Response to the comments of the reviewer

I thank the referee for their insightful comments. I have greatly benefited from the comments, and the response to all comments is provided here. All the comments of the reviewer are retyped below in bold while the responses are given in normal font, and the changes in the thesis are highlighted in yellow.

Examiner 2

Following is a detailed review of the thesis. If included am hopeful the quality thesis will improve and will help the reader to understand better.

Since the process includes an interplay of heat, mass and momentum transfer, the non-dimensional numbers are a means to know/understand which transport phenomena: are dominant in the flow and so on.

The Grashof number that has been defined is mainly a ratio of buoyant and viscous forces. I do not see how heat and mass transfer (as stated) are related here. There is a Grashof number about natural convection in mass transfer problems which happens due to concentration gradient. Is the author trying to compare the two?

The Grashof number (Gr) is used to study the relative strength of natural convection compared to forced convection. It was observed that the deposition rate increased with an increase in operating pressure. However, the uniformity index decreased at higher pressures due to buoyancy-driven flow within the reactor. At high pressures (above 250 mbar), significant circulation between the wall and substrate was observed, leading to non-uniform reactant distribution and a decrease in uniformity index. Fig. 4.6 shows the variation of Gr/Re^2 ratio at different operating pressures. The Gr/Re^2 ratio increased with increasing reactor pressure, confirming the dominance of buoyancy-driven flow at high pressures (page 66).

Secondly, the Prandtl number does not compare two different modes of heat transfer. It states the importance of heat conduction to momentum diffusion. This number has not been calculated in the whole text.

The calculation of the Péclet number involves multiplying the Reynolds number and the Prandtl number. For Pe >> 1, advection dominates mass transport compared to diffusion, as diffusion occurs over a significantly longer timescale. Therefore, in a CVD reactor convection flow is

dominant over pure diffusion. Therefore, I have not mentioned the Prandtl number (Pr) explicitly in the text.

Again, the Peclet number can be defined both for heat and momentum? It's not stated for what purpose the non-dimensional number is going to be used also it has not been calculated in the thesis.

The Prandtl number (Pr) was used to calculate the Péclet number.

If the non-dimensional numbers are provided and the main aim is to understand the flow with the help of these numbers, why have their values not been evaluated in the thesis?

I have mentioned the variation of various dimensionless numbers at different operating pressures in Fig.4.6 and explained the importance of studying the relative strength of natural convection compared to forced convection (page 66).

62 out of 95 pages in the thesis (65%) consist of an introduction and literature review!! This shows that the result part has not been written in a very detailed manner.

I have added the governing equations, boundary conditions, solution procedure, mesh independent study, hydrodynamic study, and validation of machine learning in the revised thesis as Appendix A (page no. 83-87) & Appendix C (page no. 96)

In Figure. 3.1, the geometry should portray the dimensions and the boundary conditions like inlet/wall/outlet can be mentioned. Why are the equations describing the BCs Stated in the figure? It can be written in the text.

I have added the reactor geometry with dimensions governing equations and all boundary conditions used in the simulation in Appendix A (page no. 83).

The mathematical models used have several limitations:

1. The Reynolds number of the flow has not been evaluated. Not shown in any of the Figures. So how can one know that the flow is in a laminar regime? Even if it is at low temperature at high temperature does the flow remain to be in a laminar regime? What is the justification for using a laminar flow model?

Analysis of Fig. 4.6 reveals that the Reynolds number in the reactor operates consistently in the laminar flow regime, remaining below 2100 at all operating pressures (page no. 68)

2. Are the properties of viscosity, and mass diffusivity considered to be a function of temperature? Because there is a wide range of temperature distribution present inside the reactor. What type of variation has been considered?

The viscosity and mass diffusivity were calculated as described in Appendix A (page no. 84,85). The Chapman-Enskog equation was utilized to estimate the diffusivity of the binary gas mixture, while the viscosity of the gases was determined using the kinetic theory of gases. These gaseous properties were correlated as functions of temperature and pressure and incorporated into the simulation.

3. When so many species of different compositions are present what is the justification of using an ideal gas model?

In the present study, the use of an ideal gas model is accurate because the CVD process is the following conditions are met:

i) At low pressures, gas molecules are far apart, and intermolecular forces are minimal.

ii) At high temperatures (>1000^0C), the kinetic energy of the gas molecules is significant, overwhelming the effects of intermolecular forces.

iii) Gases with small molecules, such as hydrogen and methane, are less likely to deviate from ideal behaviour due to their reduced volume and minimal intermolecular interactions (page 83-84).

This is a thesis work, not a publication so showing details of simulations like a figure depicting the grid independence test (missing in the current form) will help the reader.

The resulting inlet Reynolds number was compared, revealing a continuous increase with the grid number up to 3.5 million, as shown in Fig. 3.3. Beyond this point, the Reynolds number stabilized. Considering available computational resources, a grid number of 3,660,910 was selected for subsequent simulations where the relative increase of less than 1% in the cell values of the aforementioned variables (page no. 52-54).

