DOI: 10.2195/lj_proc_ stadlthanner_en_202410_01

Parametrization of a Simulation Model for Flexible Small Consignments Using Parallel Bayesian Optimization

Parametrierung eines Simulationsmodells für forminstabile Kleinsendungen mittels paralleler Bayes'scher Optimierung

Dominik Stadlthanner, Gabriel Leitner, Christian Landschützer

Institut für Technische Logistik
Fakultät für Maschinenbau und Wirtschaftswissenschaften
Technische Universität Graz

The increasing volume of shipments in the courier, express, and parcel (CEP) sector and the transition from traditional rectangular cardboards to small consignments with flexible packaging pose challenges for manufacturers and operators of sorting technology. Existing sorting systems often fail to meet the requirements of these changes. While physics simulations are commonly used to develop new sorting systems, they are inadequate for modeling flexible consignments. A novel simulation approach for small consignments with flexible packaging was developed recently. Our study builds on that approach by parametrizing and validating the simulation. Initially, we analyzed the movement and transit times of four different types of small consignments on three test rigs with various conveying technologies, such as belt and roller conveyors, operated at different speeds. We then created a simulation model for one of the test rigs and incorporated existing models for two consignment types. We identified the most influential parameters on the simulation results through exhaustive parameter screening. Using Bayesian optimization with batch evaluations, we determined the optimal parameter values to minimize discrepancies between the simulated and real-world movement trajectories. Finally, we validated the simulation models against additional independent real-world experiments, finding that the simulation results fell well within the range of variation observed in these experiments.

[Keywords: CEP, small consignment, polybag, simulation, MFBD, parametrization]

Steigende Sendungszahlen in der Kurier-, Express- und Paket (KEP) Branche und ein Wandel im Sendungsspektrum von quaderförmigen Kartonagen hin zu forminstabilen Kleinsendungen stellen Hersteller und Betreiber von Sortieranlagen vor große Herausforderungen. Bestehende Anlagen werden den aktuellen Anforderungen häufig nicht mehr gerecht. Während der Einsatz von Physiksimulationen zur Entwicklung neuer Anlagen heute bereits gängige Praxis ist, sind diese für die Modellierung forminstabiler Kleinsendungen ungeeignet. Kürzlich wurde ein neuartiger Simulationsansatz für forminstabile Kleinsendungen entwickelt. Die vorliegende Arbeit erweitert diesen Ansatz durch Parametrierung der Simulationsparameter und Validierung der Simulationsergebnisse. Zunächst wurden die Bewegungstrajektorien und Durchlaufzeiten von vier verschiedenen Kleinsendungstypen auf drei Versuchsständen mit verschiedenen Fördertechnik-Modulen, wie Gurt- und Röllchenförderern bei unterschiedlichen Fördergeschwindigkeiten ermittelt. Anschließend wurde ein Simulationsmodell für einen der Versuchsstände erstellt und bestehende Modelle für zwei Sendungstypen integriert. Durch eine umfassende Parameteruntersuchung wurdend die Parameter mit dem größten Einfluss auf die Simulationsergebnisse identifiziert. Mittels paralleler Bayes'scher Optimierung wurden die optimalen Parameterwerte, bei denen die Unterschiede zwischen den simulierten und den realen Bewegungstrajektorien minimiert werden, ermittelt. Abschließend wurden die Simulationsmodelle anhand zusätzlicher unabhängiger Realversuche validiert. Dabei lagen die Simulationsergebnisse deutlich innerhalb der Schwankungsbreite dieser Versuche.

[Schlüsselwörter: KEP, Kleinsendung, Polybag, Simulation, MFBD, Parametrierung]

1 INTRODUCTION

The courier, express and parcel (CEP) sector has experienced rapid growth in recent years. One of the main contributors to this development has been the ever-increasing importance of e-commerce. In Germany, the number of CEP shipments has grown by 167% between 2000 and 2023, as reported in studies commissioned by the Bundesverband Paket und Expresslogistik e. V. (BIEK) [1], [2]. There was a slight decline in shipment numbers between 2021 and 2022, mainly due to the war in Ukraine and high inflation rates in Europe, but this effect was only temporary as volumes rose again in 2023. Figure 1 provides an overview of the historical CEP shipment numbers in Germany and the forecasted figures until 2028.

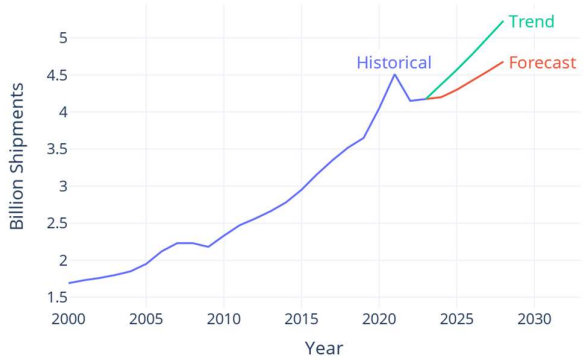

Figure 1: Historical and forecasted CEP shipment numbers in Germany until 2028 [1], [2].

Globally, this trend is even more pronounced: the number of parcel shipments has risen from 44 billion in 2014 to 161 billion in 2023, a remarkable increase of 266% [3], [4]. This trend is expected to continue both domestically and globally [2], [4].

Apart from the increase in CEP shipment numbers, the rise of e-commerce has been responsible for two additional major developments in the CEP industry.

First, there has been a shift in market share away from business-to-business (B2B) towards business-to-consumer (B2C) shipments. While B2C was only responsible for 52% of shipments in Germany in 2013, this share increased to 72% by 2023 [2]. The result of this development is a disproportionally rising number of potential delivery addresses, which in turn poses challenges for distribution center operators.

Second, there has been a shift in the CEP sector's shipment spectrum, with traditional rectangular cardboard boxes being increasingly replaced by small consignments which often feature flexible packaging materials. The reasons for this shift are manifold, but the growing importance of Asian countries in e-commerce, particularly China, which was reported by 37% of participants in a 2023 study as the country of origin for their most recent purchase [5], and the prevalence of low-value and low-weight purchases in e-commerce [6] are two of the most important factors.

Small consignments offer a cost-effective and space-efficient alternative to traditional cardboard packaging. However, they present challenges during the sorting process. They are often too large for efficient processing by mail sorters and too small for parcel sorters, which is why they are sometimes referred to as mixed-mail [7]. Additionally, they exhibit other unfavorable characteristics that complicate efficient handling, such as high variability in packaging materials, which leads to differing friction and flexural properties, and issues with contrast that hinder the visual identification of labels.

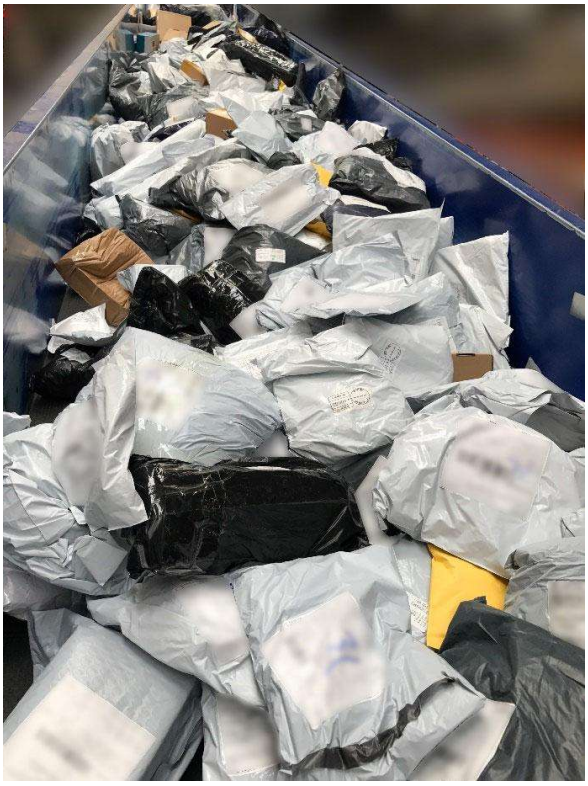

Figure 2: Mixed-mail after bulk unloading from the transportation containers onto a belt conveyor before singulation and sortation [8].

In contrast to traditional cardboard packaging, there are currently no established regulations for small consignments. The only existing definition, provided by Schadler et al. [8], categorizes small consignments and specifically plastic polybags (see Figure 2) based on their physical characteristics.

Small consignments use a variety of packaging materials, including plastic and kraft paper, some with air cushioning for added protection. This diversity in packaging materials and contents results in significant variability in physical characteristics such as flexibility, friction, size, and mass, which present challenges in handling and sorting. Currently, there are limited automated machine solutions for processing small consignments, leading to increased manual handling and higher costs. The minimum spacing required on sorters for mixed-mail items, especially with roller conveyors, is often greater than for traditional parcels to prevent missorting, which further reduces throughput and increases costs [8]. While private couriers can refuse problematic shipments, postal operators are required by the universal service obligation to deliver all types of mail and packages that meet basic requirements [9].

To improve the efficiency of mixed-mail processing, new automated material handling equipment tailored to

these needs such as the FreeFallSorter [10], a novel gravity-driven sorting concept, is necessary. Ideally, such equipment should be modular to adapt to changing market demands and packaging regulations. A key challenge in designing and testing new solutions is selecting appropriate test specimens due to the diverse nature of small consignments. Stadlthanner et al. [11] addressed this challenge by employing clustering to identify common categories of mixed-mail items, thereby simplifying the selection process.

Another challenge is the high cost and long development cycles associated with physical prototypes. Virtual prototyping has been successfully employed in various industries such as automotive and aerospace to drastically reduce development costs, and time to market and optimize products. Manufacturers of material handling equipment have also successfully adopted virtual prototyping using physics simulations [12]. However, these simulations currently only model rigid bodies, and do not account for the intricacies of small consignments with flexible packaging. To address this gap, Leitner et al. developed a simulation model for the simulation of small consignments with flexible packaging using multiple flexible body dynamics (MFBD), which combines multibody dynamics (MBD) and finite element analysis (FEA) [12], [13]. While initial tests showed that the model exhibits realistic behavior, the simulation results have yet to be validated against real-world experiments. The study identified the influence of various simulation parameters on the simulation results and determined optimal values through calibration tests, though it did not include material pairing parameters, such as friction values. Our study aims to address these gaps by introducing a novel approach to optimizing the simulation parameters for an MFBD simulation model for small consignments with flexible packaging, and by validating the simulation results against real-world experiments.

