Building Water Models Compatible With Charge Scaling Molecular Dynamics

Victor Cruces Chamorro, Pavel Jungwirth,* and Hector Martinez-Seara*

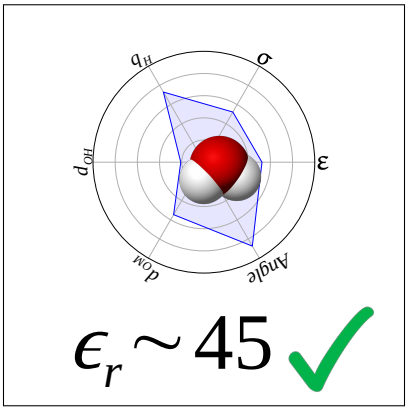
Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo nám. 2, 16610 Prague 6, Czech Republic

E-mail: pavel.jungwirth@uochb.cas.cz; hseara@gmail.com

Abstract

Charge scaling has proven to be an efficient way to account in a mean-field way for electronic polarization by aqueous ions in force field molecular dynamics simulations. However, commonly used water models with dielectric constants over 50 are not consistent with this approach leading to 'overscaling', i.e., generally too weak ion-ion interactions. Here, we build water models fully compatible with charge scaling, i.e., having the correct low-frequency dielectric constant of about 45. To this end, we employ advanced optimization and machine learning schemes in order to explore the vast parameter space of 4-site water models efficiently. As an a priori unwarranted positive result, we find a sizable range of force field parameters that satisfy the above dielectric constant constraint providing at the same time accuracy with respect to experimental data comparable with the best existing 4-site water models such as TIP4P/2005, TIP4P-FB, or OPC4. The present results thus open the way to the development of a consistent charge scaling force field for modelling ions in aqueous solutions.

TOC Graphic



Water molecules are ubiquitous in living systems and technological applications due to their physicochemical properties that make water a unique universal solvent.[1] Water thus provides an environment where life and chemistry take place by dissolving molecules and ions, allowing specific molecular and supramolecular structures, and directly contributing to stabilizing interactions and catalyzing reactions.

Force field molecular dynamics simulations (FFMD) represent a powerful tool for modeling these biological and technological processes with atomistic resolution at femtosecond to millisecond timescales. First simulations involving water data back to the early days of FFMD.[2] Consequently, the development of empirical potentials for water has been a recurrent topic in the past decades (e.g, TIPS,[3] SPC,[4] TIP3P,[5] SPC/E,[6] and TIP4P[7] models), and far from settled[8] (e.g., the more recent TIP4P/2005,[9] TIP4P-FB,[10] and 4-site OPC[11] (OPC4) models). Aqueous solutions have proven to be difficult systems to describe accurately and are thus an active area of research.[12] Even pure water behavior is not easy to model such that it accurately covers the full range of biologically relevant thermodynamic conditions.[8]

Commonly used water potentials were typically optimized to recover selected experimental or calculated data. Therefore, they reproduce these target properties at the optimization conditions, but there is no guarantee that they will also reproduce other properties or the target properties at different thermodynamic conditions. The

optimization process traditionally focuses on properties derived from the density[13] and the self-diffusion coefficient,[14] while other properties, such as the surface tension or the dielectric constant,[15] are given a secondary role or not optimized at all. It is thus not surprising that their values vary significantly between existing models[11,16] despite their physical relevance.[17]

In particular, the dielectric constant (ε_r) is an essential property dictating how interactions between charged particles are attenuated in a given medium. The dielectric constant can be approximately split into two contributions of different origins.[18]

$$\varepsilon_r = \varepsilon_N \varepsilon_e \tag{1}$$

The nuclear contribution to the dielectric constant (ε_N) accounts for the "slow" rearrangement of atomic nuclei of water molecules as a response to changes in local or external electromagnetic fields. In contrast, the electronic contribution to the dielectric constant (ε_e) accounts for the "instantaneous" response of the electronic clouds of the water molecules and can be approximated by the square of the refraction index[19] ($\varepsilon_e \approx n^2 = 1.78$).[20]

FFMD that lack polarization terms account only for the nuclei contribution of the dielectric response of the medium. One could potentially employ the computationally more demanding polarizable force fields such as Drude[21] or Amoeba[22] to capture the electronic contribution of the response. As an alternative, one can introduce the missing electronic polarization in a mean-field way denoted as the electronic continuum correction (ECC).[20,23,24] Within this approach, the system is immersed in an electronic dielectric continuum, which is mathematically equivalent to scaling the ionic charges by the inverse square root of the electronic part of the dielectric constant of the medium ($q/\sqrt{\varepsilon_e}$).

The ECC framework circumvents the problem of explicitly accounting for electronic polarization for interactions between dissolved ions or charged groups. There is, however, a catch – existing non-polarizable water models often exhibit values of dielectric constants larger than ε_N, effectively transferring (part of) the missing ε_e to ε_N. They also possess water dipole moments larger than the gas phase value (albeit typically smaller than the value in the liquid).[25] Employing currently available water models thus results in an artificial overscaling when used within the ECC approach.[23]

Within this study, we succeeded in developing a class of 4-site water models compatible with the ECC approach (i.,e, possessing $\varepsilon_r \approx 45$), which are comparable in predicting experimental observables to the best of the existing 4-site water models (possessing significantly larger values of ε_r). Considering the above constraint of a low dielectric constant, it was not clear from the onset whether such a model can be developed.

