Latent Space Physical Coupling to Improve Semantic Disentangled Representations in β-VAE Architectures

Shakil Jiwa
Department of Earth System Science
University of California, Irvine
Irvine, CA 92612
sjiwa@uci.edu

Jinyang Li
Department of Civil and Environmental Engineering
University of California, Irvine
Irvine, CA 92612
jinyal4@uci.edu

Yue Yin
Department of Biomedical Engineering
University of California, Irvine
Irvine, CA 92612
yyin17@uci.edu

Abstract

The overarching goal of this project is to provide a basis for meaningful disentangled representation learning using a modified hybrid of the physics-informed Latent Linear Adjustment Autoencoder (LLAAE) [1] and β-VAE [2] generative models. It has been shown that unsupervised disentanglement learning is impossible without inductive biases on both the model and the data [3], and the success of β-VAEs relies heavily on exploitable biases inherent to their architectures [4]; therefore work has been done to improve disentanglement trained on weakly-supervised data [5] and modified loss functions such as β-TCVAE [6] are invented to minimize the Mutual Information Gap (MIG) and encourage prior independence. Nevertheless, it is still unfeasible for many applications to attain ground truth data, and the use of modified MIG loss functions does not guarantee disentangled ground factor latent representations. In this work we attempt to upgrade the LLAAE by including a hyper-parameter β to improve disentanglement learning, modifying the loss function to learn representations that are close to linear transformations of variance-explaining ground-truth factors, and tightening the prior with an asymmetric distribution commonly used in climate science [7]. Common difficulties associated with training VAE based models include posterior collapse from stochastic sampling and instability due to the variance of the proposed loss function. We seek to demonstrate that our modified physically-supervised β-VAE can provide robust and meaningful disentangled representations to be used on many interesting downstream tasks including climate prediction using the Canadian Regional Climate Model Large Ensemble (CRCM5-LE) [8] dataset.

1 Introduction

The intensity, duration, frequency of precipitation determines how much of disposition of precipitation hits the ground, and extremes of precipitation incidence will lead to floods and droughts, which creates enormous impact on the environment and society [9]. However, the precipitation process is generally considered to be poorly represented in numerical weather/climate models [10], and shows large spatial and temporal variabilities. Therefore, how to retrieve accurate and reliable

Course project for CS27E. All authors contribute equally.

precipitation remains a challenge.

Recent advances in deep learning shows that we can train a data-driven model to simulate the precipitation through carefully designed model structures like convolutional neural networks [10], generative adversarial network [11], but it's hard to interpret the output from the input, which decreases the trust of the model in the public use. However, the deep generative models, like variational autoencoders (VAEs) [12] can generate images through the sampling of the latent space, give us the hope that we can train a proxy, like pressure to map the latent space, and further generate precipitation which has more physical meaning behind it.

Certain VAE architectures such as β-VAE [2] and CAUSAL-VAE [13] attempt to learn so-called disentangled representations of data, which map ground truth factors represented in the latent space to certain phenomena in the data. An example would be certain latent space factors at certain values being able to represent phenomena such as precipitation events and chlorophyll blooms. Training a model to learn these disentangled representations can be useful for causal analysis, climate-data downscaling, and multiple other downstream applications, however learning these models requires inductive biases, which have been previously involved using semi-supervised learning as well as analysis based on structural causal methods.

In this work we propose that the LLAAE architecture induces biases into the model by pulling the latent space of the encoder close to the latent space of a linear model which we assume through physics to represent some component of the ground-truth driving factors of the output. By slightly modifying and re-implementing the architecture of the LLAAE we seek to improve the quality of the learned disentangled representations of the data. Specficially, we achieve the following for this project:

- implementing the LLAAE using PyTorch and test our network on a small subset,
- introducing the regularization into the loss function to further improve the reconstruction,
- incorporating a novel distribution into the sampling to disentangle the latent space.

