

PAPER • OPEN ACCESS

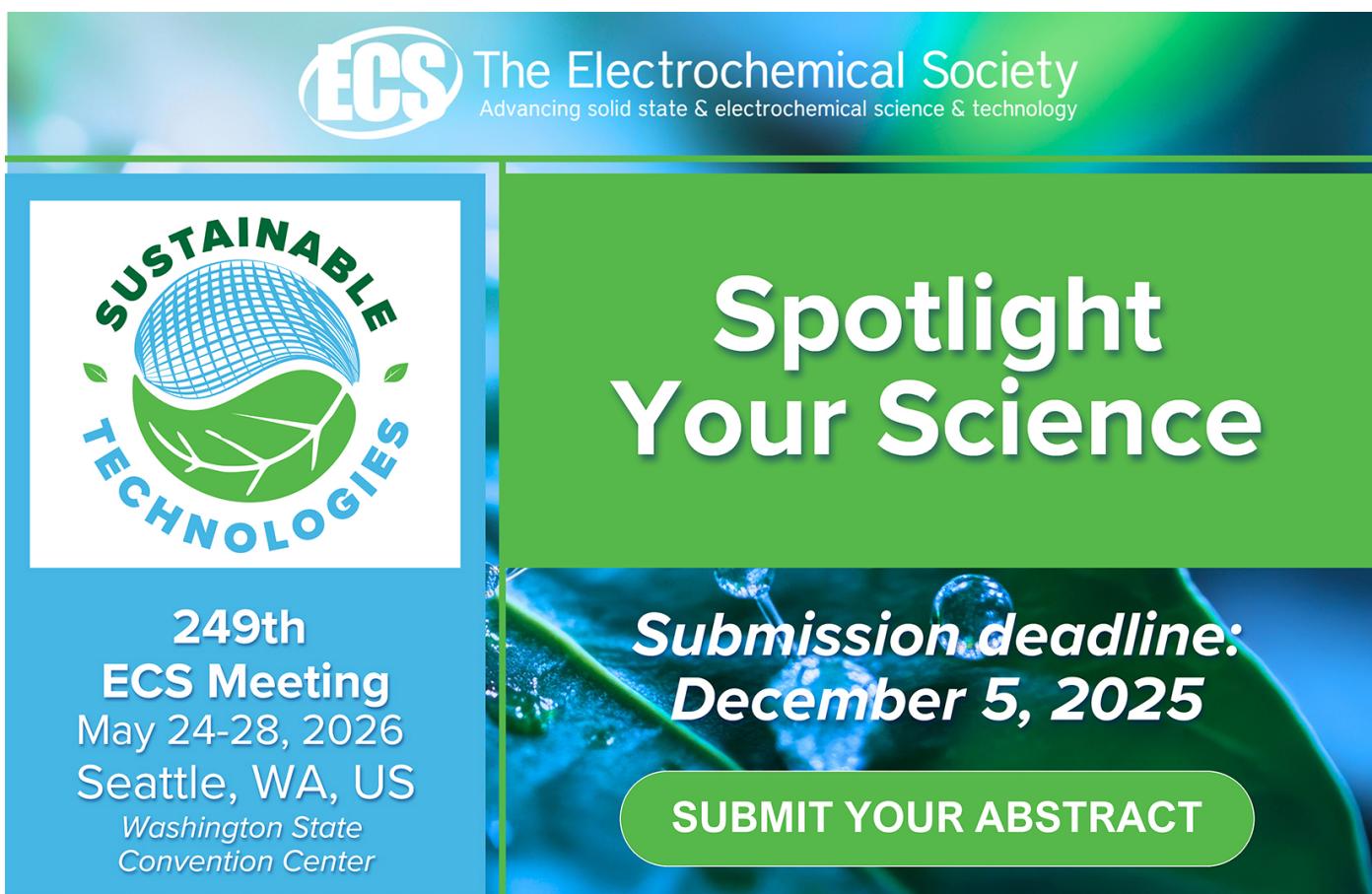
Artificial neural network assisted spectral scatterometry for grating quality control

To cite this article: Aleksi Mattila *et al* 2024 *Meas. Sci. Technol.* **35** 085025

View the [article online](#) for updates and enhancements.

You may also like

- [Virtual experiments for the assessment of data analysis and uncertainty quantification methods in scatterometry](#)
Gertjan Kok, Marcel van Dijk, Gerd Wübbeler et al.
- [Interferometric coherent Fourier scatterometry: a method for obtaining high sensitivity in the optical inverse-grating problem](#)
S Roy, N Kumar, S F Pereira et al.
- [In-line characterization of nanostructured mass-produced polymer components using scatterometry](#)
Jonas Skovlund Madsen, Lasse Højlund Thamdrup, Ilja Czolkos et al.



The advertisement features the ECS logo and the text "The Electrochemical Society Advancing solid state & electrochemical science & technology". On the left, there's a circular logo for "SUSTAINABLE TECHNOLOGIES" with a stylized globe and leaves. The main text on the right reads "Spotlight Your Science" and "Submission deadline: December 5, 2025". A green button at the bottom says "SUBMIT YOUR ABSTRACT".

**249th
ECS Meeting**
May 24-28, 2026
Seattle, WA, US
*Washington State
Convention Center*

Submission deadline:
December 5, 2025

SUBMIT YOUR ABSTRACT

Artificial neural network assisted spectral scatterometry for grating quality control

Aleksi Mattila^{1,*}, Johan Nysten¹, Ville Heikkinen¹, Jorma Kilpi¹, Virpi Korpelainen¹, Poul-Erik Hansen², Petri Karvinen³, Markku Kuittinen³ and Antti Lassila¹

¹ VTT MIKES Metrology, Espoo, Finland

² DFM Danish National Metrology Institute, Hørsholm, Denmark

³ University of Eastern Finland, Joensuu, Finland

E-mail: aleksi.mattila@vtt.fi

Received 19 January 2024, revised 26 April 2024

Accepted for publication 21 May 2024

Published 31 May 2024



Abstract

Spectral scatterometry is a technique that allows rapid measurements of diffraction efficiencies of diffractive optical elements (DOEs). The analysis of such diffraction efficiencies has traditionally been laborious and time consuming. However, machine learning can be employed to aid in the analysis of measured diffraction efficiencies. In this paper we describe a novel system for providing measurements of multiple measurands rapidly and concurrently using a spectral scatterometer and an artificial neural network (ANN) which is trained utilising transfer learning. The ANN provides values for the pitch, height, and line widths of the DOEs. In addition, an uncertainty evaluation was performed. In the majority of the studied cases, the discrepancies between the values obtained using a scanning electron microscope (SEM) and artificial neural network assisted spectral scatterometer (ANNASS) for the grating parameters were below 5 nm. Furthermore, independent reference samples were used to perform a metrological validation. An expanded uncertainty ($k = 2$) of 5.3 nm was obtained from the uncertainty evaluation for the measurand height. The height value measurements performed employing ANNASS and SEM are demonstrated to be in agreement within this uncertainty.

