

RAINNET: A LARGE-SCALE IMAGERY DATASET FOR SPATIAL PRECIPITATION DOWNSCALING SUPPLEMENTARY MATERIAL

Anonymous authors

Paper under double-blind review

1 CONTENT

In this supplementary material, we illustrate the details of introduced metrics and provide more samples of our dataset. Furthermore, we discuss the configuration of training and report more experimental results. We will publish the dataset and code on `github` soon.

2 GEOPHYSICS EQUATION

The Navier–Stokes and mass continuity equations (including the effect of the Earth’s rotation), together with the first law of thermodynamics and the ideal gas law, represent the full set of prognostic equations in the atmosphere, describing the change in space and time of wind, pressure, density and temperature is described (shown in Eq. 1–5).

Momentum Equations:

$$\begin{aligned}\frac{\partial u}{\partial t} &= -[u, v, w] \cdot \nabla u - \frac{1}{\rho} \frac{\partial p}{\partial x} + f v, & \text{Zonal} \\ \frac{\partial v}{\partial t} &= -[u, v, w] \cdot \nabla v - \frac{1}{\rho} \frac{\partial p}{\partial y} - f u, & \text{Meridional} \\ \frac{\partial w}{\partial t} &= -[u, v, w] \cdot \nabla w - \frac{1}{\rho} \frac{\partial p}{\partial z} - g, & \text{Vertical}\end{aligned}\tag{1}$$

Mass Continuity:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot ([u, v, w] \cdot \rho),\tag{2}$$

Thermo-dynamic:

$$\frac{\partial \theta}{\partial t} = -[u, v, w] \cdot \nabla \theta + \dot{Q},\tag{3}$$

Ideal Gas:

$$p = \rho R T,\tag{4}$$

Moisture equation:

$$\frac{\partial q}{\partial t} = -[u, v, w] \cdot \nabla q + \text{micro}(q).\tag{5}$$

The zonal (u), meridional (v) and vertical (w) wind speed, are driven by air pressure (p). The air pressure is driven by the mass density (ρ , determined by wind and the ingredient of air, *e.g.*, moisture) and the temperature. The moisture and air can contain heat, which forms the variable latent heat (θ). The heat can also come from other heat sources (\dot{Q}), *e.g.*, physical processes like sun radiation or chemical processes like burning coal. The moisture may also come from snowmelt or other physical processes, which is described in microphysics (*micro*) - a sub-domain of geoscience.

21 3 METRICS

22 Due to the difference between down-scaling and traditional figure super-resolution, the metrics work
 23 well under SR tasks may not be sufficient for precipitation down-scaling. By gathering the most
 24 common metrics from the meteorologic literature (for example Zhang & Yang (2004); Maraun
 25 et al. (2015); Ekström (2016); He et al. (2016); Pryor & Schoof (2020); Wootten et al. (2020)), we
 26 select and rename 6 metrics to reflect the downscaling quality: mesoscale peak precipitation error
 27 (MPPE), cumulative precipitation mean square error (CPMSE), heavy rain region error (HRRE),
 28 cluster mean distance (CMD), heavy rain transition speed (HRTS) and average miss moving degree
 29 (AMMD). These 6 metrics can be separated as reconstruction metrics: MPPE, HRRE, CPMSE,
 30 AMMD, and dynamic metrics: HRTS and CMD.

31 The MPPE ($mm/hour$) is calculated as the difference of top quantile between the generated/real
 32 rainfall dataset which considering both spatial and temporal property of mesoscale meteorological
 33 systems, *e.g.*, hurricane, squall. This metric is used in most of these papers (for example Zhang
 34 & Yang (2004); Maraun et al. (2015); Ekström (2016); He et al. (2016); Pryor & Schoof (2020);
 35 Wootten et al. (2020) suggest the quantile analysis to evaluate the downscaling quality).

36 The CPMSE ($mm^2/hour^2$) measures the cumulative rainfall difference on each pixel over the time-
 37 axis of the test set, which shows the spatial reconstruction property. Similar metrics are used in
 38 Zhang & Yang (2004); Maraun et al. (2015); Wootten et al. (2020) calculated as the pixel level
 39 difference of monthly rainfall and used in He et al. (2016) as a pixel level difference of cumulative
 40 rainfall with different length of record.

41 The HRRE (km^2) measures the difference of heavy rain coverage on each time slide between gen-
 42 erated and labeled test set, which shows the temporal reconstruction ability of the models. The
 43 AMMD ($radian$) measures the average angle difference between main rainfall clusters. Similar
 44 metrics are used in Zhang & Yang (2004); Maraun et al. (2015); Wootten et al. (2020) as rainfall
 45 coverage of a indefinite number precipitation level and used in He et al. (2016); Pryor & Schoof
 46 (2020) as a continuous spatial analysis.

