



HAL
open science

Toward mechanistic models “augmented” by machine learning. Example of a drying simulation data set exploited by a Physics Informed Neural Network

Patrick Perre

► To cite this version:

Patrick Perre. Toward mechanistic models “augmented” by machine learning. Example of a drying simulation data set exploited by a Physics Informed Neural Network. *Drying Technology*, 2024, pp.1 - 15. 10.1080/07373937.2024.2411582 . hal-04759083

HAL Id: hal-04759083

<https://hal.science/hal-04759083v1>

Submitted on 29 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Towards mechanistic models "augmented" by machine learning. Example of a drying simulation data set exploited by a physics informed neural network

Patrick Perré ^{a,b},

5 ^aUniversité Paris-Saclay, CentraleSupélec, LGPM, Centre Européen de Biotechnologie et de Bioéconomie (CEBB), 3
rue des Rouges Terres 51110 Pomacle, France

^bUniversité Paris-Saclay, CentraleSupélec, LGPM, Gif-sur-Yvette, France

Abstract

Mechanistic modeling of drying is well-established since several decades. Based on fundamental
10 balance equations and driven by relevant material parameters, it can predict the entire process,
including configurations that were not observed before. Besides, thanks to their ability to tackle
non-linear and dynamics problems, approaches based on Machine Learning (ML) based are capa-
ble of coping with complex situations even better than mechanistic modeling. The main drawback
of mechanistic models is their complexity as operational tools, namely in providing the whole set
15 of product characteristics, while the main drawback of ML tools is its restriction to the domain
paved by the data set. This paper summarizes the physics of the mechanistic formulation and
then presents the different possibilities of coupling the mechanistic and ML approaches to obtain
an "augmented" mechanistic model. The idea is to merge the advantages of both worlds. Different
strategies can be imagined:

- 20 • A full coupled method (hybrid model)
- A fully decoupled method
- A cascade coupling

The second part of the paper gives examples of a fully decoupled approach: the mechanistic model
is used to populate a data set, which is further exploited by a neural network. The originality of
25 the work is to compare a classical neural network with a physics informed neural network.

Keywords: coupled transfers, drying rate, hybrid model, neural network, soft sensor.

1. Introduction

Drying is a complex process involving many coupled and non-linear thermal, physical,
mechanical, chemical, and biological phenomena. Due to the coupling between heat and
30 mass transfer, spatial fields of moisture content (MC) and temperature develop inside
the product. The drying quality depends primarily on transfers as the moisture content
and temperature fields trigger shrinkage, thermo- and hydro-activation, thermal degra-
dation, internal pressure... For example, the transfers are also responsible for the chemical
degradation due to the coupled effect of hydric and thermal activation over time, for the
35 final MC heterogeneity, for the memory of mechanical behavior, for the development or
inhibition of biological activity. Deformation, internal stress, product failure (collapse,
checking) and poor re-hydration are among the major challenges encountered in drying
quality. Solid mechanics is needed to address these aspects. The driving force for stress
development is the shrinkage induced by water removal. During drying, the presence of

Email address: patrick.perre@centralesupelec.fr (Patrick Perré)

40 internal stresses induces creep and/or plastic deformation, that vary in space and time. The mechanical behavior also depends on temperature and moisture content. All these aspects give rise to a complex and intricate coupling, explaining stress reversal and case-hardening. Many aspects impact the quality and cost of the drying process: drying, time, energy consumption, product quality, impact of using intermittent energy... Most aspects
45 are contradictory and compromised are not easy to define, prior or during the process. Modeling these complementary and strongly coupled aspects (transfer, mechanics and additional phenomena) provides a comprehensive and in-depth view of the process and allows optimization/adaptation to current situations. This is particularly welcome in the current context of high energy prices, intermittent energy, and the need to save food and
50 raw materials.

This complexity explains why machine learning is increasingly used in drying, to predict drying traits such as moisture content or drying time, but more often to tackle more complex aspects such as product quality [1, 2, 3, 4, 5, 6]. These works mainly rely on the use of ML instead of mechanistic modeling.

55 On the contrary, this document aims to promote strategies capable of using ML in synergy with mechanistic modeling, to proposed an "augmented" mechanistic model, in the spirit of what is proposed in [7]. Among all phenomena involved to get a comprehensive mechanistic model of drying [8], this paper is focused on coupled heat and mass transfers. These intricate transfer are first outlined to better illustrate the possibilities offered
60 by machine learning (ML). In this section, we will concentrate on the potential combined contributions of ML to MM, i.e. different pathways able to merge the best of the two worlds. Among the different possibilities, examples of fully decoupled approach is proposed: generation of a data set by MM and subsequent exploitation of this data set by ML: classical Neural Network and Physics Informed Neural Network.

65 2. The comprehensive macroscopic formulation

The foundations of coupled transfer in porous media have been laid by the pioneering works undertaken in the 50s and 60s [9, 10, 11]. The set the comprehensive set of macroscopic equations governing coupled heat and mass transfer in porous media was then rigorously derived by Whitaker [12] using the method of volume averaging [13, 14]. The
70 development of volume-averaged transport equations requires the existence of a representative elementary volume (REV) V sufficiently large to smooth the microscopic fluctuations and sufficiently small to avoid macroscopic variations. The full set of equations [15] adapted to the case of hygroscopic products [16] can be considered as the state-of-the-art of the macroscopic formulation. It includes three state variables (for example moisture
75 content, temperature, and total pressure) and can describe most drying processes, including those which give rise to an internal over-pressure (high temperature drying, vacuum drying, microwave drying). This comprehensive formulation can be found in [8]. For a better understanding, the different modes of moisture transfer embedded in these equations are described in figure 1.

80 The three basics modes of migration inside a solid are induced by a gradient of MC:

- Liquid migration driven by capillary pressure,
- Binary diffusion in the gaseous phase, a crossed diffusion of water vapor and air,

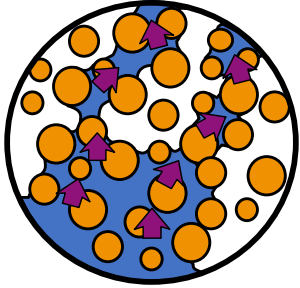
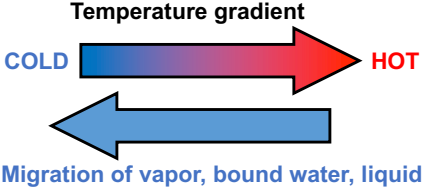
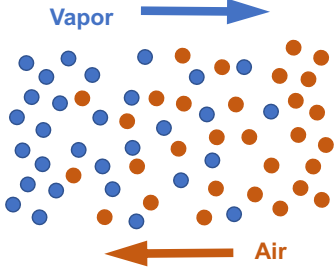
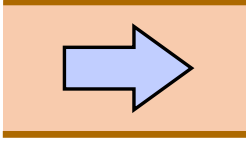
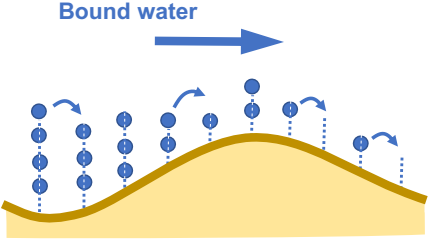
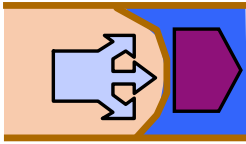
Gradient of moisture as driving force	Other driving forces
<p data-bbox="379 340 647 371">Capillary migration</p> <p data-bbox="245 389 782 510">Liquid water migrates from zones with large MC, hence large menisci to zones with low MC, where suction is larger due to smaller menisci.</p> 	<p data-bbox="954 340 1203 371">Thermo-migration</p> <p data-bbox="810 389 1347 577">Temperature increases water vapor migration (effect of T on P_{vs}), lowers capillary pressure (effect of T on surface tension) and activates bound water. Whatever the nature of water, moisture always migrates from hot zones to cold zones.</p> 
<p data-bbox="405 851 625 882">Binary diffusion</p> <p data-bbox="245 900 782 990">In the gaseous phase, binary diffusion takes place. The driving force is the gradient of molar or mass fraction.</p> 	<p data-bbox="944 851 1216 882">Convection of vapor</p> <p data-bbox="810 900 1347 1088">In high temperature configurations, the internal temperature can be higher than the external boiling point of water. A gradient of total gaseous pressure develops, which is an efficient driving force to enhance vapor transport (together with air) by Darcy's law.</p> 
<p data-bbox="363 1299 667 1330">Bound water diffusion</p> <p data-bbox="245 1348 782 1500">Bound water molecules can jump from one sorption site to the other. This is a statistical behavior activated by both temperature and MC (the molecules mobility increases for less bound molecules).</p> 	<p data-bbox="944 1299 1216 1330">Convection of liquid</p> <p data-bbox="810 1348 1347 1568">In the presence of a gradient of total gaseous pressure (high temperature configuration) and when MC is still high, the gaseous pressure can act on menisci to drive liquid water. The driving force is the gradient of liquid pressure (gaseous pressure - capillary pressure).</p> 

Figure 1: The different physical mechanisms able to drive moisture in a solid that are embedded in the comprehensive macroscopic formulation (adapted from [8]).