2 Method

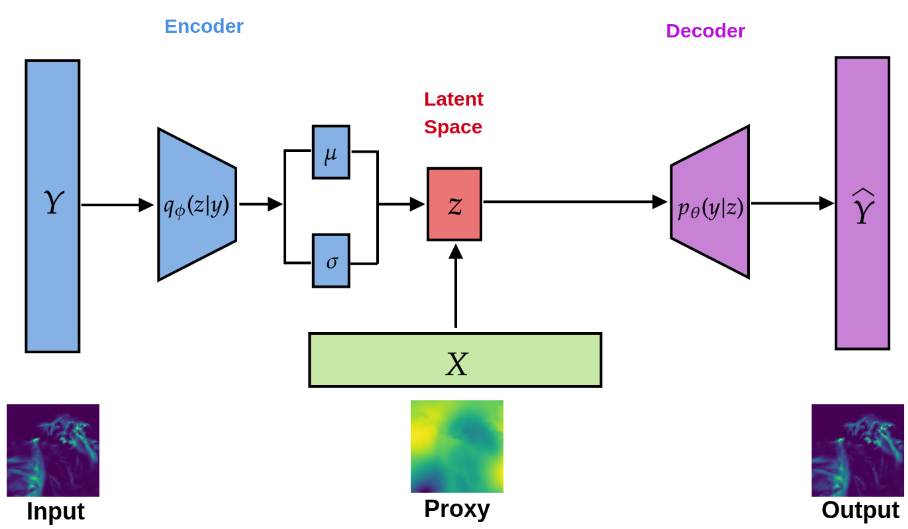

Figure 1: Architecture of LLAAE.

Building on Latent Linear Adjustment Autoencoder v1.0 [1], the precipitation model takes as input (Y) gridded precipitation (input to the encoder of the VAE) and sea level pressure data (X) (input to the physics latent encoder), and outputs reconstructions of precipitation data. The archaiture of

LLAAE model is shown in Figure 1. The physics latent encoder transforms X through a learned linear function into latent space variables Z_x. The physics latent space is trained to resemble the encoder latent space. In short, the LLAAE algorithm has two loss terms that are optimized alternatively:

$$\mathcal{L}_1 = \left\|\mathbf{Y} - \hat{\mathbf{Y}}\right\|_2^2 + \alpha \left\|\mathbf{Y} - \hat{\mathbf{Y}}_{\mathbf{X}}\right\|_2^2 + \beta \mathcal{L}_{KL}$$

$$\mathcal{L}_2 = \|e(\mathbf{Y}) - h(\mathbf{X})\|_2^2$$

In order to add further inductive biases to the dataset and in an attempt to better represent the latent space as ground-truth data, the encoder samples of the latent space were drawn using an augmentation of the reparameterization trick using the Generalized Extreme Value (GEV) distribution, which as been shown to be well representative of climate data at tail-ends of distributions [7].

The GEV distribution is parameterized by a location, scale, and skewness parameter and the expectation and variance of the distribution are given:

$$\mathrm{E}(X) = \mu + (g_1 - 1)\,\frac{\sigma}{\xi} \quad for\ \xi < 1$$

$$Var(X) = \left(g_2 - g_1^2\right)\frac{\sigma^2}{\xi^2}$$

where $g_k = \Gamma(1 - k\xi), k = 1, 2, 3, ...$ and $\Gamma(t)$ is the gamma function.

Samples z from the latent space are drawn then using the reparameterized:

$$Z = E(x) + Var(x) \cdot error$$

where error is normally distributed around 0 with variance 1. The location and variance parameters are based on neural network outputs representing $E(X)$ and $Var(X)$, noting that the skewness is selected manually (in our case to be -0.5) due to difficulty in training which is common in training the skewness parameter for the GEV distribution [7]. Lastly, regularization parameters α and β can be added to the loss of latent space and KL-divergence for the variational loss function of the LLAAE in order to attain disentangled representations. We try different regularization term to achieve optimal results. At the end of training when the input and encoder is removed, we obtain a convontional neural network that takes the proxy predict the output, as shown in Figure 2.

