AI-Guided Noise Reduction for Urban Geothermal Drilling

Daniel Ladwig[1], Martin Spitznagel[1], Jan Vaillant[1,2], Klaus Dorer[1], and Janis Keuper[1]

[1] Institute for Machine Learning and Analytics (IMLA)
Offenburg University, Germany
firstname.lastname@hs-offenburg.de
[2] Herrenknecht AG

Abstract. Urban geothermal energy production plays a critical role in achieving global climate objectives. However, drilling operations in densely populated areas generate significant noise pollution, posing challenges to community acceptance and regulatory compliance. This research presents an artificial intelligence-driven approach to dynamically reduce noise emissions during geothermal drilling. We integrate Deep Reinforcement Learning (DRL) with generative neural network models to provide real-time recommendations for optimal drilling parameters. Specifically, the Drill-LSTM model forecasts future machine states, while the Sound-GAN framework predicts sound propagation based on varying operational conditions. These models feed into a DRL-Agent that learns to balance drilling efficiency with noise minimization. Additionally, an interactive assistance system GUI presents predictions, forecasts, and recommendations to human operators, facilitating informed decision-making. Our system demonstrates significant potential in reducing noise levels, enhancing operational efficiency, and fostering greater acceptance of urban geothermal projects. Future work will focus on refining the models and validating the system in real-world drilling scenarios.

Keywords: Geothermal Drilling, Noise Reduction, Deep Reinforcement Learning, Generative Models, AI-Assisted Control

1 Introduction

Urban geothermal energy production is increasingly recognized as a pivotal component in achieving global climate goals. However, the deployment of geothermal technology in urban areas is not without challenges, particularly the significant noise pollution generated during drilling processes. Current methods for mitigating this issue are largely manual and often inadequate for maintaining noise levels within legal urban limits. In densely populated areas, continuous deep drilling operations required for geothermal energy can severely disrupt local communities. Legal requirements often cap noise levels at 35dB during nighttime, posing a substantial challenge given the 24/7 operational needs of these projects. Current solutions, including temporal shifting of operations and physical barriers, provide limited relief.

This research introduces an application of artificial intelligence to overcome the constraints of traditional noise reduction techniques in geothermal drilling. By integrating Deep Reinforcement Learning (DRL) with generative neural network models, we dynamically suggest drilling parameters based on continuous feedback. Our system utilizes two models trained on real-world data: one forecasting noise outcomes and the other predicting drilling scenarios. A DRL model uses these simulations to learn optimal drilling strategies that minimize noise while maintaining drilling efficiency. The system's performance is planned to be further refined through real-world application, aiming to ensure its effectiveness across various urban geothermal sites.

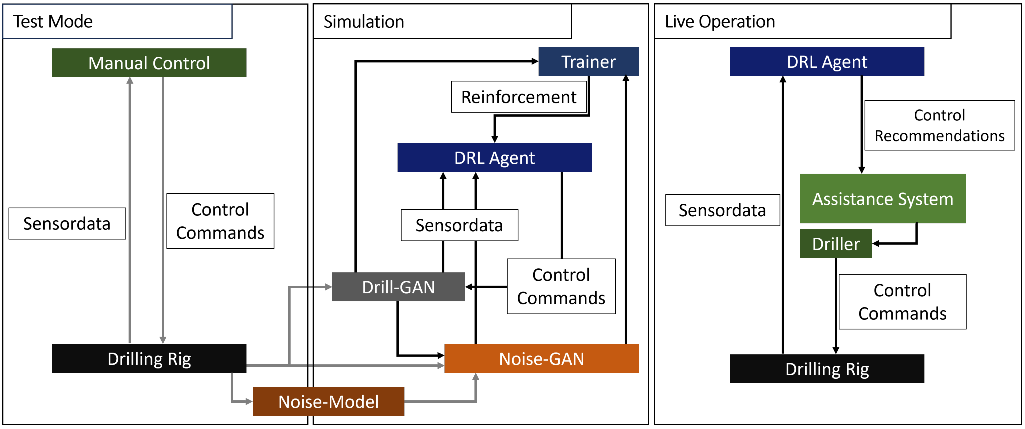


Fig. 1: Architecture of the AI models during the development stages Test Mode, Simulation and Live Operation.

