
UKRI
Science and Technology Facilities Council
ISIS Neutron and Muon Source
Welcome

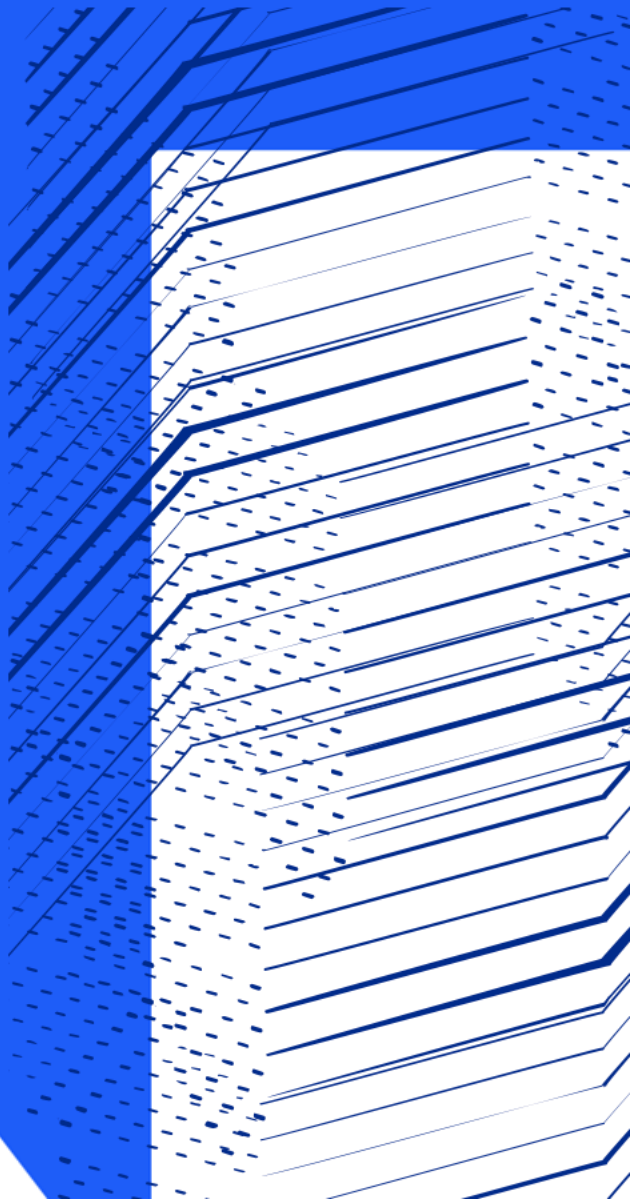
Science and Technology Facilities Council

ISIS Neutron and Muon Source

SasView:

A 'Swiss Army Knife' for SAS Data Analysis

Dr Stephen King



Outline

1 What is Data Analysis?
And what it is not!

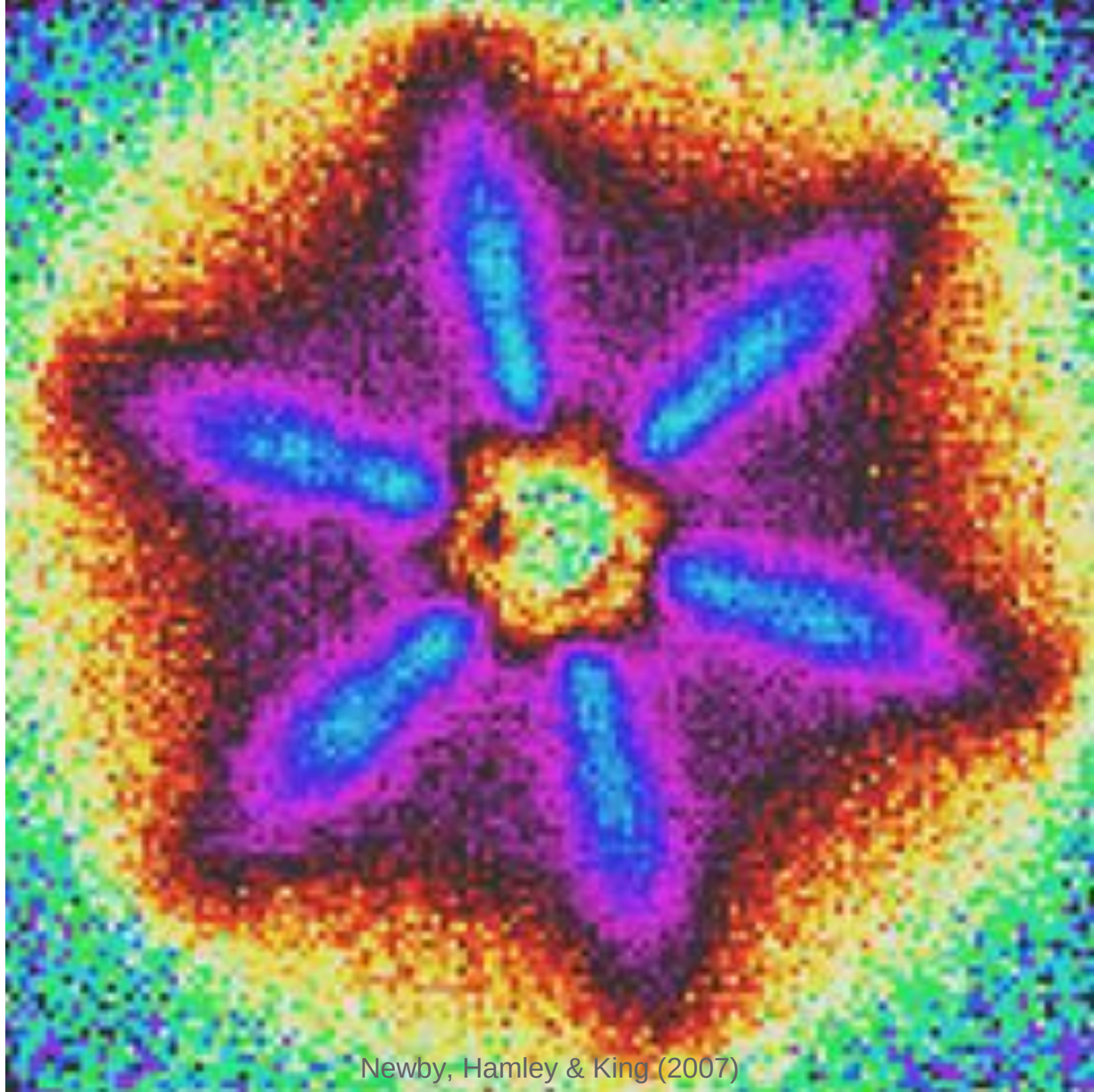
2 Data Formats
A possibly dull, but important, consideration!

