



HAL
open science

Uncertainty quantification and global sensitivity analysis of double-diffusive natural convection in a porous enclosure

Mohammad Mahdi Rajabi, Marwan Fahs, Aref Panjehfouladgaran, Behzad
Ataie-Ashtiani, Craig T Simmons, Benjamin Belfort

► **To cite this version:**

Mohammad Mahdi Rajabi, Marwan Fahs, Aref Panjehfouladgaran, Behzad Ataie-Ashtiani, Craig T Simmons, et al.. Uncertainty quantification and global sensitivity analysis of double-diffusive natural convection in a porous enclosure. *International Journal of Heat and Mass Transfer*, 2020, 162, pp.120291. 10.1016/j.ijheatmasstransfer.2020.120291 . hal-03452433

HAL Id: hal-03452433

<https://hal.science/hal-03452433v1>

Submitted on 26 Nov 2021

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

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



Uncertainty quantification and global sensitivity analysis of double-diffusive natural convection in a porous enclosure

Mohammad Mahdi Rajabi^{a,*}, Marwan Fahs^b, Aref Panjehfouladgaran^a, Behzad Ataie-Ashtiani^{c,d}, Craig T. Simmons^d, Benjamin Belfort^b

^a Civil and Environmental Engineering Faculty, Tarbiat Modares University, PO Box 14115-397, Tehran, Iran

^b LHyGES, Univ. de Strasbourg/EOST/ENGEEES, CNRS, 1 rue Blessig, 67084 Strasbourg, France

^c Department of Civil Engineering, Sharif University of Technology, PO Box 11155-9313, Tehran, Iran

^d National Centre for Groundwater Research and Training, College of Science and Engineering, Flinders University, GPO Box 2100, Adelaide, SA 5001, Australia

ARTICLE INFO

Article history:

Received 21 February 2020

Received in revised form 6 July 2020

Accepted 4 August 2020

Available online xxx

Keywords

Double-diffuse convection

Porous media

Uncertainty propagation analysis

Global sensitivity analysis

Polynomial chaos expansion

ABSTRACT

In this paper, detailed uncertainty propagation analysis (UPA) and variance-based global sensitivity analysis (GSA) are performed on the widely adopted double-diffuse convection (DDC) benchmark problem of a square porous cavity with horizontal temperature and concentration gradients. The objective is to understand the impact of uncertainties related to model parameters on metrics characterizing flow, heat and mass transfer processes, and to derive spatial maps of uncertainty and sensitivity indices which can provide physical insights and a better understanding of DDC processes in porous media. DDC simulations are computationally expensive and UPA and GSA require large number of simulations, so an appropriate strategy is developed to reduce the computational burden. The approach is built on two pillars: (a) an efficient numerical simulator based on the Fourier series method that generates training data, and (b) polynomial chaos expansion (PCE) meta-models that are trained using the simulator data, and then replace the numerical model in UPA and GSA. Assuming that the Rayleigh number (R_a), the solutal to thermal buoyancy ratio (N_b) and the Lewis number (L_e) are the uncertain input variables, the results of UPA show that the zones of high temperature and concentration variability are located in the regions where the flow is mainly driven by the buoyancy effects. GSA indicates that N_b is the most influential parameter affecting the temperature and concentration fields, followed respectively by R_a and L_e . For heat-driven flow cases ($N_b > -1$), the concentration field is more influenced by L_e than R_a . For deeper understanding of uncertainty propagation, we estimate the bias introduced by replacing uncertain parameters by deterministic values. The resulting spatial maps of the difference between deterministic output and stochastic mean show that a deterministic approach leads to different zones where the temperature, concentration and velocity fields can be either overestimated or underestimated. The conclusions drawn in this work are likely to be helpful in different applications involving DDC in porous enclosures leading to convective circulation cells.

© 2020

1. Introduction

Free convection or buoyancy driven flow caused by density variations in saturated porous media, has been studied extensively in the literature due to its wide applicability. These density variations may occur due to gradients in the fluid composition or temperature. Simultaneous occurrence of both gradients causes a type of flow known as double-diffuse convection (DDC), thermosolutal or thermohaline flow [38,40]. Instances of DDC in porous media are widespread, ranging from large scale problems in geothermal engineering [31], CO₂ sequestration [25,26], nuclear waste disposal [22] and enhanced recovery of

petroleum reservoirs [6], to small scale problems encountered in alloy solidification [9], fluidized beds and fuel cells [2]. DDC in porous media is often simulated by employing Darcy momentum conservation law with variable density, in conjunction with heat and mass transfer equations. Due to nonlinearity, DDC problems have no general analytical solutions and are often simulated using numerical models.

Use of numerical models has become widespread in the study and design of physical systems involving DDC in porous media. However, the structure and input parameters of these models are almost always prone to uncertainty resulting from simplifying assumptions, poor knowledge of underlying mechanisms, data insufficiency, natural stochasticity, etc. These sources of uncertainty propagate through the model and lead to uncertainty in model outputs. As stated by several previous studies (e.g. [37,65]), the resulting model output uncertainties have negative effects on model reliability in practical applications, and this strengthens the need for a more in depth study of the critical

* Corresponding author.

E-mail address: mmrajabi@modares.ac.ir (M.M. Rajabi)

issue of simulation under uncertainty. Addressing this issue involves at least two interrelated aspects: first, quantification of model output uncertainties resulting from the propagation of input uncertainties through the numerical model, and second, allocation of model output uncertainties to different sources of uncertainty in the model inputs [14,42,54]. The first aspect is known as uncertainty propagation analysis (UPA), and the second aspect requires global sensitivity analysis (GSA). GSA and UPA should be performed concurrently, as both are essential parts of the model development process in reliability analysis, robust design optimization, data-worth analysis, and risk assessment.

The majority of studies on DDC in porous media use a scenario-based approach to analyze the effect of inputs variations on model outputs (e.g. [17,30,46,49, 50, 64]), and despite the importance of the subject, few studies have addressed formal UPA and GSA for various types of DDC problems in porous media, or even done so in the broader context of DDC in bulk fluids. Key studies in this regard are reviewed in Table 1. Le Maitre et al. [32] discussed different techniques of UPA for compressible flows. Ganapathysubramanian and Zabarar [18] analyzed the effects of random parameters on natural convection in bulk-fluids. The same problem has been addressed in Venturi et al. [57] by assuming perturbed boundary conditions. Based on Monte Carlo method. Shome et al. [53] performed UPA for mixed convection in a circular tube. Fajraoui et al. [14] conducted UPA for natural convection in porous media. A parameter sensitivity analysis is presented in Shirvan et al. [52] for heat transfer in a porous solar cavity receiver. Shahane et al. [48] used deep neural networks to perform UPA for natural convection in a 3D box cavity. This review shows that existing studies are limited to purely thermal natural convection. Most of these studies in both porous media and bulk fluids, investigated scalar variables characterizing the overall heat transfer flux and velocity predictions, such as the average Nusselt number and maximum velocities. Hence, information on the spatial variability of model output uncertainties and sensitivity indices are not available, though similar studies have been done in Fajraoui et al. [14]. So, there are gaps in the investigation of UPA and GSA for problems of DDC in porous media. This study tends to address these gaps by performing detailed UPA and variance-based GSA on the widely adopted DDC benchmark problem of a square porous cavity with horizontal temperature and concentration gradients. The objective is to understand the impact of uncertainties re-

lated to model parameters on metrics characterizing flow, heat and mass transfer processes, and to derive spatial maps of uncertainty and sensitivity indices which can provide physical insights and a better understanding of DDC processes in porous media.

Perhaps the most important hindrance in the way of UPA and GSA of DDC problems is the computational cost. Due to nonlinearity and high dimensionality of DDC models, the unit cost of a forward simulation could be very high, depending on the time and space scale. Even for small time and space scales, simulation of DDC at high Rayleigh numbers requires dense computational grid and small time steps size, which contribute to a high CPU time. GSA and UPA both involve repetitive simulations by the numerical model, and hence the large number of DDC simulations required for obtaining accurate solutions may become computationally infeasible. To overcome the computational challenge, we (a) employ a highly efficient simulator based on the Fourier series solution [49] which allows for accurate solution with reduced degree of freedom, and (b) use the numerical model to train and validate polynomial chaos expansions (PCEs). The PCEs then replace the numerical model in UPA and GSA computations [43].

The structure of the present study is as follows. In Section 2, we describe our DDC problem and the governing equations of the mathematical model used in its simulations. Section 3 describes the UPA and GSA procedure and how PCEs are used in this context. In Section 4 the results of the UPA and GSA are presented with regards to each of the output quantities of interest (QoIs) and physical insight are provided. Finally, the summary and conclusions are provided in Section 5.

2. Problem framework

2.1. Problem statement and assumptions

Numerical models of DDC in porous media have been used in many industrial and environmental problems under realistic configurations. However in most theoretical studies, DDC in porous media is studied using the hypothetical problem of a porous enclosure. It is a very popular benchmark for DDC numerical codes (see the review by Corcione et al. [10]). Part of the popularity of this benchmark problem stems from its simplicity in terms of geometry and boundary conditions, and the fact that there are many published studies regarding its numerical solution using a variety of numerical methods [49]. This problem has been

Table 1
Review of literature on UPA and GSA for natural convection in bulk fluids and porous media.

| Reference | Type of problem | Uncertain input (s) | Uncertain output (s) | Method |
|---|--|---|--|--------------------------|
| Le Maitre et al. [32] | 2D (SC-HG), TNC, BF (Zero-Mach-number flows) | BT (GD) | T and q fields, \overline{Nu} | UPA (PCE) |
| Ganapathysubramanian, and Zabarar, [18] | 2D (SC-HG),TNC, BF/SPM | BT, BSR, ϵ (GD) | T , q and p fields | UPA (PCE) |
| Venturi et al., [57] | 2D (SC-HG), TNC, BF | BT (GD) | \overline{Nu} | UPA (MCS), GSA (Sobol) |
| Shome [53] | 3D (circular tube), TNC, BF, | R_a ($10^5 - 10^7$), P_r ($0 - 1000$), H_f ($100 - 1000$) | $\overline{Nu} \cdot f$ | MSC |
| Fajraoui et al. [14] | 2D (SC-HG), TNC, SPM | R_a ($0 - 1000$), r_k ($0 - 1$), α_L , τ ($0.1-1$, $0.01-0.1$), $\partial k_{x,y}$ ($0 - 4$) | T field, q_{max} | UPA (PCE) |
| Shirvan et al. [52] | 2D (inclined SC), TNC, SPM | R_a ($10^4 - 10^6$), D_a ($10^{-5} - 10^3$), θ ($0 - 90^\circ$), δ ($\frac{1}{3} - 1$), ϵ ($0 - 1$) | \overline{Nu} Nu | GSA (Sobol) GSA (RSM) |
| Shahane et al., [48] | 2D(SC-HG)/3D(cube), TNC, BF | P_r (GD ($\mu = 7.5$, $\sigma = 0.02\mu$)), R_a (GD ($\mu = 10^5 - 10^6$, $\sigma = 0.02\mu$)), BT (GD) | Nu , q and T fields | UPA (PCE/DNN) |

2D: two-dimensional, 3D: three dimensional, SC: square cavity, HG: horizontal gradients
TNC: temperature-induced natural convection, BF: Bulk fluid, SPM: Saturated porous media

D_a : Darcy number, R_a : Rayleigh number, P_r : Prandtl number, BT: boundary temperatures, H_f : heat flux, k : permeability, $\partial k_{x,y}$: Spatial rates of change in permeability, r_k : permeability anisotropy ratio, α_L , τ : dispersion coefficients, θ : inclination angles, δ : dimensionless porous substrate thickness, ϵ : Emissivity, BSR: boundary surface roughness, ϵ : porosity, p : pressure
GD: Gaussian distribution, \overline{Nu} : mean Nusselt number, q_{max} : maximum velocities, T : temperature, f : friction factors, q : velocity

PCE: Polynomial Chaos expansion, DNN: Deep Neural Network, RSM: response surface methodology, MCS: Monte Carlo Simulation, Sobol: Sobol sensitivity indices

the topic of GSA in Fajraoui et al. [14], but for purely thermal natural convection (without solute transport). To the author's knowledge, GSA and UPA for DDC in porous enclosure have not been assessed in the literature.

DDC in porous enclosure has been investigated for a variety of enclosure geometries (e.g. trapezoidal as in [35], circular tube as in [53]), orientations [4], gradient directions [21,36] and boundary conditions (e.g. [27]). Here we focus on the most widely adopted DDC benchmark problem in the relevant literature. As demonstrated in Fig. 1, we consider a square domain of length H [L] filled with a saturated porous medium. It is assumed that the third dimension of the enclosure is large enough so that the fluid flow and heat and mass transports are two-dimensional. We assume Dirichlet boundary conditions for temperature (T [Θ]) and concentration (C [ML^{-3}]) at the vertical walls, and thus have constant values of T_L and C_L on the left wall, and T_R and C_R on the right wall. The left side wall has a higher temperature and a higher concentration than the right wall, hence we have $T_L > T_R$ and $C_L > C_R$. The top and bottom boundaries of the problem domain are thermally adiabatic and impermeable, and are represented by Neumann boundary conditions ($\frac{\partial T}{\partial y} = \frac{\partial C}{\partial y} = 0$). Gravity acts in the negative y -direction which is orthogonal to the direction of the temperature/solute gradients. The velocity components in x and y Cartesian coordinates are represented with v_x and v_y [ML^{-1}]. Impervious boundary conditions are imposed on all walls, so that we have $v_x = 0$ on the vertical walls, and $v_y = 0$ on the horizontal surfaces. The porous media is assumed to possess isotropic and homogenous thermo-physical properties, and is in local thermal and compositional equilibrium with the saturating fluid. Viscous dissipation and porous medium inertia are not considered, and the Soret and Dufour effects are neglected. The fluid is assumed to be Newtonian and incompressible, and its flow in the cavity is steady-state laminar and complies with Darcy's law. The Boussinesq approximation is considered to facilitate the numerical solution procedure. Validity of this approach for natural convection in porous media is discussed in Fahs et al. (2019).

