Computers and Chemical Engineering 140 (2020) 106971

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/compchemeng

Global uncertainty-sensitivity analysis on mechanistic kinetic models: From model assessment to theory-driven design of nanoparticles

M. Reza Andalibi[a,b,1,*], Paul Bowen[b], Agnese Carino[a], Andrea Testino[a]

[a] Paul Scherrer Institute (PSI), ENE-LBK-CPM, 5232 Villigen-PSI, Switzerland
[b] École polytechnique fédérale de Lausanne (EPFL), STI-IMX-LMC, 1015 Lausanne, Switzerland

ARTICLE INFO

Article history:
Received 2 January 2020
Revised 12 May 2020
Accepted 11 June 2020
Available online 13 June 2020

Keywords:
Reactive crystallization
Population balance
Direct quadrature method of moments
Uncertainty/sensitivity analysis
Nanoparticles
Calcium silicate hydrate (C-S-H)

ABSTRACT

Recently, we developed a population balance framework describing the precipitation of calcium-silicate-hydrate, a key nanomaterial in the construction industry and with potential applications in biomedicine, environmental remediation, and catalysis. In this article, we first refine our computational workflow by developing a more efficient and robust method for the solution of the moment-transformed population balance equations. Then, we generalize our framework by coupling to PHREEQC, a widely used open-source speciation solver, to enhance the adaptability of the framework to new systems. Using this improved computational model, we perform global uncertainty/sensitivity analysis (UA/SA) to understand the effect of variations in the model parameters and experimental conditions on the properties of the product. With the specific surface area of particles as an example, we show that UA/SA identifies the factors whose control would allow a fine-tuning of the desired properties. This general approach can be transferred to other nanoparticle synthesis schemes as well.

1. Introduction

Calcium-silicate-hydrate ($CaO\text{-}SiO_2\text{-}H_2O$ or C-S-H for short) is the most important phase formed during the hydration of cementitious materials (Scrivener et al., 2015). Aside from its key role in the construction industry (Scrivener et al., 2015), C-S-H has recently found diverse applications in environmental clean-up (Shao et al., 2018; Zhang et al., 2015; Zhao et al., 2014), biomedicine (Jiao, 2016; Wu et al., 2013; Zhu and Sham, 2014), and even catalysis (Sheng et al., 2019; Xia et al., 2019). In the biomedical field, for instance, it offers good bioactivity, biocompatibility, and biodegradability (Wu et al., 2013; Zhu and Sham, 2014). Besides these characteristics, the inherent nanostructured construct of C-S-H, provides high surface areas, and its relatively low-cost preparation warrants further research for applications where interfaces play a major role (Shao et al., 2018; Wu et al., 2013).

Recently, we developed a formalism to model the nucleation and growth of C-S-H using a population balance equation (PBE) framework (Andalibi et al., 2018). The theoretical framework was fitted to the experimental data collected on the precipitation of a synthetic C-S-H with Ca:Si = 2, prepared under controlled conditions resembling the process of cement hydration (in terms of temporal supersaturation ratio) (Andalibi et al., 2018; Kumar et al., 2017). We estimated the optimal values for the unknown model parameters and explained procedures for the extraction of various output information from the simulation. Additionally, we assessed the merit of our computations by comparing the optimal physical parameters and various outputs against the literature data, wherever available (Andalibi et al., 2018).

Here, we build on our previous work and implement two pivotal refinements to improve the simulation speed, robustness, and generality. Specifically, we replace our *ad hoc* equilibrium solver in the previous work with PHREEQC, a popular freely-available tool widely used for thermodynamic speciation calculations (Parkhurst and Appelo, 2013). This allows for a more straightforward adaptation to new precipitation scenarios and opens up the possibility of utilizing the large thermodynamic databases already included within the software (Parkhurst and Appelo, 2013). Additionally, we employ the direct quadrature method of moments (DQMOM) for the solution of the PBE, which has several advantages over our previously used QMOM approach (Dale et al., 2017; Haderlein et al., 2017; Marchisio and Fox, 2005). We give a detailed derivation of DQMOM and relevant subtleties critical to the robust and reliable performance of the method.

* Corresponding author.
E-mail address: mra55@cam.ac.uk (M.R. Andalibi).
[1] Present address: Department of Chemical Engineering and Biotechnology, West Cambridge Site, Philippa Fawcett Drive, Cambridge, CB3 0AS, United Kingdom

https://doi.org/10.1016/j.compchemeng.2020.106971

Nomenclature

Symbols

$\boldsymbol{A}$	Coefficient matrix in DQMOM equation
$\boldsymbol{A}_1$, $\boldsymbol{A}_2$	Constituent matrices of $\boldsymbol{A_{nm}}$
$\boldsymbol{A}_{1,\ nm}$, $\boldsymbol{A}_{2,\ nm}$	Constituent matrices of
$\boldsymbol{A_{nm}}$	Similar to $\boldsymbol{A}$ but with discretization nodes expressed in nm
$\boldsymbol{a}$	Vector of a_α terms for all the discretization nodes
a_r	Crystallite aspect ratio
a_α	Time-derivative of w_α
$\boldsymbol{b}$	Vector of b_α terms for all the discretization nodes
b_α	Time-derivative of ς_α
$b_{\alpha,\ nm}$	Time-derivative of $\varsigma_{\alpha,\ nm}$ (discretization nodes expressed in nm)
$\boldsymbol{d}$	Vector of moment source term
$\boldsymbol{d_{nm}}$	Similar to d but with discretization nodes expressed in nm
$\boldsymbol{d}^{\boldsymbol{G}}_{\boldsymbol{nm}}$	Contribution of growth to $\boldsymbol{d_{nm}}$
$\boldsymbol{d}^{\boldsymbol{N}}_{\boldsymbol{nm}}$	Contribution of nucleation to $\boldsymbol{d_{nm}}$
$\boldsymbol{d}^{\boldsymbol{Volume}}_{\boldsymbol{nm}}$	Contribution of reaction volume changes to $\boldsymbol{d_{nm}}$
F	Cumulative distribution function of the scaler model output
g	Kinetic order of growth
J_{hom}	Rate of primary homogeneous nucleation
J_{sec}	Rate of true catalytic secondary nucleation
k	Order of NDF moments
k_r	Growth rate coefficient
L	Crystallite characteristic size
L_α	Crystallite characteristic size of discretization node α
$L_{\alpha,\ nm}$	Crystallite characteristic size of discretization node α in nm
$\bar{L}_c$	Average crystallite thickness
$\bar{L}_{nm}$	Average of the crystallite characteristic sizes at different nodes (abscissas) in nm
$\bar{L}_p$	Average particle edge length
L^*_{hom}	Critical nucleus size of primary nucleation in m
L^*_{sec}	Critical nucleus size of secondary nucleation in m
M	Number of uncertain parameters subject to UA/SA
m_k	Moment of NDF of order k
N	Overall number of the nodes in DQMOM or the base sample size in UA/SA
n	Number density function (NDF) (for crystallites)
n_e	Molar abundance of element e
n_p	Particle number concentration
$\boldsymbol{P}$	Preconditioning matrix
r	Number of finite differences in Elementary Effect Test
S	Source term embracing all the solid formation/transformation processes
S_G	Source term for crystal growth
S_i	PAWN sensitivity index or VBSA main effect for input parameter i
S_N	Source term for nucleation
S_{Ti}	Total-order sensitivity index for input parameter i in VBSA
$\bar{S}_k$	Moment source term (Eq. 12)
$\bar{S}^G_k$	Moment source term for growth
$\bar{S}^{Volume}_k$	Moment source term for changes in the system volume (flow systems)
SSA_c or $SSA_{Crystallite}$	Specific surface area of crystallites (m^2/g solid)
SSA_p or $SSA_{Particle}$	Specific surface area of particles (m^2/g solid)
t	Time
V	Volume of the reaction suspension
$V(y)$	Unconditional output variance (in VBSA)
w_α	Wight of discretization node α
$\boldsymbol{X}$	Input sample matrix for sensitivity analysis
$\boldsymbol{X_A}$, $\boldsymbol{X_B}$, $\boldsymbol{X_C}$	Input sample matrices constituting the Sobol' sequences
$\boldsymbol{x}$	Vector of model input factors (UA/SA)
y	Scaler model output (UA/SA)
Greek letters	:
α	Discretization node index
$\boldsymbol{\alpha}$	Vector composed of a_α and b_α terms for all the discretization nodes
$\boldsymbol{\alpha_{nm}}$	Similar to $\boldsymbol{\alpha}$ but with discretization nodes expressed in nm
γ	Interfacial tension
Δ^j_i	Perturbation in EET
δ	Dirac delta function
δ'	First derivative of the delta function
μ_{EET}	EET sensitivity index
σ	Cohesion energy
σ_{EET}	Standard deviations of finite differences in EET
ς_α	Weighted abscissa of discretization node α ($\equiv w_\alpha L_\alpha$)
$\varsigma_{\alpha,\ nm}$	Weighted abscissa of discretization node α with sizes expressed in nm ($\equiv w_\alpha L_{\alpha,\ nm}$)

Having this improved simulation framework, we assess the behavior of the C-S-H precipitation model by applying global uncertainty/sensitivity analysis (UA/SA) with different model parameters as the source of uncertainty. The propagation of uncertainty into different model outputs such as crystallite dimensions, particle edge length, specific surface areas, and precipitation yield is examined thoroughly using three different methods, namely, PAWN (derived from the developers names, Pianosi and Wagener) (Pianosi and Wagener, 2018, 2015), Elementary Effect Test (Morris, 1991; Saltelli et al., 2007), and variance-based sensitivity analysis (VBSA) (Saltelli et al., 2010, 2007). The application of these complementary methods enables unambiguous appraisal of the model performance, which in turn facilitates complexity reduction, *i.e.*, by fixing uninfluential parameters to reasonable values. This also allows for a more robust calibration during the regression to experimental data (Pianosi et al., 2016; Saltelli et al., 2007). Additional implications of such global UA/SA concerns the robustness of model predictions in response to different sources of uncertainty/variability in the model parameters (Pianosi et al., 2016; Saltelli et al., 2007). Besides these outcomes, our work provides the first example of UA/SA on a kinetic model of precipitation, and thus can serve as a benchmark for future studies in this direction (refer to (Iversen and Sin, 2019; Öner et al., 2020) for more recent examples of UA/SA applied to chemical process systems).

Once we obtained a comprehensive understanding about the model structure, we implement another UA/SA on a model of reduced complexity and incorporate uncertainty from different experimental conditions. The goal is to aid the design of nanoparticulate products by gaining insight from computer experiments. Often, optimal operating conditions for a synthesis protocol are found using one-at-a-time (OAT) experimental designs (Saltelli et al., 2007; Wu et al., 2013). An example of this practice in the preparation of synthetic C-S-H is the work by Wu *et al.* where ultrathin particles possessing surface areas as high as 505 m^2/g were obtained by tuning different experimental conditions such as reactant concentration, addition rate, aging time, synthesis temperature, and the solvent used for washing the precipitate during filtration. Besides covering a limited space of the possible input conditions, this approach overlooks the probable interaction effects between various combinations of the inputs. The latter could produce drastically different behavior compared to when only one input parameter is changed at a time (Saltelli et al., 2007). A global UA/SA circumvents these limitations and offers an inexpensive alternative to examine a wide range of operating conditions varied in an all-at-a-time fashion (Öner et al., 2020; Pianosi et al., 2016; Saltelli et al., 2007). With this approach, we propose practical recommendations in order to improve the properties of the final product. As an example, we demonstrate the key influence of reagent addition rate, in a well-mixed semi-batch reactor, on the accessible specific surface area of the final product. This can be further compounded with adjustments in the solution chemistry to obtain a product with distinctly higher specific surface area.

2. Theory

The overall computational workflow for PBE modeling of precipitation processes is explained in detail in our recent articles (Andalibi et al., 2018; Carino et al., 2018, 2017) and other literature (Haderlein et al., 2017; Peng et al., 2015; Schroeder et al., 2014; Schwarzer et al., 2006). Therefore, in this section we will focus on the developments brought forward by the current work. First, we will review the essential characteristics of the precipitation system to be studied, and enumerate the limitations of our previous work. Then, we will explain the application of DQMOM to solve the PBE model for a well-mixed system with crystallite size as the internal coordinate. Finally, we will present the underlying idea behind different sensitivity measures employed to assess the input-output relationships in the overall coupled thermodynamic-kinetic framework.

2.1. Population balance equation and its solution using DQMOM

Consider a population of crystallites whose size distribution is evolving over time. The number density function (NDF; n) describing this distribution is a function of time (t) and crystallite characteristic size (L), and can be approximated by a discretized distribution as

$$n(L,t) = \sum_{\alpha=1}^{N} w_\alpha(t)\delta(L - L_\alpha(t)) \tag{1}$$

where α denotes the discretization nodes with weights w_α at sizes (or abscissas) L_α, N is the overall number of the nodes, and δ is the Dirac delta function (Marchisio and Fox, 2013, 2005). For a homogeneous system (namely, one with uniformity across the physical space), the temporal evolution in NDF can be expressed using the so-called population balance equation (PBE)

$$\frac{\partial n}{\partial t} = S(L) \tag{2}$$

where S is a source term embracing all the solid formation/transformation processes such as nucleation, growth, and aggregation, inflows and outflows of crystallites, and possible changes in the volume of reaction liquor. In the kinetic modeling of precipitation processes, the PBE is solved along with differential equations written for mass balances (*viz.*, the conservation of elements inside the reactor).

Substituting Eq. (1) in Eq. (2) (the time-dependences are dropped for simplicity)

$$\sum_{\alpha=1}^{N} \frac{dw_\alpha}{dt}\delta(L - L_\alpha) - \sum_{\alpha=1}^{N} w_\alpha \frac{dL_\alpha}{dt}\delta'(L - L_\alpha) = S(L) \tag{3}$$

where $\delta'(L - L_\alpha)$ is the first derivative of the generalized delta function (Arfken et al., 2012). Now, defining the weighted abscissas

$$\varsigma_\alpha = w_\alpha L_\alpha \tag{4}$$

we will have the following ODE upon substitution in Eq. (3)

$$\sum_{\alpha=1}^{N} \frac{dw_\alpha}{dt}\delta(L - L_\alpha) - \sum_{\alpha=1}^{N} \left(\frac{d\varsigma_\alpha}{dt} - L_\alpha \frac{dw_\alpha}{dt}\right)\delta'(L - L_\alpha) = S(L) \tag{5}$$

Defining

$$\begin{aligned} a_\alpha &\equiv \frac{\partial w_\alpha}{\partial t} \\ b_\alpha &\equiv \frac{\partial \varsigma_\alpha}{\partial t} \end{aligned} \tag{6}$$

followed by substitution in Eq. (5), gives

$$\sum_{\alpha=1}^{N} a_\alpha\left[\delta(L - L_\alpha) + L_\alpha\delta'(L - L_\alpha)\right] - \sum_{\alpha=1}^{N} b_\alpha\delta'(L - L_\alpha) = S(L) \tag{7}$$

Applying the moment transformation defined as

$$m_k(t) \equiv \int_0^\infty L^k\ n(t,L)dL \cong \sum_{\alpha=1}^{N} w_\alpha(t)L_\alpha^k \tag{8}$$

and knowing (Arfken et al., 2012)

$$\int_0^\infty L^k\delta(L - L_\alpha)\ dL = L_\alpha^k \tag{9}$$

$$\int_0^\infty L^k\delta'(L - L_\alpha)dL = -kL_\alpha^{k-1} \tag{10}$$

we transform Eq. (7) by multiplying both sides by L_α^k and integrating over $[0, \infty)$ yielding

$$(1-k)\sum_{\alpha=1}^{N} L_\alpha^k a_\alpha + k\sum_{\alpha=1}^{N} L_\alpha^{k-1} b_\alpha = \int_0^\infty L^k S(L)dL \tag{11}$$

Now, defining the moment source term $\bar{S}_k$

$$\bar{S}_k \equiv \int_0^\infty L^k S(L)dL \tag{12}$$

Eq. (11) can be recast into a matrix form as (boldface symbols are vectors and matrices)

$$A\alpha = [\boldsymbol{A}_1, \boldsymbol{A}_2]\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\end{bmatrix} = d \tag{13}$$

where

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ -L_1^2 & \cdots & -L_N^2 \\ \vdots & \vdots & \vdots \\ (1-k)L_1^k & \cdots & (1-k)L_N^k \\ \vdots & \vdots & \vdots \\ 2(1-N)L_1^{2N-1} & \cdots & 2(1-N)L_N^{2N-1} \end{bmatrix}_{2N\times N} \tag{14}$$

$$\boldsymbol{A}_2 = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 1 \\ 2L_1 & \cdots & 2L_N \\ \vdots & \vdots & \vdots \\ kL_1^{k-1} & \cdots & kL_N^{k-1} \\ \vdots & \vdots & \vdots \\ (2N-1)L_1^{2N-2} & \cdots & (2N-1)L_N^{2N-2} \end{bmatrix}_{2N\times N} \tag{15}$$

$$\alpha = [a_1, a_2, \ldots, a_N, b_1, b_2, \ldots, b_N]^T = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b} \end{bmatrix}_{2N\times 1} \tag{16}$$

and for a well-mixed homogeneous system (namely, no dependence on physical space)

$$d = \left[\bar{S}_0, \bar{S}_1, \ldots \bar{S}_k, \ldots, \bar{S}_{2N-1}\right]^T \tag{17}$$

Note that since $\boldsymbol{A}$ is not dependent on w_α it can be defined even for nodes with zero weight.

