SNOW CLUE: PREDICTING GLOBAL SNOW DEPTH

Elias Baumann
Looking for PhD :)
e.baumann.study@gmail.com

Lucas Feldmann
School of Human Sciences
University of Osnabrück
lufeldmann@uos.de

Dirk Gütlin
Institute of Cognitive Science
University of Osnabrück
dguetlin@uos.de

January 17, 2021

1 Introduction

1.1 Context

Often overshadowed by studies on ice cover, the global snow cover and its seasonal changes play an important role in Earth's climate system. Much like ice, the bright snow reflects a high amount of energy from the sun away from Earth and acts as a natural radiation shield. Seasonal changes of snow cover have played an integral role in the Earth's ecosystem. In winter, snow provides cover for both hunter and prey [1] while melting snow in the warmer season acts as an important freshwater source not only for wildlife, but humans too [2].

Due to anthropogenic climate change, these seasonal snow cover changes have become more extreme in recent decades. A study from the National Oceanic and Atmospheric Administration (NOAA) revealed a slight increase of snow depth over the winter, but a much larger decrease of it during the summer [3]. In line with these findings, a NOAA article identified shrinking spring snow covers due to earlier onset of snow melt [4]. When rainfall hits the snow cover, snow melting is accelerated and the natural water buffer that is snow is quickly emptied. Moreover, in extreme cases this can lead to snow-related flooding. The increasingly extreme nature of snow cover changes and projected global warming suggest that snow-related floods might happen more frequently in the future.

1.2 Goals

To combat the drastic effects of the snow cover changes, SnowClue pursues two central goals. First, we want to raise awareness for the issue of climate change in general, specifically for its effects on the global snow cover. Second, we aim to predict global snow depths in the future in order to identify key points that influence snow depth. Given the short amount of time during the Hackathon, we came up with two use cases that combine both goals.

1.3 Use Cases

In order to make both use cases available to a wide audience we built a small web application that provides an interface to the predictions generated by our models of snow cover. The web page is available at `https://cutt.ly/snowclue`.

1.3.1 Will There Be Snow on Christmas?

This is a question often asked in the days leading up to the Christmas holidays. We tried to build a model that answers this question further into the future. While the scientific benefit of answering this question might not be high, we though it can be useful for the realization of our first goal - raising awareness.

In the web application, users are prompted to select a target year and a location for the prediction. Additionally, they chose one of three Shared Socioeconomic Pathways (SSPs): SSP1, SSP3 or SSP5. For simplification, these scenarios are renamed to best-case (SSP1), medium-bad (SSP3) and worst case (SSP5). Information regarding the chosen SSP is then displayed to the user. After clicking on the prediction button, the predicted snow height of around the Christmas holidays at the selected location in the given year is then displayed to the user along with an impulse to read more about impacts of climate change.

1.3.2 Global Snow Cover

Another way to raise awareness for climate change is to show its impact directly and on a global scale. Global snow levels are not only a good indicator of global warming, but also are a very tangible measure of climate change impact. We trained multiple purely convolutional Encoder-Decoder models to predict global snow levels in the future. Based on these predictions, we embedded a visualization of global snow cover up until the end of 2039 within our web application. Our application allows to investigate true snow levels (1950-2020) and predicted snow levels (2020-2039) for any date within the given time range. Additionally, the global snow cover can be compared to their according levels in 1950.

The user is asked to select a target date and SSP scenario. He or she may also enable or disable the comparison feature. The prediction button then generates a world map with the predicted (or historical if selected date is in the past) snow covers and additionally shows the difference to 1950 if desired.

