

# SCIENTIFIC REPORTS

OPEN

## Films based on crosslinked TEMPO-oxidized cellulose and predictive analysis via machine learning

Merve Özkan, Maryam Borghei, Alp Karakoç, Orlando J. Rojas  & Jouni Paltakari

We systematically investigated the effect of film-forming polyvinyl alcohol and crosslinkers, glyoxal and ammonium zirconium carbonate, on the optical and surface properties of films produced from TEMPO-oxidized cellulose nanofibers (TOCNFs). In this regard, UV-light transmittance, surface roughness and wetting behavior of the films were assessed. Optimization was carried out as a function of film composition following the “random forest” machine learning algorithm for regression analysis. As a result, the design of tailor-made TOCNF-based films can be achieved with reduced experimental expenditure. We envision this approach to be useful in facilitating adoption of TOCNF for the design of emerging flexible electronics, and related platforms.

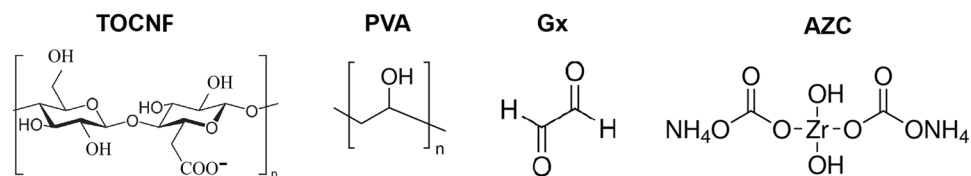
The increased demands of the future green bioeconomy have triggered efforts to substitute petroleum-based products with environmentally friendly alternatives<sup>1</sup>. In this aspect, nanocellulose based materials have shown promise for various applications, mainly owing to the possibility of obtaining transparent, flexible and recyclable materials<sup>2</sup>. However, any development needs demonstration of competitive performance relative to that of conventional substrates based on polyethylene terephthalate (PET) or polyethylene naphthalate (PEN)<sup>1,3</sup>. In this regard, films made of TEMPO-oxidized cellulose nanofibers (TOCNF) have promise, given their optical transparency achieved by the nanofibrils that carry carboxylic groups. Compared to unmodified cellulose nanofibers, TOCNF is more negative and displays a more extensive degree of fibrillation. Moreover, post treatments have been applied to further improve the properties of TOCNF films, for example, by impregnation in aqueous MgCl<sub>2</sub>, CaCl<sub>2</sub>, AlCl<sub>3</sub> and FeCl<sub>3</sub> electrolyte solutions, which have been found to increase water resistance via ion-exchange. Interfibrillar crosslinking via Mg<sup>2+</sup> and Al<sup>3+</sup> ions has been noted to exhibit 80% transmittance (600 nm) of thin films (15 ± 1 μm)<sup>4</sup>. Another application, the treatment with reactive sizing, via alkylketene dimer, reduced the hydrophilicity of TOCNF films without compromising their transparency. For instance, compared to the initial water contact angle of unmodified films (WCA ~47°), the WCA of AKD-modified TOCNF films increased significantly (~94°)<sup>5</sup>.

In this work, a film-forming biopolymer, polyvinyl alcohol (PVA), was combined with TOCNF networks, offering an alternative to the conventional cellulosic films in terms of biodegradability and stability<sup>6</sup>. PVA is available in different forms, depending on the degree of hydrolysis or deacetylation and molar mass and it is used in diverse applications ranging from paper coatings<sup>7</sup>, textile sizing<sup>8</sup>, agriculture engineering<sup>9</sup> to clinical research<sup>10</sup>. Recently, PVA was used as a reinforcing agent to produce water-stable films of TEMPO/NaClO<sub>2</sub> oxidized cellulosic nanofibrils<sup>11</sup> and chemical crosslinking was provided through esterification of COOH in TOCNF with hydroxyl groups of PVA.

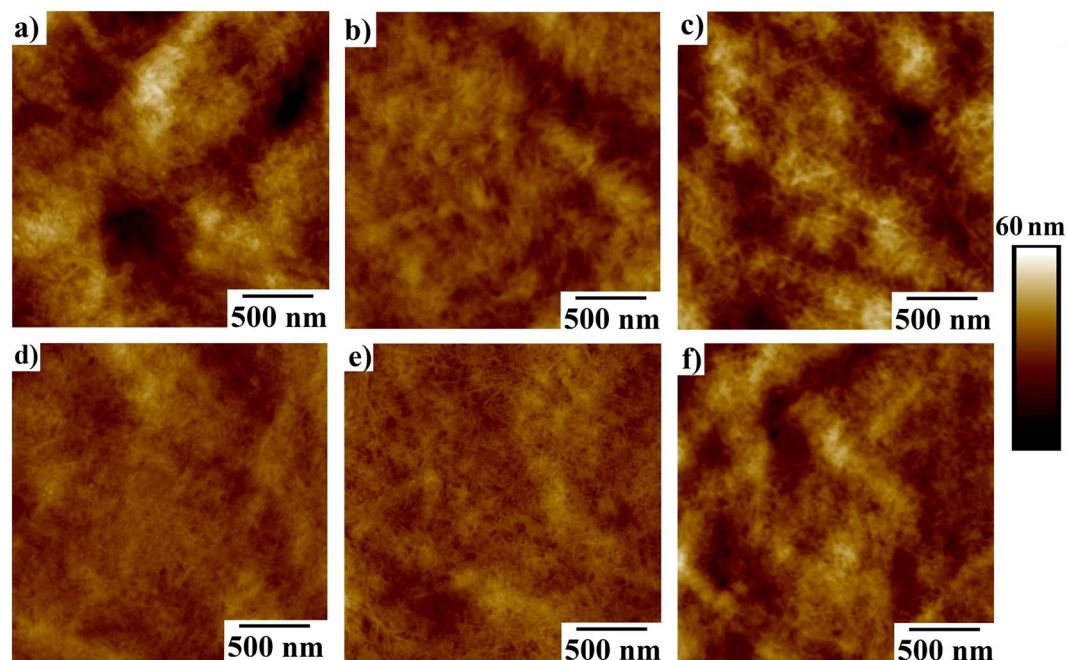
We investigate the influence of glyoxal (Gx) and ammonium zirconium carbonate (AZC), two common organic and inorganic crosslinkers, respectively, on the properties of TOCNF and composite TOCNF-PVA films. Glyoxal is widely used in textile, paper, cosmetic, health and leather industries as viscosity modifier, strength enhancer for paper, softener for less wrinkled textiles or biocide<sup>12</sup>. Ammonium zirconium carbonate (AZC), on the other hand, interacts with the hydroxyl and carboxyl groups of the fibrils and creates electrostatic charges via protonation in water. AZC is safe for food-packaging applications<sup>13</sup> and it is a crosslinker for PVA and cellulose in pulp and paper industries<sup>14</sup>. For instance, starch and cellulose, in the presence of AZC, form hydrogen bonding (fiber-starch, fiber-fiber, and starch-starch)<sup>15</sup>. The chemical structures of TEMPO-oxidized cellulose (TOCN), PVA, Gx and AZC are included in Fig. 1.

