# RainNet: A Large-Scale Imagery Dataset and Benchmark for Spatial Precipitation Downscaling Supplementary Material


**Xuanhong Chen**[1*]    **Kairui Feng**[2*]    **Naiyuan Liu**[3]    **Bingbing Ni**[1†]    **Yifan Lu**[1]
**Zhengyan Tong**[1]    **Ziang Liu**[1‡]

[1]Shanghai Jiao Tong University    [2]Princeton University    [3]University of Technology Sydney

{chen19910528, yifan_lu, 418004}@sjtu.edu.cn
{kairuif}@princeton.edu
{naiyuan.liu}@student.uts.edu.au
{ziang_liu}@brown.edu


## 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. Our dataset is available at `https://neuralchen.github.io/RainNet/`.

## 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:*

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

*Mass Continuity:*

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

*Thermo-dynamic:*

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

*Ideal Gas:*

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

---

[*]Equal contribution.

[†]Bingbing Ni is the corresponding author.

[‡]Work done during undergraduate student at Shanghai Jiao Tong University.


36th Conference on Neural Information Processing Systems (NeurIPS 2022).


*Moisture equation:*

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

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 ($\rho$, 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 ($\theta$). The heat can also come from other heat sources ($Q$), *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.

## 3 Metrics

Due to the difference between down-scaling and traditional figure super-resolution, the metrics work well under SR tasks may not be sufficient for precipitation down-scaling. By gathering the most common metrics from the meteorologic literature (for example [15, 10, 2, 6, 11, 14]), we select and rename 6 metrics to reflect the downscaling quality: mesoscale peak precipitation error (MPPE), cumulative precipitation mean square error (CPMSE), heavy rain region error (HRRE) , cluster mean distance (CMD), heavy rain transition speed (HRTS) and average miss moving degree (AMMD).These 6 metrics can be separated as reconstruction metrics: MPPE, HRRE, CPMSE, AMMD, and dynamic metrics: HRTS and CMD.

The MPPE ($mm/hour$) is calculated as the difference of top quantile between the generated/real rainfall dataset which considering both spatial and temporal property of mesoscale meteorological systems, *e.g.*, hurricane, squall. This metric is used in most of these papers (for example [15, 10, 2, 6, 11, 14] suggest the quantile analysis to evaluate the downscaling quality).

The CPMSE ($mm^2/hour^2$) measures the cumulative rainfall difference on each pixel over the time-axis of the test set, which shows the spatial reconstruction property. Similar metrics are used in [15, 10, 14] calculated as the pixel level difference of monthly rainfall and used in [6] as a pixel level difference of cumulative rainfall with different length of record.

The HRRE ($km^2$) measures the difference of heavy rain coverage on each time slide between generated and labeled test set, which shows the temporal reconstruction ability of the models. The AMMD ($radian$) measures the average angle difference between main rainfall clusters. Similar metrics are used in [15, 10, 14] as rainfall coverage of a indefinite number precipitation level and used in [6, 11] as a continuous spatial analysis.

As a single variable dataset, it is hard to evaluate the dynamical reconstruction ability of different models. So here we introduce the first order variables to evaluate the dynamical property of down-scaling results. Similar approaches are suggested in [10, 2, 11]. The CMD ($km$) physically compares the location difference of the main rainfall system between the generated and labeled test set, which could be also understand as the RMSE of the first order derivative of precipitation data on spatial directions.The HRTS ($km/hour$) measures the difference between the main rainfall system moving speed between the generated and labeled test set which shows the ability for models to capture the dynamic property, which could be also understand as the RMSE of the first order derivative of precipitation data on temporal direction.Similar metrics are suggested in [10, 2, 11] as auto-regression analysis and differential analysis.

