

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Masterthesis Informatik

---

**Quantifying the Intertropical  
Convergence Zone using wind  
convergences**

---

*Author:*

Lea Elsemüller

*Supervisors:*

Dr. Bedartha Goswami

Prof. Dr. Martin Butz

November 3, 2021

## **Selbständigkeitserklärung**

Hiermit erkläre ich, dass ich diese schriftliche Abschlussarbeit selbstständig verfasst habe, keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe.

Tübingen, 03. November 2021

Lea Elsemüller

## Abstract

The Intertropical Convergence Zone is an area characterized by high precipitation within a narrow belt around the equator. The ITCZ migrates annually to the warmer hemisphere, and the extent to which it ventures away from the equator varies from year to year and across the different oceans and continents. These variations drastically affect rainfall and droughts in the equatorial area and beyond. Till now, various approaches have been proposed to quantify the ITCZ, e.g. based on maximum precipitation or energy budgets. However, a robust quantifier of the actual convergence of surface winds around the equator is still lacking. Here I propose a method of quantifying the ITCZ mid-location based on surface wind divergences. The latitudinal ITCZ mid-position was defined as the minimum convergence over longitudes and validated by comparing the given latitude to the ITCZ location as given by existing ITCZ position proxies. I used the given latitudes to train different LSTMs to predict future mid-locations of the ITCZ. I also looked at characteristics of the ITCZ width to learn more about the influence of wind fields on the extent of the ITCZ. I used a ConvLSTM network to graphically analyze the ITCZ belt. My results reveal the interannual variability and trends in the ITCZ in the last half century. The LSTM structures are feasible to predict the ITCZ mid-location for a short period of time. The ConvLSTM structure gave first results for a graphical analysis of characteristics like the width of the ITCZ.

## Zusammenfassung

Die Intertropische Konvergenzzone (ITCZ) ist eine schmales Gebiet rund um den Äquator, das sich durch hohe Niederschläge auszeichnet. Diese Zone wandert jährlich und ihre Ausdehnung variiert von Jahr zu Jahr. Diese Variationen bestimmen die Regenfälle und Dürren in der Äquatorregion und darüber hinaus. Bislang gibt es verschiedene Ansätze um die ITCZ zu quantifizieren, z. B. auf der Grundlage von Niederschlagsmaxima oder Energiebudgets. In dieser Arbeit schlage ich eine Definition für die ITCZ basierend auf Winddivergenzen vor. Die mittlere Position der ITCZ wurde als Breitengrad mit minimaler Divergenz definiert. Diese Definition habe ich mit anderen ITCZ Variablen verglichen. Mit verschiedenen Zeitreihen für die mittlere ITCZ-Position habe ich ein Long Short Term Memory (LSTM) trainiert, um die zukünftigen Positionen vorherzusagen. Weiterhin habe ich ein convolutional LSTM (Conv LSTM) auf den Divergenzwerten trainiert, um mehr über den Einfluss der Divergenz auf die Ausdehnung der ITCZ zu erfahren und die ITCZ graphisch vorherzusagen. Meine Ergebnisse zeigen die jährliche Variabilität und die Trends in der ITCZ während des letzten halben Jahrhunderts auf. Die LSTMs sind gut dafür geeignet, die mittlere Position der ITCZ für einen kurzen Zeitraum vorherzusagen. Die ConvLSTM-Struktur lieferte erste Ergebnisse für eine grafische Analyse der Breite der ITCZ.

## Acknowledgements

Completing this thesis would not have been possible for me without the opportunity and continuous support and guidance from my supervisor Dr. Bedartha Goswami and the entire research group Machine Learning in Climate Science. Thank you for the opportunity to learn about climate science and to expand my knowledge in the use of machine learning in a practical research field. It has helped me a lot to get feedback from the whole research group and to regularly discuss the results. I would also like to acknowledge Prof. Dr. Martin Butz as the second supervisor of this thesis, and I am very grateful for his support and feedback.

# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
<b>2</b>	<b>Recent work</b>	<b>8</b>
2.1	Intertropical Convergence Zone . . . . .	8
2.2	ITCZ quantification . . . . .	9
<b>3</b>	<b>Data</b>	<b>11</b>
3.1	Study location and dataset . . . . .	11
<b>4</b>	<b>Method</b>	<b>12</b>
4.1	Estimating mid-latitudes of ITCZ . . . . .	12
4.2	Artificial neural networks . . . . .	13
4.2.1	Recurrent neural networks . . . . .	13
4.2.2	Long Short-Term Memory . . . . .	13
4.2.3	Conv LSTM . . . . .	14
4.3	Predicting ITCZ mid-location using LSTMs . . . . .	15
4.3.1	Data preparation . . . . .	15
4.3.2	Single day predictions . . . . .	16
4.3.3	Multi day predictions . . . . .	17
4.4	Predicting ITCZ proxy values using ConvLSTMs . . . . .	18
<b>5</b>	<b>Results</b>	<b>19</b>
5.1	Estimating mid-latitudes of ITCZ . . . . .	19
5.2	Predicting ITCZ mid-location using LSTMs . . . . .	22
5.2.1	Single day predictions . . . . .	22
5.2.2	Multi day predictions . . . . .	23
5.3	Predicting values over time using Conv LSTMs . . . . .	24
<b>6</b>	<b>Discussion</b>	<b>27</b>
<b>7</b>	<b>Appendix</b>	<b>32</b>

## List of Figures

1	Monthly seasonal migrations of the ITCZ . . . . .	8
2	Monthly divergence and precipitation over latitudes in 2020 for -110 W	11
3	Input and Architecture LSTM for single day predictions . . . . .	16
4	Architecture LSTM multi day predictions . . . . .	17
5	Input and architecture ConvLSTM . . . . .	18
6	Daily divergence and precipitation values . . . . .	20
7	ITCZ mid-latitudes comparison . . . . .	21
8	Prediction results latitudes of minimum divergence . . . . .	22
9	Results prediction latitudes of minimum divergence, multi day pre- dictions . . . . .	23
10	Correlation predicted latitude and observed latitude, minimum diver- gence, multi day predictions . . . . .	24
11	Visual results for divergence for the ConvLSTM . . . . .	25
12	Correlation results ConvLSTM . . . . .	26
13	Anomaly map . . . . .	26
14	Anomalys timeseries . . . . .	27
15	Average seasonal divergence and precipitation . . . . .	32
16	Average monthly divergence and precipitation 1 . . . . .	33
17	Average monthly divergence and precipitation 2 . . . . .	34
18	Seasonal migrations based on daily averaged data . . . . .	35
19	Prediction results latitudes of maximum precipitation 1 . . . . .	36
20	Prediction results latitudes of maximum precipitation 2 . . . . .	37
21	Correlation predicted latitude and observed latitude, maximum pre- cipitation 1, multi day predictions . . . . .	38
22	Correlation predicted latitude and observed latitude, maximum pre- cipitation 2, multi day predictions . . . . .	39
23	Timeseries multi day predictions precipitation 1 . . . . .	40
24	Timeseries multi day predictions precipitation 2 . . . . .	41

## List of Tables

1	Pearson correlation values between -110°W and -100°E . . . . .	20
2	Pearson correlation latitudes between -110°W and -100°E . . . . .	21
3	Pearson correlation values between 160°E and 150°E . . . . .	35
4	Pearson correlation latitudes between 160°E and 150°E . . . . .	35

# 1 Introduction

The Intertropical Convergence Zone (ITCZ) is defined as a zone of low pressure running around the entire globe at the equator, where the trade winds of the northern and southern hemispheres flow together. The ITCZ region is characterized by a warm and humid climate. Large parts of the world's tropical rainforests are located in the ITCZ. The ITCZ expresses itself as a narrow band of strong convection around the equator with associated heavy precipitation (Waliser and Gautier, 1993; Schneider et al., 2014). It can be defined as the convergence of the moist trade winds near the surface, which leads to an ascent of air masses, condensation and precipitation (Waliser and Gautier, 1993; Schneider et al., 2014). Around the globe, enormous air masses are transported along the equator. In the ITCZ, the surfaces of land and water heat up more than anywhere else on earth, as solar radiation is strongest here (Schneider et al., 2014). Warm air rises rapidly and pulls "supplies" from the northern and southern hemispheres behind it.

The ITCZ migrates seasonally and its expansion and mid-position has influence on the climate in the tropics and subtropics (Waliser and Gautier, 1993; Philander et al., 1996). An accurate description of the seasonal different mid-positions of the ITCZ over time can be used to identify patterns and predict future expansion of the ITCZ.

This thesis proposes a method to quantify the ITCZ mid-position and its migrations with a fundamental and intuitive definition using surface wind data and wind divergence. I used surface wind and divergence data from ERA5 reanalysis data at 0.25° grid resolution as a proxy for calculating the ITCZ mid-position in the Pacific area. I compared this definition to definitions based on precipitation as a proxy. To predict the ITCZ mid-location I calculated the latitudes of minimum divergence and maximum precipitation and then used a long short term (LSTM) neural network for prediction. Furthermore, I also trained a convolutional LSTM (Conv LSTM) to get predictions of the ITCZ band as such and to get predictions of further meridional extensions of the ITCZ.

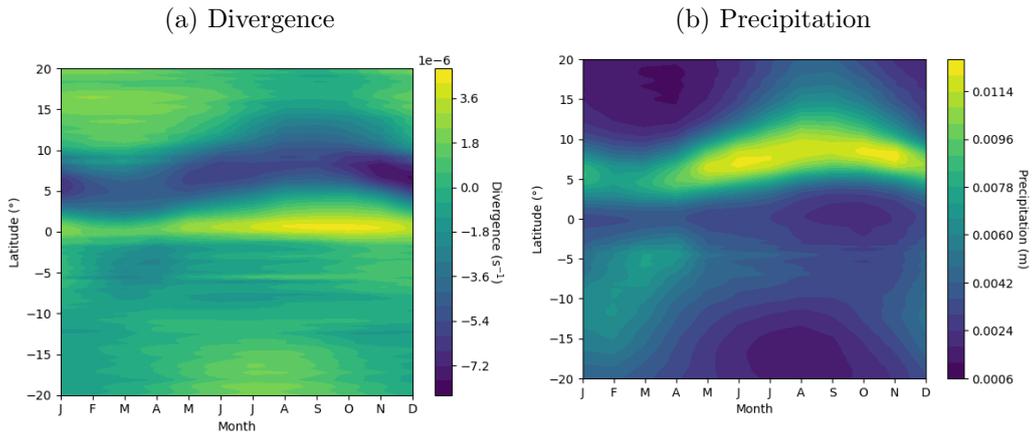


Figure 1: Seasonal migrations of the ITCZ based on a) monthly averaged divergence data and b) monthly averaged precipitation data over the Pacific longitudinal averaged between  $130^{\circ}\text{E}$  and  $-80^{\circ}\text{W}$  averaged over the years 1979-2020.

## 2 Recent work

### 2.1 Intertropical Convergence Zone

The trade winds or doldrums are known for centuries by sailors and scientists as reliable winds which blow east to west, north and south of the equator (Hadley, 1735). The trade wind systems of the two hemispheres tend to converge in the equatorial area, which leads to one of the definitions of the ITCZ as the convergence zone of the trade winds (Barry and Chorley, 2009).

The ITCZ is responsible for 32% of global precipitation (Kang et al., 2018). The amount of precipitation is much higher than the moisture coming from the ocean surface below the ITCZ, so much of the vapor which is necessary for the convection must be supplied by the moist and warm converging wind systems (Holton and Hakim, 2013; Schneider et al., 2014). Warm air masses ascent, cool down and condensate in the central area of the ITCZ which leads to heavy precipitation (Schneider et al., 2014).

The ITCZ shapes climate and society in the tropics and subtropics. It has major influence on the East Asian Monsoon (Yancheva et al., 2007) but also has a large impact on the rainfall in the tropics and on the anticyclonic subtropics. Current studies show that the ITCZ has influence on global radiation budget, which has impact on global temperature and precipitation (Su et al., 2017; Byrne and Schneider, 2018). However, the ITCZ width is narrowing as a result of global warming (Byrne and Schneider, 2016a).

The ITCZ and the Hadley circulation form a seasonal circle (Waliser and Gautier, 1993; Oort and Yienger, 1996). The ITCZ migrates seasonally between  $5^{\circ}\text{S}$  and  $15^{\circ}\text{N}$  (Waliser and Gautier, 1993). Figure 1 shows the monthly averaged annual cycle

of the ITCZ mid-location based on divergence and precipitation over the Pacific. Maps of the seasonal and monthly shifts of the ITCZ based on divergence and precipitation as well as the seasonal migrations based on average daily data can be found in the appendix. Overall it can be seen that the ITCZ shifts sinusoidal to the north from spring to late summer and shifts back south from late summer to spring independent of the used proxy. The further expansion to the north can be explained by warmer sea surface temperatures in the Northern Hemisphere, larger land masses and differences in coastal geometry (Philander et al., 1996). It has been shown that the migrations of the ITCZ mid-location are nearly sinusoidal (Waliser and Gautier, 1993). The ITCZ migrates annually to the warmer hemisphere, and the extent to which it ventures away from the equator varies from year to year and across the different oceans and continents (Schneider et al., 2014). These variations drastically affect rainfall and droughts in the equatorial area and beyond.

On a geological timescale, the ITCZ was subject to major changes (Arbuszewski et al., 2013; Haug et al., 2001). Since the last glacial maximum, the Atlantic ITCZ migrated latitudinal away from the relational cooler hemisphere responding to different climate conditions (Arbuszewski et al., 2013). In South America regional shifts in precipitation were linked to shifts in the mean latitude of the Atlantic ITCZ which might be connected by Pacific-based climate variability (Haug et al., 2001). In today's climate factors like the presence of polar ice cover or temperature variations on high latitudes can also cause shifts of the ITCZ (Broccoli et al., 2006). These shifts are again linked to shifts in trade winds and in the Hadley circulation.

## 2.2 ITCZ quantification

Today different approaches to quantify the ITCZ are used. In this thesis I focus on approaches for the variation of ITCZ mid-location and on ITCZ width. There is no fixed definition to identify the mid-location of the ITCZ but there are some promising approximations based on different proxy variables. Previous work has focused on energy constraints (Bischoff and Schneider, 2014), cloud temperature (Waliser and Gautier, 1993), high precipitation (Gu et al., 2005) but also on wind and divergence (Barry and Chorley, 2009; Berry and Reeder, 2014; Žagar et al., 2011).

