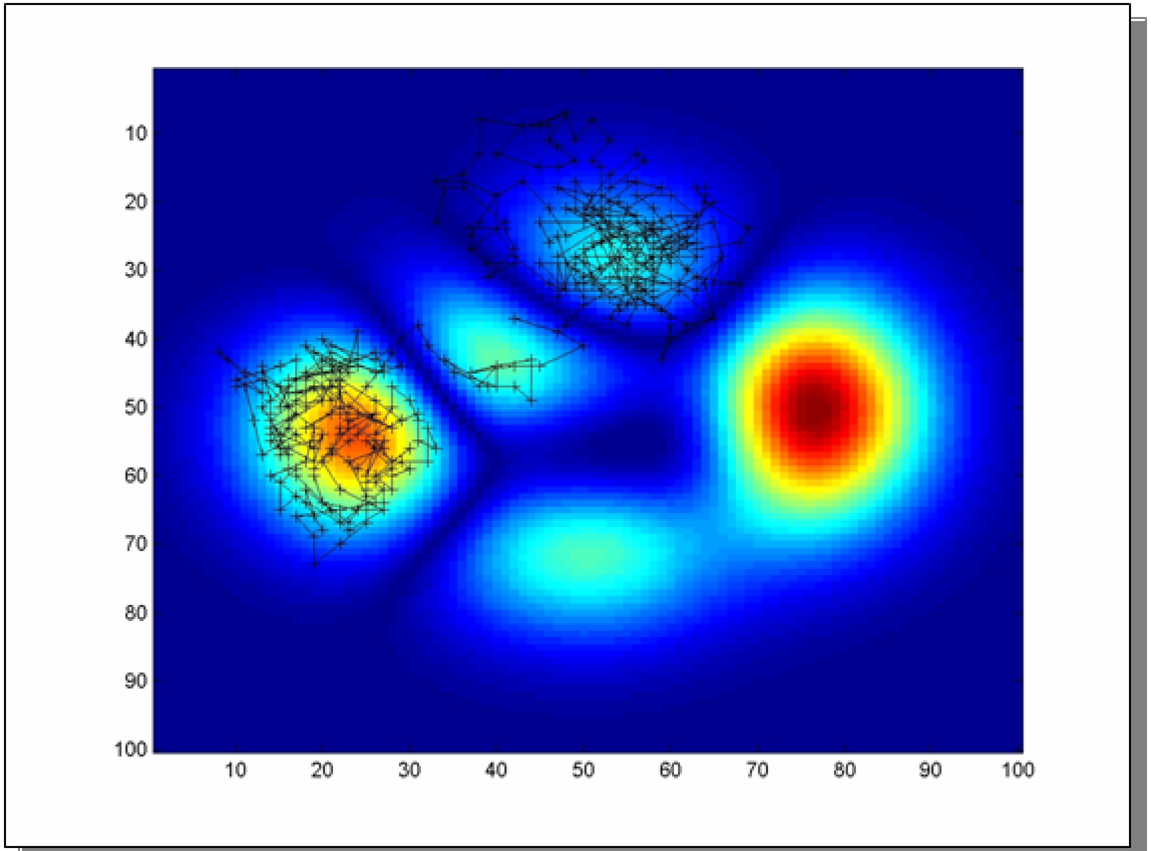
… knowing that the computation of the forward problem is expensive and computational power is finite!

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Random walks - Metropolis

First modifications:

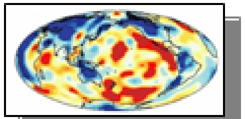
Limit the algorithm to look in the neighborhood of the present point. This is called **near neighbor sampling**.



Here we allowed the walker to move only within 10% of the total size of the model Space.

