

Analysis of Membrane Process Model from Black Box to Machine Learning

Agnar Alfons Ramel

Faculty of Physical Sciences Mathematics, University of Iceland, Reykjavík, Iceland
Aalfonsramel32@gmail.com

ArticleInfo

Journal of Machine and Computing (<http://anapub.co.ke/journals/jmc/jmc.html>)

Doi : <https://doi.org/10.53759/7669/jmc202202001>

Received 15 April 2021; Revised form 15 July 2021; Accepted 12 October 2021; Available online 05 January 2022.

©2022 The Authors. Published by AnaPub Publications.

This is an open access article under the CC BY-NC-ND license. (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Abstract – The membrane processes include the complex frameworks, typically integrating various physio-chemical aspects, and the biological activities, based on the systems researched. In that regard, the process modeling is essential to predict and simulate the process and the performance of membranes, to infer concerning the optimum process aspects, meant to analyze fouling developments, and principally, the controls and monitoring of processes. Irrespective of the real terminological dissemination such as Machine Learning (ML), the application of computing instruments to the processes of model membrane was considered in the past are insignificant from the scholarly perspective, not contributing to our knowledge of the aspects included. Irrespective of the controversies, in the past two decades, non-mechanistic and data-driven modeling is applicable to illustrate various membrane process, and in the establishment of novel tracking and modeling approaches. In that regard, this paper concentrates on the provision of a custom aspect regarding the use of Non-Mechanistic Modeling (NMM) in membrane processing, assessing the transformations endorsed by our experience, accomplished as a research segment operational in the membrane process segment. Furthermore, the guidelines are the problems for the application of the state-of-the-art computational instruments Membrane Computing (MC).

Keywords – Membrane Computing (MC), Machine Learning (ML), Non-Mechanistic Modeling (NMM), Projection to Latent Structures (PLS), Artificial Neural Networks (ANNs)

I. INTRODUCTION

In the sector of computer science, Membrane Computing (MC) [1] concentrates on the discovery of novel computational frameworks from the research of biological cells, certainly of the cell membrane. It represents a sub-task of establishing a cellular framework. MC deals with the parallel and distributed computing frameworks, processing multiset of the symboling object in the localized mode. Therefore, evolution guidelines permit from the evolving items to be compressed into the compartment illustrated by cell membranes. The communication between a compartment and the ecology plays a fundamental role in the entire process. The several forms of membrane systems are referred to as P systems following Gheorghe Păun's acceptance of the framework in the late 90s [2]. A fundamental component of the P systems is the structure of the membrane that could be a hierarchical basis of the membrane, as for the cell, or the overall membranes (integrate in the graph nodes), as for the neural net or tissues. The P system is typically defined in a graphical manner using drawings.

Statistical frameworks may be built to represent the filtering processes (– for example, permeable, selectively) using well-known physiological equations describing the mechanics included in the membrane system when applicable to membrane technology. Some theories attempt to represent factors that cannot be measured scientifically (or are impossible to measure) and connect them with measurable parameters (typically efficiency metrics) in order to reveal the system' mechanical properties. These systems are said to be mechanical or determinism when they are governed by physical rules. Although there are other approaches to categorize statistical formulas, only the mechanical and non-mechanistic classifications will be considered in this paper.

Non-Mechanistic Modeling (NMM) may be employed in membrane technologies to characterize the motivation and achievement of the networks in addition to mechanistic theories based on established physiological constants [3]. These frameworks depend on experimental results and are often referred to as model parameters; nevertheless, they are also referred to as mathematical analyses when they are constructed using statistical procedures. Non-mechanistic frameworks are considered as data-driven frameworks that are developed using computer techniques and are based on computational formulas. Because the mathematical connection (which is often unknown or concealed in sophisticated procedures) is not motivated by physiological insights, these non-mechanistic systems sometimes are known as black-box frameworks. Nevertheless, words like chemometrics, machine learning and deep learning are presently applied to signify the application of computational and mathematical techniques, and they are widely acknowledged as useful in a variety of fields.

Multiparametric NMM tools typically include Principal Components Analysis (PCA) [4] or the various variants attached to it, the multi-linear instability such as the Projection to Latent Structure (PLS) stagnation (i.e. coefficient of determination regression), and Artificial Neural Networks (ANNs), among others. Machine Learning (ML) approaches, regardless of the modeling tools employed, involve computing techniques, a large quantity of information, statistical measures to evaluate accuracy, and the correlation of a range of constraints (i.e. inputs) in order to project fundamental variables that should be

represented (i.e. outputs). In order to create the designs, such approaches also need tuning (training the formalism of the computational techniques), whereby the approaches are learned based on the experimental findings (i.e. outlet and inlet parameters have to be verified empirically). Following the tuning, a stage of verification (also known as testing) is necessary using a fresh series of laboratory information (not used for tuning). After that, the algorithms may be utilized to project outputs based on the input variable.

