Research Article | Vol. 27, No. 16 | 5 Aug 2019 | OPTICS EXPRESS A1030
Optics EXPRESS

Training artificial neural network for optimization of nanostructured VO_2-based smart window performance

IGAL BALIN,[1,4,*] VALERY GARMIDER,[1,4] YI LONG,[2,3] AND IBRAHIM ABDULHALIM[1,3]

[1]*Department of Electro-Optics and Photonics Engineering, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel*
[2]*School of Materials Science and Engineering, Nanyang Technological University, 50 Nanyang Avenue, 639798, Singapore*
[3]*Singapore-HUJ Alliance for Research and Enterprise (SHARE), Nanomaterials for Energy and Energy-Water Nexus (NEW), Campus for Research Excellence and Technological Enterprise (CREATE), 138602, Singapore*
[4]*These authors contributed equally*
[*]*igalb@post.bgu.ac.il*

Abstract: In this work, we apply for the first time a machine learning approach to design and optimize VO_2 based nanostructured smart window performance. An artificial neural network was trained to find the relationship between VO_2 smart window structural parameters and performance metrics-luminous transmittance (T_{lum}) and solar modulation (ΔT_{sol}), calculated by first-principle electromagnetic simulations (FDTD method). Once training was accomplished, the combination of optimal T_{lum} and ΔT_{sol} was found by applying classical trust region algorithm on the trained network. The proposed method allows flexibility in definition of the optimization problem and provides clear uncertainty limits for future experimental realizations.

1. Introduction

Intelligent architectural windows, or "smart windows", are a promising approach to reduce the energy consumption of buildings, which contribute up to 40% of the world's energy usage [1]. Generally, these windows operate by blocking a large fraction of the sunlight on hot days, while transmitting solar energy in cold weather [2]. Through temperature-responsive solar energy modulation, thermochromic smart windows are key components of green buildings that have been extensively studied in the recent years [3]. Vanadium dioxide (VO_2) is a thermochromic material that changes its optical behavior at a critical transition temperature (T_c) of 341°K (68°C), where a metal to insulator (MIT) phase transition is accompanied with abrupt decrease in the Infrared (IR) transmittance while maintaining transmittance in the visible range [4,5]. In this manner, VO_2-based thermochromic windows can largely block solar thermal energy inflow at high temperatures without compromising brightness. At the same time, the transmission of visible light is nearly unaffected during the phase transition. The modification of optical properties during the MIT made VO_2 an attractive candidate for thermochromic smart window applications. However, the practical use of VO_2 thin films for energy-efficient windows has long been hampered by a tradeoff between luminous transmittance (T_{lum}) and solar energy modulation (ΔT_{sol}) [6] as a function of thickness. In general, the practical architectural windows require T_{lum} larger than 60% while concurrently maintaining moderately high solar modulation (ΔT_{sol}) [7].These requirements could be achieved utilizing properly optimized nanostructured VO_2 thin films [8,9].

Conventional design approach of nanophotonic structures is based on non-systematic optimization, in which the design is performed by hand-tuning of small number of structural

#361234 https://doi.org/10.1364/OE.27.0A1030
Journal © 2019 Received 27 Feb 2019; revised 28 Mar 2019; accepted 28 Mar 2019; published 10 Jul 2019

parameters and computing the optical response using electromagnetic simulations (i.e. FDTD or similar method) [10,11]. However, in this approach the conclusions are highly dependent on the structure under investigation. Consequently, a trial-and-error process is inevitable. The problem becomes more severe when the number of parameters increases.

Combination of electromagnetic simulations with advanced optimization methods such as genetic algorithm [12], particle swarm optimization [13,14] etc. can introduce more systematic and faster approach for finding the optimal solution. However, these methods have significant drawbacks: (a) computationally expensive as the nanostructure size and complexity grow, (b) missing the flexibility of re-definition of the optimization problem if this step is required and (c) cannot be used for further investigation of the proposed design, such as uncertainty analysis etc.

