

Predicting EHL film thickness parameters by machine learning approaches

Max MARIAN^{1,*}, Jonas MURSAK², Marcel BARTZ², Francisco J. PROFITO³, Andreas ROSENKRANZ⁴, Sandro WARTZACK²

¹ Department of Mechanical and Metallurgical Engineering, School of Engineering, Pontificia Universidad Católica de Chile, Santiago 6904411, Chile

² Engineering Design, Friedrich-Alexander-University Erlangen-Nuremberg (FAU), Erlangen 91058, Germany

³ Department of Mechanical Engineering, Polytechnic School of the University of São Paulo, São Paulo 05508-030, Brazil

⁴ Department of Chemical Engineering, Biotechnology and Materials (DIQBM), FCFM, Universidad de Chile, Santiago 8370456, Chile

Received: 24 November 2021 / Revised: 09 March 2022 / Accepted: 21 April 2022

© The author(s) 2022.

Abstract: Non-dimensional similarity groups and analytically solvable proximity equations can be used to estimate integral fluid film parameters of elastohydrodynamically lubricated (EHL) contacts. In this contribution, we demonstrate that machine learning (ML) and artificial intelligence (AI) approaches (support vector machines, Gaussian process regressions, and artificial neural networks) can predict relevant film parameters more efficiently and with higher accuracy and flexibility compared to sophisticated EHL simulations and analytically solvable proximity equations, respectively. For this purpose, we use data from EHL simulations based upon the full-system finite element (FE) solution and a Latin hypercube sampling. We verify that the original input data are required to train ML approaches to achieve coefficients of determination above 0.99. It is revealed that the architecture of artificial neural networks (neurons per layer and number of hidden layers) and activation functions influence the prediction accuracy. The impact of the number of training data is exemplified, and recommendations for a minimum database size are given. We ultimately demonstrate that artificial neural networks can predict the locally-resolved film thickness values over the contact domain 25-times faster than FE-based EHL simulations (R^2 values above 0.999). We assume that this will boost the use of ML approaches to predict EHL parameters and traction losses in multibody system dynamics simulations.

Keywords: machine learning; elastohydrodynamic lubrication; film thickness; support vector machine; Gaussian process regression; artificial neural network

1 Introduction

Reducing friction and wear losses in highly loaded lubricated tribo-contacts of machine elements or mechanical components is essential for developing energy-efficient and reliable systems [1–3]. In particular, the modeling of concentrated, elastohydrodynamically lubricated (EHL) contacts (see Fig. 1), in which local elastic deformation of the rubbing surfaces and hydrodynamic fluid film formation are superimposed,

is comparatively complex and computationally expensive [4]. For a sufficiently accurate and computationally efficient incorporation of EHL effects into higher-level multibody system dynamics simulations, analytically solvable approximating equations are generally used to estimate integral fluid film parameters [4].

To generalize the EHL calculations and their results, non-dimensional similarity groups were introduced. Dowson and Higginson [5, 6] defined a group of

* Corresponding author: Max MARIAN, E-mail: max.marian@ing.puc.cl

Nomenclature

b_H	Half Hertzian contact width: $b_{H,2D} = \sqrt{(4 \cdot F_N \cdot E' \cdot R) / 2 \cdot \pi \cdot l}$, $b_{H,3D} = \sqrt[3]{(3 \cdot F_N \cdot E' \cdot R) / 8}$	u_m	Entrainment speed: $u_m = (u_1 + u_2) / 2$
C	Box constraint	W	Load parameter
C	Hooke's law elasticity matrix	x	Cartesian spatial coordinate
d	Test field distance	x_n	Training point
E	Equivalent Young's modulus	X	Normalized cartesian spatial coordinate
E'	Reduced Young's modulus: $E' = 1/2[(1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2]$	$x_{ij,LHD}$	Latin hypercube design element
E_i	Young's modulus	$x_{ij,LHS}$	Latin hypercube sampling element
f	Optimization function	y	Cartesian spatial coordinate
F_N	Normal load	\bar{y}	Arithmetic mean of the target variables
G_k	Kernel function	\hat{y}	Approximation value
G	Material parameter	Y	Normalized cartesian spatial coordinate
h_b	Base function	z	Cartesian spatial coordinate
h	Lubricant film thickness	Z	Normalized cartesian spatial coordinate
H	Normalized lubricant film thickness	Z_r	Random number
h_c	Central lubricant film thickness	α_n	Nonnegative multiplier
H_c	Central film thickness parameter	α_n^*	Nonnegative multiplier
h_{min}	Minimum lubricant film thickness	α_p	Pressure-viscosity coefficient
H_{min}	Minimum film thickness parameter	β_b	Base function coefficients
h_0	Rigid body distance	β	SVM parameter
J	Objective function	γ	Penalty function
k	Covariance function	δ	Elastic deformation
l	Contact length	$\bar{\delta}$	Normalized elastic deformation
L_g	Lagrange dual formulation optimization problem	ε_L	Lagrangian small-strain tensor
L	Viscosity parameter	ε	Permitted error
L_ε	Linear ε -insensitive loss function	η	Fluid viscosity
m	Mean function	η_0	Base viscosity
M	Load parameter	$\bar{\eta}$	Normalized fluid viscosity
n_f	Number of factors	θ	Fractional film content
n_s	Number of simulations	θ_k	Kernel function hyperparameter
N	Number of observations	κ	Ellipticity parameter
p	Fluid pressure	ξ	MaxiMin factor
P	Normalized fluid pressure	ξ_n	Slack variable
p_H	Hertzian contact pressure: $p_{H,2D} = (3 \cdot F_N) / (2 \cdot \pi \cdot b_H^2)$, $p_{H,3D} = (2 \cdot F_N) / (\pi \cdot l \cdot b_H)$	ξ_n^*	Slack variable
R	Effective radius: $R = 1 / (R_1^{-1} + R_2^{-1})$	ρ	Fluid density
R^2	Coefficient of determination	ρ_0	Base density
R_i	Radius	$\bar{\rho}$	Normalized fluid density
\mathbf{U}	Displacement vector	σ^2	Error variance
U	Velocity parameter	ν	Equivalent Poisson's ratio
u_i	Surface velocity	ν_i	Poisson's ratio
		ψ	Penalty factor
		Ω	Solution domain
		Ω_c	Contact domain

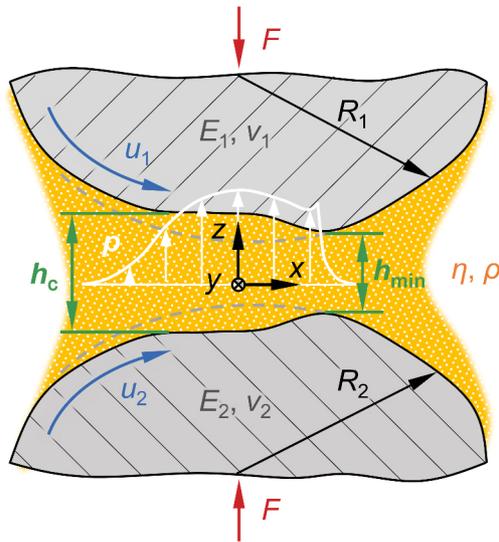


Fig. 1 Schematic of an EHL infinite 2D line contact with relevant input variables. Reproduced with permission from Ref. [4], © The authors 2020.

material, velocity, load, and lubricant film parameters, and proposed curve-fitted regression formulas (hereinafter referred to as “proximity equations”) relating those parameters from numerical simulation results. Blok and Moes [7, 8] proved that these four parameters could be transformed into three independent parameters, thus introducing a load and a viscosity parameter in addition to the fluid film parameter. Later, Johnson [9] proposed film thickness parameters for elasticity and a pressure viscosity coefficient, which, in turn, can be derived from the Blok/Moes notation. More recently, Habchi et al. [10] suggested using the Weissenberg number, the Nahme-Griffith number, a limiting shear stress-pressure coefficient, and a thermo-viscous regime indicator. However, the non-dimensional groups proposed by Dowson/Higginson and Blok/Moes are the most widely used ones in EHL literature [4, 11].

To estimate the minimum lubricant film thickness in an EHL infinite 2D line contact, Dowson and Higginson [5, 6] proposed an analytically solvable regression equation as a function of velocity, material, and load parameters. Based upon more advanced isothermal EHL simulations, these parameters have been modified [7, 12, 13] and extended by various authors to cover 3D circular and elliptical point contacts and estimate the central film thickness [14–21]. Due to the transformability, the Blok/Moes parameters

can be used to calculate the EHL film parameters. However, since these are only proximity formulas with a limited validity range, Johnson [9] differentiated four regions depending on the viscosity and material behavior. Besides the classical EHL regime, proximity equations were suggested for the hydrodynamic as well as for two semi-elastohydrodynamic regimes with rigid bodies and iso-viscous fluid behavior, respectively. Finally, Moes et al. [19, 22] developed further approximations based upon precise EHL simulations with validity over a wider range of concentrated contacts.

In these correlations, thermal effects, limited oil supply, and shear-thinning fluid behavior have not been considered. One way to account for these aspects is to apply correction factors to the EHL film thickness calculations [4]. Various authors have proposed thermal correction factors [23–27], each with a different range of validity, to adjust the central or minimum EHL film thickness. Similar approaches have also been developed to adjust the EHL film thickness for starvation [28–30], fluid compressibility [31, 32], non-Newtonian fluid behavior [33–35], and surface roughness effects under mixed lubrication conditions [36–38].

For more fundamentals and details about non-dimensional groups, film thickness equations, and applicable correction factors, the interested reader is referred to the recent review article by Marian et al. [4]. The latter [4, 39] also hypothesized that machine learning (ML) or artificial intelligence (AI) algorithms provide opportunities to predict relevant EHL film parameters more accurately and effectively. Currently, ML and AI methods are receiving growing attention in the field of tribology [40] and involve the development of computing systems that are able to learn from training data (input) and build/refine experience-based models to predict a certain behavior (output) [41]. ML algorithms can be categorized as supervised, unsupervised and reinforcement learning, whereby the selection of suitable approaches is highly task-dependent [40, 41]. For instance, support vector machine (SVM) represents an object set by a vector within a vector space. Hyperplanes within the space are used to separate the data points. Kernel functions are used to transform the vector space into

any higher-dimensional space so that interlaced vector groups are linearly separable [42, 43]. In the context of tribology, SVMs have been successfully applied to tailor composite materials [44, 45] as well as in the area of drive technology [46, 47] and manufacturing [48].