- Bound water diffusion, resulting from statistical jumps of water molecules bound to the solid

Apart from these MC gradient-triggered moisture fluxes, additional mechanisms might be of huge importance:

- Thermo-migration refers to a flux of moisture induced by a gradient of temperature. This is very likely to occur in a solid during drying. Whatever the state of water (gaseous, bound or liquid), the flux is always against the temperature gradient (moisture migration from hot to cold),
- High temperature configuration ($T > 100^{\circ}\text{C}$ or vacuum drying), for which a gradient of total pressure develops and gives rise to an additional, very efficient, driving force. This can be obtained by a high external temperature or by volumetric heating (high frequency or microwave heating for example).

All these phenomena, together with the coupling with thermal transfer generated through the latent heat of evaporation, act together in space and in time during drying. This heat and mass coupling is part of the complexity of the drying process and need care in terms of computational methods. Over the decades, the tremendous increase of computer power together with the use of efficient mathematical solvers (full Newton-Raphson, exponential integrator, Jacobian-free Newton-Krylov framework) completely changed the situation: the CPU time needed to solve the 3-variable model in 1D, dropped from one day to 0.1 second! Full 3D models and dual-scale configurations (a microscopic special field at each point of the macroscopic mesh) were progressively achieved (see [8] for further detail).

Most products exhibit a reduction of size, shrinkage, when water is removed. During drying, the local shrinkage, as defined by free shrinkage induced by the local MC value, does not satisfy the compatibility conditions (the strain cannot be integrated into a displacement field). Consequently, an additional strain field, related to a stress field, develops for the total field to be compatible. In turn, this stress level produces additional strain tied to the constitutive law of the product: viscoelastic creep, mechano-sorptive creep, plastic deformation... These additional strain fields act on the drying stresses and a complex process takes place over time. These aspects, with their different levels of coupling with heat and mass transfer, must be kept in mind as they are of crucial importance to the quality of the drying process. These mechanisms are well-documented in the literature [17, 18, 19, 20, 21] and will not be discussed further here.

3. Combining Machine Learning and mechanistic modeling

By its predictive capability, the mechanistic modeling of drying is already a tool for innovation in drying. However, despite its great potential, the state-of-the-art of the simulation of solids drying presents some limitations for its use as an operational tool:

1. The need for physical and mechanical characterization (all parameters involved in the formulation) makes it very difficult to extend the model to different products.
2. The need to include additional features for predicting the drying quality (chemical degradation, color, collapse, biological and microbiological activity...).

3. In certain configurations, non-local equilibrium requires a dual-scale modeling framework to be implemented [22].
4. The lack of standard software, to solve this complex physics limits the extension of modeling and simulation in practice.
- 130 5. The computational time, namely for 3D simulation or dual-scale modeling (dual-scale medium or packed-beds/beds of particles).

This explains that such tools are still rarely used in industry for process optimization and control-command. However, the situation is evolving. For example, the code *TransPore* was recently used to identify moisture transfer properties from drying tests [23], addressing point 1, or to adapt wood drying schedules to intermittent energy [24]. We believe that the fast-growing field of data science will be an excellent driver for the dissemination of mechanistic modeling in industry. Over the past decades, thanks to its ability to tackle non-linear and dynamics problems, machine learning, namely neural networks (NN), was amazingly successful in several fields such as autonomous cars, face recognition and weather forecasting without solving any physical equation. ML has also spread to the drying field [2, 4]. However, in most cases, ML is used instead of mechanistic modeling. In this sense, ML is capable of coping with complex situations, provided that the training data set is large enough to cover the situations of interest: as main drawback, the predictive capability is restricted to the domain paved by the data set. Rather than using ML instead of mechanistic modeling, we believe that both areas are now mature enough to benefit from the best of both worlds: a predictive model capable of adapting to different products. Such a digital twin approach would work both offline or online [7]. In this hybrid approach, the mechanistic model would remain at the heart of the interactions (Fig. 2).

150 ML-based models are easily operational and can be used in process control, based on an existing data set or real-time sensors, or both (arrows 2 of Fig. 2). As main drawback, the predictive capability of pure ML-based models is limited to situations than have been observed many times before. On the contrary, even though they remain complex as operational tools, mechanistic models can simulate situations that have never been encountered before: they can be used both for control/command and innovation (arrows 3 of Fig. 2). Drying is a process for which the product quality depends on the whole process history and is therefore difficult to predict different scenarios without a mechanistic approach. This is the reason why both areas should work together. For example, energy concerns are likely to increase in the future. Drying control will have to account for the availability of energy, in terms of cost and quantity. These ever-changing situations could be handled by hybrid models, combining mechanistic modeling and ML. The mechanistic core will allow several scenarios to be predicted and tested *in silico* by ML algorithms to choose the best one, dynamically. The quality of real-time information is crucial. ML excels at interpreting complex information gathered in real time during the process, such as spectral information, computer vision or combination of several variables, known as soft sensors. Applied to drying, ML can extract real-time information such as drying time, temperature, moisture content, color, deformation or product quality. The combination of mechanistic modeling and ML can be organized in different ways:

- 170 1. **A fully coupled method (hybrid model):** ML exploits real-time information for on-line tuning of the mechanistic model. The model becomes more and more accurate with time (arrows 5 of Fig. 2) and can be used to test several strategies regarding en-

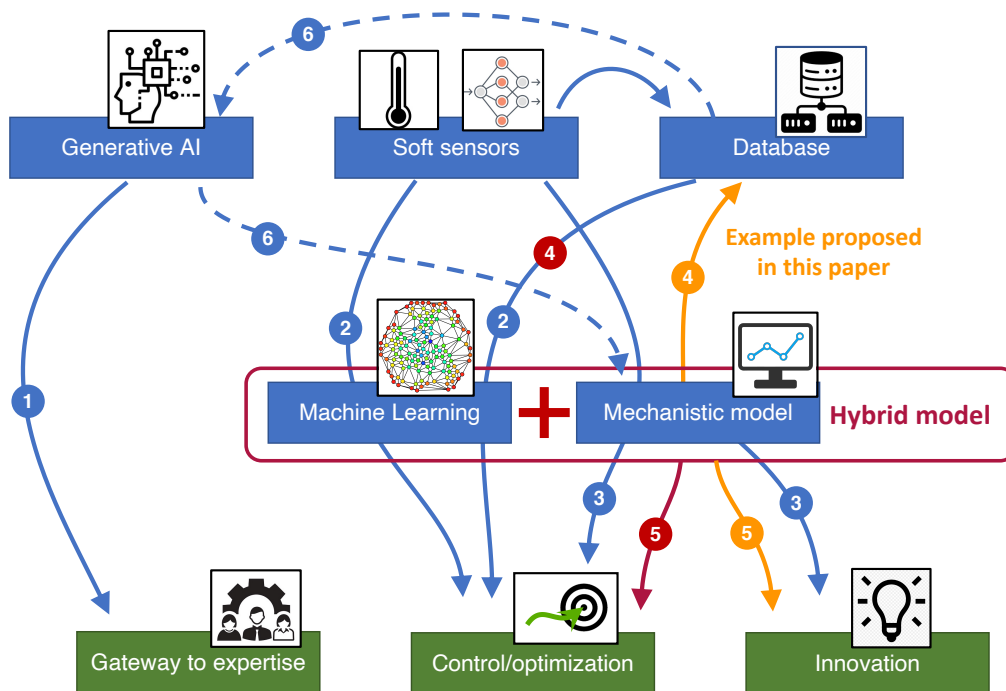


Figure 2: Synoptic diagram of the benefits of AI in the domain of drying for three fields of interest: gateway to a first expertise, control/optimization of an existing process and innovation (adapted from [25])

ergy demand, drying time and product quality and select the best condition based on multi-objective optimization.

- 175 2. **A fully decoupled method** (arrows 4 of Fig. 2): The prediction capability of the mechanistic model is used to generate a data set by numerous simulations covering a wide range of conditions. The final approach is a black box, but a black box trained on a mechanistic data set. The generation of the data set can be very CPU-intensive but, once done, the ML model trained on this data set is very fast.
- 180 3. **A cascade coupling**: The mechanistic model is “augmented” by ML used to extend the outputs of the mechanistic model. For example, the prediction of product quality by a mechanistic model (drying stress, thermal degradation, product shape, stress reversal, color change, biological activity, etc.) is by far more challenging than that of moisture content or temperature. Yet, the prediction of the final product quality by ML is likely to be more relevant if inferred from the evolution of temperature and moisture content fields over time than directly from the drying conditions.
- 185

In the not too far horizon, we could also imagine generative AI being powerful enough to revolutionize this complex coupling between mechanistic modeling and data science. Generative AI will certainly be able to derive a set of partial differential equations governing the problem from a large data set, then generating the code capable of solving it in a predictive way and using this simulation tool in a dynamic loop, considering real-time information and the forecast of exogenous parameters (arrows 6 or Fig. 4). Then, the initial data set would be dynamically populated by further real observations or simulated configurations.

For illustrative purposes, the following sections show how a neural network can be used to exploit a data set generated by mechanistic modeling.

4. Generation of a data set using the mechanistic model

By accounting all transfers phenomena in computationally solved in a coupled way, a mechanistic model is able to produce a comprehensive description of the physics of drying: all variable fields, including temperature, moisture content and internal pressure can be simulated over time in 1D, 2D or 3D [26]. Figure 3 illustrates the coupling between variables that arise during high temperature (which triggers internal vaporization) of a strongly anisotropic material (softwood here).

