

# Dynamic modeling of ultrafine particle classification in tubular bowl centrifuges

Sandesh A. Hiremath<sup>a\*</sup>, Marco Gleiß<sup>b</sup>, and Naim Bajcinca<sup>a</sup>

<sup>a</sup> Department of Mechanical and Process Engineering, University of Kaiserslautern-Landau (RPTU), Kaiserslautern, Germany

<sup>b</sup> Institute of Mechanical Process Engineering and Mechanics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

\* Corresponding Author: [sandesh.hiremath@rptu.de](mailto:sandesh.hiremath@rptu.de)

---

## ABSTRACT

Ultrafine or colloidal particles are widely used in industry as aerogels, coatings, filtration aids or films and require a defined particle size. In this case, tubular centrifuges are suitable for particle classification due to the high g-forces. Yet, the design and optimization of tubular centrifuges requires numerous pilot experiments, which is time-consuming and costly. Altogether, design of an efficient centrifuge is a challenging task. To handle this limitation, this study presents a dynamic model for the real-time simulation of the dynamic behavior during particle classification for a pilot-scale tubular centrifuge. First, we discuss the new continuous mathematical model of a previously published multi-compartment model by Winkler et al. [1]. The coupled PDE-ODE model not only provides a better interpretation and fine-grained details of the dynamics but also provides tools for numerically stable algos. The latter is especially crucial for control synthesis tasks for which model predictive control method is commonly used. We validate the model by quantitatively comparing simulation results for product loss and qualitatively comparing grade efficiency, and sediment build-up with experimental data collected from pilot-scale experiments.

---

**Keywords:** Modelling and Simulations, Process Design, Nanoparticle Classification, PDE

## INTRODUCTION

The classification of colloidal and nano particles is becoming increasingly important in various sectors of the process industry. Examples include the semiconductor industry, catalyst production, production of active pharmaceutical ingredients, separating proteins or the pigment industry [2,3]. Compared to established methods such as microfiltration or field-flow fractionation, tubular centrifuges are particularly suitable for separating large quantities of nanoparticles in diluted slurries by semi-continuous centrifugation [4-6]. Conventional pilot-scale tubular centrifuges achieve extremely high g-forces up to  $C = 80,000g$

The design of tubular centrifuges is mostly based on the sigma theory which is a stationary black-box model to predict cut-size of particles dependent on process conditions and centrifuge geometry [7]. This approach disregards separation and the real flow behavior, which influences the centrifuge operation significantly [8]. Consequently, it is necessary to carry out numerous tests with machines on a pilot-scale. The outcome of the time- and cost-intensive procedure serves as input for a basic centrifuge design. However, there are often problems with the design of centrifuges, because products are currently not in stock or not available in large quantities, which is usually the case for the synthesis of

ultrafine particles.

The computing power available today allows the simulation of tubular centrifuges on different time and length scales, which results in a very detailed process understanding [9]. Difficulties in realistic simulation of tubular centrifuges arise from the properties of the disperse and liquid phases. In particular, the sedimentation and rheological behavior have a decisive influence on the separation result.

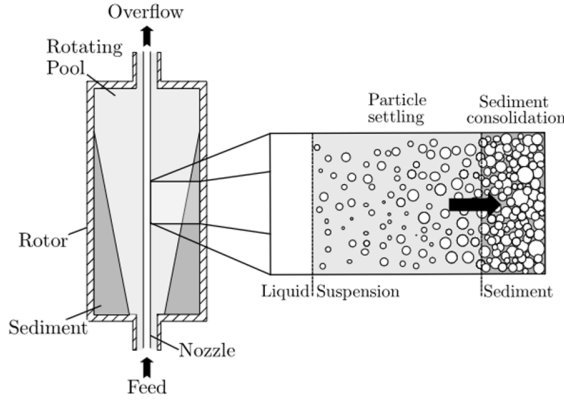
However, CFD simulations are time-consuming and require high computing power. In contrast to that, multi-compartment models are used for real-time process simulation of tubular centrifuges. It is therefore possible to extend this type of model and integrate it into advanced control strategies such as model predictive or stochastic control. The driving force is to reduce development time, respond to growing customer demands, and conserve resources and energy.

This study presents a new mathematical formulation of a previously presented multi-compartment model [1] to simulate particle separation and sediment build-up in tubular centrifuges. The contribution firstly introduces a continuous model for predicting the dynamics sedimentation in a tubular bowl centrifuge and mathematically prove the wellposedness of the model. Subse-

quently, we validate the model predictions with the experimentally obtained data. Here we formulate an optimal design problem and solve it using suitable neural network implementation.

## PROCESS DESCRIPTION

Tubular centrifuges are semi-continuously operating solid bowl centrifuges. **Figure 1** schematically shows the tubular centrifuge design and physical behavior of the separation process in the rotor. Main parts of the centrifuge are a slim cylindrical rotor and a feed nozzle. A direct drive at the top generates the rotor rotation. The suspended solids are pumped axially into the rotor via the inlet nozzle at the bottom of the rotor. An overflow weir forms a rotating liquid pool. The difference in density between the solid and the liquid causes a phase separation. Solid particles have a larger density compared to the liquid and settle towards the inner wall of the rotor and accumulate there as a saturated sediment. The machine can be used for classification if the process conditions of volume flow and speed are selected so that only some of the particles are deposited in the rotor.

