JetWeb: a WWW interface and database for Monte Carlo tuning and validation

J M Butterworth, S Butterworth, B Waugh
presented by B Waugh at MC4LHC workshop, CERN,
2003-07-30

Outline of talk

- Physics motivation
- Software technology
- Object model
- Demonstration
- Example application
- Planned improvements
- Conclusion

Physics motivation

- Need to understand hadronic final state for current and future measurements.
- Can't predict everything from first principles: need to choose e.g. underlying event model/parameters.
- Tuning to fit one data set can result in a poor description of other data.
- Better to tune to many data sets simultaneously.
- Want to:
  - automate comparisons as far as possible
  - avoid duplicating effort (and CPU usage)

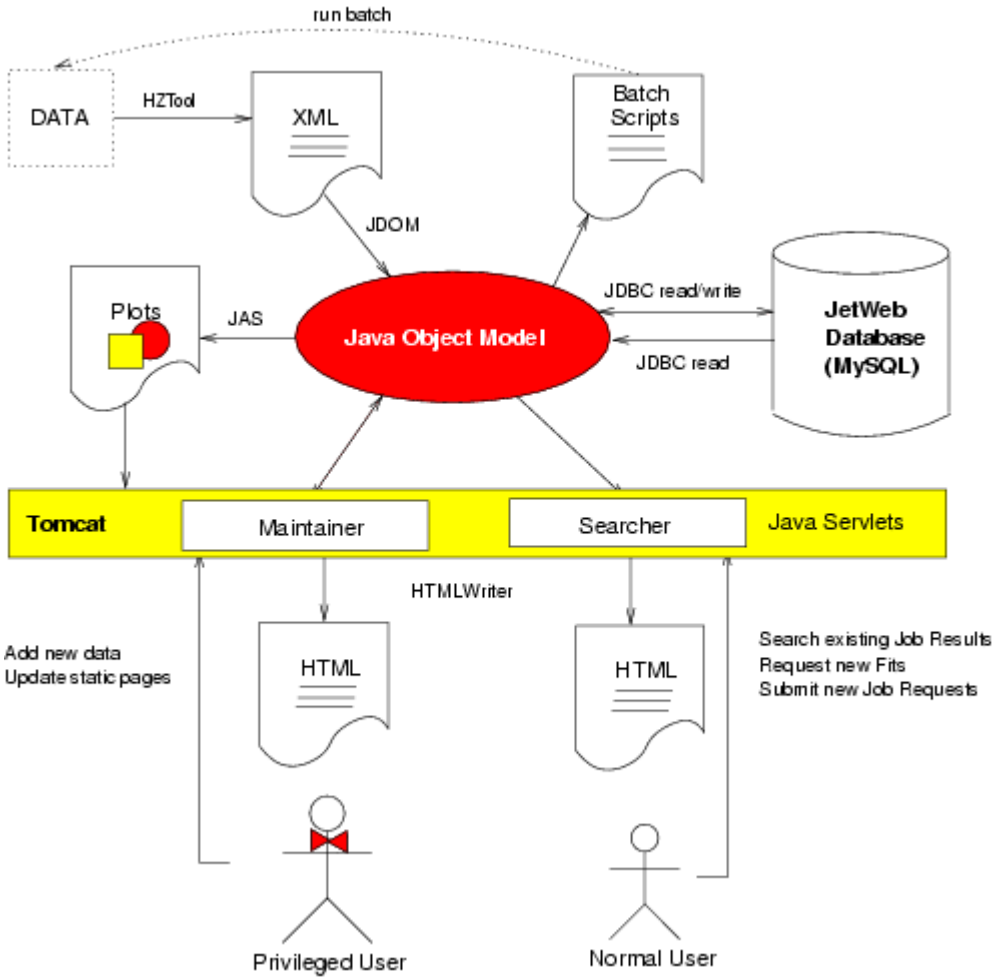
HzTool

- Fortran library: fills HBOOK histograms from generated events to compare with measurements.
- Developed in HERA Workshop:
  - J. Bromley et al., Future Physics at HERA, vol 1, 611-612
- Routine written in Fortran for each analysis.
- Range of data already included: H1, ZEUS, UA5, OPAL, CDF, D0 (in JetWeb version).
- More data being added, more wanted.