In the first part of our study, we conducted real-world experiments using four different test rigs along with four consignment types, based on the clustering results by Stadlthanner et al. [11]. The test rigs feature various conveying technologies, such as belt conveyors and roller conveyors, operated at different speeds. We analyzed these experiments using computer vision to track the trajectories of the test specimens and photoelectric sensors to measure transit times. In the second part of the study, we utilized existing simulation models developed by Leitner et al. [11] for two consignment types. We then modeled the relevant conveying technology for one of the test rigs and incorporated the consignment models into this setup. We performed parameter screening using factorial designs to identify parameters with statistically significant effects on the simulation results. We then used Bayesian optimization with batch evaluations to determine the optimal parameter values, aiming to minimize the discrepancies between the simulation results and real-world experiments. Finally, we validated the simulation models against additional real-world experiments not used in the optimization process.

Our study bridges the gap between current virtual prototyping capabilities and the advanced simulation models needed to meet the market demands posed by the prevalence of small consignments, thereby facilitating the development of more efficient material handling solutions.

2 METHODS

This study was conducted using a systematic multi-stage approach consisting of the following stages.

1. **Real-world experiments** (section 2.1): First, we conducted and analyzed real-world experiments, using three test rigs and four consignment types. The data gathered in this process served as the foundation for the subsequent parametrizations and validations of the simulation models.
2. **Simulation model** (section 2.2): Next, we created a simulation model for one of the test rigs and incorporated existing models for two consignment types.
3. **Parameter screening** (section 2.3): We then performed a screening of the simulation model parameters. This stage aimed to narrow down the parameter space and focus on the factors that have the largest impact on different performance indicators.
4. **Parametrization** (section 2.4): Building on the results of the parameter screening, we employed Bayesian optimization to find parameter values that minimize the difference between the simulation and real-world experiment results for different target values.
5. **Validation** (section 2.4.3): Finally, we validated the optimized simulation models against independent real-world experiments to ensure the generalizability of our models.

2.1 Real-World Experiments

2.1.1 Test Consignments

We used the four types of consignments shown in Figure 3 for our experiments. These consignment types are the result of an extensive cluster analysis by Stadlthanner et al. [11].

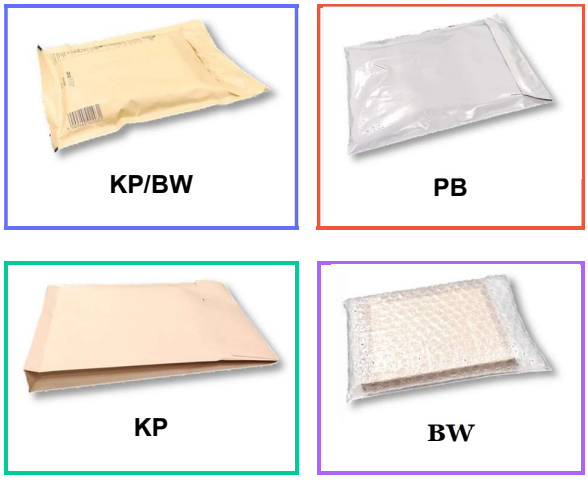

Figure 3: Overview of the consignment types kraft paper/bubble wrap (KP/BW), polybag (PB), kraft paper (KP) and bubble wrap (BW) used in this study [13].

The dimensions of the consignments and their contents as well as the mass of each consignment type are summarized in Table 1. Here, the column "consignment type" refers to different packaging types. KP/BW has an outer kraft paper (KP) and an inner bubble wrap (BW) layer, whereas consignment types KP and BW only have a single layer of the corresponding material. Polybag (PB) refers to a plastic bag made of low-density polyethylene (LDPE) and is the only consignment type that has a flexible (non-rigid) content.

Further details on the structure and manufacturing process of these consignments can be found in [13].

2.1.2 Experimental Setup

In our study we used the three test rigs "Belt-Belt" (E1), "Flow splitter" (E2), and "Induction" (E3). Each test rig was used for a series of experiments, in which the conveying speeds and the consignment types described in the previous section were varied. The speeds were set to two (for test rigs E2 and E3) and three (for test rig E1) predefined levels, respectively.[1] Each test case[2] was repeated ten times, using ten distinct test specimens of the corresponding consignment type to avoid wear-related effects.

Test rig E1, illustrated in Figure 4, is a serial combination of two belt conveyors, with a mean conveying speed of 0.62 m/s for the first conveyor for each experiment and 0.81 m/s, 1.00 m/s, and 1.57 m/s for the second one for the speed settings 1, 2 and 3, respectively. The test consignments were placed at an angle of 30° at a fixed position on the first conveyor.

Test rig E2, depicted in Figure 5, consists of conveyor belt modules and a flow splitter module – a matrix of pivoting roller units – and comprises two tests. In E2a, the test specimens were conveyed in a straight direction, while in E2b the test consignments were diverted onto a strip conveyor aligned at a 30° angle to the flow splitter module. The speed settings were 0.6 m/s and 1.0 m/s, respectively, for all conveyer modules in both E2a and E2b.

Table 1: Test consignments used in this study [13]. KP: Kraft paper, BW: Bubble wrap, PB: Polybag.

Consignment Type	Length [cm]	Width [cm]	Height [cm]	Length content [cm]	Width content [cm]	Height content [cm]	Mass [g]
KP/BW	27.5	20	2	16.5	12.5	2	140
PB	31.5	22.5	2	31	22.5	2	180
KP	32	25	3	31	22.5	3	380
BW	20	15	2	16.5	12.5	2	80

[1] We used a hand tachometer to compare the set speeds with the actual speeds, with deviations in the single-digit percentage range being measured depending on the setting. The speeds stated in this section refer to the measured values. While the speeds were largely constant, we were able to detect short-term deviations of up to 3% from the measured mean values.

[2] A test case is a given combination of test rig, test consignment and speed setting and comprises ten experiments with identical conditions.

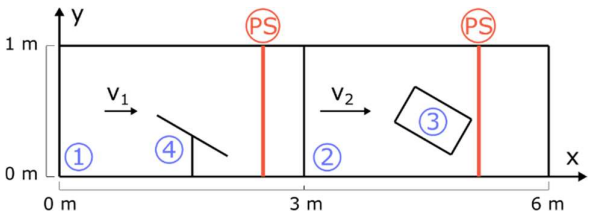

Figure 4: Schematic diagram of test rig E1 (not to scale). There are two belt conveyors (1 and 2) with the conveying speeds v_1 and v_2, respectively. Test consignments (3) are placed on conveyor (1) using a guide (4). Two photoelectric sensors (PS) are used to measure the transit times of test consignments between two pre-defined points.

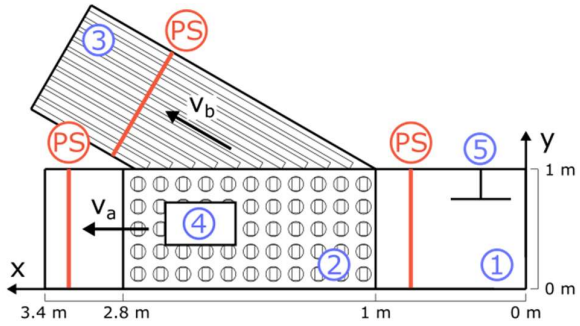

Figure 5: Schematic diagram of test rig E2 (not to scale). Test consignments (4) are placed on a belt conveyor (1) using a guide (5). In E2a, the test specimens are conveyed in a straight direction across the flow splitter module (2). In E2b, the specimens are diverted onto a strip conveyor (3). Photoelectric sensors (PS) are used to measure the transit times of test consignments between two pre-defined points.

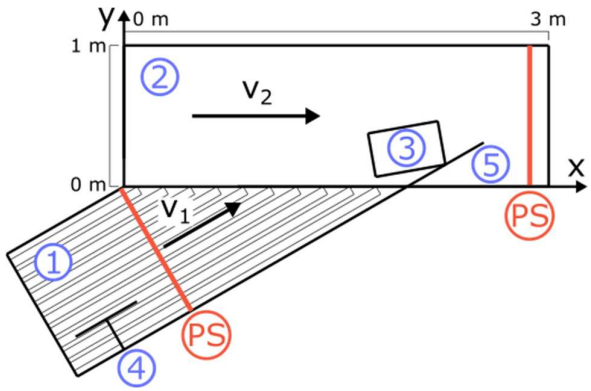

Figure 6: Schematic diagram of test rig E3 (not to scale). Test consignments (3) are placed on a strip conveyor (1) using a guide (4). A sheet metal plate (5) is used to guide the consignments towards the center of the belt conveyor (2). Photoelectric sensors (PS) are used to measure the transit times of test consignments between two pre-defined points.

Test rig E3, shown in Figure 6, consists of a strip belt conveyor and a conventional belt conveyor. The strip belt conveyor is 2 cm higher than the belt conveyor and is aligned at a 30° angle to it. For practical reasons, it was necessary to mount a metal plate to guide the test specimens toward the center of the belt conveyor. The conveying speed of the strip belt conveyor was 1.64 m/s for both speed settings, while that of the belt conveyor was 1.56 m/s and 1.86 m/s for speed settings 1 and 2, respectively.

We measured the transit times of the test specimens using SICK RAY26P-24162330A00 photoelectric sensors. We also utilized a downward-facing GoPro HERO 11 camera, which was mounted at a height of approximately 2 m above the conveyor level to record video footage of the experiments with a resolution of 5312 by 2988 pixels and a frame rate of 60 frames per second. We then postprocessed the videos in Python using the OpenCV library [14] and extracted movement trajectories. To achieve this, we attached AprilTags [15], a two-dimensional visual fiducial system commonly used in robotics, to the consignments to uniquely identify each consignment and track its x- and y-coordinate as well as its orientation during the conveying process. Another set of four AprilTags, placed along the side walls of the conveying equipment at known positions served as the coordinate system and was used to transform the image coordinates to real-world coordinates, with the use of a homography matrix. We validated the accuracy of this method in a controlled environment with AprilTags placed in various known locations, finding a maximum measurement error of 3 mm, representing a remarkable level of accuracy.