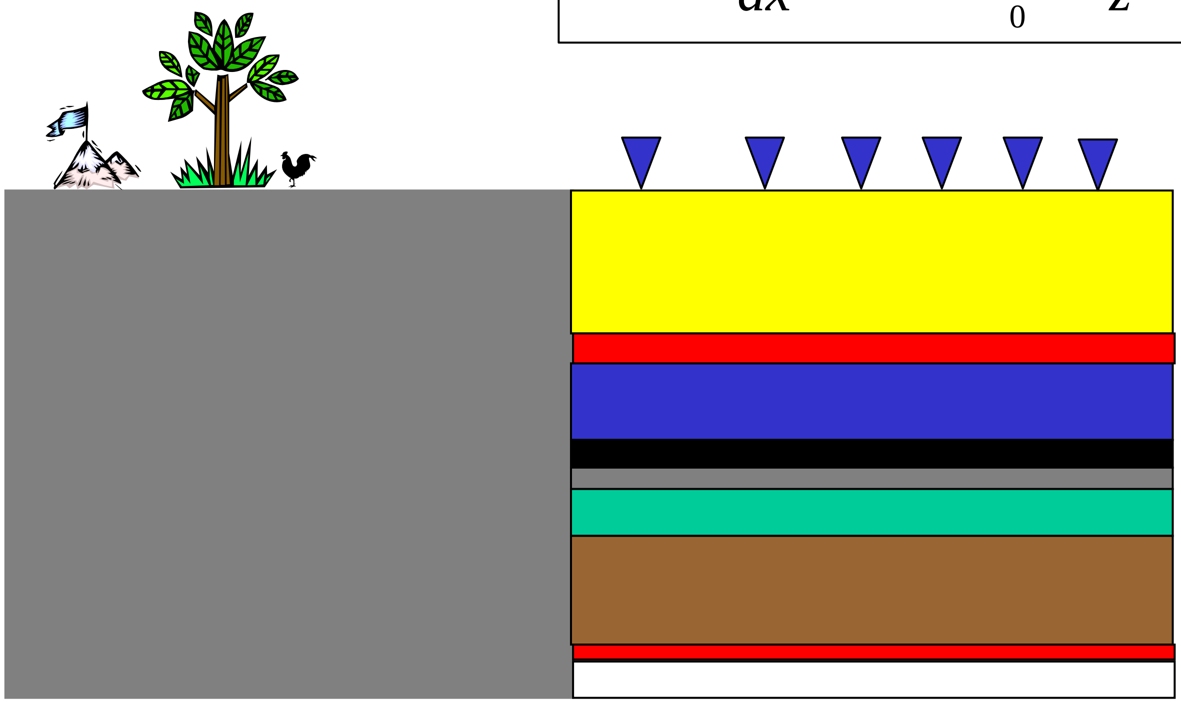
The program used was mc_neigh.m and the relevant Parameter neigh=0.1.

Monte Carlo method: gravity



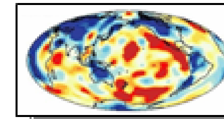
The forward problem: The gravity gradient at the surface is given by:

$$d(x) = \frac{dg}{dx}(x) = 2G\int_0^\infty dz \, \frac{z\Delta\rho(z)}{z^2 + x^2}$$



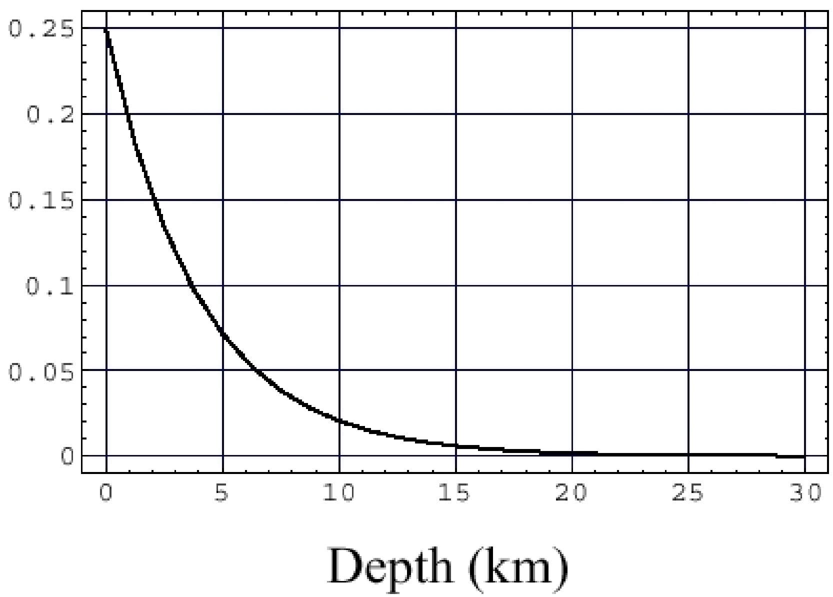
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Monte Carlo sampling of prior information

