A meta analysis of the status of AI in environmental computer science

Lukas Sinnwell[1], Julian Bruns [2], Matthias Budde[3], Andreas Abecker [4]

Abstract: Artificial Intelligence is hyped as one of the key enablers of the future which will help to solve key challenges of humanity. The research field which already addresses many of those challenges is Environmental (Computer) Science. Therefore, it seems like a natural fit to combine both fields. AI can provide algorithms and methods to ease the computations and extend the existing (simulation) models of the environmental sciences. The environmental sciences already have in-depth knowledge of the problems at hand, can evaluate and interpret both the relative importance of input data as well as the results. However, to achieve this synergy, a strong foundation and knowledge of previous work is needed. This work aims to contribute to this foundation by providing a data-driven overview of AI-based research activities in the field of environmental computer science.

Keywords: AI; Computational Sustainability, Environmental Computer Science, Meta Analysis

1 Introduction

Most of today's medium- to long-term problems are environmental problems. 8 out of the 17 United Nation's Sustainable Development Goals (SDGs) [SD18] are directly connected to environmental issues. Environmental Computer Science (ECS) plays part in solving these problems [Fi17], as it deals with modelling, understanding, and ultimately predicting environmental processes by using methods of information technology [Hi95]. However, given the complex, and often non-linear causalities in natural processes, building these models and generating usable results, i.e. predictions, is a difficult task. This is where Artificial Intelligence (AI) could make a difference, as it offers a different approach to data modeling than conventional methods do. But while AI postulates itself as a üniversal field"[RN10], it is not yet fully part of the tool set of the ECS field[Go19]. And while there are many publications, projects and visions, such as [Li21] in the proceedings of this conference last year, as well as even broader calls for actions such as the idea of computational sustainability [Go09] - an enhanced integration of Computer Science (CS), Mathematics and Statistics in a quest for a more effective use of natural resources - there does not yet exists an overview of the state of AI in ECS. To cite the German Ministry of Environment: There is "[. . .] no publication that allows a comprehensive interpretation

[1] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe lukas.sinnwell@disy.net
[2] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe julian.bruns@disy.net
[3] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe matthias.budde@disy.net
[4] Disy Informationssysteme GmbH, Ludwig-Erhard-Allee 6, 76131 Karlsruhe andreas.abecker@disy.net

of the research activities in the context of environment and sustainability - neither on a European nor on an international level (status March 2019)" [Je19]. Hence, the aim of this work is to provide a first data-driven overview of the current state of the art of AI-related work in ECS. While there are many surveys, theses and textbooks regarding AI in general [RN10][Ch20], its status[Mo20] as well as future [Ka20][Sc19]. And while there are specific surveys for the use of AI in specific sub fields of the ECS, such as [MD00] or [Sh18], their approach is guided by the applicability to their field and not about overall status of AI in their sub field.

We aim to achieve this by (1) Gather a corpus of publications with regard to ECS and AI, (2) Mapping the publications to their respective research fields and (3) evaluate the methodological approaches of the 5 most cited publications in each field in regard to 12 criteria derived from the AI literature.

2 Publication corpus

To get an impression of active fields in AI, and AI and ECS in particular, we used the records from Web of Science (WoS) to gather the different publications. WoS groups each publication into specific categories and rank these by the number of records per category. These are not assigned exclusively. A first generic query for just AI and its native fields returns about 200.000 results. After investigating the top 82 results, we refined our query to focus on publications with an environmental science background, see fig. 1. This query returns all records in which one of the words marked in purple above is present in either title, abstract or keywords. These are basic AI terms, in order to include as many records as possible. We chose those four terms as they encompass the most undisputed AI terms in regard to methods. The second part of the query (after AND) specifies the desired WoS categories. Every record on WoS is assigned at least one of more than 250 categories. We then selected all categories from the field of environmental science which belongs to the top 100 AI categories. Eight categories re-main using this approach, namely Environmental Sciences, Water Resources, Geosciences Multidisciplinary, Remote Sensing, Green Sustainable Science Technology, Geography Physical, Agriculture Multidisciplinary and Environmental Studies. Overall, the query returned 18.254 records, as of November 2020. This is slightly less than 10% of the overall AI literature found in WoS. The addition of further terms would deviate from the data driven approach used in this paper.