3 Options for Data Analysis
Where to start!

4 The SasView Collaboration
And how you can contribute!

5 Demonstrations of SasView
Including Model-Fitting 101

6 Questions?



Newby, Hamley & King (2007)

Science and Technology Facilities Council

ISIS Neutron and Muon Source

What is Data Analysis?

Two quotes…

"Data analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making."

Wikipedia

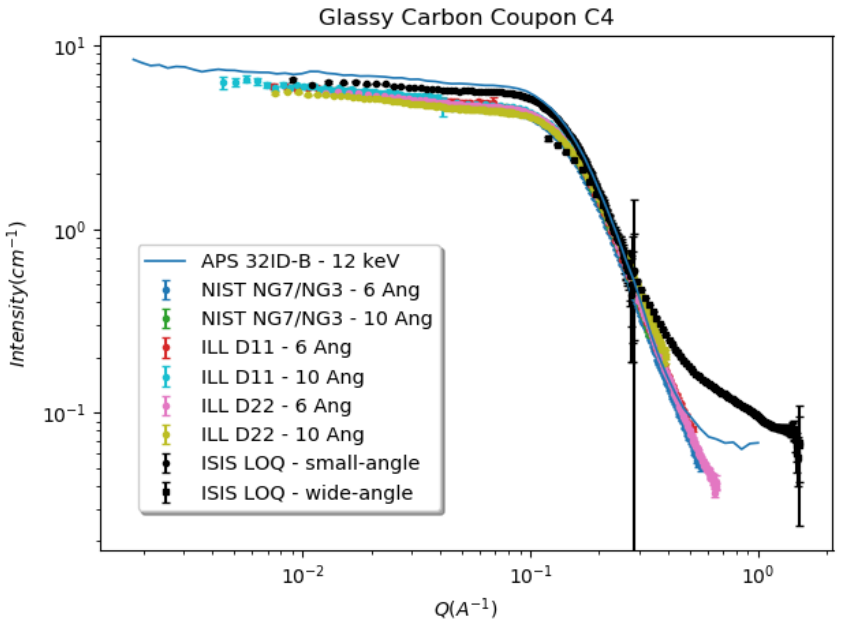
"…a way of drawing inductive inferences from data and distinguishing the signal (the phenomenon of interest) from the noise (statistical fluctuations) present in the data…"

Shamoo & Resnik (OUP, 2003)

Key:
outcomes
methods
issues

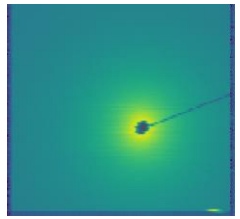
Experimental Data Flow

Reduced data should be reproducible anywhere!



Instrument-*dependent* data

Instrument-***in****dependent* data

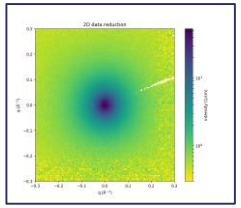


Data Acquisition System

'raw' data file (NXS)

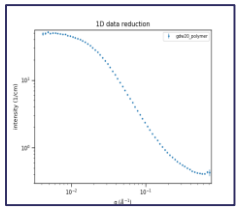
Data Reduction Programs (e.g. *Mantid*)

reduced data file (1D: HDF / XML / TXT; 2D: HDF / TXT)



Data Analysis Programs

Paper / Thesis



http://www.cansas.org

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Reduced Data Formats

Possible Types of Data

Real data:

- the scattering data; 1D (*intensity vs Q*) / 2D (*intensity vs* $Q_{parallel,x}$ *vs* $Q_{perpendicular,y}$)
- the associated uncertainties
- maybe *transmission vs* λ from TOF-SANS

Meta data:

- identifying date, time, user, instrument, sample, reduction software & version, …

Non-neutron data:

- parameters ‘scanned’ during the course of a measurement (e.g. T, P, $\dot{\gamma}$,θ , etc)
- simultaneous ancillary measurements (e.g. DLS correlation functions)
- simulation data

Possible Types of Data Format

‘Minimalist’ (e.g. ‘plain text’, TXT, DAT, CSV*):

- typically just contains some numbers + a header if you are lucky!

Self-describing (e.g. XML, CanSAS1D):

- uses ‘tags’ to order and describe the data/metadata to an agreed structure

Hierarchical (e.g. HDF, NXcanSAS):

- also self-describing, but fully portable (no outside information needed)
- can contain a mix of data types (‘a file-system-in-a-file’)

SasView inputs many of these (and more)!

* Beware the semi-colon ; !

The CanSAS Standards

CanSAS1D

- v1.0 released May 2009
- v1.1 released Mar 2013
- for 1D *reduced* data only
- XML-based (.xml)
- still human-readable

http://www.cansas.org/formats/canSAS1d/1.1/doc/

NXcanSAS

- v1.0 released Jan 2017
- for nD *reduced* data
- HDF5-based (.h5)
- NeXus-structured
- not human-readable

http://www.cansas.org/wgwiki/index.php/NXcanSAS

https://www.nexusformat.org/

https://www.hdfgroup.org/downloads/hdfview/ (free!)