Here we propose solving the design problem of VO_2 based nanostructured smart window as an example for application of machine learning methods combined with FDTD simulations and standard optimization algorithms. In this method, an artificial neural network (ANN) model was developed to predict the values of T_{lum} and ΔT_{sol} functions, (derived from FDTD simulations) on the basis of structural parameters. A standard Lagrange multipliers algorithm then uses the trained neural network prediction model to find an optimal combination of structural parameters for which T_{lum} and ΔT_{sol} are maximized. There are two main advantages in this approach: (a) the neural network is used to replace the computationally expensive electromagnetic simulations in the optimization loop, greatly reducing design time, (b) neural network is trained only once. The training phase in the neural network is a non- recurring effort, meaning once the data set is learned, the query phase is quasi instantaneous. This is in clear departure with the standard optimization methods where for each query the whole parameter space is searched [15]. The concept of using ANNs in optics and photonics has received a significant degree of exposure during the last few years. For example, it was used for designing nanophotonic and plasmonic structures [15,16] optimization of organic thin film solar cells [17] and photonic crystal nanocavities [18], modeling of plasmonic transmission lines [19] and light scattering from multilayer nanoparticles [20], spatial resolution improvement of optical microscopy [21], phase recovery and holographic image reconstruction [22]. In all these works the ANNs were implemented in its classical form, in which standard ANN training via optimization is (from a probabilistic perspective) equivalent to maximum likelihood estimation (MLE) for the weights. This approach is not well designed to model uncertainty associated with the predictions they make [23]. To overcome this limitation, we implement Bayesian regularization in our ANN. In this way we can specify priors to inform and constrain our models and get uncertainty estimation in form of a posterior distribution.

2. Structural model and optimization problem definition

2.1 Structural model and FDTD simulations

The structural model used in this work is one-dimensional subwavelength VO_2 grating structure supported by fused silica substrate (shown in Fig. 1). The chosen grating shape is parabolic-a bioinspired choice mimicking the natural broadband antireflection structures that give rise to an approximately linear refractive index gradient [24]. The structural parameters used for optimization are grating pitch w, parabola base width w_0 and height z.

The FDTD simulations were performed using commercial software (FDTD solution, Lumerical Inc. Vancouver, Canada). The entire system suspended in air/vacuum, and the incident beam modelled as a plane wave propagating along the z-axis. In all the simulations of the present study, a broadband plane wave source is propagated from the top to the bottom and at incident angle $\theta = 0$. Bloch and perfectly matched layer (PML) boundary conditions are imposed for x- and z- coordinates, respectively. The dielectric dispersion profiles of the materials were fitted by the multi-coefficient model that relies on an extensive set of basis functions. The simulation wavelength bandwidth and optical constants of the materials were

Optics EXPRESS

selected for operation across the spectral range 300 nm to 2500 nm, accounting for 99.2% of AM 1.5 integrated solar irradiance [25]. Fields monitors were placed at fixed z position below and above the VO_2 grating structure to detect the beam transmission and reflection respectively as a function of wavelength. The other parameters were set as follows: a simulation time of 100 fs with an auto-shutoff parameter of 10^{-5}, a mesh accuracy of 5 (i.e., 22 mesh points per wavelength), and mesh refinement algorithm set to "conformal variant 1" allowing for a nonuniform mesh over the FDTD domain. Data from these field monitors was used to calculate broadband reflectance and transmittance coefficients. The performance metrics of VO_2 based smart window T_{lum} and ΔT_{sol} were calculated using Eqs. (1):

$$T_{lum} = \frac{\int_{\lambda=380nm}^{780nm} \phi_{lum} T(\lambda) d\lambda}{\int_{\lambda=380nm}^{780nm} T(\lambda) d\lambda}; T_{sol} = \frac{\int_{\lambda=300nm}^{2500nm} AM1.5(\lambda) T(\lambda) d\lambda}{\int_{\lambda=300nm}^{2500nm} AM1.5(\lambda) T(\lambda) d\lambda}$$

$$\Delta T_{sol} = T_{sol,semi} - T_{sol,metal} \tag{1}$$

where T(λ) represents the measured film transmittance, Φ_{lum} - photopic luminous efficiency of the human eye [26], and the air mass (AM1.5) solar irradiance spectrum distributions [25] were used as weighting functions for the wavelength dependent transmittance coefficients, semi and metal stands for semiconducting and metallic state of VO_2, respectively. Without loss of generality, we choose the light polarization to be parallel to the grating grooves (TM polarization) and all calculations of the performance metrics were done for this polarization as well. The proposed design can be easily extended to two dimensions, which will prevent the polarization dependence of T_{lum} and ΔT_{sol}.