Implementation Prototypes: The project is structured into three distinct prototypes, each illustrated in Figure 1. In *Test Mode*, generative models simulate sound propagation, drilling rig behaviour and predict drilling progress from specific states and commands. In the *Simulation* phase, a DRL model uses input from generative models to learn the control of the simulated drilling rig through reinforcements to provide operation recommendations, which then are validated against real-world scenarios. Finally, in *Live Operation*, the DRL system operates as an assistant that recommends strategic control changes during the drilling process. The human operators stay in control, providing oversight and evaluating AI-recommended commands. The system will be further improved by continuously learning from actual data and human feedback.

2 Related Work

Generative models for complex physical problems. Integrating physical principles into generative models is a rapidly advancing field. Models such as PUGAN [1] and FEM-GAN [2] have successfully merged GANs with physical modeling, enhancing performance in environments governed by complex physical laws. Physics-guided GANs have notably improved efficiency and precision in areas like fluid dynamics and structural system identification by incorporating physics-based loss functions and simulations [3, 4]. Additionally, machine learning has made significant strides in understanding and predicting physical interactions, as demonstrated by models that grasp the dynamics of block towers beyond simple memorization [5], applications such as fall detection through body part tracking [6], and the generation of physically plausible human animations [7]. Furthermore, physics-guided AI approaches, including grey-box models, have effectively incorporated physical laws into model training, enhancing performance and reliability [8, 9].

Generative models and deep reinforcement learning. Integrating generative models with deep reinforcement learning was first introduced by [10] with a case study in the domain of near-field/far-field communication. This work mentions how generative AI could improve deep reinforcement learning considering data and policy. Simulating DRL environments of real-world scenarios with generative models could open up more applications for DRL.

3 Assistance System

The predictions, forecasts and recommendations of the developed models are presented to the driller in a continuously updated assistance system GUI. Figure 2 shows a Demo of the application appearance. The interface is separated into three areas, the status area (top), the analysis area (middle) and the recommendation area (bottom). The status area displays the two most important indicators: the rate of penetration and the sound level. This allows a quick overview on the current state. The analysis area shows the historical, the current and forecast trends of the monitoring variables. These dataplots help to explain a given recommendation to adjust the current machine control parameters and assists the driller in the decision to apply the recommendation or to decline. In particular, it shows the forecast of the machine behaviour with and without accepting the recommendation. The recommendation area displays new control recommendations given by the deep reinforcement learning agent and also allows to give feedback on their quality. A central middleware retrieves the machine data, preprocesses them for the various models, awaits results and synchronously sends updates to the UI.