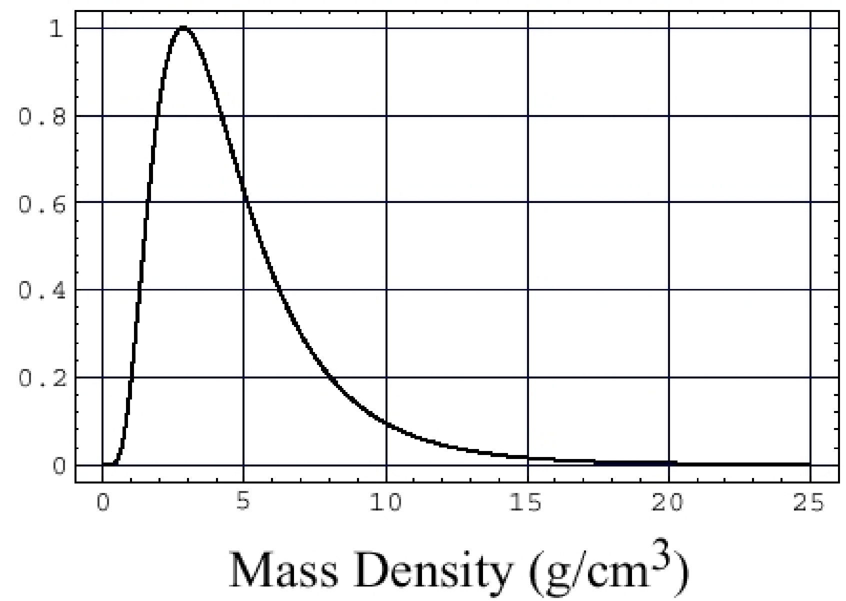


Let us look at the outcome of this process: **the prior probabilites**

Layer tickness (km). l_0=4km



Density (g/cm^3). ρ_0=3.98g/cm^3, σ=.58

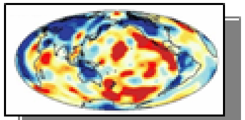


$$f(l) = \frac{1}{l_0}\exp\left(\frac{l}{l_0}\right)$$

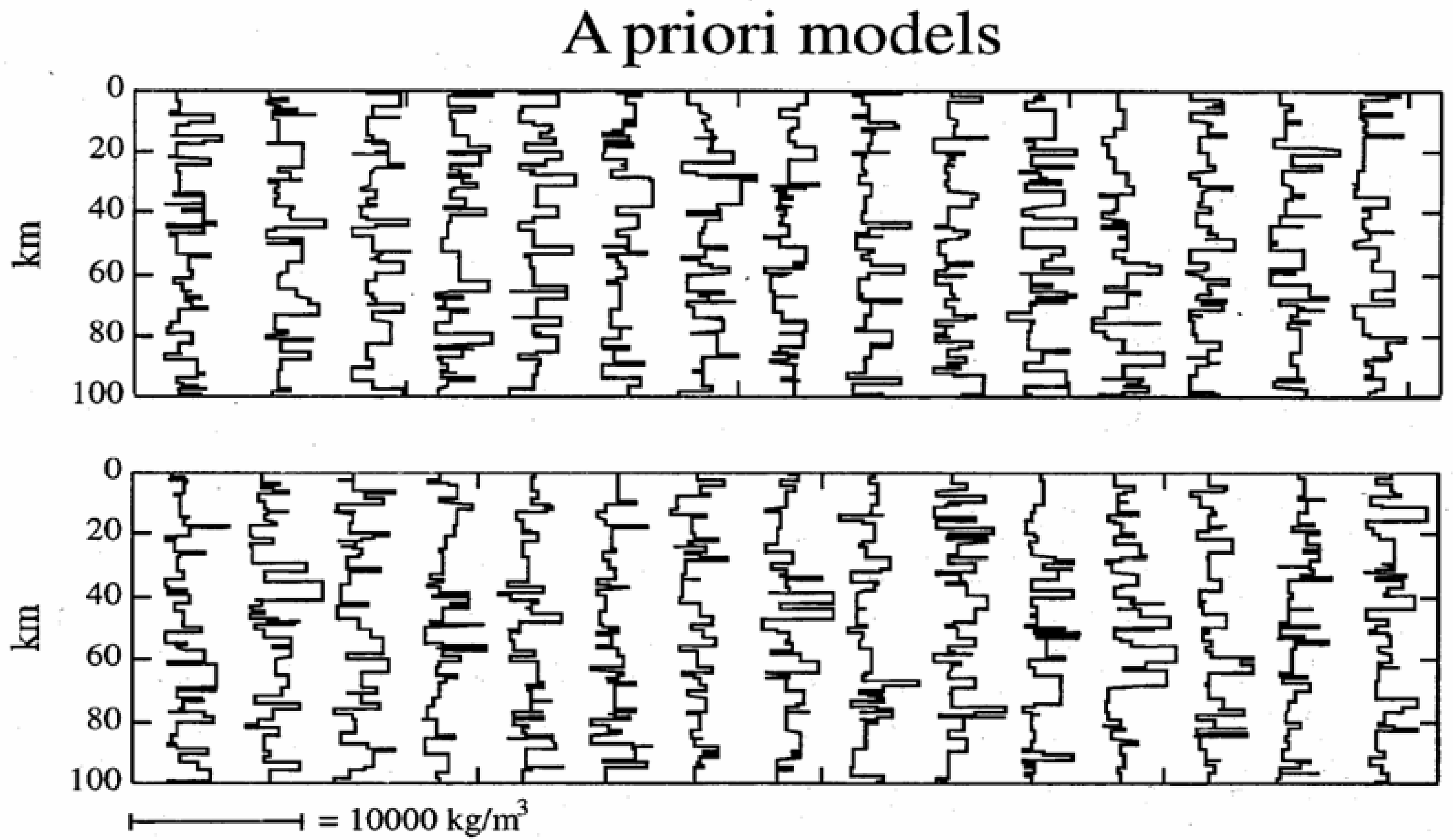
$$g(\rho) \propto \frac{1}{\rho}\exp\left(-\frac{1}{2\sigma^2}\left(\log\frac{\rho}{\rho_0}\right)^2\right)$$