In this study, TOCNF, PVA, Gx and AZC were combined in different compositions to investigate their effects on the films in terms of surface and optical properties. Obtained data set was trained to predict other possible

Department of Bioproducts and Biosystems, School of Chemical Technology, Aalto University, Espoo, Finland. Correspondence and requests for materials should be addressed to M.Ö. (email: [merve.ozkan@aalto.fi](mailto:merve.ozkan@aalto.fi))



**Figure 1.** Basic chemical structure of TOCNF and PVA, Gx and AZC.



**Figure 2.** AFM images of (a) C, (b) CG10, (c) CA10, (d) CP10, (e) CP10G10 and (f) CP10A10 film surfaces.

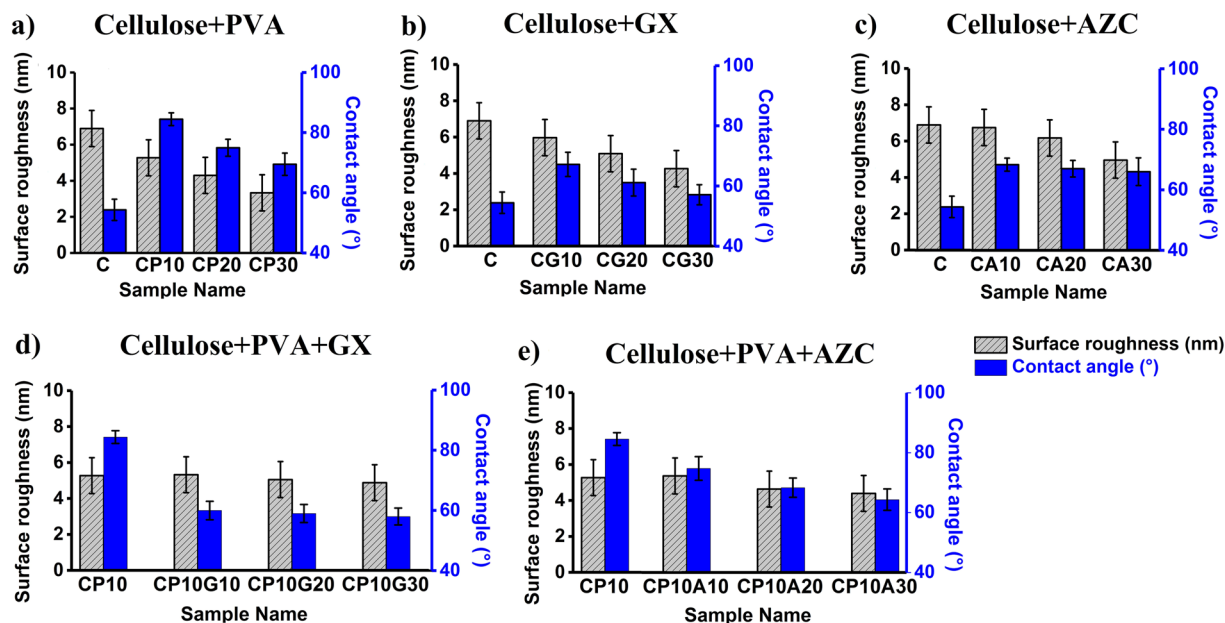
outcomes from generated samples by means of a machine learning algorithm. In particular, the random forest method was used given its power in data mining, identification, analysis and prediction. Especially, the method is popular among the researchers in the fields of finance, biology and chemistry due to its data handling efficiency and potential for capturing linear and non-linear interaction between input and output data<sup>16</sup>.

## Results

**Surface Analysis.** Modification of the film surface opens to door to create new alternative substrates for various applications, such as, printed electronics<sup>17</sup>. Films consisting of different additives (PVA, Gx and AZC) and TOCNF were introduced in Methods chapter. Surface AFM images of C, CG10, CA10, CP10, CP10G10 and CP10A10 films are shown in Fig. 2a–f, respectively. Films prepared with TOCNF reveal a randomly distributed nanofibrillar structure (Fig. 2a)<sup>5</sup>. The incorporation of AZC and Gx resulted in a more homogenous surface topology, Fig. 2b,c. Likewise, introduction of PVA (CP10) resulted in a more uniform network. One reason might be that the cellulose fibrils form a tighter structure<sup>18</sup>. It can be inferred that addition of Gx to the CP10 enhanced distribution of fibrils leading to a more uniform surface, while incorporation of AZC into the CP10 caused more aggregation and uneven surfaces. This might be attributed to the smaller molecular size of Gx compared to AZC, with higher ratio of carboxylic group to the total weight resulting into stronger chemical bonds with the hydroxyl groups of PVA and cellulose.

The water contact angle of the samples as well as the surface roughness values of the film surfaces are presented in Fig. 3. It can be seen that incorporation of PVA improved significantly the contact angle of TOCNF films, from 54° (C) to 85° (CP10) (Fig. 3a). In the literature, the WCA of TOCNF films have been reported to be around 50°<sup>5</sup> and its nanometer surface roughness was attributed to the compactness of the cellulose network, unlike traditional paper (5–10 μm)<sup>1</sup>. The contact angle of a water drop at the air-drop-solid interface depends on the topography and surface energy of the surface<sup>19</sup> as well as the interactions between the three phases. It can be inferred that the covalent bonding between the carboxyl groups of TOCNF and the hydroxyl groups in PVA, results in smoother surfaces, as observed in the AFM images. However, a lowered WCA of 70° was measured at PVA content of 30%, probably due to the excess unreacted OH that interacts with water.