Furthermore, artificial neural network (ANN) is a flexible and the most widely used approach [40] with successful applications in the field of composite materials [44, 49–51], drive technology [52–56], manufacturing [48, 57], surface engineering [58–60], and lubricant formulation [61, 62]. ANN is inspired by natural brains' architecture and involves a number of simple but highly interconnected information processing elements (neurons) [63], see Fig. 2. Transfer or propagation functions determine the neurons' network input based on the output weighting, whereas the calculation of the outputs is achieved by activation functions considering a threshold value [41]. During training, the weightings and thresholds are adjusted to optimize the overall prediction quality. The topology or architecture of ANNs, i.e., how many neurons are parallelly arranged in each layer and how many hidden layers are between the input and output layer, needs to be tailored for the respective application, whereas overfitting must be avoided. ANN has also already been used in the context of predicting the

behavior of EHL contacts. Even though there are first physics-informed ML approaches to predict the behavior of lubricated contacts [64], most of the work done so far was data-driven based upon designs of experiments (DoE) [40, 65]. Otero et al. [66] trained an ANN with 20 neurons in a single hidden layer to predict the coefficient of friction in micro-textured EHL contacts under various operating conditions and dependent on the textures' dimensions and patterns. Thereby, the underlying data was obtained from experimental ball-on-disk experiments (mini-traction machine). Marian et al. [67, 68] demonstrated that numerical EHL simulations could also be utilized for training approximation or meta-models. Nevertheless, ML approaches have not yet been adapted to predict lubricant film parameters in EHL contacts as a function of standard input variables, as is also accomplished by the well-known proximity equations [4, 39], see Fig. 2.

In this context, this contribution is based on the hypothesis that ML approaches can predict relevant EHL film parameters such as the central film thickness h_c and the minimum lubricant gap h_{\min} in 2D line and in 3D circular point contacts more efficiently than sophisticated simulation models and with higher accuracy and flexibility than analytically solvable proximity equations based upon non-dimensional

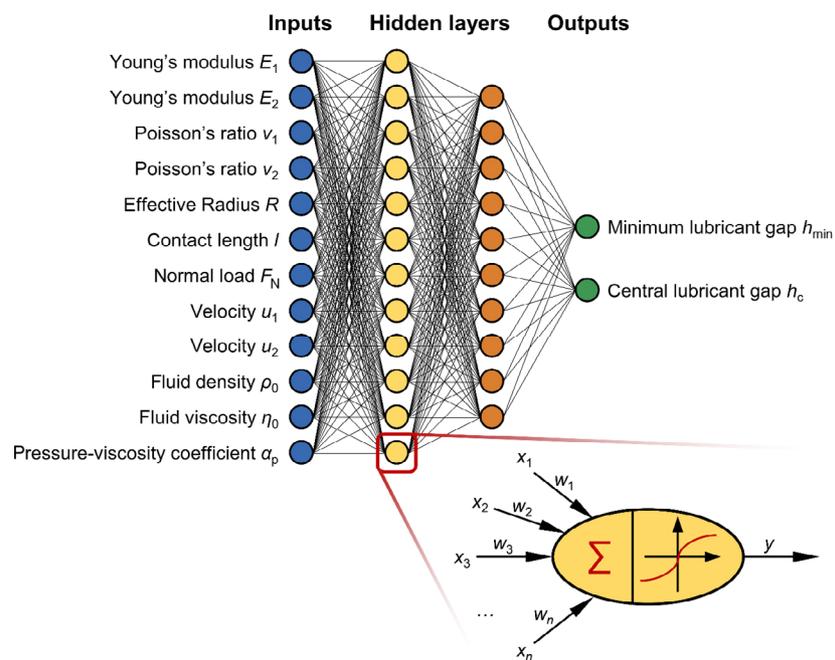


Fig. 2 Schematic of the correlation between EHL contact parameters and film thickness parameters using ANN.

groups. Thereby, the following research questions will be addressed:

1) How accurate are the predictions of SVMs, Gaussian process regression (GPR), and ANNs compared to the established analytically solvable proximity equations?

2) Does using all available input parameters (see Fig. 2) instead of non-dimensional groups to train the ML approaches affect the prediction accuracy?

3) How do the ANN architecture and structure affect prediction quality?

4) What is the influence of the size of the data base on the prediction quality?

5) Is it possible to determine locally resolved lubricant film thickness distributions, i.e., the lubricant gap as a function of the contact length $h(x)$, with sufficient accuracy and more efficiently than by EHL contact simulation?

2 Theory and methods

To answer these questions, analytically solvable proximity equations (Section 2.1) are compared with SVM, GPR as well as ANN (Section 2.4) based upon designs of experiments (Section 2.2) and data sets generated by EHL simulations (Section 2.3) with respect to their prediction quality of EHL film parameters. The overall approach is illustrated in Fig. 3.

2.1 Analytically solvable proximity equations

The proximity equations used to predict lubricant film parameters in EHL infinite 2D line or 3D circular point contacts are based on the most widely used non-dimensional similarity group proposed by Dowson and Higginson [5, 6], as summarized in Table 1. Thereby, E' is the reduced Young's modulus, R the effective radius, l the contact length, η the fluid viscosity, α_p the pressure-viscosity coefficient, u_m the entrainment velocity, and F_N the normal load. These parameters can be transformed into the notation from Blok and Moes [7, 8], see Table 2.

Given their widespread use, practicability and applicability to the overall scope covered within this contribution (see Section 2.2 and Refs. [4, 11]), the proximity equations derived by Dowson and Toyoda [14]

Table 1 Non-dimensional parameters according to Dowson/Higginson [5, 6].

Velocity parameter	$U = \frac{\eta_0 \cdot u_m}{E' \cdot R}$	(1)
Material parameter	$G = \alpha_p \cdot E'$	(2)
Load parameter	$W_{2D} = \frac{F_N}{E' \cdot l \cdot R}$	(3)
	$W_{3D} = \frac{F_N}{E' \cdot R^2}$	(4)
Film thickness parameter	$H_c = \frac{h_c}{R}$	(5)
	$H_{min} = \frac{h_{min}}{R}$	(6)

Table 2 Blok/Moes non-dimensional parameters according to Blok/Moes [7, 8].

Load parameter	$M_{2D} = W_{2D} (2 \cdot U)^{\frac{1}{2}}$	(7)
	$M_{3D} = W_{3D} (2 \cdot U)^{\frac{3}{4}}$	(8)
Viscosity parameter	$L = G (2 \cdot U)^{\frac{1}{4}}$	(9)
Film thickness parameter	$H_{min} = h_{min} \cdot U^{\frac{1}{2}}$	(10)

$$H_{c,2D} = 3.06 \cdot G^{0.56} \cdot U^{0.69} \cdot W_{2D}^{-0.1} \quad (11)$$

Dowson and Higginson [5, 6]

$$H_{min,2D} = 1.6 \cdot G^{0.6} \cdot U^{0.7} \cdot W_{2D}^{-0.13} \quad (12)$$

and Hamrock and Dowson [16]

$$H_{c,3D} = 2.69 \cdot G^{0.53} \cdot U^{0.67} \cdot W_{3D}^{-0.067} \cdot (1 - 0.61 \cdot e^{-0.73 \cdot \kappa}) \quad (13)$$

$$H_{min,3D} = 3.63 \cdot G^{0.49} \cdot U^{0.67} \cdot W_{3D}^{-0.073} \cdot (1 - e^{-0.68 \cdot \kappa}) \quad (14)$$

which represent polynomial regression equations, were used to estimate the central and minimum lubricant film thicknesses. Within this study, circular point contacts and, therefore, the same radii in x - (R_x) and in y -direction (R_y) were assumed for the ellipticity parameter

$$\kappa = 1.03 \cdot \left(\frac{R_x}{R_y} \right)^{0.64} \quad (15)$$

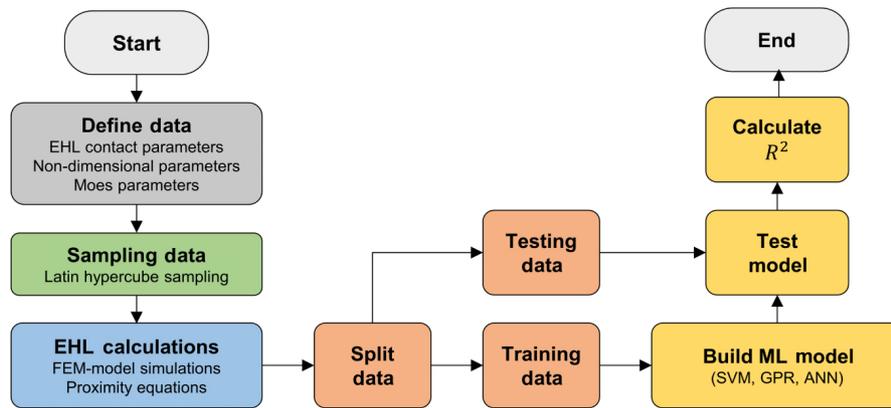


Fig. 3 Flowchart of the ML methodology proposed in the present contribution.

2.2 Design of experiments

To generate a sufficient data base to determine correlations between input and output variables with minimum computational effort, statistical DoE methods can be employed to systematically plan and evaluate computer experiments. In addition to full and partial factorial, central composite, or Box-Behnken, designs tailored to the characteristics of numerical simulation and statistical modeling can be used. To generate a broad spectrum of approximation or meta-models, a Latin hypercube design (LHD) or Latin hypercube sampling (LHS) from the group of equally distributed test fields are particularly suitable. Thereby, the data points are distributed so that they fill the factor space as evenly as possible and provide information about almost every region of the factor space with little computational effort (small number of simulations n_s) despite many factors n_f . The LHS is based on a modification of a generated LHD, which represents an $n_s \times n_f$ matrix containing random permutations of the numbers $\{1, 2, 3, \dots, n_f\}$ within its columns. The LHS elements are generated by subtracting a random number between zero and one $Z_r [0, 1)$ from each LHD element $x_{ij,\text{LHD}}$ and then dividing this value by the number of trial points [69]:

$$x_{ij,\text{LHS}} = \frac{x_{ij,\text{LHD}} - Z_r [0, 1)}{n_s} \quad (16)$$

Unlike other DoEs, the number of data points is directly specified in the LHS. The resulting test field can then be transformed to the desired factor space. In this work, the desired factor space is delimited

by the minimum and maximum values of the EHL parameters, as summarized in Table 3. The ranges of these values were based on typical machine elements that operate under hard EHL conditions, such as rolling bearings or gears [70], and were chosen to cover the parameter space without substantial changes to the individual simulation models (e.g., in terms of numerical stabilization, see Section 2.3) and with as many converged calculations as possible. Thus, the ranges of the non-dimensional Dowson/Higginson parameters (U, G, W) and Blok/Moes parameters (M, L), as shown in Table 3, could be covered. Thereby, the conventional cases of EHL infinite 2D line and 3D circular point contacts were sampled with 1,500 and 1,000 data sets, respectively. The former was able to include more data points due to the lower computational effort. Moreover, it was investigated for the 2D case to what extent a reduced data base (600, 300, and 100 data points) affects the prediction quality. According to Johnson et al. [71], the distances between the data points can be used to assess the quality of the test field regarding uniform distribution and freedom of correlation. The minimum distance between the individual test points was maximized for a suitable and uniformly distributed LHS test field. Considering all distances of the test field d , the MaxiMin criterion

$$\text{MaxiMin} = \left[\sum_{1 \leq i < j \leq n_s} d(x_i, x_j)^{-\xi} \right]^{\frac{1}{\xi}} \quad (17)$$

was obtained, with the integer, positive, and application-dependent factor ξ [69]. Within the scope of this work, the MATLAB's Statistics and Machine

Table 3 Factor space with minima (min.) and maxima (max.) values of various inputs of infinite 2D line and 3D circular EHL point contacts.

