Machine Learning Assisted Screening of 2D Materials for Water Desalination

Pikee Priya, Thanh C. Nguyen, Anshul Saxena, N. R. Aluru[#]

Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

[#]Corresponding author: aluru@illinois.edu

Abstract

There exists vast expanse of data in literature which can be harnessed for accelerated design and discovery of advanced materials for various applications of importance – for example, desalination of sea water. Here, we develop a machine learning (ML) model, training it with ~260 molecular dynamics (MD) computation results to predict the desalination performance of 2D membranes that exist in literature. The desalination performance variables of water flux and salt rejection rates are correlated to 44 material features related to the chemistry of the pores and the membranes along with applied pressure, salt concentration, partial charges on the atoms, geometry of the pore and the mechanical properties of the membranes. We used the ML model to screen 3814 structurally optimized 2D materials for maximum water flux and salt rejection rates from the literature. We found some candidates that perform 3 times better than the more popularly known 2D materials of graphene and MoS_2. This result is verified using data obtained from MD simulations performed on several 2D membranes. Such validated statistical frameworks using literature data can be very useful in guiding experiments in the field of functional materials for varied applications.

1. Introduction

With the increasing global crisis for potable water in different regions of the world, newer water desalination techniques are the need of the hour.[1–3] The various oceans and seas contain about 97% of the world's water which is undrinkable, and only a few percent of the rest are drinkable. Seawater contains 40-60 g/l of salt while potable water should have less than 1 g/l, the requirement being more stringent sometimes[4]. A possible solution to this crisis is efficient desalination of seawater. However, the prevalent desalination techniques suffer from important drawbacks such as large energy footprints, high material and capital costs, and poor performance[3,4]. Reverse osmosis (RO) is an energy-efficient desalination technique with an energy intake of 1.8 kWh/m^3 of water compared with an average ∼5 kWh/m^3 in the 1990s.[1,3,5] It is also the most prevalent technique only behind distillation which is a high energy consuming process and not practical for countries with low fuel reserves. Approximately half of the current desalination plants use RO technologies[6] where a pressure gradient drives the seawater through a semi-permeable membrane. This makes water to preferentially move across from higher salt concentration to lower salt concentration with a specified flux and a certain salt rejection efficiency.

Although the energy efficiency of industrial RO processes is increasing, it is still much smaller than the thermodynamic minimum of ~0.5-1kWh/m^3 (25-45 g/l of NaCl)[3,7] for 50% water recovery. Along with the challenges relating to energy, RO also faces challenges relating to membrane material[8,9]. The membrane materials used for RO have evolved over time from cellulose-based to zeolites to polymer-based thin films which are commonly used today. While the polymer-based semi-permeable membranes provide higher fluxes with salt rejection efficiency comparable to zeolites, they still lag behind the optimum performance goal of having fluxes as high as CNT-based filters and salt rejection rates as good as zeolites[6]. In addition, the commonly used polymer, polysulfone, suffers from surface degradation due to the formation of polyamide. This can be prevented by applying a cross-linking polymer at the surface, but the resulting polyamide is also not resistant to Cl and other oxidizing environments[7], which easily attack its amine group. Alternative techniques like deposition of nano-structured zeolites on the surface had limitations to the maximum salinity that it could treat[6]. Older RO technologies of polymer-based filtration to purify seawater are giving way to the newer techniques using advanced materials. Seawater desalination using smart materials like nano-porous two-dimensional (2D) materials

such as graphene, graphene oxide and MoS_2 have been reported in literature as a viable solution[10–12]. However, the water fluxes obtained at desired high salt rejection rates using these materials are too low to make the technology adaptable for industrial scale as shown in Figure 1(a). Therefore, material and design optimization of these membranes is required to reach the optimum desalination performance goal of high salt rejection rates and water fluxes.

There exist numerous experimental[13] and computational[12,14] studies of 2D materials used for water desalination. Some experiments using multiple layers of 2D materials have been performed in the past, for 2D materials like graphene, graphene derivatives and MoS_2[15–18]. Computational studies on single layer 2D membranes, on the other hand are abundant and have been done for the commonly known materials. Classical molecular dynamics (MD) simulations have often been used in the past to report the systematic effects of pore size, pore chemistry, and applied hydrostatic pressure on desalination performance. We, therefore, collected 257 such computational data points of water flux and salt rejection rate at the corresponding pressures, salt concentrations, atomic partial charges, pore and membrane sizes, for common 2D materials like graphene, graphene oxide, molybdenum disulfide (MoS_2), molybdenum diselenide ($MoSe_2$), hexagonal boron nitride and carbon nitride (C_2N). The complete dataset from literature with references is provided in the supplementary dataset. With the recent advent of Materials Genome Initiative, we also have access to vast expanse of data on 2D materials in open databases. These sources of data along with those reported in journals can be used for developing statistical methods to optimize the chemistry of these 2D materials for optimum performance.

The pore structure and the charge of pore atoms are very important in determining the selectivity of ions passing through the pore[12,19]. For example, as reported with MoS_2[12], water flux through the pore is higher when Mo atoms form the pore edge than the S atoms. Therefore, in this study, we use the structure, chemistry and atomic partial charges of the pores and the membranes provided in the literature to build a set of feature list and correlate them to the water flux and salt rejection efficiency, through a well-tested and validated machine learning (ML) model. The method adopted in this work is discussed in detail in Section I of Supplementary Information. A complete list of features used for machine learning is provided in Section II of the Supplementary Information. The model once built is then used to screen the best candidates for desalination with a higher water flux along with a high salt rejection rate. High-throughput computational studies

from the past for discovering new 2D materials candidates using structural optimization, are then used to screen for optimum desalination performance and understand the underlying design guidelines. These materials with their crystal structures, calculated electrical and mechanical properties, thermodynamic and dynamic stability can be found in the related open database[20]. The flow of data for exploring the chemical space for 2D materials is shown in Figure 1(b) and is also described in detail in the Methods section. The problem is complex as it involves screening 2D chemistries of type $\prod A_i^j B_k^l \; where: \; i,j,k,l = 0,1,\ldots.\infty$ (i and k being the stoichiometric constants for the given element and j and l being the element number) with A being metallic and B, a non-metallic element, candidates for which from the periodic table are shown in Figure 1(c). We have also studied the effect of functionalization with the constituent atoms or elements extrinsic to the 2D material system. In addition, the predictions from the ML model are verified using the data obtained from MD simulations for some of the 2D materials.

