

An Expert System for Automated Modeling of Small- Angle Neutron Scattering Data

William Heller

Richard Archibald

Robert Patton

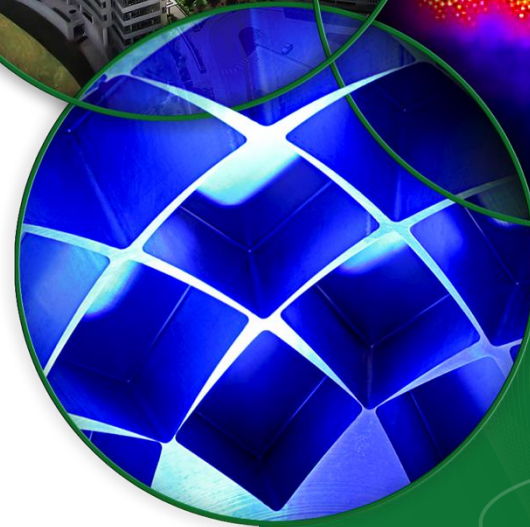
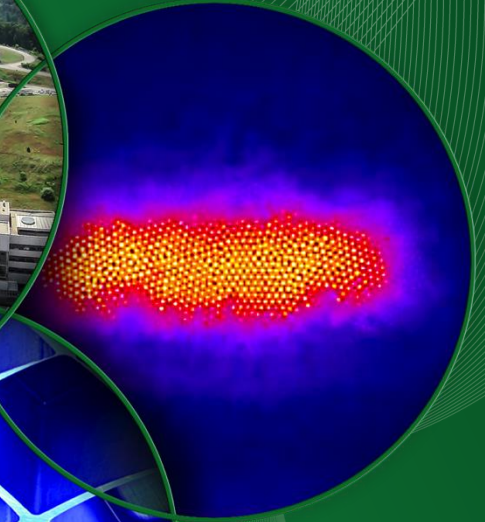
Mathieu Doucet

Steven Young

Robert French

March 21, 2016

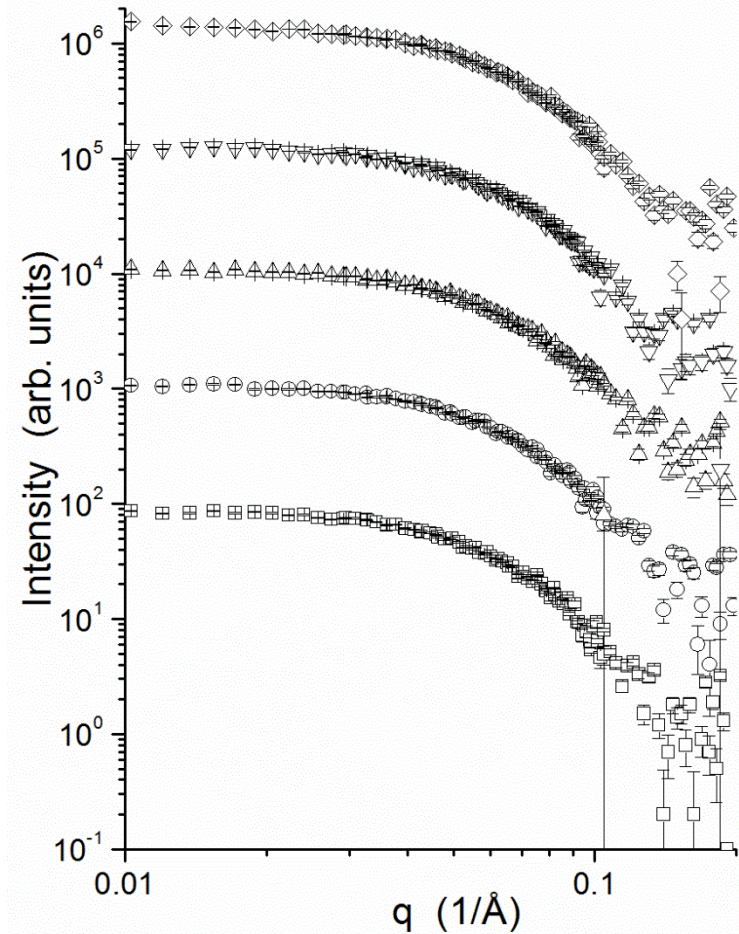
ORNL is managed by UT-Battelle
for the US Department of Energy



The Problem

- SANS data is not information rich
 - Size, shape, bulk statistical properties
 - Models with different physical meanings can fit the data equally well
- The NCNR SANS data analysis software package has well over 50 models implemented

Where does one start?

