

# A Regression-based Predictive Model Hierarchy for Nonwoven Tensile Strength Inference

Dario Antweiler, Jan Pablo Burgard, Marc Harmening, Nicole Marheineke, Andre Schmeißer, Raimund Wegener and Pascal Welke

**Abstract** Nonwoven materials, characterized by a random fiber structure, are essential for various applications including insulation and filtering. An industrial long-term goal is to establish a framework for the simulation-based design of nonwovens. Due to the random structures, simulations of material properties on fiber network level are computational expensive. We propose a predictive model hierarchy for inferring an important material property – the nonwoven tensile strength behavior. The model hierarchy is built using regression-based approaches, including linear and polynomial models, which provide interpretable results. This allows for significant speedup (six orders of magnitude) over the conventional simulations, while achieving good prediction results ( $R^2 = 0.95$ ). The proposed models open the application to nonwoven material design, as they provide accurate and cost-effective surrogates for predicting material properties. In this way, our work serves as a proof of concept.

---

Dario Antweiler  
Fraunhofer IAIS, Schloss Birlinghoven, 53757 Sankt Augustin, Germany, e-mail: [dario.antweiler@iais.fraunhofer.de](mailto:dario.antweiler@iais.fraunhofer.de)

Jan Pablo Burgard, Marc Harmening, Nicole Marheineke  
Trier University, Universitätsring 15, 54296 Trier, Germany, e-mail: [{burgardj,harmening,marheineke}@uni-trier.de](mailto:{burgardj,harmening,marheineke}@uni-trier.de)

Andre Schmeißer, Raimund Wegener  
Fraunhofer ITWM, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany, e-mail: [{andre.schmeisser,raimund.wegener}@itwm.fraunhofer.de](mailto:{andre.schmeisser,raimund.wegener}@itwm.fraunhofer.de)

Pascal Welke  
University of Bonn, Friedrich-Hirzebruch-Allee 8, 53115 Bonn, Germany, e-mail: [welke@cs.uni-bonn.de](mailto:welke@cs.uni-bonn.de)

## 1 Introduction

The efficient prediction of material properties based on production parameters is a common goal for many industrial applications. This includes the nonwoven airlay manufacturing, which serves as practical basis for this work. Nonwovens are characterized by a random fiber structure that is usually bonded by thermal, chemical, or mechanical procedures. Their low-cost production makes them a suitable choice for many fabrics, such as filters, insulation materials or hygiene products [7]. Predicting nonwoven properties from production parameters enables nonwoven material design by providing insight into the effects of individual parameters. In order to avoid costly experimental testing, this mainly involves simulation-based approaches, which, however, often suffer from high computational effort [33]. More recently, machine learning approaches have gained ground in this field, as they allow comparatively efficient predictions, see [2]. In particular, the integration of prior knowledge into the training process, termed “Informed Machine Learning”, proved to be beneficial in terms of training speed and quality of final predictive models [31]. This work demonstrates the use of machine learning approaches for predicting nonwoven material properties. We focus, as an example, on the tensile strength behavior of airlay fabricated nonwovens (see Fig. 1(a)-(c)), for which we develop and propose a predictive model hierarchy driven by simulation data. With this goal in mind, we begin with a brief discussion of related literature in the field of nonwoven tensile strength simulations and machine learning approaches, and then explain the novelty and the structure of this work.

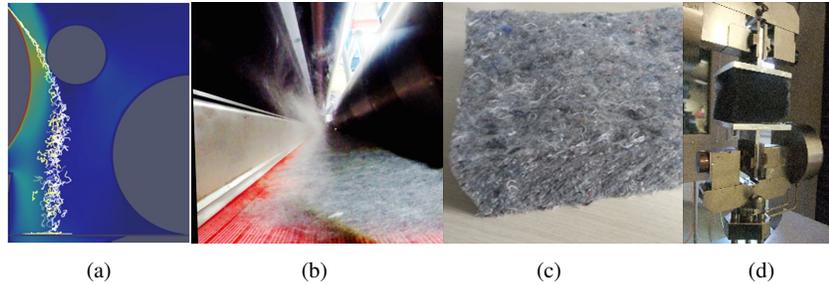


Fig. 1: Nonwoven airlay manufacturing and property testing: (a) simulated fiber dynamics and laydown in turbulent airflow (process of Airlay-K12 by machine manufacturer AUTEFA Solutions), (b) fiber laydown zone, (c) final product, (d) tensile strength test for a material sample (photo by IDEAL Automotive). Image adapted from [14] with kind permission of the authors.

## 1.1 Literature Overview on Nonwoven Modeling and Related Machine Learning Approaches

Mapping production parameters to the tensile strength behavior of nonwovens requires the simulation of the underlying production process and the mechanical behavior of the resulting fiber structure. There are many approaches to virtual generation of fiber structures, coming from statistical analysis and stochastic geometry [22, 28] or three-dimensional volume imaging covering microscopy and X-ray tomography [10, 23]. However, the challenge is to model the underlying production process, which was done by Gramsch *et al.* [14] for nonwoven airlay production. The authors introduced a chain of mathematical models coupled by parameter identification to deal with the computational complexity that arises from several thousand airlaid fibers in a complex machine geometry. The models cover a highly turbulent fiber suspension flow, a stochastic surrogate for the fiber laydown on a moving conveyor belt, and a bonding process mimicking the thermobonding. The suitability of such model hierarchies for virtual nonwoven generation is topic in [33].

Various approaches in the literature deal with the simulation of the mechanic behavior of nonwovens. A common procedure is to treat the nonwoven material as a continuum, which allows the use of finite element methods [8, 12]. In these approaches, the behavior of individual fibers is not considered, but instead knowledge of the statistical fiber orientation is incorporated to account for the randomness in the material web. In contrast, there are approaches that consider the mechanical behavior at the fiber network level, cf. [15, 18]. Kufner *et al.* [18] described the material's structure as an elastic Cosserat network. As resolving the behavior of the individual beam-type fibers in an industrial-size virtual material sample is too complex, additional homogenization techniques are necessary [19, 26]. Harmening *et al.* [15] modeled the fiber structure as truss with nonlinear elastic behavior and reduced the applied stress to the forces at the individual fiber joints, which mainly determine the nonwovens' tensile strength behavior. A problem-tailored data reduction strategy and a singularly perturbed regularization approach enable simulations with industrial-size samples. The approach underlying this work handles the problem-inherent multiscales (interplay of deterministic structural effects at macro-scale and random fiber alignment at micro-scale) and realizes the randomness in the fiber structure generation by Monte-Carlo simulations.

In the field of machine learning, there is much literature on woven material prediction, while modeling approaches for nonwovens are rare. Early work used simple neural network architectures, which prevent any interpretability of the results, to predict the strength of yarns [6] or worsted fabrics [11]. [1] investigated neural networks and linear regression for predicting the tensile strength of woven fabrics, limiting the work to seven training samples. Eltayib *et al.* [9] used linear regression to predict tear strength of fabrics based on yarn count, yarn tenacity and fabric liner density. The approach in [27] deploys multiple regression models to predict different material properties of woven fabrics, but heavily relies on huge datasets and extensive manual feature selection by domain experts. Due to the high computational cost of generating training data and due to their specialization on weaving features, these

approaches cannot be applied to property prediction of nonwovens. For nonwovens, Rahnama *et al.* [25] proposed a feed-forward neural network based on a numerical propagation model to compare heat and moisture propagation through different nonwoven fabrics. Chen *et al.* [4] integrated simple logical rules developed by domain experts into a neural network to predict elongation at break, but the paper is limited to a single test example. Investigating nonwoven features and developing a stretch algorithm, we employed linear regression for tensile strength prediction in [2], which yields promising, accurate and interpretable results.

Aside from (non-)woven manufacturing, there are other works that address the prediction of material properties from production parameters. Related to our work are those that integrate prior knowledge about the underlying physical mechanics into the data, the model architecture, or the loss term being optimized. For example, Karpatne *et al.* [17] integrated physical knowledge about feature dependencies into a neural network as additional loss terms. Lu *et al.* [20] presented an approach in which knowledge about underlying material mechanics was incorporated into a machine learning approach as algebraic formulas. However, with the handcrafted neural network architectures, they are not able to provide interpretable results. Recent research on combining machine learning and simulation approaches in a more general context can be found in [30, 31]. Within the proposed taxonomy, our approach can be contextualized as the integration of (i) algebraic equations and (ii) simulation results from scientific and domain knowledge sources.

## 1.2 New Regression-based Predictive Model Hierarchy

The model-based simulation framework underlying this work goes back to Gramsch *et al.* [14] for virtual fiber structure generation and to Harmening *et al.* [15] for tensile strength computation. Its evaluation yields a tuple consisting of (utilized) production parameters, a random fiber graph and an associated stress-strain curve indicating the relationship between strain and stress during nonwoven’s tensile strength testing. To account for the randomness in the fiber structure generation, Monte-Carlo simulations are required, which multiply the already high time requirements. This makes nonwoven material design impossible in practice and motivates a predictive surrogate. Following our ideas and strategies developed in [2], we propose a new regression-based model hierarchy for the prediction of the nonwovens’ stress-strain behavior from production parameters (see Fig. 2). Once trained, the regression models are characterized by efficient evaluations allowing for significant speedup, while providing good, interpretable results, as we will show.

The *tensile strength model-based simulation framework* (TSS-model) at the top of the model hierarchy is built on a first principle-oriented model chain. It serves as ground truth for predictions and provides the required datasets for machine learning. By considering linear regression, two approaches have been proposed in [2] that allow to circumvent the high computational effort associated to the TSS-model: The *fiber graph feature-based predictive model* (FGF-model) samples multiple fiber

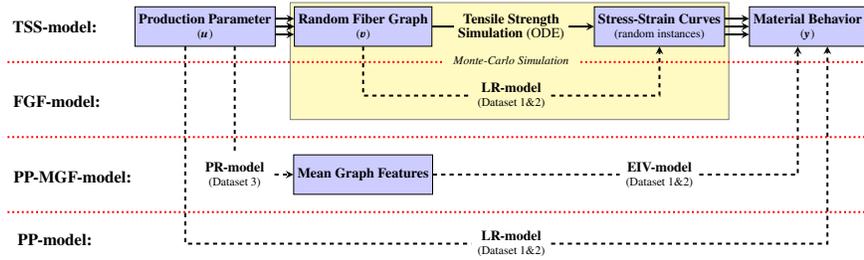


Fig. 2: Predictive model hierarchy: Mappings from the production parameters to the associated tensile strength behavior. Predictive relations are indicated with dashed lines and simulations procedures with solid lines. For predictions, we employ linear regression (LR), polynomial regression (PR) and an errors-in-variables model (EIV).

graphs, extracts associated graph and stretch features, and uses them to predict the stress-strain curve for each fiber graph. The *production parameter-based predictive model* (PP-model) predicts directly the mean stress-strain curve based solely on production parameters. The FGF-model provides better predictions, but Monte-Carlo simulations are necessary to derive expectations and variances from individual fiber graph features. This is accompanied by a computational overhead required to generate random fiber graph samples. The fact that the purely linear PP-model performs worse suggests some nonlinear relationships between production parameters and fiber graph features. In this work we introduce the novel *production parameter and mean graph feature-based predictive model* (PP-MGF-model) as a compromise between the established ones. The PP-MGF-model intercepts the nonlinearities by predicting the mean graph features using polynomial regression. Then, these (artificial) features are used as additional explanatory variables for predicting the stress-strain curves with a linear (errors-in-variables) model, in order to recover the good quality of the FGF-model. Its main advantage is that the model provides a good predictive quality without requiring Monte-Carlo simulations.