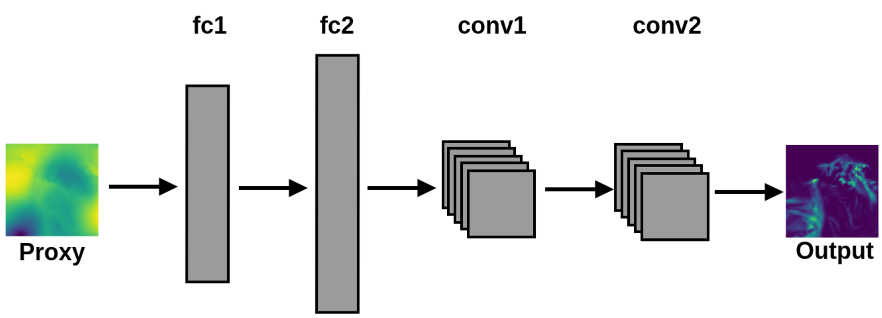

Figure 2: A convolutional neural network as the product of training.

3 Dataset

We use the same dataset as described in [1], which comes from [8]. We focus on the precipitation as the target variable, and the SLP is regarded as a proxy for the linear model. Both data were retrieved from the Coordinated Regional Climate Downscaling Experiment – European Domain (EURO-CORDEX) (WCRP, 2015), which covers the region of 27N – 72N, 22W – 45E with 0.44°. Both the precipitation and the SLP are resized to 128*128.

4 Experiments

4.1 Latent linear adjustment with VAE

In Figure 3 we show the Reconstruction from the input as well as the prediction using the proxy after 200 epochs of training.

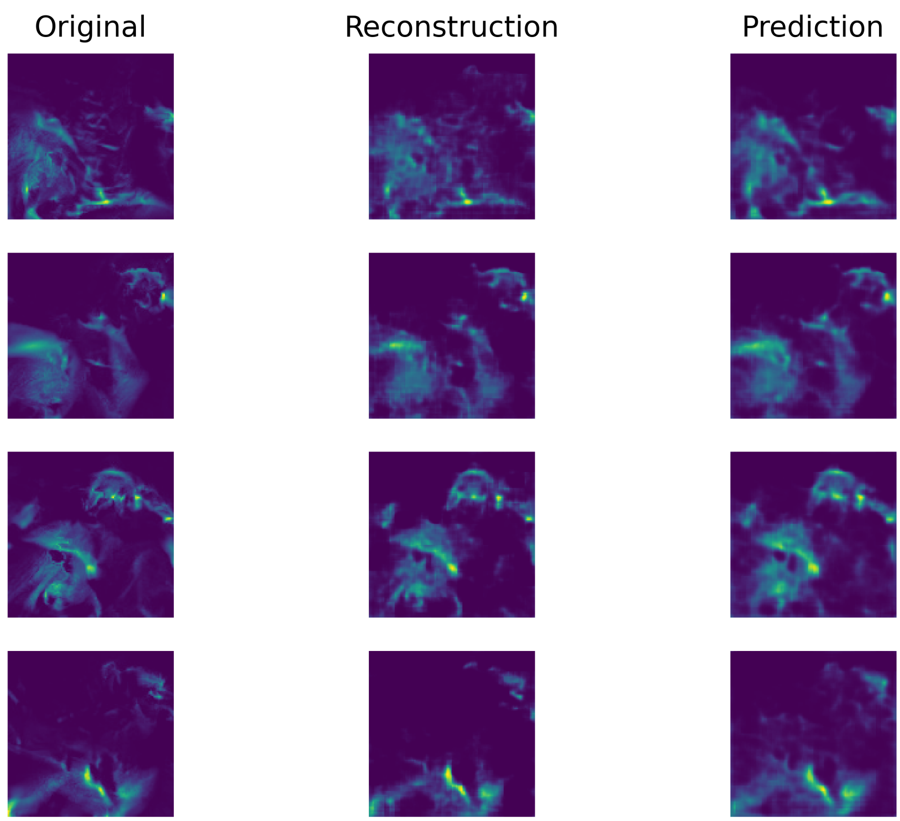

Figure 3: Sample output after training LLAAE.

In Figure 4 we show the traces of all the loss functions for both training and testing of LLAAE.