Using a simpler example (1D drying at low temperature), this section illustrates the potential of a fully decoupled approach. A driver was written to populate a data set with the code *TransPore*. In this example, a 20-mm thick board of softwood was simulated with varying conditions: Temperature (T), relative humidity (RH) of the air and initial moisture content (MC_{ini}). These values were generated randomly using a uniform distribution in selected ranges:

- Training data set 1: $20\text{ }^{\circ}\text{C} < T < 60\text{ }^{\circ}\text{C}$; $20\text{ } \% < \text{RH} < 80\text{ } \%$; $40\text{ } \% < MC_{ini} < 80\text{ } \%$,
- 210 • Test data set: $T = 50\text{ }^{\circ}\text{C}$ and $20\text{ } \% < \text{RH} < 90\text{ } \%$; $20\text{ } \% < MC_{ini} < 100\text{ } \%$.

The conditions of the test data set were intentionally slightly different. 400 simulation runs were generated to populate each data set. Fig. 4 proposes a graphical representation of these data sets.

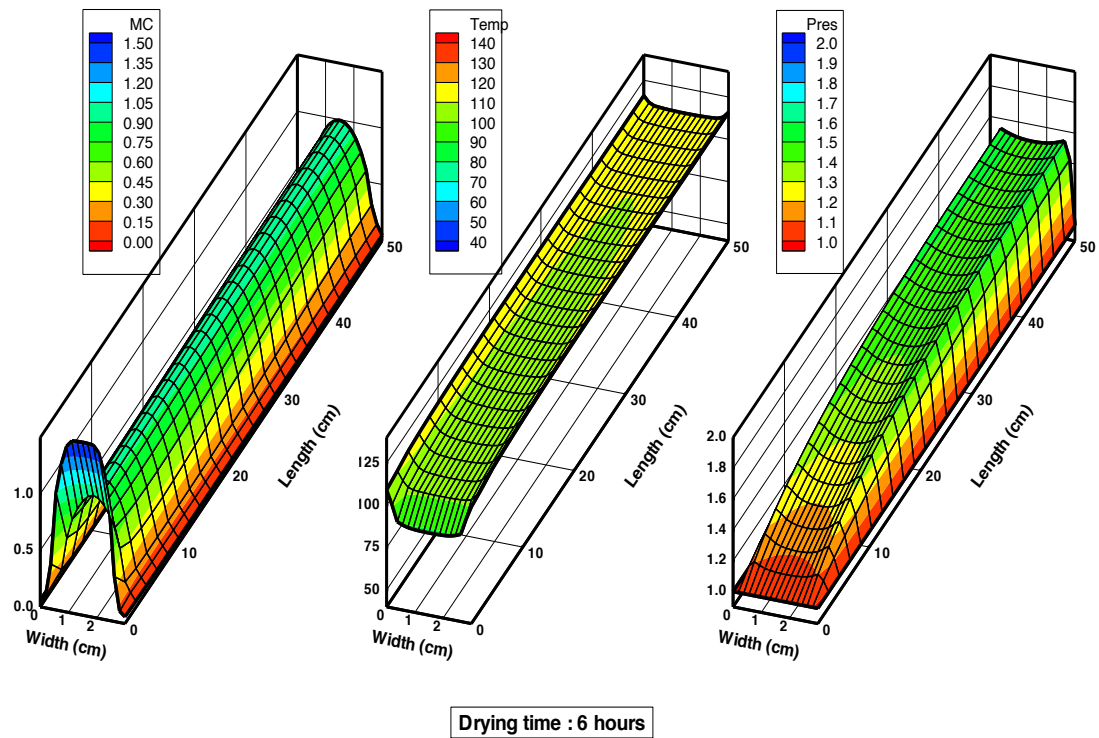


Figure 3: Potential of mechanistic modeling. 2D simulation of high temperature drying of wood. The internal overpressure developed when the product temperature is above the boiling point triggers a significant liquid movement along the very permeable longitudinal direction, which supplied the end-piece with liquid [8]

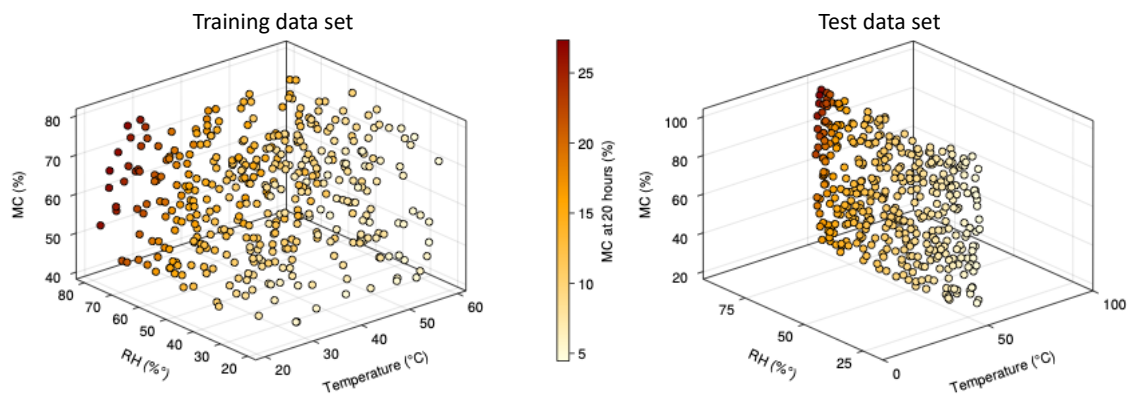


Figure 4: The two data sets populated by the mechanistic simulation tool *TransPore*. Left) Training data set 1 (400 runs with random variations of temperature, RH and initial MC) and Right) Test data set (400 runs at 50°C with random variations of RH and initial MC)

215 For the sake of example, the simulation results are plotted in figure 5 for two contrasted tests. The average MC, surface MC and core MC are plotted as blue lines. For these tests at low temperature convective drying and with a diffusion mainly as bound water migration in the hygroscopic domain, the temperature remains quasi uniform in the thickness. for this reason, only the surface temperature is represented as a red line. The first example

220 is a test with severe conditions : $T = 80 \text{ }^\circ\text{C}$; $\text{RH} = 20 \%$ and $\text{MC}_{ini} = 60 \%$. With these conditions, the drying kinetics is very fast. The temporal evolution of the surface temperature indicates that there is no constant drying period. This fact is confirmed by the very fast drop of MC at the surface. In this case, the average MC is already very close to equilibrium at 10 hours of drying.

225 The second example is a test with mild conditions : $T = 50 \text{ }^\circ\text{C}$; $\text{RH} = 80 \%$ and $\text{MC}_{ini} = 90 \%$. The kinetics is consistently much slower than the previous test, with 30 hours needed to reach the equilibrium, against 10 hours in the previous one. The temperature evolution depicts a long plateau at the wet bulb temperature: it lasts more than 15 hours. Then, the product enters the hygroscopic domain and the temperature gradually increases to

230 eventually reach the dry bulb temperature when the product is at its equilibrium MC. In this case, the MC values at 10 and 20 hours are still far from equilibrium.

The evolution of the MC profiles over time are depicted in figure 6. As expected, sharp profiles are obtained with the severe drying conditions and more flat profiles for mild conditions. It is interesting to notice that the solution of coupled heat and mass transfer is

235 able to capture the condensation at the beginning of drying, the time for the surface temperature to become higher than the dew point. This is explained by the initial temperature of $25 \text{ }^\circ\text{C}$ set in the simulations. Consistently, the quantity of condensate water is larger in the case of a very humid air.

Although these simulations are very rich in information, the following examples of Neural Network will be focused on a very reduced information: the MC values computed

240 at 10 and 20 hours of drying by *TransPore*. The results section will show that, although very coarse, this information is capable of producing excellent results thanks to the large number of calculations.

5. Classical Neural Network

245 The neural network was implemented in *Julia* using the package *Flux* [27, 28]. A dense network of 4 layers (2 hidden layers, see Fig. 7), with sigmoid curve as activation function, was used. The hidden layers contain respectively 5 and 3 neurons. Trials proved that 70 000 epochs is a good compromise to obtain a low value of the objective function, with no further decrease beyond this value. Temperature, Relative Humidity and Initial MC

250 are the three input values and MC at 10 and 20 hours are the output values.

Figure 8 plots the MC values predicted by NN at 10 hours (green markers) and 20 hours (red markers) as function of the MC simulated by the mechanistic model. For the training data set (left), the prediction is excellent, but the test data set (right) depicts serious

255 discrepancies for the largest values of MC. This is explained by the situations of the test data set that were not covered by the training data set: RH up to 90 % instead of 80 % and initial MC up to 100 % instead of 80 %. Which such high values of both RH and MC_{ini} , MC values as high as 60 % were simulated at 10 hours of drying. However, NN limits these values at around 40 % (the maximum values of the training data set). This explains the