Irrespective of the global use of aspects such as Artificial Intelligence (AI) and Machine Learning (ML), only some years back, the application of the algorithmic approach to present linkage between various parameters during the modeling of the membrane technologies was not significantly famous, and was typically ignored as not being significant from the systematic standpoint, insufficient physical and chemical significance, and not making contributions to the knowledge of a phenomenon involved. The purpose of this paper is to provide a subjective and historic viewpoint on the use of NMM in membrane technology. Moreover, the current study aims to demonstrate how sophisticated quantitative techniques (non-mechanistic designs) can be used to design membrane technology in order to solve practical issues (like monitoring system and fouling advancement), as well as the benefits of doing so in circumstances where mechanism of action designs are insufficient due to a lack of knowledge of the bio/physical occurrences. To achieve the purpose of this contribution, this paper has been arranged as follows: Section II focusses on Non-Mechanistic Modeling (NMM). Section III focusses on the concerns of the “Black Box” modeling and precautions considered when using the approach. Section IV evaluates the hybrid modeling technique while Section V focusses on the Non-Mechanistic Modeling (NMM) as a learning approach. Lastly, Section VI draws conclusions to the research and recommends future directions.

II. NON-MECHANISTIC MODELING (NMM)

The amount of work done utilizing traditional mechanical modeling of membrane technologies is enormous [5], and it includes a variety of membrane kinds and membrane activities. In truth, these frameworks and modeling systems are valuable and indispensable; but there are times when understanding of physiological events and interactions is insufficient to construct effective and usable frameworks. Moreover, when working with huge and complicated volumes of data arising from the surveillance of a membrane system, computational methods are required to extract relevant information that can be utilized in an informative manner. NMM seeks to uncover the framework occurrences and behaviors that are intrinsically present in the empirical observations, while mechanical modeling aims to describe the network (and the experimental results) depending on a priori knowledge.

Diverse multivariate approaches may be employed with various purposes in multiple regression frameworks, which tries to investigate data sets made of numerous variables that define the systems. Numerous mathematical methods are offered for data extraction and NMM, and they may be used alone or in conjunction, based on the statistics provided and the modeling aims. Unmonitored multivariate data analytics approaches try to detect variation within data in order to discover trends and/or groupings in the sources of data. For instance, the Principal Component Analysis (PCA) represents a non-supervised modeling instrument, capable of transforming the initial variable in an independent linear integration, known as the Principle Component (PC) [6]. The PCA potentially decomposes the data matrix “ x ” into products of two novel matrices with minimized dimensionality, scores matrix “ c ”, with similar row numbers and the columns as PC, and the loading matrix denoted by “ P^T ”, with similar column number “ x ”, in addition of “ ϵ ” as the noise matrices, incorporating the data variance not illustrated by PCs employed during the process of decay: $x = c.p^t + \epsilon$

In the multi-variant regression approach, more than one variable is connected to the prediction of the output variable, in a similar manner of regression between distinct variables. The multi-variate regression instruments could integrate both the linear and the non-linear functionalities correlating the inputs with the outputs and are referred to as the supervised modeling instruments, because of their capability to train in a simultaneous manner of x and y , correlating the different datasets, and focussing on the prediction of the y matrix (outputs)). The Projection to Latent Structure (PLS) modeling represents a multi-variate regression approach using the liner correlation (multi-linear) based on PCA, whereby the variable spaces are minimized and the covariances between matrix “ x ” and matrix “ y ” maximized. In that regard, redundancies in the output and input data is remove, permitting high robustness whenever there is noise and collinearity in the experimental datasets.

When PLS modeling generates multi-linear regression analyses between distinct sources, additional nonlinear statistical methods may be used to forecast the outcomes. Artificial Neural Networks (ANNs) are well-known techniques that are structured in vertices (inputs, concealed layer(s), and outcomes) to simulate brain synapses, each node containing a mathematical algorithm [7]. The neural network design is built on patterns, and ANNs modeling may be used to show complicated nonlinear relationships among intake and outcomes. When employing spectroscopy methods for monitoring systems, which might result in massive quantities of information, NMM can be effective in extracting some relevant information from a complicated data collection (the spectra). The spectra recorded in these circumstances are not always straightforward to interpret, particularly when the system being observed is complicated and prone to disruption, necessitating the use of data-mining algorithms to quickly extract useful data. The benefit of utilizing computer techniques to extract data from huge data sets is evident, and users generally embrace them since they allow users to retrieve information that would otherwise be impossible to analyse.

Using mixed populations to evaluate an Extractive Membrane Bioreactors (EMBs) for the decomposition of the chlorine chemicals, In the early 2000s, scientists focused on an integration of the Artificial Neural Networks (ANNs) with 2D fluorescent spectroscopy that generates a 3D spectral Excitation-Emission Matrix (EEM) by a screen of distinct emission and absorption of frequencies in a concurrent manner. There are various metrics to track the membrane technologies integrating the biochemical functionalities to potentially evaluate both the biological and membrane performance and

phenomena stimulated by interactions of biochemical medium within the membranes (e.g. fouling and biofilm development). As a result, Goulas, Damicelli, and Hilgetag [8] employed Artificial Neural Networks (ANNs) to obtain data on biofilm growth at the surface of the membrane from complicated fluorescent matrices recorded in real time throughout the EMB procedure. Fig. 1 shows how 2D fluorescent spectroscopic and operational data are used as input data to Artificial Neural Networks (ANNs) to forecast processing efficiency in an Extractive Membrane Bioreactor (EMB).

