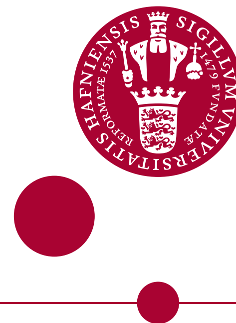
UNIVERSITY OF COPENHAGEN



Replacing libraries in scatterometry

Madsen, Jonas Skovlund Moller; Jensen, Soren Alkaersig; Nygard, Jesper; Hansen, Poul Erik

Published in:
Optics Express

DOI:
10.1364/OE.26.034622

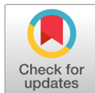
Publication date:
2018

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Madsen, J. S. M., Jensen, S. A., Nygard, J., & Hansen, P. E. (2018). Replacing libraries in scatterometry. *Optics Express*, *26*(26), 34622-34632. https://doi.org/10.1364/OE.26.034622

Download date: 09. apr.. 2020



Research Article | Vol. 26, No. 26 | 24 Dec 2018 | OPTICS EXPRESS 34622

Optics EXPRESS

Replacing libraries in scatterometry

JONAS SKOVLUND MØLLER MADSEN,[1,2,*] SØREN ALKÆRSIG JENSEN,[1] JESPER NYGÅRD,[2] AND POUL ERIK HANSEN[1,3]

[1]*DFM A/S DK-2970, Hørsholm, Denmark*
[2]*Niels Bohr Institute, University of Copenhagen DK-2100 Copenhagen, Denmark*
[3]*peh@dfm.dk*
[*]*jsm@dfm.dk*

Abstract: Diffraction gratings have a wide array of applications in optics, diagnostics, food science, sensing, and process inspection. Scattering effects from defects can severely degrade the performance of such gratings. In this paper, we consider three classes of defects: Two classes introduced at the grating/air interface, as a change in line heights, and one class introduced as a sinusoidal variation of the grating/substrate interface. The scattering properties of the gratings are modelled using rigorous coupled wave analysis, and defects are approximated with a new semi-analytical model and a neural network. The new methods make it possible to avoid the time consuming library generation/search strategy commonly used in scatterometry. The method does not introduce new numerical parameters, and therefore no new parameter correlations. This work enables improved grating reconstruction, especially of non-diffracting short pitch gratings. It is found that two of the defect classes can be adequately described by the semi-analytical model, while the third defect is accurately reconstructed by a neural network. The network is demonstrated to be faster than a library search and more versatile for related structures.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Periodic nanostructures can be used to enhance or add desired functions [1,2]. In fabrication of these nanostructures, product defects of a few nanometers are hard to avoid. As the sizes of the produced nanostructures shrink, the relative sizes of these defects increase. For NIL on hard surfaces linewidth and height defects of tens of nanometers have been reported [3].

Furthermore, defects are likely to be even more prominent when fabrication moves out of the laboratory and into the mass-production industry [4]. These defects can affect the functionality of the nanostructures, for example in their interaction with light [5].

Scatterometry is an optical technique already utilized in the semiconductor industry, where a diffraction grating is reconstructed from an optical signal by inverse modelling [6]. Rigorous coupled wave analysis (RCWA) [7] is the common workhorse for scatterometry modelling due to its speed, convergence and relatively simple implementation. In RCWA, the diffraction grating is approximated by rectangular slabs, and Maxwell's equations are solved by coupling the boundary conditions between the slabs [8].

Conventionally, angular resolved scatterometry has been deployed in the semiconductor industry [9]. Moving to wavelength-resolved scatterometry, one can reduce the complexity of the measurement hardware by removing the need for goniometric detector setups, since all information can be collected at a single point. Furthermore, as the period of the functional nanostructures decreases, the number of observable (non-evanescent) diffraction orders decreases as well. When all the optical information is extracted from the zeroth order (specular reflection), the reconstruction of complex structures such as multiple trapezes, additional material layers and rounded corners, tend to become ambiguous due to correlations between the added parameters. In the worst cases, one could end up reconstructing a structure different from the physical sample without knowing it. It is therefore relevant to investigate how the

#347873 https://doi.org/10.1364/OE.26.034622
Journal © 2018 Received 11 Oct 2018; revised 22 Nov 2018; accepted 22 Nov 2018; published 20 Dec 2018

zeroth diffraction order from a complex structure can be approximated by simpler models without the dreaded ambiguity.

In this paper, RCWA is used for forward calculation of the diffraction orders from a grating. However, when describing complex structures such as roughness and defects, a large amount of parameters is needed to describe the grating. Since the RCWA solver is called once for each permutation of parameters, this ends up being very time-consuming [10]. Furthermore, one could end up with a strong correlation between the parameters, making the reconstruction unstable [11]. Based on this, it is desirable to retain a low number of parameters in the modelling without losing too much information.