### 1.3 Structure

The structure of this work is based on the regression-based predictive model hierarchy depicted in Fig. 2. Section 2 outlines the TSS-model, by introducing the first principle-oriented model chain, and lays the foundations for predictions. Section 3 discusses the FGF-model and the PP-model originating from Antweiler *et al.* [2] and presents a performance study with focus on predictive quality. Section 4 introduces the new PP-MGF-model, which is numerically investigated in comparison to the established ones. Finally, Section 5 concludes with a general discussion and an outlook to future work.

## 2 First Principle Oriented Model Chain for Dataset Generation

The TSS-model is a first principle-oriented model chain that covers fiber graph generation and tensile strength simulation, see [14, 15]. It maps from an input set of 28 (production) parameters to a random stress-strain curve instance as output. In this work, we restrict to practically relevant production processes that are characterized by 4 parameters and refer to them as 4-parametric (production) process class. The resulting stress-strain curves obey a similar behavior that motivates a 2-parametric labeling. We refer to this restriction as stress-strain curve class. To improve the predictions in machine learning we consider additional fiber graph features. In this section we explain the TSS-model (Section 2.1) and introduce input (production parameters  $\mathbf{u}$ , Section 2.2), output (stress-strain characteristics  $\mathbf{y}$ , Section 2.3), and auxiliary variables (random fiber graph features  $\mathbf{v}$ , Section 2.4), before we describe the generation of the datasets used for training and testing our regression models in Section 2.5. Note that readers with focus on the predictive models may skip this rather technical and mathematically extensive section and think of it as a black box for data generation.

### 2.1 Fiber Graph Generation and Tensile Strength Simulation

The TSS-model involves a stochastic fiber lay-down model (A) with graph generation (B) and an ordinary differential system for tensile strength testing (C). The model parameters are specified in Section 2.2.

(A) A nonwoven material is the image of fibers deposited onto a moving conveyor belt. Consider a cubic reference material volume  $\mathcal{V}_R$  over the nonwoven height  $H$  with base area  $w_R^2$  and let  $T_R$  be the time needed to produce it. A deposited fiber of length  $L$  is identified with the lay-down time  $T$  and the planar coordinates  $(X, Y)$  of one of its end points. It contributes to  $\mathcal{V}_R$ , if  $X - x_B(T) \in [-w_R/2, w_R/2]$  is satisfied, where  $x_B$  accounts for the motion of the conveyor belt. In the three-dimensional web a fiber is modeled in terms of the curve  $\boldsymbol{\eta}^{(X, Y, T)} : [0, L] \rightarrow \mathcal{V}_R$ ,

$$\begin{aligned} d\boldsymbol{\eta}_s &= \mathbf{R}(\boldsymbol{\eta}_s \cdot \mathbf{e}_x + x_B(T)) \cdot \boldsymbol{\tau}_s ds, & \boldsymbol{\eta}_0 &= (X - x_B(T))\mathbf{e}_x + Y\mathbf{e}_y + r(X)\mathbf{e}_z \\ \mathbf{R}(x) &= \frac{1}{\sqrt{1+r'(x)^2}} [\mathbf{I} + (\sqrt{1+r'(x)^2} - 1)\mathbf{e}_y \otimes \mathbf{e}_y + r'(x)(\mathbf{e}_z \otimes \mathbf{e}_x - \mathbf{e}_x \otimes \mathbf{e}_z)] \\ r(x) &= H \int_{-\infty}^x g(\bar{x}) d\bar{x} \end{aligned}$$

with  $X$  distributed according to the lay-down probability density function  $g$  as well as  $Y \sim \mathcal{U}([-w_R/2, w_R/2])$  and  $T \sim \mathcal{U}([0, T_R])$  uniformly distributed. The system above is based on the stochastic Stratonovich differential system

$$d\xi_s = \boldsymbol{\tau}_s ds, \quad d\boldsymbol{\tau}_s = -\frac{1}{B+1} [\boldsymbol{\Pi}_s(B) \cdot \nabla \Sigma(\xi_s) ds + A \boldsymbol{\Pi}_s(\sqrt{B}) \circ d\mathbf{w}_s]$$

with unit tensor  $\mathbf{I}$ , projection  $\mathbf{\Pi}_s(x) = \mathbf{n}_{1,s} \otimes \mathbf{n}_{1,s} + x \mathbf{n}_{2,s} \otimes \mathbf{n}_{2,s}$  as well as  $\xi_0 = 0$  and  $\tau_0$  uniformly distributed in the unit circle spanned by  $\mathbf{e}_x$  and  $\mathbf{e}_y$ . The stochastic lay-down model for position and orientation  $((\xi, \tau) : [0, L] \rightarrow \mathbb{R}^3 \times \mathbb{S}^2)$  with unit sphere  $\mathbb{S}^2 \subset \mathbb{R}^3$  describes the path of a deposited fiber onto the  $\mathbf{e}_x$ - $\mathbf{e}_y$  plane. In the modeling for the fiber tangent  $\tau$ , the drift term prescribes the characteristic coiling behavior with the potential  $\Sigma$ , while the white noise term with the Wiener process  $(w : [0, L] \rightarrow \mathbb{R}^3)$  and the amplitude  $A$  accounts for fluctuations in the lay-down process. Anisotropic behavior is indicated by the parameter  $B \in [0, 1]$  with the local orthonormal triad  $\{\tau, \mathbf{n}_1, \mathbf{n}_2\}$ ,  $\mathbf{n}_1 \in \text{span}\{\mathbf{e}_x, \mathbf{e}_y\}$ . The typical nestling behavior of the fiber on the ramp-like contour surface of the nonwoven is modeled by the curve  $\eta$ . The contour line  $r$  of the fiber material in machine direction is described by means of the joint probability density function  $g$  of the deposited material. A fiber end point lies on the associated contour surface and the fiber orientation is aligned to it due to the local rotation  $\mathbf{R}(x) \in \text{SO}(3)$ .

(B) Our considerations are restricted to the embedded test material volume  $\mathcal{V} \subset \mathcal{V}_R$  with smaller base  $w^2$ ,  $w = w_R - 2L$ , to exclude lateral boundary effects. The random fiber web is consolidated by adhesive joints resulting from thermobonding. Let  $\eta_h$  denote the discretized fiber, i.e., a set of discrete fiber points. An adhesive joint  $\mathbf{a}$  to be formed between two fibers  $\eta_h$  and  $\tilde{\eta}_h$  is modeled as

$$\mathbf{a} = \frac{1}{2}(\mathbf{q}^* + \tilde{\mathbf{q}}^*)$$

$$\text{if } \|\mathbf{q}^* - \tilde{\mathbf{q}}^*\|_2 < \kappa, \quad (\mathbf{q}^*, \tilde{\mathbf{q}}^*) = \underset{(\mathbf{q}, \tilde{\mathbf{q}}) \in \eta_h \times \tilde{\eta}_h}{\text{argmin}} \|\mathbf{q} - \tilde{\mathbf{q}}\|_2$$

with contact threshold  $\kappa > 0$ . The adhesive joint takes the place of the fiber points in contact within the respective fibers. As the minimizer might be not unique, we use the first minimizer found for practical reasons. Since the fibers lie rather straight, cf. [14], we assume at most one contact between each fiber pair. If more fibers are involved in a contact, the resulting adhesive joint is centered between the respective fiber points in contact. The resulting adhered fiber structure is considered as a connected graph  $G = (V, E)$  with the nodes  $V$  representing adhesive joints as well as fiber ends and the edges  $E$  indicating fiber connections between them. The graph is supplemented by the node positions  $\mathbf{p}_0 : V \rightarrow \mathbb{R}^3$  and the edge-associated fiber lengths  $l : E \rightarrow \mathbb{R}_{\geq 0}$ .

(C) The tensile strength test is modeled as differential system on the node positions  $\mathbf{p} : V \times [0, 1] \rightarrow \mathbb{R}^3$ , initialized with  $\mathbf{p}(\cdot, 0) = \mathbf{p}_0$ ,

$$\mathbf{p}(v, 0) = \mathbf{p}_0(v), \quad \forall v \in V_l, \quad \mathbf{p}(v, t) = \mathbf{p}_0(v) + t h \mathbf{e}_3, \quad \forall v \in V_u$$

$$\varepsilon \partial_t \mathbf{p}(v, t) = \sum_{e \in \delta(v)} \mathbf{f}_e^v(t), \quad \forall v \in V \setminus (V_l \cup V_u)$$

$$\mathbf{f}_{e=\{v, v'\}}^v(t) = \frac{\mathbf{p}(v', t) - \mathbf{p}(v, t)}{d(e, t)} N \left( \frac{d(e, t) - l(e)}{l(e)} \right)$$

with  $\delta(v) \subset E$  incident edges of node  $v$ . For fixed lower face  $V_l$ , the upper face  $V_u$  of the fiber structure is linearly shifted away in (vertical)  $\mathbf{e}_3$ -direction (with maximal

displacement  $h > 0$ ). In the interior nodes of the graph the acting traction forces are balanced by a friction term with  $\varepsilon > 0$ . The force amplitude  $N$  depends on the relative strain of the fiber connection  $e$  with respect to its length  $l(e)$ , where  $d(e, t)$  denotes the Euclidean distance between its endpoints,  $d(e, t) = \|\mathbf{p}(v, t) - \mathbf{p}(v', t)\|_2$ . It reflects Hooke's law in the stretched state and is taken as zero in the unstretched state. The characterizing stress-strain relation for the fiber structure (with initial height  $H$ ) is then given by  $(\varepsilon(t), T(\mathbf{p}(\cdot, t)))$ ,  $t \in [0, 1]$ ,

$$\varepsilon(t) = \frac{h}{H}t, \quad T(\mathbf{p}(\cdot, t)) = - \sum_{v \in V_u} \sum_{e \in \delta(v)} \mathbf{f}_e^v(t) \cdot \mathbf{e}_3.$$

## 2.2 Production Process Class

An airlaid nonwoven typically consists of two fiber types for which the TSS-model has 28 input parameters in total: Each fiber type is characterized by length  $L_f$ , line density  $(\rho A)_f$ , cross-sectional weighted elasticity modulus  $(EA)_f$  and lay-down probability density  $g_f$  considered as normally distributed  $g_f \sim \mathcal{N}(\mu_f, \sigma_f^2)$ ,  $f = 1, 2$ . The joint probability density is then  $g = \beta_n g_1 + (1 - \beta_n) g_2$  with fiber number fraction  $\beta_n$  determined by mass fraction  $\beta$ . For technical reasons, we use a compact support  $\text{supp}(g) = [x_l, x_r]$ . The production plant is characterized by conveyor belt width  $b$  and speed  $v_B$  as well as mass rate  $\dot{m}$ . The nonwoven sample is specified by height  $H$  and width  $w$ . Production time  $T_R$ , trace curve  $x_B$  and number of deposited fibers per type  $n_f$ , for  $f = 1, 2$ , are resulting quantities. The laydown is parameterized regarding diffusion  $A$ , anisotropy  $B$  and bending potential  $\Sigma$  expressed by the three standard deviations  $\sigma_x, \sigma_y, \sigma_z$  in  $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ -directions. The bonding considers fiber discretization length  $\Delta s$  and contact threshold  $\kappa$ . The strength test is parameterized by adhesive thickness  $z$  for upper and lower structure faces, friction-associated regularization  $\varepsilon$  as well as traction function  $N$  with a regularization parameter  $\delta$ . Note that the displacement  $h$  in the strength test belongs to the input quantities.

