HANDS-ON TUTORIALS

10 January 2018 - 9:00-12:00

AstroFit 2

DIAMONDS™
BAYESIAN SOFTWARE

TUTORIAL 1

POLYNOMIAL FIT

1. **Polynomial fit**
2. Gaussian fit
3. Multi-linear fit

AstroFit 2

DIAMONDS™ BAYESIAN SOFTWARE

DEFINING THE APPLICATION

POLYNOMIAL FIT

Polynomial function model

$$f(x) = b + a_1 x + a_2 x^2 + a_3 x^3 + \ldots$$

Parameter vector

$$\boldsymbol{\theta} = [b, a_1, a_2, a_3, \ldots]$$

Dataset

x	y	σ_y
0.00	30.70	20.00
1.00	31.17	20.00
2.00	34.37	20.00
3.00	31.52	20.00
4.00	33.77	20.00
5.00	37.90	20.00
6.00	40.40	20.00
7.00	22.65	20.00
8.00	32.80	20.00
9.00	39.60	20.00
10.00	28.74	20.00
11.00	38.10	20.00
12.00	43.79	20.00
13.00	43.70	20.00
14.00	33.57	20.00
15.00	50.51	20.00
16.00	52.40	20.00
17.00	51.77	20.00
18.00	53.07	20.00
19.00	58.37	20.00
20.00	53.11	20.00
21.00	50.83	20.00
22.00	48.49	20.00
23.00	56.22	20.00
24.00	55.70	20.00
25.00	52.11	20.00

covariates (independent variable)

observations (dependent variable)

observations uncertainties

Astro Fit 2

DIAMONDS™ BAYESIAN SOFTWARE

The likelihood adopted is a Normal likelihood because we are assuming that observations have Gaussian distributed uncertainties.

COMPILE THE CODE

POLYNOMIAL FIT

- Open your terminal and
  cd ~/Diamonds/tutorials/polynomial_fit/

- Open **PolynomialFit.cpp** and copy the compiling line based on your OS

  Linux OS

  g++ -o PolynomialFit PolynomialFit.cpp -L../../build/ -I../../include/ -ldiamonds -std=c++11

  Mac OS

  clang++ -o PolynomialFit PolynomialFit.cpp -L../../build/ -I ../../include/ -l diamonds -stdlib=libc++ -std=c++11 -Wno-deprecated-register

- For the moment do not execute the code

For any issues in the compilation of the code, please check the installation guide page of the code at: https://fys.kuleuven.be/ster/Software/Diamonds/installation-guide

EXECUTE LINEAR FIT

POLYNOMIAL FIT

- The code you just compiled works for any dimension but you need to make sure that the dataset you provide is meaningful and that the prior hyper parameters are in a number equal to the dimensionality of the model you want to fit
- We start by performing the polynomial fit of a simple linear regression
- 2 free parameters: offset + slope
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/polynomial_fit/
- To run this tutorial execute via terminal type
  ./PolynomialFit linear_data.txt prior_hyperParameters_linear.txt