Training data	Input	Unit	EHL 2D line contact		EHL 3D point contact	
			min.	max.	min.	max.
EHL contact parameters	$E_{1,2}$	GPa	200	440	200	440
	$\nu_{1,2}$	—	0.30	0.35	0.30	0.35
	R	m	0.0075	0.02	0.05	0.10
	l	m	0.0025	0.01	—	—
	F_N	N	750	1,500	100	1,000
	$u_{1,2}$	m/s	0.025	0.4	0.025	5
	ρ_0	kg/cm ³	850	1,100	850	1,100
	η_0	Pa·s	0.005	0.05	0.01	0.2
	α_p	Pa ⁻¹	1.25×10^{-8}	2.50×10^{-8}	1.75×10^{-8}	2.50×10^{-8}
Non-dimensional parameters	G	—	2,750	8,350	2,400	8,360
	U	—	1.9×10^{-14}	1.2×10^{-11}	7.5×10^{-14}	1.5×10^{-10}
	W	—	1.1×10^{-5}	3.6×10^{-4}	3.0×10^{-8}	3.3×10^{-5}
Moes parameters	M	—	4.6	935	4.4	1,000
	L	—	1.3	17	0.6	6.2

Learning Toolbox was used to generate an LHS optimized according to the MaxiMin criterion.

2.3 EHL modeling

Numerical EHL modeling was done by applying the steady-state isothermal Reynolds differential equation [72]:

$$\underbrace{\frac{\partial}{\partial x} \left(\frac{\rho \cdot h^3}{12 \cdot \eta} \frac{\partial p}{\partial x} \right)}_{\text{Poiseuille term}} + \frac{\partial}{\partial x} \left(\frac{\rho \cdot h^3}{12 \eta} \frac{\partial p}{\partial y} \right) - \underbrace{\frac{\partial}{\partial x} \left(\theta \cdot \rho \cdot h \cdot \frac{u_1 + u_2}{2} \right)}_{\text{Couette term}} = 0 \tag{18}$$

in a slightly modified notation for hydrodynamics, whereby p is the pressure, h the lubricant film thickness, η the viscosity, ρ the density, θ the fractional film content, u_1 and u_2 the surface velocities (see Fig. 1). The fluid was assumed to be compressible and piezo-viscous with the density following Dowson and Higginson [73]:

$$\rho(p) = \rho_0 \cdot \left(1 + \frac{0.6 \cdot p}{1 + 1.7 \cdot p} \right) \tag{19}$$

as well as the dynamic viscosity following Roelands [74]:

$$\eta(p) = \eta_0 \cdot e^{\left[\ln(\eta_0) + 9.67 \right] \cdot \left[-1 + \left(1 + \frac{p}{1.96 \times 10^8 \text{ Pa}} \right)^{\frac{\alpha_p \cdot 1.96 \times 10^8 \text{ Pa}}{\ln(\eta_0) + 9.67}} \right]} \tag{20}$$

This is an important prerequisite to be consistent with the assumptions employed to obtain the EHL film thickness proximity equations (see Section 2.1). Cavitation was addressed by a mass-conserving penalty formulation of the fractional film content [75]

$$\theta(p) = e^{-\gamma(p) \cdot p^2} \tag{21}$$

where $\gamma(p)$ is zero if $p < 0$ and otherwise a sufficiently high algebraic number ξ . The lubricant gap equation

$$h(x, y) = h_0 + \frac{x^2}{2 \cdot R} + \frac{y^2}{2 \cdot R} + \delta \tag{22}$$

was comprised by the rigid body motion h_0 , a quadratic approximation of the undeformed geometry and the elastic deformation δ . The latter was calculated by solving the steady-state linear elasticity equation for an equivalent elastic body with negligible body forces, which can be expressed in the contract matrix (Voigt) notation as

$$\nabla[\mathbf{C} \cdot \boldsymbol{\varepsilon}_L(\mathbf{U})] = 0 \quad (23)$$

where \mathbf{C} is the generalized Hooke's law elasticity matrix, $\boldsymbol{\varepsilon}_L$ the contracted Lagrangian small-strain tensor, and \mathbf{U} the displacement vector. The surface displacement is admitted being the normal component U_z , i.e.,

$$\delta = |U_z| \quad (24)$$

Furthermore, the equivalent body was assumed to be of a homogenous isotropic material with equivalent Young's modulus

$$E = \frac{E_1^2 \cdot E_2 \cdot (1 + \nu_2)^2 + E_1 \cdot E_2^2 \cdot (\nu_1 + 1)^2}{[E_1 \cdot (1 + \nu_2) + E_2 \cdot (\nu_1 + 1)]^2} \quad (25)$$

and equivalent Poisson's ratio

$$\nu = \frac{E_1 \cdot \nu_2 \cdot (1 + \nu_2) + E_2 \cdot \nu_1 \cdot (1 + \nu_1)}{E_1 \cdot (1 + \nu_2) + E_2 \cdot (1 + \nu_1)} \quad (26)$$

by applying the linear elasticity equation. The integral of the hydrodynamic pressure over the contact domain Ω_c balanced the normally applied load to satisfy the force equilibrium:

$$\int_{\Omega_c} p \, dx \, dy = F_N \quad (27)$$

To solve the EHL problem and to ensure good conditioning, the relevant variables were normalized on Hertzian (subscript H) or reference values (subscript 0):

$$\begin{aligned} X &= \frac{x}{b_H}, Y = \frac{y}{b_H}, P = \frac{p}{p_H}, H = \frac{h \cdot R}{b_H^2}, \\ \bar{\delta} &= \frac{h \cdot R}{b_H^2}, \bar{\eta} = \frac{\eta}{\eta_0}, \bar{\rho} = \frac{\rho}{\rho_0} \end{aligned} \quad (28)$$

The numerical solution scheme was based on the full-system approach [76], whose overall procedure is depicted in Fig. 4(a). After reading the inputs, initial values were determined following the Hertzian theory to define an initial guess for the elastic deformation solution. Subsequently, the Reynolds equation was solved in a weak formulation on the contact domain Ω_c and fully (strongly) coupled with the calculation

of the elastic deformation in the solution domain Ω based upon FEM. Especially for higher pressures, instabilities can occur in the solution of the Reynolds equation. This is because convection-diffusion equations converge to local oscillations of the solution variables when solved with Galerkin FEM in convection-dominated problems [77]. Partially, computational stability can be improved by finer discretization and higher-order approximation functions. In this work, 5th and 7th order were used for the 2D and 3D models, respectively. Moreover, other stabilization methods can serve as artificial diffusion by introducing additional terms into the transport equations. In this study, the residual-based (consistent) stabilized Galerkin least squares (GLS) method [78] and the inconsistent method of isotropic diffusion (ID) [79] were used, whereby care was taken to minimize the influence of the stabilization on the numerical solution. Regarding the boundary conditions, zero pressure (Dirichlet) was applied at the contact domain's (Ω_c) in- and outlet. Furthermore, zero displacements on the bottom, the hydrodynamic pressure as normal stress on the top (Ω_c), and free boundary with zero normal and shear stresses on the remaining borders were applied as boundary conditions of the elastic body (Ω). The domains differed between the 2D infinite line contact (Fig. 4(b)) and the 3D circular point contact case (Fig. 4(c)) and were chosen sufficiently large to avoid numerical starvation, thus ensuring that the results correspond to the ones predicted by an infinite elastic half-space approach. The 2D domain was discretized by triangular elements and the 3D domain by tetrahedral elements with refinements in the contact center of the upper surface. Symmetry boundary conditions were used for the 3D case to reduce computational efforts. The interested reader is referred to Refs. [77, 80, 81] for more details about the fundamentals of FEM applied to EHL problems and its implementation within the software COMSOL Multiphysics.

2.4 Machine learning

Various ML regression methods, including SVM, GPR, and ANN, were employed to predict EHL film parameters using the MATLAB's Machine Learning and Deep Learning Toolbox [82]. All data (input and

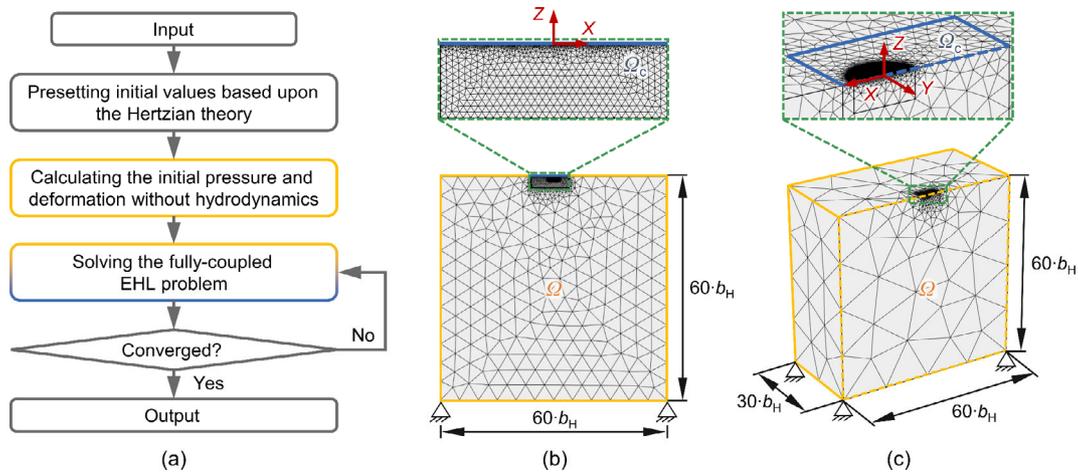


Fig. 4 (a) Schematic of the numerical solution of the EHL model and computational domains with meshing and boundary conditions for (b) infinite 2D line and (c) 3D circular point contacts.

output values) were normalized to values between -1 and 1 using the MATLAB built-in function “*mapminmax*”. The ML models were developed using training data to provide the best possible predictions for unknown test data. For this purpose, the data sets (see Section 2.2) were divided into 85% for training and 15% for testing. Due to its superior expressiveness [83], the coefficient of determination defined as

$$R^2 = \frac{\text{sum of squared regression}}{\text{total sum of squares}} = \frac{\sum (\hat{y} - \bar{y})^2}{\sum (y - \bar{y})^2} \quad (29)$$

was used to evaluate the prediction quality. The above equation represents the proportion of the variation in the response variable y explained by the independent input variables, whereas \bar{y} is the arithmetic mean of the target variables, and \hat{y} is the approximation value [84, 85]. The prediction quality of the ML approaches was compared after training with the original input data of the EHL simulations (please refer to Fig. 1 and Table 3) and after training with the non-dimensional parameters U, G, W , and M, L (please refer to Section 2.1 and Table 3), respectively.

