



Data-analysis and Inverse Methods in Astronomy

MCMC in practice – tools, convergence and results

Antti Penttilä

Homepage <http://wiki.helsinki.fi/display/53834>



MCMC in practice

I will try to demonstrate the typical case of an MCMC ‘workflow’ with an example problem. Special emphasis is on

- Model formulation
- Choosing the a priori
- Graphical tools in OpenBUGS
- The convergence of a MCMC sampler
- The results – how can we extract useful information from the a posteriori samples



Asteroid polarization

Our sample problem deals with the behavior of the linear polarization ratio of asteroids, and especially of three asteroid 'classes', C-, S- and M-types.

Degree of linear polarization (hereafter 'polarization', P), is computed from the reflected flux *in the scattering plane* (defined by Sun-asteroid-Earth) Z_{\parallel} , and the flux *parallel to scattering plane* Z_{\perp} by

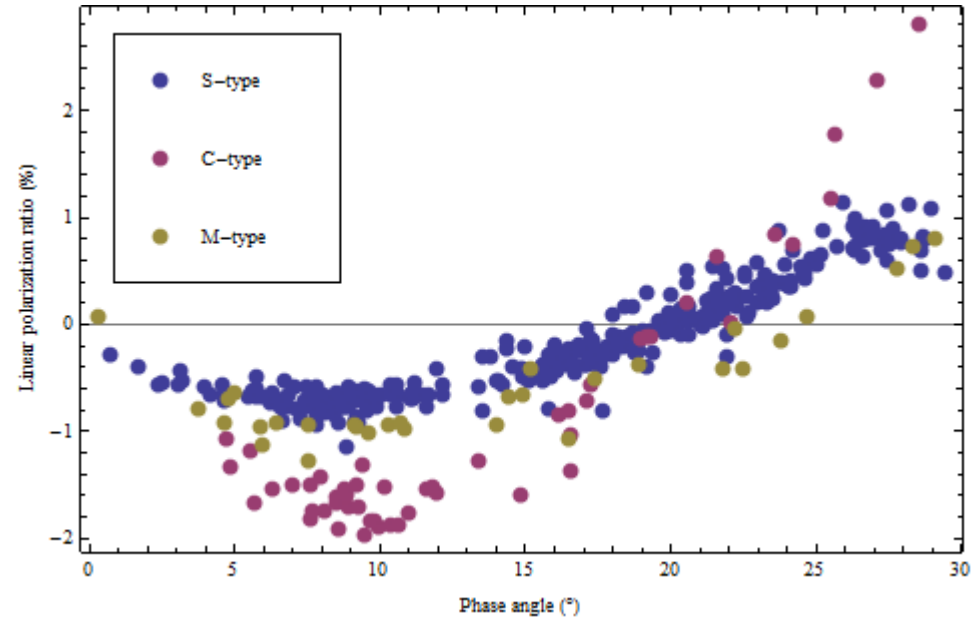
$$P(\alpha) = \frac{Z_{\perp} - Z_{\parallel}}{Z_{\perp} + Z_{\parallel}}$$

as a function of the phase angle α (angle between Sun-asteroid-Earth).



Asteroid 'classes' are based on the Tholen taxonomic system, which uses the broad band spectral properties and sometimes the asteroid albedo in classifying the object. The three largest classes are

- C-type, dark carbonaceous objects
- S-type, silicate-rich (i.e. stony) objects
- M-type, metallic objects