At this juncture, let us elaborate on some technical issues associated with DQMOM and their resolution. For nanoparticle formation and transformation systems, matrix $\boldsymbol{A}$ can be extremely ill-conditioned, complicating the accurate solution of the linear system in Eq. (13) (Dale et al., 2017). The first reason for this ill-conditioning can readily be understood by looking at the rows of matrix $\boldsymbol{A}$ which are composed of abscissas raised to nonnegative integers. Working with SI units and with crystallite sizes in the order of nanometers (10^{-9} m), the rows of matrix $\boldsymbol{A}$ will assume very different scales spanning several orders of magnitude (*e.g.*, 44 and 35 orders of magnitude variability in $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$, respectively for N=3). To alleviate this problem, we can follow the abscissas in nanometers. Thus, defining $L_{\alpha,nm} \equiv L_\alpha \times 10^{-9}$ and substitution in Eq. (11)

$$\begin{aligned}(1-k)\sum_{\alpha=1}^{N} L_{\alpha,nm}^k a_\alpha \times 10^{-9k} + k\sum_{\alpha=1}^{N} L_{\alpha,nm}^{k-1} b_\alpha \times 10^{-9(k-1)} \\ = \int_0^\infty L^k S(L)dL\end{aligned} \tag{18}$$

Moreover, substitution in Eq. (6) gives $b_{\alpha,\,nm}$ as

$$b_\alpha \equiv \frac{d(w_\alpha L_\alpha)}{dt} = \frac{d\left(w_\alpha L_{\alpha,nm} \times 10^{-9}\right)}{dt} = 10^{-9} \times b_{\alpha,nm} \tag{19}$$

Inserting this unit adjusted variable into Eq. (18) yields

$$(1-k)\sum_{\alpha=1}^{N} L_{\alpha,nm}^k a_\alpha + k\sum_{\alpha=1}^{N} L_{\alpha,nm}^{k-1} b_{\alpha,nm} = 10^{9k}\int_0^\infty L^k S(L)dL \tag{20}$$

which in matrix form reads

$$\boldsymbol{A_{nm}\alpha_{nm}} = [\boldsymbol{A}_{1,nm}, \boldsymbol{A}_{2,nm}]\begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b_{nm}} \end{bmatrix} = \boldsymbol{d_{nm}} \tag{21}$$

with

$$\boldsymbol{A}_{1,nm} = \begin{bmatrix} 1 & \cdots & 1 \\ 0 & \cdots & 0 \\ -L_{1,nm}^2 & \cdots & -L_{N,nm}^2 \\ \vdots & \vdots & \vdots \\ (1-k)L_{1,nm}^k & \cdots & (1-k)L_{N,nm}^k \\ \vdots & \vdots & \vdots \\ 2(1-N)L_{1,nm}^{2N-1} & \cdots & 2(1-N)L_{N,nm}^{2N-1} \end{bmatrix}_{2N\times N} \tag{22}$$

$$\boldsymbol{A}_{2,nm} = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 1 \\ 2L_{1,nm} & \cdots & 2L_{N,nm} \\ \vdots & \vdots & \vdots \\ kL_{1,nm}^{k-1} & \cdots & kL_{N,nm}^{k-1} \\ \vdots & \vdots & \vdots \\ (2N-1)L_{1,nm}^{2N-2} & \cdots & (2N-1)L_{N,nm}^{2N-2} \end{bmatrix}_{2N\times N} \tag{23}$$

$$\boldsymbol{\alpha_{nm}} = [a_1, a_2, \ldots, a_N, b_{1,nm}, b_{2,nm}, \ldots, b_{N,nm}]^T = \begin{bmatrix} \boldsymbol{a} \\ \boldsymbol{b_{nm}} \end{bmatrix}_{2N\times 1} \tag{24}$$

$$\boldsymbol{d_{nm}} = \left[\bar{S}_0, \bar{S}_1 \times 10^9, \ldots \bar{S}_k \times 10^{9k}, \ldots, \bar{S}_{2N-1} \times 10^{9(2N-1)}\right]^T \tag{25}$$

Note that the Eqs. (21-25) are the same as Eqs. (13-17) but expressed in terms of abscissas in nm. In our experience, this scaling procedure is an inevitable step in the application of DQMOM to the process of nanoparticle formation. Yet another reduction in the condition number of matrix $\boldsymbol{A_{nm}}$ can be readily achieved by preconditioning which makes the convergence of the iterative solution more robust and faster (Press et al., 2007). This is particularly important when dealing with particulate processes that give rise to sharp changes in the particle phase space (*e.g.*, nucleation or aggregation) (Dale et al., 2017). Here, we applied a left preconditioning using a diagonal matrix with main diagonal elements

$$P_{ii} = \frac{\bar{L}_{nm}^{i-1} + \bar{L}_{nm}^{i-2}}{2}; i = 1, \ldots, 2N \tag{26}$$

where $\bar{L}_{nm}$ is the average of the abscissas in nm

$$\bar{L}_{nm} = \frac{1}{N}\sum_{\alpha=1}^{N} L_{\alpha,nm} \tag{27}$$

Therefore, instead of solving the system of linear equations in Eq. (21) we solve

$$\boldsymbol{P}^{-1}\boldsymbol{A_{nm}\alpha_{nm}} = \boldsymbol{P}^{-1}\boldsymbol{d_{nm}} \tag{28}$$

With $\boldsymbol{P}$ being diagonal, its inverse $\boldsymbol{P}^{-1}$ can trivially be obtained *via* inverting the main diagonal elements (Arfken et al., 2012).

After these considerations, solving Eq. (28) at each time step of integrating the set of ordinary differential equations (ODE set composed of PBE + mass balances) yields $\boldsymbol{a}$ and $\boldsymbol{b_{nm}}$. For the ODE solvers to work efficiently, the dependent variables have to be properly scaled (Shampine et al., 2003). Here, having the weights in the range of 10^{20} crystallites.m^{-3} (or higher), both $a_\alpha = \frac{\partial w_\alpha}{\partial t}$ and $b_{\alpha,nm} = \frac{\partial (w_\alpha L_{\alpha,nm})}{\partial t}$ can be extremely large, especially during the burst of nucleation (nucleation rates are in excess of 10^{16} crystallites.m^{-3}.s^{-1} for typical model parameters (Andalibi et al., 2018)). To bring these values to the order of unity, we solved the ODE set for $\log_{10}$ of the weights and weighted abscissas

$$\frac{d(\log_{10} w_\alpha)}{dt} = \frac{dw_\alpha}{dt} \times \frac{1}{w_\alpha \ln 10} \tag{29}$$

$$\frac{d(\log_{10} \varsigma_{\alpha,nm})}{dt} = \frac{d\varsigma_{\alpha,nm}}{\partial dt} \times \frac{1}{\varsigma_{\alpha,nm} \ln 10} \tag{30}$$

where $\varsigma_{\alpha,\,nm} \equiv w_\alpha L_{\alpha,\,nm}$.

Now, let us describe the constituents of the source term $\bar{S}_k$. For molecular growth (which could be size-dependent (Andalibi et al., 2018)) in a homogenous system we have (Marchisio and Fox, 2005)

$$S_G(L) = -\frac{\partial}{\partial L}[n(L,t)G(L)] \tag{31}$$

Therefore,

$$\bar{S}_k^G = \int_0^\infty L^k S_G(L)dL = k\int_0^\infty L^{k-1} G(L) \times n(L,t)dL \tag{32}$$

which was obtained using integration by parts (Arfken et al., 2012). Now, using Eq. (1)

$$\bar{S}_k^G = k\sum_{\alpha=1}^{N} w_\alpha L_\alpha^{k-1} G(L_\alpha) \tag{33}$$

In fact, this is the N-point quadrature approximation of the growth source term for moment order k (Andalibi et al., 2018; Marchisio and Fox, 2005). In matrix format and for abscissas in nanometers

$$\boldsymbol{d_{nm}^G} = \begin{bmatrix} 0 \\ \sum_{\alpha=1}^{N} w_\alpha G(L_\alpha) \times 10^9 \\ \vdots \\ k\sum_{\alpha=1}^{N} w_\alpha L_\alpha^{k-1} G(L_\alpha) \times 10^{9k} \\ \vdots \\ (2N-1)\sum_{\alpha=1}^{N} w_\alpha L_\alpha^{(2N-2)} G(L_\alpha) \times 10^{9(2N-1)} \end{bmatrix} = \boldsymbol{diag}\left(1, 10^9, \ldots, 10^{9k}, \ldots, 10^{9(2N-1)}\right) \times \boldsymbol{A_2} \times diag(w_1, w_2, \ldots, w_N) \times \begin{bmatrix} G(L_1) \\ G(L_2) \\ \vdots \\ G(L_N) \end{bmatrix} \tag{34}$$

where ***diag***() denotes a diagonal matrix.

For nucleation, the source term is (Marchisio and Fox, 2005)

$$S_N(L) = J_{hom}\delta\left(L - L_{hom}^*\right) + J_{sec}\delta\left(L - L_{sec}^*\right) \tag{35}$$

where J_{hom} and J_{sec} are the rates of primary homogeneous and true secondary nucleation processes (crystallites.m^{-3}.s^{-1}) while L_{hom}^* and L_{sec}^* are the respective critical nucleus sizes (in m). Therefore, the corresponding moment source term ($\bar{S}_k^N$) is

$$\bar{S}_k^N = \int_0^\infty L^k S_N(L) dL = {L_{hom}^*}^k J_{hom} + {L_{sec}^*}^k J_{sec} \tag{36}$$

Again, in matrix form

$$\boldsymbol{d_{nm}^N} = diag\left(1, 10^9, \ldots, 10^{9k}, \ldots, 10^{9(2N-1)}\right) \times \begin{bmatrix} 1 & 1 \\ L_{hom}^* & L_{sec}^* \\ \vdots & \vdots \\ {L_{hom}^*}^k & {L_{sec}^*}^k \\ \vdots & \vdots \\ {L_{hom}^*}^{2N-1} & {L_{sec}^*}^{2N-1} \end{bmatrix} \times \begin{bmatrix} J_{hom} \\ J_{sec} \end{bmatrix} \tag{37}$$

Finally, the source term for changes in the system volume (for flow systems) reads

$$\bar{S}_k^{Volume} = -m_k \frac{d(\ln V)}{dt} \tag{38}$$

where V is the volume of the reaction suspension. Again, in matrix form

$$\boldsymbol{d_{nm}^{Volume}} = -\frac{d(\ln V)}{dt} diag\left(1, 10^9, \ldots, 10^{9k}, \ldots, 10^{9(2N-1)}\right) \times \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_{2N-1} \end{bmatrix} \tag{39}$$

Another complication arises from the fact that PBE only tracks the evolution of crystallites. Therefore, an additional differential equation is required to account for the time variation of particle number concentration (n_p; particles.m^{-3}) (Andalibi et al., 2018; Carino et al., 2017). Again, with n_p adopting very large values we solve the ODE for $\log_{10}$ of n_p

$$\frac{d(\log_{10} n_p)}{dt} = \frac{dn_p}{dt} \times \frac{1}{n_p \ln 10} \tag{40}$$

$$\frac{dn_p}{dt} = -n_p \frac{d(\ln V)}{dt} + J_{hom} \tag{41}$$

Besides the PBE set and the ODE for n_p, we have to solve for mass balances over elemental abundances (n_e) inside the solution. Since in the current precipitation system the molar amounts are in mmol range the derivatives are defined in terms of mmol.s^{-1} to bring them closer to the order of unity. The rate of precipitate formation therefore provides the coupling between kinetics and thermodynamic speciation calculation at each time step of integrating the ODE set (Andalibi et al., 2018; Haderlein et al., 2017; Schroeder et al., 2014). We will discuss the estimation of the thermodynamic driving force for precipitation later in Section 3.1.

2.2. Description of precipitation system and limitations of our previous work

The precipitation system studied here is the formation of synthetic C-S-H with Ca:Si ratio 2 (Andalibi et al., 2018; Kumar et al., 2017). The precipitate is composed of nanofoils that are a few nm thick and in the order of 100 nm wide (Fig. 1(a) and (b)). These two-dimensional nanoparticles are made up of highly defective crystallites, of a thickness typically below 10 nm, which are arranged with liquid crystalline-type orientational order (Fig. 1 (b) and (c)). Recently, we proposed a pathway (Fig. 1 (d)) for the formation of this nanoparticulate material and tested that by regressing the experimental kinetic data using a computational model based on population balance equation (PBE) modeling. The framework included primary nucleation, true catalytic secondary nucleation, and molecular growth, and accounted for the time evolution of the precipitation driving force by applying thermodynamic equilibrium to the reactions among the aqueous species (local equilibrium assumption (Andalibi et al., 2018; Haderlein et al., 2017)). From a mathematical perspective, the framework consisted of a set of ordinary differential equations (ODEs) written for the dynamic evolution in the moments of crystallite size distribution (the PBE part) and elemental amounts in the system (the mass balances). Our computational model showed very good plausibility in terms of the goodness of fit, the consistency of the regressed model parameters with respect to the knowledge from the literature, and reasonable mechanistic and kinetic predictions. This includes, for instance, the predicted size of crystallites and particles which were compatible with previous experimental and theoretical observations, or the invariably undersaturated state of solution with respect to portlandite similar to experimental observations (Andalibi et al., 2018).

In the computational framework mentioned above, the PBE set was solved using the quadrature method of moments (QMOM) (Andalibi et al., 2018; Marchisio et al., 2003). Although QMOM is a reliable and popular method for this task (Marchisio and Fox, 2005; Schroeder et al., 2014; Silva et al., 2010), a widely used variation called direct quadrature method of moments (DQMOM) offers several advantages. For instance, QMOM requires a moment inversion algorithm at every time step to find the discrete approximation to the size distribution. This is often an ill-conditioned problem and reduces the computational efficiency of the method drastically (Haderlein et al., 2017; Marchisio and Fox, 2013; Silva et al., 2010). On the contrary, DQMOM directly follows the discrete abscissas and weights approximating the size distribution, and employs commonly used numerical methods such as matrix inversion instead of moment inversion algorithms. This

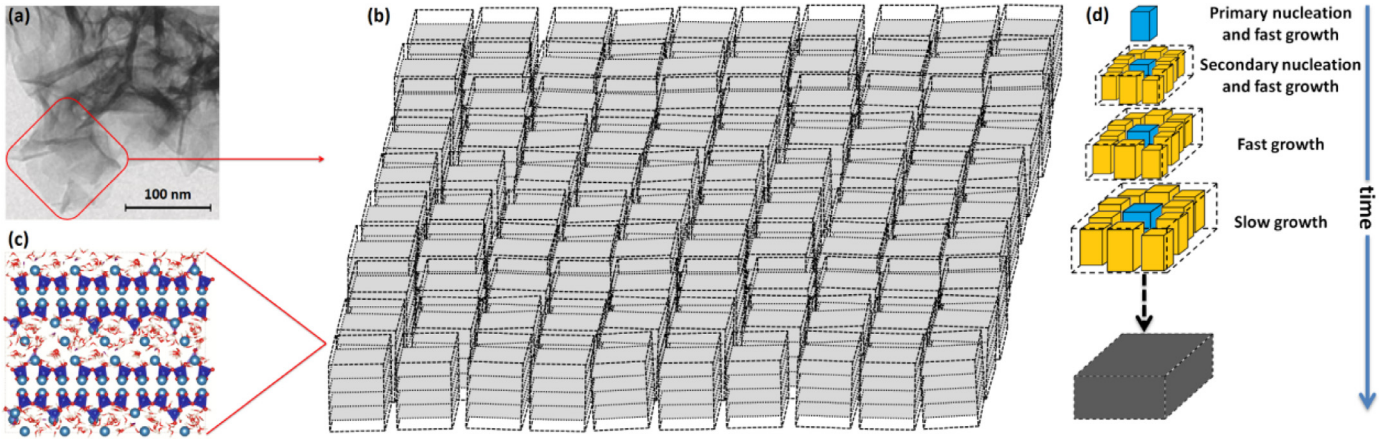

Fig. 1. Summary of synthetic C-S-H precipitation system studied here. (a) Transmission electron micrograph of C-S-H particles with foil like morphology; (b) schematic representation of C-S-H nanofoils composed of defective crystallites nematically ordered in two dimensions; (c) internal structure of C-S-H crystallites from atomistic simulations (Kumar et al., 2017; Kunhi Mohamed et al., 2018); (d) the proposed precipitation pathway for synthetic C-S-H of Ca:Si=2 (Andalibi et al., 2018). Adapted with permission from Ref. (Andalibi et al., 2018) Copyright 2018 The Royal Society of Chemistry.

makes the method more convergent and extremely fast (Dale et al., 2017; Haderlein et al., 2017; Silva et al., 2010), to an extent that Haderlein *et al.* (Haderlein et al., 2017) have employed it for the development of "flow-sheeting" software tools. Other benefits of DQMOM over QMOM are its more efficient coupling to fluid dynamics and straightforward extension to multivariate distributions (namely, with more than one internal coordinate) (Marchisio and Fox, 2013, 2005). The latter is particularly important in the case of C-S-H as oftentimes the precipitation leads to variable composition solid solutions (Lothenbach et al., 2019) necessitating the application of at least two internal coordinates, namely, size and composition (Andalibi et al., 2018).

