

# Benchmarking Simulations against the ThermoML Database: Neat Liquid Densities and Static Dielectrics

Kyle A. Beauchamp<sup>\*,1</sup> Julie M. Behr<sup>\*,1</sup> Patrick B. Grinaway,<sup>1</sup> Bas,<sup>1</sup>  
Kenneth Kronlein,<sup>2</sup> Michael R. Shirts,<sup>3,\*</sup> and John D. Chodera<sup>1,†</sup>

<sup>1</sup>Memorial Sloan-Kettering Cancer Center, New York, NY, USA

<sup>2</sup>NIST

<sup>3</sup>Department of Chemical Engineering, University of Virginia, Charlottesville, VA 22094-0471

(Dated: December 3, 2014)

## I. ABSTRACT

Useful atomistic simulations require accurate depictions of solvent. Simple experimental observables, such as density and static dielectric constants, offer straightforward targets for evaluating forcefield quality. Here we examine the possibility of benchmarking atomistic models against the NIST ThermoML database of physicochemical measurements, which curates thousands of density, dielectric, and other measurements. We present a detailed benchmark of the GAFF forcefield against measurements extracted from ThermoML and discuss the extent of available data for neat liquids. We show that empirical polarizability models correct systematic biases inherent in predicting dielectric constants with fixed-charged forcefields. Combining our dataset with the Virtual Chemistry benchmark set provides an extensive benchmark suite for liquid properties.

## II. INTRODUCTION

Intro

## III. RESULTS

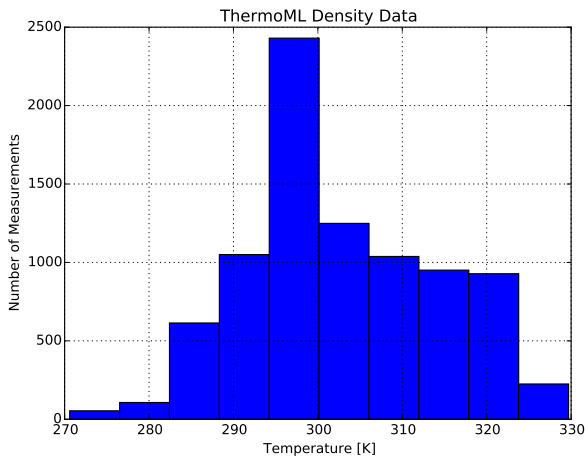
### A. Neat Liquid Measurements in ThermoML

To assess the feasibility of benchmarking organic molecule forcefields against ThermoML, we performed a number of queries to summarize the data content of ThermoML. Our aim is to explore neat liquid data with functional groups relevant to drug-like molecules. We therefore applied the following series of filters: Druglike elements (H, N, C, O, S, P, F, Cl, Br), Heavy Atom Count ( $\leq 10$ ), Temperature [K] ( $270 \leq T \leq 330$ ), and Pressure [kPA] ( $100 \leq P \leq 102$ ). After applying these filters, we also assume that all pressures within this range are one atmosphere. We also assume that temperatures can be rounded to one decimal place. These approximations

Filter	Mass Density	Static Dielectric
0. Full	308248	4170
1. Druglike Elements	48909	847
2. Heavy Atoms	35465	773
3. Temperature	18374	609
4. Pressure	8247	399
5. Aggregate T, P	3396	382

TABLE I. ThermoML Statistics

are motivated by common data entry errors; for example, an experiment performed at water’s freezing point at ambient pressure might be entered as either 101.325 kPA or 100 kPA, with a temperature of either 273 K or 273.15 K.



### B. Benchmarking GAFF against ThermoML: Mass Density

Mass density has been widely used as a critical ingredient for parameterizing and testing forcefields, particularly the Lennard Jones parameters. We therefore used the present ThermoML compilation as a benchmark of the Generalized Amber Force Field.

### C. Benchmarking GAFF against ThermoML: Static Dielectric

As a measure of the electronic medium, the static dielectric constant of neat liquids provides a critical bench-