Since unsteady state solutions are obtained has an implicit or explicit method been employed? What is the value of delta t that has been used?

The time steps used in the simulations are given in the revised thesis (page no. 53).

The thesis should present some validation results of the SVM model utilized on some well-defined smaller physical problem, which can be a part of the Appendix.

Validation of the SVM model and Nelder-Mead Algorithm with a well-defined problem is provided in Appendix C (page no. 96).

In the Results section.

In Figure 4.1 variation in temperature is not vivid. The minimum temperature of 353 K is only at the inlet whereas most of the reactor is at a much higher temperature. The scaling needs to be redone so that the temperature variation becomes evident. Also, the three figures must be in the same temperature range for easy comparison.

I have replaced this figure with a new one which shows the temperature variation inside the reactor. In the revised thesis it is added as Figure 4.2. The temperature range scale was also kept constant for better visibility (page 64).

"the temperature profile is axisymmetric and uniform, except at the bottom plate" The bottom plate temperature distribution cannot be seen at all. Ideally, sections need to be shown separately.

I replaced Fig. 4.1 with a more detailed figure showing temperature variations at the reactor inlet (page 64).

In Fig. 4.2 4.6 C_2H_4 is not visible at all. Has it been plotted? Replot so that all constituents can be visualized.

The concentration of C_2H_4 is minimal, hence it does not appear in the figure. The concentration (or mole fraction) values were measured as cross-sectional averages at the given horizontal plane, and only the species with a molar fraction $>10^{-4}$ are plotted. Because of its low concentration, it contributes very little to the rate of deposition.

In Fig. 4.3 why in the caption is written predicted deposition rate, it has been obtained from numerical simulations. There is no prediction here? Is there?

I replaced the word predicted with simulated in the text (page 65).

In Fig. 4.4, why are there 3 cross-sections of substrates shown for mole fraction than 4? Which 3 are you talking about?

In the thesis variations of the mole fraction of C_2H_2 at three cross-sections were taken. They are in the middle of two substrates in a horizontal plane (page 67).

4.3 Effect of total flow rate

The total flow rate was changed in the range of 10-30 slm" What unit is slm?

slm is a standard unit of volumetric flow rate. I have mentioned this in the revised thesis (page 69).

In Fig. 4.7 draw temperature profiles also for different flow rates.

The temperature profiles for flow rates of 10 slm and 30 slm are depicted in Fig. 4.8 (page 70).

Fig. 4.8 to compare the variation of mole fraction of CH_4 and C_2H_2 species along the reactor length for a) 0.5, b) 0.7, and c) 0.9 inlet mole fraction of CH_4 contour range must be the same.

I presented the counter of the mole fraction of CH_4 and C_2H_2 in the same range in the revised thesis (page 73).

"However, the fidelity of the CFD simulation that included the temperature-dependent physical properties was only shown in Fig. 4.9 for brevity. It has been observed that the CFD simulation that considers temperature-dependent physical properties is ~10 % more accurate than the simulation with constant temperature-physical properties".This being a thesis, it requires that the results be presented in totality and not in brief. The effect of physical variables being dependent on temperature on hydrodynamics, mass transfer and heat transfer needs to be shown. Could be included in the Appendix if not in the main text of the chapters.

The effect of thermal diffusion on the velocity profile (hydrodynamics) and concentration of methane (mass transfer) is shown in Appendix Das Fig. D1 and D2 (page no. 97).

In Fig. 4.9 why are the values from simulations underpredicted?

CFD simulations of commercial reactors often underpredict actual performance due to several factors, including oversimplified geometries, inaccurate chemical kinetics, and numerical errors. These errors can arise from discretization limitations, solver algorithm shortcomings,

and experimental challenges related to multiphase flows and coupled phenomena (page no. 62).

What is the definition of the Uniformity Index? That being the main quantity that is measured has never been defined.

It is defined in the glossary (page no.98). The average rate of deposition (RD) and uniformity index (UI) was calculated as (page no. 48):

$$\overline{RD} = \frac{\int RD dA}{A}$$

$$UI = \frac{\int |\mathrm{RD} - \overline{RD}| \, dA}{A}$$

The linear regression results can also be plotted to show the betterment brought about by the SVM Model.

The performance of various statistical and machine learning models predicting deposition rate is added in Table 4.3 (page no. 77).

The CFD results with the optimum conditions from Table 4.4 need to be plotted showing how the values have improved from the presented values.

The revised thesis now includes Fig. 4.16, which compares the uniformity index to the original operating parameters under optimal conditions from Table 4.4 (page no. 78).

Can experiments be done under these conditions? In reality, do the values improve?

I suggested the optimum conditions for the commercial CVD reactor to our project partner. They evaluated the optimum conditions and noticed an improvement in the performance of the CVD reactor for carbon coating.