2.2. Mathematical model and Fourier series solution

The coupled fluid flow, mass and heat transfer can be written in the following non-dimensional form [25, 49, 50]:

$$\frac{\partial v_x^*}{\partial x^*} + \frac{\partial v_y^*}{\partial y^*} = 0 \tag{1}$$

$$v_x^* = -\frac{\partial p^*}{\partial x^*} \tag{2}$$

$$v_y^* = -\frac{\partial p^*}{\partial y^*} - R_a [N_b C^* + T^*] \tag{3}$$

$$v_x^* \frac{\partial T^*}{\partial x^*} + v_y^* \frac{\partial T^*}{\partial y^*} - \left(\frac{\partial^2 T^*}{\partial x^{*2}} + \frac{\partial^2 T^*}{\partial y^{*2}} \right) = 0 \tag{4}$$

$$v_x^* \frac{\partial C^*}{\partial x^*} + v_y^* \frac{\partial C^*}{\partial y^*} - \frac{1}{L_e} \left(\frac{\partial^2 C^*}{\partial x^{*2}} + \frac{\partial^2 C^*}{\partial y^{*2}} \right) = 0 \tag{5}$$

where, T^* , C^* , x^* , y^* , p^* , v_x^* and v_y^* are non-dimensional temperature, concentration, horizontal and vertical Cartesian coordinates, fluid pressure and velocity components, respectively. The above system of equations is governed by three dimensionless parameters, namely the porous thermal Rayleigh number R_a , the solutal to thermal buoyancy ratio N_b and the Lewis number L_e . These parameters are defined as follows [63]:

$$R_a = \frac{\rho_0 g k H \beta_T \Delta T}{\mu \alpha} \tag{6}$$

$$N_b = \frac{\beta_C \Delta C}{\beta_T \Delta T} \tag{7}$$

$$L_e = \frac{\alpha}{D} \tag{8}$$

where ρ_0 [ML^{-3}] is reference fluid density, g [LT^{-2}] is acceleration due to gravity, k [L^2] is permeability, β_T [Θ^{-1}] and β_C [$M^{-1}L^3$] are the thermal and solute volumetric expansion coefficients respectively,

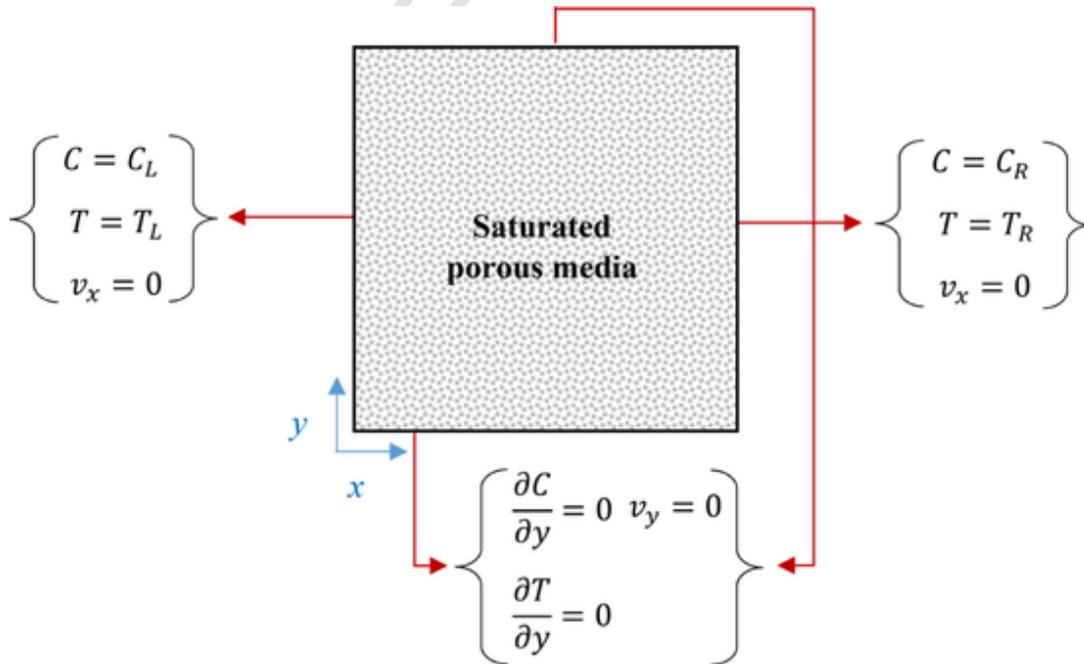


Fig. 1. Schematic diagram of the problem with description of boundary conditions.

$\Delta T = T_L - T_R$ and $\Delta C = C_L - C_R$ are the temperature and concentration difference between the left and right walls, μ [$ML^{-1}T^{-1}$] is dynamic viscosity and α [L^2T^{-1}] and D [L^2T^{-1}] are the thermal and molecular diffusivities

The porous thermal Rayleigh number (R_a) is the ratio of buoyancy and thermal diffusion forces; and high R_a implies higher buoyancy compared to viscous force. The buoyancy ratio (N_b) represents the relative strengths of the thermal and solute buoyancy forces. If $N_b > 0$, both thermal and solute gradient are cooperating each other; $N_b = 0$ represents pure heat-driven flows, and $N_b < 0$ shows opposing thermal and solute buoyancy forces. For thermal and solute double diffusive convection (i.e. thermohaline) N_b should be negative ($\beta_C > 0$ and $\beta_T < 0$). This assumption is considered in this work. The Lewis number (L_e) signifies the relative importance of thermal to solute molecular diffusion. When L_e is less than unity, the mass transfer process is dominant. L_e controls the concentration and temperature fields, and these two fields are only identical in case $L_e = 1$ [11].

In UPA and GSA, the performance of the method used to solve the governing equations is critical as repetitive simulations of the forward problem are required. In this work, the governing equations are solved using the series Fourier spectral method, as in Fahs et al. [13], [2016] and Shao et al. [49]. Thus, the governing equations are reformulated using the stream function, and the temperature and concentration are assumed as primary variables which are approximated by double Fourier series that satisfy the boundary conditions. The spectral system is simplified by expressing the spectral velocity field in terms of temperature and concentration. The Fourier series coefficients of concentration and temperature are calculated by solving the final nonlinear spectral system using an advanced nonlinear solver from the IMSL library (more details can be found in [49]). The advantages of using the Fourier series method are multiple. This method has a high convergence rate, allowing the ability to provide high accurate solution with reduced number of degree of freedom (Kopriva, 2009). The Fourier series method allows for significant CPU-time saving, as the system is solved directly as steady state. Finally, in standard numerical solutions (finite element or finite difference), the evaluation of the metrics used for UPA and GSA requires specific post-treatment of model output, which can introduce numerical error. With the Fourier series method, all these metrics are evaluated analytically without any approximation, using the Fourier series expansions. The Fourier series code has been validated by comparison against previously published works in Shao et al. [49].

3. Theoretical framework and methodology

3.1. Uncertainty propagation analysis (UPA)

UPA involves quantification of mode output uncertainty resulting from the propagation of input uncertainties through the model [7]. Monte Carlo simulation (MCS) is the most popular method for UPA in computational fluid dynamics [42]. MCS error is in the order of $\left(1/\sqrt{n_{MC}}\right)$ where n_{MC} is the number of Monte Carlo samples [8].

Hence achieving an acceptably small level of error often requires a large number of deterministic model simulations, which may be computationally difficult. A surrogate modeling approach based on PCE has been adopted in the current study to alleviate the computational challenge of performing MCS for UPA.

3.2. Polynomial chaos expansion (PCE)

Non-intrusive PCE is a kind of surrogate model that is known to be ideal for problems with low to moderate number of uncertain inputs. Details about PCE can be found in several papers including Le Maître and Knio, [33] and Xiu, [58]. Here we only provide a brief overview.

PCE characterizes the stochastic simulation output y as a series in the form of the following equation [19]:

$$y = \sum_{i=0}^{\infty} \alpha_i \psi_i(\zeta) \quad (9)$$

Every element of this series is the product of a deterministic coefficient α_i and a stochastic component $\psi_i(\zeta)$, where ζ is a random variable with a certain probability distribution and $\psi_i(\zeta)$ is a polynomial function of order i that satisfies the orthogonality condition. The optimal choice of the polynomial type for $\psi_i(\zeta)$ depends on the probability distribution for ζ , and this choice is often made using the Askey scheme [59]. The Legendre polynomials are used in this study as the probability distributions characterizing the uncertain inputs are assumed to be uniform. The deterministic PCE coefficients α_i can be estimated by non-intrusive methods which treat the physics of the system as a black-box, and use the known input-output pairs as training dataset for the estimation of α_i . Hence, non-intrusive PCEs are known as data-driven meta-models [45]. One particularly popular non-intrusive method for the estimation of PCE coefficients is the regression method. This method involves the following steps: (1) choosing a set of n_r regression points (each denoted by ϵ_k) from the probability space of the random input variable(s) using deterministic or random sampling methods, (2) employing these regression points to perform the same number of numerical simulations (we represent the model outputs for these n_r simulations with $y(\epsilon_k)$), and (3) estimating the PCE coefficients by solving the following least square optimization problem [44]:

$$\min \sum_{k=1}^{n_r} \left[y(\epsilon_k) - \sum_{i=0}^g \alpha_i \psi_i(\epsilon_k) \right]^2 \quad (10)$$

Note that the PCE term in Eq. (10) is truncated to g terms in order to make the computations practical. Estimation of each output QoI requires a PCE meta-model, and hence in the case of multivariate outputs (T^* , C^* , v_x^* and v_y^*), a PCE must be constructed for each points of the grid. In this study, PCE meta-models are constructed using the python package 'ChaosPy' [15].

3.3. Global sensitivity analysis (GSA)

In this paper, we perform GSA using a variance-based method. The most commonly used measures to quantify variance-based sensitivity are Sobol indices [24]. Their popularity stems at least from two facts: (1) they do not rely on any assumption regarding the linearity or monotonous behavior of the physical model and can hence be used for GSA of complex models [5], and (2) their interpretation is relatively easy [14]. The main notion of variance-based sensitivity analysis methods (including Sobol indices) is to measure the contribution of an input or a group of input variable to the output variance. We assume $y = f(x)$ to be a square-integrable function with mutually independent inputs $x = \{x_1, \dots, x_n\}$. The variance of y (denoted by $V(y)$) can be decomposed to (Sobol, 2001):

$$V(y) = \sum_{i=1}^n V_i + \sum_{i_1=1}^n \sum_{i_2=i_1+1}^n V_{i_1 i_2} + \sum_{i_1=1}^n \sum_{i_2=i_1+1}^n \sum_{i_3=i_2+1}^n V_{i_1 i_2 i_3} + \dots + V_{i_1 \dots i_n} \quad (11)$$

where V_i represents the variance of y with respect to the individual effect of each single input x_i , and $V_{i_1 \dots i_k}$ shows the variance of y with respect to the interaction of $\{x_{i_1}, \dots, x_{i_k}\}$. Eq. (11) can be normalized through division by $V(y)$. By doing so, the first-order or main-effect Sobol sensitivity index of an input x_i is defined as [34]:

$$S_i^1 = \frac{V_i}{V(y)} = \frac{V_{x_i}(E_{x_{-i}}(yx_i))}{V(y)} \quad (12)$$

where x_{-i} indicates all model inputs expect for x_i , $E()$ denotes the expected value, and S_i^1 is the sole contribution of x_i to the variance of y neglecting its interactions with other inputs. The rest of terms in Eq. (12), divided by $V(y)$, represent higher order indices measuring sensitivity to interactions of inputs. The overall contribution of a parameter x_i to the output variance, which includes both its individual effect and its interactions with the other input parameters, is described by the total-order or total-effect Sobol sensitivity index. It is defined as:

$$S_i^T = 1 - \frac{V_{-x_i}(E_{x_i}(y^{x_{-i}}))}{V(y)} \quad (13)$$

The first-order and total sensitivity indices of Eqs. (12) and (13) are typically estimated via Monte Carlo integration (see [41]), but the process might become computationally impractical if a single model evaluation is time-consuming. To solve this problem, Sobol sensitivity indices can be computed from PCEs, either analytically from the expansion coefficients (see Sudret, 2008), or by using PCE as a meta-model in Monte Carlo integration. This study employs the second approach by employing the python package ‘SALib’ [23].

4. Results and discussions

We assume that R_ϕ , L_e and N_b are the uncertain input variables subject to UPA and GSA. As common in literature (e.g. [49, 50, 56, 28]), we assume that the range of variability for R_ϕ is from 10 to 500, and L_e is from 1 to 5, using uniform probability distribution in both cases. Note that these values are physically plausible. For $N_b < -1$, the convective flow is mostly solute-driven (termed as case ‘SD’) while $N_b > -1$ typifies opposing flows that are largely heat-driven (called case ‘HD’). There is a sharp discontinuity in the temperature and concentration fields as well and in heat and mass fluxes between these cases. This discontinuity cannot be easily captured with the surrogate model based on polynomial interpolation. Thus, we split the rang of variability of N_b into two uniform distributions (represented by *unif*(...)) separated by a point of discontinuity at $N_b = -1$. These are *unif*[-5, -1) for ‘SD’ cases, and *unif*(-1, 0] for ‘HD’ cases. The uncertainties in these inputs can be attributed to the imperfect characterization of the saturating fluid and the porous domain (notably hydraulic conductivity, thermal conductivity of solid grains and porosity), as well as the validity of the mathematical model used to calculate the effective properties (average volume relation). Uniform distributions are often adopted to signify a general knowledge of the expected range of input parameters, and are preferred unless relevant information about parameters is available to justify other distribution types [16]. The uniform distribution is commonly used in sensitivity analysis when the main objective is to understand model behavior [39]. The assumption of uniform distributions is commonly used in applications where uncertainties are related to porous media properties ([14] and reference therein)

We consider all three uncertain parameters as ‘independent’ random variables. Other input variables (such as the size of the cavity H) are considered to be deterministic. The output QoIs in UPA and GSA are dimensionless temperature (T^*) and concentration (C^*); vertical and horizontal dimensionless velocity components (v_x^* and v_y^*) and their respective maximum values ($v_{x, max}^*$ and $v_{y, max}^*$); and average Nusselt (\overline{Nu}) and Sherwood (\overline{Sh}) numbers. The last two output QoIs are defined as [50]:

$$\overline{Nu} = \left| \int_0^1 \frac{\partial T^*}{\partial x^*} \Big|_{x^*=0} dy^* \right| \quad (14)$$

$$\overline{Sh} = \left| \int_0^1 \frac{\partial C^*}{\partial x^*} \Big|_{x^*=0} dy^* \right| \quad (15)$$

\overline{Nu} represents dimensionless thermal diffusive flux across the higher temperature wall, and \overline{Sh} denotes the solute diffusive flux at the higher concentration wall. Together these eight output variables provide a full quantitative picture of the fluid circulation and the heat and mass fluxes in the problem domain. In the subsequent sections, $v_{x, max}^*$, $v_{y, max}^*$, \overline{Nu} and \overline{Sh} are collectively called scalar output QoIs. Spatial variations of T^* , C^* , v_x^* and v_y^* provide information on the spatial variability of model output uncertainties.