\* michael.shirts@virgina.edu

† jchodera@mskcc.org

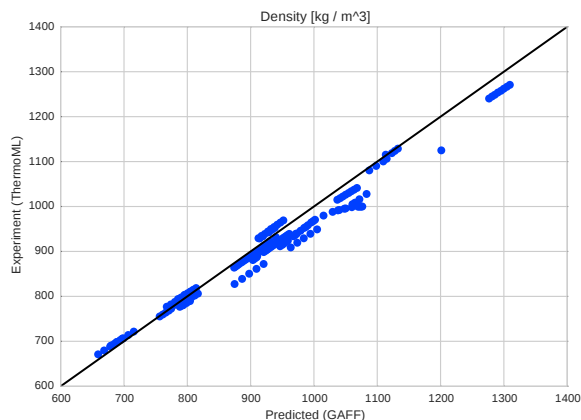
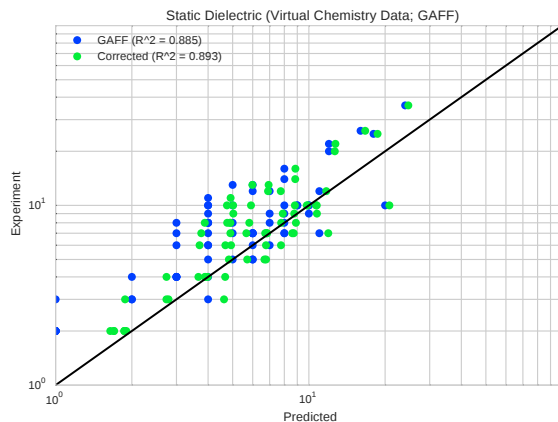
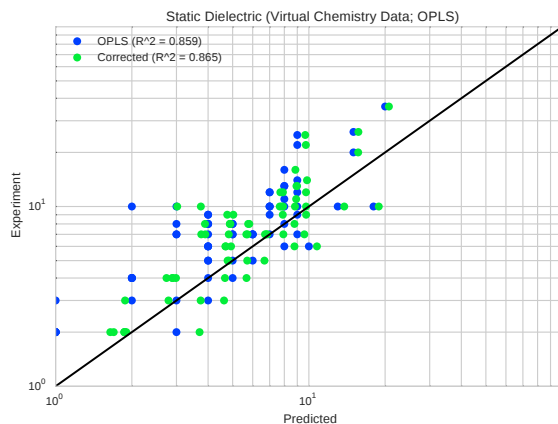


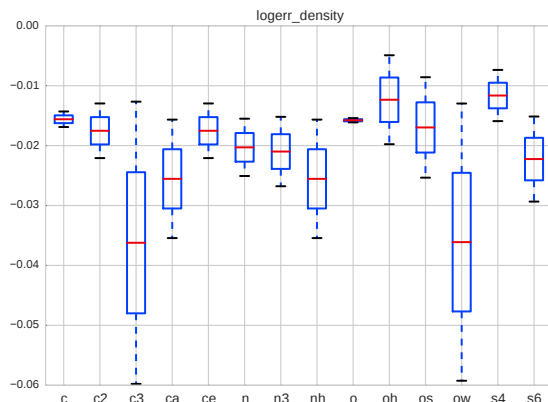
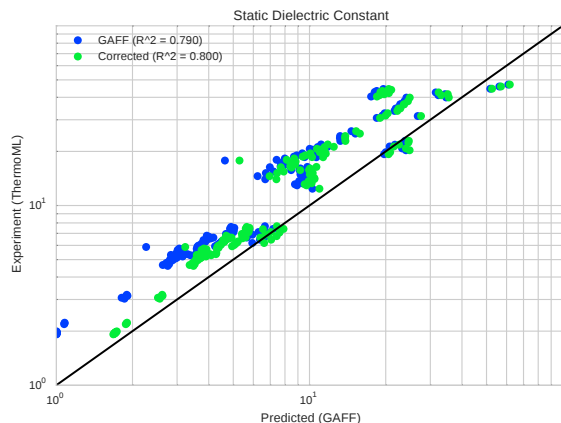
FIG. 1. Density Benchmark, ThermoML+GAFF