Our target 4-site water models are fully defined by 6 parameters, see Table 1. Similarly, as in the TIP4P family of models, these are the Lennard-Jones parameters (i.e., σ and ε) on the oxygen atom (with no explicit van der Waals terms on the hydrogens), the charge on each of the hydrogen atoms (q_H) (that also defines the charge on the dummy atom $q_M = -2q_H$) and the intramolecular parameters. Namely, these are the oxygen–hydrogen (d_{OH}) and oxygen–dummy atom (d_{OM}) distances and the hydrogen-oxygen-hydrogen angle (θ). Note that the dummy atom is placed at the bisector of the angle θ in the direction toward the hydrogen atoms.

Table 1: Optimized parameters with boundaries and seeding values.

Parameter	Units	Boundaries	Initial
σ	nm	0.3050–0.3250	0.3150
ε	kJ/mol	0.5000–1.0000	0.7500
q_H	e^-	0.3500–0.7000	0.5500
d_{OH}	nm	0.0900–0.1000	0.0960
d_{OM}	nm	0.0120–0.0180	0.0150
θ	deg	100.00–110.00	105.00

On the technical side, developing an empirical force field is a computationally expensive and time-consuming endeavor, primarily due to the large number of simulations required for testing extensive sets of parameters. For us to effectively tackle water force field development, we need a framework that reduces the number of simulations ultimately performed while still being able to localize the optimal regions of the parameter space. To this end, we have developed an automated framework that efficiently avoids sampling sub-optimal regions of parameter space using a combination of artificial intelligence (AI) tools and other advanced optimization methods (Figure 1).

To avoid any bias and to critically evaluate our developed framework for sampling the parameter space, we do not explicitly assume any concrete relationship between parameters and target properties when starting the optimization process (although such constraints could be easily incorporated). Under such conditions, random walkers (RW) are useful for an initial sampling of the parameter space and for gathering information about their relationship with target physical properties. Additionally, RW improves simulation stability because it uses the previous step molecular configuration as a starting point for new simulations. This is an important feature when using an automatic framework. RW simulations started from the parameters presented in Table 1. In addition, boundaries were set to keep the water geometry and physical properties within reasonable limits, see Table 1. The resulting parameter space is large enough to encompass both good as well as less optimal regions without enforcing initial biases while simultaneously avoiding sampling of physically unreasonable regions. In particular, to sample the present 6-dimensional space, we performed using RW 1000 simulations at 300 K and 1 bar, exploring a relatively wide ε_r range.

Once the parameter space is sparsely sampled by the above approach, a second phase begins where we optimize the process of generation of parameters using the differen-