Since the parameters  $(\rho A)_1, (\rho A)_2, \beta = \beta_1/\beta_2, \dot{m}$  and  $b$  only occur in the quantities  $\alpha_f = \beta_f \dot{m} / ((\rho A)_f L_f b)$ , for  $f = 1, 2$ , indicating the number of fibers for each type deposited per second and meter in cross direction on the conveyor belt, three parameters can be eliminated. Making the model dimensionless with nonwoven sample width  $w$ , conveyor belt speed  $v_B$  and elasticity modulus  $(EA)_1$  reduces the set of input parameters by further three. The resulting dimensionless numbers are mainly formulated as ratios, cf., Table 1. Note that the strength test is stated in dimensionless form to incorporate the friction-associated (dimensionless) regularization parameter  $\varepsilon \ll 1$  that ensures a unique solution.

In this work, we consider a 4-parametric production process class. The process class is motivated from the industrial test setting in [14]: We adopt all industrial values – except for  $\mathbf{u} = (\hat{\alpha}, \hat{\sigma}, \hat{\sigma}_y, \hat{\kappa}) \in \mathbb{R}_+^4$ . These four inputs affect the fiber amount in the nonwoven (sample), the fiber laydown behavior and the bonding (i.e., fiber graph topology). By varying them in a certain regime, a broad variety of practically

Table 1: Characteristic dimensionless input parameters for TSS-model. Values for an industrial airlay process (mixture of solid (PES) and bi-component (PES/PET) fibers in plant K12, cf., scenario in [14]). Referential values in SI units:  $w = 1.0 \cdot 10^{-2}$  m,  $v_B = 3.3 \cdot 10^{-2}$  m/s,  $(EA)_1 = 1.0$  N.

Description	Symbol	Value
fiber length	$L_1/w, L_2/L_1$	5.5, 1.0
fiber number	$\alpha_1 w^2/v_B, \alpha_2/\alpha_1$	1150, 0.65
elasticity modulus	$(EA)_2/(EA)_1$	1.0
lay-down pdf mean	$\mu_1/w, \mu_2/w$	0, 0
lay-down pdf std	$\sigma_1/w, \sigma_2/\sigma_1$	2.0, 1.0
support joint lay-down pdf	$x_l/\sigma_1, x_r/\sigma_1$	-5.0, 5.0
nonwoven sample height	$H/w$	6.0
bending potential (std)	$\sigma_y/w, \sigma_x/\sigma_y, \sigma_z/\sigma_y$	2.0, 0.75, 0.075
diffusion	$A\sqrt{\sigma_y}$	$2.8 \cdot 10^{-2}$
anisotropy	$B$	$3.0 \cdot 10^{-1}$
fiber discretization	$\Delta s/w$	$2.75 \cdot 10^{-2}$
contact threshold	$\kappa/w$	$2.6 \cdot 10^{-2}$ (calibrated)
adhesive thickness at faces	$z/w$	$6.0 \cdot 10^{-2}$
friction regularization	$\varepsilon$	$1 \cdot 10^{-7}$
traction regularization	$\delta$	$1 \cdot 10^{-4}$

Table 2: Input  $\mathbf{u}$  (4-parametric production process class) for machine learning. Parameter ranges for dataset used in ML approach and respective values in industrial scenario, Table 1. The values of all other parameters (ratios) are taken from Table 1.

Symbol	Range	Industrial Value	Effect
$\hat{\alpha} = \alpha_1 w^2/v_B$	[1000, 1515]	1150	amount of fibers
$\hat{\sigma} = \sigma_1/w$	[1.0, 5.0]	2.0	laydown behavior
$\hat{\sigma}_y = \sigma_y/w$	[1.0, 5.0]	2.0	laydown behavior
$\hat{\kappa} = \kappa/w$	$[2.8, 3.0] \cdot 10^{-2}$	$2.6 \cdot 10^{-2}$	bonding

relevant airlay scenarios are covered, see Table 2 for the parameter ranges underlying our dataset for machine learning. Note that the larger chosen  $\hat{\kappa}$  ensures a stronger bonding and hence a denser fiber structure than in the industrial test case.

### 2.3 Stress-Strain Curve Class

The stress-strain curves of the nonwovens obtained by the 4-parametric production process class show a similar pattern and allow for a 2-parametric labeling,  $\mathbf{y} = (\alpha, \beta)$ . The observed output curves are constant at a stress close to zero up to a threshold value  $\alpha$  of applied strain, above which they increase quadratically with coefficient  $\beta$ , see Fig. 3. The behavior results from more and more fibers coming under strain and thus contributing to the tensile strength, neglecting plastic effects and fiber tearing.

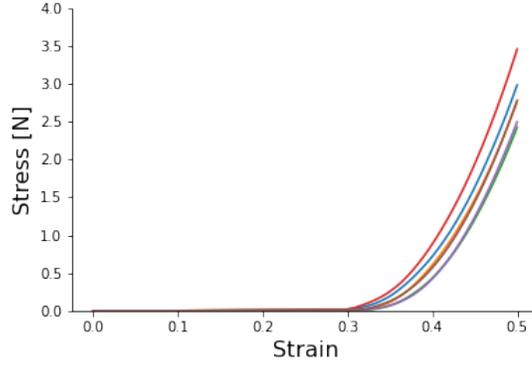


Fig. 3: Stress-strain curves obtained for fixed parameter setting by TSS-model.

Hence, we model the relation between strain and stress for a nonwoven sample by

$$T_{\mathbf{y}}(\epsilon) = \begin{cases} 0, & \epsilon < \alpha \\ \beta(\epsilon - \alpha)^2, & \epsilon \geq \alpha \end{cases}, \quad \mathbf{y} = (\alpha, \beta) \in \mathbb{R}_+^2 \quad (1)$$

where  $\epsilon$  refers to the relative strain applied to the sample and  $T_{\mathbf{y}} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  to the resulting reacting force.

The approximation of the stress-strain curve by the constant-quadratic ansatz enables a straightforward machine learning modeling approach with only two output parameters  $\mathbf{y} = (\alpha, \beta)$  as labels for prediction – instead of a complex output curve. The general tensile strength behavior can be characterized using the joint distribution of  $\alpha$  and  $\beta$ . To draw conclusions about the randomness of the material, the constant-quadratic ansatz can be used to compute, for example, the mean stress and the associated variance at individual strain levels.

## 2.4 Fiber Graph Features

The use of fiber graph features for predicting tensile strength has turned out to be advantageous in machine learning. According to Antweiler *et al.* [2] we use two groups of features: *topological graph features* representing the fiber structure connectivity which likely affects the nonwoven's tensile strength and *stretch features*, which are obtained by a heuristic stretching algorithm based on elongation of the nonwoven samples, allowing only vertical displacements of the nodes and no strain on the individual fibers. The identification of the features and the stretching algorithm originate from [2].

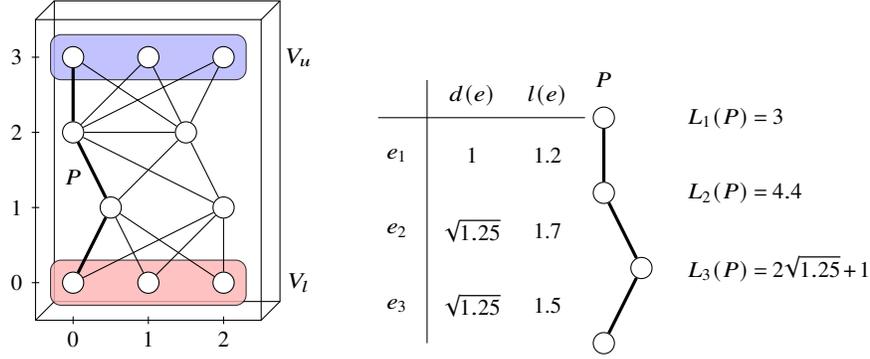


Fig. 4: Illustration of some graph features. Fiber graph with  $|V| = 10$  nodes,  $|E| = 17$  edges. Face sets  $V_u$  and  $V_l$  are colored in blue, and red, respectively. For path  $P$ , length variants  $L_1$ ,  $L_2$  and  $L_3$  are based on edge-wise fiber lengths  $l(e)$  and Euclidean lengths  $d(e)$ . A minimum cut  $C_{\min}$  separates all nodes above the value 1.5 from all nodes below that value,  $|C_{\min}| = 4$ .

**Topological Graph Features.** The graph feature set contains the numbers of nodes  $|V|$  and edges  $|E|$ , maximum node degree  $d_{\max} = \max_{v \in V} \delta(v)$ , total fiber lengths  $L_{\text{fiber}} = \sum_{e \in E} l(e)$  as well as the numbers  $|V_u|$ ,  $|V_l|$  of upper and lower face nodes. Moreover, to encode the graph connectivity several path and length-associated features are considered, see Fig. 4. Let  $L_1(P) = \text{len}(P)$  denote the edge count,  $L_2(P) = \sum_{e \in P} l(e)$  and  $L_3(P) = \sum_{e \in P} d(e)$  the fiber and Euclidean lengths for a path  $P$ . Of interest are the shortest paths connecting the upper and lower faces – in terms of edge count  $P_1$  and fiber length  $P_2$ , i.e.,  $P_1 = \text{argmin} \{L_1(P_{uv}) \mid u \in V_l, v \in V_u\}$  and  $P_2 = \text{argmin} \{L_2(P_{uv}) \mid u \in V_l, v \in V_u\}$ , so we include  $L_1(P_1)$ ,  $L_2(P_2)$  and  $L_3(P_2)$  to the feature set. In addition, we consider mean, median and sum of differences between fiber and Euclidean length over all edges  $\{l(e) - d(e) \mid e \in E\}$  and the size of a minimum cut  $C_{\min}$ , i.e., edge set with minimum cardinality disconnecting  $V_u$  from  $V_l$  when removed.