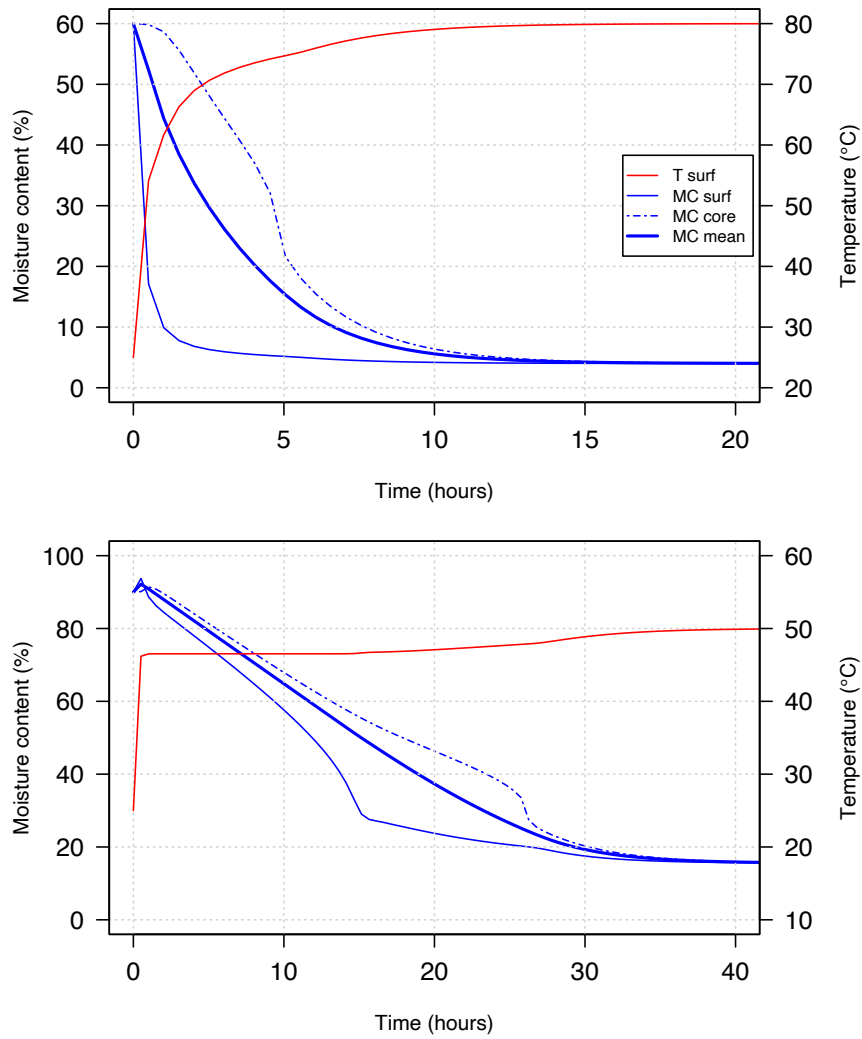


Figure 5: Two contrasted examples of mechanistic simulation. Top) Severe conditions ($T = 80\text{ }^{\circ}\text{C}$; $\text{RH} = 20\%$ and $\text{MC}_{ini} = 60\%$) Bottom) Mild conditions ($T = 50\text{ }^{\circ}\text{C}$; $\text{RH} = 80\%$ and $\text{MC}_{ini} = 90\%$).

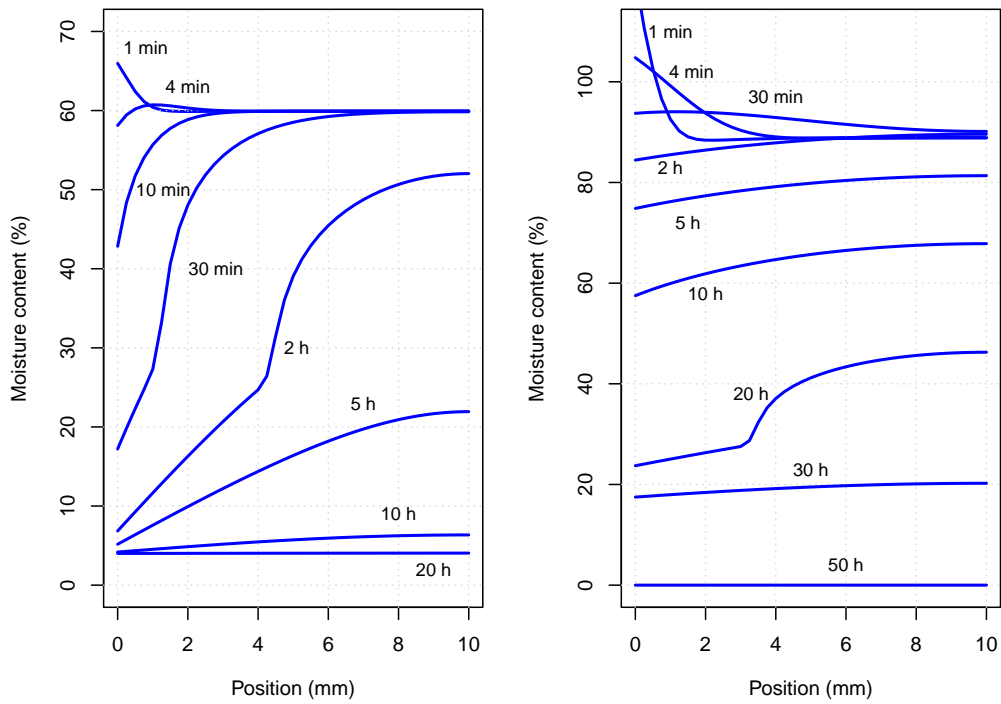


Figure 6: The MC profile computed for the two contrasted examples of mechanistic simulation. Left) Severe conditions (T = 80 °C; RH = 20 % and MC_{ini} = 60 %) Right) Mild conditions (T = 50 °C; RH = 80 % and MC_{ini} = 90 %).

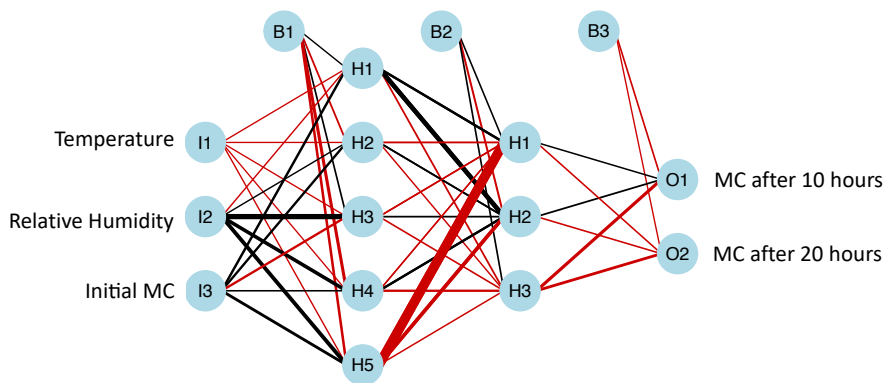


Figure 7: The neural network (2 hidden layers of logistic functions) after training; black lines indicated positive values and red line negative values. Line thickness indicates absolute value.

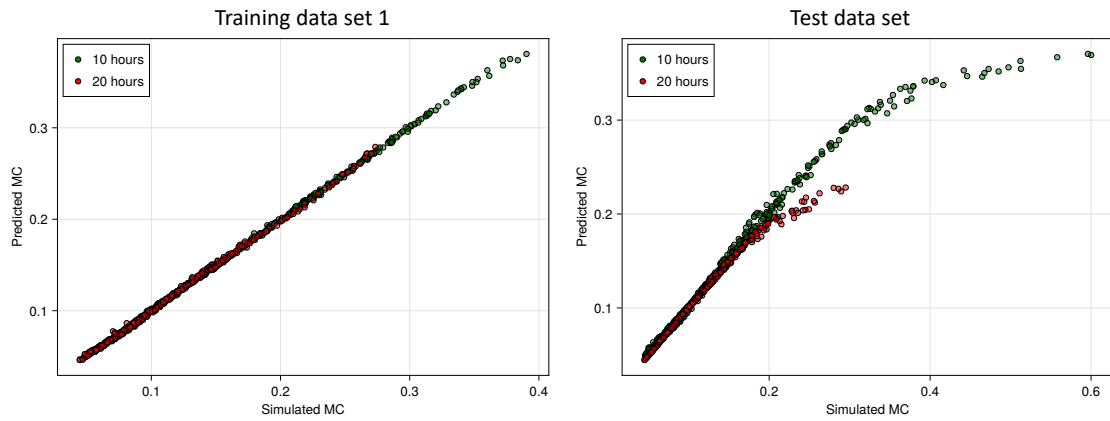


Figure 8: Performances of the Neural Network using data set 1. Left) Training set and Right) Test data set.

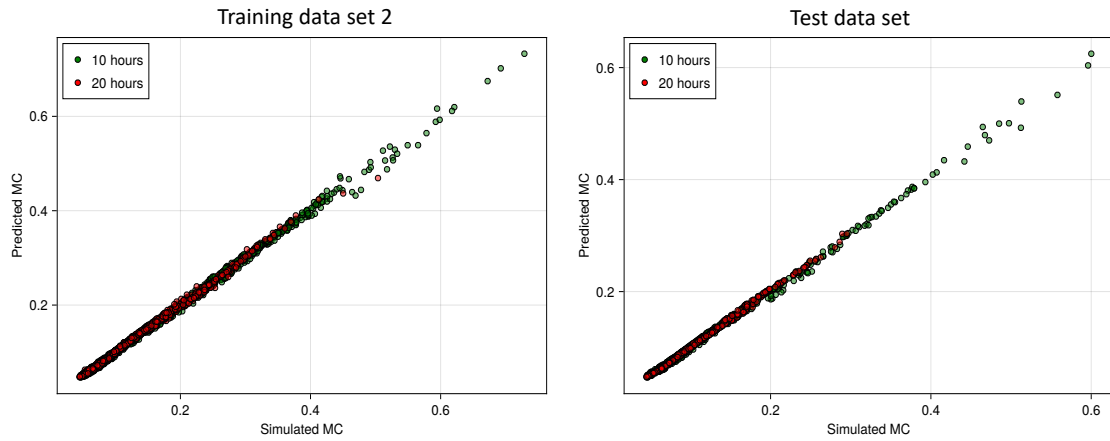


Figure 9: Performances of the Neural Network using data set 2. Left) Training data set and Right) Test data set.