Keywords: scatterometry, transfer learning, inverse modelling, diffraction, gratings

1. Introduction

The photonics industry is steadily improving the performance of innovative projectors, displays, waveguides and augmented reality (AR) headsets. AR relies on diffractive optical element (DOE) components such as wave guide couplers to create a simulated environment that annotates or enhances the real environment so that the user can experience them as

one environment [1]. A fast and reliable quality control for mass manufacturing of DOEs is needed. The measurands of interest in DOEs include realised period, line width, height, side wall angles, and edge roundings of the produced gratings. Additionally, the edge roughness and undercut can be interesting for quality control. Conventional characterisation techniques, such as atomic force microscopy (AFM) and scanning electron microscopy (SEM) are too slow, cumbersome, or otherwise unsuitable for use in industrial mass production environments.

Scatterometry is a non-destructive technique that analyses the scattering of light from periodic structures to extract information about the physical and dimensional properties of the sample. Scatterometers measure the properties of light that has interacted with a diffractive structure. The most

* Author to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

common categories of scatterometers are angular and spectral scatterometers [2].

Angular scatterometers measure the angle and intensity of incoming and/or outgoing light. The angles of the diffraction orders can be used to calculate the period of the grating [3, 4]. Spectral or spectroscopic scatterometers measure the spectral intensity of light that is reflected or transmitted through a sample [5]. It is very fast, and able to withstand vibrations and measure even moving samples [6], but its downside is that the analysis of data can be difficult and laborious as complex inverse calculations or comparisons to library data are required.

Spectral scatterometry is a fast method but does not measure the parameters of the sample directly. Usually, the measured spectrum is compared to a theoretical signature with known grating parameters [2]. This is often done using library search or non-linear regression. These methods have their own issues. The library search requires a large set of data, and non-linear regression can be slow. An alternative to these more traditional methods for solving the inverse problem is to use machine learning [7–9]. It can be used as a more efficient alternative for solving problems that would require complex and time-consuming computation.

An artificial neural network (ANN) was developed and trained to determine the grating parameters from the diffraction efficiencies measured using a spectral scatterometer. Studies where ANNs have been used in the context of scatterometry include [7–15]. These span a time period from 1999 up to 2022, and the increase in computational capacity during the period is clearly visible. The earlier studies had small ANNs with a single hidden layer, [7] has a deeper network structure with three hidden layers, and the training set size has increased from about 1000 to 400 000. In [9] the ANN structures are more complex, with about 20 layers, and their work utilises angle resolved spectroscopy. The use of transfer learning described in this study is a continuation in this research branch.

A scatterometer similar to a system described by Madsen *et al* [16] was constructed to demonstrate the capabilities of an ANN in measuring the grating parameters of DOEs. The ANN training was mostly performed using simulated data, and transfer learning was utilised to fine-tune the model. Finally, a comparison to reference samples and a preliminary uncertainty evaluation was performed for the period, step height, and line width.

2. Samples used for training and validation

A set of 100 diffraction gratings were manufactured for development and validation of the ANN assisted scatterometry. Within the set several parameters were varied to provide large enough range for the ANN. Four wafers of fused silica (SiO_2) were produced. 25 square-shaped DOEs were etched on each wafer, as figure 1 shows. They are arranged in a square pattern with 5 gratings on each axis. The period is varied along one axis, and the line width along the other. The step height is varied between the wafers. The different values of the period are 660 nm, 670 nm, 675 nm, 678 nm, and 680 nm. The widths of

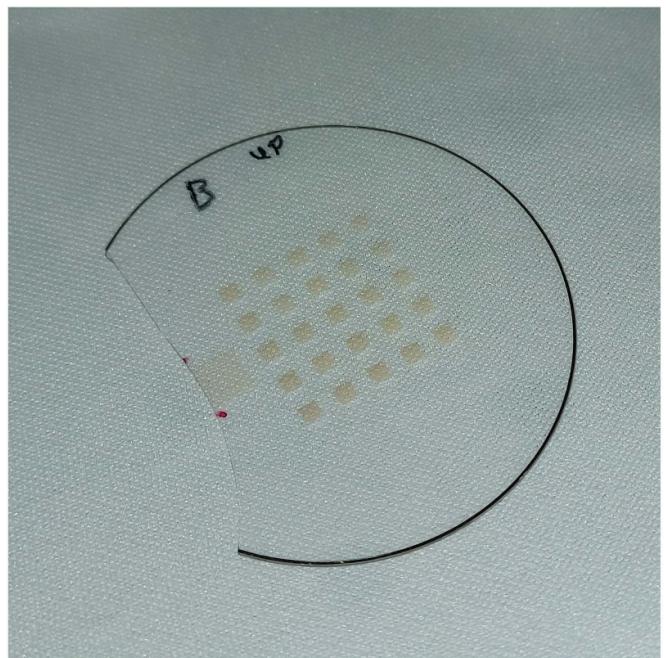


Figure 1. One of the sample wafers.

the grooves are 300 nm, 310 nm, 320 nm, 330 nm, and 340 nm. The step heights are about 220 nm, 280 nm and 300 nm. Size of a single grating is 2 mm by 2 mm. Unfortunately, one of the sample wafers was broken, so only 75 different DOEs were used in this study.

In the manufacturing process of the samples the SiO_2 wafers were first coated with a thin (30 nm) chromium layer, after which AR-P 6200 resist by Allresist GmbH was spin-coated on top. Then the gratings were exposed to electron beam lithography. The resist was then developed using ethyl-3 ethoxypropionate (EEP) and rinsed with isopropyl alcohol and water. Following this, the chromium layer was etched using ICP-RIE process with a chlorine-based recipe. Subsequently the SiO_2 was also etched using the ICP-RIE process (CHF₃-based recipe). Finally, the remaining resist and chromium residues were removed using dry etching. The geometry of the DOEs on the wafers was designed so that it was possible to cut the wafer in two pieces, where the larger piece would contain the DOEs, and the smaller piece contains the same grating geometries for the purpose of SEM measurements of the parameters.