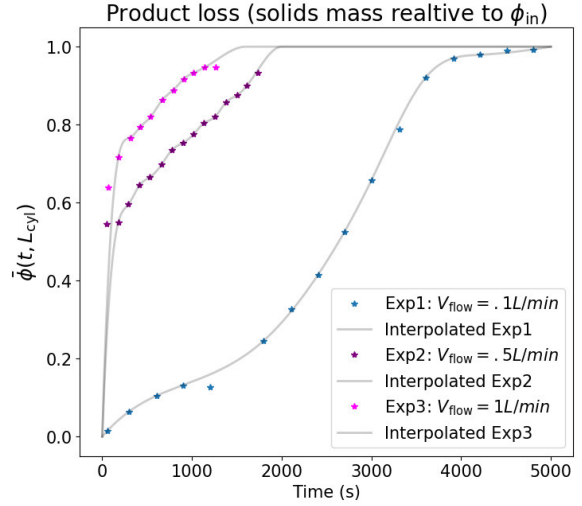


**Figure 1.** Schematic representation of a tubular centrifuge with feed, overflow and different states of the solid-liquid mixture along the radial position of the rotor.

## EXPERIMENT SETUP

Here we shall describe the experimental step and the procedure for obtaining the data. The tests were carried out on a Z11 tubular centrifuge from Carl Padberg Zentrifugenbau GmbH. Please refer to Table 1 for the geometric dimensions. Hydrophilic fumed silica particles named Aerosil 200 from Evonik were used as the test product. The average particle diameter is in the range of 100 nm. For the tests, the suspension was stirred and dispersed with a sonotrode. The rotational speed was set to 40,000 rpm. Using this setup, 3 experiments were performed by varying the feed rate i.e. volume flow, namely for  $V_{\text{flow}} \in \{0.1, 0.5, 1\} \frac{\text{L}}{\text{min}}$ . Based on this, for each experiment, the product loss (or relative solids mass)  $\bar{\phi}_{\text{out}}(t) = \bar{\phi}(t, L_{\text{cyl}})$  at the end point of the centrifuge bowl (cylinder) at different time points is measured. The product loss  $\bar{\phi}_{\text{out}}$  which describes the ratio of the solids mass of the concentrate to the solids mass of the feed, thus taking values between 0

and 1, serves a key indicator for the performance of the bowl centrifuge and thus is suitable for model validation. The raw measured data  $y = \{y^1, y^2, y^3\}$ ,  $y^i = (y^i_{t_k})$  amounts up to roughly 100 data points in total and is temporally sparse. As a preprocessing step we interpolate each sequence separately to obtain a dense continuous representation. Both the measured and the interpolated data are shown in **Figure 2**.



**Figure 2.** Measured solids mass (relative to  $\phi_{\text{in}}$ ) at the cylinder endpoint  $L_{\text{cyl}}$  for varying input flow rates. The starred markers are the raw data while the line depicts the interpolated values.

## MATHEMATICAL MODEL

The dynamic model is based on a set of partial differential equations aimed to calculate the behavior of sedimentation and sediment build-up within the investigated tubular bowl centrifuge. It is specified as follows

$$\partial_t \phi = v(t, x) \cdot \partial_x \phi + v(t, x) \cdot (\psi_t - \phi), t > 0, x \in (0, L_{\text{cyl}}] \quad (1.1)$$

$$\phi(0, x) = \phi_0(x), x \in [0, L_{\text{cyl}}], \phi_t(t, 0) = \phi_{\text{in}}, t \geq 0 \quad (1.2)$$

$$\partial_t V_{\text{sed}}(t, x) = \phi \cdot V_{\text{flow}} \cdot E, t > 0, \forall x \quad (1.3)$$

$$V_{\text{sed}}(0) = V_0 \quad (1.4)$$

$$\partial_x \psi(t, x) = -E \cdot \psi, x > 0, \forall t \quad (1.5)$$

$$\psi_t(0) = \phi(t, 0) \cdot (1 - E(t, 0)), \forall t \quad (1.6)$$

$$\partial_x q(t, x, \ell) = q \cdot P(E, T), x > 0, \forall t \quad (1.7)$$

$$q(t, 0, \ell) = q_{\text{in}} \cdot (P(E(t, 0), T(t, 0, \ell)) + 1) \quad (1.8)$$

where the coupling variables are defined as

$$P(E, T) = \left( \frac{E - T}{1 - E} \right) \quad (1.9)$$

$$R_s(V_{\text{sed}}) = \sqrt{\left( R_d^2 - V_{\text{sed}}(t) \frac{1}{\pi \ell_d} \right)} \quad (1.10)$$

$$V(V_{\text{sed}}) = \ell_d \pi \cdot (R_s^2(V_{\text{sed}}) - R_w^2) \quad (1.11)$$

$$v(V_{\text{sed}}) = \frac{V_{\text{flow}}}{V(V_{\text{sed}})} \quad (1.12)$$

$$T(\phi, V_{\text{sed}}) = \frac{R_s^2(V_{\text{sed}})}{R_s^2(V_{\text{sed}}) - R_w^2} \left[ 1 - e^{\left( -\alpha(\phi) \frac{\ell^2 \omega^2 (R_s^2 - R_w^2) \pi \ell_d}{V_{\text{flow}}} \right)} \right] \quad (1.13)$$

$$\alpha(\phi) = \frac{\delta \rho \cdot h(\phi, t)}{\eta_\ell} \quad (1.14)$$

$$E(t, x) = \int_{\ell_{\text{min}}}^{\ell_{\text{max}}} T \cdot q \cdot d\ell \quad (1.15)$$