2. Methods

2.1. Machine Learning

Computationally determined water flux and salt rejection rates per unit pore area for a specific applied pressure were manually extracted using WebPlotDigitizer[21], from 16 different journal publications. These were for 2D materials like graphene, graphene oxide, MoS_2, hexagonal BN and C_2N with atomic configurations (functionalization) at the pore edge. The water flux and salt rejection rates were correlated to 44 different features which are chemical, electrical and structural features of the membrane and the pore. While some of the features, such as, salt concentration, pore area, membrane area, applied pressure and atomic partial charges were taken from the published work on desalination using 2D materials, other features like band gap, bond angles, bond lengths, membrane thickness, atomic density of the membrane and pore were searched for in the open databases, for optimized DFT structures of the 2D materials.

With these data in hand, we evaluated the accuracy of several ML methods using 80% and 10% of the data as the training and testing set, respectively, and the remaining 10% for cross-validation. The total number of computational samples used for ML model was 257. For all the models we also used grid search with 5 k-fold cross validation to optimize the hyperparameters. This procedure has a single parameter called k that refers to the number of groups that a given data sample is to be split into. This is done so that every split dataset serves as a test set for that iteration

(carries for k iterations). ML package, Scikit-learn[22] was used for the purpose. For the linear regression models, we evaluated various regularization methods such as LASSO[23] (least absolute shrinkage and selection operator), ridge regression[24] and elastic net[25], but regularization worsened the predictive performance. However, k-nearest neighbor[26] regression with 5 nearest neighbors, delivered improvement over these methods. Moreover, we found that the random forest[27] (RF) algorithm with XG-Boosting[28] provided even better predictive performance. Random forest algorithms operate by randomly selecting a subset of features and constructing decision trees based on the limited data, which are then averaged out for the final prediction. XG-Boosting uses a sequence of RF models learning sequentially from the previous model. By combining several of these decision tree models on the data subsets, an accurate prediction of the desalination performance could be made without overfitting. A comparison of these methods is provided in Supplementary Section III. The testing-set root mean absolute error (MAE) using the XG-Boost model is 4.2 % with a coefficient of determination value (R^2) of 0.81, and the MAE value is 2.4 % with R^2 of 0.97 for the cross validation-set for salt rejection (%). For $\log_2$(water flux (l/cm^2/day)), MAE is 0.2 for both the testing and cross-validation sets, with a coefficient of determination R^2 value of 0.99 and 0.98 for the testing and cross validation sets, respectively. The coefficient of determination and the mean absolute error is given by, $R^2 = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$ and $MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$ where y_i is the ground truth, $\hat{y}_i$ is the predicted value and $\bar{y}$ is the average of the sample. The ML model is then used to predict the output water fluxes and salt rejections for 3814 2D materials from the Computational 2D Materials Database, https://cmrdb.fysik.dtu.dk/c2db/ [20]. Since the partial atomic charges on the atoms of these membranes were unknown, we optimized a different ML model to predict the charges on the atoms based on features discussed in detail in Supplementary Section IV. The effect of pore functionalization with a metallic or non-metallic element was also studied. The 2D materials were next screened for thermodynamic stability provided by Haastrup et al.[20] in the literature. The predictions made by the model are also verified using MD simulations as described in the next section.

2.2. MD Simulations

Pressure-driven transport calculations of saltwater through nanoporous 2D materials were performed with MD simulations using LAMMPS package[29]. As shown in Figure S5.1, Section V

of the Supplementary Information, the simulation domain consists of a saltwater (feed) reservoir and a pure water (permeate) reservoir, separated by a nanoporous membrane in the middle. Some of the 2D membranes simulated include C_2, C_2H_2, Mo_2O_4, Ti_2O_4, CrO_2, FeO_2, V_2O_4, CoO_2, $Au_2Cl_2N_2$, $F_2N_2Ti_3$, ZnH_2O_2, CuH_2O_2, and $Cl_4O_4Ti_4$. The 2D membrane areas are approximately 16 nm^2. Due to the differences in crystal structure of the materials, the exact area of each membrane is slightly different from 16 nm^2. Similarly, the pore area of each membrane is approximately 50 $Å^2$. A pore in a membrane was created by removing some of the membrane atoms. The pore area is defined as the total empty space surrounded by the atoms on the pore edge. A pressure-driven flow was created by applying different constant forces on the two rigid pistons on either side of the membrane. At a specified pressure difference, the left piston moves in the z-direction pushing the water molecules and ions along with it, whereas the right piston moves downstream. This method of creating pressure-driven flows through nanoporous membranes is well established in the literature[12,14]. In this study, a pressure difference of 100 MPa was created by applying a pressure of 100 MPa plus 1 atm on the left piston and a pressure of 1 atm on the right one. Initially, there are 2511 water molecules and 1680 water molecules in the feed and the permeate reservoirs, respectively. The initial distances from the left and the right pistons to the middle of the membrane are 45 Å and 35 Å, respectively. For the chosen salinity of 50 g/l, the feed reservoir was initially assigned 33 Na^+ and 33 Cl^- ions.

Simple point charge extended (SPC/E) model was used to simulate water molecules[30]. SHAKE algorithm was used to keep the O-H bond length and H-OH angle at 1.0 Å and 109.47°, respectively [31]. Truncated Lennard-Jones (LJ) 12-6 potential with a cutoff distance of 1.2 nm was used to model the intermolecular interactions (van der Waals) of salt ions, oxygen atoms in water, and membrane atoms. Oxygen-oxygen interaction values were taken from the SPC/E model[30]. LJ interaction parameters of sodium and chloride ions in aqueous solutions were modeled using GROMOS force fields[32]. The LJ interaction parameters and charges for the different membrane atoms were taken from the Reaxff force fields[33]. The interaction parameters between the remaining dissimilar atoms were obtained using Lorentz-Berthelot rule (Supplementary Information, Section V). For all atomic species having charge, short-range Coulombic interactions were computed with a cut-off distance of 1.0 nm, while long-range electrostatic interactions were computed using particle-particle, particle-mesh (PPPM) method[34].

The energy of simulation system was initially minimized with a stopping tolerance of 10^{-6}. Next, an equilibrium simulation with a constant number of atoms, pressure, and temperature (NPT) ensemble was performed at 1 atm and 300 K. The membrane was fixed at its original position during the equilibration, while piston atoms could change their positions as rigid objects. The NPT simulation preserved the thermodynamic state of water at a density of 1 g/cm^3 and temperature of 300 K. Next, a constant number of atoms, volume, and temperature (NVT) non-equilibrium simulation using the Nosé-Hoover thermostat was performed for 15 ns with a pressure difference of 100 MPa to study water transport through the membrane. In these simulations, the membrane was still fixed at its original position, while the rigid pistons gradually adjust their positions due to the applied pressure difference. Water flux was calculated based on the average rate of water molecules passing through the membrane, and the selected area (membrane area or pore area). A detailed calculation method is provided in Section V of the Supplementary Information.