JetWeb

- J Butterworth and S Butterworth:
  Comput. Phys. Commun. 153 (2003) 164-178.
- Builds on functionality of HzTool.
- Database of measurements, predictions and fits.
- Web front end allows:
  - search database
  - submit jobs to generate MC with chosen parameters
- Uses: Java servlets, MySQL relational database, XML, Java Analysis Studio (JAS).
- HzTool jobs can be submitted to a standard batch farm or to a grid.

JetWeb overview



Object model

- Java classes correspond (more or less) to tables in database.
- Data: papers contain plots contain points.
- Model = generator/version + parameters
  - several models can be equivalent for particular beams, e.g. proton PDF irrelevant to e^+e^- data, so introduce class *logparms* = *model* + *beams*
- Predicted plots/points
- Fit: comparison between data and model
  - Normalized to fit selected high-Et data
  - Chi-squared calculated for various data sets

Using JetWeb

- Start at http://jetweb.hep.ucl.ac.uk/
- Search the database: possibilities include
  - select MC generator and parameters, and test description of a range of data;
  - select data set of interest and find models giving a good description of chosen data;
  - more generally: restrict some MC parameters while allowing others to vary, and compare predictions with all available data or a subset.
- Use ready-made searches under *Selected Results*.
- Best fit pages are automatically updated.

Searching the database

- Select generator(s) of interest
- Restrict parameters as required, or leave them free
  - Common: ptmin, underlying event model, PDFs, kt
  - Generator specific (e.g. Pythia: MSTP, PARP)
- Choose order of results (e.g. find best fit to charm data)
- *Get results!* (May take minutes for very general search.)
- If nothing in DB matches your request, you can submit a request to generate some new MC.

Applications

- Tune and validate MC models.
- Select suitable parameters for use in measurements at existing colliders.
- Predict QCD background at Future Linear Collider (J Butterworth and M Wing).
- Work on LHC underlying event tuning underway.
  - see Arthur's talk tomorrow for more details

FLC study (from talk by M Wing)