Here,  $R_d$ ,  $R_w$ ,  $\rho_\delta$  and  $\eta_\ell$  are the fixed model parameters. These variables have the following application relevant interpretation  $R_w$  is the max or weir radius,  $\rho_\delta$  is the density difference between solid and liquid phase and  $\eta_\ell$  is the dynamic viscosity of liquid. The inputs, i.e., the control variables of the system are the fluid flow rate  $V_{\text{flow}}$  and the drum angular velocity  $\omega$ . Finally, the two main observed or the measure quantities of the process are the solids fraction  $\phi$  at the boundary point  $L_{\text{cyl}}$  for all times  $t \geq 0$ . Altogether, the introduced model is a novel adaptation of the compartmental model introduced in [1].

**Assumption (A).** In order to show wellposedness of the proposed model (1.1) - (1.15) we shall make the following assumptions:

1.  $T, E \in [0, 1]$ ,  $R_s - R_w \geq R_{\text{crit}} > 0$ ,  $0 \leq h(\phi, t) \leq k_h$
2.  $0 \leq v \leq k_v$ ,  $0 \leq V_{\text{sed}}, V \leq k_v$ ,  $\|T\|_\infty \leq k_T$ ,  $\|E\|_\infty \leq k_E$ ,
3.  $\phi_0(x) = \phi_{\text{in}} \forall x \leq 0$ ,  $v(t, x) = \psi(t, x) = 0, \forall x \leq 0, t \geq 0$ ,
4.  $v \in C^1(\mathbb{R})$  such that

$$\begin{aligned} \|v\|_{C^1} &\leq k_v \\ \|v(v^1) - v(v^2)\|_{C^1} &\leq k_v \|v^1 - v^2\|_{C^1} \end{aligned} \quad (2.1)$$

5.  $T \in C^1(\mathbb{R}^2; \mathbb{R})$  such that

$$\begin{aligned} \|T\|_{C^1} &\leq k_T \\ \|T(p^1, v^1) - T(p^2, v^2)\|_{C^1} &\leq k_{T,p} \|p^1 - p^2\|_{C^1} \\ &\quad + k_{T,v} \|v^1 - v^2\|_{C^1} \end{aligned} \quad (2.2)$$

6.  $E \in C^1(\mathbb{R}^3; \mathbb{R})$  such that

$$\begin{aligned} \|E\|_{C^1} &\leq k_E \\ \|E(p^1, v^1, r^1) - E(p^2, v^2, r^2)\|_{C^1} &\leq k_{E,p} \|p^1 - p^2\|_{C^1} \\ &\quad + k_{E,v} \|v^1 - v^2\|_{C^1} + k_{E,r} \|r^1 - r^2\|_{C^1} \end{aligned} \quad (2.3)$$

7.  $P \in C^1(\mathbb{R}^3; \mathbb{R})$  such that

$$\begin{aligned} \|P\|_{C^1} &\leq k_P \\ \|P(t^1, e^1) - P(t^2, e^2)\|_{C^1} &\leq k_{P,T} \|t^1 - t^2\|_{C^1} \\ &\quad + k_{P,e} \|e^1 - e^2\|_{C^1} \end{aligned} \quad (2.4)$$

**Theorem 1.** Let  $\mathcal{T}_{s,t}: C^1 \rightarrow C^1$  be an evolution operator defined as  $\mathcal{T}_{s,t}f = e^{-\mu_{s,t}t} f(x - \mu_{s,t}(x))$ ,  $\mu_{s,t}(x) = \int_s^t v(r, x) dr$ ,  $\forall f, v \in C^1$

Then under the Assumption (A), the solution to system of equations (1.1) - (1.4) is given by  $\phi, V_{\text{sed}}$  taking the following form:

$$\phi(t, x) = \mathcal{T}_{0,t}\phi_0(x) + \int_0^t \mathcal{T}_{s,t}[(v\psi)(s, x)] ds \quad (2.5)$$

$$V_{\text{sed}}(t, x) = V_0 + V_{\text{flow}} \int_0^t E(s, x) \phi(s, x) ds \quad (2.6)$$

Correspondingly, the solution to equations (1.5) - (1.8) is given by  $\psi, q$  which takes the following form:

$$q(\cdot, x, \cdot) = q_0(\cdot, 0, \cdot) e^{\int_0^x P(\cdot, y, \cdot) dy}, \psi(\cdot, x, \cdot) = \psi_0(\cdot, 0) e^{-\int_0^x E(\cdot, y) dy}.$$

**Proof.** First let us consider the system (1.1)-(1.4). Then, for a fixed  $\phi_0, \phi_{\text{in}}, v$  and  $\psi$  as per (A.3), we first observe that (2.5) satisfies (1.1). Then for  $t, s \in [0, T]$ , using (2.1) and letting  $\|\cdot\| := \|\cdot\|_{C^1}$  we get that

$$\begin{aligned} \|\mu_{s,t}^1 - \mu_{s,t}^2\| &\leq \int_s^t \|v^1 - v^2\| dr \leq |t - s| \|v^1 - v^2\| \\ \|\mathcal{T}_{s,t}^1 f - \mathcal{T}_{s,t}^2 f\| &= \|e^{-\mu_{s,t}^1 t} f(x - \mu_{s,t}^1(x)) - e^{-\mu_{s,t}^2 t} f(x - \mu_{s,t}^2(x))\| \\ &\leq |e^{-\mu_{s,t}^1 t}| \|f(x - \mu_{s,t}^1(x)) - f(x - \mu_{s,t}^2(x))\| + \\ &\quad \|f(x - \mu_{s,t}^2(x))\| \|e^{-\mu_{s,t}^1 t} - e^{-\mu_{s,t}^2 t}\| \\ &\leq (\|f'\|_{C^1} + 1) \|\mu_{s,t}^1 - \mu_{s,t}^2\|_{C^1} \\ &\leq (\|f'\|_{C^1} + 1) \|v^1 - v^2\|_{C^1} |t - s|. \end{aligned}$$

