Opinions on power grid infrastructure planning with investment risks

Kun Sheng[1], Chelun Xie[1], Jian Sha[2]*, Zhuo Zeng[3], Yuyou Liu[2], Chuanli Lei[2], Liang Chen[2] and Yujie Wu[3]

[1]State Grid Hunan Electric Power Company Limited, Changsha, China, [2]State Grid Hunan Electric Power Company Limited Economic and Technical Research Institute, Changsha, China, [3]College of Electrical and Information Engineering, Hunan University, Changsha, China

1 Introduction

Power grid infrastructure planning schemes are usually affected by various investment risks in the construction duration, such as financial risks (Ammar and Eling, 2015; Zhao et al., 2022), power demand changes (Li et al., 2022; Yang et al., 2022), and extreme climates (Cao et al., 2022). Due to these investment risks, the actual completion date of power grid infrastructure projects may be delayed, and thus the electricity demand for load growth and renewable energy integration cannot be well satisfied (Dang et al., 2019). A few investigations have been reported in the literatures (Pan et al., 2022; Wu et al., 2022) to accurately predict the infrastructure project durations. For example, various prediction models with neural networks and multiple linear regression methods based on historical infrastructure planning data were presented for project duration forecasting (Pan et al., 2022). However, the investment risks affecting the planned infrastructure projects are neglected, and these risks are important factors for the construction duration prediction problem. Therefore, this paper aims to offer insightful opinions and discussions on the power grid infrastructure planning and project duration prediction problems with investment risks.

The main contributions of this paper are twofold as listed: 1) a multi-factor risk evaluation model was formulated based on program evaluation and review technique (PERT) for power grid infrastructure planning under financial risk, extreme climates, and electricity demand changes. Various curve features are extracted from historical infrastructure project schedules using Tsfresh tool, and vital risk features are identified based on a feature filtering method; 2) a project duration prediction method is proposed based on Bayesian neural network (BNN) considering multi-factor investment risks to formulate the occurrence probability of infrastructure project delay, and hence the deferred duration of different infrastructure projects can be quantified for smart grid planning.

2 Risk factors affecting power grid infrastructure planning

In the implementation process of power grid infrastructure planning, there are various internal and external risk factors such as financial risks, electricity demand variations, and extreme weather events (Ammar and Eling, 2015). First, power grid infrastructure planning is greatly affected by financial risks, such as the movement of interest rates, exchange rates, and cash flows. Second, due to the grid integration of massive electric vehicles with charging

behavior uncertainties, electricity demand variations are inevitable (Castillo et al., 2022). Third, there are various weather risks such as hurricanes, cold snaps, rainstorms, heat waves, and thunderstorms battered power grid infrastructure planning (Cao et al., 2022). The occurrence of significant infrastructure project delay may lead to severe financial losses even if the infrastructure project itself does not sustain physical damage. As a result, based on the PERT, to quantify investment risk factors, the impact of each risk factor can be determined through Monte Carlo simulation (Quah and Quek, 2007; Cheng et al., 2021). In summary, the triangular distribution can be given for a multi-factor risk evaluation model to formulate the occurrence probability of financial risks, electricity demand variations, and extreme climates in this paper.

3 Feature extraction of infrastructure project schedules

Infrastructure project schedules can be characterized by investment and construction completion rate curves. Hence, Tsfresh and feature filtering methods are used to identify vital sequential features affecting power grid infrastructure planning from investment and construction completion rate curves (Ehya et al., 2022). Based on hypothesis testing, the feature filtering method is applied to evaluate the correlation between the project duration T and the sequential feature x extracted using the Tsfresh method. First, at the significance level α of 0.05, the null hypothesis indicates that the sequential feature x is not relevant to the project duration T and should not be retained. Second, each sequential feature x is scored by formula (1). The closer scores are to 1, the higher the occurrence probability of infrastructure project delay will be (Chen et al., 2022). Finally, the threshold Q_α can be obtained based on the significance level and the number of infrastructure projects. If $Q \geq Q_\alpha$, the null hypothesis is rejected, and the sequential feature x is retained. Otherwise, the null hypothesis is accepted, and the sequential feature x is removed. The correlation degree between project duration T and sequential feature x is calculated as follows:

$$Q(x,T) = \begin{cases} \max\left|F_{1,n}(T) - F_{0,n}(x)\right|, & \text{if } x \text{ is a Boolean type variable} \\ \dfrac{2[N_1(x,T) - N_2(x,T)]}{n(n-1)}, & \text{if } x \text{ is a continuous}-\text{valued type variable,} \end{cases}$$

$$(1)$$

where $F_{0,n}(x)$ and $F_{1,n}(T)$ are two cumulative distributions of the extracted curve feature and infrastructure project duration; n is the number of extracted curve features; $N_1(x,T)$ and $N_2(x,T)$ are the number of positively correlated element pairs and the number of negatively correlated element pairs, respectively.