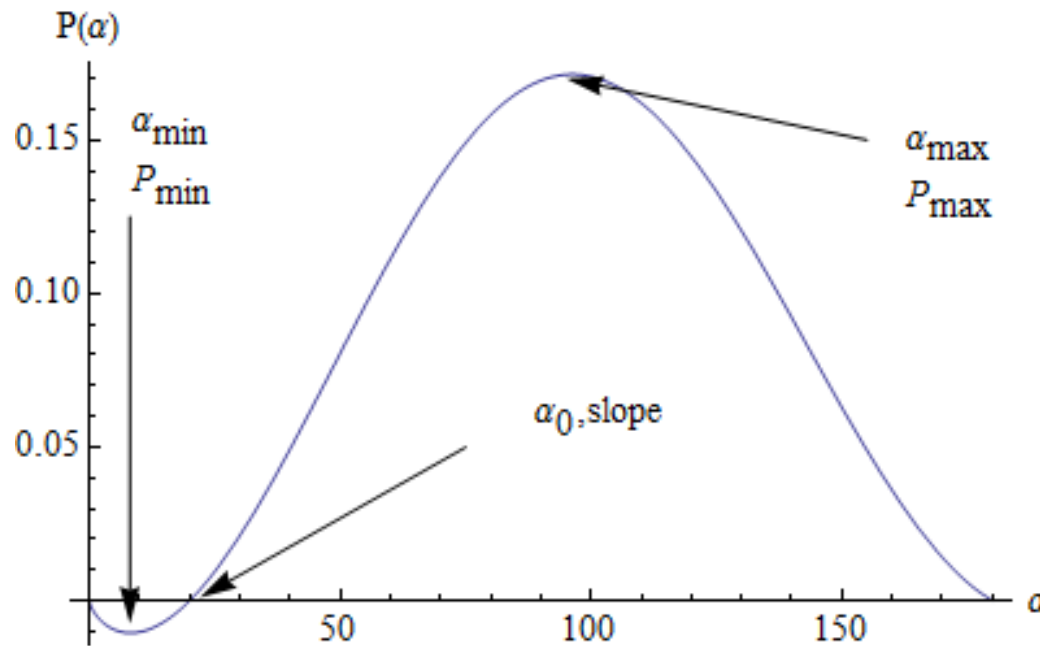
Our studies have shown that the polarization reflects the same differences between the classes as the spectra+albedo, and can be used to study the surface material properties.



Model for polarization

The so-called 'trigonometric function', suggested by K. Lumme

$$P(\alpha) = b \sin(\alpha)^{c_1} \cos\left(\frac{1}{2}\alpha\right)^{c_2} \sin \alpha - \alpha_0$$



Study interactively in *Mathematica*...



Regression model

The parameter estimation for $P(\alpha)$ is a non-linear regression problem. When we assume independent, identically distributed errors for every observation i , our *probability model* is

$$y_i = P(\alpha_i; b, \alpha_0, c_1, c_2) + \epsilon_i \sim N(P(\alpha_i; b, \alpha_0, c_1, c_2), \sigma^2)$$

So, we know the likelihood of the model, $L(\alpha, b, \alpha_0, c_1, c_2, \sigma^2; y)$. For Bayesian analysis we just need to plug in the a priori for parameters, and we have the a posteriori

$$D(b, \alpha_0, c_1, c_2, \sigma^2 | y) \propto D_{pr}(b, \alpha_0, c_1, c_2, \sigma^2) L(b, \alpha_0, c_1, c_2, \sigma^2; y)$$



Bayesian regression tools

There are some software tools that can form the Bayesian samplers (i.e. MCMC chains) for the beforementioned a posteriori automatically. One of these is the OpenBUGS (BUGS – Bayesian inference Using Gibbs Sampler).

All we have to do is to define the a priori distributions and their *hyperparameters* (parameters of the a priori distributions for actual parameters).

$$D(b, \alpha_0, c_1, c_2, \sigma^2 | y) \propto D_{pr}(b, \alpha_0, c_1, c_2, \sigma^2) L(b, \alpha_0, c_1, c_2, \sigma^2; y)$$