The ITCZ and the ITCZ mid-latitude are originally defined as the convergence of the Northern and Southern trade winds (Barry and Chorley, 2009). Convergence is defined as negative divergence, which is defined as the rate at which air is spreading out horizontally from a point. Therefore, divergence of a wind field  $\vec{V}^2$  is defined as the sum of the change of the u-component of the wind in the longitudinal direction and v-components of wind in latitudinal direction for a specific location.

$$\text{div } \vec{V}^2 = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \quad (1)$$

For values of  $\text{div } \vec{V}^2 > 0$  we speak of divergence of air masses. For values of  $\text{div } \vec{V}^2 < 0$  we speak of convergence of air masses.

In addition to the ITCZ mid-latitude, the ITCZ meridional extent plays an important role for climate in the tropics. Compared to the ITCZ mid-latitude this quantification hasn't received so much attention yet, although its importance for climate in the tropics is high (Byrne and Schneider, 2016a). ITCZ width influences the extent of rain in the tropics and there are references which connects a narrowing of the ITCZ to a warming climate (Byrne and Schneider, 2016b). The ITCZ width is defined as the latitudinal distance between the latitude of the time mean ascending branch of the Hadley circulation and the latitude of the time mean descent (Byrne and Schneider, 2016a).

There have been several approaches to simulate and predict ITCZ shifts globally and in specific regions. This is investigated a lot in the field of climate models like CMIP5 and CMIP6, with focus on the simulation of the ITCZ and the influence of shifts in the ITCZ on global climate (Xiang et al., 2017; Brown et al., 2013; Narsey et al., 2020). But there are also approaches of spatial analysis and prediction of the ITCZ using p-splines and Gaussian Markov Random Fields (Greco et al., 2018).

In this thesis I focused on lower tropospheric wind divergences as a proxy for the ITCZ mid-location. I compared this method of quantifying the ITCZ mid-location to approaches based on precipitation maxima (Adam et al., 2016a,b). The idea of a connection of convective rainfall and divergence is not new and has been proven in many areas (Watson and Blanchard, 1984). Figure 2 shows the anticorrelation between divergence and precipitation for monthly averaged precipitation and divergence values for the year 2020 in an area between 20°N and -20°S for longitude -110°W. The precipitation Maxima and the divergence Minima are clearly visible and anti-correlated. I used this anticorrelation between precipitation and divergence to compare the mid-location given by minimum divergence to the mid-location given by maximum precipitation. With the latitudes given by minimum divergence and maximum precipitation I trained three different LSTMs to get predictions of future mid-locations of the ITCZ.

ITCZ width changes with further climate warming (Byrne and Schneider, 2016b). To get further information about the ITCZ characteristics I trained a ConvLSTM to get a graphical prediction of future divergence and precipitation values to get information about future extends of the ITCZ.

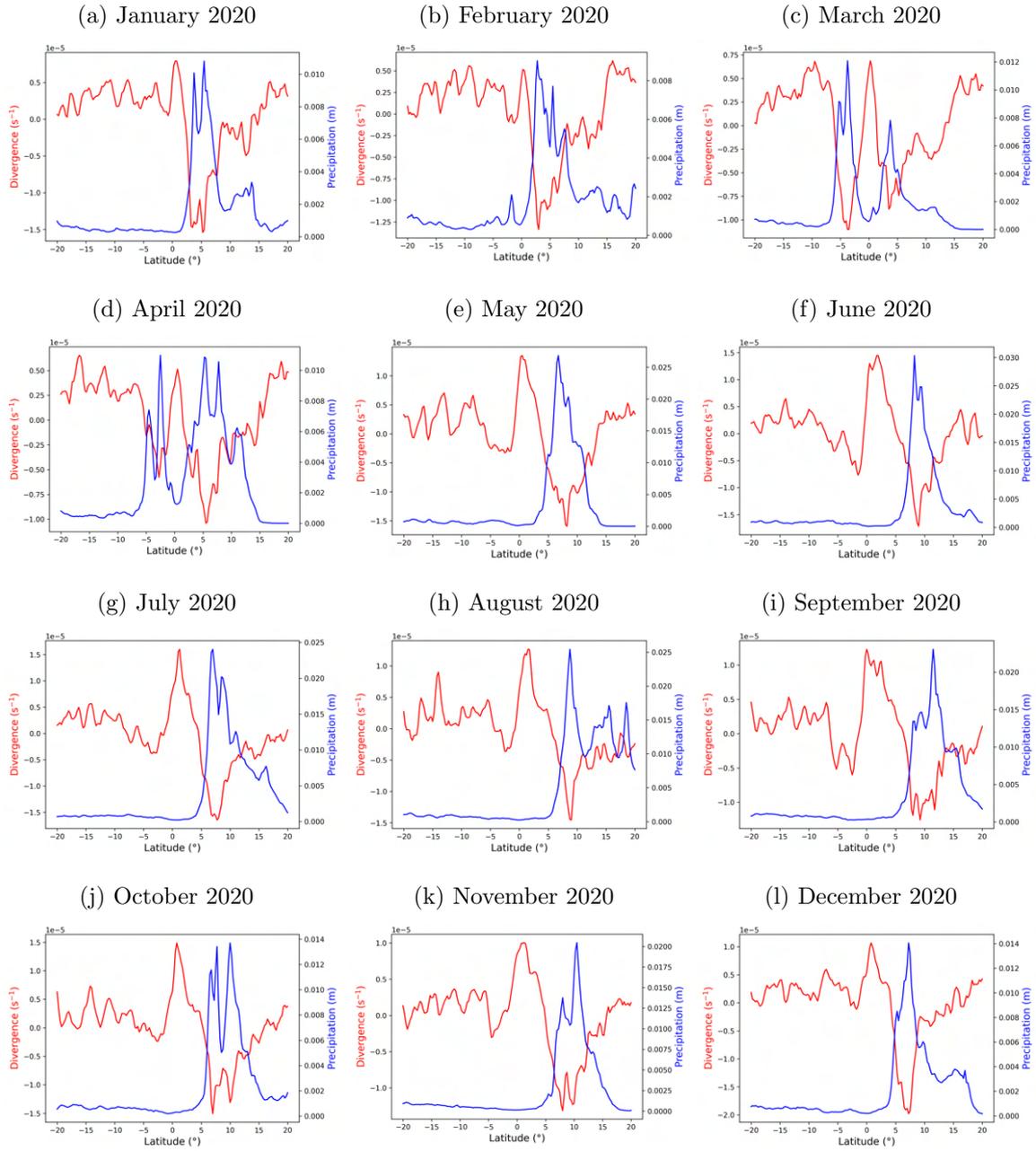


Figure 2: Monthly divergence and precipitation in 2020 for longitude  $-110^{\circ}\text{W}$  and different latitudes between  $20^{\circ}\text{N}$  and  $-20^{\circ}\text{S}$ .

### 3 Data

#### 3.1 Study location and dataset

In this thesis, I used ERA5 reanalysis data with a resolution of  $0.25\text{ deg} \times 0.25\text{ deg}$  for the period of 1979-2020 for precipitation and horizontal divergence at 1000 hPa pressure level. The values of precipitation are taken from the ERA5 hourly dataset on single levels from 1979 to present. The values of divergence are taken from the

ERA5 hourly dataset on pressure levels from 1979 to present.

Precipitation is reported in meter (m) over a 0.25 deg x 0.25 deg grid. Divergence is reported as the rate at which air is spreading out horizontally from a point, per square meter ( $s^{-1}$ ). Negative values of the divergence are called convergence. For plotting issues, I also used u and v components of the wind. The u-component of the wind is the eastward component of the wind given as horizontal speed of air, which is moving towards the east ( $ms^{-1}$ ). A negative sign indicates a westward movement. The v-component of the wind is the northward component of the wind also given as horizontal speed of air, which is moving towards the north ( $ms^{-1}$ ). A negative sign indicates a southward movement.

I used hourly data sampled four times a day and averaged it to daily data as well as to monthly data. I decided to limit the research area to the Pacific between  $130^{\circ}E$  to  $-90^{\circ}W$ . For predictions I limited this area to  $-110^{\circ}W$  to  $-100^{\circ}W$  because of high anticorrelation between divergence and precipitation.

## 4 Method

### 4.1 Estimating mid-latitudes of ITCZ

For calculating an index for the mid-latitudes of the ITCZ I used daily data between  $-110^{\circ}W$  and  $-100^{\circ}W$  and  $-20^{\circ}S$  and  $20^{\circ}N$  and for comparison between  $150^{\circ}E$  and  $160^{\circ}E$  and  $-20^{\circ}S$  and  $20^{\circ}N$ .

I followed three different approaches and compared the results. First I calculated the mid-latitude of the highest precipitation by taking the latitude of the maximum time weighted precipitation  $P$  over a specific longitude range.

$$\phi_{max} = argmax P \quad (2)$$

Second I implemented the index given by Adam et al., which also defined the ITCZ mid-latitude as the latitude of maximum precipitation (Adam et al., 2016a,b). Here the maximum precipitation is calculated as an expected latitude, which was weighted by the 10th power of the area weighted precipitation  $P$  which was integrated between  $-20^{\circ}S$  and  $20^{\circ}N$  on a latitudinal scale (Adam et al., 2016a). For comparison I called the latitude given by this definition  $\phi_{adam}$ .

$$\phi_{adam} = \frac{\int_{-20^{\circ}S}^{20^{\circ}N} \phi ((\cos(\phi)P)^{10})}{\int_{-20^{\circ}S}^{20^{\circ}N} ((\cos(\phi)P)^{10})} \quad (3)$$

Finally, for getting the latitude of the minimum divergence I calculated the minimum of the divergence  $D$ .

$$\phi_{min} = argmin D \quad (4)$$

For comparing the different latitude calculations I used the Pearson correlation between each two of the latitude estimates  $\phi_{min}$ ,  $\phi_{max}$  and  $\phi_{adam}$  given here as X

and Y

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

where  $\bar{x}$  and  $\bar{y}$  are the means of the compared latitude samples.

## 4.2 Artificial neural networks

In this study I predicted time series of latitudes and time series of spatial distributed values using neural networks. A neural network (NN) is built up of many units which are known as neurons and are organized in layers (Goodfellow et al., 2016). Many real-world problems are too complicated to be solved by a direct mapping from one function to another. This problem can be tackled by deep neural networks which splits the initial complicated mapping in a series of simpler mappings which each are described by a different layer in the model. The initial layer in a NN is used to process the input and the final layer produces the output of the network. In between there are hidden layers which extract features from the data.

To predict future ITCZ mid-latitudes which are only distributed over time, I used long short term memories (LSTM) which are a special type of recurrent neural network (Hochreiter and Schmidhuber, 1997). To predict the actual divergence and precipitation values which are spatially and temporally distributed I used convolutional long short term memories (Conv LSTM) (Xingjian et al., 2015).

### 4.2.1 Recurrent neural networks

Recurrent neural networks (RNN) are neural networks which are used for processing and predicting sequential data (Rumelhart et al., 1986). A RNN implements feedback connections or cycles and share parameters across different layers of a model to generalize over sequences (Goodfellow et al., 2016). Using feedback connections makes it possible to store information over sequences in form of activations (Elman, 1990). The activations of the previous time step are then used as inputs to the network at the current time step and influence the predictions. Compared to a simple feed forward neural network which is trained on independent data points RNNs do not need independent data. Sequential data X can be distributed over time T giving us a time series  $x^{(1)}, \dots, x^{(T)}$  where one data point depends on the previous point. One challenge in RNNs is to learn long-term dependencies using backpropagation (Goodfellow et al., 2016). Gradients tend to vanish when propagated over many time steps or on the opposite sometimes explode (Bengio et al., 1994). LSTMs give a solution to this problem by introducing a memory cell with self-loops to maintain a state over time and nonlinear gating units to regulate information flow in and out of the cell (Hochreiter and Schmidhuber, 1997).

### 4.2.2 Long Short-Term Memory

LSTMs introduce LSTM cells which have a self-loop to solve the problem with propagating back the error (Hochreiter and Schmidhuber, 1997). The LSTM cells have the same input and the same output as a normal RNN (Goodfellow et al., 2016). But they introduce internal mechanisms which are called gates to control information

flow. These gates determine which information is important and should be taken further along and which information should be forgotten. The input of a LSTM cell  $x_t$  is computed by a normal artificial neuron unit. Its value is propagated forward to the state unit  $c_t$  if the input gate  $i_t$  allows it. The input gate is computed with a sigmoid unit to get a value between 0 and 1.

$$i_t = \sigma(b_i + U_i x_t + W_i h_{t-1}) \quad (6)$$

$h_t$  describes the current hidden layer vector, which contains the output of all the LSTM cells.  $b_i$  are the biases,  $U_i$  are the input weights and  $W_i$  describe the recurrent weights for the input gate. The state unit has a linear self-loop which is delayed by one time step and feeds back the previous state to the state unit. The weight of the state unit is controlled by the forget gate  $f_t$  (for time step  $t$  and cell  $i$ ). The forget gate sets the weight of the state unit between 0 and 1 by applying a sigmoid function  $\sigma$ .

$$f_t = \sigma(b_f + U_f x_t + W_f h_{t-1}) \quad (7)$$

$b_f$  are the biases,  $U_f$  are the input weights and  $W_f$  describe the recurrent weights for the forget gate. This forget function is now used to update the state unit  $c_t$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(b_c + U_c x_t + W_c h_{t-1}) \quad (8)$$

Here  $b$ ,  $U$  and  $W$  describe the biases and the input and recurrent weights into the LSTM cell.  $\odot$  stands for the pointwise multiplication. The state unit can also be used as one time step delayed input to gating units.

$$h_t = \tanh(c_t) \odot o_t \quad (9)$$

$$o_t = \sigma(b_o + U_o x_t + W_o h_{t-1}) \quad (10)$$

$b_o$  are the biases,  $U_o$  are the input weights and  $W_o$  describe the recurrent weights for the output gate. The output or hidden state vector  $h_t$  can be controlled and shut off by the output gate  $o_t$ . The output gate is also controlled by a sigmoid function while the actual output is calculated by using a hyperbolic tangent function. (Hochreiter and Schmidhuber, 1997; Gers et al., 2000; Goodfellow et al., 2016).