3. Spectral scatterometer

The method utilised in this work was spectral reflectance scatterometry. Different wavelengths of light are diffracted differently from a grating depending on the parameters of the grating. The spectral diffraction efficiency, η_T , in transmission mode can be calculated from the spectral intensities measured by the scatterometer using formula

$$\eta_T = \frac{I_S - I_{BG}}{I_R - I_{BG}}, \quad (1)$$

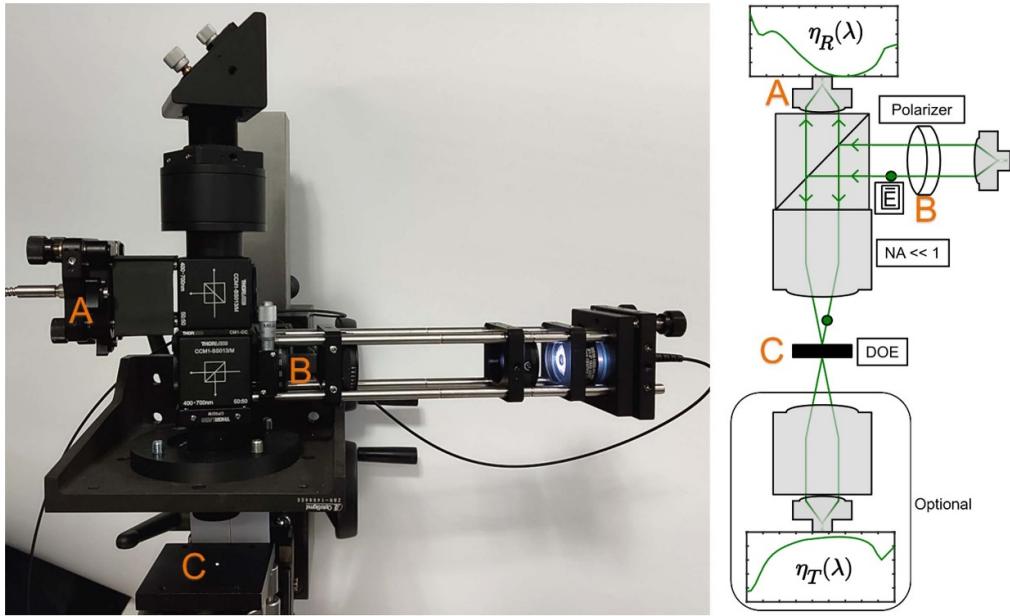


Figure 2. Scatterometer hardware and a diagram of its components. Component marked with A is the fibre-coupled spectrometer, B marks the polariser, and C shows the sample holder. The constructed scatterometer has an additional beam splitter for a viewfinder, so the sample position can be inspected visually.

where I_x are the different spectral intensities. I_S is measured from the grating, I_{BG} is the black background, and I_R is the intensity measured from a blank region of the substrate [5]. In reflectance mode, the reflectance R of the substrate needs to be taken into account, and the equation becomes

$$\eta_R = R \frac{I_S - I_{BG}}{I_R - I_{BG}}. \quad (2)$$

The spectral diffraction efficiencies can then be used to analyse the samples.

Figure 2 depicts the scatterometry hardware and a schematic of the components. The scatterometer uses a Prizmatix FC-LED-BBW50 fibre-coupled white LED light source. Figure 3 illustrates the spectrum of this light source. The chosen optical fibre has a diameter of 1 mm, and length of 1.5 m. The light is collimated using a condenser lens and directed through a pair of apertures to an adjustable polariser. Then the polarised light passes through a beam splitter to an Olympus PLN 4X microscope objective which focuses the beam to the sample, where it is reflected back through the objective into a fibre coupler and a fibre to an Ocean Optics Flame-T-VIS-NIR spectrometer. The optical fibre for the spectrometer is 2 m long and the diameter is 200 μm . The design of the scatterometer allows it to be used for both reflected and transmitted beam measurements. The measurements were performed employing a Python program.

This scatterometer should only capture 0th order diffractions, so the collection angle of the objective should be chosen according to the period of the measured grating and the cutoff wavelength. The constraint on the period d can be approximated using the formula

$$d \lesssim \frac{\lambda_{\min}}{2n \sin(2\theta_{NA})}, \quad (3)$$

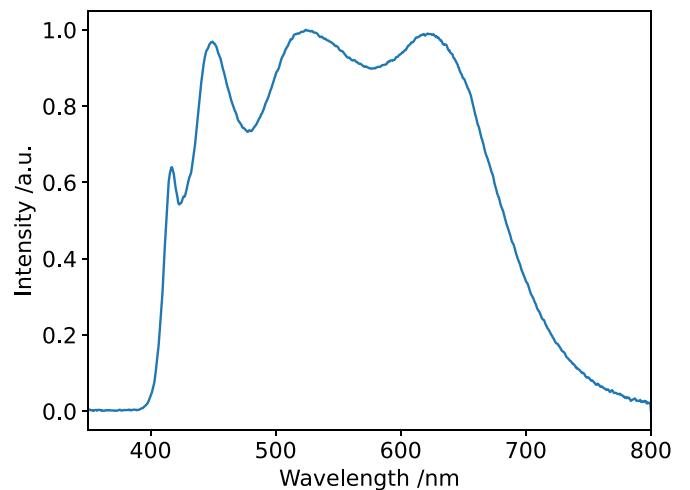


Figure 3. Spectrum of the light source.

where λ_{\min} is the minimum wavelength, n is the refractive index of the incident material (air), θ_{NA} is the collection angle of the objective [16]. Numerical aperture in air is $NA = \sin(\theta_{NA})$. In our system the numerical aperture is 0.1, so the period of a grating should be smaller than about 1000 nm.

The different components of the scatterometer need to be adjusted for optimal results. The position of the collimation lens, alignment of the polariser, alignment of the fibre mounts, focus distance of the sample, and the placement of the sample all have an effect on the measured spectra. The positioning of the sample and the alignment of the fibres were chosen so that the measured total intensity is the highest. The collimation lens was placed according to the properties of the lens. The light was polarised to TE polarisation relative to the sample.

