A REGION-BASED MACHINE LEARNING APPROACH FOR SELF-DIAGNOSIS OF A 4D DIGITAL THERMAL TWIN

E. Strauss*, D. Bulatov

Fraunhofer IOSB, Gutleuthausstraße 1, 76275 Ettlingen, Germany
eva.strauss@iosb.fraunhofer.de

Commission IV, WG IV/9

KEY WORDS: Digital Twin, Thermal Imagery, Thermal Simulation, Classification

ABSTRACT:

In this paper, we explore the applicability of machine learning as a tool for self-diagnosis of 4D digital twins with a focus on simulated surface temperatures. Generation of digital twins involves abstractions, simplification, and the closed-world assumption. Hence, performing thermal simulation in order to obtain surface temperatures involves not only mathematical modeling of the physical phenomena, but also temporal uncertainties on external conditions. To identify the types of simulation inaccuracies, our proposed method is based on thermal image comparison, i.e. the corresponding measured thermal image and the simulated thermal image resulting from the 4D digital twin. First, a statistical necessary condition is defined to obtain regions of interest in the simulated image. Second, after manual labeling of these regions into the two inaccuracy classes, we conduct a detailed feature analysis and subsequently train our Random Forest classifier. The results show a good separability of the two classes despite the limited training data, allowing to achieve values of overall accuracy around 93.5%.

1. INTRODUCTION

1.1 Motivation

Digital Twins are virtual representations of the past, present, and sometimes even future state of the real-world objects, as well as processes, relationships, and behaviors between them. In the world of geo-informatics, digital twins not only provide a better perception of the scene, but are increasingly supporting urban planners and decision-makers in exploring the built environment (Yan et al., 2019). The idea here is to help understand the performance of the model, which allows to save costs and improve the quality of life of people that are supposed to inhabit these systems. The digital twin concept builds on older techniques and technologies for object-based data models by adding realism, interactive user experience, and high-resolution 3D and 4D models of underlying systems. For urban representation, the physical form of the city is represented in 3D models, providing a graphic visualization of the city's physical elements (Shahat et al., 2021). This means: land-cover classification, detection and geometric reconstruction of buildings, trees, and other important types of objects (Bulatov et al., 2014). At the same time, adding weather information, it is possible to obtain a prediction for the surface temperature at any given time and any given spot, replacing cost-intensive and time-consuming punctual or aerial temperature measurements, and possibly providing more detailed resolutions (Bulatov et al., 2020). This is an example of a 4D digital thermal twin, which could be used for example in urban monitoring and planning for analysis of urban heat islands. However, especially in the case of necessity to collect information from noisy, multi-source sensor data, the process of creation of a 4D digital thermal twin is challenging. Likely, we can classify the error sources in the numeric output provided by such a digital twin into two main types. The first is proper to the three-dimensional scene reconstruction. Sometimes, there are inaccuracies arising from simplifications and abstractions (Chong et al., 2021), and sometimes, the closed world assumption is violated, which means that in the real scene, there are classes additional to those we supposed to be present during modeling. The second type of errors comes from the temporal component. Again according to (Chong et al., 2021), we can count to this category the scenario uncertainty arising from weather or building occupancy conditions, but also the imperfection of simulation models, which inherit probabilistic noise or numerical inaccuracies.

Generally, large thermal inaccuracies might be easily identifiable by a human observer, while subtle deviations are hidden to the human eye, though not less important. For example, slight differences in assumed and actual weather condition on the one hand, and missed buildings, on the other hand, can equally lead to inaccuracies of the same order of magnitude, though evolving from two separate causes, as outlined above. Both of these types in subtle inaccuracies are important for the simulation evaluation. The former, rising from temporal uncertainties and here referred to as the *anomalous* type, appears hard to control. The latter, stemming from the controllable sources and in the following referred to as *explainable*, directly provides concrete points of action for improvement of the 4D digital twin. This is why the ability to identify and classify these two types of errors is an important tool for self-diagnosis.

1.2 Related Works

When it comes to evaluation and analysis of 4D Digital Twins, the consecutive steps of their generation are usually considered separately. Corresponding validation metrics may assemble ground truth data and draw conclusions about the precision of, for example, the land-cover classification (Budde et al., 2021, Häufel et al., 2019), vectorization and meshing respectively (Lafarge and Mallet, 2012, Blaha et al., 2016), or surface temperatures. These can generally be evaluated in two ways: by comparing selected points in the 4D digital twin with punc-