More details about the metrics and their equations are given in supplementary materials. One metrics group (MPPE, HRRE, CPMSE, AMMD) mainly measures the rainfall deviation between the generated precipitation maps and GT. The other group (HRTS and CMD) mainly measures the dynamic deviation of generated precipitation maps. In order to further simplify the application of indices, we abstract them into two weighted and summed metrics: Precipitation Error Measure (PEM) and Precipitation Dynamics Error Measure (PDEM). We first align the dimensions of these two groups of metrics respectively. The first group of metrics (MPPE, HRRE, CPMSE, AMMD) is normalized, weighted and summed to get the precipitation error measure (PEM). According to [3], all the metrics are transferred to Percent Bias (PBIAS) to couple with expert opinions for metrics weighting. The original definition of PBIAS is the bias divided by observation, as $PBIAS = |Q_{model} - Q_{observed}|/|Q_{observed}|$. Here we rewrite the original metrics to PBIAS by dividing the metrics with annual mean observations of the original variables (AMO), as $PBIAS_i^{PEM} =$


2


$|Metrics_i^{PEM}|/|AMO_i^{PEM}|$. The metrics then are ensembled to a single metric (PEM) with equal weight, as $PEM = \sum_i w_i \cdot PBIAS_i^{PEM}$. Following the same procedure, we 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 [15] and 20-return period in VALUE [10] 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 [15, 10, 2, 6, 11, 14] 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:

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

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 [15] and number of threshold exceedances in [10]. Here we select the threshold by the definition of heavy rain. We name this metric following its physics meaning.

Similar metrics are used in [15, 10, 14] as rainfall coverage of a indefinite number precipitation level and used in [6, 11] 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:

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

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 [15, 10, 14] calculated as the pixel level difference of monthly rainfall and used in [6] 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.

3

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:

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

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

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

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

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:

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

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

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

HRTS and CMD are comparable to auto regression analysis in VALUE (but not in CLIMDEX). These are the bias on first order regression on temporal and spatial dimensions. We use these metrics to reflect the dynamic property of the downscaling results and name these metric following its physics meaning. Similar metrics are suggested in [10, 2, 11] as auto-regression analysis and differential analysis.

## 4 Dataset Details

We show more precipitation maps in proposed dataset. In order to display the dynamic characteristics of the precipitation map more conveniently, we extract the precipitation maps of 4 periods and make them into short videos. These short videos are organized in the attachment of the supplementary materials (fold: "Dataset_Video").

## 5 Extra Results of RainNet

### 5.1 Detailed Network Structure

The structure of vanilla network in our proposed framework is given in Fig. 1. We employ 6 Residual-in-Residual Dense Blocks (RRDB) [13] in our downscaling backbone and 3 RRDB in the implicit dynamic estimation module network.