Another limitation of our previous work is related to the manner of applying the local equilibrium assumption. There, we calculated the supersaturation ratio at fixed time steps selected depending on how fast the precipitation consumes the precursors (Andalibi et al., 2018). Therefore, supersaturation ratio was an externally calculated quantity provided to the ODE function (the function calculating the derivatives as a function of time and dependent variables (Shampine et al., 2003)). This is very similar to the approach adopted by Myerson *et al.* who only recalculated the supersaturation ratio when there was significant (more than 0.1%) change in the amount of the precipitate (Peng et al., 2015). This approach was employed to minimize the computational burden of speciation calculations. Additionally, similar to the work by Haderlein *et al.* (Haderlein et al., 2017) and Galbraith and Schneider (Galbraith and Schneider, 2014), a bespoke speciation solver was developed to expedite the overall simulation (Andalibi et al., 2018). Even though our equilibrium solver and that by Haderlein *et al.* are developed in a general fashion, this practice limits the applicability of the developed tools, as it requires setting up a thermodynamic database for every single new scenario. Instead, there are a number of powerful freely-available thermodynamic solvers with huge databases already implemented, such as PHREEQC and GEMS (Kulik et al., 2013; Parkhurst and Appelo, 2013). Therefore, in this work we will present a general-purpose protocol for the coupling of PHREEQC to the PBE simulations of precipitation processes with the former imbedded within the ODE function.

2.3. Uncertainty/sensitivity assessment (UA/SA)

To quantitatively apportion the variability in the output of the PBE simulation to various sources of uncertainty in the input (factors), we applied model-independent sensitivity analysis using three popular methods: PAWN (Pianosi and Wagener, 2018, 2015), Elementary Effect Test or the method of Morris (EET) (Morris, 1991; Saltelli et al., 2007), and variance-based sensitivity analysis (VBSA) using quasi-Monte Carlo samples generated by the method of Sobol' (Saltelli et al., 2010, 2007). In this section, we will briefly explain these concepts to facilitate the comprehension of the results. The interested reader is referred to the relevant literature for a more in depth discussion (Pianosi et al., 2016; Saltelli et al., 2007).

As Saltelli *et al.* (Saltelli et al., 2007) argued, UA/SA consists in the examination of uncertainty in parameters (input factors) propagating through a mathematical model all the way to the model outputs (Saltelli et al., 2007). One way to do this task is through Monte Carlo analysis, wherein a set of row vectors are generated by sampling the input variability space of different model parameters. The accumulation of these sets gives an input matrix with each row corresponding to a set of model parameters whose introduction to the model allows a single simulation run. Therefore, any UA/SA requires an input sample matrix. When the model describes a physicochemical process, the input factors can broadly be classified in two groups. The first group includes the operating conditions (*e.g.*, temperature and reactant concentrations) whose tuning can potentially change the course of the process and the properties of the final product. The second group consists of the uncertain model parameters such as those in the kinetic equations (*e.g.*, growth rate coefficient and interfacial tension).

In order to generate the input sample matrix described earlier, low-discrepancy Sobol' sequences were constructed by first generating an input sample $\boldsymbol{X}$ of size $2N \times M$, where N is the *base* sample size, and M denotes the number of uncertain parameters (*e.g.*, $M = 5$ for five model parameters subject to SA). Sample $\boldsymbol{X}$ was generated using Latin hypercube sampling (LHS) strategy (Pianosi et al., 2016, 2015; Saltelli et al., 2007). This sample was then resampled to build three matrices $\boldsymbol{X_A}$, $\boldsymbol{X_B}$, and $\boldsymbol{X_C}$, where $\boldsymbol{X_A}$ and $\boldsymbol{X_B}$ are simply the first and last N rows of $\boldsymbol{X}$, respectively, while $\boldsymbol{X_C}$ is a block matrix of M recombinations of $\boldsymbol{X_A}$ and $\boldsymbol{X_B}$

$$\boldsymbol{X_C} = \begin{bmatrix} \boldsymbol{X_{C,1}} \\ \boldsymbol{X_{C,2}} \\ \vdots \\ \boldsymbol{X_{C,M}} \end{bmatrix} \tag{42}$$

where $\boldsymbol{X_{C,\,i}}$ is an $N \times M$ matrix whose columns are all taken from $\boldsymbol{X_B}$ except for the i^{th} column which is taken from $\boldsymbol{X_A}$ (Saltelli et al., 2007). The resampling is done because of the particular way the

estimation approach is designed to calculate the sensitivity indices in VBSA *via* a quasi-Monte Carlo method (see Eqs. (4.21) and (4.23), and the discussion therein, in (Saltelli et al., 2007)). Once we have $\boldsymbol{X_A}$, $\boldsymbol{X_B}$, and $\boldsymbol{X_C}$, the computational model should be run with all the rows in sample matrices as input model parameters giving a total of $N \times (M+2)$ sample outputs.

The first SA method applied here is PAWN, a density-based (or moment-independent) method recently developed by Pianosi and Wagener (Pianosi and Wagener, 2018, 2015). The central idea behind this method is to compute sensitivity through variations in the cumulative density function (CDF) of the output, induced by fixing one input factor. In practice, this is achieved *via* estimating the divergence between unconditional output CDF, namely that generated by varying all the input factors, and the conditional CDF generated by fixing an individual factor to a prescribed value. Several values within the input variability space can be assigned to the prescribed value, a practice referred to as multiple conditioning, to generate a number of divergence values that can be aggregated in some kind of statistic (Pianosi et al., 2016; Pianosi and Wagener, 2018, 2015). In PAWN, the divergence is expressed in terms of Kolmogorov-Smirnov (KS) statistics which is the maximum vertical distance between the conditional and unconditional CDFs (Pianosi and Wagener, 2018, 2015). In mathematical terms, defining the input-output relation as $y = f(\boldsymbol{x})$ (with $\boldsymbol{x} = [x_1, x_2, \ldots, x_M]^T$ a vector of M input factors and y a scaler output), the PAWN sensitivity index for the i-th input factor is defined as (Pianosi and Wagener, 2018)

$$S_i = \underset{x_i}{stat}\ \underset{y}{max} \left| F_y(y) - F_{y|x_i}(y|x_i) \right| \tag{43}$$

In this equation, $F_y(y)$ and $F_{y|x_i}(y|x_i)$ denote the unconditional and conditional CDFs of the output y, respectively, and $\underset{x_i}{stat}$ represents a statistic (mean, median, or maximum) that can be selected by the user over the multiple conditioning intervals. Furthermore, the term $\underset{y}{max} |F_y(y) - F_{y|x_i}(y|x_i)|$ is merely the KS statistic described earlier (consult Fig.1 of (Pianosi and Wagener, 2018) for a concise schematic summary of the ideas behind PAWN).

PAWN, and moment-independent methods in general, are particularly useful in case of highly skewed or multimodal output distributions. In such cases, variance is not an adequate proxy of uncertainty and variance-based methods (see below) can no longer be applied (Pianosi and Wagener, 2018). Another advantage of these methods is that they can be estimated from generic samples, that is, without requiring tailored sampling strategies (Pianosi et al., 2016; Pianosi and Wagener, 2018). Therefore, in this work we use the samples generated as described earlier to estimate the average and maximum of KS statistics calculated over 10 conditioning intervals (Noacco et al., 2019). To distinguish influential and uninfluential input factors, following Khorashadi Zadeh *et al.* (Khorashadi Zadeh et al., 2017) we artificially introduced a dummy input factor that does not appear in the model and, thus has no impact on the output. Therefore, the sensitivity index (maximum of KS statistic) corresponding to this factor defines the threshold for parameter screening (Khorashadi Zadeh et al., 2017; Noacco et al., 2019).

The second SA method used here is the Elementary Effect Test (Morris, 1991; Saltelli et al., 2007). In this case, the idea is to correlate model sensitivity with the effect of perturbing the input factors—one at a time—on the model output. An example of this approach is to estimate (*e.g.*, by finite differences) the partial derivatives with respect to different model parameters at their nominal values. In this form, the method is computationally very cheap but only provides local sensitivity information (Pianosi et al., 2016). A global extension of this technique is to compute perturbations from multiple points within the input variability space, followed by aggregating them in some type of statistic. The most popular method in this group uses the average of r finite differences (also known as Elementary Effects or EEs) as the sensitivity measure (μ_{EET}) (Pianosi et al., 2016; Saltelli et al., 2007). Here, a refined measure taking the absolute values of EEs is used to avoid cancelation due to sign differences (Campolongo and Saltelli, 1997). Therefore for the i-th input factor we have (Pianosi et al., 2016)

$$\mu_{EET,i} = \frac{1}{r}\sum_{j=1}^{r} \frac{f\left(x_1^j, \ldots x_i^j + \Delta_i^j, \ldots x_M^j\right) - f\left(x_1^j, \ldots x_i^j, \ldots x_M^j\right)}{\Delta_i^j} c_i \tag{44}$$

where Δ_i^j is the perturbation applied to the i-th parameter in the estimation of the j-th elementary effect and c_i is an appropriate scaling factor. Beside the average of EEs, standard deviations (σ_{EET}; the standard deviation of the terms inside the summation in Eq. (44)) provide information about the degree of interaction between the parameters and/or their level of nonlinearity (Pianosi et al., 2016; Saltelli et al., 2007).

The last method employed in the current study is the variance-based SA (VBSA) (Saltelli et al., 2010, 2007). This method assumes that the output variance is an indicator of its uncertainty and the contribution of each input factor to this variance is a measure of sensitivity. This technique handles nonlinear and non-monotonic functions/computational models, as well as those exhibiting interactions between their factors. Besides, it is able to capture the influence of each factor's full-range of variation (Nossent et al., 2011; Saltelli et al., 2010; Yang, 2011). Perhaps, the biggest drawback of this method is the large number of simulation runs it requires for convergence (Pianosi et al., 2016; Yang, 2011). A workaround for this limitation can be the application of meta-models (Al et al., 2019), although at the expense of approximating the original computational model (Morris and Moore, 2015). Aside from computational aspects, another limitation of VBSA is that, variance is not a meaningful gage for highly skewed or multimodal output distributions and hence, for such situations VBSA indices are not appropriate measures of sensitivity anymore (Pianosi et al., 2016; Pianosi and Wagener, 2018).

In VBSA, typically two types of indices are defined, first-order and total-order. First-order indices (also known as main effects, S_i) measure the direct contribution of individual input factors to the variance of output distribution. Equivalently, this can be thought of as the reduction in the output variance achievable by fixing inputs one at a time (Becker and Saltelli, 2015). Mathematically, we have (Pianosi et al., 2016)

$$S_i = \frac{V_{x_i}[E_{\boldsymbol{x}_{\sim i}}(y|x_i)]}{V(y)} = \frac{V(y) - E_{x_i}[V_{\boldsymbol{x}_{\sim i}}(y|x_i)]}{V(y)} \tag{45}$$

where operator $E_{\boldsymbol{x}_{\sim i}}$ calculates the expected value of the model output y over all possible values of input parameters except for x_i which is fixed (conditional mean; $\boldsymbol{x}_{\sim i}$ is the vector of all parameters but x_i), operator V_{x_i} represents the variance calculated over all possible values for x_i, and $V(y)$ is the unconditional output variance (Saltelli et al., 2010, 2007). In a model where output variability is only a result of main effects (lack of interactions between inputs), $\sum_{i=1}^{M} S_i = 1$ and the model is said to be additive. Nevertheless, in complex computational models, this is rarely the case and main effects do not sufficiently describe the output variability. Considering the high computational expense, particularly for larger M, that has to be incurred to estimate all the interaction effects, one may calculate total-order sensitivity indices, S_{Ti}, which embrace the main effect as well as all the interactions (of any order) involving the input factor x_i (Becker and Saltelli, 2015; Homma and Saltelli, 1996)

$$S_{Ti} = 1 - \frac{V_{\boldsymbol{x}_{\sim i}}[E_{x_i}(y|\boldsymbol{x}_{\sim i})]}{V(y)} = \frac{E_{\boldsymbol{x}_{\sim i}}[V_{x_i}(y|\boldsymbol{x}_{\sim i})]}{V(y)} \tag{46}$$

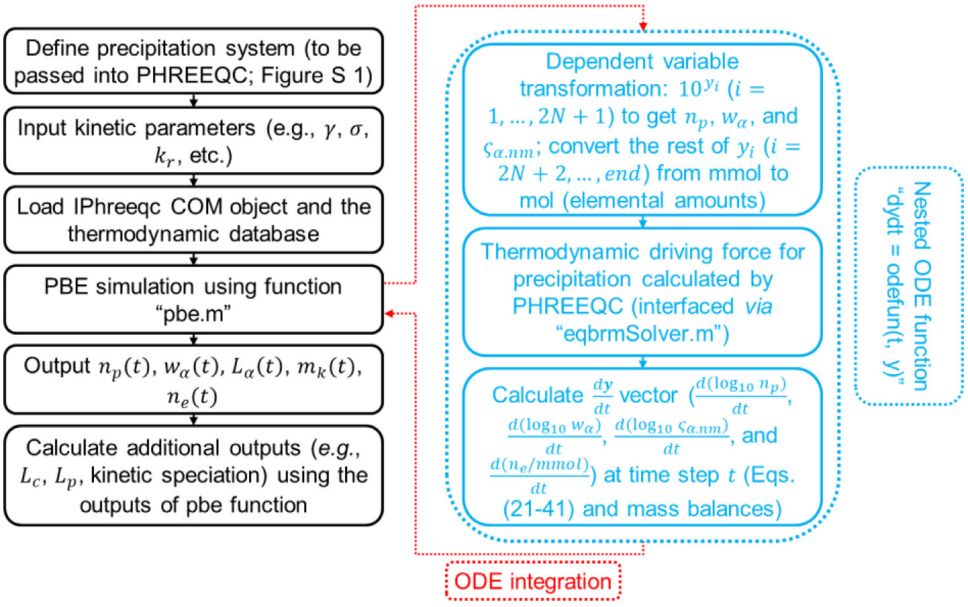

Fig. 2. The overall algorithm for the PBE simulation of a precipitation process. Black boxes comprise the main workflow backbone (*e.g.*, "demo.m" script in the SI), blue represents the content of the ODE function, and red refers to the integration of the ODE set using an appropriate solver (MATLAB's ode15s in this case (Shampine et al., 2003)). All the symbols are defined in the main text.(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Here, $V_{\boldsymbol{x}_{\sim i}}[E_{x_i}(y|\boldsymbol{x}_{\sim i})]$ is the variance due to all main effects and interactions not involving x_i and thus $\frac{V_{\boldsymbol{x}_{\sim i}}[E_{x_i}(y|\boldsymbol{x}_{\sim i})]}{V(y)}$ is the main effect of $\boldsymbol{x}_{\sim i}$ (Pianosi et al., 2016). Therefore, the remaining variance would be due to all terms that include x_i (Becker and Saltelli, 2015; Homma and Saltelli, 1996). Considering the nature of total effect indices, they are particularly suited for parameter screening as having zero total effect is a necessary and sufficient condition for an input parameter to be uninfluential (Pianosi et al., 2016).

3. Implementation

3.1. *Thermodynamic speciation* via *coupling to PHREEQC*

PHREEQC is a freely-available geochemical reaction solver capable of simulating a variety of processes including solid-liquid-gas equilibria, surface complexation, ion exchange, and much more (Parkhurst and Appelo, 2013). Aside from its carefully developed internal database, there are many comprehensive third-party databases including Cemdata18, specifically developed for cementitious systems (Lothenbach et al., 2019). With such a broad range of capabilities and extensive thermodynamic infrastructure, coupling PBE simulations with PHREEQC opens up new avenues in the practical and facile application of this powerful method to the understanding and design of particulate processes. Such coupling is facilitated by an already developed module called IPhreeqc, which enables interfacing with different scripting languages such as MATLAB and Python *via* Microsoft COM (component object model) (Charlton and Parkhurst, 2011; Wissmeier and Barry, 2011). Very recently, a number of articles have been published reporting the coupling of PHREEQC with PBE simulations (Dietemann et al., 2019; Rehage et al., 2019; Zhang et al., 2018). Nonetheless, to the best of our knowledge none of these publications provided the corresponding computer code and procedures for the implementation. Here, we developed a function (the file "eqbrmSolver.m" in the SI) that provides a general interface for PHREEQC speciation calculations in MATLAB. Briefly, information about the solution chemistry (*e.g.*, different compounds and their concentrations) and experimental conditions (such as temperature or gasses at constant partial pressure in equilibrium with solution) are provided to the interface, which in turn passes the data into PHREEQC solver *via* the COM object. These inputs are provided using keywords in a fashion similar to PHREEQC syntax (Figure S 1). This allows the simulation of precipitation in practically unlimited number of systems and scenarios without the necessity to rewrite the speciation code and/or its database. The information passed as the outputs of speciation calculation are the mass of water solvent, solution density, elemental concentrations, species concentrations and activities, pH, ionic strength, and saturation indices ($SI = \log_{10}(\frac{IAP}{K_{sp}})$ with IAP being the ionic activity product (Parkhurst and Appelo, 2013)) with respect to different solid phases.

