

Langzeitanalysen von Monsunregendaten

-

Ein explorativer Ansatz

3.07.2020



Alexander Prinz

Studiengang: Data Science (M. Sc.)
Fachbereich: Informatik und Sprachen

Hochschule Anhalt

Anhalt University of Applied Sciences

Einführung in das Thema

- Monsunregen und Ozeanoberflächentemperatur
- Datensätze

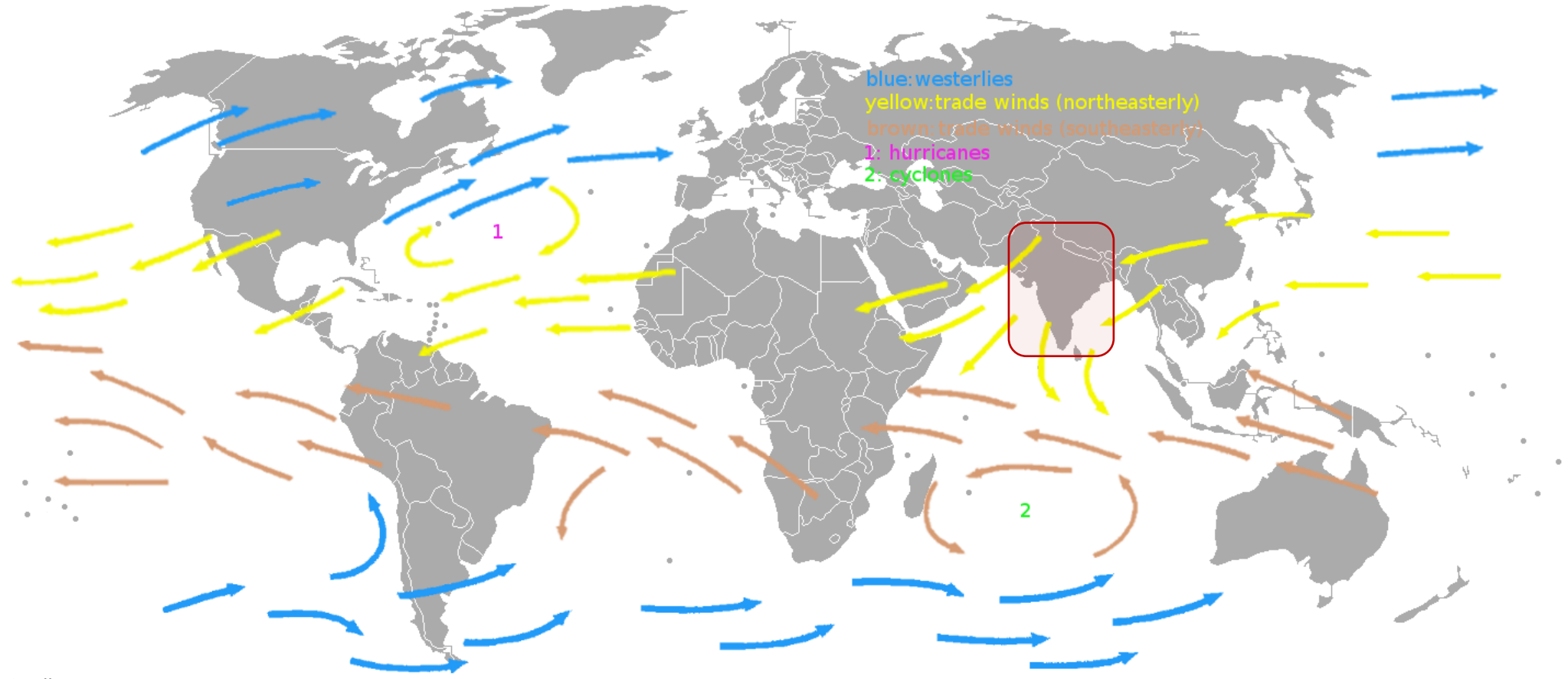
Vorgehen

- Vorprozessieren der Daten
 - ~ selektieren der relevanten Daten aus den Datensätzen
- erste Visualisierung der Vorprozessierten Daten
 - ~ Jahresniederschlagsgrafiken
 - ~ Zeitreihen
 - ~ Hypothesen
- Nachprozessierung der Daten nach Hypothesenannahme
 - ~ kubische Spline-Regression zur Datenglättung
- Fourier-Transformation zur Bestimmung von Periodizität als eine der Hypothesen
- Ozeanoberflächentemperaturabhängigkeit
- Lineare Regression als Langzeittrendmodell