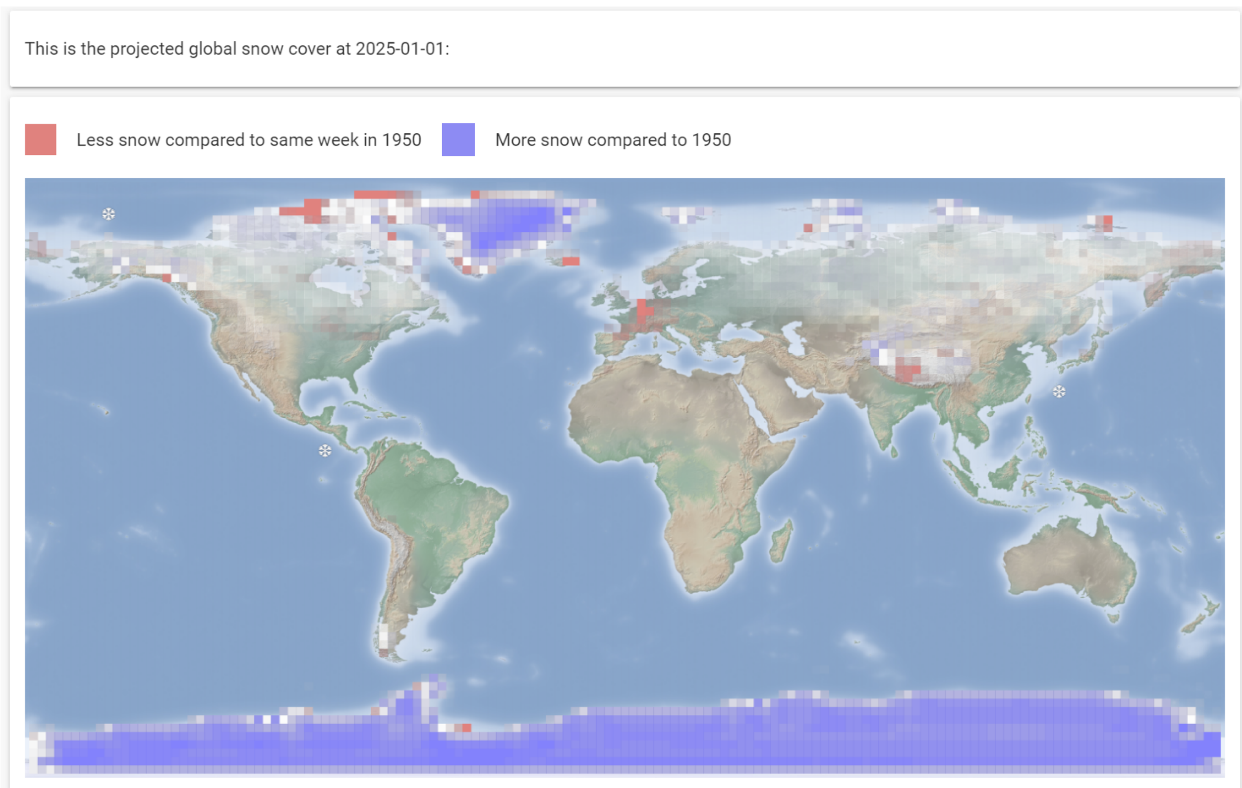


Figure 1: Exemplary output of snow cover prediction web app

2 Data

2.1 Historical Climate Data: ERA5

We used data from from the ERA5 global reanalysis [5] as historical training data. Specifically, we used the 2m temperature (t2m) and the total precipitation (tp) as features in our model and the snow depth (sd) as training target. In order to prepare the ERA5 data, we first combined the different variables into one larger dataset before downsampling from daily to weekly temporal resolution. Additionally, we added the week of the year as a data feature. We did this data preparation for all three available spatial resolutions (one, 2.5 and 5 degrees) of the ERA5 dataset.

2.2 Elevation Data: ETOPO5

In addition to the historical climate data from ERA5, we constructed elevation as an additional feature for our model. Elevation data was gathered from the ETOPO5 dataset [6]. We then mapped the elevation data to the latitude and longitude coordinates of the ERA5 dataset for all three resolutions.

2.3 Climate Projections: CMIP6

For projections of temperature (near_surface_air_temperature) and precipitation (precipitation_flux), we referred to the CMIP6 climate projection dataset. Specifically we used the SSP126, SSP370 and SSP585 pathways of the Geophysical Fluid Dynamics Laboratory (GFDL) model [7]. We downsampled this data to a weekly resolution and interpolated the spatial resolution to match that of the chosen ERA5 resolution. Thus, we feed our model projected temperature and precipitation in the same spatial and temporal resolution as the training data during the prediction stage.