3. Results and Discussion

3.1. Features important for water desalination from ML model

The screened 2D materials were of several stoichiometries and had different atomic structures as described in detail in the C2DB open database literature[20]. These include a single or multiple atoms of the membranes like A, AB, AB_2, ABC, AB_3, A_2BC_2 and AB_2C_4 where A/B/C can be either metallic or non-metallic elements. Here, we try to classify them based on the metallic and non-metallic element or elemental groups that each 2D material consists of, as we find that the features related to the position of the atom in the periodic table has a major effect on the desalination performance as also seen in Figure 2(a-c). From the feature importance bar chart in Figure 2(a), we find that the membrane average atomic number, partial charge on the atoms, and average pore atomic radius play an important role in the estimation of water flux and salt rejection rates (along with the operational properties involved in RO such as applied pressure, salt concentration, pore and membrane area not shown in the figure). Figure 2(a) also shows the effect of these features on water flux and salt rejection rate during desalination based on the collected computational data from literature. These factors are found to be important with high “feature importance”, which is a measure of the “out of bag” error, i.e., the error not including one of the features during implementation of the ML algorithm. However, there remains a lot of chemical space that remains unexplored with reasonable water fluxes and salt rejection rates which would lie at the right top corner of the contour plots shown in Figure 2(a). Based on the data from

literature, it seems lower membrane atomic numbers and intermediate membrane charges lead to higher desalination performance. The direct trends of desalination performance in the contour plots are not very clear because of simultaneous dependence on other factors/features. We try to generalize these trends again using MD simulations which will be discussed in Section 3.4.

Owing to the importance of position of the atoms in the membrane on the periodic table in determining desalination performance as seen in Figure 2(a), we chose lower atomic number H, C, N, chalcogenides and halogens to study the effect of pore functionalization. We also study metallic functionalization with transition element atoms.

3.2. Effect of Pore Functionalization

The functionalization of pores is of immense importance and has been studied by researchers in the past. Fang et al.[19] reported that increasing the charges at the pore from 0 to -3e enhances the water transport, while further increasing charges to -6e, reduces the water transport in porous graphene oxide membranes. The ML model feature importance list also points towards the importance of partial charges on the atoms in general. The detailed feature list with the importance is provided in Tables S2.1 and S2.2 in Supplementary Information Section II. We functionalized the pores of the 3814 membranes provided in the C2DB database[20] with 8 non-metallic (Cl, Fl, Br, I, O, H, C and N) and 8 metallic (Mo, Ru, V, Cr, Ti, Fe, Co and Mn) atoms. The results are shown in Figure 2(b) for non-metallic and 2(c) for metallic elements. Metals are known to be more hydrophilic than non-metals, and this numerical experiment helps us observe the difference it brings to the desalination performance.

The non-metallic functionalization increased the water fluxes with the highest flux for pores functionalized with C and H. However, the salt rejection rates were reduced below the minimum value for the unfunctionalized pores indicated by the bottom line of the grey box in Figure 2(b), more so for some of the hydrogenated pores. Metallic functionalization, on the other hand, improved the salt rejection rates and decreased the minimum for water flux shown in Figure 2(c). Comparing the two figures (2(b) and (c)), we find some of the carbonated and hydrogenated pores lie in the top right corner of Figure 2(b) with higher water fluxes and salt rejection rates.

3.3. Improved Class of 2D materials for Desalination

Figure 3(a) shows predicted classes of 2D materials with their desalination performance. Graphene, transition metal dichalcogenides and nitrides are classes that have been reported in the literature to perform well in desalinating seawater[12–14]. The chemical space that has not yet been reported is the halides complexes, oxide complexes and the lithiates which we find to be an important class of 2D materials for good water desalination performance. Many of these new candidates for desalination fall in the class of MXenes which are inorganic 2D membranes, for example, transition metal carbides, nitrides and nitro-carbides. The halide membranes although thermodynamically stable may dissolve in water due to the high electronegativity difference of the metallic and non-metallic atoms. The class that the ML model found promising and was also verified through MD simulations is the 2D metal oxides and its complexes. This class is increasingly being synthesized and studied these days due to its applications in electronics and renewable energy [35–38]. Its application in desalination has not been reported yet.

The findings of this study have been verified using MD simulations. Figure 3(b) shows the comparison of ML water fluxes for several candidates with MD calculations. A reasonable match between the MD and ML water fluxes both with respect to membrane area and that with respect to pore area are shown. Our reported fluxes for some of the oxides (FeO_2, CoO_2, CrO_2, V_2O_4, Ti_2O_4 and Mo_2O_4), oxide-complexes (CuH_2O_2, ZnH_2O_2 and $Cl_4O_4Ti_4$) and halide complexes ($Au_2Cl_2N_2$, $F_2N_2Ti_3$) are up to 3 times higher than those of graphene or MoS_2, which results in reduced work done that is required for desalination. This also implies increase in the efficiency of desalination which is the ratio of work done for desalination to the minimum work required. In addition, oxygen has turned out to be the most potent chalcogen that can be used with transition metal when it comes to desalination, even better than Se and S. MoS_2 and $MoSe_2$ have been studied for desalination[12,39] but Mo_2O_4 is found to out-perform graphene and perform as good as the MoS_2 as found by our MD results. Other period II elements like C and N with transition metals as 2D materials also perform well as seen in Figure 3(a).

3.4. MD based analysis

Figure 4(a) shows the effect of the top two features predicted by the ML model on water flux. The region of higher water fluxes indicated by the inset is further investigated using MD simulations, shown in Figure 4(b). Lower atomic numbers in the range of 5-30 for the 2D

membrane have the highest water flux when the partial charges in the atoms are in the range of 0.3-0.6e which is within the range of the atomic charges in H and O atoms in the water molecules. The charges for H and O atoms used in the MD simulations were 0.4238 and -0.8476, respectively (SPC/E water model). This is also shown in detail in Figure 4(c) and (d). Smaller atomic numbers or smaller radii of the atoms mean thinner 2D membranes which enhance the water flux. This can be simply understood based on the Hagen-Poiseuille equation: $Q = \frac{\pi R^4 \Delta p}{8\mu\delta}$, where, Δp is the applied pressure difference, δ is the membrane thickness, R is the radius, ΔC is the salt concentration difference, and μ is the viscosity of the liquid.