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Options for SAS Data Analysis

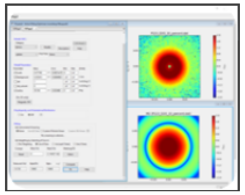
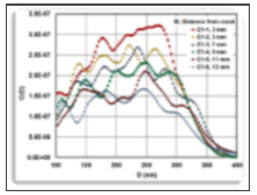
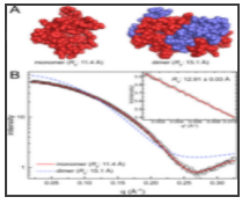
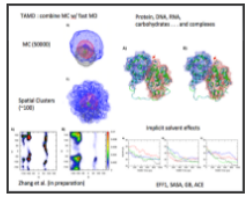
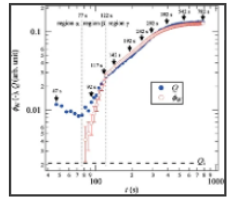
Many Options!

But:

- not every option is needed every time!
- often one option will suffice

Influences:

- what you want to know
- the data quality
- your field/speciality

Model-Fitting Methods	Real-Space Methods	Ab-Initio Methods	MC/MD Methods	Other Methods
 *Example of 2D model-fitting using the SasView application*	 *Cavity size distributions in a steel weldment as derived from SANS 10.1179/1743284714Y.0000000577*	 *Ab-initio modelling of polcalcin constrained by SAXS 10.1002/pro.3376*	 *MC & TAMD modelling of proteins constrained by SANS 10.1016/j.jmgm.2017.02.010*	 *Time evolution of the invariant during crystallisation of P4MP1 10.1038/pj.2012.204*
This approach uses iterative optimisation to match the calculated scattering from a model function describing the scattering objects to the measured scattering data. Each iteration one or more physical parameters describing the model (e.g. concentration, size, scattering length density) are adjusted.	This approach uses mathematical transformations (e.g. Fourier Transforms) to convert the measured scattering data in reciprocal-space (i.e. in Q-space) into a function in real-space. Typical outputs are density correlation functions, volume fraction distributions, and size distributions.	This approach uses iterative optimisation to match the calculated scattering from assemblies of spheres or from a 3D 'shape envelope function' to the measured scattering data. Each iteration the number and/or position of the spheres, or the curvature of the envelope function, is adjusted.	This approach uses iterative optimisation in combination with Monte Carlo (MC) and/or Molecular Dynamics (MD) techniques or RRT searches to match a calculated 'atomistic level' structure for the scattering objects to the measured scattering data.	Other approaches to data analysis may involve identifying, for example: any Q-dependencies in the measured data, particular patterns in the Q-values of any peaks present, asymptotic extrapolations, calculation of the integral under the measured data (the 'invariant'), or the intensity at Q=0.

https://www.isis.stfc.ac.uk/Pages/SANSdataanalysisOverview.aspx

Science and Technology Facilities Council

ISIS Neutron and Muon Source

The SasView Collaboration

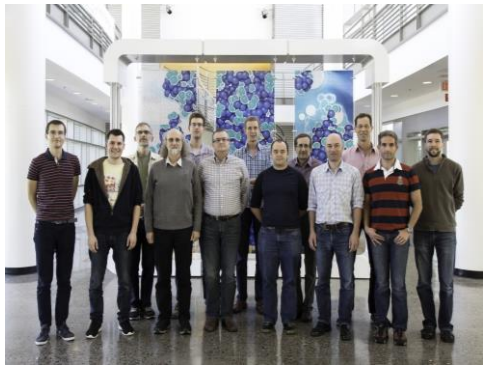
Status of SasView

- 2006; originates in NSF *DANSE* project
- 2013; transitions into a community project

- ~40 contributors from 9 organisations (~15 active at any one time)

- twice monthly zoom calls
- regular ‘camps’ & ‘hackathons’

- small management team:
  Paul Butler (NIST), Mathieu Doucet (ORNL), Andrew Jackson/Wojciech Potrzebowski (ESS), Steve King (ISIS)

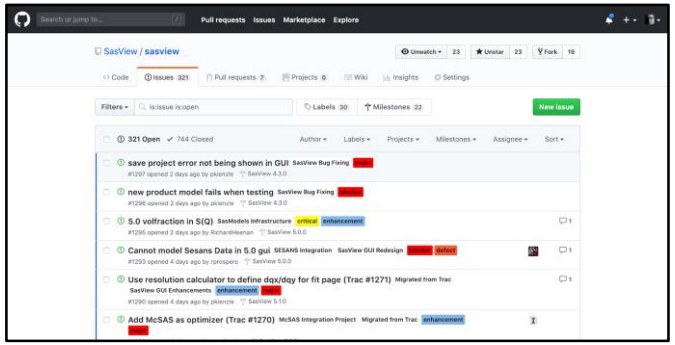


Open, Collaborative, Community **Development**

Code is open source and publicly hosted at Github
Released under BSD 3-clause license