	Minimum	Maximum
a_1	0.5	2.0
b	3.0	50.0

AstroFit 2

DIAMONDS™ BAYESIAN SOFTWARE

RESULTS FOR LINEAR FIT

POLYNOMIAL FIT

2D GRID-FIT
LN(Evidence): - 400.42583

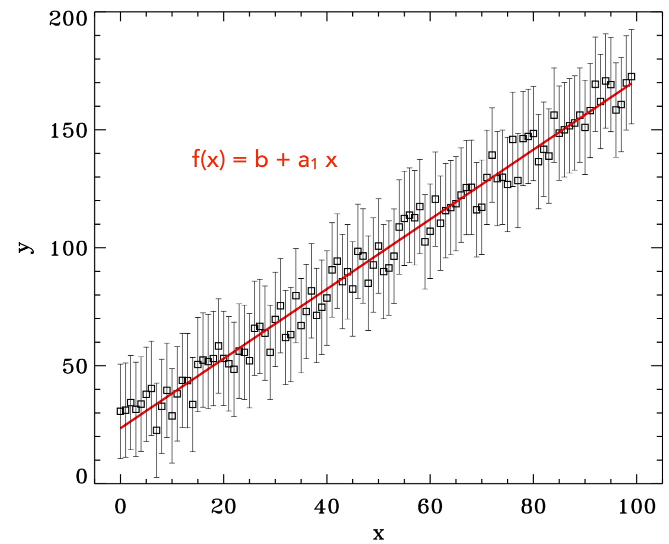
2D GRID-FIT
Slope (Median): 1.48 ± 0.07

2D GRID-FIT
Offset (Median): 24 ± 4

2D DIAMONDS
LN(Evidence): - 400.382

2D DIAMONDS
Slope (Median): 1.48 ± 0.07

2D DIAMONDS
Offset (Median): 23 ± 4



EXECUTE QUADRATIC FIT

POLYNOMIAL FIT

- Perform the polynomial fit of a quadratic regression
- 3 free parameters: offset + 2 slopes
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/polynomial_fit/
- To run this tutorial execute via terminal type
  ./PolynomialFit quadratic_data.txt prior_hyperParameters_quadratic.txt

	Minimum	Maximum
a_1	-0.1	2.0
a_2	-0.05	0.3
b	2.0	60.0

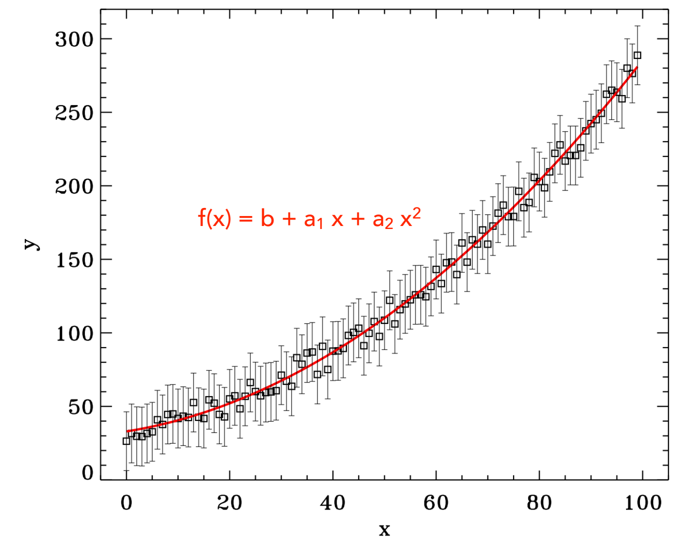
RESULTS FOR QUADRATIC FIT

POLYNOMIAL FIT

3D GRID-FIT
LN(Evidence): - 404.507

3D GRID-FIT
Slope 1 (Median): 0.6 ± 0.3

3D DIAMONDS
LN(Evidence): - 404.506



Slope1: 5.605823820e-01 GRID: 0.55681893

Slope2: 1.964364641e-02 GRID: 0.019789934

Offset: 3.300671086e+01 GRID: 32.885627

EXECUTE CUBIC FIT

POLYNOMIAL FIT

- Perform the polynomial fit of a cubic regression
- 4 free parameters: offset + 3 slopes
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/polynomial_fit/
- To run this tutorial execute via terminal type
  ./PolynomialFit cubic_data.txt prior_hyperParameters_cubic.txt