Einführung in das Thema

Monsunregen und Ozeantemperatur



Quelle:
[https://de.wikipedia.org/wiki/Passat_\(Windsystem\)](https://de.wikipedia.org/wiki/Passat_(Windsystem))



Einführung in das Thema

Datensätze

Niederschlagsmengen

- Quelle: Kaggle > „rainfall in india 1901-2015.csv“
- CSV Datei mit folgender Struktur:

```
SUBDIVISION, YEAR, JAN, FEB, MAR, APR, MAY, JUN, JUL, AUG, SEP, OCT, NOV, DEC, ANNUAL, Jan-Feb, Mar-May, Jun-Sep, Oct-Dec
ANDAMAN & NICOBAR ISLANDS, 1901, 49.20, 87.10, 29.20, 2.30, 528.80, 517.50, 365.10, 481.10, 332.60, 388.50, 558.20, 33.60, 3373.20, 136.30, 560.30, 1696.30, 980.30
ANDAMAN & NICOBAR ISLANDS, 1902, 0.00, 159.80, 12.20, 0.00, 446.10, 537.10, 228.90, 753.70, 666.20, 197.20, 359.00, 160.50, 3520.70, 159.80, 458.30, 2185.90, 716.70
ANDAMAN & NICOBAR ISLANDS, 1903, 12.70, 144.00, 0.00, 1.00, 235.10, 479.90, 728.40, 326.70, 339.00, 181.20, 284.40, 225.00, 2957.40, 156.70, 236.10, 1874.00, 690.60
ANDAMAN & NICOBAR ISLANDS, 1904, 9.40, 14.70, 0.00, 202.40, 304.50, 495.10, 502.00, 160.10, 820.40, 222.20, 308.70, 40.10, 3079.60, 24.10, 506.90, 1977.60, 571.00
ANDAMAN & NICOBAR ISLANDS, 1905, 1.30, 0.00, 3.30, 26.90, 279.50, 628.70, 368.70, 330.50, 297.00, 260.70, 25.40, 344.70, 2566.70, 1.30, 309.70, 1624.90, 630.80
...
```

Ozeanoberflächentemperatur

- Quelle: Kaggle > „elnino.csv“
- CSV Datei mit folgender Struktur:

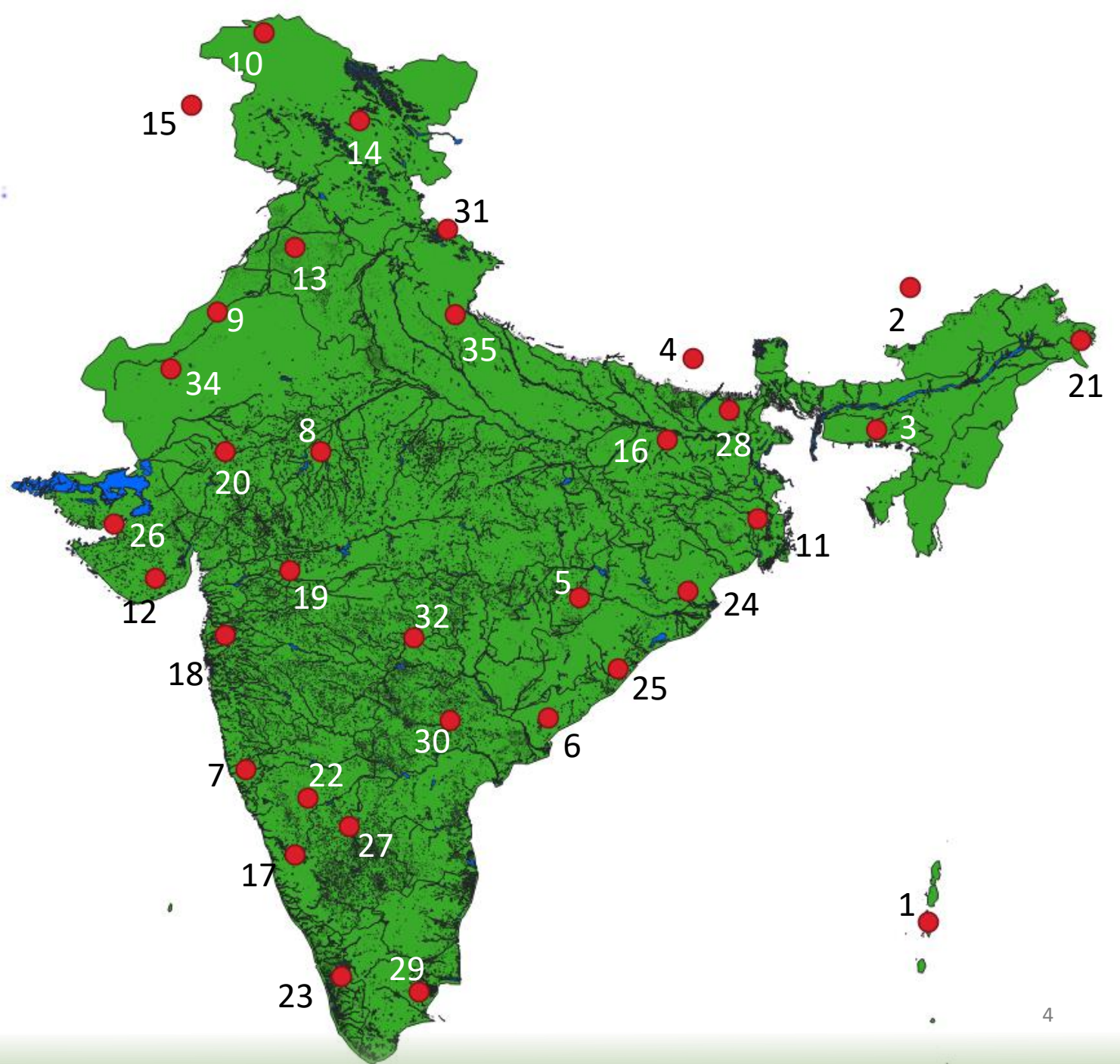
```
Observation, Year, Month, Day, Date, Latitude, Longitude, Zonal Winds, Meridional Winds, Humidity, Air Temp, Sea Surface Temp
1, 80, 3, 7, 800307, -0.02, -109.46, -6.8, 0.7, ., 26.14, 26.24
2, 80, 3, 8, 800308, -0.02, -109.46, -4.9, 1.1, ., 25.66, 25.97
3, 80, 3, 9, 800309, -0.02, -109.46, -4.5, 2.2, ., 25.69, 25.28
4, 80, 3, 10, 800310, -0.02, -109.46, -3.8, 1.9, ., 25.57, 24.31
5, 80, 3, 11, 800311, -0.02, -109.46, -4.2, 1.5, ., 25.3, 23.19
...
```



Einführung in das Thema

Messstandorte in Indien

1. ANDAMAN & NICOBAR ISLANDS
2. ARUNACHAL PRADESH
3. ASSAM & MEGHALAYA
4. BIHAR
5. CHHATTISGARH
6. COASTAL ANDHRA PRADESH
7. COASTAL KARNATAKA
8. EAST MADHYA PRADESH
9. EAST RAJASTHAN
10. EAST UTTAR PRADESH
11. GANGETIC WEST BENGAL
12. GUJARAT REGION
13. HARYANA DELHI & CHANDIGARH
14. HIMACHAL PRADESH
15. JAMMU & KASHMIR
16. JHARKHAND
17. KERALA
18. KONKAN & GOA
19. MADHYA MAHARASHTRA
20. MATATHWADA
21. NAGA MANI MIZO TRIPURA
22. NORTH INTERIOR KARNATAKA
23. ORISSA
24. PUNJAB
25. RAYALSEEMA
26. SAURASHTRA & KUTCH
27. SOUTH INTERIOR KARNATAKA
28. SUB HIMALAYAN WEST BENGAL & SIKKIM
29. TAMIL NADU
30. TELANGANA
31. UTTARAKHAND
32. VIDARBHA
33. WEST MADHYA PRADESH
34. WEST RAJASTHAN
35. WEST UTTAR PRADESH



