Deep Learning for Spatiotemporal Anomaly Forecasting: A Case Study of Marine Heatwaves

Ding Ning [1] **Varvara Vetrova** [1] **Karin Bryan** [2] **Sébastien Delaux** [3]

Abstract

Spatiotemporal data have unique properties and require specific considerations. Forecasting spatiotemporal processes is a difficult task because the data are high-dimensional, often are limited in extent, and temporally correlated. Hence, we propose to evaluate several deep learning-based approaches that are relevant to spatiotemporal anomaly forecasting. We will use marine heatwaves as a case study. Those are observed around the world and have strong impacts on marine ecosystems. The evaluated deep learning methods will be integrated for the task of marine heatwave prediction in order to overcome the limitations of spatiotemporal data and improve data-driven seasonal marine heatwave forecasts.

1. Motivation: Challenges in Spatiotemporal Data

Geophysical data have unique spatiotemporal properties, which require specific considerations when used to build a machine learning-based prediction system. While machine learning approaches are commonly used to perform forecasting tasks, spatiotemporal forecasting has its own specificities that have to be accounted for when selecting a suitable approach. The selection process requires understanding the features of spatiotemporal data, which are high-dimensional, often are limited in extent, and temporally correlated. Below we will review the characteristics of spatiotemporal data in detail.

High dimensionality: Environmental data are indexed by up to three dimensions in space and one in time. Additionally, a single climate event is usually caused by multiple factors. Either observations or reanalysis datasets commonly contain tens to hundreds of variables that describe a single system. Domain knowledge can help select and hence reduce the number of variables to use as predictors in a machine learning algorithm. For example, in a recent study of marine heatwaves in the northwest Atlantic, the heatwave in 2012 was primarily driven by atmospheric forcing and the heatwave in 2017 was driven by offshore oceanic forcing (Gawarkiewicz et al., 2019), indicating that the roles of different types of forcing vary from event to event. However, without domain knowledge, determining the right combination of inputs from many variables is not so straightforward. With the increasing demand for more interpretable machine learning models, the trend has been to introduce more variables as predictors into a model, which leads to high dimensionality in the input. For example, in El Niño–Southern Oscillation (ENSO) forecasts, besides sea surface temperature and heat content anomalies (Ham et al., 2019), is it possible to include more inputs into the model, such as zonal wind anomalies (Lai et al., 2018) and sea level pressure anomalies (Manatsa et al., 2008)? Hence, selecting an appropriate number of variables from spatiotemporal data remains an open research question.

Limited data: The training of a robust deep learning model requires as much heterogeneous data as possible in order for the model to learn the entire possible range of behavior of the system. However, spatiotemporal data are often limited in extent. Temporally, data collection has evolved a lot with the advances of physical equipment and technology, and the quality and extent of collected environmental data have increased greatly in the mid-to-late 20$^{\text{th}}$ century. Meanwhile, unlike the big data that is related to human activities which are growing quickly, such as health records, surveillance videos, and social media activities, new geophysical data are generated relatively slowly.

Spatially, collection is limited to the physical environment on Earth and the resolution might be rather low. For example, in the NCEP Global Ocean Data Assimilation System (GODAS) dataset (from 1980 to 2020), the spatial resolution of the dataset is 1 by 0.333 degrees (360 longitudes between 0.5 and 359.5 and 418 latitudes between -74.5 and 64.499) (Behringer & Xue, 2004). This degree of spatial resolution

[1]School of Mathematics and Statistics, University of Canterbury, Christchurch, New Zealand [2]Environmental Research Institute, University of Waikato, Hamilton, New Zealand [3]MetOcean Solutions, Meteorological Service of New Zealand, Raglan, New Zealand. Correspondence to: Ding Ning <ding.ning@pg.canterbury.ac.nz>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

may not be enough to capture localized climatological phenomena, and the temporal resolution may not be sufficient for training a robust deep learning model. The temporal and spatial limitations of collecting data are unlikely to be mitigated by simply modifying the intervals.

Transfer learning, including one-shot learning (Li et al., 2006) and zero-shot learning (Palatucci et al., 2009), from a similar learning task has proven its feasibility in the spatiotemporal space (Zhang et al., 2021; Buckchash & Raman, 2021). However, how to find a similar large dataset and/or an appropriate pre-trained model for transfer learning remains a problem.