**Stretch Features.** The stretch features obtained from the stretching algorithm by Antweiler *et al.* [2] provide information about the nonwoven behavior under vertical tensile loading. A graph  $G$  with node positions  $\mathbf{p}$  and edge-associated fiber lengths  $l$  is called a *valid instance* if the following length constraint is satisfied,

$$l(\{v, w\}) \geq \|\mathbf{p}(v) - \mathbf{p}(w)\|_2 = d(\{v, w\}) \quad \forall \{v, w\} \in E.$$

The stretching algorithm (Algorithm 1) addresses the question of how far the nonwoven sample can be stretched vertically without stretching any fibers, i.e., maximizing the sum of height coordinates of the upper face nodes  $V_u$ , while fixing the positions of the lower face nodes  $V_l$  and keeping the instance valid. For computational reasons

**Algorithm 1** Graph Stretching Algorithm

---

Input: a valid instance  $(G, \mathbf{p}, l)$  and  $V_l \neq \emptyset$ 

Output: a valid instance  $(G, \tilde{\mathbf{p}}, l)$  that maximizes the objective of ZStretch

---

- 1: set  $\tilde{\mathbf{p}}(v) = \mathbf{p}(v) \quad \forall v \in V$
  - 2: set  $V_\perp = V_l$  and  $B = \mathcal{N}(V_\perp)$  where  $\mathcal{N}(\cdot)$  refers to the neighbor nodes
  - 3: **for**  $v = \operatorname{argmin}_{w \in B} \max \operatorname{Move}(w, V_\perp)$  **do**
  - 4:     pop  $v$  from  $B$
  - 5:      $\tilde{p}_3(v) = \max \operatorname{Move}(w, V_\perp) + p_3(v)$
  - 6:     add  $v$  to  $V_\perp$
  - 7:      $B = B \cup \mathcal{N}(v) \setminus V_\perp$
- 

the algorithm assumes that the fiber nodes (outside  $V_l$ ) can only move freely in the vertical (third) dimension while their horizontal position is fixed. Given a valid instance  $(G, \mathbf{p}, l)$  and  $V_l \subseteq V$  the *ZStretch* problem reads:

$$\begin{aligned}
& \max \sum_{v \in V_u} \tilde{p}_3(v) \\
\text{subject to:} \quad & \tilde{\mathbf{p}}(v) = \mathbf{p}(v) && \forall v \in V_l, \\
& \tilde{p}_1(v) = p_1(v) && \forall v \in V, \\
& \tilde{p}_2(v) = p_2(v) && \forall v \in V, \\
& d(\{v, w\}) \leq l(\{v, w\}) && \forall \{v, w\} \in E,
\end{aligned}$$

where  $p_i = \mathbf{p} \cdot \mathbf{e}_i$ ,  $i = 1, 2, 3$ , denote the spatial coordinates. The optimization problem certainly ignores many real-world structure properties, e.g., fiber intertwining, or the fact that fiber nodes can in reality move in all three dimensions to allow further stretching of the nonwoven sample in the third dimension. But due to its simplicity it can be solved in  $O(|E| \log(|V|))$  run time by Algorithm 1. As a result, a lower bound to the maximum movement of any fiber node in vertical direction is determined. We use mean, standard deviation, median, maximum, and sum of the differences between initial and optimized upper face node positions as stretching features.

As extension, stretching of the individual fibers up to a multiple of their lengths is incorporated by weakening the length constraint to  $\tilde{l}(e) = cl(e) \geq d(e)$  for some  $c > 1$ . For increasing values of  $c$  and a fixed graph, this provides a nonlinear behavior of the average vertical positions of the upper face nodes. We particularly determine the stretch features for various length factors,  $c \in \{1, 1.05, 1.1, \dots, 1.5\}$ , see Table 3.

## 2.5 Dataset

The machine learning dataset is generated using the TSS-model. The combinations of input production parameters are randomly selected from a range that yields reasonable fiber structures, cf., Table 2. While the generation of fiber graphs and

**Algorithm 2** maxMove SubroutineInput: a node  $v \in V$  and  $V_\perp \subseteq V$ .Output: the largest  $h$  such that  $\tilde{\mathbf{p}}(v) = \mathbf{p}(v) + h\mathbf{e}_3$  satisfies  $\|\tilde{\mathbf{p}}(v) - \mathbf{p}(w)\|_2 \leq l(\{v, w\})$   
 $\forall w \in \mathcal{N}(v) \cap V_\perp$ 

- 1: **for all**  $\{v, w\} \in E$  **do**
- 2:   find the largest  $h$  s.t.
- 3:    $\tilde{\mathbf{p}}(v) = \mathbf{p}(v) + h\mathbf{e}_3$  satisfies
- 4:    $\|\tilde{\mathbf{p}}(v) - \mathbf{p}(w)\|_2 \leq l(v, w) \quad \forall w \in \mathcal{N}(v) \cap V_\perp$

Table 3: Input features for regression models.

Set	Symbols	Description
param	$\mathbf{u}$	four parameters for production process
graph	$ V $	number of nodes
	$ E $	number of edges
	$d_{\max}$	maximum node degree
	$L_{\text{fiber}}$	total fiber lengths
	$ V_u $	number of upper face nodes
	$ V_l $	number of lower face nodes
	$L_1(P_1)$	minimal edge count of all paths from $V_u$ to $V_l$
	$L_2(P_2)$	minimal fiber lengths of all paths from $V_u$ to $V_l$
stretch	$L_3(P_2)$	Euclidean length of weighted shortest path $P_2$
	$D_1, D_2, D_3$	{mean, median, sum} of differences between edge-wise fiber and Euclidean lengths
	$ C_{\min} $	size of minimum edge cut separating $V_u$ and $V_l$
stretch	$S_1^c, S_2^c, S_3^c, S_4^c, S_5^c$	{mean, std, median, max, sum} of stretching distance for $c \in \{1, 1.05, 1.1, \dots, 1.5\}$

accompanying features is fast, the computation of the stress-strain curves is very time-consuming as it requires solving large-scale dynamical systems on the individual fiber structure samples. To account for the systems' stiffness, we employ an implicit Euler scheme with variable step size control. For the resulting nonlinear equation systems, we use an exact Newton method with analytical Jacobian and Armijo's line search. An explicit Euler step provides a suitable initial guess for warm start. The ODE-solver typically requires between 24 and 48 hours for a single instance, making it the bottleneck for building datasets. Both the fiber graph generation as well as the tensile strength simulations are performed in parallel on a machine with 88 CPU cores (Intel(R) Xeon(R) CPU E5-2699 v4 @ 2.20GHz) and 792 GB RAM running Ubuntu 18.04.6 using Matlab (R2019a).

For 43 parameter combinations, we generate 25 sample graphs each, totaling 1.075 graphs. On average, each graph contains 51.507 nodes (standard deviation  $\pm 2.182$ ) and 198.744 edges ( $\pm 29.996$ ). We randomly select six of our parameter combinations and compute the 25 stress-strain curves associated to the graphs (set 1, fully labeled), while for all other combinations we compute only a single stress-strain curve for one of the corresponding samples (set 2, single labeled) because of the high

Table 4: Composition of dataset.

	Set 1: fully labeled	Set 2: single labeled	Set 3: unlabeled	Total
Graphs	6×25	37×25	2.000 ×1	3.075
Stress-strain curves	6×25	37×1	–	187

cost of the ODE-solver. The dataset thus includes 187 supervised/monitored samples ( $6 \times 25$  samples +  $37 \times 1$  sample) across 43 different parameter combinations. The graphs and corresponding stress-strain curves serve as ground truth examples for supervised learning. Additionally, set 3 (unlabeled) contains 2.000 graphs for 2.000 parameter combinations, cf., Table 4. Given an unseen parameter combination, our goal is to predict the average behavior as well as a range of deviation of the resulting stress-strain curves as close as possible to the ground truth. The data that we generated and used for the results shown in this paper is available for download at <https://github.com/pwelke/random-nonwoven-fibers>.

### 3 Linear Regression-based Predictive Models

This section deals with the two multivariate linear regression models recently proposed by Antweiler *et al.* [2]: the PP-model (production parameter-based) and the FGF-model (fiber graph feature-based). We explain the underlying modeling ideas (Section 3.1) and discuss the advantages of the model variants by means of a performance study (Section 3.2). For this purpose, we investigate the goodness of fit for the prediction of the mean stress-strain curves as an example. Both models aim at avoiding the high computational effort associated to the TSS-model. In the following the TSS-model is represented by the random field  $\mathcal{S} : \mathcal{U}_{ad} \times \Omega \rightarrow \mathcal{Y}$  with set of admissible production parameters  $\mathcal{U}_{ad} \subset \mathbb{R}^n$ ,  $n = 4$  (cf., Table 2) and set of stress-strain curve parametrizations  $\mathcal{Y} \subset \mathbb{R}^r$ ,  $r = 2$ , cf., (1).

#### 3.1 Linear Regression and Monte Carlo Simulations

**PP-model.** The PP-model directly relates the production parameters to the stress-strain curve parametrizations using multiple multivariate linear regression, cf., [16]. Given  $k \in \mathbb{N}$  observation pairs  $\{(\mathbf{u}_i, \mathbf{y}_i)\}_{i=1}^k$ , consisting of input (production parameters)  $\mathbf{u}_i \in \mathbb{R}^n$  and output (random stress-strain curve parametrizations)  $\mathbf{y}_i = \mathcal{S}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^r$ , the model assumes the relation

$$\mathbf{y}_i = \mathbf{b}_P + \mathbf{B}_{P,1}^T \mathbf{u}_i + \boldsymbol{\varepsilon}_{P,i} \quad \text{for } i = 1, \dots, k,$$

where the errors  $\varepsilon_{p,i} : \Omega \rightarrow \mathbb{R}^r$  account for the stochastic nature of the tensile strength simulation framework. They are assumed to be independent and identically distributed (i.i.d.) with  $E[\varepsilon_{p,i}] = \mathbf{0}$  and  $\text{Cov}[\varepsilon_{p,i}] = \Sigma_{\mathbf{p}} \in \mathbb{R}^{r \times r}$ . The task is to identify the unknown intercept  $\mathbf{b}_{\mathbf{p}} \in \mathbb{R}^r$  and regression coefficients  $\mathbf{B}_{\mathbf{p},1} \in \mathbb{R}^{n \times r}$ . The regression model for  $\mathbf{B}_{\mathbf{p}} = [\mathbf{b}_{\mathbf{p}}, \mathbf{B}_{\mathbf{p},1}^T]^T \in \mathbb{R}^{1+n \times r}$  can be summarized as

$$\mathbf{Y} = \mathbf{U}\mathbf{B}_{\mathbf{p}} + \mathbf{E}_{\mathbf{p}},$$

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1^T \\ \mathbf{y}_2^T \\ \vdots \\ \mathbf{y}_k^T \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 1 & \mathbf{u}_1^T \\ 1 & \mathbf{u}_2^T \\ \vdots & \vdots \\ 1 & \mathbf{u}_k^T \end{bmatrix}, \quad \mathbf{E}_{\mathbf{p}} = \begin{bmatrix} \varepsilon_{\mathbf{p},1}^T \\ \varepsilon_{\mathbf{p},2}^T \\ \vdots \\ \varepsilon_{\mathbf{p},k}^T \end{bmatrix}$$

with response  $\mathbf{Y} \in \mathbb{R}^{k \times r}$ , design matrix  $\mathbf{U} \in \mathbb{R}^{k \times (1+n)}$ , and error matrix  $\mathbf{E}_{\mathbf{p}} \in \mathbb{R}^{k \times r}$ . By the assumptions on the individual errors  $\text{Cov}[(\mathbf{E}_{\mathbf{p}})_{\cdot,i}, (\mathbf{E}_{\mathbf{p}})_{\cdot,j}] = (\Sigma_{\mathbf{p}})_{i,j} \mathbf{I}_k$  holds true, for  $i, j = 1, \dots, r$  and the identity  $\mathbf{I}_k \in \mathbb{R}^{k \times k}$ . Thus, the individual observations are independent, but correlations between the responses are allowed.