3 Models

3.1 Network Building Blocks

We briefly motivate the major building blocks used in our neural network models:

Convolutional Neural Networks Convolutional neural networks (CNN) use learned kernels with weights to convolve over images reducing the overall number of necessary parameters for large inputs. By using the convolution operation, applied kernels always take neighbouring pixels into account when generating an output thereby including spatial information. Considering the distribution of snow over the globe, neighbouring longitude latitude degree squares are a good indicator for snow depth or a lack thereof.

Recurrent neural networks Recurrent neural networks (RNN) such as Long Short Term Memory recurrent neural networks (LSTM) incorporate a hidden state which is kept but modified over multiple input steps. By keeping and deciding to forget other information, a LSTM can retain information from many steps in the past to make more accurate predictions including long term trends in non-stationary data. LSTMs have successfully been used for timeseries forecasts such as temperature or rainfall predictions in the past [8] and will be used in some experiments in our work to include trends and potential information from e.g. summer which could be relevant for the following winter.

3.2 Predicting global snow cover using Encoder-Decoder model of convolutional and transpose convolutional layers

While convolutional layers yield state-of-the-art performance for image recognition tasks, transpose convolutional layers are frequently used in image generation [9]. Convolutional Neural Networks can not exclusively be used for image processing, but also have their application to other dimension-consistent problems where filtering processes are required [10]. For our prediction of global snow covers, we were interested to reconstruct a grid cell 2D-representation of the global snow sheet cover, based on previous 2D-representations of global snow sheet, temperature and precipitation. For this purpose we employed an Encoder-Decoder model of 3D convolutional and transposed convolutional layers. All layers were activated with a relu function and used batch normalization between layers. Multiple models were trained, each using all global longitude and latitude data, and a time range of two years as a prediction. The models only differed by the time they looked into the future, ranging from immediate predictions (next day), one year, two years, three years, five years, and 10 years.

3.3 Predicting Snow Depth with a Sequence2Sequence convolutional LSTM

With decent baseline models, guaranteeing results, we explore some other ideas on how to do forecasting for multiple frames as well as additional information such as uncertainty measures. We develop a second more experimental model in which we used a sequence to sequence (S2S) model [11] originating from language modeling to forecast an entire sequence $[\mathbf{y}_t, ..., \mathbf{y}_t + k]$ of length k at once instead of repeatedly forecasting a specific timestep ahead. S2S models have the advantage of preventing a model from learning trivial solutions e.g. by simply letting the input pass through unchanged. This is particularly problematic on data where changes are more subtle such as weekly environmental data. Instead, an S2S model encodes an input sequence into a single hidden state output $\mathbf{h}_t$ and a single cell state $\mathbf{c}_t$, both of which are output of a LSTM. This encoding is then passed to a second LSTM network as initial state to predict the target sequence. Therefore, the encoder has to encode all important information from past steps and from input into a

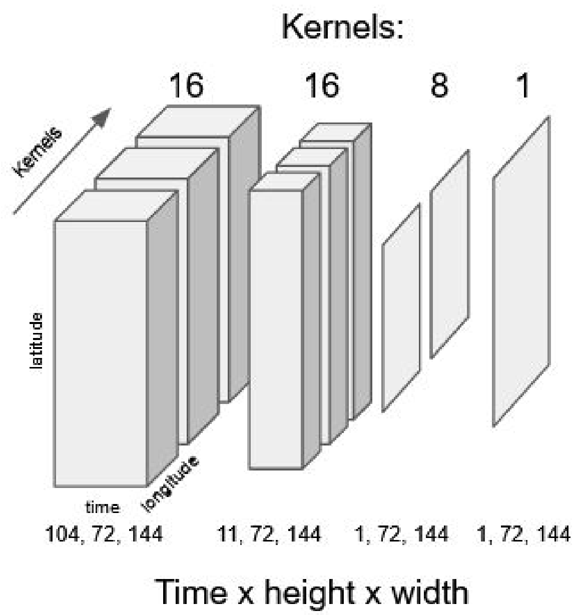


Figure 2: Architecture of our DeConv convolutional Encoder-Decoder model.

single step encoding. The decoder recieves no input or a different input such as just time steps. This sort of model can naturally be extended to work with images such as matrices consisting of longitude latitude degree squares with snow depth relevant data. This sort of model has also been applied to video frame prediction [12, 13] but with limited success when rapid movement occurs in a scene. In this work, we replace the LSTM with a ConvLSTM [14] which replaces all fully connected operations by convolutions and can therefore make use of such matrix inputs with reduced parameters. We argue that in our aggregated data, no rapid changes occur from frame to frame making a ConvLSTM a reasonable candidate in attempting to forecast snow depth multiple steps ahead.

