

# Neuronalen Netzwerke in der Klimamodellierung zur Prognose eines Zeitschritts

## Vorhersage der 2m-Temperatur

Tobias Machnitzki<sup>1</sup> und Finn Burgemeister<sup>1</sup>

<sup>1</sup>Meteorologisches Institut  
Universität Hamburg

23. Mai 2018

# Übersicht

- 1 Projektplanung
  - Ziele und Methodik
  - Tools
  - Struktur
  
- 2 Lorenz
  - Datengrundlage
  - Neuronales Netzwerk
  - Ergebnisse
  
- 3 MPI-ESM-LR
  - Datengrundlage
  - Neuronales Netzwerk
  - Optimierungsansätze
  - Ergebnisse
  
- 4 Zusammenfassung

# Ziele und Methodik

## Zielfragestellung

Ist es möglich Klimamodelle mit Hilfe neuronaler Netzwerke hinreichend genau zu approximieren?

## Ziele

- Prognose einer Feldvariable eines Klimamodells (zum Beispiel 2m-Temperatur)
- Einbezug weiterer Features (zum Beispiel mehrerer Ensembles)

## Methodik

- 1 Auswahl eines neuronalen Netzwerks
- 2 Verstehen des ausgewählten neuronalen Netzwerks anhand eines einfachen Modells
- 3 Anwendung auf ein Ensemble einer Variable eines Klimamodells
- 4 Anwendung auf mehrere Ensembles einer Variable eines Klimamodells
- 5 Optimierung

# Tools

- Python
- Keras - The Python Deep Learning library
- Tensorflow - An open source machine learning framework for everyone
- scikit-learn - Machine Learning in Python
- Testsystem „Breeze“:
  - 64 x Intel Xeon @ 2.7 Ghz
  - 256 GB Arbeitsspeicher

# Struktur

- Github: [https://github.com/fiburg/bdp\\_cnn](https://github.com/fiburg/bdp_cnn)
- Objektorientierte Programmierung - vier Klassen:

## Model

Neuronales Netzwerk  
mit Methoden zur  
Initialisierung und  
Preparation der  
Daten

## DataHandler

Aus- und Eingabe  
der Modellergebnisse  
in NETCDF

## Evaluater

Evaluation der  
Modellergebnisse

## Scaler

Skalierung der Daten

# Lorenzmodell

## Lorenz 96<sup>1</sup>

$$\frac{dx_i}{dt} = (x_{i+1} - x_{i-2})x_{i-1} - x_i + F$$

- Lösung der Differenzialgleichungen
- Idealisierung eines hydrodynamischen Systems <sup>2</sup>
- Python Code von Tobias Finn

---

<sup>1</sup>E. Lorenz (1996): Predictability – A problem partly solved. Seminar on Predictability, Vol. I, ECMWF.

<sup>2</sup>A. Karimi, M.R. Paul (2010): Extensive Chaos in the Lorenz-96 Model

# Lorenzmodell

$F = 8$

$i = 1 \dots 40$

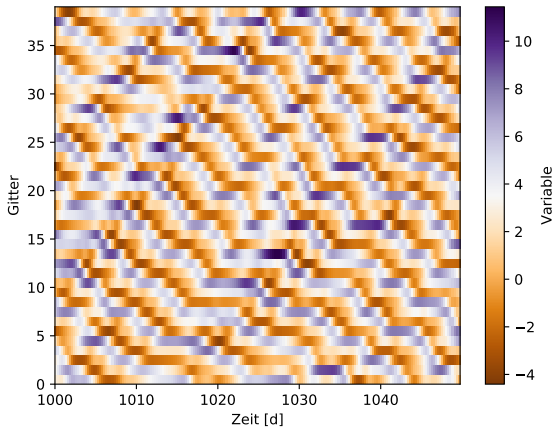


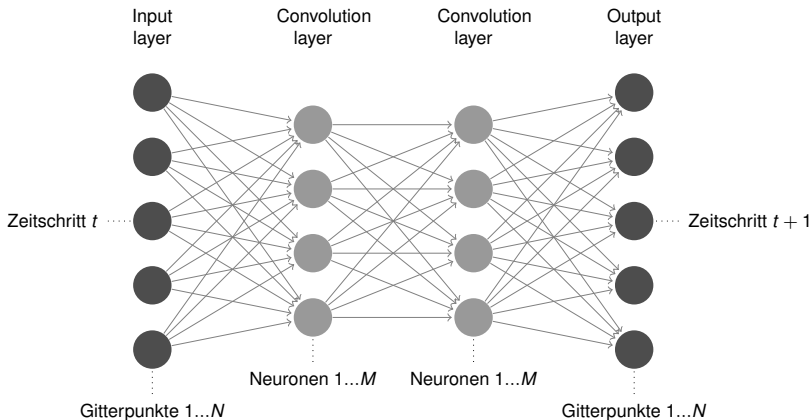
Abbildung: Beispielergebnis mit dem Lorenz Modell.