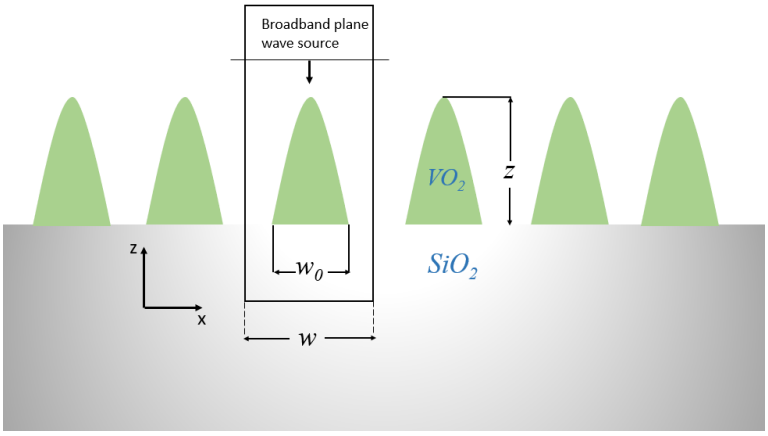

Fig. 1. VO_2 grating structure with specified design parameters.

2.2 Optimization problem

The general aim of the present study is to find the optimal set of VO_2 grating structure parameters such as both luminous transmission T_{lum} and solar modulation ΔT_{sol} are maximized. Taking into account the findings from previous studies in which the maximal luminous transmission values achieved are varying between 70 and 78% [8,9,27], we choose to formulate two separate constrained optimization problems:

$$I.\ find \max_{w,w_0,z} \Delta T_{sol} : subject\ to\ T_{lum}(w, w_0, z) > 75\% \tag{2}$$

Optics EXPRESS

$$II.\ for\ w_0 \in \left[w_{0,min}, w_{0,max} \right] find \max_{w,z} \Delta T_{sol}\ : \\ subject\ to\ T_{lum}\left(w, w_0, z\right) > 75\% \tag{3}$$

Optimization problem I was defined for identification of optimal parameters in the whole configuration space. Optimization II was performed to explore further the configuration space at the vicinity of global solution found by the combined algorithm in Optimization Problem I.

3. Bayesian ANN training

3.1 Bayesian ANN brief overview

Artificial neural networks are parametrized regression models of the relationship between input and output parameters [28]. In our case, neural networks relate VO_2 subwavelength grating structural parameters to resultant Figure of Merits (e.g. T_{lum}, ΔT_{sol}). Typically, ANN operation consists of sequential application of "layers", each one constituted of linear mapping followed by nonlinear scalar functions. The weights and biases of linear mappings in all layers are the parameters to be "learned" from data given appropriately defined merit of the discrepancy between data and the model (i.e. ANN). Traditional learning algorithms are designed to seek parameter values that minimize this discrepancy merit. The learning process is usually conducted by dividing the data into training, validation and testing sets. This is done to select the model and to ensure that the trained network is not biased towards the training data it has seen. An alternative approach is the use of Bayesian (probabilistic) formulation. A remarkable advantage of this formulation is the natural incorporation of prior information in the learning process to regularize ANN architecture complexity with intention to prevent overfitting, i.e. convergence at neural networks representing unrealistic relation which generalizes poorly to unobserved cases [29,30]. This advantage will be clearly manifested in investigation of subwavelength structures having resonant response such as plasmonic nanoparticles, metamaterials etc. Validation and testing sets are not required for learning reliability proof. Instead, Bayesian formulation is capable of assigning confidence intervals for the obtained predictions, which are the direct measures of predictions' reliability.

Bayesian approach provides estimation of quantities in terms of their expectation value with respect to posterior probability density function. This requires calculation of integrals in many dimensions (as many as ANN weights and biases), which are intractable analytically. Monte Carlo Integration (MCI) is the common approach to address this issue [29,30]. The general idea of MCI is follows:

1. Draw statistically independent samples of ANN weights and biases from the underlying posterior distribution – a learning stage.
2. Estimate expectation value of ANN prediction as arithmetic mean over its values obtained for distinct ANN samples.

In our specific case, we estimate optimal grating structural parameters and their uncertainty as first and second moments of the posterior distribution given the data obtained from electromagnetic simulations. In section 3.2 we elaborate on the implementation details of the entire optimization procedure.

3.2 Data set generation and ANN training

The optimization process starts with inputs sampling and outputs evaluation at the sampled points. The input is defined by VO_2 grating structural parameters set $\mathbf{X} = (W, W_0, z)$ whose search domain was limited to upper and lower bounds specified in Table 1. The logic behind the pitch and base width high bound choice is prevention of diffraction which is undesirable in smart window based application [31]. The lower bound is limited based on practical limitations of existing nanofabrication equipment. The grating height is limited in order to

Optics EXPRESS

prevent both interference-based effects and losses which might deteriorate the smart window performance.

The output parameters are smart window performance metrics-luminous transmission T_{lum}, solar modulation ΔT_{sol}, and reflection coefficient R_{av} (reflection averaged in visible spectral range) which are obtained for each of the parameters set from the spectra generated by FDTD simulation.

