Seasonal rainfall forecast downscaling based on single image super resolution

Minzhe Chen

A thesis submitted for the degree of
Bachelor of Advanced Computing (Honours)
The Australian National University

January 2021

Except where otherwise indicated, this thesis is my own original work.

Minzhe Chen
14 January 2021

to my family and my supervisor

Acknowledgments

I would like to express my gratitude to my supervisor Dr Warren Jin. Firstly, he provided me with great opportunity to work on this project about seasonal weather forecast downscaling and deep learning. Dr Jin introduced this project and background knowledge to me when I started this project, and then his supervision, guidance, and encouragement were significant and helped me move towards the goal step by step. Remote learning is challenging, but our weekly online meeting provided me with a clear plan and great suggestions. When I encountered difficulties, his ideas and suggestions were always creative and helpful. Dr Jin also helped me with data analysis, which led to great breakthrough. Moreover, I was greatly influenced by his experience on research and academic writing. In summary, my supervisor Dr Warren gave me great help and supervision throughout this project.

I am also thankful to other students in the downscaling group, including Weifan Jiang, Quning Zhu and Rui Wang. Their experience and previous work gave me great help and inspiration. We exchanged ideas and help each other to make improvement.

I also want to thank the stuff of NCI, National Computational Infrastructure. They always answered my questions and fixed the problems effectively. My project with so much computation cannot be completed without their help.

Moreover, I thank my parents who helped me financially and spiritually. They always advised me to concentrate on research and encouraged me to believe myself and overcome the difficulties.

Finally, I would like to thank all my teachers, especially who taught me coding skills and machine learning knowledge, and friends, such as Ruiqi Chen, Weiliang Cao and Jiazong Gong, who helped me with thesis writing and encouraged me during the COVID-19 outbreak.

Abstract

Accuracy climate forecast can produce significant economic values. The resolution of current Australian climate prediction system ACCESS-S1 is too low, thereby post-processing climate forecast is important and necessary. However, traditional downscaling methods are not time efficient. Recent development of convolutional neural network provides a new way of climate postprocessing. However, the previous deep learning based Australian rainfall forecast downscaling model was not accurate for long-time prediction. In this project, I developed an adapted Very deep Super Resolution Downscaling (VDSRd2) model, which is a combination of Very Deep Super Resolution (VDSRd) and multiple-input downscaling model. My model uses precipitation and geopotential height as inputs to generate high-resolution rainfall ensemble forecast. An adapted structure of residual learning was used to extract features from geopotential height data. Leave-one-year-out cross-validation is used to evaluate the performance of my model. The results suggest that my model can predict more accurate precipitation ensemble forecasts than other downscaling methods. The results also demonstrate the significant skill of geopotential height for precipitation forecast in Australia. Moreover, my model can generate output images with better quality than previous super resolution based downscaling model, and time efficiency of VDSRd2 is better than traditional methods. This model can be used for Australian precipitation forecast downscaling, especially for long leading time prediction.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Thesis Statement

Compared to previous methods, my adapted very deep super resolution downscaling (VDSRd2) model with geopotential height data can improve the accuracy of ensemble forecasts in Australia, especially when the leading time is large.

1.2 Introduction

Climate and precipitation has significant influence on human activities and various weather-sensitive industries like agriculture, mining, energy, and infrastructure. Extreme rainfall may cause damage or economic loss. On the other hand, reliable and accurate forecast have positive influence via improving productivity or profitability in various ways such as finance, agribusiness, and health (Merryfield et al., 2020).Therefore, people always try to study the pattern of weather changes in depth and predict impactful weather events to adjust their activities accordingly.

Australia, a country with great agriculture and mining industry, made great effort to seasonal climate forecast to reduce the loss caused by impactful weather events and make profit for industries and residents. Bureau of Meteorology established a new multi-week to seasonal climate prediction system called ACCESS-S1 (Australian Community Climate and Earth-System Simulator-Seasonal) [Hudson et al., 2017]. It includes a variety of weather variables like precipitation, maximum temperature, geopotential height etc. Based on 23-year climate data, ACCESS-S1 enhanced the performance of seasonal weather forecast. However, the current resolution of ACCESS-S1 is 60 km, which would be too spare for individuals or even for industries such as agriculture and mining. On the other hand, Barra dataset, which can provide am atmospheric high-resolution regional reanalysis for Australia, can provides a target data to enhance the resolution of ACCESS-S1 forecast [Su et al., 2019]. To enhance the resolution and accuracy of current prediction system, many Australian scientists have made great efforts on seasonal forecast downscaling in Australia, and many traditional methods had great performance. For example, climatology predicts high resolution forecast by learning from historical data, and Calibration method uses quantile mapping to downscale the climate forecast [Maraun, 2013].

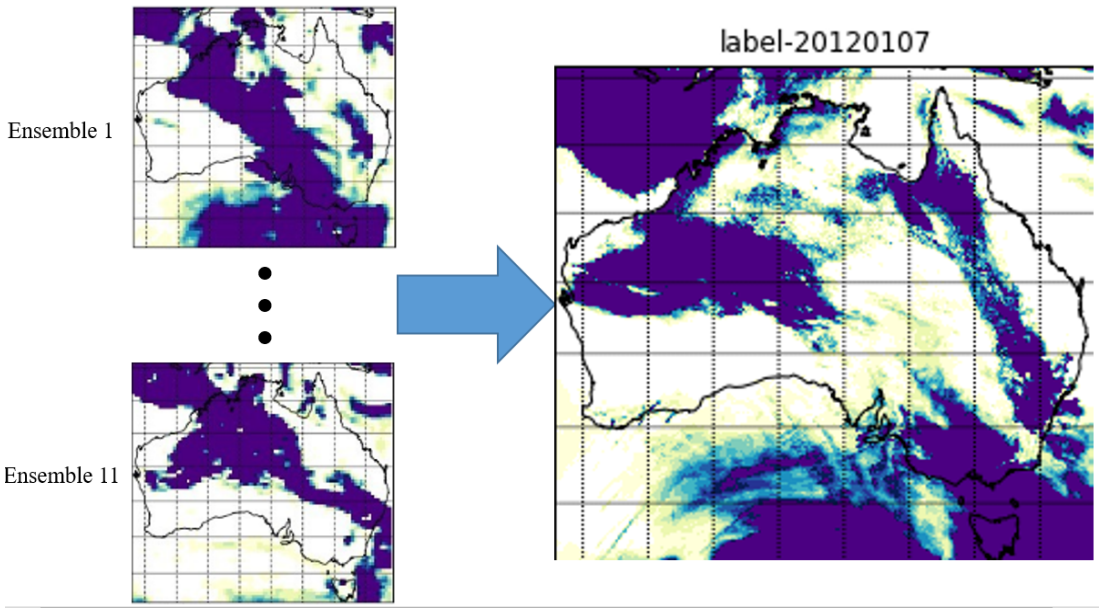

Figure 1.1: Different ensemble members share the same target HR image

In recent years, deep learning and convolutional neural network significantly improve the performance of single image super resolution (SISR) techniques. For example, a very deep super resolution neural network used residual learning and deep neural network to improved the image quality of the output images [Kim et al., 2016]. Because of its great performance and efficiency, SISR techniques have attracted attention of scientists and have been used in a wide range of industries. Regarding low resolution climate forecast as an image, various weather features like rainfall and geopotential height as different channels of the image, the resolution of such image can be enhanced by super resolution method, and high-resolution rainfall image can be generated.

Downscaling SISR model is different from traditional SISR model. For downscaling, each channel of climate image is not a colours channel with fixed range, but climate data with limitless range and sparse distribution, which could be more difficult than traditional super resolution. Moreover, climate downscaling model not only needs to enhance the image resolution but reduces the difference between forecast and reanalysis data. Moreover, as shown in the figure 1.1, there are eleven ensemble forecast members in training data, and all of them share the same high-resolution target, thereby the difference between all ensemble forecasts and ground truth value should be reduced. In other words, the downscaling model is not evaluated by the similarity of a pair of images, but by continuous ranked probability score (CRPS), which shows the reliability and accuracy of ensemble forecast over all ensemble members [Grimit et al., 2006].

Recently, SISR based downscaling method has been attempted and achieved good performance. For example, Jiang [2020] proposed a very deep super resolution downscaling (VDSRd) model. However, this model used only rainfall images as input so it only improves the image quality but cannot predict accurately for longer leading time. Moreover, VDSRd did not consider extreme rainfall situations. For example, the amount of rainfall around the Pacific will change dramatically during El Nino and La Nina periods [Philander, 1985]. Therefore, precipitation of those periods would be difficult to predict.

It is discovered that there is strong relationship between rainfall and other resolvable weather features such as temperature and pressure, so such resolvable weather features can play an important role in rainfall forecast. Recent study indicates that other weather features especially geopotential height can improve the accuracy of rainfall prediction in the United States of America [Pan et al., 2019]. Therefore, deep neural network with multiple climate features as inputs may improve the precipitation forecast performance and generate more accurate high-resolution precipitation forecast in Australia.

The aim of my project is to train a neural network to learn the underlying relationship between low resolution climate images and high-resolution precipitation image, and then generate accurate high resolution seasonal precipitation ensemble forecasts. The objectives are to enhance the resolution, and more importantly, reduce the difference between precipitation ensemble forecasts and actual observations. The method used is deep learning based SISR, which enhances the resolution of rainfall images to improve the accuracy of precipitation forecasts. To have more accurate ensemble forecast, other climate variables which could be helpful to predict the rainfall are also used. Therefore, I combined VDSRd model with pre-processed geopotential height images to reduce the difference between low-resolution rainfall ensemble forecasts and high-resolution reanalysis data. Geopotential height is found to be helpful in precipitation forecast, thereby I adopted the structure of residual learning to use geopotential height features to improve precipitation prediction. A new VDSR downscaling (VDSRd2) model is proposed for Australian precipitation forecast downscaling. In summary, the main contributions of my project are

- Present a new VDSRd2 model, which is a combination of VDSR downscaling model and multiple-input downscaling model
- Design a adapted structure of residual learning to use geopotential height features in VDSRd2
- First attempt of using geopotential height data in deep learning based Australian rainfall forecast postprocessing
- Demonstrate that geopotential height data has significant skill in precipitation forecast downscaling

- Generate high-resolution forecast that is more accurate than other downscaling methods

1.3 Thesis Outline

There are totally five chapters in the thesis. Chapter 2 includes the motivation and related work of climate forecast downscaling and SISR. Datasets used in this project are also introduced in this chapter, Chapter 3 demonstrates climate image pre-processing methods and the structure of my VDSRd2 model, Chapter 4 introduces the platforms and methodology of my experiments, Chapter 5 shows the evaluation results of my model, including typical output image demonstration and the evaluation of single image quality and ensemble forecast accuracy. The accuracy includes the forecast on whole Australia and 50 rainfall stations, and Chapter 6 discusses the conclusion of the thesis and potential future work.

Chapter 2

Background and Related Work

This chapter includes the motivation and related work of climate forecast downscaling and single image super resolution (SISR). The significant economic value of climate forecast downscaling will be discussed. Datasets used in this project, including low-resolution Australian climate forecast dataset and reanalysis dataset, will be introduced. Then I will discuss SISR techniques and analyse how to use SISR in climate forecast postprocessing. Current downscaling models and their limitations are also included.

Section 2.1 will analyse the importance of climate forecast downscaling. The main idea is that climate forecast has great economic values, but current prediction system is Australia has low resolution.

Section 2.2 includes the related work of weather forecast, downscaling and SISR. The datasets used in the project will be introduced, and the current methods of downscaling and SISR will be analysed.

2.1 Motivation

In this part, I will discuss the importance of seasonal climate forecast downscaling with SISR, especially for rainfall. Firstly, in section 2.1.1 I will discuss the importance of climate forecast and the limitation of current prediction system in Australia, which can show the necessity of downscaling. Then, in section 2.1.2 I will analyse the drawbacks of previous downscaling models and show the motivation of my project.

2.1.1 Climate Forecast

Climate forecast can benefit people in a variety of ways. Predicting extreme weather events can ensure people's safety and health, and early forecast of precipitation can also help people and industries to make profit.

Table 2.1: Potential value of forecast per year [CIE, 2014]

Industry	Potential value of forecast (AUD)	Industry value (AUD)	proportion
Construction	192	79851	0.2
Electricity	2.3	16566	0.01
Coal mining	68	20852	0.33
Oil and gas	93	20363	0.46
Transport	5	22824	0.02
Water supply	28	10550	0.27
Agriculture	1567	21429	7.31

Extreme climate changes can have negatively impact on people, but accurate and reliable forecast can help people prevent it in advance. Firstly, extreme weather events such as storm and hail may cause damage to the buildings and property, and even threaten people's safety. When there are extreme weather events, people tend to hide from it to protect themselves and their belongings, or even migrate in advance to avoid the damage. Therefore, weather forecast and early warning are essential and significant. Without accurate weather forecast, people cannot prepare and protect themselves in advance, which may lead to terrible consequence, including financial loss or even deaths. Moreover, weather can impact people's health. To be specific, when people are exposed to rain or cold environment, their migrate will decrease, which increases the risk of getting the virus. Fortunately, a reliable weather forecast can allow people to prepare in advance and greatly enhance their health.

Besides preventing people from extreme weather events, climate forecast can also help people and industries make more profits. This is because there are a variety of weather-sensitive industries such as agriculture, energy and mining, and weather forecast can allow them to make a early plan to maximize their output. As shown in table 2.1, the potential value of forecast is about 2000 million Australian dollars per year and improve agriculture by more than 7% [CIE, 2014]. If the climate forecast is accurate and reliable, farmers can decide what to grow to maximize their output. According to the studies from last 40 years, the economic contribution of seasonal climate forecast is approximately 60 million Australian Dollars for cotton and 40$ for wheat per year in Australia [Parton et al., 2019].

In summary, reliable and accurate forecast can benefit people in various ways such as finance and health [Merryfield et al., 2020]. Therefore, people always try to predict the weather.

2.1.2 Downscaling

In this part, I will discuss the importance of downscaling, and the drawbacks of current downscaling model to show the motivation of adapting SISR model for seasonal rainfall forecast downscaling.

Bureau of Meteorology has operated on low-resolution seasonal climate forecast system named ACCESS-S1 [Hudson et al., 2017]. It had great enhancement compared with previous prediction systems. However, its resolution is 60km, which means that each data item in the data set represents the average rainfall forecast of a 60km*60km area. Forecast on such a large area will lead to inaccuracy because the climate varies significantly in it. Therefore, low-resolution forecast may not be reliable and beneficial for local people. For example, a farmer may have a farm whose length is much smaller than 60km, or the farm is just on the edge of one data point, then the farmer can only get little information from the forecast because the data point does not exactly predict the climate of his farm. There are also other industries such as mining may have the same issue, and people working in weather-sensitive industries will make less profit because of the inaccurate and unreliable climate forecast [Merryfield et al., 2020]. If the resolution can be enhanced, more local people tend to get the climate forecast exactly on their position, which can be more accurate and reliable. Therefore, downscaling the current climate prediction is necessary and important.