### 4.2.3 Conv LSTM

The LSTM approach performs good for sequential data like the ITCZ mid-latitude. For predicting the divergence values however I needed to extend the model to spatio-temporal sequences. Convolutional neural networks (CNN) are a method which works well on learning images which are also spatial distributed (LeCun et al., 1989). The divergence values can be interpreted as sequential images. A possible approach of combining the sequential power of an LSTM and the spatial abilities of a CNN is using a convolutional LSTM (ConvLSTM) which was first introduced for precipitation nowcasting (Xingjian et al., 2015). The advantage of using a ConvLSTM is that this model also take spatial correlations into account. In a ConvLSTM all inputs  $x_1, \dots, x_t$ , all cell outputs  $c_1, \dots, c_t$ , all hidden states  $h_1, \dots, h_t$  and gates  $i_t$ ,  $f_t$  and  $o_t$  are 3D tensors where the first dimension is sequential and the last two are the spatial

dimensions. To get the future state of a cell, a convolutional operator is applied in the state-to-state and input-to-state operations.

$$i_t = \sigma(b_i + U_i \star x_t + W_i \star h_{t-1}) \quad (11)$$

$$f_t = \sigma(b_f + U_f \star x_t + W_f \star h_{t-1}) \quad (12)$$

$$o_t = \sigma(b_o + U_o \star x_t + W_o \star h_{t-1}) \quad (13)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(b_c + U_c \star x_t + W_c \star h_{t-1}) \quad (14)$$

$$h_t = \tanh(c_t) \odot o_t \quad (15)$$

$\star$  stands for the convolutional operator and  $\odot$  for the pointwise multiplication.

### 4.3 Predicting ITCZ mid-location using LSTMs

To predict the ITCZ mid-location I used LSTMs over the area of -100°W and -110°W and -20°S and 20°N. I first calculated the latitude of the highest precipitation by the definitions given by Equation (2) and Equation (3) to get two datasets of latitudes of highest precipitation over time. I also calculated the latitude of the lowest divergence by the definition given by Equation (4) to get a dataset of the latitudes of lowest divergence over time.

#### 4.3.1 Data preparation

Divergence and precipitation are strongly fluctuating temporally but also spatially. For calculating and predicting the mid-position of the ITCZ I am interested in the latitudes for a specific longitude range. To get less fluctuating spatial distributed values, I took the arithmetic mean of precipitation and divergence over longitudes over the area of interest. To get rid of some of the temporal fluctuations, I used a mean filter (also known as uniform filter) of size 7 over the temporal dimension for both variables. The final dataset for the range of interest for calculating the latitude of the ITCZ mid-position is therefore a 2D Tensor with a temporal and a latitudinal dimension.

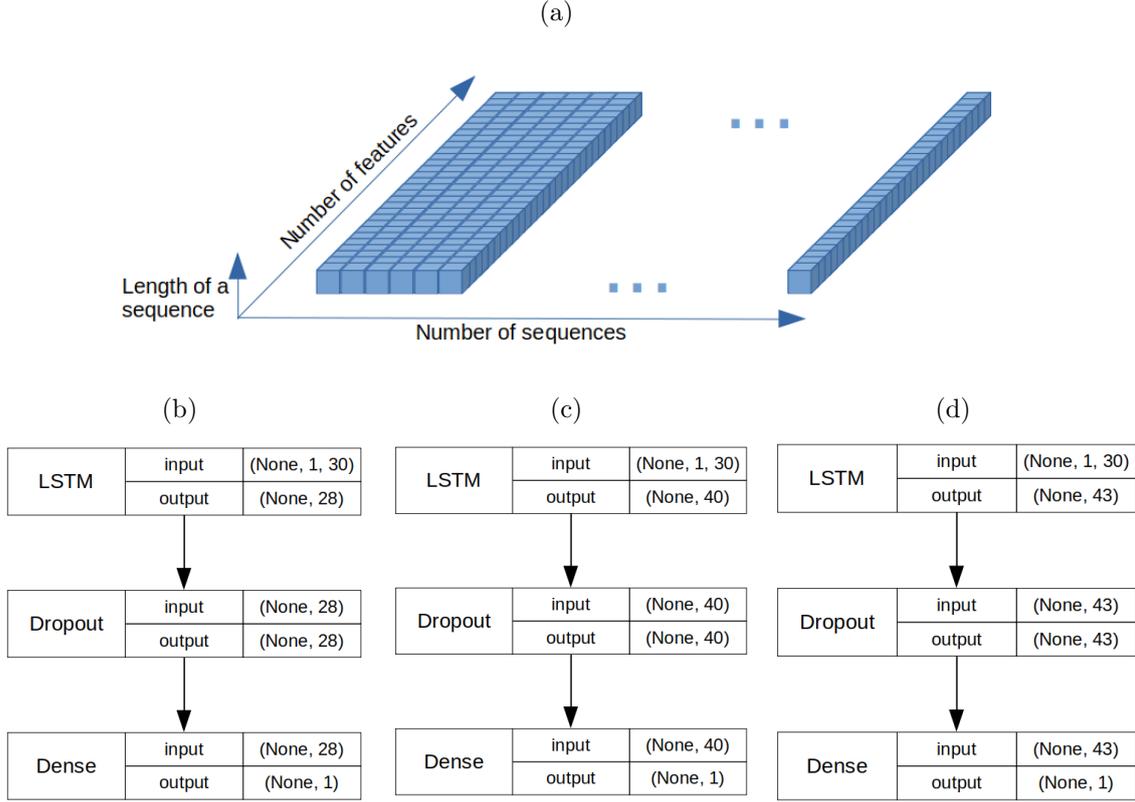


Figure 3: a) Structure of the input to the LSTM trained for single day predictions b) LSTM model structure for single day predictions for the latitude of minimum divergence c) LSTM model structure for single day predictions for the latitude of maximum precipitation d) LSTM model structure for single day predictions for the latitude of maximum precipitation as given by Equation (3)

### 4.3.2 Single day predictions

For the single day predictions, I used the daily dataset between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and between  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$ . I took the arithmetic mean over all longitudes and used a 7-day mean filter over time. I then calculated the latitudes of maximum precipitation using Equation (2) and Equation (3) as well as the latitude of minimum divergence as given by (4). I then trained three LSTMs to predict the latitudes of maximum precipitation or minimum divergence respectively for a single day. Generally the input of an LSTM needs to take the shape of a 3D array. The x-axis represents the number of sequences, here the time dimension over the years 1979 until 2020 on a 1-day scale. In total, I got 13717 sequences for training using 90 % of the total available data for training. I held back 10 % of the data for testing. The y-axis represents the length of a sequence. In this case, this is only 1 day. The z-axis describes the number of time steps (or features) which needed to be taken into account to predict the next time step. I set this to 30, so the last 30 days are used to predict the next day. An overview of the input shape is illustrated in Figure 3a. The output of the LSTM is the predicted latitude of maximum precipitation or minimum divergence for the next time step.

The structure of the LSTM can be found in Figure 3b. First a LSTM layer is used

to output hidden states for each time step input. After the LSTM layer, a dropout layer is applied for regularization and 20 % of the hidden states are probabilistically removed as inputs to the next layer. This has the advantage that the network becomes more robust and overfitting is prevented (Srivastava et al., 2014). Finally, a dense layer is applied to compare the predictions to the original values. The number of neurons is determined by using an autocorrelation function in advance. The autocorrelation function can be used to calculate if a time series is still dependent on its past and therefore I was using as many neurons as necessary to satisfy this requirement. As loss function, I selected Mean Square Error (MSE)

$$MSE = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (16)$$

where n is the number of training examples,  $Y_i$  is the actual value and  $\hat{Y}_i$  is the predicted value by the LSTM. I used the Adam optimizer for learning rate optimization (Kingma and Ba, 2014) with a learning rate of 0.001. To compare the results of the prediction to the true values I used Pearson correlation again.

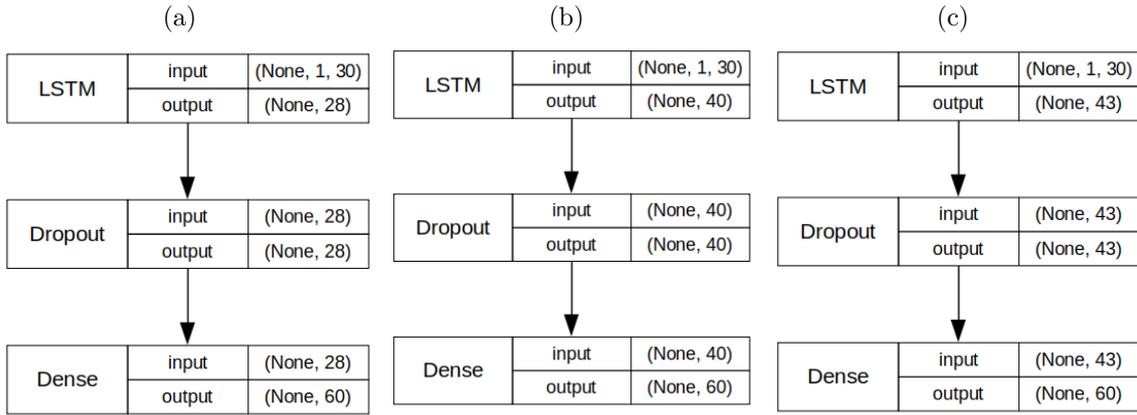


Figure 4: Structure of the LSTM trained for multi day predictions based on a) latitude of minimum divergence as given by Equation (4) b) latitude of maximum precipitation as given by Equation (2) and c) latitude of maximum precipitation as given by Equation (3).

### 4.3.3 Multi day predictions

For the multi day predictions, I also used the daily dataset between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$ . I again used the arithmetic mean over all longitudes and used the 7 days mean filter over time. The longitudes were again calculated according to the known procedures given by Equation (2), Equation (3) and Equation (4). As input I used the same data structure as illustrated in Figure 3a with the last 30 days used to predict the next days. However, I now wanted not only one time step as output but a much longer period of time. Therefore, I adjusted the last layer of the network to output a longer sequence as prediction. Here I set the length to 60 days to get a prediction for the next two months. The architecture of the three LSTMs

can be found in Figure 4. As loss I again used the MSE loss as given in Equation (16). As optimizer the Adam optimizer was applied again.

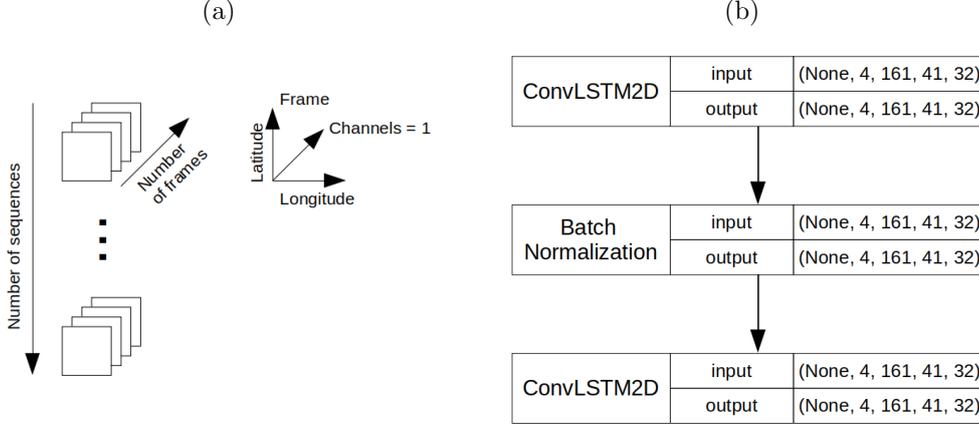


Figure 5: a) Input structure to the ConvLSTM. The input is a 5D Tensor with the shape (number of sequences, number of frames, number of longitudes, number of latitudes, number of channels) where the number of frames is fixed to 4, the longitudes are between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  with a  $0.25^{\circ}$  resolution, the latitudes are between  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  also on a  $0.25^{\circ}$  grid resolution and the number of channels is fixed to 1. b) shows the architecture of the ConvLSTM.

#### 4.4 Predicting ITCZ proxy values using ConvLSTMs

Up to now I focused on predicting the ITCZ mid-location as a latitude. To further predict characteristics like the ITCZ width I applied a ConvLSTM network for a graphical prediction of the actual divergence and precipitation values. For that purpose I again used the daily dataset between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$ . For training and testing I limited the dataset to 10 years time from 2011 until 2020. Concerning the ConvLSTM I used daily data and applied a 7-day mean filter over time before training and predictions.

ConvLSTMs are a combination of CNNs and LSTMs which are able to simultaneously process spatial and temporal information (Xingjian et al., 2015). I used this to predict the next four frames based on the last four frames for a time sequence. The input structure to the ConvLSTM can be found in Figure 5a. I applied a very basic architecture for the ConvLSTM with only two ConvLSTM layers. In between the ConvLSTMs layers I applies one layer Batch Normalization which is used to normalize the layer input (Ioffe and Szegedy, 2015). The architecture for the ConvLSTMs can be found in 5b. For the ConvLSTM layers I used a kernel size of  $3 \times 3$  and a filter size of 32. For training I again used the Adam optimizer and MSE loss.

## 5 Results

### 5.1 Estimating mid-latitudes of ITCZ

A visual description of monthly anticorrelation can be found in Figure 2 for the single longitude  $-110^{\circ}\text{W}$  for one year. I found an anticorrelation for larger areas in daily data as well. There is a moderate anticorrelation between the original values of precipitation and divergence in the area of  $-100^{\circ}\text{W}$  and  $-110^{\circ}\text{W}$ . The only moderated anticorrelation can be explained due to the time and spatial delay between divergence and precipitation. Air masses have to arise first (divergence) before they start to condensate. The anticorrelation gets a bit higher, when I exclude data points with zero precipitation. I found a strong anticorrelation when taking the longitudinal mean and an even higher anticorrelation when taking longitudinal mean and 7 days mean filter over time. For the exact results please refer to table 1 and Figure 6. The values for anticorrelation are lower for areas with larger landmasses or areas which are closer to landmasses. An overview of the anticorrelation values for an area closer to landmasses can be found in the appendix.

After calculating mid-latitudes I looked at the correlation of the calculated latitudes for precipitation and divergence. I found strong correlation between divergence and precipitation latitudes and very strong correlation between the two precipitation indexes. For exact results please refer to table 2 and Figure 7. Again I got lower results for areas nearer to landmasses or with larger landmasses included, so I stucked to areas with mostly sea and restricted the prediction area on 10 longitudes between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$ . An overview of the correlation between latitudes for an area closer to landmasses can be found in the appendix.