	Minimum	Maximum
a_1	0.5	3.0
a_2	0.02	0.14
a_3	-0.005	0.004
b	2.0	50.0

Astro Fit 2

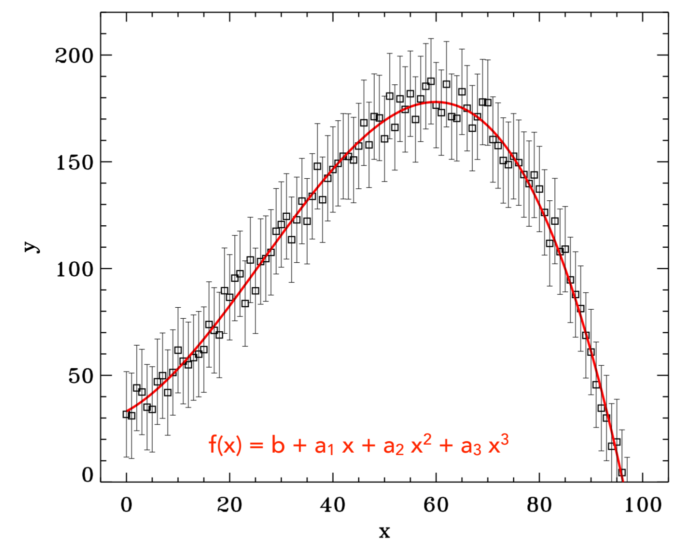
DIAMONDS™ BAYESIAN SOFTWARE

RESULTS FOR CUBIC FIT

POLYNOMIAL FIT

4D GRID-FIT
LN(Evidence): -407.1

4D DIAMONDS
LN(Evidence): -406.5



Slope1: 1.327562778e+00
Slope2: 7.692502852e-02
Slope3: -9.800494422e-04
Offset: 3.311938583e+01

TUTORIAL 2

GAUSSIAN FIT

1. Polynomial fit
2. **Gaussian fit**
3. Multi-linear fit

Astro Fit 2

DIAMONDS™
BAYESIAN SOFTWARE

DEFINING THE APPLICATION
GAUSSIAN FIT

Multi-dimensional Gaussian function model

$$G(x_i) = \prod_i \frac{1}{\sqrt{2\pi} s_i} \exp\left[- \left(\frac{x_i - \bar{x}_i}{s_i} \right)^2 \right]$$

Parameter vector

$$\boldsymbol{\theta}_i = [x_i, s_i]$$

DEFINING THE APPLICATION
GAUSSIAN FIT

Dataset

y	σ_y	x_1	x_2	x_3	…
0.00523753	0.002	3.607	11.806		
0.00212213	0.002	8.478	9.489		
0.00244753	0.002	-8.359	10.374		
0.00000906	0.002	-7.027	13.607		
0.00000142	0.002	8.611	13.834		
0.01491920	0.002	0.896	8.529		
0.00000001	0.002	-9.263	5.114		
0.00000114	0.002	8.678	6.122		
0.02622205	0.002	-2.316	9.186		
0.00929480	0.002	6.220	9.870		
0.00017350	0.002	7.427	7.424		
0.00013327	0.002	7.078	7.248		
0.00730352	0.002	-6.628	9.732		
0.00000471	0.002	1.567	5.740		
0.00172052	0.002	8.951	9.909		

observations (dependent variable)

covariates (independent variable)

observations uncertainties

DIAMONDS™ BAYESIAN SOFTWARE

There should be as many covariates as the dimensionality of the Gaussian we want to fit.
The dimensionality of the Gaussian has not to be confused with the dimensionality of the fitting problem, which has 2 free parameters for each dimension of the Gaussian, i.e. for each covariates.
The likelihood adopted is a Normal likelihood because we are assuming that observations are Gaussian distributed.

COMPILE THE CODE

GAUSSIAN FIT

- Open your terminal and
  cd ~/Diamonds/tutorials/gaussian_fit/

- Open **GaussianFit.cpp** and copy the compiling line based on your OS

  Linux OS
  g++ -o GaussianFit GaussianFit.cpp -L../../build/ -I../../include/ -ldiamonds -std=c++11

  Mac OS
  clang++ -o GaussianFit GaussianFit.cpp -L../../build/ -I ../../include/ -l diamonds -
  stdlib=libc++ -std=c++11 -Wno-deprecated-register

- For the moment do not execute the code

EXECUTE 1D GAUSSIAN FIT

GAUSSIAN FIT

- Perform the Gaussian fit in 1D, meaning that you have only one set of covariates
- 2 free parameters: 1 mean + 1 standard deviation
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/gaussian_fit/
- Copy the file **NSMC_configuringParameters_1D.txt** to **NSMC_configuringParameters.txt**
- To run this tutorial execute via terminal type
  ./GaussianFit gaussian_1D.txt prior_hyperParameters_1D.txt