Having known the importance of climate forecast downscaling in Australia, many scientists have done tremendous work to enhance the resolution and accuracy of seasonal climate forecast. For example, climatology method analysed historical climate data and predicted high-resolution forecast base on the climatology knowledge. Huge amount of data from previous 23 years and 242 ensemble members were used. Moreover, calibration used quantile mapping to enhance the resolution and had great result when the leading time is small [Maraun, 2013]. However, these traditional downscaling methods are not time efficient, which means that they are computationally expensive and require a great amount of data and knowledge. For example, as shown in the table 4.1, quantile mapping needs 0.3 second for each grid point and there are totally 886*691 gird points, which means this method need about 100 hours to downscale for whole season in Australia. By contrast, deep learning method will need only about 12 hours for training and 0.08 hour for prediction, which is much faster than the traditional methods such as quantile mapping. Therefore, recent study attempted to use deep learning based method to downscale the climate forecast in Australia. Moreover, SISR had great performance in recent years. Because climate forecast can be regarded as images, SISR can be used in seasonal rainfall forecast downscaling to enhance the resolution of the forecast. However, the difference is that possibility forecast is not to improve the quality of single image, but to predict a climate image base on 11 ensemble members from ASSCEE-S1 dataset.

Another gap from previous research is that other related resolvable climate variables are not used for rainfall forecast, but it was shown that other climate features like geopotential height is helpful to predict the precipitation in America [Pan et al., 2019]. Therefore, using other features may improve the accuracy of Australian seasonal rainfall forecast.

Table 2.2: Downscaling time of different methods

Method	Training time	Downscaling speed	Downscaling time
Deep learning (VDSRd)	12 h	0.03 s per image	0.08 h
Quantile mapping	$\sim$	0.3 s per grid point	8.3 h
ECCP	<1 h per point	0.45 s	61.1 h

Table 2.3: Terms of ACCESS-S1 dataset

Term	**Description**	**Values**
Variable	Climate variables	Precipitation, etc.
Prediction period	Target period of prediction	Monthly, Weekly, Daily
Initialization Date	The date when forecast is made	1,9,17,25 for each month
Target date	The predicted date of forecast	Every day
Leading time	Difference between initialization and target date	0-217, 1-7 for training
Ensemble	Different forecasts for the same date	1-11

2.2 Related work

2.2.1 Weather Forecast

2.2.1.1 ACCESS-S1

Australian Bureau of Meteorology established a new version of seasonal weather prediction system called Australian Community Climate Earth-System Simulator – Seasonal (ACCESS–S1)[Hudson et al., 2017]. This system includes state-of-the-art physics parameterisation schemes and has shown great enhancement compared to the previous model called POAMA based on the evaluation of the climate data from 23 years. The major improvement is that ACCESS-S1 has much smaller biases of the mean state of the climate, and it also has a good prediction of El Niño phenomenon. However, one major shortcoming of ACCESS-S1 is that its resolution is too low, and each climate prediction represent the mean climate data of 60km, which can make the forecast inaccurate and unreliable. Therefore, seasonal forecast downscaling can greatly improve the accuracy of weather forecast in Australia.

ACCESS-S1 system contains a variety of models such as atmospheric model and ocean models. These models can predict different climate variables such as temperature and precipitation. Its prediction methods are based on the UK Met Office GloSea5-GC2 seasonal prediction system, but ACCESS-S1 proposed a new strategy of ensemble generation which allows a larger ensemble size. For each weather feature, there are totally eleven ensemble members which indicates eleven different predicted cases. The accuracy of the ensemble forecast can be evaluated by the continuous ranked probability score (CRPS) [Grimit et al., 2006].

The atmospheric model describes the climates of Australian region with a resolution of approximately 60 km, which is used as the low-resolution input data in this project. There are three types of forecasts including daily, weekly and monthly

prediction. This model includes a variety of weather features such as precipitation, temperature, and sea level pressure. The time range is from 1990 to 2012 totally 23 years. For each year, there are totally 48 initialization dates, 4 in each month. The leading time, which represent the time difference between predicting date and target date, is from 0 to 217. For example, if the initialization date is 7th January 2000, and the leading time is 5, then this data represents the forecast of 7th January 2000 made 5 days prior to that day. Terms of ACCESS-S1 are shown in the table 2.3.

All climate forecast data is cropped in a rectangle range with Australia in the centre, and then the rainfall forecast is like a grayscale map, which is regarded as a low-resolution image. For geopotential height data, there are 20 pressure levels, and each level has a geopotential height forecast. Different levels of geopotential height may have different impact and information for precipitation forecast. Recent research discovered that the geopotential height around 850 Hectopascal (hPa) has significant influence of rainfall, and it can be used for rainfall forecast downscaling [Pan et al., 2019]. Therefore, geopotential height forecast is also used as the second channel of input image in this project.

2.2.1.2 BARRA

The Bureau of Meteorology Atmospheric high resolution Regional Reanalysis for Australia (BARRA) [Su et al., 2019] is a climate reanalysis dataset. As shown in figure 2.1, its region includes Australia, New Zealand, Southeast Asia, and surrounding islands. The resolution of BARRA is approximately 12km, which is four times smaller than the resolution of ACCESS-S1[Hudson et al., 2017]. Moreover, there is no ensemble members in BARRA dataset, it has only one rainfall data, which can be regarded as a grayscale image for each date. Therefore, the cropped Australian region of BARRA rainfall data is used as high-resolution target data, and all eleven ensemble members are used to predict a high-resolution rainfall forecast, and the output should be as closed to BARRA rainfall data as possible.

2.2.1.3 Probabilistic Forecast

For possibility forecast, because there are eleven ensemble members in ACCESS-S1 dataset and one image in BARRA data set, eleven high-resolution ensemble predictions will be produced after downscaling. Therefore, a criterion must be used to evaluate how good the possibility forecast is. The continuous ranked probability score (CRPS) can provide a proper way to evaluate probabilistic ensemble forecast with a single metric [Grimit et al., 2006]. The CRPS score is defined as follow.

$$CRPS(F, x) = \int_{-\infty}^{\infty} [F(y) - \mathbf{1}(y \geq x)]^2 dy \tag{2.1}$$

Where

$$\mathbf{1}(X) = \begin{cases} 1 & , X = Ture \\ 0 & , X = False \end{cases} \tag{2.2}$$

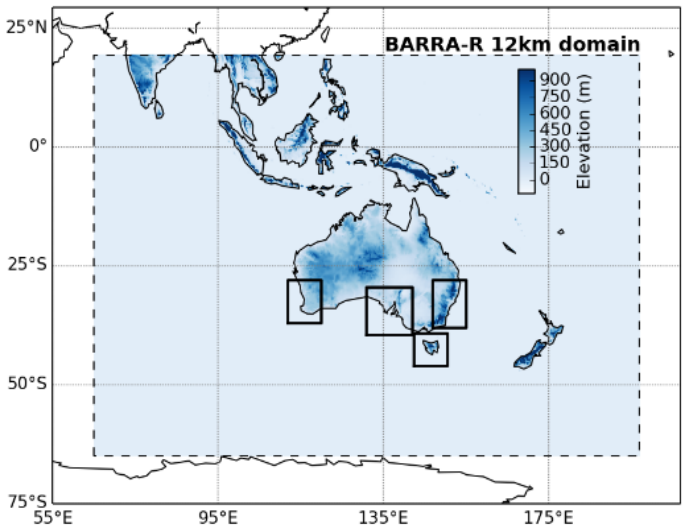

Figure 2.1: The domain of Barra dataset [Su et al., 2019]

and y and x are the predicted and ground truth values of a data point. F is the cumulative distribution function of forecast distribution [Grimit et al., 2006]. If the CRPS is small, there is little difference between the observation and the prediction, then probabilistic ensemble forecast is more reliable.

To estimate the CRPS, the skill score based on CRPS is defined as

$$CRPSSS = \frac{CRPS - CRPS_{ref}}{0 - CRPS_{ref}} = 1 - \frac{CRPS}{CRPS_{ref}} \tag{2.3}$$

where $CRPS_{ref}$ if the CRPS score of a reference climatology based probabilistic forecast.

If the CRPS skill score is larger, the probabilistic ensemble forecast is better, and if the CRPS skill score is positive, the forecast will be better than the reference on average.

2.2.1.4 El Niño and La Niña

El Niño and La Niña are two phases of southern Pacific when the temperature changes of sea surface lead to instability of atmosphere [Philander, 1985]. Australia is on the west side of the south Pacific, thereby Australia is significantly influenced by El Niño and La Niña and the precipitation during that periods may be extreme and hard to predict. It is observed that there were strong La Niña events around 2011 in Australia, thereby 2010 and 2011 were one of the wettest periods on record for Australia, and several flooding events happened during that period, which caused significant damage to local people [aus]. Therefore, downscaling and predicting pre-

cipitation of that period should be challenging and difficult.

2.2.2 Single Image Super Resolution

Single image super resolution (SISR) is enhancing the resolution of the low-resolution image to improve the image quality and get a high-resolution image. In other words, if the resolution of image is higher, the image tends to be more clear, and SISR is to divide one pixel into several pixels to make image clearer. Suppose SR and LR are the output high-resolution image and low-resolution input image, the SISR function can be defined as

$$SR = SISR(LR, \theta) \tag{2.4}$$

Where θ is the parameters. The size of HR should be larger than LR, which means HR image has more pixels. For climate forecast images, if one pixel is divided into several data points, then the scale of each data point is reduced, thereby enhancing the resolution of the images can downscale seasonal precipitation forecast.

2.2.2.1 Traditional SISR methods

Before the use of deep learning techniques, there were many traditional SISR methods which had good performance. For example, a traditional method of SISR is to use interpolation, which estimates the intermediate value between original image pixels to make the image smoother. Keys [1981] summarized that the basic interpolation function can be written as the follow formulation.

$$I(x) = \sum_k c_k u(\frac{x - x_k}{h}) \tag{2.5}$$

Where I is the interpolation function, x_k represents the interpolation node, h is the sampling increment, u is the kernel of interpolation, and parameter c_k depends on the sampled data. However, this method cannot have very good performance for normal images where there are many edges and corners.

Before using deep learning techniques, a traditional method called Adjusted Anchored Neighbourhood Regression (A+), which combined Anchored Neighbourhood Regression and Simple function, got great performance of improving image quality [Timofte et al., 2014].

2.2.2.2 Deep learning based SISR

From 2014, the use of convolutional neural network significantly improved the performance of SISR. Given an image and a kernel, for each pixel at (m,n) in the image, the convolution computation can be expressed as

$$I_{(m,n)} = \sum_i \sum_j k_{(i,j)} * I_{(m-i,n-j)} \tag{2.6}$$

Where K is kernel and I is the image. $k_{(i,j)}$ represents the elements at (i,j) in the kernel. The kernel will slide over the image with a given step length and compute the result to produce a map. Padding may be used to address the edge issue.

Convolutional layer is a layer that do convolution computation over input image and output feature maps. The size and number of the kernels are hyper-parameters which should be defined in advance, and the value of the kernel will be learned through backpropagation. ReLU layer is very often used with convolutional layer in the neural network. It is a layer to compute the result of the ReLU activation function on each pixel in the image:

$$f(x) = max(0, x) \tag{2.7}$$

. Convolutional layer can extract features from the image, and ReLU layer can remove the meaningless negative value. With the use of those two layers, convolutional neural network had great performance on computer vision and SISR.

Deep learning based SISR method started to attract attention from people. The first proposed model is Super resolution convolutional neural network (SRCNN) [Dong et al., 2015], which was the first successful convolutional neural network that was used as a SISR model. The structure of SRCNN is relatively simple. It has only three convolutional layers and two ReLU layers between them. The first convolutional layer is used to extract features from the upsampled input image. Then the second convolutional layer will produce high level feature vectors. The last layer reconstructs the high-resolution output image with the feature map. SRCNN achieved impressive performance. Therefore, deep learning based SISR became popular and many scientists started to study this field.

After that, because of the development of hardware and deep learning techniques, great improvement was made by using more advanced deep learning techniques and increasing the depth of the network. Residual learning, which uses skip connections between layers, was used in SISR to avoid gradients vanishing, thereby the deeper neural network could be trained by gradient descent. A Very Deep Super Resolution [Kim et al., 2016] model used residual learning with more than ten convolutional layers. Instead of predicting the whole new image, residual learning allows the model to learning the difference between the low-resolution image and high-resolution image, so the model can focus on the significant parts of the image such as edges and corners. Moreover, the larger number of convolutional layers allows the model to learning more complex patterns from the dataset. Residual channel attention network [Zhang et al., 2018] and second order attention network [Dai et al., 2019] used recursive residual design and channel attention mechanism to improve the super resolution performance. The recent RCAN and SAN showed slightly better performance in traditional SISR, but for climate forecast downscaling, climate images generally do not have clear edges and corners, and the channel is not like RGB channels but precipitation and other climate variables, thereby the attention

Figure 2.2: The overall structure of SRCNN and VDSR [Anwar et al., 2019], the greed rectangles represent convolutional and ReLU layers

network may not have better performance in climate forecast downscaling [Jiang, 2020].

2.2.3 Downscaling

ACCESS-S1 is a climate forecast system which has great enhancement to previous models, and it became the weather forecast model for Bureau of Meteorology in Australia [Hudson et al., 2017]. However, its resolution is only approximately 60km, thereby many scientists tried to downscale and enhance the resolution. Downscaling is to reduce the range of area represented by each data point in the climate forecast data. Regarding climate data as images, downscaling is the same as enhancing the resolution of the precipitation forecast image.

2.2.3.1 Interpolation

A straightforward method of enhancing image resolution is interpolation. Compared with normal image, this method works better for climate data because climate images are relatively smooth, which means there are few edges or corners in the climate image. Therefore, bicubic interpolation can improve the image quality of climate forecast to some extent. However, climate forecast downscaling is not only improving the image quality but to predict a high-resolution forecast based on all eleven ensemble members, which means that the difference between each forecast and target data should be reduced. Therefore, the performance of bicubic interpolation cannot be perfect because it does not predict the amount of rainfall according to the knowledge. To address this problem, more advanced traditional methods using climatology knowledge were proposed.

2.2.3.2 Traditional postprocessing downscaling methods

A traditional method called cross-validated climatology forecast used historical data to downscale the forecast. To predict a high-resolution weather forecast on a date, the climatology model would study what happened on the same date in previous years with a 11-day slide window, and analyse its relationship using knowledge of climatology [Li and Jin]. It totally uses 242 ensemble members to learn the relationship

and predict high-resolution rainfall. Its performance is impressive and this model is generally used as the reference model for skill score computation.

