Controllable Generation for Climate Modelling

Tackling Climate Change with Machine Learning Workshop, NeurIPS 2022

Moulik Choraria[1,*] Daniela Szwarcman[2] Bianca Zadrozny[2]

Campbell D. Watson[3] Lav R. Varshney[1]

[1]: University of Illinois at Urbana Champaign [2]: IBM Research, Brazil [3]: IBM Research, NY

[*]: Corresponding author: moulikc2@illinois.edu

Motivation

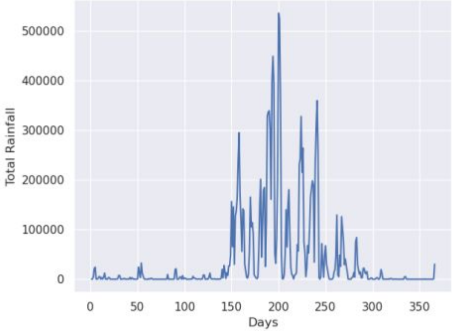
- Modelling of climate change patterns and extreme events of immense interest, to understand and develop strategies to mitigate societal impacts
- Downstream impact evaluation often requires high-resolution climate scenarios, where traditional stochastic weather generators become suboptimal and resource-intensive
- On the other hand, Machine learning based generative models have achieved tremendous attention in modelling complex distributions, including weather data
- Our goal is to leverage machine learning based models for weather generation, with focus on controllability for accurate user-guided modelling

Task

- Precipitation modelling using the CHIRPS [1] rainfall data set of daily precipitation fields at a 0.05 degree resolution, spanning 1981–2021.

- For our task, consider bounding box of 6.4 × 6.4 degrees (128 × 128 pixels) with chosen region geographically corresponding to central india, focussing on monsoon season

Fig.1: Depiction of selected geographical area (left) and its corresponding rainfall distribution over an year (right)



[1]: "The climate hazards infrared precipitation with 162 stations—a new environmental record for monitoring extremes," - Funk et al, 2015.

Method

- We consider the Generative Adversarial Networks [2] (GANs) for modelling precipitation

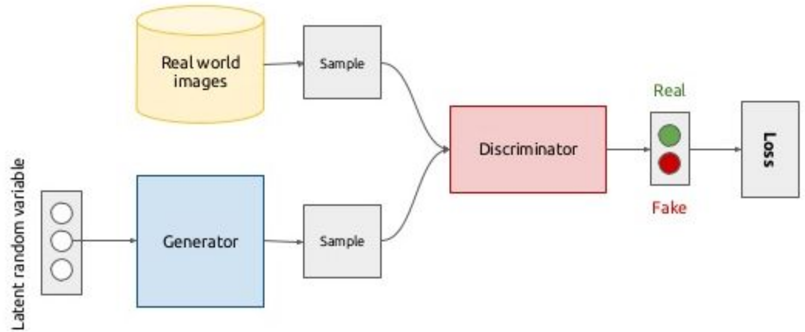


Fig.2: Depiction of the GAN learning framework [4]

- Generator and discriminator parametrized as Resnet [3]-based deep networks

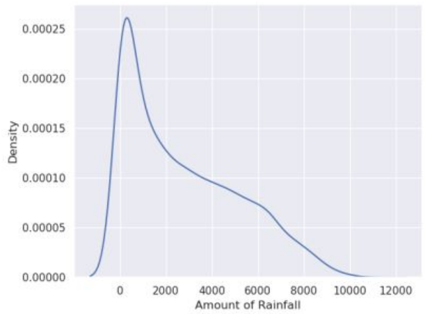
[2]: "Generative adversarial networks,"- Goodfellow et al, 2014.
[3]: "Deep Residual Learning for Image Recognition" - He et al., 2015
[4]: Image Reference: "Revisiting Recent and Current Anomaly Detection based on Machine Learning in Ad-Hoc Networks" - Wang et al, 2019

Method

- Sampling in standard GANs uncontrolled, concentrates around mode of zero/low rainfall samples. To obtain diverse samples, we use conditioning within the GAN framework

- **Conditioning Variable:** Obtain the distribution of total rainfall per precipitation field, discretize into a histogram with n bins, use the bin membership (m ϵ {1, 2…n}) as conditioning variable

- Serves as pseudo-measure for “extremeness”, higher bin equals higher total rainfall. Further, we combine conditioning with under/over-sampling for better representation of each bin