* Corresponding author

tual temperature measurements in the real world, which generally are time-consuming and probably cost-intensive, or by comparison to large-scale thermal imagery. For the latter, the 4D digital twin has obviously to be brought into proper format. (Xiong et al., 2016) evaluated their thermal simulation method against punctual recordings of several thermistors, but also performed a basic approach to thermal infrared rendering from a selected perspective and a subsequent qualitative comparison to real thermal images. (Kottler et al., 2019) proceeded similarly, but performed an orthogonal projection of the 4D digital twin instead of infrared rendering to compare the relative temperature distribution to an aerial thermal image. (Guo et al., 2018) merely focused on thermal imagery comparison and used a more sophisticated approach to thermal infrared rendering by taking into account environmental conditions. (Bulatov et al., 2020) pursued both evaluation methods, yet applied an even more sophisticated approach for their large-scale evaluation. They chose a physics-based model of the thermal radiance, which is detected by the imaging sensor, and carried out the thermal infrared rendering by Monte Carlo ray-tracing. At this point, all approaches have in common that there is a strong focus on large simulation inaccuracies and no conclusions on how the two type of errors, as mentioned above, have impact on the overall simulation accuracy, i.e. there is no potential for feedback on the simulation itself. Finally, self-diagnosis routines of digital twins are already being used in civil engineering, medical research and smart manufacturing. According to (Xu et al., 2019), they are necessary for successful cooperation between real and virtual entities and self-adjustment, forecasting performance evolution, and self-adaption to new working conditions. Referring to as intelligent development phase, a self-diagnosis module based on artificial intelligence is trained using digital twin output and the real world data. The appropriate repairs and redesigns can be implemented and, at the same time, error diagnosis models embedded in the digital twin will allow improving its performance and predicting actual and future failures. In remote sensing, however, error diagnostics of digital twins have not yet found entry to such a broad extent. For example, flood risk evaluation system of (Lyu et al., 2019) is supposed to effectuate early warnings of inundation risks in urban systems. However, due to the low frequency of the disasters' occurrence, the reliability of this digital twin is hard to be proved in practice.

1.3 Contribution

In this paper, we wish to explore how machine learning can contribute to the correct classification of inaccuracies in the 4D digital thermal twin focusing on the simulated surface temperatures itself. A necessary condition is when these simulated temperatures deviate from the measured temperatures, whereby the metric to measure subtle deviations must be defined properly. For further classification, a sufficient condition is required. This is where machine learning comes into play. Because of our necessary condition, the number of training examples is limited. Thus, we have to choose our features deliberately and apply a classifier that is less sensitive to overfitting.

In the following, Section 2 gives details about the 4D digital thermal twin that is evaluated by our proposed method presented in Section 3. In Section 4, a presentation of our data generation process together with a feature analysis is followed by the main classification results. Our conclusion is summarized is Section 5.

2. PRELIMINARIES

The generation of the 4D digital thermal twin involves, in the first instance, the 3D digital twin generation which is enriched with semantic and material information. Subsequently, the numerical calculation of surface temperatures has to be carried out. Thirdly and lastly, to compare the simulation to an aerial thermal infrared (TIR) image, just like we aim to, the 4D digital twin has to be brought into a proper format comparable to measured aerial surface temperatures. To carry out these steps, we follow the procedure of (Bulatov et al., 2020). Since our main focus lies on the thermal simulation and its deviations from reality, we briefly summarize the chosen mathematical model for surface temperature estimation. Given the three-dimensional mesh of the 4D digital twin, a surface temperature per triangle is calculated. Conductive, convective and radiative heat transfer are involved. Given a finite virtual depth d of the surface element, the temporal-spatial heat transfer for each triangle can be described as

$$c_v \rho d \frac{\partial T}{\partial t} = S + R + I + A \tag{1}$$

where
c_v = specific heat capacity at constant volume
ρ = material density
d = virtual triangle depth
T = temperature
t = time
S = short-wave radiative heat
R = long-wave radiative heat
A = convective heat
I = conductive heat

Material related parameters, namely c_v, ρ, a (solar albedo) within the radiative term S, ϵ (thermal emissivity) within R, and k (thermal conductivity) within I, may differ between triangles as they are read from a corresponding database depending on the assigned material label from the preceding land cover and roof material classification. The short-wave radiative term S reflects the terrestrial heating due to solar radiation, cp. (Duffie and Beckman, 2013), while the long-wave radiative heat R models the heat exchange with the surrounding atmosphere (Lienhard IV and Lienhard V, 2019). Both depend on weather-related parameters, such as cloud coverage, air temperature or relative humidity, which are picked from weather servers. The convective heat A is formulated as a very simple empirical and linear model depending on two heat transfer coefficients, for natural and forced convection respectively (Groth, 2009), and the wind velocity assumed to be constant over the whole 4D digital twin. The conductive heat term I is, in our case, reduced to one dimension for orthogonal heat conduction following (Malaplate et al., 2007), i.e. heat transferred from the surface into the ground or object. Further physical phenomena such as lateral heat conduction along the surface, heat transmission and latent heat are not considered as they probably have low impact on the simulation accuracy while unreasonably increasing the simulations' runtime. Finally, Equation (1) is an initial value problem, which can be solved using Euler method. As initial temperature, we chose the inner boundary layer temperature which is material-dependent and thus read from the material database. After initally simulating a 3-days period to overcome heat inertia, the simulation was continued for another period of 24 hours to reach the point in time corresponding to the measured thermal image.

3. METHODOLOGY

Our proposed method concerns two steps implying firstly regions-of-interest (ROI) extraction by a statistical measure and secondly classification of ROI into explainable and anomalous by a machine learning approach. To do so, some pre-processing of the simulation result as described in Section 2 has to be carried out.