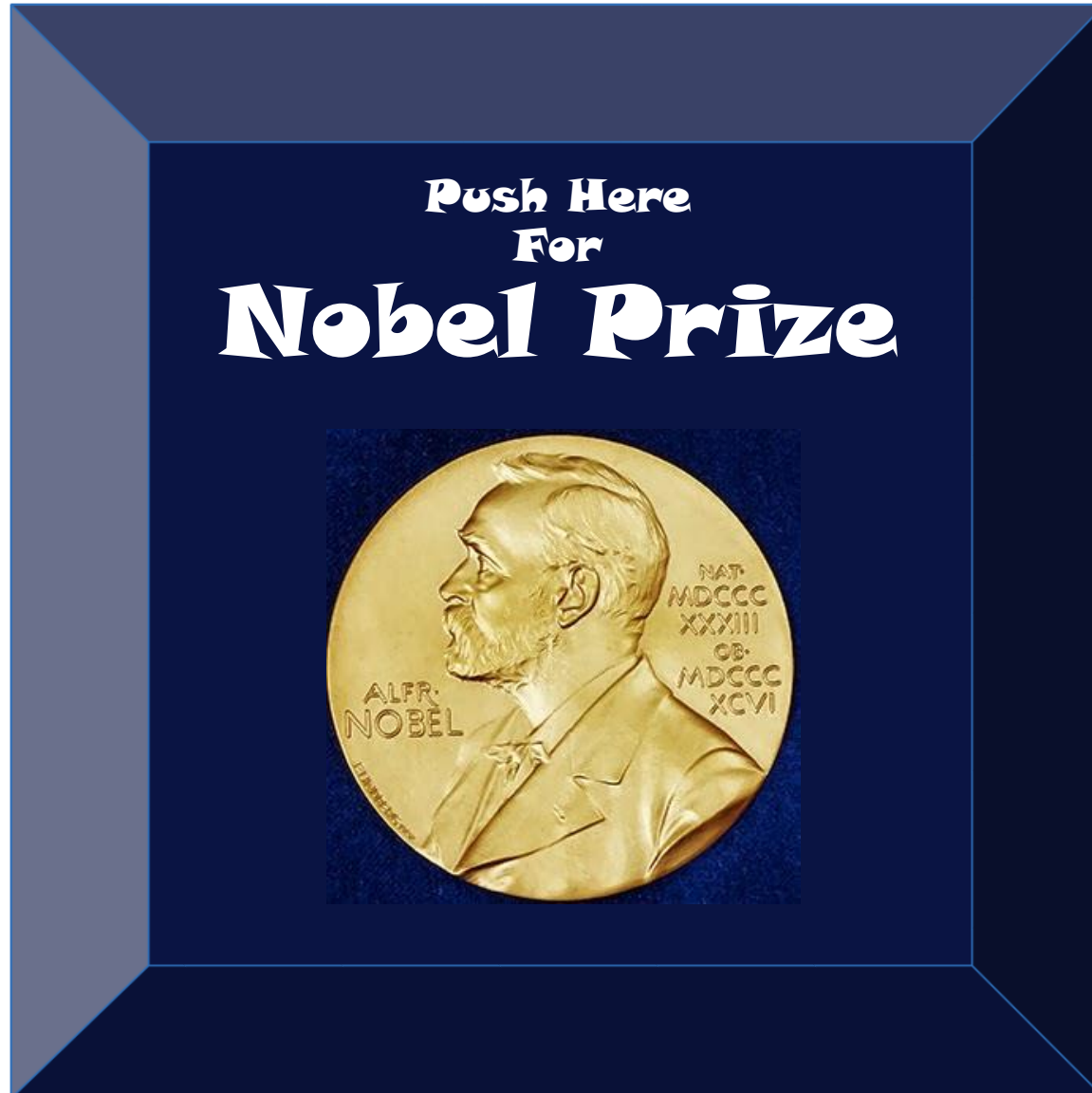


The User Community

- The users of SANS are quite often materials experts rather than technique experts
 - Many come from fields that often do not require much mathematical training
 - Many are not regular users
 - SANS will provide once piece of the puzzle in their broader project

Assistance with data analysis is often required from the instrument scientist who helped them with their experiment

What users seem to want...