Temporal correlations: Spatiotemporal data are spatially-correlated between nearby locations and temporally-correlated between adjacent timestamps (Atluri et al., 2018). The common assumption that data is independent and identical distributed (iid) does not hold true. Spatial representations can be learned by machine learning, but temporal correlations are problematic. In turn, non-iid distribution of data could lead to difficulties in the training and evaluation process of machine learning models. For example, standard goodness-of-fit measures such as cross-validation are not appropriate in this case. In this regard, spatiotemporal data are very similar to videos and other sequence-type data, where relatively strong temporal correlations are present. In this sense, we could use deep learning-based video processing methods to handle spatiotemporal data. In addition, the sliding window and the expanding window methods could be utilized for model evaluation.

2. Seasonal Marine Heatwave Forecasts

Marine heatwaves: Marine heatwaves (MHWs) are observed around the world and have strong impacts on marine ecosystems; such impacts include shifts in species ranges, local extinctions, and can have a follow-on economic impact on seafood industries (Hobday et al., 2016). The devastating impact on marine ecosystems caused by MHWs brings irreversible loss of species or foundation habitats (Oliver et al., 2019), for example, mass coral bleaching and substantial declines in kelp forests and seagrass meadows (Holbrook et al., 2020). MHWs also affect aquaculture businesses and area-restricted fisheries because of the change of the distribution of sea life and then their production (Hobday et al., 2018), such as mussel, oyster and salmon farms. Accurately foreseeing MWHs, 6 months in advance for instance, and preparation for these potential climate impacts, such as collecting samples for repopulation, DNA sampling, and adjusting production beforehand, have both positive ecological and socioeconomic implications.

Anthropogenic climate change is expected to cause an increase in both the intensity and frequency of MHWs. There are multiple definitions of MHWs, and we use the following one that has been widely accepted: MHWs are discrete prolonged anomalously warm water events in a particular location and are generally identified following as a period of at least five consecutive days for which seawater temperature is warmer than the 90^{th} percentile based on a 30-year historical baseline period (Hobday et al., 2016). Such warm water events can be characterized using metrics, such as maximum temperature (Berkelmans et al., 2004), temperature anomaly (Smale & Wernberg, 2013), degree heating days (Maynard et al., 2008), etc. Even with the same metrics, different datasets may provide substantially different MHW information (Hobday et al., 2016). Dynamical prediction systems have been widely used for MHW forecasts (Merryfield et al., 2013; Saha et al., 2014; Vecchi et al., 2014). For example, a case study of the California Current System (CCS) MHW of 2014-2016 uses 8 global coupled climate prediction systems to predict the MHW up to 8 months ahead (Jacox et al., 2019); in this case, 2 of the 4 phases are predicted well by dynamic models but 2 others are missed. Hence, as a complement to these numerical models, deep learning-based models can possibly improve the MHW's predictability.

Datasets: We selected several datasets in our explorative analysis. The datasets provide spatiotemporal information that is associated with MHWs in particular. Besides the global data, we are also interested in spatiotemporal data in a local area around New Zealand. The datasets related to this research are:

- NCEP Global Ocean Data Assimilation System (GODAS) (Behringer & Xue, 2004), global numerical data from 1980 to present, with a spatial resolution of 1 by 0.333 degrees.
- Simple Ocean Data Assimilation (SODA) (Giese & Ray, 2011), global numerical data from 1870 to 2010, with a spatial resolution of 0.5 by 0.5 degrees.
- Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2012), global numerical data from 1861 to 2100, with a spatial resolution of 5 by 5 degrees.
- Moana Project Backbone v2 (O'Callaghan et al., 2019), numerical data around New Zealand from 1993 to 2017, with a spatial resolution of approximately 5 kilometers and a 3-hourly output period.

For the time scale, we are planning to use the monthly average fields as we are aiming at seasonal predictions. In the next section, we select and briefly review some deep learning methods for forecasting on spatiotemporal data.