Three issues are discussed in the results section. As PCEs are used as a meta-model to perform UPA and GSA, we first investigate their validity by comparing their results with those of the numerical model. Secondly, we present the outcome of UPA based on the results of the QoIs at mean values and the corresponding standard deviations. Finally, we discuss and provide some physical explanations of the GSA results which aims at identifying the key parameters affecting the QoIs. We also employ PCEs to analyze marginal effects of model parameters on the output QoIs.

4.1. Validation of PCE results

In this study, PCE meta-models are constructed by employing a training dataset of size $q = 150$, which is consistent with similar previous studies such as Fajraoui et al. [14]. An independent set of 50 sample points is used for model validation. These samples are obtained by employing an optimized Latin hypercube sampling (OLHS) strategy based on the centered L_2 -discrepancy (CLD) criterion and the enhanced stochastic evolutionary (ESE) optimization algorithm (see [42] for a brief description). The design of experiment for selection of regression points is performed using the R package ‘DiceDesign’.

Different orders of the polynomials are used for different output QoI. For each QoI, the order of polynomial (O_p) is chosen by comparing PCEs of different orders from $O_p = 0$ to 7, and the PCE order with the lowest normalized root mean square error (NRMSE) value for the validation dataset is chosen. Here, seven is the maximum value of O_p in which the number of PCE coefficients (which equals to $\binom{O_p + n_v}{n_v}$, where n_v is the number of variables [55]) remains below the number of training points ($n_r = 150$).

In order to pave the way for the use of PCE meta-models in UPA and GSA, the performance of the constructed PCEs has been assessed through the use of both statistical diagnostics and graphical methods. Table 2 presents NRMSE for the estimation of various scalar QoIs. NRMSE is estimates as follows:

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^n (Es_{num}(y_i) - Es_{PCE}(y_i))^2}{n_s}}}{\max(Es_{num}(y_i)) - \min(Es_{num}(y_i))} \quad (16)$$

where $Es_{num}(y_i)$ is the numerical model outputs for a specific QoI (denoted by y_i), $Es_{PCE}(y_i)$ is the respective PCE estimation, and n_s is the number of sample points. The NRMSE values in Table 2 are relatively low, indicating that the PCE meta-models provide satisfactorily accurate estimations of the input-output relationships. Furthermore, Fig. 2a–d shows scatter plots of Es_{num} versus Es_{PCE} for the validation dataset. Results shown in these figures pertain to the average Nusselt and Sher-

Table 2
PCE NRMSE error for the estimation of scalar QoIs.

| Case | \overline{Nu} | \overline{Sh} | $v_{x, max}$ | $v_{y, max}$ |
|------|-----------------|-----------------|--------------|--------------|
| SD | 0.082 | 0.129 | 0.031 | 0.010 |
| HD | 0.090 | 0.134 | 0.019 | 0.001 |

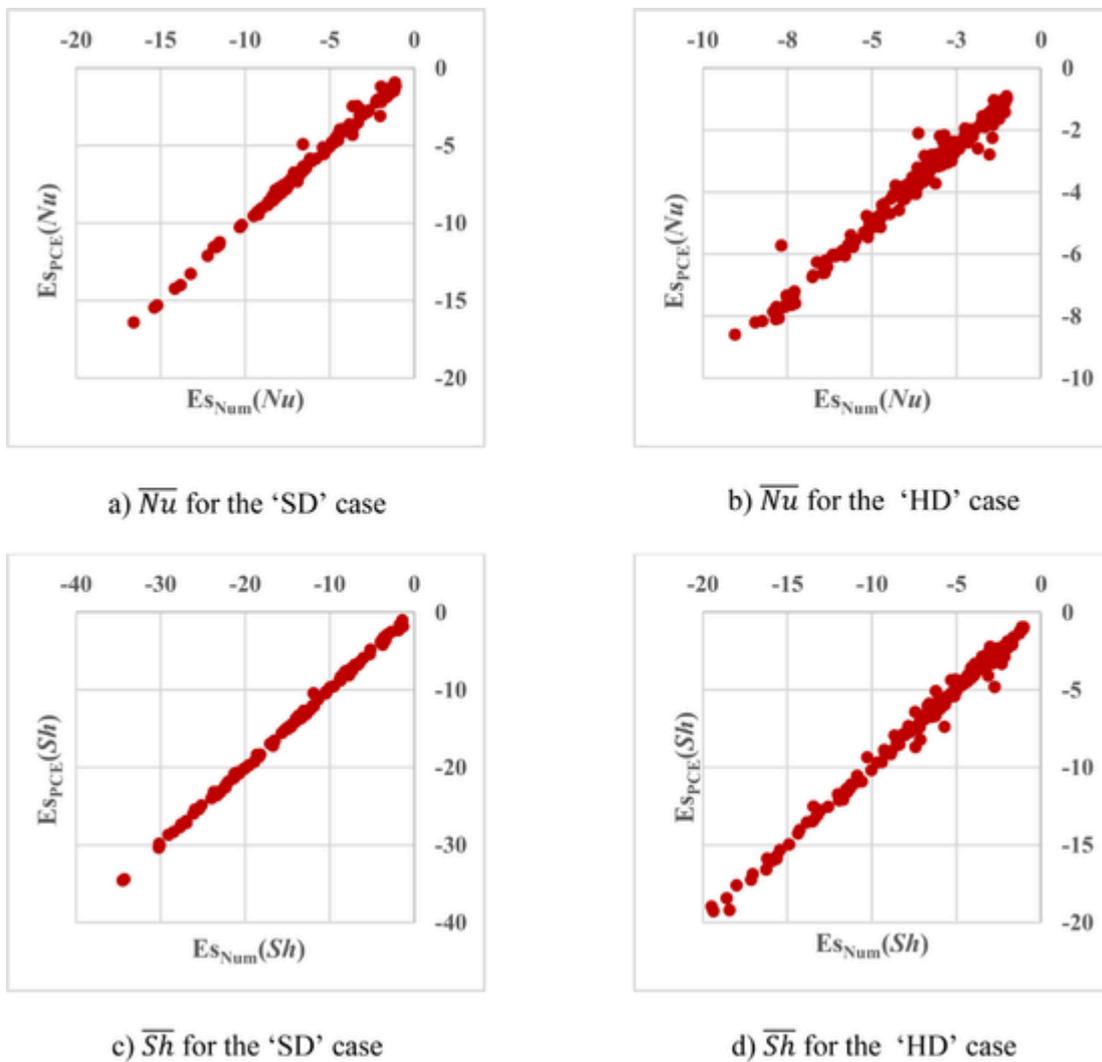


Fig. 2. Pairwise comparison of simulated outputs of the original numerical model and the PCE meta-models for (a, b) average Nusselt number and, (c, d) average Sherwood number. 'SD' represents the solute-driven case ($N_b \in [-5, -1]$) and 'HD' represents the heat driven case ($N_b \in (-1, 0]$).

wood numbers for cases 'SD' ($N_b < -1$) and 'HD' ($N_b > -1$). The clouds in all four sub-figures clearly look aligned with the bisector and are hence satisfactory. These scatter plots are presented as examples, and plots of other output QoI demonstrate similar behavior.

4.2. Uncertainty propagation analysis results

The verified PCE meta-models are subsequently used in MCS (with $n_{MC} = 1,000$ based on convergence analysis) to estimate the mean (μ_{PCE}), standard deviation (σ_{PCE}) and probability distribution function (PDF) of the output QoIs. The spatial distribution of mean temperature, concentration and velocity components are presented as Fig. S1 in the **Supplementary Material**. The results of mean values reflect the general behavior of double diffusive heat and mass transfer in a square porous cavity [20,49]. For temperature, they are in full agreement with the results obtained from stochastic simulations of Fajraoui et al. [14]. This further confirms the accuracy of our PCE meta-models. The PDFs of T^* , C^* , v_x^* and v_y^* in three exemplary observation points as well as for $v_{x,max}^*$, $v_{y,max}^*$, \overline{Nu} and \overline{Sh} are presented in Figs. S2 and S3 of the **Supplementary Material**. Some PDFs resemble a normal distribution and this is mostly the case in areas of the domain with low standard deviations of the output QoIs. But others are skewed to the right or left. None of the PDFs have multiple peaks, and they are all unimodal.

We present the results of the UPA analysis in two sub-sections. In Section 4.2.1, the standard deviations resulting from UPA are analyzed, and then in Section 4.2.2 the outcome of UPA is compared with those of deterministic simulations.

4.2.1. Standard deviations of the QoIs: spatial maps spatial maps of T^* , C^* , q_x^* and q_y^*

Fig. 3a and b shows the spatial distribution of standard deviation for dimensionless temperature estimates in the 'SD' and 'HD' cases, respectively. For both cases, we observe symmetry around the center point, with the lowest standard deviations pertaining to the thermal boundary layers at the side walls. The thermal boundary layer is a region where the fluid temperature is directly subjected to heating or cooling from the boundary wall, and the low standard deviations in this region are due to the assumption of constant (deterministic) temperature in the side walls. We also observe a zone with low standard deviation at the center of the flow eddy where fluid rotates at a relatively slow pace. As convective flow and thermal gradient are weak in this zone, heat transfer processes are insignificant. Thus, the temperature distribution is insensitive to model parameters. For the 'SD' case, the largest standard deviations are located around the top-left and bottom-right corners. In these zones, the convective flow and the thermal gradient exhibit their highest amplitudes (see Fig. S1 in **Supplementary**

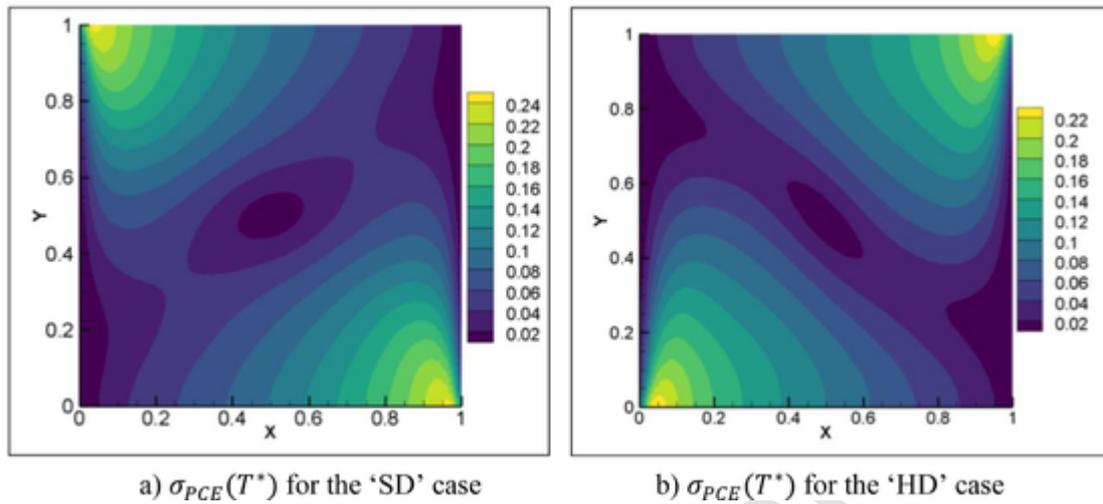


Fig. 3. Spatial distribution of the standard deviation for dimensionless temperature.

Material). Hence, uncertainties on model parameters can considerably affect the temperature distribution in these zones. Following the same logic, we can understand why the largest standard deviations are located around the bottom-left and top-right in the 'HD' case (Fig. 3b). The zone of high temperature variability is more spread in the 'HD' case than 'SD' case. As thermal conduction is more powerful than solute molecular diffusion, the solute-driven convective flow is more intense than thermal-driven flow. The results of the 'HD' case concerning the standard deviation of the temperature are in agreement with Fajraoui et al. [14]. When the convective flow is more intense, the thermal solute layers shrink and become more confined to the walls. This could explain the difference in the extent of high variability zones between the 'HD' and 'SD' cases.

All observations reported on the variability of the temperature are valid for the concentration distribution (Fig. 4a and b). This concerns the spatial distribution of the zones of high and low variability as well as the difference between the spatial distribution of the 'SD' and 'HD' cases. However, despite the similarities, the spatial distribution of standard deviations for concentration and temperature do not entirely coincide. Fig. 4b shows that in the 'HD' case, the zone of low concentration variability is mostly limited to the boundary layer toward the vertical walls, and contrary to temperature (Fig. 3b) the zone of low concentration variability at the center of the domain is relatively small. This indicates that in the 'HD' case, except for the boundary layers

around the vertical walls, the uncertainties on parameters affect the concentration distribution everywhere in the domain, but of course not at the same magnitude.

Concerning the horizontal velocity component (v_x^*), as shown in Fig. 5a and b, the zones of high standard deviation are located along the horizontal walls. The highest variability can be observed at the top-left and bottom-right corners for the 'SD' case and at the opposite corners for the 'HD' case. This is understandable as horizontal flow occurs mainly along the horizontal walls. For the 'SD' case, the convective flow is upward along the left wall. Due to the boundary conditions and mass conservation, v_x^* is well correlated with the upward convective flow which is considerably sensitive to the input parameters. In the same manner, we can explain the high variability of (v_x^*) at the bottom-right corner. It is also worth mentioning, that as for concentration and temperature, the zone of high variability of v_x^* is more spread in the 'HD' case than in 'SD' case. As explained in the previous sub-section, this is related to the difference between molecular diffusion and thermal conduction that yields a solute-driven convective flow that is more intense than the heat-driven flow. Same observations can be reported for the variability of the vertical velocity component (v_y^*) along the vertical walls (Fig. 6a and b).