With the generated data at hand, we trained ANN model constituted of 3 inputs (W, W_0, z) and 3 outputs (T_{lum},, ΔT_{sol}, R_{av}). Training was performed by Markov Chain Monte-Carlo (MCMC) sampling of ANN weights and biases [29]. The entire procedure is described in detail in the Appendix (section 6).

Table 1. VO_2 structural parameters input space used for ANN training

Parameter	W	$W_0 = fW$ (f<1)	z
Lower Bound (nm)	35	0	0
Upper Bound (nm)	180	180	240

Markov Chain Monte Carlo (MCMC) is utilized to draw statistically independent samples of ANN weights and biases from the underlying posterior distribution [29]. As a result, we obtained different ANN realizations corresponding to the obtained weights/biases sets. For each one of these ANN realizations we solve the optimization problem to obtain typical VO_2 grating structure parameters. Then we average over all the obtained structure parameters sets. Their standard deviation provides measure of prediction uncertainty [29–32]. The entire optimization procedure is depicted on Fig. 2.

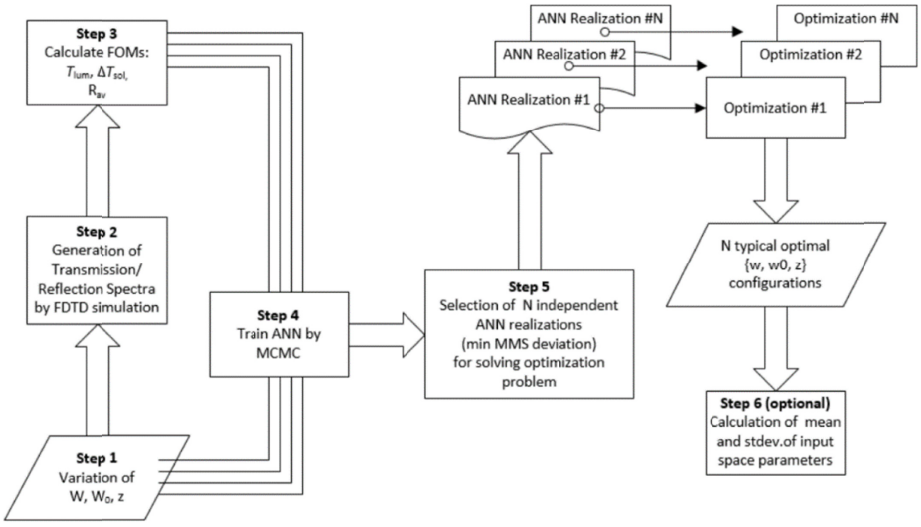

Fig. 2. Flow chart describing the ANN aided optimization procedure.

In the broad context, we performed an interpolation for T_{lum} and ΔT_{sol} functions by means of ANN. The advantage in using ANN compared to simpler approaches such as polynomic interpolation is its ability of prediction. As a result, ANN consider much less nodes than any other method of interpolation and less computational time as a consequence.

Optics EXPRESS

4. Results and discussion

Following ANN training process accomplishment, the last 20 ANN realizations are chosen for optimization. Trust region algorithm is used to solve the constrained optimization separately for each one of the ANN realizations in order to estimate the uncertainty of the optimal parameters prediction. The result for optimization problem I is shown in Fig. 3. It can be clearly seen that every optimization parameter set for the chosen ANN realizations is localized around the mean value within 3% standard deviation. This finding is an evidence for high reliability of the proposed method. Despite the sparse training set used for ANN training, the proposed algorithm managed to represent properly the relationship between grating parameters and FOMs. It should be noted that the proper training set spacing can be identified in the hindsight only once there is prior knowledge about the functional dependence of the performance metrics, which is not available in general. Based on uncertainty introduced by Bayesian method, a proper decision can be made regarding the necessity of additional data incorporation if this step is required.

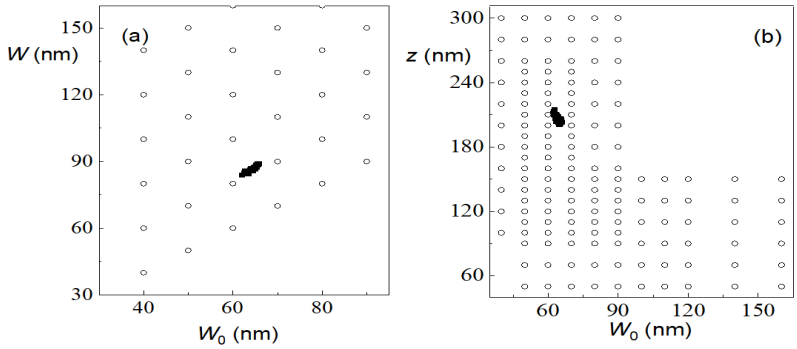

Fig. 3. Optimization problem I results for ANN input parameters. Visualization is done in W_0W (a) and zW_0 (b) projection planes. The empty circles denote points in input space used for ANN optimization and black circles denote optimization results for each ANN realization.