We then used the transit times and movement trajectories to select a reference experiment for each test case. To achieve this, we compared the differences in transit times, the translational trajectories of the consignment centers in the conveying plane, and the angular changes over time between all ten experiments of a given test case using distance matrices. The entries of the distance matrix $\boldsymbol{D}_t$ with respect to the transit times are defined by:

$$D_{\text{time}}(i,j) = \left(T_i - T_j\right)^2,$$

where T_i denotes the transit time of consignment associated with the experiment i. To quantify the distances between pairs of translational trajectories and pairs of angular changes, we used dynamic time warping (DTW) [16], an algorithm based on dynamic programming for measuring the (dis-)similarities between time series. Let $\text{DTW}(\boldsymbol{X}_i, \boldsymbol{X}_j)$ denote the DTW-distance between the time series of $\boldsymbol{X}_i(t)$ and $\boldsymbol{X}_j(t)$, where $\boldsymbol{X}_i(t)$ represent the two-dimensional coordinates of the center of the consignment associated with experiment i over time t. In the case of multi-dimensional time series, DTW requires the specification of a local or pointwise distance function – Euclidean distance in our case. We defined the entries of the distance matrix $\boldsymbol{D}_{\text{DTW}_{x,y}}$ with respect to translatory movement:

$$D_{\text{DTW}_{x,y}}(i,j) = \left(DTW\left(\boldsymbol{X}_i, \boldsymbol{X}_j\right)\right)^2 \qquad (1)$$

The squaring of the DTW-values was done to put a stronger emphasis on large deviations between time series. Similarly, for the rotatory motion, the entries of the distance matrix $\boldsymbol{D}_{\mathrm{DTW}_\theta}$ are defined as

$$D_{\mathrm{DTW}_\theta}(i,j) = \left(DTW(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)\right)^2,$$

where $\boldsymbol{\theta}_i(t)$ denotes the angular orientation around the z-axis[3] of the consignment associated with experiment i over time t. Since $\boldsymbol{\theta}_i(t)$ is one-dimensional, the local distance function for $\mathrm{DTW}(\boldsymbol{\theta}_i, \boldsymbol{\theta}_j)$ is trivially given by the absolute difference between two scalar values. DTW distances were computed using the Python implementation of dtw-Package [17], originally implemented for R.

Each distance matrix was then normalized by dividing all entries by the largest entry of the corresponding matrix, yielding the normalized distance matrices $\widetilde{\boldsymbol{D}}_{\mathrm{time}}$, $\widetilde{\boldsymbol{D}}_{\mathrm{DTW}_{x,y}}$ and $\widetilde{\boldsymbol{D}}_{\mathrm{DTW}_\theta}$. We then combined these matrices into a single combined distance matrix $\boldsymbol{D}_{\mathrm{combined}}$ by summing the respective normalized matrices:

$$\boldsymbol{D}_{\mathrm{combined}} = \widetilde{\boldsymbol{D}}_{\mathrm{time}} + \widetilde{\boldsymbol{D}}_{\mathrm{DTW}_{x,y}} + \widetilde{\boldsymbol{D}}_{\mathrm{DTW}_\theta}$$

Finally, to identify the reference experiment i^*, we calculated the vector of row sums[4] $\boldsymbol{s}$ of $\boldsymbol{D}_{\mathrm{combined}}$:

$$\boldsymbol{s}(i) = \sum_j \boldsymbol{D}_{\mathrm{combined}}(i,j)$$

Here, i^* is the row corresponding to the minimum value of $\boldsymbol{s}$:

$$i^* = \underset{i}{\operatorname{argmin}}\, \boldsymbol{s}(i)$$

The translatory and rotatory movements, $\boldsymbol{X}_{i^*}(t)$ and $\boldsymbol{\theta}_{i^*}(t)$ respectively, serve as target values for the parametrization of the simulation parameters.

2.2 Simulation Model

The advent of sophisticated computational techniques has significantly enhanced our ability to simulate and analyze complex mechanical systems. Among these techniques, multi flexible body dynamic (MFBD) is particularly noteworthy for its integration of the strengths of multibody dynamics (MBD) and finite element analysis (FEA). This approach allows for the detailed modeling of both the overall motion and the internal deformations of flexible consignments interacting with various sorting and conveying facilities.

2.2.1 Consignment Models

Leitner et al. [12] have proposed a bottom-up strategy for the modeling of flexible consignments in a step-by-step manner using the simulation software RecurDyn [18] and have described the modeling possibilities of various types of consignments. In a follow-up study [13], Leitner et al. further developed their approach to model consignments based on the clustering results by Stadlthanner et al. [11]. The authors also provide a detailed analysis of the various parameters and their effects on various simulation results.

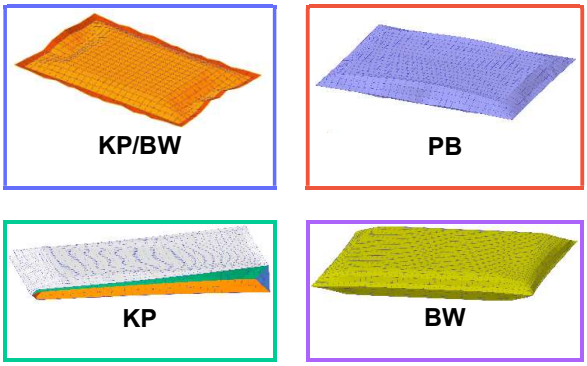

Figure 7: Simulation models for the consignment types kraft paper/bubble wrap (KP/BW), polybag (PB), kraft paper (KP) and bubble wrap (BW) used in this study [13].

The consignment contents can be categorized as either rigid or flexible, which influences their modeling approach. Rigid contents, found in consignments KP/BW, KP, and BW, maintain their structural integrity during transportation, making them suitable for rigid body modeling. Each model is adjusted for density and geometric parameters and is represented as simple cuboids with rounded edges. In contrast, flexible contents, such as those in consignment PB, require a more sophisticated approach. The use of finite volume elements allows for the effective simulation of material flexibility, thereby capturing deformations and movements under varying loads [13]. The packaging layers are modeled as finite shell elements, connected at their edge nodes, and pulled apart by external forces. The contents are then placed between the layers and the external forces are removed causing the packaging to enclose the contents through contacts, thus completing the consignment models, as shown in Figure 7. These consignment models can then be used in various experimental setups. A more detailed description of the modeling of the consignments and the various parameters can be found in [13]. Our study uses models for consignment types KP/BW and PB.

[3] Coordinate systems according to Figures Figure 4 to Figure 6.

[4] Since all distance matrices are symmetric, column sums could also be used.

2.2.2 Experimental Setup

Due to time and budget restrictions, the simulations in our study are limited to the E3 test rig setup. Future research will focus on simulating the other test rigs. The simulation model for test rig E3 is made up of flexible conveyor belts, metal sheets, and rollers. Shell elements are used to model the conveyor belts. The belts are driven by rollers, with one driven roller and one passive roller per belt. In addition to conveyor belts and rollers, numerous plates are required to complete the model. These are modeled as rigid bodies and placed appropriately to complete the overall models. Figure 8 shows the simulation model for test rig E3 with consignment type PB.

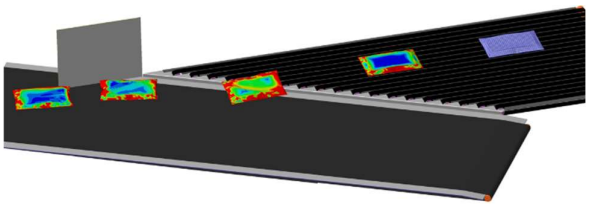

Figure 8: Simulation model of test rig E3 showing the movement of consignment type PB at different time steps.

2.3 Parameter Screening

The simulation models described in section 2.2 contain a wide range of physical parameters. While a small number of parameters such as Poisson's ratio were set to values from the literature or left at default values, a total number of 16 and 18 parameters were initially investigated for consignment types KP/BW and PB[5], respectively. To limit the search space for the parametrization, we conducted a three-stage parameter screening using the response values $\mathrm{DTW}_{x,y}$ and DTW_{θ}. These values represent the DTW distances between the translatory and rotatory movements of the reference experiment $\boldsymbol{X}_{i^*}(t)$ and $\boldsymbol{\theta}_{i^*}(t)$ respectively, and the corresponding simulation results. In the first stage, we employed a two-level fractional factorial design of resolution IV for the 16 parameters associated with test rig E3 and consignment type KP/BW at speed setting 2 to eliminate any insignificant parameters. In the second stage, we employed a two-level full factorial design using the remaining parameters, to more accurately quantify their main effects as well as their interaction effects on the response values. We then used the results of this screening stage to inform our decision on the parameters used for the parametrization of test rig E3 and KP/BW consignment type. In the final stage, we conducted another two-level full factorial design, incorporating all statistically significant parameters from the second stage as well as the two additional parameters for the PB consignment type as the basis for the parameter selection for the E3 test rig and the PB consignment type.

It should be noted that the selection of parameters for parametrization always represents a compromise between computational effort and model accuracy and it therefore rarely makes sense to blindly use all statistically significant parameters for parametrization. For this reason, we set an upper limit of four parameters for the parametrization in this study due to the high computational cost of the simulations and selected the parameters based on the effect sizes in stages two and three.

Table 2 offers an overview of the parameters and corresponding parameter ranges analyzed in stages two (KP/BW) and three (PB) of the parameter screening process. While the parameters E_{Co} and ζ_{Co} don't exist in the simulation model for consignment type KP/BW due to its rigid content, the parameters k_{MePa} and $\mu^{\mathrm{s}}_{\mathrm{MePa}}$ were excluded from the full factorial design in stage three since their effect sizes in stage two were negligible.