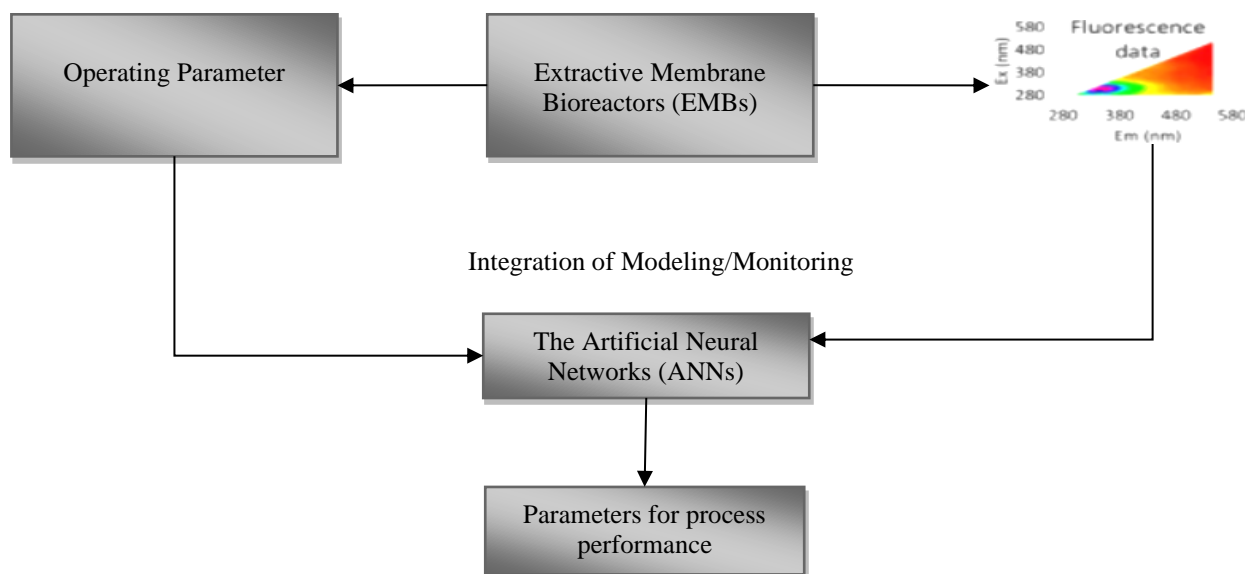


Fig 1. 2D fluorescent spectroscopic and operational data for data input ANNs and EMB

In the early twenty-first century, deep learning and, in specific, Artificial Neural Networks (ANNs) were novel analysis devices in bioengineering and clinical applications, especially appealing because of the complexities of the biochemical systems, the massive numbers of compounds integrated, and the non-linear implications of the biomedical processes. In [9], spectroscopic methods were also being intensively studied as analytic and fingerprint tools to describe biomedical material. Authors in [10] investigated 2D fluorescent spectrometers as a method for detecting specific microbial activity. As a result, it's not unexpected that Wolf and colleagues expanded the use of these identical technologies to study membrane technology including biological reactions. Section III presents an evaluation of the “black box” modeling and the precautions considered when using it.

III. “BLACK BOX” MODELING AND PRECAUTIONS

While using NMM to deconvolute spectral information has obvious benefits, using the same modeling methods to connect operational factors with effectiveness measures has aroused some reservations among scientists. The most obvious issues pertain to the statistical connections' credibility, because they are not established on the physical aspects and require tunings with the experimental findings, making them significantly vulnerable to perturbation in the tuning situations (an extrapolation capacity of these frameworks is limited). The issues with data preparation and selection before the process of learning or tuning, the choice of the internalized structures (latent variable with PLS or the ANNs nodes) of the computational algorithms, and the best validation of the correlation attained could amount to overfitted frameworks with poor predictive capability, or misleading correlation, whereby the random valuation is considered incorrectly correlated. Nonetheless, by following good modeling techniques, non-expert users may readily avoid and analyze typical mistakes in tuning and validation.

The first step in using non-mechanistic frameworks is to choose an appropriate data set to utilize for NMM, as illustrated in Fig. 2. The data collection should be indicative of the mechanism under investigation, with important factors (variables) used to define the system's inputs and outputs. In addition, each parameter should have sufficient variation to enable the statistical framework to “train” on the algorithmic parameters (i.e. the manner in which variations in the chosen parameters associate with the process conditions). In order to prevent input variables with larger values (regardless of variance recorded) having greater influence in the modeling framework than the other variables, an extra phase is typically essential to ensure that an input employed has a similar beginning weight. Even though other scaling procedures could be applied, this phase is typically fulfilled to standardize the scope of every parameter “x”.

The standard value “x” is determined with respect to the standard deviation and mean found for every parameter. Therefore, the variance over every input variable is developed whenever the correlated to the transformations in the performance of the system. Before the tuning of the modeling system, it is also fundamental to ensure that solid testing findings set is issued for the framework's maximum verification. The datasets could be selected randomly from the initial datasets or could be gathered afterwards. Non-mechanistic methods are viable only within calibrated circumstances and

include enough variation to examine the theory in all areas of validation, therefore it should include data from the same set of scenarios as the validation set.