47 As a single variable dataset, it is hard to evaluate the dynamical reconstruction ability of different
 48 models. So here we introduce the first order variables to evaluate the dynamical property of down-
 49 scaling results. Similar approaches are suggested in Maraun et al. (2015); Ekström (2016); Pryor &
 50 Schoof (2020). The CMD (km) physically compares the location difference of the main rainfall sys-
 51 tem between the generated and labeled test set, which could be also understand as the RMSE of the
 52 first order derivative of precipitation data on spatial directions. The HRTS ($km/hour$) measures the
 53 difference between the main rainfall system moving speed between the generated and labeled test set
 54 which shows the ability for models to capture the dynamic property, which could be also understand
 55 as the RMSE of the first order derivative of precipitation data on temporal direction. Similar metrics
 56 are suggested in Maraun et al. (2015); Ekström (2016); Pryor & Schoof (2020) as auto-regression
 57 analysis and differential analysis.

58 More details about the metrics and their equations are given in supplementary materials. One met-
 59 rics group (MPPE, HRRE, CPMSE, AMMD) mainly measures the rainfall deviation between the
 60 generated precipitation maps and GT. The other group (HRTS and CMD) mainly measures the dy-
 61 namic deviation of generated precipitation maps. In order to further simplify the application of
 62 indices, we abstract them into two weighted and summed metrics: Precipitation Error Measure
 63 (PEM) and Precipitation Dynamics Error Measure (PDEM). We first align the dimensions of these
 64 two groups of metrics respectively. The first group of metrics (MPPE, HRRE, CPMSE, AMMD)
 65 is normalized, weighted and summed to get the precipitation error measure (PEM). According to
 66 Gupta et al. (1999), all the metrics are transferred to Percent Bias (PBIAS) to couple with expert
 67 opinions for metrics weighting. The original definition of PBIAS is the bias divided by observa-
 68 tion, as $PBIAS = |Q_{model} - Q_{observed}| / |Q_{observed}|$. Here we rewrite the original metrics to
 69 PBIAS by dividing the metrics with annual mean observations of the original variables (AMO), as
 70 $PBIAS_i^{PEM} = |Metrics_i^{PEM}| / |AMO_i^{PEM}|$. The metrics then are ensembled to a single metric
 71 (PEM) with equal weight, as $PEM = \sum_i w_i \cdot PBIAS_i^{PEM}$. Following the same procedure, we
 72 then ensemble the second group of dynamic metrics (HRTS and CMD) to a single metrics PDEM.

Here we would clarify the calculation in detail and compare the metrics select here with other metrics sets. These metrics follow different formulations under the test set time length T and test set area size A :

Mesoscale peak precipitation error (MPPE) This metric measures the ability for the down-scaling models to capture the mesoscale peak precipitation. The mesoscale large weather/meteorological events are happening on a scale of $200km \times 200km$, such as hurricane or squall. The ability of capturing this metrics would help improve the flood prediction, as the precipitation events at this scale could stimulus large flooding. By measuring 1/1000 quantile of precipitation ($5000km^2$) over temporal and spatial, we could capture the precipitation at the mesoscale weather events.

This metric is comparable to R99p in CLIMDEX Zhang & Yang (2004) and 20-return period in VALUE Maraun et al. (2015) by definition.

MPPE is a similar index to R99p (relatively R99.2p). And 20-year-return-period in VALUE is R99.7p in the language of CLIMDEX. Because this is a downscaling dataset, we change that to the RMSE of R99.2p, and name that as MPPE following its physics meaning.

This metric is used in a variety of literature (for example Zhang & Yang (2004); Maraun et al. (2015); Ekström (2016); He et al. (2016); Pryor & Schoof (2020); Wootten et al. (2020) suggest the quantile analysis to evaluate the downscaling quality).

Heavy rain region error (HRRE) This metric measures the difference between the reconstructed dataset and the real high-resolution observations of the heavy rain region. The heavy rain is defined by $56mm/day$, which is a conventional benchmark for heavy rain in weather prediction (America: $50.8 - 76.2mm/day$; Japan, India and China: $50 - 75mm/day$; European: $40 - 60mm/day$). This metrics is formed by:

$$HRRE = \left(\frac{1}{T} \sum_t (A_{HR}(P > 56, t) - A_{GT}(P > 56, t)^2)^{0.5} \right),$$

where HR means the high-resolution data and GT is the generated data. Average miss moving degree (AMMD) measures the ability for model to capture the temporal direction of heavy rain, which is obtained by recording the center of heavy rain on each frame and record the directional difference.