2.4 Parameter Optimization

2.4.1 Optimization Method Selection

The task of finding a set of parameter values for a physics simulation that best approximates real-world behavior is an optimization problem in which the difference between a real-world experiment and a corresponding simulation is to be minimized with respect to one or several target values. Over the years, a myriad of optimization algorithms belonging to different classes have been proposed, all of which have their strengths and weaknesses, making it difficult to select a suitable algorithm for a particular optimization problem. For this reason, numerous attempts have been made to create taxonomies for different optimization algorithms. One of the most comprehensive taxonomies is the one proposed by Stork et al. [19], which distinguishes between exact[6], hill-climbing, trajectory-based, population-based, surrogate-based, and hybrid methods.

[5] Consignment type PB has a flexible content that requires the specification of the additional parameters Young's modulus and damping ratio.

[6] Exact methods are reviewed in [19] but not part of the proposed taxonomy. Due to the nature of the optimization problem associated with the parametrization of simulation models, these methods will not be explored in our study.

Table 2: Parameters investigated using two full factorial designs for consignment types kraft paper/bubble wrap (KP/BW) and polybag (PB). The underlying physical models implemented in the simulation software RecurDyn are summarized in [13].

	Parameter	Description	Unit	Range (default value)	
				KP/BW	PB
1	k_{CoPa}	Stiffness coefficient: content/packaging	N/mm	100 – 1000 (500)	100 – 1000 (500)
2	k_{MePa}	Stiffness coefficient: metal/packaging	N/mm	100 – 1000 (611)	-
3	$\mu^{\mathrm{d}}_{\mathrm{BePa}}$	Dynamic friction coefficient: belt/packaging	-	0.3 – 0.6 (0.45)	0.3 – 0.6 (0.45)
4	$\mu^{\mathrm{d}}_{\mathrm{MePa}}$	Dynamic friction coefficient: metal/packaging	-	0.15 – 0.35 (0.25)	0.15 – 0.35 (0.25)
5	μ^{s}_{MePa}	Static friction coefficient offset: metal/packaging	-	0 – 0.3 (0.15)	-
6	E_{Pa}	Young's modulus: packaging	N/mm²	200 – 1600 (1000)	50 – 200 (100)
7	E_{Co}	Young's modulus: content	N/mm²	-	0.05 – 0.5 (0.08)
8	ζ_{Co}	Damping ratio: content	-	-	0.01 – 0.1 (0.04)

Hill climbing algorithms are local search algorithms, based on greedy exploitation with minimal exploration, and have fast convergence to a local optimum. Well-known examples are the Nelder-Mead method [20] and various gradient descent and stochastic gradient descent algorithms [19].

Trajectory-based algorithms are similar to hill-climbing algorithms but place a stronger emphasis on exploration by using information from successive function evaluations. These algorithms iteratively develop an initial solution, accepting worse solutions in some circumstances to avoid local optima. Simulated annealing [21] and tabu search [22] are prominent examples of this class [19].

Population-based approaches, in contrast to trajectory-based approaches, develop multiple solutions simultaneously, often utilizing search strategies inspired by processes found in nature. In some cases, two or more solutions may be combined to create new solutions. Due to their distributed nature and strong emphasis on exploration, these approaches often require a relatively high number of function evaluations and tend to be slow to converge. Well-known examples include particle swarm optimization [23] and genetic algorithms [19].

Surrogate-based optimization aims to estimate the behavior of the target function by fitting a regression or surrogate model to data obtained from function evaluations. This is particularly useful for functions that are expensive to evaluate, as surrogate evaluations are usually several orders of magnitude faster. Initially, the function is sampled at various points using sampling strategies such as factorial designs or Latin hypercube sampling (LHS) [24], and a surrogate model is constructed from these observations. This surrogate model then identifies new candidate points for evaluation, typically by optimizing an acquisition function. The purpose of an acquisition function is to solve the exploration-exploitation dilemma by combining uncertainty (exploration) and predicted performance (exploitation) of surrogate model predictions into a single value. Common acquisition functions include expected improvement (EI) [25] and upper confidence bound (UCB) [26]. Due to the high computational performance of surrogate model predictions, sophisticated optimization algorithms can be used when searching for new candidate points. In each iteration, one or more new candidate points are evaluated by the original function, updating the surrogate model. This process continues until a stopping criterion is met. The effectiveness of surrogate-based methods depends on the choice of surrogate model, acquisition function, and the associated

hyperparameters [19]. While this description mainly applies to Bayesian optimization (BO), arguably the most important sub-category of the surrogate-based optimization class, other surrogate-based approaches may follow different steps.

Methods that are based on the combination of concepts from one or more of the optimization classes outlined above are labeled hybrid approaches in the guideline. Since the class of hybrid approaches is large and heterogeneous, it will not be discussed in more detail in this study.

In addition to the algorithm taxonomy, the authors in [19] propose an algorithm selection guideline, which is shown in Figure 10.

The optimization problem in our study has the following properties:

- black box function with complex nonlinear fitness landscape and multiple local optima
- multiple objectives (minimizing both $\mathrm{DTW}_{x,y}$ and DTW_{θ} with respect to the reference experiment)
- expensive evaluations[7]
- noisy observations due to numerical noise and stochastic effects
- constrained continuous input variables
- very small evaluation budget

Applying the algorithm guideline to these properties, it is obvious that most algorithm classes are not suitable for the given problem. Surrogate-based methods are the most appropriate optimization class for the problem at hand. Bayesian optimization and similar surrogate-based approaches are commonly used in parametrization and calibration tasks in other domains such as discrete element method (DEM) simulations (e. g. [27], [28], [29], [30], [31]). For this reason, we opted for Bayesian optimization in our study. The following section describes the details of the algorithm we used.

2.4.2 Bayesian Optimization with Batch Evaluations

The surrogate-based approaches discussed in section 2.4.1 apply to most forms of BO. Figure 9 illustrates the concepts underlying BO.

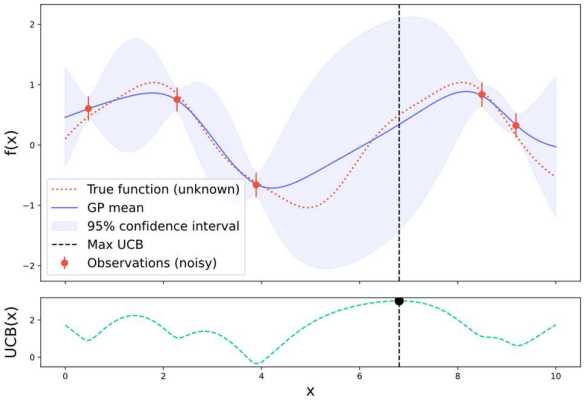

Figure 9: Bayesian optimization illustration: The Gaussian process (GP) was fitted to (noisy) observations from function evaluations. The bottom plot shows the Upper Confidence Bound (UCB) acquisition function, with its maximum indicating the next candidate point.

A black-box function is evaluated at various points, with observations being noisy, as indicated by the error bars. Next, a surrogate model—a Gaussian process in this case—is fitted to these observations. Depending on the chosen kernel and its hyperparameters, a Gaussian process can model uncertain observations. Consequently, the model's predictions at previously observed points may differ from the actual observed values, and the variance remains greater than zero, as shown by the non-vanishing confidence interval areas at the observation points. The bottom plot shows the acquisition function, which is to be maximized. The input value corresponding to the maximum of the acquisition function is the next candidate point. Upper Confidence Bound (UCB) is a commonly used acquisition function as it is simple and has been shown to perform well in benchmarks [26]. It is defined as

$$\mathrm{UCB}(x) = \mu(x) + \lambda\sigma(x), \qquad (2)$$

where $\mu(x)$ and $\sigma(x)$ are the mean and standard deviation of the surrogate model at sample point x, respectively and λ is a hyperparameter, that can be tuned to balance exploration (σ) and exploitation (μ) explicitly.

The search in BO is inherently sequential, since the surrogate model is updated every time new candidate points are evaluated. However, numerous attempts have been made to parallelize BO by evaluating multiple points simultaneously (refer to [32], [33], [34]). In this study, we employed a parallelizing approach based on k-means clustering of the acquisition space, which has demonstrated excellent performance in benchmark studies [35], [36], [37].

[7] Simulating the movement of a consignment on test rig E3, with a simulated time of under two seconds, takes about five hours.

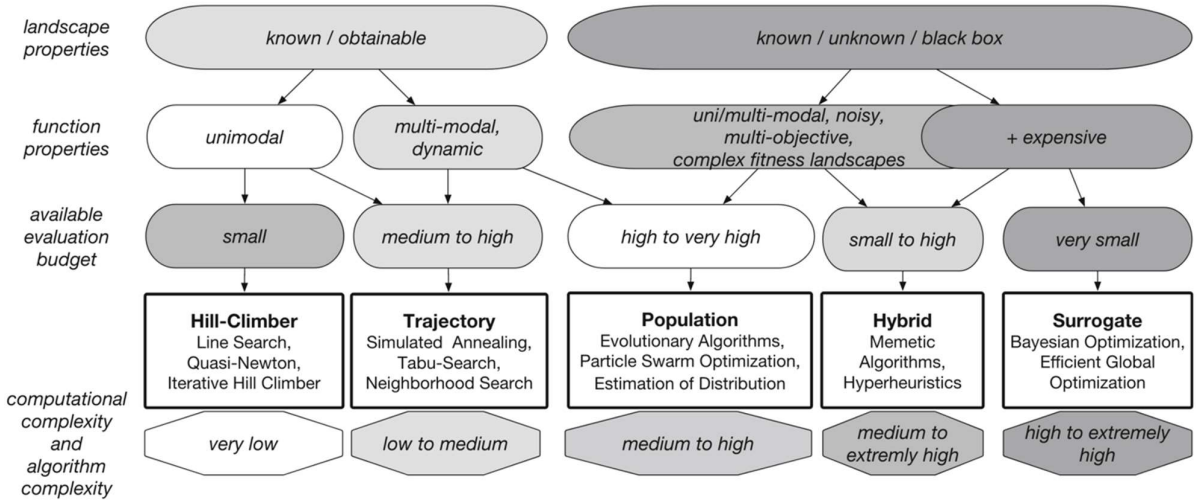

Figure 10: Algorithm selection guideline, mapping function properties, and evaluation budget to suitable algorithm classes [19].