Quantile Mapping (QM) [Maraun, 2013] is another significant method which has great performance when the leading time is not large. It learned the matching of an empirical distribution from an 11-day slide window.

Extended copula postprocessing (ECPP) used Copula function to derive likelihood estimation and stimulate the ensemble forecast [Li and Jin]. The framework of copula model can be formulated as follows.

$$F(X_o, X_f) = C\{F_o(X_o), F_f(X_f)\} \tag{2.8}$$

Where C is the copula function and the Xo and Xf is the bivariate distribution function between true and raw forecast rainfall from a physical climate model. Then this can be used to estimate the likelihood of samples by multiplying all possibility from difference cases. After estimation, the ensemble forecast can be generated by using the quantile function and the stimulated value.

A problem of tradition methods is that, as shown in table4.1, the computation costs are expensive, thereby it is not efficient to use traditional methods to predict a high-resolution forecast. For example, QM will need 0.3 second to predict a gird point and the total number of gird point is about 90,000. To predict a high-resolution forecast in Australia, this method requires about 100 hours, which might be too expensive for daily forecast. Therefore, more efficient deep learning based downscaling methods became popular in recent years.

2.2.3.3 Deep learning based Downscaling

Recently deep learning techniques was greatly improved and widely used. Deep learning based methods have achieved great performance in many fields such as computer vision and SISR. SISR methods based on convolutional neural network can efficiently improve the image quality, which is much better than the performance of traditional SISR methods such as A+ . Therefore, deep learning based SISR methods started to be used in a variety of fields such as medical imaging and climate forecast downscaling. A VDSR model was used to downscale the seasonal climate forecast in Australia, which was much more efficient than traditional methods. It was trained by the data of leading time 0-7, from 1990 to 2011, and its downscaling result was impressive when the leading time is less than 7 for CRPS validation of 2012 [Jiang, 2020]. However, this deep learning based model has worse CRPS result when the leading time is larger than the leading time used for training, and its validation of other years may be worse because of some extreme rainfall caused by El Nino [Philander, 1985].

2.2.3.4 Downscaling with geopotential height

There is significant difference between original SISR on general images and its application in climate forecast downscaling. This is because the aim of downscaling is probabilistic ensemble forecast, thereby the model should not only improve the image quality by enhancing the resolution but need to use all ensemble members to predict high-resolution rainfall images that are closed to the target image. The result is not only evaluated by Peak signal-to-noise ratio (PSNR), but by CRPS and CRPS skill score, which can show a possibility that the ensemble members are reliable.

Because there is difference between the forecast and the target image, the model should learn how to improve the forecast while enhancing the resolution, thereby other weather features might be helpful. For example, geopotential height is found to be useful in rainfall downscaling [Pan et al., 2019]. Geopotential height is similar to geometric height but adjusted with gravity. Given an elevation h, geopotential is defined as

$$\Phi(h) = \int_0^h g(\phi, z)dz \tag{2.9}$$

where $g(\phi, z)$ is gravitational acceleration under altitude ϕ and geometric elevation z[Minzner et al., 1976]. And geopotential height is defined as

$$Z_g = \frac{\Phi(h)}{g_0} \tag{2.10}$$

where g_0 is the standard gravity at mean sea level. Geopotential height thereby shows the variation of gravity with the altitude and latitude.

A study in the united states used geopotential height from three levels (500, 850, 1000 hPa) as input to improve the daily precipitation prediction with convolutional neural network [Pan et al., 2019]. Four channels of climate data were stacked as an input image to go through convolutional and ReLU layer, and then performed max pooling, followed by a dense layer to reconstruct the high-resolution image. Their result indicates that geopotential height can contribute to rainfall downscaling, thereby using geopotential height may also improve the performance of deep learning based Australian seasonal forecast downscaling model. Therefore, in my project, I combined VDSR with this multiple-input downscaling model, and downscale the precipitation forecast in Australia.

2.3 Summary

Because climate changes have significant impact on human activities, accurate and reliable weather forecast is valuable in modern society. It benefits weather-sensitive industries, especially agriculture, approximately 7 percent of total agricultural value come from accurate weather forecast [CIE, 2014; Merryfield et al., 2020].

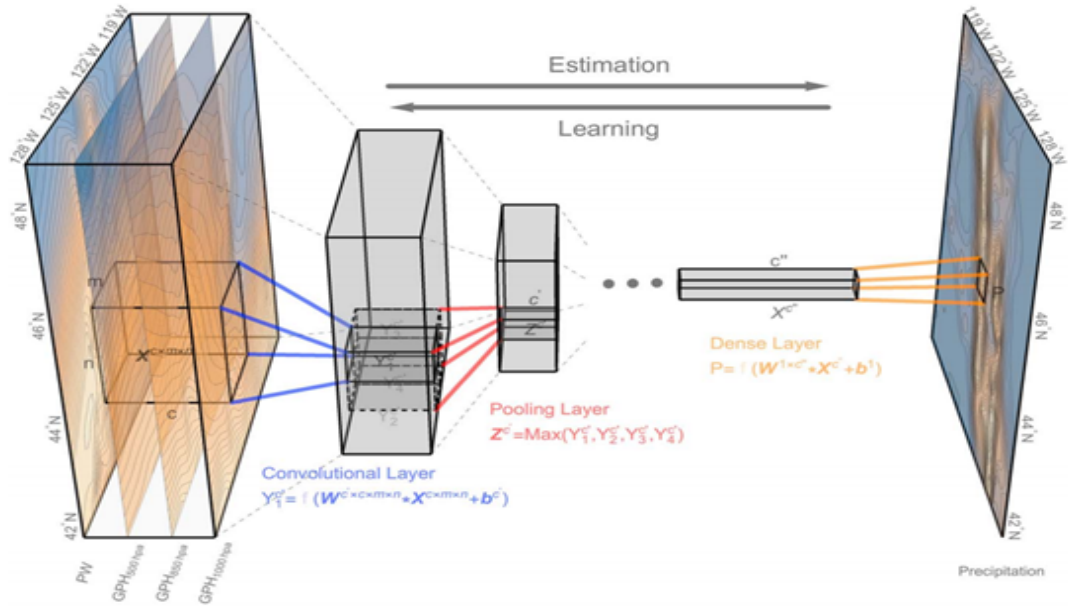

Figure 2.3: Rainfall downscaling model structure with geopotential height [Pan et al., 2019]

Having realized the importance of weather forecast, Bureau of Meteorology started to operate on a new weather prediction system called ACCESS-S1, which has eleven ensemble members and show great improvement compared with previous system [Hudson et al., 2017]. However, the current resolution of ACCESS-S1 is 60km, which is too low. Therefore, seasonal climate forecast downscaling and postprocessing is necessary and important. On the other hand, BARRA dataset contains high-resolution reanalysis climate data, which can be used as label of the ensemble forecast of ACCESS-S1.

Traditional postprocessing method climatology used historical data of the same date with 11-day sliding window to predict a high-resolution rainfall image [Li and Jin]. Climatology is used as reference model of this project. Moreover, other traditional methods such as quantile mapping [Maraun, 2013] and Extended copula postprocessing [Li and Jin] also achieved great result in climate forecast downscaling when the leading time is small. However, traditional postprocessing approaches are computationally expensive, because each of them requires at least 8 hours to downscale a rainfall forecast of whole Australia.

Single image super resolution (SISR) is a technique to enhance the resolution of images. Upsampling and interpolation are the easiest approaches of SISR and downscaling. Super Resolution Convolutional Neural Network (SRCNN) [Dong et al., 2015] was the first successful deep learning based SISR model. After that, residual

learning and attention mechanism was used to improve the performance of SISR model. For example, very deep Super Resolution model uses more than ten convolutional layer to extract deeper features of the upsampled input image, and its skip connection allows the model to learn only the difference between the input and target.

Regarding seasonal climate forecast as images, SISR can be used to downscale the forecast. But the main difference is that downscaling model should not only enhance the resolution but reduce the difference between forecasts and the corresponding ground truth value and generate eleven high-resolution ensemble forecast. The result is evaluated by the continuous ranked probability score (CRPS), which is a good metric of probabilistic forecast [Grimit et al., 2006]. Jiang [2020] attempted to used VDSR model to downscale the rainfall forecast in Australia, the performance was promising but worse than Climatology after 7-day leading time in 2012 validation. Study in America showed that geopotential height around 850 hPa can improve the postprocessing performance of deep learning based model [Pan et al., 2019]. Therefore, combining VDSR downscaling model and geopotential height data may improve the result in Australian seasonal climate forecast downscaling.

Chapter 3

Design and Implementation

According to the previous work, we understood the importance of seasonal climate forecast and postprocessing. However, the previous VDSRd model [Jiang, 2020] did not have much improvement compared to traditional methods such as climatology. The reason is that although the deep convolutional neural network enhanced image resolution, there is still significant difference between the precipitation forecast and reanalysis data. Therefore, a new approach must be made to learn the pattern of rainfall and improve the reliability of forecast. The new model should not only improve the image quality evaluated by Peak signal-to-noise ratio (PSNR) but should have better ensemble forecasts evaluated by the continuous ranked probability score (CRPS). To improve the accuracy of rainfall forecast, other climate variables such as geopotential height was helpful in American downscaling model [Pan et al., 2019]. Therefore, I proposed a combination of the VDSR downscaling model with geopotential height input, to produce more accurate seasonal precipitation prediction.

A new model was designed as the combination of the Very Deep Super Resolution downscaling model and multiple-input downscaling model. The overall structure of my downscaling method is

1. Normalization and upsampling

2. Feature extraction and residual learning

3. Image reconstruction

Section 3.1 I will discuss the methods of data pre-processing, including processing input and label images, and normalization method.
Section 3.2 introduces the design and structure of my VDSRd2 model. The overall structure will be described, followed by detailed functions of each layers in the network.
Section 3.3 lists the training details such as hyper-parameters and loss function.

```
|-- access-s1
|   |-- model : atmosphare model
|       |-- variable: precipitation
|       |   |-- forecast period: daily forecast
|       |   |   |-- date: 1990-2011
|       |   |   |   |-- ensemble: 01-11
|       |   |   |   |   |-- leading time: 1-7
|       .   .   .
|       |-- variable: geopotential height
|       |   |--  forecast period: daily forecast
|       |   |   |-- date: 1990-2011
|       |   |   |   |-- ensemble: 01-11
|       |   |   |   |   |-- leading time: 1-7
```

Figure 3.1: File structure from ACCESS-S1 used for training

3.1 Data Pre-Processing

The input data is low-resolution climate forecast images from ACCESS-S1 dataset, the output is high-resolution forecast images. The label is high resolution rainfall images from BARRA reanalysis dataset. For each date, both input and output contain eleven ensemble members, but there is only one label image.

This part will introduce the input, output, label, and pre-processing methods including normalization and upsampling.

3.1.1 Input Data

The input is low-resolution climate forecast images. The image contains two channels, which are the precipitation and geopotential height forecast of the same date. The forecast is projected on the Australian map to form an image, which means the value of each pixel is the forecast value of the corresponding area.

The training data is daily forecast climate images from atmosphere model, ACCESS-S1 dataset. Precipitation and geopotential height forecast images are used. Time range for training data is about 22 years, with one year left for validation. For each year, there are 48 initialization dates. The leading time of ACCESS-S1 is 0-217, but only 1-7 is used for training. For each climate variable, there are 11 ensemble members and all of them is used for training. The file structure is shown in the figure 3.1. For each forecast image, it has a initialization date, which means the date when the forecast has been done, and a target date of the forecast. For example, on 1st January, if a forecast of tomorrow is made, then the leading time is 1, the initialization date is 1st January, and the target date is 2nd January. The relationship can be formulated as follow.

$$Target\ date = Initialization\ date + leading\ time \tag{3.1}$$

Before entering the neural network, the climate forecast data from ACCESS-S1 dataset will be rearrange as images. Then two weather variables of the same date,

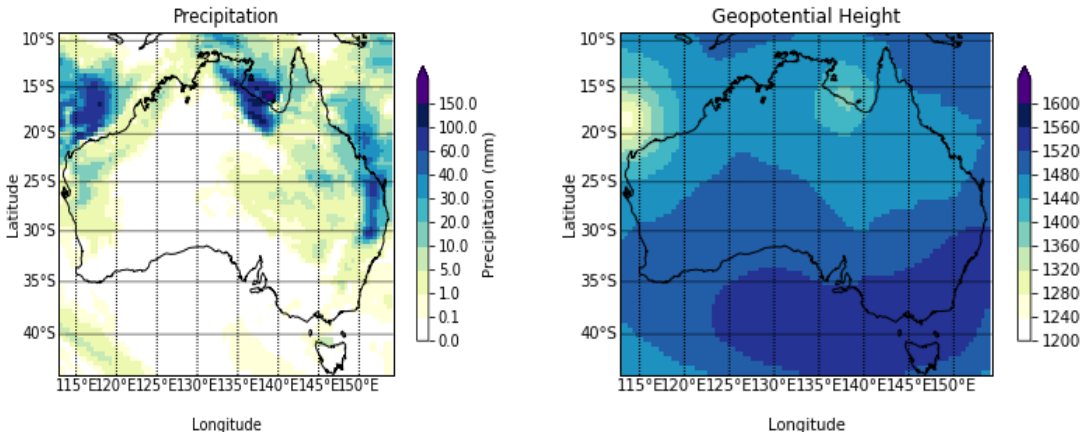

Figure 3.2: Demo of two channels of low-resolution input image, daily forecast of 25/01/2012

ensemble and leading time will be stacked together as two channels of the input image. The sample images of those two variables are shown in figure 3.2. Moreover, the image will be upsampled, which means that they are resized into the same size as the high-resolution image. In this way, the model can directly learn the pattern of the same size and position. Although the input image is upsampled and resized, it is still unclear and the forecast is not accurate, the model will try to enhance its image quality and forecast accuracy.

3.1.2 Output and Label

The output is high resolution precipitation forecast. Its resolution is approximately 12 km which is the same as label images. The label is the rainfall reanalysis data from BARRA dataset. Different from the input and output data, label does not have ensemble or leading time. There is only one image for each target date, which is the precipitation map of that day. Therefore, for each date, as shown in figure1.1, all ensemble members share the same label image, which means 11 input image will generate 11 high resolution output images, and the 11 output images will be compared to the same label. Moreover, for each forecast, the label is the precipitation of the target date, which means the forecast from different initialization date may also share the same label depending on their leading time.

3.1.3 Normalization

For traditional SISR, images have three channels red, blue and green, and the range of all channels are the same, 0-255, thereby the knowledge from each channel are similar and the computer vision model is easy to be trained. However, in climate forecast downscaling model, the first channel represents the predicted amount of rainfall and the second is the geopotential height. These two channels have significantly different range and average values. The range of precipitation is about 0-800

mm per day, while the range of geopotential height is around 1200-1600, so the average values of two channels are not balanced. In this case, the network tends to learn much more information from geopotential height than precipitation, thereby the useful knowledge from the variable with smaller range will be lost, which might have negative impact on the training and prediction performance.