This metric is comparable to R20mm in CLIMDEX Zhang & Yang (2004) and number of threshold exceedances in Maraun et al. (2015). Here we select the threshold by the definition of heavy rain. We name this metric following its physics meaning.

Similar metrics are used in Zhang & Yang (2004); Maraun et al. (2015); Wootten et al. (2020) as rainfall coverage of a indefinite number precipitation level and used in He et al. (2016); Pryor & Schoof (2020) as a continuous spatial analysis.

Cumulative precipitation mean square error (CPMSE) This metric represents the ability for model to capture the spatial difference of precipitation over a long time, which is usually considered in climatology. Through long time observation, we use this metric to lay out the impact of miss alignment issue and focus on the climatology and spatial rainfall estimation. This metrics is formed by:

$$CPMSE = \frac{1}{T \cdot A} \sum_{ij} \left(\sum_t P_{HR}(i, j, t) - \sum_t P_{GT}(i, j, t) \right)^2)^{0.5}.$$

This metrics in comparable to PRCPTOT (definitely the same definition) in CLIMDEX and "mean" in VALUE. We name this metric following its physics meaning.

Similar metrics are used in Zhang & Yang (2004); Maraun et al. (2015); Wootten et al. (2020) calculated as the pixel level difference of monthly rainfall and used in He et al. (2016) as a pixel level difference of cumulative rainfall with different length of record.

Cluster mean distance (CMD) and Heavy rain transition speed (HRTS) The CMD measures the distance between the main rainfall clusters between the generated dataset and the high-resolution.

This metric blocks the rainfall quantity estimation error and focuses on spatial difference on each time slide. For each frame, we first calculate out the area size of the heavy rain in HR dataset. We mark the contour of heavy rain in HR dataset as $f_{HR}(x, y, t)$. The area of this contour is marked as A_{Rain} . Then we find the contour of generated rainfall dataset with the same area size as in $f_{HR}(x, y, t)$ and mark the contour as $f_{GT}(x, y, t)$. We calculate out the heavy rain contour difference between HR and GT dataset under 2-norm. The metric could be calculated as:

$$CMD = (\sum_t \frac{1}{T \cdot A_{Rain}} \langle \iint [x, y] f_{HR}(x, y, t) - [x, y] f_{GT}(x, y, t) dx dy \rangle)^{0.5},$$

105 in which \oint is the area integration; $\langle \rangle$ means the self inner product.

106 To further calculate this value, we need to discrete this value as:

$$CMD = (\sum_t \frac{1}{T \cdot A_{Rain}} \langle \sum_i \sum_j [i, j] f_{HR}^D(i, j, t) - [i, j] f_{GT}^D(i, j, t) \rangle)^{0.5},$$

107 where f_*^D becomes 1 when the lattices are on the boundary of the contours; otherwise it would be 0.

The HRTS measures the ability for model to capture the dynamics (transition speed) of heavy rain. For each frame, we first calculate out the area size of the heavy rain in HR dataset. We mark the contour of heavy rain in HR dataset as $f_{HR}(x, y, t)$. Then for the last frame in HR dataset, we find the contour of generated rainfall dataset with the same size. We also do this for the generated dataset. We calculate out the heavy rain contour difference between this and last frame. Then we compare the difference of HR and GT under 2-norm. This metrics actually shows the order-1 property of dynamics which is shown in main text Eq. 1 - the wind blowing effect. The metrics could be calculated as:

$$HRTS = (\sum_t \frac{1}{T \cdot A_{Rain}} \langle \iint [x, y] f_{HR}(x, y, t) - [x, y] f_{GT}(x, y, t) dx dy \rangle)^{0.5}.$$

108 To further calculate this value, we need to discrete this value as:

$$HRTS = (\sum_t \frac{1}{T \cdot A_{Rain}} \langle \sum_i \sum_j [i, j] (f_{HR}(i, j, t) - f_{HR}(i, j, t-1)) - [i, j] (f_{GT}(i, j, t) - f_{GT}(i, j, t-1)) \rangle)^{0.5}.$$

109 HRTS and CMD are comparable to auto regression analysis in VALUE (but not in CLIMDEX).
 110 These are the bias on first order regression on temporal and spatial dimensions. We use these met-
 111 rics to reflect the dynamic property of the downscaling results and name these metric following its
 112 physics meaning. Similar metrics are suggested in Maraun et al. (2015); Ekström (2016); Pryor &
 113 Schoof (2020) as auto-regression analysis and differential analysis.