2.4.3 Proposed Parametrization Method

Figure 11 provides a flowchart of the algorithm used in our study. For the optimization loop (dashed rectangle in Figure 11) we used the Python library Mango[8] [36], [37], which is partially based on the algorithm proposed by Groves and Pyzer-Knapp [35].

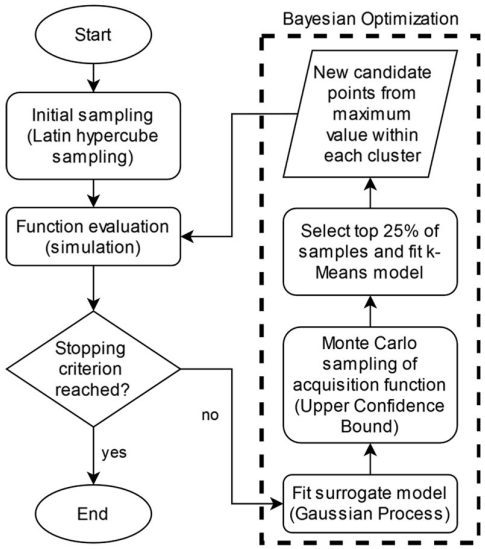

Figure 11: Flowchart of the optimization approach used in this study.

The optimization starts with an initial sample. We used LHS with a sample size of 16. Next, 16 copies of the simulation model were created, and the model parameters were adjusted accordingly. These simulations were run in parallel and postprocessed, once completed. The cost function to be minimized is a weighted sum of $\widetilde{\mathrm{DTW}}^2_{x,y}$ and $\widetilde{\mathrm{DTW}}^2_{\theta}$.

$$c = w_{x,y}\widetilde{\mathrm{DTW}}^2_{x,y}\left(i^*, \mathrm{sim}_{\boldsymbol{p}}\right) + w_{\theta}\widetilde{\mathrm{DTW}}^2_{\theta}(i^*, \mathrm{sim}_{\boldsymbol{p}}) \quad (3)$$

Here $\widetilde{\mathrm{DTW}}_{x,y}(i^*, \mathrm{sim}_{\boldsymbol{p}})$ and $\widetilde{\mathrm{DTW}}_{\theta}(i^*, \mathrm{sim}_{\boldsymbol{p}})$ represent the normalized DTW distances between the simulation time series $\boldsymbol{X}_{\mathrm{sim}_{\boldsymbol{p}}}(t)$ and $\boldsymbol{\theta}_{\mathrm{sim}_{\boldsymbol{p}}}(t)$ with parameter set $\boldsymbol{p}$ and the corresponding values of the reference experiment i^*. We used the same normalization factors as in the normalization of the distance matrices described in section 0 and set the weights $w_{x,y}$ and w_{θ} to 1. We excluded segments of the conveying systems from the DTW comparisons where the test consignments interacted solely with a single belt conveyor without any obstacles. Our focus was on the "interesting" sections of the technology, ensuring that comparisons were limited to areas with more complex interactions. Small safety margins around these "interesting" sections were included to capture inertia-induced movement in the analysis. Furthermore, in the DTW comparisons, we considered only the relative changes and not the absolute values of $\boldsymbol{X}(t)$ and $\boldsymbol{\theta}(t)$ with respect to the initial positions and orientations. This approach minimized the impact of variations in the initial placement of the test consignments on the resulting DTW distances.

[8] Source code available on GitHub [38].

In the next step of the algorithm, a surrogate model is initialized from the points of the initial sample and the corresponding cost values.[9] In our study, the surrogate model was a Gaussian process with a Matérn kernel where the hyperparameter ν, which controls the smoothness of the resulting model, was set to 5/2. Matérn kernels are a popular choice due to their high flexibility [39] and good performance in benchmark studies [40]. For this we used the Gaussian process implementation of the Python library Scikit-learn [41]. It was necessary to allow for noisy observations in the Gaussian process to avoid numerical problems and overfitting. This was done by adjusting the regularization parameter alpha. Experimentation showed that setting alpha to 0.35^2, where 0.35 is the standard deviation of the noise, represents a good balance between model accuracy and generalization.

The algorithm then uses Monte Carlo sampling to maximize the UCB acquisition function. For each sample point, the surrogate model makes a prediction that yields the mean cost value and standard deviation, which are used to calculate the UCB according to eq. (2). The 25% of sample points with the highest UCB values are then selected and clustered using k-means, with $k = 4$ in our case. The clustering separates the sample points into spatially distinct regions. For each cluster, the sample point with the largest UCB value is chosen as a new candidate point, and all four candidate points are then simulated in parallel and the surrogate model is updated with the new simulation results. This process continues until a stopping criterion is reached.

In our study, we implemented a stopping criterion based on validation experiments, inspired by the training protocols of artificial neural networks. After every five iterations, the optimization is paused to perform a validation experiment. This involves simulating a validation run with the best parameters identified so far but at a speed setting not used for the parametrization. If the validation is successful, the optimization process terminates; otherwise, it continues for an additional five iterations. Our study considers three validation criteria:

1. **Transit time**: The transit time of the simulated consignment must fall within the range observed in the real-world experiments.
2. **$\mathrm{DTW}_{x,y}$**: The simulation experiment is added to the set of real-world experiments and the distance matrix is computed using the $\mathrm{DTW}_{x,y}$ metric as defined in eq. (1). For each experiment, the mean distance to all other experiments is calculated. Validation is considered successful if the mean distance of the simulation is not greater than the largest mean distance of the real-world experiments.
3. **DTW_{θ}**: This criterion is similar to the previous one but uses the DTW_{θ} metric instead of $\mathrm{DTW}_{x,y}$.

Using a stopping criterion based on validation experiments offers three primary advantages over criteria directly tied to the optimization process. First, the optimization halts as soon as a valid parameter set is identified, limiting unnecessary iterations to a maximum of four. Second, it circumvents the need to define arbitrary boundaries such as objective value thresholds, total iteration runs, or patience windows[10]. Third, while termination criteria based on optimization processes can be susceptible to overfitting, our approach helps mitigate this risk by allowing for earlier termination.

3 Results and Discussion

3.1 Real-World Experiments

Depending on the test rig and speed setting, we observed statistically significant differences in the transit time distributions between the different consignment types (refer to Figure 13). We utilized pairwise Mann-Whitney U tests and adjusted the p-values using the Benjamini-Hochberg procedure [42] to control the false discovery rate (FDR) at 5%. The results indicate that the PB consignment type consistently exhibited the slowest or second-slowest median transit times across all test rigs and speed settings. This consignment type is distinguished by its flexible packaging and limp contents, features that are considered problematic by manufacturers and operators of sorting systems. These findings confirm the perceived issues with polybags as reported by experts in the field and are in line with prior findings by Schadler et al. [8] regarding transit time delays for consignments with flexible packaging on flow-splitting equipment. Our study extends these previous findings by incorporating additional test rigs and varying conveyor speeds.

The BW consignment type, the smallest and lightest in this study, achieved the fastest transit times on test rig E1. This behavior can be partly explained by the small dimensions of this type, which facilitate a quicker transition from the first to the second belt, resulting in faster acceleration to belt 2 speed, compared with the other consignment types. However, this speed advantage does not exist or is reversed in the other test rigs, especially at speed settings 2 and 3, where there is also a comparatively large variance in transit times for BW. In test rigs E2a and E2b, both featur-

[9] For practical reasons we maximized $-c$ rather than minimizing c in our approach.

[10] Number of iterations to wait before the optimization stops if no progress is made.

ing a flow splitter module, this variance is caused by uncontrolled collisions of BW consignments with flow splitter rollers, resulting in chaotic jumping behavior observed at high conveying speeds.

Finally, in test rig E3, transit times are primarily influenced by the degree of rotation a consignment experiences. Consignment type BW is particularly prone to rotation, often being rotated by more than 180°, while consignment type KP, on the other extreme, experiences minimal rotation during the conveying process. The rotation depends on the consignment's physical properties and the relative angle at which it collides with the sheet metal plate. Upon collision, consignments either slide along the sheet metal or flip around, with the extent of rotation significantly impacting their transit times.

3.2 Parameter Screening

We conducted an initial parameter screening for test rig E3 and consignment type KP/BW using a fractional factorial design of resolution IV (see section 2.3). Among the 16 parameters included in the fractional factorial design k_{CoPa}, k_{MePa}, μ^{d}_{BePa}, μ^{d}_{MePa}, $\mu^{\mathrm{s}}_{\mathrm{MePa}}$ and E_{Pa} exhibited statistically significant effects ($p < 0.05$). These parameters were subsequently analyzed using a full factorial design. Figure 14 shows the standardized effect sizes of these factors and their two-factor interactions from the full factorial design for the response values $\mathrm{DTW}_{x,y}$ (top) and DTW_{θ} (bottom). The effect sizes, sorted from largest to smallest, are limited to the 10 largest effects. The parameters k_{MePa} and μ^{s}_{MePa}, along with any interactions involving them, did not show statistically significant effects, while the other four parameters did exhibit significant effects. Despite its significant impact on $\mathrm{DTW}_{x,y}$, we excluded k_{CoPa} from the parameter optimization due to slow convergence when k_{CoPa} was included. The parameter screening for consignment type PB was based on the screening results for consignment type KP/BW. We conducted a single full factorial design involving the two parameters E_{Co} and ζ_{Co}, unique to the consignment type, along with the four significant factors of the previous full factorial design and adjusted parameter levels (see Table 2). The results of this design, presented in Figure 15, show significant effects for the parameters μ^{d}_{BePa}, μ^{d}_{MePa}, E_{Pa}, E_{Co} and ζ_{Co}. Due to the small effect sizes of ζ_{Co} and its interactions, we excluded this parameter from the parameter optimization to limit the size of the search space.