In the presented work we, show how scatterometric data can be used to detect different classes of defects. We simulate the multi-spectral zeroth order diffraction from three classes of defects, described in the next section, with varying magnitudes of the introduced defect. This is done by describing the grating by a unit cell containing multiple grating lines, called a supercell. We then investigate how these defects can be characterized without using a conventional library search. We present a semi-analytical method based on the total integrated scatter model [12] to incorporate the effect of scattering from defects into the scatterometric reconstruction. This method requires only a single numerical calculation describing the perfect grating. The total integrated scatter model is combined with the simplest structure form of RCWA, a rectangular grating described by a period, height and width, to make a semi-analytical model. We show that this model accurately predicts the scatterometric signature of the defects on the sample. The model can be used to reduce the dimensionality of the scatterometry library, and thus simplify the analysis process. It is stressed that the semi-analytical model suggested could just as well be used for an angle resolved spectrum.

Furthermore, we shown how a neural network can be developed and used to analyze defects in place of a library search strategy. Neural networks have previously been deployed to select a library for scatterometry [13,14] or replace the library [15,16]. These works aimed to characterize gratings created from simple unit cells, where all grating lines are assumed identical. We demonstrate that we can use the neural network to characterize defects from the most complex supercell and achieve good performances on similar defects.

2. Simulations

The diffraction efficiency, η, is defined as the zeroth diffracted order with respect to the incoming light [8]. Numerically calculated diffraction efficiencies, η_{Num}, for the structures with different defects are found using RCWA. All structures modelled in the examples shown here are constructed by imposing the defects on perfect line gratings in silicon. The perfect line gratings are described by their period, Δ, height, h, their width, w. The simulated gratings have the following parameters: Δ = 2 μm, h = 0.7 μm and w = 1 μm. In addition, periodic defects are added to theses gratings by describing the grating as a super cell consisting of ten unit cells of the perfect grating, this super cell then has a period of Γ = 20 μm.

Three classes of defects are examined by incorporating them into the structure of the supercell:

1) A simple defect where the first of the ten grating lines are higher than the others, see Fig. 1(a).

   This was observed in our previous work with injection molded nanostructures, where some grating lines had additional material on top [17].

2) A sinusoidal defect in the height over the grating, see Fig. 1(b). This defect can occur in bottom-up fabrication [18].

Optics EXPRESS

3) A perfect grating on a sinus shaped substrate, see Fig. 1(c). This defects could arise from a grinding step in the substrate preparation, or could be purposely introduced as in ref [19].

The magnitude of the defects is described by the parameter d shown in Fig. 1. Simulations are performed using incoming light polarized along the grating lines (TE), with an incident angle θ_I of 70 degrees with respect to the grating normal, and a wavelength, λ, ranging from 250 nm to 850 nm. In order for the numerical supercell calculations to converge, a large number of diffraction orders, defining the truncation of the Fourier series in the RCWA calculations [19], needs to be retained in the calculations, which in turn makes the calculations very time-consuming. In the presented work, over 400 orders are reatined in the RCWA calculations.

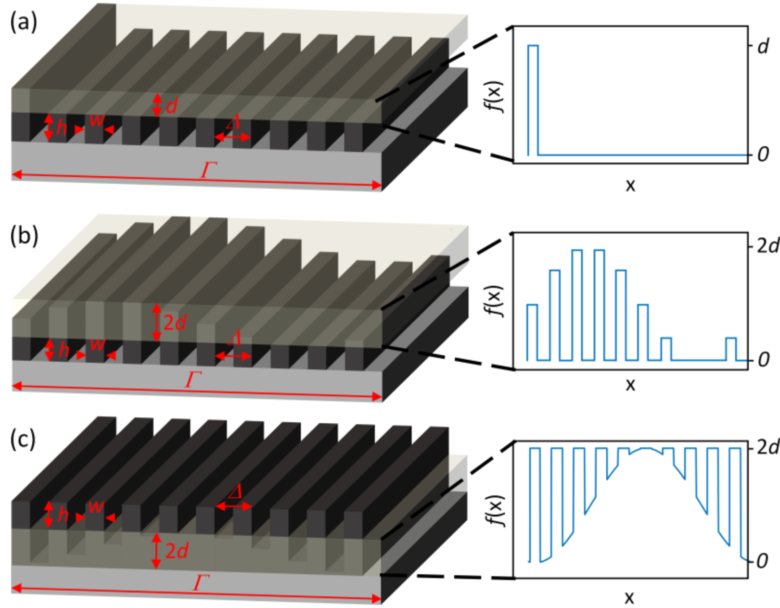


Fig. 1. Sketches of the different grating imperfections examined. The denoted parameters are the height of the perfect grating, h, the width of the grating, w, and the size of the imperfection, d, the period of the simple grating, Δ, and the period of the supercell, Γ. The marked volume shows where the defect is introduced and the plot to the left shows the profile of the defect used in the semi-analytical model. (a) Simple grating defect. (b) Sinusoidal grating defect. (c) A perfect grating on a sinusoidal substrate. All imperfections have been exaggerated for clarity.