114 4 DATASET DETAILS

115 We show more precipitation maps in proposed dataset. In order to display the dynamic characteris-
 116 tics of the precipitation map more conveniently, we extract the precipitation maps of 4 periods and
 117 make them into GIFs. These GIFs are organized in the attachment of the supplementary materials.

118 5 EXTRA RESULTS OF RAINNET

119 5.1 DETAILED NETWORK STRUCTURE

120 The structure of vanilla network in our proposed framework is given in Fig. 1. We employ 6
 121 Residual-in-Residual Dense Blocks (RRDB) Wang et al. (2018) in our downscaling backbone and 3
 122 RRDB in the implicit dynamic estimation module network.

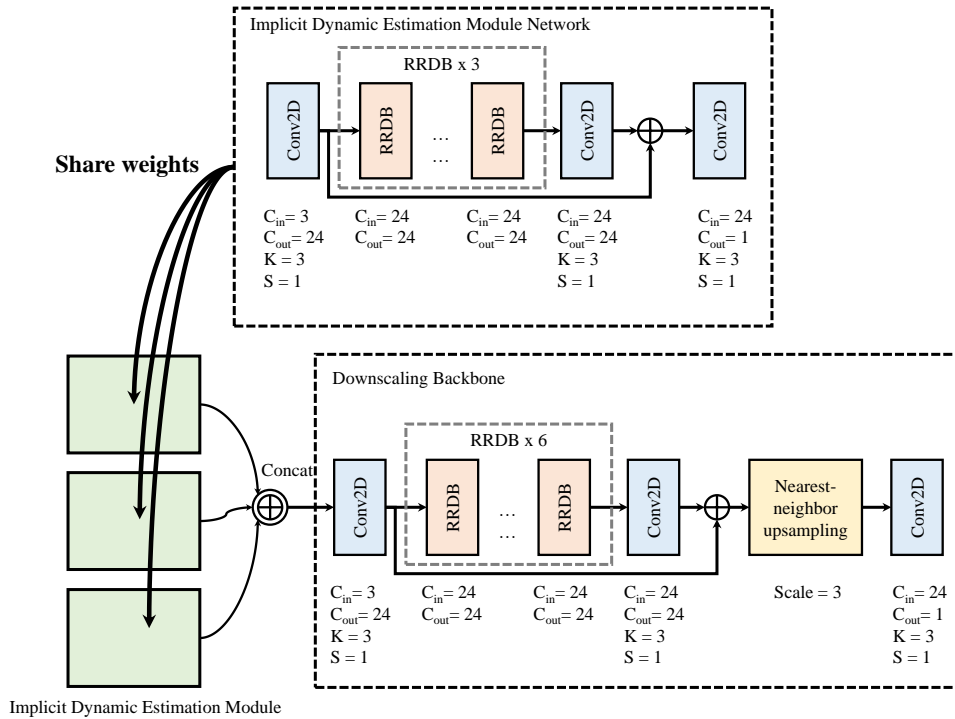


Figure 1: The details structure of vanilla network in our proposed model.

5.2 TRAINING DETAILS

We select 13 algorithms as benchmark: Bicubic Keys (1981), SRCNN Dong et al. (2016), SRGAN Ledig et al. (2017), EDSR Lim et al. (2017), ESRGAN Wang et al. (2018), DBPN Haris et al. (2018), RCAN Zhang et al. (2018), SRGAN-V, EDSR-V, ESRGAN-V, RBPN Haris et al. (2019), EDVR Wang et al. (2019) and Kriging. These implementations are derived or adapted from publicly available code provided by the authors. Since all these methods process three-channel pictures by default, we modify the number of input channels of these models (the precipitation map in our proposed dataset are all single-channel). According to our task, we also adjust the hype parameters of these models for better performance.

5.3 EXTRA RESULTS

We randomly pick 6 sets of results and show them in Fig. 2~7. In addition, we extract the downscaling results (our proposed method) of 5 periods and make them into GIFs. These GIFs are organized in the attachment of the supplementary materials.

REFERENCES

- Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016.
- Marie Ekström. Metrics to identify meaningful downscaling skill in wrf simulations of intense rainfall events. *Environmental Modelling & Software*, 79:267–284, 2016.
- Hoshin Vijai Gupta, Soroosh Sorooshian, and Patrice Ogou Yapo. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of hydrologic engineering*, 4(2):135–143, 1999.