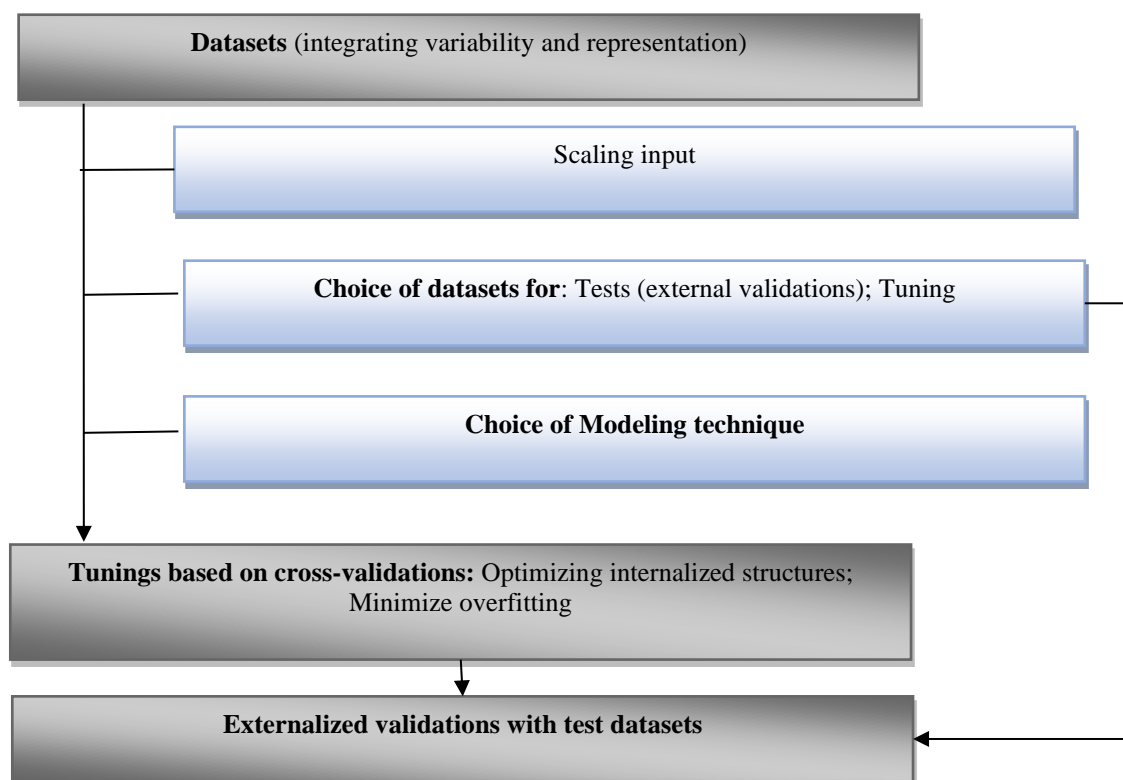


Fig 2. Representation of the summarized NMM process

A crucial stage is deciding which modeling tools are best for the job. Preventing non-linear connections when the network is regular or there is insufficient information, for example, increases the likelihood of excellent modeling by reducing the risk of data overfitting. Additionally, cross-validation may be used to evaluate the tuning technique. The “Leave One Batch Out” and the “Leave One Out” approaches are often used in cross-validation, with the goal of determining how well various subgroups of the calibrated set operate in estimating the information “left out”. The cross-validation aspect is employed in refining the interior structures of the approach and minimize the overfitting of data by running it consecutively through the whole calculation data set.

Aside from the need of exterior verification using the sample set of data, it is critical to precisely analyze the degree of fitting and estimate capabilities of the algorithms that have been created. As a result, utilizing appropriate variables to infer the accuracy of the designs is another crucial factor when employing NMM. Mean errors for alignment, cross-validation, and testing set (typically computed using root average squared errors: RMSEP, RMSECV, RMSEC), and the match qualities, as evaluated by the coefficient of correlations (R^2) for similar datasets, should also be considered. Additional quantitative measures (e.g., the Akaike data criteria) may also be used to evaluate the validity of systems, albeit their use typically necessitates bigger data sets and beginning data with specified mathematical behaviour, like Gaussian dispersion.

A typical critique of such systems i.e. “black box” techniques, do not add to the body of knowledge of the physical occurrence linked to the operation, and have a restricted applicability. The study of features utilized in computing methods, on the other hand, may be used to learn from the connection discovered. This may be done by evaluating the formulas generated by tools like PLS regression (also known as grey box analysis since it generates formulas) or by doing a sensitive assessment of the intake factors and output variables.

The influence of each input on the ANNs may be determined by doing a sensitivity analysis on each output value before averaging all observations. The standardized proportion among the output results and the input variables is used to determine sensitivities for each quantity. Knowing which inputs are significant in predicting outcomes, along with their corresponding weights in the model, gives you a mechanistic understanding of the processes you're modeling. In reality, many literatures aimed at modeling membrane technology also revealed this sort of fundamental understanding, offering more than a good tool for projection. Section IV below focussed on hybrid modeling architectures.

IV. HYBRID MODELING

Hybrid modeling is the use of both mechanical and non-mechanistic frameworks to simulate a mechanism or activity. In optimization methods, the way each kind of functions (mechanism of action or non-mechanistic) interactions might alter depending on the modeling goals. Hybrid model architectures are often divided into two categories: serial and parallel (see

Fig. 3). When the non-mechanistic framework is normally employed to predict the input essentialized by scientific frameworks in serial arrangement (see Fig. 3a), in series arrangement, both the non-mechanical and mechanistic frameworks are integrated by an impulse. In general, the parallel arrangements, NMM anticipates the mechanistic framework's deviance, and the mechanical modeling is modified to enhance output predictions by integrating both frameworks. The logic behind such a hybrid modeling technique is that the particles from the physiological approach are not noisy; they contain information that the non-mechanistic theory may find. Moreover, the returns from the physiological framework could be defined based on similar "x" as a mechanistic framework or/and novel input "w, which could signify a missed information (see Fig. 3b).