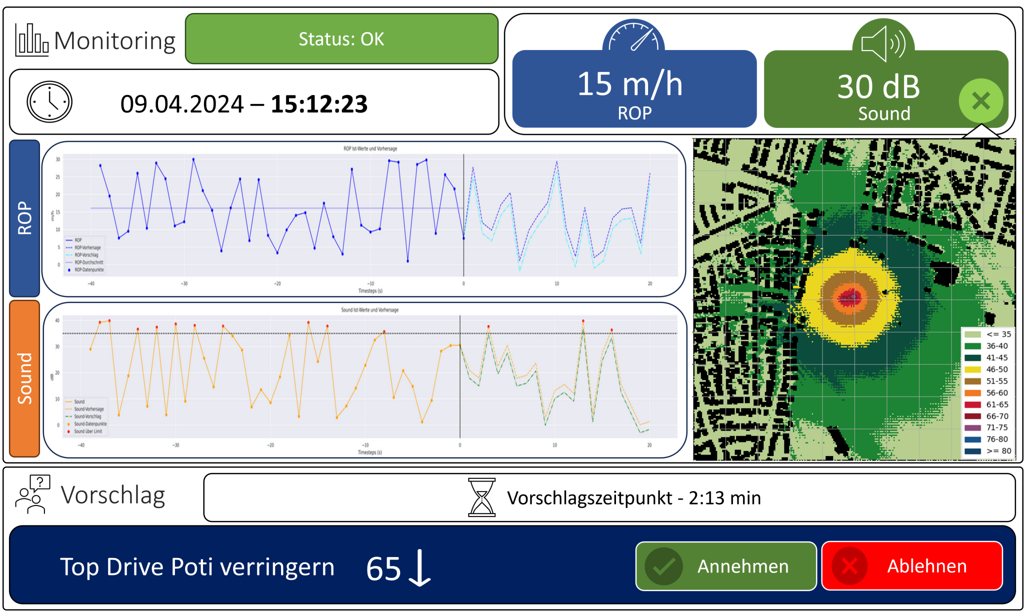


Fig. 2: Assistance system GUI Mockup

4 Drill-LSTM

Accurate prediction of machine states is vital for informed decision-making in drilling operations. The Drill-LSTM model serves as the foundational component for simulating machine behaviour, enabling the DRL agent to understand and anticipate the outcomes of various control actions. In this study, we used a sequence-to-sequence encoder-decoder Long Short-Term Memory (LSTM) network, to predict future machine states based on defined action features. Our dataset comprises about 700 features collected throughout the entire drilling procedure, with one feature intentionally shifted to represent human actions that control machine behaviour. The training dataset encompasses one month of drilling operations, while the test dataset covers one week.

Data preprocessing involved standardizing the logging interval to a fixed 60-second interval and normalizing features. For the forecasting task, the model was trained to predict the next time step based on a defined history window, with current experiments focusing on 1-step forecasting. The model was trained for 50 epochs using the Mean Squared Error (MSE) loss.

Figure 3 visualizes the forecasting performance of the Drill-LSTM model on four anonymised machine features. The model successfully captures general trends. These results establish a foundation for future multi-step forecasting, which will require architectural enhancements to maintain prediction accuracy over longer horizons. The main purpose of the forecasting model is to serve as a machine model for the assistance system, enabling the prediction of future machine states based on potential user actions.

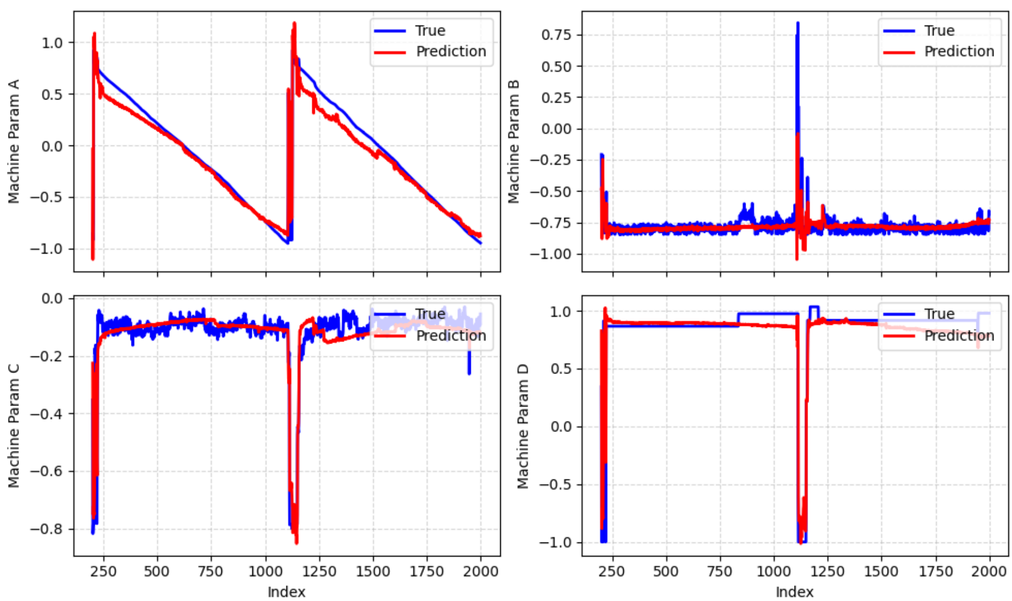


Fig. 3: Forecasting performance of the sequence-to-sequence LSTM model on four anonymised machine features. The true machine parameter values (blue lines) are compared with the predicted values (red lines) over the test dataset.