3.3 Parameter Optimization and Validation

3.3.1 Optimization Results

For the parametrization of the simulation models associated with test rig E3 and consignment types KP/BW and PB, we employed the approach detailed in section 2.4.2. LHS with a sample size of 16 was used for the initial sampling. In the Bayesian optimization process, we utilized a batch size of four, meaning that each iteration comprised four parallel simulations with varying parameters. Figure 12 illustrates the progression of the best objectives (as defined by the cost function in eq. (3)) found by the optimization for both consignment types.

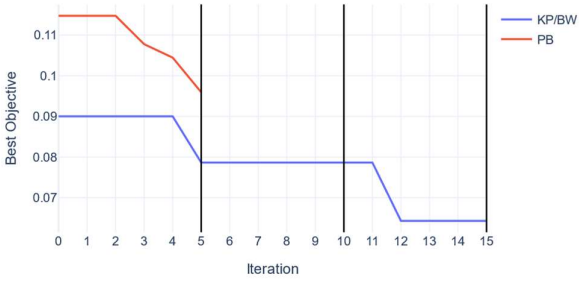

Figure 12: Best objectives found by iteration for consignment types KP/BW and PB. Validation experiments, indicated by the vertical lines, were performed every five iterations. The optimization was stopped once the validation was successful.

After every five iterations, validation experiments were conducted using the best parameters found until this point. Note that in the case of KP/BW, the validation was omitted after the tenth iteration since no new best solution had been found. The optimization process was halted once all three validation criteria (as defined in section 2.4.2) were satisfied. For consignment type KP/BW the best parameter set was identified after twelve iterations and validated after 15 iterations. Considering the initial sample size of 16, a batch size of four and the two validation experiments the total number of simulations performed was 78 in this case. For consignment type PB, valid parameters were identified after just five iterations, resulting in a total of 37 simulations. The computations were performed on an AMD Ryzen 9 7950X 16-core CPU, requiring approximately 120 hours for consignment type KP/BW and 52 hours for consignment type PB. The identified parameter sets for both consignment types are summarized in Table 3.

Table 3: Optimized parameter values for consignment types KP/BW and PB with test rig E3.

Parameter	Unit	Value	
		KP/BW	PB
$\mu^{\mathrm{d}}_{\mathrm{BePa}}$	-	0.300	0.565
$\mu^{\mathrm{d}}_{\mathrm{MePa}}$	-	0.346	0.350
E_{Pa}	N/mm²	506	196
E_{Co}	N/mm²	-	0.358

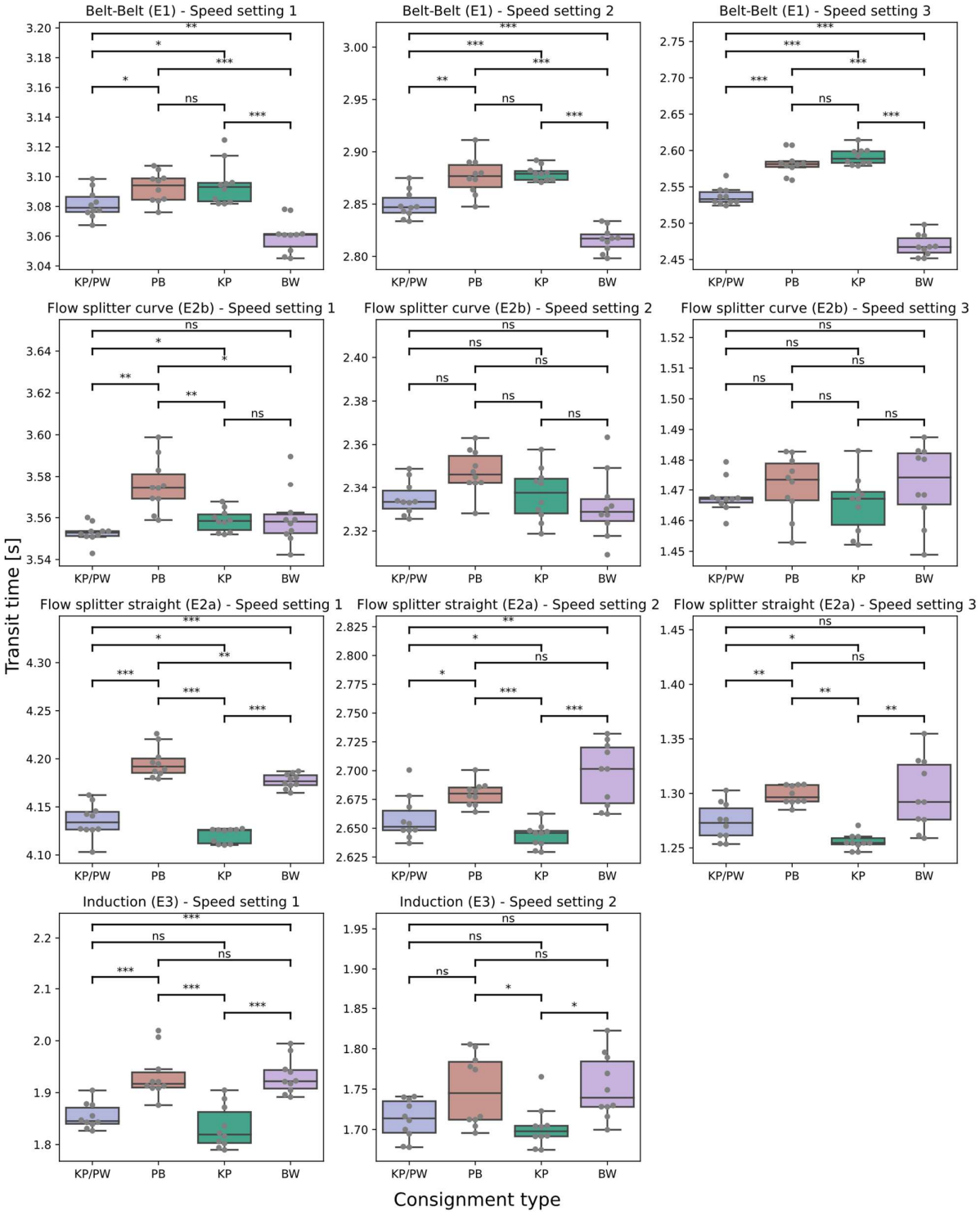

Figure 13: Comparison of transit time distributions between different consignment types by test rig and speed setting for the real-world experiments. The comparisons were performed using pairwise Mann-Whitney U tests with Benjamini-Hochberg correction [42] to control the false discovery rate at 5%.

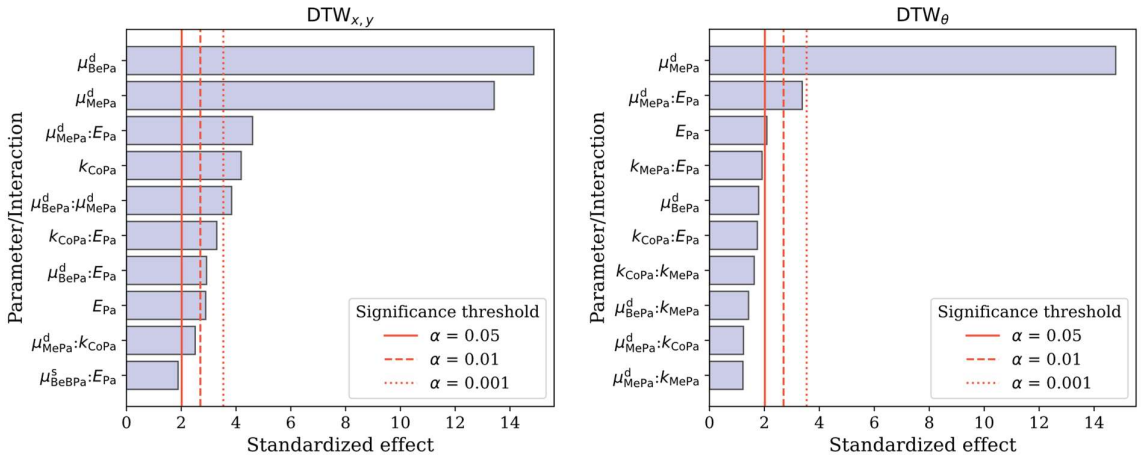

Figure 14: Pareto plots of standardized effects with respect to the objectives $DTW_{x,y}$ $(left)$ and DTW_θ $(right)$ for the simulation of test rig E3 with consignment type KP/BW.

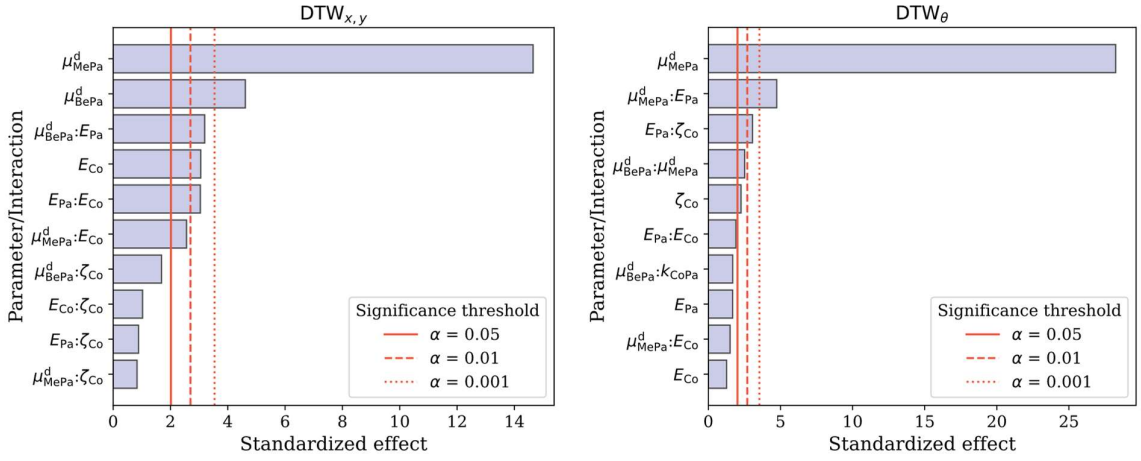

Figure 15: Pareto plots of standardized effects with respect to the objectives $\mathrm{DTW}_{x,y}$ $(left)$ and DTW_θ $(right)$ for the simulation of test rig E3 with consignment type PB.