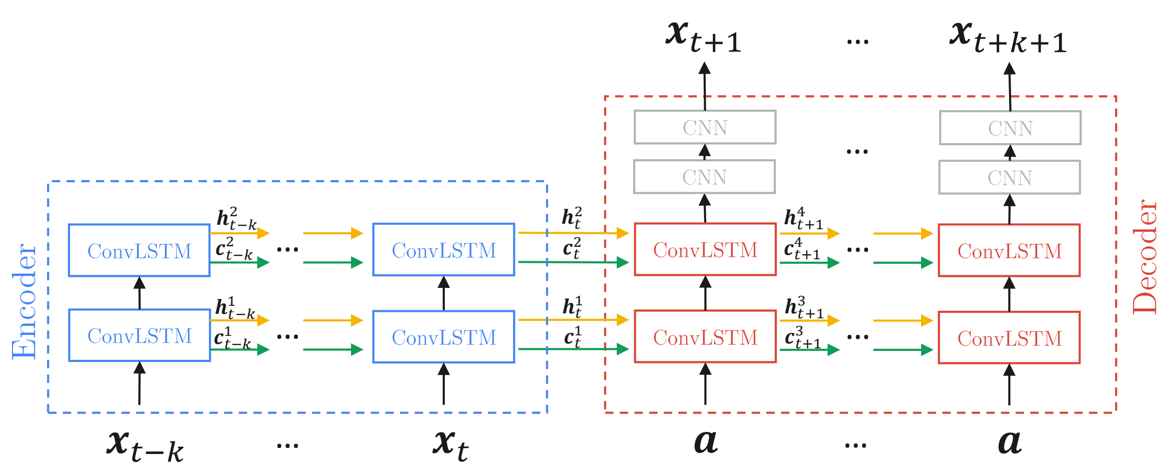


Figure 3: Overview of the S2S convolutional LSTM model

The model consists of two layers of convolutional LSTMs in both encoder and decoder with 52 filters each. For the final output, two CNN layers with 52 and 3 filters corresponding to the 3 target variables are used to generate the output logits for every timestep. In this model we predict future sequences $[\mathbf{x}_{t+1}, ..., \mathbf{x}_{t+k+1}]$ using past sequences $[\mathbf{x}_{t-k}, ..., \mathbf{x}_t]$ for three input variables: snow depth, surface temperature and total precipitation. Predicting all variables together ensures that all input data is used in creating the embedding. Additionally, the decoder recieves the constant altitude $\mathbf{a}$ of the predicted area at every sequence step $[\mathbf{a}, ..., \mathbf{a}]$. Data is resampled such that every timestep represents 2 Weeks. The model ist trained for 10000 steps using adam optimizer, 5e-4 learning rate and using sum of squared distances loss. This loss is frequently used in image regression problems and has the advantage that individual pixels are considered in the loss.

3.3.1 Providing uncertainty estimates using dropout

Neural Networks produce a deterministic output based on an input. During inference, for any given input, the output will always be the same value with no information on how certain the network is on the result. Recently, a number of approaches have been proposed to acquire uncertainty estimates from neural networks [15]. One of the strategies consists of repeatedly sampling a network with dropout during inference to generate an samples of an approximate posterior. The induced randomness from removing 10-20% of neurons gives a neural network properties similar to gaussian process. The samples of the approximate posterior can then be used to derive a normal distribution by calculating mean and variance and the result is a distribution for every value predicted by the neural network. Particularly for forecasting environmental data, snow depth in this case, individual distributions allow for a breadth of possible outcomes and can be used to define best-case and worst case scenarios.