4

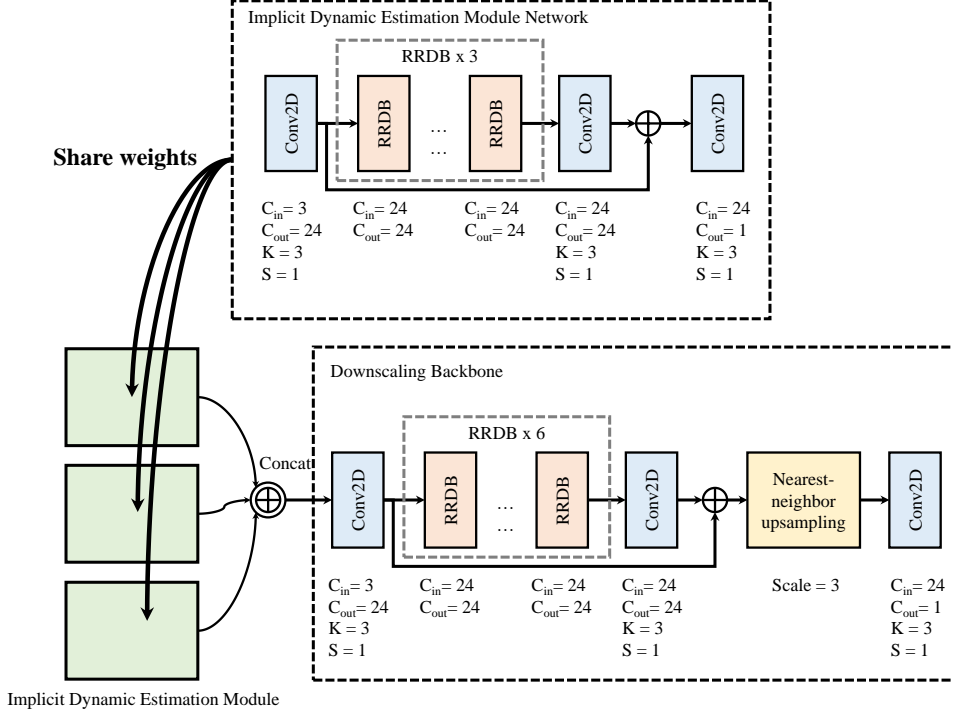



Figure 1: The details structure of vanilla network in our proposed model.

## 5.2 Training Details

We select 13 algorithms as benchmark: Bicubic [7], SRCNN[4] [1], SRGAN[5] [8], EDSR[6] [9], ESR-GAN[7] [13], DBPN[8] [4], RCAN[9] [16] and 5 VSR frameworks (*i.e.*, SRGAN-V, EDSR-V, ESRGAN-V, RBPN[10] [5], EDVR[11] [12] 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 down-scaling results (our proposed method) of 5 periods and make them into short videos. These short videos are organized in the attachment of the supplementary materials (fold: "Results_Video" and "Ours_Results_Video").

## References


[1] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2016.


---

[4]https://github.com/yjn870/SRCNN-pytorch

[5]https://github.com/leftthomas/SRGAN

[6]https://github.com/sanghyun-son/EDSR-PyTorch

[7]https://github.com/xinntao/ESRGAN

[8]https://github.com/alterzero/DBPN-Pytorch

[9]https://github.com/yulunzhang/RCAN

[10]https://github.com/alterzero/RBPN-PyTorch

[11]https://github.com/xinntao/EDVR

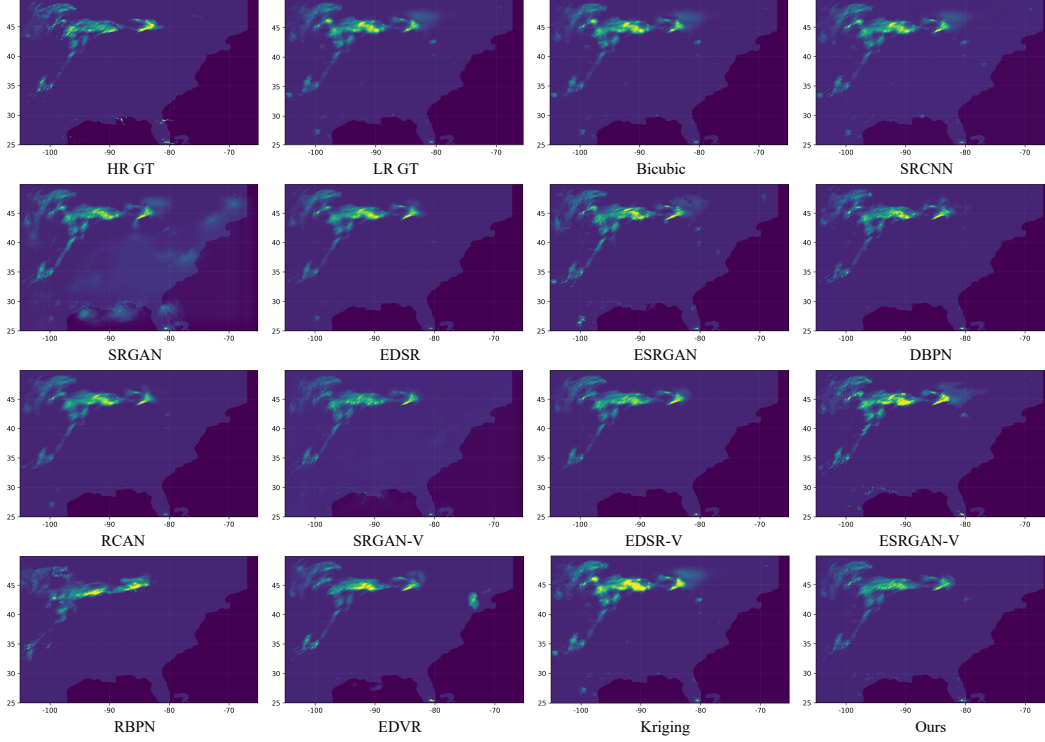

5


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.