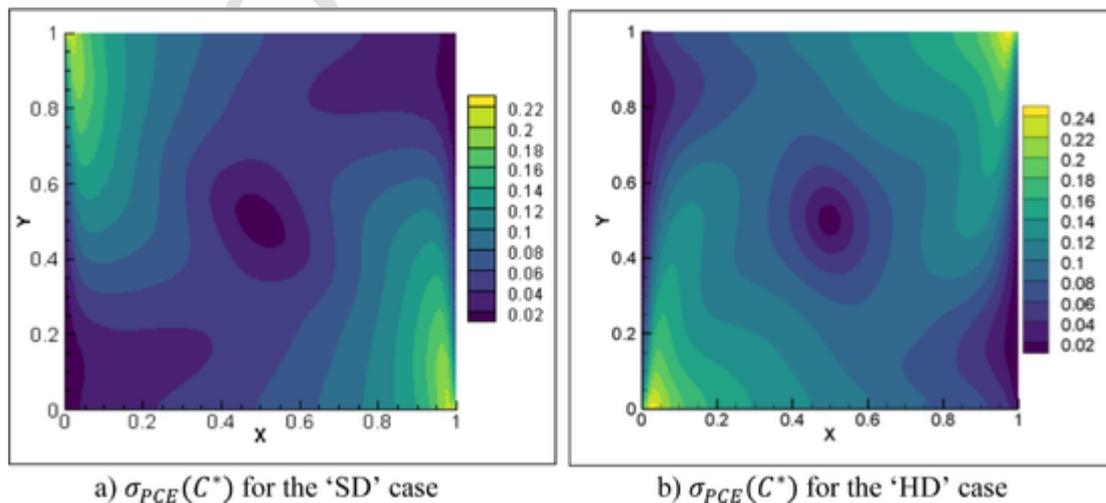


Fig. 4. Spatial distribution of the standard deviation for dimensionless concentration.

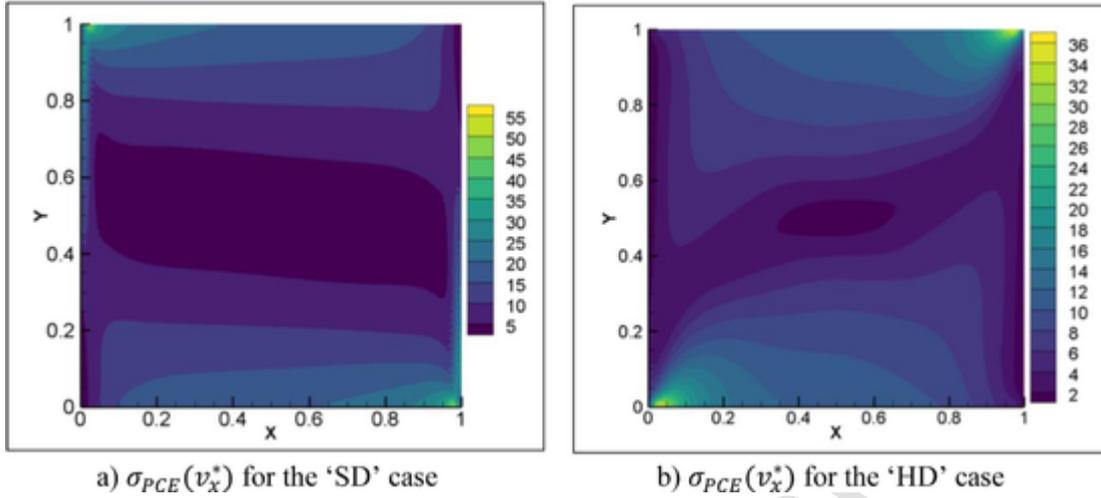


Fig. 5. Spatial distribution of the standard deviation for non-dimensional horizontal velocity.

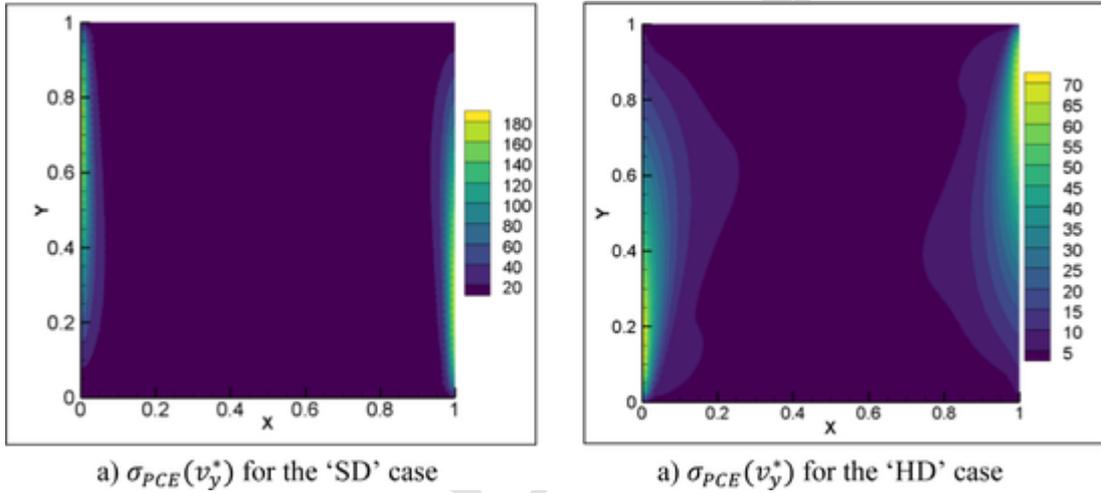


Fig. 6. Spatial distribution of the standard deviation for non-dimensional vertical velocity.

4.2.2. Comparison of deterministic and stochastic outputs

For a deeper understanding of uncertainty propagation, we employ a similar approach as in Shahane et al. [48] to investigate the difference between deterministic model outputs and the stochastic mean estimates ($D_{PCE}(y_i) - \mu_{PCE}(y_i)$). This difference is termed as DDSM (difference between deterministic output and stochastic mean). It is calculated based on two independent estimates of the same output (QoI), both obtained using PCEs. For the deterministic case, PCEs are used to obtain single estimates of the QoIs using the deterministic values of the inputs ($R_a = 255, L_e = 3$ and $N_b = -3$ for the 'SD' case and $N_b = -0.5$ for the 'HD' case). For the stochastic mean, PCEs are employed in the framework of MCS as previously described. Hence, DDSM is meant to provide an estimate of the bias introduced as a result of assuming deterministic values for parameters that are inherently uncertain.

Spatial maps of DDSM are given in Fig. 7. As demonstrated, regions of highest DDSM coincide with regions of highest standard deviation for all four variables T^*, C^*, v_x^* and v_y^* ; which makes sense as these zones are considerably affected by the uncertainties in parameters. It is observed that the DDSM values are on the same order, but generally less than the respective stochastic standard deviations. The general patterns of spatial variability for DDSM in our study are in agreement with

those for natural convection in bulk fluids developed by Shahane et al. [48].

While standard deviation distribution indicates the zones of high variability, DDSM distribution allows for identifying where a QoI can be overestimated (positive values in Fig. 7) or underestimated (negative values in Fig. 7) in a deterministic approach. For example, Fig. 7a shows that the use of deterministic values without considering the uncertainties on parameters leads to an underestimation of the temperature at the top part of the domain and overestimation in the bottom part, in the 'SD' case. The opposite is true in the 'HD' case (Fig. 7b), but the corresponding zones are shaped diagonally. The same can be observed for concentration (Fig. 7c and d). More complex DDSM distributions can be seen for v_x^* and v_y^* (Fig. 7e-h).

Table 3 presents the mean and standard deviations of the scalar QoIs, along with their respective deterministic simulation outputs. Comparison of mean values resulting from UPA with the deterministic outputs, shows that a deterministic approach creates 4% to 16% bias in the estimation of the scalar QoIs. The table also shows relatively high standard deviations for the QoIs, with coefficients of variation ($= |\sigma/\mu| \times 100$) equal to 50–85%. With the deterministic parameters, the model would overestimate heat and mass fluxes, as respectively estimated by $\overline{Nu} \cdot \overline{Sh}$, as well as the maximum horizontal velocity ($v_{x,max}^*$) in both 'SD' and 'HD' cases. However, by neglecting uncertainties, the

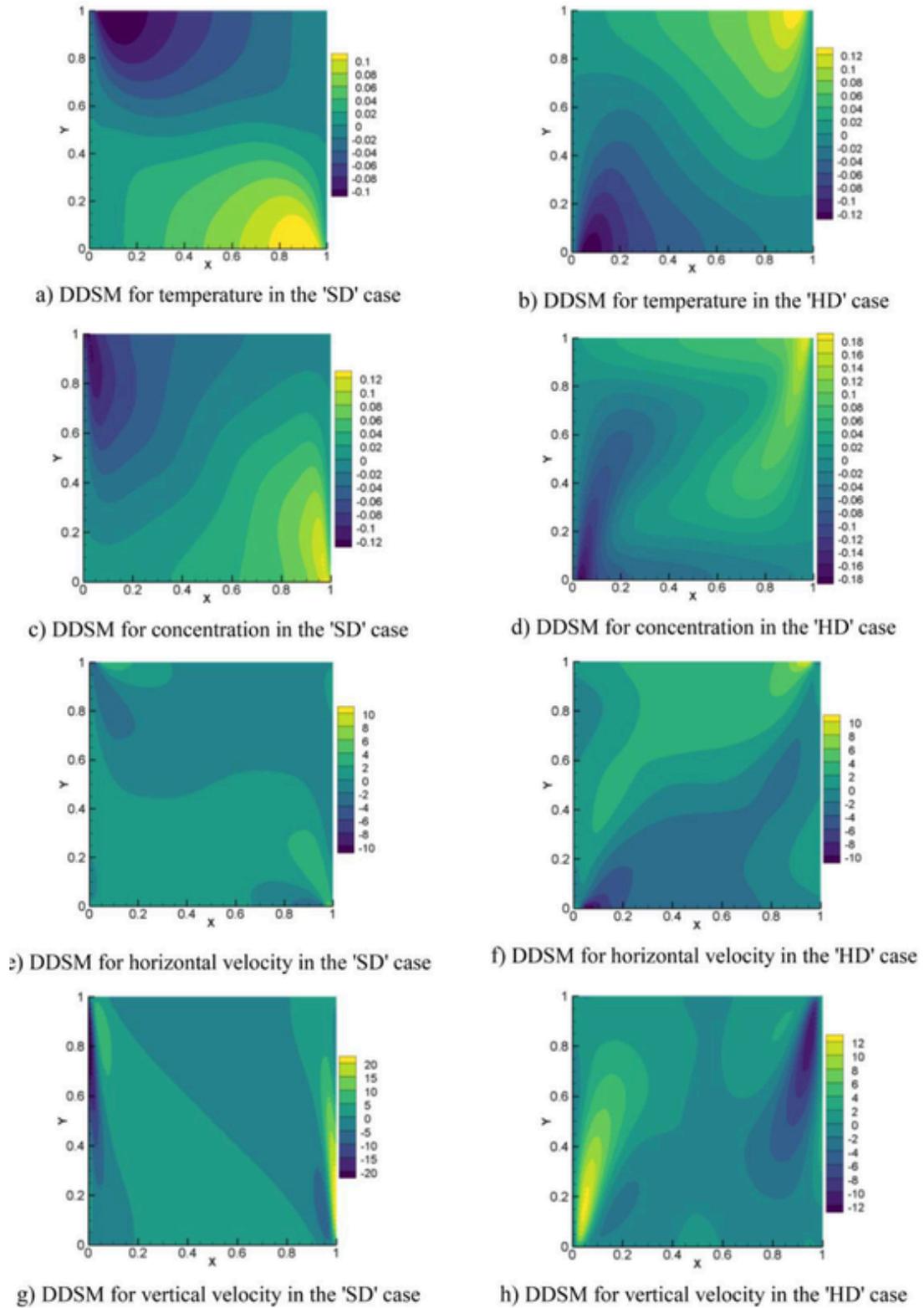


Fig. 7. Difference between deterministic and stochastic mean (DDSM) for temperature (a, b), concentration (c, d), horizontal velocity (e, f) and vertical velocity (g, h).

Table 3
Comparison of deterministic and stochastic outputs.

| Case | \overline{Nu} | | \overline{Sh} | | $v_{x,max}$ | | $v_{y,max}$ | | | | | |
|------|-----------------|----------|-----------------|----------|-------------|----------|-------------|----------|--------|--------|--------|--------|
| | stoch | det | Stoch | det | Stoch | det | Stoch | det | | | | |
| | μ | σ | μ | σ | μ | σ | μ | σ | | | | |
| SD | 5.12 | 3.87 | 5.81 | 14.00 | 7.46 | 15.9 | 95.93 | 57.35 | 102.47 | 269.86 | 191.68 | 289.72 |
| HD | 4.13 | 2.09 | 4.32 | 7.19 | 4.63 | 8.31 | 50.68 | 32.54 | 54.56 | 82.14 | 70.07 | 70.68 |

stoch: stochastic model outputs.

det: deterministic model outputs.

model overestimates the maximum vertical velocity ($v_{y,max}^*$) in the 'SD' case and underestimate it in the HD 'case'.

4.3. Sensitivity analysis results

In this sub-section, we present and analyze the results of GSA with the intention of understanding how the uncertainty in the outputs of the model can be allocated to different sources of uncertainty in model inputs.

4.3.1. What are the most influential parameters that affect T^* , C^* , v_x^* and v_y^* ?

Fig. 8 shows the boxplots of first-order and total Sobol indices for T^* , C^* , v_x^* and v_y^* with respect to R_a , L_e and N_b . In these boxplots, the central mark indicates the median value of Sobol indices for all model elements, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme values of the Sobol indices. Based on the median value and the interquartile range (IQR) of the first-order and total Sobol indices in Fig. 8a and b, the most influential parameter on temperature is N_b , and temperature is clearly more sensitive to R_a than L_e in both the 'SD' and 'HD' cases. In both cases, N_b is also the top contributing parameter to the concentration field (see Fig. 8c and d). For the 'SD' case, the concentration distribution is more sensitive to R_a than L_e . For the 'HD' case, as shown in Fig. 8d, the total Sobol indices show that L_e is more important than R_a while the first-order Sobol indices for the two parameters are similar. This means that for the 'HD' case, the high sensitivity to L_e is related to interactions with other parameters. For horizontal velocity (v_x^*) in the 'SD' case (Fig. 8e), N_b is the most influential parameter, followed by R_a and then L_e . But in the 'HD' case, horizontal velocities are most sensitive to R_a , followed by N_b and L_e (Fig. 8f). Vertical velocities are highly sensitive to both N_b and L_e in the 'SD' case, and in contrast to horizontal velocity, v_y^* is more sensitive to L_e than R_a . For the 'HD' case, N_b and R_a are the most important parameter influencing v_y^* (Fig. 8h).