3 Clustering of the field

Based on this corpus of publications, we are able to built a network of the publications based on their title, keywords, categories as well as citation links. This allows us to cluster the overall publications and then map each publication to a cluster. We use the tool VOSviewer [VEW13] to perform these tasks. We use the association strength as a measure for the

```
(TS = ("Artificial Intelligence") OR TS = ("Machine Learning") OR
TS = ("Deep Learning") OR TS = ("Artificial Neural Network*"))
AND
(WC = ("ENVIRONMENTAL SCIENCES") OR WC = ("WATER RESOURCES") OR WC = ("GEO-
SCIENCES MULTIDISCIPLINARY") OR WC = ("REMOTE SENSING") OR
WC = ("GREEN SUSTAINABLE SCIENCE TECHNOLOGY") OR
WC = ("GEOGRAPHY PHYSICAL") OR WC = ("AGRICULTURE MULTIDISCIPLINARY") OR
WC = ("ENVIRONMENTAL STUDIES"))
```

Fig. 1: Query for Web of Science to gather all publications for analysis

mapping and similarity of the publications to each other. We then applied the Smart Local Moving Algorithm (SLM) from VOSviewer to create the different clusters. However, there is no inherent selection criteria for the optimal number of clusters in the algorithm or software. We used a variant of the so-called ëlbow criteria often used in the k-means clustering - to decide the different parameter for the clustering. This results in seven different classes of cluster and it can be seen in fig. 2.

Having established suitable parameters for mapping and clustering, the problem of labelling the clusters emerges. VOSviewer does not label the clusters natively. However, VOSviewer allows to save a created mapping into a map-file and a net-file. The map-file contains all items, their coordinates, information (title, authors, source), and most important for this, the cluster they belong to. The different terms used for the mpassing of the items for each cluster are extracted from the map-file into single, tab delimited text files. These files can in turn be analysed in VOSviewer to extract frequent term via NLP. After extracting all the titles into a separate file, VOSviewer detects just over 10.000 different terms. We select only 25 most frequent domain terms for each cluster. Generic Terms like ’model’, ’prediction’, etc. are not meaningful in this context and therefore are removed beforehand. By performing this analysis we are able to name the cluster based on their content and most often used terms. Fig. 3shows the top 5 terms per cluster and the cluster names derived. We chose those terms based on existing knowledge about the fields and our pre-existing domain knowledge as well as discussions with environmental scientists.

4 Methodological Evaluation

Based on the seven different clusters and their mapped publications we are now able to investigate more in-depth the approach to AI in the different sub fields. We investigate the five most cited publications in each field. Reviews and surveys were removed in order to analyse publications carrying out a specific study. Experience has shown that AI often only appears marginally in reviews and is one of many explored methods. This way, the use of AI and the methods are defined more clearly and are more accessible. In this context, it is

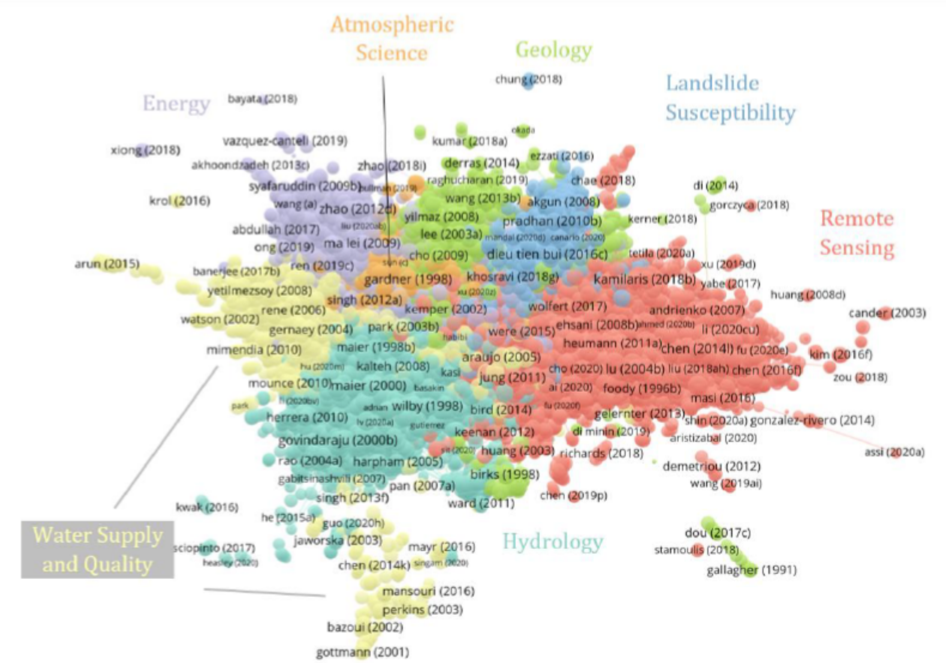

Fig. 2: Map of the WoS dataset created in VOSviewer using 0.30 resolution and a minimum cluster size of 300. The biggest cluster (red) contains 4839 paper.