Define a priori for these parameters

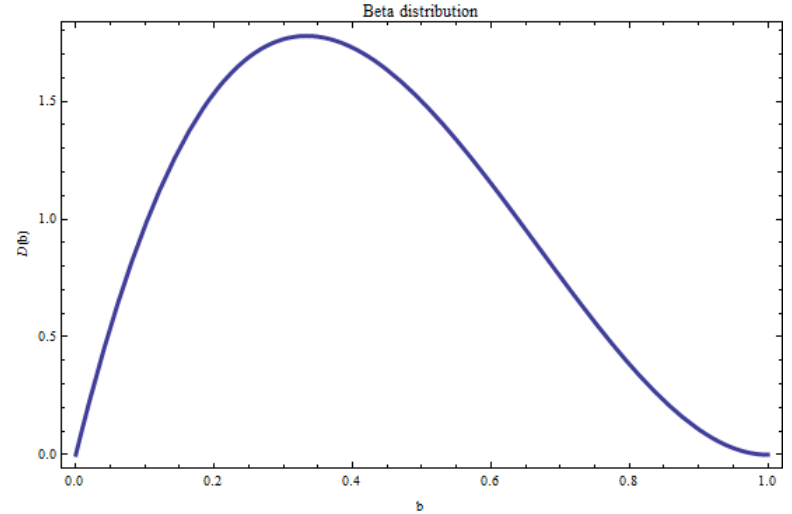
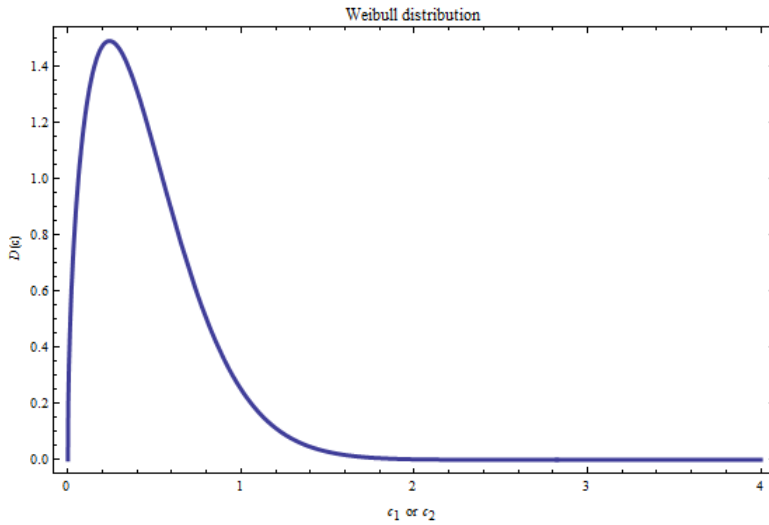


...a priori's for

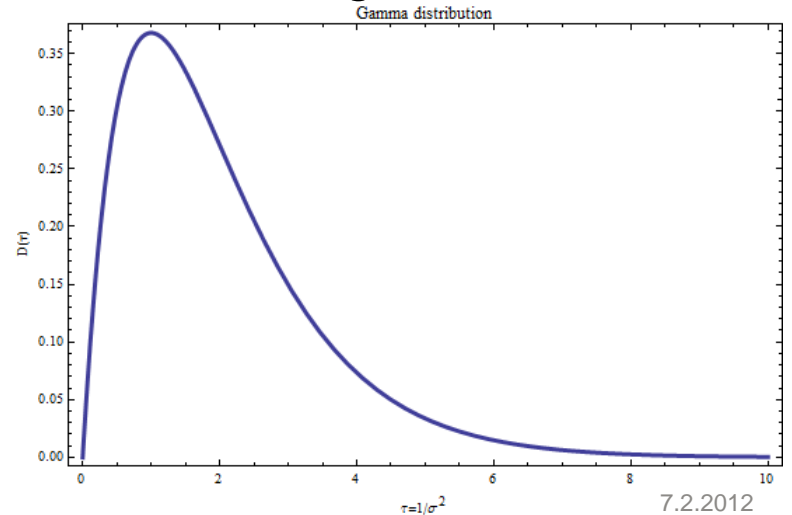
$$P(\alpha) = b \sin(\alpha)^{c_1} \cos\left(\frac{1}{2}\alpha\right)^{c_2} \sin \alpha - \alpha_0$$

c_1 and c_2

b



$1/\sigma^2$ in regression model



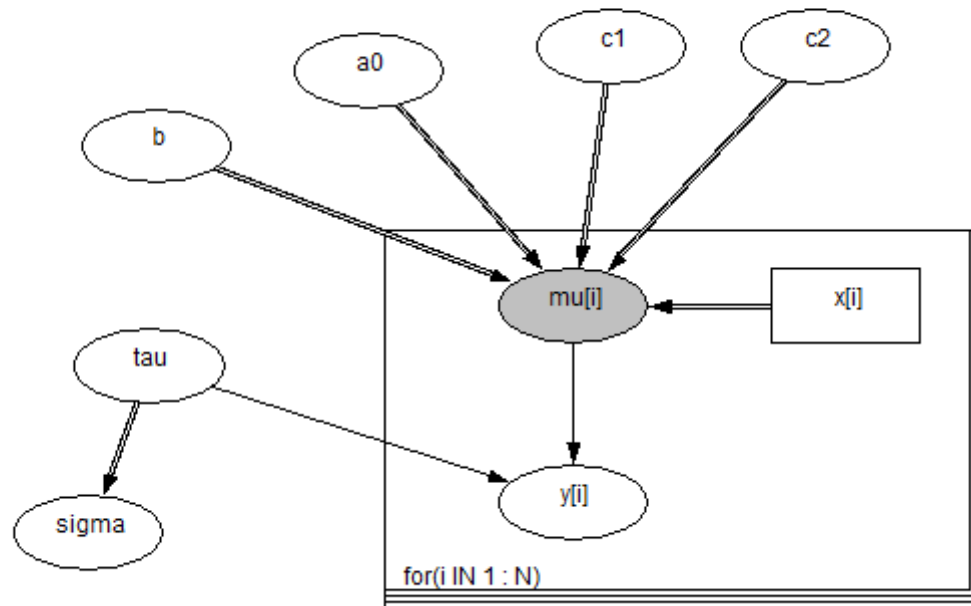
Study interactively in *Mathematica*...