5 Sound-GAN

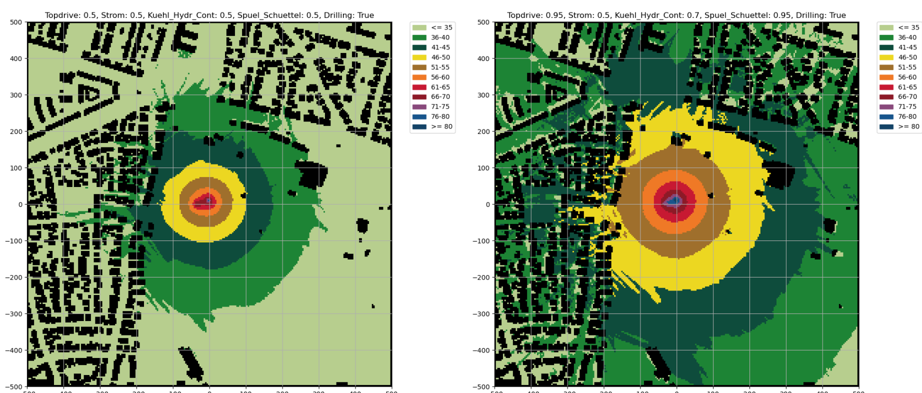
Understanding and predicting sound propagation is essential for mitigating noise pollution during drilling operations. The Sound-GAN framework leverages generative models to efficiently simulate sound distribution, providing real-time feedback that the DRL agent can use to recommend noise-reducing actions to human operators.

Table 1: Model vs. Simulation Performance Comparison for Single Sample Processing. The complex source is a single test sample for a more complex source with 28 descriptive sound signal sources for the simulation, illustrating how the processing time increases significantly with more complex signal sources. It is important to note that this analysis may not provide a completely fair assessment from a theoretical perspective, as no efforts were made to optimize the simulation codes for GPU execution.

Model - Condition	Mean Runtime (s)	Std. Dev. (s)
UNet	0.0126	0.0012
GAN	0.0095	0.0012
Diffusion	4.1560	0.0061
Simulation Single Source - 3rd Order Reflections	186.2295	16.8491
Simulation Machine Setup - Complex Source	540.0000	-

Simulation vs. Generative AI: Building on our previous work, where we evaluated the effectiveness of generative models for predicting sound propagation [11], we generated over 15,000 data samples using the *NoiseModelling v4* framework [12], compliant with *CNOSSOS-EU* standards [13]. Each sample was defined by unique drilling parameters. This simulation data serves as the foundation for three different generative image-to-image models: **GANs** based on [14], **UNets** [15] and **DDPM diffusion** models [16]. Generative models significantly outperform traditional sound propagation simulations in processing speed, achieving up to a $50k$ factor improvement in runtime, with a mean absolute error of 0.55 dB in their predictions.

Simulation Setup: The dataset for this study was generated using the Noise Modelling Framework v4 [12], We systematically altered five key machine parameters, each representing distinct components that influence the noise distribution around a stationary drilling machine. The parameters were anonymised, with load values ranging from 1.0, representing the maximum dB output of a component, to 0.5, indicating a -20 dB linear reduction. Each simulation was based on the drilling machine's fixed location and initial noise measurements.



(a) Simulation with low load settings

(b) Simulation with high load settings

Fig. 4: Comparison of simulated sound propagation maps under different machine settings: (a) low load and (b) high load operational states.