Consequently, we get that

$$\begin{aligned} \|\phi^1 - \phi^2\| &\leq \|\mathcal{T}^1 \phi_0 - \mathcal{T}^2 \phi_0\| + \int_0^t \|\mathcal{T}^1 v^1 - \mathcal{T}^2 v^2\| ds \\ &\leq ((k_{\phi_0} k_v + 1) + (k_{\phi_0} k_v + 1) k_{\psi} k_v T + k_T k_{\psi}) \|v^1 - v^2\| T \\ &\leq k_{\phi} k_v \|V_{\text{sed}}^1 - V_{\text{sed}}^2\| T. \end{aligned}$$

Next, we see that for fixed  $\phi$  and  $E$  (2.6) satisfies (1.3)-(1.4). Then following similarly as above we get that  $V_{\text{sed}}(\phi)$  satisfies

$$\|V_{\text{sed}}^1 - V_{\text{sed}}^2\| \leq k_{V_{\text{flow}}} k_E \|\phi^1 - \phi^2\| T$$

Combining the two, via Picard's successive approximation we obtain a sequence  $(\phi^n)_{n \in \mathbb{N}}$  such that

$$\|\phi^{n+1} - \phi^n\| \leq T k_{\phi} k_v k_{V_{\text{flow}}} k_E \|\phi^n - \phi^{n-1}\|.$$

Now, taking

$$0 < T < \frac{1}{k_{\phi} k_v k_{V_{\text{flow}}} k_E}$$

we obtain that the mapping  $\phi^n \mapsto \phi^{n+1}$  is a contraction mapping in  $C^1$ . Thus, applying Banach's fixed point theorem we obtain the existence of a unique local solution on  $[0, T]$ . Thus, for fixed  $(q, \psi)$  we get a unique solution  $\phi, V_{\text{sed}} \in C^1$  to (1.1) - (1.4) on  $[0, T]$ .

The situation is similar with the inequalities

$$\begin{aligned} \|E^1 - E^2\| &\leq k_E \|q^1 - q^2\| \\ \|f_q^1 - f_q^2\| &\leq k_P \|E^1 - E^2\| \leq k_P k_E \|q^1 - q^2\|. \end{aligned}$$

for  $(\psi, q)$  we have that

$$\begin{aligned} \|q^1 - q^1\| &\leq T e^{k_P T} \|f_q^1 - f_q^2\| \\ \|\psi^1 - \psi^2\| &\leq k_\ell \|q^1 - q^2\|. \end{aligned}$$

Now using the same argument as above, we obtain that the mapping  $q^n \mapsto q^{n+1}$  forms a contraction mapping for small enough  $T > 0$ . Thus  $q^n \rightarrow q$  in  $C^1$  as  $n \rightarrow \infty$ . Consequently, the sequence  $(\psi^n)_{n \in \mathbb{N}}$  obtained via the mapping  $q^n \mapsto \psi^n$  is also convergent with  $\psi^n \rightarrow \psi$  in  $C^1$  as  $n \rightarrow \infty$ . Thus, for fixed  $(\phi, V_{\text{sed}})$  we obtain a local solution  $(q, \psi)$  to (1.5) - (1.8) on  $[0, T]$ ,  $T > 0$ . Finally, using the inequalities (2.1) - (2.4) from (A) we can also obtain that the mapping

$$(\phi^n, V_{\text{sed}}^n, q^n, \psi^n) \mapsto (\phi^{n+1}, V_{\text{sed}}^{n+1}, q^{n+1}, \psi^{n+1})$$

is a contraction for a small enough  $T > 0$ . ■

Let  $X = [\phi, V_{\text{sed}}, \psi, q]^T$  denote the state vector, let  $\zeta = [\phi_0, V_0, q_{\text{in}}, \phi_{\text{in}}, V_{\text{flow}}, \omega, R_d, R_w, \rho_\delta, \eta_\ell]^T$  denote the given fixed parameters of the model and let  $\theta = [\kappa, h]$  be the free unknown model parameter vector that needs to be estimated. Then based on the Theorem 1 we have that for every given  $(\zeta, \theta)$  there exists a mapping  $(\zeta, \theta) \mapsto \mathcal{S}(\zeta, \theta)$  such that  $\mathcal{S}(\zeta, \theta) = X^\theta$  solves the system (1.1) - (1.15). This is to say that  $\mathcal{S}(\zeta, \theta)$  is the solution operator of the system, which is dependent on the parameter tuple  $(\zeta, \theta)$ . Let  $y = (y_k)_{k \in \mathbb{N}}$  be the set of measurements and let  $Y_t = H X_t^\theta = \phi_{\text{out}}(t) = \phi(t, L_{\text{cyl}})$  be the measurement model. In order to validate the model, we formulate the following optimal design problem (ODP) provides the mechanism to find a suitable  $\theta$  that fits the data