A linear, unbiased estimator of  $\mathbf{B}_{\mathbf{p}}$  is the well-known least-squares estimator  $\widehat{\mathbf{B}}_{\mathbf{p}} = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{Y}$ . Considering the decomposition  $\widehat{\mathbf{B}}_{\mathbf{p}} = [\widehat{\mathbf{b}}_{\mathbf{p}}, \widehat{\mathbf{B}}_{\mathbf{p},1}^T]^T$  yields the predictor

$$\widehat{\mathbf{y}}_{\mathbf{p}}(\mathbf{u}) = \widehat{\mathbf{b}}_{\mathbf{p}} + \widehat{\mathbf{B}}_{\mathbf{p},1}^T \mathbf{u}, \quad \text{with } \widehat{\mathbf{y}}_{\mathbf{p}} : \mathbb{R}^n \rightarrow \mathbb{R}^r, \quad (2)$$

that maps the production parameters to the associated mean stress-strain curve parametrization. If the errors are assumed to be multivariate normally distributed, i.e.,  $\varepsilon_{p,i} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{p}})$ , the maximum likelihood estimator of the covariance matrix  $\Sigma_{\mathbf{p}}$  is given by

$$\widehat{\Sigma}_{\mathbf{p}} = \frac{1}{k} \widehat{\mathbf{E}}_{\mathbf{p}}^T \widehat{\mathbf{E}}_{\mathbf{p}}, \quad \widehat{\mathbf{E}}_{\mathbf{p}} = \mathbf{Y} - \mathbf{U} \widehat{\mathbf{B}}_{\mathbf{p}}, \quad (3)$$

where  $\widehat{\mathbf{E}}_{\mathbf{p}}$  are the residuals between actual observation and prediction.

The PP-model approximates the TSS-model as  $\mathcal{S} \approx \widehat{\mathcal{S}}_{\mathbf{p}}$ , which is specified by  $\widehat{\mathcal{S}}_{\mathbf{p}}(\mathbf{u}, \cdot) \sim \mathcal{N}(\widehat{\mathbf{y}}_{\mathbf{p}}(\mathbf{u}), \widehat{\Sigma}_{\mathbf{p}})$  for all  $\mathbf{u} \in \mathcal{U}_{ad}$ . Note that  $\widehat{\mathbf{y}}_{\mathbf{p}}$  only predicts the mean stress-strain curve parametrization. However, if we use the additional distributional assumptions, we can resample multiple stress-strain curve parametrizations and insert them in the constant-quadratic ansatz (1). Averaging over the resulting curves yields then a prediction of the mean stress-strain curve.

**FGF-model.** The FGF-model, unlike the PP-model, builds on predicting the stress-strain curve parametrizations for individual fiber graphs. This requires the generation of random fiber structure samples, from each of which  $m \in \mathbb{N}$  features (i.e., combinations of production parameters, topological graph and stretch features, as listed in Table 3) are extracted. We view fiber graph generation and feature extraction as a random field  $\mathcal{M} : \mathcal{U}_{ad} \times \Omega \rightarrow \mathbb{R}^m$ . Then, the FGF-model relates production parameters and (fiber graph) features to the associated stress-strain curve parametrizations for which we again consider a multiple multivariate linear regression model.

Given  $k \in \mathbb{N}$  observation tuples  $\{(\mathbf{u}_i, \mathbf{v}_i, \mathbf{y}_i)\}_{i=1}^k$ , consisting of production parameters  $\mathbf{u}_i \in \mathbb{R}^n$ , random features  $\mathbf{v}_i = \mathcal{M}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^m$  and associated<sup>1</sup> stress-strain curve parametrizations  $\mathbf{y}_i = \mathcal{S}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^r$ , it reads

$$\mathbf{y}_i = \mathbf{b}_F + \mathbf{B}_{F,1}^T \mathbf{u}_i + \mathbf{B}_{F,2}^T \mathbf{v}_i + \boldsymbol{\varepsilon}_{F,i} \quad \text{for } i = 1, \dots, k,$$

with intercept  $\mathbf{b}_F \in \mathbb{R}^r$  and regression coefficients  $\mathbf{B}_{F,1} \in \mathbb{R}^{n \times r}$ ,  $\mathbf{B}_{F,2} \in \mathbb{R}^{m \times r}$ . The assumptions on the errors are the same as those of the PP-model, with covariance matrix  $\text{Cov}[\boldsymbol{\varepsilon}_{F,i}] = \boldsymbol{\Sigma}_F$ . However, the errors are here motivated as simple regression errors and not as sampling errors, as the FGF-model describes the input-output behavior of the deterministic tensile strength simulations. The model can be summarized as

$$\mathbf{Y} = \mathbf{W} \mathbf{B}_F + \mathbf{E}_F,$$

with  $\mathbf{B}_F = [\mathbf{b}_F, \mathbf{B}_{F,1}^T, \mathbf{B}_{F,2}^T]^T$  and  $\mathbf{W} = [\mathbf{U}, \mathbf{V}]$  using  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_k]^T$ . Again, the task is to identify intercept and regression coefficients, for which the unbiased linear least-squares estimator is given by  $\widehat{\mathbf{B}}_F = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{Y}$ . Thus, we obtain the linear predictor

$$\widehat{\mathbf{y}}_F(\mathbf{u}, \mathbf{v}) = \widehat{\mathbf{b}}_F + \widehat{\mathbf{B}}_{F,1}^T \mathbf{u} + \widehat{\mathbf{B}}_{F,2}^T \mathbf{v}, \quad \text{with } \widehat{\mathbf{y}}_F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^r, \quad (4)$$

that maps a given set of production parameters and fiber graph features to the stress-strain curve parametrization associated to the respective fiber graph.

The FGF-model approximates the TSS-model by the coupling of the predictor  $\widehat{\mathbf{y}}_F$  with the random field  $\mathcal{M}$ , i.e.,  $\mathcal{S} \approx \widehat{\mathcal{S}}_F$  where  $\widehat{\mathcal{S}}_F(\mathbf{u}, \cdot) \sim \widehat{\mathbf{y}}_F(\mathbf{u}, \mathcal{M}(\mathbf{u}, \cdot))$  for all  $\mathbf{u} \in \mathcal{U}_{ad}$ . As we have no analytical insights in the behavior of  $\mathcal{M}$ , this coupling has to be treated as a stochastic black box. To obtain a predictor of the mean stress-strain curve we have to conduct Monte-Carlo simulations where we repeatedly sample fiber graphs, predict their stress-strain curves and average over the results.

*Remark 1* The case where only fiber graph features are used for predictions in the FGF-model can be covered by choosing  $\widehat{\mathbf{B}}_{F,1} = \mathbf{0}$ . The corresponding least-squares estimator is  $[\widehat{\mathbf{b}}_F, \widehat{\mathbf{B}}_{F,2}] = (\widetilde{\mathbf{V}}^T \widetilde{\mathbf{V}})^{-1} \widetilde{\mathbf{V}}^T \mathbf{Y}$  with design matrix  $\widetilde{\mathbf{V}} = [\mathbf{1}, \mathbf{V}] \in \mathbb{R}^{k \times (1+m)}$  containing an extra column of ones to account for the intercept term.

*Remark 2* We state the closed form solution of the least-square estimator for convenience. In practice, we avoid solving the ill-conditioned normal equations and instead solve the associated least-squares optimization problem via pseudo inverse by means of a singular value decomposition.

---

<sup>1</sup> The explicit usage of a fixed production parameter combination  $\mathbf{u}_i$  together with a fixed probabilistic state  $\omega_i$  emphasizes that graph features and stress-strain curve parametrization are obtained from the same fiber graph sample (by simulation and feature extraction).

### 3.2 Numerical Results

**Experimental Setting.** To assess the predictive quality for inferring the mean stress-strain curves, we perform a leave-one-out cross-validation (LOOCV) across all 43 production parameter combinations contained in the labeled datasets 1&2 (Table 4). In each run, we separate the data into a training set containing the samples of 42 parameter combinations and a test set containing the samples of the remaining parameter combination. Hence, the test set always contains 25 fiber graph samples, where either one (set 2) or all of them (set 1) are labeled with a stress-strain curve parametrization. The training set is used to fit the PP-model and the FGF-model. For the latter one, we compare different combinations of the feature groups (listed in Table 3). During inference, the fitted models use the production parameters without/with fiber graph features as input to predict the mean stress-strain curve parametrizations (PP-model) or the stress-strain curve parametrizations associated to individual fiber graphs (FGF-model). For the FGF-model, the stress-strain curves of the 25 fiber graphs in the test set are reconstructed using the constant-quadratic ansatz (1). Averaging these curves yields a mean stress-strain curve prediction. For the PP-model, the procedure differs slightly. As it only predicts mean stress-strain curve parametrizations we use the covariance estimate (3) on top of the predicted mean parametrization (2) to resample 1.000 stress-strain curve parametrizations. Then, averaging over the associated stress-strain curves that are reconstructed by means of the constant-quadratic ansatz provides a prediction of the mean stress-strain curve.

For each production parameter combination, we compare the predicted mean curve to the ground truth curve. Using parameter combinations with multiple training samples (set 1), we compare the means of the predicted and the ground truth curves. For the single-sample parameter combinations (set 2) we take all 25 (mainly unlabeled) graph samples and check how much the ground truth curve of the single labeled sample deviates from the mean of the predicted curves that our model produces. For assessment, we use the test set to calculate the coefficient of determination,  $R^2$ , and the adjusted coefficient of determination,  $\bar{R}^2$ , between the means of predicted and ground truth curve evaluations. This provides a measure of the model fit that is independent of the strain. Thus, to compute the  $R^2$  values, we evaluate each curve (predicted and ground truth) at  $K$  ( $=1.000$ ) equally distanced strain points in the interval  $[0, 0.5]$  and take the means over the values at each strain point. In every run, this yields the predicted means  $\hat{y}_1, \dots, \hat{y}_K$  as well as the observed means  $\bar{y}_1, \dots, \bar{y}_K$ , from which we compute the (adjusted) coefficient of determination through

$$R^2 = 1 - \frac{\sum_{i=1}^K \bar{y}_i - \hat{y}_i}{\sum_{i=1}^K \bar{y}_i - \bar{\bar{y}}} \quad \text{and} \quad \bar{R}^2 = 1 - (1 - R^2) \frac{K - 1}{K - m - 1},$$

where  $\bar{\bar{y}} = \sum_{i=1}^K \bar{y}_i$  and  $m$  is the number of features used by the predictive model. Given the variability of the samples within the same parameter combination, this validation provides a robust estimation of the model quality. While the  $R^2$  value is a default evaluation score for regression tasks, we supplement it with the  $\bar{R}^2$  value which penalizes for larger numbers of selected attributes within a model. Furthermore,

Table 5: Regression results for the baseline, the constant-quadratic ansatz, the PP-model and the FGF-model. Listed are the medians of the  $R^2$  and  $\bar{R}^2$  values observed during the LOOCV as well as the OTLoss.