[2] Marie Ekström. Metrics to identify meaningful downscaling skill in wrf simulations of intense rainfall events. *Environmental Modelling & Software*, 79:267–284, 2016.

[3] 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.

[4] Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deep back-projection networks for super-resolution. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. IEEE Computer Society, 2018.

[5] Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Recurrent back-projection network for video super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 2019.

[6] Xiaogang He, Nathaniel W Chaney, Marc Schleiss, and Justin Sheffield. Spatial downscaling of precipitation using adaptable random forests. *Water resources research*, 2016.

[7] Robert Keys. Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, 1981.

[8] Christian Ledig, Lucas Theis, Ferenc Huszar, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew P. Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, and Wenzhe Shi. Photo-realistic single image super-resolution using a generative adversarial network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*. IEEE Computer Society, 2017.

[9] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *2017 IEEE Conference on Computer*



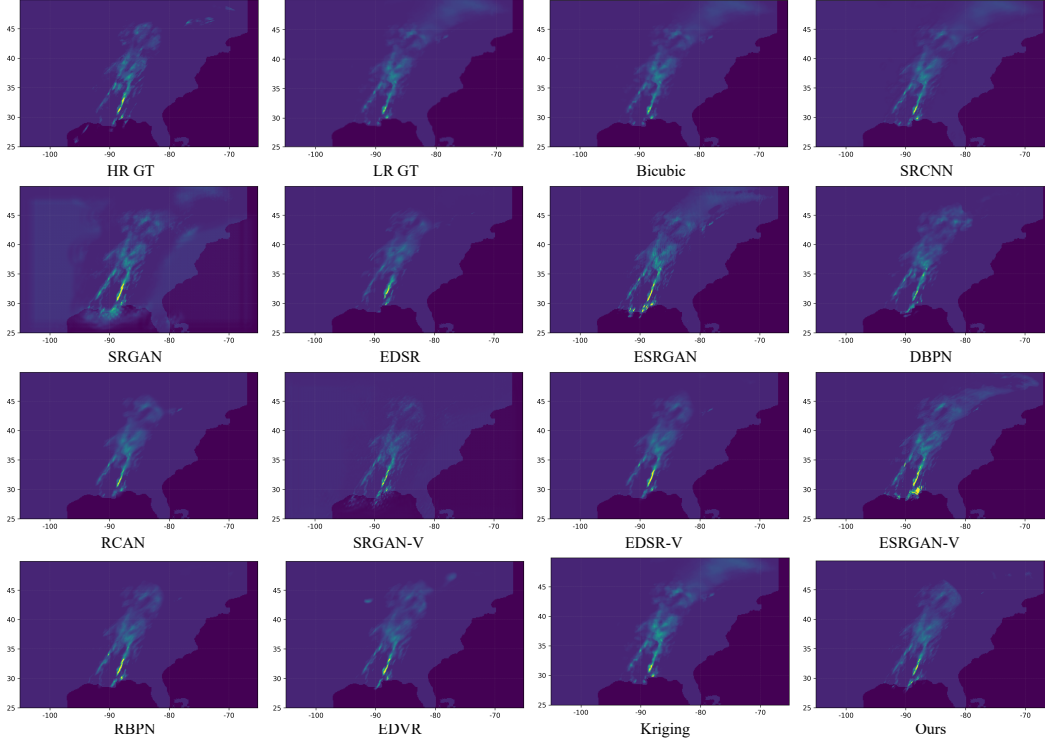
6


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.