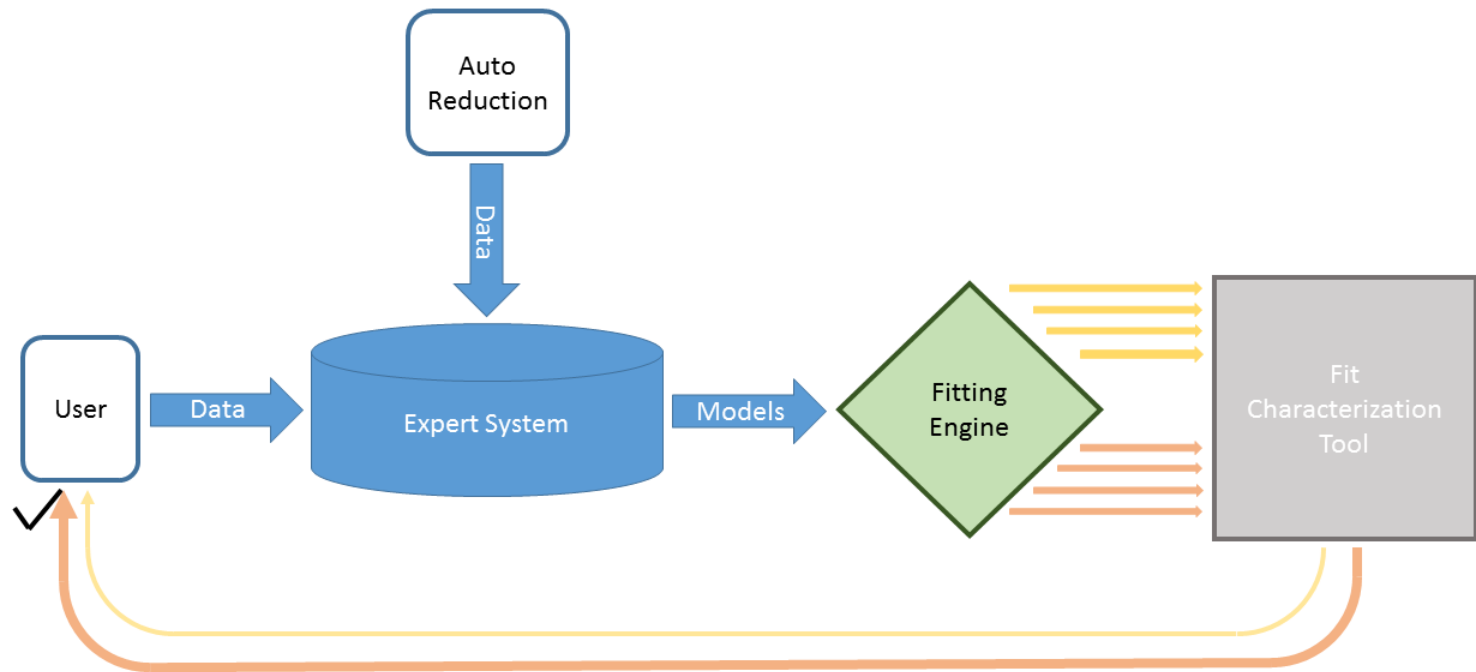


The Idea

- Facilitate the process of modeling SANS data by using Machine Learning and High-Performance Computing
 - Provide guidance to the non-expert, “casual” users
 - Provide robust tools for fitting data that automate the process
 - Provide a tool that probes the potential ways to fit the data that also characterizes how broad the space of solutions that fits the data is

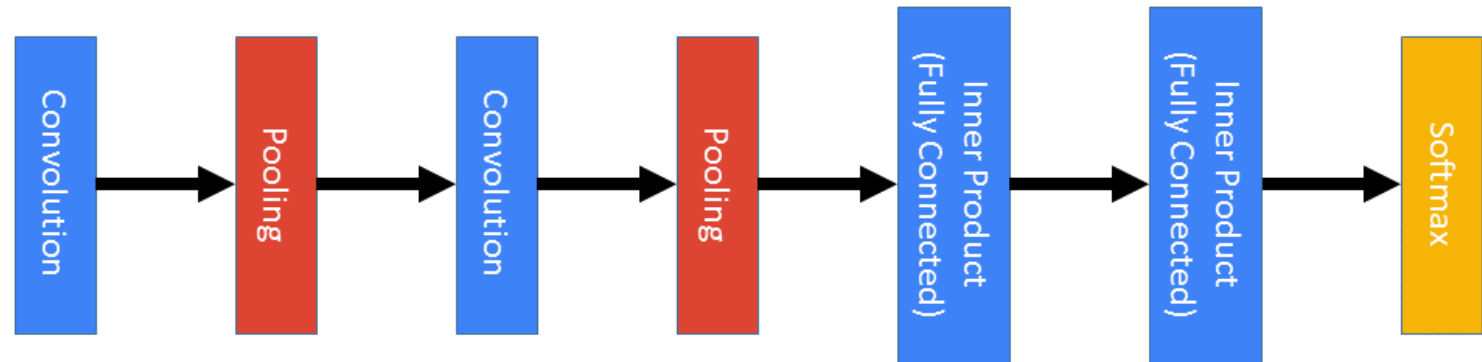
- The goal is to help improve beamline productivity

The Idea



The Expert System

- Use Deep Learning to do what an instrument scientist does – look at a plot of the data as a picture



- The initial implementation is a convolutional neural network
- Training data are models calculated using SASView with Gaussian noise added
- Researching what distinguishing features exist in the data that could improve the training of the system

The Fitting Engine

- The Expert System tells the Fitting Engine what models to try to use to fit the data
- The Fitting Engine is envisioned to run on a “cluster” or a HPC system in a headless configuration
 - Fit the data many times with each model
- We want to build upon use is SASView & Bumps
 - SASView has engaged the community and involves several user facilities
 - A large library of models in the SASView set
 - Bumps has several fitting engines
 - Excellent documentation of the models

The Fit Characterization Tool

- Unfortunately, there are often several models that can fit a data set reasonably
- Even worse, there can be distinct parameter sets for a single kind of model that can fit data equally well
- We want to develop tools for measuring these parameter spaces to help distinguish which models are better than others
 - Filter the output returned to the user...while being wary of “garbage in, garbage out”

Acknowledgements

Research sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U. S. Department of Energy (Project-8235)