The final dataset comprises over 15,000 data points, each depicting a unique combination of machine settings while keeping the noise source stationary across all simulations. Figure 4 provides a comparison of sound propagation maps under different operational states, with (a) illustrating low load settings and (b) depicting high load settings, both color-coded according to dB levels. This simulation setup allows the model to capture how varying machine loads affect the spatial distribution of noise.

Sound Prediction Results: Building on this simulation setup, we evaluated the performance of three models — UNet, GAN, and DDPM — in predicting sound propagation from a fixed grayscale OpenStreetMap (OSM) input, where buildings are represented by black pixels and open spaces by white pixels. The task was framed as a conditional image-to-image translation, where each model received the same OSM image as input, and the objective was to predict the interpolated sound distribution map based on the machine parameter settings.

Each model architecture was trained for 50 epochs, with 20% of the dataset held out for testing. The UNet and Diffusion models were trained using Mean Squared Error (MSE) loss, while the GAN employed a combination of Binary Cross Entropy (BCE) loss and L1 loss. The evaluation of the models was based on two key metrics: Mean Absolute Error (MAE), which quantifies the average magnitude of prediction errors, and Weighted Mean Absolute Percentage Error (wMAPE), which penalizes incorrect predictions behind and inside of buildings.

Table 2: Evaluation of all architectures on the LoS and NLoS tasks using MAE and wMAPE metrics.

Model	**MAE**	**wMAPE**	**LoS MAE**	**NLoS MAE**	**LoS wMAPE**	**NLoS wMAPE**
UNet	0.70	12.78	0.58	0.85	5.32	21.87
GAN	**0.48**	**3.42**	**0.32**	**0.68**	**1.93**	**5.24**
DDPM	1.19	24.94	1.07	1.35	14.23	37.98

Results: The results of the evaluation are summarized in Table 2, which shows the performance of each model for both Line of Sight (LoS) and Non-Line of Sight (NLoS) regions. The GAN model consistently outperformed the other architectures, achieving the lowest MAE and wMAPE scores in both LoS and NLoS conditions. To further analyze the performance of the GAN model, we generated a heatmap of the wMAPE across the test dataset, as shown in Figure 5. This heatmap provides a pixel-wise visualization of the prediction error. The highest errors occur in NLoS areas.

In addition to the quantitative analysis, Figure 6 provides a visual comparison of the predicted sound distributions from all three models for a single datapoint. The GAN model shows the most accurate spatial distribution of sound. By contrast, the Diffusion model exhibits larger deviations with visible over-prediction, especially around building edges and occluded spaces.

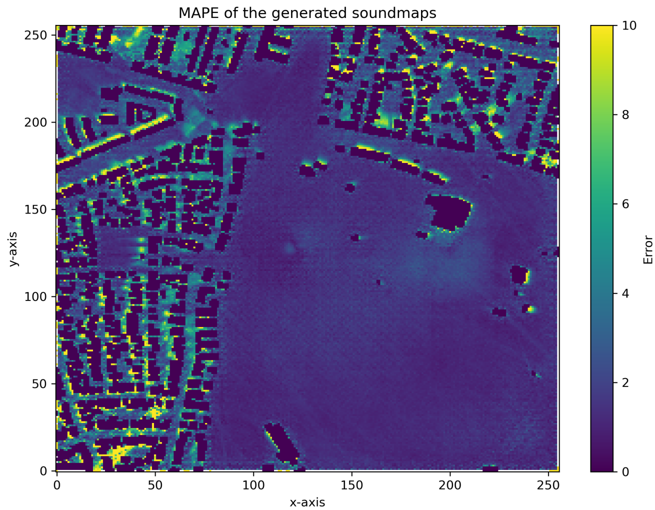


Fig. 5: Heatmap of the Mean Absolute Percentage Error (MAPE) over the entire dataset for the UNet prediction.