cluster 1	cluster 2	cluster 3	cluster 4	cluster 5	cluster 6	cluster 7
sentinel	river	solar radiation	adsorption	soil	emission	landslide susceptibility mapping
remote sensing	river basin	building	wastewater	rock	city	gis
forest	basin	energy	aqueous solution	blast	air quality	landslide
hyperspectral image classification	reservoir	energy consumption	treatment	earthquake	ozone concentration	landslide susceptibility assessment
remote sensing image	temperature	photovoltaic system	water	mine	ozone	landslide susceptibility
Remote Sensing	*Hydrology*	*Energy*	*Water Supply & Quality*	*Geology*	*Atmospheric Science*	*Landslide Susceptibility*

Fig. 3: Top 5 domain terms per cluster and cluster names derived from this.

also interesting how the studies were carried out and what steps and information were given. We use the most cited publications as our selection criteria as the number of citations can be used as measure of their importance for the field. Either by their innovative character, foundation of the use of AI in the field or by relevance of the topic for the overall field. It also enhances the chance that it is indicative for the field. However, this approach has a trade-off. We implicitly provide more weight to older publications and thereby ignore more recent developments and improvements in the fields. We argue that this trade-off is acceptable for the aim of this paper to provide a first overview for our reasons stated before.

For the final, in-depth analysis we had to perform several small changes from our selection of publications for the clusters. For the "Water Supply and Quality"cluster the five papers with the most citations all revolve around the topic of climate change. However the majority of papers in this cluster, deal with topics evolving around water, waste water, and chemistry. While they are often relevant in the context of climate change, they address issues beyond that in most cases. Since the most cited papers would not be adequate representation of this cluster, only the most cited publication is chosen in this cluster with explicit connection to climate change. The selection is filled up with the next most cited publication, which are thematically address a broader spectrum and provide a more representative view of the cluster. Tab. 1 provides an overview of the selected publications for each cluster.

Cluster	Most cited paper
Landslide Susceptibility	[PL10], [Yi09], [Bu16], [PPG12], [Le04]
Remote Sensing	[HDT02], [Ro12], [Ch14], [Ch16] , [CZH16]
Hydrology	[HGS95], [MH96], [VBB13], [Wa09], [Na04]
Geology	[KS02], [Zo08], [KS07], [SMJ02], [Br00]
Energy	[Br10], [Gu12], [Gh09], [MRH98], [Ka09]
Atmospheric Science	[AWBAA05], [So07], [GD99], [Fe15], [GC06]
Water Supply	[Ar05], [HKH04], [Ma16], [YD08], [NSD04]

Tab. 1: Overview of papers selected for analysis grouped by clusters.

To evaluate those 35 publications, we developed a set of methodological criteria to compare those publications objectively on their approach to AI. Those criteria are based on the common steps in an AI workflow commonly used in the literature or in practical AI projects, for example proposed in a similar form by [Za19] for human centred AI models. The criteria are extended by characteristics proposed by [YBL11], [Yu20] and [RS20]. 4 shows the used criteria in this publication to evaluate the publications. We divided the criteria in the general criteria for statistical and bibliographic aspects and in methodological criteria, which are the focus in this work. The 12 different criteria allow the reader to quickly identify the key information of the AI model building process. While additional criteria, such as the type of data, the preprocessing and so on could be added, we decided to focus on the most common and overarching criteria which can be investigated objectively. They allow to investigate key means of reproducibility of the studies as well.