Hence, more than 200 vital sequential features consisting of dynamic time-warping (DTW) distance, linear least-squares regression, time lagged cross correlation (TLCC), Spearman correlation coefficient, and enclosed area between investment and construction completion rate curves can be identified based on the feature filtering method for the project duration prediction method. Furthermore, in order to investigate the coordination degree of investment and construction completion schedules, the curve coordination degree of sequential features is proposed as follows:

$$K = \lambda_1 \cdot x_{DTW} + \lambda_2 \cdot x_{TLCC} + \lambda_3 \cdot x_\rho + \lambda_4 \cdot x_{AR}, \qquad (2)$$

where x_{DTW} denotes DTW distance between investment and construction completion rate curves; x_{TLCC} represents the degree of the construction completion rate curve lagging behind the investment completion rate curves; x_ρ denotes the Spearman correlation coefficient; x_{AR} denotes the enclosed area between investment and construction completion rate curves; and $\lambda_1 - \lambda_4$ denote corresponding weights which can be obtained by the coefficient of the variation method. In this article, their values are 27.4%, 25.9%, 18.7%, and 28.1%, respectively.

Meanwhile, it can be found that infrastructure projects with critical features, including higher voltage levels, larger planning capacities, location in backcountry areas, and hysteretic investment, tend to result in higher delay risks.

4 Deep learning-driven project duration prediction with investment risks

Due to various investment risk factors affecting the power grid infrastructure planning, the project duration prediction problem can be transformed into a probabilistic forecasting model. Consequently, a deep learning-driven Bayesian method (Jia et al., 2021) is proposed to formulate the occurrence probability distribution of infrastructure project delay and quantify the deferred duration of different projects under diverse investment risks. The BNN algorithm can learn from the historical data of project duration and respond to changes in future variables with credibility. With conducting deep learning model training, the internal relationship between risk factors and infrastructure duration from historical data can be found efficiently.

The BNN structure is made up of input layer, hidden layer, and output layer. The hidden layer is a probability layer, while the weight ω in the probability layer obeys Gaussian distribution with mean μ and variance δ (Thiagarajan et al., 2022). The BNN-driven power grid infrastructure project duration prediction process with investment risks can be divided into the following steps:

- implementing the tri-layer BNN structure to formulate a deep learning-driven project duration prediction model with identified sequential features and investment risk factors as input data.
- training the BNN prediction model for exploring the non-linear relationship between sequential feature input data and project duration to determine the posterior distribution of weights.
- using historical project delay data of power grid infrastructure planning to verify the validity of the proposed model and obtain the duration delay probability distribution results.

With the proposed project duration prediction model, the power grid infrastructure projects data $D = (X, T)$ are given to train the project duration prediction model. X is the input data consisting of project labels, investment risk factors, curve coordination degree, and sequential features extracted from investment and construction
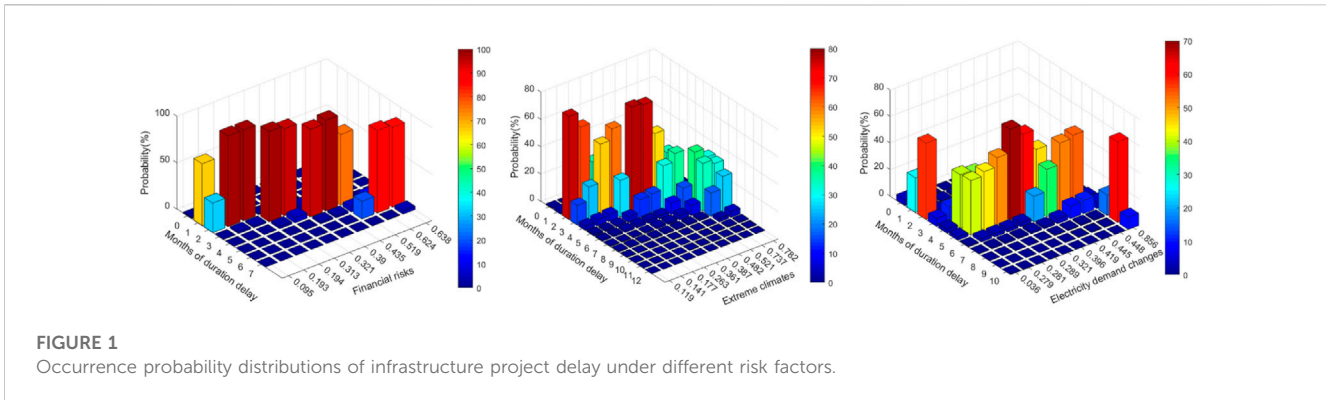
FIGURE 1
Occurrence probability distributions of infrastructure project delay under different risk factors.