2.4.1 Support vector machines

The linear epsilon-insensitive SVM (ϵ -SVM) regression was implemented in MATLAB’s Machine Learning and Deep Learning Toolbox [82]. The goal was to find

a function

$$f(x) = x' \beta + b \quad (30)$$

for each training point x_n of a multi-variate set of N observations that was as flat as possible through formulation as a convex optimization problem and deviated from the response value y_n by a value smaller than ϵ [82]. The slack variables ξ_n and ξ_n^* are included to deal with otherwise infeasible constraints results in the objective function (primal formula)

$$J(\beta) = \frac{1}{2} \beta' \beta + C \sum_{n=1}^N (\xi_n + \xi_n^*) \quad (31)$$

constrained by [82, 86]:

$$\begin{aligned} \forall n : y_n - (x_n' \beta + b) &\leq \epsilon + \xi_n, \\ \forall n : (x_n' \beta + b) - y_n &\leq \epsilon + \xi_n^*, \\ \forall n : \xi_n &> 0, \\ \forall n : \xi_n^* &> 0 \end{aligned} \quad (32)$$

with the box constraint C , a numeric value to control the penalty imposed on observations outside the ϵ -margin. The linear ϵ -insensitive loss function is described by [82]:

$$L_\epsilon = \begin{cases} 0 & \text{if } |y - f(x)| \leq \epsilon \\ |y - f(x)| & \text{otherwise} \end{cases} \quad (33)$$

To computationally simplify this optimization problem,

its Lagrange dual formulation

$$L_g(\alpha) = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) G(x_i, x_j) + \varepsilon \sum_{i=1}^N (\alpha_i - \alpha_i^*) + \sum_{j=1}^N y_j (\alpha_j - \alpha_j^*) \quad (34)$$

with the nonnegative multipliers α_n and α_n^* for each observation x_n , was used and subjected to the constraints [82]

$$\begin{aligned} \sum_{n=1}^N (\alpha_n - \alpha_n^*) &= 0, \\ \forall n : 0 \leq \alpha_n &\leq C, \\ \forall n : 0 \leq \alpha_n^* &\leq C \end{aligned} \quad (35)$$

The parameter

$$\beta = \sum_{n=1}^N (\alpha_n - \alpha_n^*) x_n \quad (36)$$

was described as a linear combination of the training observations, and the function

$$f(x) = \sum_{n=1}^N (\alpha_n - \alpha_n^*) G_k(x_n, x) + b \quad (37)$$

was used to predict new values which depend only on the support vectors [82]. Thereby, $G_k(x_j, x_k)$ was a linear

$$G_k(x_j, x_k) = x_j' x_k \quad (38)$$

Gaussian

$$G_k(x_j, x_k) = e^{(-\|x_j - x_k\|^2)} \quad (39)$$

or polynomial

$$G_k(x_j, x_k) = (1 + x_j' x_k)^q, \text{ with } q = \{2, 3, \dots\} \quad (40)$$

semidefinite kernel function to map x to a higher dimensional space [82]. Finally, the Karush–Kuhn–Tucker (KKT) complementarity conditions were applied to find optimal solutions [82]:

$$\forall n : \alpha_n (\varepsilon + \xi_n - y_n + f(x_n)) = 0,$$

$$\forall n : \alpha_n^* (\varepsilon + \xi_n^* - y_n - f(x_n)) = 0,$$

$$\forall n : \xi_n (C - \alpha_n) = 0,$$

$$\forall n : \xi_n^* (C - \alpha_n^*) = 0 \quad (41)$$

The minimization problem was solved by an iterative single data algorithm (ISDA) [82, 87].

2.4.2 Gaussian process regression

A GPR non-parametric probabilistic model was implemented in MATLAB’s Machine Learning and Deep Learning Toolbox [88]. The GPR explained the response by latent variables $f(x_i)$ from a Gaussian process (GP) and explicit base functions. The covariance function of the latent variables described the smoothness of the response, and base functions $h_b(x)$ transferred the inputs into a p -dimensional space. A GP is a set of random variables with Gaussian distribution, and was defined by its mean $m(x)$ and kernel parametrized or hyper-parametrized covariance function $k(x, x')$ [88]

$$f(x) \sim GP[m(x), k(x, x', \theta_k)] \quad (42)$$

With the error variance σ^2 and the p -by-1 vector of the base function coefficients β_b , the response of the GPR was modeled as [88]:

$$P(y_i | f(x_i), x_i) \sim N(y_i | h(x_i)^T \beta_b + f(x_i), \sigma^2) \quad (43)$$

During training, the MATLAB built-in function “*fitrgp*” estimated the base function coefficients β_b , the noise variance σ^2 , and the hyperparameter θ_k of the kernel function [88].

2.4.3 Artificial neural networks

Multi-hidden layer feedforward backpropagation ANNs were trained by the MATLAB’s app “*nnstart*” utilizing the Levenberg-Marquart algorithm [89–91]. Thereby, the neuron weights were adjusted layer-by-layer, starting backwards with the connection to the output layer [92]. 10% of the training data was used for validation to avoid overfitting. At first, the most favorable architecture of the ANNs was determined, for which the number of neurons per layer varied between 10 and 20, while the number of hidden layers was assumed to lay between zero and three. If not explicitly mentioned, the hyperbolic tangent function

$$\tanh(x) = \frac{2}{1 + e^{-2x}} - 1 \quad (44)$$

was used as activation function, linear identity

$$\text{identity}(x) = x \quad (45)$$

as output function and

$$\text{net}_j = \sum_k \omega_{kj} \cdot o_k \quad (46)$$

as propagation function [92, 93]. After having identified the best architecture, the prediction was compared to training with sigmoid (logistic):

$$\text{sigmoid}(x) = \frac{2}{1 + e^{-x}} \quad (47)$$

and identity activation functions [93]. The results were compared with data obtained after training with reduced data sets (600, 300, and 100 data points) to study the influence of the number of training data.

In the FEM simulations used for data generation, the global central and minimum lubricant film thicknesses and the spatially resolved field quantities

(e.g., fluid pressure, surface elastic deformation, lubricant film gap) were calculated. To use the ANNs to predict the local lubricant film thickness in the EHL infinite 2D line contact, the position along the contact length ($-1.2 \leq X \leq 1.2$) was used as an additional input variable. With a resolution of 86 uniformly distributed points in this length domain, a total number of 128 914 data sets were obtained, which were divided into training, test and testing data.

3 Results and discussion

3.1 EHL simulations

Representative distributions of the lubricant film gap and the hydrodynamic pressure along the contact length of the infinite 2D line contact for four exemplary cases (lower M and higher L , higher M and higher L , lower M and lower L , and higher M and lower L) are shown in Fig. 5. The presented diagrams illustrate typical characteristics of EHL contacts, such as the elastic flattening in the contact center and the additional constriction in the lubricant gap near the contact outlet. Compared to the Hertzian

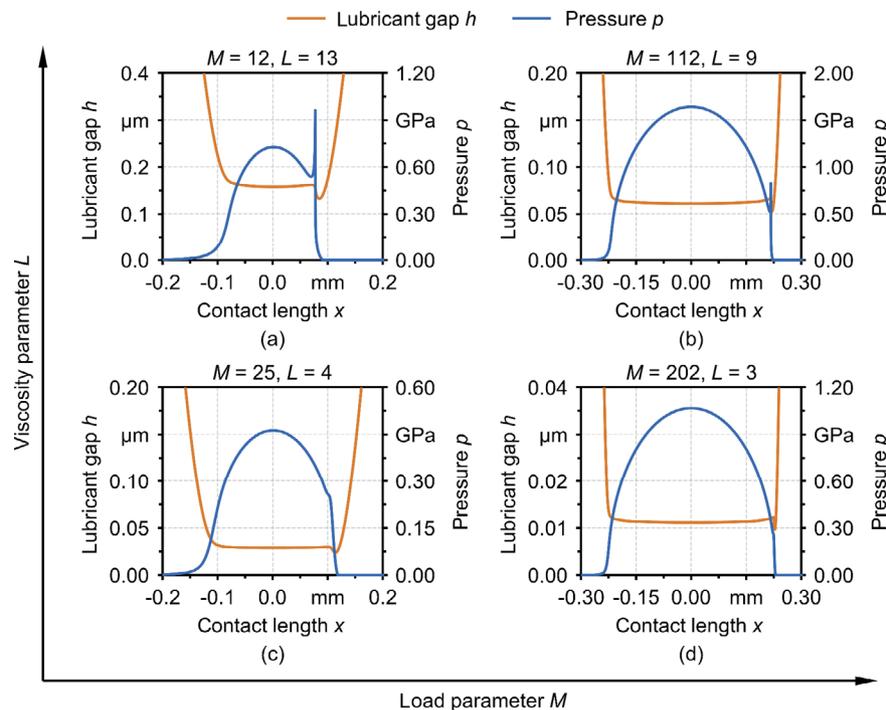


Fig. 5 Representative distributions of the lubricant film and the hydrodynamic pressure along the contact length for the infinite 2D line contact with four exemplary cases including (a) lower M and higher L , (b) higher M and higher L , (c) lower M and lower L , and (d) higher M and lower L .

theory, the fluid pressure distribution increased earlier due to lubricant compression at the contact inlet and displayed the additional Petrusевич peak at the contact outlet. The same holds for the 3D circular point contact (not shown for spatial reasons), whereby the film thickness showed the characteristic horseshoe shape. Generally, a higher M led to a fluid pressure distribution approaching the Hertzian theory and, with constant L , to a decrease in the lubricant film thickness. A higher L with constant M , in turn, increased the magnitude of the Petrusевич spike and the fluid film thickness. From the performed calculations within the LHS, 1,499 converged data sets for the 2D case and 764 for the 3D case were ultimately obtained to compare with the proximity equations (Section 3.2.1) and train the ML models (Section 3.2.2).