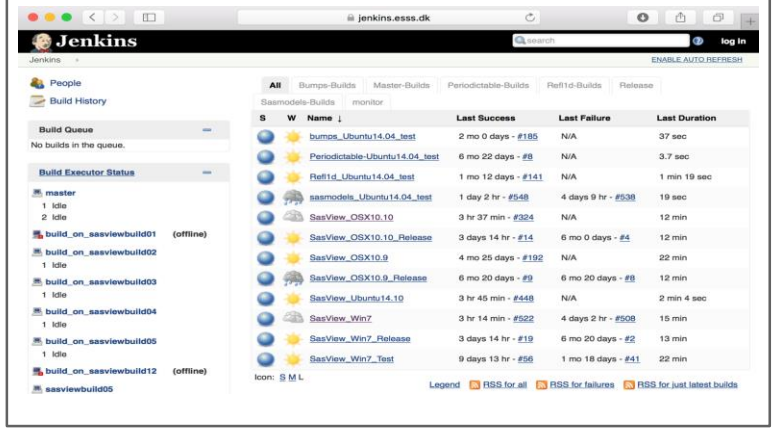
Rolling **5 Year Roadmap**

Code Hosting, **Issue Tracking** & Developer Wiki on Github



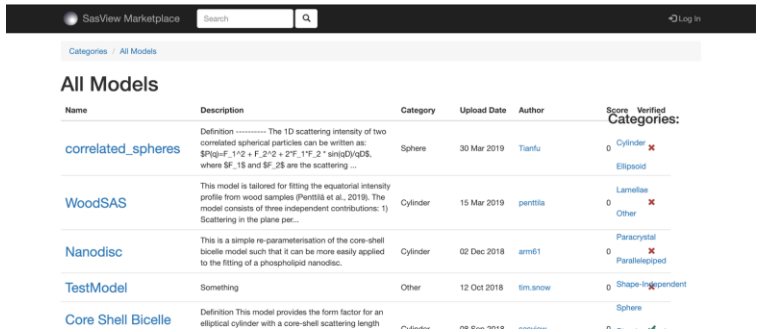
https://github.com/SasView

Automated Builds



http://build.sasview.org/

Model Marketplace for Users to share their models



http://marketplace.sasview.org

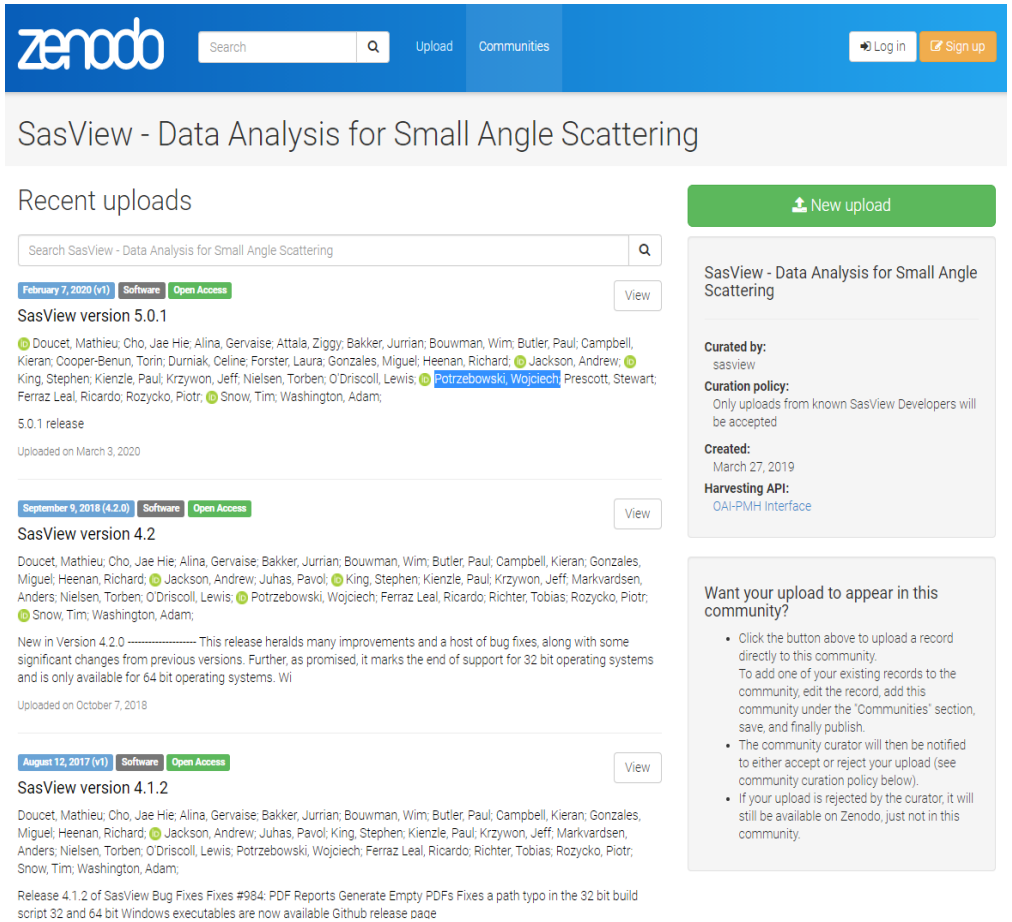
python™

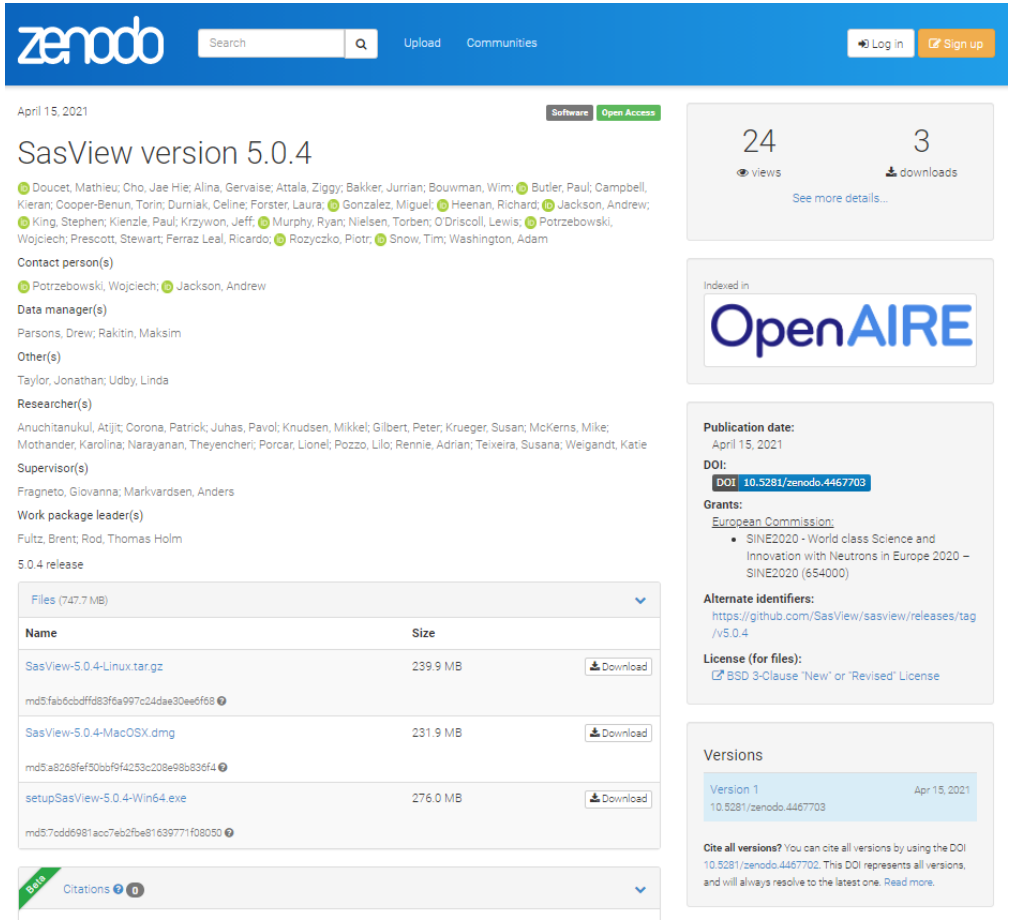
https://www.sasview.org

matplotlib