completion rate curves. Project labels include project location, planning capacity, and theoretical duration. While T denotes the duration of power grid infrastructure projects. The triangular distribution parameters of financial risks, extreme climates and electricity demand variations are (0, 0.3, 1), (0, 0.2, 1), and (0, 0.4, 1), respectively.

Variational inference is introduced into network training to obtain the posterior distribution $P(\omega|D)$ of BNN weights suitable for duration prediction under different investment risks (Zhang et al., 2019). An existing simple distribution $q(\omega|\theta)$ is introduced to approximate the weight posterior distribution $P(\omega|D)$ by minimizing the difference between the two distributions. The weight distribution can be optimized by minimizing KL divergence, as follows:

$$\mathcal{L} = \arg\min_{\theta} KL[q(\omega|\theta)\|P(\omega|D)]$$
$$= \arg\min_{\theta} \mathbb{E}_{q(\omega|\theta)}\left[\log\left[\frac{q(\omega|\theta)}{P(D|\omega)P(\omega)}\right]\right], \qquad (3)$$

where $\theta = (\mu, \sigma)$ and each weight ω_i obeys a standard normal distribution (μ_i, σ_i); $P(\omega)$ is the prior distribution of the weight under different investment risks.

In this paper, historical power grid infrastructure projects in Hunan, China, are presented to validate the effectiveness of the proposed project duration prediction method with multi-factor investment risks. The posterior distribution of weights in the BNN-based duration prediction model varies with the probability of risk occurrence. Then, the occurrence probability distributions of project duration delay under different investment risks can be extracted from a large amount of sampling data from Monte Carlo analysis, and the resulting project delay probability distributions are shown in Figure 1.

In Figure 1, it can be found that number of delay months is increased with the growing investment risk. On the other hand, the delay months of infrastructure projects under financial risk usually have a small standard deviation, and the project delay probability under extreme climate is the highest. The results show that the number of delay months can reach 10 months, and this is because electricity demand changes will lead to variation in design processes and technologies.

5 Discussion and conclusion

Based on the statistical analysis of practical infrastructure projects data in Hunan, China, the following are the key findings of this study: 1) the project delay is more likely to occur in backcountry, and the delay rate can be up to 43%; 2) risk of infrastructure project delay rises as the voltage level increases. The delay rates with voltage levels 35 kV, 110 kV and 220 kV are 26%, 29%, and 35%, respectively; 3) timely investment will accelerate the construction rate of infrastructure projects, thereby reducing the probability of delay; 4) the larger the planning capacity is, the lower the degree of timely investment will be. Risk of infrastructure project delay decreases as the curve coordination degree increases; and 5) further research will focus on the diversity of investment risk factors and combined effects of multiple risk factors on power grid infrastructure planning.

Author contributions

Conceptualization, KS and CX; data curation, JS and YL; formal analysis, ZZ; methodology, YW; visualization, CL and LC; writing—original draft, JS and ZZ; writing—review and editing, KS.

Funding

Conflict of interest

Authors KS and CX were employed by State Grid Hunan Electric Power Company Limited. Authors JS, YL, CL, and LC were employed by State Grid Hunan Electric Power Company Limited Economic and Technical Research Institute.