The changes in the measured diffraction efficiencies were investigated by changing the positions and angles of the components. The changes were a few millimetres for the position of the collimation lens, aperture size, and focus distance. For the polarisation angle, sample tilt, and fibre angles the changes were a few degrees. The largest effects were caused by moving the collimation lens, which caused the diffraction efficiency curve to shift up or down in the wavelength range between 500 nm and 550 nm. Changing the alignment of the fibres mostly changes the intensity evenly across the spectrum. Variations in the other components did not create noticeable effects, at least in such small ranges of variations. The of these investigations were used for determining the sensitivity coefficients for the uncertainty evaluation.

Once the adjustments have been set, performing the measurements is very fast. The integration time of the spectrometer is set to 8 ms. Even with such a short integration time, the beam width needs to be restricted using the apertures to prevent the spectrometer from reaching saturation. The sample holder alignment is the only thing that needs to be adjusted during measurements, and everything else should stay fixed. The transfer learning will be performed using data measured using these adjustments, and if there are changes to the device, the accuracy of the ANN will suffer. Transfer learning takes into account the possible flaws of the device. As long as the errors are not dependent on the samples but only the instrument, the ANN should perform well.

4. Scatterometer measurements

The measurements for the spectral intensities were done once from a blank section of the wafer for the reference intensity, and once from a beam dump for the background intensity. Then all the DOEs were measured in order without redoing the reference or background measurements. They were saved in hdf5 format, and the files contain the wavelengths, intensities for the reference, the background, and the DOEs, and the diffraction efficiencies calculated for each grating. Figure 4 shows examples of the intensity spectra required to calculate the diffraction efficiency calculated using equation (2).

The measurements were conducted in reflectance mode, so only the light reflected from the sample was recorded. When the beam of light hits the wafer, a portion of it passes through it. Some light is reflected back from the bottom surface of the wafer, causing unwanted interference effects that make the measured spectrum unusable. To prevent the internal reflections, the wafers were attached to a black backplate using index matching gel before measurements.

The samples were positioned manually during the measurements. To ensure the correct rotation angle for TE polarisation, the reflected higher diffraction orders were aligned with a reference mark in the system, as shown in figure 5. The correct lateral position was determined using the measurement software. When the measured reflection mode diffraction efficiency is the smallest, the beam of light is not outside the edges of the grating. The spot size is smaller than the gratings, so when the sample is moved the diffraction efficiency spectrum

should not change while the spot is within the area of the grating. Measured data was filtered using a boxcar filter with a width of about 2 nm to reduce the noise.

The measurements were repeated in order to provide more data for the training of the neural network. There had been minor changes in the adjustments of the scatterometer, so the two sets of spectra were slightly different. In addition, some variation is introduced by noise and the positioning done by hand.

5. Creation and training of the ANN

To quickly analyse scatterometric data and provide grating characteristics, an ANN was developed. The training was performed in two phases. In the first phase, only simulated data was used. As figure 6 shows, the match between simulations and experiments is fairly good but not perfect, so an ANN trained using only simulated data can produce significant deviations in the .

Transfer learning can be used to tune the ANN to a new instrument, sample, or variable [17]. If the major parts of the system, excluding the sample tray, were changed or moved, the measurements and the ANN transfer learning would have to be performed again in order to get the best from the ANN. As the transfer learning tunes the ANN for the specific instrument, all the deviations caused by the positioning of the fixed components are taken into account and negated.

5.1. Simulated training data

Preliminary training data was created using Rigorous Coupled Wave Analysis (RCWA). It is a powerful computational tool for analysing the interaction between light and periodic structures by solving Maxwell's equations [18]. RCWA approximates the shape of the structures using stacked rectangular shapes, as figure 7 illustrates, taking into account the sidewall angles, roundness of the top and bottom corners, and other features that the structures might have.

About 200 000 spectral diffraction efficiencies were simulated with grating parameters of a similar magnitude as parameters of the reference samples. The ranges of chosen parameters are described in table 1. In addition to these, a smaller set of simulations was done with 2 nm increments for the period and line width, but the side wall angles and corner radii were constant 90° and 15 nm respectively.

Each simulated spectrum contained the diffraction efficiency for 256 wavelengths evenly distributed between 400 nm and 750 nm. The simulated gratings consist of 15 segments. 6 taller segments for the sidewalls, and 9 segments to model the rounding of the bottom corners. The simulations assume a plane wave light source at transverse electric (TE) mode and angle of incidence of 0° from vacuum to SiO₂.

5.2. Transfer learning reference data

The gratings were characterised using Zeiss LEO Gemini 1550 scanning electron microscope (SEM). Test gratings with identical parameters to the reference gratings were etched

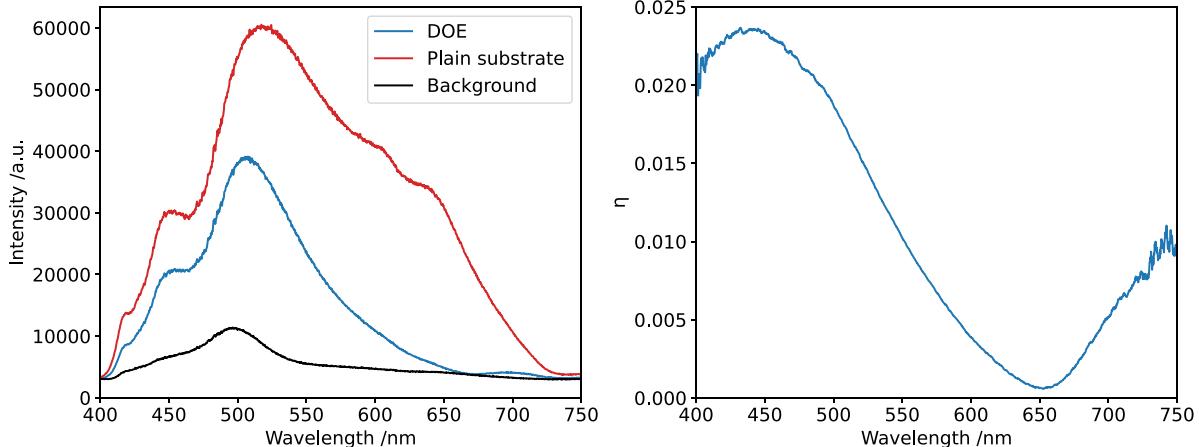


Figure 4. Left: measured spectra from the DOE, plain substrate, and the background. Right: the calculated diffraction efficiency.