Another significant problem is that the range from different date is not the same. For traditional images in SISR, the range is fixed 0-255 for all images, but the total amount of rainfall can vary significantly from date to date. For example, the maximum value of precipitation forecast of a day may be 300, but it could be 600 in other dates. Consequently, downscaling model cannot project data onto range 0-1, because it the maximum value may be unknown, and the true amount of rainfall will be lost. What is more important is that if all precipitation images are made into range 0-1, the average amount of rainfall of training images will be the same for all dates, which is not the real picture and important information will be lost.

Because the aim is to predict precipitation, and geopotential height data is used to improve the accuracy of rainfall prediction, the range and mean values of precipitation channel is much more important than that of geopotential height channel. Therefore, to balance and normalize the two channels, the geopotential height channel is projected on to the range of precipitation, and precipitation channel uses the original forecast values from ACCESS-S1 dataset. The formulation of channel normalization is shown as follow.

$$gh^* = max(PR) * \frac{gh}{max(GH) - min(GH)} \tag{3.2}$$

Where gh* and gh represents the new value and original value of each pixel in geopotential height channel, PR and GH is the whole precipitation channel, and original geopotential height channel. In this way, both channels share the same range and may have similar contribution to rainfall prediction.

3.2 Deep Neural Network Structure

The overall structure of multiple-input VDSR downscaling model is shown in feature 3.3. This model takes two upsampled low resolution images as input, which includes precipitation and geopotential height daily forecast. For each 2-channel input, it predicts a residual image, which is the difference between low-resolution rainfall forecast and label images, and finally perform element-wise addition to add the upsampled low resolution precipitation image onto the residual image to generate a predicted super resolution image. Two channels are used for training and feature extraction, but only precipitation forecast is used for final addition. In this way, the model does not need to predict the whole rainfall forecast but predict the difference between low-resolution forecast and the ground truth value. The high-

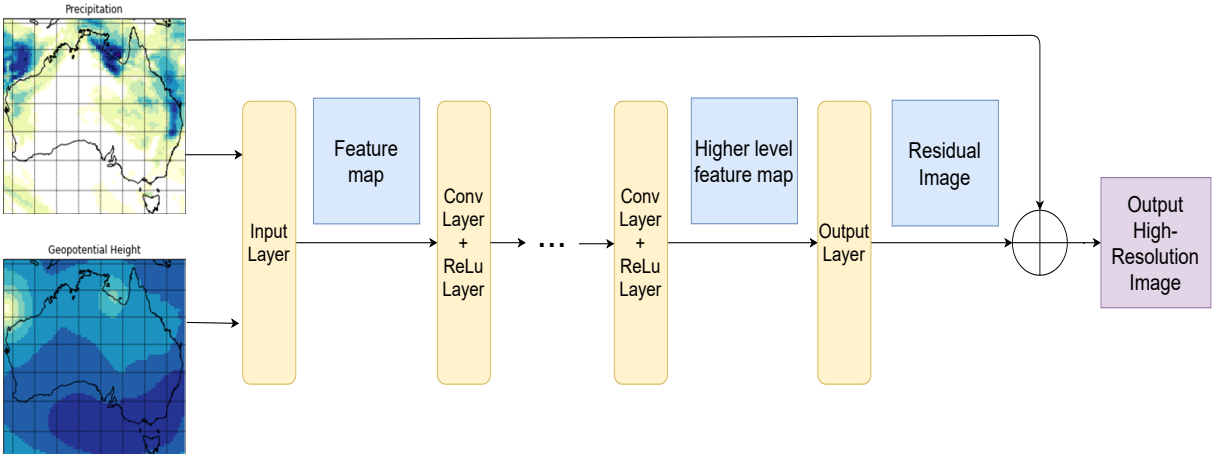

Figure 3.3: The structure of VDSR downscaling model, where $\oplus$ represents element-wise matrix addition, orange blocks are layers of neural network, blue rectangles are intermediate images, input and output images are on left and right hand side.

level formulation can be written as

$$SR = PR_{lr} + F(PR_{lr}, ZG_{lr}) \tag{3.3}$$

where SR is output high resolution precipitation forecast, F represent all operations in the whole neural network, PR_{lr} and ZG_{lr} are upsampled low-resolution forecast of precipitation and geopotential height.

My model has a new adopted version of residual learning. The traditional residual learning model adds residual to the input image to get the result. However, the input for my rainfall forecast downscaling model contain both rainfall and geopotential height images, thereby it would be meaningless to add both input images to the residual. Instead, although both channels are used to extract features, only precipitation input image is added to the residual image. In this way, residual learning can improve the performance, and features from geopotential height data can improve precipitation prediction.

As shown in feature 3.3, the model mainly has three part: input layer, intermediate feature extraction layer and output layer. The input layer takes two-channel input image extract features with 64 filters. Then the intermediate layers generate higher level feature maps. Finally, the output layer will convert the feature map into a residual image, which is the predicted difference between input and target image. The residual image will be added to the input precipitation image to output a super-resolution rainfall forecast. All procedures and intermediate outputs are shown in figure 3.4. Next, these three parts will be introduced one by one and the methods of feature extraction and image reconstruction will be discussed.

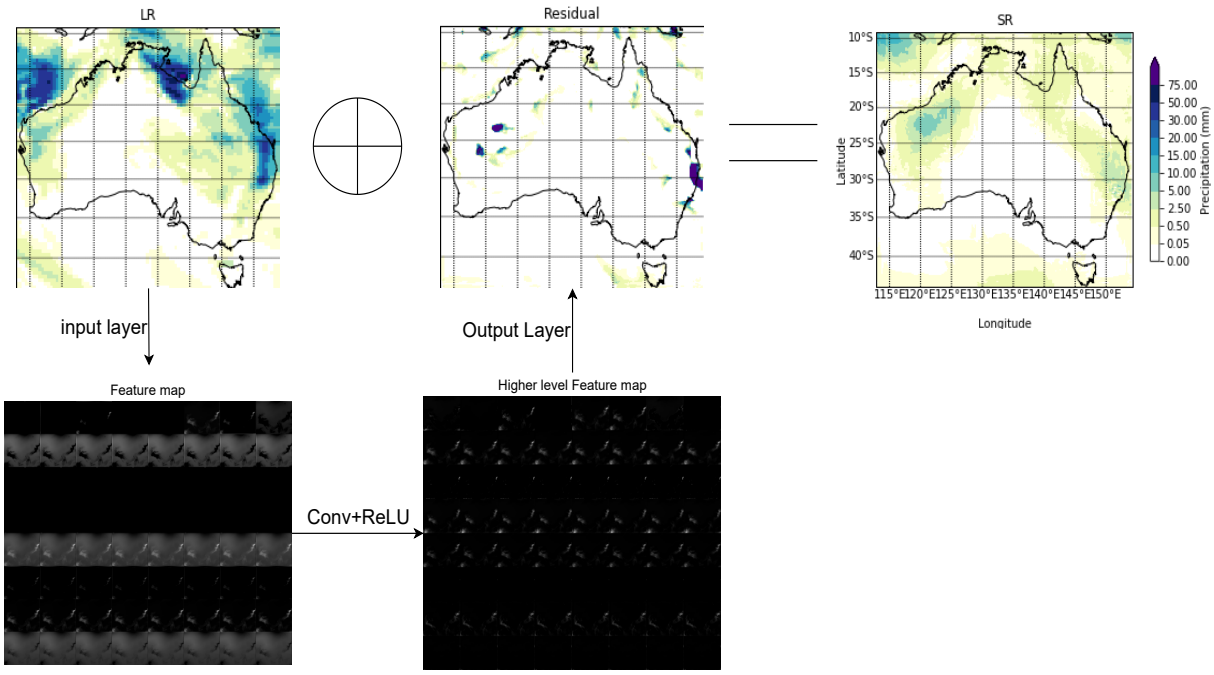

Figure 3.4: Visualization of input, intermediate outputs and output. The first image is input image, then 64 feature maps are generated, and then Convolutional and ReLU layers extract 64 higher level feature maps, and finally the residual image is constructed, which is added to the input precipitation forecast to get a super resolution forecast. Values of feature maps are rearranged to (0,255) for visualization.

3.2.1 Input Layer

As shown in figure 3.3, the 2-channel input image will firstly go through an input layer. This layer has a convolutional layer and a ReLU layer. The convolutional layer has 64 kernels, and will produce 64 first level feature maps. Then the ReLU layer will perform ReLU function to remove the meaningless negative values from the feature maps. The operation can be formulated as

$$FM_0 = F(PR_{lr}, ZG_{lr}) = ReLU(Conv(PR_{lr}, ZG_{lr})) \tag{3.4}$$

where FM_0 is the first level feature maps generated by input layer, and ReLU() and Conv() are ReLU and convolutional layer that perform the ReLU function 2.7 and 2-dimensional convolution 2.6.

The kernel size is set to be 3x3. Padding and the step length are 1. Therefore, the size of each feature map is the same of the size of high-resolution image. Suppose the size of input image is (2*m*n), then the size of the feature map generated by input layer will be (64*m*n). (In this project, the size of training image is 316*376.) It will convert the two-channel image into a low level feature map as the second image shown in figure 3.4. Most of the values in the feature maps are 0 after ReLU layer, and some maps are similar.

This input layer can extract basic features from the 2-channel input and generate 64 feature maps. It also allows the following intermediate layers to extract higher level features.

3.2.2 Intermediate Blocks

There are totally 18 intermediate blocks between the input and output layers. All of them are identical and each of them consists of a convolutional layer, which extract deeper features, and a ReLU layer, which removes the negative values. Each convolutional layer has 64 kernels to produce 64 feature maps and takes 64 feature maps from previous block as input. Therefore, the operation of each intermediate block is the same, which can be written as

$$FM_n = B(FM_{n-1}) = ReLU(Conv(FM_{n-1})) \tag{3.5}$$

where FM_n represents the nth level feature map, B is the operation of intermediate block. Therefore, the result of operation of all intermediate blocks can be written as

$$FM_{18} = B^{18}(FM_0) \tag{3.6}$$

where

$$B^n(x) = \begin{cases} B(B^{n-1}(x)) & , n > 1 \\ B(FM_0) & , n = 1 \end{cases} \tag{3.7}$$

Table 3.1: Training settings

Epoch	Learning rate	Bacth size	Loss function	Optimization	momentum
50	0.0001	4-32	L1	SGD	0.9

and FM_0 represents the first level feature map generated by the input layer.

These intermediate blocks can extract higher level features. The large numbers of layers and kernels allow the model to learn complex patterns. These complex patterns may improve the performance of SISR and precipitation forecast. An eighteenth level feature map is shown in the figure3.4, and this will be used as input of output layer to generate the residual image.

3.2.3 Output Layer

Output layer is a convolutional layer that converts 64 high level feature maps into a residual image as shown in figure 3.4. That is to use discovered complex patterns to predict the difference between unsampled low-resolution forecast and the target image. Finally, the residual image will be added to the upsampled precipitation input image to predict a super resolution precipitation forecast.

3.3 Training Details

Table 3.1 lists some hyper-parameters of training. The number of epochs is 50, which is approximately the time of convergence. Learning rate is small because the network is very deep, large learning rate may cause vanishing/exploding gradients problem [Bengio et al., 1994]. Optimization method is Stochastic Gradient Descent (SGD) with 0.9 momentum. Loss function is L1Loss which is defined as

$$L_1 = mean(||SR - HR||) \tag{3.8}$$

where SR is the predicted super resolution image and HR is the ground truth high-resolution image. The L1Loss measures the mean absolute difference between values of all pixels.

3.4 Summary

Figure 3.3 shows the overall structure of the VDSR downscaling 2-channel model with geopotential height. As shown in equation 3.3, the input is upsampled 2-channel image. Both channels are used for feature extraction but only precipitation channel is added to predicted residual image to output a super resolution precipitation forecast.

To predict the residual image, the neural network has three parts, including input layers, intermediate blocks and output layer. The input layer converts 2-channel input image into 64 level feature maps. Then intermediate blocks use previous feature map as input to extract higher level feature maps. All 18 blocks are identical, and each block has 64 kernels of size 3*3. The final feature map is sent to the output layer, and the output layer use convolution to generate a residual image that predicts the difference between input and label image. All convolutional layers in the network share the same kernel size of 3*3 and same padding of 1 and slide length of 1.

The overall formulation of the VDSR 2 channel downscaling model can be written as

$$SR = PR_{LR} + Conv(B^{18}(ReLU(Conv(PR_{LR}, ZG_{LR})))) \tag{3.9}$$

where SR is the output super resolution image, B^{18} is a composition function defined as equation 3.7, PR_{LR} and ZG_{LR} are precipitation and geopotential height forecast from input image, Conv and ReLU are convolutional and ReLU layers.

The training setting is listed in table 3.1. Training data is the data of leading time 1-7. One year of data was left for validation, and all other years' data was used for training. After designing and training the model, validation is necessary to evaluate its performance.

Chapter 4

Experimental Methodology

In this chapter, I will introduce the platforms of experiment, including software and hardware platforms. Moreover, experimental methodology such as cross-validation and evaluation methods like CRPS and T-test will be discussed.

Section 4.1 and section 4.2 will list the software and hardware platforms of training and testing the model.

Section 4.3 will introduce the data used for training and testing, and leave-one-year-out cross-validation method. 1997, 2010 and 2012 are selected for validation. The reason why these three years are selected is provided.

Section4.4 will discuss the evaluation metrics for image quality and ensemble forecast accuracy. It also includes statistic analysis methods like T-test.

4.1 Software platform

Software flatform is Python 3.6. Important packages is shown in the table 4.1. The model was trained and evaluated based on Pytorch package. All climate data, including input forecast and label, is in .nc format, thereby netNDF4 package was used to read climate data. All codes are in python script files. Other packages like pandas were used for visualization and demonstration.

Package name	Version	Description
numpy	1.18.4	Data storage and processing
netCDF4	1.5.3	Read .nc files
torch	1.5.1	Neural network training
tqdm	4.48.2	Iteration

Table 4.1: Software environment in our evaluation.

Table 4.2: Hardware details of Gadi from NCI

CPU	Intel(R) Xeon(R)Cascade Lake Platinum 826
Number of CPU	36
CPU clock rate	2.9GHz
CPU cache	35.75MB
GPU	Nvidia(R) Tesla Volta(TM) V10
Number of GPU	3
GPU memory	32GB

4.2 Hardware platform

All the code was running on the newest super computer called Gadi from National Computational Infrastructure (NCI). The computation resource used for this project is list in the table 4.2

4.3 Leave-one-year-out Cross-validation

The time range of climate forecast in ACCESS-S1 dataset is from 1990 to 2012. Leave-one-year-out cross-validation was used to evaluate the performance of the downscaling model. Data from one year was used for testing and all the other data was used for training. In this way, forecast of all different seasons and target dates could be used for testing.