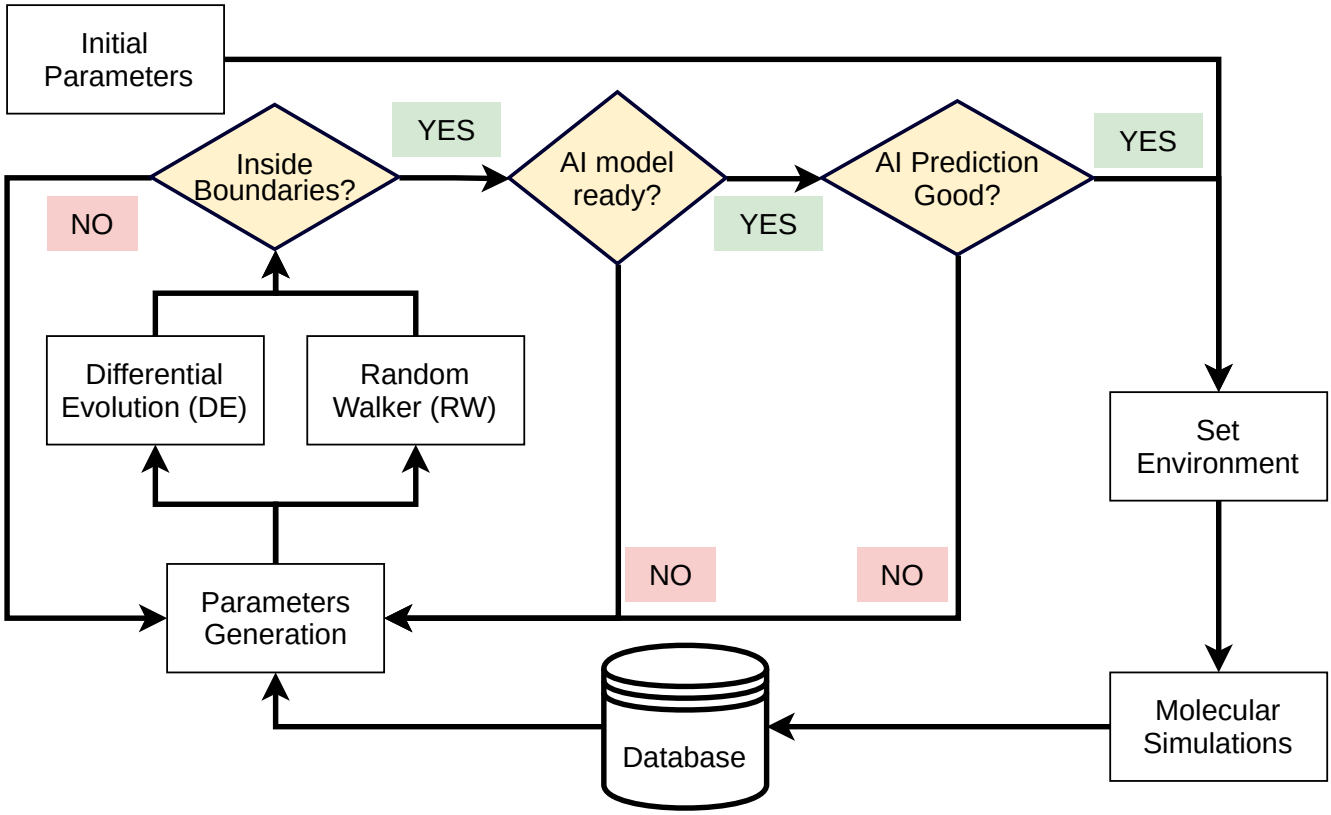


Figure 1: Scheme of the program routine used to generate new parameters.

tial evolution (DE) algorithm.[26] This algorithm generates new parameter sets or points as a linear combination of the parameters from the best points of the available population, i.e., the previously obtained points. Such a method efficiently parallelizes the optimization process while simultaneously improving the sampling capacity, which is crucial when dealing with high-dimensional problems such as force field development. The price to pay is that such parameter generation, being stochastic in nature, does not ensure that the new candidate is necessarily better than the existing points. Only when the generated candidate improves the quality of the parameters in terms of the accuracy of the simulated target properties is used in subsequent steps by DE. The optimization process is finished once the population of parameters has reached the desired convergence. In this work, this corresponds roughly to 4500 sets of parameters. Considering that one needs to ultimately test the generated parameters by performing FFMD and that DE may generate (particularly at the beginning of the optimization process) points which far from optimal regions, there is a need for further streamlining the whole process.

The parameter convergence can be significantly accelerated, i.e., the number of simulations needed to be performed can be reduced if we introduce a method that estimates the output results for the DE suggested parameter sets or points without actually running the simulations. As shown in Figure 1, we can use a mapper function to predict the outcome of the candidate such that if the predicted outcome is worse than a pre-defined value of the target cost function, the program skips the actual simulation and directly generates a new candidate. The mapper function used in this work is a fully connected multilayer neural network. The input layer vector contains all our parameters normalized from 0 to 1. The ReLu activation function is used in the four hidden layers connected by a dropout layer with a rate of 0.10, each layer having 40 nodes which cannot have a bigger norm than 5.0. A linear activation function is used for the output layer. Finally, the neural network is trained using early stopping such that we avoid possible overfitting while conserving the prediction capacity of the neural network.[27]

In this work, we build the neural networks used as mapping functions employing the data obtained from all simulations performed so far. As creating neural networks is very fast compared to performing simulations, they are recreated whenever new data is available, i.e., when new simulations are performed. While initially the neural network's performance is not yet optimal, even at this point, it is often sufficient to discriminate bad points. Also, the fact that good points are occasionally wrongly rejected does not affect the convergence significantly since these can be sampled at a later time as the prediction capability of the neural network improves upon being trained with an increasing amount of data. More data also reduces overfitting, which would otherwise negatively impact the prediction capabilities of the neural network. Using the finally obtained well-performing neural network, an efficient refinement algorithm described in the Supporting Information was used to increase the sampling capacity further, improving the obtained water models. A point is considered

better than a previous one when it lowers a cost function that expresses the weighted difference between reference and simulation values for our selected target experimental properties, see Table 2. Note that for the diffusion constant D_{OW}, we have scaled the experimental value[14] used for comparison to adjust to the effect of the finite size of the simulated unit cell.[28] For optimization of the dielectric constant, the cost function (C_{ECC}) considers as ε_r only the nuclear contribution to the experimental dielectric constant to be compatible with the ECC approach.[23] Otherwise, ε_r is not included in the cost function (C_G). Our cost function reads as

$$C_X = 10 \sum_{i=1}^{N} f_i \cdot w_i, \tag{2}$$

where f_i and w_i are the loss function and weight for each property, respectively. Here, we use mean absolute percentage error (MAPE), normalized to 1, as a loss function to calculate the deviation between a given simulation property and experiments. Being a percentage-based metric, it is scale-independent, making it useful for comparing the accuracy of properties on different scales. The cost function is a weighted average, see equation 2, where the weights are normalized to sum to 1.