3.1 Preprocessing

Considering the last step of the thermal image generation as proposed in (Bulatov et al., 2020), we rather do not rely on thermal radiance rendering but perform an orthogonal projection of the 4D digital twin, resulting in a top-down view of the surface temperatures. This simulated thermal image displays these temperatures with sharp edges, in contrast to the actual measured image. Therefore, we apply Gaussian blur. Eventually, the well-known behavior of metallic surfaces in aerial measured thermal images, i.e. non-contact temperature measurements, is taken care of. Such metallic surfaces exhibit a low thermal emissivity, leading to low radiation of heat. Thermal infrared sensors, however, measure this radiation as reference for the surface temperature under the assumption of very high emissivity, thus yielding a falsely, very low temperature estimation for metallic surfaces. Those temperatures will appear as strong outliers on a measured thermal image. To address this issue in our case, the temperature distribution of the measured image was examined, following approximately a Gaussian distribution. Thereof we deduce and apply an effective cut-off of temperatures below 20 degree Celsius. With this, obvious outliers referring to metallic surfaces are suppressed. Finally, the simulated thermal image is now adequately comparable to the given measured aerial image.

3.2 Regions-of-Interest determination

To determine the ROIs, a proper condition needs to be carefully defined. Aiming for inaccuracies identification, pixel-wise differences of the simulated and measured thermal image are determined first. Simply thresholding the absolute deviations at this point would exclude the sought-after subtle inaccuracies which appear within the tails of the distribution of the image differences ΔT between T in (1) and the measured thermal image. Therefore, we follow (Burkard et al., 2020) by determining the corresponding mean value μ and standard deviation σ of ΔT and thereof generating a binary mask following the Gaussian outlier detection. For each pixel p of the difference image, we check whether

$$|\Delta T(p) - \mu| > 3\sigma. \tag{2}$$

Connected pixels fulfilling this condition are summarized to continuous ROIs. Figure 1 exemplarily displays the mean-substracted difference image by the left-hand side of equation 2 and the resulting binary masks, together with the corresponding orthophoto. In this figure, rows 1 to 3 each shows anomalous temperature distributions in the area of the identified ROI. Possible cause might be temporal changes in ground moisture or temporal occlusions, i.e. by dynamic objects such as vehicles. Rows 4 and 5, on the contrary, show two examples of the explainalbe type of inaccuracies. They are explainable by taking into account the same sensor data that had been used for the digital twin generation. Therefrom it can be deduced that the ROIs are placed around small buildings, with grayish roofing, that were missing in the digital twin, i.e. the digital twin was incomplete at these locations.

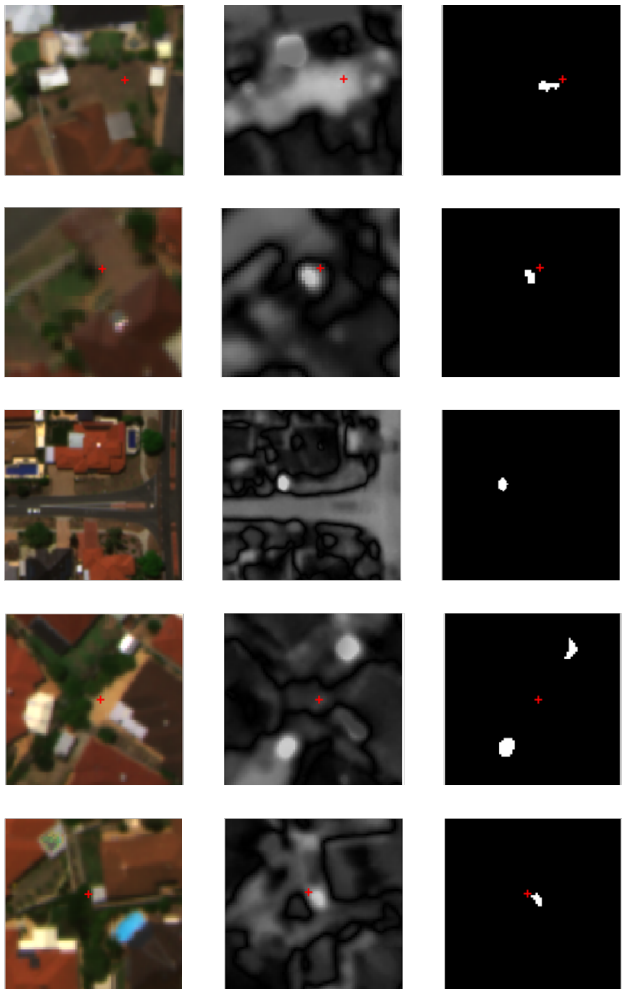

Figure 1. Examples of regions of interest: Orthophoto (left), mean-substracted difference image (middle), binary mask of regions of interest (right), labeled as anomaly (rows 1-3) and explainable (rows 4 and 5)