The first year used for validation is 2012, which is the most recent year in the dataset. Therefore, precipitation and climate situations of 2012 may be the most similar to the climate nowadays, thereby the performance in 2012 may be similar to its actual performance in predicting the rainfall nowadays. For 2012 validation, the model was trained by data from 1990 to 2011 and evaluated by data of 2012.

Besides 2012, 1997 and 2010 were also left for cross-validation. This is because 2010 and 1997 are strong El Niño and La Niña periods in Australia, which means that the precipitation in Australia was extreme and unusual, and it should be hard to predict the amount of rainfall [aus; Philander, 1985]. Therefore, cross-validation on these two years can test the performance under unusual circumstance, which is important for weather perdition because great economic values can be made if the extreme weather events can be predicted.

For 1997 validation, data from 1997 was used for test, and the training data was from 1990-1-1 to 1996-12-31, and from 1998-8-1 to 2012-12-31. The first seven months of 1998 was not included in training set because the leading time is 0-217, and as shown in figure 4.1, if the initialization date is the last date in 1997, the target date can be in first 7 months of the next year. Therefore, about one and a half years data will be used in test so it should not be included in the training set. This is the same

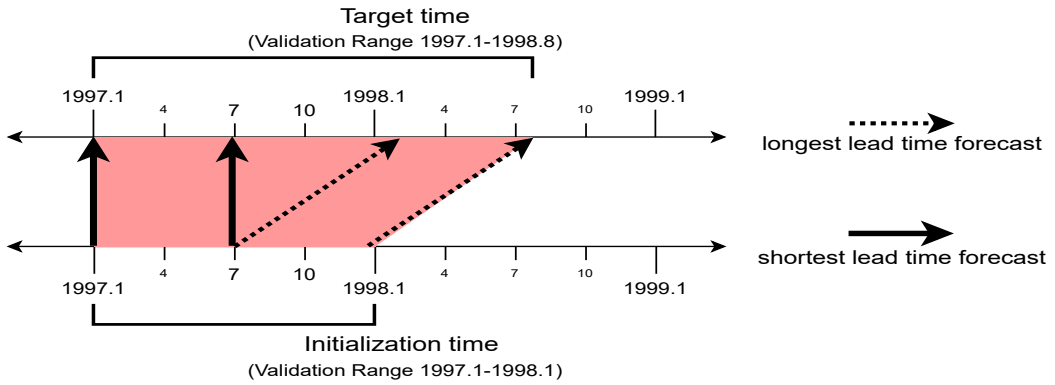

Figure 4.1: The initialization and target time of validation data for 1997

for 2010 validation, where the data from first 7 months in 2011 was not used for training.

Training data use leading time 1-7, because forecast from this leading time range has relatively the best skill.

4.4 Evaluation

The model of this project is a weather forecast downscaling model based on SISR. The objective is to enhance the resolution of the precipitation forecast and improve the accuracy over all ensemble members. Therefore, the evaluation contains two parts: the image quality, which compares each output to its label, and forecast accuracy, which evaluates the overall accuracy of all generated super resolution ensemble members. For seasonal rainfall forecast downscaling, the ensemble forecast accuracy, which is evaluated by CRPS, is the most important measurement.

4.4.1 Image Quality

For SISR model, the output super resolution image should be as similar as possible to the label image. Its performance is good when the difference between the pair of images is small.

Peak signal to noise radio (PSNR) is an important and widely used measurement of the difference between images. It shows a radio of the maximum value against the noise in the image. If the radio is large, then the image quality is better. The formulation of PSNR can be written as

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left(\frac{MAX^2}{MSE} \right) \\ &= 20 \cdot \log_{10} \left(\frac{MAX}{\sqrt{MSE}} \right) \end{aligned} \tag{4.1}$$

where the MAX is the maximum possible value in the image and MSE stands for

mean squared error of two-dimension images defined as

$$MSE = \frac{1}{mn}\sum_{i}^{m}\sum_{j}^{n}[SR(i,j) - HR(i,j)]^2 \tag{4.2}$$

For traditional RGB images, the maximum value could be 255, which is a known and fixed value. However, for rainfall forecast image, the maximum value is unknown, or it is infinity. To address this problem, an approximate value 1000 is used as the maximum value of daily precipitation(mm). This is based on the observation that the maximum value over last 23 years is no more than 900.

4.4.2 Ensemble Forecast Accuracy

The accuracy of ensemble forecast can be evaluated by the continuous ranked probability score (CRPS) [Grimit et al., 2006]. Base on the equation2.1, each data point can have a CRPS score on each initialization date and each leading time, and the mean CRPS score of forecast map of a date t can be formulated as

$$CRPS(t) = \frac{1}{mn}\sum_{m}\sum_{n}\int_{-\infty}^{\infty}\left[F\left(SR_{t,m,n}\right) - \mathbf{1}\left(SR^{g}_{t,m,n} \geq HR^{g}_{t,m,n}\right)\right]^2 d(SR_{t,m,n}) \tag{4.3}$$

where m and n is the size of image, t contains leading time and initialization time. To evaluate the CRPS, the mean value over all date in the year is used, so a means CRPS score is computed for each leading time.

Climatology model is used as the reference forecast. According to equation 2.3, the CRPS skill score of each pixel is computed by

$$CRPSSS = 1 - \frac{CRPS}{CRPS_{clim}} \tag{4.4}$$

where CRPS and $CRPS_{clim}$ are the element wise CRPS scores of my deep learning based downscaling model and climatology model, and all CRPS scores have a corresponding leading time.

The model with higher CRPS score tends to have better accuracy. If the skill score is positive on average, then the new model is better than climitology model. If the reference model is the same, the model with higher mean skill score tends to be better.

Because sometimes the CRPS and skill scores of different models may be similar, T-test was also used to determine whether there are significant difference between model performance.

Chapter 5

Results

This chapter contains the test results, including ensemble forecast accuracy, image quality and time efficiency. The results come from validation of three years, and the most detailed evaluation is done for 2012 forecast, because it is the most recent data in ACCESS-S1 dataset. All the results from my model is compared to other climate downscaling model such as climatology, QM and VDSRd, and the performance of my model is analysed by statistic methods like T-test. Some typical results are visualized.

Section 5.1 is about ensemble forecast accuracy. The most significant measurement is the CRPS skill score, which shows the accuracy of super resolution ensemble forecast in comparison to climatology model. Results includes the CRPS of whole Australia and 50 rainfall stations [Li and Jin].

Section 5.2 evaluates the similarity between label image and output image. PSNR and typical output images are computed and demonstrated to show the quality of single output image.

Section 5.3 is contains time efficiency and the change of training loss. Time efficiency includes training time and predicting time.

Section 5.4 includes CRPS results of using other climate variables like temperature. Those combinations were attempted but unsuccessful, because the forecast accuracy is worse than two-input model that uses only geopotential height and precipitation.

5.1 Ensemble Forecast Accuracy

The accuracy of ensemble forecast is evaluated by CRPS and skill score. CRPS of the whole Australia and 50 rainfall stations are computed. The region of the whole Australia is the same as training and label images. 50 stations are 50 representative rainfall stations from different climate regions in Australia. The positions of 50 stations are shown in the figure 1 and their details are listed in the table 1. The CRPS skill score excludes prediction on the ocean, because the rainfall on the ocean is dif-

ficult to predict and not included in QM prediction.

This section contains result from 2012 and El Niño and La Niña periods. The results are compared with other models such as QM and VDSRd. T-test is used to compare and analyse the performance form different models. Some results are visualized on Australian map to analyse the performance of different region.

5.1.1 Precipitation Forecast for 2012

5.1.1.1 Whole Australia forecast

The confidence band graph shows the confidence of different downscaling models in comparison to the reference climatology model. As shown in the figure 5.6, for validation on 2012, the overall skill of my model is overall slightly better than climatology model, especially when leading time is small. On the other hand, other traditional postprocess methods like interpolation and quantile mapping are worse than the reference, they can have positive skill scores only in first several days. The median value lines from these models are compared in figure 5.2. Deep learning base models have better skill than traditional methods. The single input VDSRd model has greater skill for small lead time forecasts, but my model overall has a better skill, especially when the leading time is greater than 7. Moreover, my downscaling model is more stable in predicting 2012 precipitation, because the size of shades, which represent the 50% and 90% confidence intervals are smaller than that of other models.

The visualization of average CRPS and skill score values are shown in figure 5.3. These maps demonstrate the downscaling performance of different locations. The rainfall on the surrounding ocean is extremely hard to predict for postprocess models, thereby the CRPS is high and skill score is negative on average. The CRPS score is high on the east and north coast. The possible reason is that the amount of rainfall on ocean is large and hard to predict. For CRPS skill score, my model outperforms climatology in southern and northeast parts of Australia but has worse skill on the northwest and Tasmania. The performance of predicting southeastern area is similar. Overall, my prediction model is more accurate than the reference in mainland Australia.

5.1.1.2 50 Stations result

Figure 5.4 demonstrates the mean CRPS skill scores of 2012 precipitation prediction on 50 rainfall stations. The performance of my model is better than the reference climatology model, as the mean skill score of my model is 0.011, which means that my model has about 1% improvement against climatology model for 2012 50 stations rainfall forecast. All other downscaling models has negative mean skill scores. Quantile mapping and VDSRd have good skill in first seven days but only my model can remain positive afterwards.

Figure visualizes the mean 50 stations CRPS skill score result. Most stations have

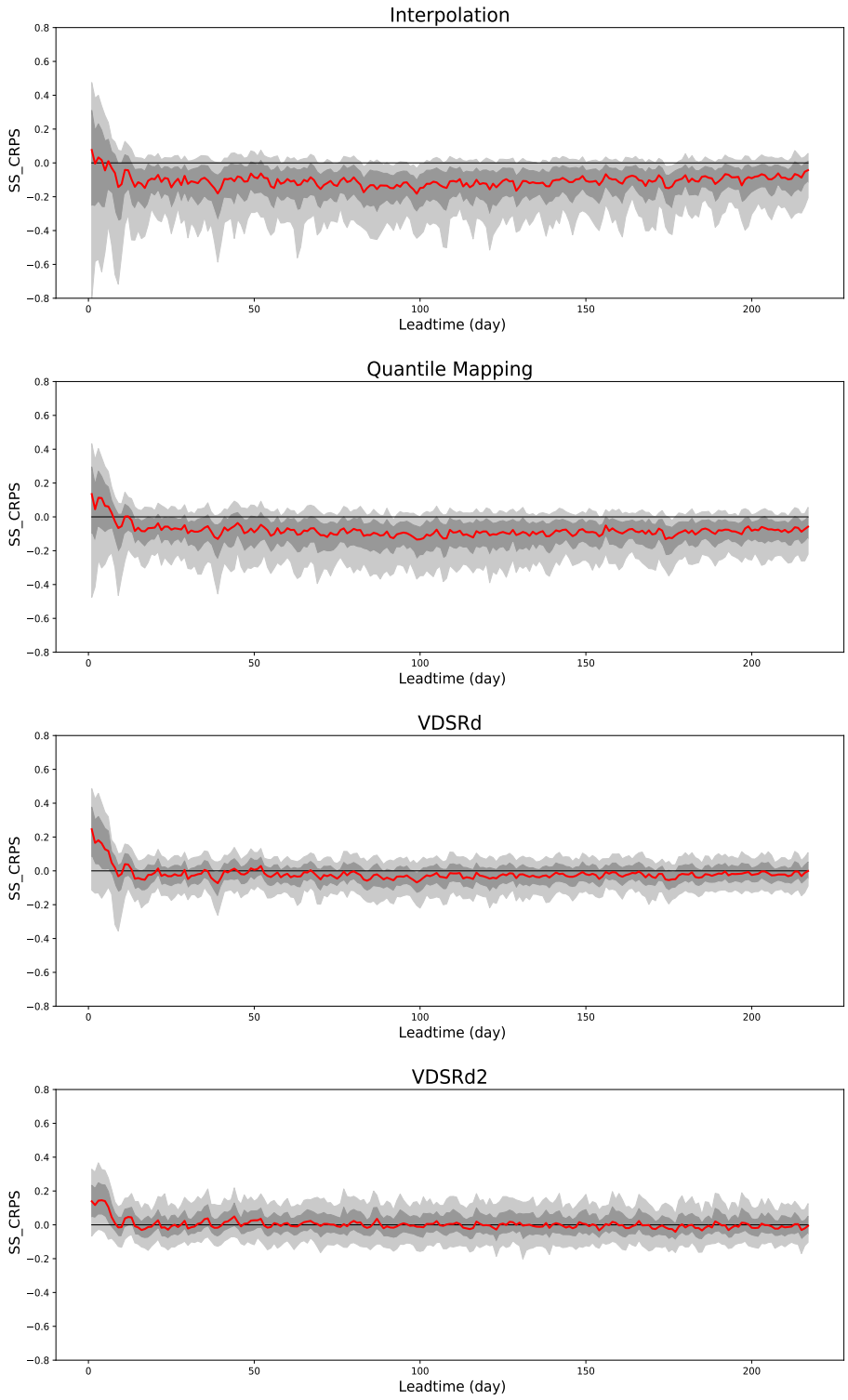

Figure 5.1: CRPS skill score of interpolation, QM, VDSRd and VDSRd2 models in 2012 precipitation forecast of all Australia. The light grey shade indicates the 90% confidence interval of CRPS SS, the dark grey shade indicates the 50% confidence interval, and the red line indicates the median

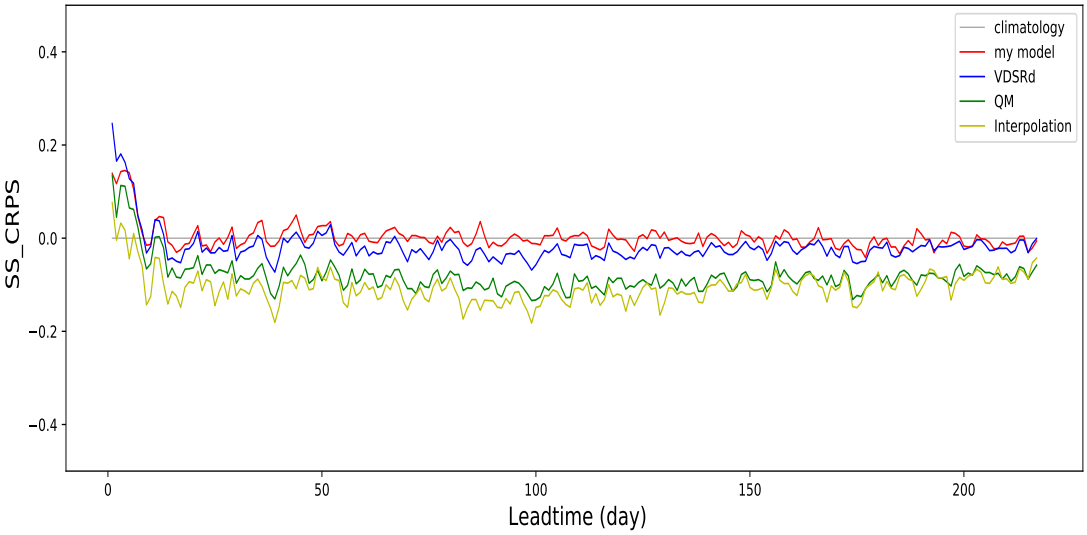

Figure 5.2: Median values of CRPS skill scores from my model, VDSRd, QM and interpolation on 2012 precipitation forecast.