In the current study, we employed the Cemdata18 database (Lothenbach et al., 2019) for all the aqueous reactions and the density and solubility product of the precipitate (C-S-H with Ca:Si=2) were taken from our previous paper (Andalibi et al., 2018).

3.2. *Overall simulation workflow*

Fig. 2 summarizes the steps involved during the simulation of a precipitation process (implemented in MATLAB R2019a). An example simulation is presented in the "demo.m" script provided in the SI. All the core PBE simulations are handled by "pbe.m" function while the output provided by this function suffices for the calculation of any other output of interest. For instance, knowing the moments allows one to calculate different crystallite and particle characteristics such as size and specific surface area. Similarly, knowing the amount of water solvent (input to "pbe.m") and temporal mole amounts of all the elements, one can back calculate the full speciation using the function "eqbrmSolver.m".

For a typical scenario that simulates 24 hours of precipitation with known model parameters (from our previous work (Andalibi et al., 2018)) on an ordinary HP laptop, with dual-core Intel® Core™ i5-4310 M CPU @ 2.70 GHz 2.70 GHz processor and 8.00 GB of RAM, the run time was 80 seconds. This is roughly half the time it took using our previous PBE code (which was using QMOM and bespoke speciation solver). It is worth noting that in the current version, in contrast to our previous code, the specia-

tion calculations are embedded within the ODE function. In other words, the calculation is performed at every time step selected by the ODE solver, which makes the total number of such calculations much greater in number than in our previous code. Therefore, the real speed up due to the application of DQMOM (in the modified form introduced here) in place of QMOM is much larger, consistent with previous reports (Haderlein et al., 2017; Silva et al., 2010) (see SI Section 1 for further details on the implementation of the workflow).

3.3. Problem setting and implementation of UA/SA

The UA/SA presented in this study was mostly implemented using the SAFE package, an open-source MATLAB toolbox that includes various functions for the generation of input samples, estimation and evaluation of sensitivity indices, and extensive visualization tools (Pianosi et al., 2015). Following our previous work the target of uncertainty/sensitivity analysis, that is, the PBE model for the precipitation of synthetic C-S-H, has five unknown parameters: interfacial tension (γ), cohesion energy (σ), growth rate coefficient (k_r), kinetic order of growth (g), and crystallite aspect ratio (a_r) (Andalibi et al., 2018). Below we will discuss the feasible uncertainty domain for each of these parameters.

The nominal value of interfacial tension from our previous regression to experimental data was estimated to be in the range of 0.05-0.06 J.m^{-2} (Andalibi et al., 2018). Preliminary numerical experiments proved that 0.065 is a suitable upper bound as larger values gave unrealistically hindered nucleation. Under the conditions of interest the surface charge density of C-S-H particles lies in the range of 0.4 C.m^{-2} (Plassard et al., 2005). From Figure 8.2 in Mersmann's handbook this could bring about 0.025 J.m^{-2} variation in the interfacial tension between C-S-H and liquid solutions. Therefore, a lower bound of 0.04 J.m^{-2} was assumed for γ. Again, numerical experiments showed that smaller γ would result in unrealistically early nucleation (at very low supersaturation ratios) or even spinodal decomposition (Kashchiev, 2003).

Theoretical considerations dictate that the relative cohesion energy between an already precipitated solid substrate and secondary nuclei is in the range 0-2 (Andalibi et al., 2018; Marchisio, 2009; Mutaftschiev, 2001; Testino et al., 2005). Preliminary tests, however, indicate that in our system of interest values larger than unity would give extremely small effective interfacial tensions for secondary nucleation, particularly when the interfacial tension is close to the lower bound set earlier. Additionally, a value of $\sigma/\gamma = 2$, which corresponds to coherent interfaces or epitaxial growth, hardly happens in precipitation from liquid solutions because of ions and solvent molecules adsorbed onto the surface of the substrate (Marchisio, 2009). This is compounded with the extremely defective nature of C-S-H crystallites hampering the formation of interfaces with matching lattices (Andalibi et al., 2018; Kumar et al., 2017). All things considered, a pragmatic upper bound of 1 was selected for σ/γ.

In our previous work, we fitted values in the order of 10^{-9} to the parameter k_r and 2 for parameter g (Andalibi et al., 2018). For the sake of UA/SA, values within one order of magnitude around 10^{-9} were considered. As suggested by Marino *et al.* (Marino et al., 2008), to sample the variability space more uniformly, considering the variation of k_r over two orders of magnitude, $\log_{10} k_r$ was preferred for the sampling. As for kinetic order of growth we sampled in the range 1-3 which is the typical variation range covering rough growth, dislocation-controlled mechanisms and surface nucleation regimes (Mersmann, 2001; Söhnel and Garside, 1992).

Finally, the ratio of crystallite edge length to its thickness regressed in our previous work was 0.5 (Andalibi et al., 2018). Therefore, an input range 1/3 to 1 was considered for the possible variability. For all the input factors uniform probability distribution functions were considered for the generation of the sample (Noacco et al., 2019).

Concerning the model outputs of interest, the size of crystallites and particles, their specific surface areas (SSA), C-S-H precipitation yield (conversion with respect to equilibrium composition), solution pH, and saturation index with respect to portlandite, all quantities after 12 hours of precipitation, are qualitatively examined (uncertainty analysis). Subsequently, quantitative assessment in the form of global sensitivity analysis was performed on three selected outputs: crystallite thickness ($\bar{L}_c$), particle edge length ($\bar{L}_p$), and specific surface area of particles (SSA_p). These are of special practical relevance and they can be compared to experimentally measured values (Andalibi et al., 2018; Jiao, 2016; Richardson, 2004; Wu et al., 2013; Zhang et al., 2015). We started with a base sample size of N = 2000 (corresponding to 14,000 input sample points) and extended the $\boldsymbol{X}$ matrix using the SAFE toolbox function "AAT_sampling_extend" to assess the convergence behavior of different sensitivity measures. Throughout this work, bootstrapping over 1000 resamples has been used to estimate the 95% confidence bounds on all the SA indices (Noacco et al., 2019; Pianosi et al., 2015).

As a last note regarding the implementation of UA/SA, the application of the input matrix generated as discussed in Section 2.3 is straightforward in the case of PAWN and VBSA methods. On the other hand, to apply EET using the Sobol' sample, the input and output were converted to the format required by EET functions. This was done using the function "fromVBSAtoEET" in the SAFE toolbox, which rearranges the matrices $\boldsymbol{X_B}$, $\boldsymbol{Y_B}$, $\boldsymbol{X_C}$, and$\boldsymbol{Y_C}$ so that they can be used to calculate EET indices from a radial design (Pianosi et al., 2015; Saltelli et al., 2007). With this approach, the number of sampling points (r) would be equal to the Sobol' base sample size (N).

4. Results and discussion

4.1. Uncertainty analysis with model parameters as input factors

In this section, we employ visual tools (scatter plots and histograms) to appraise the propagation of uncertainty from model parameters to different model outputs. Fig. 3(a) shows the histograms for average crystallite thickness and edge length ($\bar{L}_c$ and $a_r\bar{L}_c$, respectively) with probability normalization while Fig. 3(b) portrays the corresponding results for particle edge length ($\bar{L}_p$). From Fig. 3(a) we can see that the crystallite thickness and edge length are typically a few nm, consistent with previous reports for different C-S-H products (Andalibi et al., 2018; Gatty et al., 2001; Kumar et al., 2017; Richardson, 2004; Wu et al., 2013; Zhang et al., 2004). Particle edge length, instead, is typically 1-2 orders of magnitude larger and its distribution assumes a very long tail spanning up to a few μm (Fig. 3(b)) (Scrivener et al., 2019).

Concerning the specific surface area (*viz.*, surface area per unit mass of precipitate), the hierarchical structure of the solid C-S-H gives rise to two types of surfaces. One is the overall area of the external surfaces of crystallites (*i.e.*, neglecting the internal crystallite structure; $SSA_{Crystallite}$), and the other only considers the external surface of particles neglecting the surfaces embedded within the bulk of the particles ($SSA_{Particle}$). Therefore, by definition $SSA_{Particle} \leq SSA_{Crystallite}$ with the equality happening in the absence of secondary nucleation and aggregation (in other words, when every crystallite is a particle by itself). $SSA_{Crystallite}$ is calculated from the zeroth moment of crystallite size distribution (which is directly available from PBE simulations) and $\bar{L}_c$ while we can estimate $SSA_{Particle}$ from n_p and $\bar{L}_p$ (see the Electronic Supplementary Information in Ref. (Andalibi et al., 2018) for the estimation of $\bar{L}_p$ using geometrical considerations). Fig. 3(c) shows the scatter plots for $SSA_{Particle}$ *vs.* $SSA_{Crystallite}$ and Fig. 3(d) pro-

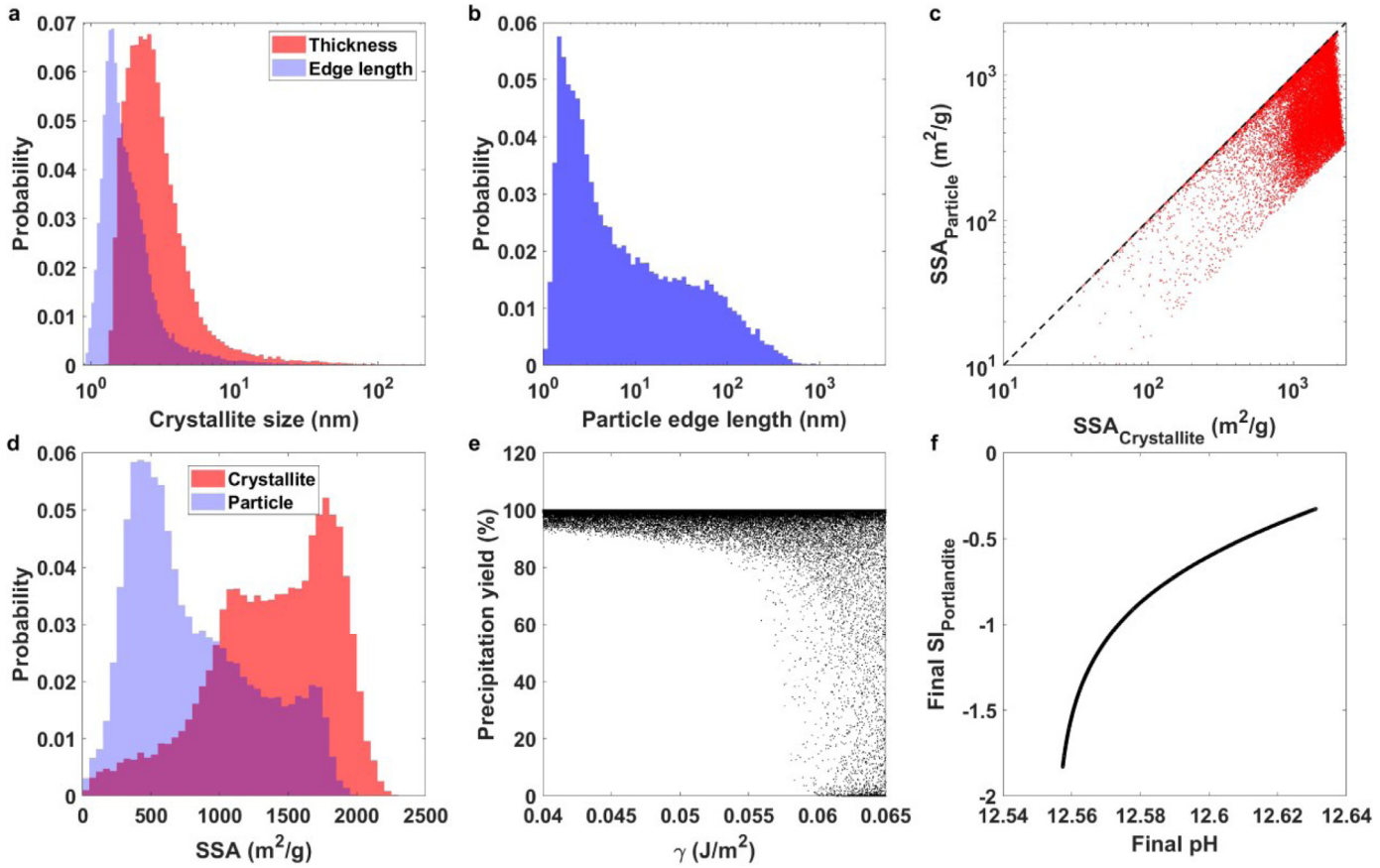

Fig. 3. Results of uncertainty analysis with model parameters as input factors. (a) Probability-normalized histograms of crystallite thickness and edge length; (b) probability-normalized histogram of particle edge length; (c) specific surface area (SSA) of particles *vs.* that of crystallites (the black dashed line is the identity line); (d) probability-normalized histograms of crystallite and particle SSA; (e) scatter plot of precipitation yield *vs.* interfacial tension; (f) dependence of final saturation index with respect to portlandite on final solution pH.

vides the corresponding histograms. In Fig. 3(c), we can easily see that the condition $SSA_{Particle} \leq SSA_{Crystallite}$ holds (no data above the identity line) confirming the correct performance of our PBE simulations. Another important constraint is that $SSA_{Crystallite}$ has to always be smaller than or equal to the total surface area of C-S-H building units. Assuming all the calcium in solution is converted to C-S-H and knowing the surface area of a single building unit (1.1297×10^{-18} m^2; estimated from C-S-H molar volume (Andalibi et al., 2018) and considering a cubic shape, which has the highest surface area among cuboids with similar volume), the theoretical upper bound for $SSA_{Crystallite}$ would be 6014 m^2/g. As we can see in Fig. 3(c) we are well within this constraint. From Fig. 3(d), with the presumed input variability space, the typical values for $SSA_{Crystallite}$ and $SSA_{Particle}$ are a few tens up to around 2000 m^2/g.

Another output of PBE simulation is the precipitation yield calculated as

$$Precipitation\ yield\ (\%) = \frac{V_0 \times c_{Ca,0} - V_{end} \times c_{Ca,end}}{V_{end} \times c_{Ca,SLE}} \times 100 \quad (47)$$

where V_0 and V_{end} refer to the initial and final volume of the reaction medium, respectively, $c_{Ca,\,0}$ and $c_{Ca,\,end}$ are the corresponding Ca concentrations, and $c_{Ca,\,SLE}$ is the Ca concentration at the end if precipitation reaches equilibrium. From Fig. 3(e), there is a strong relation between γ and precipitation yield, wherein the yield can drop significantly at higher γ. This is anticipated because larger γ hinders the onset of primary nucleation (Andalibi et al., 2018). It is worth noting that much weaker correlations, if at all, are observed with other input parameters (Figure S 2).

Another simulation output is the saturation index (*SI*) with respect to portlandite, a solid phase that competes with C-S-H for precursor ions during the precipitation (Andalibi et al., 2018; Kumar et al., 2017; Richardson, 2004). In Fig. 3(f) we have plotted this quantity *vs.* the solution pH at the end of the precipitation (after 12 h). From this plot, an unambiguous correlation is visible, where *SI* increases monotonically with pH. In our previous work, we showed that under the examined operating conditions the system was always undersaturated with respect to portlandite, consistent with experimental observations (Andalibi et al., 2018; Kumar et al., 2017). Nevertheless, according to Fig. 3(f) at higher pH values portlandite may precipitate along with C-S-H. Indeed, our simulations with nominal model parameters (Andalibi et al., 2018) but at a higher inflow NaOH concentration (*e.g.*, 4 times the value reported in (Andalibi et al., 2018) giving a final pH of 13.2; Figure S 14(a)) showed that the system does become supersaturated with respect to portlandite in line with certain experiments (refer to SI Section 2 for further discussion) (Kumar, 2017; Richardson, 2004).