Open, Collaborative, Community Development

DOI for each release



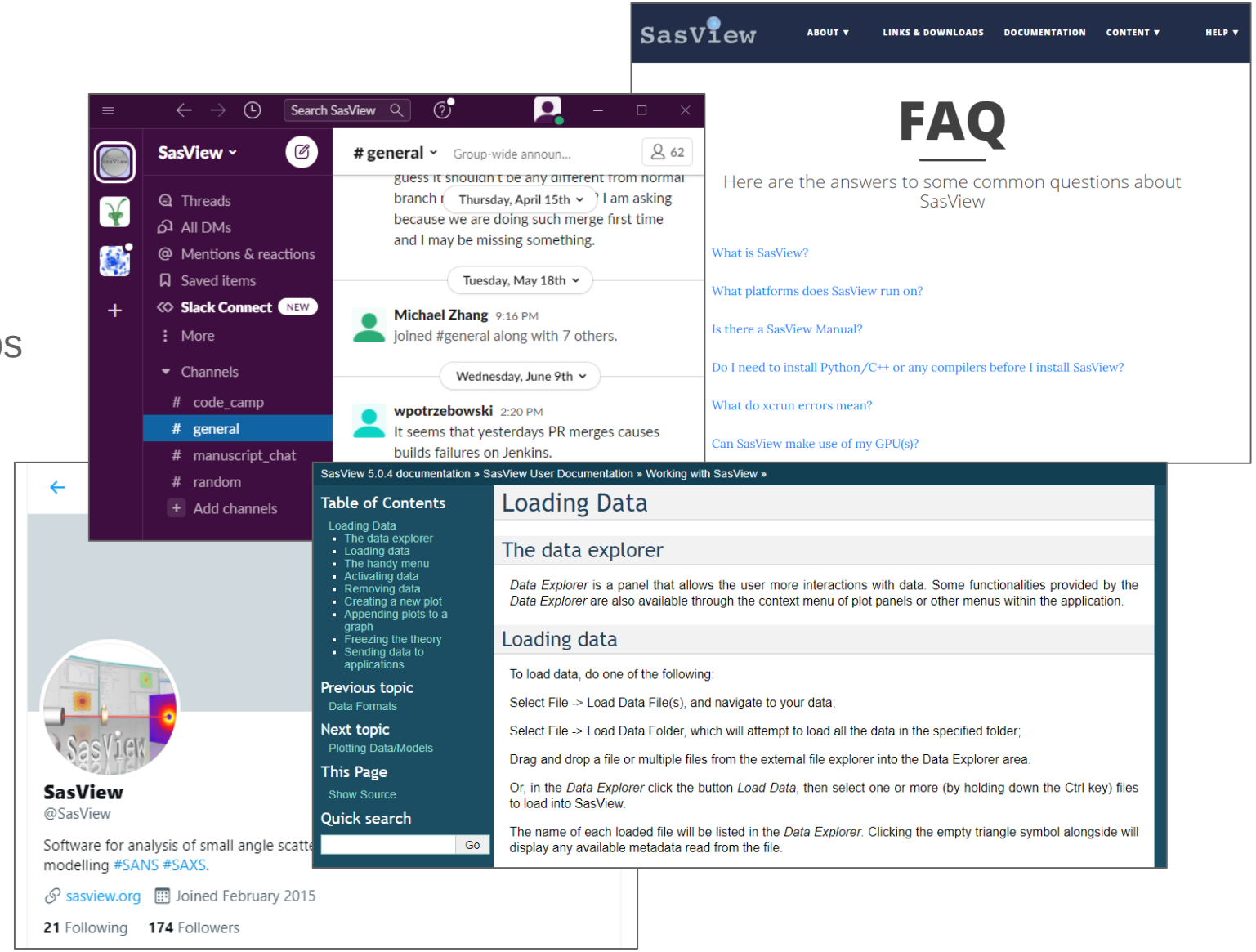


https://zenodo.org/communities/sasview-analysis

Open, Collaborative, Community Learning

- **Website**
- **Documentation**
  - in-program & online
- Written Tutorials
- Video Tutorials (YouTube)
- Taught Courses
  - scattering schools/workshops
  - university courses
- Bootcamps & regional workshops
- e-Learning
- **Slack**
- **Twitter**
- (Marketplace)