The water flux through the nano pore is determined by both the average density and velocity of water molecules in the pore. Higher is the charge in the membrane, more is the membrane-water interaction, leading to higher average densities. However, higher interactions also reduce the average velocity of the water passing through the pores. This is seen in Figure 5(e) and schematically represented in Figure 5(f). For a brief illustration, the average densities, and velocities of water for three 2D membranes, graphene (C_2), V_2O_4 and $Cl_4O_4Ti_4$ are further discussed. Figure 5(a)-(c) show the 2D density distribution of water inside the pores, respectively. The average density of water follows the trend as 0.39 g/cm^3 (C_2) < 0.6 g/cm^3 (CuH_2O_2) < 0.86 g/cm^3 (V_2O_4). The maximum positive charge in the membranes follows the same trend as 0 e (C_2) < 1.1 e ($Cl_4O_4Ti_4$) < 1.23 e (V_2O_4). This reveals the important role of Coulombic interactions in determining the average density of water inside the pore. Similar membrane-water interactions affecting the water fluxes in carbon nanotubes have also been reported earlier[40,41]. However, the average velocity follows a different trend of 2.06 m/s (C_2) < 2.34 m/s (V_2O_4) < 3.79 m/s ($Cl_4O_4Ti_4$). V_2O_4 has higher charges leading to higher average water density in the pore but lower average velocity as compared to $Cl_4O_4Ti_4$, making the water flux higher for $Cl_4O_4Ti_4$. The graphene with no charges on the membrane has both lower average density and lower average velocity. Figure 5(d) shows the average barrier energy experienced by a water molecule passing the three pores. It is seen that C_2 hinders the passage of water through the pore the most while $Cl_4O_4Ti_4$ does the least, explaining the trend in average velocity. We, therefore, need intermediate charges in the range of that of the atoms in the water molecules for a combination of better average density with average velocity.

4. Conclusions

A machine learning model correlating the computationally determined water flux and salt rejection rates to the applied pressure, salt concentrations, pore area, membrane thickness, radius of the pore and 39 other features related to the chemical, electrical and structural features of the membrane and its pore, has been developed. Among these features, pore atomic number and maximum positive charge in the membrane are found to have maximum effect on the desalination performance. The model is used to screen 3814 2D materials listed in open databases. Having transition metal at the pore, is found to enhance salt rejection rates while non-metals like halogens and chalcogens increase the water flux. 2D transition metal oxides and their complexes are predicted to be effective for desalination. We report candidates with 3 times better water fluxes than the state of art 2D membranes which has also been validated through MD simulations. Higher partial charges on the membrane atoms lead to higher water-membrane interactions and enhanced average water density but reduced average velocities.

Acknowledgement

The simulations were performed using the Extreme Science and Engineering Discovery Environment (XSEDE) (supported by NSF Grant no. OCI1053575) and Blue Waters (supported by NSF awards OCI-0725070, ACI-1238993 and the state of Illinois).

Data Availability

The data collected from literature and the features used for ML algorithms in this work are available on figshare link as supplementary dataset:
https://figshare.com/s/68a8c5945e3d94e2b73e

Code Availability

The ML scripts will be made available on request to the corresponding author.