Now let us examine the mapping of input uncertainty to three selected outputs $\bar{L}_c$, $\bar{L}_p$, and SSA_p. From Fig. 4, we see that variability in parameter g has almost no effect on any of the outputs (as implied by the uniformity of the scattered points and the lack of pattern (Pianosi et al., 2016)). Concerning other input factors, however, the relative degree of uncertainty propagation depends on the output. From Fig. 4 (a-c, e), parameters γ, σ/γ, $\log_{10}k_r$, and a_r are all influential with respect to $\bar{L}_c$ as the output, with σ/γ having less impact compared to others. Consulting Fig. 4(f-h, j), uncertainty in σ/γ clearly has the highest effect on the variability of

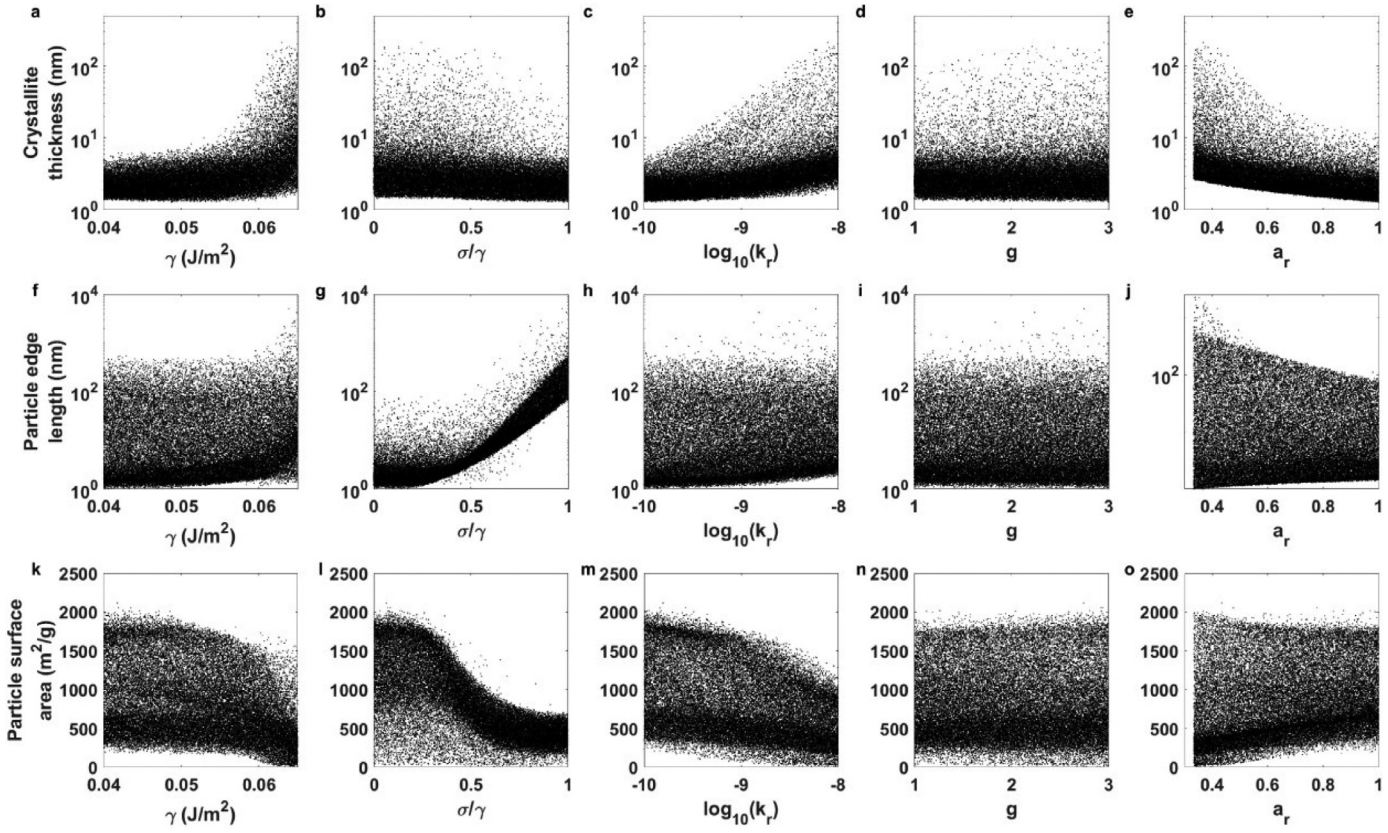

Fig. 4. Scatter plots of crystallite thickness (a-e), particle edge length (f-j), and particle surface area (k-o) *vs.* different input model parameters (base sample size 20,000).

$\bar{L}_p$ (strong pattern formed by the scattered data points) while γ, $\log_{10}k_r$, and a_r have much less of an impact. A similar argument applies to SSA_p (although to a lesser extent) with the output being much more sensitive to σ/γ (Fig. 4(k-m, o)). In the next section, we will present the quantitative assessment of sensitivity with respect to different model parameters.

4.2. Sensitivity analysis with model parameters as input factors

Fig. 5 summarizes the PAWN sensitivity indices calculated with $\bar{L}_c$, $\bar{L}_p$, and SSA_p as the outputs (obtained from a sample of size 140,000; see SI Section 2 for the convergence analysis of the indices). The top row is the mean KS statistic across the ten conditioning intervals while the bottom row present the maximum of KS. The latter can also be used to identify the influential and uninfluential input factors by comparing the sensitivity indices to that of a dummy variable (which has no effect on the model outcome) (Khorashadi Zadeh et al., 2017; Noacco et al., 2019). From both indices, we can clearly verify the minimal effect of g on the studied model outputs consistent with the conclusions made from the scatter plots as discussed earlier (see previous section and Fig. 4). Indeed, taking the maximum KS statistic as the sensitivity measure, it is barely higher than the value estimated for the dummy input (Fig. 5(d-f)). With $\bar{L}_c$ as the model output, the rest of the parameters are all influential with σ/γ being slightly above the dummy variable, and γ, $\log_{10}k_r$, and a_r exhibiting quite similar higher influences (Fig. 5(a,d)). With $\bar{L}_p$ and to a lesser extent SSA_p, variability in σ/γ has the highest impact on the output uncertainty. Except for the uninfluential factor g, the rest of model parameters have similar impact over these two outputs (Fig. 5(b,c,e,f)).

Fig. 6 summarizes the SA results using the EET method (see SI Section 2 for the convergence analysis of the indices). We have plotted the average of absolute values for the EEs (μ_{EET}) against the standard deviations of EEs normalized by their respective averages (σ_{EET}/μ_{EET}). We will refer to the latter as the coefficient of variation (C.V.) although strictly speaking C.V. is obtained using the average of signed EEs (and not the absolute values) (Menberg et al., 2016; Morris, 1991; Saltelli et al., 2007). Again, consistent with the results obtained from the scatter plots (Fig. 4) and PAWN (Fig. 5), we observe very little effect from g on different outputs (Fig. 6; μ_{EET} for g is invariably much smaller than that of the most influential factor). With $\bar{L}_c$ as the model output, σ/γ is identified as the second least influential parameter while close values are predicted for the other inputs (similar to PAWN; Fig. 6(a) and Fig. 5 (a,d)). Along the same lines, with $\bar{L}_p$ and to a smaller degree SSA_p, σ/γ has the highest impact on the output with sensitivity indices being 48 and 12 times that of g, respectively (Fig. 6(b,c)).

Another observation from our EET analysis is related to the level of nonlinearity in model parameters and/or the degree of interaction between them. Following a method proposed by Garcia Sanchez *et al.* (Garcia Sanchez et al., 2014), C.V. values can be classified in four regions [0, 0.1], [0.1, 0.5], [0.5, 1], and $>$ 1. These regions correspond to almost linear, monotonic, almost monotonic, and markedly non-monotonic and/or interacting parameters, respectively. From Fig. 6(a,b) (and Figure S 5(a,b)), we see that with $\bar{L}_c$ and $\bar{L}_p$ as outputs all the parameters exhibit a high degree of nonmonotonicity and/or interaction. With SSA_p, however, although parameters γ and g exhibit highly nonlinear and/or interactive behavior, the output is almost monotonic with respect to parameters σ/γ, $\log_{10}k_r$, and a_r (Fig. 6(c) and Figure S 5(c)).

Fig. 7 summarizes the SA results based on the variance-based method of Sobol'. Here, considering the sluggish convergence of the sensitivity measures (in particular, the total effects for $\bar{L}_c$ and $\bar{L}_p$, Figure S 6(d,e); see SI Section 2 for the convergence analysis of the indices), we have further explored the possible application of

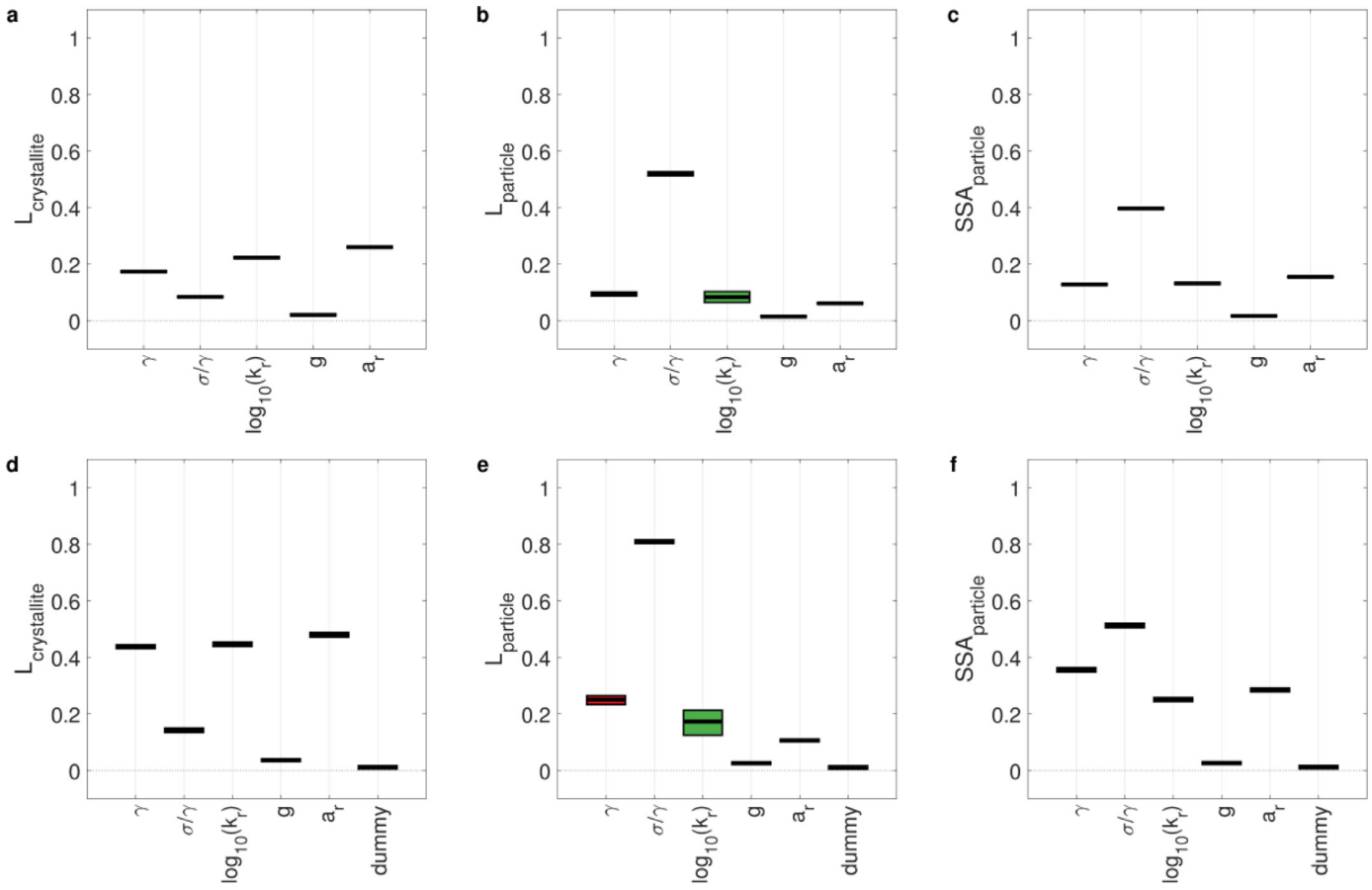

Fig. 5. PAWN sensitivity indices in the form of mean (a-c) and maximum (d-f) of the KS statistic, with 95% confidence intervals obtained from bootstrapping, for crystallite thickness (a,d), particle edge length (b,e), and particle surface area (c,f) as the outputs (sample size of 140,000).

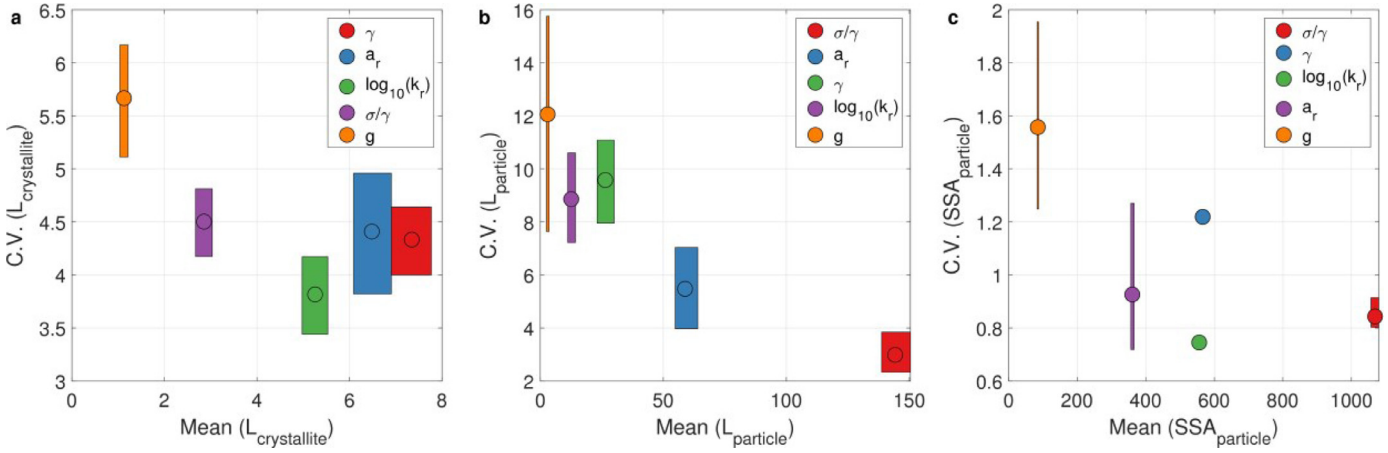

Fig. 6. Sensitivity indices obtained using EET with crystallite thickness (a), particle edge length (b), and particle surface area (c) as the outputs (sample size of 120,000). The results are presented as the mean of Elementary Effects plotted against their coefficients of variation (all the 95% confidence intervals are estimated by bootstrapping).

output transformation. Therefore, rank transformation (*i.e.,* replacing the numerical value of the output with its ranking within the sample) and $\log_{10}$ transformation, which have been very popular in the literature (Homma and Saltelli, 1996; Saltelli and Sobol', 1995), are included in Fig. 7(d-f) and Fig. 7 (g-i), respectively.

