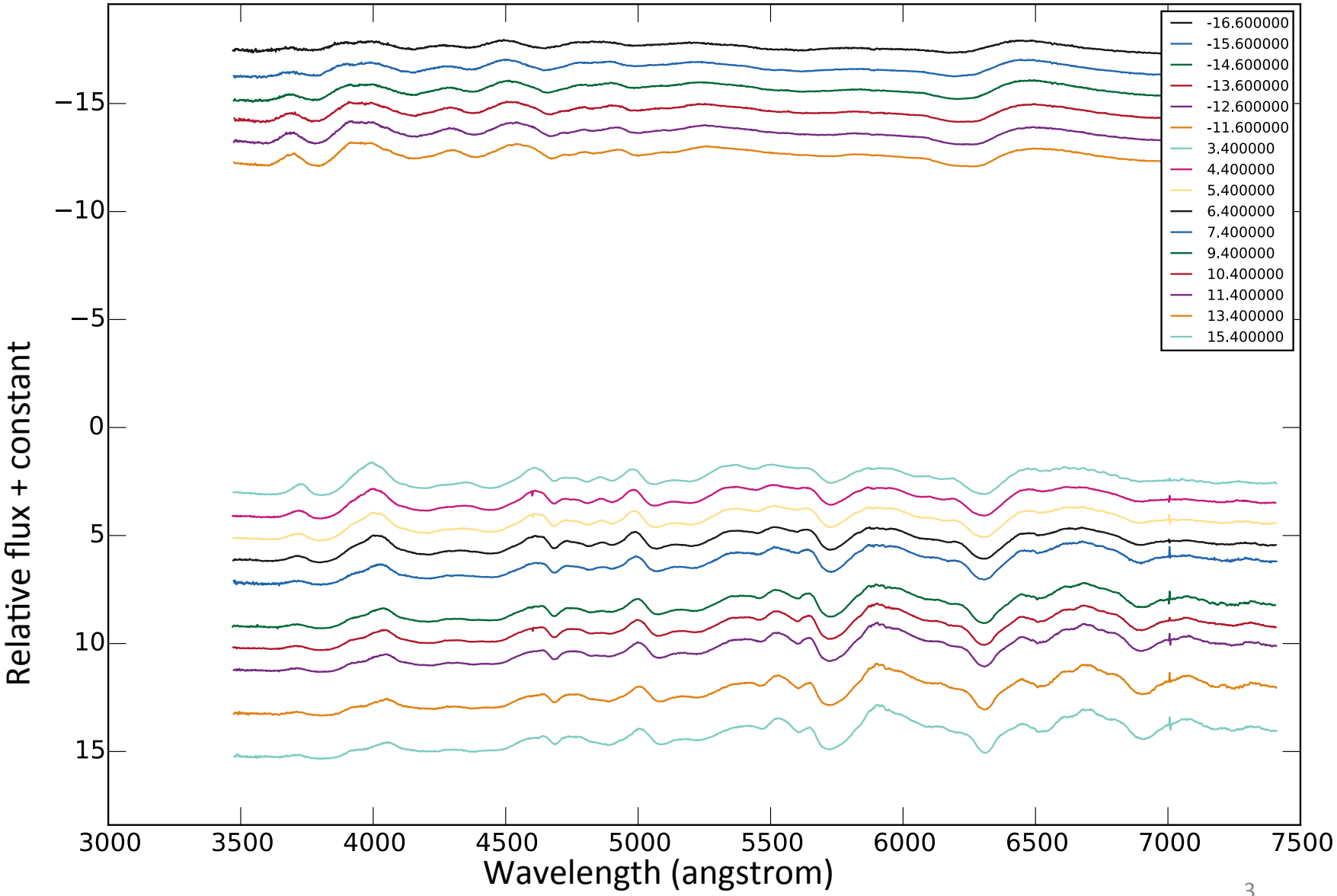
Aim

- To **interpolate** spectra of supernovae (SNe) in the time domain

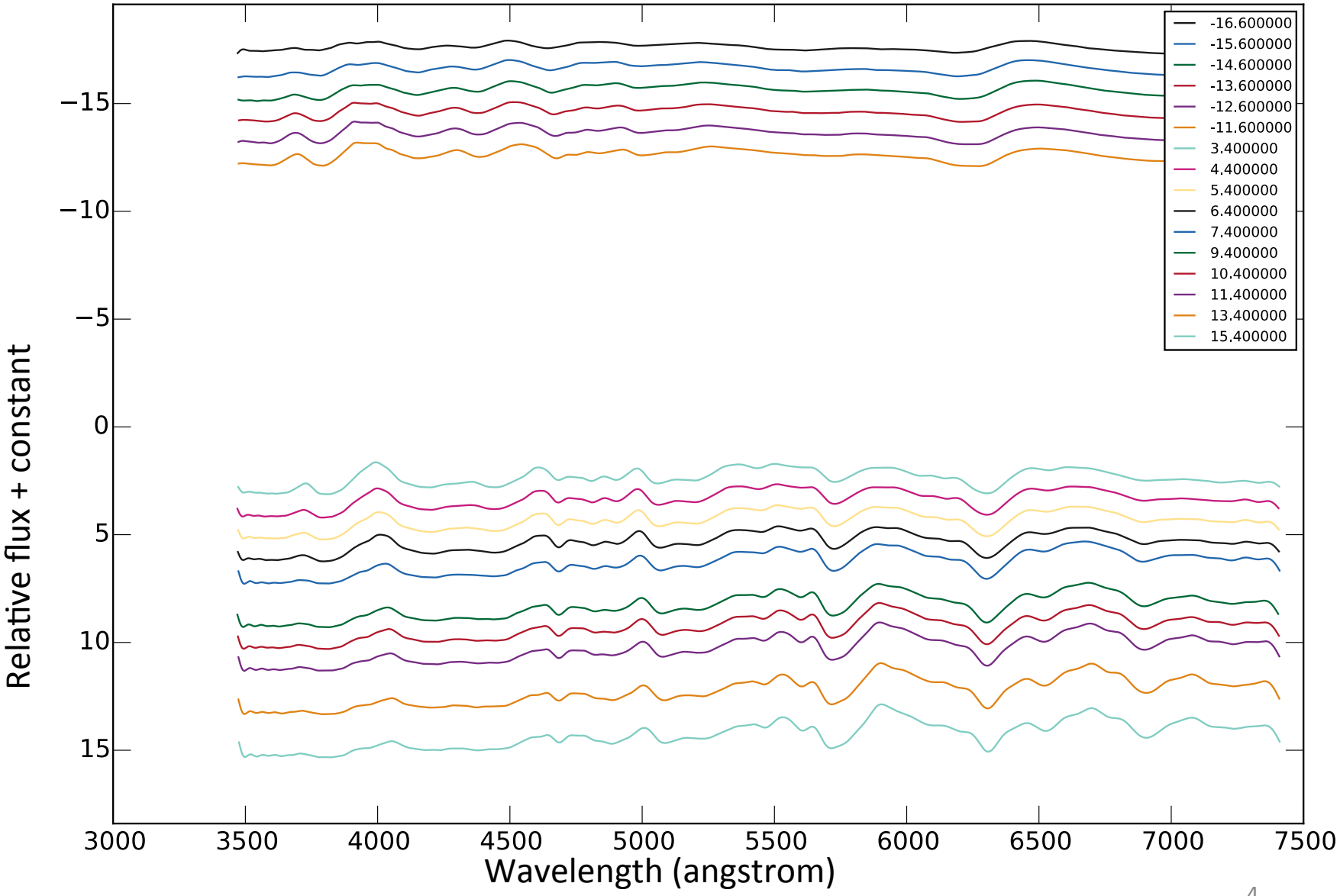
Motivation

- To construct **average spectra** and the corresponding standard deviations as a function of time
  - To characterize **spectral diversity** within each SN subtype
  - To be used in **photometric classification** for the future SN surveys
- To guide the **cadence** of taking SN spectra for the future SN surveys

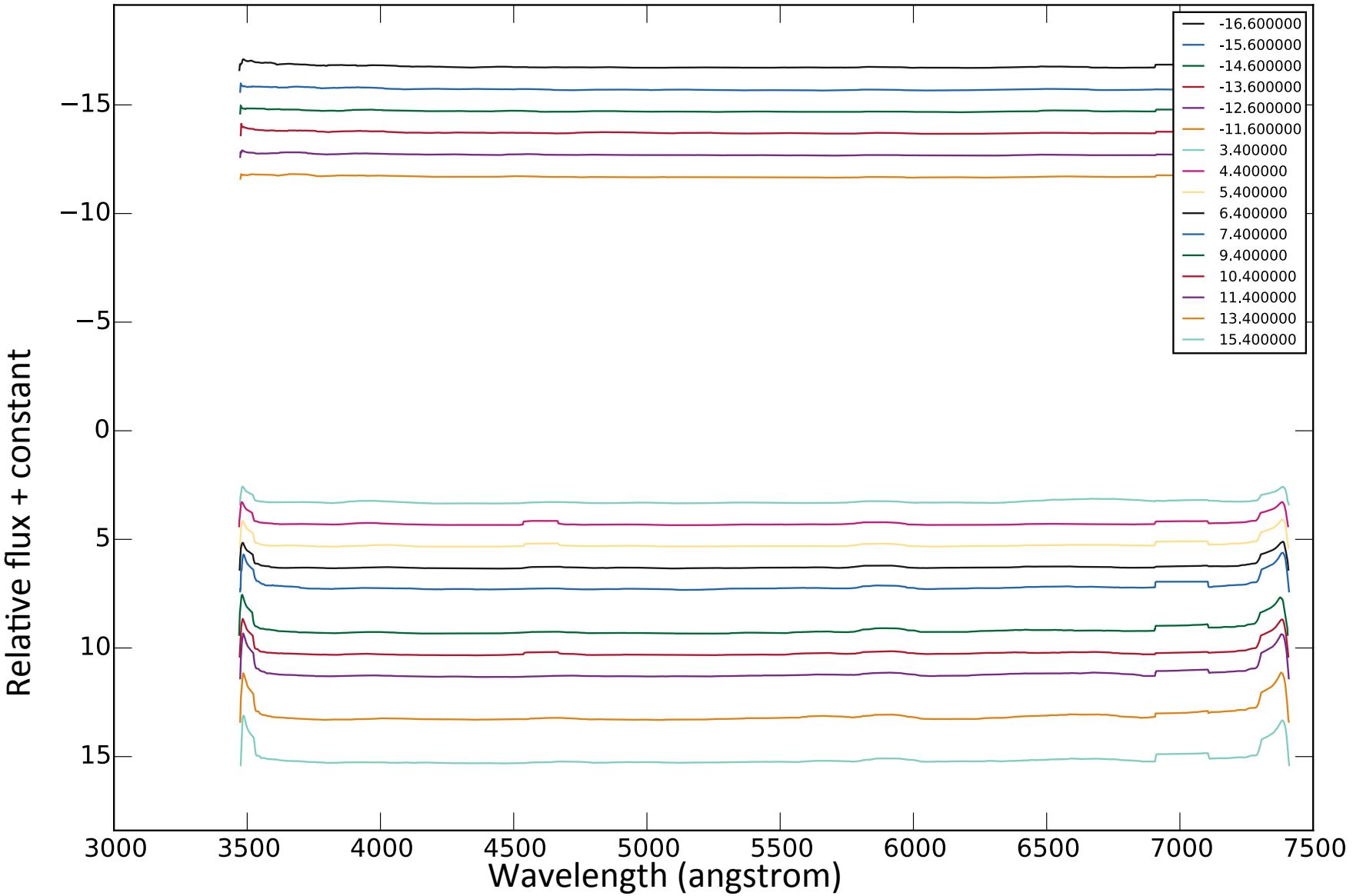
Data set – spectra of a SN



FFT-smoothed spectra



Error spectra



Method

- Gaussian Process Regression
  - A non-parametric method (compared with polynomial fit etc.)
  - Good at modeling noise
  - Gaussian random variables are convenient for many analytical manipulations
  - Kernel-based
  - A Bayesian regression algorithm
- *Sklearn* in Python
  - Three parameters
    - Regression model; Kernel; theta0 (correlation length)