Incorporation to TOCNF of 10% crosslinkers (Gx or AZC), improved the hydrophobicity but to a limited extent compared to PVA (Fig. 3a–c). The highest addition level of Gx, 30%, reduced the WCA, bringing it back to



**Figure 3.** Surface properties of films made of: (a) TOCNF and PVA, (b) TOCNF and Gx, (c) TOCNF and AZC, (d) TOCNF, PVA and Gx, (e) TOCNF, PVA and AZC.

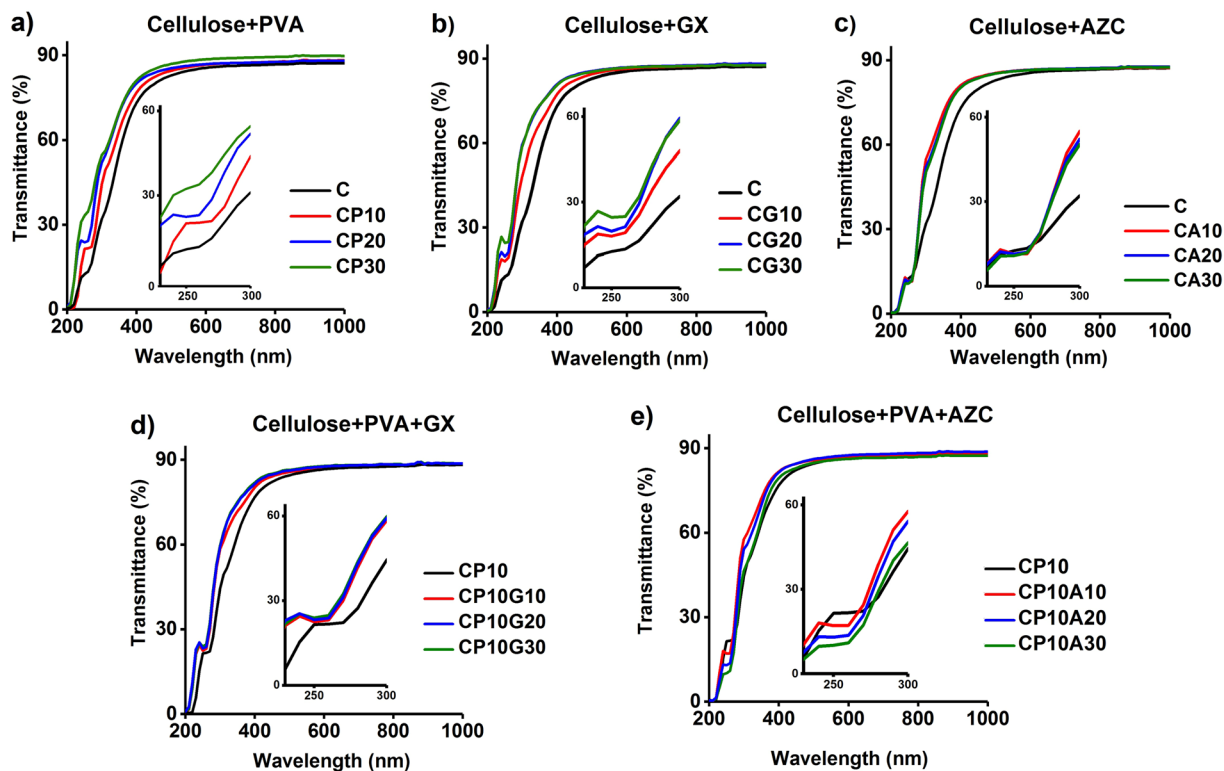
the value of pristine cellulose while the amount of AZC did not influence the WCA. Interestingly, when Gx was added to CP10, the contact angle was reduced drastically. Compared to Gx, AZC had a reduced impact on the hydrophobicity of CP10. It is proposed that the addition of PVA and crosslinkers mostly decreased the porosity of the samples, making it more resistant to water wetting. A contact angle of around  $102^\circ$  was observed for the PVA film due to a compact structure, strong hydrogen bonding between OH and the presence of few hydroxyl groups on the film surface (not presented in Fig. 3).

**Optical Properties.** Filling the pores and voids between the cellulose network as well as avoiding aggregation lead to reduced films porosity and thus less light scattering and higher transmittance<sup>1</sup>. In this study, films with thickness and density of  $40 \pm 5 \mu\text{m}$  and  $1.55 \pm 0.19 \text{ g/m}^3$  were obtained, respectively. Figure 4 indicates 84.6–88% transmission of the incident light at 550–1000 nm, which is in good agreement with literature<sup>5</sup>. Similar trends in the infrared region (700–1000 nm) as well as along the visible region (400–700 nm) were observed in the spectra of each sample. The incorporation of the different additives into the cellulose network altered mainly the characteristics of the films in the ultraviolet (UV) region (200–400 nm). An enhanced light transmittance of films in this region is of importance for some UV optoelectronic devices, such as light emitting diodes (LEDs)<sup>20</sup> and photodiodes<sup>21–23</sup>.

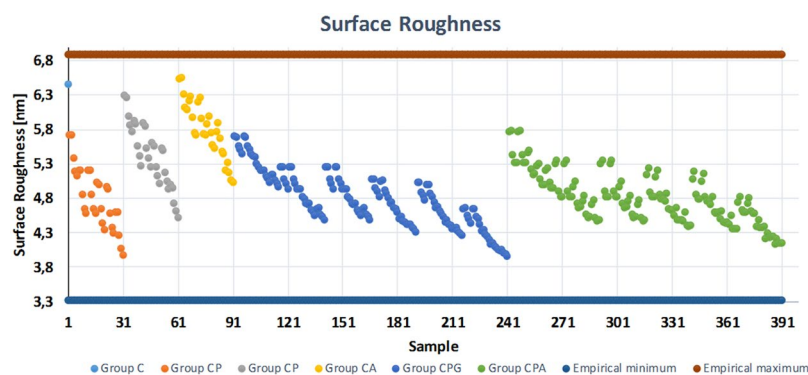
For pure TOCNF film, a high transmittance value (84.6%) was observed at 600 nm and the incorporation of the additives enhanced the transmittance, reaching 88% maximum transmittance for CP30 sample. On the other hand, the shoulder at around 250 nm increased by incorporation of PVA (Fig. 4a) and Gx (Fig. 4b), suggesting the presence of more aldehyde groups at C6 due to the esterification reaction<sup>2</sup>. In contrast, introducing AZC to TOCNF did not influence the shoulder intensity, but enhanced the transmittance at 250–500 nm (Fig. 4c). Interestingly, introducing AZC to TOCNF leads to a lower intensity of the absorbance shoulder compared to the Gx-crosslinked film (Fig. 4b,c). The effect of crosslinkers on the CP10 samples was slightly different compared to that for the pristine cellulose with one additive only. In the former case, Gx did not affect significantly the peak intensity or spectra whereas the peak intensity became slightly lower as AZC was added (Fig. 4d,e).