$$\theta^* = \underset{\theta \in \Xi}{\operatorname{argmin}} J(\theta; \zeta, y) \quad (4)$$

$$\text{s.t. } X^\theta = \mathcal{S}(\zeta, \theta)$$

where,  $J(\theta; \zeta, y) := \sum_k \|H X_{t_k}^\theta - y_k\|^2$

**Table 1:** Fixed model parameters

Constant	Value	Constant	Value
$V_{\text{flow}}$	0.1, 0.5, 1	$\omega$	4187 s <sup>-1</sup>
$R_w$	0.0093 m	$L_{\text{cyl}}$	0.18 m
$\rho_\delta$	1202 kg/m <sup>3</sup>	$\ell_{\text{min}}$	1 × 10 <sup>-8</sup> m
$\eta_\ell$	9.9810 <sup>-4</sup> Pas	$\phi_0$	0.0
$a_1$	0	$a_2$	1
$q_0(\ell)$	$a_2 + \frac{a_1 - a_2}{1 + (\ell/a_3)^6}$	$\ell_\phi$	0.0063 m
$R_d$	0.0215 m	$V_0$	0.0 m <sup>3</sup>
$\phi_{\text{in}}$	0.005	$a_3$	7.68 × 10 <sup>-8</sup>
$\ell_{\text{max}}$	5 × 10 <sup>-7</sup> m	$\ell_V$	0.4070 m

## IMPLEMENTATION

To solve the ODP (4) requires providing certain input  $\xi = (\zeta, y)$  based on which a  $\theta^*$ , i.e. a solution, is obtained as output. Thus, the algorithm for solving the ODP (4) can be seen as a mapping  $\xi \mapsto \theta^*(\xi)$ . Let  $\mathcal{X}$  denote the input space for the ODP (4), with  $\mathcal{X} \ni \xi$ , our task in this section is to create a map  $\xi \mapsto \Psi(\xi)$  such that  $\Psi(\xi) \approx \theta^*(\xi)$  to the ODP. We represent  $\Psi(\xi)$  as parameterized estimator  $\mathbb{R}^q \ni \hat{\theta} \mapsto \Psi(\xi; \hat{\theta})$ ,  $q \gg 0$ , such that  $\Psi(\xi; \hat{\theta}) = \theta^*(\xi)$  for nearly every  $\xi \in \mathcal{X}$ . A weaker criterion is to attain  $\mathbb{E}^X[\Psi(\xi; \hat{\theta})] = \mathbb{E}^X[\theta^*(\xi)]$ . This leads to the statistical optimization problem (SP)

$$\hat{\theta} = \underset{\hat{\theta} \in \mathbb{R}^q}{\operatorname{argmin}} \mathbb{E}^X[J(\Psi(\xi; \hat{\theta}); \xi)] \quad (5)$$

Since evaluating the expectation in (5) is difficult, we consider empirical measure over a finite set of input samples given as  $\mathcal{X} \rightarrow \mathbb{X} = \{\xi^k\}_{k=1 \dots M}$ . Consequently, the SP takes the following discrete SP takes the following form

$$\hat{\theta}^M = \underset{\hat{\theta} \in \mathbb{R}^q}{\operatorname{argmin}} \frac{1}{M} \sum_{k=1}^M J(\Psi(\xi^k; \hat{\theta}); \xi^k).$$

Further considering after temporal discretization of the objective function  $J$  we get

$$\begin{aligned} \hat{\theta}^M &= \underset{\hat{\theta} \in \mathbb{R}^q}{\operatorname{argmin}} \frac{1}{M} \sum_{k=1}^M \sum_{i=1}^N \|H X_{t_i}^k - y_i^k\|^2 \\ &= \underset{\hat{\theta} \in \mathbb{R}^q}{\operatorname{argmin}} \mathbb{L}(\hat{\theta}; \mathbb{X}). \end{aligned} \quad (6)$$

Since each input sample  $\xi^k$  involves a sequence data, the predicted output of the model  $\Psi$  is also a sequence. This means for a sample  $\xi^k$  involve a sequence of length  $N$  obtained at time  $t_k$

$$\Psi(\xi^k; \hat{\theta}) = \theta_{t_k:t_{k+N-1}}^k = [\theta_{t_k}^k, \dots, \theta_{t_{k+N-1}}^k]^\top$$

and

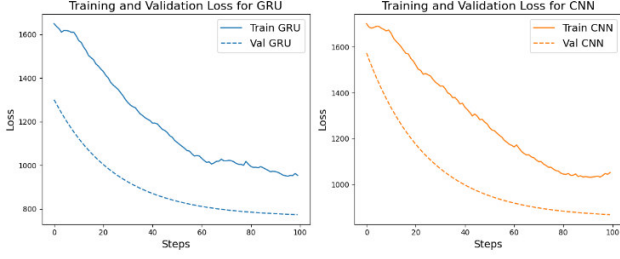
$$X^{\theta^k} = X^k(\theta_{t_k:t_{k+N-1}}^k) = X(\Psi(\xi^k; \hat{\theta})) = X^{\Psi(\xi^k; \hat{\theta})}$$

is the predicted state of the process under the parameter  $\theta_{t_i:t_{i+N-1}}^k$ . Following this, the parameterized optimal design function  $\Psi(\cdot; \hat{\theta})$  is implemented as a deep neural network (DNN) whose internal parameter  $\hat{\theta}$  is determines the solution to SP (6). Obtaining  $\hat{\theta}$  involves the algorithmic step called learning which uses stochastic approximation schemes to computationally determine the appropriate minimizer  $\hat{\theta}$  of the SP (6). Prior to discussing about the training procedure, we shall next discuss about the network architecture for the  $\Psi$  followed by a short discussion about dataset generation.