Result – parameter range

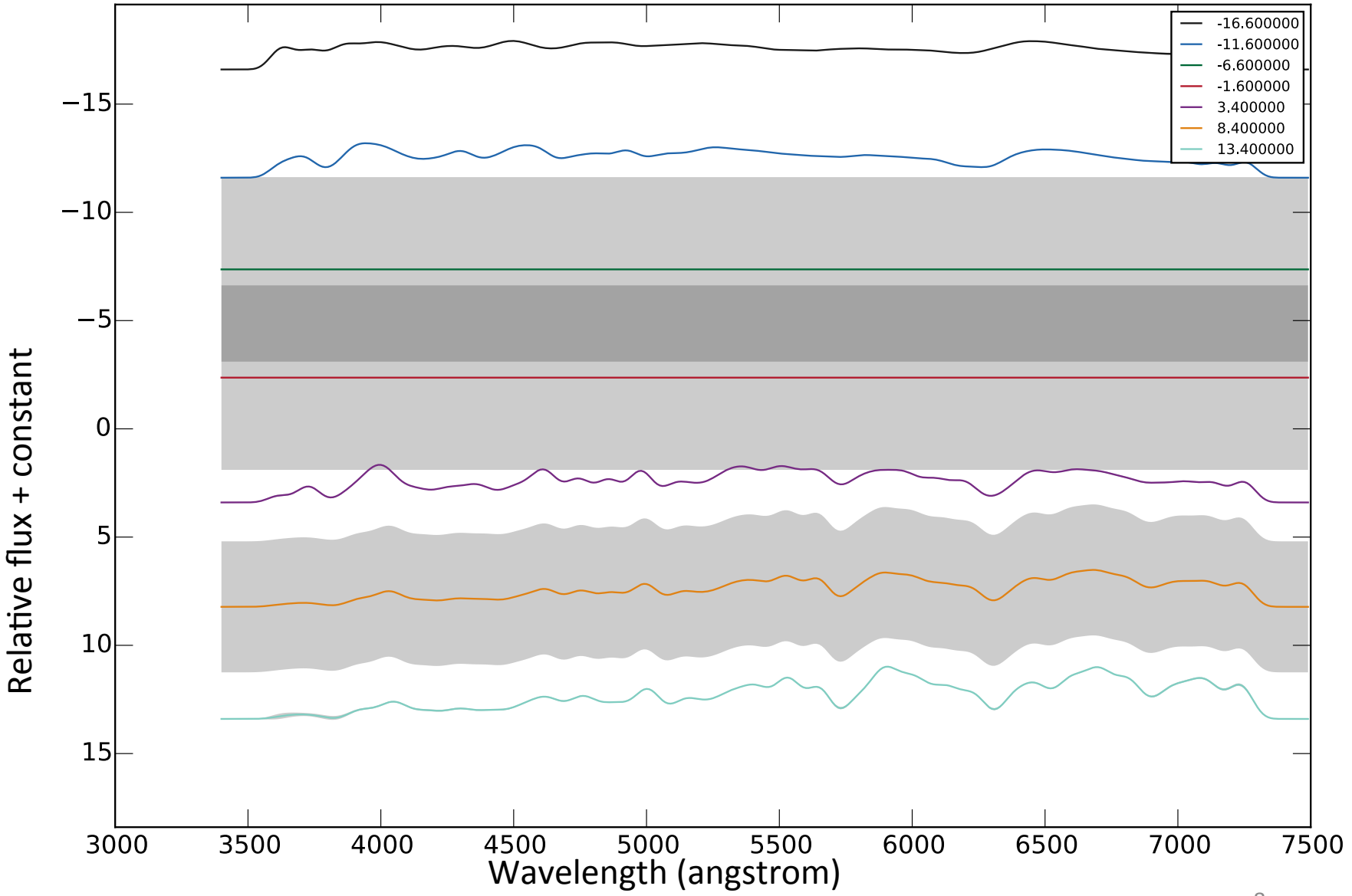
- Range of theta0

regression model / kernel	constant	linear	quadratic
Linear	0.1-10	0.1-10	0.1-10
Cubic	59.1750-59.1751	59.1750-59.1751	59.1750-59.1751
Absolute exponential	0.1-10	0.1-10	0.1-10
Squared exponential	100.80365872 - 100.80365873	100.80365872 - 100.80365873	100.80365872 - 100.80365873