It is worth noting that our choice of Bayesian framework was made with intention to demonstrate the capability to obtain reliable and valuable results in situations where only small data set (less than 1000 examples) is available for training, which is typical if costly system simulations are involved to produce the data. If no such limits are imposed on data generation process, much larger data sets can be exploited in conjunction with standard learning algorithms. Uncertainty estimation, however, is out of the scope of these methods capability, and one should look for alternative measures of reliability proof.

In foresight, instead of solving constrained optimization problem an alternative approach of ANN utilization for structure design problems would be the following: given certain requirements imposed on FOM by practical application, execute a variant of reinforced learning to seek all possible sets of geometrical structure parameters for which the resultant FOM are expected to comply with these requirements. This might be useful in robustness and tolerance analysis of the designed structure.

The optimization results for problem II are presented in Fig. 4. Figure 4(a) and 4(b) show the optimal grating pitch W and height z and Fig. 4(c) and 4(d) show the corresponding solar modulation, ΔT_{sol} and semiconducting state averaged reflection R_{av} as a function of parabola base width W_0. The optimal duty cycle W_0/W remains constant however the aspect ratio (z/W_0) shows a monotonic decrease as the base width W_0 increases. As a result, the effective area of VO_2 per unit cell $S \propto (W \cdot z)$ decreases. The decrease in effective area is expected to reduce the absorption in metallic state resulting in decrease of solar modulation $\Delta T_{sol.}$ It can be seen in Fig. 4(c) that both effective area and the optimal solar modulation indeed show similar

functional dependence. The maximal solar modulation is found at the same set of input parameters for both optimization problems I and II. These findings are an additional evidence for consistency and reliability of the chosen optimization method.

The parabola base range in which ΔT_{sol} is maximized and average reflection minimized is identical for all chosen ANN realizations (see Fig. 4(c) and 4(d) for details). Despite the intrinsic losses in thin VO_2 film in both semiconducting and metallic state, the anti-reflective properties of parabolic grating profile allow achieving less than 1% averaged reflection in the visible wavelength range and consequent improvement of luminous transmission T_{lum}.

We performed a sanity check of the global optimization results – the output optimal parameters were compared with the output parameters of single FDTD simulation run having the same input parameters as the mean value in the optimal input parameters set. According to the results presented in Table 2, the deviation between performance FOMs of single simulation vs optimized results is less than 3%.

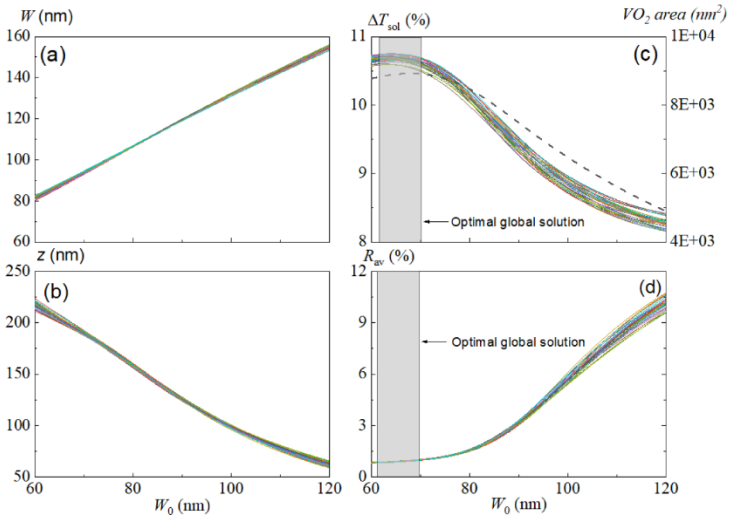

Fig. 4. Optimization problem II results visualization for ANN input parameters: (a), (b) and output parameters: (c), (d). Solid lines denote different ANN realizations optimization results in (a)-(d). The dash line in (c) denotes parabola area within one grating period.

The origin of optimal solar modulation can be understood by attentive analysis of FDTD simulation raw spectra generated with optimal set of input parameters, shown in Fig. 5.