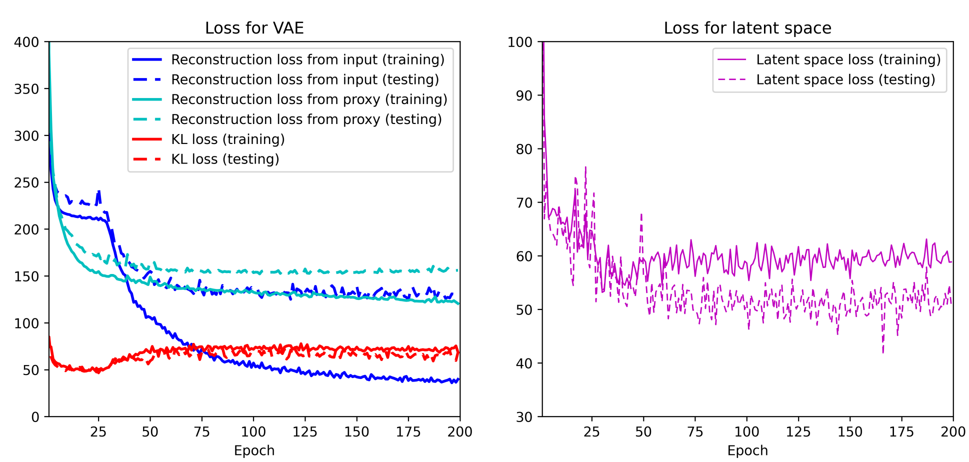

Figure 4: Traces for loss functions.

4.2 Latent linear adjustment with β-VAE

We experiment with different α and β values to showcase that reconstruction and prediction can be further improved by adding the regularization term. In general, a high α and a low β result in better reconstruction and latent space match.

Table 1: Comparison of loss with α and β adjusted

	α=1, β=1	α=5, β=0.2	α=5, β=0.5	α=2, β=0.5
Reconstruction loss from input (training)	39.81	23.08	30.81	30.73
Reconstruction loss from input (testing)	132.77	132.10	131.45	123.84
Reconstruction loss from proxy (training)	120.57	115.10	122.93	126.60
Reconstruction loss from proxy (testing)	156.08	158.16	158.45	154.21
KL loss (training)	68.64	112.98	84.47	80.66
KL loss (testing)	68.53	122.63	76.80	80.72
Latent space loss (training)	58.87	38.20	54.86	58.60
Latent space loss (testing)	50.67	30.30	47.48	46.83

4.3 LLAAE with GEV distribution

Using the GEV distribution to represent the latent space as well as using a latent linear adjustment as inductive biases we were able to find meaningful disengtangled state representations of different precipitation event shapes.