# Trainingsdatenerzeugung

- 100 Jahre Simulation
- Output alle 6h  
⇒ 146.000 Zeitschritte
- 2/3 Trainingsdaten, 1/6 Validierung, 1/6 Testdaten  
⇒ 97.000 Trainingszeitschritte

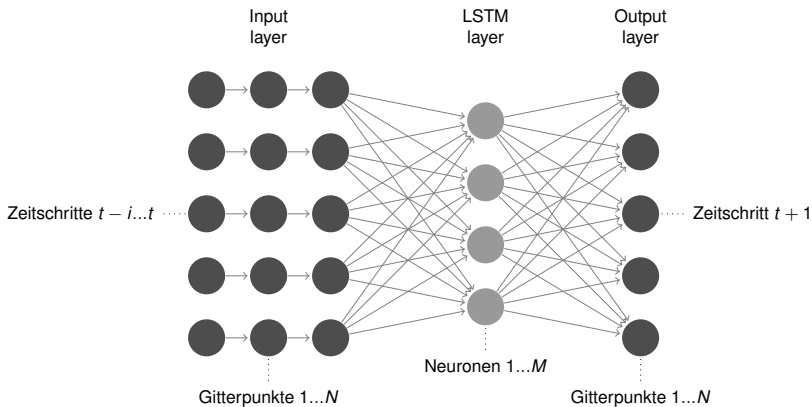


# Convolutional Neural Network (CNN)



**Abbildung:** CNN mit zwei „Convolution Layer“. Für die Erhaltung der Gittergröße wurden zyklische Randbedingungen angenommen (Padding, Reflection)

# Long short-term memory (LSTM)

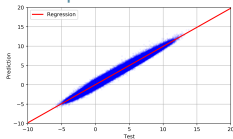


**Abbildung:** LSTM mit einem „Long Short-Term Memory layer“.

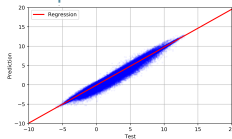
# Ergebnisse

## CNN

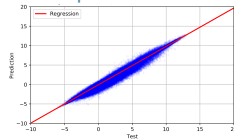
1 Epoche



5 Epochen

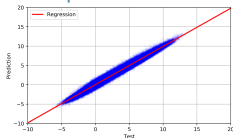


10 Epochen

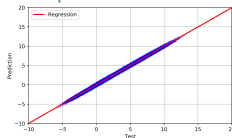


## LSTM

1 Epoche



5 Epochen



10 Epochen

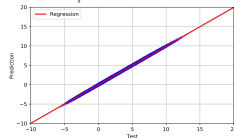


Abbildung: Vergleich zwischen CNN und LSTM.

# Datengrundlage

- Max-Planck-Institute Earth System Model (MPI-ESM-LR) aus Coupled Model Intercomparison Project phase 5 (CMIP5)<sup>3</sup>
- idealisiertes Experiment mit einer jährlichen Erhöhung der CO<sub>2</sub>-Konzentration<sup>4</sup>
- Nutzen der 2m-Temperatur aus dem Bodendatenset<sup>5</sup>
- Grobes Gitter - 96 Breitengrade und 192 Längengrade)
- Grobe zeitliche Auflösung - Monatsmittel
- Zeitraum - 1850 bis 2005 mit insgesamt 1872 Monaten
  - ⇒ 1165 Trainingszeitschritte
  - ⇒ 205 Validierungszeitschritte
  - ⇒ 502 Testzeitschritte

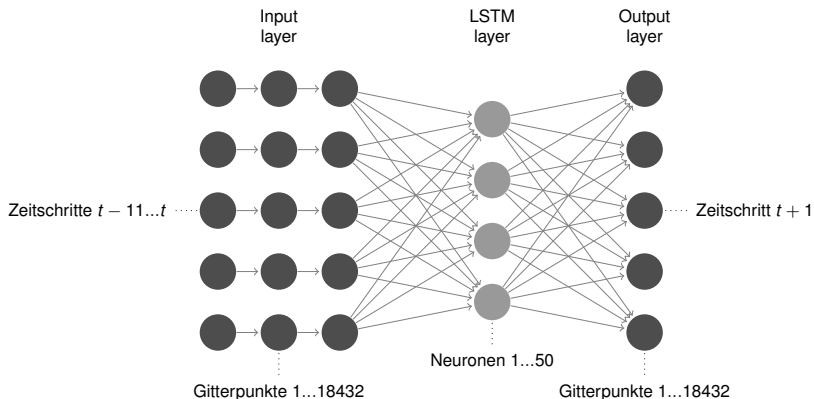
---

<sup>3</sup>Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., ... Glushak, K. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3), 572-597.