4.3.2. The spatial distribution of sensitivity indices

For the sake of brevity, we only present the maps of the total Sobol indices in the paper, and the respective maps of the first-order indices are provided in the **Supplementary Material**.

Figs. 9 and 10 show the spatial maps of total Sobol indices with respect to uncertainty in R_a , L_e and N_b for temperature (T^*) and concentration (C^*) estimations, respectively. For the 'SD' case, the zones of high sensitivity of T^* to N_b are located near the top-right and bottom-left corners (Fig. 9e). In these zones the locations of the isotherms are highly affected by the convective flow which is mainly governed by N_b . At the opposite corners (top-left and bottom-right), the isotherms are close to each other and their locations are impacted by the boundary conditions. The zones of high sensitivity of T^* to R_a and L_e have lemniscate shapes located at the center of the domain (Fig. 9a and c). Along

the vertical walls, the heat transfer by convection is upward, thus the horizontal penetration of isotherms is mainly related to thermal diffusion which is represented by R_a and L_e . As for the temperature field, the zones of high sensitivity of C^* to N_b are located at the top-right and bottom-left corners where the positions of the concentration contours are mainly controlled by the convective flow (Fig. 10e). However, for the concentration field, a lemniscate-shaped zone of high sensitivity appears at the center of the domain. High sensitivity of C^* to R_a and L_e are observed in the corners and in lemniscate-shaped zones at the center (Fig. 10a and c).

For the 'HD' case, the zones of high sensitivity of T^* to L_e and N_b are the horizontal reverse of those for the 'SD' case (Fig. 9d and f), which makes sense due to the reversal of the flow direction. For the sensitivity of T^* to R_a , the spatial pattern in the 'HD' case are completely different from the 'SD' case and the zones of high sensitivity appear along the vertical walls relatively far from the corners (Fig. 9b). For C^* , the zones of high sensitivity to N_b are the horizontally reverse of those for the 'SD' case due to flow reversal from counter-clockwise to clockwise (Fig. 10f). Sensitivity maps of C^* to R_a and L_e have different patterns from those in the 'SD' case (Fig. 10a and d).

It's noteworthy that the spatial distribution of the Sobol indices for T^* with respect to R_a for the 'HD' case of DDC are not similar to the ones for pure thermal convective. Comparison of Fig. 9b with the corresponding results in Fajraoui et al. [14] demonstrates that the zones of high sensitivity do not coincide. This shows that even if convective processes are dominated by heat gradient, temperature distribution is still impacted by the mass transport processes.

Figs. 11 and 12 illustrate the maps of total Sobol indices for the horizontal and vertical components of the velocity field. For v_x^* and in the 'SD' case, significant sensitivity to N_b can be observed along the top and bottom surfaces, which is logical, as the velocity is mainly horizontal in these parts of the domain (Fig. 11e). However, the zone of highest sensitivity of v_x^* to N_b is shaped like a lemniscate at the center of the domain away from the boundary layer, where the flow is vertical. In the 'HD' case, highest sensitivity of v_x^* to N_b is observed around the top-left and bottom-right corners (Fig. 11f), and a sharp zone of significant sensitivity expands along to the first bisector. In both the 'SD' and 'HD' cases, the sensitivity maps of v_x^* to R_a and N_b are complementary with each other. Sensitivity of v_x^* to L_e is limited to the center of the domain in the 'HD' case (Fig. 11d).

For v_y^* , the zone of highest sensitivity to N_b is observed at the center of the domain, while significant sensitivity can also be observed along the vertical walls. This is true for both the 'SD' and 'HD' cases (Fig. 12e and f). In the 'SD' case, the highest values of $S_{L_e}^T(v_y^*)$ are expanded along the vertical walls outside the vertical boundary layers (Figs 11c and d). High sensitivity of v_y^* to R_a can be observed near the vertical walls. But the highest zone of sensitivity to R_a is shaped like a lemniscate expanded along the diagonal.

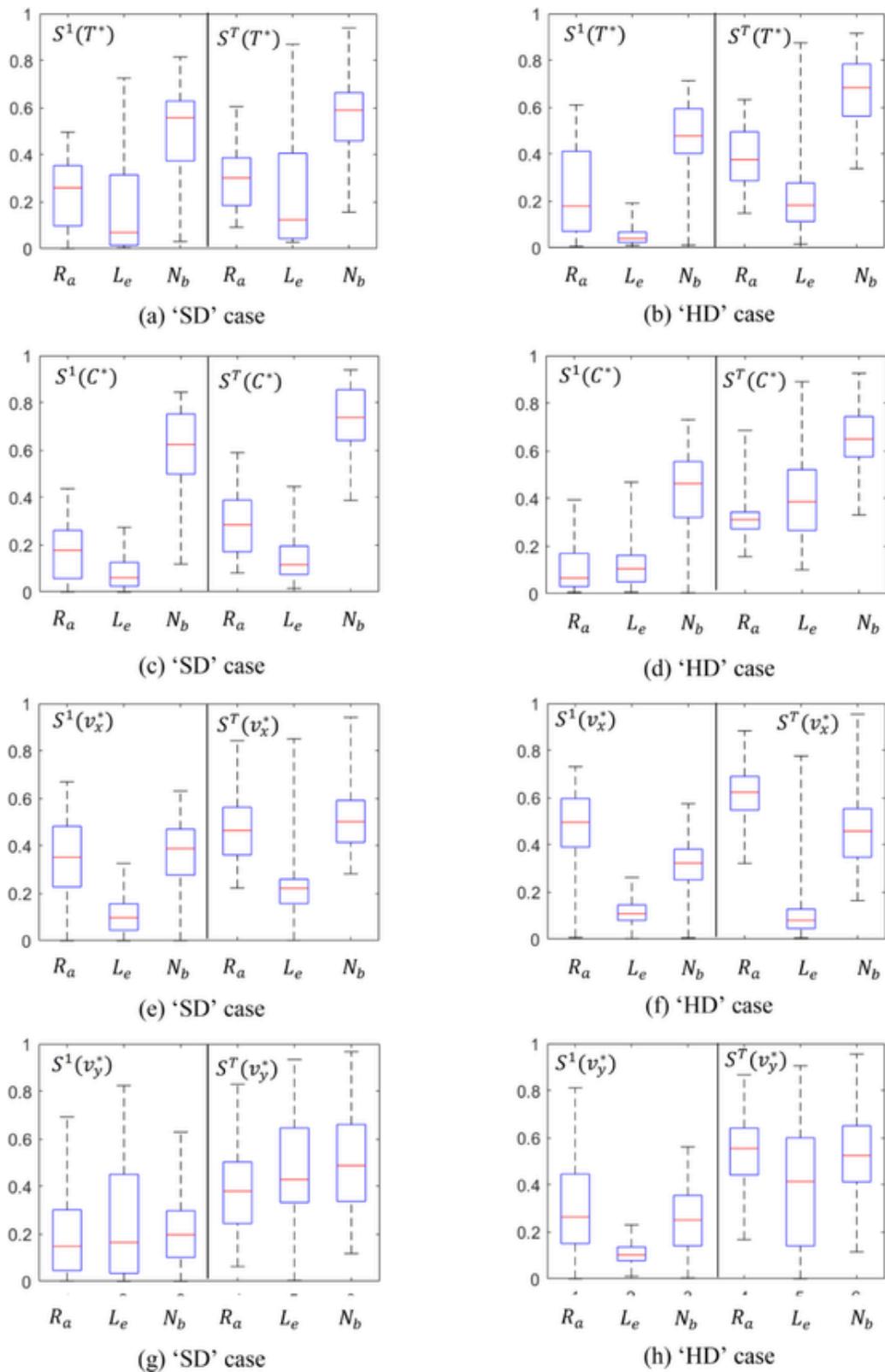


Fig. 8. Boxplots of first-order and total Sobol indices for (a, b) temperature and (c, d) concentration, (e, f) horizontal and (g, h) vertical velocity.

4.3.3. First-order vs. total sobol indices

Maps of the first-order Sobol indices (presented in the **Supplementary Material**) and those of the total Sobol indices show very similar patterns in the majority of cases. Some discrepancy between the two

maps is found for the sensitivity of C^* to R_a . While the highest total Sobol indices ($S_{Ra}^T(C^*)$) are located within a lemniscate zone at center of the domain, the highest first-order Sobol indices ($S_{Ra}^1(C^*)$) are located near the top-left and bottom-right corners. There are also discrepancies between the two maps for the sensitivity of v_x^* and v_y^* to L_e in the

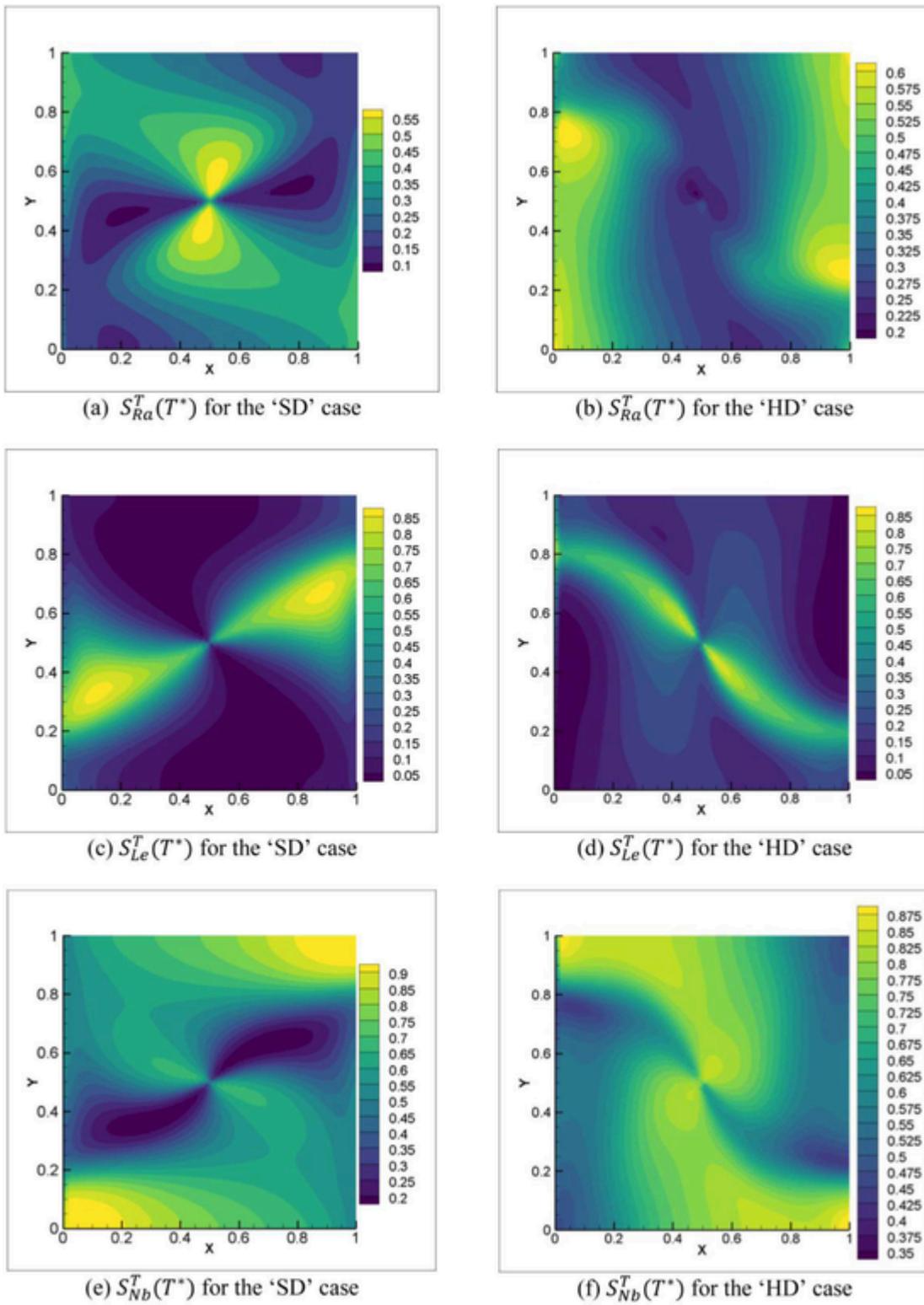


Fig. 9. Spatial distribution of total Sobol sensitivity indices for temperature.

'HD' case. In these instances, interactions between parameters do not affect only the magnitude of sensitivity, but also its spatial distribution.

As expected, the absolute values of the total indices are higher than those of the first order indices, indicating significant interaction effects between parameters. Table 4 shows the percent increase in the spatial average of total Sobol indices, compared to the associated first-order Sobol indices. As demonstrated, the differences between the two in-

stances is generally more pronounced in the 'HD' case. We observe the biggest differences for Sobol indices of L_e in both the 'SD' and 'HD' cases, except for C^* and v_y^* in 'SD' where the biggest differences pertain to R_a and N_b respectively.