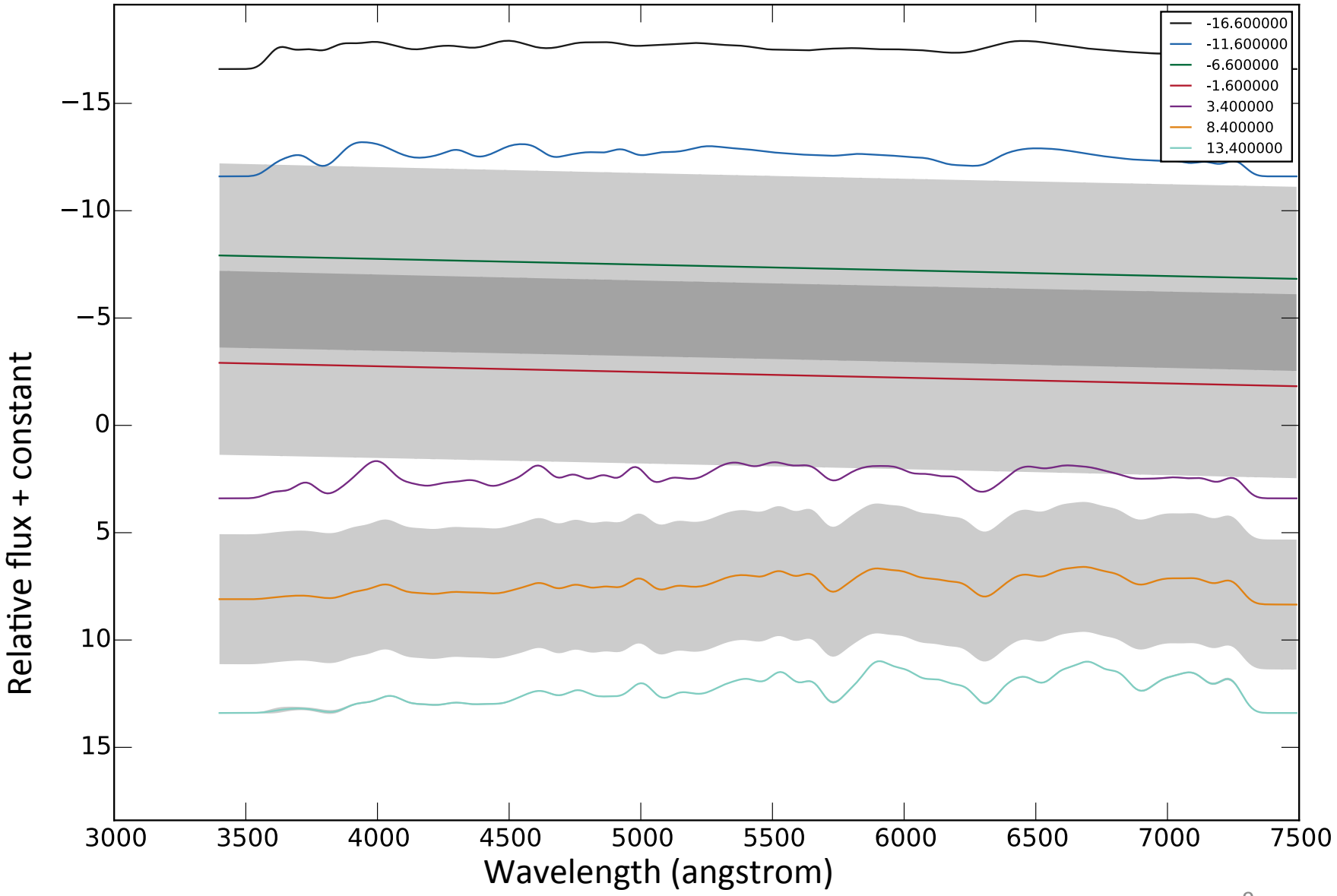
failed

failed

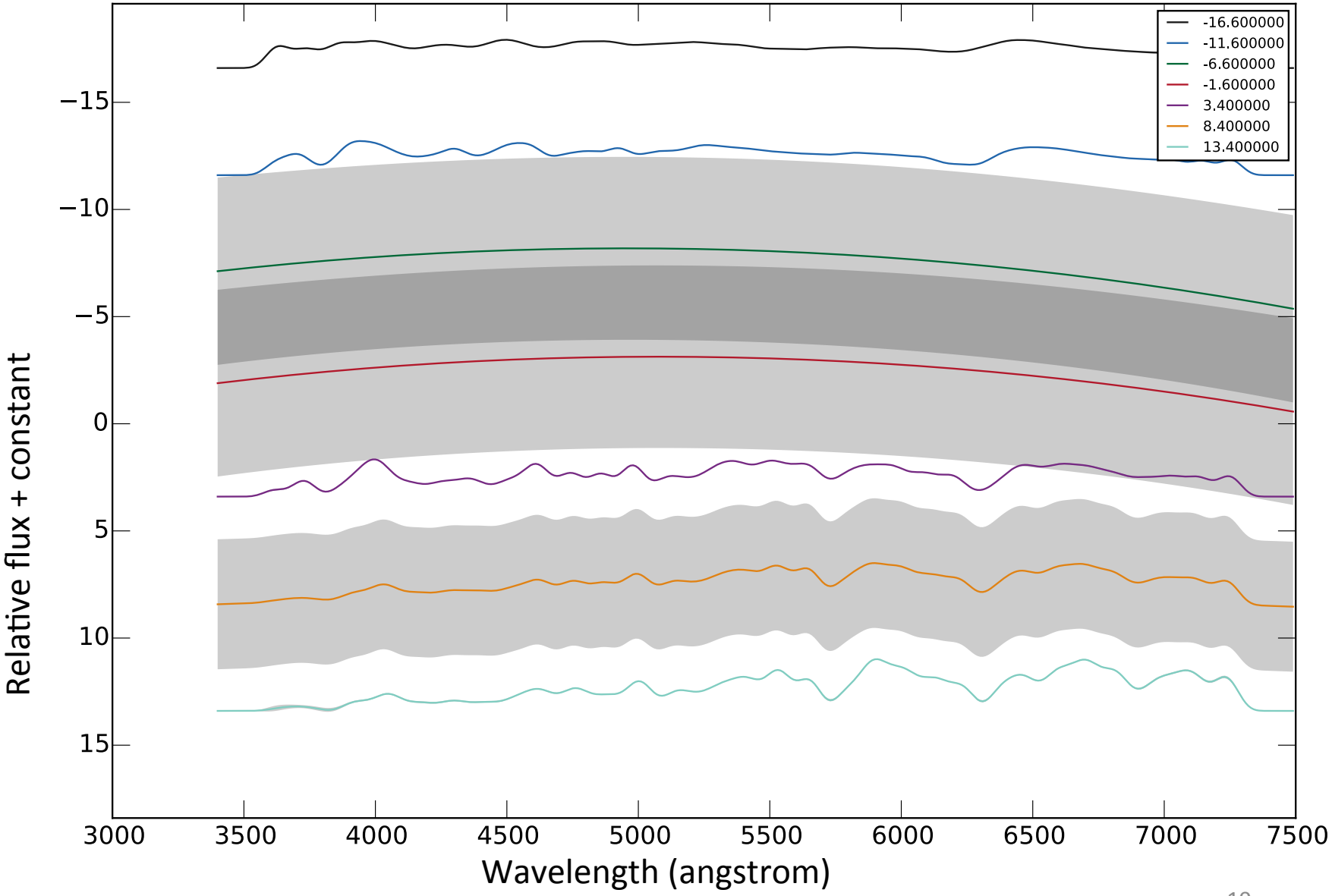
Examples of failed interpolation constant, SE, 100.80365873



Examples of failed interpolation
linear, SE, 100.80365873



Examples of failed interpolation
quadratic, SE, 100.80365873



Result – cross validation

- KF=10
  - Kernel: Absolute exponential

Regression model	theta0	R2 score
constant	0.1	0.9722+/-0.0503
constant	1	0.9720+/-0.0604
constant	10	0.6487+/-0.4452
linear	0.1	0.9722+/-0.0503
linear	1	0.9717+/-0.0617
linear	10	0.6173+/-0.4856
quadratic	0.1	0.9717+/-0.0526
quadratic	1	0.9741+/-0.0566
quadratic	10	0.8079+/-0.2175

Highest score

Result – cross validation

- KF=10
  - Kernel: Linear

Regression model	theta0	R2 score
constant	0.1	0.9726+/-0.0474
constant	1	0.9493+/-0.1171
linear	0.1	0.9727+/-0.0473
linear	1	0.9471+/-0.1466
quadratic	0.1	0.9719+/-0.0501
quadratic	1	0.9408+/-0.1638

Highest score

Result – cross validation

- Leave data at one phase out
  - Kernel: Absolute exponential

Regression model	theta0	R2 score
constant	0.1	0.9829+/-0.0419
constant	1	0.9828+/-0.0361
linear	0.1	0.9829+/-0.0419
linear	1	0.9826+/-0.0378
quadratic	0.1	0.9826+/-0.0442
quadratic	1	0.9820+/-0.0471

Highest score

Highest score

Result – cross validation

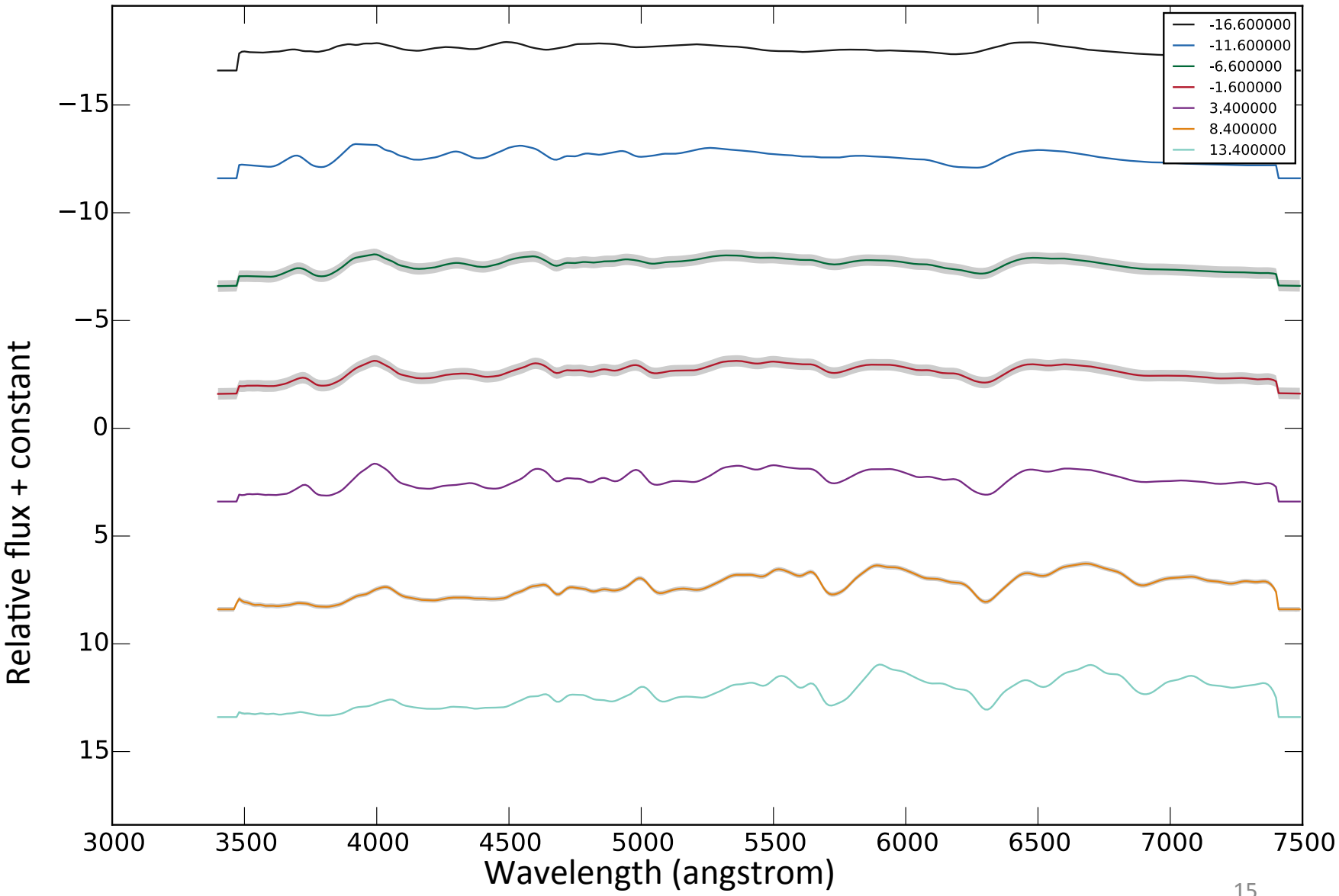
- Leave data at one phase out
  - Kernel: Linear

Regression model	theta0	R2 score
constant	0.1	0.9831+/-0.0412
constant	1	0.9764+/-0.0538
constant	10	0.2791+/-0.3153
linear	0.1	0.9831+/-0.0412
linear	1	0.9757+/-0.0623
linear	10	0.2275+/-0.6156
quadratic	0.1	0.9828+/-0.0437
quadratic	1	0.9688+/-0.0912
quadratic	10	0.6369+/-0.1760