3. Deep Learning for Spatiotemporal Anomaly Forecasting

In recent years, various deep learning-based approaches have been proposed, and among them, some were implemented for spatiotemporal data specifically and others were applied for this type of data. Detecting and forecasting anomalies in spatiotemporal data require consideration of the following aspects. Firstly, the type of machine learning task needs to be determined. For example, if the anomalies are unknown phenomena or have not previously occurred, then the learning task is unsupervised. On the other hand, if anomalies have been observed, such as MHWs, the learning type is generally supervised or semi-supervised, depending on the conditions of data and availability of labels. Secondly, the distribution of anomalies is usually imbalanced (Pang et al., 2021), for instance, accounting for less than 5 percent of data. The data imbalance makes anomalies unlikely to be detected and predicted using standard approaches and requires specific data pre-processing methods. Thirdly, there could be multiple anomaly classes within one dataset (Pang et al., 2021), such as detecting tropical cyclones and atmospheric rivers using the Community Atmosphere Model v5 (CAM 5.0) dataset (Neale et al., 2010). Such heterogeneity makes unsupervised learning more difficult. Fourthly, domain knowledge reveals that teleconnections influence anomalies (Cachay et al., 2020). Therefore, methods that tackle such relational data structure are needed.

In this section, we select and review some methods of interest for this research. These methods are applied to the case study of MHWs but are potentially useful outside of the MHW scope, for general spatiotemporal anomaly forecasts.

CNNs: Previous studies have demonstrated that CNNs could be used to forecast high-dimensional spatiotemporal data and outperform some state-of-the-art numerical models. For example, CNN-based models provide improved predictions for ENSO up to 1.5 years in advance, using sea surface temperature and heat content anomalies as input (Ham et al., 2019). Emulating a simple general circulation model (GCM), CNNs have been shown to successfully perform weather forecasting up to 14 days in advance with both input and output being high-dimensional (40 channels and 2048 grid points) (Scher, 2018). However, there are some drawbacks of CNNs, especially in the climate domain.

- CNNs that are designed to process 2D planar images may not suit spherical data which is connected with the globe.
- CNN-only models have no specific methods to tackle temporal correlations.
- CNNs have difficulty modeling teleconnections (large-scale atmospheric patterns) (Cachay et al., 2020).
- CNNs are slow in the training process.

In order to handle these drawbacks, some models have been proposed for modifying the convolutional layers (such as using spherical CNNs (Cohen et al., 2018)) or adding another type of layer (such as CNN-CapsNets (Chattopadhyay et al., 2020)) in a CNN architecture, or finding outperforming alternatives to CNNs (such as GNNs (Cachay et al., 2020)).

Spherical CNNs: CNNs are designed to perform learning and forecasting tasks on 2D planar images, while as one of its variants, spherical CNNs are invented to detect patterns regardless of the rotation over a sphere, i.e. rotation equivariance (Cohen et al., 2018). Spherical CNNs are a potential tool to learn information from global geophysical data. Furthermore, graph-based spherical CNNs improve the efficiency of spherical CNNs while keeping rotation equivariance by sampling a sphere as a graph, and have been applied to global extreme climate event segmentation (Defferrard et al., 2019).

CNN-LSTMs: Deep learning for video understanding and prediction is another area of interest that might be appropriate for use in spatiotemporal forecasting. CNN long short-term memory networks (CNN-LSTMs) are common baseline models for video prediction. For example, a CNN-LSTM-based predictive coding network has outperformed the previous CNN-LSTM models on the Rotating Faces Dataset (Singular Inversions) and the CalTech Pedestrian Dataset (Dollár et al., 2009; Lotter et al., 2016). Related to the climate domain, CNN-LSTMs are applied to forecast short-term global solar irradiance (Zang et al., 2020; Gao et al., 2020). Such networks could tackle temporal correlations and can be utilized in our research.

CapsNets: As an alternative to CNNs, capsule neural networks (CapsNets) can better model hierarchical relationships (Sabour et al., 2017). The structures called "capsules" reuse output from several lower capsules and then form more stable representations for higher capsules (Hinton et al., 2011). CapsNets have demonstrated their abilities to forecast extreme weather patterns up to 5 days ahead using mid-tropospheric large-scale circulation patterns (Z500) (Chattopadhyay et al., 2020).