3.3 Regions-of-Interest classification

We aim to separate the ROIs into the two classes *explainable* and *anomalous* as described previously. As manual separation is presumably tremendous, we employ a supervised machine learning approach. Given the co-registered sensor data and thereof deviated information from the 4D digital twin, the pixel-wise data inside the ROI are summarized to form segment-based features. Namely, the pixel-wise input data is given by a multi-spectral aerial image which includes the wavebands red, green, blue, near-infrared (NIR) and red-edge (RE), an aerial thermal image, i.e. measured surface temperatures, and the normalized digital surface model (NDSM), normalized difference vegetation index (NDVI), normalized difference water index (NDWI), the planarity map of the area and the semantic land-cover segmentation which were determined during the 4D digital twin generation. Additionally, an orthogonal projection of the 3D digital twin is performed to get a second pixel-wise class image differing from the land-cover segmentation due to constraints during meshing. This information will later be particularly helpful in training data generation. For each pixel-wise feature, except the land-cover segmentation result and projec-

ted class image, the mean and standard deviation over each of the ROI is determined as segment-based predictor. For the semantic classes, the most frequent class label over the ROI is determined.

Finally, Random Forest (Breiman, 2001) was chosen for the classification task due to its simplicity, flexibility, and the option to evaluate given features with out-of-bag estimates. Random Forest, which is built up on decision trees, allows to categorize given data by means of hyperrectangles and is stable towards redundant features. Also, Random Forest is able to deal with high number of features even in the case of a low number of training samples. However, Random Forest does overfit. Possible countermeasures are the introduction of a cost term to penalize a split (Hastie et al., 2011), pruning the trees using cross validation (Breiman et al., 1984), or using a stopping criterion like a minimum sample size per node. The latter was chosen in our approach due to its simplicity. As we will outline in the following, this choice of classifier is evaluated and verified by our feature analysis.

4. EXPERIMENT

4.1 Test Site

The simulation as described in Section 2 had been carried out for a 2×1km area in the City of Melville, Australia. The urban residential zone is characterized by mostly low-rise or two-level buildings with tiled or metallic roofs, yards and courtyard entrances, a few bigger nonresidential buildings, several grasland areas, and geo-typical vegetation such as palm trees. Figure 2 shows the simulated thermal image which results from the final 4D digital twin along with the corresponding orthophoto and measured thermal image. The mean difference and standard deviations are given by $\mu = 2.321$ and $\sigma = 2.515$, and are used in our necessary condition according to Equ. (2). The simulation achieved an overall root-mean-square error of 3.745.

4.2 Data Generation

Applying the necessary condition on the image differences yielded 154 ROIs which were thereupon manually labeled. During labeling, the differences in pixel-wise semantic land-cover segmentation and projected semantic object classification often played an important role, given that small building structures had been removed in the process of the 3D digital twin creation. However, those appeared to create a significant TIR signature. Since the source of error of corresponding ROIs is known, they naturally are not marked as anomalies. Furthermore, NIR, RE, NDVI and NDWI played a minor role within the manual labeling process.

Eventually, it was found that 132 out of 154 ROIs can be ascribed to the class of explainable inaccuracies, since they deviate from image misregistration, semantic classification error, or other known sources of error within the procedure of the 4D digital twin generation. The other 22 ROIs remained as anomalies whose source of deviation could not be resolved.

4.3 Feature analysis

Given the segment-based features designed as introduced in Section 3.3, the corresponding scatterplot matrices were computed, see Figure 6. This standard tool allows for a first insight into the features' utility. As can easily be seen, most of the feature spaces hardly display clustering. Mean measured and simulated temperatures show two clusters; however, they are not correlated to our ground truth classification. Further sensor data such as NIR and RE appear to not allow any kind of separation within the regarded feature space. Yet, considering the NDSM and planarity feature space, the regions clearly show distinguishable distributions of the two classes. From the high planarity and low NDSM values of the anomalous regions, we conclude that mainly ground areas contribute to them. This is confirmed by the scatterplot of the land-cover segmentation map feature where the classes 1, 3 and 5 correspond to street, grass and soil respectively, whereas 2 and 4 corresponds to building and high vegetation. At last, the scatterplot histograms of standard deviations show narrow distributions around low values for anomalous regions, whereas others display a rather broad distribution.

In conclusion, the scatterplot matrix reveals that the separation of the two classes in the feature space is preferably carried out by hyperrectangles than hyperplanes. Also, reliable feature selection based on this analysis would be demanding, therefore we select each available feature for the subsequent classification task. Thereby, the number of features is high in comparison to the number of training samples, as outlined in the preceeding section. Taking these findings into account, Random Forest appears to be the most appropriate classifier for this task. To prevent our Random Forest model from overfitting, we apply a stopping criterion as outlined in Section 3.3. Due to the low amount of training samples, we chose a minimum node size of 2 samples. During experimentation, it was found that this number yielded best results on the test data.

In the following Sec. 4.4, the classification results by training and testing the Random Forest classifier are presented. However, since Random Forest allows further feature analysis, we introduce these analysis results here in advance. Figure 3 displays the predictor importance yielding from out-of-bag-score evaluation from Random Forest. Clearly, the importance of both the features planarity and NSDM are confirmed, yet overtopped by the standard deviations of the NIR and RE feature.