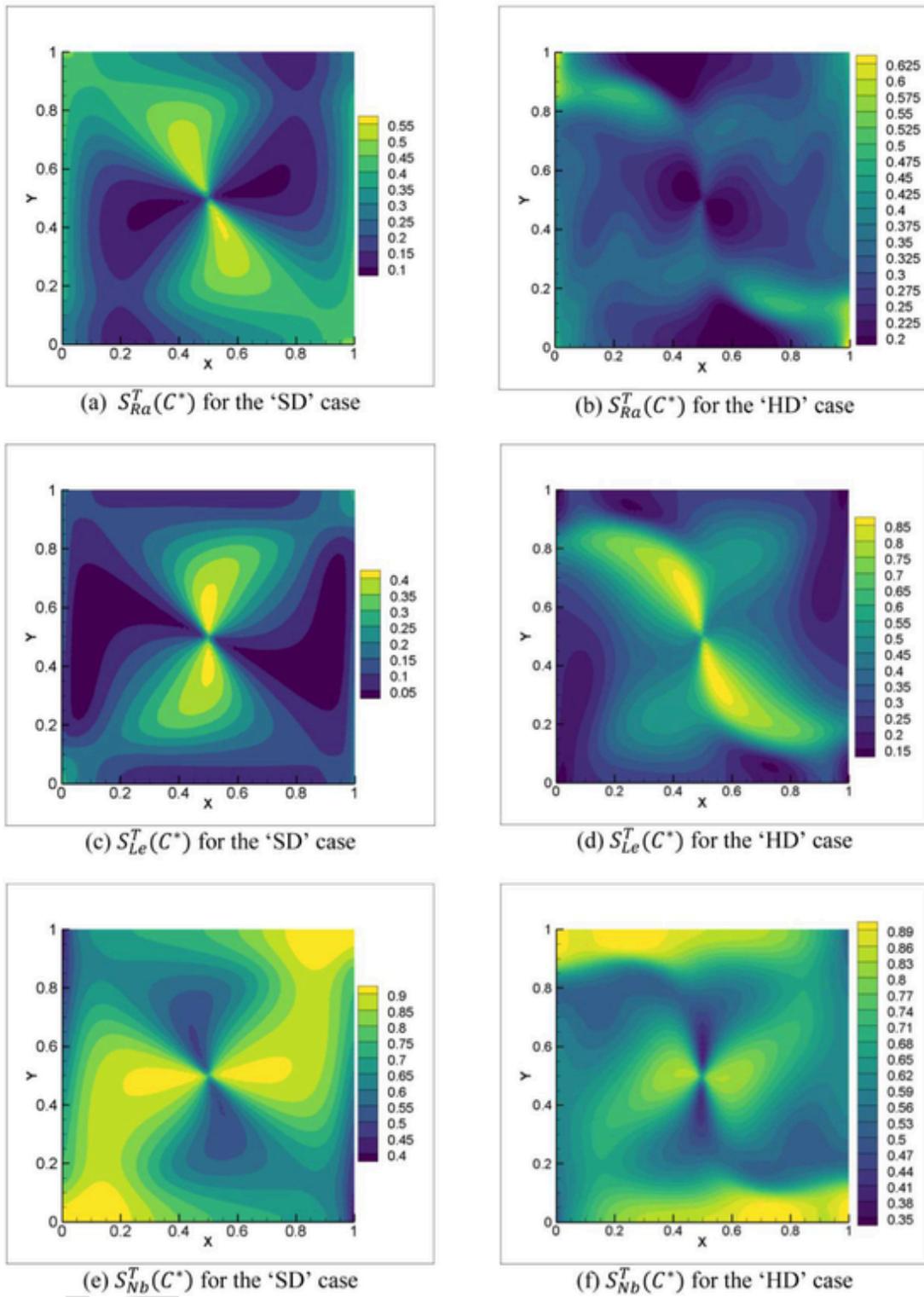


Fig. 10. Spatial distribution of total Sobol sensitivity indices for concentration.

4.3.4. Sensitivity indices for the scalar QoI

Fig. 13 illustrates the total and first-order Sobol indices for the scalar QoI as bar-plots. Fig. 13a shows that the variability of the Nusselt number is mainly related to R_a and then to N_b , for both the 'SD' and 'HD' cases. However, in contrast to temperature distribution, R_a is more influential than N_b . For the 'SD' case, the results in Fig. 13a are

consistent with the previous results regarding the spatial maps of the Sobol indices (Fig. 9a) and standard deviation (Fig. 3c). In fact, \overline{Nu} represents the diffusive heat flux at the hot wall. Thus, it is related to the temperature distribution near this wall. As shown in Fig. 9a, the zones in the vicinity of the hot wall, where the temperature field is highly sensitive to N_b and L_e , are located in the bottom part of the domain where the variability of temperature field is insignificant. How-

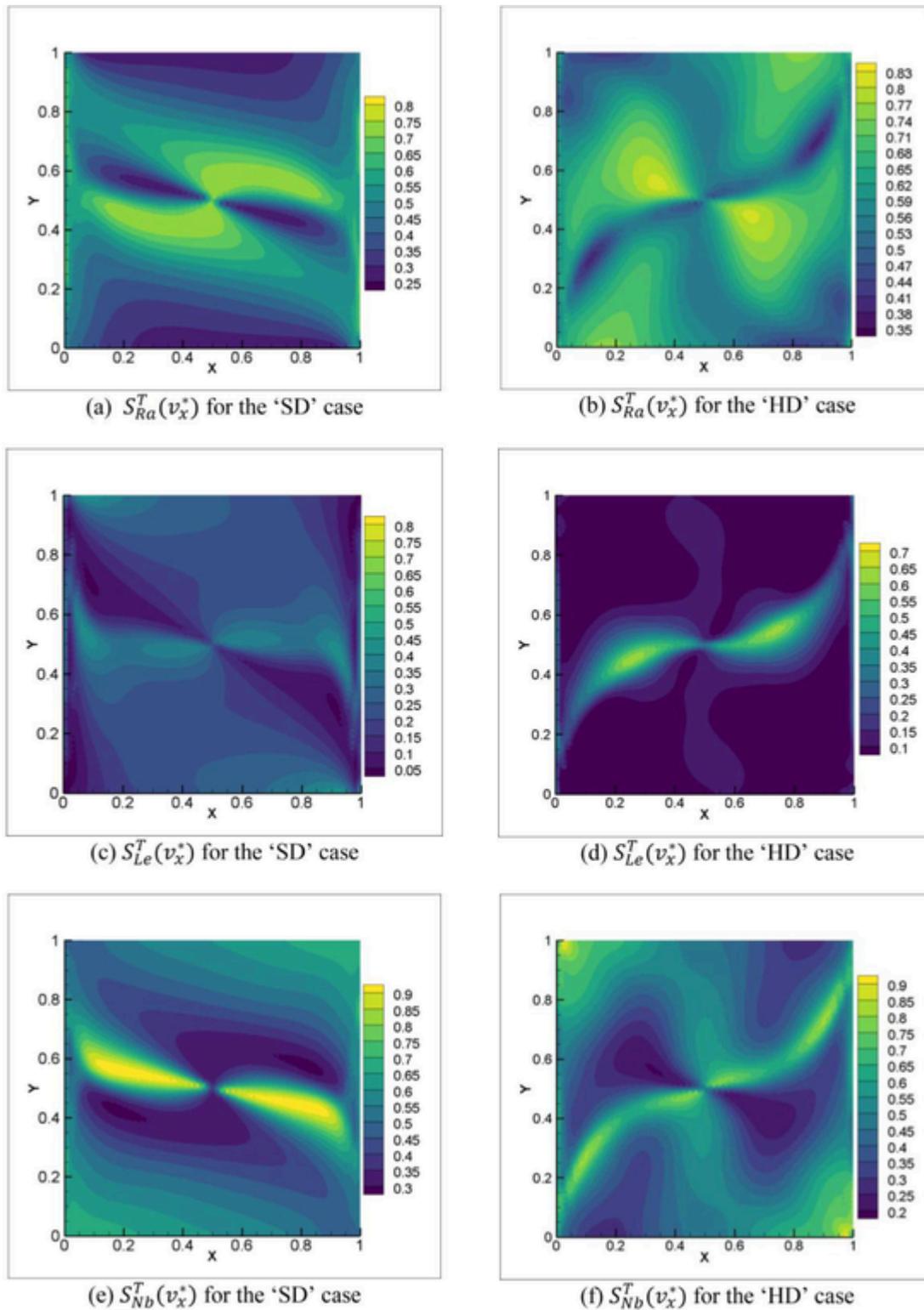


Fig. 11. Spatial distribution of total Sobol sensitivity indices for horizontal velocity.

ever, the zone of high sensitivity to R_a , toward the hot wall, is located in the upper part of the domain where temperature exhibits high variability. This explains the high sensitivity of \overline{Nu} to R_a . The same logic can be followed to explain the variation of the Sobol indices in the 'HD' case (Fig. 13b) using Figs. 9b and 3d. Fig. 13a shows slight interactions between parameters as first-order and total Sobol indices are almost close.

For \overline{Sh} , the most influential parameter is N_b , followed by R_a and L_e . This is also in full agreement with the sensitivity analysis of the concentration distribution and its standard deviation as in Figs. 10a and 4c (for the 'SD' case) and Figs. 10b and 4d (for the 'HD' case). Fig. 13 shows that \overline{Nu} is more sensitive to R_a than \overline{Sh} while the latter is more sensitive to N_b than \overline{Nu} . Interactions between parameters for \overline{Sh}

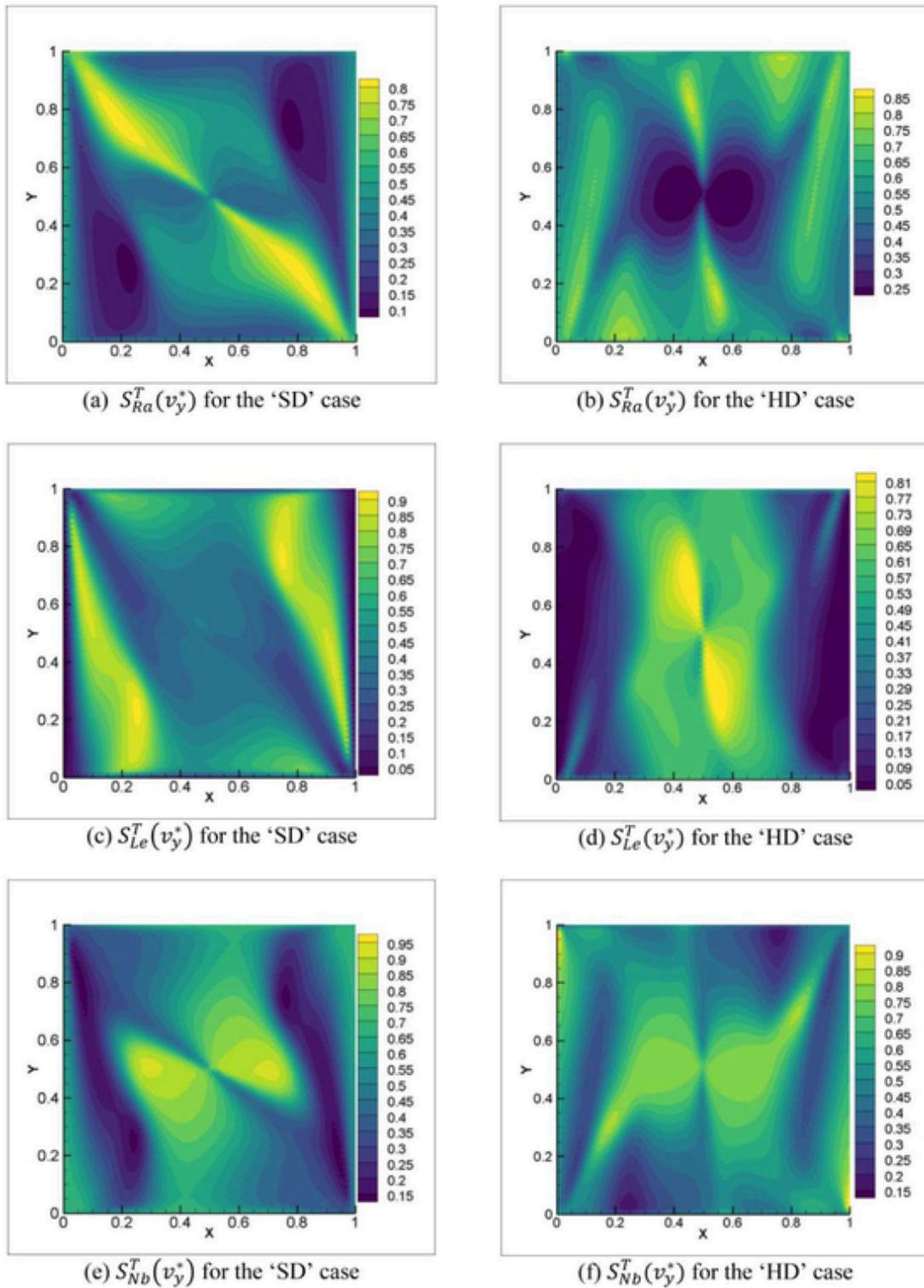


Fig. 12. Spatial distribution of total Sobol sensitivity indices for vertical velocity.

is smaller than \overline{Nu} . The maximum velocity components ($v_{x,max}^*$ and $v_{y,max}^*$) behave in a similar fashion regarding parameters (Fig. 13e–h). For both $v_{x,max}^*$ and $v_{y,max}^*$, the most important parameters are R_a and then N_b , which makes sense as these parameters control the convective flow. L_e is almost a non-influential parameter. More interactions between parameters can be observed for $v_{x,max}^*$ and $v_{y,max}^*$ than \overline{Nu} and \overline{Sh} .