$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$

Monte Carlo sampling of prior information



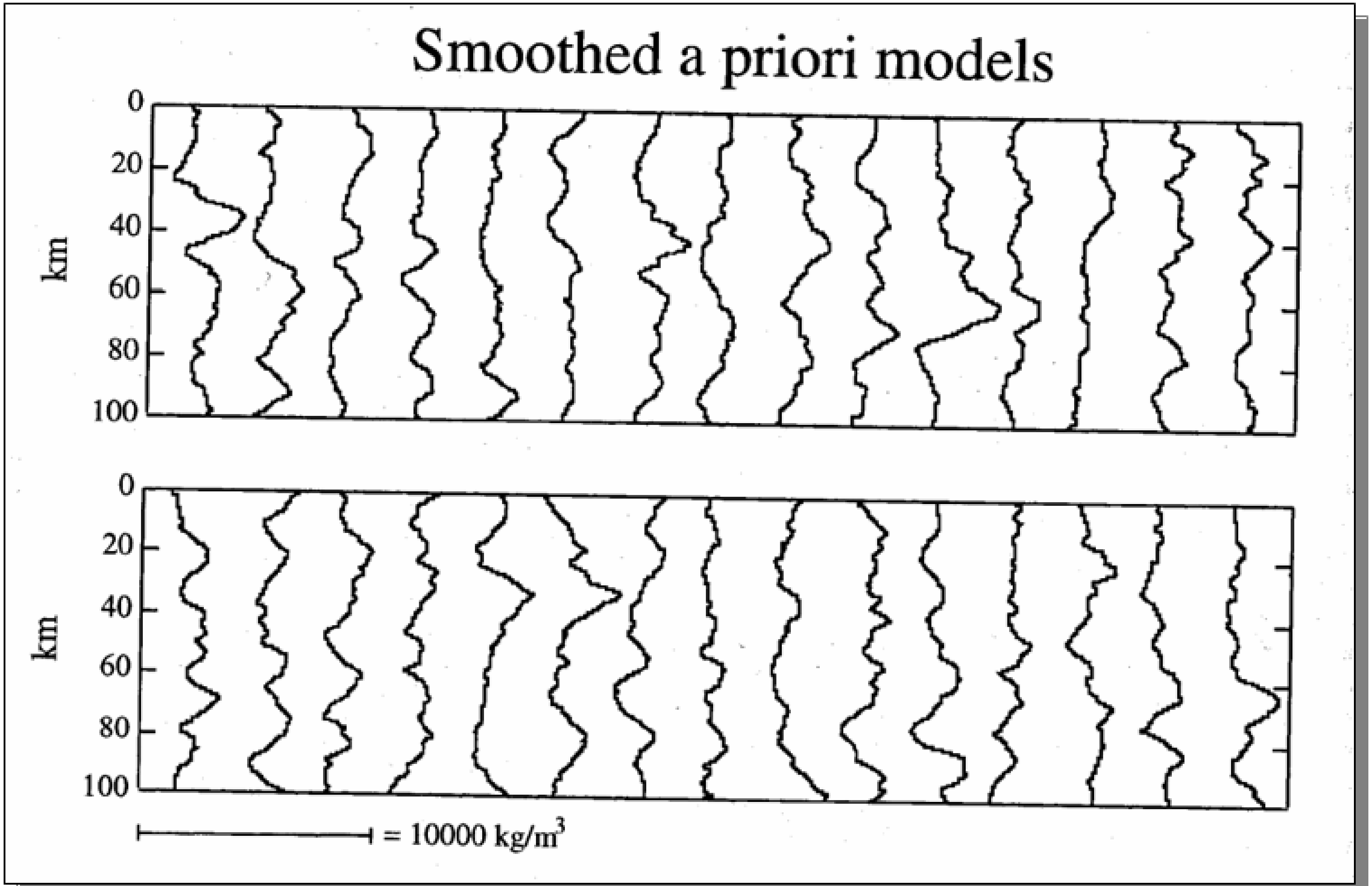
… a **random walk** through the prior probabilities produce models that look like this:



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

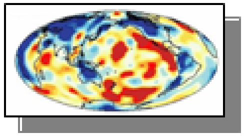
Monte Carlo sampling of prior information

... we do not expect that the fine layering is well resolved, which is why it makes sense to look a smoothed models …

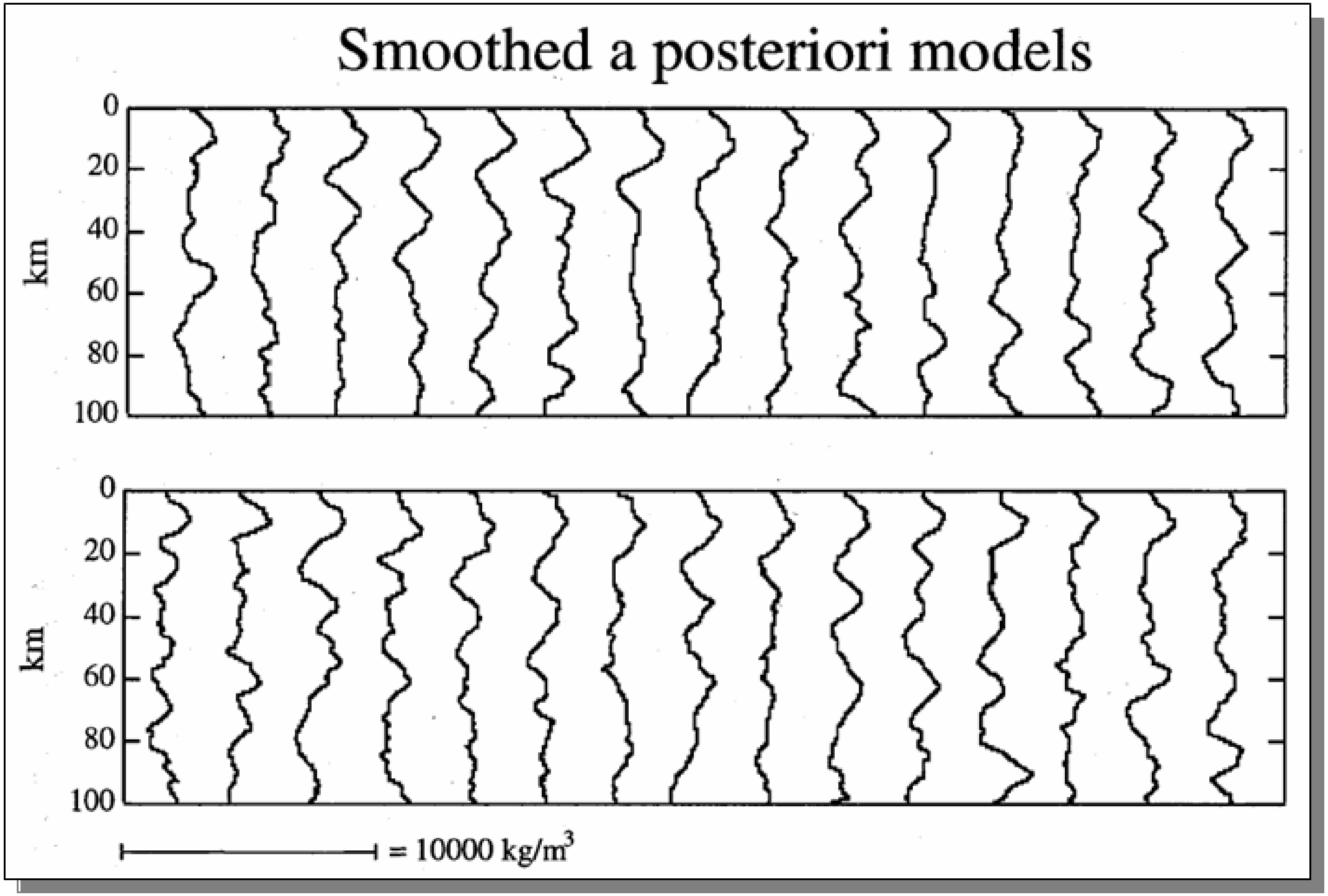


$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gravity: posterior random walk

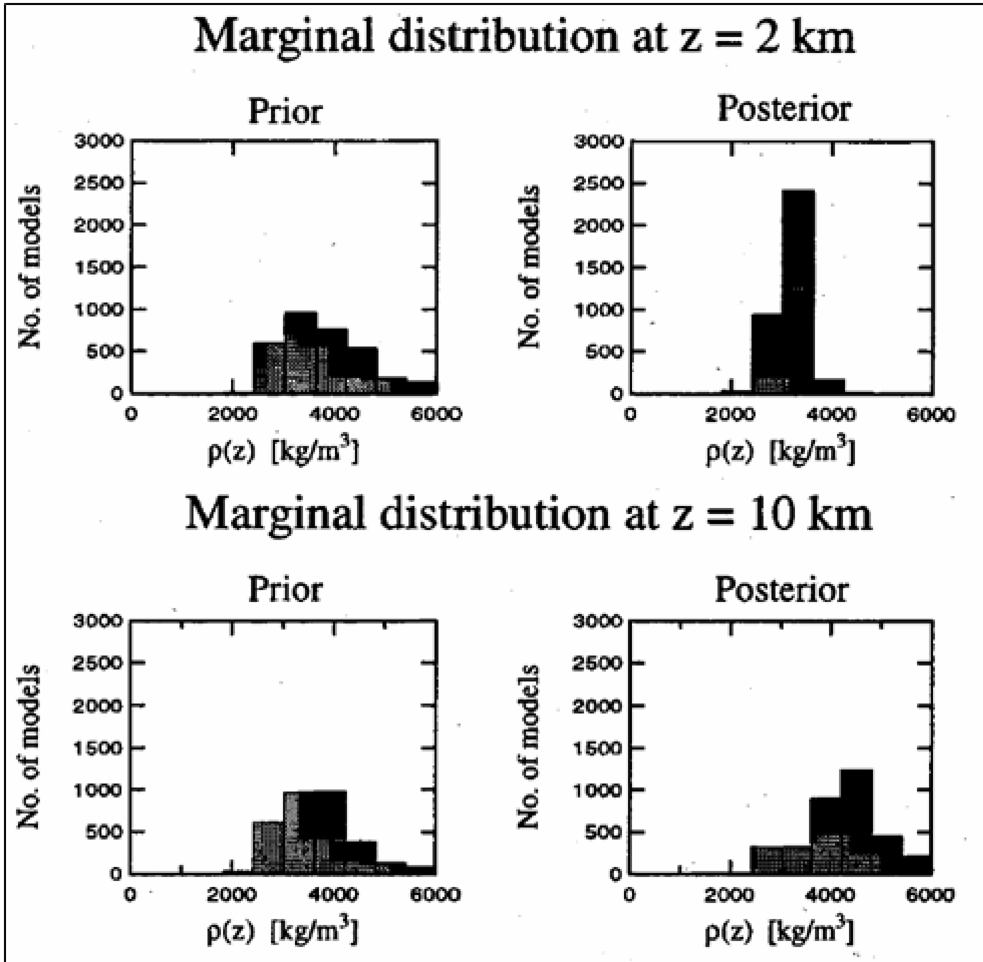


The random walk through the *a posteriori probability* leads to the models:



$$\sigma(d,m) = k \frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gravity: posterior probability



Note:

These models are samples of the a posteriori probability density function. They represent the state of information we have on our Earth model. With these samples we can now ask questions like:

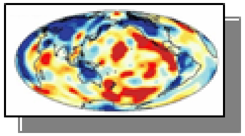
What is the value for the mass density at depth z=2km or z=20km and how well is it constrained? We only have to calculate the marginal probabilities to answer this questions.