*Vision and Pattern Recognition Workshops, CVPR Workshops 2017, Honolulu, HI, USA, July 21-26, 2017*. IEEE Computer Society, 2017.

[10] Douglas Maraun, Martin Widmann, José M Gutiérrez, Sven Kotlarski, Richard E Chandler, Elke Hertig, Joanna Wibig, Radan Huth, and Renate AI Wilcke. Value: A framework to validate downscaling approaches for climate change studies. *Earth's Future*, 3(1):1–14, 2015.

[11] SC Pryor and JT Schoof. Differential credibility assessment for statistical downscaling. *Journal of Applied Meteorology and Climatology*, 59(8):1333–1349, 2020.

[12] Xintao Wang, Kelvin CK Chan, Ke Yu, Chao Dong, and Chen Change Loy. Edvr: Video restoration with enhanced deformable convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[13] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 0–0, 2018.

[14] Adrienne M Wootten, Elias C Massoud, Agniv Sengupta, Duane E Waliser, and Huikyo Lee. The effect of statistical downscaling on the weighting of multi-model ensembles of precipitation. *Climate*, 8(12):138, 2020.

[15] Xuebin Zhang and Feng Yang. Rclimdex (1.0) user manual. *Climate Research Branch Environment Canada*, 22, 2004.

[16] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII*, Lecture Notes in Computer Science. Springer, 2018.



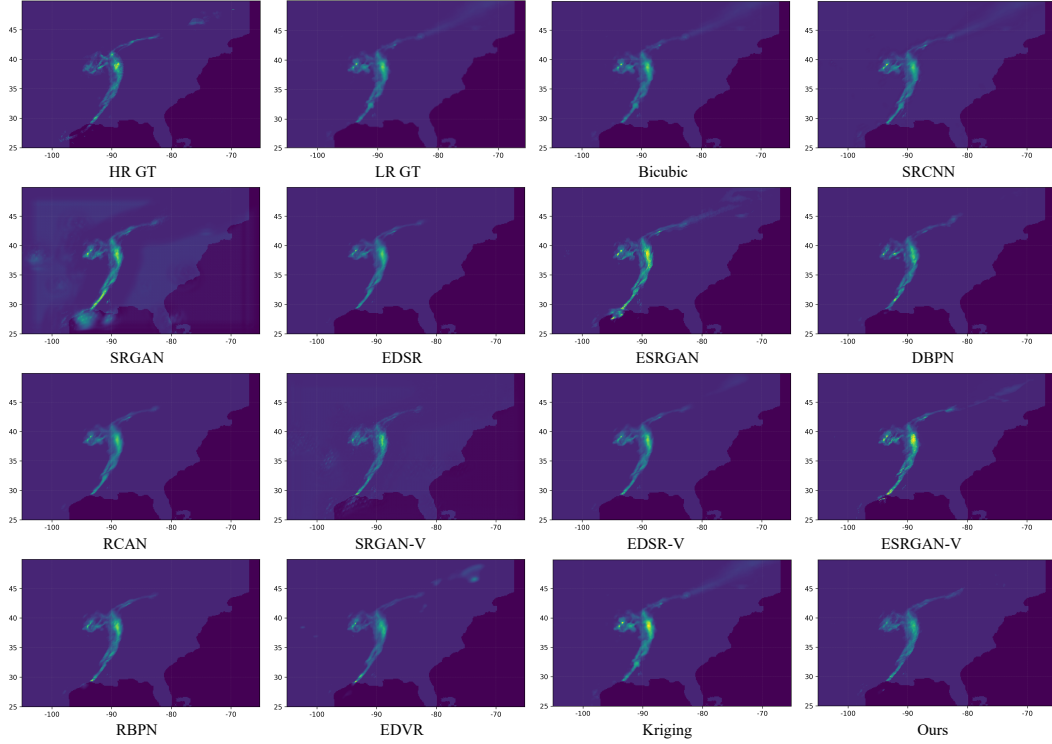
7


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.