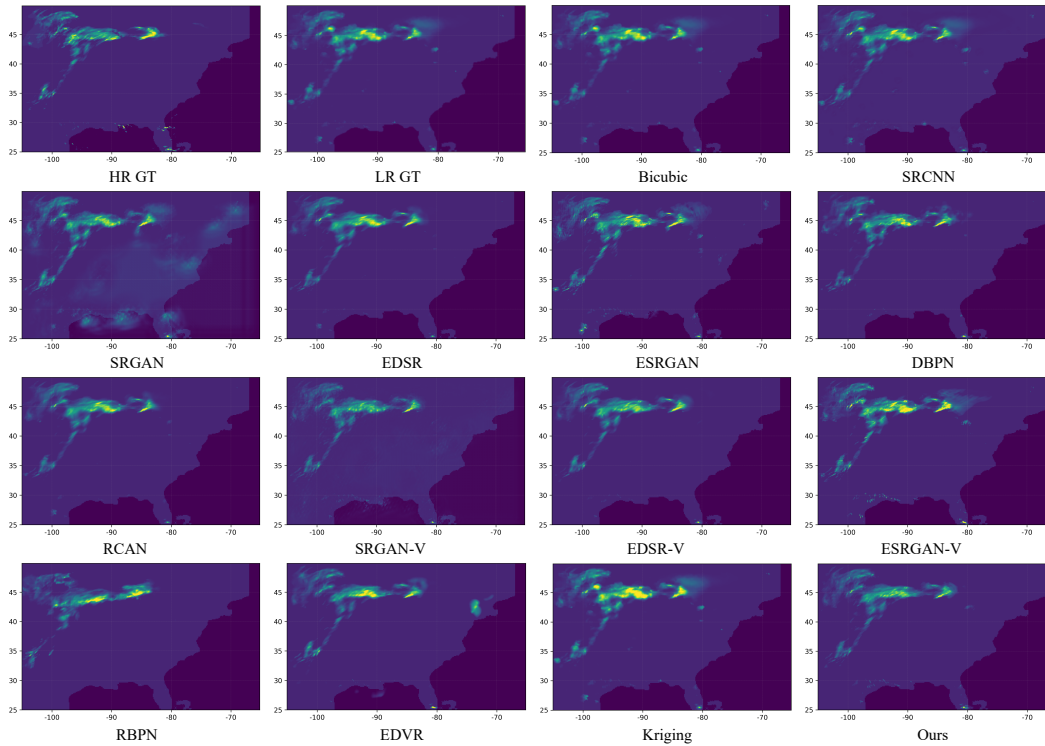


Figure 2: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 546-th hour in September 2010). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

- 144 Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deep back-projection networks
145 for super-resolution. In *2018 IEEE Conference on Computer Vision and Pattern Recognition,*
146 *CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. IEEE Computer Society, 2018.
- 147 Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Recurrent back-projection net-
148 work for video super-resolution. In *IEEE Conference on Computer Vision and Pattern Recogni-*
149 *tion, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE,
150 2019.
- 151 Xiaogang He, Nathaniel W Chaney, Marc Schleiss, and Justin Sheffield. Spatial downscaling of
152 precipitation using adaptable random forests. *Water resources research*, 2016.
- 153 Robert Keys. Cubic convolution interpolation for digital image processing. *IEEE transactions on*
154 *acoustics, speech, and signal processing*, 1981.
- 155 Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro
156 Acosta, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-
157 realistic single image super-resolution using a generative adversarial network. In *2017 IEEE*
158 *Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July*
159 *21-26, 2017*. IEEE Computer Society, 2017.
- 160 Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep resid-
161 ual networks for single image super-resolution. In *2017 IEEE Conference on Computer Vision*
162 *and Pattern Recognition Workshops, CVPR Workshops 2017, Honolulu, HI, USA, July 21-26,*
163 *2017*. IEEE Computer Society, 2017.
- 164 Douglas Maraun, Martin Widmann, José M Gutiérrez, Sven Kotlarski, Richard E Chandler, Elke
165 Hertig, Joanna Wibig, Radan Huth, and Renate AI Wilcke. Value: A framework to validate
166 downscaling approaches for climate change studies. *Earth’s Future*, 3(1):1–14, 2015.

