VALOR DUNE ND analysis for the 2nd pass-through

Costas Andreopoulos, *Steve Dennis*, Lorena Escudero

DUNE ND TF Monthly
April 1, 2016

UNIVERSITY OF LIVERPOOL

Summary

- The VALOR Interaction Systematic Model
- Validation of the current VALOR DUNE ND fit:
  - The current version resembles what will be used for the second pass-through.
- VALOR in the DUNE ND Tool-chain:
  - What we expect as input.
  - What we intend to provide as output.

The VALOR Cross-section Model

1. **Cover all modelling aspects**

Studies to define binning ongoing

- 6 correlated CCQE parameters, 3 kinematic bins each for Nu/NuBar.
- 12 correlated CC1pi parameters 3 kinematic bins each for Nu/NuBar and charged/neutral pions.
- 1 MEC normalisation systematic
- 1 Other CC resonance (eg 1-gamma) systematic
- 6 CC DIS (>1 pion) systematics, 3 kinematic bins each for Nu/NuBar
- 1 NuE/NuMu normalisation systematic
- 1 Nu/NuBar normalisation systematic
- 1 NC normalisation
- 1 Coherent normalisation
- 1 Pion and Nucleon mean free path systematic
- 1 Pion charge exchange fraction systematic
- 1 Pion absorption & multi-nucleon knockout systematic
- 1 Pion inelastic fraction systematic
- 1 Hadronization systematic for events containing Etas etc

Total: 37 systematics

Not final but a comprehensive list for the 2nd pass-through, even more complexity can be added later

Prefit Cross-Section Uncertainty Generation with GENIE

Considering 40 interaction modelling systematics, we plan to **feed into the VALOR DUNE/ND analysis a 40 x 40 matrix of interaction modelling pre-fit errors**

Tweak internal model-dependent parameters in GENIE

- MA
- MV
- EB
- kF
- etc

Using simple GENIE reweight tools →

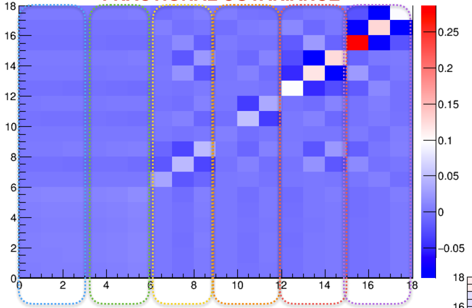
Translate into effect in the correlated model-independent parameters used in the VALOR/ND fit

- 6 correlated CCQE parameters, 3 kinematic bins each for Nu/NuBar.
- 12 correlated CC1pi parameters 3 kinematic bins each for Nu/NuBar and charged/neutral pions.
- etc

Construct covariance matrix by tweaking all internal parameters at the same time (using a normal distribution with their 1σ error inside GENIE), reweighting the number of events to obtain the effect in terms of the model-independent parameters

Example Prefit Error Matrix

FRACTIONAL COVARIANCE



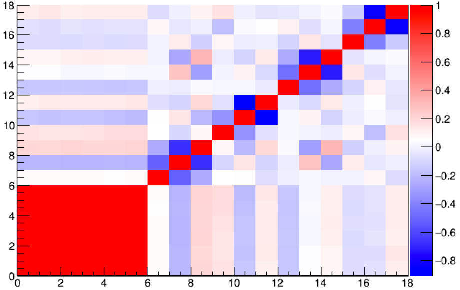
0: ν CCQE y_{reco} (bin1)
1: ν CCQE y_{reco} (bin2)
2: ν CCQE y_{reco} (bin3)
3: anti ν CCQE y_{reco} (bin1)
4: anti ν CCQE y_{reco} (bin2)
5: anti ν CCQE y_{reco} (bin3)
6: ν CC1piC y_{reco} (bin1)
7: ν CC1piC y_{reco} (bin2)
8: ν CC1piC y_{reco} (bin3)
9: anti ν CC1piC y_{reco} (bin1)
10: anti ν CC1piC y_{reco} (bin2)
11: anti ν CC1piC y_{reco} (bin3)
12: ν CC1pi0 y_{reco} (bin1)
13: ν CC1pi0 y_{reco} (bin2)
14: ν CC1pi0 y_{reco} (bin3)
15: anti ν CC1pi0 y_{reco} (bin1)
16: anti ν CC1pi0 y_{reco} (bin2)
17: anti ν CC1pi0 y_{reco} (bin3)

y_{reco} binning different for each category and flavour

Quick test!