Table 2: Reference properties and functional parameters used in the optimization process. $\rho_{1\mathrm{bar}}$ are density values at 1 bar from 260 K to 360 K every 20 K. ε_r is the relative permittivity according to the ECC approach.[23] D_{OW} is the experimental self-diffusion coefficient of water at 300 K and 1 bar accounting for our simulation of 832 water molecules using Hummer-Yeh periodic boundary conditions correction.[28] rdf_{1p} and rdf_{1h} are the position and height of the first oxygen–oxygen RDF peak. The weights of the properties, ensuring a balanced sampling of all the properties, are normalized to sum to 1. [a] The weights correspond to C_{ECC}.

	Values	Units	Loss function	Weight[a]
$\rho_{1\mathrm{bar}}$	Table S2	kg/m^3	MAPE	0.667
ε_r	44.5	—	MAPE	0.111
D_{OW}	2.16E-5[b]	cm^2/s	MAPE	0.111
rdf_{1p}	0.280	nm	MAPE	0.0555
rdf_{1h}	2.58	—	MAPE	0.0555

All FFMD simulations for the optimization process were performed using the GROMACS2019 molecular dynamics package.[29] The number of water molecules in the cubic simulation box is 832. This number was chosen because it is small enough for an efficient optimization process but large enough (i.e., minimum unit cell size of 2.70 nm) to fulfill the minimum image convention and the corresponding cutoffs. Namely, we employed an interaction cutoff of 1.2 nm for the Particle Mesh Ewals (PME)[30] and the PME Lennard-Jones schemes that take into account the long-range electrostatic and van der Waals interactions. We used the leapfrog algorithm with a timestep of 2.0 fs and a total simulation time of 21 ns. The first nanosecond was considered equilibration and skipped for the analysis. The isothermic-isobaric (NpT) ensemble was enforced using the Nosé-Hoover thermostat[31] with a relaxation time of 1.0 ps and the Parrinello-Rahman barostat[32] with a compressibility of 5E-5 bar^{-1} and a relaxation time of 5.0 ps.

The results of the optimization process are summarized in Figure 2. A total of 1343 parameter sets were generated within the optimization process possessing ε_r values between 40 and 50, i.e., very close to the value of 45 fully compatible with the ECC approach. From these, there is a sizable region in the parameter space with an acceptably small deviation from experiments ($C_{ECC} < 0.7$). This region includes 791 points. For comparison, a widely used 3-site model TIP3P possesses a much larger value of $C_G = 1.973$. To further illustrate the performance of these points, we categorize them in two additionally constrained regions with $C_{ECC} < 0.5$ and $C_{ECC} < 0.3$. As discussed below, the latter corresponds to models with performance comparable to that of current state-of-the-art 4-site water force fields. The optimal ECC water force field region with $C_{ECC} < 0.3$ occupies a well defined region of parameters $\sigma \approx [0.315 - 0.316]$ nm, $\varepsilon \approx [0.65 - 0.825]\ kJ/mol$, $q_H \approx [0.51 - 0.64]\ e^-$, $d_{OH} \approx [0.90 - 1.0]$ nm, $d_{OM} \approx [0.135 - 0.180]$ nm, and $\theta \approx [106^\circ - 110^\circ]$. Also, note that the 50 best-performing models are spread fairly evenly in this optimal region. This suggests a rather flat cost-optimal region in the parameter space compatible with ECC. An extended view of the sampled parameter space as a function of the resulting cost function is presented in Figure S2.

To contextualize our optimal region, we compare its performance to that of existing state-of-the-art 4-site water models simulated under the same conditions (the empty black symbols in Figure 2 correspond to TIP4P/2005 (○), OPC4 (□), and TIP4P-FB (△)). Two of these water models (i.e., TIP4P/2005 and TIP4P-FB) are of a fixed gas phase geometry, while OPC4 and our models optimize the water geometry parameters (see Table 3 for a complete list of their parameters). Note that σ and ε values and the charges q_H of all these water models fall within a narrow region for $C_{ECC} < 0.7$, which seems to be highly preserved (especially for σ) for water models.[16] The bond parameters d_{OH} and d_{OM} of these models also fall within the optimal region with the exception of d_{OM} for TIP4P-FB that is 30 % smaller. Finally, the θ parameters of these models are at the edge of our optimal region. In summary, our results demonstrate that despite the constraint of the target ε_r compatible with ECC, the optimal parameter region is sizable and robustly defined.