**Classification and Regression Analysis with the Random Forest Method.** Target physical properties of films are obtained usually after through tedious experimental work and analysis. A computational tool can possibly reduce such explorations by predicting the characteristics of hybrid films. In this study, prediction and regression analyses were carried out with 391 hypothetical inputs (films with different amounts and combination of TOCNF, PVA, Gx and AZC) as well as three outputs: surface roughness, contact angle and optical transmittance, as presented in vertical axes of Figs 5–7, respectively. The horizontal axes of these graphs consist of the inputs, which were sorted according to pre-defined rules as explained below (Rule 1 and Rule 2). By analyzing the experimental results and learning from the experimental data, possible substrate outcomes were predicted within the investigated/experimental data range by using the random forest method thus, extrapolation is beyond the scope of this work.

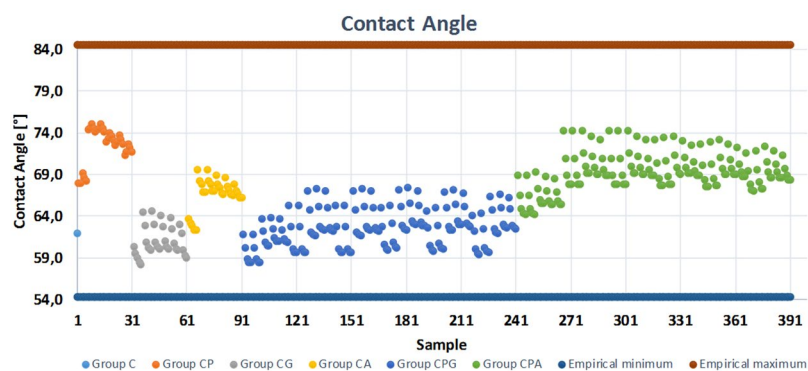
In each case, the experimental minimum and maximum values were used to constrain the predicted outputs. A neat cellulose film was selected as sample 1 and represented in the x-axes of the graphs correspondingly on the left most side. Samples 2–91 were assigned to the two-component systems (CP, CG and CA, in the presentation order along the x-axes) while samples 92–391 were designed as three-component system (CPG and CPA, respectively). Tables 1 and 2 explain the rule of two-component systems (Rule 1) and three-component systems (Rule 2) using CP and CPG categories as examples, respectively. Similarly, the rules can be applied to the rest of the input space.



**Figure 4.** Light transmittance of films made of: (a) TOCNF and PVA, (b) TOCNF and Gx, (c) TOCNF and AZC, (d) TOCNF, PVA and Gx, (e) TOCNF, PVA and AZC.



**Figure 5.** Surface roughness prediction of samples with different cellulose, PVA and crosslinker ratio.



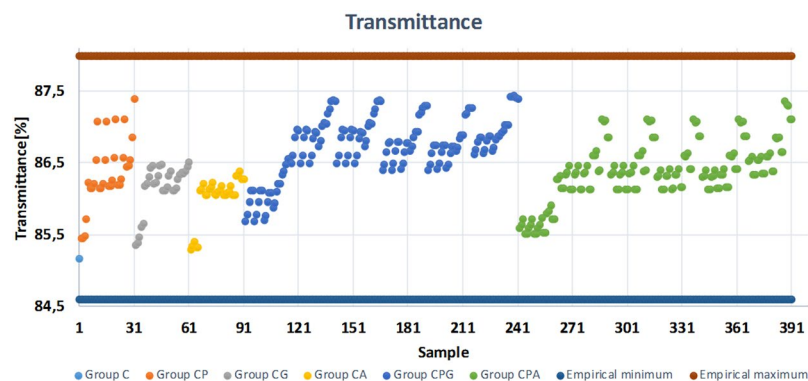
**Figure 6.** Contact angle prediction of samples with different cellulose, PVA and crosslinker ratio.

Sample	Sector	C	Pva	Rule 1
2–6	1	0.95	0.057–0.285	Input space of CP is divided into six sectors. In each sector, C is constant and it decreases as the number of sector increases. On the other hand, PVA increases from 0.057 to 0.285 in five steps inside a sector and the sector number increases when PVA reaches to 0.285.
7–11	2	0.893	0.057–0.285	
12–16	3	0.836	0.057–0.285	
17–21	4	0.779	0.057–0.285	
22–26	5	0.722	0.057–0.285	
27–31	6	0.665	0.057–0.285	

**Table 1.** Cellulose and PVA content (grams) in samples 2–31 (CP). Here, for CP sample set,  $G_x = AZC = 0$ . Rule 1 is also applied to CG and CA films presented in samples 32–61 and samples 62–91, respectively.

Sample	Sector	C	PVA	Gx	Rule 2
92–116	1	0.95	0.057–0.285	0.057–0.285	Input space is divided into six main sectors as in Rule 1. Differently, third component ( $G_x$ ) increases from 0.057 to 0.285 while the other two remains constant through five samples. When $G_x$ reaches to 0.285 PVA increases to the next value and $G_x$ starts from 0.057 again. Sector number increases when PVA reaches to 0.285.
117–141	2	0.893	0.057–0.285	0.057–0.285	
142–166	3	0.836	0.057–0.285	0.057–0.285	
167–191	4	0.779	0.057–0.285	0.057–0.285	
192–216	5	0.722	0.057–0.285	0.057–0.285	
217–241	6	0.665	0.057–0.285	0.057–0.285	

**Table 2.** Cellulose, PVA and  $G_x$  content (grams) in samples 92–241 (CPG). Here, for CPG sample set,  $AZC = 0$ . Rule 2 is also applied to CPA films presented in samples 242–391.