4.4 Simulation Error Classification

From the generated dataset, cf. Sec. 4.2, we chose 60% to be used for training and 40% for testing, aiming for a relative high fraction of anomalies in each subset. Since our chosen classifier is prone to imbalanced datasets, the training data was further reduced to impose balance between the two classes. Table 1 summarizes the counts of the ROIs in the overall labeled dataset, the training data and the test data.

The Random Forest was trained multiple times with different numbers of decision trees. Best results in classification were achieved by 64 trees, as can be seen from the evaluation metrics displayed in Figure 4. In Figure 5, some examples of the testsamples which were correctly and falsely classified, respectively, are shown. The example of an anomly shows a temperature variation within a ground segment. A probable, but not definitely determined cause might be ground moisture or violation of the closed-world assumption. The example on the second column displays the case of an explainable ROI. As can be seen from the measured and thermal images, misregistration is present which yields the ROI. Our classifier was, in this case, able to classify correctly. However, see the third example on the third column, there is some misregistration too, and the ROI highlights the border area between the street and gras, as can

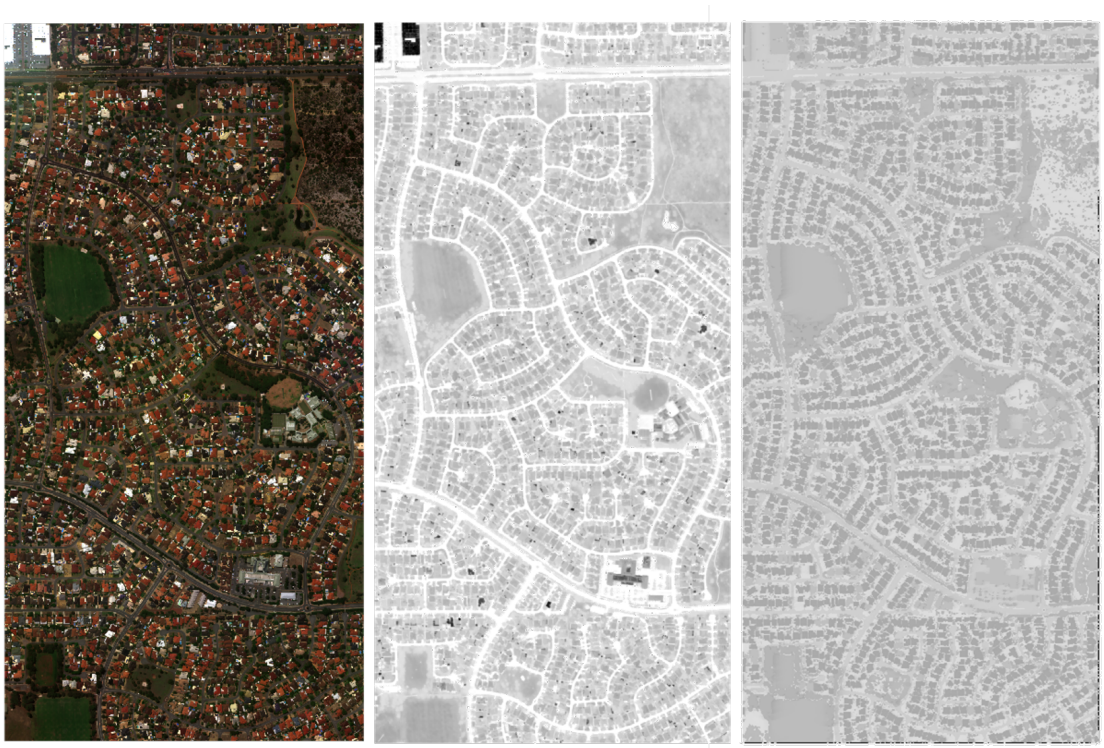

Figure 2. Orthophoto (left), measured (middle) and simulated temperatures (right) in white-hot representation.