Table 1: Pearson correlation values for divergence and precipitation for different data configurations between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$ .

	Pearson correlation divergence/ precipitation
Original data	-0.363
Precipitation $> 0$	-0.372
Precipitation $> 0.001$	-0.352
Precipitation $> 0.002$	-0.304
Longitudinal mean	-0.601
Longitudinal mean and 7 days mean filter time	-0.735

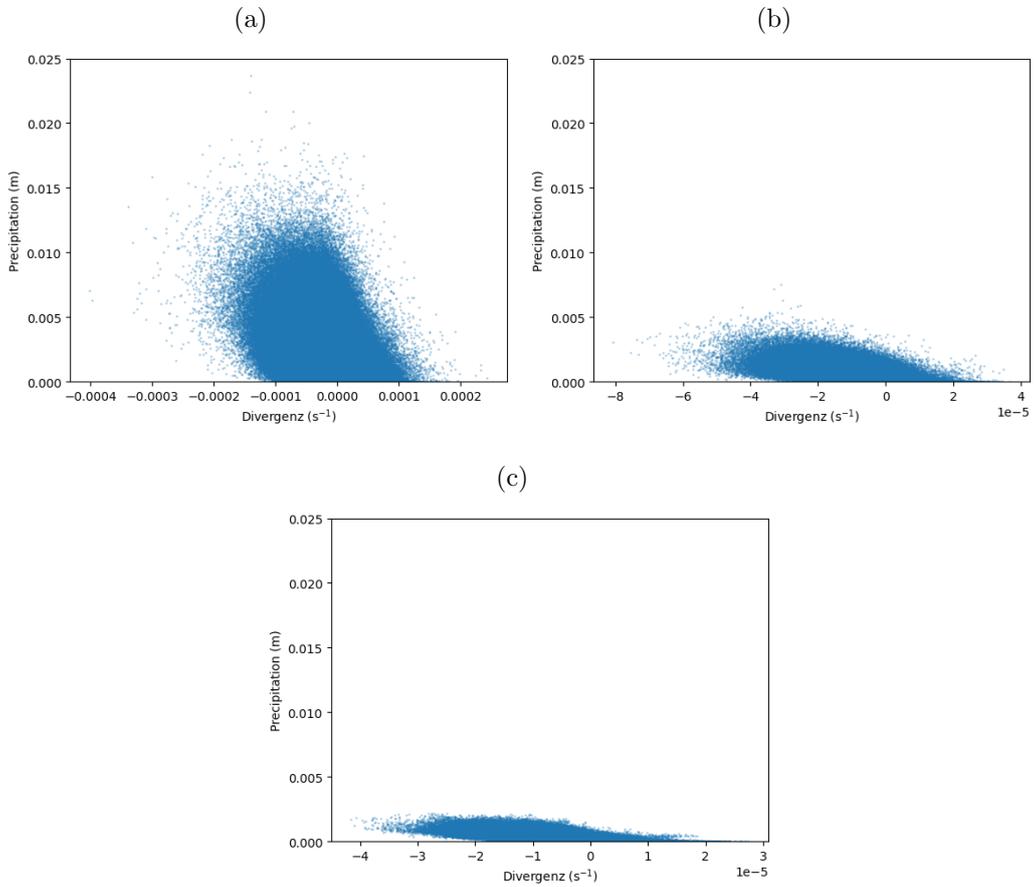


Figure 6: Daily divergence and precipitation values between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  for a) the original dataset, b) the dataset after taking the mean over longitudes and c) the dataset after taking the mean over longitude dimension and taking a mean filter of size 7 over time dimension.

Table 2: Pearson correlation values for latitudes of maximum precipitation as given by Equation (3), maximum precipitation as given by Equation (2) and minimum divergence as given by Equation (4) between  $-100^{\circ}\text{W}$  and  $-110^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$ .

	Adam index	Maximum precipitation	Minimum divergence
Adam index	1	0.941	0.709
Maximum precipitation	0.941	1	0.701
Minimum divergence	0.709	0.701	1

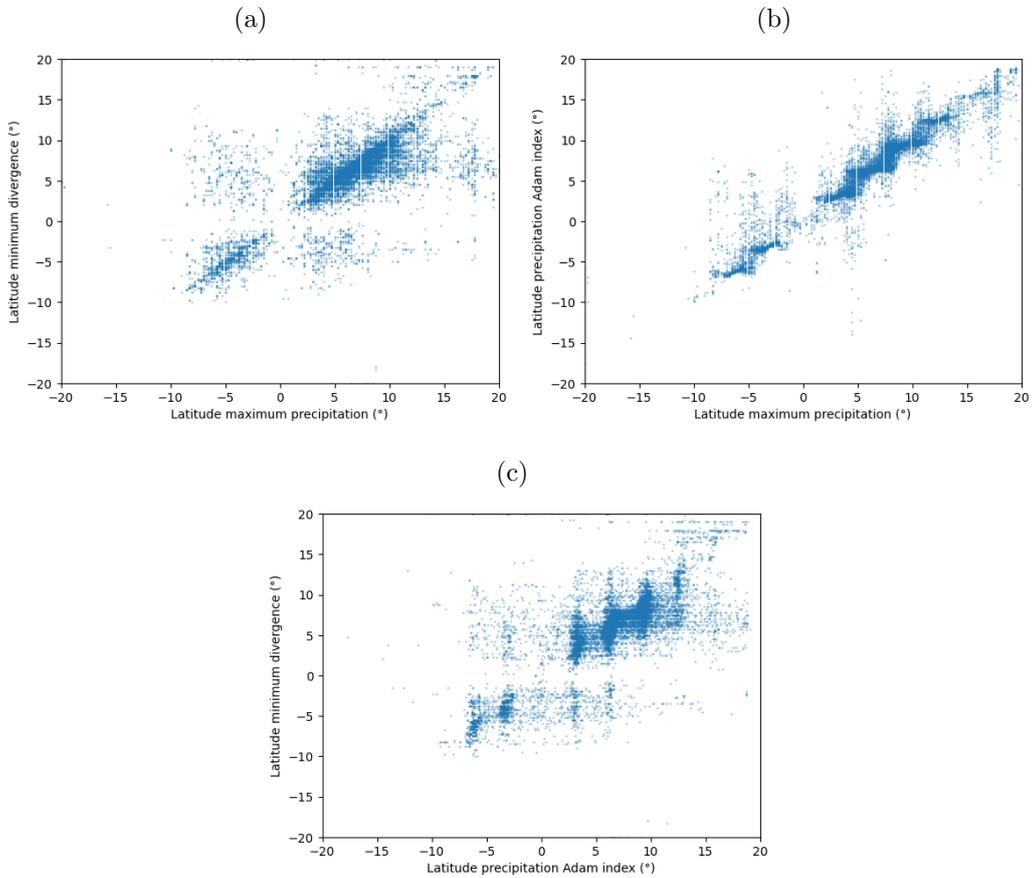


Figure 7: ITCZ mid-latitudes comparison between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  for a) maximum precipitation and minimum divergence, b) maximum precipitation and maximum precipitation calculated by Equation (3) and c) maximum precipitation and minimum divergence.

## 5.2 Predicting ITCZ mid-location using LSTMs

### 5.2.1 Single day predictions

The single day predictions gave good results for all three LSTMs. I got good results for training 30 epochs with a batchsize of 64 for predicting the latitudes of minimum divergence. The results of the prediction and a comparison of the predicted and observed latitudes can be found in Figure 8. The Pearson correlation between the latitude of observed data and the latitude of predicted data was 0.887. I got even better results for the maximum precipitation as given by the Adam index. Here I got a Pearson correlation between the latitude of observed data and the latitude of predicted data of 0.936. For the LSTM trained on the latitude of maximum precipitation as given by Equation (3) I got an Pearson correlation between the latitude of observed data and the latitude of predicted data of 0.905. The results for the single day predictions for precipitation can be found in the appendix.

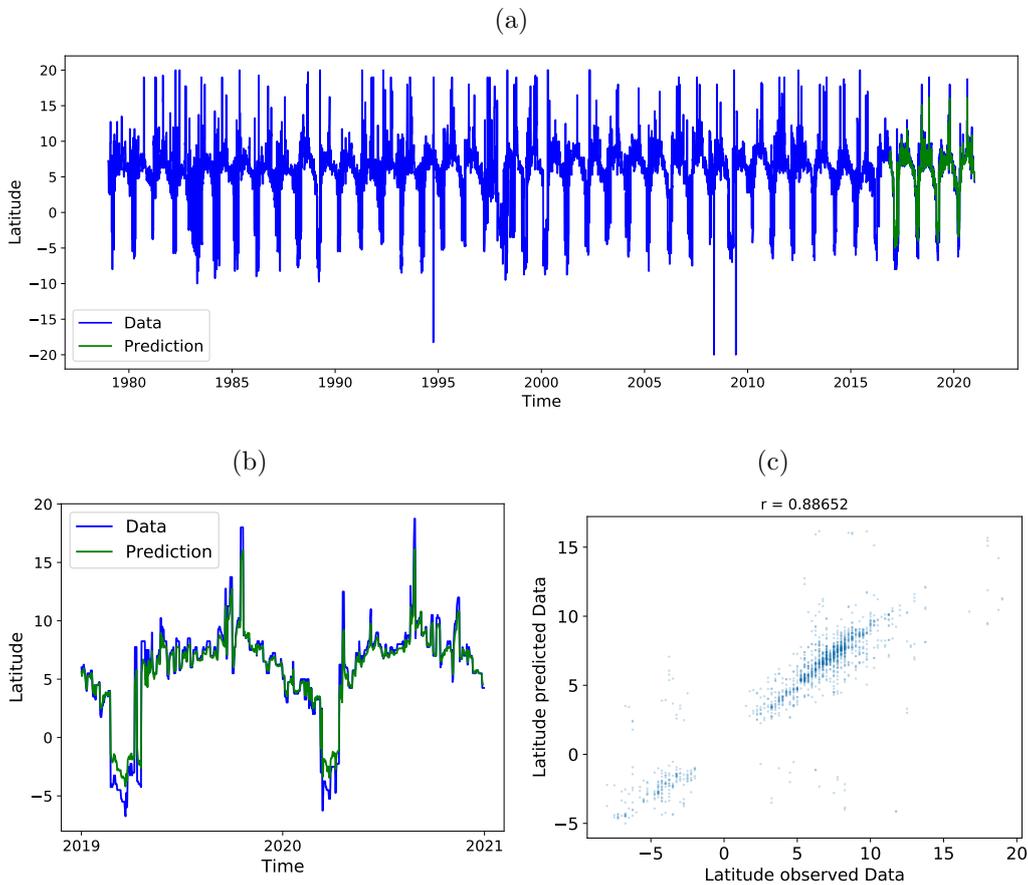


Figure 8: Latitude of minimum divergence predictions for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  trained 30 epochs with a batchsize of 64. a) shows the training data and the predictions for the time span of 40 years from 1979 until 2020. b) Shows the original data and the prediction for 2 years from 2019 until 2021. c) Shows the correlation between the latitude of observed data and the latitude of predicted data for the complete area and time used for predictions.

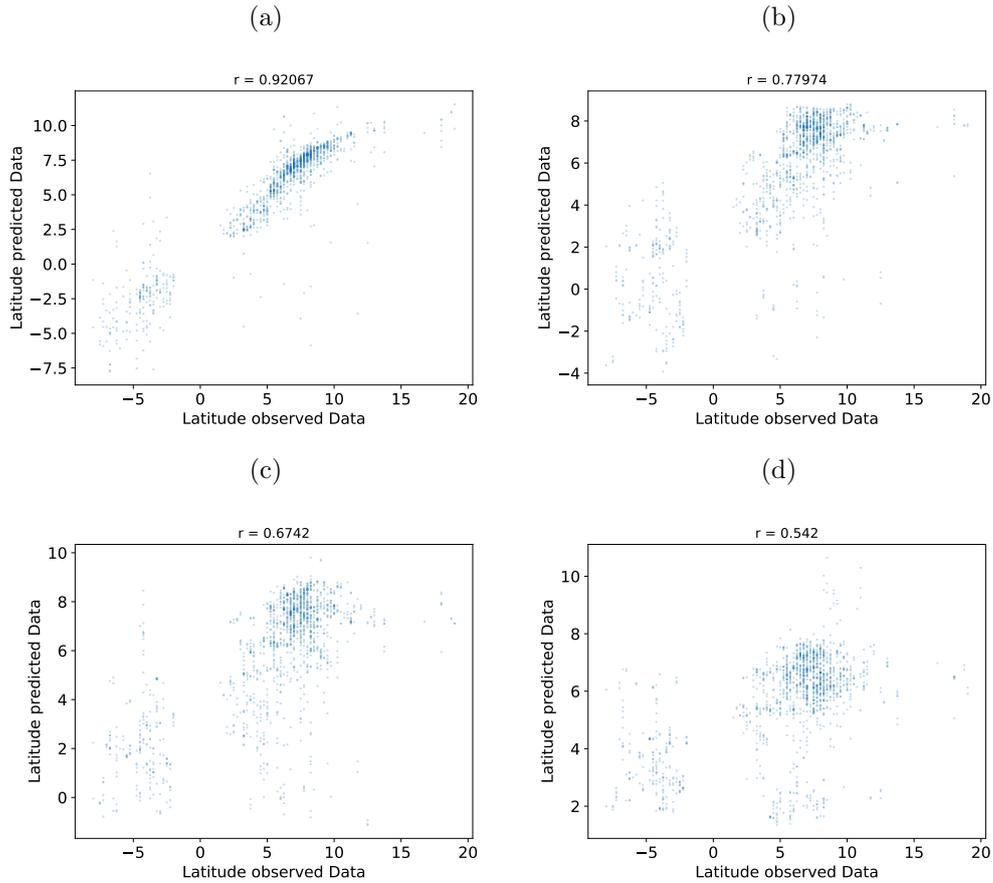


Figure 9: Latitude of minimum divergence predictions for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different timesteps ahead predictions trained 400 epochs with a batchsize of 4.