3.2 Prediction of EHL parameters

3.2.1 Analytically solvable film thickness equations

As shown in Table 4, the coefficients of determination of the analytically solvable proximity equations featured good prediction quality despite the DoE slightly exceeding these equations' validity range, see Table 3 and Ref. [4]. All data sets were used for the calculation, which reduces the influence of individual errors and outliers. The R^2 values were in a similar range as the fits between proximity equations and experimental film thickness measurements using an

optical ball-on-disk tribometer as reported by van Leeuwen [11]. However, the comparison with calculated values derived from EHL simulations (Fig. 6) revealed that the film thicknesses tended to be slightly overestimated by the proximity equations, especially for larger lubricant gaps. This can be attributed to the partly different modeling aspects (e.g., the rheological and cavitation models) between the EHL simulations employed here and the original data used to obtain the proximity equations.

3.2.2 Machine learning

3.2.2.1 Global EHL film parameters

The R^2 values of the central and minimum film thicknesses for the infinite 2D line and 3D circular point contacts against testing data as predicted by SVM and GPE methods are summarized in Table 5 and graphically compared in Fig. 7. With the original input data (see Fig. 1 and Table 3), the values calculated with EHL simulations were accurately predicted by both SVM and GPR methods (Figs. 7(a)–7(h)). The coefficients of determination with values larger than 0.99 (Table 5) reached higher values than the analytically solved proximity equations (Section 3.2.1). An overestimation, as seen for the analytically solved proximity equations, was not verifiable. Comparing both ML approaches, GPR proved to be slightly superior to SVM (Table 5).

Table 4 Coefficients of determination of the analytically solved proximity equations for the central and minimum film thickness for the infinite 2D line and 3D circular point contacts, respectively.

	EHL 2D line contact		EHL 3D point contact	
	h_c	h_{min}	h_c	h_{min}
R^2	0.996	0.999	0.991	0.986

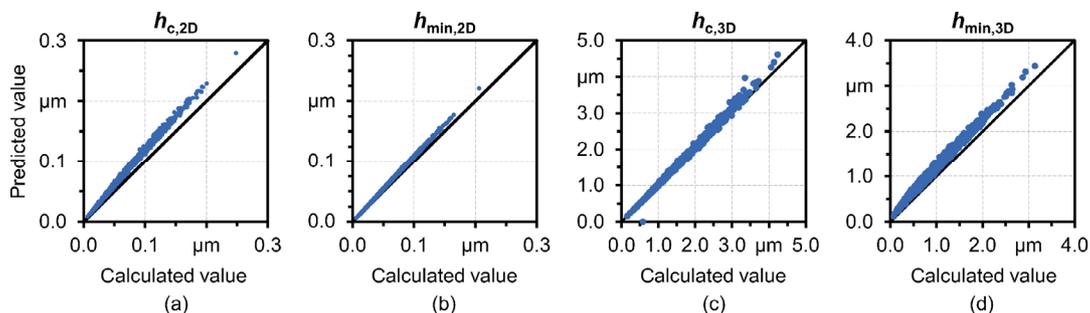


Fig. 6 Central and minimum film thickness predicted from the proximity equations (vertical axis) versus the calculated ones from EHL simulations (horizontal axis) for the (a, b) infinite 2D line and (c, d) 3D circular point contacts.

Table 5 Coefficients of determination of the SVM and GPR predictions for the central and minimum film thicknesses for the infinite 2D line and 3D circular point contacts against testing data after training with the original EHL inputs, the non-dimensional parameters U, G, W , or M, L .

R^2	EHL 2D line contact				EHL 3D point contact			
	SVM		GPR		SVM		GPR	
Training data	h_c	h_{min}	h_c	h_{min}	h_c	h_{min}	h_c	h_{min}
Original EHL inputs	0.997	0.997	1.000	1.000	0.996	0.997	0.998	0.998
U, G, W	0.797	0.794	0.100	0.091	0.875	0.901	0.013	0.901
M, L	0.732	0.722	0.732	0.723	0.680	0.740	0.678	0.739

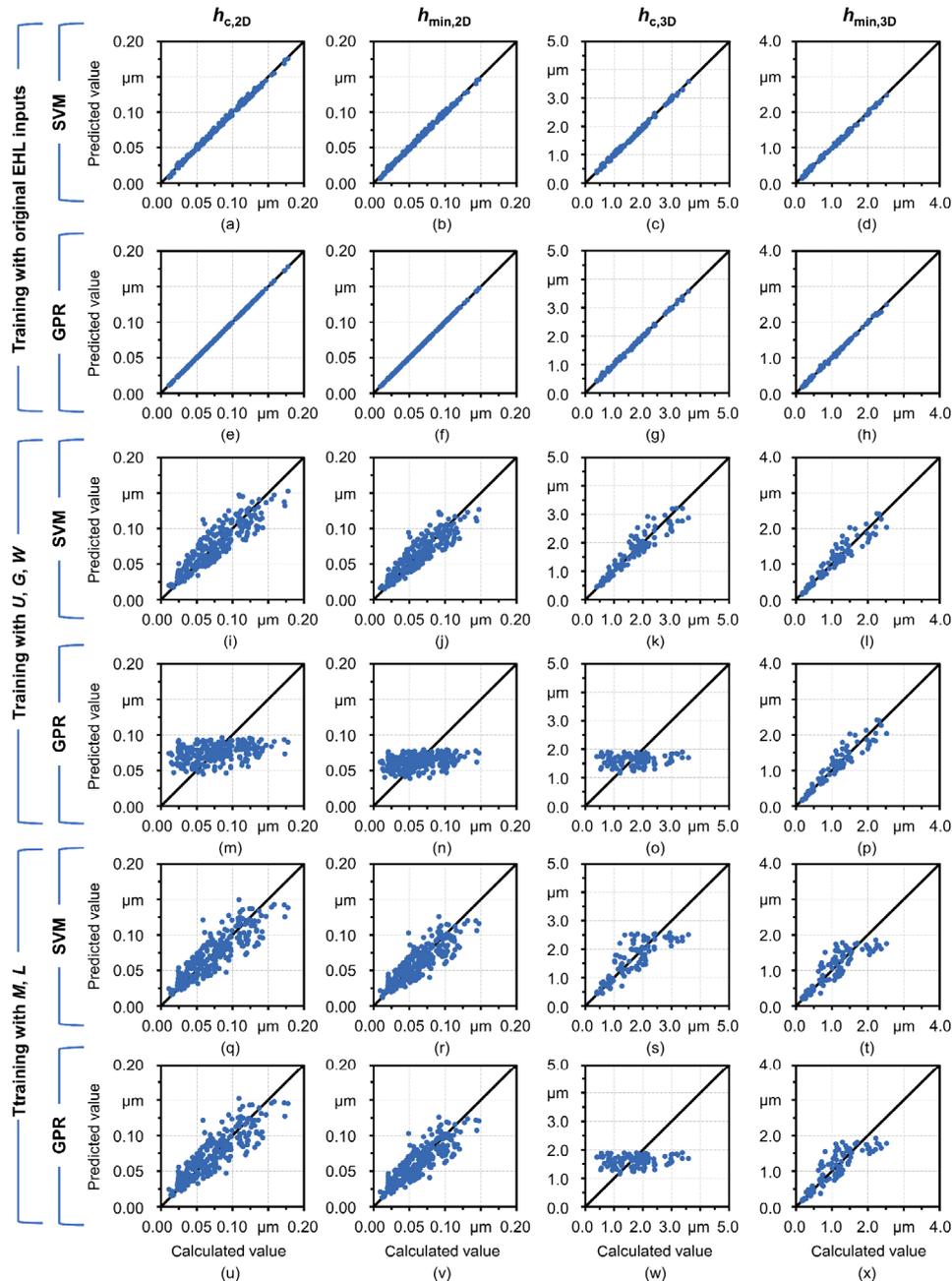


Fig. 7 Predicted versus calculated values (testing data) of the central and minimum film thicknesses for the infinite 2D line and 3D circular point contacts using (a–d, i–l, q–t) SVM and (e–h, m–p, u–x) GPR trained with (a–h) the original EHL inputs, (i–p) the non-dimensional parameters U, G, W , or (q–x) M, L .

When using the non-dimensional groups U , G , W , or M , L instead of the original input data for training purposes, the values were predicted with significantly lower accuracy (Figs. 7(j)–7(l), 7(p)–7(v), and 7(x)) or even totally insufficiently (Figs. 7(m)–7(o) and 7(w)). The R^2 values fell well below 0.9 and almost zero for some cases (Table 5). The substantially inferior prediction quality can be traced back to redundant input variables for different output variables. Since similar dimensionless parameters can result from different original input data, it is possible that similar dimensionless variables can lead to different film parameters (Table 6). These inconsistencies in the data base led to problems in the training of SVM and, particularly, GPR (Table 5). A possibility to overcome this issue is the use of the dimensionless film thickness H as output of the regression equations for training the ML approaches without prior dimensionalization.

Table 6 Exemplary points from the data base for the infinite 2D line contact with inconsistency in the film thickness when using the non-dimensional parameters G , U , W , or M , L instead of the original input data.

	Input	Unit	#110	#1073	#1338
EHL contact parameters	E_1	GPa	236	298	307
	E_2	GPa	204	217	206
	ν_1	—	0.33	0.31	0.34
	ν_2	—	0.34	0.35	0.30
	R	mm	8.1	11.5	18.7
	l	mm	8.1	5.8	3.1
	F_N	N	1,133	1,145	778
	$u_{1,2}$	m/s	0.36	0.16	0.16
	$u_{1,2}$	m/s	0.26	0.36	0.31
	ρ_0	Pa·s	897	870	1,069
	η_0	g/cm ³	0.03	0.04	0.05
	α_p	Pa ⁻¹	1.5×10^{-8}	1.4×10^{-8}	1.7×10^{-8}
Non-dimensional parameters	G	—	3,695	3,980	4,585
	U	—	4.7×10^{-12}	3.5×10^{-12}	2.2×10^{-12}
	W	—	7.0×10^{-5}	6.1×10^{-5}	4.9×10^{-5}
Moes parameters	M	—	23.0	23.0	23.0
	L	—	6.5	6.5	6.7
Film thickness	h_c	μm	0.080	0.097	0.131
	h_{\min}	μm	0.067	0.082	0.110

The unambiguous prediction of the film thickness becomes possible by using only dimensionless parameters as input of the ANN and the dimensionless film thickness H as output since the latter is unique. Subsequently, the film thickness can be dimensionalized in the post-processing using the radius R .

The coefficients of determination for the prediction of the central and minimum film thicknesses for the infinite 2D line and the 3D circular point contacts against testing data using ANNs with 10, 12, 15, or 20 neurons in one, two, or three hidden layers are summarized in Tables 7 and 8. Thereby, the original input values have been used for training. It can be observed that the EHL film parameters were predicted with high accuracy (R^2 values above 0.99) by some ANN configurations. While the number of neurons played a less dominant role, the number of hidden layers had a decisive impact on the prediction quality. Therefore, the lowest R^2 values resulted in only one hidden layer, followed by three and finally two hidden layers with the highest accuracy. Regarding the number of neurons, the configurations with more than 10 neurons proved to be favorable. Since 12 neurons in 2 hidden layers gave the best results in the more complex 3D circular point contact (Table 8), this configuration was kept constant for the following predictions. The comparison of the calculated and predicted central and minimum film thicknesses for this configuration is depicted in Fig. 8. The prediction was more precise than that of the analytically solvable proximity equations (Section 3.2.1) and comparable to SVM or GPR (Table 5). The 3D circular point contact case featured higher R^2 values against the testing data compared to the 2D line contact, which may be due to the considerably larger data base for the latter and thus some overfitting.