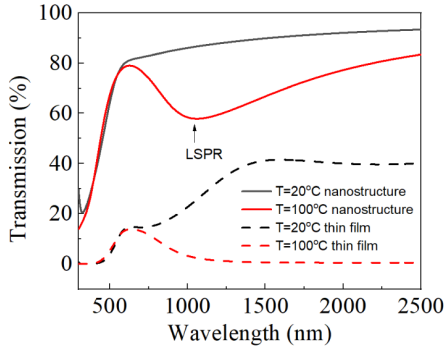

Fig. 5. Transmission spectra generated by FDTD simulation for VO_2 grating structure having optimal parameters: W_0 = 64 nm, W = 86 nm and z = 208 nm (solid lines) and VO_2 film having same thickness. Black and red color denote semiconducting and metallic phase, respectively.

Optics EXPRESS

The localized plasmon resonance (LSPR) is excited at ~1 μm wavelength in metallic phase. This excitation is the major contributing factor for absorption in VO_2. It can be seen that in order to achieve an optimal combination of maximal luminous transmission and solar modulation, the excitation of LSPR should be located as close as possible to the lower edge of NIR wavelength range. This requirement is achieved by constant duty cycle of VO_2 grating and identified by our optimization algorithm (Fig. 4(a)). Overall, based on Table 2 it can be seen the optimal performance achieved in this study (T_{lum} = 75.1% and ΔTsol = 10.7 are comparable with the results achieved in the field of VO_2 based smart windows [8,9,27]. However, in this work the approach for finding an optimal solution is robust and not configuration/system-dependent.

Table 2. Sanity check for global optimization results vs single FDTD simulation

Input/Output Parameters	Model	Single FDTD simulation
W_0 (nm)	64 ± 3	64
W (nm)	86 ± 4	86
Z (nm)	208 ± 8	208
T_{lum} (%)	75.1 ± 0.1	74.8
ΔT_{sol}(%)	10.7 ± 0.1	10.9

5. Conclusions

This study presents a novel approach for performance optimization of nanostructured VO_2 smart window based on machine learning between structural parameters and performance metrics. This approach is extremely fast, e.g., once the ANN is trained, the optimization problem can be solved within 200 ms on laptop computer (2.3 GHz Intel Core i5-6200U processor with 8 GB RAM) whereas single FDTD simulation run requires several minutes. The usage of ANN allows flexibility in optimization problem re-definition without executing additional sets of electromagnetic simulation. The introduction of Bayesian regularization into ANN allowed estimating the uncertainty of optimization results, which is crucial for future experimental realization.

Furthermore, using the same machine learning approach can be used not only for optimization of nanostructure parameters but also for intrinsic properties of VO_2 itself such as bandgap engineering via doping for achieving reduction of VO_2 transition temperature, control of the hysteresis width and sharpness, control of T_{lum} and ΔT_{sol} in thin films, etc. The combination of structural and intrinsic properties optimization using ANNs will necessarily lead towards achieving designs that will overcome the obstacles for practical implementation of thermochromic smart window coatings.

6. Appendix: ANN training specification

6.1 ANN architecture and hierarchical model description

The ANN model F(x,w) we used in our study is given by:

$$F(x,w) = w_{ho} \tanh(w_{ih} \cdot x + b_h) + w_{io} \cdot x + b_0 \quad (4)$$

where w_{ih} are weights on connections from inputs to hidden units, b_h are hidden biases, w_{ho} are weights on connections from hidden to output units, w_{io} are weights on connections from inputs to outputs and b_o are output biases. These parameters comprise the ANN weights and biases w = (w_{ih}, b_h, w_{ho}, w_{io}, b_o). The ANN model we used constituted of 3 inputs x =

(W,W_0,z), one hidden layer with 12 units with "tanh" activation function, 3 outputs y = ($T_{lum,}$, ΔT_{sol}, R) and 9 skipped connections from inputs to outputs.

Bayesian rule gives the posterior distribution for ANN weights and biases w and the hyper-parameters in hierarchical manner:

$$p(w,\beta,\tau|x,y) \sim p(y|x,w,\beta)\cdot p(w|\tau)\cdot p(\beta) \tag{5}$$

where $p(y|x,w,\beta)$ ~Normal($F(x,w)$, $\beta^{-1/2}$) is the likelihood, $p(w|\tau)$ ~Normal(0, $z^{-1/2}$) is the prior for weights and biases, $p(\tau)$ ~Gamma(μ,α) and p(β) ~Gamma(μ,β) are the priors for the hyperparameters τ (precisions of weights and biases) and β (outputs' noise precisions).To reflect the vague prior knowledge regarding the hyper-parameters, the scale μ and the shape α parameters were set to 10 and 0.02 respectively.