### 5.2.2 Multi day predictions

For the multi day predictions I trained again three LSTMs on maximum precipitation latitudes and minimum divergence latitudes. I trained the LSTMs to achieve good results for 2 months ahead predictions (60 days). Compared to the single day predictions the results dropped with longer predictions ahead. For the LSTM trained on the latitude of minimum divergence I got the best results for training 200 epochs with a batchsize of 4. The 60 days ahead predictions still got a Pearson correlation between latitudes of observed data and latitudes of predicted data of 0.533. The Pearson correlation for one day ahead predictions has been 0.923, 15 days ahead it already dropped to 0.757 and 30 days ahead it went down to 0.649. The prediction results for divergence for different days ahead can be found in Figure 9 and Figure 10. I got similar results for the two other proxies for the latitude of maximum precipitation. The corresponding results can be found in the appendix. For the latitude of maximum precipitation as given by Equation (3) I got a Pearson correlation of 0.966 for predicting one day ahead, for 15 days ahead it dropped to 0.795, for 30 days ahead it went further down to 0.716 and for 60 days it was 0.547. For the latitude of maximum precipitation as given by Equation (2) I got a

Pearson correlation of 0.938 for predicting one day ahead. For 15 days ahead I got a correlation of 0.769, for 30 days ahead it dropped further down to 0.718 and for 60 days ahead predictions Pearson correlation was 0.572. The corresponding results can be found in the appendix.

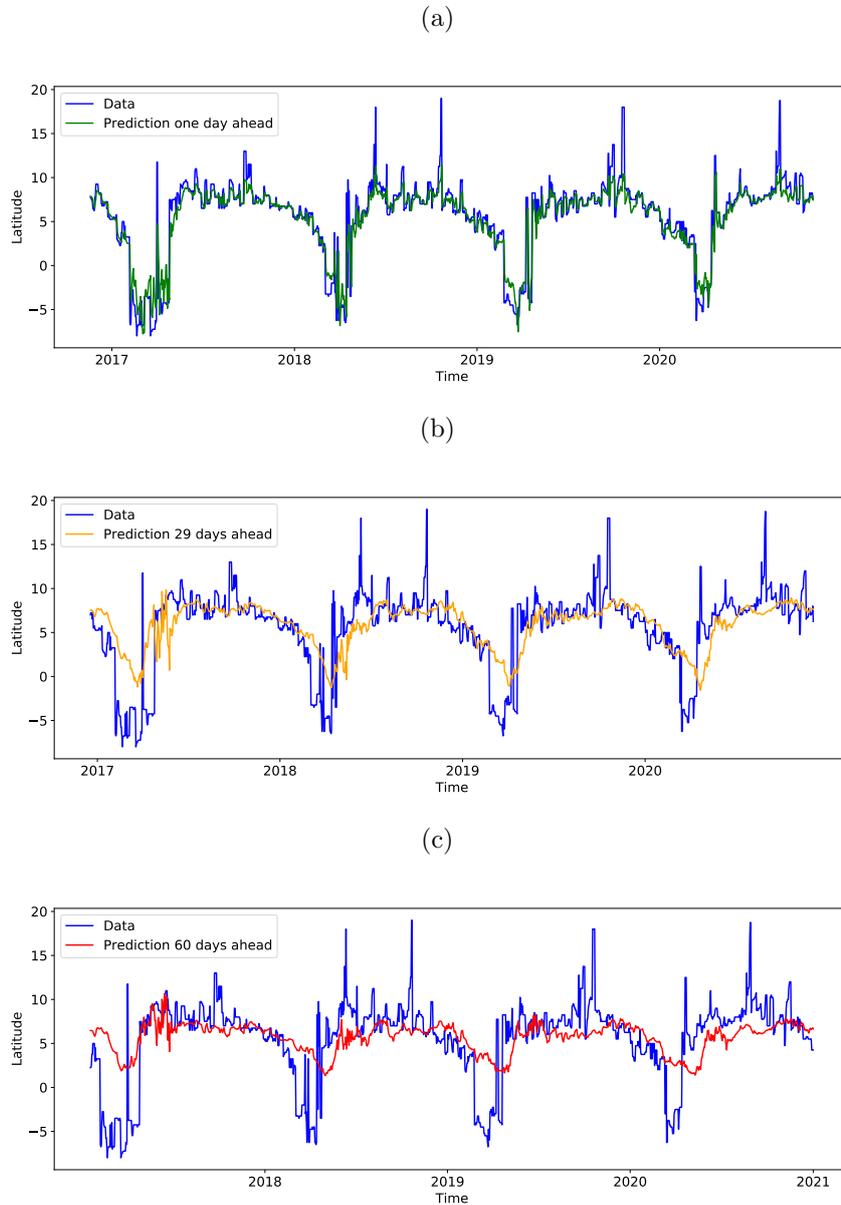


Figure 10: Latitude of minimum divergence predictions for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different timesteps ahead predictions trained 400 epochs with a batchsize of 4.

### 5.3 Predicting values over time using Conv LSTMs

For predicting the actual divergence values over time I trained a ConvLSTM to predict four frames. I used 7-day mean filtered data for the time dimension. I got acceptable first results for the graphical predictions with a good correlation

between true and predicted values. Figure 11 shows the graphical results for four days predictions compared to the original values. The correlation for the different frames can be found in Figure 12. The ConvLSTM was able to capture similar structures compared to the real divergence belt in the ITCZ however the width and the exact shapes are not yet correctly captured. Interestingly I got a better correlation between the results for four frame ahead predictions compared to one frame ahead predictions. Future research should certainly take another look on this and continue to work on this approach and to try different architectures to get more accurate graphical results.

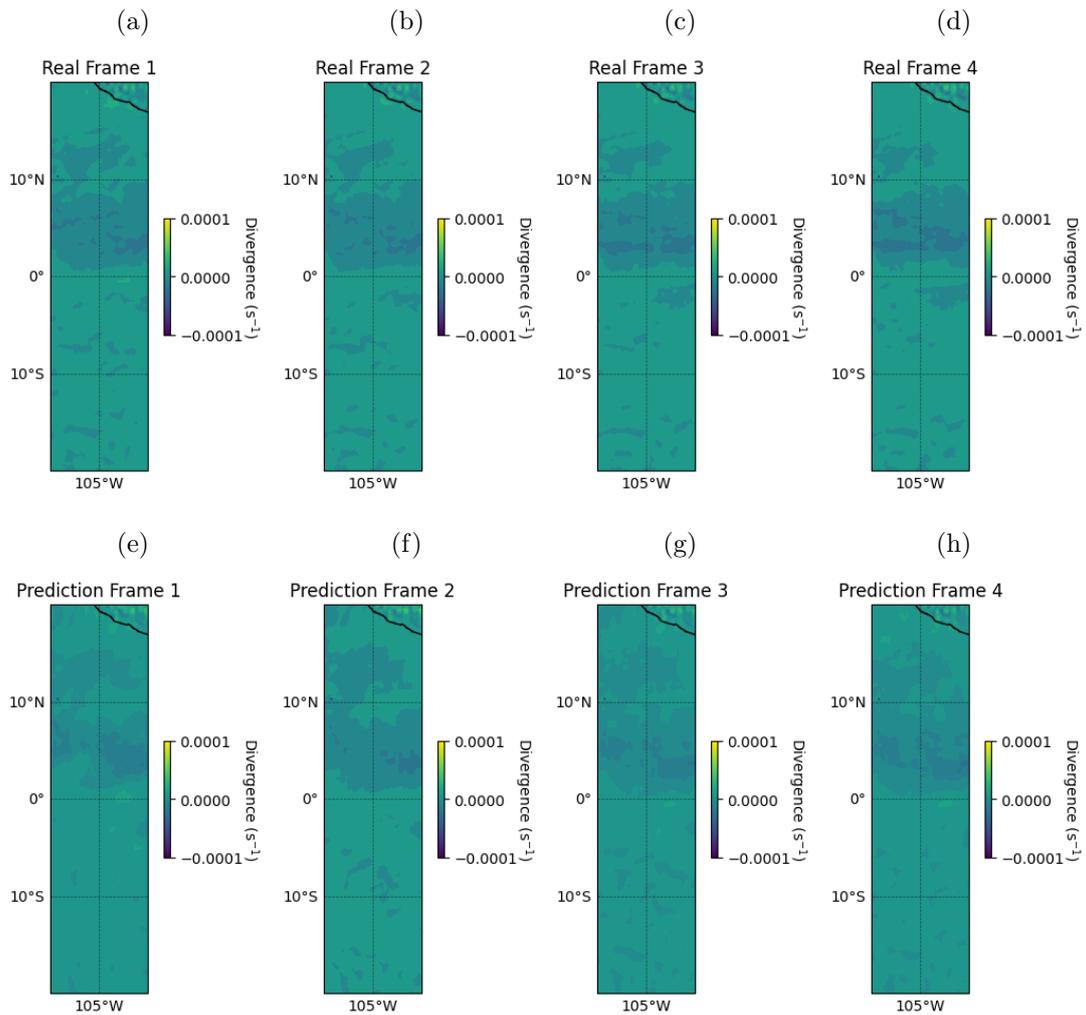


Figure 11: Visual results for the ConvLSTM trained on divergence for frames based on data beginning 1st January 2011 until 4th January 2011 a) - d) show the real values for that time period, e)-h) show the predicted values for the same time period.

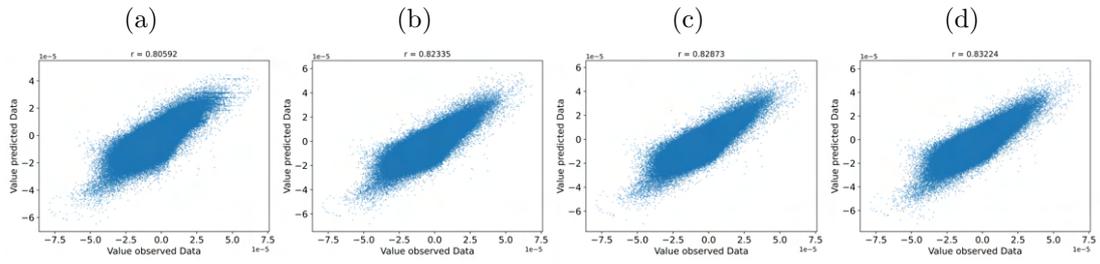


Figure 12: Correlation results for divergence for the ConvLSTM for different frames based on data beginning 1st January 2011 until 4th January 2011. a) shows results for frame 1, b) for frame 2, c) for frame 3 and d) for frame 4.

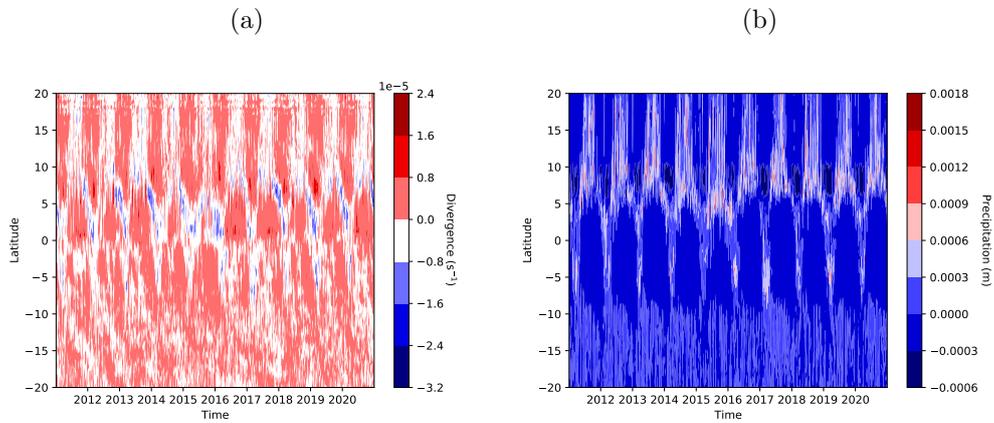


Figure 13: Anomaly map for a) divergence and b) precipitation for the longitudinal mean in the area  $-110^{\circ}\text{W}$  to  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for the years 2011 to 2020.

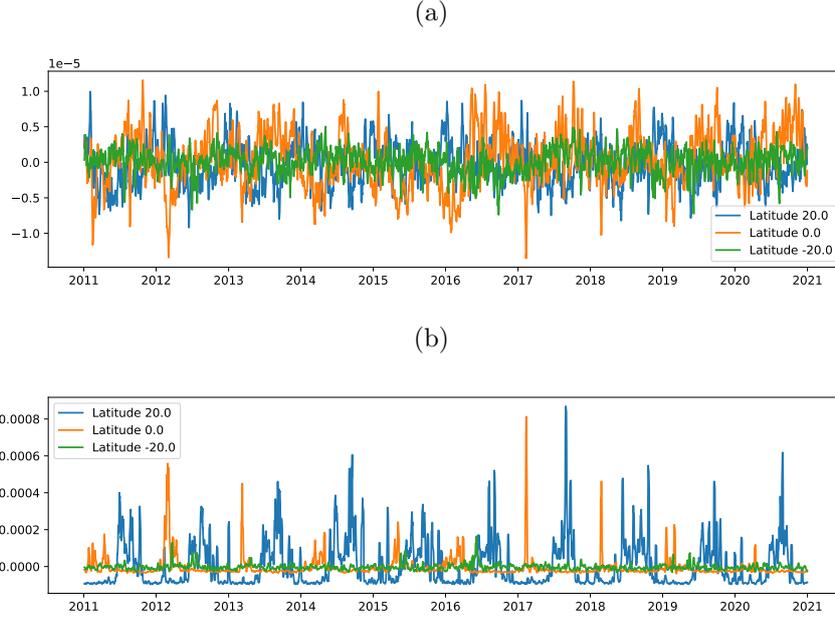


Figure 14: Anomaly timeseries for a) divergence and b) precipitation for the longitudinal mean in the area  $-110^{\circ}\text{W}$  to  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for the years 2011 to 2020.

## 6 Discussion

In this thesis I proposed a method to calculate the latitudinal ITCZ mid-location based on divergence. Therefore, I used this definition to predict the time series of ITCZ mid-location to predict shifts in the ITCZ. Furthermore, I also used the original time series of divergence values to predict the actual values over time using ConvLSTMs to predict the expansion of the ITCZ graphically.

