MATERIALS SIMULATION

The shape of water

Credit: krystiannawrocki/Getty

Water is a validation system for benchmarking and demonstrating the method

A combination of density functional theory (DFT), machine learning and statistical methods can be applied to make accurate predictions of the thermodynamic properties of water, as Bingqing Cheng and colleagues report in the *Proceedings of the National Academy of Sciences of the USA*.

Water provides an excellent model system for computational methods because most of its properties are well known, but for many there is no satisfying theoretical explanation. Indeed, despite its simple chemical formula, predicting the thermodynamic properties of water is all but trivial. For example, water crystallizes into two forms of ice — one with a hexagonal crystal structure and one with a cubic crystal structure — and their relative stability is very difficult to measure experimentally.

Predicting the properties of a material starting from the laws of quantum mechanics is a theorist's dream, but *ab initio* methods are computationally expensive, especially when many atoms need to be modelled at finite temperatures. To correctly account for the properties of water, the quantum mechanical nature of both electrons and nuclei have to be taken into account, together with the entropy of a complex system with many degrees of freedom. Thus, there is a trade-off between the accuracy of the description and the computational cost

associated with sampling all the relevant atomistic configurations.

Cheng and co-workers used a combination of DFT calculations, machine learning and statistical methods to substantially increase the accuracy of the results while keeping the calculations computationally manageable. Their DFT calculations provide a quantum mechanical description of the water potential energy surface, whereas the machine-learning algorithm enables the extensive exploration of the phase space by learning the atomic interactions and thus avoiding prohibitively expensive calculations that explicitly consider electrons. The efficiency of the combination enables a quantum mechanical description of the nuclei, proton disorder and anharmonic quantum nuclear fluctuations. “When I became aware of the possibility of using machine-learning techniques to mimic the quantum mechanical interactions between atoms, it struck me as a free lunch,” says Cheng. “It is an ideal tool to squeeze out more information from electronic structure calculations and expand the power of quantum mechanical methods.”

Interatomic potentials obtained with machine learning inevitably contain small errors compared with the real potential. Therefore, to promote the results to the *ab initio* level of theory, the properties of a few hundred representative

configurations were re-computed using DFT, and the residual small error of the machine-learning interatomic potentials were calculated and then corrected using state-of-the-art statistical methods.

This combination of methods enables the rigorous computation of the relative stability of the two forms of ice, as well as of hexagonal ice and water. Apart from closely reproducing many properties of water, the results also shed light on the physical origin of phenomena, such as the density anomaly of water, the floating of ice, the difference between the melting point of heavy and normal water and the six-fold symmetry of snowflakes.

“Water is a validation system for benchmarking and demonstrating the method,” explains Cheng. “Going forward, we will apply this combination of machine learning, quantum mechanics and statistical mechanics to predict the thermodynamic properties of other technologically relevant systems, including novel materials that haven't been experimentally synthesized yet.” Pharmaceutical compounds, materials for hydrogen storage, metallic alloys and hydrocarbons are among the systems to which this combination of methods can be applied.

Giulia Pacchioni

ORIGINAL ARTICLE Cheng, B. et al. *Ab initio* thermodynamics of liquid and solid water. *Proc. Natl Acad. Sci. USA* **116**, 1110–1115 (2019)