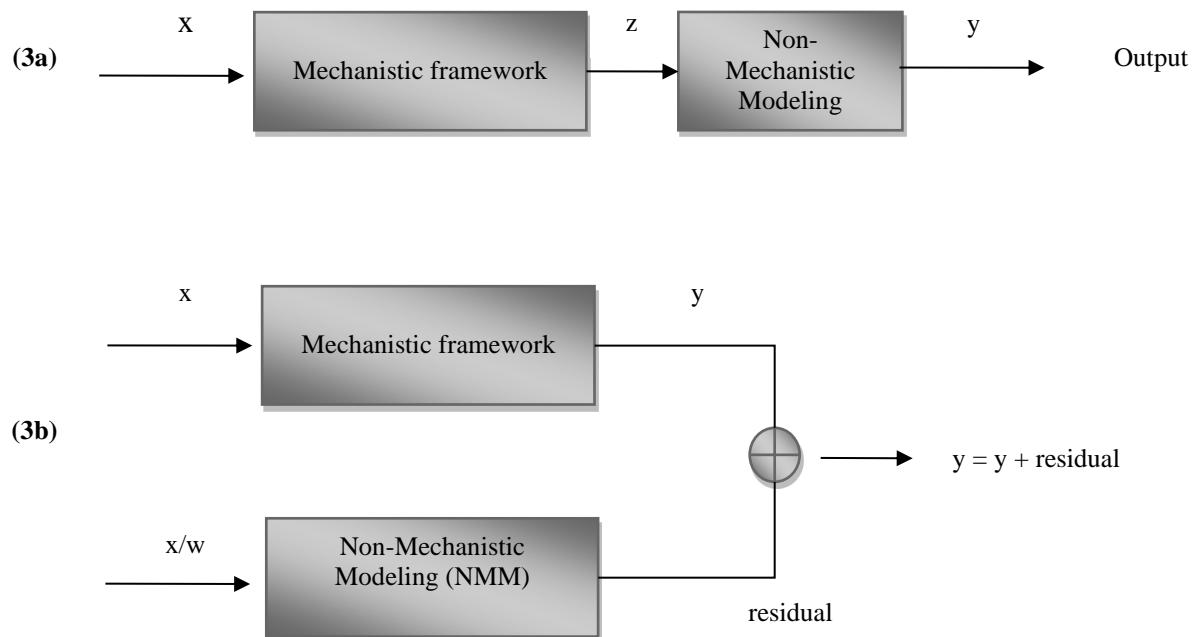


Fig 3. The hybrid modeling structures (3a) serial; (3b) parallel

To characterize the solvent transportation via nanostructured materials, [11] evaluated the application methodological, hybrid and NMM methodologies. Option membrane and solvent attributes were chosen for that study, which were used as input in the model's framework using PLS analysis, ANNs, and PCA mixed with ANNs (Fig. 4). Non-mechanistic systems with descriptive variables as inputs were employed in this concurrent hybrid technique to calculate the variances of the solution-diffusion model. Not only did the hybrid modeling improve the characterization of the experimental data (which were gathered from several editorials), but it also enabled for the identification of which factors were more important for production efficiency, namely those whose participation was not taken into account by the solution framework. As a result, the input contributions to the PLS equation were examined, and it was discovered that the physiological model utilized lacked additional significant data on polar solvents.

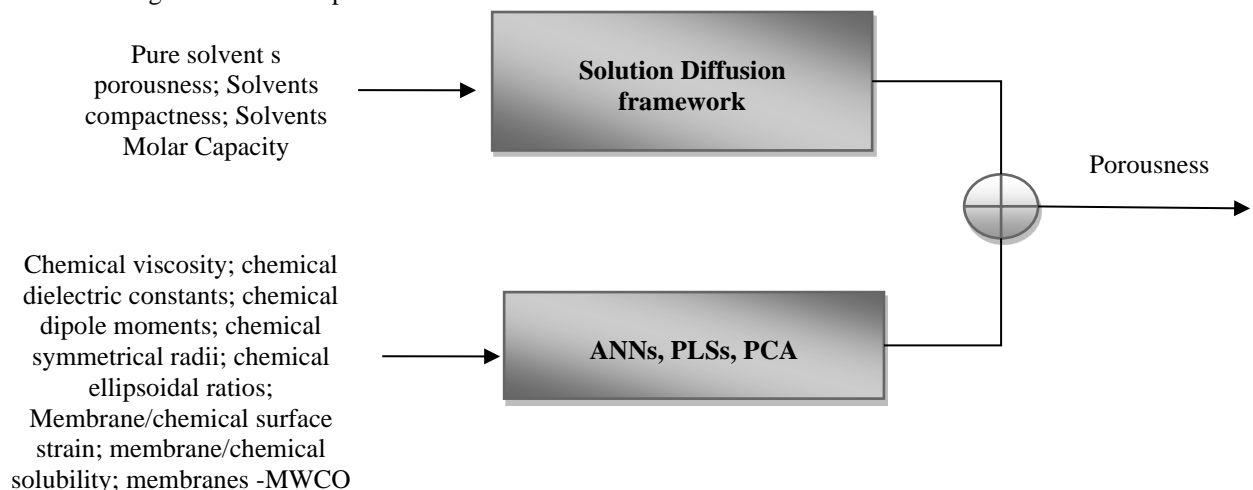


Fig 4. The hybrid modeling structures employed for the forecast of solvent permeability via nano-filtration membrane [ANNs, PLSs, PCA]