260 poor performance of this neural network for the test step (Table 1), with a maximum error
of 20.2 %. The advantage of populating the data set by a mechanistic model, rather than
by experiments, is that it can be generated again with suitable conditions. We generated
a second training data set, which includes 1000 runs with wider ranges of parameters. In
particular, this data set has a wider range and temperature and initial MC and includes
265 tests with a very high relative humidity (up to 90 %), which results in very slow drying
rate:

- Training data set 2: $20\text{ }^{\circ}\text{C} < T < 90\text{ }^{\circ}\text{C}$; $20\text{ } \% < \text{RH} < 90\text{ } \%$; $20\text{ } \% < \text{MC}_{ini} < 100\text{ } \%$,

Using this second set, the NN excels in both training and test (Fig. 9 and Table 1). The
maximum error now equals 2.93 % with a RMSE of 0.42 %.

270 Once trained, the NN is very fast and can be used to address various problems. As a case
study, it was used here to extract the combination of parameters able to produce a target
MC of 20 % (red dots of Fig. 10) or 10 % (blue dots) at 10 hours of drying. The desired
target MC values were selected from 50,000 random sets of input parameters computed

Table 1: Performances of the neural networks.

		RMSE(% MC)	Max error (% MC)
Training data set 1	Training	0.17	0.98
	Test	1.76	20.2
Training data set 2	Training	0.46	3.71
	Test	0.42	2.93

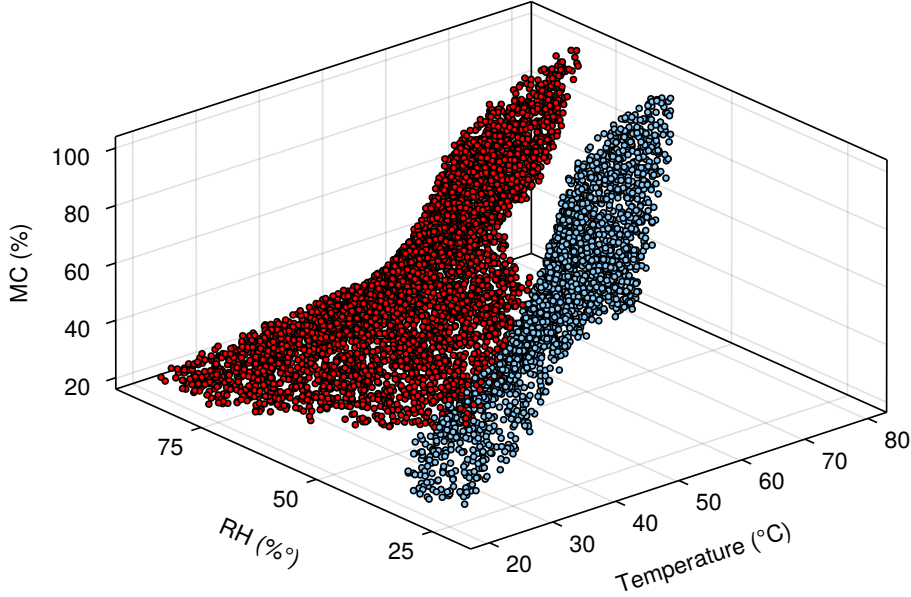


Figure 10: Using the Neural Network to determine the combination of parameters allowing the MC at 10 hours to be equal to 20 % (red dots) or 10 % (blue dots). Selection performed from 50 000 random values.

in less than one second by the trained NN. As the date set was produced by a mechanistic
 275 model, the trends reported in this figure are physically consistent. For example, as the
 initial MC approaches 20 %, the drying conditions required to obtain 20 %, which means
 almost no drying, after 10 hours of drying should be extremely mild: low temperature
 with nearly saturated conditions.

6. Physics Informed Neural Network

280 Classical neural networks perform efficiently but they don't include any physical informa-
 tion. This section gives an example of Physics Informed Neural Network (PINN). Start-
 ing from the simple case presented in the previous section, it is clear that each individual
 drying kinetics, whatever the drying conditions, follows a well-known trend explained
 by the drying rate : constant drying rate when it exists, followed by a decreasing drying
 285 rate [29, 30].

In this section, a simple physical model was embedded in the Neural Network. Instead
 of predicting the MC values at 10 and 20 hours, the NN was used to predict the degrees
 of freedom of the concept of dimensionless drying rate curve, as defined decades ago by

```

A ←  $\dot{X}_0 / (X_{cr} - X_{eq})$ 
B ←  $\min(X_{cr}, X_{ini}) - X_{eq}$ 
if  $X_{ini} \geq X_{cr}$  then
   $t_{cr} \leftarrow (X_{ini} - X_{cr}) / \dot{X}_0$ 
  if  $t \leq t_{cr}$  then
     $X(t) = X_{ini} - \dot{X}_0 \times t$ 
  else
     $X(t) = B \exp(-A(t - t_{cr})) + X_{eq}$ 
  end if
else
   $X(t) = B \exp(-A(t)) + X_{eq}$ 
end if

```

Figure 11: Algorithm coded in *Julia* to compute the drying kinetics defined by the van Meel model

van Meel [29]. The concept of dimensionless drying curve expressed the drying rate as a function of the dimensionless moisture content X^* :

$$X^* = \frac{X - X_{eq}}{X_{cr} - X_{eq}} \quad (1)$$

Where X_{eq} is the equilibrium moisture content defined by the sorption isotherm of the product and X_{cr} the critical moisture content, the value below which the drying rate is not constant any more and starts to decrease. The drying rate is then expressed as a function of X^* :

$$\dot{X} = \dot{X}_0 f(X^*) \quad (2)$$

Here, the simplest shape of f ($f(X^*) = X^*$) was used. With this function, the drying kinetics can be then obtained by analytical integration of (2). This involves several cases summarized in the algorithm of figure 11. \dot{X}_0 is supposed to be a positive number in these expressions.

The model is defined by three parameters: X_{cr} , X_{eq} and \dot{X}_0 . This simple physical model was coded as a *Julia* function and embedded in the neural network. A network similar to that of figure 7 was used, except that the last layer has 3 outputs instead of 2. These outputs correspond to the 3 model parameters. The physical model was embedded in the loss function: for each point of the data set (learning or testing), the kinetics is solved to obtain the moisture contents at the given times and compute the RMSE error to the computational values. For the sake of comparison, the learning phase parameters were the same parameters in the previous section (70,000 epochs and a convergence rate of the Adam's solver of 10^{-2}).

With the same piece of information, the quantified results show that the PINN case always performs better than the simple NN (Table 2). This is particularly true for the test phase using data set 2, for which the RMSE drops from 0.42 % to 0.35 %, and the maximum error from 2.93 % to 1.55 %. Figure 12 depicts the quality of the predicted MC versus the computed MC for the two data sets already used in section 5. Data set 1 clearly suffers from its too narrow range of conditions for a good prediction of the MC values, as depicted by the slight scattering of points still visible after the learning phase (right top sub-graph).

However, even though, the PINN exhibits much less points badly predicted in the test data set than in the case of classical NN (to be compared to figure 8), which denotes a better

Table 2: Performances of the physics informed neural networks.

		RMSE(% MC)	Max error (% MC)
Training data set 1	Training	0.45	2.07
	Test	1.18	17.2
Training data set 2	Training	0.43	2.11
	Test	0.35	1.55

potential of prediction. For data set 2, the fit is excellent for the two steps (learning and test).

Figure 13 depicts the three values of the van Meel model obtained by the PINN after the learning phase on data set 2 (1000 simulation runs). The equilibrium moisture content X_{eq} is strongly related to the relative humidity of the air flow and it is easy to recognize the sorption isotherm embedded in the computational code. It is interesting to notice that the point above this sorption curve correspond to high relative humidity values and low temperature (clear marker colors). These points are also tied to conditions with a high initial moisture content (not represented in this curve). All these conditions represent configurations that remained far from the equilibrium at 20 hours, the longest drying time provided as input data.

The critical moisture content X_{cr} is always near 100 % with a low relative humidity, hence a fast drying rate. Consistently, this value drops to about 50 % at high relative humidity, which represent very mild conditions. The constant drying rate value \dot{X}_0 depends on both relative humidity and temperature: high values at low RH and high temperature, lower values at low RH but lower temperature and, as expected, and very low values at high relative humidity (low gap between the dry and wet bulb temperatures), whatever the temperature level.

The Physics Informed Neural Network is therefore much closer to the physical behavior. It is remarkable to observe such relevant trends when supplying only the moisture content at two different times. One must keep in mind that these good trends are explained by two main facts:

- the data set was generated by a comprehensive mechanistic-based computational model,
- the information for each test is scarce (2 drying times), but the data set contains 1000 runs with random drying conditions values.