Among all the ECC compatible models in the optimal region ($C_{ECC} < 0.3$), we present here in detail one of the best performing models in terms of the cost function (C_{ECC} =0.231/C_G =0.262) while possessing a balanced structural, thermodynamic, and dynamic behavior, see Table 4. The quality of our model, which we label as ECCw2024, is comparable to that of existing 4-site models such as TIP4P/2005, OPC4, or TIP4P-FB, see Table 4 and Figure 3. This is a non-trivial result, allowing further force field development with a water model fully compatible with the ECC framework, i.e., possessing a dielectric

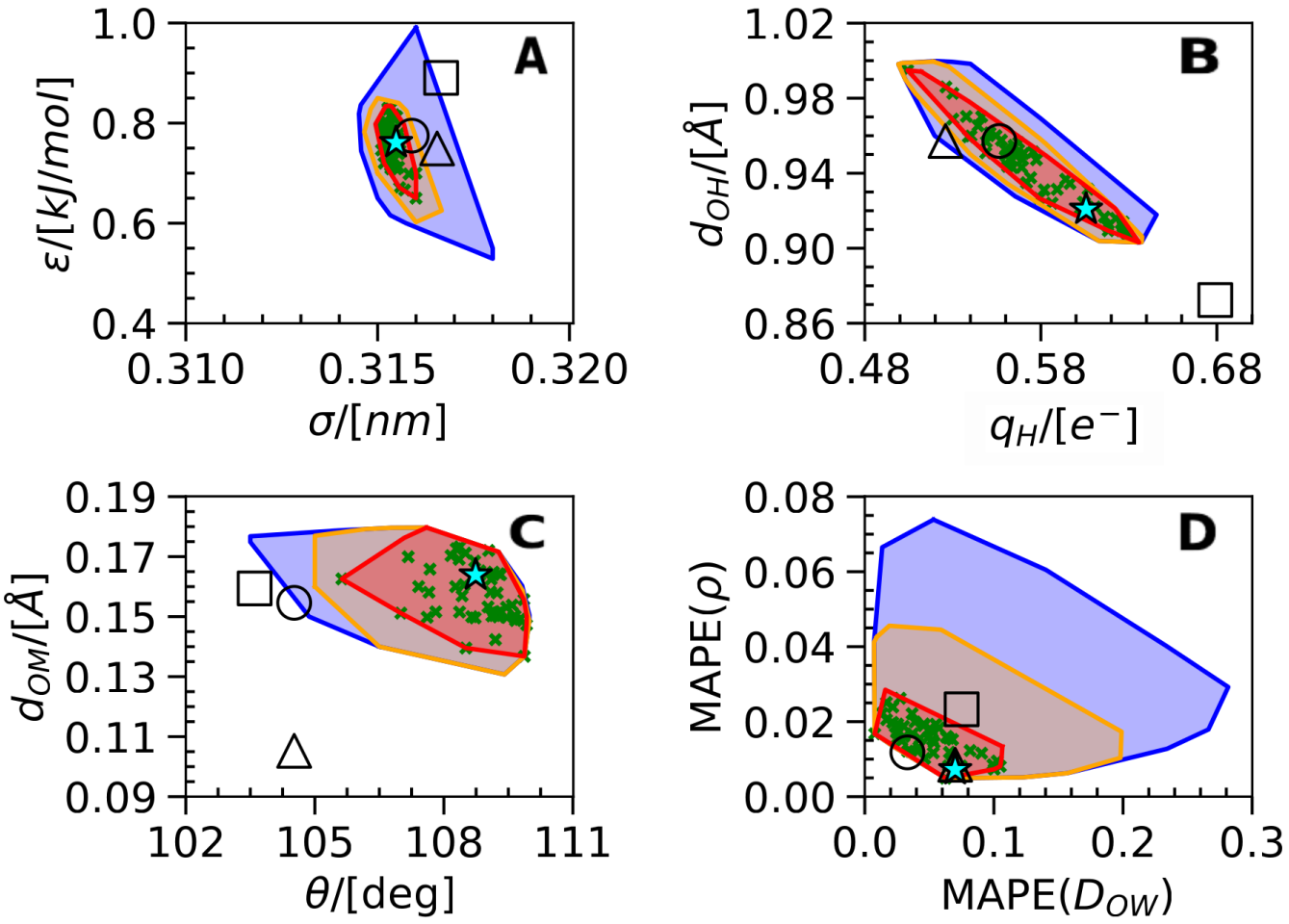


Figure 2: Minimum convex polygon (convex hull) that contains all points inside a region for selected pairs of parameters or properties: A) σ and ε, B) q_H and d_{OH}, C) θ and d_{OM}, and D) mean percentage errors of D_{OW} and ρ. The regions are defined by the scoring values points: blue ($C_{ECC} < 0.7$), orange ($C_{ECC} < 0.5$), and red ($C_{ECC} < 0.3$). The green symbols (×) are our best 50 points with ECCw2024 denoted as (★). The black open symbols correspond to TIP4P/2005 (○), OPC4 (□), and TIP4P-FB (△).

Table 3: Parameters of the water models used in this publication. TIP4P/2005 and TIP4P-FB have the gas phase molecular. OPC4 and ECCw2024 allow different molecular geometries during their optimization.

Param.	ECCw2024	TIP4P/2005	TIP4P-FB	OPC4
d_{OH}[nm]	0.092084	0.09572	0.09572	0.08724
θ[deg]	108.7392	104.52	104.52	103.60
σ[nm]	0.315480	0.31589	0.31655	0.316655
ε[kJ/mol]	0.761154	0.7749	0.74928	0.89036
$q_{\mathrm{H}}[au]$	0.605689	0.5564	0.52587	0.6791
d_{OM}[nm]	0.016388	0.01546	0.010527	0.01594
$\mu[D]$	2.167631	2.305097	2.427804	2.479542
$Q_T[D\text{Å}]$	2.444435	2.296802	2.170775	2.299607

constant of about 45. Table S2 contains the numerical values for each of the evaluated properties for these models.