CNN-AEs and GANs: Limitation in extent and incomplete labels are common problems with geophysical data, which make full-supervised learning unable to provide satisfying outcomes. CNNs combined with autoencoders (CNN-AEs) could generate new data and/or labels from existing labeled data, and perform learning on both. In an extreme weather detection task, the semi-supervised CNN-AEs are trained on a simulation from 1979 to 2005 with errors in the labeling, and outperformed the comparative fully-supervised CNNs in some aspects (Racah et al., 2016). More specifically, physical-based variational AEs and generative adversarial networks (GANs) are possible to better handle the incomplete domain knowledge.

GNNs: As a potential alternative to CNNs, graph neural networks (GNNs) process rasterized spatiotemporal information that has been transformed from Euclidean space. GNNs have the advantage of finding the hidden connections (via edges) among nodes and filtering out the non-essential spatial information outside nodes to reduce computational expense. GNNs have been proposed as a way to model teleconnections and have outperformed CNNs for multi-year ENSO forecasts (Cachay et al., 2020). GNNs also suit network-type data. Traffic forecasting (Song et al., 2020; Wu et al., 2019) and earthquake prediction (van den Ende & Ampuero, 2020) are other areas where GNNs have been employed.

4. Future Work

We propose to consider the following questions in our further research.

- How to select an appropriate number of predictors for spatiotemporal forecasting, what domain knowledge is required, and whether we can make the learning semi-supervised or unsupervised.
- How to overcome geophysical data insufficiency, and whether we can use the generative models to create additional data.
- Whether transfer learning, including one-shot learning and zero-shot learning, can be used for data insufficiency, and how to select appropriate relevant datasets and/or pre-trained models.
- How to tackle temporal correlations, and whether the video processing techniques can be used for spatiotemporal data.
- Whether we can use the reviewed deep learning methods that tackle spatial and/or temporal information, such as spherical convolutions, capsules, and graph reasoning, to improve the MHW predictability.

As a first step towards the research aims outlined above, we are planning to create a benchmark spatiotemporal dataset for MHWs, evaluate the deep learning approaches that are relevant to spatiotemporal forecasts, and propose an effective predictive model for MHWs that could outperform the existing ones such as (Jacox et al., 2019).

Acknowledgements

We would like to acknowledge that this PhD project is supported by the TAIAO project CONT-64517-SSIFDS-UOW (Time-Evolving Data Science / Artificial Intelligence for Advanced Open Environmental Science) funded by the New Zealand Ministry of Business, Innovation, and Employment (MBIE). URL https://taiao.ai/.

References

Atluri, G., Karpatne, A., and Kumar, V. Spatio-temporal data mining: A survey of problems and methods. *ACM Computing Surveys (CSUR)*, 51(4):1–41, 2018.

Behringer, D. W. and Xue, Y. Evaluation of the global ocean data assimilation system at ncep: The pacific ocean. In *Proceedings of the 8th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface*, 2004.

Berkelmans, R., De'ath, G., Kininmonth, S., and Skirving, W. J. A comparison of the 1998 and 2002 coral bleaching events on the great barrier reef: spatial correlation, patterns, and predictions. *Coral Reefs*, 23(1):74–83, 2004.

Buckchash, H. and Raman, B. Towards zero shot learning of geometry of motion streams and its application to anomaly recognition. *Expert Systems with Applications*, 177:114916, 2021.

Cachay, S. R., Erickson, E., Bucker, A. F. C., Pokropek, E., Potosnak, W., Osei, S., and Lütjens, B. Graph neural networks for improved el niño forecasting. *arXiv preprint arXiv:2012.01598*, 2020.

Chattopadhyay, A., Nabizadeh, E., and Hassanzadeh, P. Analog forecasting of extreme-causing weather patterns using deep learning. *Journal of Advances in Modeling Earth Systems*, 12(2):e2019MS001958, 2020.

Cohen, T. S., Geiger, M., Köhler, J., and Welling, M. Spherical cnns. *arXiv preprint arXiv:1801.10130*, 2018.

Defferrard, M., Milani, M., Gusset, F., and Perraudin, N. Deepsphere: a graph-based spherical cnn. In *Proceedings of the 7th International Conference on Learning Representations*, 2019.

Dollár, P., Wojek, C., Schiele, B., and Perona, P. Pedestrian detection: A benchmark. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 304–311, 2009.