Model	Feature set	Median $R^2$ $\uparrow$	Median $\bar{R}^2$ $\uparrow$	OTLoss $\downarrow$
baseline	-	0.3928	-	-
constant-quadratic	-	<b>0.999927</b>	-	-
PP-model	param	0.7967	0.7958	292.56
FGF-model	stretch	0.9730	0.9714	111.77
FGF-model	graph	0.9737	0.9733	99.65
FGF-model	param + stretch	0.9723	0.9705	85.24
FGF-model	param + graph	0.9717	0.9712	82.62
FGF-model	graph + stretch	<b>0.9760</b>	<b>0.9742</b>	<b>71.44</b>
FGF-model	param + graph + stretch	0.9778	0.9761	85.71

we perform an Optimal Transport (OT) optimization between the sets of curves embedded in  $\mathbb{R}^K$ . It computes a mapping between two sets of points, that is minimal in terms of total work, i.e. transportation of mass. For optimization, the Wasserstein distance for discrete distributions is used. In comparison to the median  $R^2$  score, the OT score penalizes substantial differences between individual predicted and actual curves to a larger degree. With this additional score, we can adequately assess the difference in distribution between prediction and ground truth curves.

As baseline, in each run of the LOOCV we also compare the (mean) ground truth curve to the curve obtained by feeding the constant-quadratic ansatz with the means of the parametrizations in the training set. Computing the corresponding  $R^2$  values yields a simple comparative value to beat. Moreover, we include a comparison of the ground truth curves to that obtained by means of the best found constant-quadratic curve fits to get an idea of the suitability of the utilized stress-strain curve model class. Corresponding code and experimental data are available at <https://github.com/pwelke/random-nonwoven-fibers> and as a reproducible run on CodeOcean <https://codeocean.com/capsule/7514050/tree/v1> [see 3].

**Results and Discussion.** The main results for the prediction of the mean stress-strain curves are illustrated in Table 5. It reports the median (adjusted) coefficients of determination,  $R^2$  and  $\bar{R}^2$ , observed during the LOOCV. Most importantly, the results show that the constant-quadratic ansatz is a well-chosen approximation for the ground truth stress-strain curves, which is expressed by a median coefficient of determination that is very close to 1. Further, we note that both, the PP-model and the FGF-model, outperform the identified baseline by a clear margin. The fiber graph feature-based approach in particular works surprisingly well and delivers significant improvements over the simple production parameter-based approach. With regard to different feature set combinations, it should be emphasized that a union of topological graph and stretch features already achieves a remarkable performance with a median coefficient of determination of  $R^2 = 0.9760$ , calculated between the mean predicted

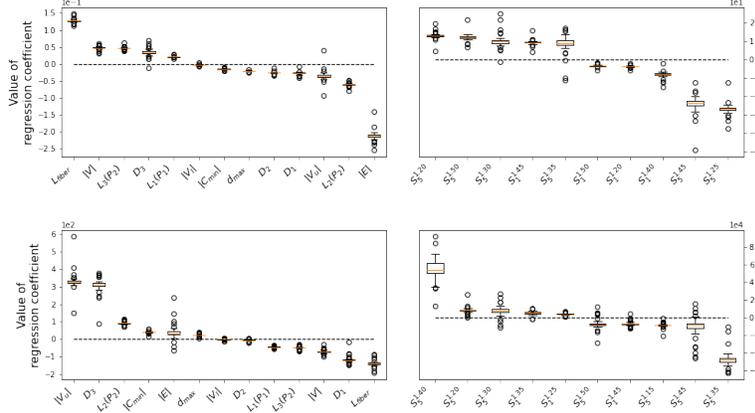


Fig. 5: Feature importance values for the FGF-model (graph + stretch) to predict  $\alpha$  (top) and  $\beta$  (bottom) with  $\mathbf{y} = (\alpha, \beta)$ . Topological graph (left) and stretch features (right). To reduce visual clutter, we display the five stretch features with biggest and smallest mean values, respectively. For reasons of comparison, the explanatory variables are scaled before training (min-max scaling).

and the mean ground truth curves (highlighted in bold in Table 5). This indicates that the topological and geometric structure of the fiber graphs already encodes much of the tensile strength behavior under vertical load. It should be noted, that we also compared a lasso and ridge regularization for the parameter estimation, leading to no significant change in results.

A major advantage of using simple regression models for prediction is the interpretability of the individual regression coefficients. In the following, we investigate the feature importance exemplarily for the FGF-model (graph + stretch) that uses the union of graph and stretch features for prediction. The regression weights observed during the 43-fold LOOCV are displayed in Fig. 5. Apparently, high impact features differ between  $\alpha$  and  $\beta$  prediction. Generally, stretch features display a large impact, especially  $S_1^c$  (mean) and  $S_5^c$  (sum) for larger values of the overstretching factor  $c$ . Examining the graph feature importance discloses the following relationships: For the prediction of  $\alpha$  the negative regression coefficient values with respect to  $|E|$  indicate that as the number of edges increases, the quadratic behavior of the stress-strain curves sets in earlier ( $\alpha$  is smaller). In line with that, the positive coefficient values for the prediction of  $\beta$  indicate that an increase in  $|E|$  also causes the quadratic incline to grow quicker ( $\beta$  is bigger). This underlines the intuition that more fiber connections result in firmer materials (higher tensile strength). Similar relationships can be observed for the maximum degree  $d_{\max}$  and the size of the minimum edge cut  $|C|_{\min}$ , as higher feature values are likely to represent a higher fiber structure connectivity. Opposed to that  $L_3(P_2)$ , the Euclidean length along the weighted shortest path in terms of fiber length, exhibits positive coefficient values for the prediction of  $\alpha$  and

negative coefficient values for the prediction of  $\beta$  (reversed effect). An explanation is that at high  $L_3(P_2)$  values, the shortest path has more leeway to be pulled apart during the tensile strength experiment without contributing to the tensile strength (lower tensile strength). The same applies to  $L_1(P_1)$ , the length of the shortest path in terms of edge count. Comparable interpretations cannot be made for all features, since some of them exhibit reciprocal relationships. Overall, the coefficients are stable over different parameter combinations, indicating a robust model fitting.

In comparison to the TSS-model, the regression models achieve a significant speedup. The time needed to compute a stress-strain curve for a sample generated by an unseen parameter combination is reduced by the FGF-model by more than three orders of magnitude, from 24 to 48 hours to two minutes per sample. As both workflows can be executed in a parallelized fashion, the speedup is of a factor greater than 1.000.

## 4 Sequential Predictive Regression Model

The prediction quality of the FGF-model is significantly better than that of the PP-model, but is brought by a costly underlying Monte-Carlo simulation procedure. Thus, the performance of the FGF-model crucially depends on the fiber graph generation. The remarkable difference in the prediction quality suggests some nonlinear relations between production parameters and fiber graph features. Such relationships are overlooked when using a purely linear model, as is the case with the PP-model. To capture the nonlinearities, we propose an intermediate multivariate polynomial regression model to infer mean topological graph features. This is a fairly straightforward approach, cf. [21, 24]. Alternatives may include multivariate adaptive regression splines (MARS) [13] or radial basis functions [32]. However, in our application, we observe already very good predictive results with polynomials of a total degree up to 5. To enable the prediction of mean stress-strain curve parametrizations we couple the intermediate model with an errors-in-variables model (Section 4.1). The quality of the resulting predictive pipeline, referred to as production parameter and mean fiber graph feature-based predictive model (PP-MGF-model), is investigated in comparison to the previously discussed models in Section 4.2. The PP-MGF-model is new, in view of the existing literature, and represents a good compromise between the efficiency of the PP-model and the predictive accuracy of the FGF-model.

### 4.1 Coupled Polynomial Regression and Errors-In-Variabels Model

The underlying assumption of the intermediate multivariate polynomial regression model is that the production parameters and the individual graph features obey a perturbed polynomial relation. Let  $\Gamma_{g,n}$  denote the set of  $n$ -dimensional multi-indices up to total degree  $g \in \mathbb{N}$  with cardinality  $l$ , i.e.,

$$\Gamma_{g,n} = \left\{ \boldsymbol{\gamma} \in \mathbb{N}_0^n : |\boldsymbol{\gamma}| = \sum_{j=1}^n \gamma_j \leq g \right\}, \quad l = |\Gamma_{g,n}| = \binom{n+g}{g}$$

and assume an arbitrary enumeration  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_l$  of the multi-indices such that  $\boldsymbol{\gamma}_1 = (0, \dots, 0)$ . Then the multivariate polynomials  $q_j : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $j = 1, \dots, m$ , are defined through

$$q_j(\mathbf{u}) = \sum_{i=1}^l c_{ij} \mathbf{u}^{\boldsymbol{\gamma}_i}, \quad \text{where } \mathbf{u}^{\boldsymbol{\gamma}_i} = \prod_{r=1}^n u_r^{\gamma_{ir}},$$

with polynomial coefficients  $c_{ij} \in \mathbb{R}$  and factor interactions between the individual production parameters  $\mathbf{u}^{\boldsymbol{\gamma}_i} \in \mathbb{R}$ . Given  $s \in \mathbb{N}$  observation pairs  $\{(\mathbf{u}_i, \mathbf{v}_i)\}_{i=1}^s$ , consisting of production parameters  $\mathbf{u}_i \in \mathbb{R}^n$  and random fiber graph features  $\mathbf{v}_i = \mathcal{M}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^m$ , the intermediate polynomial regression model assumes the relation

$$\mathbf{v}_i = \mathbf{q}(\mathbf{u}_i) + \boldsymbol{\varepsilon}_{R,i}, \quad \mathbf{q}(\mathbf{u}) = (q_1(\mathbf{u}), \dots, q_m(\mathbf{u}))^T = \mathbf{C}^T(\mathbf{u}^{\boldsymbol{\gamma}_1}, \dots, \mathbf{u}^{\boldsymbol{\gamma}_l})^T$$

with (unknown) coefficient matrix  $\mathbf{C} \in \mathbb{R}^{l \times m}$ ,  $(\mathbf{C})_{i,j} = c_{ij}$ . Analogously to the basic linear regression model, the errors  $\boldsymbol{\varepsilon}_{R,i}$  are assumed to be i.i.d. with  $\mathbb{E}[\boldsymbol{\varepsilon}_{R,i}] = \mathbf{0}$  and  $\text{Cov}[\boldsymbol{\varepsilon}_{R,i}] = \boldsymbol{\Sigma}_R$ , for  $i = 1, \dots, s$ . The task is to identify  $\mathbf{C}$  in order to simultaneously fit a multivariate polynomial of (total) degree  $g$  for each of the  $m$  (fiber graph) features. It is convenient to think of the  $l$  possible factor interactions as independent variables. This allows to reformulate the model as a multiple multivariate linear regression model, since linearity is only required with respect to the regression coefficients  $c_{ij}$ . Thus, let  $x_{ij} = u_i^{\gamma_j}$  be the set of explanatory variables, then we get

$$\mathbf{V} = \mathbf{X}\mathbf{C} + \mathbf{E}_R$$

with design matrix  $\mathbf{X} \in \mathbb{R}^{s \times l}$  ( $(\mathbf{X})_{i,j} = x_{ij}$ , containing the factor interactions), response matrix  $\mathbf{V} \in \mathbb{R}^{s \times m}$  and error matrix  $\mathbf{E}_R \in \mathbb{R}^{s \times m}$ . Note that no intercept must be included for setting up the design matrix  $\mathbf{X}$ , since  $x_{i1} = u_i^{\gamma_1} = 1$  for  $i = 1, \dots, s$ . Especially, for the case  $g = 1$  the polynomial regression model includes the classic multiple multivariate linear regression model with intercept. In view of the reformulation, an adequate estimator for  $\mathbf{C}$  is given by the least-squares estimator  $\widehat{\mathbf{C}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}$  which provides a (non-linear) predictor of the mean fiber graph features for given combinations of production parameters

$$\widehat{\mathbf{v}}(\mathbf{u}) = \widehat{\mathbf{C}}^T(\mathbf{u}^{\boldsymbol{\gamma}_1}, \dots, \mathbf{u}^{\boldsymbol{\gamma}_l})^T. \quad (5)$$