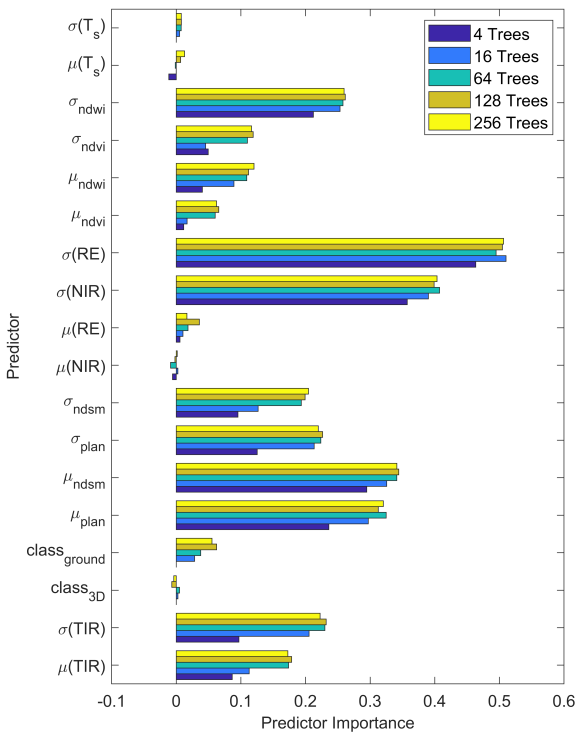

Figure 3. Predictor Importance

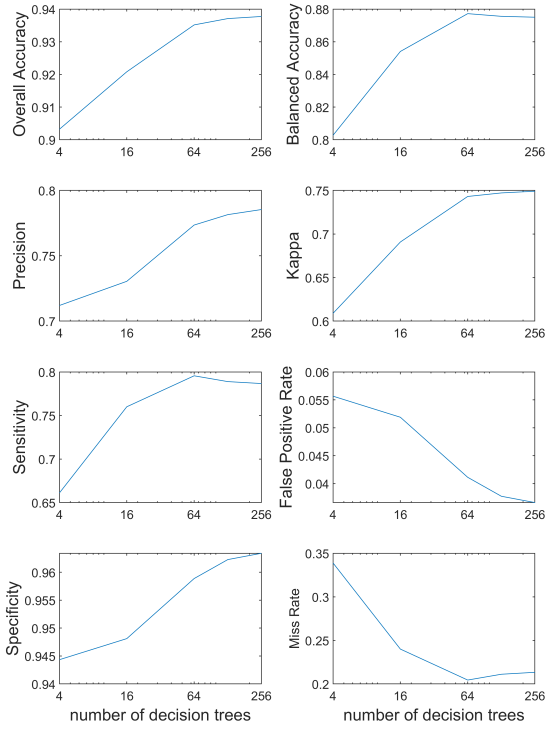

Figure 4. Evaluation Metrics

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W2-2022-265-2022 | © Author(s) 2022. CC BY 4.0 License.

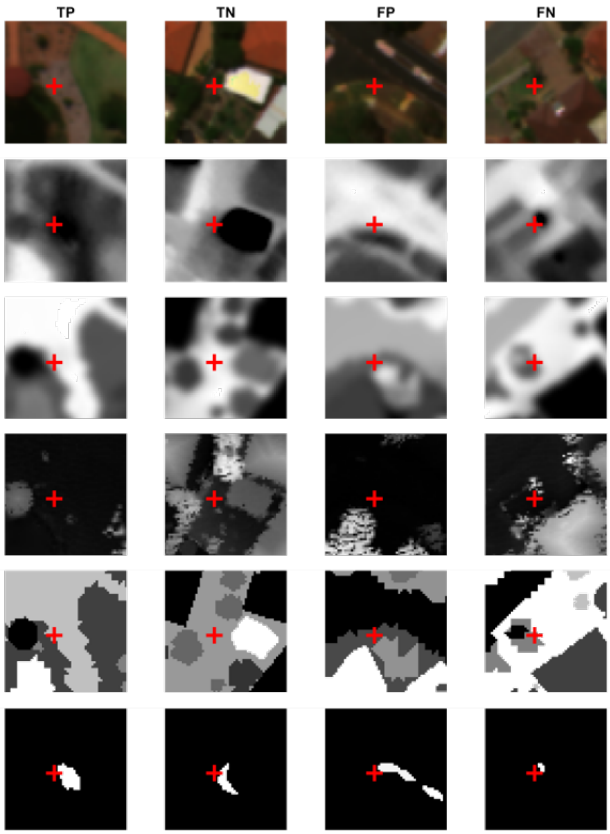

Figure 5. Examples of the test data. From left to right: correctly classified anomaly (true positive, TP), correctly classified explainable area (true negative, TN), falsely classified anomaly (false positive, FP), falsely classified explainable area (false negative, FN). The rows show the corresponding image section of the region-of-interest, from top to bottom: orthophoto, measured thermal image, simulated thermal image, NDSM, class image, binary mask of the region.

Dataset	# ROI	# normal	# anomalous
All	154	132	22
Training Data	25	13	12
Test Data	62	53	9

Table 1. Number of regions of interest (ROI) in summary, labeled as normal and as anomalous respectively, in the generated dataset and its subsets for training and testing.

OA	BA	Pr	κ	Se	FPR	Sp	M
93.5	87.8	77.4	0.74	79.6	4.1	95.9	20.4

Table 2. Evaluation of the classification results in percentage: Overall accuracy (OA), balanced accuracy (BA), precision (Pr), kappa (κ), sensitivity (Se), false positive rate (FPR), specificity (Sp), miss rate (M)

be seen from the class image. Yet the classifier assigned this ROI to be anomalous. Possible cause of failure might be that, in comparision to the correctly classified misregistration, there is no change in the NDSM. Since NDSM states an important feature, cp. our findings of Section 4.3, it is possible that our training data quantity was not sufficient to adequately train our Random Forest on misclassifications. Lastly, the example on the forth column shows a misclassified explainable region. Indeed, it appears that no cause for this ROI could be found, not even estimated.