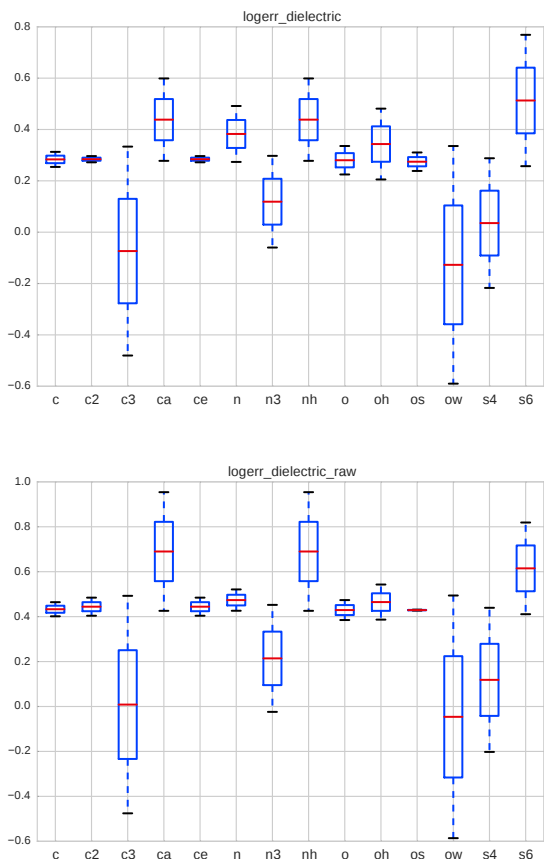
mark that is orthogonal to density and thermodynamic quantities. We therefore compare our GAFF simulations against the measurements in our ThermoML compilation. Overall, we find the dielectric constants to be qualitatively reasonable, but with clear deviations from experiment. In particular, the nonpolar organics show a clear discrepancy, with the MD predictions of 1.0 being substantially less polar than the measurements near 2.0. Because this deviation likely stems from the lack of electronic polarization, we added a simple empirical correction for polarization [? ], which leads to better agreement with experiment. A similar polarization correct was used in the developemnt of the TIP4PEW water model; however, the need is much greater for the non-polar organics, as the missing polarizability is the dominant term in the static dielectric constant. For comparison, we also applied the same empirical correction to the VirtualChemistry dataset and saw similarly improved agreement with experiment for both the GAFF and OPLS forcefields.



## IV. DISCUSSION

### A. Forcefield Accuracy Depends on Functional Group





### B. Fitting Forcefields to Dielectric Constants

Recent forcefield development has seen a resurgence of papers fitting dielectric constants as primary data [? ?]. However, a number of authors have pointed out potential challenges in constructing self-consistent fixed-charge force fields [? ]. Interestingly, a recent work by Dill [? ] pointed out that, for CCL4, reasonable choices of point charges are incapable of recapitulating the observed dielectric of 2.X, instead producing dielectric constants in the range of  $1.0 \leq \epsilon \leq 1.05$ . Suppose, for example, that one attempts to directly fit the static dielectric constants of CH<sub>4</sub>, CClH<sub>3</sub>, CCl<sub>2</sub>H<sub>2</sub>, CCl<sub>3</sub>H, CCl<sub>4</sub>. In moving from the tetrahedrally-symmetric CCL<sub>4</sub> to CCl<sub>3</sub>H, it suddenly becomes possible to achieve the observed dielectric constant of 2.X. However, in the model for CCl<sub>3</sub>H, the fixed point charges are overpolarized and account for two sources of dielectric: the

net dipole moment and the (electronic) polarizability. We hypothesize that this inconsistency in parameterization may lead to strange mismatches, where symmetric molecules (e.g. benzene, CCL<sub>4</sub>) have qualitatively different properties than closely related asymmetric molecules (e.g. toluene, CHCl<sub>3</sub>). As a first-order fix, we suggest using empirical polarization corrections before directly comparing measured static dielectric constants to fixed-charge models—particularly when examining low-dielectric solvents.

### C. ThermoML as a Data Source

Pro: Automated, Curated, Growing, Free, Paper trail  
 Cons: Requires Parsing, Data Entry Errors

## V. METHODS

### A. Simulation

Boxes of 1000 molecules were constructed using PackMol [?]. AM1-BCC charges were generated using the OpenEye toolkit, using the XYZ module with parameter ZYX. The selected conformer was then processed using antechamber in AmberTools 14. The resulting AMBER files were converted to OpenMM [?] XML files. Simulation code used libraries gaff2xml (rev), trustbutverify (rev), openmm 6.2, and MDTraj [?] 1.2.

Molecular dynamics simulations were performed in the NPT ensemble using a Langevin integrator (friction 1 / picosecond?) and a Monte Carlo barostat. Particle mesh ewald [?] was used with a long-range cutoff of 1.0??? nanometers. Simulations were continued until density standard errors reached N decimals of precision, as estimated using the equilibration detection module in pymbar 2.1 [?]. Trajectory analysis was performed using OpenMM [?] and MDTraj [?].

## VI. CONCLUSIONS

1. ThermoML is a potentially useful resource for the forcefield community 2. We have curated a subset of ThermoML for neat liquids with druglike atoms, with thousands of densities and hundreds of dielectrics 3. Empirical polarization models correct a systematic bias in comparing fixed-charge forcefields to static dielectric constants