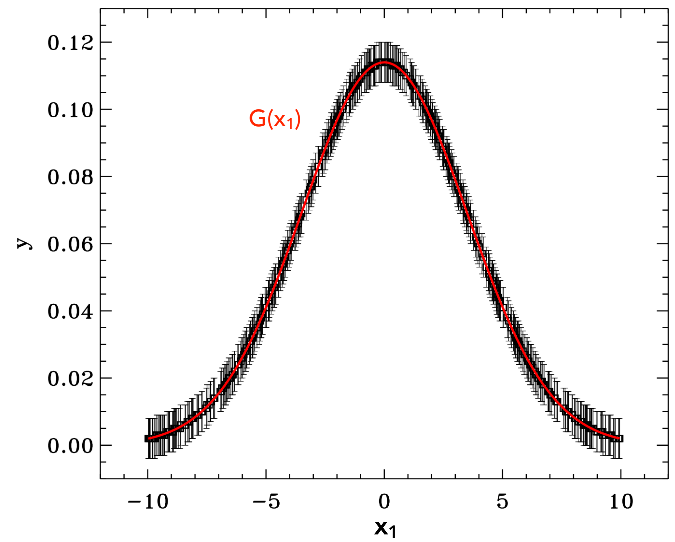
RESULTS FOR 1D GAUSSIAN FIT

GAUSSIAN FIT

2D GRID-FIT
LN(Evidence): 2088.996

2D DIAMONDS
LN(Evidence): 2088.969

$$\bar{x_1} = 0.0$$
$$s_1 = 3.5$$



What % of agreement do you have between your fitted parameters and the ones provided in the slide?

EXECUTE 2D GAUSSIAN FIT

GAUSSIAN FIT

- Perform the Gaussian fit in 2D, meaning that you have two sets of covariates
- 4 free parameters: 2 mean + 2 standard deviation
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/gaussian_fit/
- Copy the file **NSMC_configuringParameters_2D.txt** to **NSMC_configuringParameters.txt**
- To run this tutorial execute via terminal type
  ./GaussianFit gaussian_2D.txt prior_hyperParameters_2D.txt

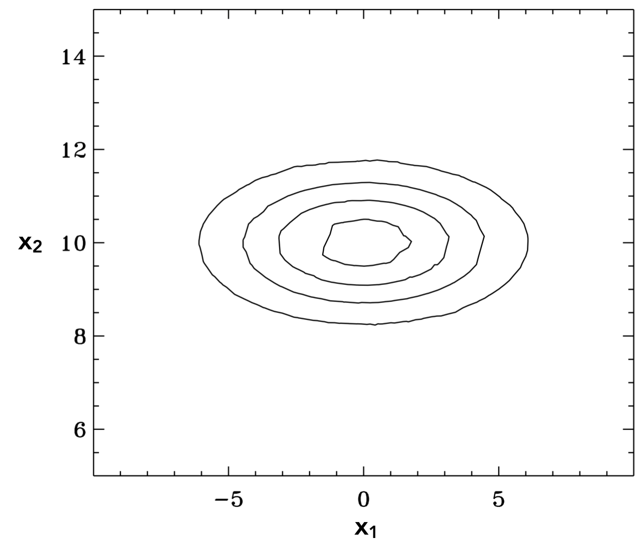
RESULTS FOR 2D GAUSSIAN FIT

GAUSSIAN FIT

4D GRID-FIT
LN(Evidence): 10524

4D DIAMONDS
LN(Evidence): 10570

$$\bar{x_1} = 0.0$$
$$s_1 = 3.5$$



$$\bar{x_2} = 10.0$$
$$s_2 = 1.0$$

What % of agreement do you have between your fitted parameters and the ones provided in the slide?

EXECUTE 3D GAUSSIAN FIT

GAUSSIAN FIT

- Perform the Gaussian fit in 3D, meaning that you have three sets of covariates
- 6 free parameters: 3 mean + 3 standard deviation
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/gaussian_fit/
- Copy the file **NSMC_configuringParameters_3D.txt** to **NSMC_configuringParameters.txt**
- To run this tutorial execute via terminal type
  ./GaussianFit gaussian_3D.txt prior_hyperParameters_3D.txt

RESULTS FOR 3D GAUSSIAN FIT

GAUSSIAN FIT

$$\bar{x}_1 = 0.0$$
$$s_1 = 3.5$$

$$\bar{x}_2 = 10.0$$
$$s_2 = 1.0$$

$$\bar{x}_3 = -5$$
$$s_3 = 0.1$$

What % of agreement do you have between your fitted parameters and the ones provided in the slide?