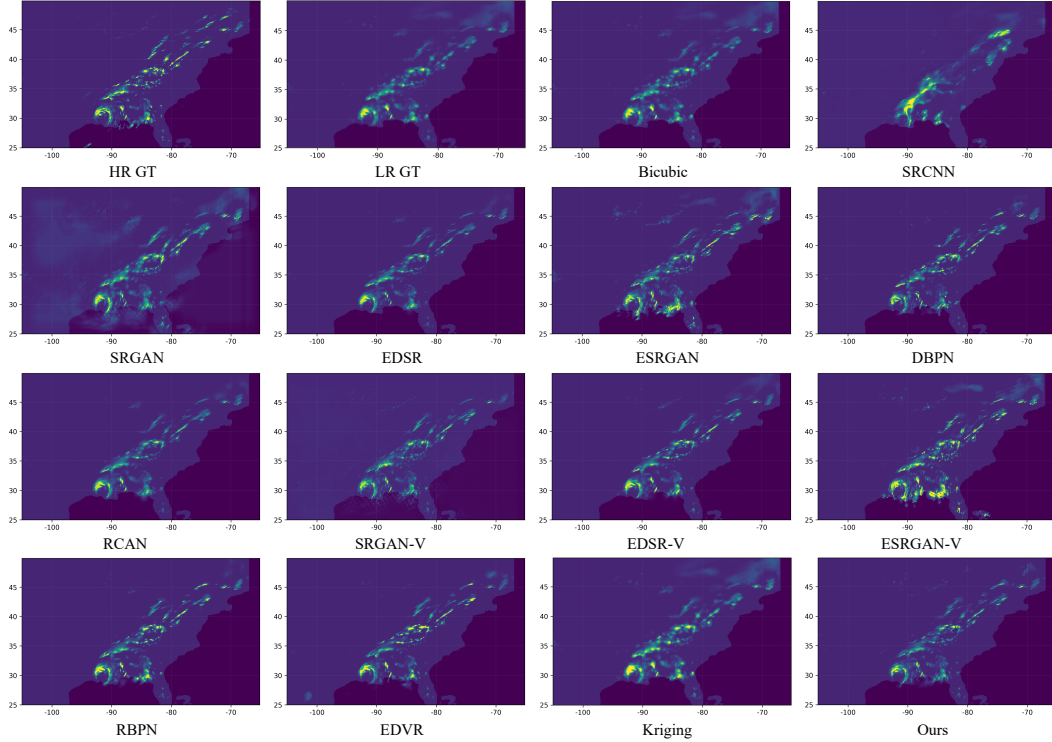



Figure 5: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 95-th hour in September 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

