canonicalH3 Model Evaluation