TUTORIAL 3

MULTI-LINEAR FIT

1. Polynomial fit
2. Gaussian fit
3. **Multi-linear fit**

AstroFit 2

DIAMONDS™ BAYESIAN SOFTWARE

DEFINING THE APPLICATION

MULTI-LINEAR FIT

Multi-linear function model and total uncertainty

$$M(x_i; a_i, b) = b + a_1 x_1 + a_2 x_2 + \ldots$$

$$s^2_{\mathrm{tot}}(\sigma_y, \sigma_i; a_i) = \sigma_y^2 + a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \ldots$$

Parameter vector

$$\boldsymbol{\theta} = [b, a_i]$$

Multi-Linear Likelihood

$$\ln \mathcal{L}(\boldsymbol{\theta}) \propto \sum_i^N \left[\frac{\ln O_i - \ln E_i(\boldsymbol{\theta})}{\sigma_i(\boldsymbol{\theta})} \right]^2$$

This application requires that the covariates correspond to a natural logarithm of a dimensionless quantity. This is verified when we want to fit a power law model depending on many variables, each one calibrated with a given exponent. In this case, each exponent of the power law will be a coefficient of the multi-linear fit when we linearize the power law model by taking its natural logarithm (see next slide for an application, and Corsaro et al. 2013, Corsaro et al. 2017b).

The total uncertainty is calculated through the Gaussian error propagation law and has a defined analytical expression that is incorporating the uncertainty on the observations and the uncertainties on each of the covariates multiplied by the associated free parameters (i.e. the exponents of the power law).

A REAL APPLICATION

MULTI-LINEAR FIT

Multi-linear function model and total uncertainty

$$M(x_i; a_i, b) = b + a_1 x_1 + a_2 x_2 + \ldots$$

$$s^2_{\mathrm{tot}}(\sigma_y, \sigma_i; a_i) = \sigma_y^2 + a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \ldots$$

$$\left(\frac{a_{\mathrm{meso}}}{a_{\mathrm{meso},\odot}}\right) = \beta \left(\frac{\nu_{\mathrm{max}}}{\nu_{\mathrm{max},\odot}}\right)^s \left(\frac{M}{M_\odot}\right)^t \mathrm{e}^{u[\mathrm{Fe/H}]}$$

$$\ln\left(\frac{a_{\mathrm{meso}}}{a_{\mathrm{meso},\odot}}\right) = \ln\beta + s\ln\left(\frac{\nu_{\mathrm{max}}}{\nu_{\mathrm{max},\odot}}\right) + t\ln\left(\frac{M}{M_\odot}\right) + u\,[\mathrm{Fe/H}]$$

$$\widetilde{\sigma}_a^2(s,t,u) = \widetilde{\sigma}^2_{a_{\mathrm{meso}}} + s^2\widetilde{\sigma}^2_{\nu_{\mathrm{max}}} + t^2\widetilde{\sigma}^2_M + u^2\widetilde{\sigma}^2_{[\mathrm{Fe/H}]}$$

Models have to be linearized, if they are not already. This is the typical case of scaling relations, namely power law models in multi-dimensions, very useful to scale quantities and widely adopted e.g. in asteroseismology. A scaling relation is a power law of different variables, with terms scaled (or normalized) by a reference value.