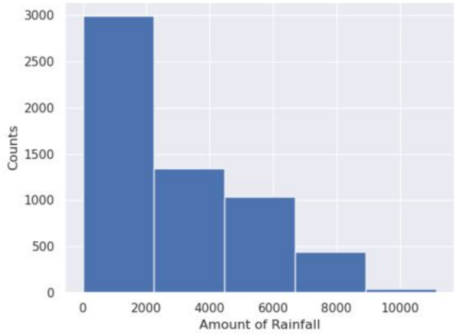
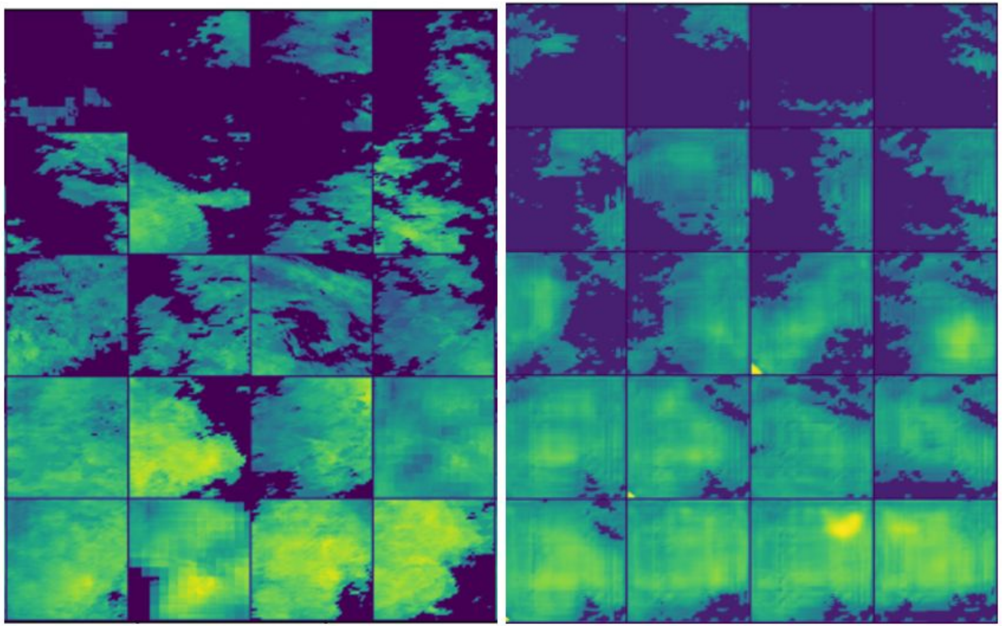


Fig.3: Distribution of total rainfall per sample (left), along with the discretized histogram (right)

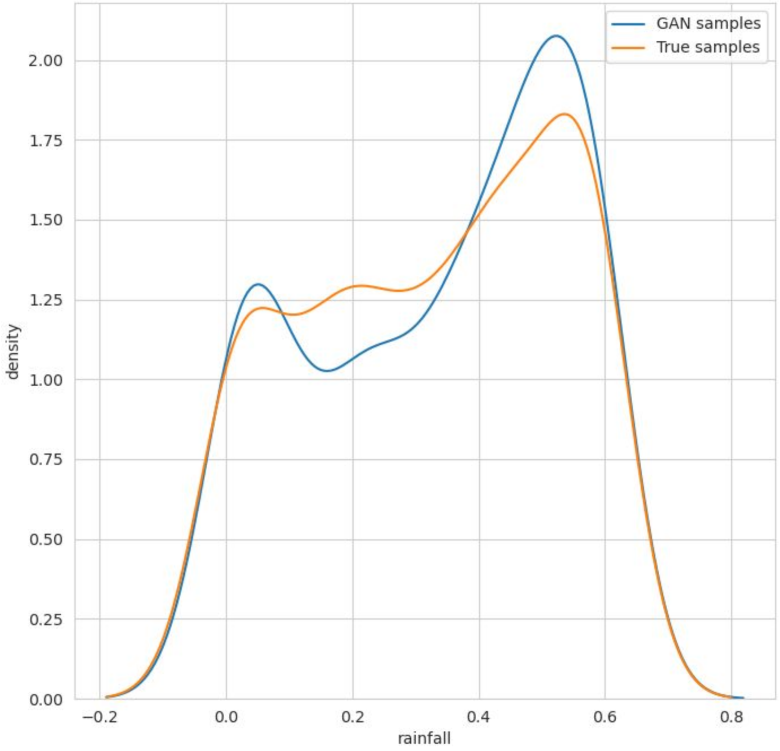
Results: Visual Comparison



(a) Training Samples (b) Generated Samples

Results: Density Evaluation

-



Future Work

- Finer grained discretization would enable more control over the generation process, including that of extreme scenarios.
- Augment learning framework with metric learning in the representation space for more explicitly distinguishing between “extremeness” of different bins
- Instead of scalar conditioning, condition via lower-resolution maps for better control, where each pixel in the map conditions/controls a sub-area in the higher resolution output samples
- For better quality and diversity of generation, explore use of other generative modelling formulations such as state-of-art diffusion models