3.4 Validation Results

The validation was performed on test rig E3, using the same experimental setup as for the parametrization, but while speed setting 1 was used for optimization, speed setting 2 was used for validation. The validation results for KP/BW are summarized in Figure 16. The chart on the left shows the distribution of transit times of all real-world experiments as well as the simulation (blue diamond). While the transit time for the simulation is at the upper end of the distribution, it is still clearly within the range of fluctuation of the real-world experiments. The center and right charts show the distributions of the mean $\mathrm{DTW}_{x,y}$- and DTW_θ-distances for all real-world experiments and the simulation experiment. For both cases, the simulation has a lower-than-average mean distance, indicating that the simulation results are in very good agreement with the real-world experiments. This fact is further emphasized when comparing the trajectories and changes in angle between the real-world experiments and the simulation (see Figure 17).

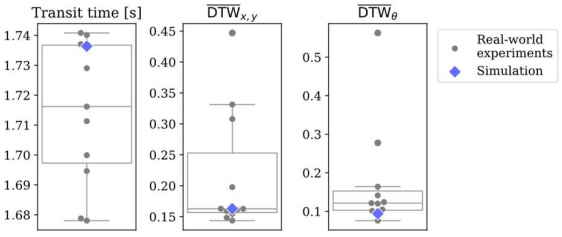

Figure 16: Boxplots of the validation experiments for the simulation of test rig E3 with consignment type KP/BW at speed setting 2. For $\overline{DTW}_{x,y}$ and $\overline{DTW}_{\theta}$ the mean DTW distance of an experiment to the other experiments was calculated.

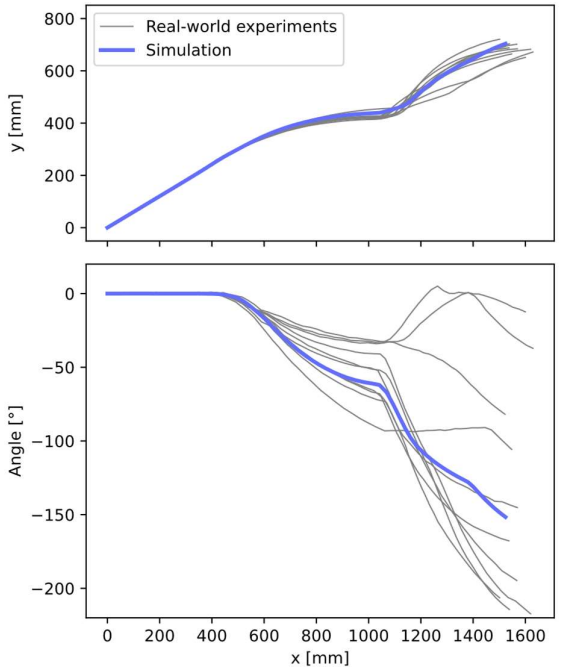

Figure 17: Changes in x- and y- y-coordinates (top chart) and changes in angle over x-coordinate (bottom chart) of real-world experiments (gray) and simulation (blue) for consignment type KP/BW on test rig E3 at speed setting 2.

The corresponding validation results for consignment type PB are presented in Figure 18 and Figure 19. Although all three validation criteria are met for this consignment type, the trajectory comparison (Figure 19, top chart) reveals some discrepancies between the real-world experiments and the simulation. Specifically, the real-world consignment contents exhibited a greater tendency to deform upon impact with the sheet metal, causing the centers to shift further to the right, a phenomenon not replicated in the simulation. It is unclear if this discrepancy arises from incorrect parameter selection (especially E_{Co}) or inaccuracies in content modeling (refer to section 2.2.1). This finding further highlights some limitations of the validation criteria employed in our study. Despite the simulation achieving a below-average mean $\text{DTW}_{x,y}$, thus satisfying this validation criterion, the trajectory comparison indicates a small yet distinct difference between simulation and real-world experiments.

Overall, while our findings indicate that the simulations generally align well with real-world data, some refinements to the validation criteria may be necessary.

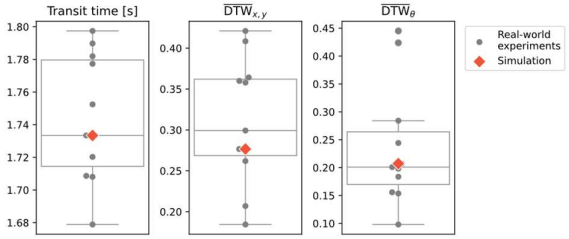

Figure 18: Boxplots of the validation experiments for the simulation of test rig E3 with consignment type PB at speed setting 2. For $\text{DTW}_{x,y}$ and DTW_{θ} the mean DTW distance of an experiment to the other experiments was calculated.

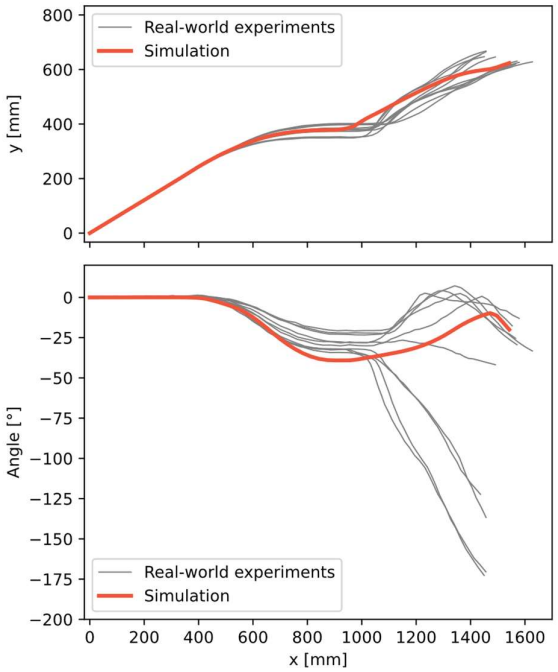

Figure 19: Changes in x- and y- y-coordinates (top chart) and changes in angle over x-coordinate (bottom chart) of real-world experiments (gray) and simulation (red) for consignment type PB on test rig E3 at speed setting 2.

4 Conclusion and Outlook

This study aimed to analyze the movement behavior of different types of small consignments on three test rigs comprising various kinds of conveying technologies. In the first part of the study, we conducted real-world experiments

to measure transit times and movement trajectories. Our findings revealed statistically significant differences in transit times between the different consignment types, which we discussed in detail. In the second part of the study, we created a simulation model for one of the test rigs and incorporated existing simulation models for two consignment types. We performed a multi-stage parameter screening to assess the impact of various simulation parameters, such as friction coefficients and flexural properties, on the simulation results. Subsequently, we utilized Bayesian optimization with batch evaluations to determine the values for the most influential parameters. The parametrized simulation models were validated by comparing their outcomes against independent real-world experiments, demonstrating that the results fell within the observed range of variation. However, for consignment type PB, a small discrepancy between the movement trajectories of the simulation and the real-world experiments was observed, which was not detected by the validation criterion used in our study. Therefore, future research should focus on refining validation criteria to robustly detect outliers in movement trajectories. The parametrization framework presented in our study represents a step towards bridging the gap between existing virtual prototyping capabilities and the capabilities needed to develop new material handling solutions tailored to the requirements of small consignments with flexible packaging. Additionally, the parameter values identified in our study can serve as starting points for further simulation tasks involving similar equipment and material combinations, though validation will be necessary in such cases.

Further research will involve simulating more test rigs and additional consignment types to expand the applicability of our findings. It remains to be seen whether the parameters found in our study will yield valid results in other setups using the same material pairings. Efforts should also be directed towards streamlining the calibration process to make it more accessible, e.g., by utilizing simpler test rigs that generalize better to other scenarios, though the current study has laid a solid foundation. Incorporating an additional objective that better captures the flexural behavior of the consignment packaging could further improve parametrization accuracy, particularly for parameters related to flexural properties. Enhancing the validation criteria and conducting additional validation experiments will be critical in ensuring the robustness of the simulation models. Finally, exploring alternative simulation approaches and frameworks with faster computation times will be beneficial to enhance efficiency and scalability in future studies.

AUTHOR CONTRIBUTIONS

For brevity, the authors' initials are used. Conceptualization, D.S., and C.L.; methodology, D.S., and G.L.; software (simulation) G.L.; software (other) D.S.; validation, D.S.; formal analysis, D.S.; investigation, D.S., and G.L.; data curation, D.S.; writing – original draft, D.S.; writing – review and editing[11], D.S., G.L., and C.L.; visualization, D.S., and G.L.; supervision, C.L.; project administration, D.S., and C.L.; funding acquisition, D.S., and C.L. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGEMENTS AND FUNDING

This research was conducted as part of the Austrian Research Promotion Agency (FFG)-funded project „Entwicklung eines Simulationsansatzes zur Analyse von Kleinsendungen (ISAAK)" (No. 891088) in collaboration with the project partners Körber Supply Chain Logistics GmbH in Konstanz and Austrian Post. We wish to express our gratitude to Körber Supply Chain Logistics GmbH for enabling us to conduct the data collection required for this research at their premises. Furthermore, we would like to thank our colleague Harald Steinkellner for his invaluable assistance during the data collection process.

LITERATURE

[1] K. Esser and J. Kurte, "KEP-Studie 2019 – Analyse des Marktes in Deutschland. Eine Untersuchung im Auftrag des Bundesverbandes Paket und Expresslogistik e. V. (BIEK)," Jun. 2019.

[2] K. Esser and J. Kurte, "KEP-Studie 2024 – Analyse des Marktes in Deutschland. Eine Untersuchung im Auftrag des Bundesverbandes Paket und Expresslogistik e. V. (BIEK)," Jul. 2024.