The remaining 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

Ammar, S. B., and Eling, M. (2015). Common risk factors of infrastructure investments. *Energy Econ.* 49, 257–273. doi:10.1016/j.eneco.2015.01.021

Cao, Y., Zhou, B., Chung, C. Y., Shuai, Z., Hua, Z., and Sun, Y. (2022). Dynamic modelling and mutual coordination of electricity and watershed networks for spatio-temporal operational flexibility enhancement under rainy climates. *IEEE Trans. Smart Grid*, 1. doi:10.1109/TSG.2022.3223877

Castillo, Z. V., de Boer, H.-S., Muñoz, R. M., Gernaat, D. E. H. J., Benders, R., and van Vuuren, D. (2022). Future global electricity demand load curves. *Energy* 258, 124741. doi:10.1016/j.energy.2022.124741

Chen, J., Yang, M., Gong, W., and Yu, Y. (2022). Multi-neighborhood guided kendall rank correlation coefficient for feature matching. *IEEE Trans. Multimedia*, 1–15. doi:10.1109/TMM.2022.3217410

Cheng, F., Lin, M., Yüksel, S., Dinçer, H., and Kalkavan, H. (2021). A hybrid hesitant 2-tuple IVSF decision making approach to analyze PERT-based critical paths of new service development process for renewable energy investment projects. *IEEE Access* 9, 3947–3969. doi:10.1109/ACCESS.2020.3048016

Dang, C., Wang, X., Shao, C., and Wang, X. (2019). Distributed generation planning for diversified participants in demand response to promote renewable energy integration. *Prot. Control Mod. Power Syst.* 7, 1559–1572. doi:10.1007/s40565-019-0506-9

Ehya, H., Skreien, T. N., and Nysveen, A. (2022). Intelligent data-driven diagnosis of incipient interturn short circuit fault in field winding of salient Pole synchronous generators. *IEEE Trans. Industrial Inf.* 18 (5), 3286–3294. doi:10.1109/TII.2021.3054674

Jia, X., Yang, J., Liu, R., Wang, X., Cotofana, S. D., and Zhao, W. (2021). Efficient computation reduction in bayesian neural networks through feature decomposition and memorization. *IEEE Trans. Neural Netw. Learn. Syst.* 32 (4), 1703–1712. doi:10.1109/TNNLS.2020.2987760

Li, X., Liu, C., Sheng, K., Wen, B., Xie, H., Zhao, W., et al. (2022). A multi-level investment demand assessment model for renewable-dominated power system planning. *Front. Energy Res.* 9. doi:10.3389/fenrg.2021.807590

Pan, Y., Chen, S., Ma, T., Gao, J., Wang, C., Liu, X., et al. "Research on project duration prediction based on multiple linear regression," in 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), Nanjing, China, May, 2022, 1606–1612. doi:10.1109/CIEEC54735.2022.9846374

Quah, K. H., and Quek, C. (2007). Mces: A novel Monte Carlo evaluative selection approach for objective feature selections. *IEEE Trans. Neural Netw.* 18 (2), 431–448. doi:10.1109/TNN.2006.887555

Thiagarajan, P., Khairnar, P., and Ghosh, S. (2022). Explanation and use of uncertainty quantified by bayesian neural network classifiers for breast histopathology images. *IEEE Trans. Med. Imaging* 41 (4), 815–825. doi:10.1109/TMI.2021.3123300

Wu, Y., Li, X., Zhang, L., Liu, C., Zhao, W., and Zhang, T. (2022). Machine learning-driven deduction prediction methodology for power grid infrastructure investment and planning. *Front. Energy Res.* 10. doi:10.3389/fenrg.2022.893492

Yang, Q., Li, Z., Chen, Y., Zhu, Y., and Dou, Q. (2022). An investment efficiency evaluation model for distribution network with distributed renewable energy resources. *Front. Energy Res.* 10. doi:10.3389/fenrg.2022.931486

Zhang, C., Bütepage, J., Kjellström, H., and Mandt, S. (2019). Advances in variational inference. *IEEE Trans. Pattern Analysis Mach. Intell.* 41 (8), 2008–2026. doi:10.1109/TPAMI.2018.2889774

Zhao, W., Liu, X., Wu, Y., Zhang, T., and Zhang, L. (2022). A learning-to-rank-based investment portfolio optimization framework for smart grid planning. *Front. Energy Res.* 10. doi:10.3389/fenrg.2022.852520