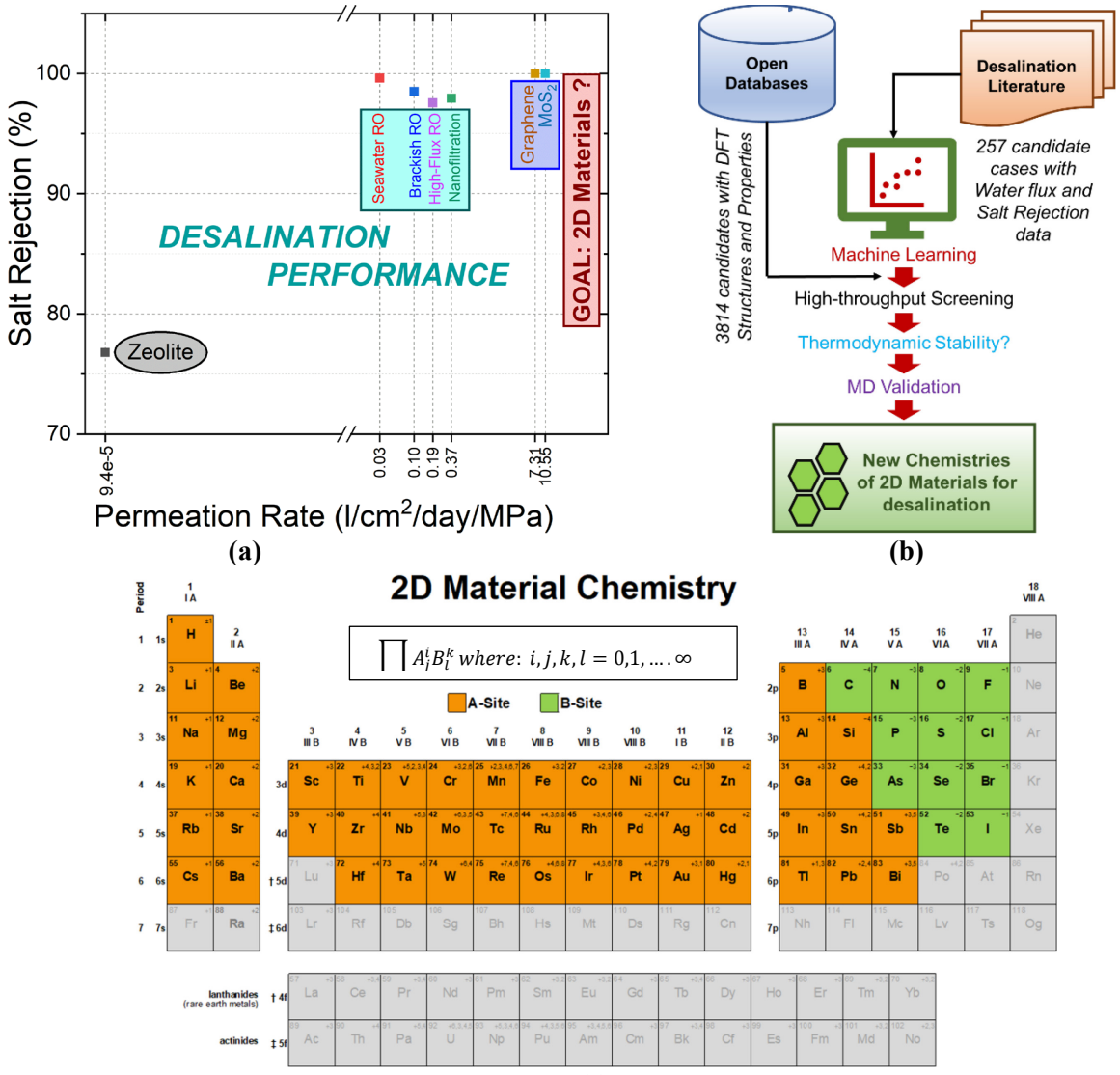


Figure 1: (a) The state of the art of desalination performance (Seawater, Brackish and High-flux Reverse Osmosis (RO), Zeolites, and some well-known 2D materials) showing the goal we want to reach with the complex chemistries of 2D materials. (b) The flow of data for screening 2D materials for high desalination performance is shown schematically. The data from desalination literature is used for developing a Machine Learning (ML) framework testing many algorithms for minimum Mean Absolute Error (MAE). This framework is used to screen 3814 2D materials provided in open Database for high desalination performance. The