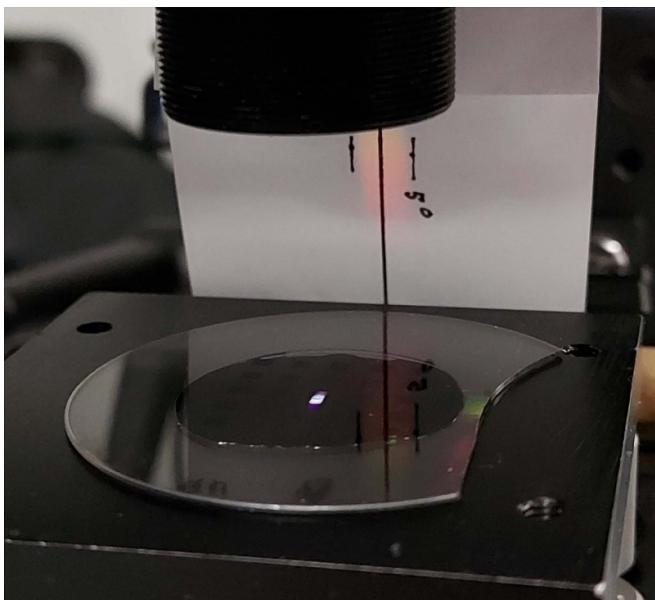


Figure 5. A sample DOE being measured in the scatterometer. The reflected light is aligned with the reference line for correct polarisation angle. The diameter of the wafer is 50 mm.

along one edge of the samples for the purposes of SEM measurements. The wafers were cut so that the test gratings were cut from the main wafer, bisecting the test wafers. The smaller parts containing the test gratings were then prepared for SEM imaging.

For the SEM characterisation the test gratings were coated in a layer of chromium with a nominal thickness of 7 nm. This causes the line width of the SEM images to be larger than it is in the reference samples. Figure 8 shows an example SEM image of the cross section of a grating structure from which the dimensions were measured.

The characterisations were performed by fitting parallel lines to match the edge locations in magnified images to reach subpixel resolution. The standard uncertainty of the SEM characterisation for the period, line width, and step height were estimated to be 2.5 nm. The magnification of the SEM was

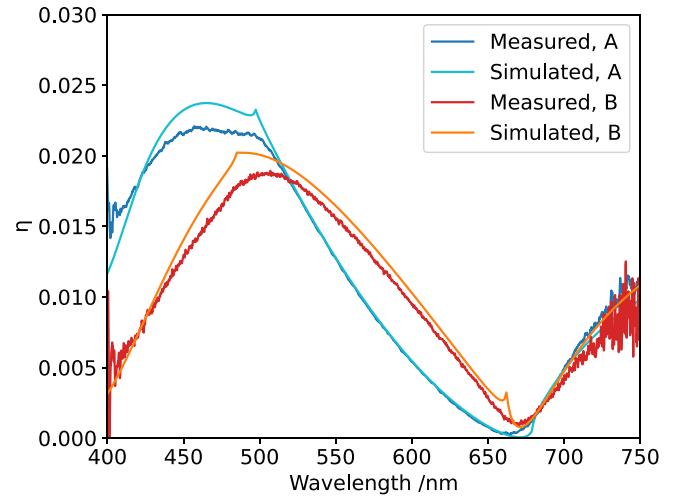


Figure 6. Comparison of simulated and measured spectra of two samples A and B. The simulated curves are created using parameters measured from the SEM images. Sample A period is 680 nm, line width is 309 nm, and height is 210 nm. Sample B period is 675 nm, line width is 351 nm, and height is 283 nm.

separately examined by traceable AFM measurements to agree at level of 0.013%.

5.3. Structure of the ANN

The ANN was created using Tensorflow [19] and Keras [20] packages for Python. The model was constructed in two phases. In the first phase the multi-layer perceptron model is created and trained using the simulated data. The second phase includes the transfer learning where the model was fine-tuned using experimental data.

In the preliminary training phase an array of training data was generated by first randomly selecting a simulated spectrum. Then a small amount of random noise was added to it, so the spectrum would better match the real-world data. Finally, the spectrum was added to the training data set. This was repeated until the set contained 1000000 spectra. This way there was no recognisable structure in the order of the

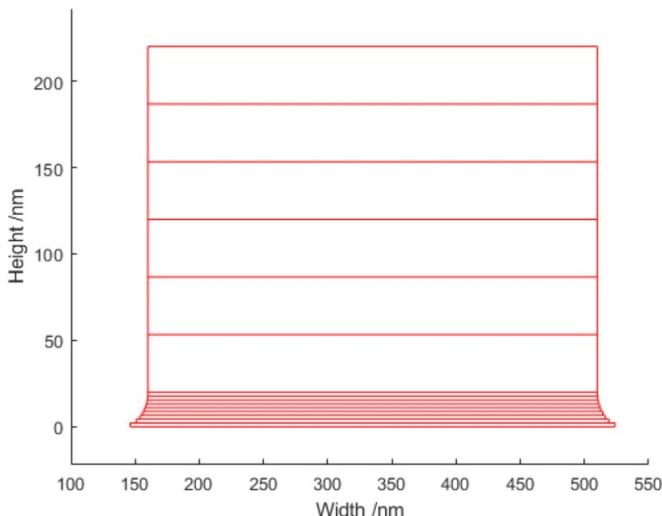


Figure 7. Grating structure constructed for RCWA simulations.

Table 1. The minimum and maximum values, and increment sizes of the grating parameters used in simulations.

Grating parameter	Minimum	Maximum	Increment size
Period	640 nm	700 nm	5 nm
Line width	300 nm	400 nm	5 nm
Height	200 nm	320 nm	2 nm
Side wall angle	87°	90°	1°
Bottom corner radius	10 nm	30 nm	10 nm

spectra, and more data was available for the training as some spectra were included multiple times with different noise.

Finally, the data set was separated into training and test data arrays which were used in the training. The spectra were also scaled using the scaler function from scikit-learn [21] package. Early stopping condition with patience value of 50 was employed to prevent overfitting. The learning rate of the ANN was fairly low, and the network was quite large, so the training required about 600 epochs until the early stopping condition was fulfilled. At this point the ANN contains only one hidden layer. The hyperparameters of this layer seem to have little effect on the final after the transfer learning is applied. A more detailed hyperparameter optimisation was done for the transfer learning phase.