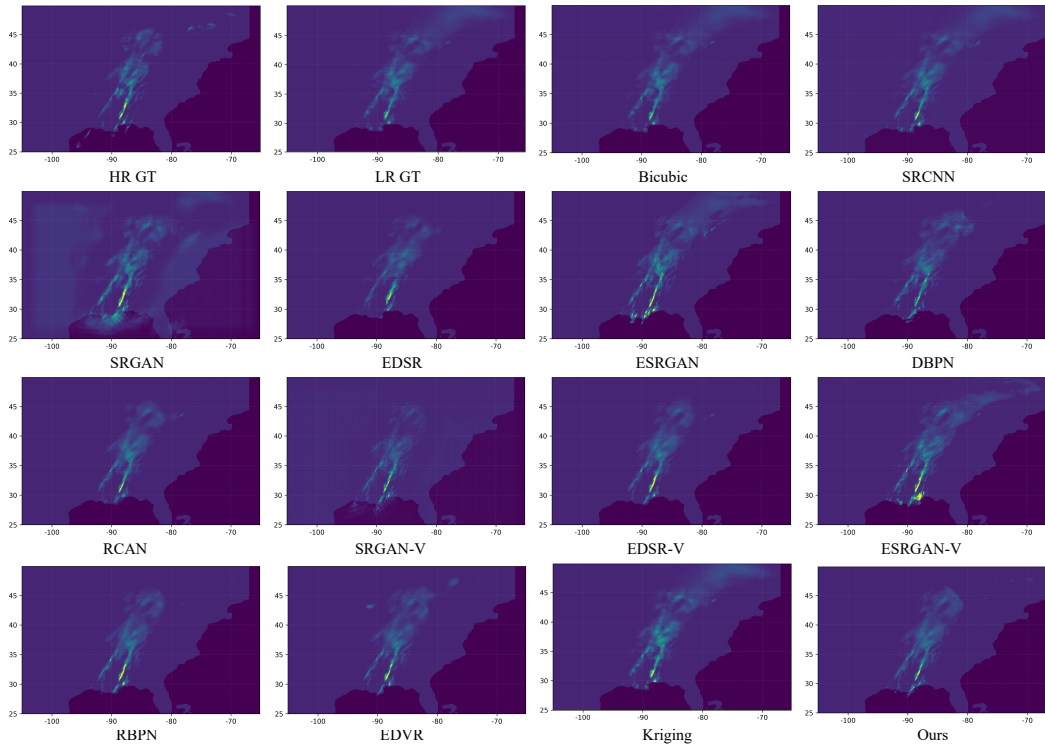


Figure 3: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 635-th hour in November 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

- 167 SC Pryor and JT Schoof. Differential credibility assessment for statistical downscaling. *Journal of*
168 *Applied Meteorology and Climatology*, 59(8):1333–1349, 2020.
- 169 Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen
170 Change Loy. Esgan: Enhanced super-resolution generative adversarial networks. In *Proceedings*
171 *of the European Conference on Computer Vision (ECCV)*, pp. 0–0, 2018.
- 172 Xintao Wang, Kelvin CK Chan, Ke Yu, Chao Dong, and Chen Change Loy. Edvr: Video restoration
173 with enhanced deformable convolutional networks. In *Proceedings of the IEEE Conference on*
174 *Computer Vision and Pattern Recognition Workshops*, 2019.
- 175 Adrienne M Wootten, Elias C Massoud, Agniv Sengupta, Duane E Waliser, and Huikyo Lee. The
176 effect of statistical downscaling on the weighting of multi-model ensembles of precipitation. *Cli-*
177 *mate*, 8(12):138, 2020.
- 178 Xuebin Zhang and Feng Yang. Rclimdex (1.0) user manual. *Climate Research Branch Environment*
179 *Canada*, 22, 2004.
- 180 Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-
181 resolution using very deep residual channel attention networks. In *Computer Vision - ECCV*
182 *2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part*
183 *VII*, Lecture Notes in Computer Science. Springer, 2018.