## NETWORK ARCHITECTURE

Since determining  $\hat{\theta}$  requires predicting the states  $X^k(\cdot)$  for random time intervals  $[t_k, t_{k+N}] \subset [0, T]$ , it is necessary to incorporate suitable solvers that is able to perform discrete numerical integration of the system (1.1)-(1.15). This motivates us to implement  $\Psi$  as a NeuralODE [10]. This means, that a suitably parameterized discrete numerical integrator (solver) is considered as a neural network and thus represents  $\Psi$ . Consequently, the core component of the network is the solver module, which we call as SysSolver. Due to the non-linear coupling of the system we implemented SysSolver using finite difference scheme using Pytorch library. Accordingly, the PDE (1.1)-(1.2) is discretized using Lax-Wenderhof method while the rest of the temporal and spatial ODEs are solved using Euler discretization method. Here, the time, space and particle size intervals are discretized in  $N$ ,  $M$  and  $L$  steps respectively. Furthermore, for each time step  $t$ , firstly the spatial ODEs (1.5) - (1.7) are solved for the known quantities  $\phi$  and  $V_{\text{sed}}$  at time  $t$  to obtain  $q$  and  $\psi$ . The latter are then used to solve for  $\phi$  and  $V_{\text{sed}}$ . Thus, given  $\zeta$  and  $\theta$ , SysSolver provides the solution  $X$  for the required time point  $T$ . Now in order to make SysSolver learnable, the free model parameter  $\theta$  is obtained using a neural network. Since  $\theta$  is composed of  $\kappa$  and  $h$  which are both time dependent and the latter additionally also being space dependent, the neural network  $\Gamma$  used to generate  $\theta$  is implemented using two different methods namely- using Gated-Recurrent-Units (GRU) and using Convolutional Neural Network (CNN). Accordingly, the input to  $\Gamma$  is given as  $\chi = (y, \phi_{\text{in}}, q_0, V_{\text{flow}}, \omega)$  and is obtained via the components of the  $\xi$  vector. Letting  $M = L$  and considering the initial value  $\phi_0$  and  $q_0$  which are space and size dependent the input to  $\Gamma$  is of the shape  $(b, 5, N, M)$  and the output is of shape  $(b, 2, N, M)$ . Based on this the GRU based network  $\Gamma$  is implemented in a two-layered stacked bidirectional model. The feature size of the hidden state is taken to be of the size 128. The hidden states at the output layer of the GRU (sub)-network are lastly passed through a fully connected linear layer followed by a RELU activation to obtain the final output. On the contrary the CNN implementation consists of five hidden convolutional layers with channel sizes 16, 32, 64, 32, 16, respectively. Each convolution layer is followed by a ReLU activation unit. Both the GRU and CNN implementations are inherently designed to handle variable-length input sequences. The GRU, by its recurrent structure, accommodates varying input lengths, while the CNN, being exclusively composed of convolutional layers, i.e. lacking fully connected layers, is also able to handle varying input sequence lengths. Following this approach,  $\Psi$  is implemented as  $\Psi(\xi; \hat{\theta}) = \text{SysSolver}(\zeta, \Gamma(\chi; \hat{\theta}))$ . Accordingly, the GRU based implementation of  $\Psi$  constitutes of roughly 13K parameters amounting to approximately 15MB of VRAM space, while the CNN based constitutes to roughly 47K parameters amounting to roughly 55MB of VRAM space.

Next, we shall briefly describe the procedure for both implementations of the network which is followed by results and conclusion sections.



**Figure 3.** Training and validation plots for two different implementations of  $\Psi$ , namely GRU (left) and CNN (right) based.

## NETWORK TRAINING

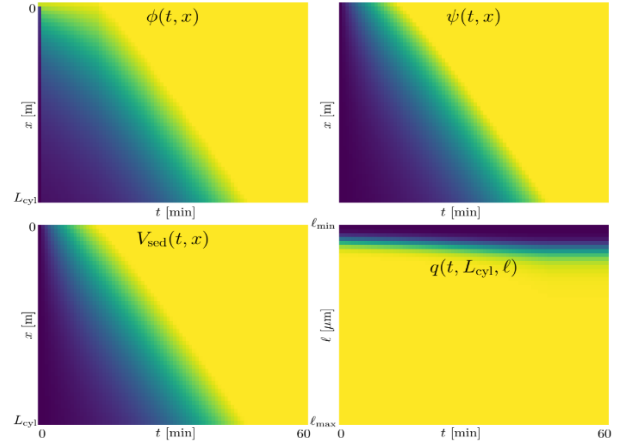
To train the model, we firstly generate a dataset using the experimental data  $\bar{\phi}_{\text{out}} = \bar{\phi}(t, L_{\text{cyl}}) = \phi(t, L_{\text{cyl}})/\phi_{\text{in}}$  for roughly 10-15 different time points. Since, these are temporally sparse, we interpolate the experimental data to obtain continuous representation as shown in **Figure 2**. Consequently, based on the different sequence lengths  $N$ , a  $k^{\text{th}}$  sample  $y_{t_k:t_k+N}^k$  at random time point  $t \in [0, T]$  drawn from experiment performed with parameter values  $\zeta^k$ . In order to obtain  $\zeta^k$  for different  $t_k$ , which requires knowing  $V_{t_k}$  and  $\phi_{t_k}$ , we used a novel bootstrapping method. In this method, first we used samples solely  $t_k = 0$  to obtain an estimate of  $\kappa_0 \approx .05$  and  $h_0 \approx 1$ . Using this we then simulated the process for the entire time interval  $[0, T]$ , thereby yielding us  $V_{t_k}$  and  $\phi_{t_k}$  for different values of  $t_k$ . This results in a dataset  $\mathbb{X} = (\xi^k)_{k \in \mathbb{N}}$ ,  $\xi^k = (y_{t_k:t_k+N}^k, \zeta^k)$ . Since, both the GRU and CNN implementations allows inputs of varying sequence lengths, we took  $N \in \{6, 12, 24, 48\}$  and trained them for 100 epochs with batch size of 50. The training was performed using Ryzen 7 CPU and GeForce RTX 5070 GPU. The obtained results are as show in **Figure 3**.