Graphical model

Graphical presentation of the model in OpenBUGS

name:	mu[i]	type:	logical	link:	identity
value:	100*b*sin(x[i]-a0)*pow(sin(x[i]),c1)*pow(cos(0.5*x[i]),c2)				





Do MCMC in OpenBUGS...

- Define model and compile with data
- Look at 1 chain, trace and convergence
- Look at 2 chains, trace and convergence
- Burn-in period, look at running quantiles
- Sample length, look at kernel densities
- Finally, take coda out and analyze...

Change to *OpenBUGS* and test the demo...

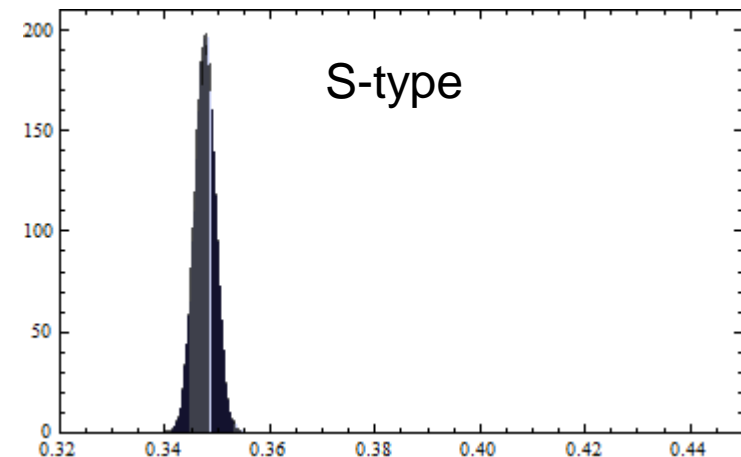
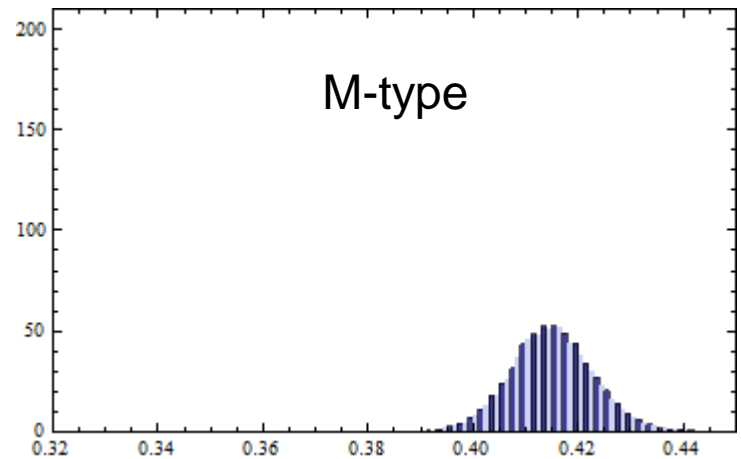
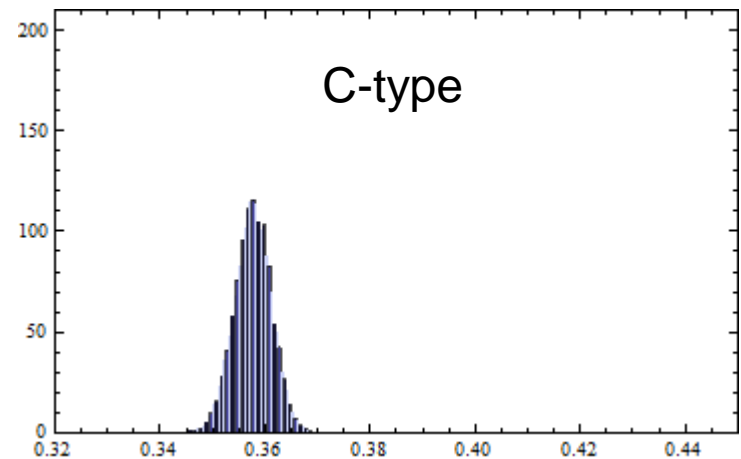