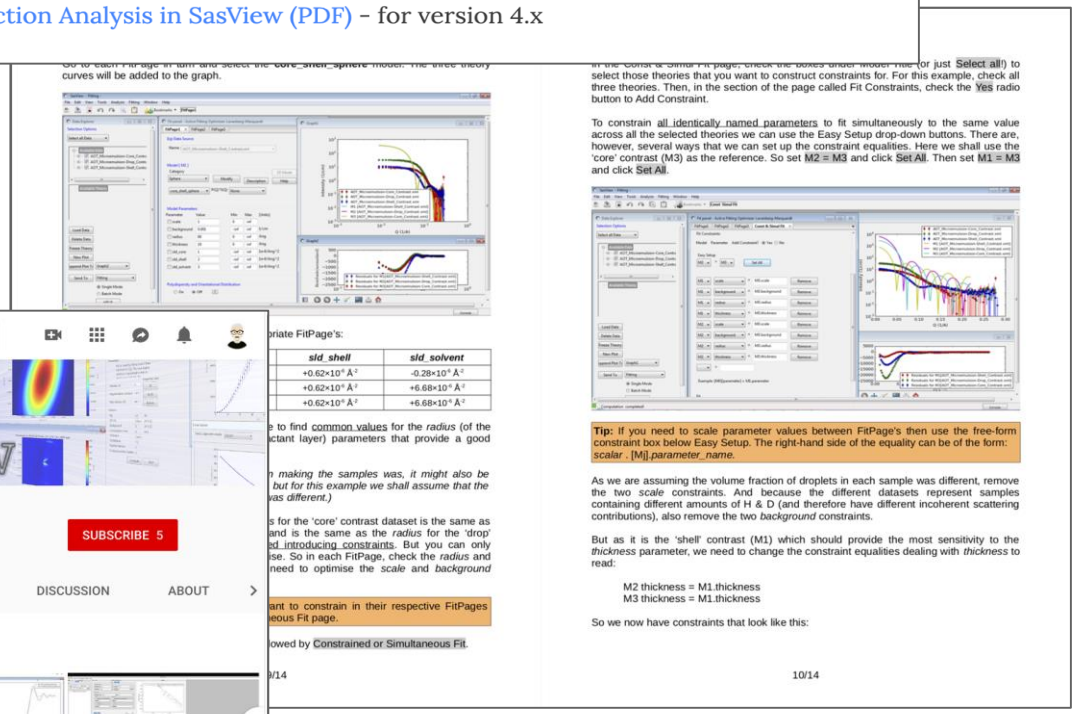
- **help@sasview.org**
- **users@sasview.org**

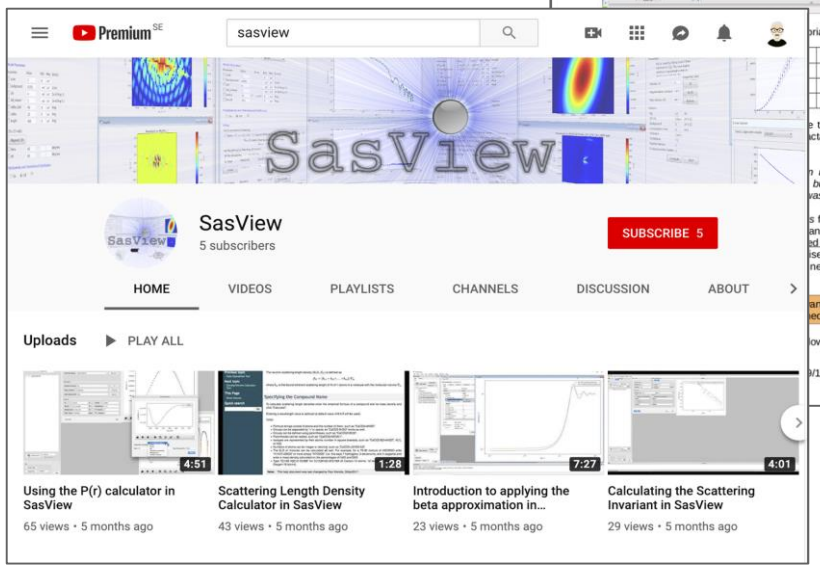


Open, Collaborative, Community Learning

- Website
- Documentation
  - in-program & online
- **Written Tutorials**
- **Video Tutorials (YouTube)**
- Taught Courses
  - scattering schools/workshops
  - university courses
- Bootcamps & regional workshops
- e-Learning
- Slack
- Twitter
- (Marketplace)