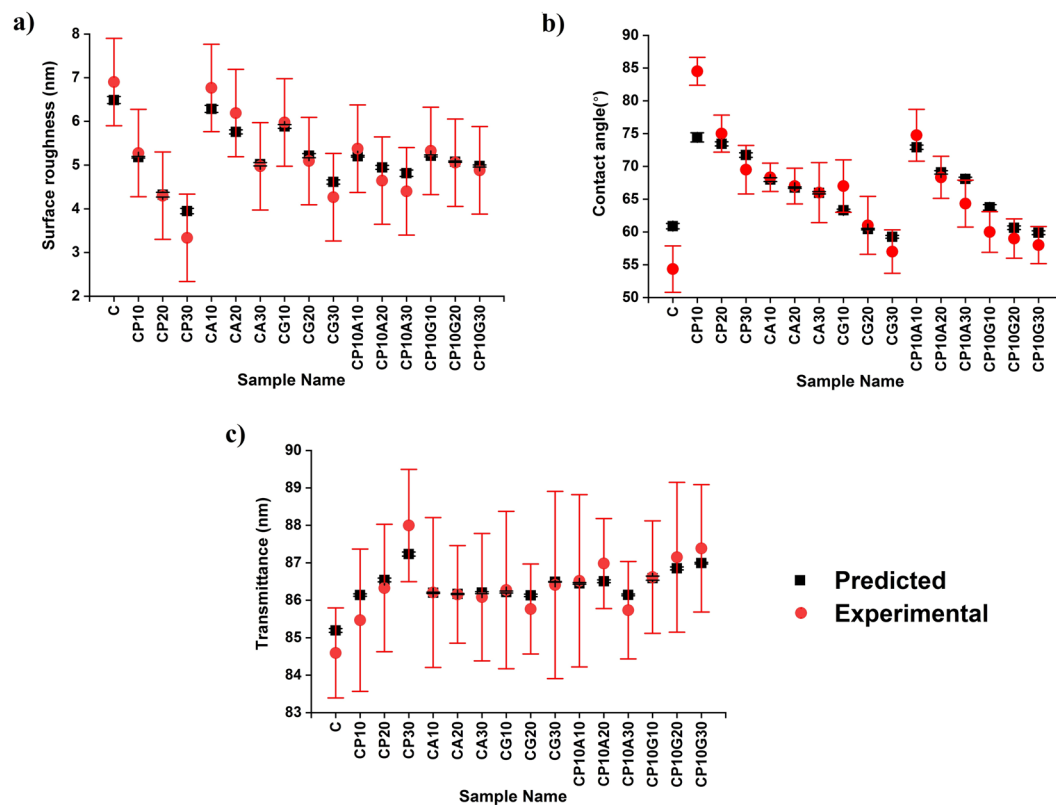


**Figure 7.** Light transmittance prediction of samples with different cellulose, PVA and crosslinker ratio.

**Accuracy of predictive model.** Given the mean absolute percentage error (MAPE) (Equation 2) in Methods section, the measure used for accuracy was calculated as 4.9, 3.9 and 0.3% for surface roughness, contact angle and transmittance outputs, respectively. The accuracy of the prediction suggests that the data training was handled with a relatively small prediction error<sup>24</sup>. On the other hand, it is well known that there is no definition of “precise” prediction in the literature<sup>25</sup>. MAPE of transmittance is almost 10 fold less compared to the other segments of data set with an apparent reason that the constraint boundaries of the data set were quite different (Figs 5–7). Here, the prediction model learns from the experimental results and predicts the corresponding outputs of not only the generated input values but also the existing input values. Therefore, comparing the experimental results with the predicted data can be used as another way to evaluate the accuracy of the prediction (Fig. 8). From Fig. 8, it can be seen that experimental data and predicted data follows similar trends for all three cases with minor differences. Here, the prediction variation of each data set was obtained by running the algorithm five times.

The surface smoothness of a substrate is an indicator of the surface porosity and compactness. According to the experiment data, the roughness of the films varies between 3.3 nm and 6.6 nm. The roughest surface was predicted to be belong to the film with  $(C, PVA, G_x, AZC) = (0.95, 0, 0, 0.114)$ , respectively. This seems logical because according to the experimental data, pristine cellulose and CA10 samples exhibited almost the same roughness values, which was around 35% more than the average of all samples measured (Fig. 8a). Additionally, 391 samples had an average surface roughness of  $4.9 \pm 0.5$  nm and the roughest surface values were found for the CA samples. On the other hand, both of the three-component systems exhibited statistically almost the same surface roughness regardless of the secondary additive,  $G_x$  or  $AZC$ .

The goniometric study of the surface with a drop determines the wetting capability of the films with the liquid<sup>26</sup>. In this regard, the proposed predictive tools can provide insights about the interactions before an actual