(Default) predictions at 500 GeV

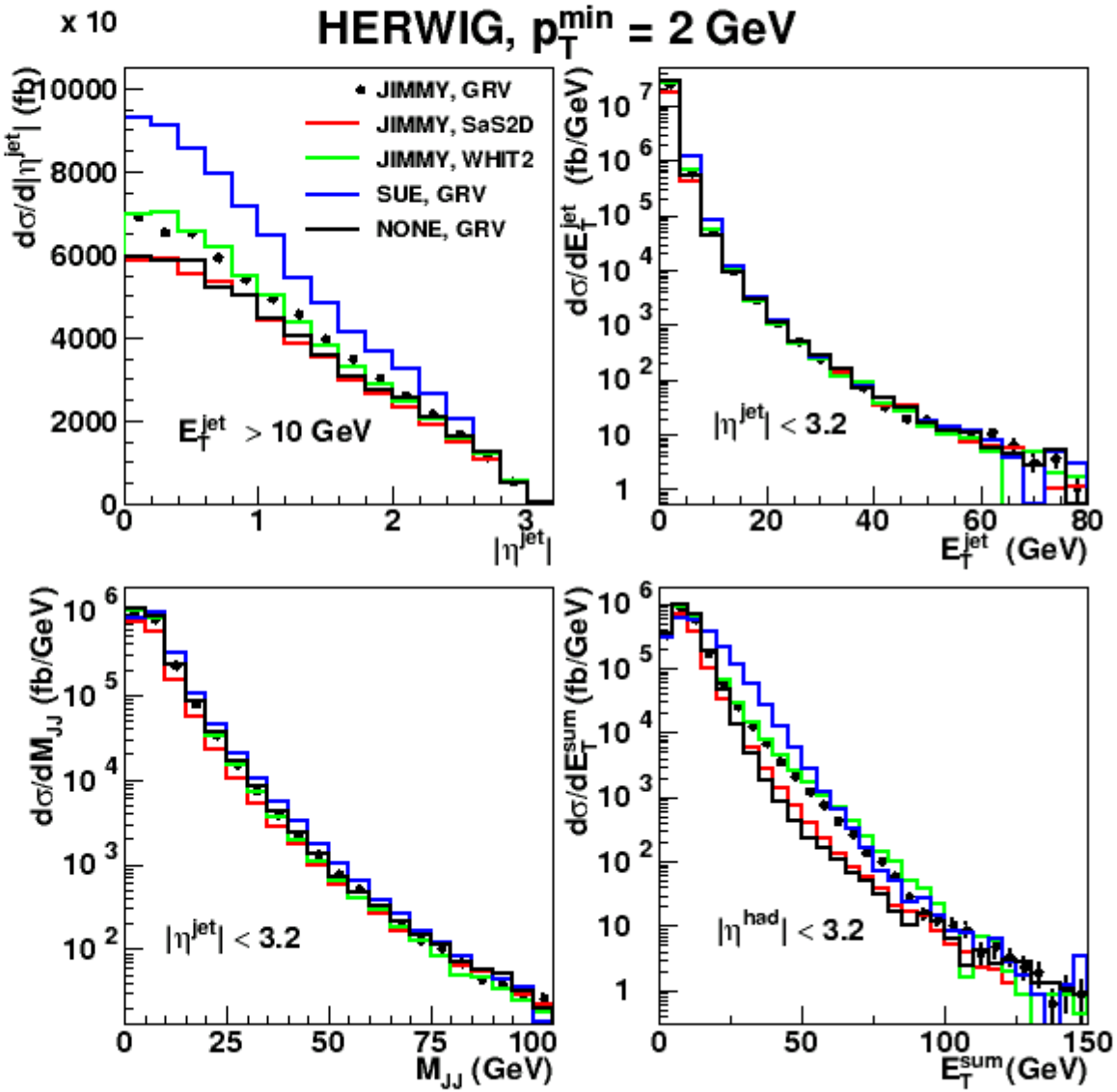
Default HERWIG prediction used with changes in underlying event and photon PDF.

All "reasonable" parameter settings.

Large spread in predictions, even at high energies.

How accurately do we know QCD production?

Not very well!



FLC study (from talk by M Wing)