3.3.2 S2S ConvLSTM Results

Due to time issues, we were unable to fully evaluate results and compare to ground truth. However, briefly visually evaluating the predictions, the model is able to capture a winter 2019 - summer 2020 change and accurately predicts snow for europes mountain ranges. To be able to run experiments on GPU, input data was limited to 60x80px resolution containing just Europe of ERA5 resolution 100 data. Snow depth has very low values ($< 0.01m$) in many parts of europe which creates issues as predicting no snow is a close to ground-truth prediction. This is reflected in the results as parts of greenland and scandinavia are set to almost consistently maximum value, whereas the rest of europe is mostly predicted to be 0m. Therefore, we log scale the snow depth and achieve better results. Preliminary results in Figure 4 show mean over 50 samples for every timestep, however every pixel actually corresponds to a distribution and can be treated as such for evaluation purposes.

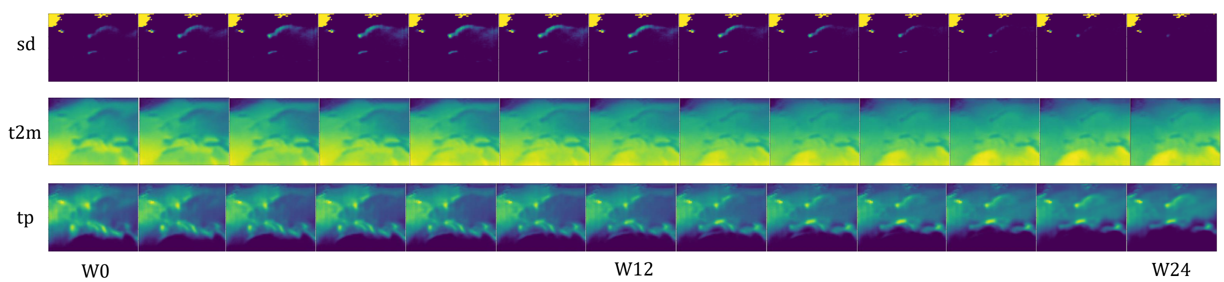
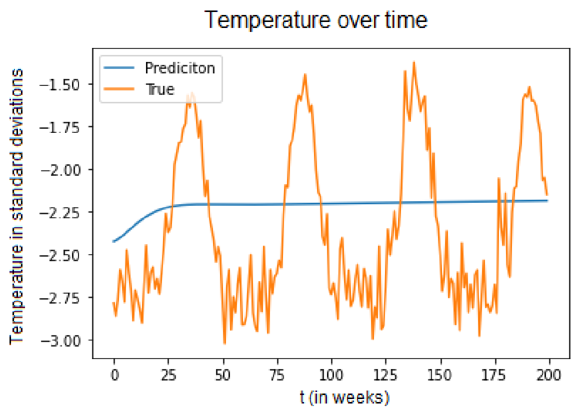


Figure 4: Mean predictions over 50 samples for every timestep. 2 Week aggregates. Forecast for beginning of 2020 using data from end of 2019. TP = Total precipitation, SD = Snow depth, T2M = surface temperature

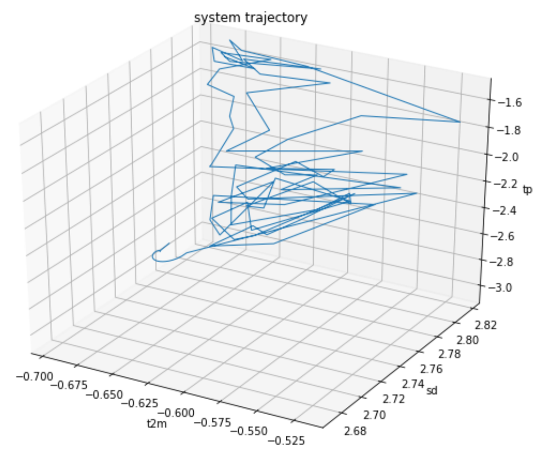
Published work attempting to forecast video frames many steps into the future utilize multiple input objectives with different encoders producing small embeddings which are in turn upsampled in an generative adversarial network-like fashion [16]. Modifying our model to include multiple encoders prediciting other datapoints such as wind and encoders with different time scales could improve prediction performance. Multiple spatial resolutions are currently also not considered but are highly valuable for image analysis [17]. Conducted experiments also showed that forecasting many frames at once ($k = 26, 52$) fails making the S2S ConvLSTM network only slightly viable for short term high resolution forecasts or long term low resolution forecasts.