Vorgehen

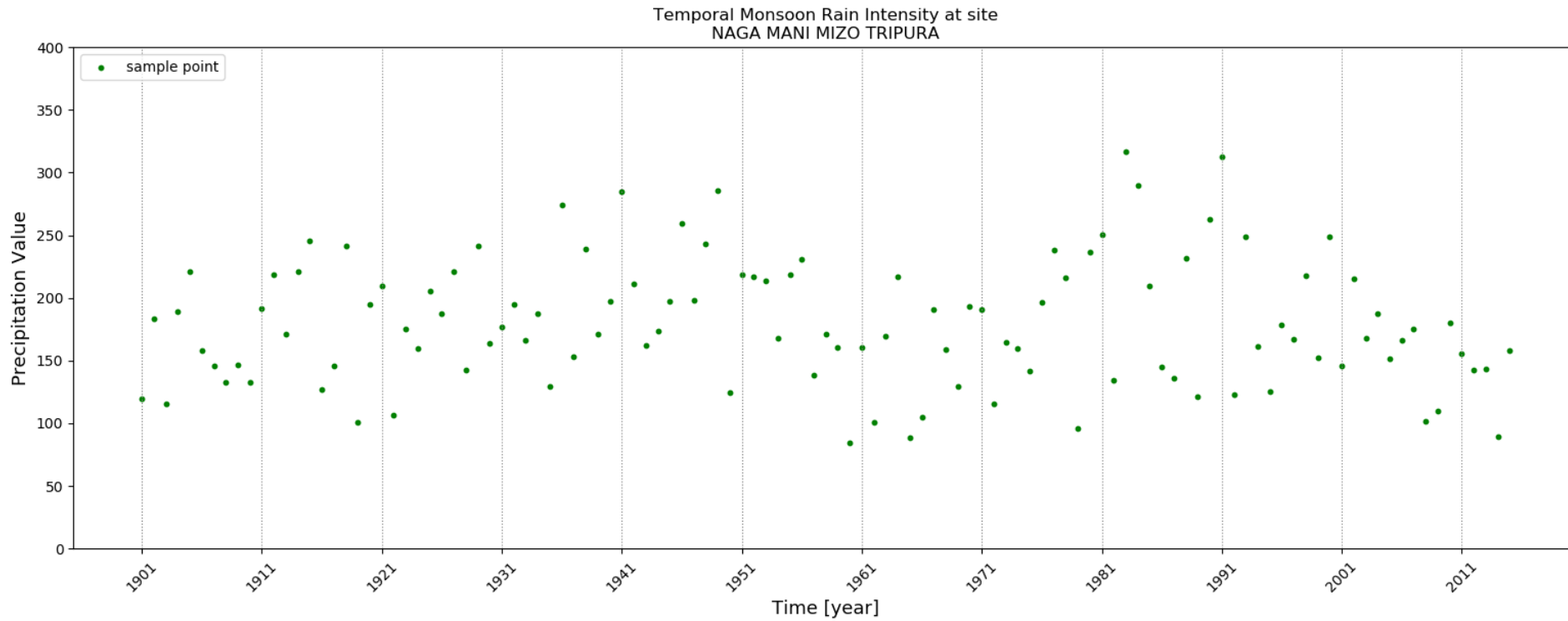
Vorprozessieren

Niederschlagsmengenzeitreihen:

- Monsunzeit zwischen April und September gemittelt
 - > anderer Zeitraum wird nicht beachtet
- ausselektieren von Zeiträumen mit NAN - Werten
 - > np.nanmean() ist keine gute Option -> Erklärung folgt