I evaluated the anticorrelation between precipitation and divergence as proxies to calculate the latitudinal ITCZ mid-location. I used the found anticorrelation to calculate the latitudinal mid-location based on different definitions on maximum precipitation and on minimum divergence and compared the results. The calculated latitudes showed good correlation so I used those latitudes to train different LSTMs. I got good results for single day predictions for all three ITCZ mid-location proxies. I was also able to predict trends in the ITCZ mid-location with multi day predictions up to 60 days ahead. I then applied a ConvLSTM on the actual divergence and precipitation values which gave me promising first results.

I was able to show that the predictions of the ITCZ mid-latitude works good for small time horizons up to 30 days but accuracy drops with longer predictions ahead. The 60 days prediction was only partially accurate. Here more work on improving the network or even combining ITCZ mid-location proxies should be done. Wind and rain are also highly variable and fluctuating on daily basis. For trend predictions of the ITCZ daily data might not be needed so it might be better to use data on weekly or even monthly level. The divergence index also might become more stable with weight-

ing it to an area similar to the calculations done by Adam et al. for precipitation as shown in Equation (3) (Adam et al., 2016a). Also there are other proxies for ITCZ mid-location as well as width which seems more stable. Schneider et al. showed that using energy budgets and defining the ITCZ mid-latitude as the energy flux equator is a convenient and stable proxy for ITCZ (Schneider et al., 2014). Comparing the divergence approach to these proxies instead of precipitation might also be interesting.

For further generalization my approach also needs to be applied to larger areas. I only tested in areas without any larger landmasses like coastlines or islands. I only tested in the Pacific region but especially when it comes to the connection of ITCZ and monsoon it would be interesting to also test in Western Atlantic or Indian Ocean.

Further work should also be done on an anomaly dataset of the ITCZ mid-location. The anomalies for the divergence and precipitation for the last 10 years are shown in Figure 13 and Figure 14. The anomalies seem to show regular patterns especially in the area of the ITCZ. An anomaly evaluation could be used to find and predict anomaly patterns in the ITCZ for example in connection to El Niño-Southern Oscillation (ENSO). This could also be used to investigate the response of the ITCZ to a warming climate.

The ITCZ has been reported to become narrower over the recent decades and climate models also predict a further narrowing of the ITCZ (Byrne et al., 2018). I mainly focused on the ITCZ mid-location but also the ITCZ width and related the ITCZ strength play a big role for hydroclimates in tropical regions (Byrne and Schneider, 2016b). It can be observed that a narrowing and strengthening of the ITCZ leads to a strengthening in precipitation in the ITCZ region as well as to less humidity in the subtropics (Byrne and Schneider, 2016b; Byrne et al., 2018; Hohenegger and Jakob, 2020). With the ConvLSTM approach I tried to give a graphical prediction of the ITCZ band over time. Predicting the LSTM using different spatio-temporal artificial neural network architectures with different ITCZ proxies might bring insights in the future structure of the ITCZ.

Extending the ConvLSTM to different heights might also bring more insights on characteristics of the ITCZ. I tried predicting the ITCZ with a two dimensional ConvLSTM but especially divergence and its relation to the Hadley cell is rather spatially three dimensional on different pressure levels. Future research could extend the ConvLSTM or different spatio-temporal prediction approaches to different pressure levels and to the full Hadley circulation.

## References

- Adam, O., Bischoff, T., and Schneider, T. (2016a). Seasonal and interannual variations of the energy flux equator and itcz. part i: Zonally averaged itcz position. *Journal of Climate*, 29(9):3219–3230.
- Adam, O., Bischoff, T., and Schneider, T. (2016b). Seasonal and interannual variations of the energy flux equator and itcz. part ii: Zonally varying shifts of the itcz. *Journal of Climate*, 29(20):7281–7293.
- Arbuszewski, J. A., Demenocal, P. B., Cl eroux, C., Bradtmiller, L., and Mix, A. (2013). Meridional shifts of the atlantic intertropical convergence zone since the last glacial maximum. *Nature Geoscience*, 6(11):959–962.
- Barry, R. G. and Chorley, R. J. (2009). *Atmosphere, weather and climate*. Routledge.
- Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, 5(2):157–166.
- Berry, G. and Reeder, M. J. (2014). Objective identification of the intertropical convergence zone: Climatology and trends from the era-interim. *Journal of climate*, 27(5):1894–1909.
- Bischoff, T. and Schneider, T. (2014). Energetic constraints on the position of the intertropical convergence zone. *Journal of Climate*, 27(13):4937–4951.
- Broccoli, A. J., Dahl, K. A., and Stouffer, R. J. (2006). Response of the itcz to northern hemisphere cooling. *Geophysical Research Letters*, 33(1).
- Brown, J. R., Moise, A. F., and Colman, R. A. (2013). The south pacific convergence zone in cmip5 simulations of historical and future climate. *Climate dynamics*, 41(7-8):2179–2197.
- Byrne, M. P., Pendergrass, A. G., Rapp, A. D., and Wodzicki, K. R. (2018). Response of the intertropical convergence zone to climate change: Location, width, and strength. *Current climate change reports*, 4(4):355–370.
- Byrne, M. P. and Schneider, T. (2016a). Energetic constraints on the width of the intertropical convergence zone. *Journal of Climate*, 29(13):4709–4721.
- Byrne, M. P. and Schneider, T. (2016b). Narrowing of the itcz in a warming climate: Physical mechanisms. *Geophysical Research Letters*, 43(21):11–350.
- Byrne, M. P. and Schneider, T. (2018). Atmospheric dynamics feedback: Concept, simulations, and climate implications. *Journal of Climate*, 31(8):3249–3264.
- Elman, J. L. (1990). Finding structure in time. *Cognitive science*, 14(2):179–211.
- Gers, F. A., Schmidhuber, J., and Cummins, F. (2000). Learning to forget: Continual prediction with lstm. *Neural computation*, 12(10):2451–2471.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. MIT press.

- Greco, F., Ventrucci, M., and Castelli, E. (2018). P-spline smoothing for spatial data collected worldwide. *Spatial Statistics*, 27:1–17.
- Gu, G., Adler, R. F., and Sobel, A. H. (2005). The eastern pacific itcz during the boreal spring. *Journal of the atmospheric sciences*, 62(4):1157–1174.
- Hadley, G. (1735). Vi. concerning the cause of the general trade-winds. *Philosophical Transactions of the Royal Society of London*, 39(437):58–62.
- Haug, G. H., Hughen, K. A., Sigman, D. M., Peterson, L. C., and Röhl, U. (2001). Southward migration of the intertropical convergence zone through the holocene. *Science*, 293(5533):1304–1308.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- Hohenegger, C. and Jakob, C. (2020). A relationship between itcz organization and subtropical humidity. *Geophysical Research Letters*, 47(16):e2020GL088515.
- Holton, J. R. and Hakim, G. J. (2013). Chapter 11 - tropical dynamics. In Holton, J. R. and Hakim, G. J., editors, *An Introduction to Dynamic Meteorology (Fifth Edition)*, pages 377–411. Academic Press, Boston, fifth edition edition.
- Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR.
- Kang, S. M., Shin, Y., and Xie, S.-P. (2018). Extratropical forcing and tropical rainfall distribution: Energetics framework and ocean ekman advection. *npj Climate and Atmospheric Science*, 1(1):1–10.
- Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551.
- Narsey, S., Brown, J., Colman, R., Delage, F., Power, S., Moise, A., and Zhang, H. (2020). Climate change projections for the australian monsoon from cmip6 models. *Geophysical Research Letters*, 47(13):e2019GL086816.
- Oort, A. H. and Yienger, J. J. (1996). Observed interannual variability in the hadley circulation and its connection to enso. *Journal of Climate*, pages 2751–2767.
- Philander, S., Gu, D., Lambert, G., Li, T., Halpern, D., Lau, N., and Pacanowski, R. (1996). Why the itcz is mostly north of the equator. *Journal of climate*, 9(12):2958–2972.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088):533–536.

- Schneider, T., Bischoff, T., and Haug, G. H. (2014). Migrations and dynamics of the intertropical convergence zone. *Nature*, 513(7516):45–53.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958.
- Su, H., Jiang, J. H., Neelin, J. D., Shen, T. J., Zhai, C., Yue, Q., Wang, Z., Huang, L., Choi, Y.-S., Stephens, G. L., et al. (2017). Tightening of tropical ascent and high clouds key to precipitation change in a warmer climate. *Nature communications*, 8(1):1–9.
- Waliser, D. E. and Gautier, C. (1993). A satellite-derived climatology of the itcz. *Journal of climate*, 6(11):2162–2174.
- Watson, A. I. and Blanchard, D. O. (1984). The relationship between total area divergence and convective precipitation in south florida. *Monthly weather review*, 112(4):673–685.
- Xiang, B., Zhao, M., Held, I. M., and Golaz, J.-C. (2017). Predicting the severity of spurious “double itcz” problem in cmip5 coupled models from amip simulations. *Geophysical Research Letters*, 44(3):1520–1527.
- Xingjian, S., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015). Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810.
- Yancheva, G., Nowaczyk, N. R., Mingram, J., Dulski, P., Schettler, G., Negendank, J. F., Liu, J., Sigman, D. M., Peterson, L. C., and Haug, G. H. (2007). Influence of the intertropical convergence zone on the east asian monsoon. *Nature*, 445(7123):74–77.
- Žagar, N., Skok, G., and Tribbia, J. (2011). Climatology of the itcz derived from era interim reanalyses. *Journal of Geophysical Research: Atmospheres*, 116(D15).

## 7 Appendix

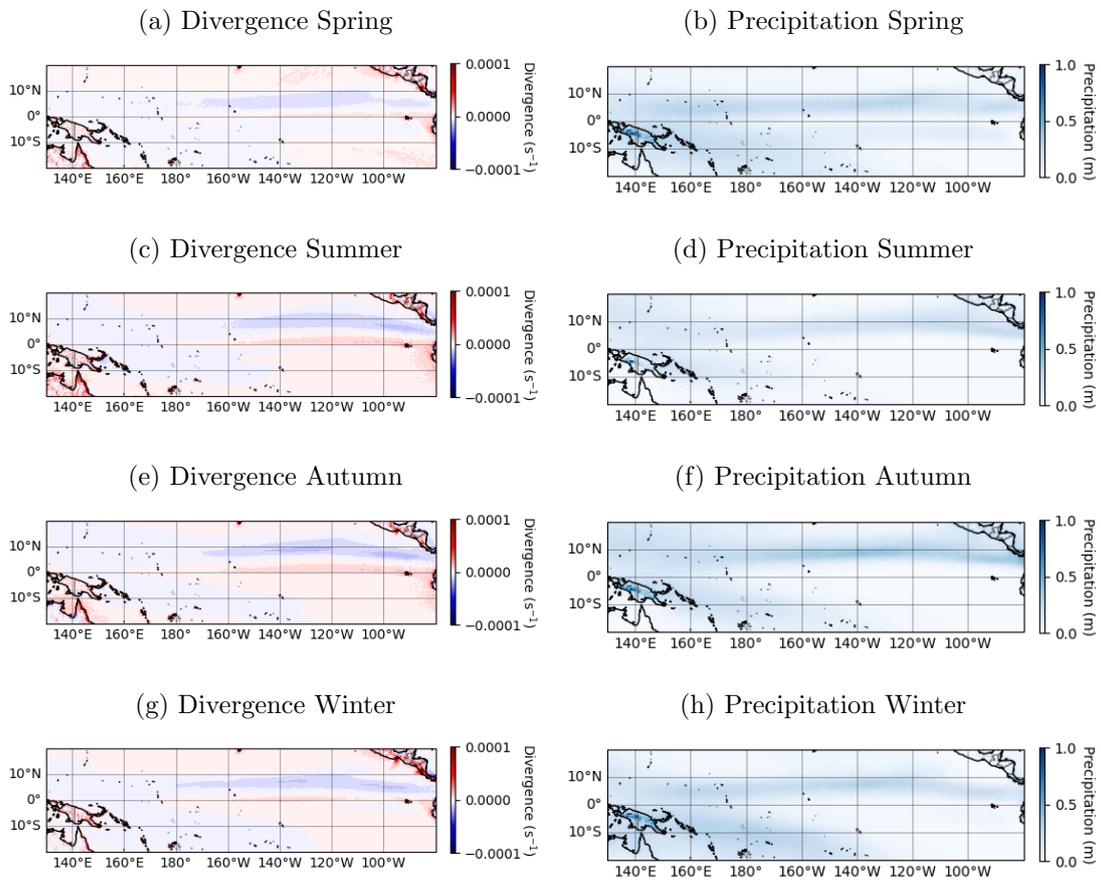


Figure 15: Average seasonal divergence and precipitation averaged over the years 1979-2020.