For the transfer learning another hidden layer was added to the model. The first hidden layer was frozen, and the model was then trained using experimental data measured using the scatterometer and SEM imaging. As the resolution of the measured spectra is greater than the resolution of the simulations, the closest matches between the simulated and measured wavelengths and their intensities were chosen for the training. An early stopping condition with patience of 100 was added to prevent overfitting. The value is so large because the learning rate is again very low. Training of the transfer learning layer can take over 2000 epochs.

An in-depth hyperparameter optimisation was conducted only for the transfer learning phase. The hyperparameters

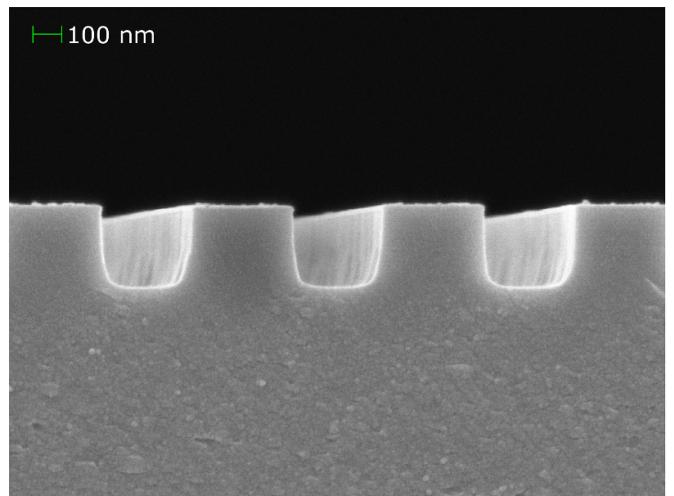


Figure 8. SEM image of a reference sample.

investigated were activation function, number of hidden layers, number of neurons in each layer, and regularisers L1 and L2. The results of the hyperparameter search indicate that the best choices for the parameters are hyperbolic tangent for activation function, one hidden layer with 512 neurons, and regulariser values L1 = 0.00003 and L2 = 0.0009.

As figure 9 shows, the final ANN consists of an input layer with 256 inputs, two hidden layers with 512 neurons each, and an output layer with 7 outputs. It was then integrated to the measurement software, so it can analyse the diffraction efficiencies right as they are measured.

6. Results

The deviations in the between SEM characterisation and the artificial neural network assisted spectral scatterometer (ANNASS) are shown in figure 10. The values are for every measurement done, and each plot contains 138 points. This describes the deviation of the ANNASS and SEM data for the measurands after transfer learning. The residual standard deviations for pitch, line width, and step height are 1.3 nm, 4.3 nm, and 2.6 nm respectively. Before transfer learning the results deviated from the SEM measurements by 1.75 nm for period, -23.47 nm for line width, and 13.73 nm for step height. The residual standard deviations are 4.37 nm, 13.45 nm, and 8.55 nm for the period, line width, and step height respectively.

The values measured from the 6 reference sample gratings that were left out from the ANN transfer learning data are shown in table 2.

6.1. Uncertainty

The following presents a simplified uncertainty estimate for the height parameter measurement of ANNASS. The uncertainty components of the ANNASS can be divided to two main

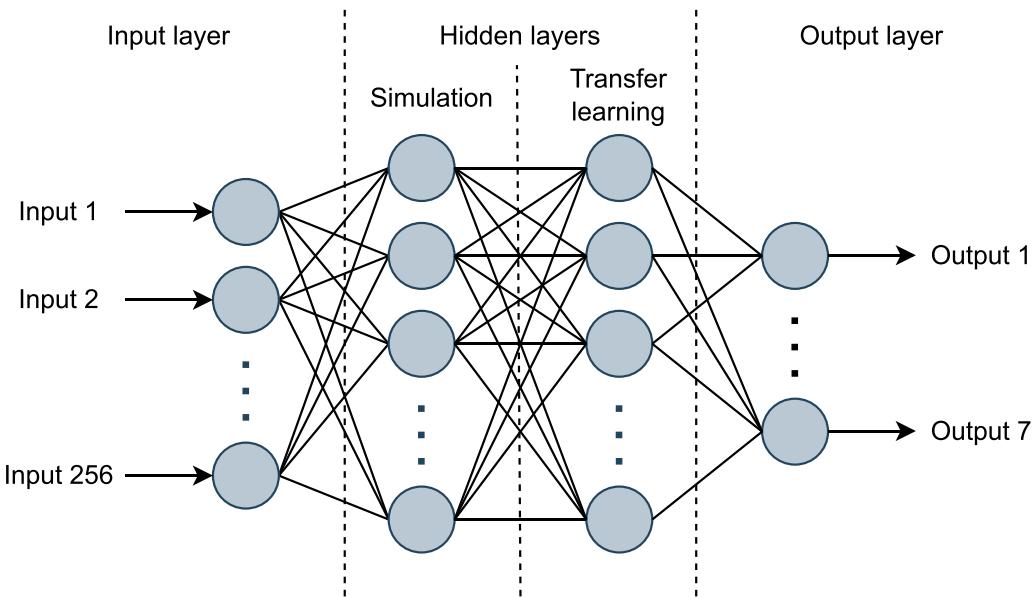


Figure 9. Diagram of the architecture of the ANN. The input parameters are the diffraction efficiencies for each wavelength, and outputs are the grating parameters.