Another major advantage of the Physics Informed Neural Network over the classical neural network is to allow the full kinetics to be predicted, instead of only the moisture content at the two drying times supplied in the data set. For the sake of example, figure 14 plots the drying kinetics of the eight first conditions of the data set. These kinetics were computed using the van Meel model with the model parameters predicted by the trained PINN (Table 3). The random conditions explains the diversity of these curves. For example, the pink curve is a test with a high initial moisture content and moderate conditions. One the contrary, the dark yellow curve is a test with very low initial content and very small drying rate. For each predicted curve, the moisture contents computed by the comprehensive model are plotted as markers of the same color. The match is always good,

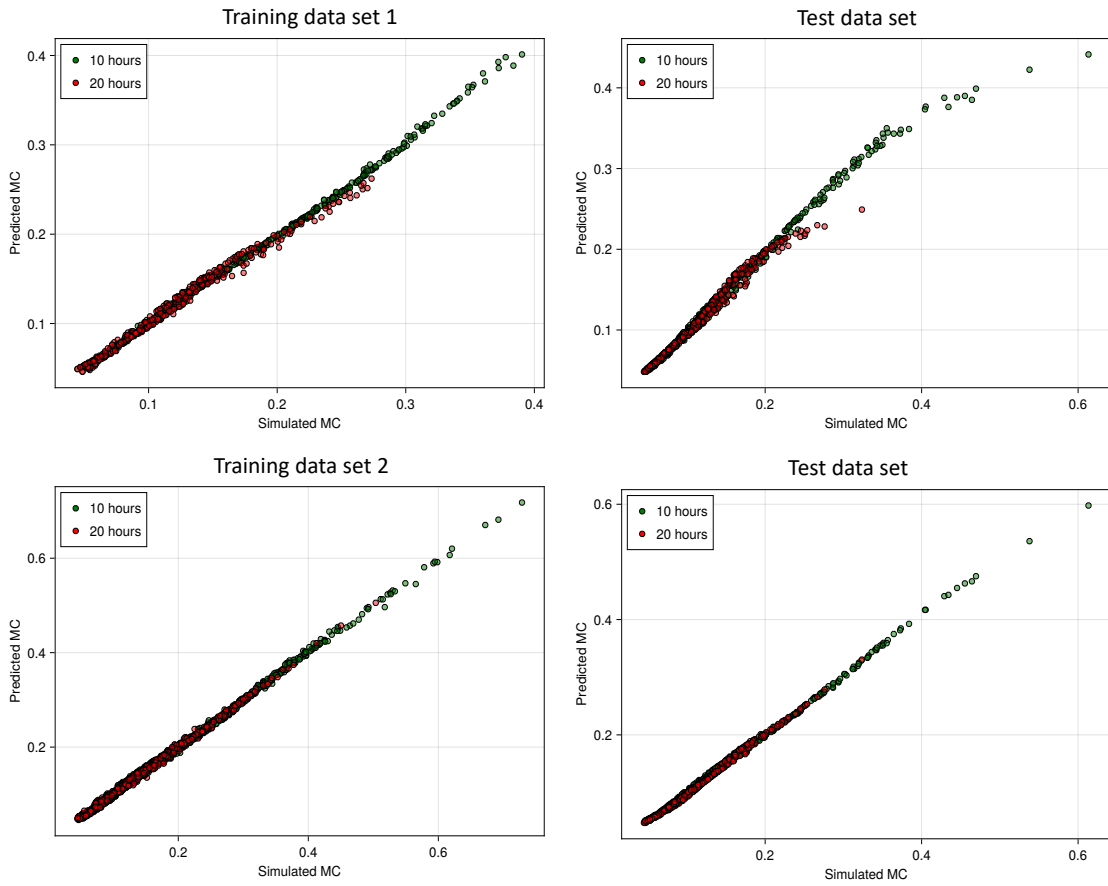


Figure 12: Performances of the Physics Informed Neural Network using data set 1 (Top) and data set 2 (Bottom). Left) Training data set and Right) Test data set.

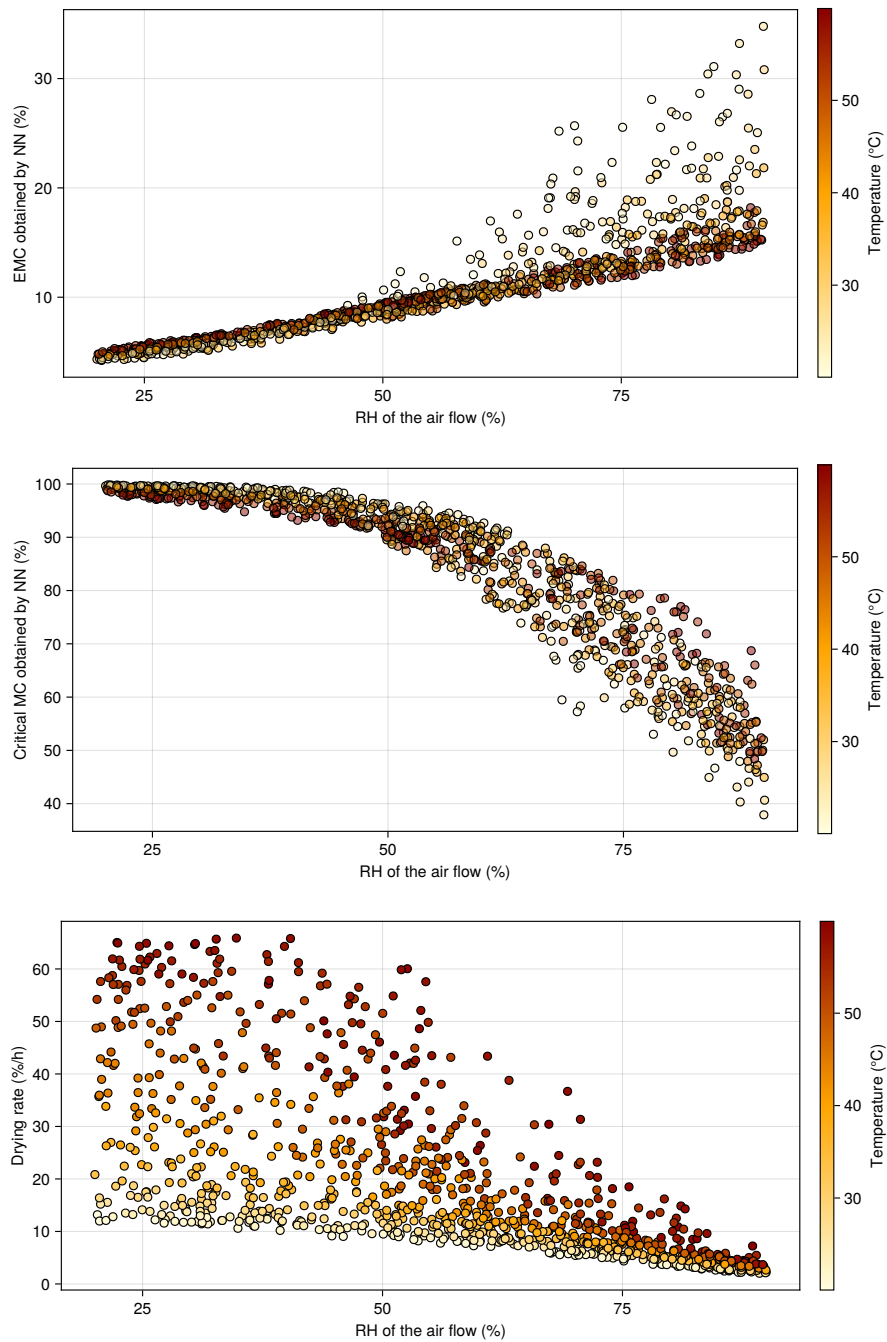


Figure 13: Parameters of the van Meel model obtained after training of the Physics Informed Neural Network. From top to bottom: equilibrium moisture content, critical moisture content and constant drying rate.

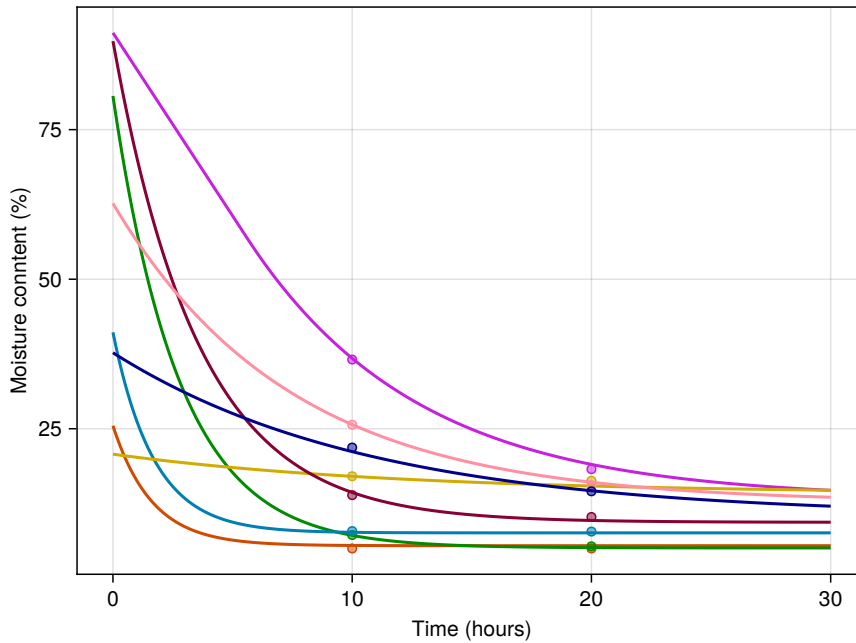


Figure 14: Drying curves computed using the trained PINN parameters for the first eight random drying conditions.

if not excellent, which confirms the global figures of table 2. Table 3 includes the values of moisture content at 10 and 20 hours: the gap is most generally less than 0.5 % with a maximum value of 0.9 % at 20 hours for test 5 (the dark yellow curve of Fig. 14). The results obtained for the test data set are not shown here, but are very similar in terms of quality. They are just less interesting due to the choice of a single temperature value.