In another membrane system, hybrid modeling was utilized to increase the modeling capabilities of the Activated Sludge Model (ASM) that is typically employed to evaluate biological efficacy in handling wastewaters, when it was implemented to a bioreactor with little extra growth to know. The Malik, Arslan, Kim, Jun, and Park in [12] employed a concurrent hybrid technique to enhance the predictions of precipitated particles, biochemical oxygen demands in contamination, and nitrite + nitrate in effluent in this contribution. ASM fell short of detailed and actual-time datasets from the runoff and operational variables since it was validated using starting circumstances (since the first 50 days of service) and then operated for 400 days without subsequent recalculation. The goal of this study was to enhance estimation accuracy without conducting time-consuming analytical experimentations. To do so, 2D fluorescent spectrophotometric information recorded instantly in effluent, wastewaters, biological media were integrated with PCA in addition to PLS modeling strategy to approximate covariances from the ASM. By adding fluorescent data received from real-time observation of the operation, the hybrid model enabled ASM to be improved. Moreover, combining operational conditions and statistical results in hybrid modeling did not enhance estimations considerably, demonstrating that data collected from 2D spectrophotometric analysis is adequate to capture the fluctuations in artificial media that the physiological model fails to detect.

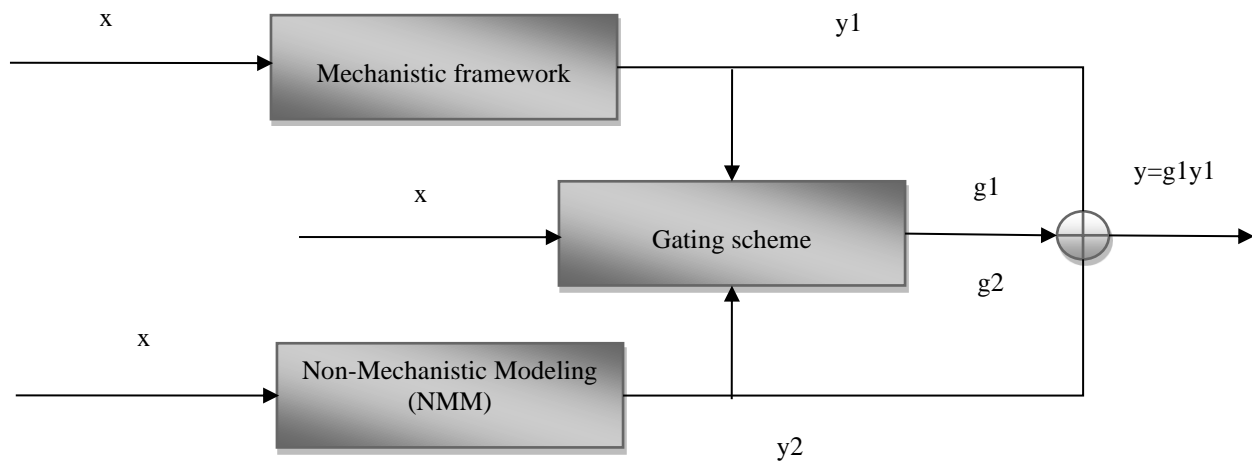


Fig 5. A competing mixture-of-experts framework with parallel hybrid modeling, where a filtering mechanism determines the contributions of each theory to the ultimate forecast.

Another study employed hybrid modeling to calculate counterion flux over the ion-exchange membranes in the biofilm reactors supported by membranes. A hybrid approach was utilized in compensation for the intricate biofilm contribution for transportation. In that analysis, a competing mixture-of-experts framework was used in conjunction with the concurrent hybrid approach, in which the contributions were mixed and the residue of the physiological transportation framework was estimated using PLS prediction (see Fig. 5). The “mixture of expert” design employed both the non-mechanistic and mechanical techniques to predict counterion mass transfers over the membrane with gating scheme determining the application of on or the other based on the contribution of an intake (the operational situation) to the outputs. Employing the parallel approach, a PLS regression analysis was capable to potentially obtained missing datasets within the physiological framework from the real operational dataset. The approximation was insufficient for particular counterions because of the residual variances integrated by the physiological model in PLS tunings. As a consequence, the restrictions of the narrative structures were circumvented, and the competing mixture-of-experts design was chosen as the optimum modeling approach in this circumstance.

As demonstrated for the membrane filtration, hybrid approaches provided a proper understanding of the systems compared to the non-mechanistic systems, not just as a result of the mechanisms of the action designs employed known physiological connection, but also due to the integration of the non-mechanistic frameworks, it is probable to identify “why” or “when” the metabolic designs stop identifying the information that is lacking within the mechanistic frameworks (via the novel input chosen or distinct correlation).

Moreover, these techniques are valuable in revealing whenever computational complexities being described is greater than the intricacy of the available mechanistic frameworks. As a result, hybrid modeling may be utilized in cases when mechanistic frameworks are insufficient to represent the system owing to the intricacy of connections or by expanding the usage of frameworks to operating settings other than those assumed by the systems (needed for the mechanistic frameworks). NMMs, on the contrary, often necessitates a considerable quantity of experimental results and has restricted extrapolating potential, both of which may be avoided by including a physiological model (even if employing simplified hypotheses) into a hybrid modeling framework.