Going into further detail, the oxygen–oxygen radial distribution functions (RDF) are presented in Figure 3A. The four water models yield very similar results, particularly within the first coordination shell. They all fit well the position of the experimental position of the first peak (0.280 nm) but overshoot its height as expected due to the lack of many-body interactions and potentially other effects.

Table 4: Results are provided as mean absolute percentage errors MAPE (%) at different thermodynamic conditions. Radial distribution function (rdf), viscosity (η), and surface tension (γ) correspond to 300 K and 1 bar. The melting point temperature is at 1 bar. Finally, the comparison between water models is done employing the cost function C_G, see equation 2, without including the relative permittivity. Lower values of C_G mean better performance of the water model (for comparison, the 3-site TIP3P model yields a very high value of $C_G = 1.973$). Note that the properties used in C_G are provided in Table 2 and that the MAPE values are normalized.

Property	ECCw2024	TIP4P/2005	OPC4	TIP4P-FB
$\rho_{1\mathrm{bar}}$	0.072	0.118	0.233	0.085
$\rho_{300\mathrm{K}}$	0.173	0.075	0.034	0.012
D_{OW}	7.0	3.1	7.3	6.6
rdf_{1p}	1.429	1.429	0.714	1.486
rdf_{1h}	25.6	24.2	21.5	25.8
η	0.106	2.456	5.982	3.420
γ	5.26	2.17	2.71	3.79
T_{melt}	6.23	8.48	10.3	11.0
C_G	0.262	0.208	0.248	0.260

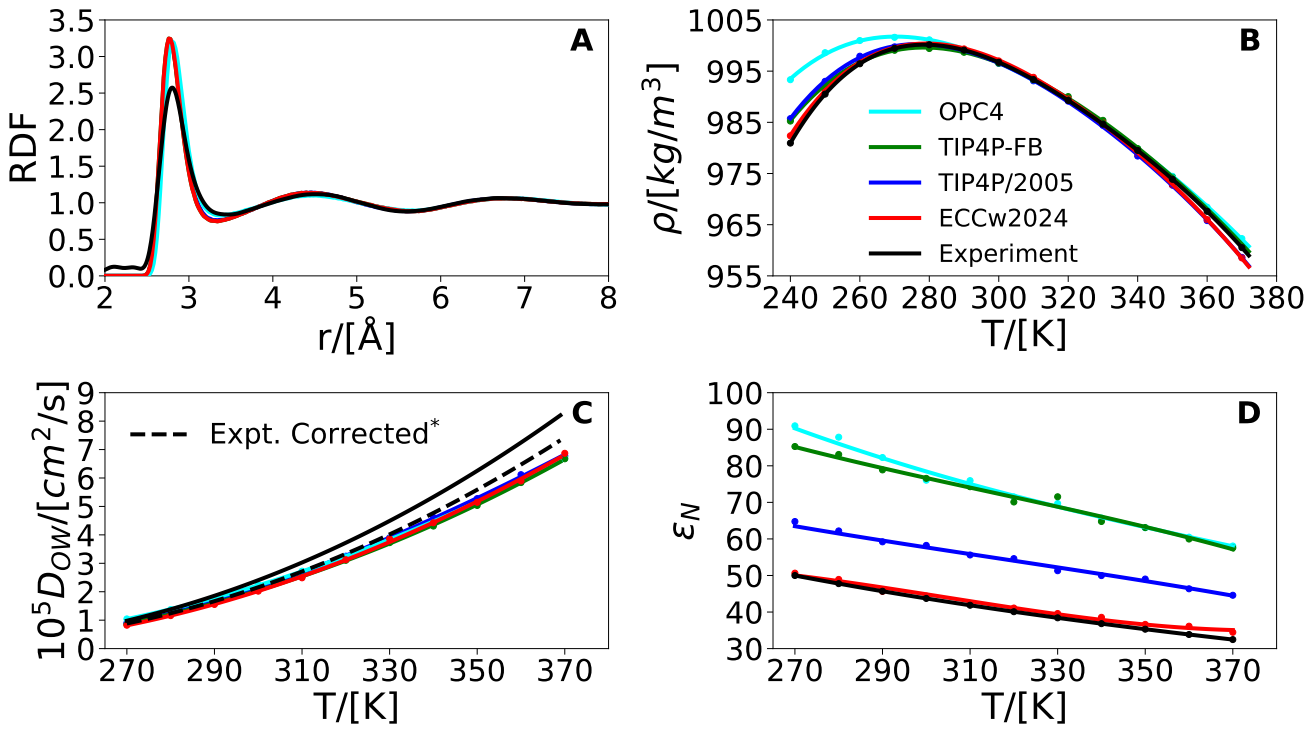


Figure 3: Water models performance in comparison with water experimental results. A) The radial distribution function[33] at 300 K and 1 bar. B) density isobar at 1 bar.[13] C) D_{OW} isobar at 1 bar.[14] D) ε_N isobar at 1 bar.[15,19] *Periodic boundary conditions correction.[28]

Figure 3B shows the temperature dependence of the density at 1 bar. All models perform well between 300 K and 340 K. In addition, our model matches the experiment within ≈2 kg/m^3 all the way to 240 K. At low temperatures, this fixes the ≈5 kg/m^3 deviations of TIP4P/2005 and TIP4P-FB (which is already substantially smaller than the deviation of OPC4 that reaches ≈13kg/m^3).