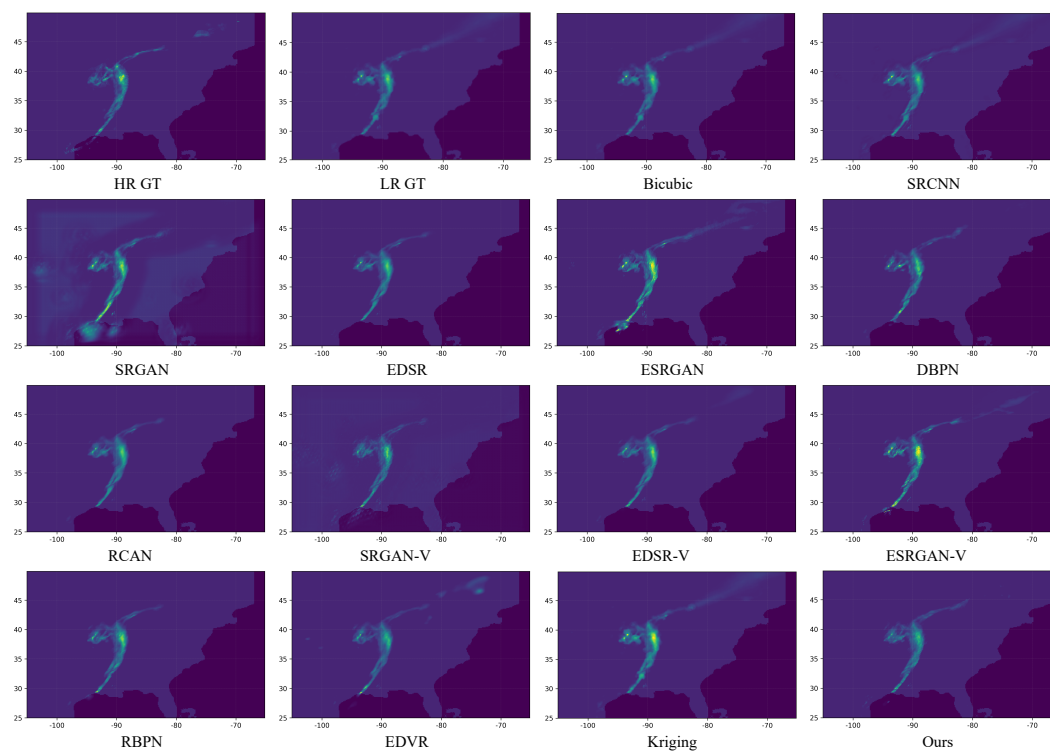


Figure 4: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 59-th hour in November 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

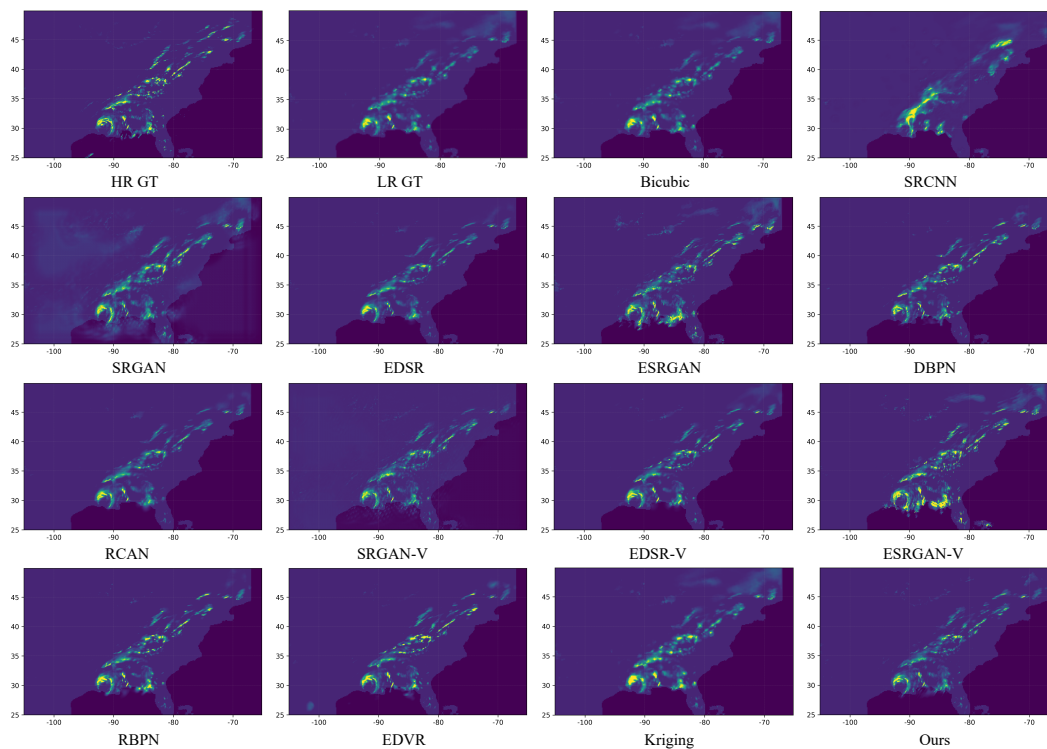


Figure 5: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 95-th hour in September 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