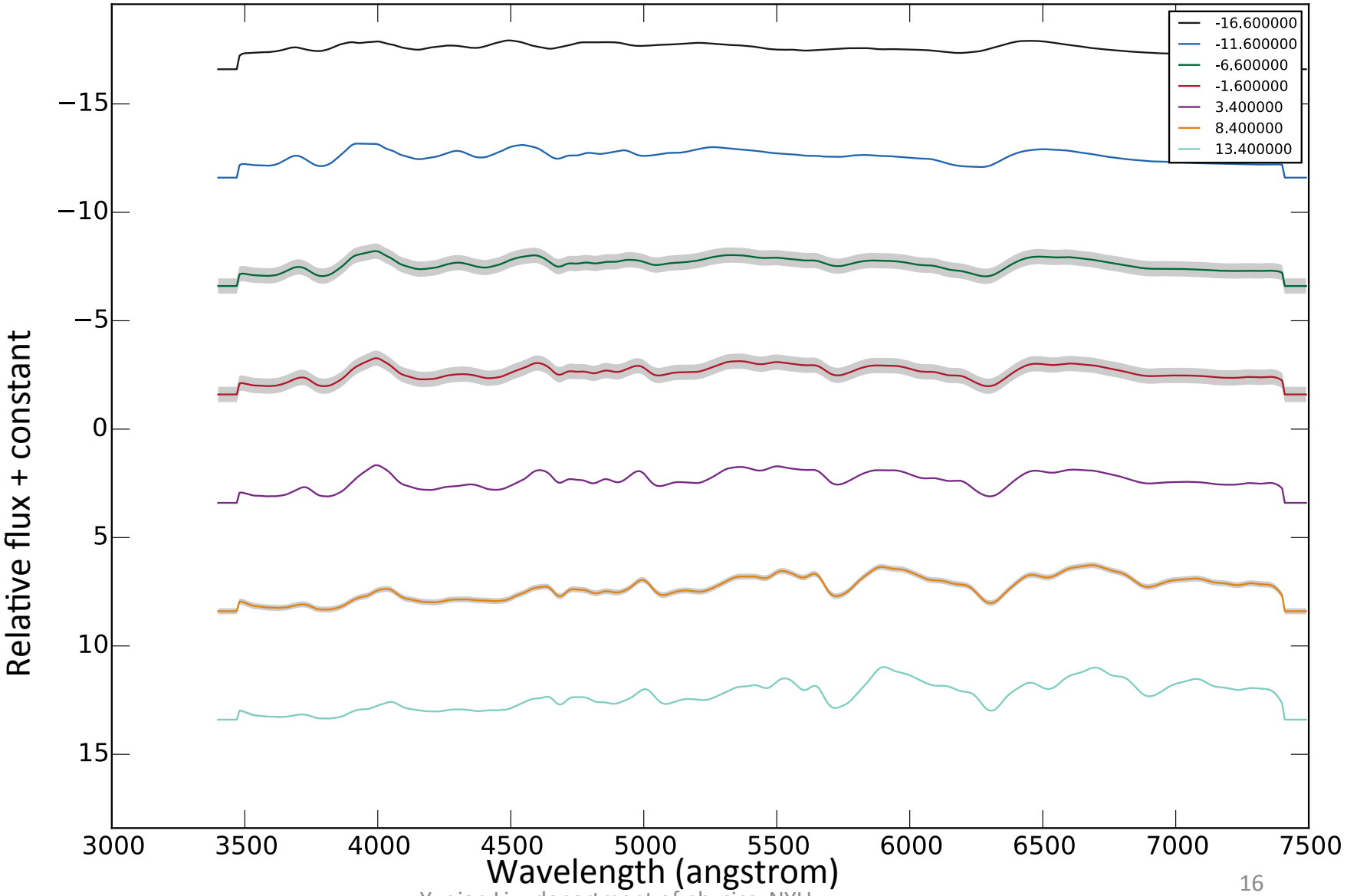
Highest score

Highest score

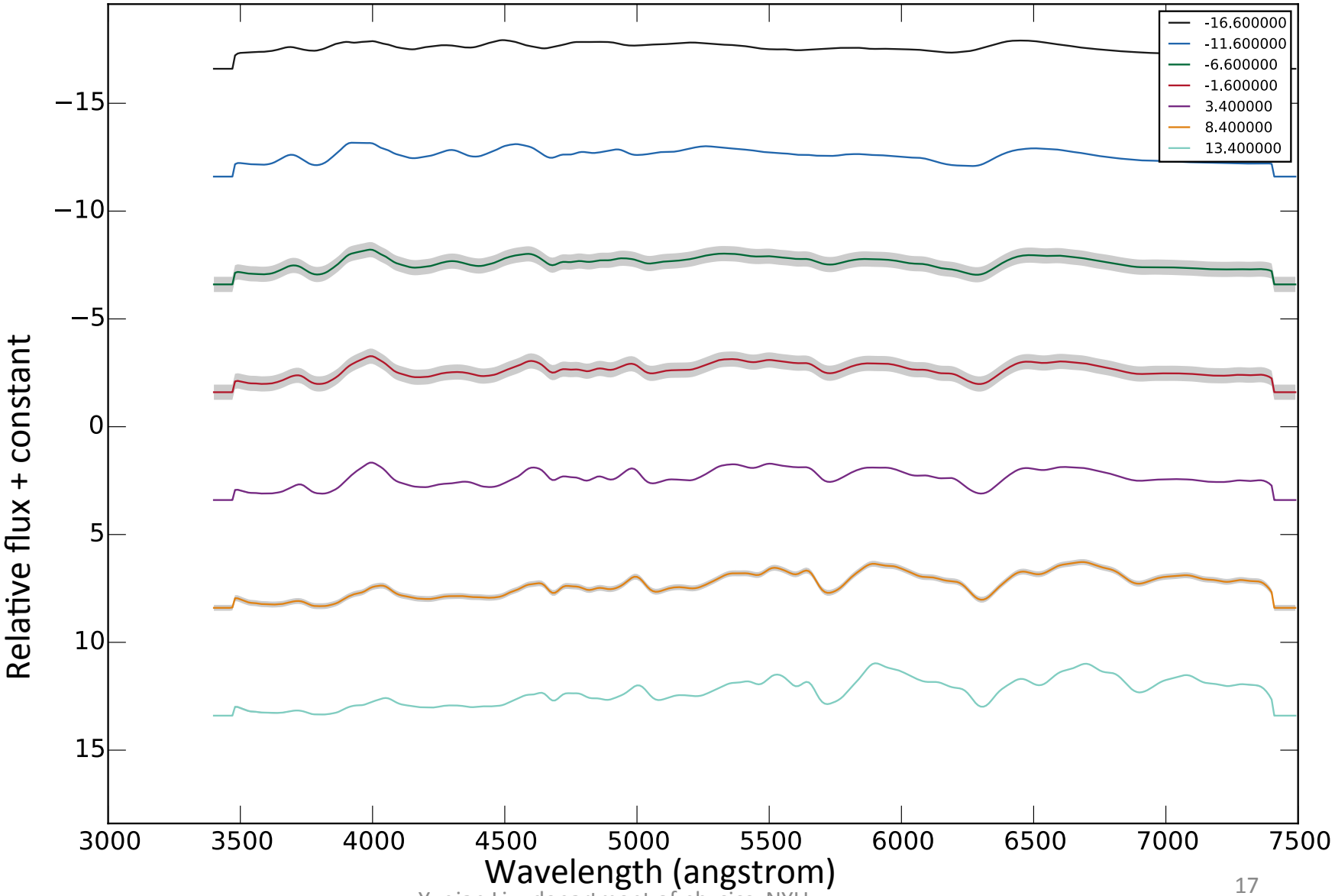
Examples of good interpolation
quadratic, AE, 1



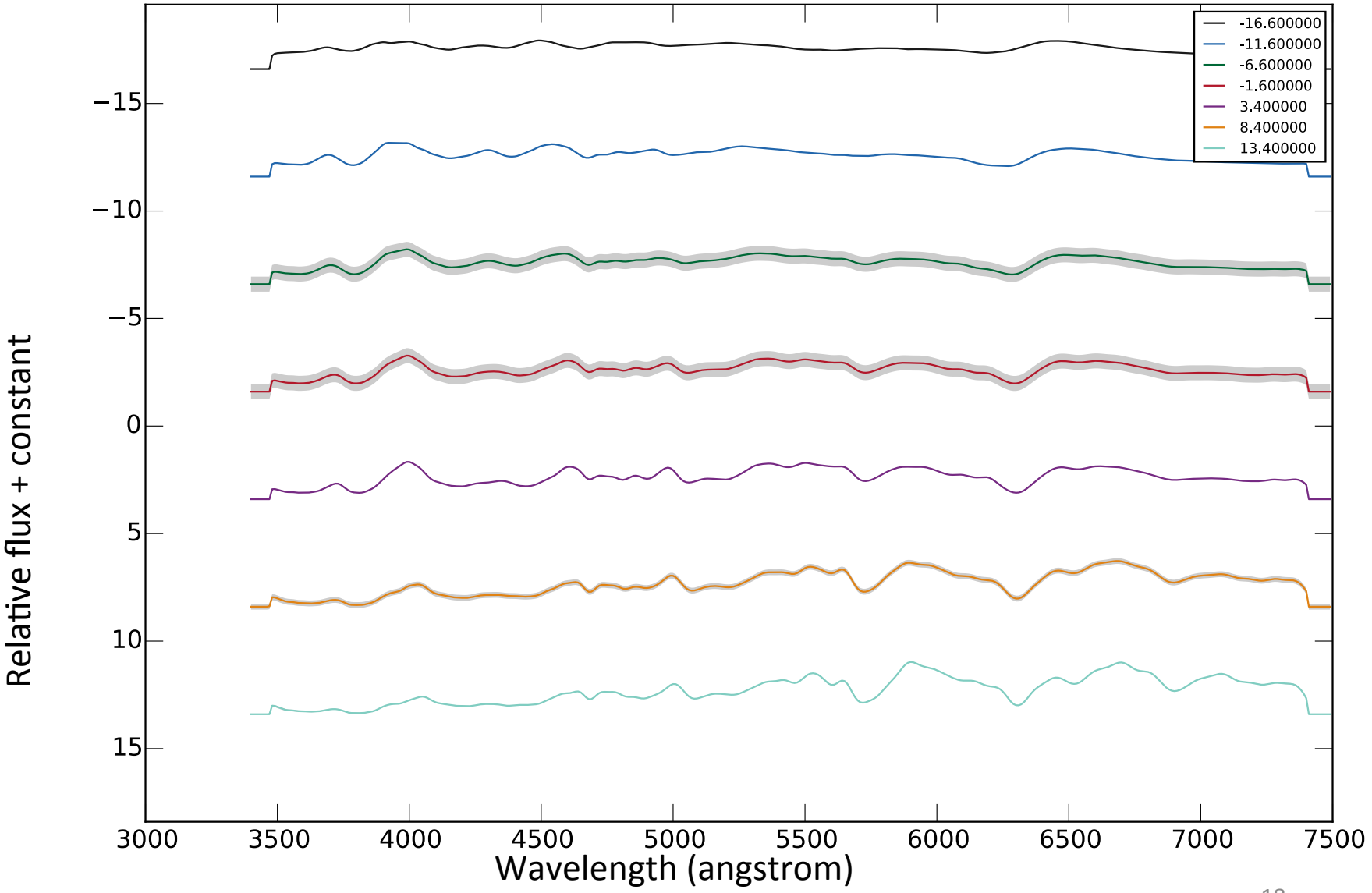
Examples of good interpolation constant, AE, 0.1