3. Semi-analytical model

The Total Integrated Scatter (TIS) originates from scalar diffraction theory [20]. The model describes what fraction of the reflected light is scattered from a rough surface. It is conventionally described as [12]:

$$TIS = \frac{R_T - R_S}{R_T} = 1 - \exp\left(-\left(4\pi \frac{\sigma \cdot \cos(\theta_I)}{\lambda}\right)^2\right) \tag{1}$$

Here, R_T, is the total reflectance, R_S, is specular reflected light, σ is the Root Mean Square roughness of the surface, also commonly known as R_q [21].

The roughness, σ, used in the TIS calculation is found as:

$$\sigma = \sqrt{\frac{1}{\Gamma}\int_0^{\Gamma} f(x)^2 dx} \tag{2}$$

where the height function *f(x)* describes the part of the grating with the introduced defect (see profile in Fig. 1). The magnitude of the defect *d* can be moved out of the integral, making it possible to solve the integral and scale with *d*, for an analytical mapping between σ and *d*. It is important to note that in the general case, σ should be the appropriately bandwidth limited surface roughness, since spatial frequencies larger than $\frac{1}{\lambda}$ produce evanescent waves and are irrelevant with regards to scattering [12]. In this paper, the variations have a low frequency, and Eq. (2) can safely be used.

Our semi-analytical model assumes that the introduced defect can be treated as a perturbation to the perfect grating described by the TIS model, and that the scattering caused by the defect and the perfect grating are uncoupled so that the defect can be treated as a perturbation. The diffraction efficiency found from the semi-analytical model, η_{SA}, is given by:

$$\eta_{SA} = \frac{I_S}{I_I} = \eta_{Grat} \cdot (1 - TIS) \tag{3}$$

where η_{Grat} is the diffraction efficiency from the rectangular grating with no defects, I_S and I_I are the intensities of the scattered and incoming light respectively. This enables us to use a single numerical calculation for the simple rectangular grating when characterizing a grating with defects. It should be noted that other computational methods than RCWA, such as finite difference time domain [22] or finite element method [23], can be used to find η_{Grat} as well.

4. Results

In order to compare the agreement between the diffraction efficiency, η_{Num}, calculated using RCWA on the supercell, and the diffraction efficiency η_{SA} from the semi-analytical model, we look at the difference between the samples with introduced defects and perfect samples.

In Fig. 2 we show: $\delta\eta_{Num} = \eta_{Num} - \eta_{Grat}$ (solid line) and $\delta\eta_{SA} = \eta_{SA} - \eta_{Grat}$ (crosses) as a function of the wavelength for the three defect classes at different magnitudes. We see that the semi-analytical model works well for the first class and exceptionally well for the second class.

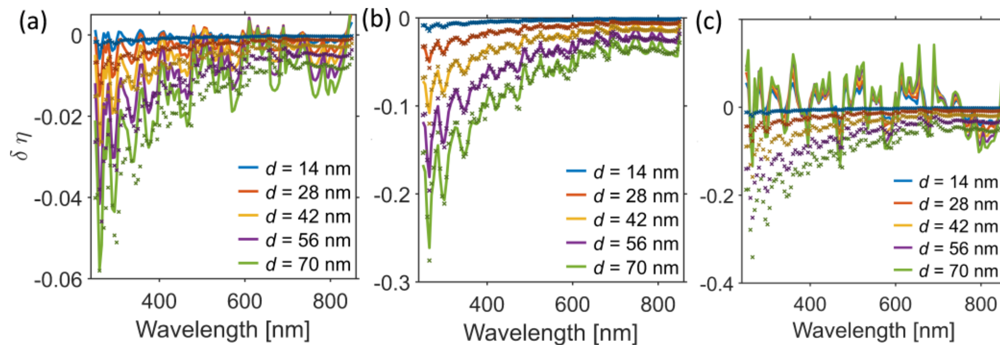


Fig. 2. Defect class 1 (top), class 2 (middle) and class 3 (bottom). Fully drawn lines show the $\delta\eta_{Num}$ while the slightly darker crosses show $\delta\eta_{SA}$. *d* is illustrated in Fig. 1. Note the different y-axes.

However, the third class is not described well by the semi-analytical model. Intuitively, this break-down of the model can be understood by looking at how the defect is introduced on the structure: For the first two classes, the defect is placed on top, but for the third, the defect

is embedded into the perfect structure. The poor result in Fig. 2(c) suggests that the signals from the two areas cannot be decoupled, and therefore cannot be approximated by the semi-analytical model. This suggests that the defects above the grating, and defects inside the grating must be treated differently.