The objective is now to predict the expected stress-strain curve parametrizations based on the production parameters and mean fiber graph features. Assuming a linear relation, which has been shown to be accurate in the case of the FGF-model, the functional relation is as follows

$$\mathbf{y}_i = \mathbf{b} + \mathbf{B}_1^T \mathbf{u}_i + \mathbf{B}_2^T \widehat{\mathbf{v}}(\mathbf{u}_i) + \boldsymbol{\varepsilon}_i, \quad \text{for } i = 1, \dots, k. \quad (6)$$

Here  $\boldsymbol{\varepsilon}_i : \Omega \rightarrow \mathbb{R}^r$  models the deviation from the mean parametrization caused by the stochastic nature of the simulation framework. To fit the relationship (6) to data, we replace the predictor  $\widehat{\mathbf{v}}$  with the variable  $\bar{\mathbf{v}}$  representing the mean features. In this context, we note that sampling data to fit the model using the TSS-model is not feasible, because the mean graph features  $\bar{\mathbf{v}}$  are not directly observable. Instead, we only have access to observations tuples  $\{(\mathbf{u}_i, \mathbf{v}_i, \mathbf{y}_i)\}_{i=1}^k$  composed of production parameters  $\mathbf{u}_i \in \mathbb{R}^n$ , random features  $\mathbf{v}_i = \mathcal{M}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^m$  and associated stress-strain curve parametrizations  $\mathbf{y}_i = \mathcal{S}(\mathbf{u}_i, \omega_i) \in \mathbb{R}^r$ . However, the fiber graph features can be thought of as perturbed realizations of  $\bar{\mathbf{v}}$ , i.e.,  $\mathbf{v}_i = \bar{\mathbf{v}}_i + \boldsymbol{\delta}_i$ . Thereby,  $\boldsymbol{\delta}_i \in \mathbb{R}^m$  represents the error of measuring  $\bar{\mathbf{v}}_i$ . Thus, in addition to the conventional errors in the regression equation, we assume errors in the explanatory variables as well. This results in the usage of the generalized errors-in-variables model [29] which assumes the relation

$$\mathbf{y}_i = \mathbf{b} + \mathbf{B}_1^T \mathbf{u}_i + \mathbf{B}_2^T \bar{\mathbf{v}}_i + \boldsymbol{\varepsilon}_i, \quad (7a)$$

$$\mathbf{v}_i = \bar{\mathbf{v}}_i + \boldsymbol{\delta}_i. \quad (7b)$$

In (7) the observable variables are  $\mathbf{y}_i$ ,  $\mathbf{u}_i$  and  $\mathbf{v}_i$ , whereas  $\bar{\mathbf{v}}_i$  is referred to as latent variable. Analogously to the multivariate linear regression model, the joint errors  $\boldsymbol{\psi}_i = (\boldsymbol{\delta}_i^T, \boldsymbol{\varepsilon}_i^T)^T$  are assumed to be i.i.d. with  $\mathbb{E}[\boldsymbol{\psi}_i] = \mathbf{0}$  and  $\text{Cov}[\boldsymbol{\psi}_i] = \boldsymbol{\Sigma}$ , for  $i = 1, \dots, k$ . Then the task is to estimate  $\mathbf{B} = [\mathbf{b}, \mathbf{B}_1^T, \mathbf{B}_2^T]^T$ . By applying the conventional least-squares estimator  $\widehat{\mathbf{B}} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{Y}$  (cf., Section 3.1), we neglect the measurement error described by (7b) during estimation. Even though it is well known that the least-squares estimator is not a consistent estimator for  $\mathbf{B}$  in the errors-in-variables model, it gives good results for prediction [5, 29].

Eventually, the coupling of the feature predictor (5) from the polynomial regression model with the fitted errors-in-variables model yields a mapping from the production parameters to the mean stress-strain curve parametrizations. It is determined by the intercept  $\widehat{\mathbf{b}}$  and the coefficient matrices  $\widehat{\mathbf{B}}_1, \widehat{\mathbf{B}}_2, \widehat{\mathbf{C}}$  according to the previous explanations and results in the (nonlinear) predictor

$$\widehat{\mathbf{y}}(\mathbf{u}) = \widehat{\mathbf{b}} + \widehat{\mathbf{B}}_1^T \mathbf{u} + \widehat{\mathbf{B}}_2^T \widehat{\mathbf{C}}^T (\mathbf{u}^{\gamma_1}, \dots, \mathbf{u}^{\gamma_l})^T, \quad \text{with } \widehat{\mathbf{y}} : \mathbb{R}^n \rightarrow \mathbb{R}^r. \quad (8)$$

A very convenient property of this coupling is that we can use different datasets for fitting the polynomial regression model and for fitting the errors-in-variables model, cf., Fig. 2. Particularly, since tensile strength simulations (computational bottleneck) are not necessary, the dataset for fitting the polynomial regression model can be chosen much larger ( $s \gg k$ ). This is appropriate in order to account for the larger number of explanatory variables.

Conclusively, to approximate the input-output behavior of the TSS-model, we need distributional assumptions for the joint error behavior. Again, we rely on a multivariate normal distribution (similar to the PP-model) and determine a covariance estimate  $\widehat{\boldsymbol{\Sigma}}$  analogously to (3). Then, the PP-MGF-model behaves as  $\mathcal{S} \approx \widehat{\mathcal{S}}$ ,

where  $\widehat{S}(\mathbf{u}, \cdot) \sim \mathcal{N}(\widehat{y}(\mathbf{u}), \widehat{\Sigma})$  for all  $\mathbf{u} \in \mathcal{U}_{ad}$ . Predicting the mean stress-strain curve requires resampling, as it is the case for the PP-model.

## 4.2 Numerical Results

**Experimental Setting.** Using dataset 3 (Table 4), we investigate the relation between production parameters and topological graph features by means of a 5-fold cross validation. Therefore, the set is randomly divided in 5 subsets, containing 400 samples each. In each run, one of the subsets is used as test set, while the remaining ones are used for training. We train the multivariate polynomial regression model for a total degree of  $g \in \{1, \dots, 5\}$ , using a least-squares estimator for fitting the regression coefficients. To assess the model quality, we compare the median adjusted coefficients of determination,  $\bar{R}^2$ , observed throughout the cross-validation.

Subsequently, we investigate the quality of the PP-MGF-model for predicting the mean stress-strain curves. To achieve a fair comparison with regard to the PP-model and the FGF-model, we again perform a leave-one-out cross-validation (LOOCV) across the 43 production parameter combinations (dataset 1&2), as described in Section 3.2. To train the polynomial regression model, we use all fiber graphs (labeled and unlabeled) associated to the training set. Thereby, we test polynomial relations of the degree  $g \in \{2, \dots, 6\}$ . To train the errors-in-variables model, we use the labeled training data only. For both models, we employ a least-squares fit. During inference on the test set, the fitted models use the production parameter combinations as input in order to predict the mean stress-strain curve parametrizations. To obtain a prediction of the mean stress-strain curve we resample 1.000 stress-strain curve parametrizations, reconstruct the associated curves using the constant-quadratic ansatz and then average over them (similar to the PP-model). In comparing the predicted mean stress-strain curves to the ground truth curves, we follow the descriptions from Section 3.2.

**Results and Discussion.** The main results of the described 5-fold cross validation are summarized in Table 6. We observe that the adjusted  $\bar{R}^2$  values, acting as a measure of model fit, peak for a degree of 3 and 4. Further increasing the polynomial degree for regression leads to a deterioration in terms of the adjusted  $\bar{R}^2$  value. Since the case of polynomial degree 1 resembles the linear model, an improvement by moving to a higher polynomial degree is apparent. The high  $\bar{R}^2$  values, which are even above 0.9 in most cases, are particularly astonishing and justify the use of a polynomial model for the mean fiber graph feature prediction.

The results regarding the 43-fold LOOCV are summarized in Table 7. The predictive results of the PP-MGF-model outperform that of the PP-model by a clear margin and almost reach the predictive quality of the FGF-model. We note that this is achieved without the need of a Monte-Carlo simulation procedure. Particularly, a polynomial fit of total degree 5 works best for the relation between production parameters and topological graph features. For higher degrees, we observe a deterioration of the  $\bar{R}^2$  value, which is probably related to overfitting. We note that we also tested fitting the errors-in-variables model by means of a generalized total-least-

Table 6: Results of the 5-fold cross-validation: Median of  $\bar{R}^2$  values for the prediction of the topological graph features and for different polynomial degrees.

Feature	Degree 1	Degree 2	Degree 3	Degree 4	Degree 5
$ V $	0.8526	0.9698	<b>0.9812</b>	0.9809	0.9741
$ E $	0.9552	0.9852	<b>0.9893</b>	0.9892	0.9868
$ V_u $	0.8465	0.9344	<b>0.9460</b>	0.9460	0.9361
$ V_l $	0.8596	0.9416	<b>0.9499</b>	0.9466	0.9350
$d_{\max}$	0.4303	0.5861	0.6241	<b>0.6352</b>	0.5567
$L_{\text{fiber}}$	0.9551	0.9823	<b>0.9883</b>	0.9880	0.9852
$L_1(P_1)$	0.8819	0.9448	<b>0.9469</b>	0.9423	0.9329
$L_2(P_2)$	0.8799	0.9403	0.9393	<b>0.9406</b>	0.9261
$L_3(P_2)$	0.8761	0.9356	<b>0.9369</b>	0.9358	0.9190
$D_1$	0.8776	0.9536	0.9680	<b>0.9687</b>	0.9437
$D_2$	0.7835	0.9292	0.9638	<b>0.9652</b>	0.8153
$D_3$	0.9460	0.9842	<b>0.9897</b>	0.9891	0.9851
$ C_{\min} $	0.8494	0.9150	0.9311	<b>0.9395</b>	0.9223

Table 7: Regression results of the LOOCV comparing the PP-model, the FGF-model and the PP-MGF-model: Median of observed  $R^2$  and  $\bar{R}^2$  values as well as OT loss.