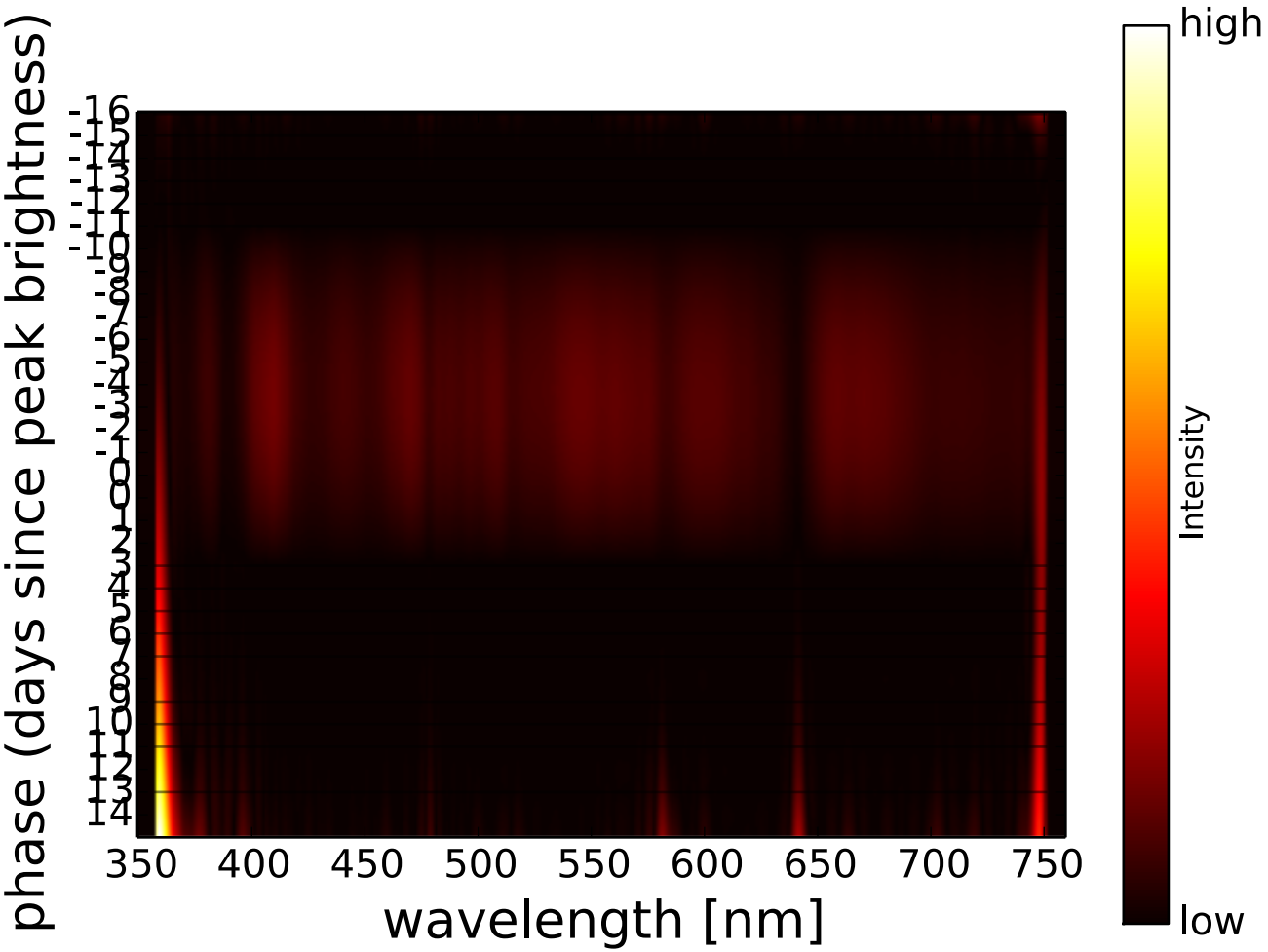
Examples of good interpolation
linear, AE, 0.1



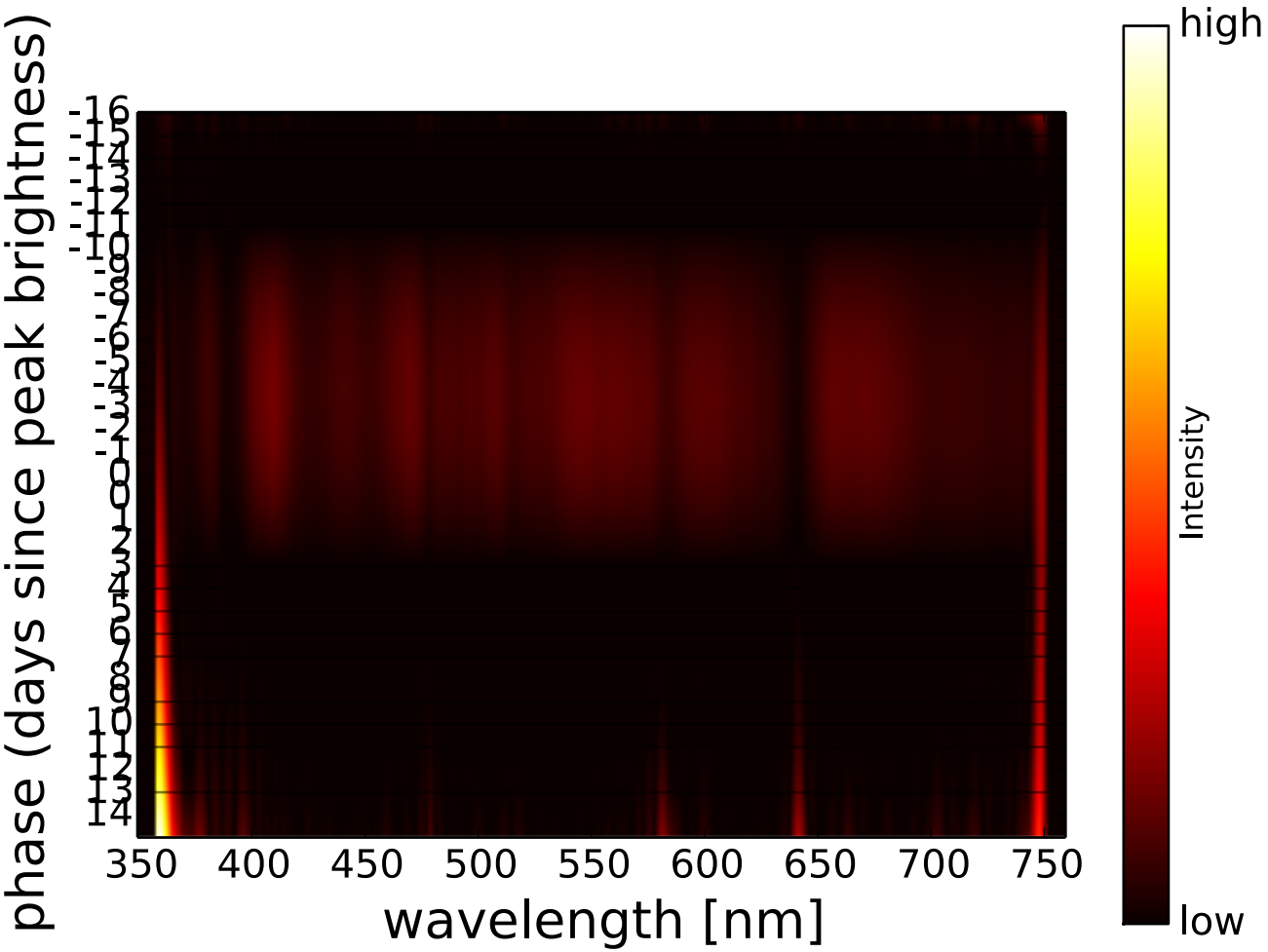
Examples of good interpolation
linear, Linear, 0.1



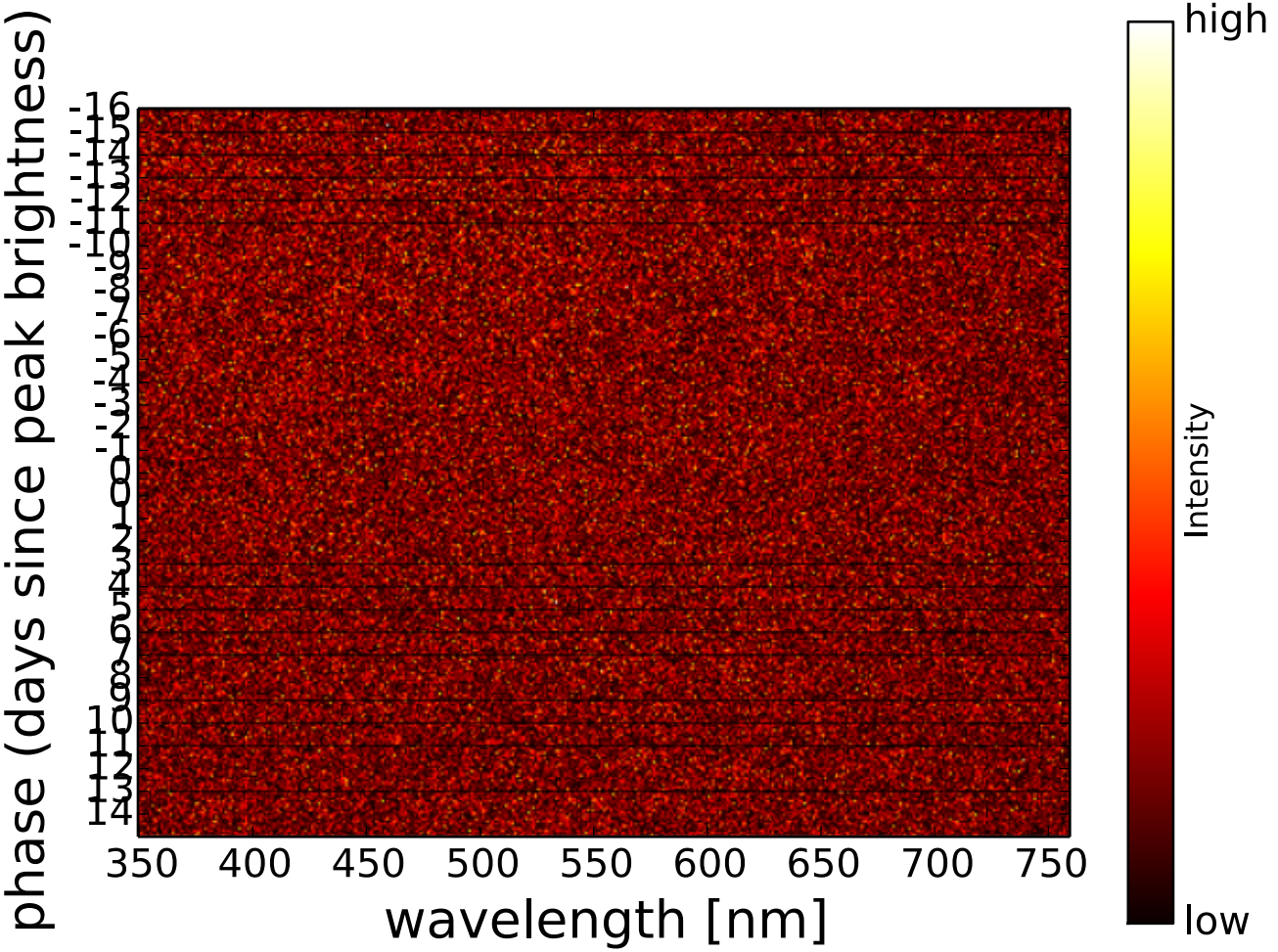
Difference plot
linear, AE, 0.1 vs. linear, Linear, 0.1