- help@sasview.org
- users@sasview.org



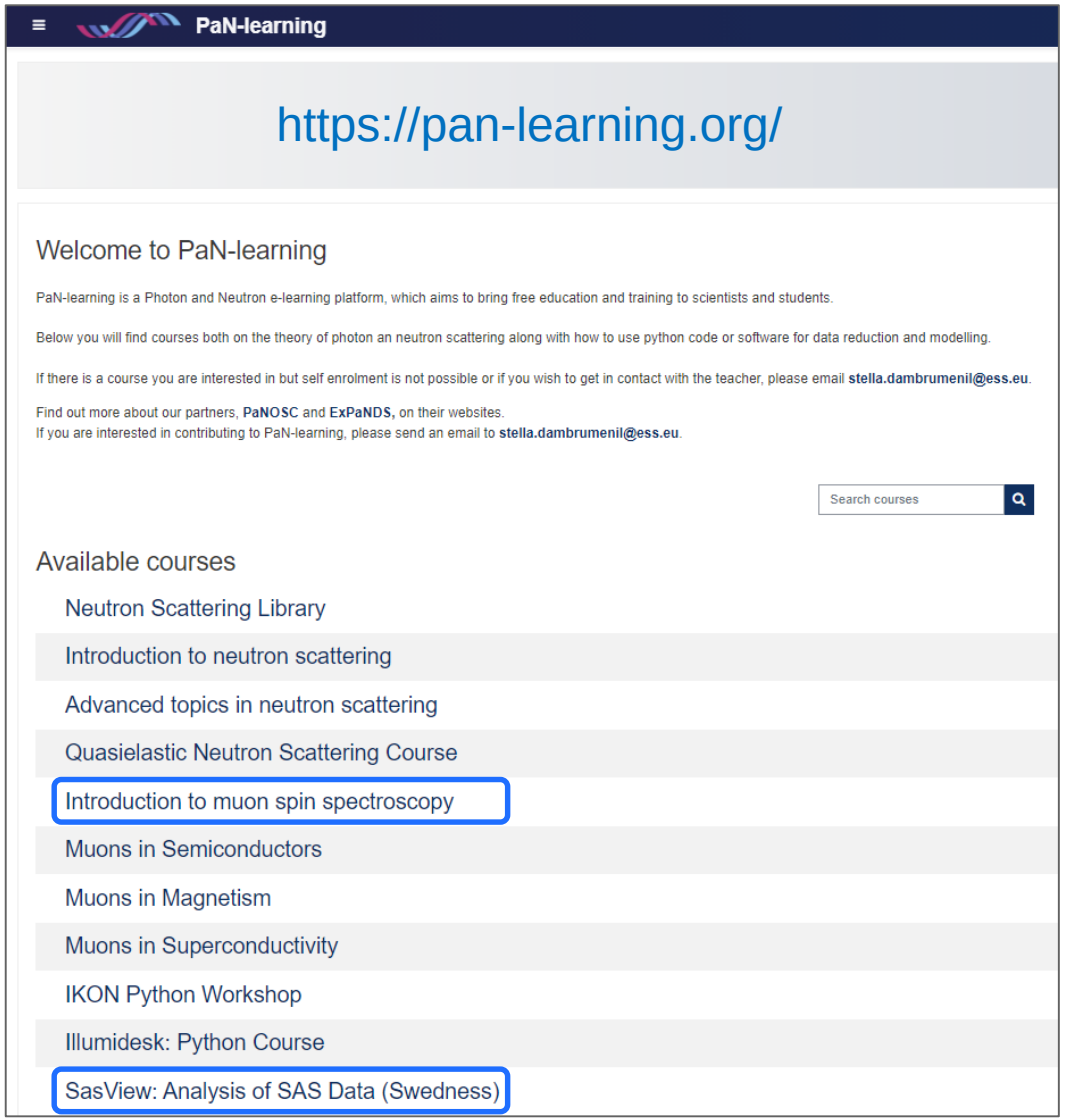


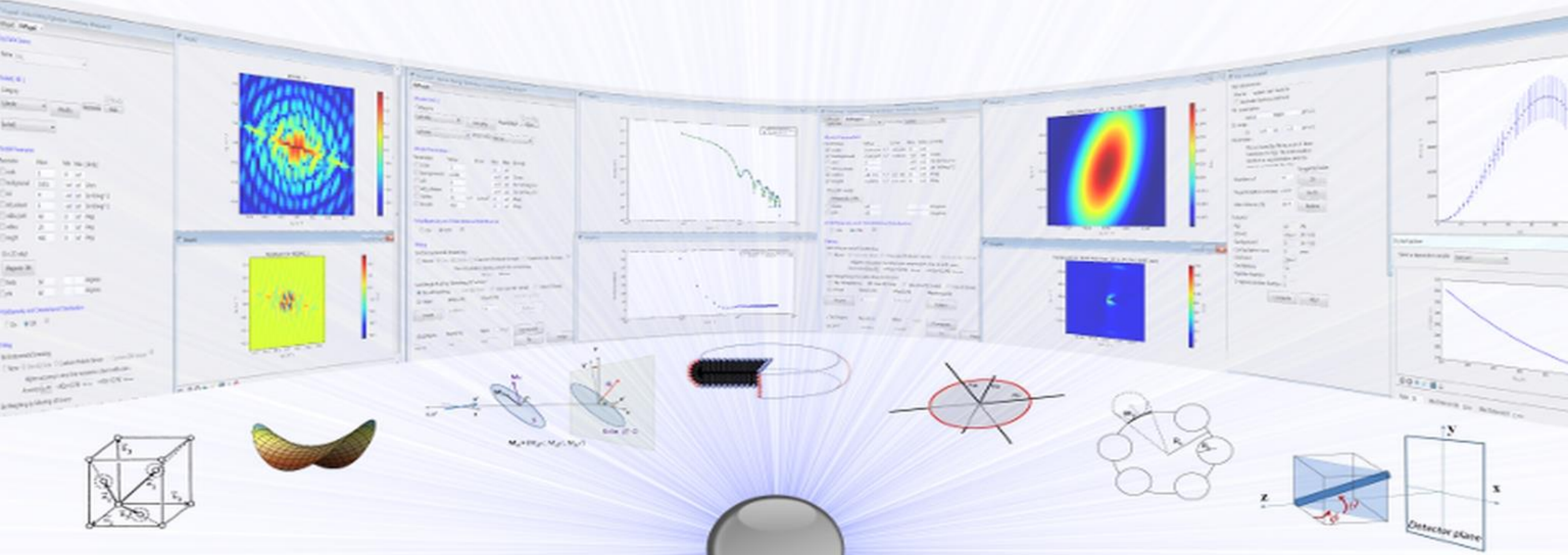
Open, Collaborative, Community Learning

- Website
- Documentation
  - in-program & online
- Written Tutorials
- Video Tutorials (YouTube)
- Taught Courses
  - **scattering schools/workshops**
  - university courses
- Bootcamps & regional workshops
- **e-Learning**
- Slack
- Twitter
- (Marketplace)

- help@sasview.org
- users@sasview.org

All the work of ISIS Sandwich Student Michael Oakley





SasView

UT ISIS NIST diamond OAK RIDGE National Laboratory ansto ILL NEUTRONS FOR SCIENCE ess TU Delft

Model-Fitting 101

- **The objective of model-fitting is to find a model solution (theory) that:**
  - describes the form of the experimental data
  - offers minimal uncertainty
  - has converged nicely
- **And above all is…**

physically realistic !!!