predictions from the ML framework are next screened for stability. The fluxes obtained are verified using Molecular Dynamics (MD) to reach new chemistries of 2D materials for desalination. (c) Periodic table showing the metallic A and non-metallic B sites possible for 2D materials screened for water desalination performance. The chemistry is of type $\prod A_j^i B_l^k$ *where*: $i, j, k, l = 0,1, \ldots . \infty$ with different combinations of metallic elements A_i and non-metallic B_k with different stoichiometries j and l.

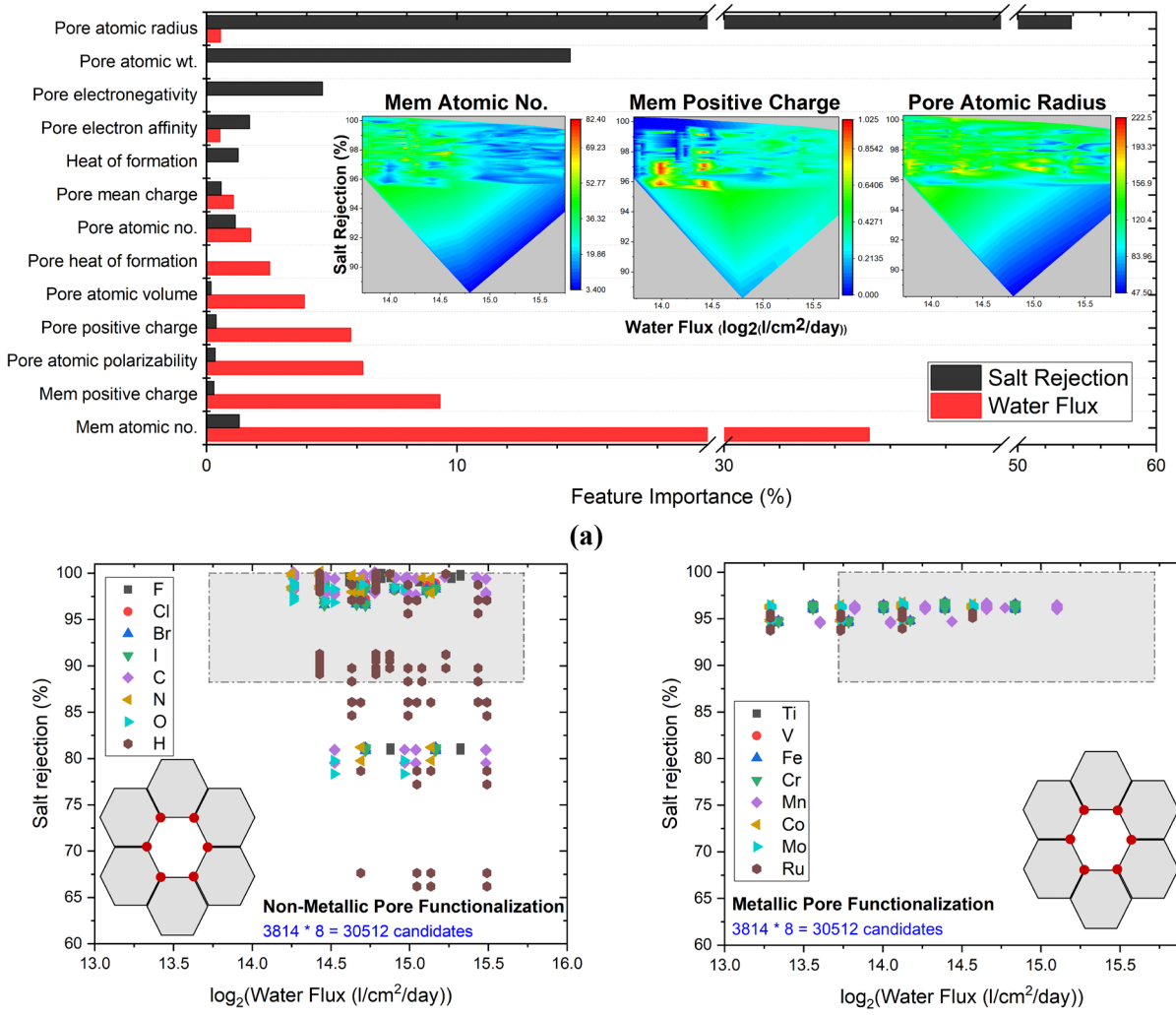
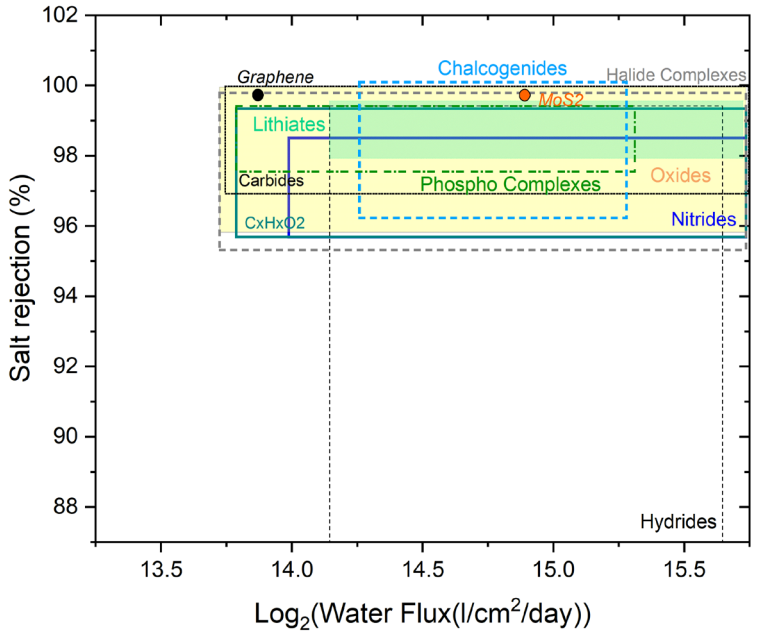
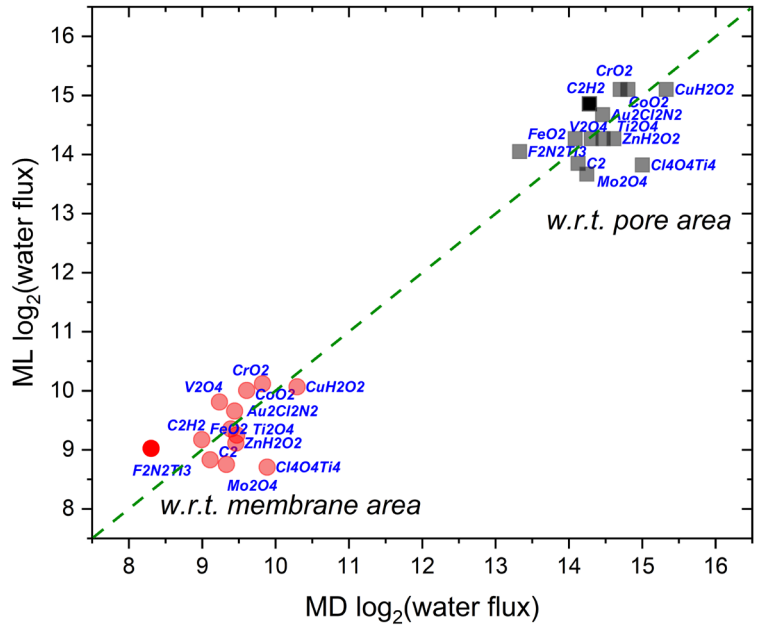