<sup>4</sup>Plesca, E., Grützun, V., Buehler, S. A. (2018). How robust is the weakening of the Pacific Walker circulation in CMIP5 idealized transient climate simulations?. *Journal of Climate*, 31(1), 81-97.

<sup>5</sup>Bereitstellung durch die Arbeitsgruppe Strahlung und Fernerkundung

# LSTM



**Abbildung:** LSTM mit einem „Long Short-Term Memory layer“ mit 50 Neuronen. Die Input- und Output-Daten sind eindimensional. Aus 12 Zeitschritten wird der folgende Zeitschritt vorhergesagt.

# Übersicht des Modells

Folgende Parameter wurden beispielhaft primär verwendet:

- 50 Neuronen
- 20 Epochen
- 12 Zeitschritte
- 128 Batchsize
- Skalierung auf einen Wertebereich
- Optimizer „Adam“ (Learning Rate 0,001)
- Aktivierungsfunktion „tanh“ (Hyperbolic tangent activation function)
- Callbacks - erste Implementierung von „ReduceLROnPlateau“

# Implementierungen

## Performance

- Angepasste Batch-Size
- Benutzung von Generatoren

## Genauigkeit

- Variation der Zeitschritte
- Training an 6 Ensemblemitgliedern
- Zufällige Reihenfolge der Input-Daten
- Anpassung der Learning Rate
- Angepasste Batch-Size<sup>6</sup>

---

<sup>6</sup>Bengio (1012): Practical recommendations for gradient-based training of deep architectures.