**Figure 8.** Experimental and predicted data: (a) surface roughness, (b) contact angle and (c) transmittance.

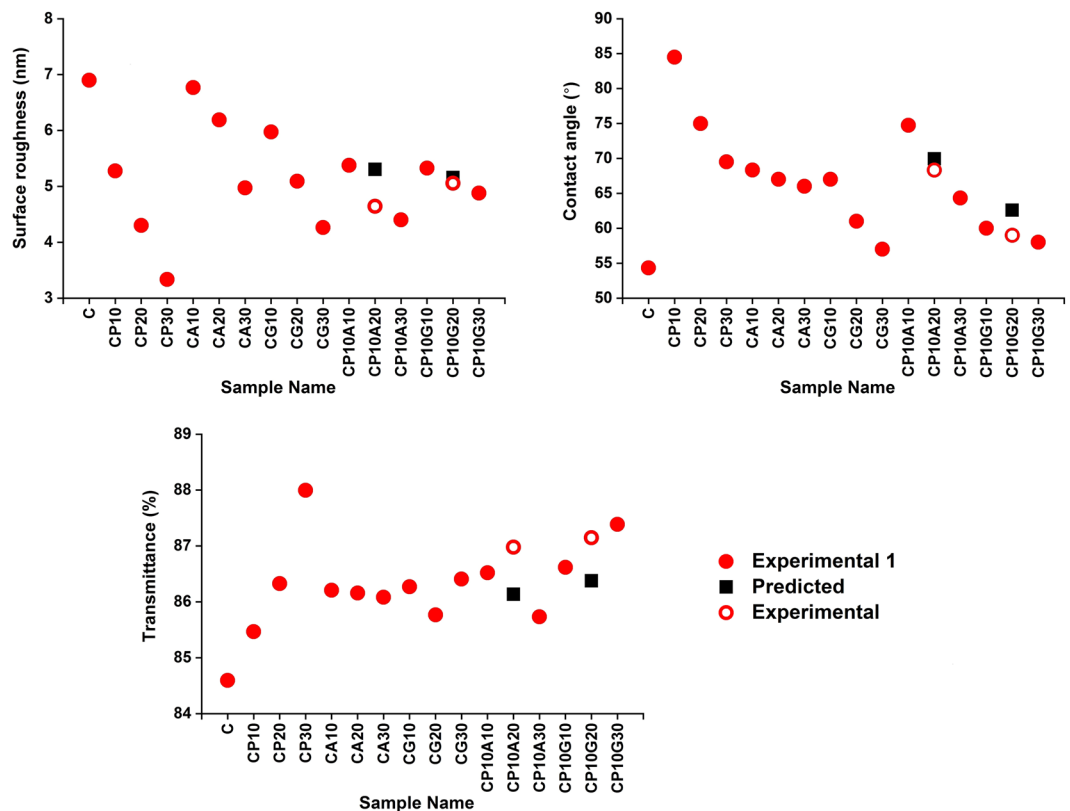
application step, for instance, a device integration process. The practical limits of the contact angle studies were between  $85^\circ$  and  $54^\circ$ . The prediction studies indicated that the highest water contact angle ( $75^\circ$ ) was observed on the films with  $(C, PVA, Gx, AZC) = (0.893, 0.171, 0, 0)$  or  $(C, PVA, Gx, AZC) = (0.836, 0.171, 0, 0)$ . In contrast, the experimental data showed that  $85^\circ$  can be obtained with CP10 having  $(C, PVA, Gx, AZC) = (0.855, 0.095, 0, 0)$ . Additionally, the lower marginal was found to be  $58^\circ$  predicted to be in the sample with  $(C, PVA, Gx, AZC) = (0.95, 0, 0.285, 0)$  whereas the neat cellulose had a contact angle of  $54^\circ$  (Fig. 8b).

The prediction analysis of the transmittance shows that the most transparent (87.5% transmittance) and opaque (85.2% transmittance) films can be obtained using  $(C, PVA, Gx, AZC) = (0.665, 0.285, 0.171, 0)$  and  $(C, PVA, Gx, AZC) = (0.95, 0, 0, 0)$ , respectively (Fig. 8c). Such a prediction is in line with the analysis carried out on the experimental results since the scattering of the cellulose network is reduced by filling the voids with the polymer and Gx, which can promote a network with less voids.

The accuracy of the method was also tested by using fewer inputs. In this case, experimental results of CP10A20 and CP10G20 were excluded and the output prediction was conducted without the inputs, CP10A20 and CP10G20 (“Experiment 1”, in Fig. 9). Thereafter, the predicted values (“Predicted”, in Fig. 9) were compared to the real outputs of CP10A20 and CP10G20 (“Experimental”, in Fig. 9). As a result, surface roughness, contact angle and transmittance of CP10A20 and CP10G20 samples were predicted with 14, 2, 1% and 2, 6, 1% errors (error =  $100 \times \text{difference between experimental value and predicted value} / \text{experimental value}$ ) for CP10A20 and CP10G20 samples, respectively. The error in this case was bigger than the previous assessment, albeit staying still in favorable amount<sup>24</sup>. Note that in order to increase experimental precision, the following procedure was followed: (1) All the samples were prepared, characteristics were measured and the experimental data of 16 types of films were obtained. (2) Data of two samples (~13% of the whole data set) was excluded from the training set. (3) Data was trained and the corresponding outputs were predicted.

## Discussions

We demonstrated an example of the convergence of information technology and materials science to achieve innovative developments based on nanocellulose films<sup>27</sup>. Hybrid films made of TEMPO-oxidized nanocellulose fibers (TOCNF), a film-forming polymer (PVA) as well as crosslinkers (glyoxal, Gx and ammonium zirconium carbonate, AZC) were prepared. Optical transmittance and surface properties of the hybrid films were investigated by means of UV-vis spectroscopy and atomic force microscopy and contact angle analysis. Depending on the material composition, given optical transmittance and surface properties were achieved. The highest transmittance was obtained blending 30 wt.% PVA, which resulted in very smooth surfaces while the highest water contact angle was obtained blending only 10 wt.% PVA into the cellulose dispersion. Aiming at the design of custom-made substrates for different applications, a “Random Forest” machine learning method was utilized to train the data by using the experimental results. The surface roughness, contact angle and transmittance properties predicted for the films containing



**Figure 9.** Experimental and predicted data: (a) surface roughness, (b) contact angle and (c) transmittance.

different components were predicted with 4.5, 3.9 and 0.3% MAPE, respectively. The introduction of such a tool for the design of hybrid films can minimize time and costs usually involved in experimental optimization. On the other hand, the prediction performance should be further evaluated with a comparative study with different methods.

## Methods

A fully hydrolyzed, water-dispersible polyvinyl alcohol (PVA, degree of hydrolysis = 99.0–99.8) with molecular weight of ~145,000 (Mowiol 28–99) was obtained from Sigma-Aldrich. Aqueous solutions of glyoxal (40 wt.%) and ammonium zirconium (IV) carbonate were also obtained from Sigma-Aldrich. Commonly, all solutions were used in 0.5% dry weight.