4.3.5. The marginal effects

We investigate the marginal effects to understand the impacts of each input parameter on the scalar QoIs. The marginal effects are calculated using the PCEs, by varying each input parameter within its range of uncertainty in a stepwise manner while keeping the other two parameters constant at their deterministic values. The results are de-

Table 4
Percent increase in the spatial average of total Sobol indices, compared to the associated first-order Sobol indices.

| QoI | Solute-driven flow ('SD' case) | | | Heat-driven flow ('HD' case) | | |
|---------|--------------------------------|-------|-------|------------------------------|-------|-------|
| | R_a | L_e | N_b | R_a | L_e | N_b |
| T^* | 27.9 | 35.6 | 13.4 | 67.2 | 664.9 | 45.5 |
| C^* | 73.8 | 69.5 | 22.5 | 211.7 | 263.5 | 56.9 |
| v_x^* | 32.3 | 105.7 | 44.0 | 31.2 | 300.1 | 48.6 |
| v_y^* | 97.6 | 84.0 | 135.0 | 81.1 | 628.7 | 109.5 |

Biggest increase in each case.

pictured in Fig. 14. Each sub-plot of Fig. 14 is based on 1000 PCE simulations. Evident similarities can be observed in the behavior of \overline{Nu} and \overline{Sh} with respect to R_a (Fig. 14a and b). But it is clear that \overline{Sh} is higher than \overline{Nu} , which is expected as thermal conduction is larger than molecular diffusion ($L_e > 1$). Increasing variations of \overline{Nu} and \overline{Sh} with respect to R_a are expected (for both 'SD' and 'HD' cases), as the increase of this parameter is accompanied by the intensification of the convective flow. In consequence, the thicknesses of the thermal and solute boundary layers reduce and the temperature and concentration gradients near the vertical walls increase which lead to higher rates of heat and mass transfer. The intensification of the convective flow also explains the increasing variation of $v_{x,max}^*$ and $v_{y,max}^*$ with increasing R_a (Fig. 14c and d). Different behaviors are observed in the 'SD' and 'HD' cases for the marginal effects of the QoIs due to variation of L_e which measures the relative importance of mass to thermal diffusion. For the 'SD' case, Fig. 14e shows that \overline{Nu} decreases with the increase of L_e . To explain this behavior, we interpret the increase of L_e as an increase of the thermal diffusivity. This would improve the heat transfer by diffusion and decreases \overline{Nu} . When L_e is high enough, heat transfer occurs mainly by diffusion, thus \overline{Nu} becomes constant. Fig. 14e indicates that, in the 'HD' case, \overline{Nu} is slightly increasing as L_e is increased. In both 'SD' and 'HD' cases, increasing variation of \overline{Sh} with respect to L_e can be seen in Fig. 14f. The increase of L_e can be interpreted as a decrease of the molecular diffusion coefficient. The mass transfer flux becomes convection dominated, which explains the increase of \overline{Sh} . Fig 14e and 1f shows that, unlike the mass flux, heat flux to the cavity does not increase as L_e is increased. It is clear from these figures that at $L_e = 1$, \overline{Nu} and \overline{Sh} are equal, given the similarity of thermal and solute diffusion processes. In general, the behaviors of \overline{Nu} and \overline{Sh} due to the variation of L_e are in line with the results reported in Al-Amiri and Khanafer [1] and Sankar et al. [47]. Fig. 14g confirms that $v_{x,max}^*$ decreases with the increase of L_e in the 'SD' case, and decreases in the 'HD' case. $v_{y,max}^*$ is almost insensitive to L_e (Fig. 14f).

Fig. 14i depicts the marginal effect of N_b on \overline{Nu} . In the 'HD' case, \overline{Nu} decreases as N_b is increased. The increase of N_b intensifies the solute convective flow which occurs in the opposite direction of the main thermally-driven convective flow. This weakens the overall convective flow and decreases the rate of heat transfer to the domain. For the 'SD' case, the intensification of the solute-dominated convective flow increases the rate of heat transfer. This explains the increase of \overline{Nu} with N_b . Similar behavior can be depicted for \overline{Sh} (Fig. 14j). The same logic can be used to explain the variation of $v_{x,max}^*$ and $v_{y,max}^*$ with respect to N_b (Fig. 14k and l). Unlike the 'SD' case, in the 'HD' case, these figures indicate flow deceleration related to the enhancement of the solute-driven flow that opposes the main flow.

5. Summary and conclusions

We consider DDC in a porous square cavity and we investigate the effects of uncertainties related to physical parameters on model outputs characterizing flow and heat and mass transfer processes. Darcy's law is used to describe fluid flow in the porous media. The hydrodynamics and thermo-physical parameters can be regrouped in three dimensionless parameters which are the thermal Rayleigh number (R_a), the buoyancy ratio (N_b) and the Lewis number (L_e). These parameters are assumed to be uncertain, and UPA and GSA are performed to understand how these uncertainties would propagate through the model and affect temperature, concentration and velocity field, as well as maximum velocity components and heat and mass fluxes as represented by the average Nusselt and Sherwood numbers, respectively.

Knowing that DDC simulations are computationally expensive and that UPA and GSA require a large number of simulations, an appropriate strategy is developed to perform UPA and GSA in an efficient manner. The problem of DDC in porous square cavity is simulated using an efficient and accurate simulator based on the Fourier series method that allows for highly accurate solutions with a reduced number of degree of freedom. Numerical simulations are used to train meta-models based on PCEs. The accuracy of the PCE surrogate models is investigated by comparing their results with those of the numerical model based on a set of simulations independent from the ones used to train the meta-model. Monte Carlo-based UPA is performed using the PCE meta-model to deal with high computational cost of DDC numerical simulations. GSA is performed using a variance-based technique via the Sobol indices using the PCE meta-models. Due to its efficiency, the developed strategy allows for investigating UPA and GSA not only for scalar average metrics but also for the spatially variable model output uncertainties and sensitivity indices.

The results of the UPA show that temperature and concentration fields are slightly impacted by uncertainties within the thermal and solute boundary layers. The zones of high temperature and concentration variability are located in the regions where the flow is mainly driven by the buoyancy effects. The zones where both temperature and concentration are highly affected by uncertainties are more spread in the thermally driven case ('HD') than the solute driven case ('SD'). The uncertainties in the parameters lead to high variability of the velocity fields in the vicinity of the walls. Vertical and horizontal velocity components are particularly affected by uncertainties in the vicinity of the vertical and horizontal walls respectively.

We estimated the bias introduced by replacing uncertain parameters by deterministic values, and presented them as spatial maps of the difference between deterministic outputs and stochastic means. The results show that the use of deterministic values leads to different zones where the temperature, concentration and velocity fields can be either overestimated or underestimated. For instance, in the 'SD' case, the temperature field is underestimated in the upper part of the domain close to the hot wall and overestimated in the lower part close to the cold wall. The opposite is true for the 'HD' case. The same behavior is observed for the concentration field. With the deterministic parameters, the model overestimates the maximum horizontal velocity and the average heat and mass fluxes into the domain. It overestimates the maximum vertical velocity in the 'SD' case and underestimates it in the 'HD' case.

GSA is used to identify the key parameters responsible for uncertainty in a certain model output. It is also employed to rank parameters according to their significance. It is clear that N_b is the most influential parameter affecting the temperature field. N_b is followed respectively by R_a and L_e . N_b is also the most important parameter controlling the concentration field. As for temperature, it is followed by R_a and L_e in the 'SD' case. For the 'HD' case, L_e is more important than R_a . The zones where the temperature field is highly sensitivity to N_b are located

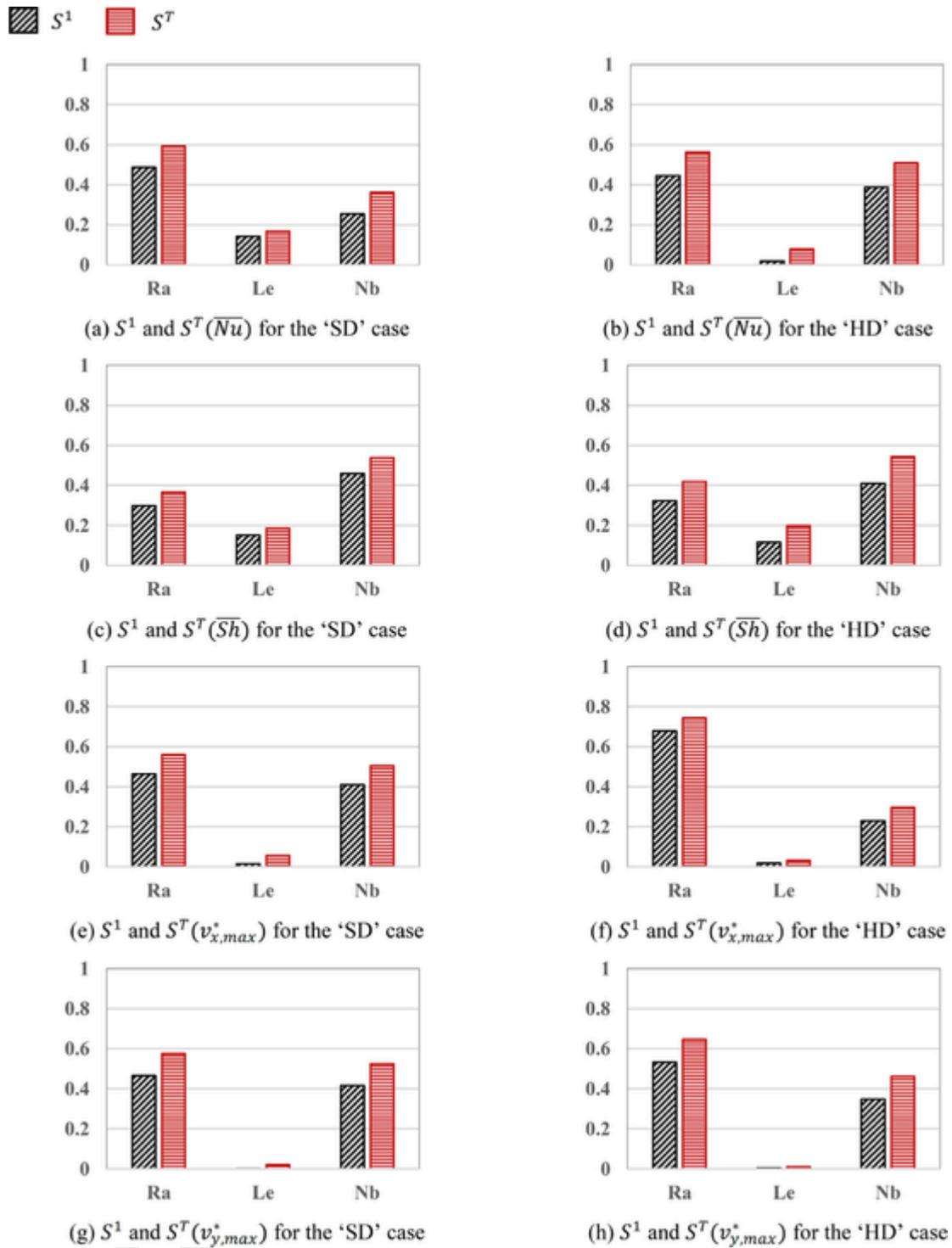


Fig. 13. First and total order Sobol sensitivity indices for the scalar output QoIs.

in top and bottom corners toward the cold and hot walls, respectively. For the concentration field, the zone of high sensitivity expands to cover a lemniscate-shaped zone at the center of the domain. The horizontal velocity component is almost equally influenced by N_b and R_a , in the 'SD' case. For the 'HD' case, R_a is the most significant parameter followed respectively by N_b and L_e . The variability of the vertical velocity is mainly attributed to N_b and L_e in the 'SD' case and N_b and R_a in the 'HD' case. The average heat flux to the domain is mainly controlled first by R_a and then N_b and it is slightly sensitive to L_e . The top con-

tributing parameter to the average mass flux is N_b and the most influential parameters for the maximum velocity components is R_a .

The conclusions drawn in this work are based on the problem of DDC in porous square cavity but are likely to be helpful in different applications involving DDC in porous enclosures leading to convective circulation cells. The results point out the capacity of the methodology based on the PCE in addressing highly nonlinear problem such as DDC. They highlight the benefits of GSA and UPA that could be helpful in risk assessment studies and safe designing procedures. The results are also helpful in domain characterization from the measurement of state

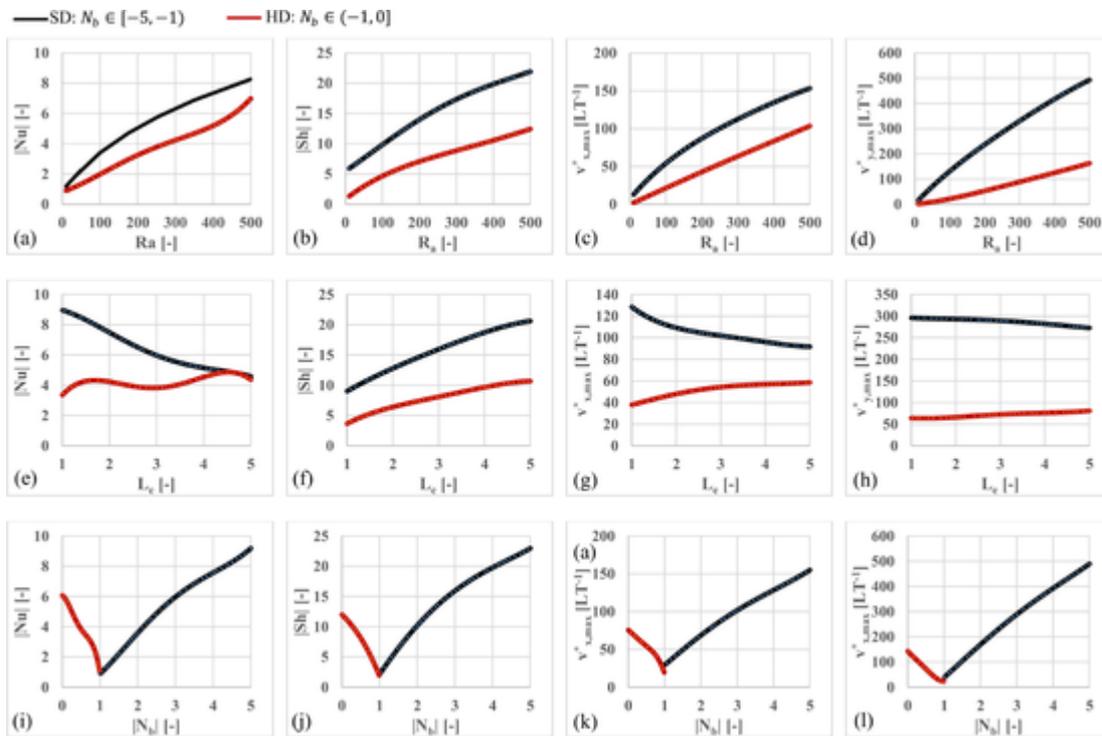


Fig. 14. Marginal effects of the input parameters on the scalar output QoLs.

variables as it is usually done in applications involving natural porous media. Further extension of this work could be the evaluation of the capacities of the proposed methodology in dealing with unstable configurations at high Rayleigh involving highly fingered states and several convective cells, and in addressing applications at large scale such as in geothermal reservoirs. For such applications, addressing uncertainties related to geological heterogeneity, temperature and concentration boundary conditions and fracture characteristics would be helpful for theoretical and practical purposes.

Uncited references

[3,12,17,20,21,25,3729,51,60–62].

CRediT authorship contribution statement

Mohammad Mahdi Rajabi: Conceptualization, Methodology, Investigation, Writing - original draft. **Marwan Fahs:** Conceptualization, Methodology, Writing - original draft. **Aref Panjehfouladgaran:** Investigation. **Behzad Ataie-Ashtiani:** Writing - review & editing. **Craig T. Simmons:** Writing - review & editing. **Benjamin Belfort:** Writing - review & editing.