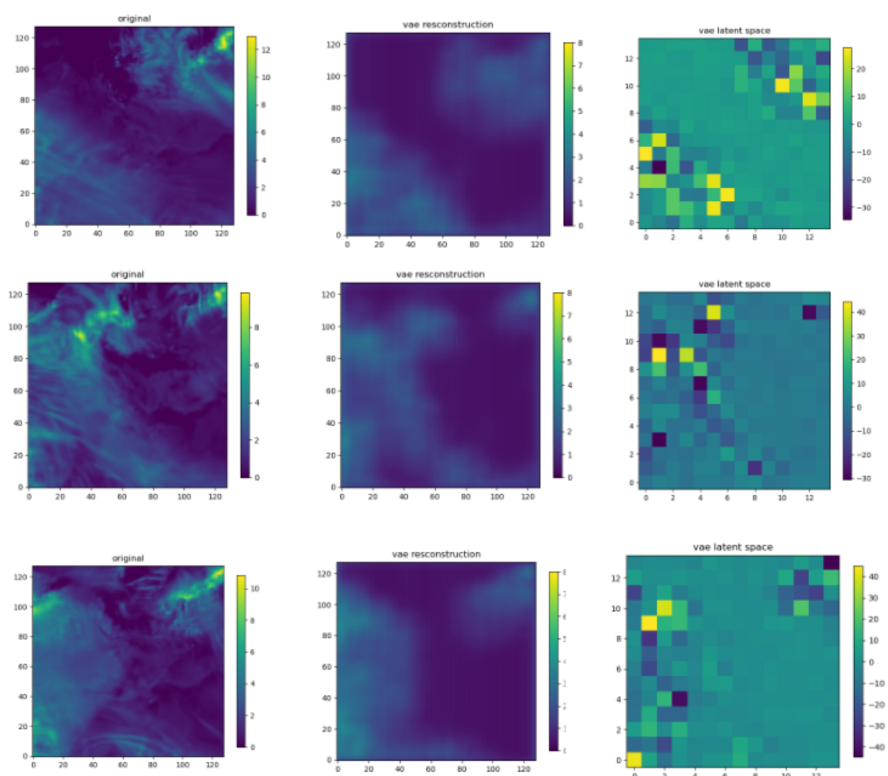

Figure 5: Sample output after training LLAAE with GEV.

It is apparent that our beta modified LLVAE loss function does not exactly train for good reconstructions, however the latent space is well representative of precipitation phenomenon. It is known that these two objectives are hard to obtain simultaneously The goal in using the latent linear adjustment

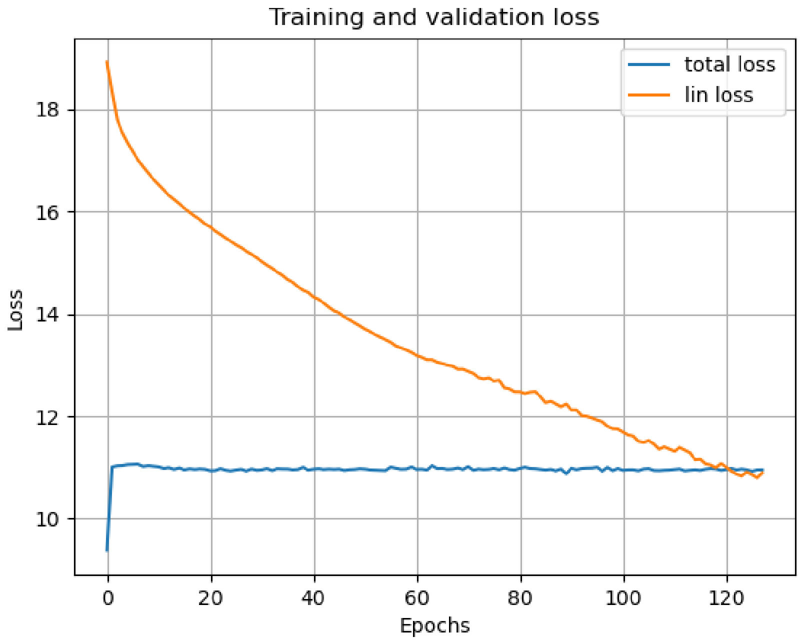

Figure 6: Training loss for LLAAE with GEV Distribution.

autoencoder with our modifications was to find downscaled representations of ground-truths that can be seen to semantically represent certain climate phenomena on a finer-resolution scale. The LLVAE pulls the latent space towards a representation of some component of precipitation, and in turn forces the encoder to represent the other components given the prescribed linear structure. The layered-CNN encoder output represents a coarse-scale resolution image of ground truth data containing the precipitation structure. The decoder captures the shape of this structure however the reconstruction fails to capture extreme values, which is somewhat expected, however, as the main objective was latent space representation rather than reconstruction for this experiment. The ability of the loss function to learn was also significantly increased through our modifications to the underlying probability distributions. It is also important to note that the resulting latent space representations showed better disentanglement over our Gaussian-Based VAE.

4.4 Biogeochemical proxy and downscaling application

We generalized the LLVAE for representing disentangled representations by training the model on ocean chlorophyll growth data. In this model, proxies were taken at a 500km nominal resolution (the same size as the latent space), and chlorophyll was taken at a 50km resolution. The proxies used were radiation; where radiation-induced growth is represented by a quadratic function with learned parameters, and temperature, which is represented by an exponential function with learned parameters. Finally, in order to generalize the framework to learn disentangled representations of other types of data through physical proxies, we trained a similar model with inputs and outputs ocean chlorophyll data and physics-based functions linking temperature and shortwave downwelling radiation to the latent space. In this model, the coarse scale temperature and downwelling radiation data are transformed into their respectively-induced growth as described in [14].

