frontiers | Frontiers in Energy Research

TYPE Opinion
PUBLISHED 16 January 2023
DOI 10.3389/fenrg.2022.1073976

Statistical machine learning techniques of weather simulation for the fishery-solar hybrid systems

Chunyu Zhang[1,2,3,4], Xueqian Fu[1,2,3,4]* and Xianping Wu[1,2,3,4]

[1]College of Information and Electrical Engineering, China Agricultural University, Beijing, China, [2]National Innovation Center for Digital Fishery, China Agricultural University, Beijing, China, [3]Key Laboratory of Smart Farming Technologies for Aquatic Animal and Livestock, Ministry of Agriculture and Rural Affairs, Beijing, China, [4]Beijing Engineering and Technology Research Center for Internet of Things in Agriculture, Beijing, China

Check for updates

OPEN ACCESS

EDITED BY
Jiehui Zheng,
South China University of Technology, China

REVIEWED BY
Zhang Hongwei,
China University of Mining and Technology, China
Yixun Xue,
Taiyuan University of Technology, China

*CORRESPONDENCE
Xueqian Fu,
fuxueqian@cau.edu.cn

SPECIALTY SECTION
This article was submitted to Energy Storage,
a section of the journal
Frontiers in Energy Research

RECEIVED 19 October 2022
ACCEPTED 04 November 2022
PUBLISHED 16 January 2023

CITATION
Zhang C, Fu X and Wu X (2023), Statistical machine learning techniques of weather simulation for the fishery-solar hybrid systems.
Front. Energy Res. 10:1073976.
doi: 10.3389/fenrg.2022.1073976

KEYWORDS

fishery-solar hybrid system, photovoltaic power, statistical machine learning, weather simulation, generative adversarial networks

1 Introduction

The economically developed and densely population-rich southeastern part of China has always been a load center for electricity. However, due to the high population density, land resources in southeastern China are tight, making it impossible to build large-scale contiguous ground-mounted photovoltaics (PV) plants as in western China (Fu, 2022a). Therefore, distributed PV has become the preferred choice in southeastern China. Combining the characteristics of coastal and wetlands of rivers and lakes, a new concept of the fishery-solar hybrid system is proposed, which is a new model of distributed PV combined with the fishery, that is, the photovoltaic panel array is set up above the water surface of the fish pond, and the water below the photovoltaic panels can be used for fish and shrimp farming, and the photovoltaic array can also provide good shading for fish farming (Fu, 2022b). Through the clean, efficient, low-carbon innovation model, we can improve the added value of land, and also achieve complementary development between multiple industries. The project is doubly beneficial to achieve the carbon peaking and carbon neutrality goal and economic development.

Fishery meteorological services play an important role in the production regulation of fishery. At the same time, fishery weather also affects PV power generation in the fishery-solar hybrid system. In other words, weather can directly affect the sources and loads of the fishery energy network, and weather sensitivity of the energy network is inevitable (Fu, 2022a). Therefore, how to accurately model fishery weather is significant. In general, traditional methods for generating weather scenarios can usually be divided into three types, which are the fixed-date method, the shifted-date method, and the bootstrap method (Xie and Hong, 2018). The fixed-date method is to learn the profile of the given weather data. The shifted-date method is based on the fixed-date method, which shifts the profile of historical weather data forward to generate weather data with shifted profile features. The bootstrap method is a computational inference method that relies on the resampling of the dataset. Most of the above methods are based on probabilistic models, which cannot adequately portray the complex and high-dimensional characteristics of weather data (Fu et al., 2022). Meanwhile, most of the above methods do not consider the correlation between weather variables (Sun et al., 2020). Compared to probabilistic models, Statistical Machine Learning (SML) based models are well-controllable and scalable. Legasa and Gutirrrez (2020) used Bayesian networks to approximate the probability distribution of observed multivariate (multi-location) weather data. Fu and Niu (2022) used nonlinear regression models, seasonal autoregressive models, and Markov chains to simulate temperature and solar radiation data respectively. Klemmer et al. (2021) proposed a GAN (Generative Adversarial Network) based method to generate spatio-temporal weather scenarios under extreme events. By combining the interpretability of probability theory with the autonomous learning capability of machine learning, the above SML method effectively deals with the simulation of weather data of high dimensional complexity (Fu et al., 2020), which is of guiding significance for us to carry out the simulation of fishery weather data.

2 Relationship between weather and fishery-solar hybrid system

Weather directly affects the sources and loads of the fishery energy network in the fishery-solar hybrid system. For sources, temperature and solar radiation can directly affect photovoltaic power generation. For loads, different meteorological conditions directly affect the physiological habits of fish. We will elaborate on both sides in Sections 2.1, Sections 2.2, respectively.