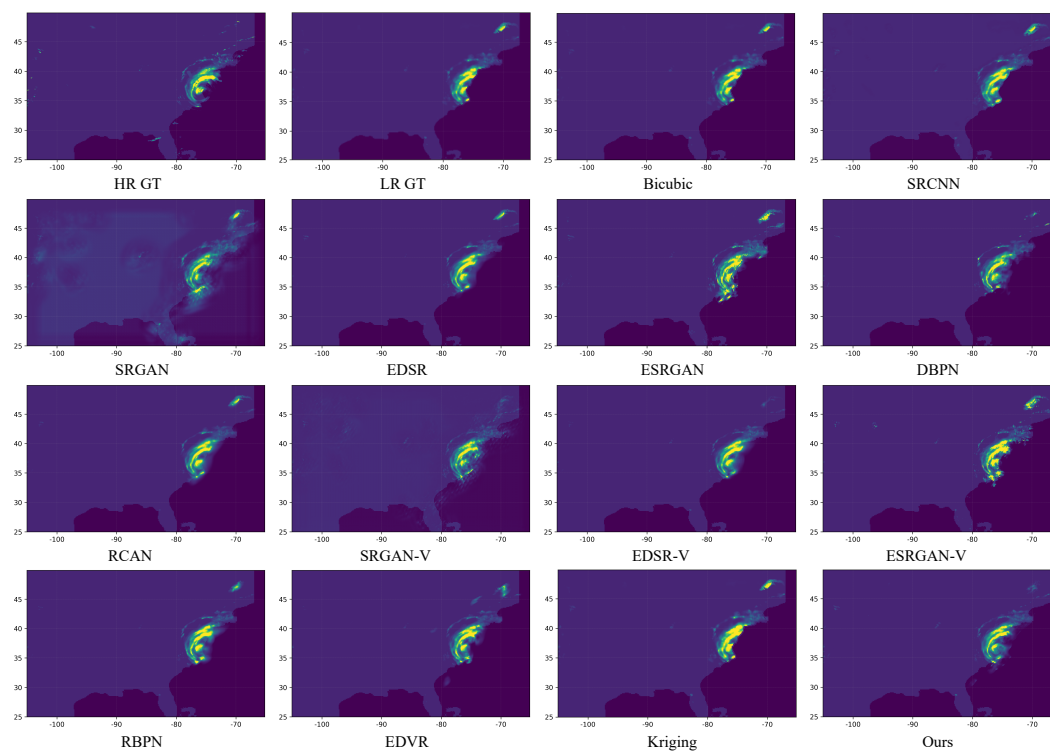


Figure 6: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 649-th hour in August 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

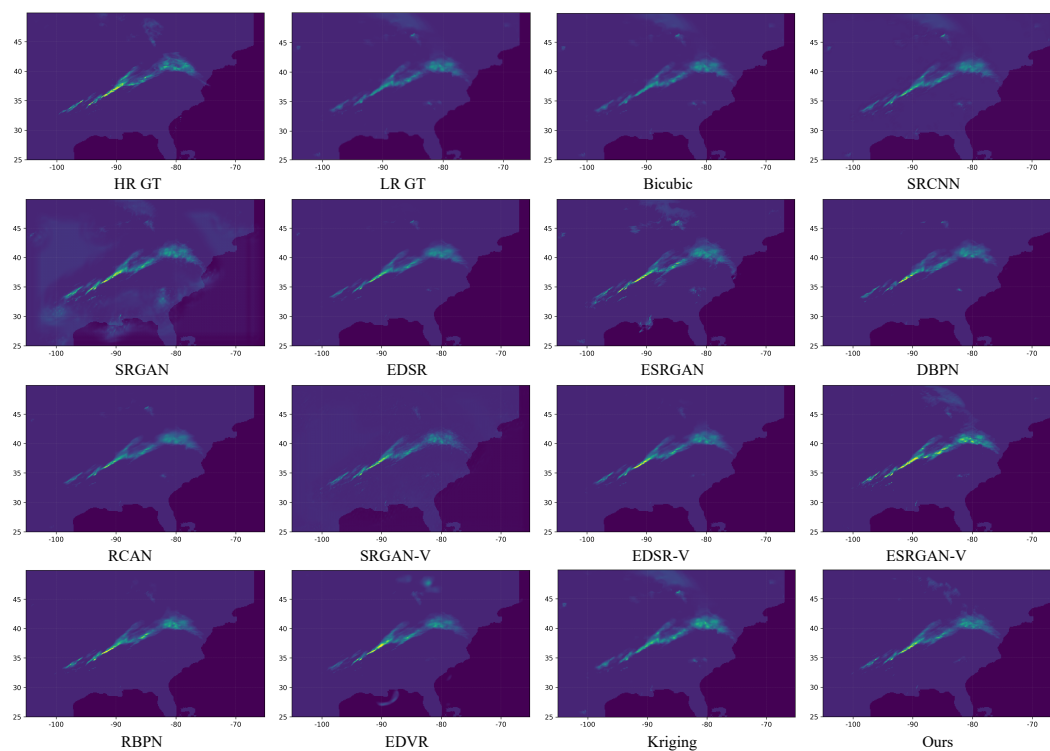


Figure 7: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 590-th hour in November 2010). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.