The aggregated results of this comparison to our criteria is shown in fig. 5. Only two studies,

General criteria
Research Area
Title
Authors
Times Cited
Year
Countries
Type
Journal
Conference
Keywords
AI methods in focus
Key notes
Outcome

Methodological criteria	
type of data	*Kind of data is used in study*
sample size	*size of dataset*
Input Selection	
Input Significance	*methods for deciding which inputs to use in model*
Input Independence	*methods for checking correlation among inputs*
Data Division	
Data Division	*method for dividing data into train / test set*
Training [%]	*Percentage of whole dataset used for training*
Test [%]	*Percentage of whole dataset used for testing*
Validation	*validation technique or percentage of dataset used for validation*
Model Building	
Hyperparameter tuning	*methods for setting HPs*
NN architecture	*if available, methods of determining NN architecture*
Evaluation & Reference	
Evaluation metrics	*metrics used for model evaluation*
Reference model	*reference for model developed in study*

Fig. 4: Criteria used for in-depth analysis.

namely [Na04] and [Zo08] met all criteria. A large portion of studies met all but one, two or three criteria. As the number of missing information grows, the number of affected papers declines. The focus of this in-depth an analysis is on completeness, transparent documentation, reproducibility, and replicability. The criteria are chosen to reflect these characteristics. By not meeting all the criteria, not all of these characteristics can be achieved by the studies. However, it is important to say that when a study does not mention certain information about the ML process, it does not automatically imply that the respective step was neglected. It might as well be that the authors deemed this information as not important enough to appear in the publications. None of the papers we investigated provide access to materials i.e., neither the data which was used nor the code or fragments of it. This means an exact replication these studies is not possible. In the following subsection we provide a more in-depth view of each criteria.

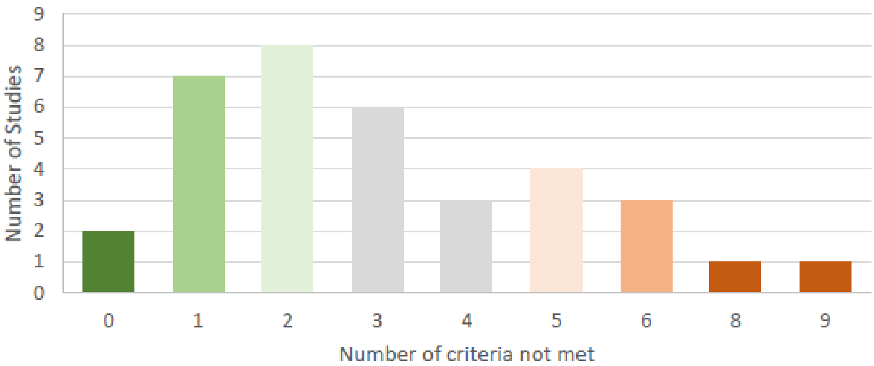

Fig. 5: Distribution of number of studies not meeting a number of criteria.

4.1 Sample size

All but three studies reported on their sample size. It ranges from 36 soil samples to 12.000 environmental drugs and chemicals.

4.2 Input significance

Concerning Input Significance, 12 studies made no clear statement, whether input significance measures were applied and if so, which ones. One has to consider that input significance can be a highly domain specific matter. For some studies, it may be the case that from a domain perspective the selection of inputs is generally agreed on. Studies that explicitly used domain knowledge for input selection were marked with ‘ad hoc’. This relates to eight studies. This leaves 15 studies which used some numerical tool to determine significance of input features; four studies used two procedures. Four studies used training models with different combinations of inputs i.e., trial and error, and three time each regression analysis, correlation analysis, and Principal Component Analysis were used.

4.3 Input independence

Input independence was neglected more often. But here one can also not rule out that some generally recognized domain knowledge was assumed to be known by the reader. Only five studies made explicit statements regarding input independence, of which three used correlation analysis, one trial and error approach and one variance inflation factor.

4.4 Data division

In terms of data division, a random division is most used among the sample studies (12 times). However, for time series forecasting, a random division of data is not advised. In these cases, the time series were divided arbitrarily at some point in time according to the chosen training and test sizes. All selected studies concerned with time series forecasting (11) used this technique. The remaining three cases used stratified random division, supervised division or a given division, respectively. Nine studies performed no data division. The percentage division ranges from a 50:50 split to 80:20 split (training/testing). Six publications provided no information about a data division and two studies provided the number of training samples but no further information.

4.5 Validation

Validation was mentioned significantly fewer times than training, testing or model build-ing related facts. Nearly half of the studies (17) did not mention this topic. It was shown that some textbooks recommend a three-way split, where a partition of the data is re-served for validation purposes. However, in recent years, the cross-validation technique gained popularity. This technique does not require a static validation set. The studies are almost evenly distributed between these two approaches with eight using a validation set and seven using cross-validation. One study used the built-in Out-of-bag Error (OOB) of the Random Forest algorithm, one used "[. . .] random samples [. . .]" [Le04], and one switched test and training set for validation [KS02].