From Fig. 7(a) we see that with $\bar{L}_c$ as the output the main effects are generally small, with σ/γ and g assuming zero and $\log_{10}k_r$ being barely above zero. On the other hand, the total effects account for the output variability signifying the nonadditive nature of the input factors (Saltelli et al., 2010, 2007; Saltelli and Sobol', 1995). This is consistent with the C.V. values observed from EET (Fig. 6(a) and Figure S 5(a)) all being larger than 1. Nevertheless, even at such a high sample size (140,000) the confidence intervals are wide and there is overlap between the total effect of σ/γ with $\log_{10}k_r$, and $\log_{10}k_r$ with g (Fig. 7(a)). Therefore, we attempted a second SA fixing $g = 2$ (which we already know is practically uninfluential) and going up to a sample size of 288,000 (base sample = 48,000). Doing that, the confidence intervals of σ/γ fall well below the other three influential parameters (γ, $\log_{10}k_r$, and a_r;

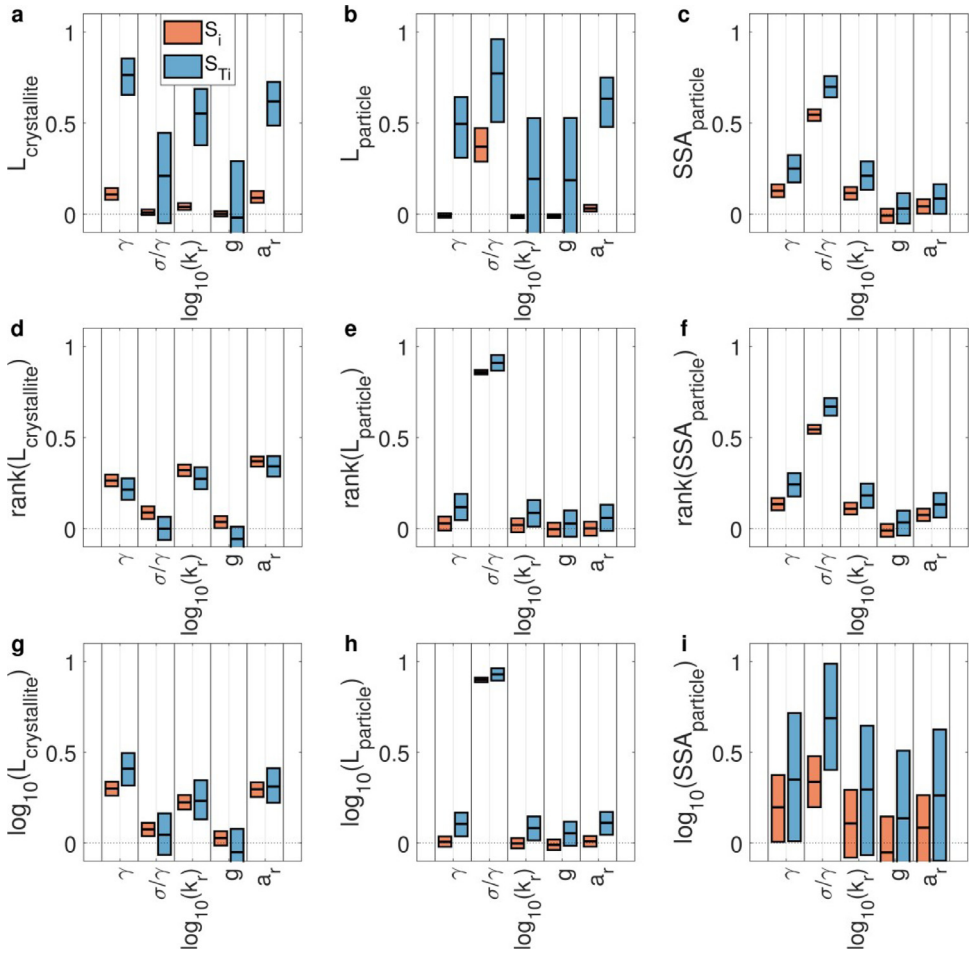

Fig. 7. Variance-based sensitivity indices for $\bar{L}_c$, $\bar{L}_p$, and SSA_p (a-c) and their rank (d-f) and $\log_{10}$ transformations (g-i) as the outputs, with a sample size 140,000 (base sample size $N = 20{,}000$; S_i and S_{Ti} are the main and total effects, respectively, and the 95% confidence intervals are estimated by bootstrapping).

Figure S 7(a)). This outcome is consistent with our previous results from PAWN and EET (Fig. 5(a,d), Fig. 6(a), and Figure S 5(a)).

Consulting Fig. 7(b), we note the same problem with $\bar{L}_p$ as with $\bar{L}_c$. This time, however, even at an input sample size of 288,000 the confidence bounds overlap significantly (Figure S 7(b)). This observation can readily be explained by looking at the probability histograms of the output, extending over three orders of magnitude (Fig. 3(b)). In other words, the complication arises from the highly skewed distribution of $\bar{L}_p$ with a skewness of 30 (compare with 10.7 for $\bar{L}_c$; Fig. 3(a)) (Pianosi et al., 2016; Pianosi and Wagener, 2018). One remedy to this problem is the application of rank (Fig. 7(e) and Figure S 7(e)) or $\log_{10}$ (Fig. 7(h) and Figure S 7(h)) transformation both making the confidence intervals very narrow. Nonetheless, we have to note that sensitivity analysis on transformations of an output cannot always be converted back the non-transformed output (Homma and Saltelli, 1996; Iman and Hora, 1990; Saltelli and Sobol', 1995). In fact, rank transformation frequently increases the relative weight of the main effects at the expense of interaction terms. Consequently, the effect of those factors influencing the output mostly by way of interaction with other parameters may be underestimated in a rank-based analysis (Saltelli and Sobol', 1995). In a similar way, in our particular example (Fig. 7(d-i) and Figure S 7(d-i)) for all the transformed outputs (both rank and $\log_{10}$) the interactions are dampened rendering the total effects almost equal to the main effects. Especially, with parameters γ, $\log_{10}k_r$, and a_r affecting $\bar{L}_p$ mainly *via* interactions (see their zero main effects in Fig. 7(b) and Figure S 7(b)), upon transformation they apparently become much less influential (that is, they adopt smaller total effects; Fig. 7(e,h) and Figure S 7(e,h)). The same complication can be traced in Fig. 7(d,g) (and Figure S 7(d,g)) because parameters mainly affect $\bar{L}_c$ by way of interactions.

Another interesting feature can be seen in variance-based indices with $\log_{10}SSA_p$ as the output (Fig. 7(i) and Figure S 7(i)). Here, in contrast to the case with untransformed output variable (Fig. 7(c) and Figure S 7(c)), the indices have very broad confidence intervals that significantly overlap and make any conclusive deduction impossible (Pianosi et al., 2016). A closer examination of probability distributions of SSA_p and its $\log_{10}$ transformation reveals that while the former is almost unimodal (Fig. 3(d)) the latter is highly multimodal (Figure S 8). Quantitatively, the Hartigan's dip test of unimodality (Freeman and Dale, 2013; Hartigan, 1985) gives p-values of 0.11 (insignificant multimodality) and 0 (significant multimodality) for the untransformed and $\log_{10}$ transformed variable, respectively. Therefore, aside from the complication in converting SA results from transformed outputs back to the original ones (Homma and Saltelli, 1996; Iman and Hora, 1990; Saltelli and Sobol', 1995), $\log_{10}$ transformation may render the output distribution multimodal limiting the applicability of VBSA to such scenarios (Pianosi et al., 2016; Pianosi and Wagener, 2015). It is worth noting that the latter problem should not happen with rank transformation, as the converted distribution is always uniform and thus unimodal (Hazewinkel, 2002).

To summarize the results presented in this section, we showed that the different outputs might have very different degrees of sensitivities with respect to various uncertain input model parameters. For example, while the average particle size is extremely sensitive

to the relative cohesion energy between the crystallites, the mean crystallite size is only weakly dependent on this parameter. This is an invaluable piece of information when calibrating the model using experimental data as it allows one to rationally decide on the methods used to collect such data. Furthermore, one has to be cautious about the interpretation of sensitivity results and always crosscheck the outcome of several independent methods to make sure the results are reasonable and conclusive (Noacco et al., 2019; Pianosi et al., 2015).

4.3. UA/SA with selected model parameters and experimental conditions as input factors

Now that we have examined the model behavior in detail, we can turn our attention to the ultimate goal of the current study, that is, the theory-driven design of nanoparticle synthesis processes. Ideally, a precipitation model should be able to explain the process as a function of experimental conditions alone. In other words, all the parameters in the theoretical framework have to be defined as a function of operating conditions such as temperature, concentrations of reagents, ionic strength, *etc*. This is not an easy task because the development of such models requires extensive and sometimes independent sets of experimental data to identify the mechanistic steps involved and calibrate the corresponding theoretical constructs. For instance, Schroeder *et al.* attempted to calibrate such a framework for the formation and polymorphic transformation of calcium carbonate (Schroeder et al., 2014). Although they accounted for different physicochemical aspects and correlated different parameters with the environmental conditions inside the reactor, limited success was achieved in reproducing the experimental data given the extremely complicated nature of the precipitation process. In the specific case of C-S-H precipitation, additional complications arise due to the nature of the precipitate usually forming a solid solution whose composition depends on the environmental conditions and may evolve as a function of time (Andalibi et al., 2018; Prieto et al., 2016; Thomas et al., 2011). Therefore, with the experimental kinetic data being scarce for synthetic C-S-H (Andalibi et al., 2018; Kumar et al., 2017), it is only possible to semi-quantitatively design the product properties as we will present in this section.

In its novel environmental (Shao et al., 2018; Zhang et al., 2015; Zhao et al., 2014), biomedical (Jiao, 2016; Wu et al., 2013; Zhu and Sham, 2014), and catalysis applications (Sheng et al., 2019; Xia et al., 2019), the accessible specific surface area of C-S-H product (*i.e., SSA_p*) is one of the most important properties of interest. Therefore, in this section we mainly focus on this characteristic, while information about crystallite and particle sizes are addressed for benchmarking against the literature data.

From the discussion in the previous sections, we found that among the model parameters g is significantly less influential and its impact is barely above the dummy variable. Additionally, our previous studies showed that the aspect ratio of C-S-H crystallites is 0.5 irrespective of mixing flow rate (Andalibi et al., 2018). Interestingly, the same aspect ratio was found for lower Ca:Si solids based on atomistic simulations (Dolado et al., 2011). Therefore, in this section we fix these parameters to nominal values $g = 2$ and $a_r = 0.5$. The rest of the model parameters were also constrained within reasonable neighborhoods of the regressed parameters taken from our previous study ($0.05 \leq \gamma \leq 0.06$; $0.5 \leq \sigma/\gamma \leq 1$; $-9 \leq \log_{10} k_r \leq \log_{10}(5 \times 10^{-9})$) (Andalibi et al., 2018). In particular, the considered ranges are similar to the confidence intervals obtained earlier (Andalibi et al., 2018), slightly expanded to compensate for the limited window in which the experimental data were collected. The judgement was based on preliminary UA where different expanded ranges were selected and the outputs were evaluated until they fell in the literature and/or physically accessible window. For instance, interfacial tensions larger than 0.06 give unrealistically hindered nucleation while values smaller than 0.05 lead to the limit of monomer-sized critical nucleus size (spinodal decomposition). This procedure allows us to account for potential variations in these parameters when the experimental conditions deviate from those under which the regression data were collected. This is critical especially for parameters highly sensitive to the environmental conditions (such as interfacial tension and cohesion energy). Fixing suach parameters to nominal values does not allow for considering the possibility of parameter variations arising from the changes in the experimental conditions and leads to less reliable conclusions.

Having the uncertainty window for the model parameters in place, we investigate the effect of four experimental variables on the outcome of precipitation, in a moderately wide neighborhood of the nominal experimental conditions we previously used for regression (Andalibi et al., 2018). This includes the precipitation temperature (10-50°C; see SI Section 1 for details on the temperature-dependence of C-S-H dissolution constant), the initial concentration of $Ca(NO3)_2$ inside the reactor (0.01-0.1 mol/kg water; the ratio of Na_2SiO_3 to $Ca(NO3)_2$ is kept constant at the original value of 0.5), the concentration of NaOH in the inflow stream (0.1-0.4 mol/kg water), and the rate of inflow stream (Q; 0.5-10 mL/min).

Fig. 8 summarizes the distributions of crystallite and particle sizes and specific surface areas. With the model parameters varying in a practically more accessible neighborhood of the values regressed to experimental data, we anticipate the distributions presented in this figure match the variations in real systems more closely. As before, the crystallite dimensions are in the range of a few nm, in agreement with experimental observations for different C-S-H products (Andalibi et al., 2018; Gatty et al., 2001; Kumar et al., 2017; Richardson, 2004; Wu et al., 2013; Zhang et al., 2004) (Fig. 8(a)). It is worth noting that the extreme values of C-S-H crystallite size can be obtained experimentally by carefully adjusting the synthesis conditions. For instance, Mehrali *et al.* obtained C-S-H crystallites as large as 13 and 25 nm in the presence of sodium dodecyl sulfate (Mehrali et al., 2014).

From Fig. 8(b) we see that the particle edge length can assume values up to a few hundreds of nm. It is worth mentioning that there are fewer reports on the size of synthetic C-S-H particles, which may coincide with the correlation/cutoff length commonly measured in SANS and SAXS and is of the same order as what we observe here (Chiang et al., 2012; Ioannidou et al., 2016; Jennings et al., 2007). TEM images also give values within the range of a few tens to a few hundreds of nm for the width of C-S-H particle (Andalibi et al., 2018; Kumar et al., 2017; Mehrali et al., 2014).

From Fig. 8(c) the condition $SSA_{Particle} \leq SSA_{Crystallite}$ can be verified. In fact, with the lower bound for σ/γ being 0.5, there is always significant contribution from secondary nucleation rendering the $SSA_{Particle}$ smaller than $SSA_{Crystallite}$. From Fig. 8(d), we note that the corresponding distributions for $SSA_{Particle}$ and $SSA_{Crystallite}$ give values in the order of 400 and 1600 m^2/g.

Fig. 9(a-g) presents the scatter plots for SSA_p as a function of different individual input factors. Among the model parameters, σ/γ appears to be the most influential factor, in line with our results in the previous sections (note the pattern formation in Fig. 9(b)). Among the experimental conditions, the addition flow rate of silicate solution seems to dominate the output variability, albeit with a lower impact when compared to σ/γ (Fig. 9(g)). Fig. 9(h) shows the colored scatter plot for these two factors with marker colors proportional to the output value. The emergence of color patterns in such a plot is a simple and intuitive tool to assess the degree of interaction between pairs of input factors

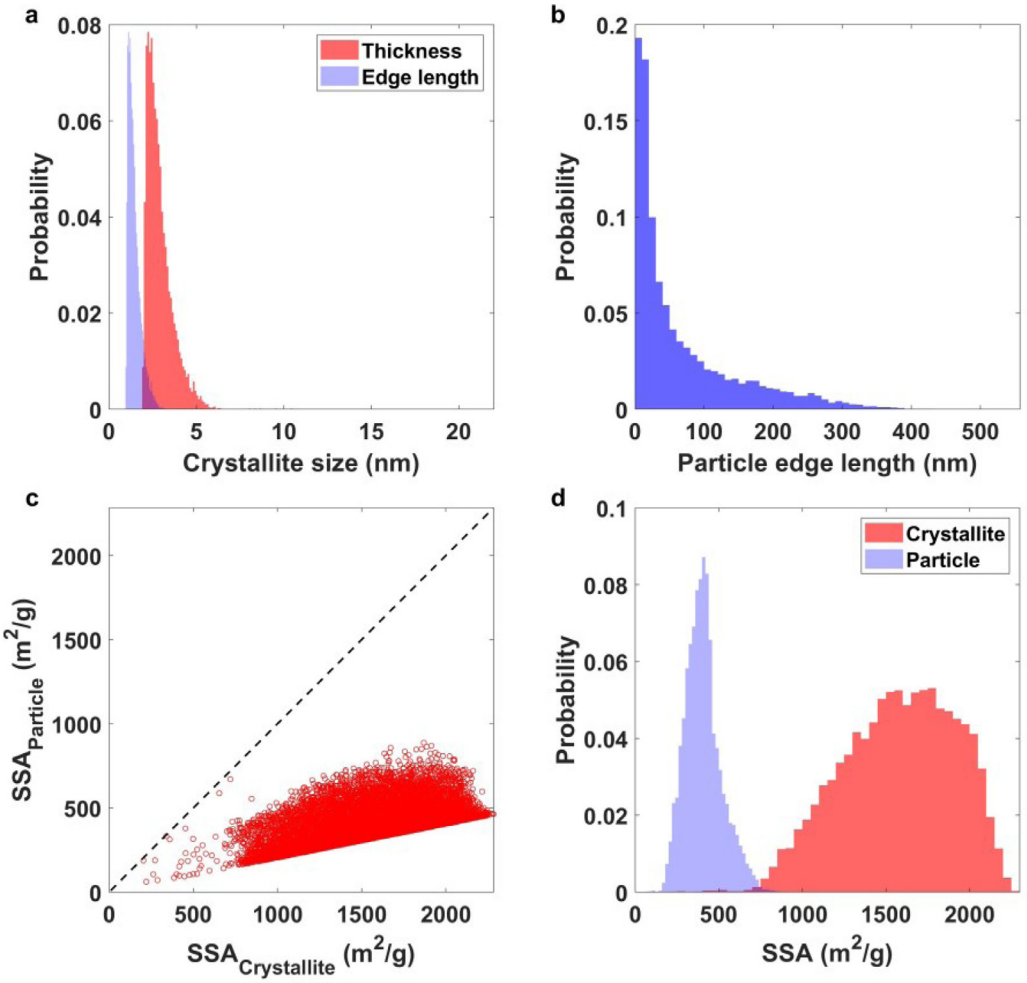

Fig. 8. Results of UA with selected model parameters and experimental conditions as the input factors. (a) Probability-normalized histograms of crystallite thickness and edge length; (b) Probability-normalized histogram of particle edge length; (c) Specific surface area (SSA) of particles *vs.* crystallites (the black dashed line is the identity line); (d) Probability-normalized histograms of crystallite and particle SSA.