Note:

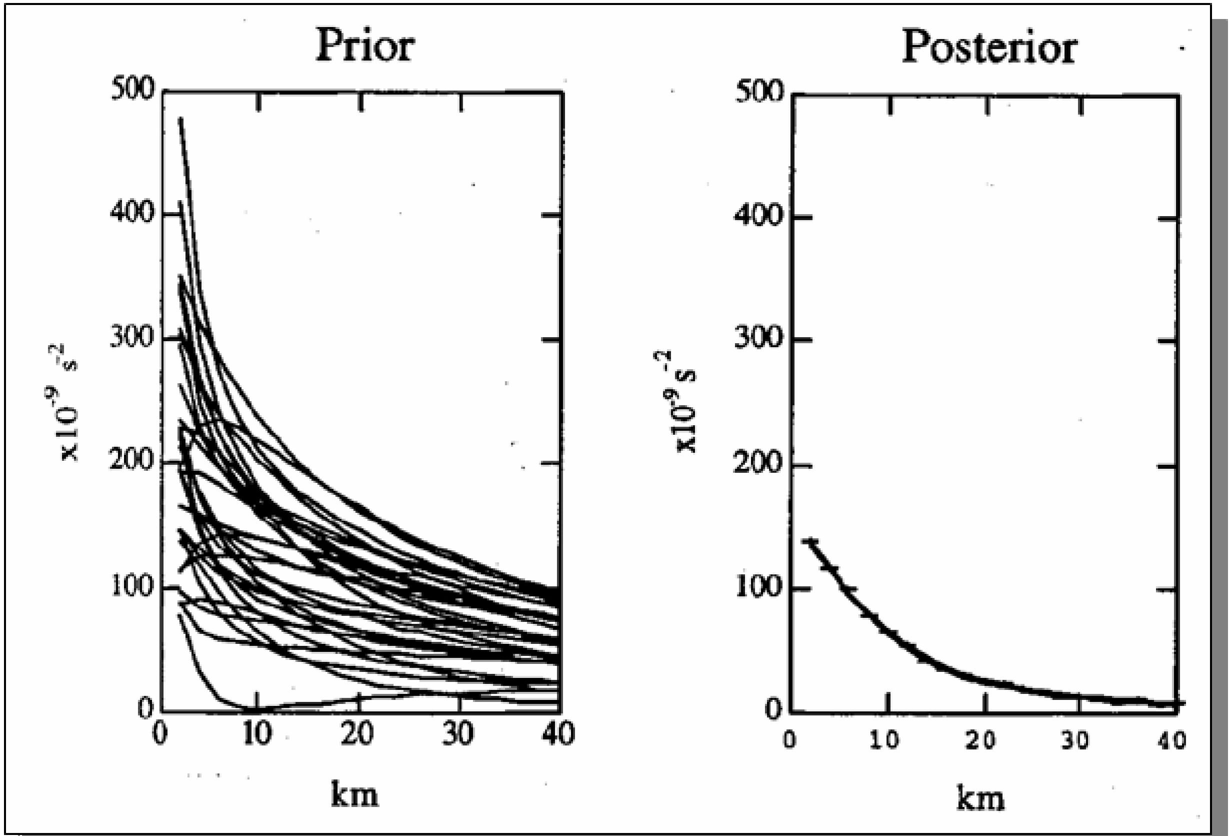
At depth 2km we seem to have clearly gained information.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gravity: posterior probability



How has the misfit of our models improved compared to the a priori models?