4.1 Defect above grating

Going back to the first two classes, we see that we have a good agreement between the RCWA and the semi-analytical model, which becomes gradually worse as the magnitude of the defect increases. This would suggest that the Semi-analytical model is valid for “moderately” rough surfaces. This is a property inherited from the TIS model. It is noted that type 1 has a lesser impact on the optical signal (notice the different y-axis), which is to be expected since the defect class 1 has a lower change in the volume of the material.

Since the total integrated scatter based model describes the second type of grating well, it can be used to find the defect magnitude d from an intensity signal using a library search approach traditional employed in scatterometry. To demonstrate this, the correct signal is taken to be the RCWA signal with- and without applied white Gaussian noise with a standard deviation of 0.5%, and the semi-analytical model is fitted by minimizing the mean square error, *MSE,* described below using nearly continuous values for the magnitude of the defect, d. These continuous values of d are denoted δ.

$$MSE(\delta, d) = \frac{1}{N}\sum_{i=1}^{N}\left(\eta_{SA}(\lambda_i, \delta) - \eta_{Num}(\lambda_i, d)\right)^2 \quad (4)$$

where N (121 in this case) is the number of wavelengths used, and $\eta_{Num}(\lambda_i, d)$ is the numerical diffraction efficiency with d locked at several values from 2% to 10% of the total grating height. $MSE(\delta, d)$ can be seen for the different defect d values in Fig. 3 We see that the best fitting solution finds the defect size from the RCWA simulations with- or without added noise. This demonstrates that for this case, the simple semi-analytical model can be used to describe the defect. Furthermore, for the noiseless RCWA, we see that the *MSE* “dip” becomes wider as the defect magnitude increases, and therefore the best solution becomes less well defined. This result shows that the semi-analytical model works best for small defect values, a property inherited from the TIS model. In the presence of noise, we see that the best fitting solution becomes worse due to the MSE reaching the noise floor. This shows, that the signal-to-noise ratio acts as a lower boundary condition for the model. If one wants to look at very small perturbations to the perfect grating, one needs a good signal to noise ratio.

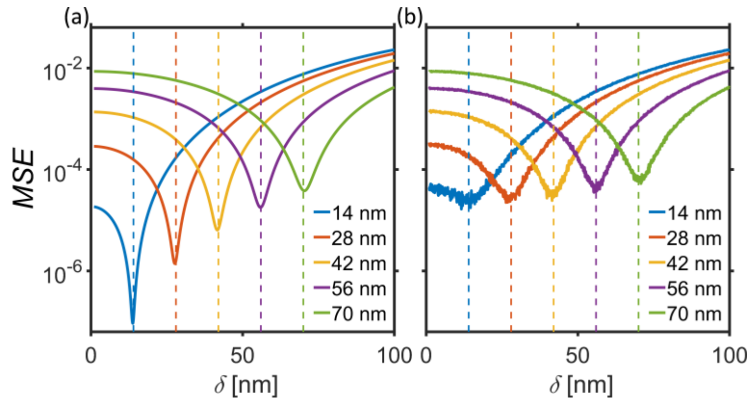


Fig. 3. Mean square error as a function of defect size used in the semi-analytical model for defect class 2 using pure RCWA (a) and RCWA with added white Gaussian noise (b). The legends and the dashed lines indicate the defect size d used in the RCWA calculations, while the fully drawn lines indicate the defect sizes used in the semi-analytical model.

Since the semi-analytical model is based on the simplest form of RCWA (a rectangle described by a single slab), it could also be used in combination with conventional library search scatterometry aimed at determining the parameters of the perfect grating. Here one could use the model to add a defect or roughness parameter to a pre-generated database without having to re-calculate a library with a higher dimensionality. In the simplest case one would only need to simulate a single structure numerically and account for other variations analytically, which is computationally much more efficient.

It would also be possible to completely avoid a library search by first assuming that the measured signal is described by η_{SA}. Then the ratio $\frac{\eta_{SA}}{\eta_{Grat}}$ could be used to analytically calculate σ from Eq. (3) and (1). Since there is a linear mapping from σ to d for each class, this could be used to analytically characterize the defect magnitude.

Before the time of computers, where exponential functions should be avoided, Eq. (1) was approximated by a firs order Taylor expansion:

$$TIS = \frac{R_T - R_S}{R_T} = 1 - \exp\left(-\left(4\pi\frac{\sigma \cdot \cos(\theta_I)}{\lambda}\right)^2\right) \approx \left(4\pi\frac{\sigma \cdot \cos(\theta_I)}{\lambda}\right)^2 \quad (5)$$

The approximated TIS expression is assumed to be valid for optically smooth surfaces, commonly described by the g-factor [24]:

$$g = 4\pi\frac{\sigma \cdot \cos(\theta_I)}{\lambda} << 1 \quad (6)$$

If g is small enough, higher order terms from the Taylor series can be safely discarded. This is thoroughly discussed in refs [25–27].