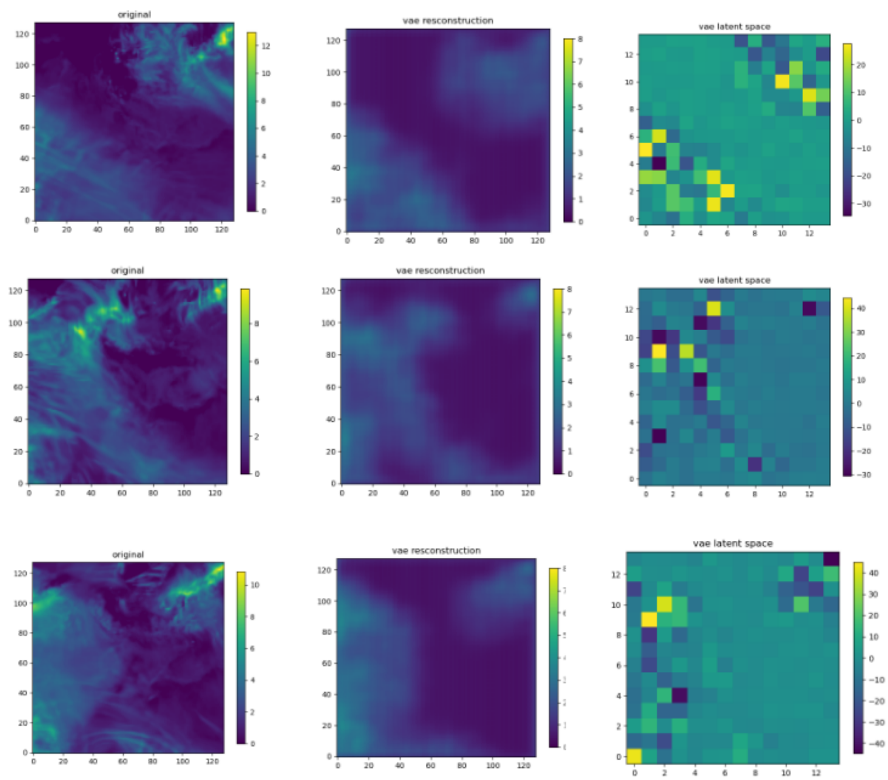

Figure 7: Sample output after training LLAAE with GEV.

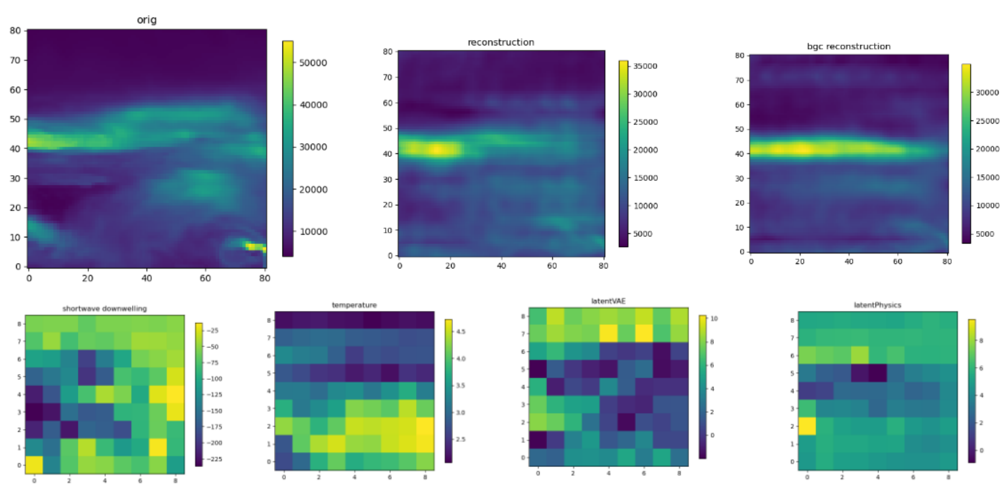

Figure 8: Sample output for biogeochemical proxy and downscaling application.