## RESULTS

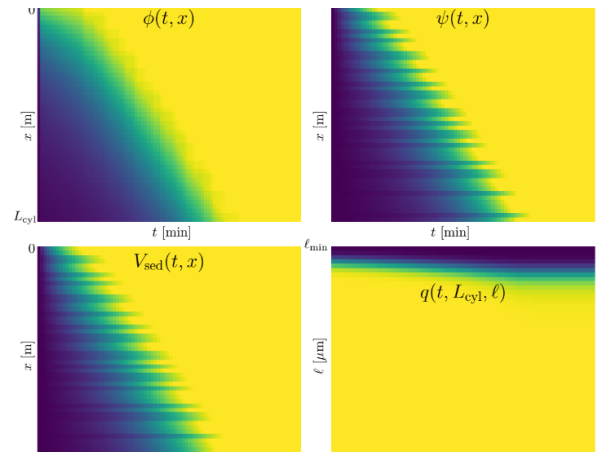
Based on the training and validation curves shown in Figure 3, both models, CNN and GRU, were able to perform the data fitting sufficiently well with training loss exhibiting significant reduction and in parallel to the validation loss. Subsequently, we tested both the models on the test set and observed that CNN model produces a smoother qualitative result while the GRU model shows signs of non-smoothness as depicted in **Figure 4**. and **Figure 5**.

Based on this we opted the CNN model to further investigate the predictions of the model. Firstly, we compared the estimated product loss with measured product loss for all three varying flow rates. Based on the top left subplot of Figure 5, we see that the estimates follow closely the measured data for lower flow rates while at higher flow rates it is over-estimating the product loss. Next, we also checked other unobserved qualitative metrics such as filling level  $U_{\text{fill}} = \frac{V_{\text{sed}}}{V}$ , grade efficiency  $T(t, L_{\text{cyl}}, \ell)$  along particle sizes at the cylinder end, separation efficiency  $E(t, x)$  along the cylinder length, cake height  $H_{\text{cake}}(x)$  along the cylinder length and finally the particle size density  $q_{\text{out}}(t, \ell)$  at the cylinder end. The plots for these functions are shown as different subplots in **Figure 6** for different time points. From plot for  $U_{\text{fill}}$  (top center) we can infer that the fill rate of the centrifuge increases with the

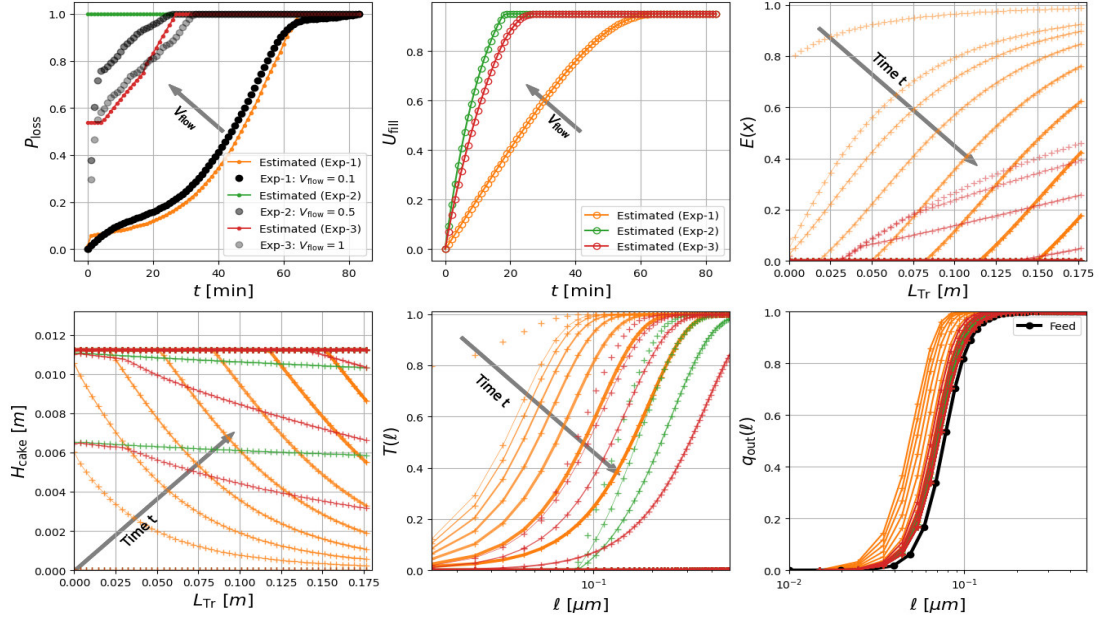
increase feed rate of the centrate. Next from the efficiency plot  $E(x)$  (top right) we can infer that separation efficiency decreases with time for each constant feed rate ( $V_{\text{flow}}$ ) and the spatial profile suggests hat the efficiency is relatively higher at the cylinder end compared to the beginning. Furthermore, at higher feed rates the efficiency is relative lower compared to the efficiency at lower feed rate. This is directly correlated to the increase in cake height  $H_{\text{cake}}$ , which is much faster for higher values of feed rates and is higher at the starting section of the cylinder. Observe that the trendline with respect to time (marked in gray) indicates that efficiency decreases with increasing time which is due to increase in cake height w.r.t time. Next looking at the separation rates  $T(t, L_{\text{cyl}}, \ell)$  we see that separation rate decreases more quickly (with respect to time) for smaller particles compared to the larger ones. This difference is even more drastic for higher feed rates which is due to the lesser residence time of smaller particles. Lastly, the plot for  $q_{\text{out}}(t, \ell)$  (bottom right) also indicates that particle density for smaller sizes increase over time at the cylinder end (seen as shift of the logistic curve to the left). Furthermore, this shift is more pronounced for smaller feed rates which can again be attributed to higher residence time of the particles.