Gao, B., Huang, X., Shi, J., Tai, Y., and Zhang, J. Hourly forecasting of solar irradiance based on ceemdan and multi-strategy cnn-lstm neural networks. *Renewable Energy*, 162:1665–1683, 2020.

Gawarkiewicz, G., Chen, K., Forsyth, J., Bahr, F., Mercer, A. M., Ellertson, A., Fratantoni, P., Seim, H., Haines, S., and Han, L. Characteristics of an advective marine heatwave in the middle atlantic bight in early 2017. *Frontiers in Marine Science*, 6:712, 2019.

Giese, B. S. and Ray, S. El niño variability in simple ocean data assimilation (soda), 1871–2008. *Journal of Geophysical Research: Oceans*, 116(C2), 2011.

Ham, Y.-G., Kim, J.-H., and Luo, J.-J. Deep learning for multi-year enso forecasts. *Nature*, 573(7775):568–572, 2019.

Hinton, G. E., Krizhevsky, A., and Wang., S. D. Transforming auto-encoders. In *Proceedings of the 21th Neural Information Processing Systems*, 2011.

Hobday, A. J., Alexander, L. V., Perkins, S. E., Smale, D. A., Straub, S. C., Oliver, E. C., Benthuysen, J. A., Burrows, M. T., Donat, M. G., Feng, M., Holbrook, N. J., Moore, P. J., Scannell, H. A., Gupta, A. S., and Wernberg, T. A hierarchical approach to defining marine heatwaves. *Progress in Oceanography*, 141:227–238, 2016.

Hobday, A. J., Spillman, C. M., Eveson, J. P., Hartog, J. R., Zhang, X., and Brodie, S. A framework for combining seasonal forecasts and climate projections to aid risk management for fisheries and aquaculture. *Frontiers in Marine Science*, 5(137), 2018.

Holbrook, N. J., Gupta, A. S., Oliver, E. C. J., Hobday, A. J., Benthuysen, J. A., Scannell, H. A., Smale, D. A., and Wernberg, T. Keeping pace with marine heatwaves. *Nature Reviews Earth & Environment*, 1(9):482–493, 2020.

Jacox, M. G., Tommasi, D., Alexander, M. A., Hervieux, G., and Stock, C. A. Predicting the evolution of the 2014–2016 california current system marine heatwave from an ensemble of coupled global climate forecasts. *Frontiers in Marine Science*, 6:497, 2019.

Lai, A. W.-C., Herzog, M., and Graf, H.-F. Enso forecasts near the spring predictability barrier and possible reasons for the recently reduced predictability. *Journal of Climate*, 31(2):815–838, 2018.

Li, F.-F., Fergus, R., and Perona, P. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4):594–611, 2006.

Lotter, W., Kreiman, G., and Cox, D. Deep predictive coding networks for video prediction and unsupervised learning. *arXiv preprint arXiv:1605.08104*, 2016.

Manatsa, D., Chingombe, W., Matsikwa, H., and Matarira, C. H. The superior influence of darwin sea level pressure anomalies over enso as a simple drought predictor for southern africa. *Theoretical and Applied Climatology*, 92 (1):1–14, 2008.

Maynard, J. A., Turner, P. J., K. R. Anthony, A. H. B., Berkelmans, R., Eakin, C. M., Johnson, J., Marshall, P. A., Packer, G. R., Rea, A., and Willis, B. L. Reeftemp: An interactive monitoring system for coral bleaching using high-resolution sst and improved stress predictors. *Geophysical Research Letters*, 35(5), 2008.

Merryfield, W. J., Lee, W.-S., Boer, G. J., Kharin, V. V., Scinocca, J. F., Flato, G. M., Ajayamohan, R. S., Fyfe, J. C., Tang, Y., and Polavarapu, S. The canadian seasonal to interannual prediction system. part i: Models and initialization. *Monthly Weather Review*, 141(8):2910–2945, 2013.

Neale, R. B., Richter, J. H., Conley, A. J., Park, S., Lauritzen, P. H., Gettelman, A., Williamson, D. L., Rasch, P. J., Vavrus, S. J., Taylor, M. A., Collins, W. D., Zhang, M., and Lin, S.-J. Description of the ncar community atmosphere model (cam 5.0). *NCAR Tech. Note NCAR/TN-486+ STR*, 1(1):1–12, 2010.