(Fitted) predictions at 500 GeV

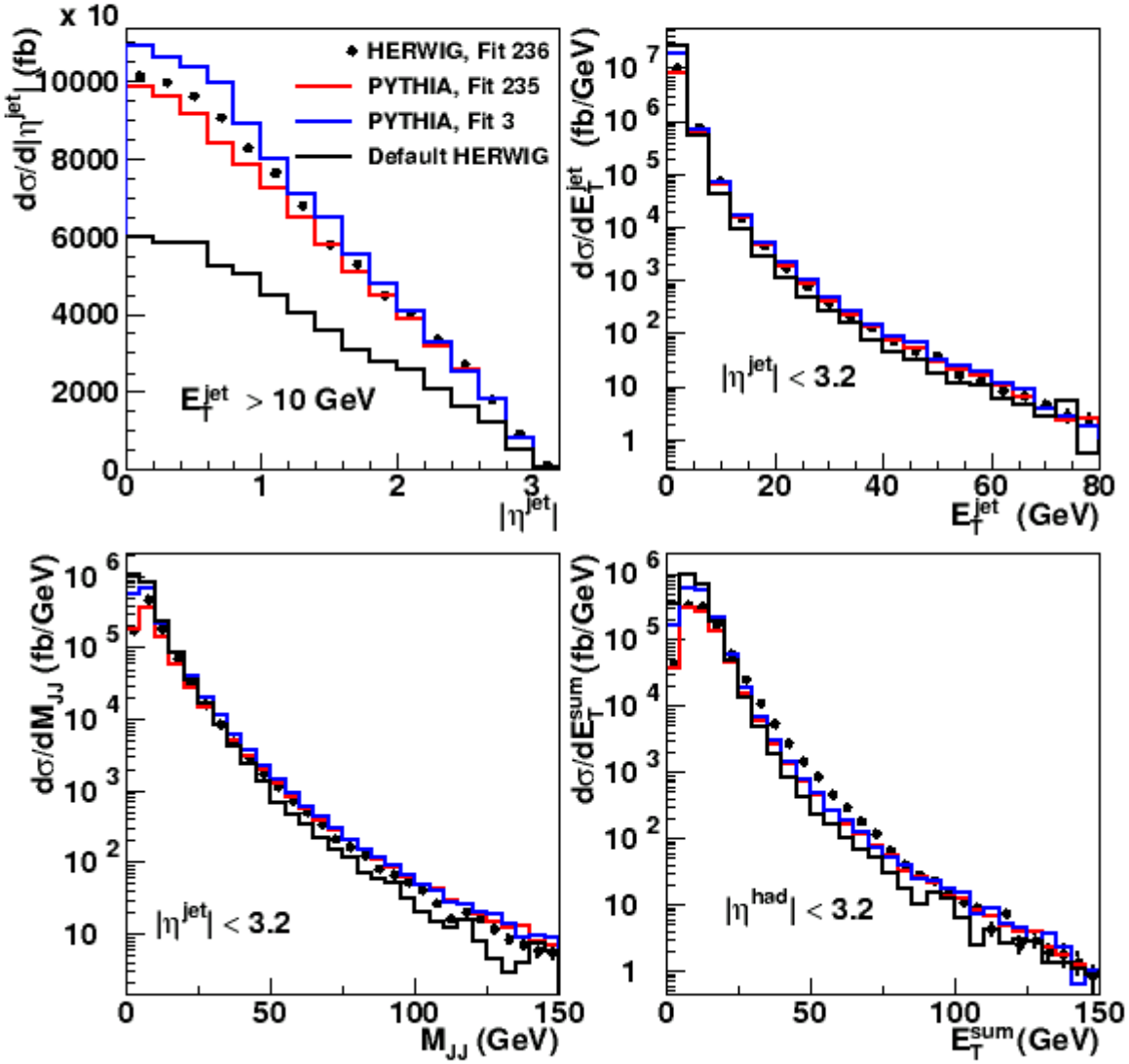
Again fits give similar results.

Spread is also reduced,
NB. predictions from two MCs

Significant differences to "default" prediction.

Predictions of QCD background known to much better accuracy.

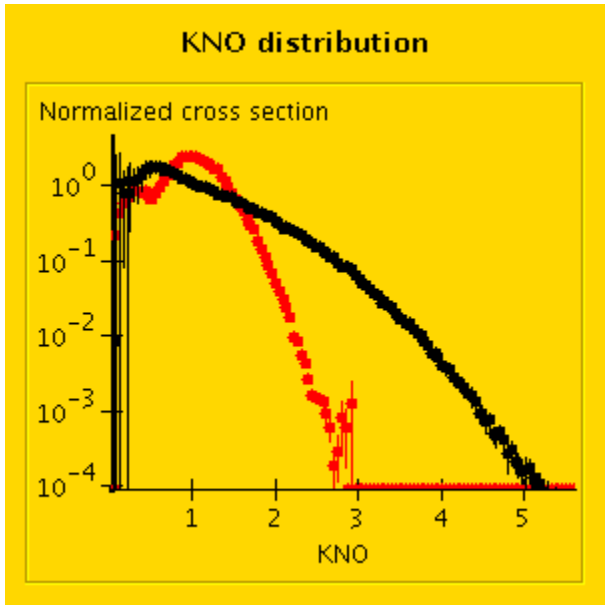
These MC settings should be used in QCD background estimates for and colliders.