DEFINING THE APPLICATION

MULTI-LINEAR FIT

Astro Fit 2

Dataset

y	σ_y	x_1	σ_{x1}	x_2	σ_{x2}	x_3	σ_{x3}
-4.2685	0.0427	-4.6324	0.0069	0.06	0.04	0.3196	0.0205
-4.2787	0.0481	-4.6393	0.0101	-0.05	0.05	0.3196	0.0205
-4.2779	0.0391	-4.4214	0.0041	0.09	0.03	0.3196	0.0205
-4.1038	0.0457	-4.3140	0.0057	0.11	0.03	0.3196	0.0205
-4.3073	0.0686	-4.7453	0.0087	0.04	0.04	0.3196	0.0205
-4.2478	0.0366	-4.4403	0.0044	0.05	0.03	0.3196	0.0205
-3.0450	0.0284	-3.1751	0.0050	-0.03	0.03	0.3196	0.0205
-4.4323	0.0316	-4.6417	0.0071	0.10	0.03	0.3196	0.0205
-3.3478	0.0211	-3.4950	0.0053	0.04	0.03	0.3196	0.0205
-4.5308	0.0535	-4.6949	0.0077	-0.18	0.03	0.3196	0.0205
-4.0661	0.0380	-4.2252	0.0037	0.14	0.03	0.3196	0.0205
-3.8391	0.0293	-3.8073	0.0028	0.23	0.04	0.3196	0.0205
-3.4750	0.0156	-3.5095	0.0023	0.04	0.02	0.3196	0.0205
-4.4092	0.0380	-4.5979	0.0054	0.09	0.03	0.3196	0.0205
-4.3313	0.0511	-4.5869	0.0064	0.11	0.03	0.3196	0.0205
-3.9492	0.0338	-4.2236	0.0061	0.36	0.04	0.3196	0.0205
-4.5298	0.0422	-4.7706	0.0074	0.02	0.03	0.3196	0.0205
-3.2577	0.0333	-3.3687	0.0027	0.10	0.02	0.3196	0.0205
-3.2755	0.0418	-3.5252	0.0038	0.03	0.03	0.3196	0.0205
-3.6345	0.0480	-3.6093	0.0032	0.08	0.04	0.3196	0.0205
-3.3588	0.0620	-3.5622	0.0040	0.05	0.03	0.3196	0.0205
-4.4802	0.0630	-4.6217	0.0119	0.12	0.04	0.3196	0.0205
-3.8960	0.0312	-3.9184	0.0031	0.09	0.04	0.3196	0.0205
-4.3965	0.0409	-4.6504	0.0185	0.13	0.06	0.3196	0.0205
-3.9060	0.0278	-3.9398	0.0031	0.08	0.03	0.3196	0.0205
-4.3687	0.0490	-4.6217	0.0093	0.09	0.04	0.3196	0.0205

observations uncertainties

covariates uncertainties

observations (dependent variable)

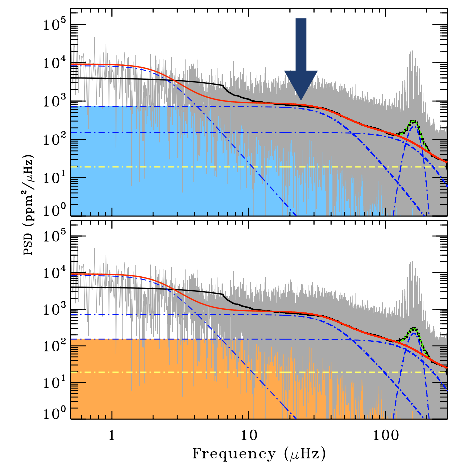
covariates (independent variable)

DIAMONDS™ BAYESIAN SOFTWARE

All the variables (y,x_i) have to be expressed as natural logarithms of an adimensional argument. This is the necessary condition to have a total uncertainty that has an analytical expression.

YOUR MODEL
MULTI-LINEAR FIT

$$\left(\frac{b_{\mathrm{meso}}}{b_{\mathrm{meso},\odot}}\right) = \beta \left(\frac{\nu_{\mathrm{max}}}{\nu_{\mathrm{max},\odot}}\right)^{s} \left(\frac{M}{M_{\odot}}\right)^{t} \mathrm{e}^{u[\mathrm{Fe/H}]}$$



The scaling relation for the characteristic frequency of the meso-granulation signal has been calibrated by Corsaro et al. 2017b using a Bayesian inference and model comparison process.

EXECUTE MULTI-LINEAR FIT

MULTI-LINEAR FIT

- E.g. perform the multi-linear fit in 2D, meaning that you have only one set of covariates
- 2 free parameters: 1 slope (nuMax) + 1 offset
- First make sure you are in the directory
  cd ~/Diamonds/tutorials/multilinear_fit/
- To run this tutorial via terminal type
  ./MultiLinearFit Corsaro17_data_M1.txt prior_hyperParameters_M1.txt
- At the end of the computation, annotate the total evidence collected, and the median estimates of the free parameters