Nonetheless, when a physiological paradigm is insufficient to represent a process and hybrid modeling is possible, the feasibility of employing purely NMM should be evaluated. Non-mechanistic approaches may actually simplify statistical connections (by directly connecting input variables with outputs) without having various algorithms to monitor the operation (as in optimization methods). As a result, before pursuing a hybrid modeling technique, the modeling work's goals should

be clearly specified, and the usage of just physiological and NMM should be examined first. Aside from their usefulness, combining non-mechanistic and mechanistic frameworks promotes user acceptance by making the mechanical principles and limits that need the addition of non-mechanistic concepts simpler to grasp. Section V focuses on the Non-Mechanistic Modeling (NMM) as learning.

V. NON-MECHANISTIC MODELING (NMM) AS LEARNING

When it comes to modeling membrane mechanisms, the need for NMM can originate from a desire to include non-traditional inputs as framework identifiers, like spectral data, online statistical results, or structures attributes requirements (as with hybrid modeling), or to deconvolute intricate, big quantities of information arising from network surveillance.

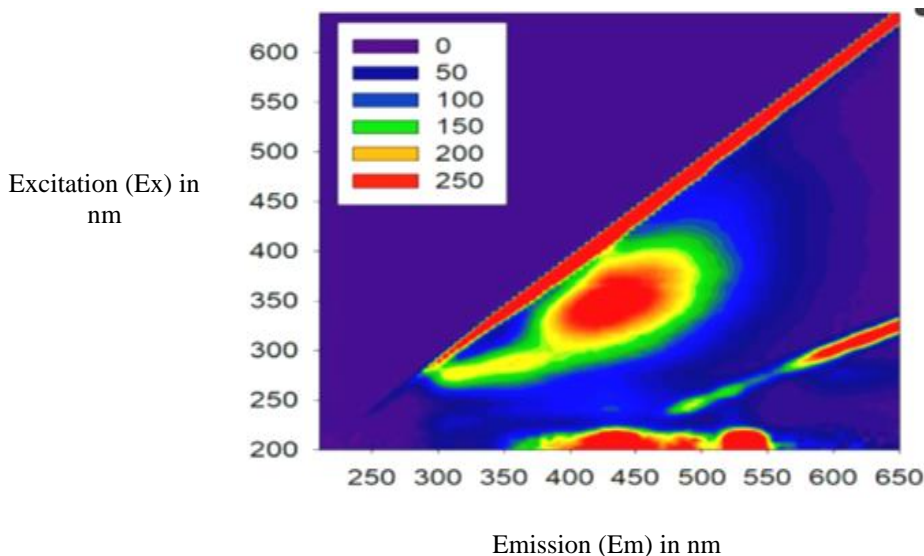


Fig 6. A contoured map of a microbial media's light excitation-emission matrix (EEM). The axes represent the frequencies of emission (Em) and excitation (Ex) in nanometers, while the color scale shows the intensity of emissions

Indeed, our foray into NMM began with the goal of employing 2D fluorescent spectrometers to track membrane operations including biological interactions. Due to the presence of various natural fluorophores, such as amino acids, NADH, and pigments, 2D fluorescence is a method that may collect a great quantity of data about natural systems without affecting the mechanism (Fig. 6). An optic probe may be used to capture fluorescent excitation–emission matrices (EEMs) to analyze the operation effectively from the medium without collecting, and the probe can also be pointed towards the surface of the membrane to gather information instantly from it. Because this approach is responsive to the existence of numerous substances that react with the fluorophores, affecting their fluorescent response, and/or with excitement and emitted light, the spectral analysis depicts not only the fluorophores available, but also other media characteristics (e.g., inner-filter effect). While 2D fluorescence can analyze a huge quantity of data, the EEMs obtained are complicated, and relevant data cannot be recovered directly owing to the many inferences present. NMM should therefore be utilized to extract useful data not just from fluorophores, but also from disruptions, which reflect the existence of other substances.

Different modeling techniques have previously been investigated when utilizing NMM to deconvolute light spectra for monitoring membranes activities. The mathematical techniques were first used to connect 2D light with chemical concentrations or process efficacy metrics, which is a more understandable approach. 2D light spectra obtained at various regions of the surface of membranes, whereby biofilm establishes, we employed as data inputs within ANNs to potentially forecast dichloroethane, ammonia and chloride within the bioreactor's liquid mediums whereas modeling an EMB [13] for a degrade of chlorine organic chemical. After considering the probabilities of the spectral subtractions and considering that non-linear classification technique could be maximum for fluorescence spectra demodulation, the usage of ANNs was chosen in this study. A similar method was utilized in Membrane Bioreactor (MBR) [14] for treating wastewater to estimate Chemical Oxygen Demand (COD) in each flow by connecting 2D light spectra collected in the permeate effluent streams and effluent permeation stream. The non-mechanistic modeling approach utilized in this investigation was PLS modeling. This instrument was selected due to its multi-linear instruments that are less intricated compared to non-linear ANNs and might potentially minimize the aspect of overfitting and issue comprehensible mathematical evaluations.