Since the approximation of the TIS model is still seeing industrial use [9], it is interesting to investigate how well the semi-analytical model fares using the first order Taylor approximation for the total integrated scatter model.

Figure 4 shows the *MSE* minimization using the approximation of Eq. (5) in the semi-analytical model. In this case we see that the semi-analytical model finds smaller defects than the numerical approach. This “deficit” seems to increase with the magnitude of the defect, corresponding to the approximation becoming worse as g gets larger. The limit where the approximation stops working could be interpreted as the point where the optically smooth criterion is no longer valid. For class 2 with $d = 30$ nm, we obtain a roughness parameter g of 0.365 and 0.107 for the lowest and highest wavelengths respectively. Again, we see the trend that the RCWA with noise finds the same solution as the noiseless RCWA.

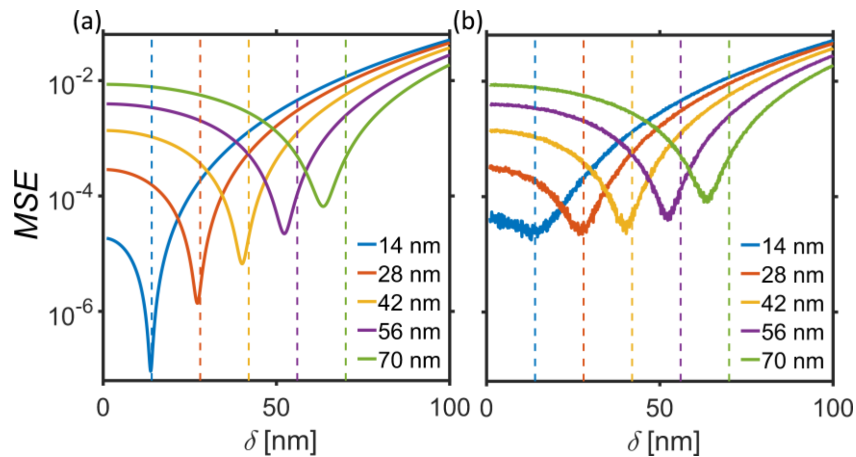


Fig. 4. Same approach as Fig. 3 but using the simplified (first order Taylor) TIS model for defect class 2. The legends and the dashed lines indicate the defect size d used in the RCWA calculations, while the fully drawn lines indicate the defect sizes used in the semi-analytical model. We see that this model already falls off for defect values above 28 nm (corresponding to 4% of the grating height).

Previous work on injection molded nanostructures [17,28] has shown that one often ends up with very little characteristic features in the wavelength resolved spectrum. In those cases, it is necessary to restrict the reconstruction to a few model parameters. The presented method can be used as a perturbation for the simpler structures and gives an idea of whether or not these defects are within a safe limit or if they might be detrimental to the desired functionality. The easy implementation would make the method much more attractive to use with existing libraries rather than recalculating a new library with a higher dimension to account for the defects.

4.2 Defect in the grating area

As we saw in the previous section, class three defects could not be well described by the semi-analytical model. The simulations do, however, clearly show that we can easily distinguish a grating on a periodic grinded (sinusoidal) surface from a grating on a plain surface, since the signal change is much larger than typical measurement uncertainties associated with scatterometry [8]. This result was also experimentally reported in ref [19]. Furthermore, it is observed that the signal changes very little with the magnitude of the defect. Looking at the solid line data in Fig. 2(c), it is clear that any model describing the effect of the third defect class would be complex. Therefore, it is decided to attempt a solution using machine learning.

A neural network has been developed using RCWA simulated data sets, as a placeholder for experimental data, with d varying from 1 to 100 nm in steps of 1 nm. Physical measurements have been simulated by adding white Gaussian noise [29] to the simulated spectra. A thousand sets of noisy data are made from each simulation, resulting in 100.000 data sets used for the network.

The network type is a multilayer perceptron [30]. Here the input data points are passed through an input layer with a node for each wavelength, 121 in total, a hidden layer with 10 nodes, and then converted to output data at the end by an output layer with a single node finding the defect magnitude d. All neurons from the hidden layer are interconnected to all nodes in input layer and the output layer through weighted transfer functions. The network is sketched in Fig. 5. When training the network, the weights of the transfer functions are adjusted in order to map a desired output from the input by minimizing a mean squared error function. The network described here is trained using a Levenberg-Marquardt algorithm [31–33]. The input

layer uses a tan-sigmoid transfer function (Tansig), and the output layer uses a linear transfer function (Purelin) [34,35].