**Film Preparation.** 2,2,6,6-tetramethylpiperidine-1-oxyl radical (TEMPO) oxidized cellulose nanofibrils were obtained at room temperature and pH 10<sup>2.5</sup>. After the mechanical disintegration in a micro fluidizer (Microfluidics M-110EH-30, Microfluidics Int., USA), 0.5 wt.% aqueous dispersion was obtained after 24 h magnetic stirring followed by centrifugation at 3000 g for 6 min. All samples were prepared from the supernatant dispersion and after mixing with PVA and/or crosslinker 75 minutes before film casting. Films made of pristine TOCNF were used as reference and, referred to as “C”. The effect of PVA on film properties was studied by addition of 10, 20 and 30 wt.% PVA to the TOCNF based on the weight of hybrid films (CP10, CP20 and CP30, respectively). For the crosslinker, 10, 20 and 30 wt.% Gx and AZC were added separately, resulting in films labeled as CG10, CG20, CG30 (for Gx) as well as CA10, CA20 and CA30 (for AZC). The effect of crosslinkers together with PVA was also investigated by applying crosslinkers to 10, 20 and 30 wt.% CP10 solution. The resulting samples are named as CP10G10, CP10G20 and CP10G30 (for Gx) and CP10A10, CP10A20 and CP10A30 (for AZC). Table 3 summarizes the sample compositions and nomenclature. For all three additives, 30% was found to be the approximate upper limit before the dispersion starts to contain bundles after preparing the hybrid dispersion.

The dispersions were casted in polypropylene Petri dishes (13.5 cm diameter) and dried in room condition. A filter paper (Schleicher & Schuell, 240 nm) was kept on the top of the petri dish in order to maintain a uniform vapor pressure during drying, resulting in a uniform film thickness. The average thickness values were calculated by taking five measurement points from each sample (one at the center and four at the edges of the samples).

**Characterizations.** The transmittance profiles of the films along 200 nm and 1000 nm wavelength frame were identified with a Perkin Elmer lambda 950 UV-Vis spectrometer. A CAM 200 optical goniometer from KSV instruments with a camera and dispenser system were used to monitor the sessile drop on the surface of the films at 50% relative humidity and 23 °C. The initial contact angle of the drop was calculated in accordance with Young’s equation by embedded image processing program. The surface morphologies of the samples were studied with an atomic force microscope (AFM), Multimode 8, including a NanoScope V controller (Bruker Corporation,

Sample Name	C (g)	PVA (g)	Gx (g)	AZC (g)
C	0.95	0	0	0
CP10	0.855	0.095	0	0
CP20	0.76	0.19	0	0
CP30	0.665	0.285	0	0
CG10	0.855	0	0.095	0
CG20	0.76	0	0.19	0
CG30	0.665	0	0.285	0
CA10	0.855	0	0	0.095
CA20	0.76	0	0	0.19
CA30	0.665	0	0	0.285
CP10G10	0.855	0.095	0.0855	0
CP10G20	0.855	0.095	0.171	0
CP10G30	0.855	0.095	0.2565	0
CP10A10	0.855	0.095	0	0.0855
CP10A20	0.855	0.095	0	0.171
CP10A30	0.855	0.095	0	0.2565

**Table 3.** Composition of the films made of TOCNF (C), PVA and crosslinkers (Gx and AZC) at different compositions (dry weight in gram).

Billerica, MA), operating in tapping mode. A NanoScope 8.15 software (Bruker) was used to process obtained images and calculate the surface roughness values. No image processing (except flattening) was performed during this step.

**Random Forest Method.** The random forest method, or more specifically Breiman–Cutler ensembles of decision trees, was used as a machine learning technique for classification and regression analysis. This approach is based on decision combination model  $g$  obtained from a sequence of models  $f$ . It can be broadly represented as:

$$g(x_1, \dots, x_n) = \sum_j f_j(x_1, \dots, x_n) \quad (1)$$

This technique uses several models to obtain predictive performance, known as model ensemble, where each  $f$  is called as a decision tree<sup>28</sup>. It is noteworthy that the models are independently formed through individual data sampling. Thus, the random forest model is efficient for numerical data handling and capturing linear and non-linear interactions between the input and output data<sup>16</sup>. Given that a priori assumption in statistical processes for establishing relationship between input and output parameters, i.e. regression analysis, is necessary, the random forest model is preferred in the current study. In the present study, for regression analysis and predictions, built-in predictor function, Predict[] with “RandomForest” method setting provided in Mathematica technical computing software, is used. The random forest method can deal with various types of variables in one data set<sup>29</sup>. On the other hand, in big data sets, the prediction accuracy drops were reported, therefore, various modifications were proposed<sup>30,31</sup>.

After applying the random forest model as a prediction method, the mean absolute percentage error (MAPE) was calculated from:

$$M = \frac{100}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (2)$$

where,  $A_t$  is the actual value and  $F_t$  is the forecast value of data set consisting of  $n$  samples<sup>32</sup>. MAPE was selected due to its simplicity and because of the fact that it applies to data sets that do not contain close to zero or negative values<sup>33</sup>.

**Data availability.** The authors declare that the data supporting the findings of this study are available within the article and Mathematica technical computing software is available in <https://www.wolfram.com/mathematica/>.

## References

- Zhu, H., Fang, Z., Preston, C., Li, Y. & Hu, L. Transparent paper: fabrications, properties, and device applications. *Energy Environ. Sci.* **7**, 269–287 (2014).
- Isogai, A., Saito, T. & Fukuzumi, H. TEMPO-oxidized cellulose nanofibers. *Nanoscale* **3**, 71–85 (2011).
- Fujisawa, S., Ikeuchi, T., Takeuchi, M., Saito, T. & Isogai, A. Superior reinforcement effect of TEMPO-oxidized cellulose nanofibrils in polystyrene matrix: optical, thermal, and mechanical studies. *Biomacromolecules* **13**, 2188–2194 (2012).
- Shimizu, M., Saito, T. & Isogai, A. Water-resistant and high oxygen-barrier nanocellulose films with interfibrillar cross-linkages formed through multivalent metal ions. *Journal of Membrane Science* **500**, 1–7 (2016).
- Fukuzumi, H., Saito, T., Iwata, T., Kumamoto, Y. & Isogai, A. Transparent and High Gas Barrier Films of Cellulose Nanofibers Prepared by TEMPO-Mediated Oxidation. *Biomacromolecules* **10**, 162–165 (2008).