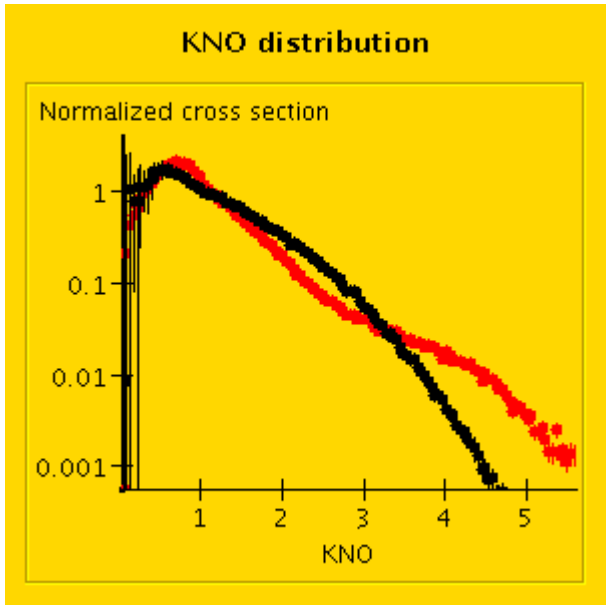
LHC underlying event tune

- Craig Buttar, Ian Dawson, Arthur Moraes (Sheffield)
- Jon Butterworth, Ben Waugh (UCL)
- Starting with Pythia, but may try other generators later
- Varying e.g.
  - ptmin and energy dependence
  - proton matter distribution (default, Gaussian, double Gaussian, different core sizes)
- Using data including
  - multiplicity (KNO) etc. from Tevatron and SPS
  - various high-pt cross sections for normalization and to check consistency