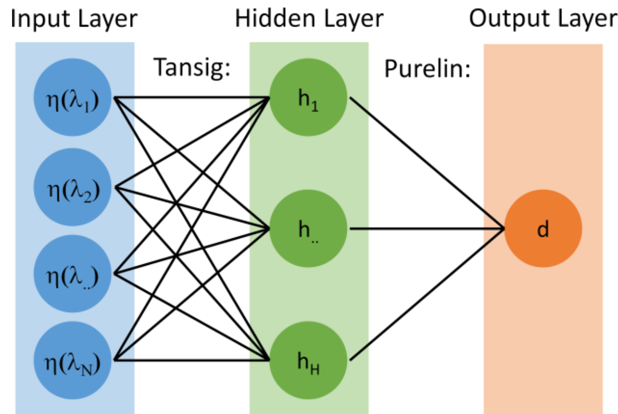


Fig. 5. Sketch of the neural network. The input layer has a node for each wavelength simulated (121). The hidden layer has a total of 10 nodes, and the output layer has a single node finding the defect magnitude d. The nodes from the input layer are connected to the hidden layer through a weighted Tansig transfer function. In the same manner, all nodes in the hidden layer are connected to the output node through a Purelin transfer function.

Once the network is trained, it can be used to predict outputs from new unseen data. In order to evaluate the networks ability to determine defects from a scatterometry measurement, a new set of measurement signals are simulated. These signals are then passed through the network, and the estimated defect is extracted. The targeted defect value, d_T, is the value of d used in the RCWA code to generate the true signal, and the found defect value, d_F, is the value of d predicted by the network. The results can be seen in Fig. 6. We see that we have an overall good agreement with a seemingly randomly distributed deviation from the targeted value. This is expected, since the signal changes very little with a change in the defect magnitude, as seen in Fig. 2(c). The found magnitudes are fitted as a linear function of the targeted magnitudes, and the correlation:

$$d_F = 0.9959 \cdot d_T \tag{7}$$

is found.

Since we have a good fit and no major outliers, it is concluded that this neural network can be used to reliably characterize class three defects. Since the prediction does not repeat calculation steps, it can be done very fast. The prediction time does not scale with the size of the data used to generate the network as opposed to a standard library approach, where the search time is directly proportional to the library size. This would make a neural network approach even more suited for inline characterization, if computation time starts to present a bottle-neck. For a quick comparison, the neural network finds the defect magnitude in 0.46 ms, while a library approach as used in ref [36]. uses roughly 0.01 ms for each generated RCWA structure (typically tens of thousands, but could easily be larger for complex structures). Both calculations were performed on a standard laptop.

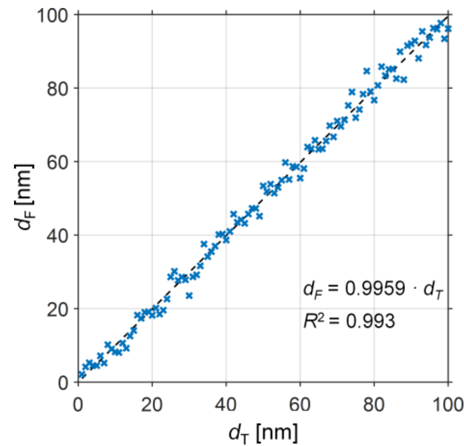


Fig. 6. Performance of the Neural network. The defect found by the neural network, d_F, is plotted vs the targeted defect, d_T, from the simulated models with added noise. The black dashed line shows the best linear fit to the data.

Defects, by definition, are not perfect. It is therefore interesting to see how the same network performs on similar, but different substrate structures. To test the developed network, we look at how well it predicts the defect magnitude from a substrate described by a bottom cut sinusoidal. Data was simulated for cuts of 12.5%, 25%, 37.5% and 50% of the total height as sketched in the upper insert of Fig. 7. This was done without any retraining of the network.

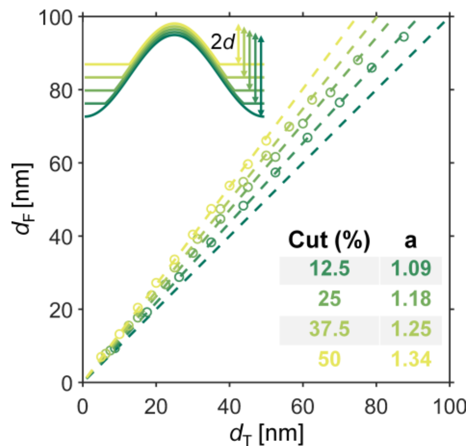


Fig. 7. Evaluation of the developed network using different substrate defects. Circles shows the magnitude found by the network, and the dashed lines show the best linear fit. The diagonal line shows the fit obtained from the uncut sinusoidal in Fig. 6 to guide the eye. Upper insert shows the simulated structures from uncut to 50% cut. The curves have been displaced for clarity. The grating lines on top of the substrate has been omitted in the illustration. The Arrows mark 2*d* for the different cut values, where d is the target value for the given structure. Lower insert: table showing the linear coefficient for the fit: $d_F = a\, d_T$.