Conclusions

A posteriori histograms for α_0 , the inversion angle.

Conclusions:

- Posterior mode – point estimate to inversion angle
- Width of distribution – error/uncertainty in the estimate



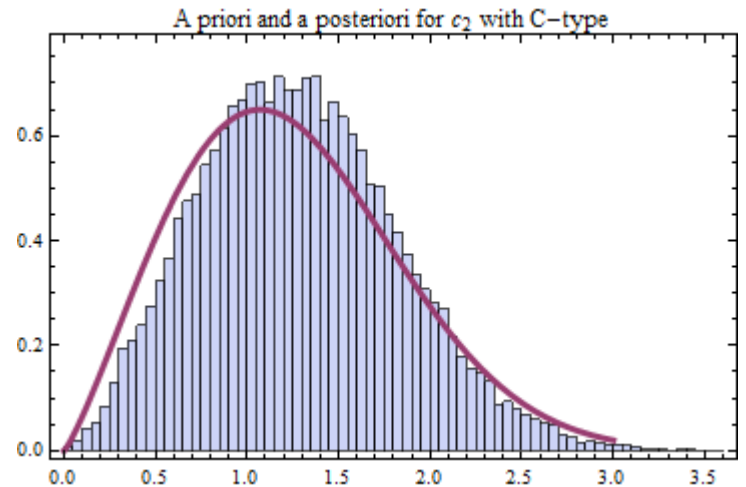
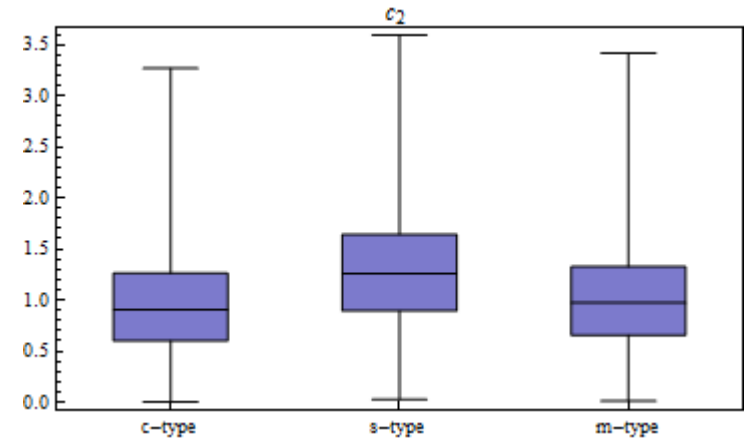


...conclusions

Do we find differences between asteroid types with some parameters?