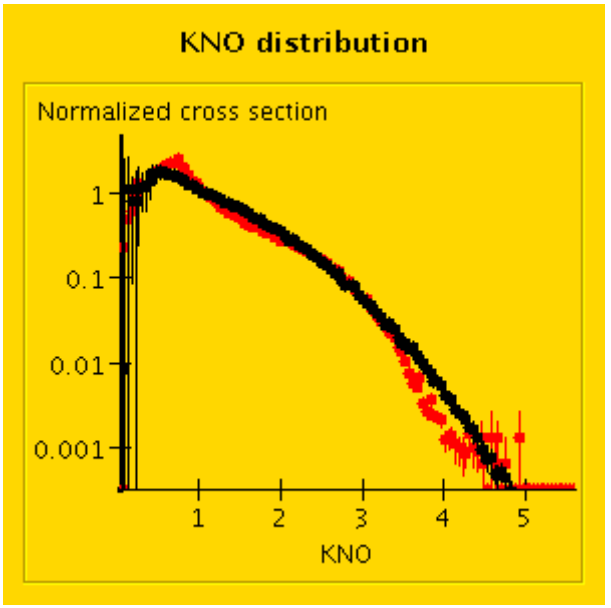
E735 KNO distribution



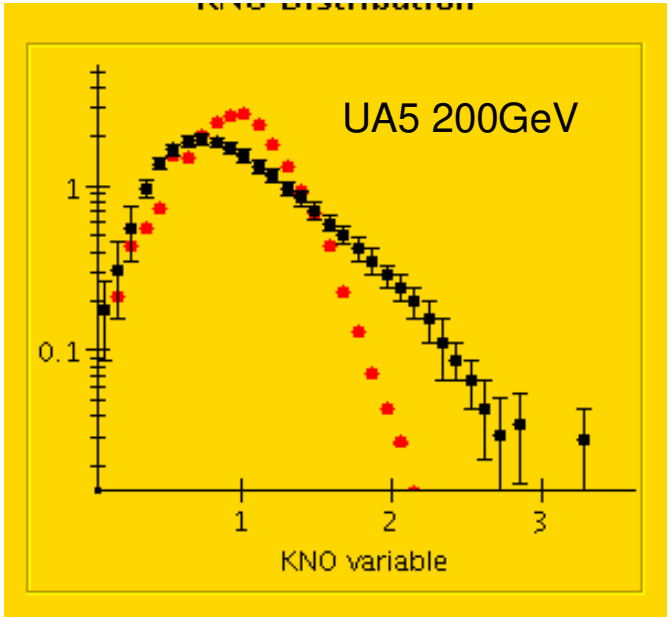
Default Pythia

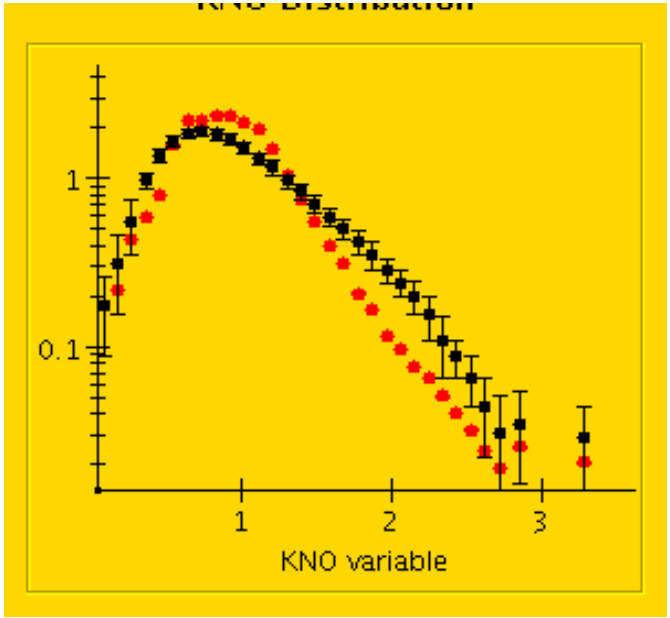


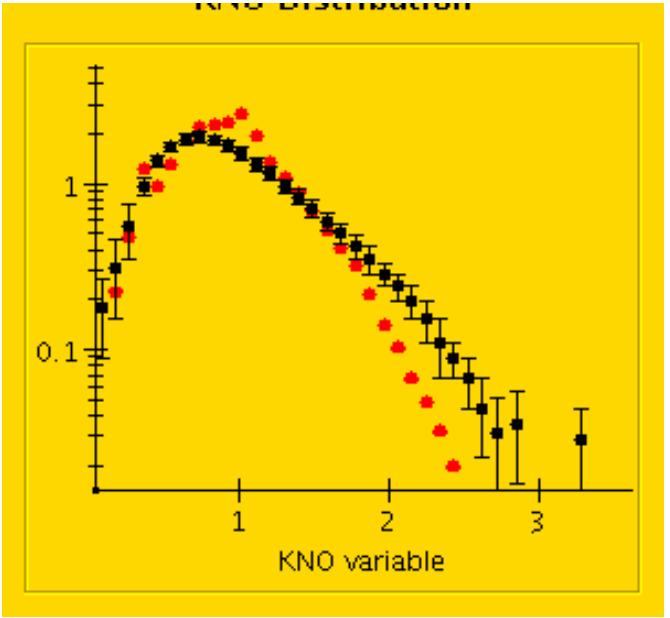
Default double Gaussian



AM tune



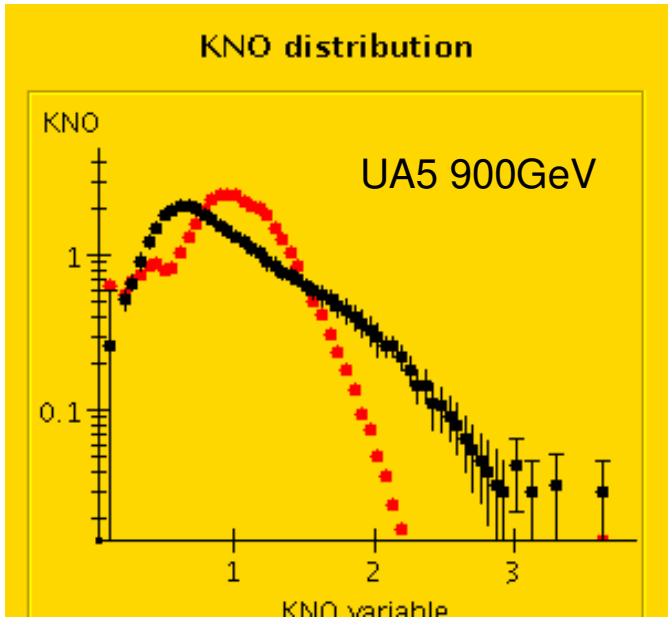


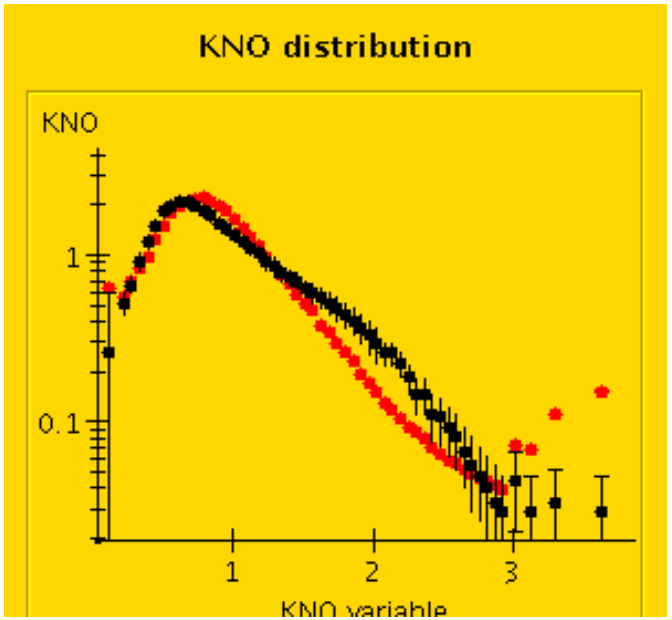


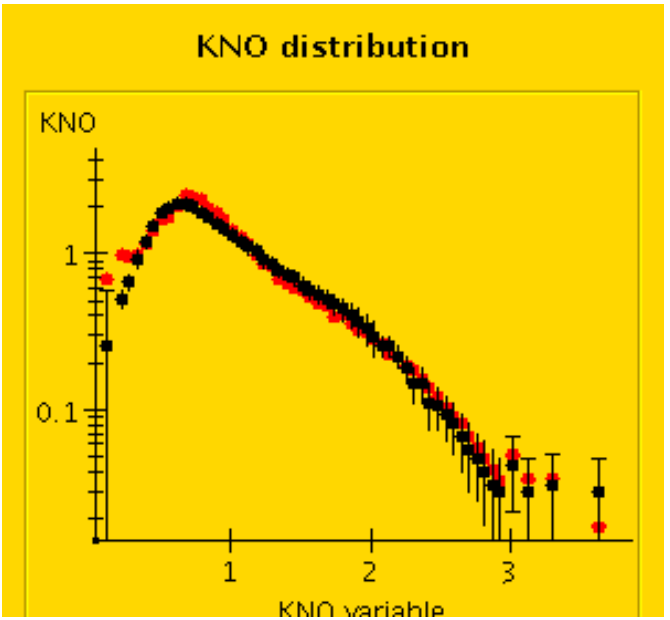
Default Pythia

Default double Gaussian

AM Tune







Development

- Add models:
  - NLO (Ridolfi and Frixione calculation)
  - Pythia 6.3, Pythia 7, Herwig++, etc.
- Integrate with HEPDATA database
  - HEPDATA being converted to RDB
  - make available as more general (grid) resource
- Update HzTool using OO methods
- More use of grid job submission
- More data!

Conclusion

- Already a usable (and used) tool.
- Development ongoing.
- Need more data, i.e. more HzTool routines.
- User feedback is also appreciated (+ and -).