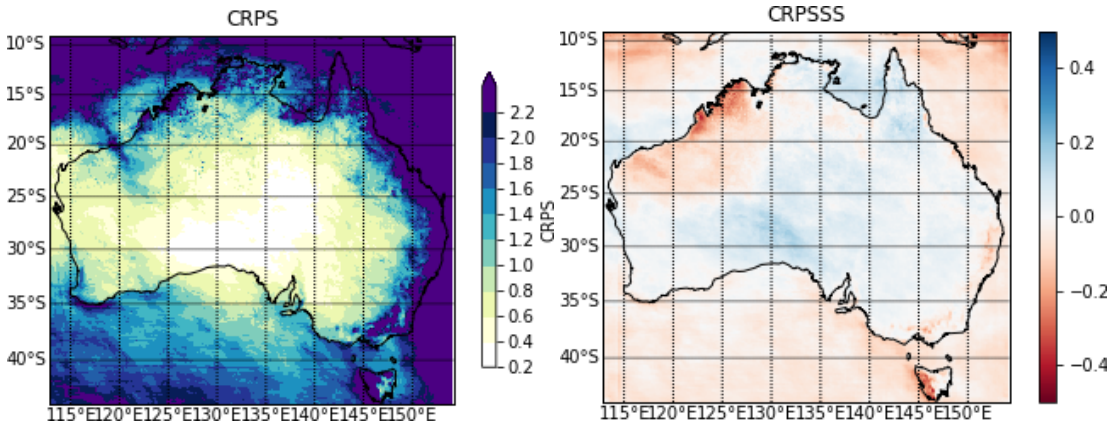

Figure 5.3: Average CRPS and skill score map of all leading time and initialization time for 2012 precipitation forecast downscaling

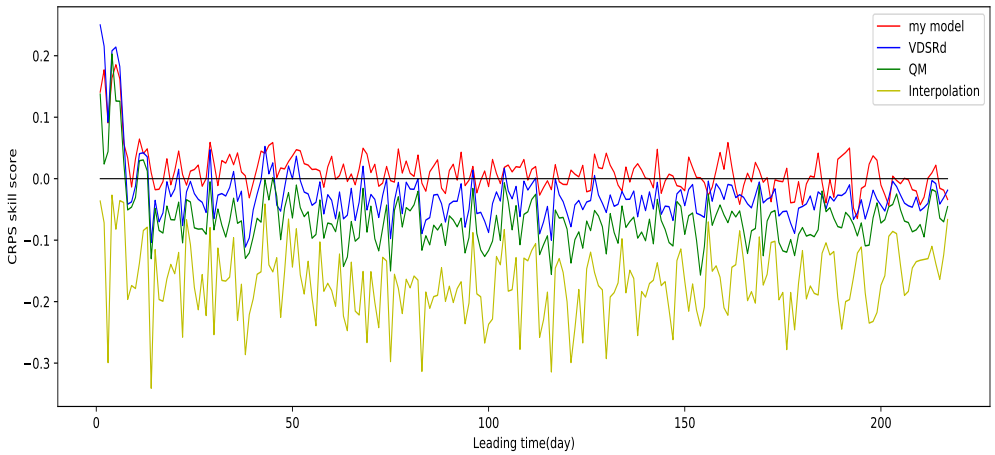

Figure 5.4: Mean CRPS skill score for 217 leading time of 2012 precipitation forecast on 50 stations.

positive skill values, which are shown by blue scatter points. Most negative results are greater than -0.02 and there are only one red points, which means they are only slightly worse than the reference model climatology. On the other hand, there are many deep blue points where the performance of my model is strongly recommended.
The stations near the coast tend to have negative mean skill score, while station that are far from the ocean have good result on average. The possible reason is that precipitation near the ocean changes dramatically and hard to predict.

5.1.2 Cross-Validation on El Niño and La Niña periods

The CRPS skill score is shown in figure 5.6. The overall skill of VDSRd2 is slightly worse than climatology model. The first reason is that these two years are strong El Niño and La Niña periods, thereby the rainfall is extreme and hard to predict compared to 2012. The second reason is that climatology method totally uses 242 ensemble members, and large ensemble size will lead to better CRPS score. Therefore, although the CRPS skill score is negative on average, my model is still promising because it uses less ensemble members and time. In comparison to other downscaling models like VDSRd, my model still has great advantages overall.

5.1.3 T-test

The difference of CRPS score is very small, especially for 2010 and 1997 prediction. T-test is used to determine whether there is significant difference between model

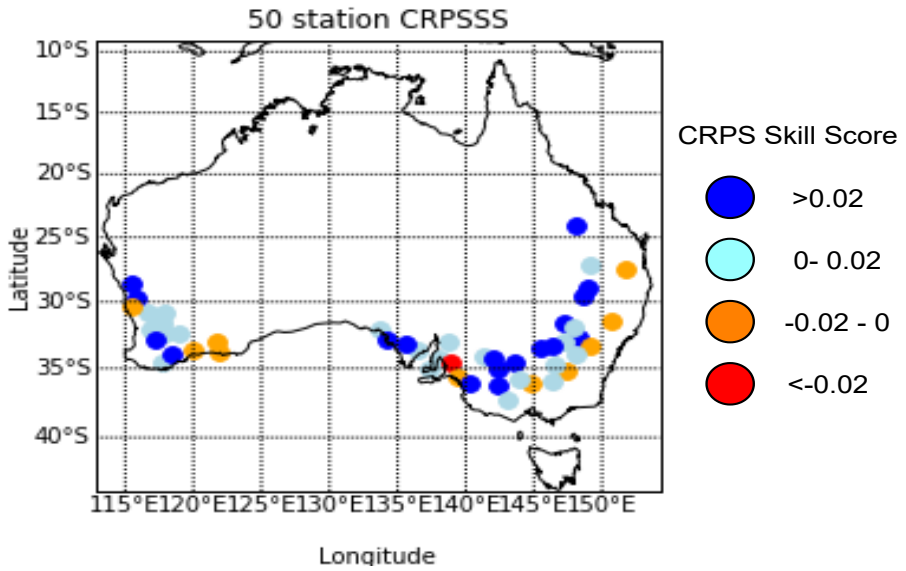

Figure 5.5: Mean CRPS skill score of 2012 precipitation forecast on 50 stations on the Australia map.

performance. I performed one-side paired t-test by assuming that my model would be better than previous VDSRd model. For 2012 forecast, the p-value of the one-side paired t-test of 76 pairs of CRPS values is 0.017, and the p-value of for 2010 prediction is smaller than 0.0001, which indicates that my model generates statistically significantly better result than VDSRd at the significant level 0.05.

5.2 Image Quality

The quality of super resolution output image can be visualized and evaluated by PSNR.

In traditional SISR, input and label are generally the images of the same thing, thereby the position and shape in the image are almost the same. However, for SISR based downscaling model, there are sometimes significant difference between forecast and label, and the model should not strictly follow the pattern from the input, but learn to predict the target precipitation. This increases the difficulty, and there might be difference between input and output.

5.2.1 Visualization

Typical result of three ensembles are shown in figure 5.7. These three inputs has the same target date 2012-1-7, and their leading time are all 3, thereby they share the

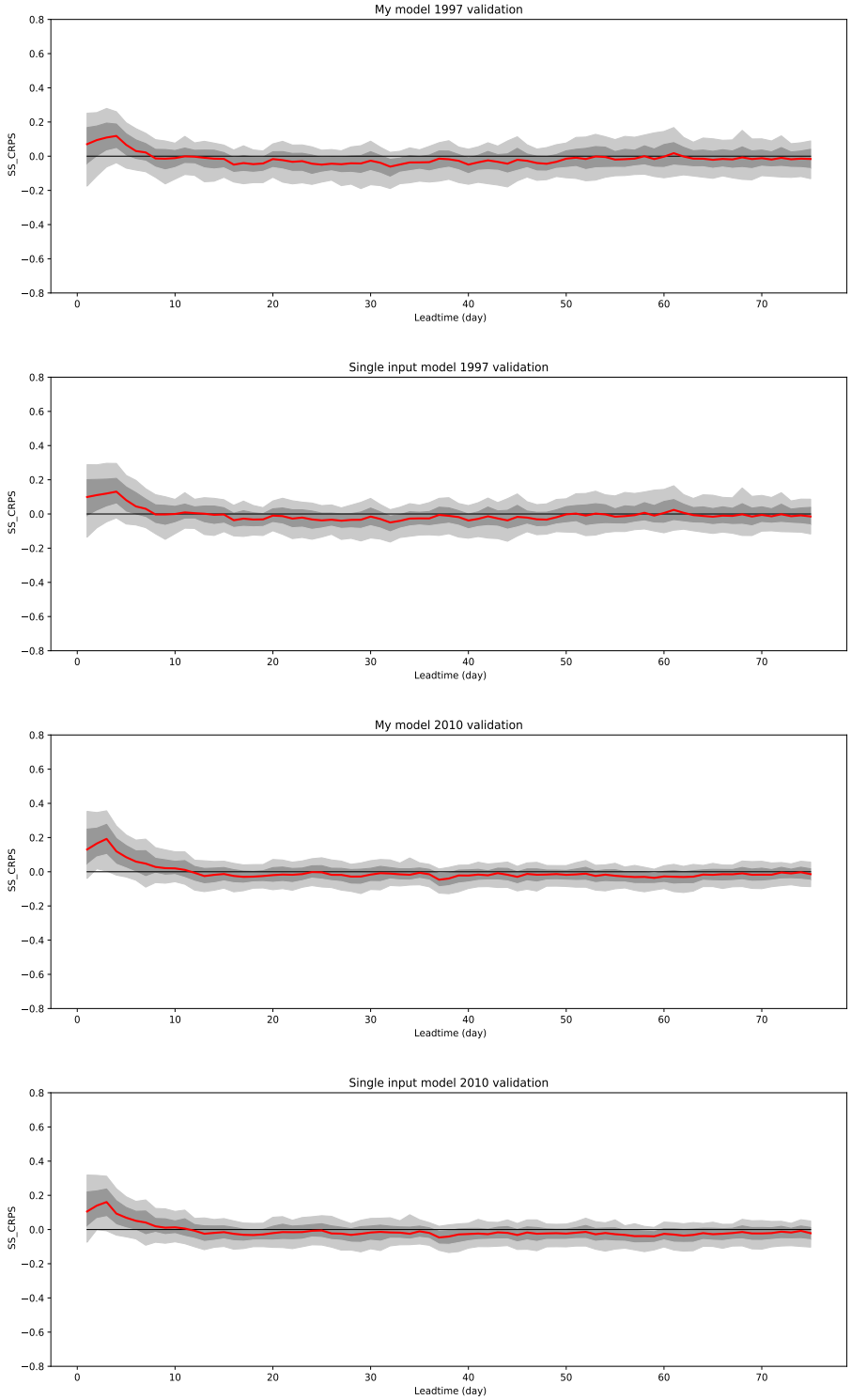

Figure 5.6: CRPS skill score of validation on 1997 (first two graphs) and 2010 (the last two graphs). The first and the third are the results of my model, and the other two are results of VDSRd

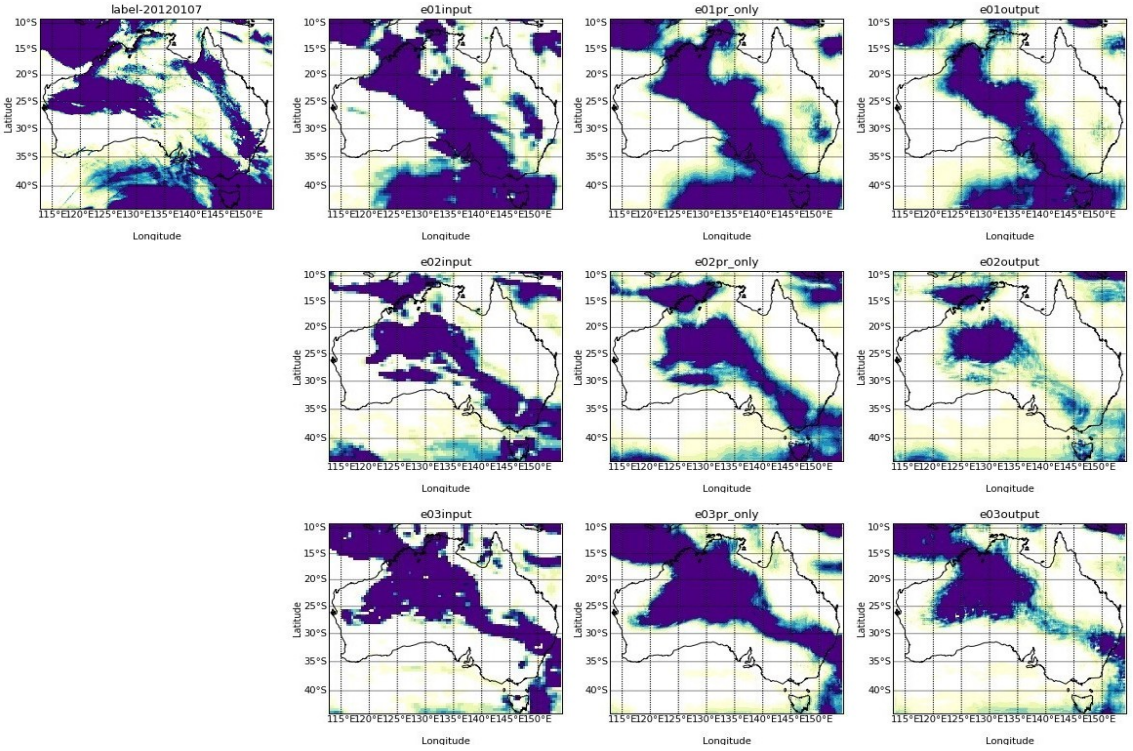

Figure 5.7: Demonstration of typical result of forecast on 2012-01-07. Images from the first to the last column are the label, three ensemble inputs, output of single input downscaling model [Jiang, 2020], and output from my model

same label.