4.6 Hyperparameter tuning

In 24 studies remarks were made regarding the tuning of these parameters. Out of the 24 which made remarks, ten were ad hoc settings, just providing information on the values that were chosen or referring to other works that used the same settings. This leaves 14 studies where a technique was used to determine the optimal hyper parameters. Here, Grid Search was the most frequent technique (five times), followed by trial and error (twice) and stepwise adjustment (twice).

4.7 NN architecture

The architecture refers to the structure of neural networks, more specifically to the determination of the number of hidden layers and hidden units used. Therefore, it applies only to studies which used NNs. 30 of the 35 publications used some kind of neural network. All gave information about the architecture itself i.e., the number of neurons and layers. Ten publications were marked as ‘ad hoc’, where no clear approach could be identified. A total number of 11 studies used a trial-and-error approach to determine network architecture, whereas three used a controlled, stepwise approach.Three studies referred to literature to justify their network architecture. Other approaches (Grid Search, predefined) were used once, respectively.

4.8 Evaluation metrics

This choice of the evaluation metric is dependent on the type of problem (classification, regression, clustering). Generally, 25 different metrics were used and only nine studies used just one metric, whereas 4 studies did not provide this information at all.

4.9 Reference model

The use of a reference is important in order to properly evaluate the result of a study. Usually, the performance of a trained model is compared to the performance of some pre-existing models, approaches, algorithms or simulation approaches. It has to be trained or computed with the same data basis and evaluated on the same metrics to allow a fair comparison. A major part of studies did so in some way, namely 27, leaving 8 studies without reference for the evaluation. Three publications developed multiple ML models of different kinds and compared them to each other. This can be seen as a form of reference, as these studies aim to find the best performing algorithm for their data and use case. Most of the other studies concentrated on a maximum of two ML methods and compared them to a simpler model or, in two cases, to common models and methods of their domain[KS07][SMJ02]. Most regression based studies used simple regression models as reference.

5 Conclusion

To survey a field as large as Artificial Intelligence in Environmental Computer Science a data driven approach is a suitable first step. We showed that there is indeed a widespread use of methods of Artificial Intelligence in Environmental (Computer) Science. There are 18.000 publications concerning these topics to be found in the databases Web of Science alone. According to our analysis the research activities of ECS in regard to AI are driven mainly by seven different fields. Based on the interconnectivity of the terms these clusters could be derived and their fields are based on the most common terms in each cluster. Finally, we analysed the five most cited publications of each field to investigate in detail their ML processes and documentation of these processes.

The way of publication suggests that the field of AIECS is predominantly populated by domain researchers who use AI as a tool. The way of publication suggests that the field of AIECS is predominantly populated by domain researchers who use AI as a tool, which is the most important finding of this work. AI in ECS is mostly seen as a set of methods, a helpful tool like GIS, Statistics or Data Visualization. It is a means to an end, a way to get the best out of one's own data. This bears a second important realization. The goal of most publications in our sample for in-depth analysis is to get a good prediction from a very specific dataset. During our literature research we often found that most researcher are mainly interested to find a good prediction for an environmental problem at a specific location. Seldom, the goal is to develop a general applicable model for a specific issue. This is entirely legitimate and is a question that naturally comes up again and again in location-related sciences. However, this signifies are divergence between research in pure Artificial Intelligence and Environmental Science. Research in pure AI aims to design models and develop algorithms and methods that are applicable as broadly as possible. On the other hand, Environmental Sciences use AI as one of many tools to achieve their actual goal – often times, the determination of variable Y in place Z. There are of course cases in which the application of AI in ECS is analogous to developments in AI.

However, this often specific application can also be seen more critical. Many if not most of the publications investigated can not be reproduced and therefore truly evaluated. The problems detected throughout this work resemble those that [FW19] describe as reproducibility crisis - studies not being reproducible and missing transparency in description of procedure. In particular, establishing guidelines and rules for data sharing and transparency could be ways to achieve this in the long term, see e.g. [FW19]. And while there are already steps taken by different agencies such as the explicit publication and citation of data sets, requirements by journals and conferences, we want to emphasize this problem.