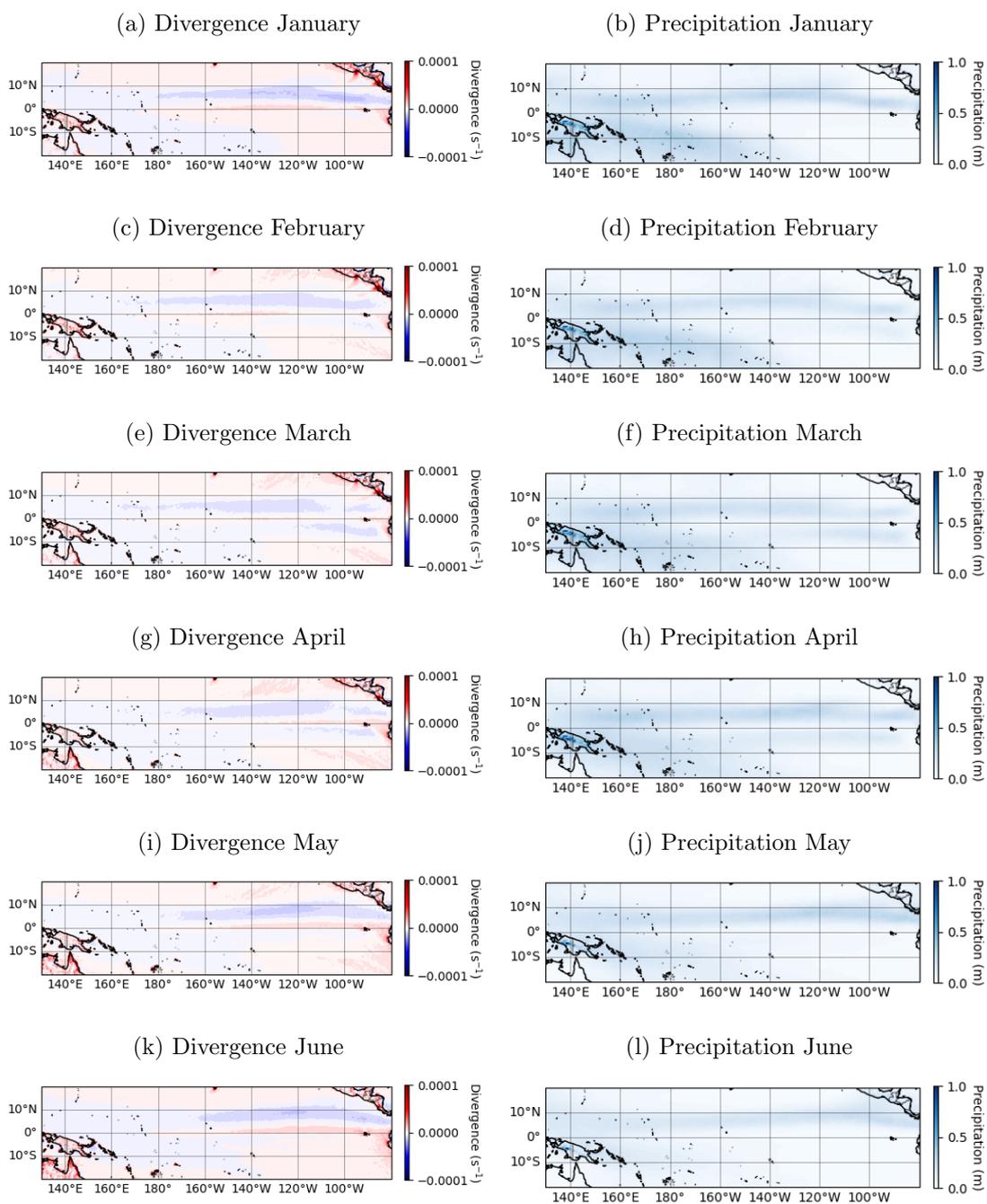


Figure 16: Average monthly divergence and precipitation averaged over the years 1979-2020 for January to June.

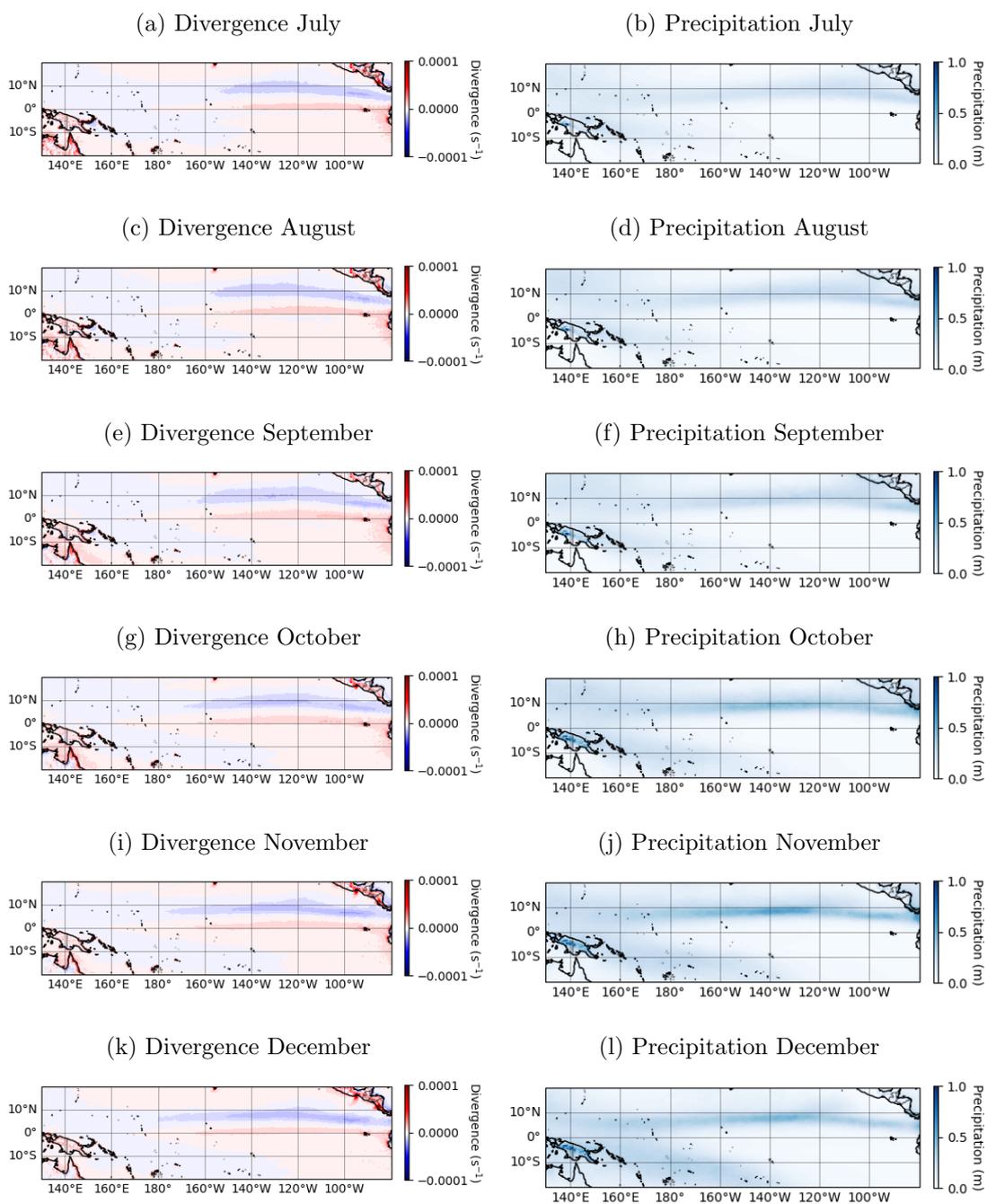


Figure 17: Average monthly divergence and precipitation averaged over the years 1979-2020 for July to December.

Table 3: Pearson correlation values for divergence and precipitation for different data configurations between 160°E and 150°E and -20°S and 20°N.

	Pearson correlation divergence/ precipitation
Original data	-0.216
Precipitation > 0	-0.204
Precipitation > 0.001	-0.275
Precipitation > 0.002	-0.31
Longitudinal mean	-0.482
Longitudinal mean and 7 days mean filter time	-0.651

Table 4: Pearson correlation latitudes for maximum precipitation as given by Adam et al. (Adam et al., 2016a), maximum precipitation and minimum divergence between 160°E and 150°E and -20°S and 20°N.

	Adam index	Maximum precipitation	Minimum divergence
Adam index	1	0.895	0.311
Maximum precipitation	0.895	1	0.311
Minimum divergence	0.311	0.311	1

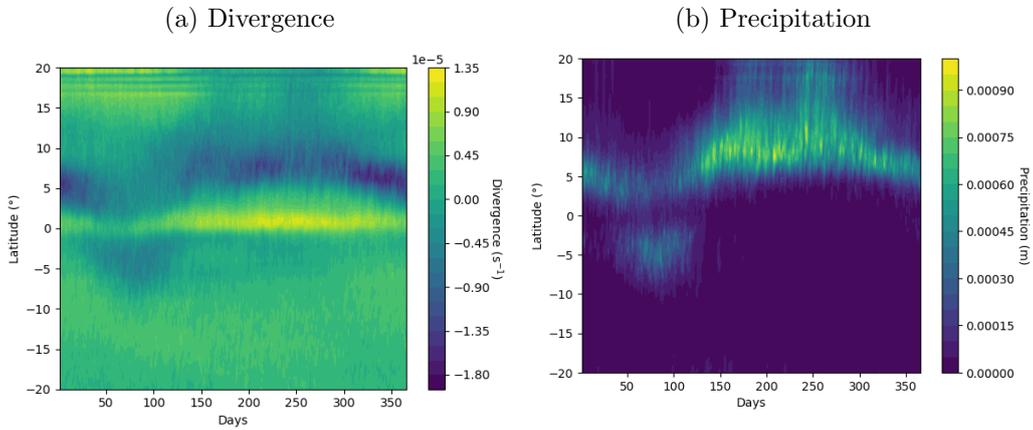


Figure 18: Seasonal migrations of the ITCZ based on a) divergence and b) precipitation. Daily averaged data averaged over the years 1979-2020. Longitudinal average over a window over the Pacific between -110°W and -100°W

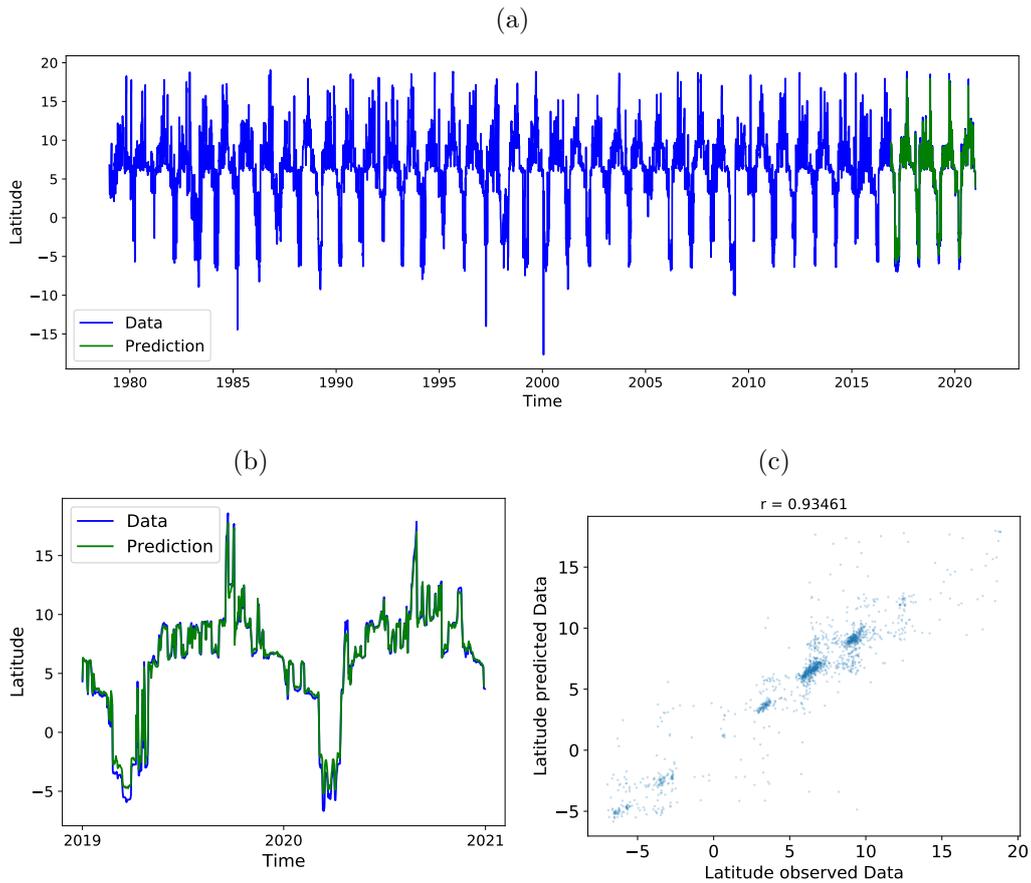


Figure 19: Latitude of maximum precipitation as given by the Adam index predictions for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  trained 20 epochs with a batchsize of 32. a) shows the training data and the predictions for the time span of 40 years from 1979 until 2020. b) Shows the original data and the prediction for 2 years from 2019 until 2021. c) Shows the correlation between the latitude of observed data and the latitude of predicted data for the complete area and time used for predictions.

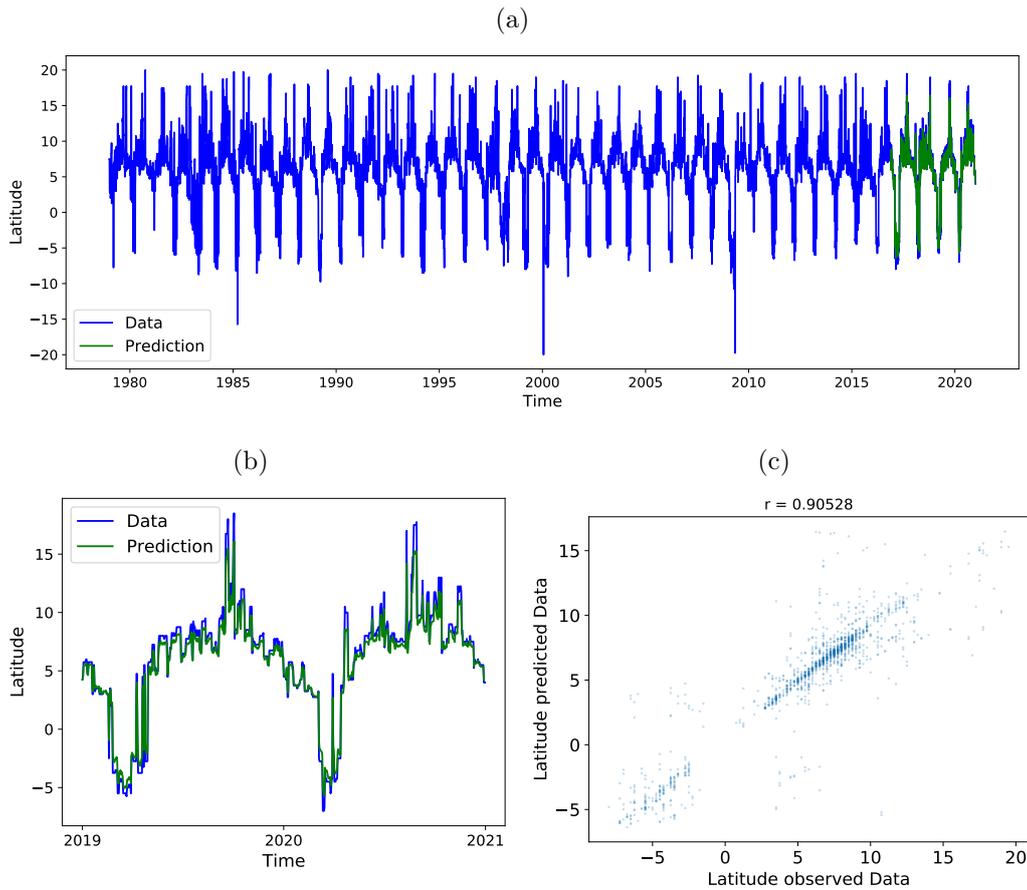


Figure 20: Latitude of maximum precipitation as given by (2) predictions for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  trained 30 epochs with a batchsize of 32. a) shows the training data and the predictions for the time span of 40 years from 1979 until 2020. b) Shows the original data and the prediction for 2 years from 2019 until 2021. c) Shows the correlation between the latitude of observed data and the latitude of predicted data for the complete area and time used for predictions.