- With simple MC events (25k ν+ 25k anti-ν) with 0-120 GeV
- Just hundred tweaks of only MA CC QE and MA CC RES
- In a preliminary binning in y_reco (smeared with a 10% resolution from true values)

CORRELATION



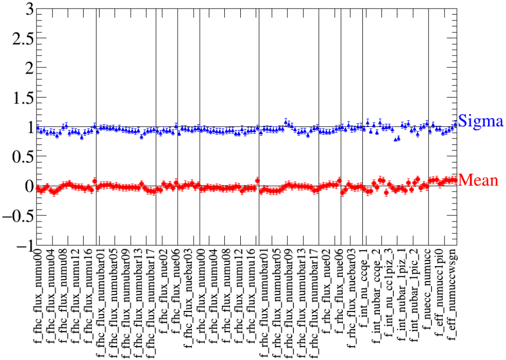
Bias Studies

- We want to prove that the VALOR ND studies give accurate constraints on the OA systematics.
- To do this:
  - Generate N toy experiments (full sets of randomised systematics). ($N \sim 200$)
  - Fit the dataset generated given that set of systematics (without statistical fluctuations).
  - Define ‘pull’ variable that over many fits should form a Gaussian with mean 0 and sigma 1.
- Definition of pull for constrained parameter f in fit i with initial parameter central value a_f and best-fit value $x_{f,i}$:

$$P_{f,i} = \frac{x_{f,i} - a_f}{\sqrt{\sigma^2_{f,prior} - \sigma^2_{f,i,fitted}}} \tag{1}$$

- For more info on our pull definitions, see L. Demortier & L. Lyons - CDF/ANAL/PUBLIC/5776 (PDF).

Biases - 21 Interaction Systematics.



Looks good!

- This uses 21 interaction systematics, 104 flux systematics, 9 detector systematics (all normalisations.)
  - 6 CCQE bins, 12 1π bins, 1 CC-other, 1 NuE, 1 NC.
  - The bins are separated ν and $\bar{\nu}$ in 3 kinematic bins (E_{true} 0-2 GeV, 2-4 GeV, 4+ GeV).
  - These is a placeholder binning. We'll use a better kinematic variable (Q^2) soon.
- For space reasons, only 1 in 4 systematics is labelled. You should able to interpolate to identify most others.

What we'd like as input.

- For us to do everything we need to do, we need MC provided in Trees with:
  - GENIE event record (where we'll get our true kinematics, mode etc).
  - Reconstructed neutrino energy (E_{reco}) - your best estimate.
  - Reconstructed inelasticity $y_{reco} = \nu/E$ - your best estimate.
  - Sample ID - which VALOR sample the events fit into using your best selection.
  - Arbitrary weight (separate tree?)
- **We'd like example files from the ND groups ASAP so we can make sure can we get everything we need from what you produce.**
- We'll also need the POT normalisation of our input MC files.
- We can also accept input covariance matrices for the following groups of parameters:
  - Flux (currently using MINERνA 104 parameter matrix).
  - Interaction. We currently generate this in-house with GENIE but could accept external ones.
  - Detector (whatever parameters we eventually need).

What we plan to produce as output.

For any given dataset, we should be able to produce:

- Prefit covariance matrix (containing unconstrained parameter correlations and uncertainties).
- Post covariance matrix (containing constrained parameter correlations and uncertainties).
- Vector of tuned (postfit) central values suggested for the far detector fit.
- If necessary, additional error matrix to represent biases from our method.
  - We'll be trying our utmost to ensure we have no biases that require us to inflate errors like this.
  - But we should be aware that it could become necessary.
- A goodness-of-fit measure (single figure - P-value?).

Overview

- Currently have a well-validated analysis with 21 normalisation interaction systematics.
- The full 37-parameter model should be feasible for the second run-through.
  - But if we don't get that validated properly in time, we have a good fall-back position.
- We need inputs! That's our biggest issue at this point.
- For the third run-through, we'll move to better detector systematic models, and smarter marginalisation of those uncertainties.

Backup

Current VALOR Sample List

- ν_μ CCQE 1-track. (code $= 1$)
- ν_μ CCQE 2-track. (code $= 2$)
- ν_μ CC $1\pi^\pm$. (code $= 3$)
- ν_μ CC $1\pi^0$. (code $= 5$)
- ν_μ CC $1\pi^0 + 1\pi^\pm$. (code $= 6$)
- ν_μ CC other. (code $= 7$)
- ν_e inclusive. (code $= 8$)
- ν_μ CC wrong-sign inclusive. (code $= 9$)
- NC inclusive. (code $= 10$)

Unused codes in current analysis (CC-inclusive $= 0$, CC $2\pi^\pm = 4$).