The vector τ is comprised of 11 precisions, which are assigned to separate groups of weights and biases as follows: one per group of connections from each input to all hidden, one to all hidden biases, one per group of connections from all hidden to each output, one per group of connections from all inputs to each output. The output biases were given a fixed precision set to 1. This choice of precision parameters assignment incorporates the idea of Automatic Relevance Determination (ARD), due to MacKay and Neal [29,32]. The vector β is comprised of 3 separate noise model precisions for each output.

6.2 Data set preprocessing

Values of input data, i.e. structure parameters, are "spherized" i.e. subtracted the mean from and normalized by standard deviation over the dataset. Since the outputs, i.e. figures of merit we deal with, $T_{lum,}$, ΔT_{sol} and R_{av}, are all positive valued quantities by definition, we apply "log" operation on entire dataset to map them onto [-∞, +∞] numerical axis. The obtained values are then "spherized". The obtained spherization bias and normalizing constants were saved for the necessary later inverse transformation of the predictions to original numerical space.

6.3 ANN training

Weights and biases are initialized by setting all to zero. Then consecutive Gibbs sampling updates of hyper-parameters and ANN weights/biases are performed as follows:

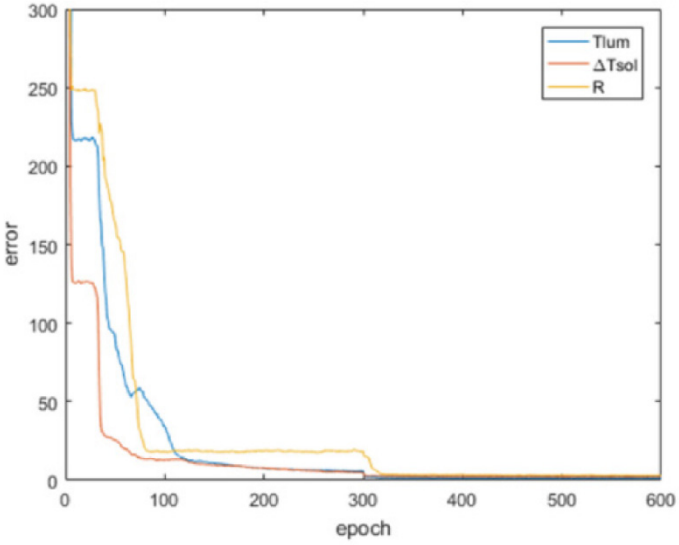

Fig. 6. Model-target MMS deviation.

Optics EXPRESS

- Hyper-parameters are drawn from conditional distribution $p(\beta,\tau|w,x,F)$ which is of the Gamma form for given current values of w.
- ANN weights and biases are sampled from conditional distribution $p(w|\beta,\tau,x,F)$ by Hamiltonian Monte Carlo [29] with number of leapfrog steps L and step size parameter η depending on training phase as specified below.

These steps have been repeated 2600 times, each time updating ANN weights/biases together with the hyperparameters, thus generating different ANN realization. Entire procedure has been divided into three phases: 1. Burn-in phase – first 300 iterations with L = 10 and η = 0.4 for convergence at approximate vicinity of the typical set. 2. Equilibration phase – next 2270 iterations with L = 300 and η = 0.3 for relaxation and the typical set approach. 3. Sampling phase – last 30 iterations with L = 300 and η = 0.3 for ANN realization samples to be utilized in optimization problem solution. Learning curves are shown in Fig. 6. The noise model precisions β during first 600 iterations are shown to demonstrate the fast convergence during burn-in phase and rapid equilibration after increase of the number of leapfrog steps to L = 300.

Funding

National Research Foundation, Prime Minister's Office, Singapore under its Campus of Research Excellence and Technological Enterprise (CREATE) program.