2.1 How weather affects fishery PV

Solar radiation directly affects the magnitude of the output power of photovoltaic cells, which are the main components of photovoltaic power generation. In addition, the power output of the PV cell is directly related to the temperature of the PV cell itself. Solar radiation and temperature are the direct factors that affect PV cells, this is because it affects the temperature drift of PV cells and thus has a great impact on the efficiency of PV cells.

The above are the direct effects of weather on PV power generation, but in the context of fishery production, there are other effects of weather on PV power generation. As the PV power modules are built on top of the water surface of the fish pond, the specific heat of water is larger than that of soil, and the distance between the PV modules is larger than that of traditional power stations, which creates good sunlight, cooling, and ventilation environment for the PV power module group. Therefore, it is conducive to prolonging the life of the PV power modules and improving the power generation efficiency of the PV power modules. However, as PV modules are always in a hygrometric environment, water vapor will easily accumulate on the PV modules and will easily erode the PV set, which will lead to a serious loss of PV power generation. As a result, we may also increase the maintenance cost of PV modules in the context of the fishery-solar hybrid system.

2.2 How weather affects electricity consumption in the fishery industry

Fish are typically poikilothermal animals. Low temperatures affect the reproduction of fish and even freeze to death. High temperatures can affect the hatching rate of fish eggs and may also lead to outbreaks of certain bacterial or viral infectious diseases. The heat pump is an important piece of equipment to ensure the constant temperature of the aquatic environment. In recent years, air-source heat pumps have been commonly used in the fishing industry because of their energy efficiency and stable operation. An air-source heat pump is a form of heat pump that uses outdoor air as the heat source, and the equipment is used to heat up the aquaculture water through heat exchange with the outdoor air. The heat pump can control the water temperature in the appropriate range, thus effectively improving the survival rate and breeding density, which strongly improves the economic benefits of farmers.

Precipitation can also have a significant impact on fishery production. Heavy rain may cause a sharp rise in the water level of ponds, increasing the risk of pond overflow. Continuous heavy rainfall may also impede convection up and down the water column, causing illness and even death of fish due to lack of oxygen. Fishermen need to use pumps to exchange water from fish ponds with water from rivers, lakes, or reservoirs so that the water level in fish ponds remains within a reasonable range, thus coping with the economic impact caused by precipitation.

The oxygen content of the water is one of the main factors affecting the growth of fish. When the oxygen content of the water is too low, most of the fish will have slow growth and even leading to morbidity or death. Therefore, oxygenators have become essential equipment for aquaculture ponds. The oxygenator increases the oxygen content of fish pond water by increasing the contact area between the water and the air that infiltrates the oxygen from the air into the water. Fishermen need to deploy oxygenators in time to reduce the impact of low water oxygen levels on fish growth.

Most of the current fishery equipment is intelligent, capable of intelligent regulation and intelligent control, thus ensuring the appropriate environment for fish growth even when the weather changes. Therefore, guaranteeing a stable power supply plays a key role in guaranteeing a normal growing water environment for fish.

3 SML model for weather simulation of fishery-solar hybrid system

We have introduced the drawbacks of traditional weather simulation algorithms in the previous section. Most of the

TABLE 1 Comparison of weather simulation algorithms.

Classification	Algorithm	References	Advantage	Disadvantage
Traditional algorithm	First-order homogeneous Markov model	Tseng et al. (2020)	Preserve seasonal characteristics of the variables	Low time resolution
	Hidden Markov model	Ahn and Steinschneider, (2019)	Weather type is introduced as a latent variable and the types fit the data well	Model is simple and cannot reproduce the complexity of the data
	Resampling method	Verdin et al. (2018)	Multi-variable weather data can be generated	Potential changes in extreme weather are difficult to obtain
	Block bootstrap method	Zhao et al. (2019)	Reproduce the spatial variability of precipitation	Generated weather scenarios rely on the data itself and block size
SML algorithm	GAN	Chan et al. (2021)	Direct learning of the distribution characteristics of the data	Prone to pattern collapse and uncontrolled process of generating data
	Wasserstein GAN	Ming et al. (2020)	Ensure the diversity of data generated by GAN.	
	Conditional GAN	Loey et al. (2020)	Make the GAN generation process controllable	
	RNN	Luo et al. (2020)	Fully learn the time-series nature of weather data	Easy to fall into gradient disappearance, long iteration time
	Transformer	Liu et al. (2021)	Learn global features of time-series data directly, fast training	

traditional modeling methods rely on sophisticated probabilistic analysis, which is highly interpretable but has limitations in the face of complex and high-dimensional data. Machine learning algorithms represented by deep learning have powerful autonomous learning capabilities, the results of weather simulation are better than traditional probabilistic models, but internally they can be approximated as a black box model with poor interpretability. As an integrator of both, SML combines the advantages and becomes a powerful tool to deal with high-dimensional complex data and can better handle the simulation of weather data.