The temperature dependence of the water self-diffusion coefficient (D_{OW}) is presented in Figure 3C. At temperatures below ≈320 K, all water models, including the present one, converge to values matching experiments, except for OPC4 that diffuses slightly faster than the other water models. At high temperatures ($T >\approx 320K$), all water models deviate from experiment in a similar way yielding a somewhat too slow dynamics. Over the whole investigated temperature range, all the models show a very similar performance with a MAPE of ≈ 7%, see Table 4, except for TIP4P/2005 with a bit smaller MAPE of 3.1%.

With good agreement with experiments that our model has been optimized against, the next question to address is whether it also predicts correctly other physical properties. To this end, we have computed a set of additional properties, namely ε_N, $\rho_{300\mathrm{K}}$, γ, η, and T_{melt} (see Table 4, Table S2, Figure 3, and Figure S4).

One property our model was not a priory optimized against is the temperature dependence of ε_N, i.e., the nuclear contribution to ε_r, see Figure 3D. Within the present non-polarizable simulations, it represents the only contribution to the dielectric constant, while as a reference it can be computed by dividing the experimental total dielectric constant ε_r at a given temperature[15,34] by the infinite frequency dielectric constant at the same temperatures.[19] The very good agreement of the present model with experiments in the whole temperature range is remarkable, particularly in comparison to the other water models (Figure 3D).

We also calculated the pressure dependence of the density at 300K ($\rho_{300\mathrm{K}}$), see Figure S4. The response to pressure of our models is slightly offset with respect to the other reference models remaining, however, within 4 kg/m^3 from experiments in the whole investigated pressure range. Together with the proper description of water densities at different temperatures, this agreement shall result in a correct description of the isobaric (κ_p) and isothermal (α_T) compressibilities. All water models yield very similar values of surface tension within 4 mN/m below the experimental value of 71.68 mN/m at 300 K (Table S2). All considered models also do a good job reproducing the viscosity of water at 300 K and 1 bar, falling slightly short of the experimental value 0.85 mPa·s with values between 0.80 mPa·s and 0.88 mPa·s (Table S2). Finally, all water models somewhat underestimate the melting point of the I_h ice. Although the present model performs the best (see Figure S3), its melting point still lies 17 K below the experimental value of 273.15 K. Note also that for the TIP4P/2005, TIP4P-FB, and OPC4 water models the reported melting points are consistent (within 3 K) with previously computed values.[35]

Overall, using the presently developed optimization framework that takes advantage of AI machinery, we were able to sample efficiently the water parameter phase space and produce a 4-site water model compatible with the ECC framework (i.e., possessing $\varepsilon_r \approx 45$) with a very good performance, which is comparable with that of currently widely employed 4-site water models such as TIP4P/2005, OPC4, or TIP4P-FB. It should be stressed that it was by no means obvious from the onset that it is at all possible to generate a non-polarizable water model with such a low value of a dielectric constant (truly reflecting only the contribution from nuclear motions) that reproduces experimental properties of liquid water so well. Most importantly, we identified a sizable region of the parameter space encompassing this model that yields high-quality ECC-compatible water models. This will allow us to perform future modifications of the water model if needed to accommodate solutes within the charge scaling ECC approach, such as simple ions or charged biomolecules (or fragments thereof).

Acknowledgement P.J. acknowledges support from the European Research Council via an ERC Advanced Grant no. 101095957. V.C.C. acknowledges the support from Charles University in Prague and the International Max Planck Research School in Dresden.

Supporting Information Available

Detailed information about calculations of properties such as the melting point and the dielectric constant ε_N, as well as the absolute optimization results for the water models employed are available in the Supporting Information. This material is available free of charge at XXX.

References

(1) Kontogeorgis, G. M.; Holster, A.; Kottaki, N.; Tsochantaris, E.; Topsøe, F.; Poulsen, J.; Bache, M.; Liang, X.; Blom, N. S.; Kronholm, J. Water structure, properties and some applications – A review. *Chemical Thermodynamics and Thermal Analysis* **2022**, *6*, 100053.

(2) Rahman, A.; Stillinger, F. H. Molecular Dynamics Study of Liquid Water. *J. Chem. Phys.* **2003**, *55*, 3336–3359.

(3) Jorgensen, W. L. Quantum and Statistical Mechanical Studies of Liquids. 10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers. Application to Liquid Water. *J. Am. Chem. Soc.* **1981**, *103*, 335–340.

(4) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13–16, 1981*; Pullman, B., Ed.; Springer Netherlands: Dordrecht, 1981; pp 331–342.