The performance of the network can be seen in Fig. 7. The parameters for the best fitting line can be seen in the inserted table. The network clearly recognizes features from the perfect sine substrates, seen by the linearity between d_T and d_F. Good linear fits are seen for all cut values. As a trend, the neural network over-predicts the amplitude, by a larger ratio for the larger cut values. This is likely caused by the network only being trained on perfect sine structures. This means that the network will match the signal to the best fitting perfect structure, which by nature has a larger d value than the corresponding cut structure. This suggests that even if we do not have a perfect sinusoidal structure, we can still estimate a substrate defect

and compare the relative substrate roughness between two samples. It has thus been demonstrated that the neural network can be used to predict a defect size for similar, but not identical, substrates. This method is believed to be more stable, than a library of RCWA signals simulated from perfect sinusoidal.

Future work will emphasize using the semi-analytical model in combination with inline characterization and further development of the neural network by adding new defect types. Here defects as line edge roughness will be of certain interest.

5. Conclusion

We have examined three different classes of defects introduced on perfect rectangular silicon gratings. The defects were introduced above the grating area and in the grating area for the two first classes and the last class respectively. A semi-analytical model has been suggested to determine the magnitude of the grating defect. For the first two classes, the defect is in agreement with a semi-analytical model based on TIS and RCWA. The third class cannot be described by the semi analytical model. This model enables defect characterization of low period non-diffracting structures. A neural network has been developed to characterize these defects. The network can accurately determine the magnitude of the defect. Both methods can be used to create simple models describing the defects without the need of additional RCWA computations, and in some cases make it possible to entirely omit a library generation and search. Future work will emphasize using the semi-analytical model in combination with inline characterization and further development of the neural network by adding new defect types. Here defects such as line edge roughness will be considered.

Funding

Danish Agency for Institutions and Educational Grants, the Quantum Innovation Center; Eurostars project E11002-OptoRough.

References

1. C. Pan, Z. Liu, Y. Pang, X. Zheng, H. Cai, Y. Zhang, and Z. Huang, "Design of a high-performance in-coupling grating using differential evolution algorithm for waveguide display," Opt. Express **26**(20), 26646–26662 (2018).
2. W. Lee and S.-J. Park, "Porous anodic aluminum oxide: anodization and templated synthesis of functional nanostructures," Chem. Rev. **114**(15), 7487–7556 (2014).
3. *13th International Conference on Metrology and Properties of Engineering Surfaces 2011: Twickenham, United Kingdom, 12 - 15 April 2011*, Journal of Physics Conference Series No. 311 (Curran, 2011).
4. M. G. S. Bernd, S. R. Bragança, N. Heck, and L. C. P. S. Filho, "Synthesis of carbon nanostructures by the pyrolysis of wood sawdust in a tubular reactor," J. Mater. Res. Technol. **6**(2), 171–177 (2017).
5. V. K. Tewary, ed., *Modeling, Characterization, and Production of Nanomaterials* Woohead Publishing Series in Electronic and Optical Materials No. Number 73 (Elsevier, 2015).
6. C. Raymond, "Overview of scatterometry applications in high volume silicon manufacturing" in (AIP, 2005), Vol. **788**, pp. 394–402.
7. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," J. Opt. Soc. Am. **71**(7), 811–818 (1981).
8. M. H. Madsen and P.-E. Hansen, "Scatterometry—fast and robust measurements of nano-textured surfaces," Surf. Topogr. **4**(2), 023003 (2016).
9. J. C. Stover, *Optical Scattering: Measurement and Analysis.* (SPIE Press, 2016).
10. S. Schröder, A. Duparré, L. Coriand, A. Tünnermann, D. H. Penalver, and J. E. Harvey, "Modeling of light scattering in different regimes of surface roughness," Opt. Express **19**(10), 9820–9835 (2011).
11. S. Rizvi, ed., *Handbook of Photomask Manufacturing Technology* (Taylor & Francis, 2005).
12. J. E. Harvey, "Total integrated scatter from surfaces with arbitrary roughness, correlation widths, and incident angles," Opt. Eng. **51**(1), 013402 (2012).
13. I. Gereige, S. Robert, S. Thiria, F. Badran, G. Granet, and J. J. Rousseau, "Recognition of diffraction-grating profile using a neural network classifier in optical scatterometry," J. Opt. Soc. Am. A **25**(7), 1661–1667 (2008).
14. I. Kallioniemi, J. Saarinen, and E. Oja, "Characterization of diffraction gratings in a rigorous domain with optical scatterometry: hierarchical neural-network model," Appl. Opt. **38**(28), 5920–5930 (1999).
15. S. Wei and L. Li, "Measurement of photoresist grating profiles based on multiwavelength scatterometry and artificial neural network," Appl. Opt. **47**(13), 2524–2532 (2008).