Similar to SVM and GPR, training the ANN with dimensionless groups instead of the original input parameters resulted in a considerable reduction of prediction quality (Table 9) due to inconsistencies in the training data set (Table 6).

The coefficients of determination of the prediction with an ANN with 12 neurons in each of two hidden layers after training with hyperbolic tangent, logistic

Table 7 Coefficients of determination for the prediction of the central and minimum film thicknesses for the infinite 2D line contact against testing data using ANNs with variable numbers of hidden layers and neurons per hidden layer.

R^2		Number of neurons per hidden layer							
		10		12		15		20	
		h_c	h_{min}	h_c	h_{min}	h_c	h_{min}	h_c	h_{min}
Number of hidden layers	1	0.987	0.987	0.979	0.978	0.978	0.9779	0.981	0.983
	2	0.690	0.690	0.988	0.988	1.000	1.000	0.999	0.999
	3	0.977	0.978	0.999	0.999	0.927	0.930	0.986	0.985

Table 8 Coefficients of determination for the prediction of the central and minimum film thicknesses for the 3D circular point contact against testing data using ANNs with variable numbers of hidden layers and neurons per hidden layer.

R^2		Number of neurons per hidden layer							
		10		12		15		20	
		h_c	h_{min}	h_c	h_{min}	h_c	h_{min}	h_c	h_{min}
Number of hidden layers	1	0.998	0.996	0.984	0.9828	0.992	0.989	0.995	0.995
	2	0.994	0.995	1.000	1.000	0.993	0.991	0.999	0.999
	3	0.978	0.971	1.000	1.000	0.978	0.970	0.999	0.999

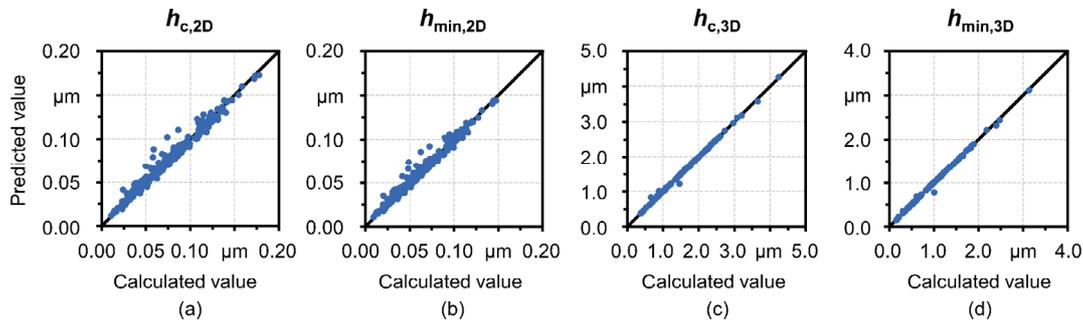


Fig. 8 Predicted versus calculated values (testing data) of the central and minimum film thicknesses for the (a, b) infinite 2D line and (c, d) 3D circular point contacts using an ANN with 12 neurons in each of the two hidden layers.

Table 9 Coefficients of determination of the prediction for the central and minimum film thicknesses for the infinite 2D line and 3D circular point contacts against testing data using an ANN with 12 neurons in each of two hidden layers after training with the original EHL inputs, the non-dimensional parameters U , G , W , or M , L .

R^2	EHL 2D line contact		EHL 3D point contact	
	h_c	h_{min}	h_c	h_{min}
Training with				
Original EHL inputs	0.988	0.988	1.000	1.000
U, G, W	0.750	0.757	0.834	0.853
M, L	0.724	0.695	0.598	0.661

Table 10 Coefficients of determination of the prediction for the central and minimum film thicknesses for the infinite 2D line and 3D point contacts using an ANN with 12 neurons in each of two hidden layers after training with hyperbolic tangent, sigmoid and identity activation functions.

R^2	EHL 2D line contact		EHL 3D point contact	
	h_c	h_{min}	h_c	h_{min}
Activation function				
tanh	0.988	0.988	1.000	1.000
Sigmoid	1.000	1.000	0.995	0.999
Identity	0.961	0.96	0.932	0.933

and identity activation functions are compared in Table 10. Please note that the output and propagation functions were kept constant and only the activation function was varied. The sigmoid activation function

and the previously employed hyperbolic tangent function led to high R^2 values beyond 0.99, while the linear identity function resulted in a slightly downgraded prediction quality.

To minimize the efforts required to generate the training data base, it is advisable to keep the number of training data as small as possible without affecting the prediction quality. The prediction coefficients of the central and minimum film thicknesses for the infinite 2D line contact with a varying number of training data are shown in Table 11. Even when the original training data set was reduced to 600 data points, the R^2 values remained at a very high level above 0.99, thus showing slightly better R^2 values compared to the 1,249 training data sets. This may imply some overfitting for the original 1,249 training data set. For 300 or 100 training data points, the coefficients of determination showed a decreasing tendency but stayed above 0.9. The deteriorated prediction accuracy due to an insufficient data base can also be seen from the comparison between the calculated and predicted values in Fig. 9. It can be concluded that for the present case with 12 varying input variables, at least 600 data points are required for the training. This also fits the data consisting of 600 sets used for the 3D circular point contact at the same number of input variables and explains its sufficient prediction quality.

Table 11 Coefficients of determination of the prediction for the central and minimum film thicknesses for the infinite 2D line contact using an ANN with 12 neurons in each of two hidden layers after training with differently sized data sets.

Number of training data sets	h_c	h_{min}
1,249	0.988	0.988
600	1.000	1.000
300	0.961	0.963
100	0.922	0.927

3.2.2.2 Local EHL film distribution

For the prediction of the local lubricant film distribution for the infinite 2D line contact by an ANN with 12 neurons in 2 hidden layers, an overall coefficient of determination of 1.000 against testing data including all data sets and positions along the contact length was obtained. Thus, the model exhibited excellent prognosis accuracy, which can be recognized in the direct comparison of the lubricant film distribution calculated by FEM and the values predicted by ANN as depicted in Fig. 10. While the two curves were congruent over almost the entire contact length, including the minimum film thickness, only minor deviations were predicted close to the upstream of the film constriction at the contact outlet (slight overshooting of the ANN prediction as enlarged detail in Fig. 10). Regarding computational time, training and execution tasks must be differentiated. The training of the ANN with the locally resolved data required several minutes. To finally compare the execution time of the FEM simulation and the ANN, the lubricant film thickness for a data set not included in the training data was calculated using both methods on the same computer. Although the ANN was trained with 86 uniformly distributed data points along the contact length, it is possible to query any desired discretization. Since the EHL simulation provided the lubricant film thickness at 8,641 positions in the relevant contact length range, the same number of values was predicted with the ML algorithm for consistency. As illustrated in Table 12, ANN was over 25 times faster than the FEM model, while the calculation time is comparable to the analytically solvable proximity

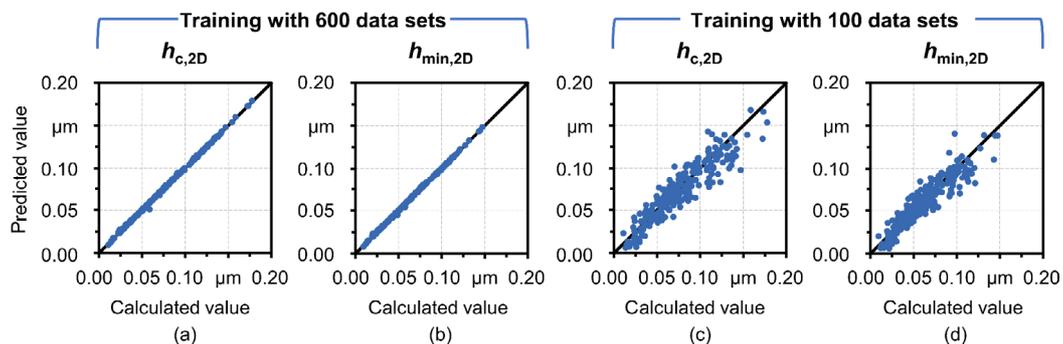


Fig. 9 Predicted versus calculated values of the central and minimum film thicknesses for the infinite 2D line contact using an ANN with 12 neurons in each of the two hidden layers after training with (a, b) 600 and (c, d) 100 data sets.

Table 12 Computational time for determining the lubricant film distribution over the contact length for a representative infinite 2D line contact using the FEM-based EHL simulation and an ANN with 15 neurons in each of the two hidden layers.

	FEM-based EHL simulation	ANN prediction
Computational time	19 s	0.74 s

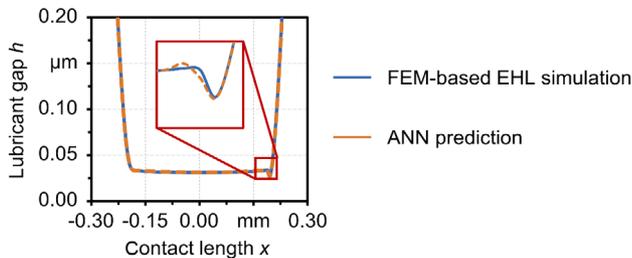


Fig. 10 Comparison of calculated (blue) and predicted (orange) lubricant film distribution over the contact length for a representative infinite 2D line contact using FEM and an ANN with 12 neurons in each of the two hidden layers.

equations solution. However, ANN offers the possibility to predict the lubricant film distribution locally. Moreover, the computational time could even be further improved when implementing the ANN in programming languages with a stronger focus on computational speed. By incorporating the ANN into multibody system simulations, their accuracy in friction and dynamics analyses or the identification of critical operating conditions could become even more precise without negatively affecting the required computational time.