"The fact, that A model fits your data, it is NOT proof that it is THE appropriate model."
Jan Ilavsky, APS

Model-Fitting 101

- **Wherever possible, select a model/parameters using *a priori* knowledge:**
  - ancillary data (EM, DLS, *etc*)
  - published literature
- **A model with more parameters is not always better than a model with fewer parameters!**
- **Simultaneously fitting too many parameters is generally a bad idea!**
  - at least, until the program is homing in on a solution

Here's why…

Model-Fitting 101

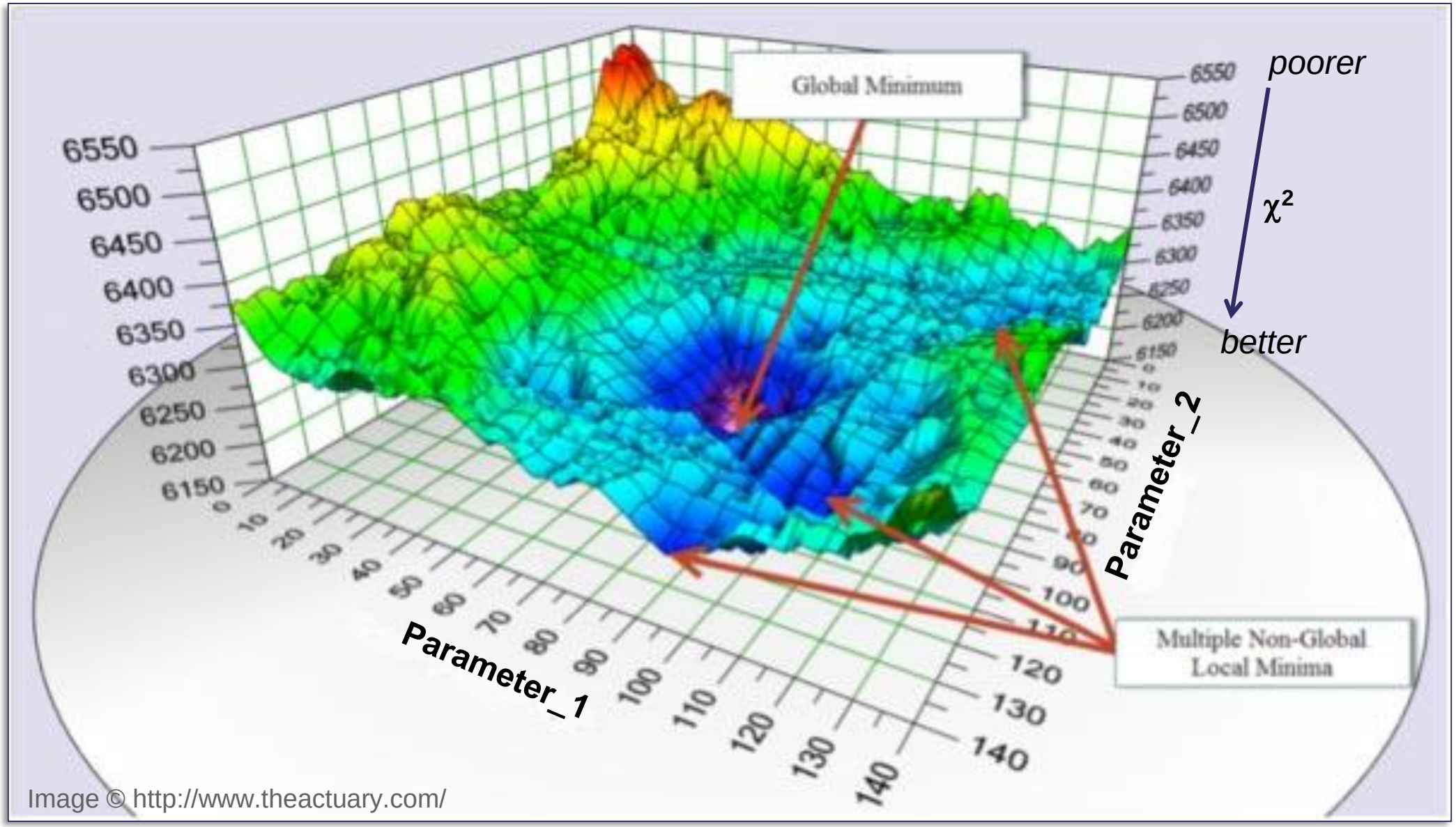


Image © http://www.theactuary.com/

Model-Fitting 101

- **Finding the global minimum → optimisation**
  - you can have speed *or* robustness, but not (generally-speaking) both!
- **For speed: → gradient-descent algorithms**
- **For robustness: → population algorithms**
  - also called ‘evolutionary’ or ‘genetic’ algorithms

    → Bayesian (probability) algorithms
  - AI methods…

SasView features 5 optimisers…

SasView Optimisers

- **For speed: → gradient-descent algorithms**
  - ***Levenberg-Marquardt***
    - the tried-and-trusted workhorse found in every model-fitting package
    - use when you have a reasonable fit in a minimum & want the best values
    - but beware it finding local minima
  - ***Quasi-Newton BFGS***

- **In-between:**
  - ***Nelder-Mead Simplex***
    - a good first choice (especially if your function is well-behaved)

- **For robustness: → population & Bayesian algorithms**
  - ***Differential Evolution***
  - ***DREAM***
    - use when you need a really robust fit &/or want to explore correlations

'bumps'

https://pypi.org/project/bumps/
Paul Kienzle, NIST

https://www.sasview.org/docs/user/qtgui/Perspectives/Fitting/optimizer.html

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Time for a Demonstration!

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Thanks to:

The members of the
SasView Development Team

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Questions?

www.sasview.org

@SasView

sasview.slack.com

e: help@sasview.org

Science and Technology Facilities Council

ISIS Neutron and Muon Source

Thank you

www.sasview.org

@SasView

sasview.slack.com

e: help@sasview.org