References

1. A. M. Omer, "Energy, environment and sustainable development," Renew. Sustain. Energy Rev. **12**(9), 2265–2300 (2008).
2. C. M. Lampert, "Smart switchable glazing for solar energy and daylight control," Sol. Energy Mater. Sol. Cells **52**(3-4), 207–221 (1998).
3. M. E. A. Warwick and R. Binions, "Advances in thermochromic vanadium dioxide films," J. Mater. Chem. A Mater. Energy Sustain. **2**(10), 3275–3292 (2014).
4. E. E. Chain, "Optical properties of vanadium dioxide and vanadium pentoxide thin films," Appl. Opt. **30**(19), 2782–2787 (1991).
5. K. A. Khan and C. G. Granqvist, "Thermochromic sputter-deposited vanadium oxyfluoride coatings with low luminous absorptance," Appl. Phys. Lett. **55**(1), 4–6 (1989).
6. S.-Y. Li, G. A. Niklasson, and C. G. Granqvist, "Thermochromic fenestration with VO2-based materials: three challenges and how they can be met," Thin Solid Films **520**(10), 3823–3828 (2012).
7. S.-Y. Li, G. A. Niklasson, and C. G. Granqvist, "Thermochromic undoped and Mg-doped VO2 thin films and nanoparticles: optical properties and performance limits for energy efficient windows," J. Appl. Phys. **115**(5), 053513 (2014).
8. A. Taylor, I. Parkin, N. Noor, C. Tummeltshammer, M. S. Brown, and I. Papakonstantinou, "A bioinspired solution for spectrally selective thermochromic VO_2 coated intelligent glazing," Opt. Express **21**(S5 Suppl 5), A750–A764 (2013).
9. C. Liu, I. Balin, S. Magdassi, I. Abdulhalim, and Y. Long, "Vanadium dioxide nanogrid films for high transparency smart architectural window applications," Opt. Express **23**(3), A124–A132 (2015).
10. A. Taflove, "Application of the finite-difference time-domain method to sinusoidal steady-state electromagnetic penetration Problems," IEEE Trans. Electromagn. Compat. **EMC-22**(3), 191–202 (1980).
11. 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).
12. J. Jiang, J. Cai, G. P. Nordin, and L. Li, "Parallel microgenetic algorithm design for photonic crystal and waveguide structures," Opt. Lett. **28**(23), 2381–2383 (2003).
13. M. Djavid, S. A. Mirtaheri, and M. S. Abrishamian, "Photonic crystal notch-filter design using particle swarm optimization theory and finite-difference time-domain analysis," J. Opt. Soc. Am. B **26**(4), 849–853 (2009).
14. B. Saghirzadeh Darki and N. Granpayeh, "Improving the performance of a photonic crystal ring-resonator-based channel drop filter using particle swarm optimization method," Opt. Commun. **283**(20), 4099–4103 (2010).
15. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," Light Sci. Appl. **7**(1), 60 (2018).
16. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," ACS Photonics **5**(4), 1365–1369 (2018).
17. M. Kaya and S. Hajimirza, "Rapid optimization of external quantum efficiency of thin film solar cells using surrogate modeling of absorptivity," Sci. Rep. **8**(1), 8170 (2018).
18. T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," Opt. Express **26**(25), 32704–32717 (2018).

Optics EXPRESS

19. R. R. Andrawis, M. A. Swillam, M. A. El-Gamal, and E. A. Soliman, "Artificial neural network modeling of plasmonic transmission lines," Appl. Opt. **55**(10), 2780–2790 (2016).
20 J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks", Sci. Adv. **4**(6), eaar4206 (2018).
21. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," Optica **4**(11), 1437–1443 (2017).
22. Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light Sci. Appl. **7**(2), 17141 (2018).
23. Y. Gal and Z. Ghahramani, "Dropout as a Bayesian approximation: representing model uncertainty in deep learning," Proceedings ICML **33**(48), 1050–1059 (2016).
24. S. J. Wilson and M. C. Hutley, "The optical properties of 'moth eye' antireflection surfaces," Opt. Acta (Lond.) **29**(7), 993–1009 (1982).
25. "Solar Spectral Irradiance: Air Mass 1.5", Available: https://rredc.nrel.gov/solar//spectra/am1.5/.
26. T. Smith and J. Guild, "The C.I.E. colorimetric standards and their use," Trans. Opt. Soc. **33**(3), 73–134 (1931).
27. Y. Zhou, Y. Cai, X. Hu, and Y. Long, "VO_2/hydrogel hybrid nanothermochromic material with ultra-high solar modulation and luminous transmission," J. Mater. Chem. A Mater. Energy Sustain. **3**(3), 1121–1126 (2015).
28. T. Marwala, *Finite Element Model Updating Using Computational Intelligence Techniques: Applications to Structural Dynamics,* 1st ed. Springer Publishing Company, Incorporated, 2010.
29. R. M. Neal, *Bayesian Learning for Neural Networks*,New York: Springer-Verlag, 1996.
30. J. Lampinen and A. Vehtari, "Bayesian approach for neural networks--review and case studies," Neural Netw. **14**(3), 257–274 (2001).
31 P. Lalanne and D. Lemercier-Lalanne, "On the effective medium theory of subwavelength periodic structures," J. Mod. Opt. **43**(10), 2063–2085 (1996).
32. D. J. C. MacKay, *Information Theory, Inference & Learning Algorithms*. New York, USA: Cambridge University Press, 2002.