As final remark, one must keep in mind that the PINN can give the van Meel parameters for any set of drying conditions covered by data set 2. Then, the drying kinetics can be computed very simply. This means that this PINN became a very fast predictive tool. This opens the door for a neural network approach able to predict much more than the drying kinetics. Keeping in mind how rich are the data generated by the mechanistic model, information such as energy consumption, thermal degradation, mechanical quality... could be added to the data set before subsequent exploitation by the PINN. For example, thermal degradation can be obtained by the time-integration of the product temperature and mechanical quality by the time-integration of the different of shrinkage between the surface and the core of the product. This requires only a simple quantification of the product temperature or the MC values at the surface and core (Fig. 5) to be supplied to the data set.

7. Conclusion

This paper intends to show how mechanistic modeling and machine learning can work jointly in the field of drying. It first summarized the physics embedded in the mechanistic formulation of drying. Then, ML, a mushrooming domain, is not proposed to be used instead of mechanistic model, but as a complement to it. The different possibilities

Table 3: The eight first random conditions of the learning data set and the corresponding data (values of the van Meel model predicted by the PINN after learning and the values of moisture contents t 10 and 20 hours, both for the comprehensive computational model and for the van Meel model using the PINN values.).

	Drying conditions			van Meel values			MC at 10 hours		MC at 20 hours	
	T	RH	X_{ini}	X_{cr}	X_{eq}	\dot{X}_0	Simul.	Pred.	Simul.	Pred.
	°C	%	%	%	%	% /h	%	%	%	%
1	58.9	83.3	91.2	56.7	13.3	6.1	36.6	36.8	18.2	19.0
2	54.8	23.4	25.5	98.1	5.5	56.9	5.0	5.5	5.0	5.5
3	43.0	24.2	80.6	99.1	5.1	33.7	7.2	7.2	5.3	5.1
4	51.6	40.4	41.1	95.4	7.6	51.4	7.9	7.7	7.8	7.6
5	30.3	78.3	20.8	74.2	14.2	5.0	17.1	17.0	16.3	15.4
6	53.7	55.7	89.8	86.6	9.4	21.5	13.9	14.4	10.2	9.7
7	45.0	76.2	62.7	68.8	12.7	7.6	25.7	25.7	15.6	16.1
8	20.6	56.0	37.7	94.1	10.4	7.8	21.9	21.2	14.5	14.6

of coupling the mechanistic and ML approaches to obtain an "augmented" mechanistic model are outlined. Different pathways are proposed and commented to merge the best of the two worlds:

- A full coupled method (hybrid model)
- 380 • A fully decoupled method
- A cascade coupling

As an illustrative example, a fully decoupled approach is proposed and discussed at the second part of the paper. To that purpose, a large data set was generated by a comprehensive mechanistic model and then explored by Neural Networks. The originality of this section is to compare a classical Neural Network (NN) and a Physics Informed Neural Network (PINN). The second one depicts a much higher accuracy and a real prediction potential. Many drying aspects can be tackled on this basis by in-depth exploitation of the exhaustive data set generated by the mechanistic model (energy consumption, thermal degradation, mechanical quality, etc.). Another interesting perspective would be to use neural ordinary differential equations (neural ODEs), particularly well suited to predicting time-dependent configurations.

Acknowledgment

Communauté urbaine du Grand Reims, Département de la Marne, Région Grand Est and European Union (FEDER Champagne–Ardenne 2014–2020, FEDER Grand Est 2021–2027) are acknowledged for their financial support to the Chair of Biotechnology of Centrale-Supélec and the Centre Européen de Biotechnologie et de Bioéconomie (CEBB).

References

- [1] M. Aghbashlo, S. Hosseinpour, A. S. Mujumdar, Application of artificial neural networks (anns) in drying technology: a comprehensive review, *Drying technology* 33 (12) (2015) 1397–1462.
- [2] Q. Sun, M. Zhang, A. S. Mujumdar, Recent developments of artificial intelligence in drying of fresh food: A review, *Critical reviews in food science and nutrition* 59 (14) (2019) 2258–2275.
- [3] E. Taghinezhad, M. Kaveh, A. Jahanbakhshi, I. Golpour, Use of artificial intelligence for the estimation of effective moisture diffusivity, specific energy consumption, color and shrinkage in quince drying, *Journal of Food Process Engineering* 43 (4) (2020) e13358.
- [4] A. Martynenko, N. Misra, Machine learning in drying, *Drying Technology* 38 (5-6) (2020) 596–609.
- [5] M. I. H. Khan, S. S. Sablani, M. Joardder, M. Karim, Application of machine learning-based approach in food drying: Opportunities and challenges, *Drying Technology* 40 (6) (2022) 1051–1067.
- [6] C. Chen, Z. Pan, An overview of progress, challenges, needs and trends in mathematical modeling approaches in food drying, *Drying Technology* 41 (16) (2023) 2586–2605.
- [7] R. Yang, J. Chen, Mechanistic and machine learning modeling of microwave heating process in domestic ovens: A review, *Foods* 10 (9) (2021) 2029.
- [8] P. Perré, R. Rémond, G. Almeida, P. Augusto, I. Turner, State-of-the-art in the mechanistic modeling of the drying of solids: A review of 40 years of progress and perspectives, *Drying Technology* 41 (6) (2023) 817–842.
- [9] O. Krischer, K. Kröll, *Die wissenschaftlichen Grundlagen der Trocknungstechnik*, Springer, Berlin (1956).
- [10] J. Philip, D. De Vries, Moisture movement in porous materials under temperature gradients, *Eos, Transactions American Geophysical Union* 38 (2) (1957) 222–232.
- [11] A. V. Luikov, Heat and mass transfer in capillary-porous bodies, in: *Advances in heat transfer*, Vol. 1, Elsevier, 1964, pp. 123–184.
- [12] S. Whitaker, Simultaneous heat, mass and momentum transfer in porous media: a theory of drying, in: J. P. Hartnett, T. F. Irvine (Eds.), *Advances in Heat Transfer*, Vol. 13, Elsevier, 1977, pp. 119–203.
- [13] J. Slattery, Flow of viscoelastic fluids through porous media, *AIChE J.* 13 (1967) 1066–1071.
- [14] W. Gray, A derivation of the equations for multiphase transport, *Chemical Engineering Science* 30 (1975) 229–233.
- [15] S. Whitaker, Coupled transport in multiphase systems: a theory of drying, in: Y. I. C. J. P. Hartnett, T. F. Irvine, G. A. Greene (Eds.), *Advances in Heat Transfer*, Vol. 31, Elsevier, 1998, pp. 1–104.
- [16] P. Perré, A. Degiovanni, Simulation par volumes finis des transferts couplés en milieux poreux anisotropes: séchage du bois à basse et à haute température, *International Journal of Heat and Mass Transfer* 33 (11) (1990) 2463–2478.
- [17] P. Perré, B. May, A numerical drying model that accounts for the coupling between transfers and solid mechanics. case of highly deformable products, *Drying Technology* 19 (8) (2001) 1629–1643.

- [18] S. J. Kowalski, K. Rajewska, Drying-induced stresses in elastic and viscoelastic saturated materials, *Chemical Engineering Science* 57 (18) (2002) 3883–3892.
- 445 [19] P. Perré, J. Passard, A physical and mechanical model able to predict the stress field in wood over a wide range of drying conditions, *Drying Technology* 22 (1-2) (2004) 27–44.
- [20] D. Mihoubi, F. Zagrouba, J. Vaxelaire, A. Bellagi, M. Roques, Transfer phenomena during the drying of a shrinkable product: modeling and simulation, *Drying Technology* 22 (1-2) (2004) 91–109.
- 450 [21] M. Vinjamur, R. A. Cairncross, Non-fickian nonisothermal model for drying of polymer coatings, *AIChE Journal* 48 (11) (2002) 2444–2458.
- [22] P. Perré, Coupled heat and mass transfer in biosourced porous media without local equilibrium: a macroscopic formulation tailored to computational simulation, *Int. J. Heat Mass Transfer* 140 (2019) 717–730.
- 455 [23] S. Ouertani, A. Stéphan, P. Perré, C. L’hostis, R. Rémond, Evaluating moisture transfer properties of wood by inverse analysis of moisture content profiles determined during drying by x-ray attenuation, *Drying Technology* 42 (1) (2024) 168–181.
- [24] A. Stéphan, P. Perré, C. L’Hostis, R. Rémond, Mechanistic-based probabilistic optimization of industrial wood drying considering energy consumption, process duration, quality and cost, *Drying Technology* 42 (7) (2024) 1178–1187.
- 460 [25] P. Perré, Artificial intelligence in the field of drying: Revolution or evolution?, *Drying Technology* 42 (4) (2024) 589–591.
- [26] P. Perré, I. W. Turner, A 3D version of Transpore: a comprehensive heat and mass transfer computational model for simulating the drying of porous media, *International Journal for Heat and Mass Transfer* 42 (1999) 4501–4521.
- 465 [27] M. Innes, E. Saba, K. Fischer, D. Gandhi, M. C. Rudilosso, N. M. Joy, T. Karmali, A. Pal, V. Shah, [Fashionable modelling with flux](#), *CoRR* abs/1811.01457 (2018). [arXiv:1811.01457](#).
- 470 URL <https://arxiv.org/abs/1811.01457>
- [28] M. Innes, Flux: Elegant machine learning with julia, *Journal of Open Source Software* (2018). [doi:10.21105/joss.00602](https://doi.org/10.21105/joss.00602).
- [29] D. van Meel, Adiabatic convection batch drying with recirculation of air, *Chemical Engineering Science* 9 (1) (1958) 36–44.
- 475 [30] R. B. Keey, *Introduction to industrial drying operations*, Pergamon press Oxford, 1978.
- [31] P. Perré, R. Remond, I. W. Turner, Comprehensive drying models based on volume averaging: Background, application and perspective, in: E. Tsotsas, A. S. Mujumdar (Eds.), *Modern Drying Technology*, Vol. 1, Wiley-VCH, 2007.