The overall accuracy reached 93.5%, with a precision of 77.4%. With a sensitivity of 79.6% and a false positive rate of 4.1%, the anomalies could be well separated from the explainable regions. Since the test data involves asymmetry in the two categories, the balanced accuracy as mean value of sensitivity and specificity is computed as well, which yields an adequate value of 87.8%. Considering the Cohen's kappa value of 0.74, the classifier achieves substantial agreement between the classification results and the ground truth. Table 2 summarizes these findings.

5. CONCLUSION

In this paper, we examined the applicability of machine learning in the process of simulation error identification and classification. We focused on the separation of errors due to closed-world assumption violation and simplification in the digital model (explainable errors) to errors raised by scenario uncertainty and simulation inaccuracies (anomalous errors). We defined a proper deviation metric for the thermal simulation results and manually labeled regions of error according to the two defined classes. During feature analysis, we concluded the elevation and planarity of the regions to be of major importance, which had been confirmed by the evaluation of the Random Forest classifier. Despite the few training data available, encouraging classification results were achieved. Identification rates of the two classes constitute 95.9% and 79.6% for explainable and anomalous errors respectively. From the training data, we conclude that the anomalous error types play a minor role in the applied thermal simulation model. However, many errors arise from simplification and closed-world assumption within the 3D digital twin. During labelling of the training data of this presented work, clear conclusions could be drawn to what needs to be improved in the 4D digital twin generation. Therefore, we wish to explore a finer classification of simulation errors in future, in order to gain deeper insight into the interference of digital and simulation model assumptions and simulation accuracy.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-X-4-W2-2022-265-2022 | © Author(s) 2022. CC BY 4.0 License.

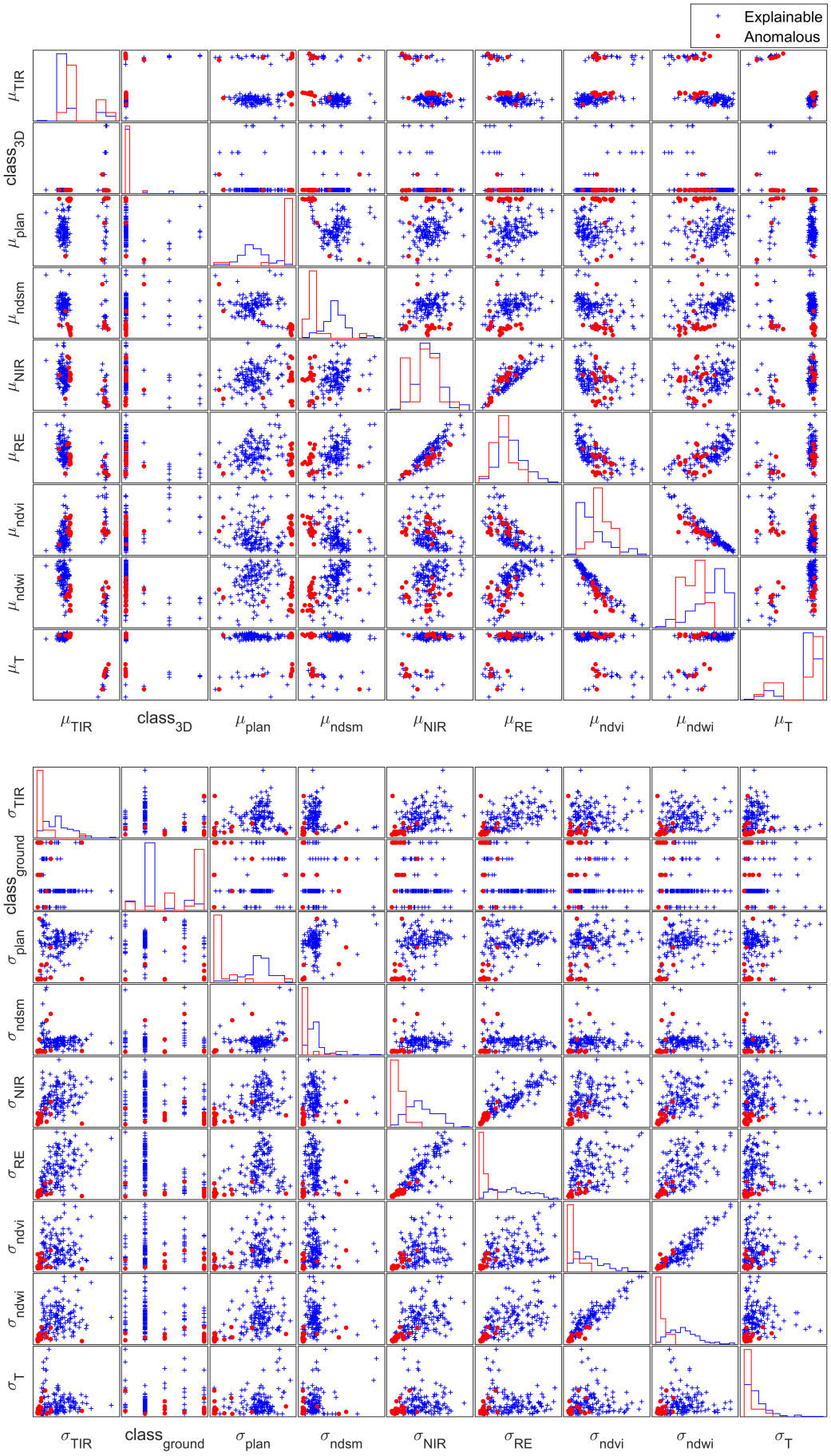

Figure 6. Scatterplot matrix of means and standard deviations of the features: measured aerial TIR image, land-cover segmentation map $class_{ground}$, projected semantic class image $class_{3D}$, planarity map *plan*, NDSM, NIR image, RE image, NDVI, NDWI, and simulated surface temperature T