This study is only a first step to provide understanding of the field and is based on a pure data based approach to identify the clusters of the field and pre-select the most cited studies in each field. While we argue that it provides a starting point in particular in regard to the difference in approaching AI for computer scientist and environmental scientists - and their differences in their goals - a more in depth survey paper which investigates this field of

AI in ECS is still missing. We only focussed on the most cited publications, which could have led to a bias to older publications. For future work, it would also be interesting to investigate the seven sub fields more in-depth as to why those fields seem to publish the majority of AI related work in the ECS field. While remote sensing can be easily deduced to be influenced by the increasing availability of remote sensing data and their huge amount of data as well as their similarity to the AI field of vision, the other fields are not so obvious for us. Other potential evaluation criteria, such as the use of the well-known KDD process or CRISP-DM, were considered, but proved to be difficult to asses in an automated manner. The aim was to first focus purely on the AI models and the core-application of those by the researchers. Further studies, which investigate the use of standardized approaches to the application of AI in their broader sense, would provide interesting insights. Finally, a more in-depth evaluation of different time periods would provide a benefit to our understanding of the evolution of the field. Almost a third of the examined publications, around 5600, were published in 2019 and 2020. This would allow to capture the rapid advances in the field. However, this was beyond the scope of this work. In our view this can only be created collaboratively by researchers from both fields to provide the overview and understanding needed to provide this overview. It could then be possible to truly leverage the benefits of AI in ECS.

Acknowledgement

This publication was done in the context ofthe project Nitrat-Monitoring 4.0 - Intelligente Systeme zur nachhaltigen Reduzierung von Nitrat im Grundwasser (NiMo 4.0), https://nimo-projekt.de/, funded by the BMU with Förderkennzeichen FKZ 67KI2048.

Bibliography

[Ar05] Araújo, Miguel B; Pearson, Richard G; Thuiller, Wilfried; Erhard, Markus: Validation of species–climate impact models under climate change. Global change biology, 11(9):1504–1513, 2005.

[AWBAA05] Abdul-Wahab, Sabah A; Bakheit, Charles S; Al-Alawi, Saleh M: Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. Environmental Modelling & Software, 20(10):1263–1271, 2005.

[Br00] Brown, Warick M; Gedeon, TD; Groves, DI; Barnes, RG: Artificial neural networks: a new method for mineral prospectivity mapping. Australian journal of earth sciences, 47(4):757–770, 2000.

[Br10] Brekken, Ted KA; Yokochi, Alex; Von Jouanne, Annette; Yen, Zuan Z; Hapke, Hannes Max; Halamay, Douglas A: Optimal energy storage sizing and control for wind power applications. IEEE Transactions on Sustainable Energy, 2(1):69–77, 2010.

[Bu16] Bui, Dieu Tien; Tuan, Tran Anh; Klempe, Harald; Pradhan, Biswajeet; Revhaug, Inge: Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides, 13(2):361–378, 2016.

[Ch14] Chen, Yushi; Lin, Zhouhan; Zhao, Xing; Wang, Gang; Gu, Yanfeng: Deep learning-based classification of hyperspectral data. IEEE Journal of Selected topics in applied earth observations and remote sensing, 7(6):2094–2107, 2014.

[Ch16] Chen, Yushi; Jiang, Hanlu; Li, Chunyang; Jia, Xiuping; Ghamisi, Pedram: Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing, 54(10):6232–6251, 2016.

[Ch20] Chowdhary, KR: Fundamentals of Artificial Intelligence. Springer, 2020.

[CZH16] Cheng, Gong; Zhou, Peicheng; Han, Junwei: Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 54(12):7405–7415, 2016.

[Fe15] Feng, Xiao; Li, Qi; Zhu, Yajie; Hou, Junxiong; Jin, Lingyan; Wang, Jingjie: Artificial neural networks forecasting of PM2. 5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmospheric Environment, 107:118–128, 2015.

[Fi17] Fisher, Douglas: A selected summary of AI for computational sustainability. In: Proceedings of the AAAI Conference on Artificial Intelligence. volume 31, 2017.

[FW19] Fidler, Fiona; Wilcox, John: Reproducibility of Scientific Results. In (Zalta, Edward N., ed.): The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Winter 2019 edition, 2019.

[GC06] Grivas, G; Chaloulakou, A: Artificial neural network models for prediction of PM10 hourly concentrations, in the Greater Area of Athens, Greece. Atmospheric environment, 40(7):1216–1229, 2006.