Despite this, many literatures highlighted how the non-mechanistic frameworks could assist in the elucidation of system processes, in addition to estimating levels in liquid media. The estimated values of the PLS multilinear equations may be employed to assess the significance of every input variable, i.e. every illumination or emitted frequency pair, and hence the contributions of each section of the fluorescence emission, in predicting each output molecule. A sensitivity analysis may be done with ANNs to determine which input parameters have the greatest impact on the output variable predictions. The application of principle components analysis improved the non-mechanistic deconvolutions of the light spectroscopic in various modes of biochemical procedures, Membrane Bioreactors (MBRs), including the Extractive Membrane Bioreactors (EMBs). Prior to connecting fluorescence information with production efficiency, this was done to compress the set of

statistics into a smaller variety of parameters (Principal Component Analysis - PCA). Regardless of the modeling method employed (ANNs or PLS), this strategy simplifies and minimizes the quantity of inputs required for model training, lowering computing effort. Additionally, since just the major variance collected in the light emission is employed in the final connections, it allows for correlation reduction and can deal with noise reduction a priori.

VI. CONCLUSION AND FUTURE DIRECTIONS

Currently, it is easy to visualize the non-mechanistic instruments, evaluate and study the results, and trust the findings from practical application due to the multiple membrane process tests, which have been successful to a while now. In that regard, Machine Learning (ML) techniques could be employed to mitigate the problems like tracking the process performances, tracking the fouling development, and optimizing the membrane process by establishing critical forecasts with respect to different input datasets. Non-Mechanistic Modeling (NMM) techniques have been indicated to have the capability to change the process of tracking datasets into process performance features in actual time, allowing for more state-of-the-art process controls. Moreover, the capability of ML approaches to integrate multiple datasets, and integrate distinct elements (e.g. hybrid modeling) is merit in the establishment of the control instruments for membrane procedures. The instruments based on modeling methods can be employed in future for digital monitoring prediction development, learning, and continuous upgrading of dynamic modeling systems. Lastly the approaches employed to collect more data from NMMs employed to membrane processes demonstrate the powers of ML to assist of process optimization.

References

- [1]. R. Ceterchi, L. Zhang, K. G. Subramanian, and G. Zhang, "Hilbert words as arrays generated with P systems," *J Membr Comput*, vol. 3, no. 3, pp. 163–169, 2021.
- [2]. M. Gheorghe, A. Păun, S. Verlan, and G. Zhang, "Membrane computing, power and complexity," in *Encyclopedia of Complexity and Systems Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2017, pp. 1–16.
- [3]. G. Wolf, J. S. Almeida, M. A. M. Reis, and J. G. Crespo, "Non-mechanistic modelling of complex biofilm reactors and the role of process operation history," *J. Biotechnol.*, vol. 117, no. 4, pp. 367–383, 2005.
- [4]. R. W. Farebrother, "Notes on the prehistory of principal components analysis," *J. Multivar. Anal.*, no. 104814, p. 104814, 2021.
- [5]. T. Singh and N. Kaur, "Layered axial force coupled membrane based metal contact single-pole quad-throw RF MEMS switch: design, RF performance and mechanical modeling," *Microsyst. Technol.*, vol. 22, no. 8, pp. 2117–2123, 2016.
- [6]. Y. Nakayama, K. Yata, and M. Aoshima, "Clustering by principal component analysis with Gaussian kernel in high-dimension, low-sample-size settings," *J. Multivar. Anal.*, vol. 185, no. 104779, p. 104779, 2021.
- [7]. O. Solopchuk and A. Zénon, "Active sensing with artificial neural networks," *Neural Netw.*, vol. 143, pp. 751–758, 2021.
- [8]. A. Goulas, F. Damicelli, and C. C. Hilgetag, "Bio-instantiated recurrent neural networks: Integrating neurobiology-based network topology in artificial networks," *Neural Netw.*, vol. 142, pp. 608–618, 2021.
- [9]. R. Silvers and M. T. Eddy, "NMR spectroscopic studies of ion channels in lipid bilayers: Sample preparation strategies exemplified by the voltage dependent anion channel," *Methods Mol. Biol.*, vol. 2302, pp. 201–217, 2021.
- [10]. L. Monico et al., "Development of a multi-method analytical approach based on the combination of synchrotron radiation X-ray micro-analytical techniques and vibrational micro-spectroscopy methods to unveil the causes and mechanism of darkening of 'fake-gilded' decorations in a Cimabue painting," *J. Anal. At. Spectrom.*, 2022.
- [11]. A. V. Mitkari, Nanostructured Thin Film Material Laboratory, Department of Physics, Govt. Vidarbha Institute of Science and Humanities, VMV Road, Amravati 444604, Maharashtra, India, A. U. Ubale, and Nanostructured Thin Film Material Laboratory, Department of Physics, Govt. Vidarbha Institute of Science and Humanities, VMV Road, Amravati 444604, Maharashtra, India, "Thickness dependent physical properties of SILAR deposited nanostructured CoS thin films," *ES Mater.Manuf.*, 2019.
- [12]. S. A. Malik, S. Arslan, H. W. Kim, J. Jun, and H. Park, "Hybrid concurrent driving technique for large touch screen panels," in *2019 International SoC Design Conference (ISOCC)*, 2019.
- [13]. L.-F. Ren, C. Liu, Y. Xu, X. Zhang, J. Shao, and Y. He, "High-performance electrospinning-phase inversion composite PDMS membrane for extractive membrane bioreactor: Fabrication, characterization, optimization and application," *J. Memb. Sci.*, vol. 597, no. 117624, p. 117624, 2020.
- [14]. Y. Ding et al., "The influence of different operation conditions on the treatment of mariculture wastewater by the combined system of anoxic filter and membrane bioreactor," *Membranes (Basel)*, vol. 11, no. 10, p. 729, 2021.