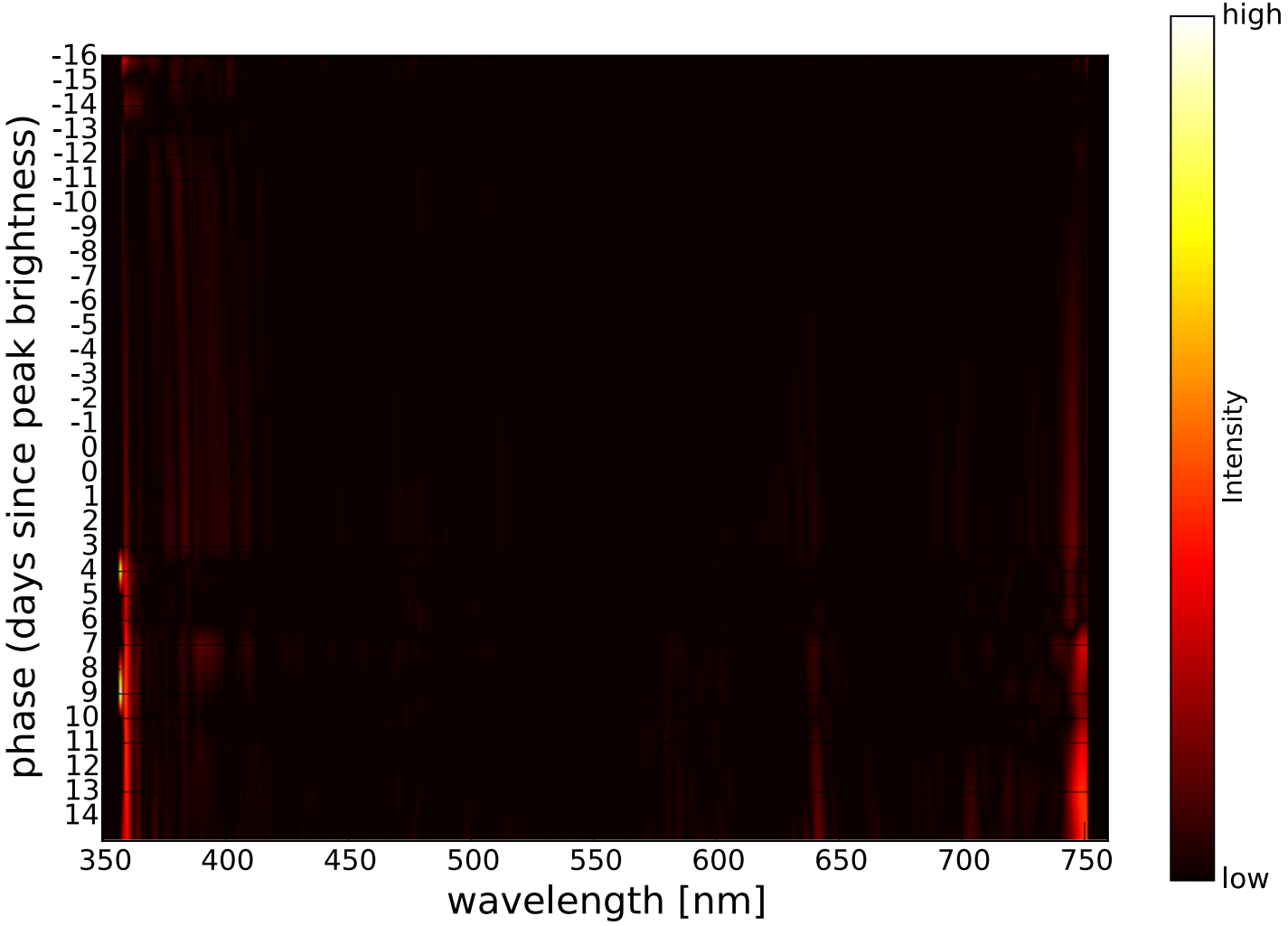
Difference plot
constant AE 0.1 vs. constant Linear 0.1



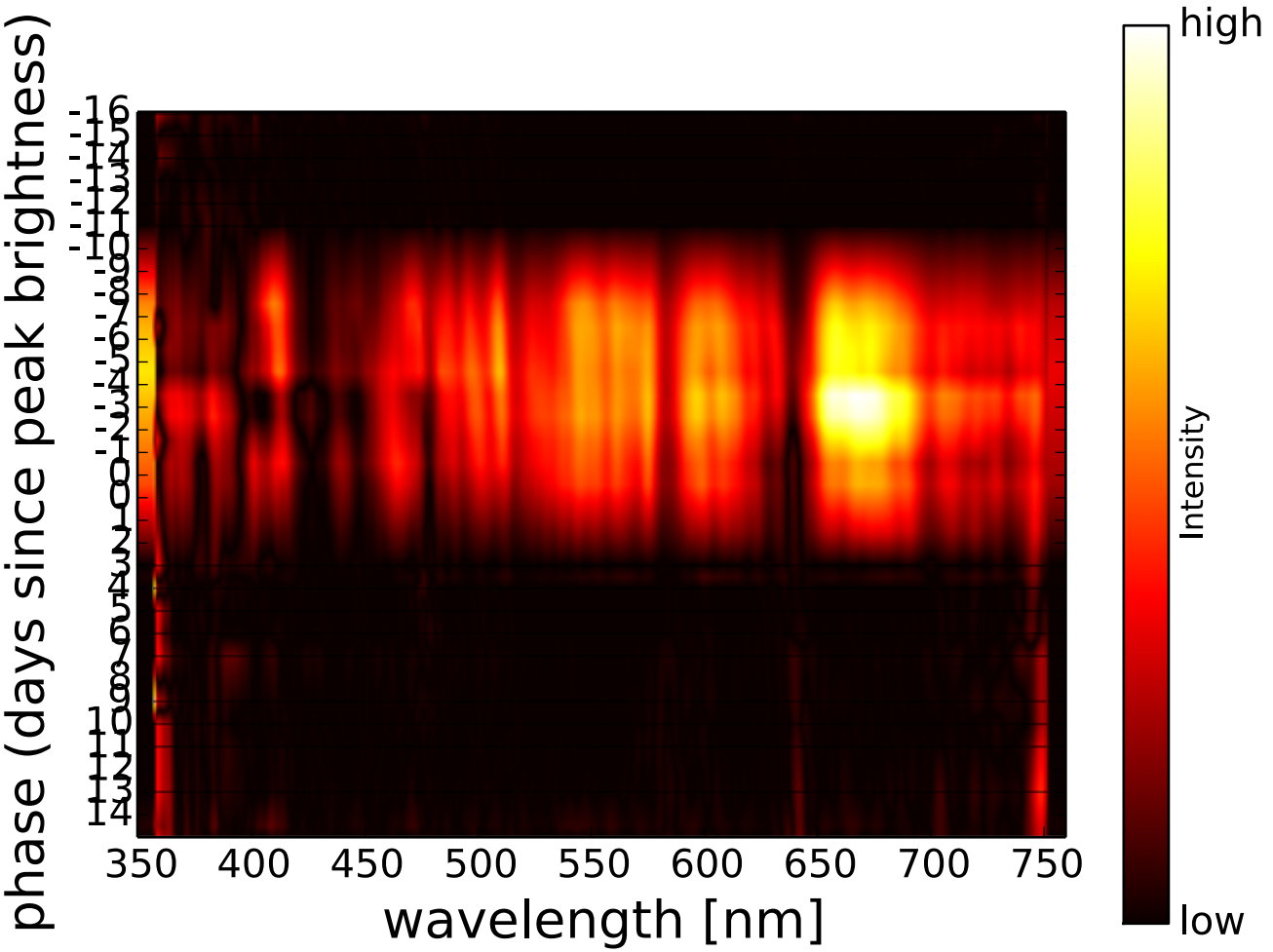
Difference plot
linear, Linear, 0.1 vs. constant, Linear, 0.1



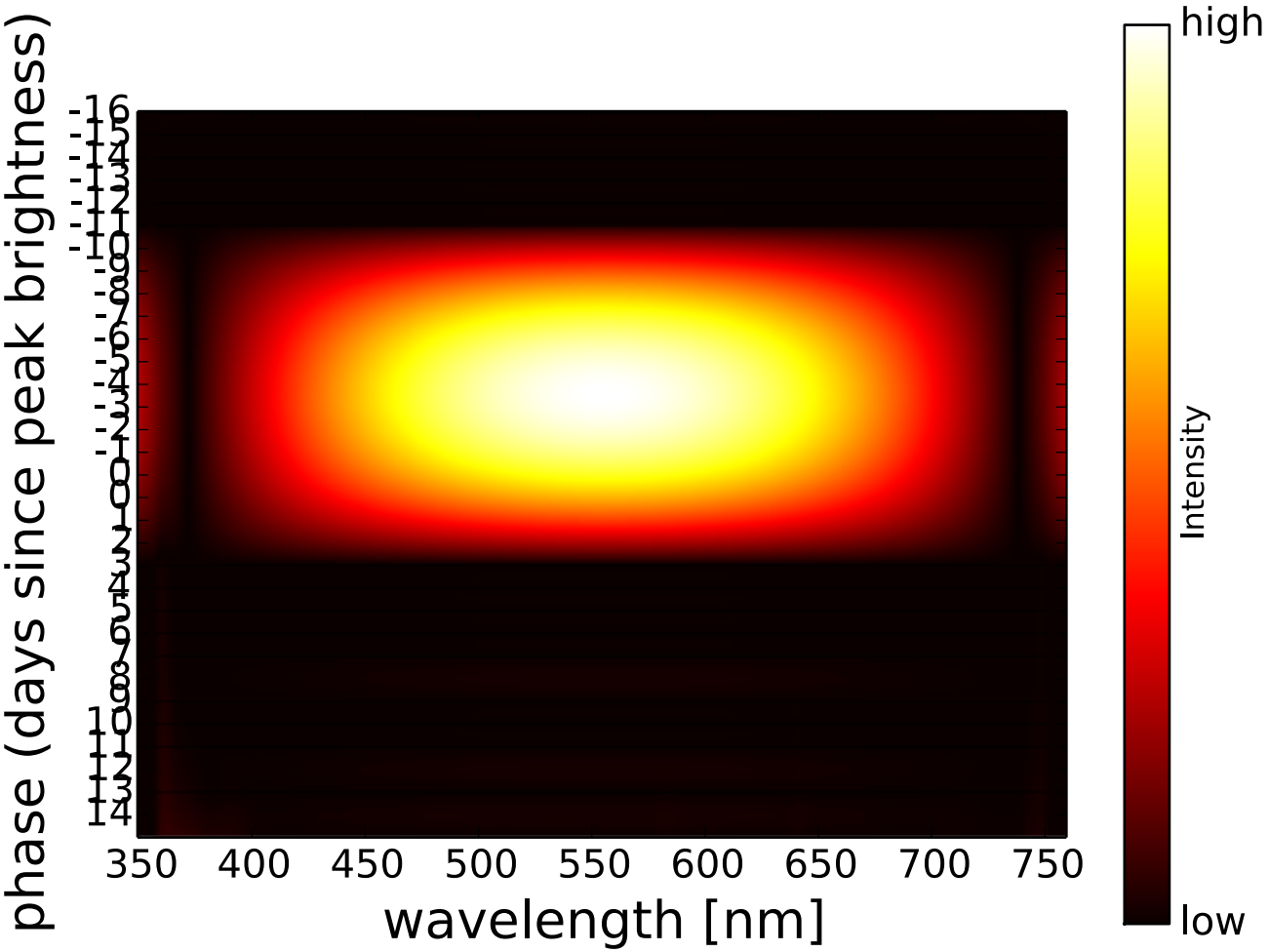
Difference plot
simple vs. constant, Linear, 0.1