16. S. Robert, A. Mure-Ravaud, and D. Lacour, "Characterization of optical diffraction gratings by use of a neural method," J. Opt. Soc. Am. A **19**(1), 24–32 (2002).
17. J. S. Madsen, L. H. Thamdrup, I. Czolkos, P.-E. Hansen, A. Johansson, J. Garnaes, J. Nygard, and M. H. Madsen, "In-line characterization of nanostructured mass-produced polymer components using scatterometry," J. Micromech. Microeng. **27**(8), 085004–085010 (2017).
18. B. D. Gates, Q. Xu, M. Stewart, D. Ryan, C. G. Willson, and G. M. Whitesides, "New approaches to nanofabrication: molding, printing, and other techniques," Chem. Rev. **105**(4), 1171–1196 (2005).
19. P.-E. Hansen, M. H. Madsen, J. Lehtolahti, and L. Nielsen, "Traceable Mueller polarimetry and scatterometry for shape reconstruction of grating structures," Appl. Surf. Sci. **421**, 471–479 (2017).
20. P. Beckmann, *The Scattering of Electromagnetic Waves from Rough Surfaces*, International Series of Monographs on Electromagnetic Waves Volume 4 (Macmillan Company, 1963).
21. D. J. Whitehouse, *Surfaces and Their Measurement*, Kogan Page Science Paper Edition (Kogan Page Science, 2004).
22. K. Yee, "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media," IEEE Trans. Antenn. Propag. **14**(3), 302–307 (1966).
23. T. Delort and D. Maystre, "Finite-element method for gratings," J. Opt. Soc. Am. A **10**(12), 2592–2601 (1993).
24. J. E. Harvey, N. Choi, and A. Krywonos, "Scattering from moderately rough interfaces between two arbitrary media," in S. A. Straka and N. Carosso, eds. (2010), pp. 77940–77950.
25. H. Davies, "The reflection of electromagnetic waves from a rough surface," Proc. IEE - Part IV Inst. Monogr. **101**, 209–214 (1954).
26. H. E. Bennett and J. O. Porteus, "Relation between surface roughness and specular reflectance at normal incidence," J. Opt. Soc. Am. **51**(2), 123–138 (1961).
27. J. C. Stover, S. A. Serati, and C. H. Gillespie, "Calculation of surface statistics from light scatter," Opt. Eng. **23**(4), 234406 (1984).
28. J. S. Madsen, S. A. Jensen, L. Nakotte, A. Vogelsang, L. H. Thamdrup, I. Czolkos, A. Johansson, J. Garnæs, T. Nielsen, J. Nygård, and P. E. Hansen, "Scatterometry for optimization of injection molded nanostructures at the fabrication line," Int. J. Adv. Manuf. Technol. **99**(9-12), 2669–2676 (2018).
29. V. Mathuranathan, *Digital Modulations Using Matlab: Build Simulation Models from Scratch* (2017).
30. M. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," Atmos. Environ. **32**(14-15), 2627–2636 (1998).
31. K. Levenberg, "A method for the solution of certain non-linear problems in least squares," Q. Appl. Math. **2**(2), 164–168 (1944).
32. D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963).
33. B. G. Kermani, S. S. Schiffman, and H. T. Nagle, "Performance of the Levenberg–Marquardt neural network training method in electronic nose applications," Sens. Actuators B Chem. **110**(1), 13–22 (2005).
34. D. Ramakrishnan, T. N. Singh, N. Purwar, K. S. Barde, A. Gulati, and S. Gupta, "Artificial neural network and liquefaction susceptibility assessment: a case study using the 2001 Bhuj earthquake data, Gujarat, India," Comput. Geosci. **12**(4), 491–501 (2008).
35. R. Hmamouchi, M. Larif, S. Chtita, A. Adad, M. Bouachrine, and T. Lakhlifi, "Predictive modelling of the LD_{50} activities of coumarin derivatives using neural statistical approaches: Electronic descriptor-based DFT," J. Taibah Univ. Sci. **10**(4), 451–461 (2016).
36. J. S. Madsen, P. E. Hansen, P. Boher, D. Dwarakanath, J. F. Jørgensen, B. Bilenberg, J. Nygård, and M. H. Madsen, "Study on microgratings using imaging, spectroscopic, and fourier lens scatterometry," J. Micro Nano-Manuf. **5**(3), 031005 (2017).

130.225.212.4 {ts '2019-01-16 08:34:32'}