4 Conclusions

This contribution demonstrated that ML approaches could predict relevant film parameters more efficiently than sophisticated EHL simulation models. Moreover, we verified that ML outperforms analytically solvable proximity equations based upon non-dimensional groups in terms of accuracy and flexibility. Based upon the presented results, we derive the following conclusions to address the introduced research questions:

1) Compared to the proximity equations according to Dowson/Toyoda/Higginson [5, 6, 14] for infinite 2D line contacts and Hamrock/Dowson [16] for 3D circular point contacts, all ML approaches (SVM,

GPR, and ANNs) reproduced the behavior calculated by the FEM-based EHL simulations more accurately. Thereby, coefficients of determinations up to 1 were verified when predicting EHL film parameters within the operating conditions covered by the training data. Other film thickness equations, e.g., from Nijenbanning et al. [19], Wolf et al. [21], or Moes [22], may also enable accurate predictions. However, the implementation of these equations requires the definition of more complicated analytical expressions to fit the simulation results, which, in turn, increases the computational effort and the proneness to errors. In this regard, the omission of a large range of dimensionless parameters represents another benefit of the ML approaches employed to efficiently predict EHL film thickness without the need to define complex proximity equations. In contrast to proximity equations, the usage of original EHL input parameters is crucial for ML approaches. Using only the parameters G , U , W , or M , L notably decreases the prediction quality. We anticipate that this aspect can be overcome by training ML approaches with dimensionless film thickness as output instead of the dimensional one.

2) The architecture of ANNs and, especially, the number of hidden layers notably influences the accuracy. For the data set of this contribution, ANNs with 12 neurons in two hidden layers were most favorable. Suitable activation functions are hyperbolic tangent and sigmoid functions, while linear identity functions deteriorated the prediction accuracy.

3) The size of the data base affects the prediction quality. For ML approaches trained using 12 input parameters, roughly 600 training data proved sufficient to reach high R^2 values without overfitting.

4) It should be noted that the trained ML approaches were not physics-informed and, therefore, cannot provide physical explanations. Although not reflected in the results, the predictions have to be considered carefully as it is possible to obtain negative film thicknesses. Nevertheless, the trained ML approaches provided meaningful results when employed for the operating conditions covered in the training data base of this article. In the future, the architecture of ANNs can be further modified to exclude unphysical results by using a strictly positive linear function (e.g., the poslin function in MATLAB) instead of linear

identity in combination with a normalization of the output data in the range of $[0, 1]$. This would help to ensure that the ANN will not predict negative film thicknesses for any input parameters.

5) In addition to global parameters, ML approaches can also predict the local lubricant film distribution. Compared to FEM-based EHL simulations, the trained ANNs achieved 25 times shorter computation times at prediction coefficients above 0.999. At present, lubrication behavior and dynamics calculations in multibody system simulations typically employ analytically solvable proximity equations and the Hertzian theory to account for fluid film formation and pressure distribution. We hypothesize that in the future, the incorporation of ML approaches such as ANNs into these higher-level simulations will add value to the engineering system design process.

6) In the future, other contact conditions and physical effects should be considered by training novel ML approaches without the need to define complex proximity equations or introduce further correction factors. Thereby, elliptical contacts, thermal effects and other rheological models should be considered. It is also conceivable that time-dependent squeeze, starvation effects, and the influences of surface roughness and solid asperity contact may be also taken into account.

Conflict of interest

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

Acknowledgements

M. Marian greatly acknowledges the support from Pontificia Universidad Católica de Chile. A. Rosenkranz gratefully acknowledges the financial support given by ANID (Chile) in the framework of the Fondecyt projects (Nos. 11180121 and EQM190057). Additionally, A. Rosenkranz acknowledges the financial support given by the VID of the University of Chile within the project U-Moderniza (No. UM-04/19).

Author contributions

M. Marian, F. J. Profito, and A. Rosenkranz conceived the idea. J. Mursak carried out numerical modeling and analyzed the data. M. Marian supervised the investigations and wrote the initial draft. F. J. Profito and A. Rosenkranz provided suggestions for the discussion. All authors have reviewed, edited, and approved the final version of the manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made.

The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- [1] Holmberg K, Erdemir A. Influence of tribology on global energy consumption, costs and emissions. *Friction* **5**(3): 263–284 (2017)
- [2] Holmberg K, Andersson P, Erdemir A. Global energy consumption due to friction in passenger cars. *Tribol Int* **47**: 221–234 (2012)
- [3] Holmberg K, Andersson P, Nylund N O, Mäkelä K, Erdemir A. Global energy consumption due to friction in trucks and buses. *Tribol Int* **78**: 94–114 (2014)
- [4] Marian M, Bartz M, Wartzack S, Rosenkranz A. Non-dimensional groups, film thickness equations and correction factors for elastohydrodynamic lubrication: A review. *Lubricants* **8**(10): 95 (2020)
- [5] Dowson D, Higginson G R. The effect of material properties on the lubrication of elastic rollers. *J Mech Eng Sci* **2**(3): 188–194 (1960)

- [6] Dowson D, Higginson G R, Whitaker A V. Elastohydrodynamic lubrication: A survey of isothermal solutions. *J Mech Eng Sci* **4**(2): 121–126 (1962)
- [7] Moes H. Discussion on Paper D1 by D. Dowson. *Proc Instn Mech Engrs* **180**: 244–245 (1966)
- [8] Moes H. Communications. In *Proc. of the Symposium on Elastohydrodynamic Lubrication*, 1965: 244–245.
- [9] Johnson K L. Regimes of elastohydrodynamic lubrication. *J Mech Eng Sci* **12**(1): 9–16 (1970)
- [10] Habchi W, Bair S, Vergne P. On friction regimes in quantitative elastohydrodynamics. *Tribol Int* **58**: 107–117 (2013)
- [11] van Leeuwen H. The determination of the pressure–viscosity coefficient of a lubricant through an accurate film thickness formula and accurate film thickness measurements. *Proc Inst Mech Eng Part J J Eng Tribol* **223**(8): 1143–1163 (2009)
- [12] Dowson D. Elastohydrodynamics. *Proc Inst Mech Eng* **182**: 151–157 (1968)
- [13] Jacobson B O, Hamrock B J. Non-Newtonian fluid model incorporated into elastohydrodynamic lubrication of rectangular contacts. *J Tribol* **106**(2): 275–282 (1984)
- [14] Dowson D, Toyoda S. A central film thickness formula for elastohydrodynamic line contacts. In *Proceedings of the 5th Leeds-Lyon Symposium on Tribology*, 1978: 60–65.
- [15] Moes H. *Lubrication and Beyond—University of Twente Lecture Notes code 115531*. Enschede, Netherlands: University of Twente, 2000.
- [16] Hamrock B J, Dowson D. Isothermal elastohydrodynamic lubrication of point contacts. Part III— Fully flooded results. *J Lubr Technol* **99**(2):264–275. (1977)
- [17] Chittenden R J, Dowson D, Dunn J F, Taylor C M. A theoretical analysis of the isothermal elastohydrodynamic lubrication of concentrated contacts. I. Direction of lubricant entrainment coincident with the major axis of the Hertzian contact ellipse. *Proc R Soc Lond A* **397**(1813): 245–269 (1985)
- [18] Evans H P, Snidle R W. The isothermal elastohydrodynamic lubrication of spheres. *J Lubr Technol* **103**(4): 547–557 (1981)
- [19] Nijebanning G, Venner C H, Moes H. Film thickness in elastohydrodynamically lubricated elliptic contacts. *Wear* **176**(2): 217–229 (1994)
- [20] Sperka P, Krupka I, Hartl M. Analytical formula for the ratio of central to minimum film thickness in a circular EHL contact. *Lubricants* **6**(3): 80 (2018)
- [21] Wolf M, Solovveyev S, Arshia F. Film thickness in elastohydrodynamically lubricated slender elliptic contacts: Part I—Numerical studies of central film thickness. *Proc Inst Mech Eng Part J J Eng Tribol* **236**(6): 1043–1055 (2022)
- [22] Moes H. Optimum similarity analysis with applications to elastohydrodynamic lubrication. *Wear* **159**(1): 57–66 (1992)
- [23] Greenwood J A, Kauzlarich J J. Inlet shear heating in elastohydrodynamic lubrication. *J Lubr Technol* **95**(4): 417–423 (1973)
- [24] Murch L E, Wilson W R D. A thermal elastohydrodynamic inlet zone analysis. *J Lubr Technol* **97**(2): 212–216 (1975)
- [25] Jackson A. A simple method for determining thermal EHL correction factors for rolling element bearings and gears. *S L E Trans* **24**(2): 159–163 (1981)
- [26] Wilson W R D, Sheu S. Effect of inlet shear heating due to sliding on elastohydrodynamic film thickness. *J Lubr Technol* **105**(2): 187–188 (1983)
- [27] Pandey R K, Ghosh M K. Thermal effects on film thickness and traction in rolling/sliding EHL line contacts—An accurate inlet zone analysis. *Wear* **192**(1–2): 118–127 (1996)
- [28] Hamrock B J, Dowson D. Isothermal elastohydrodynamic lubrication of point contacts: Part IV—starvation results. *J Lubr Technol* **99**(1): 15–23 (1977)
- [29] Wedeven L D, Evans D, Cameron A. Optical analysis of ball bearing starvation. *J Lubr Technol* **93**(3): 349–361 (1971)
- [30] Wiśniewski M. Einfluß eines begrenzten Ölangebotes auf die elastohydrodynamische Schmierung von Zahnradern. *Tribologie und Schmierungstechnik* **30**: 270–277 (1983)
- [31] Habchi W, Bair S. Quantitative compressibility effects in thermal elastohydrodynamic circular contacts. *J Tribol* **135**(1): 011502 (2013)
- [32] Venner C H, Bos J. Effects of lubricant compressibility on the film thickness in EHL line and circular contacts. *Wear* **173**(1–2): 151–165 (1994)
- [33] Bair S. Shear thinning correction for rolling/sliding elastohydrodynamic film thickness. *Proc Inst Mech Eng Part J J Eng Tribol* **219**(1): 69–74 (2005)
- [34] Jang J Y, Khonsari M M, Bair S. Correction factor formula to predict the central and minimum film thickness for shear-thinning fluids in EHL. *J Tribol* **130**(2): 024501 (2008)
- [35] Habchi W, Bair S, Qureshi F, Covitch M. A film thickness correction formula for double-Newtonian shear-thinning in rolling EHL circular contacts. *Tribol Lett* **50**(1): 59–66 (2013)
- [36] Kumar P, Jain S C, Ray S. Study of surface roughness effects in elastohydrodynamic lubrication of rolling line contacts using a deterministic model. *Tribol Int* **34**(10): 713–722 (2001)