6. Kalarikkal, N., Augustine, R., Oluwafemi, O. S., Joshy, K. S., & Thomas, S. *Nanomedicine and Tissue Engineering: State of the Art and Recent Trends Vol. 1* (eds Mohanan, D. P. et al.) Ch. 11, 365–390 (CRC Press, 2016).
7. Bock, E. F., Burdick, C. L. & Walsh, L. D. U.S. Patent No. 6,030,443. Washington, D C: U.S. Patent and Trademark Office (2000).
8. Song, T., Tanpichai, S. & Oksman, K. Cross-linked polyvinyl alcohol (PVA) foams reinforced with cellulose nanocrystals (CNCs). *Cellulose* **23**, 1925–1938 (2016).
9. Seo, J. K. et al. Nitrification performance of nitrifiers immobilized in PVA (polyvinyl alcohol) for a marine recirculating aquarium system. *Aquacultural Engineering* **24**, 181–194 (2001).
10. Kobayashi, M., Chang, Y. S. & Oka, M. A two year *in vivo* study of polyvinyl alcohol-hydrogel (PVA-H) artificial meniscus. *Biomaterials* **26**, 3243–3248 (2005).
11. Hakalahti, M., Salminen, A., Seppälä, J., Tammelin, T. & Hänninen, T. Effect of interfibrillar PVA bridging on water stability and mechanical properties of TEMPO/NaClO<sub>2</sub> oxidized cellulosic nanofibril films. *Carbohydrate polymers* **126**, 78–82 (2015).
12. Zhang, Y., Zhu, P. C. & Edgren, D. Crosslinking reaction of poly (vinyl alcohol) with glyoxal. *Journal of Polymer Research* **17**, 725–730 (2010).
13. Chen, X. F., Ren, J. L. & Meng, L. Influence of ammonium zirconium carbonate on properties of poly (vinyl alcohol)/xylan composite films. *Journal of Nanomaterials* **16**, 1–8 (2015).
14. Mikkonen, K. S., Schmidt, J., Vesterinen, A. H. & Tenkanen, M. Crosslinking with ammonium zirconium carbonate improves the formation and properties of spruce galactoglucomannan films. *Journal of Materials Science* **48**, 4205–4213 (2013).
15. Song, D., Zhao, Y., Dong, C. & Deng, Y. Surface modification of cellulose fibers by starch grafting with crosslinkers. *Journal of Applied Polymer Science* **113**, 3019–3026 (2009).
16. Yang, M., Xu, H., Zhu, D. & Chen, H. Visualizing the random forest by 3D techniques. *Internet of Things* **312**, 639–645 (2012).
17. Özkan, M. et al. Inkjet-printed platinum counter electrodes for dye-sensitized solar cells. *Organic Electronics* **44**, 159–167 (2017).
18. López-Suevos, F., Eyholzer, C., Bordeanu, N. & Richter, K. DMA analysis and wood bonding of PVAc latex reinforced with cellulose nanofibrils. *Cellulose* **17**, 387–398 (2010).
19. Wu, C. N., Saito, T., Yang, Q., Fukuzumi, H. & Isogai, A. Increase in the water contact angle of composite film surfaces caused by the assembly of hydrophilic nanocellulose fibrils and nano clay platelets. *ACS Applied Materials & Interfaces* **6**, 12707–12712 (2014).
20. Adivarahan, V. et al. Sub-milliwatt power III-N light emitting diodes at 285 nm. *Japanese Journal of Applied Physics* **41**, 435–436 (2002).
21. Brown, J. D. et al. UV-specific (320–365 nm) digital camera based on a 128 × 128 focal plane array of GaN/AlGaIn p-i-n photodiodes. *Mat. Research Society Internet Journal of Nitride Semiconductor Research* **5**(6), 1–12 (2000).
22. Brown, J. D., Li, J., Srinivasan, P., Matthews, J. & Schetzina, J. F. Solar-blind AlGaIn heterostructure photodiodes. *Materials Research Society Internet Journal of Nitride Semiconductor Research* **5**(9), 1–2 (2000).
23. Zhu, T. G., Chowdhury, U., Denyszyn, J. C., Wong, M. M. & Dupuis, R. D. AlGaIn/AlGaIn UV light-emitting diodes grown on sapphire by metalorganic chemical vapor deposition. *Journal of Crystal Growth* **248**, 548–551 (2003).
24. Antipov, E. A. & Pokryshevskaya, E. B. Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications* **39**, 1772–1778 (2012).
25. Swanson, D. A., Tayman, J. & Bryan, T. M. MAPE-R: a rescaled measure of accuracy for cross-sectional subnational population forecasts. *Journal of Population Research* **28**, 225–243 (2011).
26. Özkan, M. et al. Highly transparent cellulose substrate for dye sensitized solar cells. *Proc. 32nd European Photovoltaic Solar Energy Conference and Exhibition, 3DV.2.56*, 1353–1356 (2016).
27. Dobbs, R., Ramaswamy, S., Stephenson, E. & Viguier, S. P. Management intuition for the next 50 years. *McKinsey Quarterly* **1**, 1–13 (2014).
28. Breiman, L. Random forests. *Machine Learning* **45**, 5–32 (2001).
29. Qin, Y., Li, S., Vu, T. T., Niu, Z. & Ban, Y. Synergistic application of geometric and radiometric features of LiDAR data for urban land cover mapping. *Optics express* **23**, 13761–13775 (2015).
30. Liu, Y. Random forest algorithm in big data environment. *Computer Modelling & New Technologies* **18**, 147–51 (2014).
31. Xie, Y., Li, X., Ngai, E. W. T. & Ying, W. Customer churn prediction using improved balanced random forests. *Expert Systems with Applications* **36**, 5445–5449 (2009).
32. Hyndman, R. J. & Koehler, A. B. Another look at measures of forecast accuracy. *Int. Journal of Forecasting* **22**, 679–688 (2006).
33. O'Connell, R. T. & Koehler, A. B. Forecasting, time series, and regression: *An applied approach Vol. 4, Ch. 1, 18–19* (South-Western College Pub, 2005).

## Acknowledgements

This work is supported in part by the Finnish Cultural Foundation. MB would like to acknowledge Academy of Finland Centre of Excellence (Hyber).

## Author Contributions

M.O., M.B. and A.K. designed and executed the experiments and wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018