Vorgehen



-> 3 Hypothesen: zeitlich periodisches Verhalten, Ozeanoberflächentemperatur verhält sich ähnlich, langzeitliche Trendentwicklung (Ab- oder Zunahme der Niederschlagsmengen mit der Zeit)

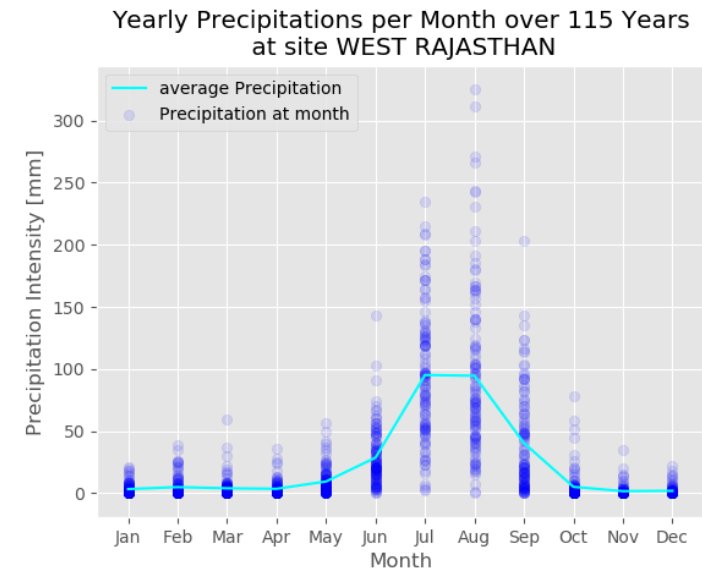
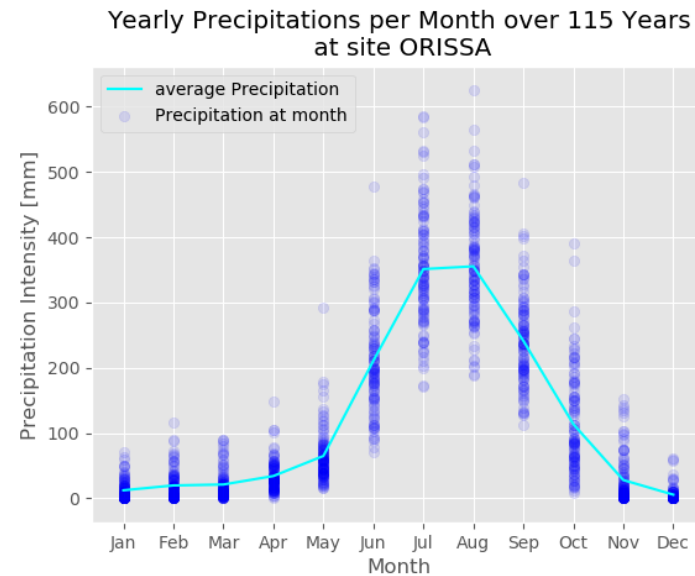
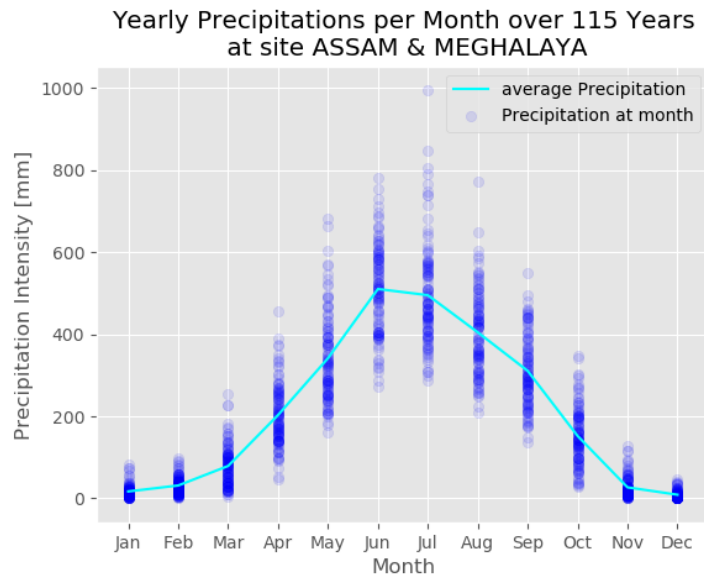


Vorgehen

Vorprozessieren

bzgl. Grafiken zum Visualisieren der jährlichen Niederschläge:

- Extrahieren der letzten Spalte welche die mittlere jährliche Niederschlagsmenge enthält.