Figure 2: (a) The effect of important features with their relative importance from the ML model (along with the operational properties involved in RO such as applied pressure, salt concentration, pore and membrane area not shown in the figure). Effect of membrane atomic number, membrane positive charge and pore atomic radius from the literature data is shown in inset contour plots; Some of the features have 0 or low feature importance as we have plotted them for both the targets of water flux and salt rejection rate, for which they might be important for one but not for the other. (b) Effect of non-metallic and (c) metallic pore functionalization on desalination performance of water flux with respect to membrane area and salt rejection from the ML model; The grey box represents the range of values for 3814 non-functionalized 2D material candidates.



(a)



(b)

Figure 3: (a) The best thermodynamically stable 2D material classes predicted by the ML model compared to the state-of-the-art RO, Nanofiltration, Zeolites and 2D membranes of graphene and MoS_2. 2D Oxide complexes, Nitro-halides and Oxide membranes perform the best for desalination; (b) Comparison of MD prediction of water flux with ML model for verification of the ML model. The average pore area is 50 $Å^2$ and membrane area is 16 nm^2. The details of the pore and membrane size is provided in the Supplementary Material Section V.

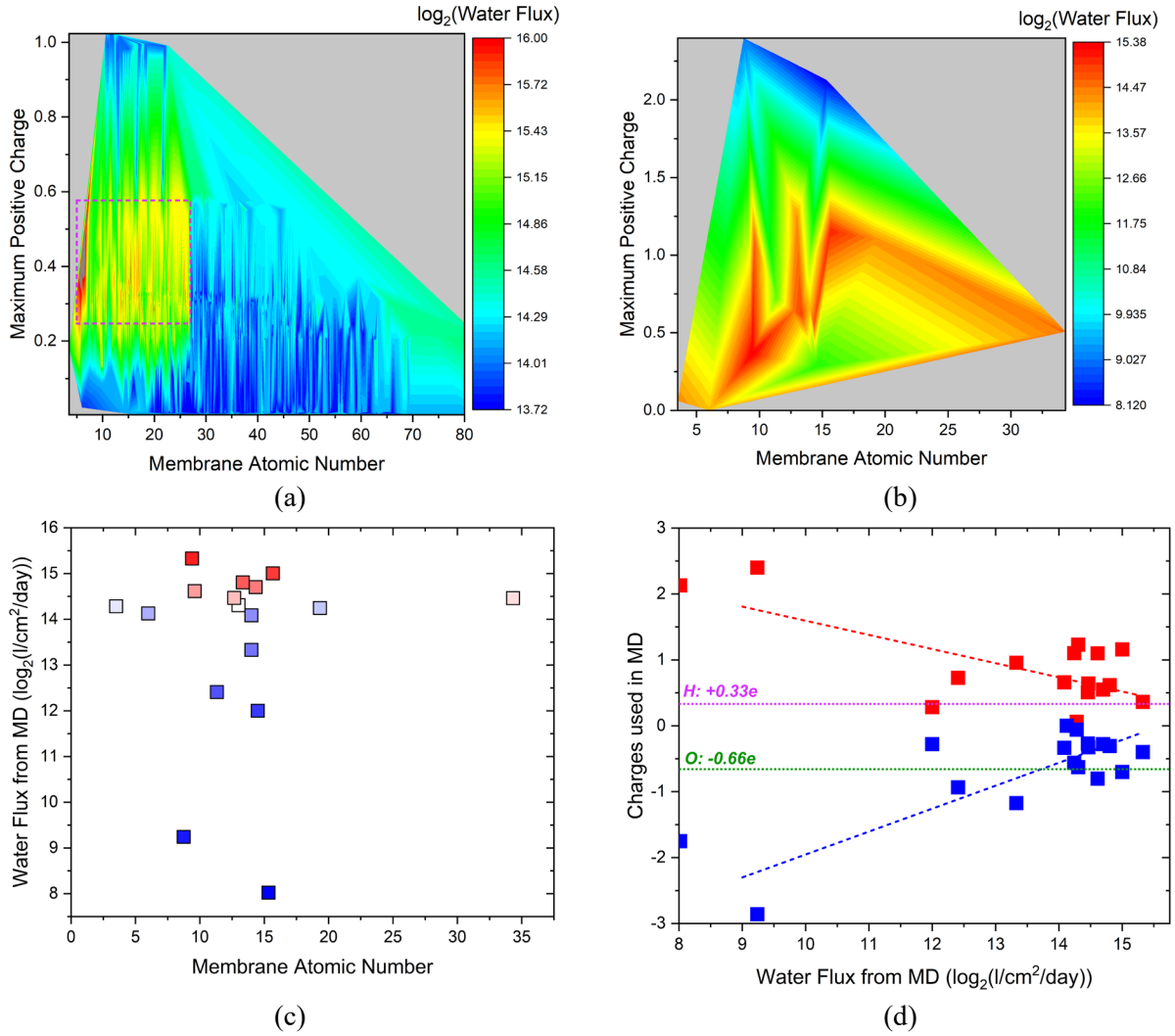


Figure 4: (a) Water Flux predicted from ML model with respect to the top two important features: membrane atomic number and maximum positive charge on the membrane. The dotted box shows the region of better desalination performance which has been probed using MD. (b) The variation of water flux, with respect to the top 2 features, calculated from MD. (c) Lower atomic numbers for the membrane unit cells showing higher fluxes. The symbol colors follow the temperature palette, red indicating higher fluxes and blue indicating lower fluxes (d) The effect of charges used in MD on water flux calculated from MD. Lower the charges on the membrane atoms better are the water fluxes indicated by the linear fits on the positive (red) and negative charges (blue) on the membrane. The charges on oxygen and hydrogen atoms in the water molecule are also indicated for reference.

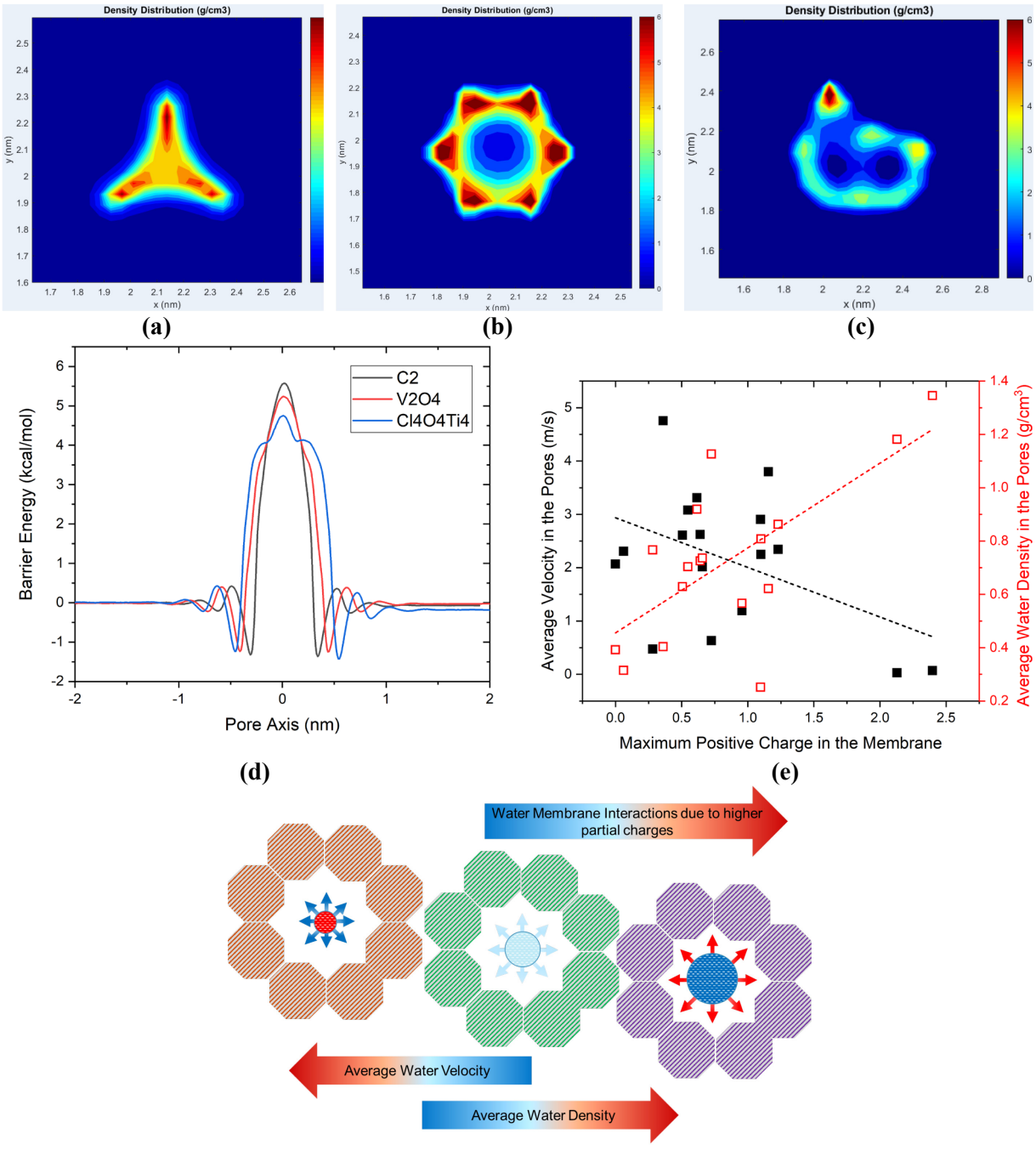


Figure 5: The time averaged water density profile for the pores of 2D membranes (a) C_2 (b) V_2O_4 (c) $Cl_4O_4Ti_4$ showing more interaction of water with the pore edges lead to more dispersed densities, higher at the pore edges and lower in the center indicating the tendency for more membrane-water interactions with increase in charges in the membrane. (d) The variation of the energy barriers for the above cases (calculation method is presented in Supplementary Information) and (e) average density in the pores (red) and the average velocity in the pores (black) calculated using MD, with maximum charge on the membrane atoms. A linear relationship of average water densities in the pore and inverse linear relationship of velocities at the pore center with maximum charge is observed. (f) The effect of higher water membrane interactions (due to higher charges) on the average water density in the pore (increases) and average water velocity during desalination (decreases) is shown.

References

1. Shannon, M. A. *et al.* Science and technology for water purification in the coming decades. *Nature* **452**, 301–310 (2008).

2. Kim, C. *et al.* Facile fabrication strategy of highly dense gadolinium-doped ceria/yttria-stabilized zirconia bilayer electrolyte via cold isostatic pressing for low temperature solid oxide fuel cells. *J. Power Sources* **415**, 112–118 (2019).