- [37] Masjedi M, Khonsari M M. On the effect of surface roughness in point-contact EHL: Formulas for film thickness and asperity load. *Tribol Int* **82**: 228–244 (2015)
- [38] Masjedi M, Khonsari M M. Film thickness and asperity load formulas for line-contact elastohydrodynamic lubrication with provision for surface roughness. *J Tribol* **134**(1): 11503 (2012)
- [39] Rosenkranz A, Marian M, Profito F J, Aragon N, Shah R. The use of artificial intelligence in tribology—A perspective. *Lubricants* **9**(1): 2 (2020)
- [40] Marian M, Tremmel S. Current trends and applications of machine learning in tribology—A review. *Lubricants* **9**(9): 86 (2021)
- [41] Bell J. *Machine Learning: Hands-On for Developers and Technical Professionals*. Hoboken: Wiley, 2014.
- [42] Müller AC, Guido S. *Introduction to machine learning with Python: A guide for data scientist*. Beijing: O'Reilly, 2016.
- [43] Schölkopf B, Smola A J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, Mass.: MIT Press, 2002.
- [44] Hasan M S, Kordijazi A, Rohatgi P K, Nosonovsky M. Triboinformatic modeling of dry friction and wear of aluminum base alloys using machine learning algorithms. *Tribol Int* **161**: 107065 (2021)
- [45] Hasan M S, Kordijazi A, Rohatgi P K, Nosonovsky M. Triboinformatics approach for friction and wear prediction of Al-graphite composites using machine learning methods. *J Tribol* **144**(1): 011701 (2022)
- [46] Yin Y, Liu X F, Huang W F, Liu Y, Hu S T. Gas face seal status estimation based on acoustic emission monitoring and support vector machine regression. *Adv Mech Eng* **12**(5): 168781402092132 (2020)
- [47] Timur M, Ayding F. Anticipating the friction coefficient of friction materials used in automobiles by means of machine learning without using a test instrument. *Turk J Elec Eng & Comp Sci*: 1440–1454 (2013)
- [48] Das B, Pal S, Bag S. Torque based defect detection and weld quality modelling in friction stir welding process. *J Manuf Process* **27**: 8–17 (2017)
- [49] Egala R, Jagadeesh G V, Setti S G. Experimental investigation and prediction of tribological behavior of unidirectional short castor oil fiber reinforced epoxy composites. *Friction* **9**(2): 250–272 (2021)
- [50] Vinoth A, Datta S. Design of the ultrahigh molecular weight polyethylene composites with multiple nanoparticles: An artificial intelligence approach. *J Compos Mater* **54**(2): 179–192 (2020)
- [51] Gangwar S, Pathak V K. Dry sliding wear characteristics evaluation and prediction of vacuum casted marble dust (MD) reinforced ZA-27 alloy composites using hybrid improved bat algorithm and ANN. *Mater Today Commun* **25**: 101615 (2020)
- [52] Subrahmanyam M, Sujatha C. Using neural networks for the diagnosis of localized defects in ball bearings. *Tribol Int* **30**(10): 739–752 (1997)
- [53] Canbulut F, Yildirim Ş, Sinanoğlu C. Design of an artificial neural network for analysis of frictional power loss of hydrostatic slipper bearings. *Tribol Lett* **17**(4): 887–899 (2004)
- [54] Senatore A, D'Agostino V, Giuda R D, Petrone V. Experimental investigation and neural network prediction of brakes and clutch material frictional behaviour considering the sliding acceleration influence. *Tribol Int* **44**(10): 1199–1207 (2011)
- [55] Aleksendrić D, Barton D C. Neural network prediction of disc brake performance. *Tribol Int* **42**(7): 1074–1080 (2009)
- [56] König F, Sous C, Chaib A O, Jacobs G. Machine learning based anomaly detection and classification of acoustic emission events for wear monitoring in sliding bearing systems. *Tribol Int* **155**: 106811 (2021)
- [57] Dewan M W, Huggett D J, Liao T W, Wahab M A, Okeil A M. Prediction of tensile strength of friction stir weld joints with adaptive neuro-fuzzy inference system (ANFIS) and neural network. *Mater Des* **92**: 288–299 (2016)
- [58] Sahraoui T, Guessasma S, Fenineche N E, Montavon G, Coddet C. Friction and wear behaviour prediction of HVOF coatings and electroplated hard chromium using neural computation. *Mater Lett* **58**(5): 654–660 (2004)
- [59] Cetinel H. The artificial neural network based prediction of friction properties of Al₂O₃–TiO₂ coatings. *Ind Lubr Tribol* **64**(5): 288–293 (2012)
- [60] Kalliorinne K, Larsson R, Pérez-Rafols F, Liwicki M, Almqvist A. Artificial neural network architecture for prediction of contact mechanical response. *Front Mech Eng* **6**: 579825 (2021)
- [61] Humelnicu C, Ciortan S, Amortila V. Artificial neural network-based analysis of the tribological behavior of vegetable oil–diesel fuel mixtures. *Lubricants* **7**(4): 32 (2019)
- [62] Bhaumik S, Pathak S D, Dey S, Datta S. Artificial intelligence based design of multiple friction modifiers dispersed castor oil and evaluating its tribological properties. *Tribol Int* **140**: 105813 (2019)
- [63] Krogh A. What are artificial neural networks? *Nat Biotechnol* **26**(2): 195–197 (2008)
- [64] Almqvist A. Fundamentals of physics-informed neural networks applied to solve the Reynolds boundary value problem. *Lubricants* **9**(8): 82 (2021)

- [65] Kügler P, Marian M, Schleich B, Tremmel S, Wartzack S. tribAIIn—Towards an explicit specification of shared tribological understanding. *Appl Sci* **10**(13): 4421 (2020)
- [66] de la Guerra Ochoa E, Echávarri Otero J, Chacón Tanarro E, Lafont Morgado P, Lantada A D, Muñoz-Guijosa J M, Sanz J M. Optimising lubricated friction coefficient by surface texturing. *Proc Inst Mech Eng C J Mech Eng Sci* **227**(11): 2610–2619 (2013)
- [67] Marian M, Grützmaier P, Rosenkranz A, Tremmel S, Mücklich F, Wartzack S. Designing surface textures for EHL point-contacts—Transient 3D simulations, meta-modeling and experimental validation. *Tribol Int* **137**: 152–163 (2019)
- [68] Wirsching S, Marian M, Bartz M, Stahl T, Wartzack S. Geometrical optimization of the EHL roller face/rib contact for energy efficiency in tapered roller bearings. *Lubricants* **9**(7): 67 (2021)
- [69] Siebertz K, van Bebbber D, Hochkirchen T. *Statistische Versuchsplanung: Design of Experiments (DoE)*. Berlin: Springer, 2010.
- [70] Gohar R. *Elastohydrodynamics*. Chichester: Halsted Press, 1988.
- [71] Johnson M E, Moore L M, Ylvisaker D. Minimax and maximin distance designs. *J Stat Plan Inference* **26**(2): 131–148 (1990)
- [72] Reynolds O. On the theory of lubrication and its application to Mr. Beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil. *Phil Trans R Soc* **177**: 157–234 (1886)
- [73] Dowson D, Higginson GR. *Elasto-Hydrodynamic Lubrication: The Fundamentals of Roller and Gear Lubrication*. Oxford: Pergamon Press, 1966.
- [74] Roelands C. Correlational aspects of the viscosity-temperature-pressure relationship of lubricating oils. Ph.D. Thesis. Delft University of Technology, 1966.
- [75] Marian M, Weschta M, Tremmel S, Wartzack S. Simulation of microtextured surfaces in starved EHL contacts using commercial FE software. *Mats Perf Charact* **6**(2): MPC20160010 (2017)
- [76] Habchi W, Eyheramendy D, Vergne P, Morales-Espejel G. A full-system approach of the elastohydrodynamic line/point contact problem. *J Tribol* **130**(2): 021501/1-9 (2008)
- [77] Habchi W. *Finite Element Modelling of Elastohydrodynamic Lubrication Problems*. Chichester, UK: John Wiley & Sons Ltd, 2018
- [78] Hughes T J R, Franca L P, Hulbert G M. A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective-diffusive equations. *Comput Methods Appl Mech Eng* **73**(2): 173–189 (1989)
- [79] Zienkiewicz OC, Taylor RL, Nithiarasu P. *The Finite Element Method for Fluid Dynamics*. 7th edn. Oxford: Elsevier Butterworth-Heinemann, 2014.
- [80] Lohner T, Ziegltrum A, Stemplinger J P, Stahl K. Engineering software solution for thermal elastohydrodynamic lubrication using multiphysics software. *Adv Tribol* **2016**: 6507203 (2016)
- [81] Tan X C, Goodyer C E, Jimack P K, Taylor R I, Walkley M A. Computational approaches for modelling elastohydrodynamic lubrication using multiphysics software. *Proc Inst Mech Eng Part J J Eng Tribol* **226**(6): 463–480 (2012)
- [82] Mathworks Matlab. Understanding support vector machine regression: Mathematical formulation of SVM regression. Available from: <https://de.mathworks.com/help/stats/understanding-support-vector-machine-regression.html>, October 08, 2021
- [83] Chicco D, Warrens M J, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput Sci* **7**: e623 (2021)
- [84] Glantz SA, Slinker BK. *Primer of Applied Regression and Analysis of Variance*. New York, NY: McGraw-Hill, 1990.
- [85] Montgomery D C, Runger G C. *Applied Statistics and Probability for Engineers 6th Edition*. Wiley, 2014.
- [86] Vapnik V N. *The Nature of Statistical Learning Theory*. New York: Springer, 2000.
- [87] Huang T-M, Kecman V, Kopriva I. *Kernel Based Algorithms for Mining Huge Data Sets: Supervised, Semi-supervised, and Unsupervised Learning*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2006.
- [88] Mathworks Matlab. Gaussian process regression models. Available from: <https://de.mathworks.com/help/stats/gaussian-process-regression-models.html>, October 08, 2021
- [89] Levenberg K. A method for the solution of certain non-linear problems in least squares. *Quart Appl Math* **2**(2): 164–168 (1944)
- [90] Marquardt D W. An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math* **11**(2): 431–441 (1963)
- [91] Wilamowski B M, Yu H. Improved computation for levenberg–marquardt training. *IEEE Trans Neural Netw* **21**(6): 930–937 (2010)
- [92] Lämmel U, Cleve J. *Künstliche Intelligenz: Wissensverarbeitung - neuronale Netze*. 5th edn. Munich, Germany: Hanser, 2020.
- [93] Bhattacharyya S. *Deep Learning. Research and Applications*. 1st edn. Boston, United States: De Gruyter, 2020.



Max MARIAN. He is an assistant professor for multiscale engineering mechanics at the Department of Mechanical and Metallurgical Engineering of Pontificia Universidad Católica de Chile. His research focuses on energy efficiency and sustainability through tribology,

with an emphasis on the modification of surfaces through micro-texturing and coatings (diamond-like carbon, molybdenum disulfide, MXenes). Besides machine elements and engine components, he expanded

his fields towards biotribology and artificial joints as well as triboelectric nanogenerators. His research is particularly related to the development of numerical multiscale tribo-simulation and machine learning approaches. He has published more than 30 peer-reviewed journal publications and is in the Editorial Boards of *Frontiers in Chemistry Nanoscience*, *Industrial Lubrication and Tribology*, *Lubricants* as well as *Tribology—Materials, Surfaces & Interfaces*. Furthermore, he is the member of the Society of Tribologists and Lubrication Engineers (STLE) and the German Society for Tribology (GfT).