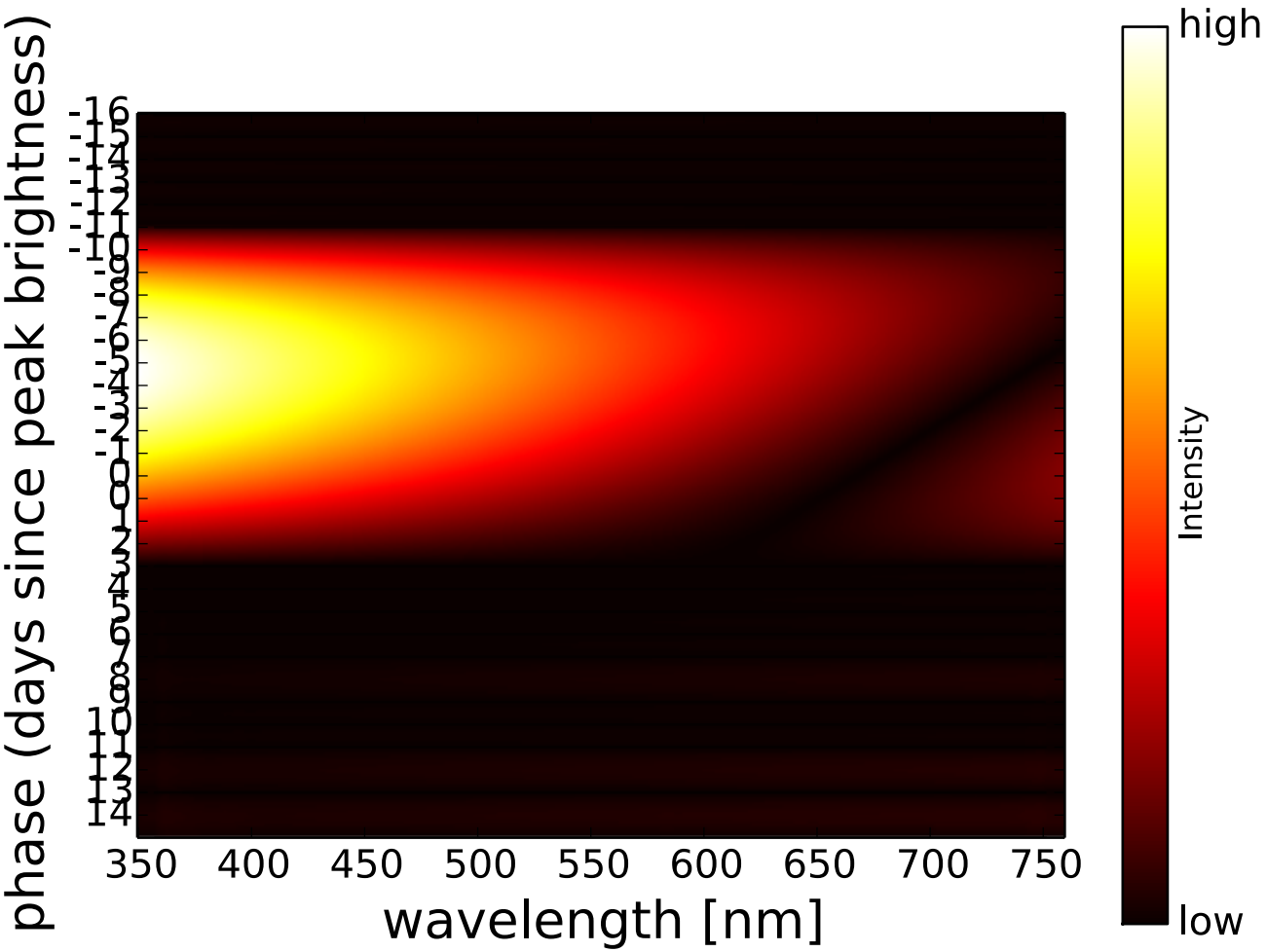
Difference plot
linear, Linear, 0.1 vs. linear, Linear, 1



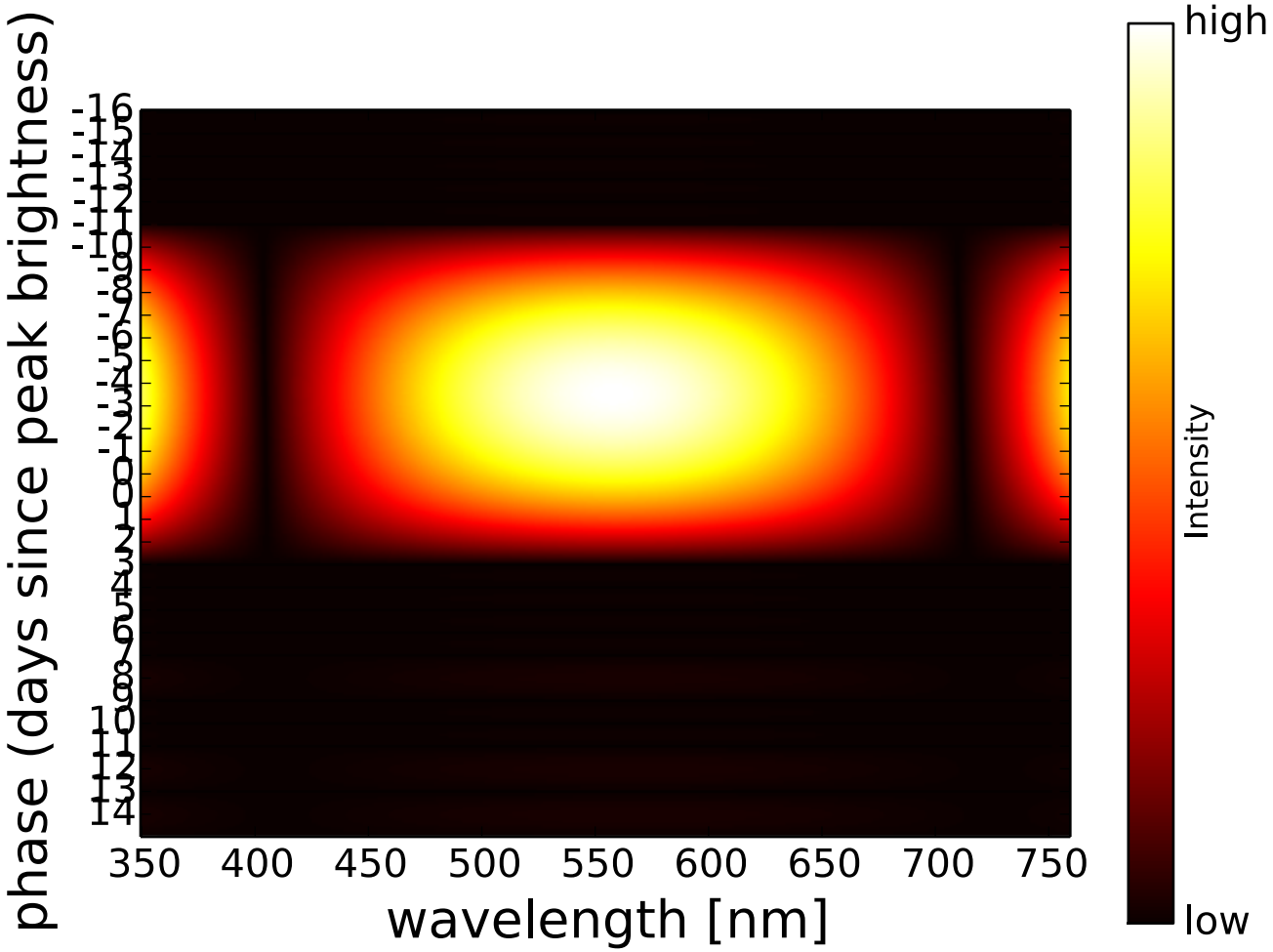
Difference plot
quadratic, AE, 0.1 vs. constant, AE, 0.1