(5) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(6) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(7) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(8) Perrone, M.; Capelli, R.; Empereur-mot, C.; Hassanali, A.; Pavan, G. M. Lessons Learned from Multiobjective Automatic Optimizations of Classical Three-Site Rigid Water Models Using Microscopic and Macroscopic Target Experimental Observables. *J. Chem. Eng. Data* **2023**, *68*, 3228–3241.

(9) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(10) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.

(11) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863–3871.

(12) Sedano, L. F.; Blazquez, S.; Noya, E. G.; Vega, C.; Troncoso, J. Maximum in density of electrolyte solutions: Learning about ion–water interactions and testing the Madrid-2019 force field. *J. Chem. Phys.* **2022**, *156*, 154502.

(13) Lide, D. R. *CRC Handbook of Chemistry and Physics*, 74th ed.; CRC Press, 2005.

(14) Holz, M.; Heil, S. R.; Sacco, A. Temperature-Dependent Self-Diffusion Coefficients of Water and Six Selected Molecular Liquids for Calibration in Accurate ^{1}H NMRPFG Measurements. *Phys. Chem. Chem. Phys.* **2000**, *2*, 4740–4742.

(15) Fernández, D. P.; Goodwin, A. R. H.; Lemmon, E. W.; Levelt Sengers, J. M. H.; Williams, R. C. A Formulation for the Static Permittivity of Water and Steam at Temperatures from 238 K to 873 K at Pressures up to 1200 MPa, Including Derivatives and Debye–Hückel Coefficients. *J. Phys. Chem. Ref. Data* **1997**, *26*, 1125–1166.

(16) Vega, C.; Abascal, J. L. F. Simulating Water with Rigid Non-Polarizable Models: A General Perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663–19688.

(17) Tempra, C.; Ollila, O. H. S.; Javanainen, M. Accurate Simulations of Lipid Monolayers Require a Water Model with Correct Surface Tension. *J. Chem. Theory Comput.* **2022**, *18*, 1862–1869.

(18) Leontyev, I. V.; Stuchebrukhov, A. A. Electronic continuum model for molecular dynamics simulations. *J. Chem. Phys.* **2009**, *130*, 085102.

(19) Schiebener, P.; Straub, J.; Levelt Sengers, J. M. H.; Gallagher, J. S. Refractive Index of Water and Steam As Function of Wavelength, Temperature and Density. *J. Phys. Chem. Ref. Data* **1990**, *19*, 677–717.

(20) Farahvash, A.; Leontyev, I.; Stuchebrukhov, A. Dynamic and Electronic Polarization Corrections to the Dielectric Constant of Water. *J. Phys. Chem. A* **2018**, *122*, 9243–9250.

(21) Lamoureux, G.; MacKerell, J., A. D.; Roux, B. A Simple Polarizable Model of Water Based on Classical Drude Oscillators. *J. Chem. Phys.* **2003**, *119*, 5185–5197.

(22) Das, A. K.; Demerdash, O. N.; Head-Gordon, T. Improvements to the AMOEBA Force Field by Introducing Anisotropic Atomic Polarizability of the Water Molecule. *J. Chem. Theory Comput.* **2018**, *14*, 6722–6733.

(23) Kirby, B. J.; Jungwirth, P. Charge Scaling Manifesto: A Way of Reconciling the Inherently Macroscopic and Microscopic Natures of Molecular Simulations. *J. Phys. Chem. Lett.* **2019**, *10*, 7531–7536.

(24) Duboué-Dijon, E.; Javanainen, M.; Delcroix, P.; Jungwirth, P.; Martinez-Seara, H. A practical guide to biologically relevant molecular simulations with charge scaling for electronic polarization. *J. Chem. Phys.* **2020**, *153*, 050901.

(25) Vega, C. Water: One Molecule, Two Surfaces, One Mistake. *Mol. Phys.* **2015**, *113*, 1145–1163.

(26) Storn, R.; Price, K. Differential Evolution – A Simple and Efficient Heuristic for Global Optimization Over Continuous Spaces. *J. Global Optim.* **1997**, *11*, 341–359.

(27) Goodfellow, I. J.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

(28) Yeh, I.-C.; Hummer, G. System-Size Dependence of Diffusion Coefficients and Viscosities from Molecular Dynamics Simulations with Periodic Boundary Conditions. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

(29) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1-2*, 19–25.

(30) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(31) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(32) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(33) Skinner, L. B.; Huang, C.; Schlesinger, D.; Pettersson, L. G. M.; Nilsson, A.; Benmore, C. J. Benchmark Oxygen-Oxygen Pair-Distribution Function of Ambient Water from X-Ray Diffraction Measurements with a Wide Q-Range. *J. Chem. Phys.* **2013**, *138*, 074506.

(34) Malmberg, C. G.; Maryott, A. A. Dielectric Constant of Water from 0 to 100 C. *J. Res. Natl. Bur. Stand.* **1956**, *56*, 1.

(35) Blazquez, S.; Vega, C. Melting points of water models: Current situation. *J. Chem. Phys.* **2022**, *156*, 216101.