GAN, a generative model that has emerged in recent years, plays an important role in the simulation of various types of data. GAN takes full advantage of the feedforward properties of neural networks to learn the distribution characteristics of the weather data directly, without explicitly specifying the probability distribution function as in the case of probabilistic weather models. Meanwhile, the neural networks in GAN can actively establish functional relationships for data, effectively solving the problem of difficulty in modeling multivariate weather data in traditional probabilistic models. Some variants of GAN can also be effectively applied in weather simulation. Traditional GAN is not pre-modeled, and this poses the problem that GAN is uncontrolled in the process of generating data. Given the scenarios of fishery-solar hybrid system applications, we may need some weather data for specific situations, such as temperature data under high-temperature conditions or solar radiation data under cloudy conditions. For this application requirement, we can use Conditional GAN (CGAN) to simulate weather data by introducing conditional variables in the modeling of both generators and discriminators. The added conditional variables can guide the data generation process (Loey et al., 2020).

For weather data, we also need to design a specific neural network based on the characteristics of the weather data (Lim and Zohren, 2021). Weather data are typically time-series data, and for this feature, we generally use Recurrent Neural Networks (RNN) or Long Short-Term Memory (LSTM) for modeling. However, in recent years, a model that has received a lot of attention in the field of natural language processing has emerged, namely the transformer. Transformer, like RNN, is a model with obvious advantages for processing time-series data, but it can directly obtain the global information of the data, while RNN requires one layer of recursion to obtain the global information, which effectively solves the criticism of slow training speed of RNN (Karita et al., 2019). Given the advantage of the fast-training speed of the transformer, the transformer can be increased to a very deep depth, so it can fully exploit the characteristics of a deep neural network and improve the realism of weather data simulation.

Table 1 shows some of the traditional probabilistic methods used for weather simulation and some SML algorithms with applications. Probabilistic models need to have a rigorous derivation process for different weather variables, the model capacity is small, and the numerical features can only capture local data features, which cannot fully portray the complex high-dimensional features of weather data. At the same time, probabilistic methods are also a major problem for modeling correlations. However, the SML model has a high capacity to learn the numerical characteristics of the weather more fully, and the neural network actively establishes relationships between the weather variables, which have certain advantages in dealing with weather data modeling. Therefore, we conclude from the comparison that SML is the main direction for our future fishery weather simulation.

4 Conclusion

As a new model with complementary advantages, the fishery-solar hybrid system has natural advantages for development in areas such as southeastern China where land resources are tight but aquaculture is well developed. The obvious advantage of the fishery-solar hybrid system is that it can make comprehensive use of land and increase the economic return per unit of land. The electricity generated by PV can also be used directly for electricity for the fishery, reducing the cost of aquaculture.

In the context of the fishery-solar hybrid system, weather conditions become an important factor affecting PV and fishery production in multiple directions. Therefore, how to model fishery weather accurately becomes an important issue for modeling and analysis of the fishery-solar hybrid system. We use the SML method to solve this problem, which combines the interpretability of probability theory and the autonomous learning capability of machine learning to effectively handle the high-dimensional complexity of weather data. Compared to traditional probabilistic models, SML can more fully capture the numerical characteristics of weather data, while generating more diverse data due to the high model capacity. In addition, we do not need to perform complex modeling work for correlations because the neural network actively makes connections between features.

In today's rapid development of artificial intelligence technology, SML technology represented by deep learning continues to push the boundaries, providing more mindfulness and new methods for weather simulation. At the same time, as computer hardware continues to be upgraded, we are able to handle more high-dimensional and large-scale weather data more easily.