3.4 Predicting snow and climate dynamics with a simple LSTM

Another approach was to predict the dynamic interactions of longitude, latitude, temperature, precipitation and snow height as a recurrent dynamical system. For this purpose, we used a model of two stacked LSTMs (64 and 32 units), combined with a fully-connected layer of 32 neurons. The intention of the model was to recreate the underlying dynamical system on all included variables, such that $states(t+1) = model(states(t))$. Although the same model previously yielded good results, e.g. on predicting the Lorenz system, the approach was quickly abandoned, since it did not seem capable of inheriting the system's properties (see Figure 5).



(a) Comparison of recurrent model prediction and true state for the system's temperature.



(b) Model trajectory in the state space of temperature (t2m), precipitation (tp) and snow level (sd).

Figure 5: Prediction results for the simple LSTM approach

4 Outlook

While the predictions of our model are currently only made for the sake of the web app, we imagine that these can also be used to predict snow melt and implement early warning systems for extreme snow-related events. Especially the prediction of snow-related flooding using our model and additional data regarding historical records of such events seems to be realizable.

Furthermore, adapting our model to incorporate real-time weather data can serve two functions. For one, we would continuously train our model. This would enable the prediction window of 20 years to move along with real time instead of being static. Second, we could use recent data to make more accurate short-term predictions. For example, the question about white Christmas should be answered with a higher accuracy when asked one month in advance rather than a year.

Additionally, we would like to derive direct counter-measures to the existing trend of more extreme snow cover anomalies using our model. One could compare the impacts of different scenarios on the snow cover and gather more information about the complex interaction of temperature, precipitation, elevation, and snow depth.

Finally, we want to further motivate those that viewed our web app. In the end, the climate crisis is a problem caused by humanity and every human should strive to combat this crisis.

References

[1] Natalie T Boelman, Glen E Liston, Eliezer Gurarie, Arjan JH Meddens, Peter J Mahoney, Peter B Kirchner, Gil Bohrer, Todd J Brinkman, Chris L Cosgrove, Jan UH Eitel, et al. Integrating snow science and wildlife ecology in arctic-boreal north america. *Environmental Research Letters*, 14(1):010401, 2019.

[2] Tim P Barnett, Jennifer C Adam, and Dennis P Lettenmaier. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438(7066):303–309, 2005.

[3] C Derksen, R Brown, L Mudryk, and K Luojus. Terrestrial snow cover. 2018.

[4] Climate change: Spring snow cover | NOAA climate.gov.

[5] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, et al. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

[6] MO Edwards. Global gridded elevation and bathymetry on 5-minute geographic grid (etopo5). *NOAA, National Geophysical Data Center, Boulder, Colorado, USA*, 1989.

[7] David J Paynter, Raymond Menzel, Stuart Freidenreich, Daniel M Schwarzkopf, Alexandra Jones, and Garrett Wright. Cmip6 gfdl dataset). 2018.

[8] Mohamed Akram Zaytar and Chaker El Amrani. Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *International Journal of Computer Applications*, 143(11):7–11, 2016.

[9] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27:2672–2680, 2014.

[10] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

[11] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. *Advances in neural information processing systems*, 27:3104–3112, 2014.

[12] William Lotter, Gabriel Kreiman, and David Cox. Deep predictive coding networks for video prediction and unsupervised learning. *arXiv preprint arXiv:1605.08104*, 2016.

[13] Chelsea Finn, Ian Goodfellow, and Sergey Levine. Unsupervised learning for physical interaction through video prediction. In *Advances in neural information processing systems*, pages 64–72, 2016.

[14] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28:802–810, 2015.

[15] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, pages 1050–1059, 2016.

[16] Ruben Villegas, Jimei Yang, Yuliang Zou, Sungryull Sohn, Xunyu Lin, and Honglak Lee. Learning to generate long-term future via hierarchical prediction. *arXiv preprint arXiv:1704.05831*, 2017.

[17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.