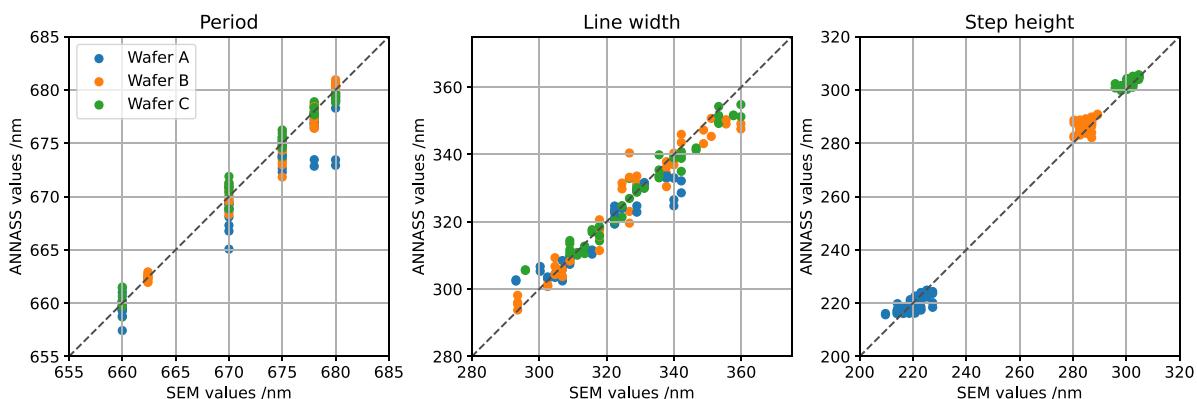


Figure 10. Differences between the period, height, and line width measured by SEM and ANNASS.

Table 2. Comparison between the results for the three measurands acquired using SEM and ANNASS for the reference samples.

Period (nm)			Line width (nm)			Step height (nm)		
SEM	ANNASS	diff.	SEM	ANNASS	diff.	SEM	ANNASS	diff.
665.1	670.0	-4.9	328.6	342.2	-13.6	224.0	223.0	1.0
676.8	678.0	-1.2	331.6	331.2	0.4	220.7	223.0	-2.3
670.5	670.0	0.5	350.3	355.5	-5.2	287.3	282.6	4.7
677.2	678.0	-0.8	330.4	337.8	-7.4	283.7	284.8	-1.1
671.1	670.0	1.1	351.5	357.7	-6.2	301.0	298.1	2.9
678.3	678.0	0.3	339.8	342.2	-2.4	304.7	304.7	0.0

parts: data uncertainty and model uncertainty. Data uncertainty includes the uncertainty of the training data measurements, and the uncertainty sources related to the instrument and the actual measurements. Instrument uncertainties include the following: spectrometer wavelength scale, spectrometer intensity linearity scale, position of the collimation lens, focus distance of the microscope, aperture size, sample angle of

incidence alignment, sample rotary alignment or polarisation angle, alignment of the optical fibres, noise, and uncertainties of the background and reference spectra.

The step height, h , can be expressed as a function with input quantities

$$h = ANN_h(S_\lambda, S_I, c, f, a, \alpha_s, \beta, \alpha_f, n, \sigma_{ANN}, D_{tl}, I_{ref}, I_{bg}). \quad (4)$$

Table 3. Uncertainty evaluation for the step height.

x_i	Quantity	$u(x_i)$	$c_i = \partial ANN_h / \partial x_i$	$u_i(l)$
S_λ	Spectrometer wavelength scale	0.10 nm	1.00	0.10 nm
S_I	Spectrometer intensity scale	0.10 nm	1.00	0.10 nm
c	Collimation lens position	0.10 mm	1.11 nm mm^{-1}	0.11 nm
f	Focus distance	0.10 mm	0.37 nm mm^{-1}	0.04 nm
a	Aperture size	0.10 mm	0.61 nm mm^{-1}	0.06 nm
α_s	Sample angular alignment	1.00°	0.14 nm°	0.14 nm
β	Sample rotational alignment	1.00°	0.18 nm°	0.18 nm
α_f	Fibre alignment	1.00°	0.25 nm°	0.25 nm
n	Noise	0.18 nm	1.00	0.18 nm
I_{bg}	Background spectrum	0.40 nm	1.00	0.40 nm
I_{ref}	Reference spectrum	0.25 nm	1.00	0.25 nm
D_{tl}	Transfer learning reference data	2.50 nm	1.00	2.50 nm
σ_{ANN}	Neural network variations	0.61 nm	1.00	0.61 nm
			standard unc.	2.65 nm
			expanded unc.	5.30 nm

The input quantities are described in table 3. The uncertainty of the h can be estimated, if the input quantity standard uncertainties and sensitivity coefficients are known, using the following formula [22]:

$$u^2(h) = \sum_{n=1}^N \left[\frac{\partial ANN_h(x_1, \dots, x_N)}{\partial x_i} u(x_i) \right]^2. \quad (5)$$

The partial derivative represents the sensitivity coefficient c_i of the quantity x_i . The input quantities, respective sensitivity coefficients, and uncertainty contributions to the total uncertainty for evaluation of the step height measurement uncertainty are detailed in table 3.

The uncertainty of the wavelength scale, S_λ , was estimated based on the calibration data of the spectrometer. The data has largest error of approximately 0.1 nm, which was chosen as the standard uncertainty of the quantity. The sensitivity coefficient of the wavelength scale is naturally 1. The uncertainty of the intensity scale, S_I , was considered negligible based on the nonlinearity error of 0.1%.

The uncertainties of the collimation lens position, c , focus distance, f aperture size, a sample angular and rotational alignments, α_s and β , and fibre alignment, α_f , were evaluated by making small changes to the adjustments of the respective parameters of the scatterometer. The effects that these changes produced on the ANN results were examined to determine the sensitivity coefficients. In addition, evaluation was done for the standard uncertainty of each of the variables. These values are listed in table 3.

Contribution from noise, n , was evaluated by repeated measurements of a grating and observing the changes in the ANNASS results. Uncertainty of the background and reference intensities, I_{bg} and I_{ref} , were evaluated by measuring the same grating but remeasuring the background intensity or the reference intensity between each measurement and observing how the ANNASS results are affected.

The uncertainty of the transfer learning data, D_{tl} , was evaluated by considering the accuracy of the detection of the grating parameter h from the SEM characterisation, taking into account calibration uncertainty and edge detection accuracy. This uncertainty has sensitivity coefficient of 1. The model uncertainty, or the uncertainty introduced by neural network variations, σ_{ANN} , was evaluated by repeating the creation and training of the ANN multiple times with same data in different order. The uncertainty was chosen as the standard deviation of the difference between SEM and ANNASS measurements for h .

The expanded uncertainty ($k = 2$) of the period is 5.94 nm, for the step height it is 5.30 nm, and for the line width it is 7.28 nm. The results depend greatly on the data used for the transfer learning. The largest source of uncertainty is clearly the accuracy of the transfer learning data, i.e. the uncertainty of the grating parameters from SEM measurements. For period and line width, the uncertainty evaluation was performed in a similar manner.