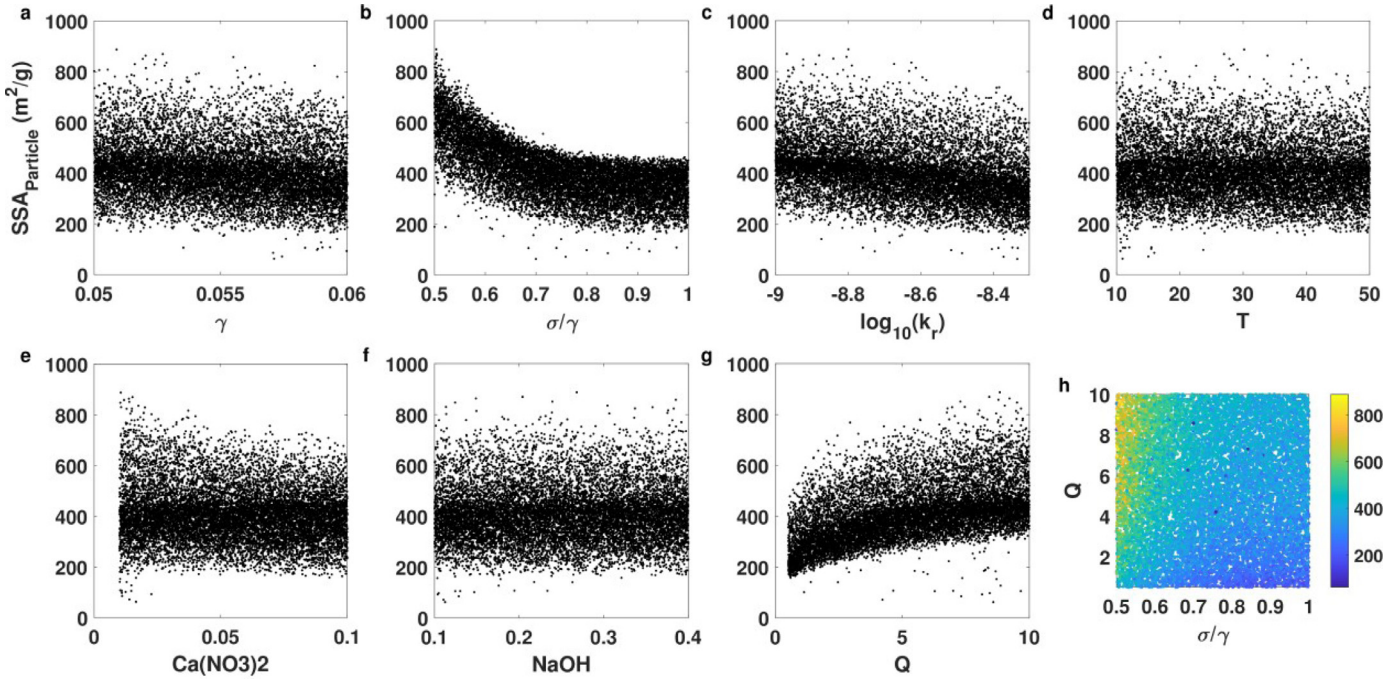

Fig. 9. Scatter plots of particle surface area *vs.* different factors in UA with selected model parameters and experimental conditions as the uncertain inputs. (a-g) Output plotted against individual inputs; (h) Scatter plot of Q *vs.* σ/γ with marker color proportional to the value of output.

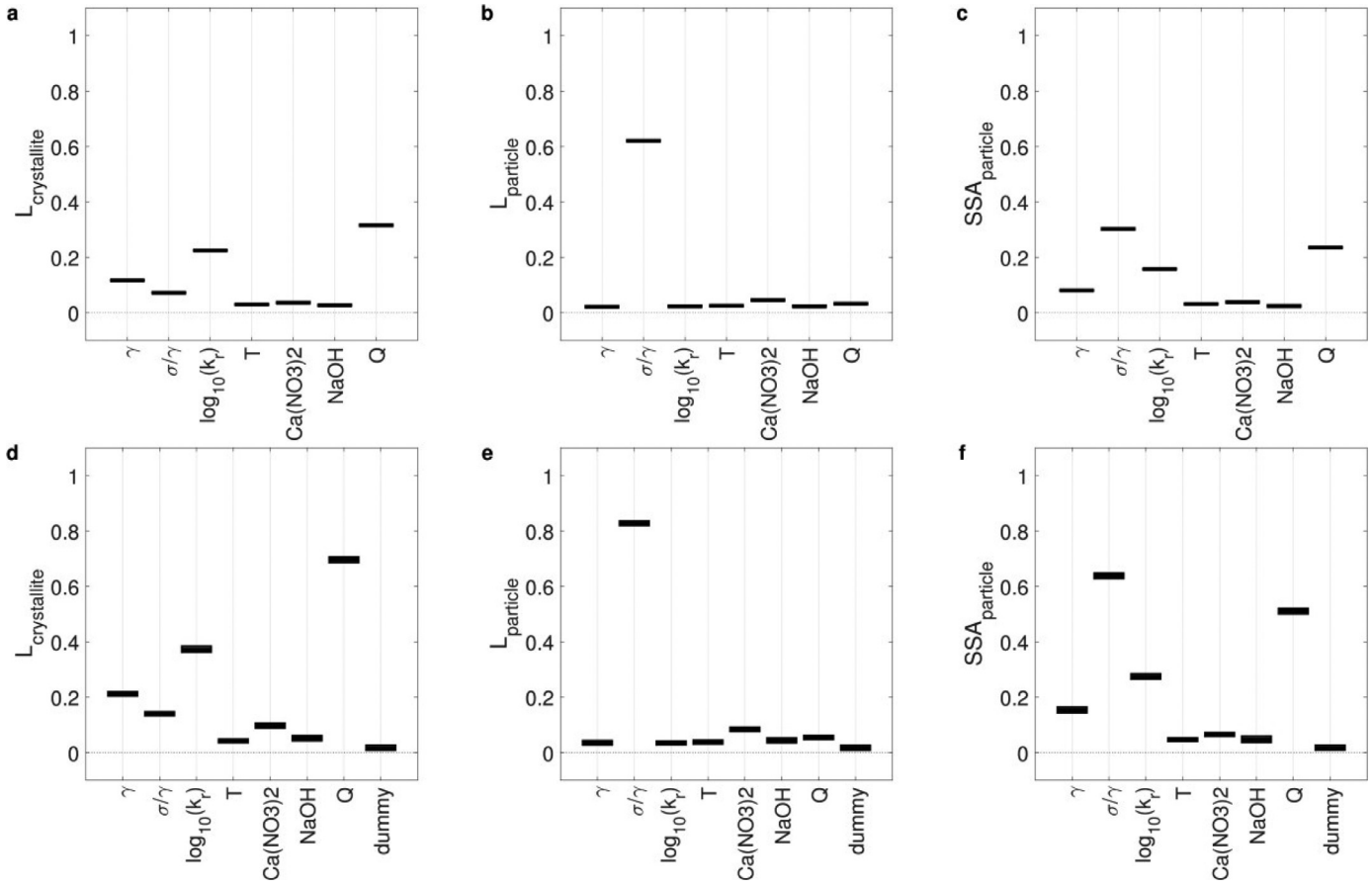

Fig. 10. Results of SA with selected model parameters and experimental conditions as the input factors. PAWN sensitivity indices in the form of mean (a-c) and maximum (d-f) of KS statistic, with 95% confidence intervals obtained from bootstrapping, for crystallite thickness (a,d), particle edge length (b,e), and particle surface area (c,f) as the outputs (sample size of 54,000).

(Pianosi et al., 2016, 2015). From Fig. 9(h) a weak pattern can be discerned (upper left region) where simultaneous occurrence of high Q and low σ/γ gives rise to exceptionally higher surface areas (see also Figure S 11(c) presenting the corresponding EET results where C.V. for the three most influential parameters are all below unity indicating weak interactions among the parameters).

For a quantitative assessment of variability propagation to model outputs, PAWN sensitivity indices were estimated for different model parameters and experimental conditions as input factors. Fig. 10 summarizes the results with $\bar{L}_c$, $\bar{L}_p$, SSA_p as the outputs (consult Figure S 10 for the convergence of PAWN indices; similar conclusions can also be obtained using EET as depicted in Figure S 11). For $\bar{L}_c$, $\log_{10}k_r$ and Q are the most influential factors (Fig. 10(a,d)). All of the experimental variables have low influences on $\bar{L}_c$ barely above the dummy index (Fig. 10(d)). The larger impact of flow rate can be understood from the fact that at higher addition rates, the supersaturation build up is larger which in turn induces more contribution from nucleation events to the overall precipitate. Put differently, higher nucleation rates give rise to larger number of crystallites among which the remaining precursor is divided giving rise to smaller crystallites (the same trend was also detected in our previous work; see Table 1 in Ref. (Andalibi et al., 2018)).

Looking at Fig. 10(b,e), $\bar{L}_p$ is most sensitive to σ/γ with the rest of input factors having minimal effects only marginally above the dummy index. Physically, this means that the relative rates of primary and secondary nucleation events determine the final particle size.

With SSA_p as the SA target, again σ/γ is the most influential parameter (Fig. 10(c,f)) compatible with our scatter plots (Fig. 9(b)). Besides, among the experimental conditions, we can distinguish a comparable dependence on Q (Fig. 10(c,f); similar inference as in scatter plot Fig. 9(g)). Conversely, the SSA_p of the product is much less sensitive (in a global sense) with respect to the other experimental variables. This is a favorable outcome as it allows for optimizing this key property by tuning the synthesis conditions. Therefore, higher surface areas can generally be obtained by increasing Q irrespective of the value other uncertain input factors assume. From a physical point of view, this can again be explained in the light of supersaturation buildup brought about by higher Q (providing the limiting reactants—Na_2SiO_3 and NaOH—faster), which favors primary nucleation over secondary nucleation (and nucleation, in general, over growth) (Dirksen and Ring, 1991). Consequently, higher particle number concentrations are obtained making the overall surface area larger.

Previously, Wu *et al.* synthesized C-S-H (of lower Ca:Si ratios) with specific surface areas ranging between 100 and 500 m^2/g, obtained by varying the synthesis conditions (Wu et al., 2013). Our results show that there is room for further improvement by increasing the addition flow rate of silicate solution although one has to optimize the design of the synthesis reactor for maximal mixing (Jiang and Braatz, 2019). This can be reinforced by the synergistic effect of lowering the cohesion energy (that is, lowering σ/γ; Fig. 9(b)) which can be induced either by increasing the relative concentration of monovalent ions (*e.g.*, adding a sodium salt to the mixture) or by working at lower pH values (Jönsson et al., 2005; Plassard et al., 2005). Of course, this conclusion only applies to

the set of synthesis conditions investigated here and carefully calibrated computational models are needed to cover scenarios that are more diverse. This includes, for instance, different reactant ratios (which typically induces variations in the Ca:Si ratio of the precipitate (Sheng et al., 2019)), alternative addition orders, different pH levels, and the inclusion of other reagents/surfactants (beside those used in the original experiments (Andalibi et al., 2018)). Still in the absence of extensive experimental data, the theory-driven approach presented in this section proves to be a promising tool to gain insight about the outcome of the process as a function of different operation conditions.

5. Conclusions

In summary, we presented a faster, more user-friendly, and more robust version of our previous PBE modeling framework (put forward in Ref. (Andalibi et al., 2018)) describing the process of precipitation from liquid solutions. This was achieved by replacing our speciation function with an interface to PHREEQC, providing access to the large databases already implemented in this popular software. Thanks to this modification, the adaptation to new precipitation scenarios is made much more straightforward and can be performed using keywords similar to the conventions used in PHREEQC, eliminating the need to prepare the database file for every new system. Another modification was the application of DQMOM, which offers several advantages in terms of speed, robustness, and adaptability over our previously implemented method (QMOM). Subtle technicalities in the implementation of DQMOM to obtain a reliable and quickly converging solution method were explained to allow replication/extension of the current work by other researchers. We also provide fully commented MATLAB codes implementing the PBE simulation workflow in the accompanying Supporting Information.

Upon developing an improved computational framework, three different global uncertainty/sensitivity analysis (UA/SA) methods were applied to understand the behavior of the model in response to uncertainty in various model parameters. For several simulation outputs, either we demonstrated the consistency of the results from different SA measures or explained the reason behind the inadequacy of the applied method. In the latter case, for instance, we presented particle edge length as an output whose highly skewed distribution limited the applicability of variance-based indices. Having a comprehensive picture of the uncertainty propagation through the model, we simplified the variability space of the model parameters and employed UA/SA as a tool for the theory-driven design of nanoparticle formation processes. Here, we showed that—within simplifying assumptions such as the constancy of C-S-H composition across the considered experimental conditions—one could take advantage of UA/SA to decide on the optimal synthesis conditions for a target property in the final product. The procedure was demonstrated using the specific surface area of particles as the output, revealing the critical role of reagent addition rate and intercrystallite cohesion energy in obtaining the desirable outcome. These results warrant similar future studies using mechanistic kinetic models with carefully calibrated parameters that are only dependent on the system specifications. Such models allow for UA/SA over a wide range of operating conditions, which subsequently provide an invaluable basis for the rational design of production units. In the particular case of C-S-H system examined here, future experimental work relating the effect of different factors—most importantly, mixing/hydrodynamics as well as additives such as polymers and/or salts—on the properties of the product could lead to the development of more predictive models, which in turn guide the synthesis more decisively. In a broader sense, the current study provides a basis for future efforts where meticulously developed computational models guide the practical implementation in the laboratory.

Declaration of Competing Interest

None.

CRediT authorship contribution statement

M. Reza Andalibi: Methodology, Software, Validation, Visualization, Formal analysis, Writing - original draft, Writing - review & editing, Visualization. **Paul Bowen:** Supervision, Project administration, Funding acquisition, Writing - original draft, Writing - review & editing. **Agnese Carino:** Software, Resources, Visualization. **Andrea Testino:** Supervision, Project administration, Funding acquisition, Writing - original draft, Writing - review & editing.

Acknowledgements

This work was supported by the Energy and Environment Research Division (ENE), Paul Scherrer Institute, Switzerland.

The authors thank Dr. Sasa Bjelic (PSI) for providing additional computational resources, Prof. Christian Ludwig (PSI and EPFL) and Dr. Debora Foppiano (formerly at PSI) for valuable discussions, Dr. Laurin Wissmeier (CSD INGENIEURE AG) and Prof. D. Andrew Barry (EPFL) for providing their MATLAB codes for interfacing with IPhreeqc, and Prof. Karen Scrivener (EPFL) for continued support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.compchemeng.2020.106971.

References

Al, R., Behera, C.R., Zubov, A., Gernaey, K.V., Sin, G., 2019. Meta-modeling based efficient global sensitivity analysis for wastewater treatment plants – An application to the BSM2 model. Comput. Chem. Eng. 127, 233–246. doi:10.1016/j.compchemeng.2019.05.015.

Andalibi, M.R., Kumar, A., Srinivasan, B., Bowen, P., Scrivener, K., Ludwig, C., Testino, A., 2018. On the mesoscale mechanism of synthetic calcium–silicate–hydrate precipitation: a population balance modeling approach. J. Mater. Chem. A 6, 363–373. doi:10.1039/C7TA08784E.

Arfken, G.B., George, B, Weber, H.-J., Harris, F.E., 2012. Mathematical Methods For Physicists, 7th ed Academic Press.

Becker, W., Saltelli, A., 2015. Design for Sensitivity Analysis, in: Handbook of Design and Analysis of Experiments. Chapman and Hall/CRC, pp. 627–673.

Campolongo, F., Saltelli, A., 1997. Sensitivity analysis of an environmental model: an application of different analysis methods. Reliab. Eng. Syst. Saf. 57, 49–69. doi:10.1016/S0951-8320(97)00021-5.

Carino, A., Ludwig, C., Cervellino, A., Müller, E., Testino, A., 2018. Formation and transformation of calcium phosphate phases under biologically relevant conditions: experiments and modelling. Acta Biomater. 74, 478–488.

Carino, A., Testino, A., Andalibi, M.R., Pilger, F., Bowen, P., Ludwig, C., 2017. Thermodynamic-Kinetic precipitation modelling. A case study: the amorphous calcium carbonate (ACC) precipitation pathway unravelled. Cryst. Growth Des. 17, 2006–2015. doi:10.1021/acs.cgd.7b00006.

Charlton, S.R., Parkhurst, D.L., 2011. Modules based on the geochemical model PHREEQC for use in scripting and programming languages. Comput. Geosci. 37, 1653–1663. doi:10.1016/J.CAGEO.2011.02.005.

Chiang, W.-S., Fratini, E., Baglioni, P., Liu, D., Chen, S.-H., 2012. Microstructure determination of calcium-silicate-hydrate globules by small-angle neutron scattering. J. Phys. Chem. C 116, 5055–5061. doi:10.1021/jp300745g.

Dale, A.L., Lowry, G.V., Casman, E.A., 2017. Accurate and fast numerical algorithms for tracking particle size distributions during nanoparticle aggregation and dissolution. Environ. Sci. Nano 4, 89–104. doi:10.1039/C6EN00330C.

Dietemann, M., Baillon, F., Espitalier, F., Calvet, R., Greenhill-Hooper, M., 2019. Amorphous magnesium silicate ultrasound-assisted precipitation in a mixing system: population balance modelling and crystallization rates identification. Powder Technol. 356, 83–96. doi:10.1016/J.POWTEC.2019.08.004.

Dirksen, J.A., Ring, T.A., 1991. Fundamentals of crystallization: kinetic effects on particle size distributions and morphology. Chem. Eng. Sci. 46, 2389–2427.