480 **Appendix A. Macroscopic formulation of coupled heat and mass transfer**

This annex describes the comprehensive set of macroscopic equations describing coupled transfer occurring during drying. This set was proposed by [12, 15] and adapted to the case of hygroscopic products with a balance of dry air required to consider the effect of internal pressure [26, 31].

485

Moisture conservation

All water phases are considered in the accumulation term (liquid, vapor and bound).

The fluxes includes the liquid velocity driven by the liquid pressure (convection and capillary), the convection of gas, the diffusion of bound water, and the binary diffusion of vapor in air.

490

$$\frac{\partial (\varepsilon_w \rho_w + \varepsilon_g \bar{\rho}_v^g + \bar{\rho}_b)}{\partial t} + \nabla \cdot (\rho_w \bar{\mathbf{v}}_w + \bar{\rho}_v^g \bar{\mathbf{v}}_g + \bar{\rho}_b \bar{\mathbf{v}}_b) = \nabla \cdot (\rho_g \mathbf{D}_{eff} \cdot \nabla \omega_v) \quad (\text{A.1})$$

Energy conservation

The enthalpy conservation includes the enthalpy of all phases in the accumulation terms, the pressure work, and all enthalpy fluxes due to convection of diffusion migrations.

$$\begin{aligned} \frac{\partial}{\partial t} (\varepsilon_w \rho_w h_w + \varepsilon_g (\bar{\rho}_v^g h_v + \bar{\rho}_a^g h_a) + \bar{\rho}_b \bar{h}_b + \varepsilon_s \rho_s h_s - \varepsilon_g p_g) \\ + \nabla \cdot (\rho_w h_w \bar{\mathbf{v}}_w + (\bar{\rho}_v^g h_v + \bar{\rho}_a^g h_a) \bar{\mathbf{v}}_g + h_b \bar{\rho}_b \bar{\mathbf{v}}_b) + \bar{\mathbf{v}}_w \cdot \nabla p_w + \bar{\mathbf{v}}_g \cdot \nabla p_g \\ = \nabla \cdot (\rho_g \mathbf{D}_{eff} (h_v \nabla \omega_v + h_a \nabla \omega_a) + \boldsymbol{\lambda}_{eff} \nabla T) + \varphi \end{aligned} \quad (\text{A.2})$$

In this equation, the term φ accounts for possible volumetric heat sources, for example in the case of high frequency, or microwave heating. The transport of enthalpy due to bound water migration must be treated with care. As the differential heat of sorption depends on the bound water content, the averaged value \bar{h}_b should be used in the time evolution (accumulation term), whereas the value at ρ_b (h_b) should be used in the migration term, as it is assumed that the less bound water molecules are those likely to migrate.

500

Air conservation

This equation simply considers the accumulation of dry air in the gaseous phase and its migration by convection and diffusion.

$$\frac{\partial (\varepsilon_g \bar{\rho}_a^g)}{\partial t} + \nabla \cdot (\bar{\rho}_a^g \bar{\mathbf{v}}_g) = \nabla \cdot (\rho_g \mathbf{D}_{eff} \nabla \omega_a) \quad (\text{A.3})$$

In these equations, the barycentric mass velocities are from the generalized Darcy's law:

$$\bar{\mathbf{v}}_g = -\frac{\mathbf{K} \mathbf{k}_g}{\mu_g} (\nabla p_g - \rho_g \nabla \psi_g) \quad (\text{A.4})$$

$$\bar{\mathbf{v}}_w = -\frac{\mathbf{K} \mathbf{k}_w}{\mu_w} (\nabla p_w - \rho_w \nabla \psi_w) \quad \text{with} \quad P_w = P_g - P_c(X, T) \quad (\text{A.5})$$

When the bound water flux is expressed using the bound water density as the driving force, the bound water flux takes the following form:

505

$$\begin{aligned} \bar{\rho}_b \bar{\mathbf{v}}_b = -\mathbf{D}_b \nabla \bar{\rho}_b = -\rho_0 \mathbf{D}_b \nabla X_b \\ \text{where} \quad X_b = \min(X, X_{fsp}) \end{aligned} \quad (\text{A.6})$$

Boundary conditions

$$\begin{aligned} \mathbf{J}_v|_{x=0^+} \cdot \mathbf{n} &= h_m c M_v \ln \left(\frac{1 - x_\infty}{1 - x_v|_{x=0}} \right) \\ \mathbf{J}_h|_{x=0^+} \cdot \mathbf{n} &= h_h (T|_{x=0} - T_\infty) \\ P_g|_{x=0^+} &= P_{atm} \end{aligned} \quad (\text{A.7})$$

The first boundary equation assumes that only water vapor is exchanged between the product and the surrounding air.

This macroscopic formulation assumes that the porous medium is locally at equilibrium:

510 **A1** the temperature is the same for all phases $T_s = T_w = T_g$

A2 the partial pressure of water vapor inside the gaseous phase is in equilibrium with moisture content X via $p_v = p_{vs}(T) \times a_w(T, X)$, where function a_w is the sorption isotherm, also called water activity, namely in food science.

Further assumptions allow this set of equations to have a more convenient form:

515 **A3** the variation in partial densities inside the REV are negligible; therefore, the intrinsic average is equal to the local value $\bar{\rho}_v^g = \rho_v$ and $\bar{\rho}_a^g = \rho_a$,

A4 the solid density is assumed to be constant $\rho_s = \text{constant}$,

A5 the moisture content X is used to consider the total amount of water present in the porous medium $\rho_0 X = \varepsilon_w \rho_w + \varepsilon_g \bar{\rho}_v^g + \bar{\rho}_b$ where $\rho_0 = \varepsilon_s \rho_s$,

520 **A6** the effective diffusivity is expressed as a function of the binary diffusivity of vapor in air: $\mathbf{D}_{eff} = \mathbf{f} \mathbf{D}_v$, where \mathbf{f} is a dimensionless diffusivity tensor (indeed, along one given direction, $f = 1/\mu$ where μ is the vapor resistance ratio used for building materials),

A7 The effect of pressure variation has a negligible contribution in enthalpy conservation

525 With these additional assumptions, the set of equations takes the following compact form:

<i>Moisture conservation</i>	
$\rho_0 \frac{\partial X}{\partial t} + \nabla \cdot (\rho_w \bar{\mathbf{v}}_w + \rho_v \bar{\mathbf{v}}_g) = \nabla \cdot (\rho_g \mathbf{f} \mathbf{D}_v \nabla \omega_v + \rho_0 \mathbf{D}_b \nabla X_b)$	(A.8)
<i>Energy conservation</i>	
$\frac{\partial}{\partial t} (\varepsilon_w \rho_w h_w + \varepsilon_g (\rho_v h_v + \rho_a h_a) + \bar{\rho}_b h_b + \varepsilon_s \rho_s h_s) + \nabla \cdot (\rho_w h_w \bar{\mathbf{v}}_w + (\rho_v h_v + \rho_a h_a) \bar{\mathbf{v}}_g)$	
$= \nabla \cdot (\lambda_{eff} \nabla T + (h_v - h_a) \rho_g \mathbf{f} \mathbf{D}_v \nabla \omega_v + h_b \rho_0 \mathbf{D}_b \nabla X_b) + \varphi$	(A.9)
<i>Air conservation</i>	
$\frac{\partial (\varepsilon_g \rho_a)}{\partial t} + \nabla \cdot (\rho_a \bar{\mathbf{v}}_g) = \nabla \cdot (\rho_g \mathbf{f} \mathbf{D}_v \nabla \omega_a)$	(A.10)

Nomenclature

Greek Symbols

	ϵ	Volume fraction	-
530	μ	Viscosity	Pa · s
	Φ	Porosity	-
	ρ	Density	kg/m ³

	φ	Volumetric heating term	W/m ³
Roman Symbols			
535	n	Normal unit vector	-
	c_p	Heat capacity	J/(K · m ³)
	D	Diffusion coefficient	m ² /s
	f	Dimensionless mass diffusivity	-
	h	Specific enthalpy	J/(K · kg)
540	h_h	Heat transfer coefficient	W/(K · m ²)
	h_m	Mass transfer coefficient	m/s
	K	Permeability	m ²
	L_v	Latent heat of vaporization	J/kg
	m	Mass	kg
545	MC	Moisture content, dry basis	-
	P	Pressure	Pa
	q_h	Flux density of heat	W/m ²
	q_v	Flux density of vapor	kg/(s · m ²)
	R	Gas constant	8.314 J/(K · mol)
550	RH	relative humidity	-
	T	Temperature	°C
	t	Time	s
	V	Volume	m ³
	v	Velocity	m/s
555	X	Moisture content, dry basis	-
Subscripts			
	a	Air	
	atm	Atmospheric value	
	eff	Effective value at the macroscopic scale	
560	eq	Equilibrium	
	g	Gaseous	
	h	Heat	
	m	Mass	
	s	Solid	
565	v	Water vapor	
	w	Wet bulb or liquid water	
Superscripts			
	$\overline{\psi}$	average of variable ψ over the REV	

Other Symbols

570	∇	Gradient
	$\nabla \cdot$	Divergence