These results provide both good latent space representations as well as provide an opportunity for downscaling climate data based on physics-based models at more coarse resolutions from the trained decoder and proxy functions. Future work here could involve training models to downscale proxy data into fine-scale growth data on a global scale.

5 Discussion and Conclusion

The sample output trained from LLAAE can be found in figure3, which the original one and reconstruction refer to the input and output from the VAE and the prediction refers to the output from the linear model passing to the decoder of VAE. We can see that he prediction is as good as the reconstruction, which enables SLP to explain the prediction of the portion of the precipitation field. Figure 4 shows the loss for training VAE and linear part. It was shown that training the data for good latent representation often led to poor reconstruction and vice versa, however in this case, both were attained by modified versions of the LLVAE. Furthermore, the model was extended to chlorophyll growth to show that physics-induced latent spaces aid in disentangled representation learning. Future work could involve looking into further downscaling applications as well as causal inference using disentangled representation.

References

[1] Christina Heinze-Deml, Sebastian Sippel, Angeline G Pendergrass, Flavio Lehner, and Nicolai Meinshausen. Latent linear adjustment autoencoder v1. 0: a novel method for estimating and emulating dynamic precipitation at high resolution. *Geoscientific Model Development*, 14(8):4977–4999, 2021.

[2] Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. 2016.

[3] Francesco Locatello, Stefan Bauer, Mario Lucic, Gunnar Raetsch, Sylvain Gelly, Bernhard Schölkopf, and Olivier Bachem. Challenging common assumptions in the unsupervised learning of disentangled representations. In *international conference on machine learning*, pages 4114–4124. PMLR, 2019.

[4] Dominik Zietlow, Michal Rolinek, and Georg Martius. Demystifying inductive biases for β-vae based architectures. *arXiv preprint arXiv:2102.06822*, 2021.

[5] Francesco Locatello, Ben Poole, Gunnar Rätsch, Bernhard Schölkopf, Olivier Bachem, and Michael Tschannen. Weakly-supervised disentanglement without compromises. In *International Conference on Machine Learning*, pages 6348–6359. PMLR, 2020.

[6] Ricky TQ Chen, Xuechen Li, Roger Grosse, and David Duvenaud. Isolating sources of disentanglement in variational autoencoders. *arXiv preprint arXiv:1802.04942*, 2018.

[7] Elisa Ragno, Amir AghaKouchak, Linyin Cheng, and Mojtaba Sadegh. A generalized framework for process-informed nonstationary extreme value analysis. *Advances in Water Resources*, 130:270–282, 2019.

[8] Martin Leduc, Alain Mailhot, Anne Frigon, Jean-Luc Martel, Ralf Ludwig, Gilbert B Brietzke, Michel Giguère, François Brissette, Richard Turcotte, Marco Braun, et al. The climex project: a 50-member ensemble of climate change projections at 12-km resolution over europe and northeastern north america with the canadian regional climate model (crcm5). *Journal of Applied Meteorology and Climatology*, 58(4):663–693, 2019.

[9] Kevin E Trenberth, Aiguo Dai, Roy M Rasmussen, and David B Parsons. The changing character of precipitation. *Bulletin of the American Meteorological Society*, 84(9):1205–1218, 2003.

[10] Baoxiang Pan, Kuolin Hsu, Amir AghaKouchak, and Soroosh Sorooshian. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55(3):2301–2321, 2019.

[11] Cunguang Wang, Guoqiang Tang, and Pierre Gentine. Precipgan: Merging microwave and infrared data for satellite precipitation estimation using generative adversarial network. *Geophysical Research Letters*, 48(5):e2020GL092032, 2021.

[12] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[13] Mengyue Yang, Furui Liu, Zhitang Chen, Xinwei Shen, Jianye Hao, and Jun Wang. Causalvae: disentangled representation learning via neural structural causal models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9593–9602, 2021.

[14] Kyle F Edwards, Mridul K Thomas, Christopher A Klausmeier, and Elena Litchman. Phytoplankton growth and the interaction of light and temperature: A synthesis at the species and community level. *Limnology and Oceanography*, 61(4):1232–1244, 2016.