Oliver, E. C., Burrows, M. T., Donat, M. G., Gupta, A. S., Alexander, L. V., Perkins-Kirkpatrick, S. E., Benthuysen, J. A., Hobday, A. J., Holbrook, N. J., Moore, P. J., Thomsen, M. S., Wernberg, T., and Smale, D. A. Projected marine heatwaves in the 21st century and the potential for ecological impact. *Frontiers in Marine Science*, 6(734), 2019.

O'Callaghan, J., Stevens, C., Roughan, M., Cornelisen, C., Sutton, P., Garrett, S., Giorli, G., Smith, R. O., Currie, K. I., Suanda, S. H., Williams, M., Bowen, M., Fernandez, D., Vennell, R., Knight, B. R., Barter, P., McComb, P., Oliver, M., Livingston, M., Tellier, P., Meissner, A., Brewer, M., Gall, M., Nodder, S. D., Decima, M., Souza, J., Forcén-Vazquez, A., Gardiner, S., Paul-Burke, K., Chiswell, S., Roberts, J., Hayden, B., Biggs, B., and Macdonald, H. Developing an integrated ocean observing system for new zealand. *Frontiers in Marine Science*, 6: 143, 2019.

Palatucci, M., Pomerleau, D., Hinton, G. E., and Mitchell, T. M. Zero-shot learning with semantic output codes. In *Proceedings of the 22th Neural Information Processing Systems*, 2009.

Pang, G., Shen, C., Cao, L., and van den Hengel, A. Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)*, 54(2):1–38, 2021.

Racah, E., Beckham, C., Maharaj, T., Kahou, S. E., and Pal, C. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. *arXiv preprint arXiv:1612.02095*, 2016.

Sabour, S., Frosst, N., and Hinton, G. E. Dynamic routing between capsules. *arXiv preprint arXiv:1710.09829*, 2017.

Saha, S., Nadiga, S., Thiaw, C., Wang, J., Wang, W., Zhang, Q., van den Dool, H. M., Pan, H.-L., Moorthi, S., Behringer, D., Stokes, D., Peña, M., Lord, S., White, G., Ebisuzaki, W., Peng, P., and Xie, P. The ncep climate forecast system. *Journal of Climate*, 27(6):2185–2208, 2014.

Scher, S. Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. *Geophysical Research Letters*, 45 (22):616–622, 2018.

Singular Inversions, I. Facegen. *http://facegen.com*.

Smale, D. A. and Wernberg, T. Extreme climatic event drives range contraction of a habitat-forming species. In *Proceedings of the Royal Society B: Biological Sciences*, volume 280, pp. 20122829, 2013.

Song, C., Lin, Y., Guo, S., and Wan, H. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 914–921, 2020.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. An overview of cmip5 and the experiment design. *Bulletin of the American meteorological Society*, 93(4):485–498, 2012.

van den Ende, M. P. and Ampuero, J. P. Automated seismic source characterization using deep graph neural networks. *Geophysical Research Letters*, 47(17):e2020GL088690, 2020.

Vecchi, G. A., Delworth, T., Gudgel, R., Kapnick, S., Rosati, A., Wittenberg, A. T., Zeng, F., Anderson, W., Balaji, V., Dixon, K., Jia, L., Kim, H.-S., Krishnamurthy, L., Msadek, R., Stern, W. F., Underwood, S. D., Villarini, G., Yang, X., and Zhang, S. On the seasonal forecasting of regional tropical cyclone activity. *Journal of Climate*, 27 (21):7994–8016, 2014.

Wu, Z., Pan, S., Long, G., Jiang, J., and Zhang, C. Graph wavenet for deep spatial-temporal graph modeling. *arXiv preprint arXiv:1906.00121*, 2019.

Zang, H., Liu, L., Sun, L., Cheng, L., Wei, Z., and Sun, G. Short-term global horizontal irradiance forecasting based on a hybrid cnn-lstm model with spatiotemporal correlations. *Renewable Energy*, 160:26–41, 2020.

Zhang, Y., Wu, X., Gach, H. M., Li, H., and Yang, D. Groupregnet: a groupwise one-shot deep learningbased 4d image registration method. *Physics in Medicine & Biology*, 66(4):045030, 2021.