Dolado, J.S., Griebel, M., Hamaekers, J., Heber, F., 2011. The nano-branched structure of cementitious calcium–silicate–hydrate gel. J. Mater. Chem. 21, 4445–4449.

Freeman, J.B., Dale, R., 2013. Assessing bimodality to detect the presence of a dual cognitive process. Behav. Res. Methods 45, 83–97. doi:10.3758/s13428-012-0225-x.
Galbraith, S.C., Schneider, P.A., 2014. Modelling and simulation of inorganic precipitation with nucleation, crystal growth and aggregation: a new approach to an old method. Chem. Eng. J. 240, 124–132.
Garcia Sanchez, D., Lacarrière, B., Musy, M., Bourges, B., 2014. Application of sensitivity analysis in building energy simulations: Combining first- and second-order elementary effects methods. Energy Build. 68, 741–750. doi:10.1016/J.ENBUILD.2012.08.048.
Gatty, L., Bonnamy, S., Feylessoufi, A., Clinard, C., Richard, P., Van Damme, H., 2001. A transmission electron microscopy study of interfaces and1. Gatty, L. et al. A transmission electron microscopy study of interfaces and matrix homogeneity in ultra-high-performance cement-based materials. J. Mater. Sci. 36, 4013–4026. doi:10.1023/A:1017938725748, 2001). matrix homog. J. Mater. Sci. 36, 4013–4026https://doi.org/.
Haderlein, M., Güldenpfennig, A., Segets, D., Peukert, W., 2017. A widely applicable tool for modeling precipitation processes. Comput. Chem. Eng. 98, 197–208. doi:10.1016/J.COMPCHEMENG.2016.12.007.
Hartigan, P.M., 1985. Algorithm AS 217: computation of the dip statistic to test for unimodality. Appl. Stat. 34, 320. doi:10.2307/2347485.
Hazewinkel, M., 2002. Unimodal distribution. Encycl. Math.
Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliab. Eng. Syst. Saf. 52, 1–17. doi:10.1016/0951-8320(96)00002-6.
Iman, R.L., Hora, S.C., 1990. A robust measure of uncertainty importance for use in fault tree system analysis. Risk Anal 10, 401–406. doi:10.1111/j.1539-6924.1990.tb00523.x.
Ioannidou, K., Krakowiak, K.J., Bauchy, M., Hoover, C.G., Masoero, E., Yip, S., Ulm, F.-J., Levitz, P., Pellenq, R.J.-M., Del Gado, E., 2016. Mesoscale texture of cement hydrates. Proc. Natl. Acad. Sci. U. S. A. 113, 2029–2034. doi:10.1073/pnas.1520487113.
Iversen, T.B., Sin, G., 2019. Global uncertainty and sensitivity analysis for robust design of a rotary kiln process, in: computer aided chemical engineering. Elsevier B.V., pp. 805–810. https://doi.org/ 10.1016/B978-0-12-818634-3.50135-1
Jennings, H.M., Thomas, J.J., Gevrenov, J.S., Constantinides, G., Ulm, F.-J., 2007. A multi-technique investigation of the nanoporosity of cement paste. Cem. Concr. Res. 37, 329–336. doi:10.1016/j.cemconres.2006.03.021.
Jiang, M., Braatz, R.D., 2019. Designs of continuous-flow pharmaceutical crystallizers: developments and practice. Cryst. Eng. Comm. 21, 3534–3551. doi:10.1039/C8CE00042E.
Jiao, W., 2016. Facile synthesis of silver nanoparticles deposited on a calcium silicate hydrate composite as an efficient bactericidal agent. RSC Adv 6, 112931–112938. doi:10.1039/C6RA24265K.
Jönsson, B., Nonat, A., Labbez, C., Cabane, B., Wennerström, H., 2005. Controlling the cohesion of cement paste. Langmuir 21, 9211–9221. doi:10.1021/la051048z.
Kashchiev, D., 2003. Thermodynamically consistent description of the work to form a nucleus of any size. J. Chem. Phys. 118, 1837–1851. doi:10.1063/1.1531614.
Khorashadi Zadeh, F., Nossent, J., Sarrazin, F., Pianosi, F., van Griensven, A., Wagener, T., Bauwens, W., 2017. Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model. Environ. Model. Softw. 91, 210–222. doi:10.1016/J.ENVSOFT.2017.02.001.
Kulik, D.A., Wagner, T., Dmytrieva, S.V, Kosakowski, G., Hingerl, F.F., Chudnenko, K.V, Berner, U.R., 2013. GEM-Selektor geochemical modeling package: revised algorithm and GEMS3K numerical kernel for coupled simulation codes. Comput. Geosci. 17, 1–24.
Kumar, A., 2017. Synthetic calcium silicate hydrates. EPFL doi:10.5075/EPFL-THESIS-7658.
Kumar, A., Walder, B.J., Mohamed, Kunhi, A., Hofstetter, A., Srinivasan, B., Rossini, A.J., Scrivener, K., Emsley, L., Bowen, 2017. The atomic-level structure of cementitious calcium silicate hydrate. J. Phys. Chem. C acs.jpcc, .7b02439 doi:10.1021/acs.jpcc.7b02439.
Kunhi Mohamed, A., Parker, S.C., Bowen, P., Galmarini, S., 2018. An atomistic building block description of C-S-H - Towards a realistic C-S-H model. Cem. Concr. Res. 107, 221–235. doi:10.1016/J.CEMCONRES.2018.01.007.
Lothenbach, B., Kulik, D.A., Matschei, T., Balonis, M., Baquerizo, L., Dilnesa, B., Miron, G.D., Myers, R.J., 2019. Cemdata18: a chemical thermodynamic database for hydrated Portland cements and alkali-activated materials. Cem. Concr. Res. 115, 472–506. doi:10.1016/J.CEMCONRES.2018.04.018.
Marchisio, D.L., 2009. On the use of bi-variate population balance equations for modelling barium titanate nanoparticle precipitation. Chem. Eng. Sci. 64, 697–708.
Marchisio, D.L., Fox, R.O., 2013. Computational models for polydisperse particulate and multiphase systems. Cambridge University Press.
Marchisio, D.L., Fox, R.O., 2005. Solution of population balance equations using the direct quadrature method of moments. J. Aerosol Sci. 36, 43–73. doi:10.1016/J.JAEROSCI.2004.07.009.
Marchisio, D.L., Pikturna, J.T., Fox, R.O., Vigil, R.D., Barresi, A.A., 2003. Quadrature method of moments for population-balance equations. AIChE J. 49, 1266–1276.
Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E., 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J. Theor. Biol. 254, 178–196. doi:10.1016/j.jtbi.2008.04.011.
Mehrali, Mehdi, Seyed Shirazi, S.F., Baradaran, S., Mehrali, Mohammad, Metselaar, H.S.C., Kadri, N.A., Bin, Osman, N.A.A., 2014. Facile synthesis of calcium silicate hydrate using sodium dodecyl sulfate as a surfactant assisted by ultrasonic irradiation. Ultrason. Sonochem. 21, 735–742. doi:10.1016/J.ULTSONCH.2013.08.012.
Menberg, K., Heo, Y., Choudhary, R., 2016. Sensitivity analysis methods for building energy models: comparing computational costs and extractable information. Energy Build 133, 433–445. https://doi.org/10.1016/J.ENBUILD.2016.10.005.
Mersmann, A., 2001. Crystallization technology handbook. CRC Press.
Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33, 161–174. https://doi.org/10.1080/00401706.1991.10484804.
Morris, M.D., Moore, L.M., 2015. Design of computer experiments: introduction and background,. In: Dean, A., Morris, M.D., Stufken, J., Bingham, D. (Eds.), Handbook of Design and Analysis of Experiments. Chapman and Hall/CRC, New York, NY, pp. 597–612.
Mutaftschiev, B., 2001. The atomistic nature of crystal growth. Springer Series in Materials Science. Springer, Berlin Heidelberg, Berlin, Heidelberg https://doi.org/10.1007/978-3-662-04591-6.
Noacco, V., Sarrazin, F., Pianosi, F., Wagener, T., 2019. Matlab/R workflows to assess critical choices in global sensitivity analysis using the SAFE toolbox. MethodsX 6, 2258–2280. https://doi.org/10.1016/j.mex.2019.09.033.
Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol' sensitivity analysis of a complex environmental model. Environ. Model. Softw. 26, 1515–1525. https://doi.org/10.1016/j.envsoft.2011.08.010.
Öner, M., Stocks, S.M., Sin, G., 2020. Comprehensive sensitivity analysis and process risk assessment of large scale pharmaceutical crystallization processes. Comput. Chem. Eng. 135, 106746. https://doi.org/10.1016/j.compchemeng.2020.106746.
Parkhurst, D.L., Appelo, C.A.J., 2013. Parkhurst, David L., and C. A. J. Appelo. Description of input and examples for PHREEQC version 3: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. No. 6-A43. US Geological Survey. Lakewood, Colo. : U.S. Dept. of the Interior, U.S. Geological Survey ; Denver, CO : Earth Science Information Center, Open-File Reports Section [distributor], 1995.
Peng, Y., Zhu, Z., Braatz, R.D., Myerson, A.S., 2015. Gypsum crystallization during phosphoric acid production: modeling and experiments using the mixed-solvent-electrolyte thermodynamic model. Ind. Eng. Chem. Res. 54, 7914–7924. https://doi.org/10.1021/acs.iecr.5b01763.
Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environ. Model. Softw. 79, 214–232. https://doi.org/10.1016/j.envsoft.2016.02.008.
Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. Environ. Model. Softw. 70, 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009.
Pianosi, F., Wagener, T., 2018. Distribution-based sensitivity analysis from a generic input-output sample. Environ. Model. Softw. 108, 197–207. https://doi.org/10.1016/J.ENVSOFT.2018.07.019.
Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environ. Model. Softw. 67, 1–11. https://doi.org/10.1016/J.ENVSOFT.2015.01.004.
Plassard, C., Lesniewska, E., Pochard, I., Nonat, A., 2005. Nanoscale experimental investigation of particle interactions at the origin of the cohesion of cement. Langmuir 21, 7263–7270. https://doi.org/10.1021/LA050440+.
Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., Cambridge University Press, 2007. Numerical recipes : the art of scientific computing.
Prieto, M., Heberling, F., Rodríguez-Galán, R.M., Brandt, F., 2016. Crystallization behavior of solid solutions from aqueous solutions: an environmental perspective. Prog. Cryst. Growth Charact. Mater. 62, 29–68. https://doi.org/10.1016/j.pcrysgrow.2016.05.001.
Rehage, H., Scherer, S., Kind, M., 2019. A steady-state precipitation model for flowsheet simulation and its application. Comput. Chem. Eng. 128, 524–537. https://doi.org/10.1016/J.COMPCHEMENG.2019.06.030.
Richardson, I.G., 2004. Tobermorite/jennite- and tobermorite/calcium hydroxide-based models for the structure of C-S-H: Applicability to hardened pastes of tricalcium silicate, β-dicalcium silicate, Portland cement, and blends of Portland cement with blast-furnace slag, metakaol. Cem. Concr. Res. 34, 1733–1777. https://doi.org/10.1016/j.cemconres.2004.05.034.
Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. 181, 259–270. https://doi.org/10.1016/j.cpc.2009.09.018.
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2007. The Primer. Global Sensitivity Analysis. John Wiley & Sons, Ltd, Chichester, UK https://doi.org/10.1002/9780470725184.
Saltelli, A., Sobol', I.M., 1995. About the use of rank transformation in sensitivity analysis of model output. Reliab. Eng. Syst. Saf. 50, 225–239. https://doi.org/10.1016/0951-8320(95)00099-2.
Schroeder, B.B., Harris, D.D., Smith, S.T., Lignell, D.O., 2014. Theoretical framework for multiple-polymorph particle precipitation in highly supersaturated systems. Cryst. Growth Des. 14, 1756–1770.
Schwarzer, H.-C., Schwertfirm, F., Manhart, M., Schmid, H.-J., Peukert, W., 2006. Predictive simulation of nanoparticle precipitation based on the population balance equation. Chem. Eng. Sci. 61, 167–181.
Scrivener, K., Ouzia, A., Juilland, P., Kunhi Mohamed, A., 2019. Advances in understanding cement hydration mechanisms. Cem. Concr. Res. 124, 105823. https://doi.org/10.1016/j.cemconres.2019.105823.
Scrivener, K.L., Juilland, P., Monteiro, P.J.M., 2015. Advances in understanding hydra-

tion of Portland cement. Cem. Concr. Res. 78, 38–56. https://doi.org/10.1016/J.CEMCONRES.2015.05.025.

Shampine, L.F., Gladwell, I., Thompson, S., 2003. Solving ODEs with MATLAB. Cambridge University Press, Cambridge https://doi.org/10.1017/CBO9780511615542.

Shao, N., Tang, S., Liu, Z., Li, L., Yan, F., Liu, F., Li, S., Zhang, Z., 2018. Hierarchically structured calcium silicate hydrate-based nanocomposites derived from steel slag for highly efficient heavy metal removal from wastewater. ACS Sustain. Chem. Eng. 6, 14926–14935. https://doi.org/10.1021/acssuschemeng.8b03428.

Sheng, K., Zeng, F., Pang, F., Ge, J., 2019. Highly dispersed ni nanoparticles on anhydrous calcium silicate (ACS) nanosheets for catalytic dry reforming of methane: tuning the activity by different ways of Ni introduction. Chem. – An Asian J. 14. asia.201900611 https://doi.org/10.1002/asia.201900611 .

Silva, L.F.L.R., Rodrigues, R.C., Mitre, J.F., Lage, P.L.C., 2010. Comparison of the accuracy and performance of quadrature-based methods for population balance problems with simultaneous breakage and aggregation. Comput. Chem. Eng. 34, 286–297. https://doi.org/10.1016/J.COMPCHEMENG.2009.11.005.

Söhnel, O., Garside, J., 1992. Precipitation: basic principles and industrial applications. Butterworth-Heinemann.

Testino, A., Buscaglia, V., Buscaglia, M.T., Viviani, M., Nanni, P., 2005. Kinetic modeling of aqueous and hydrothermal synthesis of barium titanate (BaTiO3). Chem. Mater. 17, 5346–5356.

Thomas, J.J., Biernacki, J.J., Bullard, J.W., Bishnoi, S., Dolado, J.S., Scherer, G.W., Luttge, A., 2011. Modeling and simulation of cement hydration kinetics and microstructure development. Cem. Concr. Res. 41, 1257–1278. https://doi.org/10.1016/j.cemconres.2010.10.004.

Wissmeier, L., Barry, D.A., 2011. Simulation tool for variably saturated flow with comprehensive geochemical reactions in two- and three-dimensional domains. Environ. Model. Softw. 26, 210–218. https://doi.org/10.1016/J.ENVSOFT.2010.07.005.

Wu, J., Zhu, Y.-J., Chen, F., 2013. Ultrathin calcium silicate hydrate nanosheets with large specific surface areas: synthesis, crystallization, layered self-assembly and applications as excellent adsorbents for drug, protein, and metal ions. Small 9, 2911–2925. https://doi.org/10.1002/smll.201300097.

Xia, X., Xu, Y., Chen, Y., Liu, Y., Lu, Y., 2019. The distinct catalytic behaviours of calcium silicate hydrate for the high selectivity of 2,2′-isomer in reaction of phenol with formaldehyde. Catal. Commun. 118, 15–18. https://doi.org/10.1016/J.CATCOM.2018.09.004.

Yang, J., 2011. Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis. Environ. Model. Softw. 26, 444–457. https://doi.org/10.1016/j.envsoft.2010.10.007.

Zhang, C., Xu, K., Zheng, M., Li, J., Wang, C., 2018. Factors affecting the crystal size of Struvite-K formed in synthetic urine using a stirred reactor. Ind. Eng. Chem. Res. 57, 17301–17309. https://doi.org/10.1021/acs.iecr.8b03328.

Zhang, H., Liu, Q., Wang, J., Liu, J., Yan, H., Jing, X., Zhang, B., 2015. Preparation of magnetic calcium silicate hydrate for the efficient removal of uranium from aqueous systems. RSC Adv. 5, 5904–5912. https://doi.org/10.1039/C4RA08678C.

Zhang, X., Chang, W., Zhang, T., Ong, C.K., 2004. Nanostructure of calcium silicate hydrate gels in cement paste. J. Am. Ceram. Soc. 83, 2600–2604. https://doi.org/10.1111/j.1151-2916.2000.tb01595.x.

Zhao, J., Zhu, Y.-J., Wu, J., Zheng, J.-Q., Zhao, X.-Y., Lu, B.-Q., Chen, F., 2014. Chitosan-coated mesoporous microspheres of calcium silicate hydrate: Environmentally friendly synthesis and application as a highly efficient adsorbent for heavy metal ions. J. Colloid Interface Sci. 418, 208–215. https://doi.org/10.1016/J.JCIS.2013.12.016.

Zhu, Y.-J., Sham, T.-K., 2014. The potential of calcium silicate hydrate as a carrier of ibuprofen. Expert Opin. Drug Deliv. 11, 1337–1342. https://doi.org/10.1517/17425247.2014.923399.