3. Liu, C., Rainwater, K. & Song, L. Energy analysis and efficiency assessment of reverse osmosis desalination process. *Desalination* **276**, 352–358 (2011).

4. Greenlee, L. F., Lawler, D. F., Freeman, B. D., Marrot, B. & Moulin, P. Reverse osmosis desalination: Water sources, technology, and today's challenges. *Water Res.* **43**, 2317–2348 (2009).

5. Kim, J., Park, K., Yang, D. R. & Hong, S. A comprehensive review of energy consumption of seawater reverse osmosis desalination plants. *Appl. Energy* **254**, 113652 (2019).

6. Humplik, T. *et al.* Nanostructured materials for water desalination. *Nanotechnology* **22**, (2011).

7. Elimelech, M. & Phillip, W. A. The future of seawater desalination: Energy, technology, and the environment. *Science (80-.).* **333**, 712–717 (2011).

8. Ismail, A. F., Padaki, M., Hilal, N., Matsuura, T. & Lau, W. J. Thin film composite membrane - Recent development and future potential. *Desalination* **356**, 140–148 (2015).

9. Lee, K. P., Arnot, T. C. & Mattia, D. A review of reverse osmosis membrane materials for desalination-Development to date and future potential. *J. Memb. Sci.* **370**, 1–22 (2011).

10. Suk, M. E. & Aluru, N. R. Water transport through ultrathin graphene. *J. Phys. Chem. Lett.* **1**, 1590–1594 (2010).

11. Cohen-Tanugi, D., Lin, L. & Grossman, J. C. Multilayer Nanoporous Graphene Membranes for Water Desalination. *Nano Lett.* **16**, 1027–1033 (2016).

12. Heiranian, M., Farimani, A. B. & Aluru, N. R. Water desalination with a single-layer MoS 2 nanopore. *Nat. Commun.* **6**, 1–6 (2015).

13. Surwade, S. P. *et al.* Water desalination using nanoporous single-layer graphene. *Nat. Nanotechnol.* **10**, 459–464 (2015).

14. Cohen-Tanugi, D. & Grossman, J. C. Water desalination across nanoporous graphene. *Nano Lett.* **12**, 3602–3608 (2012).

15. Xu, G. R. *et al.* Two-dimensional (2D) nanoporous membranes with sub-nanopores in reverse osmosis desalination: Latest developments and future directions. *Desalination* **451**, 18–34 (2019).

16. Hegab, H. M. & Zou, L. Graphene oxide-assisted membranes: Fabrication and potential applications in desalination and water purification. *J. Memb. Sci.* **484**, 95–106 (2015).

17. Li, H. *et al.* Experimental Realization of Few Layer Two-Dimensional MoS2 Membranes of Near Atomic Thickness for High Efficiency Water Desalination. *Nano Lett.* **19**, 5194–5204 (2019).

18. Mahmoud, K. A., Mansoor, B., Mansour, A. & Khraisheh, M. Functional graphene nanosheets: The next generation membranes for water desalination. *Desalination* **356**, 208–225 (2015).

19. Fang, C., Yu, Z. & Qiao, R. Impact of Surface Ionization on Water Transport and Salt Leakage through Graphene Oxide Membranes. *J. Phys. Chem. C* **121**, 13412–13420 (2017).

20. Haastrup, S., Strange, M., Pandey, M., Deilmann, T. & Schmidt, P. S. The Computational 2D Materials Database : high- throughput modeling and discovery of atomically thin crystals The Computational 2D Materials Database : high-throughput modeling and discovery of atomically thin crystals. *2D Mater.* **5**, 042002 (2018).

21. Rohatdi, A. WebPlotDigitizer. (2019). Available at: https://automeris.io/WebPlotDigitizer.

22. Varoquaux, G. *et al.* Scikit-learn. *GetMobile Mob. Comput. Commun.* **19**, 29–33 (2015).

23. Tishbirani, R. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58**, 267–288 (1996).

24. Morozov, V. A. Regularization of incorrectly posed problems and chois of regularization parameter. *USSR Comput. Math. Math. Phys.* **6**, 242–251 (1966).

25. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net (Journal of the Royal Statistical Society. Series B: Statistical Methodology (2005) 67 (301-320)). *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**, 768 (2005).

26. Fix, E. & Hodges, J. L. Nonparametric discrimination: Consistency properties. *Randolph Field, Texas, Proj.* 21–49 (1951).

27. Breiman, L. Random forests. *Random For.* **45**, 5–32 (2001).

28. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* **42**, 785–794 (ACM, 2016).

29. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics* **117**, 1–19 (1995).

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

31. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

32. Patra, M. & Karttunen, M. Systematic Comparison of Force Fields for Microscopic Simulations of NaCl in Aqueous Solutions: Diffusion, Free Energy of Hydration, and Structural Properties. *J. Comput. Chem.* **25**, 678–689 (2004).

33. Van Duin, A. C. T., Dasgupta, S., Lorant, F. & Goddard, W. A. ReaxFF: A reactive force field for hydrocarbons. *J. Phys. Chem. A* **105**, 9396–9409 (2001).

34. Allen, M. P. & Tildesley, D. J. *Computer simulation of liquids*. (Oxford university press, 2017).

35. Tao, H. *et al.* Two-dimensional materials for energy conversion and storage. *Prog. Mater. Sci.* **111**, 100637 (2020).

36. Yuan, J. *et al.* Room-Temperature Magnetic Order in Air-Stable Ultrathin Iron Oxide. *Nano Lett.* **19**, 3777–3781 (2019).

37. Forouzandeh, P. & Pillai, S. C. Two-dimensional (2D) electrode materials for supercapacitors.

Mater. Today Proc. (2020). doi:10.1016/j.matpr.2020.05.233

38. Huang, L. *et al.* Salt-Assisted Synthesis of 2D Materials. *Adv. Funct. Mater.* **30**, 1–27 (2020).

39. Shen, J. W. *et al.* A molecular dynamics study on water desalination using single-layer MoSe2 nanopore. *J. Memb. Sci.* **595**, 117611 (2020).

40. Joseph, S. & Aluru, N. R. Why are carbon nanotubes fast transporters of water? *Nano Lett.* **8**, 452–458 (2008).

41. Mattia, D. & Calabrò, F. Explaining high flow rate of water in carbon nanotubes via solid-liquid molecular interactions. *Microfluid. Nanofluidics* **13**, 125–130 (2012).