CRediT authorship contribution statement

Mohammad Mahdi Rajabi: Conceptualization, Methodology, Investigation, Writing - original draft. **Marwan Fahs:** Conceptualization, Methodology, Writing - original draft. **Aref Panjehfouladgaran:** Investigation. **Behzad Ataie-Ashtiani:** Writing - review & editing. **Craig T. Simmons:** Writing - review & editing. **Benjamin Belfort:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Mohammad Mahdi Rajabi and Marwan Fahs would acknowledge the support from the National School of Water and Environmental Engineering of Strasbourg (PORO6100 Grant). Behzad Ataie-Ashtiani and Craig T. Simmons acknowledge support from the National Centre for Groundwater Research and Training, Australia. Behzad Ataie-Ashtiani also appreciates the support of the Research Office of the Sharif University of Technology, Iran. The data used in this work are available on the GitHub repository: <https://github.com/fahs-LHYGES>.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijheatmasstransfer.2020.120291.

Reference

- [1] A M Al-Amiri, K M Khanafer, Numerical simulation of double-diffusive mixed convection within a rotating horizontal annulus, *Int. J. Therm. Sci.* 45 (6) (2006) 567–578.
- [2] Frano Barbir, *PEM Fuel Cells: Theory and Practice*, Academic Press, 2012.
- [3] A C Baytas, A F Baytas, D B Ingham, I Pop, Double diffusive natural convection in an enclosure filled with a step type porous layer: non-Darcy flow, *Int. J. Therm. Sci.* 48 (4) (2009) 665–673.
- [4] R Bennacer, A Tobbal, H Beji, P Vasseur, Double diffusive convection in a vertical enclosure filled with anisotropic porous media, *Int. J. Therm. Sci.* 40 (1) (2001) 30–41.
- [5] G Blatman, B Sudret, Efficient computation of global sensitivity indices using sparse polynomial chaos expansions, *Reliab. Eng. Syst. Saf.* 95 (11) (2010) 1216–1229.
- [6] E W Bolton, A Firoozabadi, Numerical modeling of temperature and species distributions in hydrocarbon reservoirs, *J. Geophys. Res.: Solid Earth* 119 (1) (2014) 18–31.
- [7] J D Brown, G B Heuvelink, Assessing uncertainty propagation through physically based models of soil water flow and solute transport, *Encyclopedia of Hydrological Sciences*, 2006.
- [8] R E Caflisch, Monte Carlo and quasi-Monte Carlo methods, *Acta Numer.* 7 (1998) 1–49.

- [9] S Chakraborty, P Dutta, The effect of solutal undercooling on double-diffusive convection and macrosegregation during binary alloy solidification: a numerical investigation, *Int. J. Numer. Methods Fluids* 38 (9) (2002) 895–917.
- [10] M Corcione, S Grignaffini, A Quintino, Correlations for the double-diffusive natural convection in square enclosures induced by opposite temperature and concentration gradients, *Int. J. Heat Mass Transf.* 81 (2015) 811–819.
- [11] V A F Costa, Double-diffusive natural convection in parallelogrammic enclosures, *Int. J. Heat Mass Transf.* 47 (14–16) (2004) 2913–2926.
- [12] M Fahs, T Graf, T V Tran, B Ataie-Ashtiani, C T Simmons, A Younes, Study of the effect of thermal dispersion on internal natural convection in porous media using Fourier series, *Transp. Porous Med.* 131 (2020) 537–568, doi:10.1007/s11242-019-01356-1.
- [13] M Fahs, A Younes, T A Mara, A new benchmark semi-analytical solution for density-driven flow in porous media, *Adv. Water Res.* 70 (2014) 24–35.
- [14] N Fajraoui, M Fahs, A Younes, B Sudret, Analyzing natural convection in porous enclosure with polynomial chaos expansions: effect of thermal dispersion, anisotropic permeability and heterogeneity, *Int. J. Heat Mass Transf.* 115 (2017) 205–224.
- [15] J Feinberg, H P Langtangen, Chaospy: an open source tool for designing methods of uncertainty quantification, *J. Comput. Sci.* 11 (2015) 46–57.
- [16] G Freni, G Mannina, Bayesian approach for uncertainty quantification in water quality modelling: the influence of prior distribution, *J. Hydrol.* 392 (1–2) (2010) 31–39.
- [17] S N Gaikwad, S Kouser, Double diffusive convection in a couple stress fluid saturated porous layer with internal heat source, *Int. J. Heat Mass Transf.* 78 (2014) 1254–1264.
- [18] B Ganapathysubramanian, N Zabararas, Sparse grid collocation schemes for stochastic natural convection problems, *J. Comput. Phys.* 225 (1) (2007) 652–685.
- [19] R Ghanem, Probabilistic characterization of transport in heterogeneous media, *Comput. Methods Appl. Mech. Eng.* 158 (3–4) (1998) 199–220.
- [20] B Goyeau, J P Songbe, D Gobin, Numerical study of double-diffusive natural convection in a porous cavity using the Darcy-Brinkman formulation, *Int. J. Heat Mass Transf.* 39 (7) (1996) 1363–1378.
- [21] N Hadidi, R Bennacer, Heat and mass transfer by natural convection in a bi-layered cubic enclosure with opposing temperature and concentration gradients, *Int. J. Therm. Sci.* 132 (2018) 534–551.
- [22] Y Hao, J J Nitao, T A Buscheck, Y Sun, Double-diffusive natural convection in a nuclear waste repository, *Nucl. Technol.* 163 (1) (2008) 38–46.
- [23] J D Herman, W Usher, SALib: an open-source python library for sensitivity analysis, *J. Open Source Softw.* 2 (9) (2017) 97.
- [24] B Iooss, A Saltelli, Introduction to sensitivity analysis, *Handbook of Uncertainty Quantification*, 2017, pp. 1103–1122.
- [25] A W Islam, M A Sharif, E S Carlson, Numerical investigation of double diffusive natural convection of CO₂ in a brine saturated geothermal reservoir, *Geothermics* 48 (2013) 101–111.
- [26] A W Islam, H R Lashgari, K Sephernoori, Double diffusive natural convection of CO₂ in a brine saturated geothermal reservoir: study of non-modal growth of perturbations and heterogeneity effects, *Geothermics* 51 (2014) 325–336.
- [27] Z Jamshidzadeh, F T C Tsai, S A Mirbagheri, H Ghasemzadeh, Fluid dispersion effects on density-driven thermohaline flow and transport in porous media, *Adv. Water Res.* 61 (2013) 12–28.
- [28] G R Kefayati, Thermosolutal natural convection of viscoplastic fluids in an open porous cavity, *Int. J. Heat Mass Transf.* 138 (2019) 401–419.
- [29] M Konz, A Younes, P Ackerer, M Fahs, P Huggenberger, E Zechner, Variable-density flow in heterogeneous porous media—laboratory experiments and numerical simulations, *J. Contam. Hydrol.* 108 (3–4) (2009) 168–175.
- [30] A V Kuznetsov, D A Nield, The effect of strong heterogeneity on the onset of convection induced by internal heating in a porous medium: a layered model, *Transp. Porous Media* 99 (1) (2013) 85–100.
- [31] M Le Lous, F Larroque, A Dupuy, A Moignard, P C Damy, Performance of an open-loop well-doublet scheme located in a deep aquitard-aquifer system: Insights from a synthetic coupled heat and flow model, *Geothermics* 74 (2018) 74–91.
- [32] O Le Maître, M T Reagan, B Debusschere, H N Najm, R G Ghanem, O M Knio, Natural convection in a closed cavity under stochastic non-Boussinesq conditions, *SIAM J. Sci. Comput.* 26 (2) (2004) 375–394.
- [33] O Le Maître, O M Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer Science & Business Media, 2010.
- [34] C Li, S Mahadevan, An efficient modularized sample-based method to estimate the first-order Sobol’ index, *Reliab. Eng. Syst. Saf.* 153 (2016) 110–121.
- [35] T R Mahapatra, P Mondal, Heatline and massline analysis due to magnetohydrodynamic double diffusive natural convection in a trapezoidal enclosure with various aspect ratios, *Int. J. Appl. Comput. Math.* 5 (3) (2019) 82.
- [36] A A Mohamad, R Bennacer, Double diffusion, natural convection in an enclosure filled with saturated porous medium subjected to cross gradients; stably stratified fluid, *Int. J. Heat Mass Transf.* 45 (18) (2002) 3725–3740.
- [37] A Mohammadi, M Raisee, Efficient uncertainty quantification of stochastic heat transfer problems by combination of proper orthogonal decomposition and sparse polynomial chaos expansion, *Int. J. Heat Mass Transf.* 128 (2019) 581–600.
- [38] A Mojtabi, M C Charrier-Mojtabi, Double-diffusive convection in porous media, *Handbook of Porous Media*, CRC Press, 2005, pp. 287–338.
- [39] H Monod, C Naud, D Makowski, Uncertainty and sensitivity analysis for crop models, *Working with Dynamic Crop Models: Evaluation, Analysis, Parameterization, and Applications*, 4, 2006, pp. 55–100.
- [40] D A Nield, A Bejan, *Double-diffusive convection, Convection in Porous Media*, Springer, Cham, 2017, pp. 473–537.
- [41] E Qian, B Peherstorfer, D O’Malley, V V Vesselinov, K Willcox, Multifidelity Monte Carlo estimation of variance and sensitivity indices, *SIAM/ASA J. Uncertain. Quant.* 6 (2) (2018) 683–706.
- [42] M M Rajabi, B Ataie-Ashtiani, Sampling efficiency in Monte Carlo based uncertainty propagation strategies: application in seawater intrusion simulations, *Adv. Water Res.* 67 (2014) 46–64.
- [43] M M Rajabi, B Ataie-Ashtiani, C T Simmons, Polynomial chaos expansions for uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations, *J. Hydrol.* 520 (2015) 101–122.
- [44] M M Rajabi, B Ataie-Ashtiani, Efficient fuzzy Bayesian inference algorithms for incorporating expert knowledge in parameter estimation, *J. Hydrol.* 536 (2016) 255–272.
- [45] M M Rajabi, Review and comparison of two meta-model-based uncertainty propagation analysis methods in groundwater applications: polynomial chaos expansion and Gaussian process emulation, *Stoch. Environ. Res. Risk Assess.* 33 (2) (2019) 607–631.
- [46] K Roy, P V S N Murthy, Soret effect on the double diffusive convection instability due to viscous dissipation in a horizontal porous channel, *Int. J. Heat Mass Transf.* 91 (2015) 700–710.
- [47] M Sankar, Y Park, J M Lopez, Y Do, Double-diffusive convection from a discrete heat and solute source in a vertical porous annulus, *Transp. Porous Media* 91 (3) (2012) 753–775.
- [48] S Shahane, N R Aluru, S P Vanka, Uncertainty quantification in three dimensional natural convection using polynomial chaos expansion and deep neural networks, *Int. J. Heat Mass Transf.* 139 (2019) 613–631.
- [49] Q Shao, M Fahs, A Younes, A Makradi, A high-accurate solution for Darcy-Brinkman double-diffusive convection in saturated porous media, *Numer. Heat Transf. Part B: Fundamentals* 69 (1) (2016) 26–47.
- [50] Q Shao, M Fahs, A Younes, A Makradi, T Mara, A new benchmark reference solution for double-diffusive convection in a heterogeneous porous medium, *Numerical Heat Transf. Part B: Fundam.* 70 (5) (2016) 373–392.
- [51] Q Shao, A Younes, M Fahs, T A Mara, Bayesian sparse polynomial chaos expansion for global sensitivity analysis, *Comput. Methods Appl. Mech. Eng.* 318 (2017) 474–496.
- [52] K M Shirvan, M Mamourian, S Mirzakanlari, R Ellahi, K Vafai, Numerical investigation and sensitivity analysis of effective parameters on combined heat transfer performance in a porous solar cavity receiver by response surface methodology, *Int. J. Heat Mass Transf.* 105 (2017) 811–825.
- [53] B Shome, Monte Carlo based uncertainty analysis for variable property mixed convection flow in a uniformly heated circular tube, *Int. J. Uncertain. Quant.* 6 (5) (2016).
- [54] R C Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, 12, SIAM, 2013.
- [55] A Tarakanov, A H Elsheikh, Regression-based sparse polynomial chaos for uncertainty quantification of subsurface flow models, *J. Comput. Phys.* 399 (2019) 108909.
- [56] O V Trevisan, A Bejan, Natural convection with combined heat and mass transfer buoyancy effects in a porous medium, *Int. J. Heat Mass Transf.* 28 (8) (1985) 1597–1611.
- [57] D Venturi, M Choi, G E Karniadakis, Supercritical quasi-conduction states in stochastic Rayleigh-Bénard convection, *Int. J. Heat Mass Transf.* 55 (13–14) (2012) 3732–3743.
- [58] D Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
- [59] D Xiu, G E Karniadakis, Modeling uncertainty in flow simulations via generalized polynomial chaos, *J. Comput. Phys.* 187 (1) (2003) 137–167.
- [60] A Younes, P Ackerer, Solving the advection-dispersion equation with discontinuous Galerkin and multipoint flux approximation methods on unstructured meshes, *Int. J. Numer. Methods Fluids* 58 (6) (2008) 687–708.
- [61] A Younes, M Fahs, S Ahmed, Solving density driven flow problems with efficient spatial discretizations and higher-order time integration methods, *Adv. Water Res.* 32 (3) (2009) 340–352.
- [62] A Younes, M Fahs, Extension of the Henry semi-analytical solution for saltwater intrusion in stratified domains, *Comput. Geosci.* 19 (6) (2015) 1207–1217.
- [63] F Y Zhao, D Liu, G F Tang, Application issues of the streamline, heatline and massline for conjugate heat and mass transfer, *Int. J. Heat Mass Transf.* 50 (1–2) (2007) 320–334.
- [64] F Y Zhao, D Liu, G F Tang, Natural convection in a porous enclosure with a partial heating and salting element, *Int. J. Therm. Sci.* 47 (5) (2008) 569–583.
- [65] Y Zhao, C Yan, X Wang, H Liu, W Zhang, Uncertainty and sensitivity analysis of SST turbulence model on hypersonic flow heat transfer, *Int. J. Heat Mass Transf.* 136 (2019) 808–820.