8

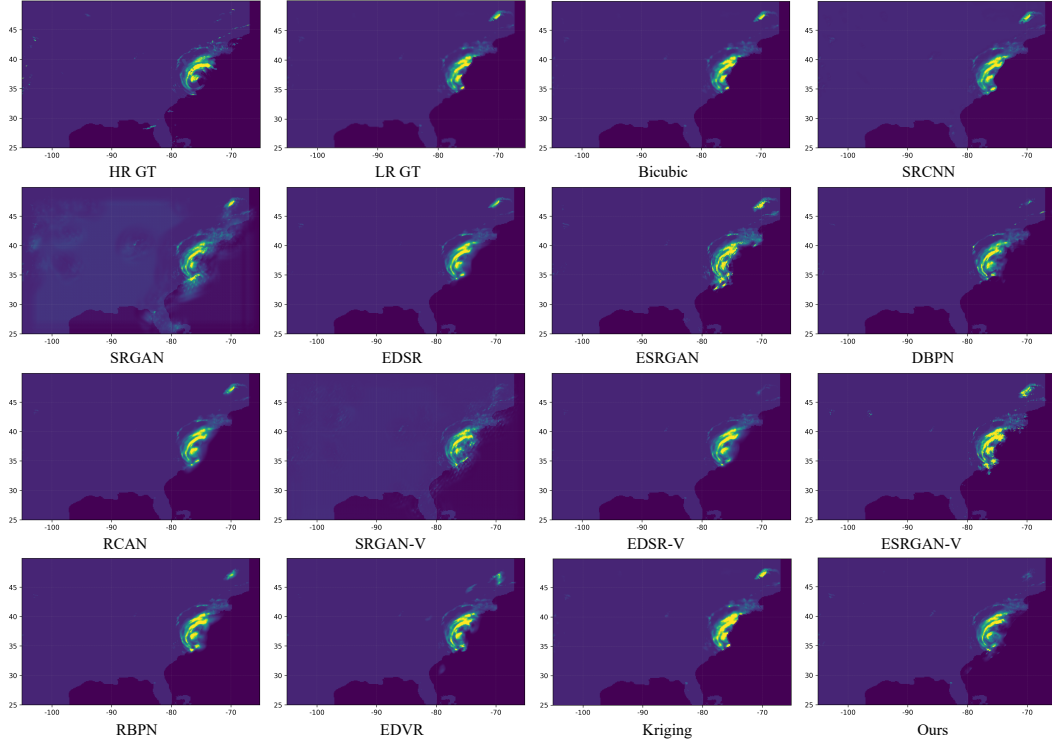



Figure 6: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 649-th hour in August 2011). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

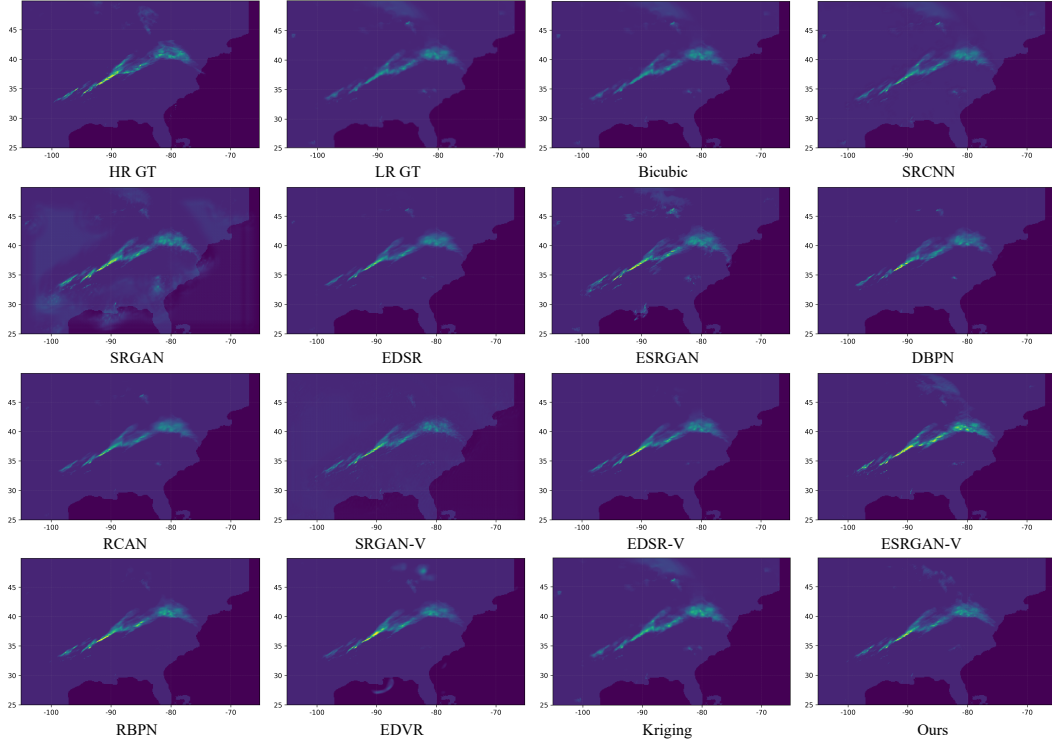



Figure 7: Visual comparison with state-of-the-art Super Resolution approaches(the specific time: the 590-th hour in November 2010). Please zoom-in the figure for better observation. Randomly picked results. Please note that the details of the precipitation map are partially lost due to file compression.

9