[GD99] Gardner, MW; Dorling, SR: Neural network modelling and prediction of hourly NOx and NO2 concentrations in urban air in London. Atmospheric Environment, 33(5):709–719, 1999.

[Gh09] Ghobadian, Barat; Rahimi, Hadi; Nikbakht, AM; Najafi, Gholamhassan; Yusaf, TF: Diesel engine performance and exhaust emission analysis using waste cooking biodiesel fuel with an artificial neural network. Renewable energy, 34(4):976–982, 2009.

[Go09] Gomes, Carla P: Computational sustainability: Computational methods for a sustainable environment, economy, and society. The Bridge, 39(4):5–13, 2009.

[Go19] Gomes, Carla; Dietterich, Thomas; Barrett, Christopher; Conrad, Jon; Dilkina, Bistra; Ermon, Stefano; Fang, Fei; Farnsworth, Andrew; Fern, Alan; Fern, Xiaoli et al.: Computational sustainability: Computing for a better world and a sustainable future. Communications of the ACM, 62(9):56–65, 2019.

[Gu12] Guo, Zhenhai; Zhao, Weigang; Lu, Haiyan; Wang, Jianzhou: Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renewable Energy, 37(1):241–249, 2012.

[HDT02] Huang, Chengquan; Davis, LS; Townshend, JRG: An assessment of support vector machines for land cover classification. International Journal of remote sensing, 23(4):725–749, 2002.

[HGS95] Hsu, Kuo-lin; Gupta, Hoshin Vijai; Sorooshian, Soroosh: Artificial neural network modeling of the rainfall-runoff process. Water resources research, 31(10):2517–2530, 1995.

[Hi95] Hilty, LM; Page, Bernd; Radermacher, FJ; Riekert, W-F: Environmental informatics as a new discipline of applied computer science. In: Environmental Informatics, pp. 1–11. Springer, 1995.

[HKH04] Hamed, Maged M; Khalafallah, Mona G; Hassanien, Ezzat A: Prediction of wastewater treatment plant performance using artificial neural networks. Environmental Modelling & Software, 19(10):919–928, 2004.

[Je19] Jetzke, Tobias; Richter, Stephan; Ferdinand, Jan-Peter; Schaat, Samer: Künstliche Intelligenz im Umweltbereich. Umweltbundesamt, 2019.

[Ka09] Karatepe, E; Hiyama, T et al.: Artificial neural network-polar coordinated fuzzy controller based maximum power point tracking control under partially shaded conditions. IET Renewable Power Generation, 3(2):239–253, 2009.

[Ka20] Kayid, Amr: The role of Artificial Intelligence in future technology. 2020.

[KS02] Kemper, Thomas; Sommer, Stefan: Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. Environmental science & technology, 36(12):2742–2747, 2002.

[KS07] Khandelwal, Manoj; Singh, TN: Evaluation of blast-induced ground vibration predictors. Soil Dynamics and Earthquake Engineering, 27(2):116–125, 2007.

[Le04] Lee, Saro; Ryu, Joo-Hyung; Won, Joong-Sun; Park, Hyuck-Jin: Determination and application of the weights for landslide susceptibility mapping using an artificial neural network. Engineering Geology, 71(3-4):289–302, 2004.

[Li21] Liesch, Tanja; Bruns, Julian; Abecker, Andreas; Hilbring, Désirée; Karimanzira, Divas; Martin, Tobias; Wagner, Martin; Wunsch, Andreas; Fischer, Thilo: Nitrat-Monitoring 4.0. INFORMATIK 2020, 2021.

[Ma16] Mayr, Andreas; Klambauer, Günter; Unterthiner, Thomas; Hochreiter, Sepp: DeepTox: toxicity prediction using deep learning. Frontiers in Environmental Science, 3:80, 2016.

[MD00] Maier, Holger R; Dandy, Graeme C: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software, 15(1):101–124, 2000.

[MH96] Minns, AW; Hall, MJ: Artificial neural networks as rainfall-runoff models. Hydrological sciences journal, 41(3):399–417, 1996.

[Mo20] Mondal, Bhaskar: Artificial intelligence: state of the art. Recent Trends and Advances in Artificial Intelligence and Internet of Things, pp. 389–425, 2020.

[MRH98] Mohandes, M; Rehman, S; Halawani, TO: Estimation of global solar radiation using artificial neural networks. Renewable energy, 14(1-4):179–184, 1998.