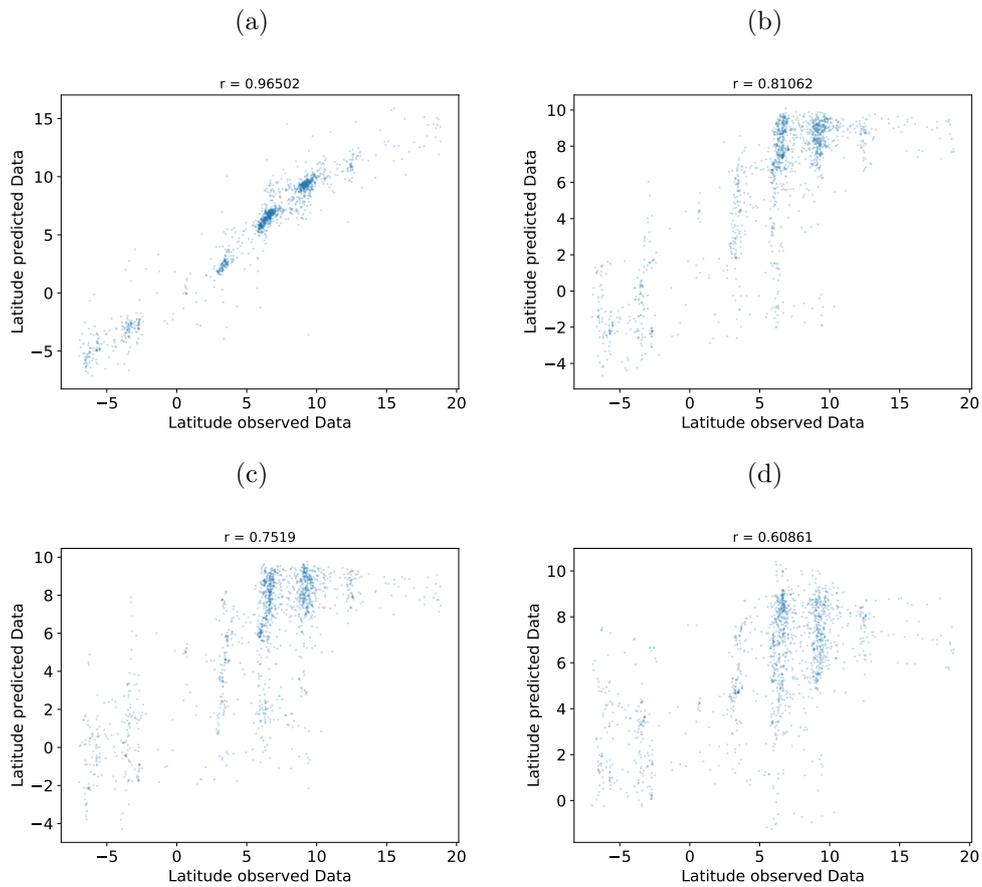


Figure 21: Latitude of maximum precipitation as given by Adam et al. (Adam et al., 2016a), for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different timesteps ahead predictions trained 400 epochs with a batchsize of 4. a) shows the correlation of the latitudes of observed data and the predicted data for the first day ahead prediction. b) shows the correlation of the latitudes of observed data and the predicted data for 15 days c) for 45 days and d)for 60 days ahead prediction.

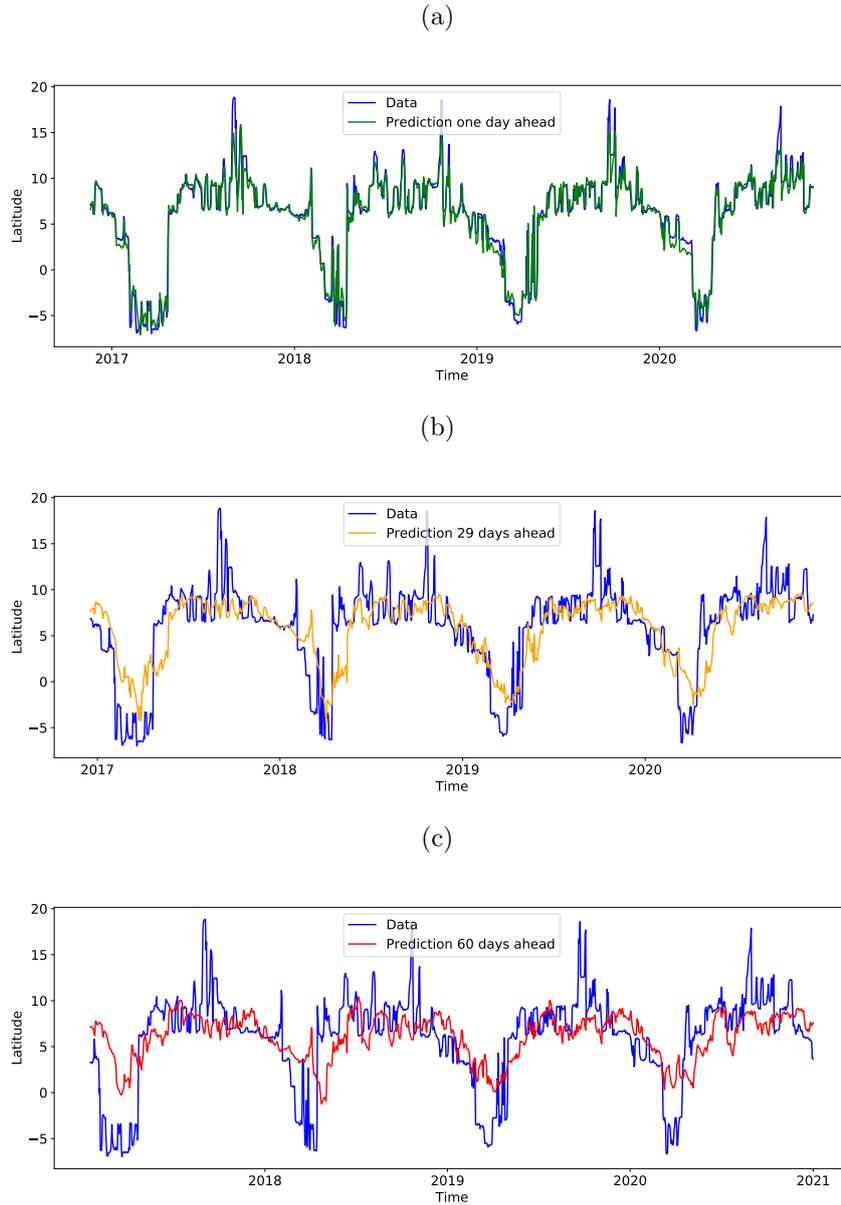


Figure 22: Latitude of maximum precipitation as given by Adam et al. (Adam et al., 2016a), for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different timesteps ahead predictions trained 400 epochs with a batchsize of 4. a) shows the predicted latitudes one day ahead, b) shows the predicted latitudes 30 days ahead and c) shows the predicted latitudes 60 days ahead.

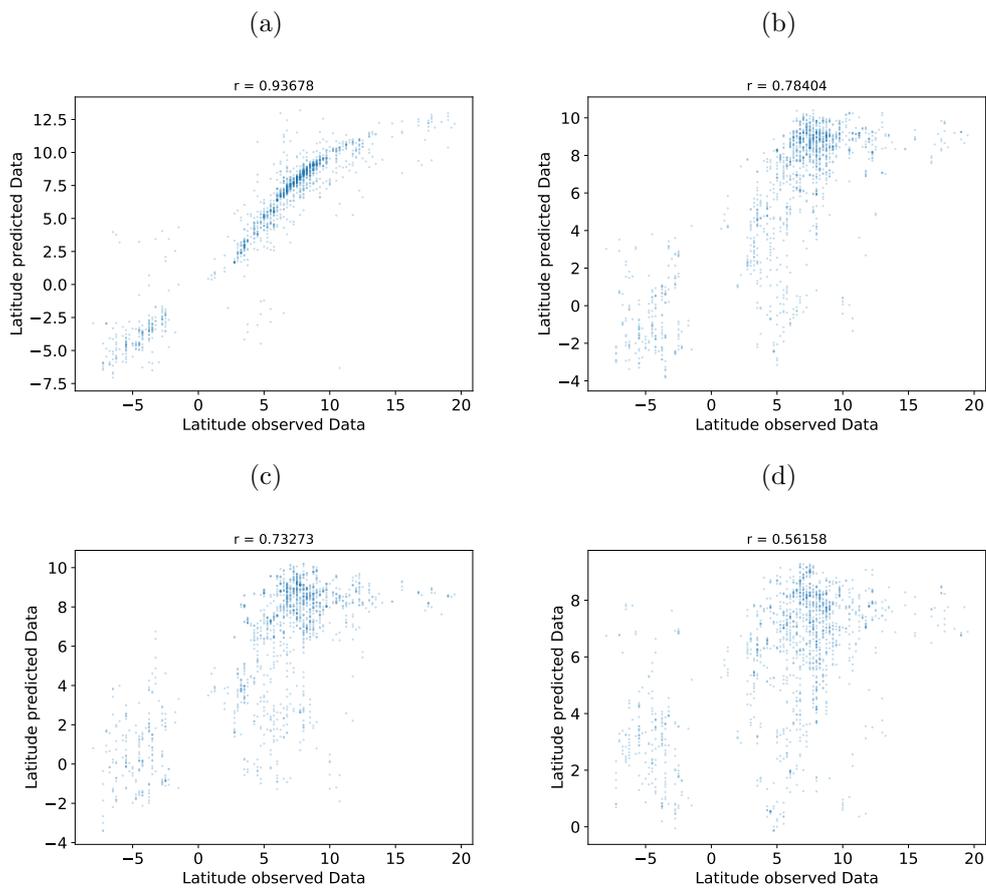


Figure 23: Latitude of maximum precipitation for area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different time steps ahead predictions trained 300 epochs with a batchsize of 4. a) shows the correlation of the latitudes of observed data and the predicted data for the first day ahead prediction. b) shows the correlation of the latitudes of observed data and the predicted data for 15 days c) for 45 days and d) for 60 days ahead prediction.

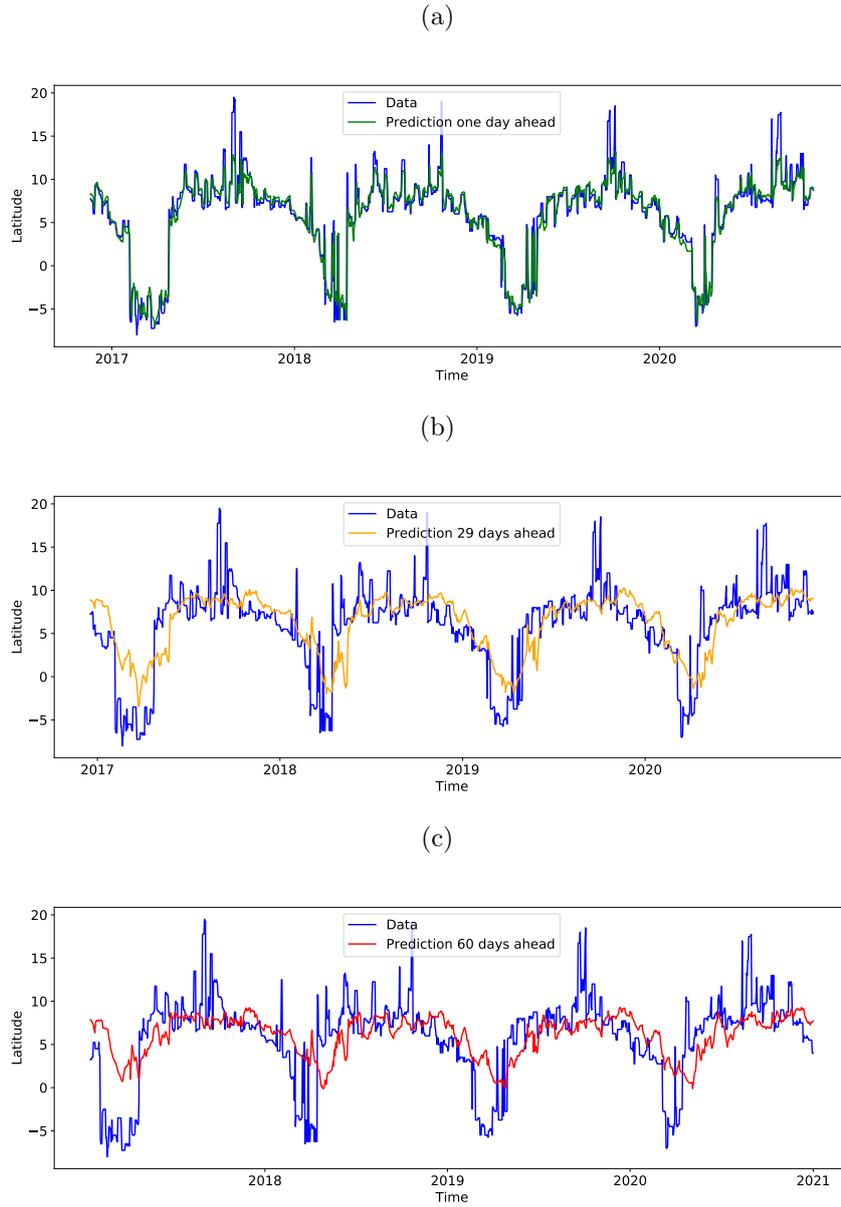


Figure 24: Latitude of maximum precipitation as given by Equation (2) for an area between  $-110^{\circ}\text{W}$  and  $-100^{\circ}\text{W}$  and  $-20^{\circ}\text{S}$  and  $20^{\circ}\text{N}$  for different time steps ahead predictions trained 300 epochs with a batchsize of 4. a) shows the predicted latitudes one day ahead, b) shows the predicted latitudes 30 days ahead and c) shows the predicted latitudes 60 days ahead.