Model	Approach	Median $R^2 \uparrow$	Median $\bar{R}^2 \uparrow$	OTLoss $\downarrow$
PP-model	param	0.7967	0.7958	292.56
PP-MGF-model	Degree 2	0.9282	0.9258	125.37
PP-MGF-model	Degree 3	0.9428	0.9396	130.01
PP-MGF-model	Degree 4	0.9427	0.9372	121.63
PP-MGF-model	Degree 5	<b>0.9584</b>	<b>0.9515</b>	<b>93.43</b>
PP-MGF-model	Degree 6	0.9572	0.9447	110.16
FGF-model	param + graph	0.9717	0.9712	82.62

squares estimation. However, this did not improve the prediction quality, for which the results presented are limited to the use of a conventional least-squares estimator.

In comparison to the TSS-model, the time needed to compute a stress-strain curve for a sample generated by an unseen parameter combination is reduced by more than six orders of magnitude, from 24-48 hours to 10 milliseconds per sample. In that, the PP-MGF-model is similar to the PP-model and more than three orders of magnitude better than the FGF-model in terms of computation time. However, note that the training of the PP-MGF-model is slightly more expensive than that of the other regression models, since it depends on a large amount of additional graph samples to fit the nonlinear relations between production parameters and graph features.

Summing up, the PP-MGF-model is cheap to evaluate and has excellent predictive quality, making it suitable for nonwoven material design. To conclude our discussion we refer to Fig. 6 which shows the predicted mean stress-strain curves of all models included in the predictive model hierarchy. Although isolated instances also led to

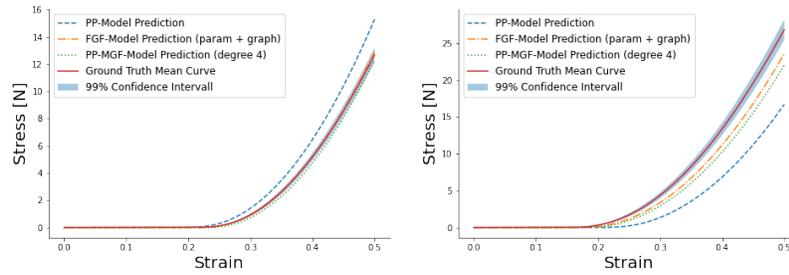


Fig. 6: Mean stress-strain curve predictions resulting from our predictive model hierarchy. The exemplarily illustrated instances are observed during the LOOCV for two fully labeled test sets (belonging to set 1).

other predictive gradations, the plots are representative of the observable results of the models and reflect the results of the numerical experiments performed.

## 5 Conclusion and Future Work

This work demonstrates the power of informed machine learning in predicting material properties. We developed a regression-based model hierarchy for predicting the tensile strength behavior of nonwovens. While direct linear regression on the production parameters lacks accuracy (PP-model) and linear regression using individual fiber graphs requires time-consuming Monte-Carlo simulations (FGF-model), a coupling of a polynomial model with a linear (errors-in-variables) model (PP-MGF-model) has proven to be a good compromise combining the best of both model variants. By reducing the computation time by several orders of magnitude, a high accuracy of the prediction results (compared to the ground truth) is achieved. Thus, the PP-MGF-model promises to be of great benefit as a surrogate model for nonwoven material design, which is a field for further work. Our approach incorporates extensive domain knowledge into the modeling process at the points of training data and hypothesis set via simulation results and algebraic equations from scientific and expert sources. To our knowledge, our approach is completely new in the context of nonwoven material design.

## References

- [1] Abou-Nassif, G. A. 2015. Predicting the Tensile and Air Permeability Properties of Woven Fabrics using Artificial Neural Network and Linear Regression Models. *Journal of Textile Science & Engineering*, 5(5).

- [2] Antweiler, D., Harmening, M., Marheineke, N., Schmeißer, A., Wegener, R., & Welke, P. 2022a. Graph-based tensile strength approximation of random nonwoven materials by interpretable regression. *Machine Learning with Applications*, **8**, 100288.
- [3] Antweiler, D., Harmening, M., Marheineke, N., Schmeißer, A., Wegener, R., & Welke, P. 2022b. Machine learning framework to predict nonwoven material properties from fiber graph representations. *Software Impacts*, **14**, 100423.
- [4] Chen, T., Li, L., Koehl, L., Vroman, Ph., & Zeng, X. 2007. A soft computing approach to model the structure–property relations of nonwoven fabrics. *Journal of Applied Polymer Science*, **103**(1), 442–450.
- [5] Cheng, C.-L., & Van Ness, J. W. 1999. *Statistical Regression with Measurement Error*. John Wiley & Sons.
- [6] Cheng, L., & Adams, D. L. 1995. Yarn Strength Prediction Using Neural Networks: Part I: Fiber Properties and Yarn Strength Relationship. *Textile Research Journal*, **65**(9), 495–500.
- [7] Das, D., & Pourdeyhimi, B. 2014. *Composite Nonwoven Materials: Structure, Properties and Applications*. Elsevier Science.
- [8] Demirci, E., Acar, M., Pourdeyhimi, B., & Silberschmidt, V. 2011. Finite Element Modelling of Thermally Bonded Bicomponent Fibre Nonwovens: Tensile Behaviour. *Computational Materials Science*, **50**(4), 1286–1291.
- [9] Eltayib, H. E., Ali, A. H. M., & Ishag, I. A. 2016. The Prediction of Tear Strength of plain weave fabric using Linear Regression Models. *International Journal of Advanced Engineering Research and Science*, **3**(11), 151–154.
- [10] Faessel, M., Delisée, C., Bos, F., & Castéra, P. 2005. 3D Modelling of random cellulosic fibrous networks based on X-ray tomography and image analysis. *Composites Science and Technology*, **65**(13), 1931–1940.
- [11] Fan, J., & Hunter, L. 1998. A Worsted Fabric Expert System: Part II: An Artificial Neural Network Model for Predicting the Properties of Worsted Fabrics. *Textile Research Journal*, **68**(10), 763–771.
- [12] Farukh, F., Demirci, E., Sabuncuoglu, B., Acar, M., Pourdeyhimi, B., & Silberschmidt, V. 2015. Mechanical analysis of bi-component-fibre nonwovens: Finite-element strategy. *Composites Part B: Engineering*, **68**, 327–335.
- [13] Friedman, J. H. 1991. Multivariate Adaptive Regression Splines. *The Annals of Statistics*, **19**(1), 1–67.
- [14] Gramsch, S., Klar, A., Leugering, G., Marheineke, N., Nessler, Ch., Strohmeyer, Ch., & Wegener, R. 2016. Aerodynamic web forming: process simulation and material properties. *Journal of Mathematics in Industry*, **6**(1), 1–13.
- [15] Harmening, M., Marheineke, N., & Wegener, R. 2021. Efficient Graph-based Tensile Strength Simulations of Random Fiber Structures. *ZAMM Journal of Applied Mathematics and Mechanics*, **101**(13).
- [16] Johnson, Richard A., & Wichern, Dean W. 2007. *Applied multivariate statistical analysis*. 6 edn. Upper Saddle River, NJ: Pearson.
- [17] Karpatne, A., Watkins, W., Read, J., & Kumar, V. 2017. Physics-guided Neural Networks (PGNN): An Application in Lake Temperature Modeling. *arXiv e-prints 1710.11431*.

- [18] Kufner, T., Leugering, G., Semmler, J., Stingl, M., & Strohmeier, C. 2018. Simulation and structural optimization of 3D Timoshenko beam networks. *ESAIM: Mathematical Modelling and Numerical Analysis*, **52**(6), 2409–2431.
- [19] Le Bris, C. 2010. Some numerical approaches for weakly random homogenization. *Pages 29–45 of: Numerical Mathematics and Advanced Applications 2009*. Springer.
- [20] Lu, Y., Rajora, M., Zou, P., & Liang, S. Y. 2017. Physics-Embedded Machine Learning: Case Study with Electrochemical Micro-Machining. *Machines*, **5**(1), 4.
- [21] Nizam, A., Rosenberg, E., Kleinbaum, D. G., & Kupper, L. L. 2013. *Applied regression analysis and other multivariable methods*. Brooks/Cole.
- [22] Ohser, J., & Mücklich, F. 2000. *Statistical Analysis of Microstructures in Materials Science*. Weinheim: John Wiley.
- [23] Ohser, J., & Schladitz, K. 2009. *3D Images of Materials Structures: Processing and Analysis*. Weinheim: Wiley-VCH.
- [24] Ostertagová, E. 2012. Modelling using Polynomial Regression. *Procedia Engineering*, **48**, 500–506.
- [25] Rahnema, M., Semnani, D., & Zarrebini, M. 2013. Measurement of the Moisture and Heat Transfer Rate in Light-weight Nonwoven Fabrics Using an Intelligent Model. *Fibres and Textiles in Eastern Europe*, **21**, 89–94.
- [26] Raina, A., & Linder, C. 2014. A homogenization approach for nonwoven materials based on fiber undulations and reorientation. *Journal of the Mechanics and Physics of Solids*, **65**, 12–34.
- [27] Ribeiro, R., Pilastrri, A., Moura, C., Rodrigues, F., Rocha, R., Morgado, J., & Cortez, P. 2020. Predicting Physical Properties of Woven Fabrics via Automated Machine Learning and Textile Design and Finishing Features. *Pages 244–255 of: Artificial Intelligence Applications and Innovations*. Springer.
- [28] Schladitz, K., Peters, S., Reinel-Bitzer, D., Wiegmann, A., & Ohser, J. 2006. Design of acoustic trim based on geometric modeling and flow simulation for non-woven. *Computational Materials Science*, **38**(1), 56–66.
- [29] van Huffel, S., & Vandewalle, J. 1991. *The Total Least Squares Problem*. Society for Industrial and Applied Mathematics.
- [30] von Rüden, L., Mayer, S., Sifa, R., Bauckhage, C., & Garcke, J. 2020. Combining Machine Learning and Simulation to a Hybrid Modelling Approach: Current and Future Directions. *Pages 548–560 of: Advances in Intelligent Data Analysis*. Lecture Notes in Computer Science, vol. 12080. Springer.
- [31] von Rüden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Walczak, M., Pfrommer, J., Pick, A., Ramamurthy, R., Garcke, J., Bauckhage, C., & Schuecker, J. 2021. Informed Machine Learning – A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems. *IEEE Transactions on Knowledge and Data Engineering*.
- [32] Walczak, B., & Massart, D.L. 1996. The Radial Basis Functions — Partial Least Squares approach as a flexible non-linear regression technique. *Analytica Chimica Acta*, **331**(3), 177–185.

- [33] Wegener, R., Marheineke, N., & Hietel, D. 2015. Virtual Production of Filaments and Fleeces. *Pages 103–162 of: Currents in Industrial Mathematics.* Springer.

# Index

- Armijo's line search, 13
- coefficient of determination, 17
- cross-validation, 17
- differential system, 7
- dynamical system, 13
- errors-in-variables model, 20
- Euler scheme, 13
- feature extraction, 16
- feature importance, 19
- graph, 7
  - graph connectivity, 10
- Hooke's law, 8
- inference, 23
- insulation material, 2
- interpretability, 19
- Monte-Carlo simulation, 3
- neural networks, 3
- Newton method, 13
- nonwoven airlay manufacturing, 2
- numerical experiment, 25
- optimal transport, 18
- parallelization, 20
- random field, 15
- regression, 4
  - linear regression, 4
  - polynomial regression, 5
- regularization, 19
- simulation, 2
- surrogate model, 3
- Wiener process, 7