Compared to the input and single input downscaling model, my model can adjust the rainfall prediction and make it closer to the label image. For example, there are little precipitation on the 30-degree latitude line, and my model predicted it successfully and reduced the amount of rainfall at that area for all ensembles.

The drawbacks of my model from this forecast is that my model may tend to reduce the amount of rainfall, thereby the predicted precipitation could be less than the actual.

5.2.2 PSNR

PSNR can measure the similarity between the label and output image. The model with higher PSNR tends to have better SISR performance. There are 11 ensemble members and 48 initialization dates, thereby there are totally 48*11=528 pairs of images per leading time. Mean values of all dates and ensembles are used to show the overall quality of predicting for each leading time.

The PSNR scores from three models are shown in figure 5.8. Deep learning based

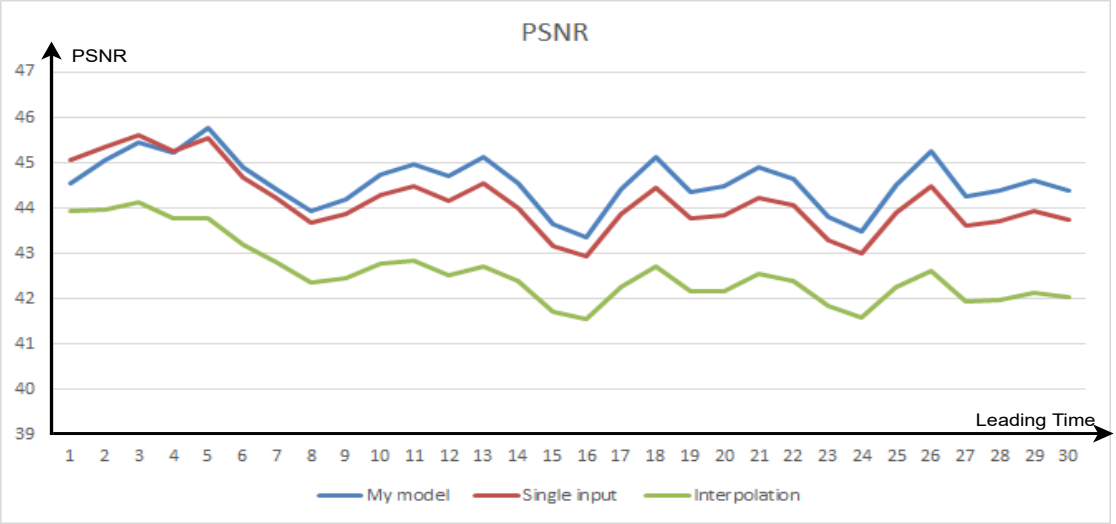

Figure 5.8: PSNR results of different leading time forecast downscaling from my model, VDSRd and interpolation.

model has significantly better result compared to the interpolation model. Compared to the previous model which has only one precipitation input, my model with geopotential height has better performance when the leading time is greater than 4, which shows that geopotential height data have positive impact on Australian precipitation forecast downscaling.

The PSNR of precipitation images is different from that of general images. Firstly, the absolute value of PSNR is significantly influenced by the quality of input forecast image. In other words, if the input climate forecast image has bad skill and accuracy, the PSNR score will be low at that day. Therefore, the PSNR result can only be used for downscaling model comparison. Moreover, the PSNR values of precipitation forecast images may be higher than that of normal images because the possible maximum value is higher. Another possible reason is that precipitation is a one-channel gray-scale image with many 0 values that are easy to predict, because some positions, such as in the desert, seldom have rainfall.

5.3 Training Loss and Time Cost

The decrease of training loss during training is shown in the figure 5.9. The model almost converge after 50 epochs and the training loss totally reduced by approximately 15%.

Training VDSRd2 model needs 50 epochs, and each epoch will need 1000 seconds, thereby the total training time is approximately 13 hours. Predicting a super resolution climate image needs about 2 seconds per images, and totally 0.13 hour

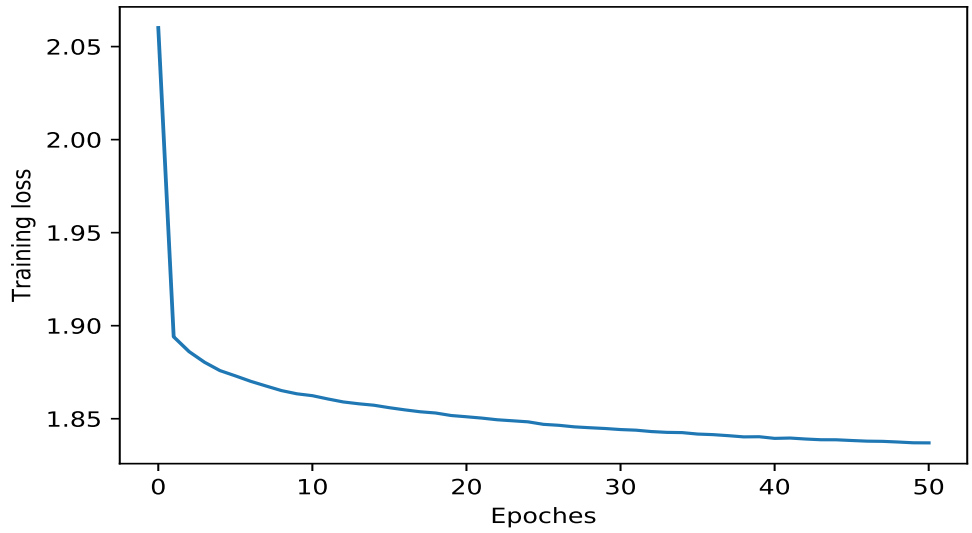

Figure 5.9: Training loss.

is needed for one-year downscaling, which is faster than traditional methods like ECPP and QM according to table 4.1, but slower than VDSRd because more weather variables are used in my model.

5.4 Result of Using Other Climate Variables

Besides precipitation and geopotential height, I also attempted to use other climate variables in my multiple-input downscaling model. However, those attempts were not successful because those new variables can bring no improvement in CRPS score evaluation. An example is a model with temperature as the third channel of input images, and its result is shown in figure 5.10. Compared to the two-channel model that uses geopotential height and precipitation, adding the third channel temperature will slightly reduce the skill, especially when the leading time is small. The possible reason may be that there is no clear relationship between temperature and precipitation. The data contains all four seasons, for example, though summer tend to have more rainfall than winter, a day in summer can have very high temperature but no rainfall, while a day in winter may have lower temperature but higher rainfall. Therefore, the final model contains only two input weather features, which are geopotential height and precipitation.

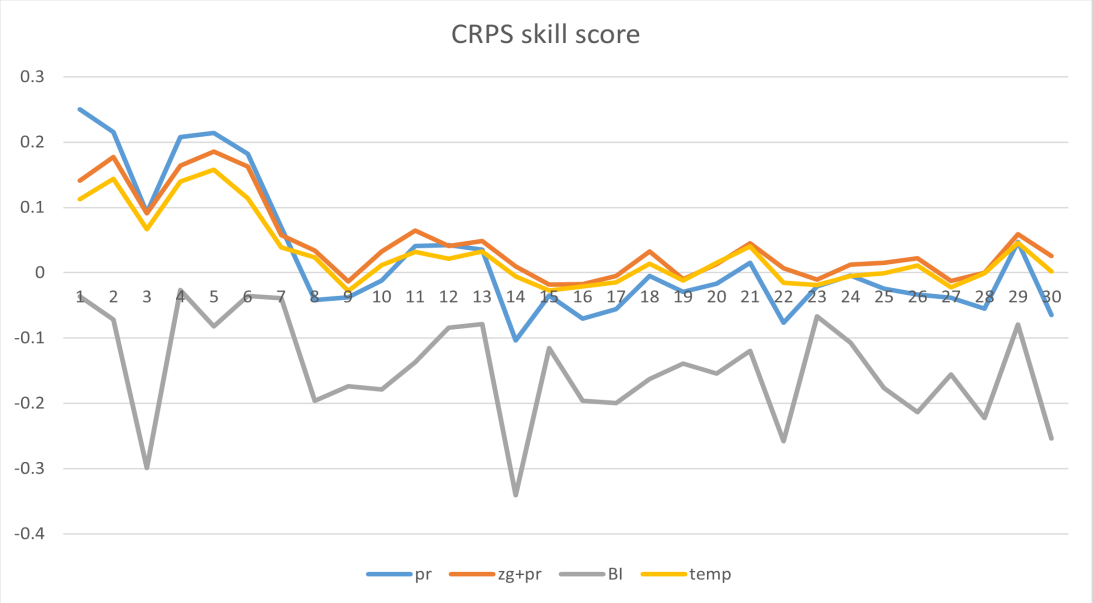

Figure 5.10: CRPS skill score of using temperature (shown as eh yellow line), in comparison with my final model with geopotential height (zg, red line), VDSRd (pr, blue line), and interpolation (BI, gray line)

Chapter 6

Conclusion

In this project, a new climate downscaling model VDSRd2 is design for Australian precipitation forecast downscaling. This model combined the previous VDSRd model and multiple-input downscaling model. A adapted residual learning structure was proposed to uses both rainfall and geopotential height as input to improve the skill and generates more accurate high resolution precipitation forecast. The result shows that geopotential height data have strong skill in improving the precipitation prediction, especially when the leading time is large.

This project shows that SISR techniques can have great performance on Australian seasonal climate ensemble forecast downscaling. This model uses SISR and residual learning to learn the difference between low-resolution and high-resolution image. It not only enhances the resolution of climate images but reduces the difference between input ensemble forecasts and the ground truth values. To improve the accuracy, features from geopotential height data is extracted to provides important knowledge for precipitation prediction. As a result, compared with previous downscaling model, VDSRd2 has great skill when the leading time is large, and have similar or slightly better result compared with climatology method that uses 242 ensemble members. This indicates the promising future of deep learning and super resolution based downscaling model.

6.1 Future Work

This section contains 3 possible future work, including ensemble prediction model, historical model, and generative model.

6.1.1 Ensemble Prediction Model

One possible future work is to develop an ensemble prediction model which uses all ensemble members as input and includes CRPS in loss function.

The aim of my project is to produce accurate high-resolution ensemble forecast, and the CRPS evaluation of my model is based on all eleven ensemble members rather than single forecast image. However, the technique I used was still based on SISR, where each input image of my model contains only one ensemble member. In

this way, the relationship between each ensemble cannot be learned, and the model can only generate high-resolution forecast based on one ensemble image.

A possible and promising idea is to include all ensemble members of the same date in each training image, and then the model uses all ensembles to predict a high resolution rainfall image. In this way, each input and label image will have one-to-one relationship rather than eleven to one. This has significant difference between a SISR model, and it can be called as an ensemble prediction model.

CRPS can evaluate the accuracy of ensemble forecast, and the major objective of the project is to have better CRPS. Therefore, if all eleven ensemble members are used in each input, the objective function of training the deep neural network can be changed to a adapted version of CRPS function. In this way, the model will be optimized to have a better CRPS score, which means a better ensemble forecast.

6.1.2 Historical Model

Climatology downscaling model usee historical data and has great performance in rainfall prediction. Therefore, we can combine deep learning based downscaling model with historical input data. However, a potential problem of this idea is that traditional climatology method uses 11 slide window and 242 ensemble members to get good result, but training deep learning model with such a large ensemble size may be difficult and computationally expensive.

6.1.3 Generative Model

Generative adversarial network (GAN) is one of the state-of-art techniques in SISR. It can produce high-resolution output image with great quality. Moreover, because precipitation is hard to predict, there are great difference between forecast image and ground truth image. GAN may be good at reducing the difference and improve the forecast accuracy of each ensemble output.

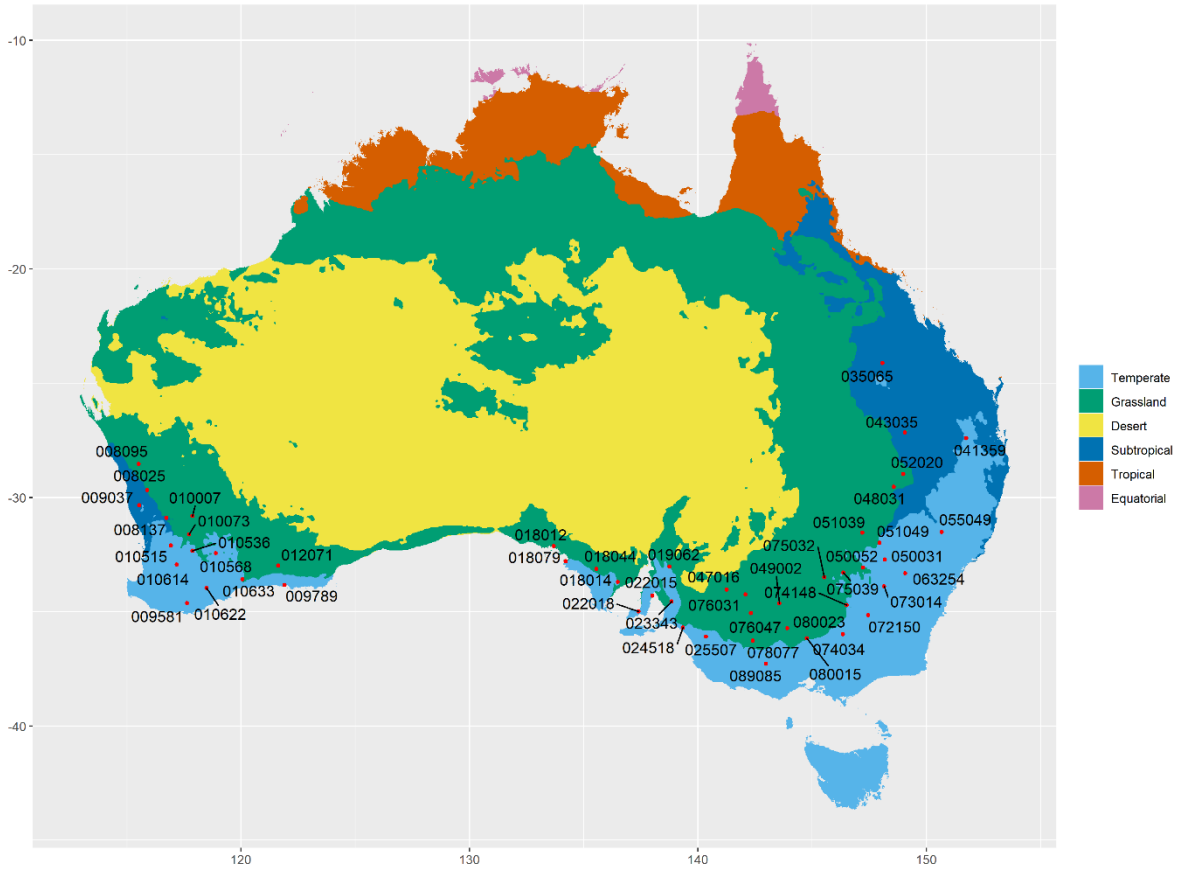

Figure 1: The locations of 50 rainfall stations in Australia.[Li and Jin]