[3] Pitney Bowes, "Parcel shipping index 2017," 2017. Accessed: Apr. 15, 2023. [Online]. Available: https://www.pitneybowes.com/content/dam/pitneybowes/us/en/shipping-index/22-pbcs-04529-2021-global-parcel-shipping-index-ebook-web-002.pdf

[4] Pitney Bowes, "Parcel shipping index 2023," 2023. Accessed: Apr. 15, 2023. [Online]. Available: https://www.pitneybowes.com/content/dam/pitneybowes/us/en/shipping-index/22-pbcs-04529-2021-global-parcel-shipping-index-ebook-web-002.pdf

[5] International Post Corporation, "Cross-border e-commerce shopper survey 2023," Survey, 2024.

[11] The writing has been elaborated with the use of ChatGPT (13.05.2024 version) by OpenAI (https://chat.openai.com/chat) for structure and stylistic text revisions and DeepL Translate by Deepl SE (https://www.deepl.com/en/translator) for translation.

[6] International Post Corporation, "Cross-border e-commerce shopper survey 2021," Survey, 2022.
[7] Siemens Logistics GmbH, "Creating new business with mixed-mail automation. Integrating small parcel processing into mail processing centers." Siemens Logistics GmbH, 2021. [Online]. Available: https://www.siemens-logistics.com/en/news/white-paper/creating-new-business-with-mixed-mail-automation
[8] M. Schadler, M. Schedler, M. Knödl, D. Prims, C. Landschützer, and A. Katterfeld, "Characteristics of 'polybags' used for low-value consignments in the mail, courier, express and parcel industry," *Logistics Journal*, p. 25, 2022.
[9] *Directive 97/67/EC of the European Parliament and of the Council of 15 December 1997 on common rules for the development of the internal market of Community postal services and the improvement of quality of service*, vol. 015. 1997. Accessed: Aug. 01, 2024. [Online]. Available: http://data.europa.eu/eli/dir/1997/67/oj/eng
[10] Schedler, Michael and Landschützer, Christian, "Methodische Entwicklung eines neuartigen Sortiersystems für Polybags," *Volume 2021*, p. Issue 17, 2021, doi: 10.2195/LJ_PROC_SCHEDLER_DE_202112_01.
[11] D. Stadlthanner, H. Steinkellner, C. Landschützer, and D. Kaever (geb. Prims), "A Hierarchical Density-Based Clustering Method Applied to Mixed-Mail in Austria," *Logistics Research*, Jul. 2024, Accessed: Jul. 29, 2024. [Online]. Available: https://doi.org/10.23773/2024_4
[12] G. Leitner, D. Stadlthanner, and A. Ortner-Pichler, "Modellierung von forminstabilen Kleinsendungen mittels Multi Flexible Body Dynamics," in *Logistikwerkstatt Graz 2024*, 2024.
[13] G. Leitner, D. Stadlthanner, and A. Ortner-Pichler, "Modeling and Parameterization of Flexible Consignments Using Multi Flexible Body Dynamics," in *33rd edition of the International Conference on Manufacturing Systems*, Bucharest, submitted 2024.
[14] G. Bradski, "The OpenCV library," *Dr. Dobb's Journal of Software Tools*, 2000.
[15] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 3400–3407. doi: 10.1109/ICRA.2011.5979561.
[16] R. Bellman and R. Kalaba, "On adaptive control processes," *IRE Transactions on Automatic Control*, vol. 4, no. 2, pp. 1–9, Nov. 1959, doi: 10.1109/TAC.1959.1104847.
[17] T. Giorgino, "Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package," *Journal of Statistical Software*, vol. 31, pp. 1–24, Aug. 2009, doi: 10.18637/jss.v031.i07.
[18] FunctionBay, Inc., "RecurDyn Overview." Accessed: Aug. 14, 2024. [Online]. Available: https://support.functionbay.com/en/page/single/2/recurdyn-overview
[19] J. Stork, A. E. Eiben, and T. Bartz-Beielstein, "A new taxonomy of global optimization algorithms," *Nat Comput*, vol. 21, no. 2, pp. 219–242, Jun. 2022, doi: 10.1007/s11047-020-09820-4.
[20] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965, doi: 10.1093/comjnl/7.4.308.
[21] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," in *Spin Glass Theory and Beyond*, vol. Volume 9, in World Scientific Lecture Notes in Physics, no. Volume 9, vol. Volume 9. , WORLD SCIENTIFIC, 1986, pp. 339–348. doi: 10.1142/9789812799371_0035.
[22] F. Glover, "Future paths for integer programming and links to artificial intelligence," *Computers & Operations Research*, vol. 13, no. 5, pp. 533–549, Jan. 1986, doi: 10.1016/0305-0548(86)90048-1.
[23] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, Nagoya, Japan: IEEE, 1995, pp. 39–43. doi: 10.1109/MHS.1995.494215.
[24] M. D. McKay, R. J. Beckman, and W. J. Conover, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979, doi: 10.2307/1268522.
[25] J. Mockus, V. Tiesis, and A. Zilinskas, "The application of Bayesian methods for seeking the extremum," in *Towards Global Optimization*, vol. 2, 2014, pp. 117–129.
[26] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design," *ICML*, 2010.
[27] M. Rackl, C. D. Görnig, K. J. Hanley, and W. A. Günthner, "Efficient Calibration of Discrete Element Material Model Parameters Using Latin Hypercube Sampling and Kriging," in *Proceedings of the VII European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS Congress 2016)*, Crete Island, Greece: Institute of Structural Analysis and Antiseismic Research School of Civil Engineering National Technical University of Athens (NTUA) Greece, 2016, pp. 4061–4072. doi: 10.7712/100016.2092.5915.
[28] C. Richter, T. Rößler, G. Kunze, A. Katterfeld, and F. Will, "Development of a standard calibration procedure for the DEM parameters of cohesionless bulk materials – Part II: Efficient optimization-based calibration," *Powder Technology*, vol. 360, pp. 967–976, Jan. 2020, doi: 10.1016/j.powtec.2019.10.052.
[29] H. Cheng, T. Shuku, K. Thoeni, P. Tempone, S. Luding, and V. Magnanimo, "An iterative Bayesian

filtering framework for fast and automated calibration of DEM models," *Computer Methods in Applied Mechanics and Engineering*, vol. 350, pp. 268–294, Jun. 2019, doi: 10.1016/j.cma.2019.01.027.

[30] S. S. Razu *et al.*, "Bayesian Calibration of Computational Knee Models to Estimate Subject-Specific Ligament Properties, Tibiofemoral Kinematics, and Anterior Cruciate Ligament Force With Uncertainty Quantification," *Journal of Biomechanical Engineering*, vol. 145, no. 071003, Mar. 2023, doi: 10.1115/1.4056968.

[31] W. Zhang *et al.*, "A numerical Bayesian-calibrated characterization method for multiscale prepreg preforming simulations with tension-shear coupling," *Composites Science and Technology*, vol. 170, pp. 15–24, Jan. 2019, doi: 10.1016/j.compscitech.2018.11.019.

[32] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the Human Out of the Loop: A Review of Bayesian Optimization," *Proc. IEEE*, vol. 104, no. 1, pp. 148–175, Jan. 2016, doi: 10.1109/JPROC.2015.2494218.

[33] R. Rubin, "New Heuristics for Parallel and Scalable Bayesian Optimization," Aug. 22, 2018, *arXiv*: arXiv:1807.00373. doi: 10.48550/arXiv.1807.00373.

[34] L. D. González and V. M. Zavala, "New paradigms for exploiting parallel experiments in Bayesian optimization," *Computers & Chemical Engineering*, vol. 170, p. 108110, Feb. 2023, doi: 10.1016/j.compchemeng.2022.108110.

[35] M. Groves and E. O. Pyzer-Knapp, "Efficient and Scalable Batch Bayesian Optimization Using K-Means," Sep. 19, 2018, *arXiv*: arXiv:1806.01159. Accessed: Nov. 30, 2023. [Online]. Available: http://arxiv.org/abs/1806.01159

[36] S. S. Sandha, M. Aggarwal, I. Fedorov, and M. Srivastava, "MANGO: A Python Library for Parallel Hyperparameter Tuning," May 22, 2020, *arXiv*: arXiv:2005.11394. doi: 10.48550/arXiv.2005.11394.

[37] S. S. Sandha, M. Aggarwal, S. S. Saha, and M. Srivastava, "Enabling Hyperparameter Tuning of Machine Learning Classifiers in Production," in *CogMI 2021, IEEE International Conference on Cognitive Machine Intelligence*, IEEE, 2021.

[38] *ARM-software/mango*. (Jun. 21, 2024). Jupyter Notebook. Arm Software. Accessed: Jul. 01, 2024. [Online]. Available: https://github.com/ARM-software/mango

[39] F. Tronarp, T. Karvonen, and S. Särkkä, "MIXTURE REPRESENTATION OF THE MATÉRN CLASS WITH APPLICATIONS IN STATE SPACE APPROXIMATIONS AND BAYESIAN QUADRATURE," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, Sep. 2018, pp. 1–6. doi: 10.1109/MLSP.2018.8516992.

[40] R. Martinez-Cantin, "Funneled Bayesian Optimization for Design, Tuning and Control of Autonomous Systems," *IEEE Transactions on Cybernetics*, vol. 49, no. 4, pp. 1489–1500, Apr. 2019, doi: 10.1109/TCYB.2018.2805695.

[41] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[42] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

Dipl.-Ing. Dominik Stadlthanner, B.Sc., University Assistant at the Institute of Logistics Engineering, Graz University of Technology.

Dipl.-Ing. Gabriel Leitner, B.Sc., University Project Assistant at the Institute of Logistics Engineering, Graz University of Technology.

Assoc.Prof. DI Dr.techn. Prof.h.c. Christian Landschützer, Professor at the Institute of Logistics Engineering, Graz University of Technology.

Address: Institute of Logistics Engineering, Graz University of Technology, Inffeldgasse 25/E, 8010 Graz, Austria, www.itl.tugraz.at