Author contributions

CZ wrote the article. XF conceived the general idea of the article. XW reviewed the article. All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This study is supported by the National Natural Science Foundation of China under grant 52007193 and The 2115 Talent Development Program of China Agricultural University.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahn, K., and Steinschneider, S. (2019). Seasonal predictability and change of large-scale summer precipitation patterns over the northeast United States. *J. Hydrometeorol.* 20, 1275–1292. doi:10.1175/JHM-D-18-0252.1

Chan, E., Monteiro, M., Kellnhofer, P., Wu, J., and Wetzstein, G. (2021). "pi-GAN: Periodic implicit generative adversarial networks for 3D-aware image synthesis," in 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, une 20 2021 to June 25 2021 (IEEE), 5795–5805. doi:10.1109/CVPR46437.2021.00574

Fu, X., Guo, Q., and Sun, H. (2020). Statistical machine learning model for stochastic optimal planning of distribution networks considering a dynamic correlation and dimension reduction. *IEEE Trans. Smart Grid* 11, 2904–2917. doi:10.1109/TSG.2020.2974021

Fu, X., and Niu, H. (2022). Key technologies and applications of agricultural energy internet for agricultural planting and fisheries industry. *Inf. Process. Agric.* doi:10.1016/j.inpa.2022.10.004

Fu, X. (2022b). Statistical machine learning model for capacitor planning considering uncertainties in photovoltaic power. *Prot. Control Mod. Power Syst.* 7 (1), 5–13. doi:10.1186/s41601-022-00228-z

Fu, X. (2022a). Viewpoints on the experiences and challenges of fishery energy internet. *Front. Energy Res.* 10, 884920. doi:10.3389/fenrg.2022.884920

Fu, X., Zhang, C., and Wu, X. (2022). Statistical machine learning model for uncertainty analysis of photovoltaic power. *Front. Energy Res.* 10, 956543. doi:10.3389/fenrg.2022.956543

Karita, S., Chen, N., Hayashi, T., Hori, T., Inaguma, H., Jiang, Z., et al. (2019). "A comparative study on transformer vs RNN in speech applications," in 2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Sentosa, Singapore, 14-18 December 2019 (IEEE), 449–456. doi:10.1109/ASRU46091.2019.9003750

Klemmer, K., Saha, S., Kahl, M., Xu, T., and Zhu, X. (2021). Generative modeling of spatio-temporal weather patterns with extreme event conditioning. arXiv, abs/2104.12469. doi:10.48550/arXiv.2104.12469

Legasa, M. N., and Gutiérrez, J. M. (2020). Multisite weather generators using bayesian networks: An illustrative case study for precipitation occurrence. *Water Resour. Res.* 56, e2019WR026416. doi:10.1029/2019WR026416

Lim, B., and Zohren, S. (2021). Time-series forecasting with deep learning: A survey. *Phil. Trans. R. Soc. A* 379, 20200209. doi:10.1098/rsta.2020.0209

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). "Swin transformer: Hierarchical vision transformer using shifted windows," in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, Canada, Oct 11, 2021 (IEEE), 9992–10002. doi:10.1109/ICCV48922.2021.00986

Loey, M., Manogaran, G., and Khalifa, N. M. (2020). A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. *Neural comput. Appl.* 2020, 1–13. doi:10.1007/s00521-020-05437-x

Luo, Y., Chen, Z., and Yoshioka, T. (2020). “Dual-path RNN: Efficient Long sequence modeling for time-domain single-channel speech separation,” in ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, May 4-8, 2020 (IEEE), 46–50. doi:10.1109/ICASSP40776.2020.9054266

Ming, Z., Tong, L., Rui, Z., Yahui, T., Mingjing, T., Leilei, L., et al. (2020). Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification. *Inf. Sci.* 512, 1009–1023. doi:10.1016/j.ins.2019.10.014

Sun, M., Feng, C., and Zhang, J. (2020). Probabilistic solar power forecasting based on weather scenario generation. *Appl. Energy* 266, 114823. doi:10.1016/j.apenergy.2020.114823

Tseng, S., Chen, C., and Senarath, S. U. (2020). Evaluation of multi-site precipitation generators across scales. *Int. J. Climatol.* 40, 4622–4637. doi:10.1002/joc.6480

Verdin, A., Rajagopalan, B., Kleiber, W., Podestá, G. P., and Bert, F. E. (2018). A conditional stochastic weather generator for seasonal to multi-decadal simulations. *J. Hydrology* 556, 835–846. doi:10.1016/J.JHYDROL.2015.12.036

Xie, J., and Hong, T. (2018). Temperature scenario generation for probabilistic load forecasting. *IEEE Trans. Smart Grid* 9, 1–1687. doi:10.1109/TSG.2016.2597178

Zhao, Y., Nearing, M. A., and Guertin, D. P. (2019). A daily spatially explicit stochastic rainfall generator for a semi-arid climate. *J. Hydrology* 574, 181–192. doi:10.1016/J.JHYDROL.2019.04.006