**Figure 4.** Plots of the state variable obtained via the predicted parameters from CNN based  $\Psi$  for  $V_{\text{flow}} = 0.1$ .



**Figure 5.** Plots of the state variable obtained via the predicted parameters from CNN based  $\Psi$  for  $V_{\text{flow}} = 0.1$ .



**Figure 6:** Model validation for different performance indices of the tube centrifuge. Top left: Product loss vs time. Top center. Filling level vs. time. Top right. Separation efficiency vs. axial position of the rotor. Bottom left: Sediment height vs. axial position of the rotor. Bottom center: Grade efficiency vs. particle diameter. Bottom right. Density distribution function at the centrate vs. particle diameter.

The estimated values for the free parameters are  $h(t, x) \approx 1.1$  and  $\kappa(t, x) \approx .01$ . Thus, the model predicts the free parameters as constants whose values are in close agreement with the bootstrapping values obtained in the pre-training phase. Altogether, these observations indicate that, based on the sensor data,  $\Psi$  is able to produce parameters that qualitatively and quantitatively mimics the behavior of a tubular centrifuge thereby validating the dynamical behavior of the mathematical model (1.1) - (1.19).

## CONCLUSION

In this work we have proposed a novel continuous mathematical model for the dynamics of tubular bowl centrifuge. Based on that we first showed the wellposedness of the model which then paved way for simulations followed by experimental validation. For the validation of the model, we made use of NeuralODE approach where the solver is an integral part of the network. The free model variables are obtained via the parameterized part of the network which was implemented as GRU and CNN, among which the latter showed good qualitative and quantitative results, thus serving as the basis for building a DNN based observer.

## ACKNOWLEDGEMENTS

This work was funded by DFG (German Research Foundation) in the priority program- SPP2364 – grant no: 504852622

## REFERENCES

1. Winkler M, Rhein F, Nirschl H, Gleiss M. Real-time modeling of volume and form dependent nanoparticle fractionation in tubular centrifuges. *Nanomaterials*, 12(18):3161, 2022.
2. Narayanan R, El-Sayed MA. Shape-dependent catalytic activity of platinum nanoparticles in colloidal solution. *Nano letters*, 4(7):1343–1348, 2004

3. Kowalczyk B, Lagzi I, Grzybowski BA. Nanoseparations: Strategies for size and/or shape-selective purification of nanoparticles. *COCIS*, 16(2):135–148, 2011.
4. Seo YJ, et.al., Advantages of point of use (pou) slurry filter and high spray method for reduction of cmp process defects. *Microelectronic engineering*, 70(1):1–6, 2003.
5. Bendixen N, Losert S, et.al. Membrane particle interactions in an asymmetric flow field flow fractionation channel studied with titanium dioxide nanoparticles. *Journal of Chromatography A*, 1334:92–100, 2014.
6. Spelter LS, Nirschl H. Classification of fine particles in high-speed centrifuges. *Chemical engineering & technology*, 33(8):1276–1282, 2010.
7. Ambler CM. The theory of scaling up laboratory data for the sedimentation type centrifuge. *J. Biochem. Microbiol*, 1(2):185–205, 1959.
8. Gleiss M, Nirschl H. About modeling and optimization of solid bowl centrifuges. *KONA Powder and Particle Journal*, 41:58–77, 2024.
9. Hammerich S, Gleiß M. An efficient numerical approach for transient simulation of multiphase flow behavior in centrifuges. *Chem. Eng. Technol.*, 41(1):44–50, 2018.
10. Chen RTQ, et.al. Neural ordinary differential equations. *Neural Information Processing Systems (NIPS'18)*. Curran Associates Inc., Red Hook, NY, USA, 6572–6583.

© 2025 by the authors. This is an open access article under the creative commons CC-BY-SA licensing terms. Credit must be given to creator and adaptations must be shared under the same terms. See <https://creativecommons.org/licenses/by-sa/4.0/>