[Na04] Nayak, Puma C; Sudheer, KP; Rangan, DM; Ramasastri, KS: A neuro-fuzzy computing technique for modeling hydrological time series. Journal of Hydrology, 291(1-2):52–66, 2004.

[NSD04] Nadal, M; Schuhmacher, M; Domingo, JL: Metal pollution of soils and vegetation in an area with petrochemical industry. Science of the total environment, 321(1-3):59–69, 2004.

[PL10] Pradhan, Biswajeet; Lee, Saro: Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling. Environmental Modelling & Software, 25(6):747–759, 2010.

[PPG12] Pourghasemi, Hamid Reza; Pradhan, Biswajeet; Gokceoglu, Candan: Application of fuzzy logic and analytical hierarchy process (AHP) to landslide susceptibility mapping at Haraz watershed, Iran. Natural hazards, 63(2):965–996, 2012.

[RN10] Russell, Stuart J; Norvig, Peter: Artificial Intelligence: A Modern. Approach, 2nded., PrenticeHall, December2002, 2010.

[Ro12] Rodriguez-Galiano, Victor Francisco; Ghimire, Bardan; Rogan, John; Chica-Olmo, Mario; Rigol-Sanchez, Juan Pedro: An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing, 67:93–104, 2012.

[RS20] Raghu, Maithra; Schmidt, Eric: A survey of deep learning for scientific discovery. arXiv preprint arXiv:2003.11755, 2020.

[Sc19] Schuchmann, Sebastian: Analyzing the Prospect of an Approaching AI Winter. PhD thesis, PhD thesis, 05 2019, 2019.

[SD18] SDG, UN: , Sustainable development goals, 2018.

[Sh18] Shen, Chaopeng: A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resources Research, 54(11):8558–8593, 2018.

[SMJ02] Shahin, Mohamed A; Maier, Holger R; Jaksa, Mark B: Predicting settlement of shallow foundations using neural networks. Journal of Geotechnical and Geoenvironmental Engineering, 128(9):785–793, 2002.

[So07] Sousa, SIV; Martins, Fernando Gomes; Alvim-Ferraz, MCM; Pereira, Maria C: Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environmental Modelling & Software, 22(1):97–103, 2007.

[VBB13] Valipour, Mohammad; Banihabib, Mohammad Ebrahim; Behbahani, Seyyed Mahmood Reza: Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of hydrology, 476:433–441, 2013.

[VEW13] Van Eck, Nees Jan; Waltman, Ludo: VOSviewer manual. Leiden: Univeristeit Leiden, 1(1):1–53, 2013.

[Wa09] Wang, Wen-Chuan; Chau, Kwok-Wing; Cheng, Chun-Tian; Qiu, Lin: A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. Journal of hydrology, 374(3-4):294–306, 2009.

[YBL11] Yue, Tao; Briand, Lionel C; Labiche, Yvan: A systematic review of transformation approaches between user requirements and analysis models. Requirements engineering, 16(2):75–99, 2011.

[YD08] Yetilmezsoy, Kaan; Demirel, Sevgi: Artificial neural network (ANN) approach for modeling of Pb (II) adsorption from aqueous solution by Antep pistachio (Pistacia Vera L.) shells. Journal of hazardous materials, 153(3):1288–1300, 2008.

[Yi09] Yilmaz, Işık: Landslide susceptibility mapping using frequency ratio, logistic regression, artificial neural networks and their comparison: a case study from Kat landslides (Tokat—Turkey). Computers & Geosciences, 35(6):1125–1138, 2009.

[Yu20] Yu, Manzhu; Bambacus, Myra; Cervone, Guido; Clarke, Keith; Duffy, Daniel; Huang, Qunying; Li, Jing; Li, Wenwen; Li, Zhenlong; Liu, Qian et al.: Spatiotemporal event detection: a review. International Journal of Digital Earth, 13(12):1339–1365, 2020.

[Za19] Zaldivar, Andrew; Hutchinson, Ben; Spitzer, Elena; Raji, Inioluwa Deborah; Vasserman, Lucy; Mitchell, M.; Barnes, Parker; Wu, Simone Sanoian McCloskey; Gebru, Timnit, eds. Model Cards for Model Reporting, 2019.

[Zo08] Zorlu, K; Gokceoglu, C; Ocakoglu, F; Nefeslioglu, HA; Acikalin, SJEG: Prediction of uniaxial compressive strength of sandstones using petrography-based models. Engineering Geology, 96(3-4):141–158, 2008.