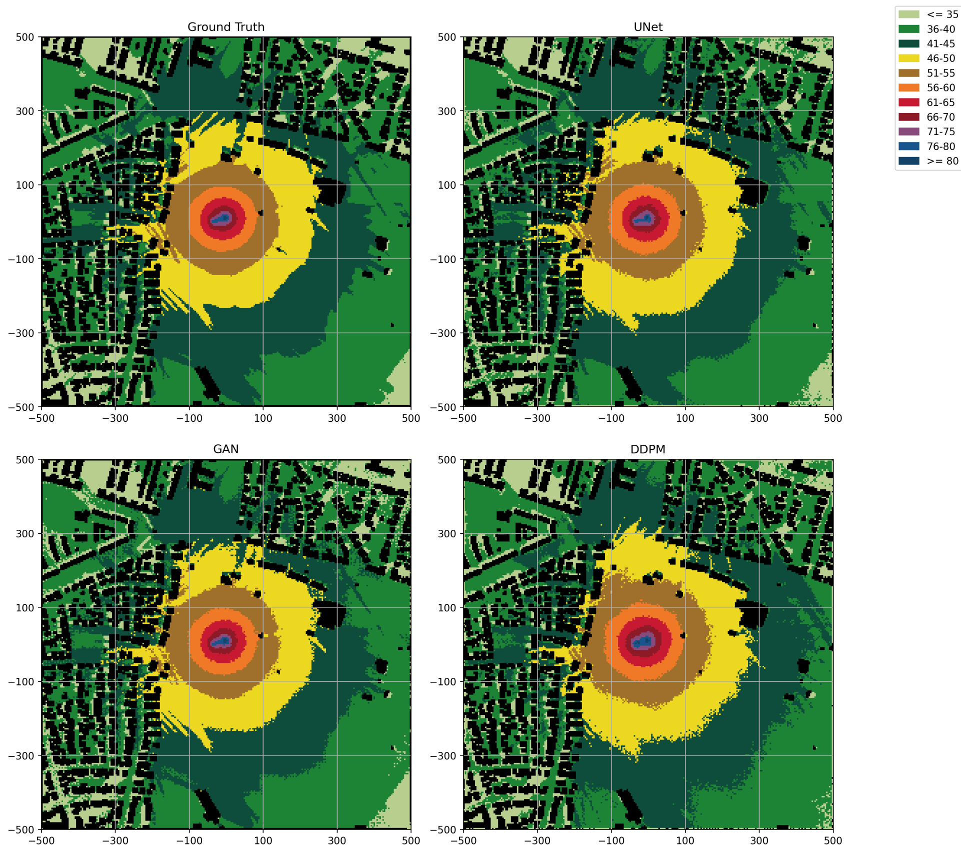


Fig. 6: Comparison of predicted sound distributions for a single datapoint across three models: UNet, GAN, and DDPM. The ground truth simulation is shown in the top left.

6 DRL-Agent

Recommending beneficial control changes requires an effective control strategy. Deep reinforcement learning will be utilized to train an agent to fulfill this task. Since training on the real drilling machine is no reasonable option, a simulation of the machine (DRILL-LSTM) is being developed to emulate the necessary machine behavior to train on. To be able to monitor and react to the resulting noise of a machine state after applying control changes on the simulated machine, a model named NOISE-GAN is being developed. These two models build the prerequisite for developing a DRL-Agent, which is a part of future work. The combination of generative models with deep reinforcement learning is promising especially for domains where training an agent on the real machine is not feasible.

The second crucial issue to approach is how to recommend control changes in a timely manner that are processable by a human operator. A deep reinforcement agent usually operates with a high frequency and has direct access to the environment to quickly react to changes when necessary. Since the control of the drilling machine still depends on a human operator to make the final decision, the drilling recommendations have to fulfill special requirements. The recommendations have to be as sparse as possible to not annoy or distract the driller but with the most impact within a reasonable timeframe. Current recommendations have to be monitored to be dismissed if the elapsed time since occurring or machine state changes turn them ineffective. The upcoming challenge is to produce recommendations for humans based on a high frequency observing DRL-Agent.