# Zeitliche Variabilität - Schritt 1

BIAS = 0,2734 K

STD = 2,1808 K

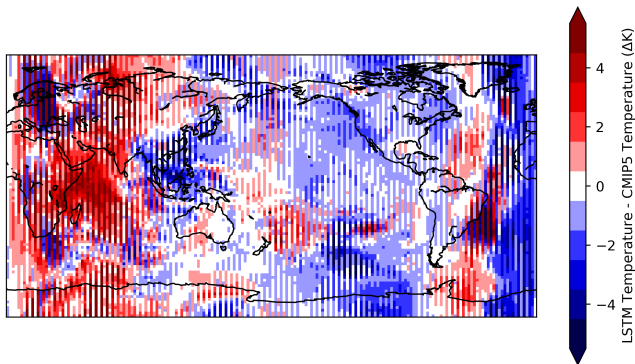


Abbildung: Zeitschritt 13



## Zeitliche Variabilität - Schritt 2

BIAS = 0,7461 K

STD = 2,5067 K

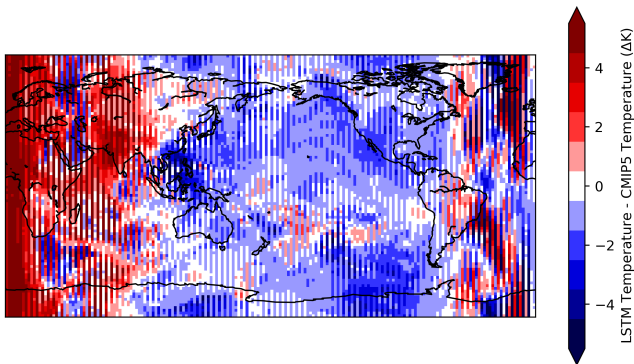


Abbildung: Zeitschritt 14

## Zeitliche Variabilität - Schritt 3

BIAS =  $-0,1174$  K

STD =  $2,4723$  K

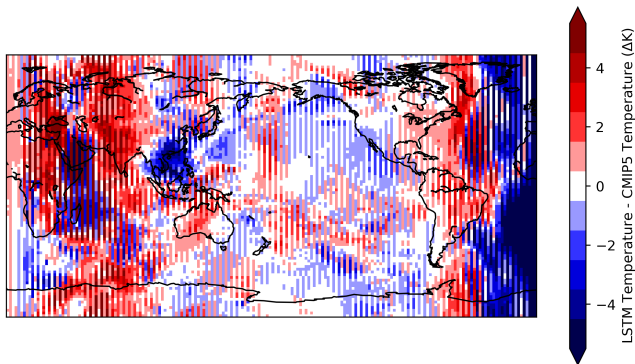


Abbildung: Zeitschritt 15

# Zeitliche Variabilität - Schritt 4

BIAS = 0,5362 K

STD = 2,3549 K

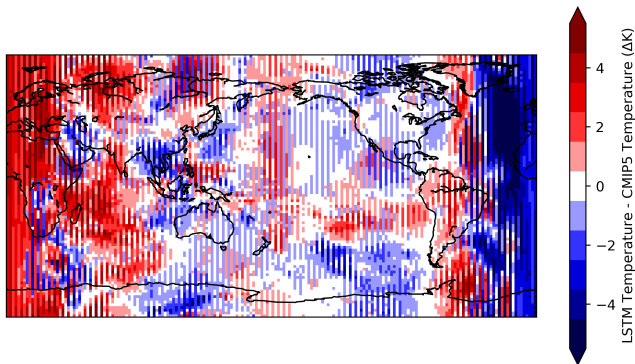


Abbildung: Zeitschritt 16

# Zeitliche Variabilität - Schritt 5

BIAS = 0,2156 K

STD = 1,9186 K

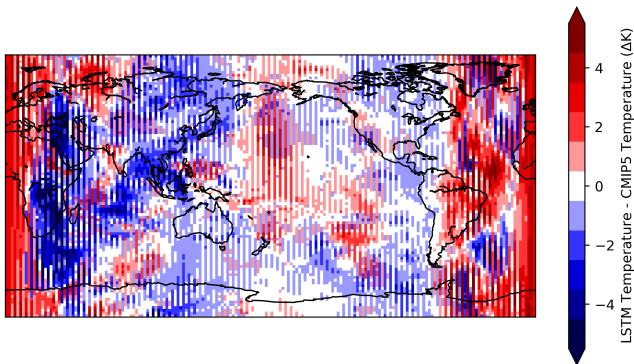


Abbildung: Zeitschritt 17

# Verteilung der Modelldifferenzen

BIAS = 0,2197 K

STD = 1,1240 K

Laufzeit = 24,4 min

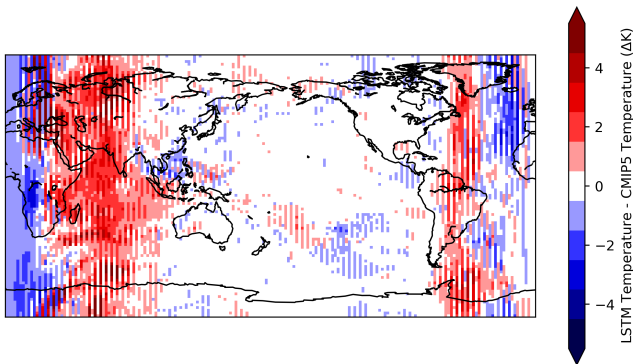


Abbildung: Differenzen der Temperaturen zwischen LSTM und CMIP5 Daten gemittelt über 40,75 Jahre

# Korrelation

RMSE = 2,4980 K

CORR = 0,9929

Laufzeit = 24,4 min

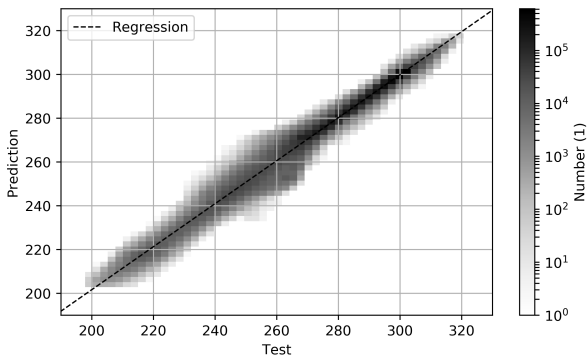


Abbildung: 2D-Histogramm der vorhergesagten und wahren Temperatur.

# Fazit

- Nicht Erfüllung des Courant-Friedrichs-Lewy Kriterium<sup>7</sup> der Datengrundlage
- Vorhersage des einzelnen Zeitschrittes nicht belastbar
- Annäherung der Statistik, nicht der Physik
- Statistik belastbar:
  - RMSE  $\approx$  2,5 K
  - CORR  $>$  0,99
- Laufzeit der Vorhersage einzelnen Zeitschrittes  $<$  5 s
- Fehlende Vergleichbarkeit des Projektes

---

<sup>7</sup>Trivellato und Castelli (2013): On the Courant–Friedrichs–Lewy criterion of rotating grids in 2D vertical-axis wind turbine analysis.

# Mögliche Optimierungsansätze

## Performance

- GPUs
- Effizientere I/O
- Empirische Batch-Size Vergrößerung

## Genauigkeit

- Abbruch des Trainings bei Genauigkeit
- Empirische Anpassung der Veränderung der Learning Rate
- Implementierung zyklischer Randbedingungen
- Nutzung weiterer Variablen
- Nutzung weiterer Ensemble Member
- Simulation eines Modells mit kleinerem Zeitschritt



# Neuronalen Netzwerke in der Klimamodellierung zur Prognose eines Zeitschritts

## Vorhersage der 2m-Temperatur

Tobias Machnitzki<sup>1</sup> und Finn Burgemeister<sup>1</sup>

<sup>1</sup>Meteorologisches Institut  
Universität Hamburg

23. Mai 2018