The c_2 seems the same for all

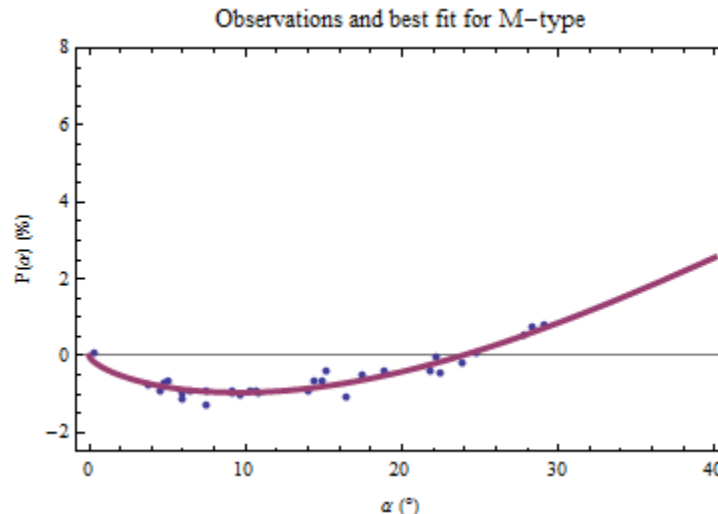
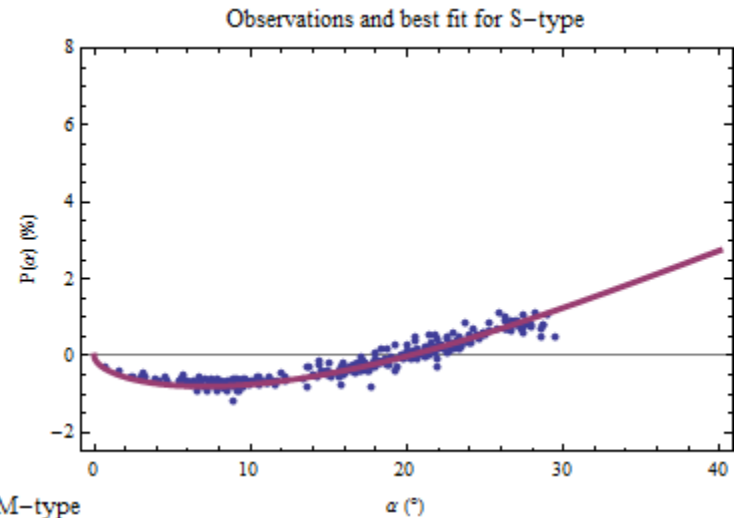
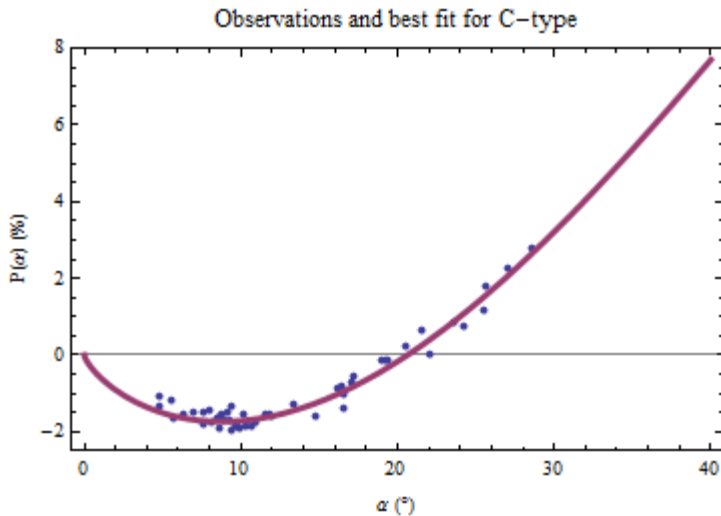
Can c_2 even be estimated? Compare a priori and a posteriori (for C-type for example)





...conclusions

The model parameters are not the only interesting thing. What about the model fit and polarization properties?

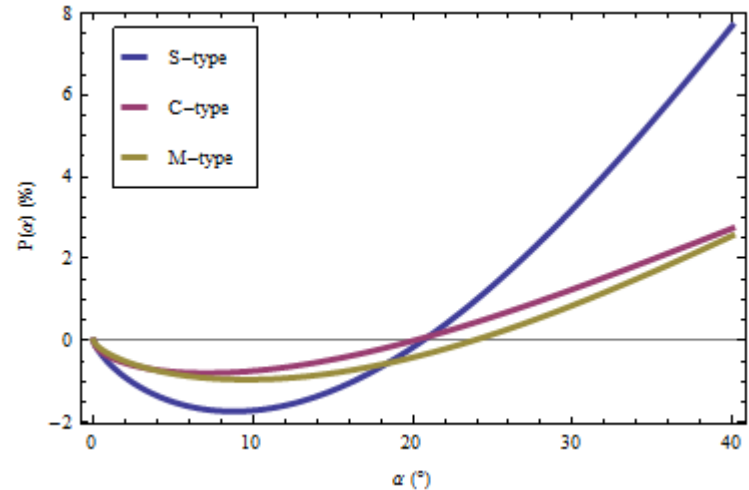




...conclusions

One interesting polarization feature – the slope at inversion. This is related to the asteroids albedo by (experimental) slope-albedo relation.

All other kinds of inference/tables/figures possible, just rely on the a posteriori samples.



Slope at inversion