RESULTS MULTI-LINEAR FIT

MULTI-LINEAR FIT

	s	t	u	lnβ
$\mathcal{M}_{b,1}$	$0.954^{+0.007}_{-0.008}$	–	–	$0.03^{+0.03}_{-0.03}$
$\mathcal{M}_{b,2}$	$0.917^{+0.008}_{-0.009}$	$0.20^{+0.02}_{-0.02}$	–	$-0.19^{+0.03}_{-0.03}$
$\mathcal{M}_{b,3}$	$0.889^{+0.009}_{-0.009}$	–	$-0.52^{+0.03}_{-0.03}$	$-0.16^{+0.03}_{-0.03}$
$\mathcal{M}_{b,4}$	$0.898^{+0.012}_{-0.014}$	$-0.38^{+0.06}_{-0.06}$	$-1.15^{+0.12}_{-0.10}$	$0.10^{+0.06}_{-0.07}$

M1 (nuMax) LN(Bayesian Evidence): -237.298

M2 (nuMax+Mass) LN(Bayesian Evidence): -129.275

M3 (nuMax+FeH) LN(Bayesian Evidence): -45.368

M4 (nuMax+Mass+FeH) LN(Bayesian Evidence): -12.596

What is the favored model for this application? Run all the tutorials for the multi-linear fit (four in total) and check the Bayesian evidence for each of the models you fitted. Then compute the Bayes' factors and identify the best model.

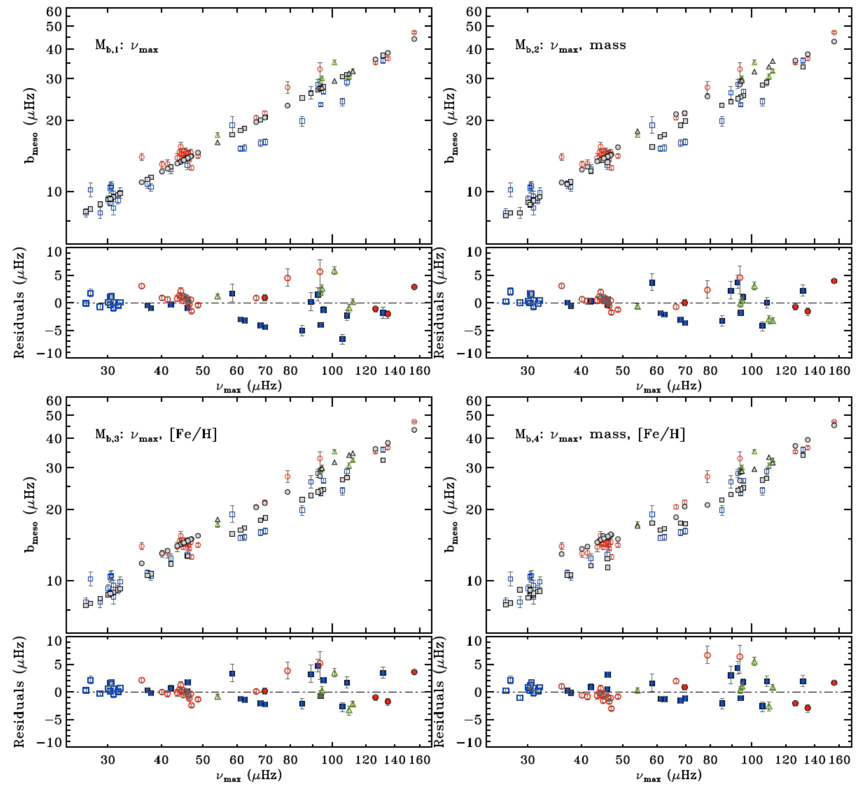
The values provided below are computed from multi-dimensional grid fits so can be considered accurate.

M1 LN(Bayesian Evidence): -237.29763497
M2 (Mass) LN(Bayesian Evidence): -129.27513818
M3 (FeH) LN(Bayesian Evidence): -45.36780603
M4 LN(Bayesian Evidence): -12.59555190

Astro Fit 2

RESULTS MULTI-LINEAR FIT

MULTI-LINEAR FIT



DIAMONDS™ BAYESIAN SOFTWARE

What is the favored model for this application?

M1 LN(Bayesian Evidence): -237.29763497
M2 (Mass) LN(Bayesian Evidence): -129.27513818
M3 (FeH) LN(Bayesian Evidence): -45.36780603
M4 LN(Bayesian Evidence): -12.59555190