Difference plot
linear, AE, 0.1 vs. constant, AE, 0.1



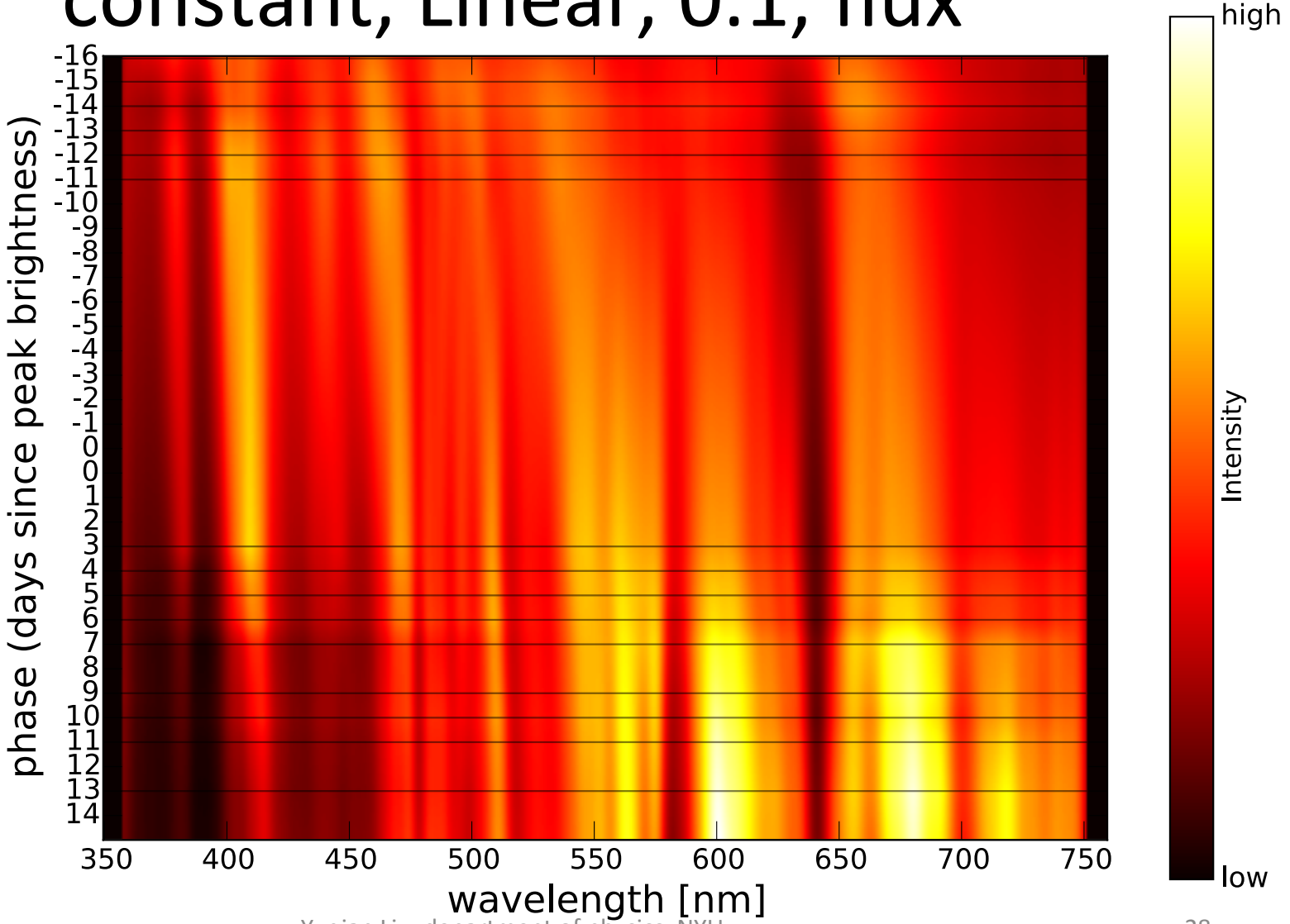
Difference plot
quadratic, AE, 1 vs. constant, AE, 1



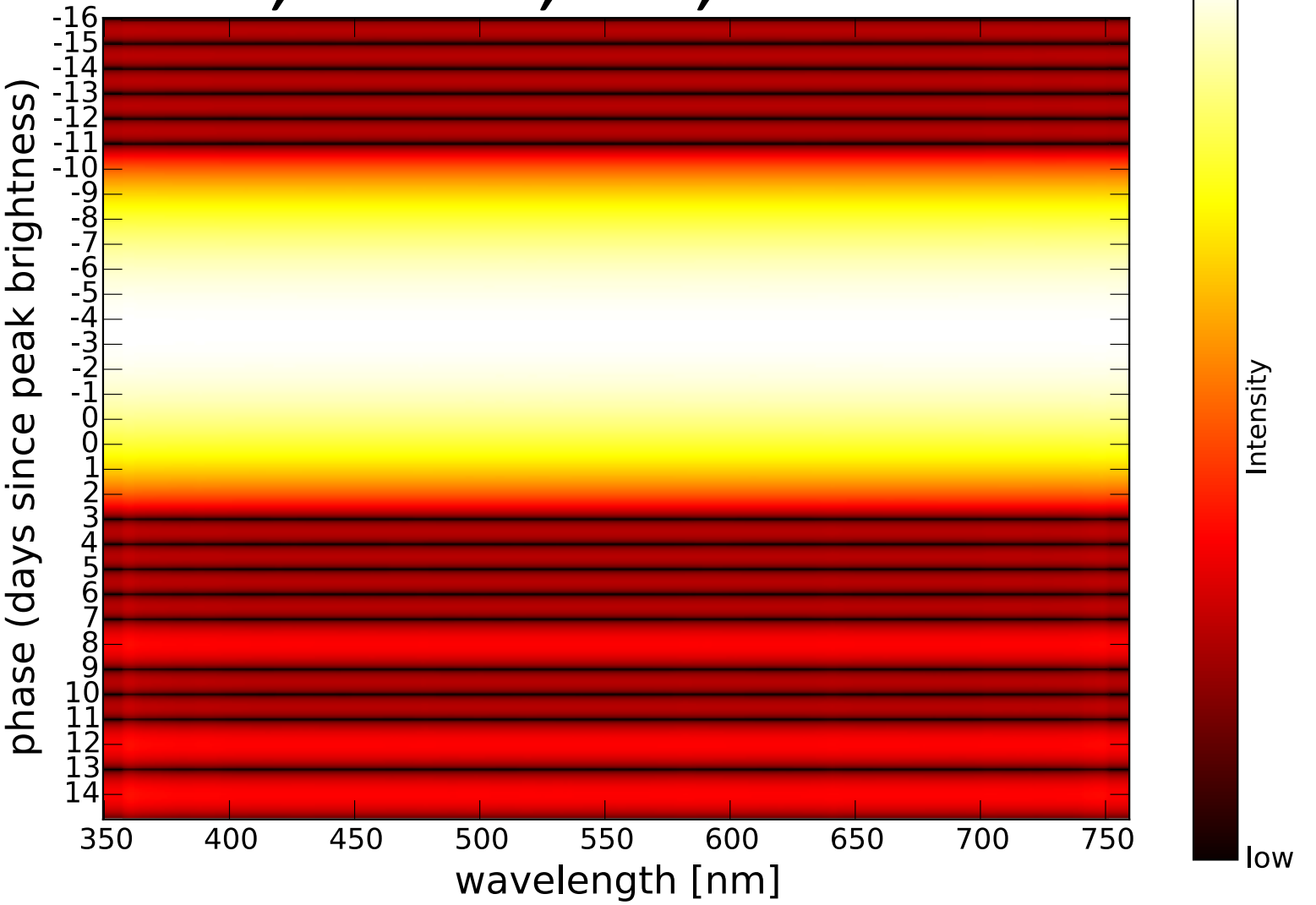
conclusion

- Constant/Linear as regression models, linear kernel, 0.1 as the correlation length work best in this case.

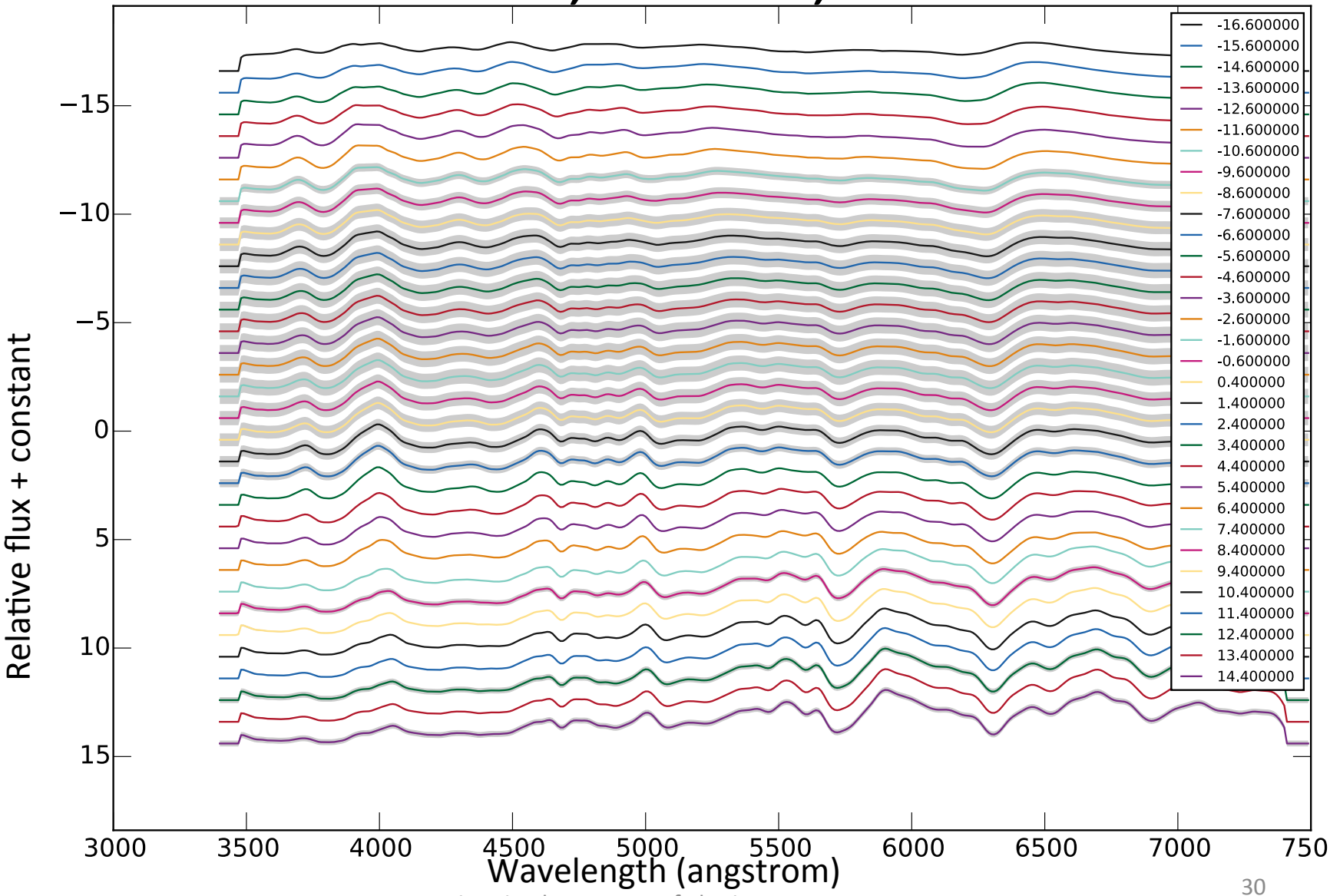
Best interpolation constant, Linear, 0.1, flux



Best interpolation constant, Linear, 0.1, flux error



Best interpolation constant, Linear, 0.1



Next steps

- Read chapter 4 & 5 of **Gaussian Process for Machine Learning** (Rasmussen & Williams 2006) to learn more about determining the covariance function and hyper-parameters
- Apply GP to other SNe, where the regression model, kernel, and theta0 are determined by cross validation
- Based on spectra sampling of individual SNe and the corresponding highest CV score, try to find a better cadence of taking SN spectra for the future SN surveys