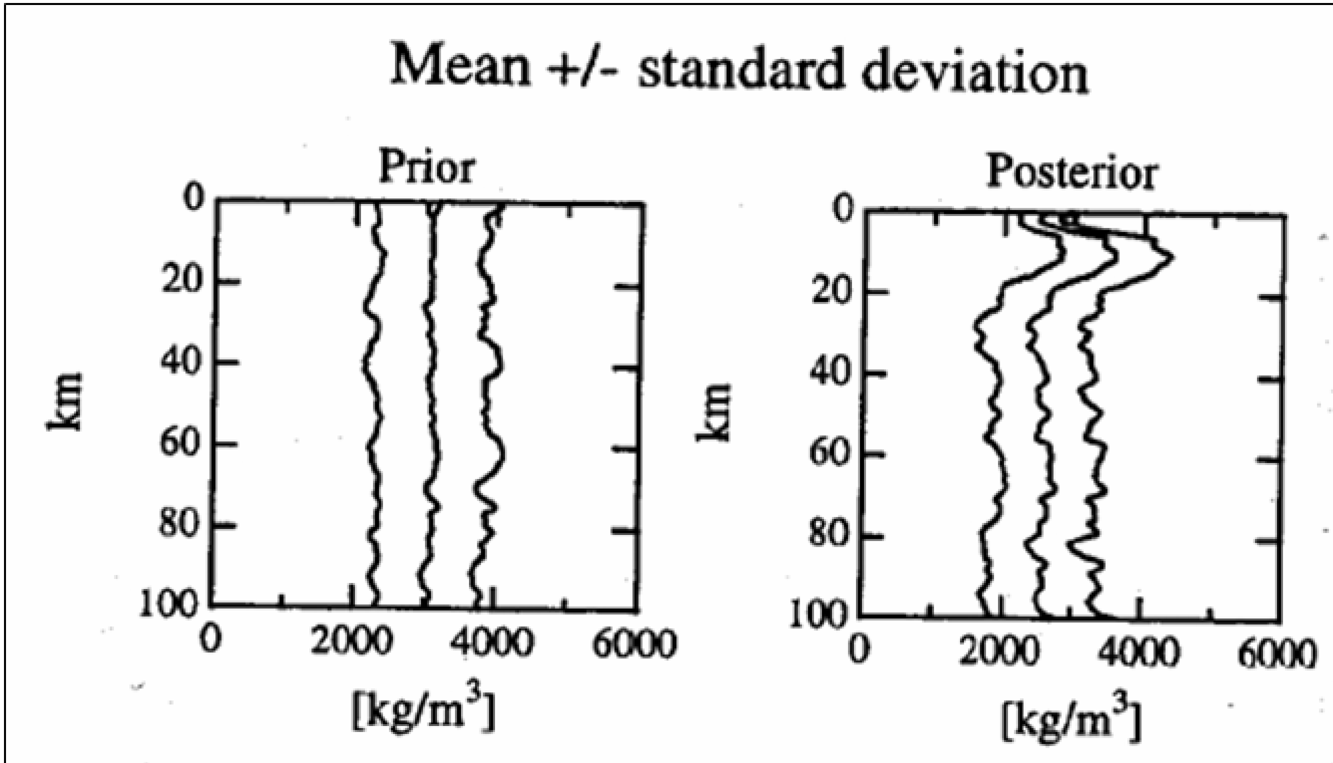


The misfit is almost perfect for all our a posteriori models but again we hit on the particular gravity problem that many **very different** models explain the data!

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Gravity: posterior probability

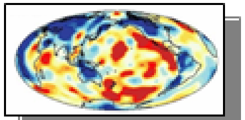
What are the mean values and standard deviations of the density as a function of depth?



Here we clearly see that we gave gained information in the top 20 km !

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Summary



Can we use some of those concepts in seismic waveform inversion problems?

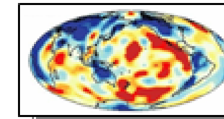
Can we decribe in a quantitative way prior information on our Earth models

- **To search for good starting models**
- **around some final models to investigate uncertainties**

Can we visualize at the same time Earth models AND their uncertainties?

….

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$



Gravity: experimental uncertainties

The measued data are assumed to be contaminated by random, uncorrelated noise. To make it a little more complicated, the errors are assumed to come from two processes with difference variances σ_i and relative probabilities (expressed through a):

$$f(\varepsilon) = \frac{a}{\sqrt{2\pi}\sigma_1}\exp\left(-\frac{\varepsilon^2}{2\sigma_1^2}\right) + \frac{1-a}{\sqrt{2\pi}\sigma_2}\exp\left(-\frac{\varepsilon^2}{2\sigma_2^2}\right)$$