7. Conclusion

In this paper is described a spectral scatterometer for photonics industry quality control purposes. The instrument utilises machine learning for spectral analysis and is capable of providing results for several measurands simultaneously and very quickly. The most important measurands in typical samples are height, pitch, line width, and side wall angles. A novel concept was demonstrated by employing transfer learning in the training of an artificial neural network. The benefits of using an ANN for inverse solving are that it is faster than traditional methods, and if the simulations do not exactly match the measurements, transfer learning will increase the accuracy of the inverse solving. Additionally, once the ANN is trained, there is no need for large libraries of simulated diffraction efficiencies.

The deviation of the results for the grating parameters by SEM and ANNASS was under 5 nm in most cases, with the

largest deviations for line width. In addition, a metrological validation was made using independent reference samples. The uncertainty evaluation for the measurand height, h , resulted in the value of 5.3 nm for the expanded uncertainty. Within this uncertainty, the deviation of h values between the results obtained using machine learning and SEM measurements are shown to be in agreement.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.23729/7a9bbd75-7b8a-43d5-8c55-489f96e718b7> [23].

Acknowledgments

We thank Heikki Astola for his assistance regarding transfer learning and hyperparameter optimisation.

This Project (20IND04 ATMOC) has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

This work was supported by Academy of Finland Flagship Programme, Photonics Research and Innovation (PREIN), Finland, (Decision No. 320167).

ORCID iDs

- Aleksi Mattila  <https://orcid.org/0009-0008-6047-6189>
 Johan Nysten  <https://orcid.org/0000-0001-6294-4267>
 Ville Heikkinen  <https://orcid.org/0000-0002-9208-5501>
 Jorma Kilpi  <https://orcid.org/0000-0003-2932-4298>
 Virpi Korpelainen  <https://orcid.org/0000-0001-9123-8521>
 Poul-Erik Hansen  <https://orcid.org/0000-0002-1664-8763>
 Petri Karvinen  <https://orcid.org/0000-0002-3251-2524>
 Markku Kuittinen  <https://orcid.org/0000-0002-5334-6844>
 Antti Lassila  <https://orcid.org/0000-0002-6991-7082>

References

- [1] Xiong J, Hsiang E-L, He Z, Zhan T and Wu S-T 2021 Augmented reality and virtual reality displays: emerging technologies and future perspectives *Light Sci. Appl.* **10** 216
- [2] Calaon M, Madsen M and Leach R 2018 *Scatterometry* (Springer) pp 1–5
- [3] Husu H, Saastamoinen T, Laukkanen J, Siitonens S, Turunen J and Lassila A 2014 Scatterometer for characterization of diffractive optical elements *Meas. Sci. Technol.* **25** 044019
- [4] Wurm M, Pilarski F and Bodermann B 2010 A new flexible scatterometer for critical dimension metrology *Rev. Sci. Instrum.* **81** 023701
- [5] Garnaes J, Hansen P-E, Agersnap N, Holm J, Borsetto F and Kühle A 2006 Profiles of a high-aspect-ratio grating determined by spectroscopic scatterometry and atomic-force microscopy *Appl. Opt.* **45** 3201–12
- [6] Madsen M and Hansen P-E 2016 Scatterometry—fast and robust measurements of nano-textured surfaces *Surf. Topogr. Metrol. Prop.* **4** 023003
- [7] Liu S, Chen X, Yang T, Guo C, Zhang J, Ma J, Chen C, Wang C, Zhang C and Liu S 2022 Machine learning aided solution to the inverse problem in optical scatterometry *Measurement* **191** 110811
- [8] Madsen J S M, Jensen S A, Nygård J and Hansen P E 2018 Replacing libraries in scatterometry *Opt. Express* **26** 34622
- [9] Li T *et al* 2021 Photonic-dispersion neural networks for inverse scattering problems *Light Sci. Appl.* **10** 154
- [10] Kallioniemi I, Saarinen J and Oja E 1999 Characterization of diffraction gratings in a rigorous domain with optical scatterometry: hierarchical neural-network model *Appl. Opt.* **38** 5920
- [11] Kallioniemi I and Saarinen J 1999 Optical scatterometry with neural network model for nondestructive measurement of submicron features *Proc. SPIE* **3743** 33–40
- [12] Robert S, Mure-Ravaud A and Lacour D 2002 Characterization of optical diffraction gratings by use of a neural method *J. Opt. Soc. Am.* **19** 24–32
- [13] Gereige I, Robert S, Thiria S, Badran F, Granet G and Rousseau J J 2008 Recognition of diffraction-grating profile using a neural network classifier in optical scatterometry *J. Opt. Soc. Am. A* **25** 1661
- [14] Wei S and Li L 2008 Measurement of photoresist grating profiles based on multiwavelength scatterometry and artificial neural network *Appl. Opt.* **47** 2524
- [15] Tchéré M G, Robert S, Fawzi Z S and Bayard B 2019 Reconstruction of a complex profile shape by weighting basic characterization results for nanometrology *Appl. Opt.* **58** 6118–25
- [16] Madsen M H, Hansen P-E, Zalkovskij M, Karamehmedović M and Garnaes J 2015 Fast characterization of moving samples with nano-textured surfaces *Optica* **2** 301–6
- [17] Astola H, Seitsonen L, Halme E, Molinier M and Lönnqvist A 2021 Deep neural networks with transfer learning for forest variable estimation using Sentinel-2 imagery in boreal forest *Remote Sens.* **13** 2392
- [18] Moharam M G, Grann E, Pommert D and Gaylord T K 1995 Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings *J. Opt. Soc. Am. A* **12** 1068–76
- [19] Abadi M *et al* 2015 TensorFlow: large-scale machine learning on heterogeneous systems (available at: tensorflow.org)
- [20] Chollet F *et al* 2015 Keras (available at: <https://keras.io>)
- [21] Pedregosa F *et al* 2011 Scikit-learn: machine learning in Python *J. Mach. Learn. Res.* **12** 2825–30
- [22] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML 2008 *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement* JCGM 100:2008 (Joint Committee for Guides in Metrology) (available at: www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf/cb0ef43f-baa5-11cf-3f85-4dc86f77bd6)
- [23] Mattila A 2024 Artificial neural network assisted spectral scatterometry for grating quality control *Fairdata* (<https://doi.org/10.23729/7a9bbd75-7b8a-43d5-8c55-489f96e718b7>)