Table 1: List of information of 50 rainfall stations.[Li and Jin]

	StationCode	StationName	State	Soil	Cultivar	Lat	Lon	climate_region	climate_charac	annual_rainfall	rainday_prop
1	008025	CARNAMAH	WA	Northern Tenosol	Endure	-29.6886	115.8872	Grassland	hot (summer drought)	337.793357933579	0.22259254850613
2	008095	MULLEWA	WA	Northern Tenosol	Derrimut	-28.5367	115.5142	Grassland	hot (summer drought)	332.831686703964	0.187120580883228
3	008137	WONGAN HILLS	WA	Northern Kandosol	Endure	-30.8917	116.7186	Subtropical	distinctly dry summer	388.057433638853	0.24163790024997
4	009037	BADGINGARRA RESEARCH STN	WA	Northern Tenosol	Correll	-30.3381	115.5394	Subtropical	distinctly dry summer	515.740090465421	0.290203547196762
5	009581	MOUNT BARKER	WA	Southern Sodosol	Bolac	-34.625	117.6361	Temperate	distinctly dry (and warm) summer	651.682061659326	0.472800857040828
6	009789	ESPERANCE	WA	Southern Sodosol	Endure	-33.83	121.8925	Temperate	distinctly dry (and warm) summer	609.242828234734	0.365908820378526
7	010007	BENCUBBIN	WA	Northern Kandosol	Wyalkatchem	-30.8081	117.8603	Grassland	hot (summer drought)	305.942209260802	0.242114034043566
8	010073	KELLERBERRIN	WA	Northern Chromosol	Wyalkatchem	-31.6183	117.7217	Grassland	warm (summer drought)	310.047970479705	0.207951434353053
9	010515	BEVERLEY	WA	Northern Chromosol	Endure	-32.1083	116.9247	Temperate	distinctly dry (and hot) summer	410.597845494584	0.25116057612189
10	010536	CORRIGIN	WA	Southern Sodosol	Wyalkatchem	-32.3292	117.8733	Temperate	distinctly dry (and hot) summer	352.530651112963	0.235091060588025
11	010568	HYDEN	WA	Southern Kandosol	Derrimut	-32.4419	118.8983	Temperate	distinctly dry (and hot) summer	365.230270205928	0.245327937150339
12	010614	NARROGIN	WA	Southern Chromosol	Endure	-32.9342	117.1797	Temperate	distinctly dry (and hot) summer	448.740149982145	0.285323175812403
13	010622	ONGERUP	WA	Southern Sodosol	Correll	-33.9644	118.4889	Temperate	distinctly dry (and warm) summer	395.30442804428	0.354005475538626
14	010633	RAVENSTHORPE	WA	Southern Sodosol	Endure	-33.5803	120.0458	Grassland	warm (persistently dry)	446.676407570527	0.344601833115105
15	012071	SALMON GUMS RES.STN.	WA	Southern Sodosol	Correll	-32.9869	121.6239	Grassland	warm (persistently dry)	379.176824187597	0.287108677538388
16	018012	CEDUNA AMO	SA	Northern Calcarosol	Wyalkatchem	-32.1297	133.6976	Grassland	warm (summer drought)	270.993274610165	0.242114034043566
17	018014	CLEVE	SA	Southern Calcarosol	Endure	-33.7011	136.4937	Temperate	distinctly dry (and hot) summer	400.552850851089	0.36424235210094
18	018044	KYANCUTTA	SA	Southern Calcarosol	Derrimut	-33.1337	135.5521	Grassland	warm (summer drought)	299.924770860612	0.255564813712653
19	018079	STREAKY BAY	SA	Southern Calcarosol	Endure	-32.7963	134.2116	Temperate	distinctly dry (and hot) summer	370.817581240329	0.302225925485061
20	019062	YONGALA	SA	Southern Calcarosol	Wyalkatchem	-33.0276	138.756	Grassland	warm (persistently dry)	362.849363171051	0.269610760623735
21	022015	PRICE	SA	Southern Calcarosol	Bolac	-34.2971	138.0014	Grassland	warm (summer drought)	338.775264849423	0.270443994762528
22	022018	WAROOKA	SA	Southern Calcarosol	Bolac	-34.9906	137.3995	Temperate	distinctly dry (and warm) summer	459.106653969765	0.357576478990596
23	023343	ROSEDALE (TURRETFIELD RESEARCH CENTRE)	SA	Southern Calcarosol	Bolac	-34.5519	138.8342	Temperate	distinctly dry (and hot) summer	480.743363885252	0.32019997619331
24	024518	MENINGIE	SA	Southern Tenosol	Endure	-35.6902	139.3375	Temperate	distinctly dry (and warm) summer	457.820616593263	0.3928103797167
25	025507	KEITH	SA	Southern Tenosol	Endure	-36.098	140.3556	Temperate	distinctly dry (and warm) summer	436.214319723842	0.336626592072372
26	035065	SPRINGSURE COMET ST	QLD	Northern Vertosol	Wyalkatchem	-24.123	148.0856	Subtropical	moderately dry winter	712.338709677419	0.190096417093203
27	041359	OAKEY AERO	QLD	Northern Vertosol	Wyalkatchem	-27.4034	151.7413	Temperate	no dry season (hot summer)	578.682061659326	0.216402809189382
28	043035	SURAT	QLD	Northern Dermosol	Wyalkatchem	-27.1591	149.0702	Grassland	hot (persistently dry)	569.39739316748	0.192596119509582
29	047016	LAKE VICTORIA STORAGE	NSW	Southern Calcarosol	Wyalkatchem	-34.0438	141.2676	Grassland	warm (persistently dry)	258.367515771932	0.190096417093203
30	048031	COLLARENEBRI (ALBERT ST)	NSW	Northern Vertosol	Wyalkatchem	-29.5407	148.5818	Grassland	hot (persistently dry)	564.566063563861	0.179859540530889
31	049002	BALRANALD (RSL)	NSW	Southern Calcarosol	Wyalkatchem	-34.6398	143.561	Grassland	warm (persistently dry)	339.987441971194	0.218069277466968
32	050031	PEAK HILL POST OFFICE	NSW	Southern Chromosol	Wyalkatchem	-32.7235	148.1902	Temperate	no dry season (hot summer)	578.317105106535	0.220568979883347
33	050052	CONDOBOLIN AG RESEARCH STN	NSW	Southern Chromosol	Wyalkatchem	-33.0664	147.2283	Grassland	warm (persistently dry)	448.166646827759	0.234019759552434
34	051039	NYNGAN AIRPORT	NSW	Northern Chromosol	Wyalkatchem	-31.5495	147.1961	Grassland	hot (persistently dry)	486.30895131532	0.180573741221283
35	051049	TRANGIE RESEARCH STATION AWS	NSW	Northern Vertosol	Wyalkatchem	-31.9861	147.9489	Temperate	no dry season (hot summer)	499.169325080348	0.207951434353053
36	052020	MUNGINDI POST OFFICE	NSW	Northern Vertosol	Wyalkatchem	-28.9786	148.9899	Grassland	hot (persistently dry)	512.755267230092	0.17735983811451
37	055049	QUIRINDI POST OFFICE	NSW	Northern Vertosol	Wyalkatchem	-31.5086	150.6792	Temperate	no dry season (hot summer)	693.995298178788	0.249137007499107
38	063254	ORANGE AGRICULTURAL INSTITUTE	NSW	Southern Chromosol	Wyalkatchem	-33.3211	149.0828	Temperate	no dry season (warm summer)	936.708784668492	0.422687775264849
39	072150	WAGGA WAGGA AMO	NSW	Southern Chromosol	Correll	-35.1583	147.4575	Temperate	no dry season (hot summer)	569.884001904535	0.27734793476967
40	073014	GRENFELL (MANGANESE RD)	NSW	Southern Chromosol	Correll	-33.8934	148.1523	Temperate	no dry season (hot summer)	621.607903820974	0.281037971670039
41	074034	COROWA AIRPORT	NSW	Southern Chromosol	Endure	-35.9887	146.3574	Temperate	no dry season (hot summer)	556.91935483871	0.261992619926199
42	074148	NARRANDERA AIRPORT AWS	NSW	Southern Chromosol	Wyalkatchem	-34.705	146.514	Grassland	warm (persistently dry)	430.409772646114	0.229139388168075
43	075032	HILLSTON AIRPORT	NSW	Southern Chromosol	Wyalkatchem	-33.4915	145.5248	Grassland	warm (persistently dry)	398.128496607547	0.176050470182121
44	075039	LAKE CARGELLIGO AIRPORT	NSW	Southern Calcarosol	Wyalkatchem	-33.2832	146.3706	Grassland	warm (persistently dry)	420.825318414474	0.188668015712415
45	076031	MILDURA AIRPORT	VIC	Southern Calcarosol	Wyalkatchem	-34.2358	142.0867	Grassland	warm (persistently dry)	286.690751101059	0.184977978812046
46	076047	OUYEN (POST OFFICE)	VIC	Southern Calcarosol	Wyalkatchem	-35.0682	142.3125	Grassland	warm (persistently dry)	323.247232472325	0.224854184025711
47	078077	WARRACKNABEAL MUSEUM	VIC	Southern Vertosol	Bolac	-36.2614	142.405	Grassland	warm (persistently dry)	380.523687656231	0.275919533388882
48	080015	ECHUCA AERODROME	VIC	Southern Sodosol	Derrimut	-36.1647	144.7642	Grassland	warm (persistently dry)	404.732472324723	0.309129865492203
49	080023	KERANG	VIC	Southern Vertosol	Bolac	-35.7236	143.9197	Grassland	warm (persistently dry)	378.846625401738	0.23925723128199
50	089085	ARARAT PRISON	VIC	Southern Sodosol	Endure	-37.2769	142.9786	Temperate	no dry season (warm summer)	555.381323651946	0.409117962147363

Bibliography

La niña – detailed australian analysis. http://www.bom.gov.au/climate/enso/lnlist/. (cited on pages 10 and 30)

Anwar, S.; Khan, S.; and Barnes, N., 2019. A deep journey into super-resolution: A survey. *arXiv preprint arXiv:1904.07523*, (2019). (cited on pages xiii and 13)

Bengio, Y.; Simard, P.; and Frasconi, P., 1994. Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5, 2 (1994), 157–166. (cited on page 26)

CIE, T. C. f. I. E., 2014. Analysis of the benefits of improved seasonal climate forecasting for agriculture. http://www.climatekelpie.com.au/Files/MCV-CIE-report-Value-of-improved-forecasts-non-agriculture-2014.pdf. (cited on pages xv, 6, and 15)

Dai, T.; Cai, J.; Zhang, Y.; Xia, S.-T.; and Zhang, L., 2019. Second-order attention network for single image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 11065–11074. (cited on page 12)

Dong, C.; Loy, C. C.; He, K.; and Tang, X., 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38, 2 (2015), 295–307. (cited on pages 12 and 16)

Grimit, E. P.; Gneiting, T.; Berrocal, V. J.; and Johnson, N. A., 2006. The continuous ranked probability score for circular variables and its application to mesoscale forecast ensemble verification. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 132, 621C (2006), 2925–2942. (cited on pages 2, 8, 9, 10, 17, and 32)

Hudson, D.; Alves, O.; Hendon, H. H.; Lim, E.-P.; Liu, G.; Luo, J.-J.; MacLachlan, C.; Marshall, A. G.; Shi, L.; Wang, G.; et al., 2017. Access-s1 the new bureau of meteorology multi-week to seasonal prediction system. *Journal of Southern Hemisphere Earth Systems Science*, 67, 3 (2017), 132–159. (cited on pages 1, 7, 8, 9, 13, and 16)

Jiang, W., 2020. High resolution seasonal climate forecast using deep-learning methods. (cited on pages xiv, 3, 13, 14, 17, 19, and 40)

Keys, R., 1981. Cubic convolution interpolation for digital image processing. *IEEE transactions on acoustics, speech, and signal processing*, 29, 6 (1981), 1153–1160. (cited on page 11)

Kim, J.; Kwon Lee, J.; and Mu Lee, K., 2016. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1646–1654. (cited on pages 2 and 12)

Li, M. and Jin, H. Development of a postprocessing system of daily rainfall forecasts for seasonal crop prediction in australia. (cited on pages xiv, xv, 13, 14, 16, 33, 47, and 48)

Maraun, D., 2013. Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate*, 26, 6 (2013), 2137–2143. (cited on pages 1, 7, 14, and 16)

Merryfield, W. J.; Baehr, J.; Batté, L.; Becker, E. J.; Butler, A. H.; Coelho, C. A.; Danabasoglu, G.; Dirmeyer, P. A.; Doblas-Reyes, F. J.; Domeisen, D. I.; et al., 2020. Current and emerging developments in subseasonal to decadal prediction. *Bulletin of the American Meteorological Society*, 101, 6 (2020), E869–E896. (cited on pages 6, 7, and 15)

Minzner, R. A.; Reber, C.; Jacchia, L.; Huang, F.; Cole, A.; Kantor, A.; Keneshea, T.; Zimmerman, S.; and Forbes, J., 1976. Defining constants, equations, and abbreviated tables of the 1975 us standard atmosphere. (1976). (cited on page 15)

Pan, B.; Hsu, K.; AghaKouchak, A.; and Sorooshian, S., 2019. Improving precipitation estimation using convolutional neural network. *Water Resources Research*, 55, 3 (2019), 2301–2321. doi:https://doi.org/10.1029/2018WR024090. https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR024090. (cited on pages xiii, 3, 7, 9, 15, 16, 17, and 19)

Parton, K. A.; Crean, J.; and Hayman, P., 2019. The value of seasonal climate forecasts for australian agriculture. *Agricultural Systems*, 174 (2019), 1–10. (cited on page 6)

Philander, S., 1985. El niño and la niña. *Journal of the Atmospheric Sciences*, 42, 23 (1985), 2652–2662. (cited on pages 3, 10, 14, and 30)

Su, C.-H.; Eizenberg, N.; Steinle, P.; Jakob, D.; Fox-Hughes, P.; White, C. J.; Rennie, S.; Franklin, C.; Dharssi, I.; and Zhu, H., 2019. Barra v1.0: the bureau of meteorology atmospheric high-resolution regional reanalysis for australia. *Geoscientific Model Development*, 12, 5 (2019), 2049–2068. doi:10.5194/gmd-12-2049-2019. https://gmd.copernicus.org/articles/12/2049/2019/. (cited on pages xiii, 1, 9, and 10)

Timofte, R.; De Smet, V.; and Van Gool, L., 2014. A+: Adjusted anchored neighborhood regression for fast super-resolution. In *Asian conference on computer vision*, 111–126. Springer. (cited on page 11)

Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; and Fu, Y., 2018. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 286–301. (cited on page 12)