ACKNOWLEDGEMENTS

The authors would like to thank the City of Melville, in particular Petra Helmholtz (Curtin University) for providing the research data of the test site.

REFERENCES

Blaha, M., Vogel, C., Richard, A., Wegner, J. D., Pock, T., Schindler, K., 2016. Large-scale semantic 3D reconstruction: An adaptive multi-resolution model for multi-class volumetric labeling. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 3176–3184.

Breiman, L., 2001. Random forests. *Machine learning*, 45(1), 5–32.

Breiman, L., Friedman, J., Stone, C. J., Olshen, R. A., 1984. *Classification and regression trees*. CRC press.

Budde, L., Bulatov, D., Iwaszczuk, D., 2021. Identification of misclassified pixels in semantic segmentation with uncertainty evaluation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2021, 441-448.

Bulatov, D., Burkard, E., Ilehag, R., Kottler, B., Helmholz, P., 2020. From multi-sensor aerial data to thermal and infrared simulation of semantic 3D models: Towards identification of urban heat islands. *Infrared Physics & Technology*, 105, 103233.

Bulatov, D., Häufel, G., Meidow, J., Pohl, M., Solbrig, P., Wernerus, P., 2014. Context-based automatic reconstruction and texturing of 3D urban terrain for quick-response tasks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93, 157–170.

Burkard, E., Bulatov, D., Kottler, B., 2020. Towards detection of thermal anomalies in large urban areas using simulation. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 43.

Chong, A., Gu, Y., Jia, H., 2021. Calibrating building energy simulation models: A review of the basics to guide future work. *Energy and Buildings*, 253, 111533.

Duffie, J. A., Beckman, W. A., 2013. *Solar engineering of thermal processes John A. Duffie, William A. Beckman*. 4th ed. edn, John Wiley.

Groth, C., 2009. *FEM für Praktiker: Temperaturfelder: Basiswissen und Arbeitsbeispiele zu FEM-Anwendungen der Temperaturfeldberechnung; Lösungen mit dem Programm ANSYS; mit CD-ROM und zahlreichen Beispielen*. 3, expert Verlag.

Guo, S., Xiong, X., Liu, Z., Bai, X., Zhou, F., 2018. Infrared simulation of large-scale urban scene through LOD. *Optics express*, 26(18), 23980–24002.

Hastie, T., Tibshirani, R., Friedman, J. H., 2011. *The elements of statistical learning : data mining, inference, and prediction*. 2. ed., corr. at 5. print. edn, Springer, New York, NY.

Häufel, G., Bulatov, D., Helmholz, P., 2019. Statistical analysis of airborne imagery combined with GIS information for training data generation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W16, 111-118.

Kottler, B., Burkard, E., Bulatov, D., Haraké, L., 2019. Physically-based thermal simulation of large scenes for infrared imaging. *International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, INSTICC, 53–64.

Lafarge, F., Mallet, C., 2012. Creating Large-Scale City Models from 3D-Point Clouds: A Robust Approach with Hybrid Representation. *International Journal of Computer Vision*, 99(1), 69–85.

Lienhard IV, J. H., Lienhard V, J. H., 2019. *A Heat Transfer Textbook*. 4th ed. edn, Dover Publications, Inc.

Lyu, H.-M., Shen, S.-L., Zhou, A., Yang, J., 2019. Perspectives for flood risk assessment and management for mega-city metro system. *Tunnelling and Underground Space Technology*, 84, 31–44.

Malaplate, A., Grossmann, P., Schwenger, F., 2007. Cubi: a test body for thermal object model validation. *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVIII*, 6543, International Society for Optics and Photonics, 654305.

Shahat, E., Hyun, C. T., Yeom, C., 2021. City digital twin potentials: A review and research agenda. *Sustainability*, 13(6), 3386.

Xiong, X., Zhou, F., Bai, X., Xue, B., Sun, C., 2016. Semi-automated infrared simulation on real urban scenes based on multi-view images. *Optics express*, 24(11), 11345–11375.

Xu, Y., Sun, Y., Liu, X., Zheng, Y., 2019. A digital-twin-assisted fault diagnosis using deep transfer learning. *IEEE Access*, 7, 19990–19999.

Yan, J., Zlatanova, S., Aleksandrov, M., Diakite, A., Pettit, C. J., 2019. Integration of 3D objects and terrain for 3D modelling supporting the digital twin. 14th 3D GeoInfo Conference, 24-27 September 2019, Singapore.