7 Conclusions

In this work we have shown that the prediction of sound propagation can be sped up by a factor of 50.000 when using generative AI models to do the prediction. This can be used to effectively train a DRL model to efficiently - in the sense of speed and noise - drill in urban areas. The so reduced sound emissions from geothermal drilling operations will enhance community tolerance and broader acceptance of urban geothermal energy projects. This contributes to the economic viability of such projects and their potential impact on achieving climate targets. Integrating AI to control noise in geothermal drilling presents a transformative solution to one of the most pressing challenges facing urban renewable energy projects, promising not only to reduce noise pollution effectively but also to streamline drilling operations. The advancements in generative AI could open ways to solve optimization problems with DRL where training on the real environment is not an option. In future work the practicality of the presented integration of DRL and the simulated environment with generative AI will be tested on real world data. Furthermore the efficiency of DRL as an assistance system in human processable speed is yet to prove in a real world scenario.

Acknowledgements

The authors acknowledge the financial support by the German Federal Ministry of Education and Research (BMBF) in the program “Forschung an Fachhochschulen in Kooperation mit Unternehmen (FH-Kooperativ)” within the joint project ”KI-Bohrer” under grant 13FH525KX1.

References

1. Cong, R., Yang, W., Zhang, Weiand Li, C., Guo, C., Huang, Qingmingand Kwong, S.: Pugan: Physical model-guided underwater image enhancement using gan with dual-discriminators. IEEE Transactions on Image Processing (2023)
2. Argilaga, A.: Fem-gan: A physics-supervised deep learning generative model for elastic porous materials. Materials (2023)
3. Kim, B., Azevedo, V.C., Thürey, N., Kim, T., Gross, M., Solenthaler, B.: Deep fluids: A generative network for parameterized fluid simulations. Computer Graphics Forum **38** (2018)
4. Yu, Y., Liu, Y.: Physics-guided generative adversarial network for probabilistic structural system identification. Expert Systems with Applications **239** (2024) 122339
5. Lerer, A., Gross, S., Fergus, R.: Learning physical intuition of block towers by example (2016)
6. Li, W., Azimi, S., Leonardis, A., Fritz, M.: To fall or not to fall: A visual approach to physical stability prediction (2016)
7. Zhang, Y., Li, W., Gao, Y., Li, S.: Physdiff: Physics-guided human motion diffusion model. arXiv preprint arXiv:2106.03704 (2021)
8. Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V.: Physics-guided neural networks (pgnn): An application in lake temperature modeling. arXiv preprint arXiv:2212.02500 (2022)
9. Chen, Y., Wu, Q., Sui, Y., Wang, Y., Zhu, S.C.: Physically grounded vision-language models for robotic manipulation. arXiv preprint arXiv:2309.02561 (2023)
10. Sun, G., Xie, W., Niyato, D., Mei, F., Kang, J., Du, H., Mao, S.: Generative ai for deep reinforcement learning: Framework, analysis, and use cases (2024)
11. Spitznagel, M., Keuper, J.: Urban sound propagation: a benchmark for 1-step generative modeling of complex physical systems (2024)
12. Bocher, E., Guillaume, G., Picaut, J., Petit, G., Fortin, N.: Noisemodelling: An open source gis based tool to produce environmental noise maps. ISPRS International Journal of Geo-Information **8**(3) (2019) 130
13. Kephalopoulos, S., Paviotti, M., Anfosso-Lédée, F.: Common noise assessment methods in europe (cnossos-eu). Technical Report EUR 25379 EN, Publications Office of the European Union, Luxembourg (Luxembourg) (2012) JRC72550.
14. Isola, P., Zhu, J., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. CoRR **abs/1611.07004** (2016)
15. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. CoRR **abs/1505.04597** (2015)
16. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. CoRR **abs/2006.11239** (2020)