Vorgehen

Vorprozessieren

Ozeanoberflächentemperatur:

selektieren nach dem Messort mit der dichtesten Entfernung an Indien

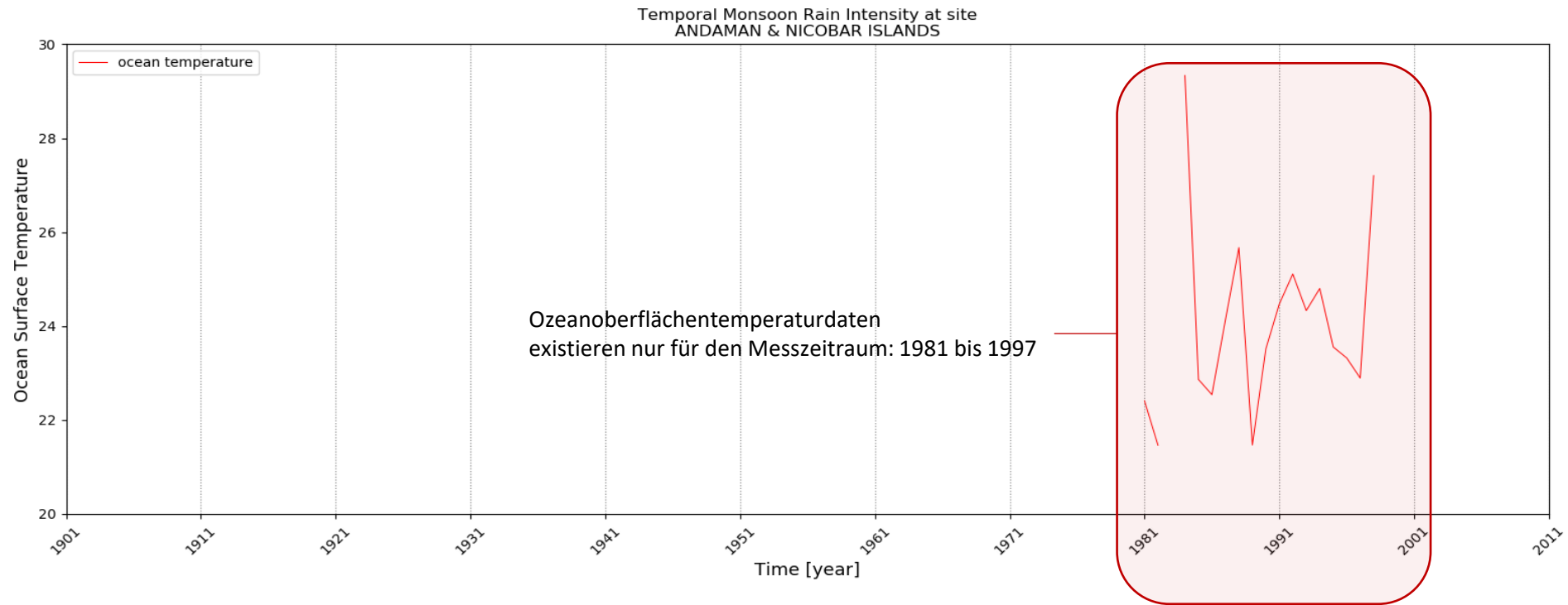


Vorgehen

Vorprozessieren

Ozeanoberflächentemperatur:

selektieren nach dem Messort mit der dichtesten Entfernung an Indien



Vorgehen

Nachprozessieren

Niederschlagsmengenzeitreihen:

- > Glättung der Daten durch kubische Spline Regression visualisiert das angenommene periodische Verhalten besser und ist wichtiger Vorbereitungsschritt für die weitere Analyse via FOURIER-Transformation

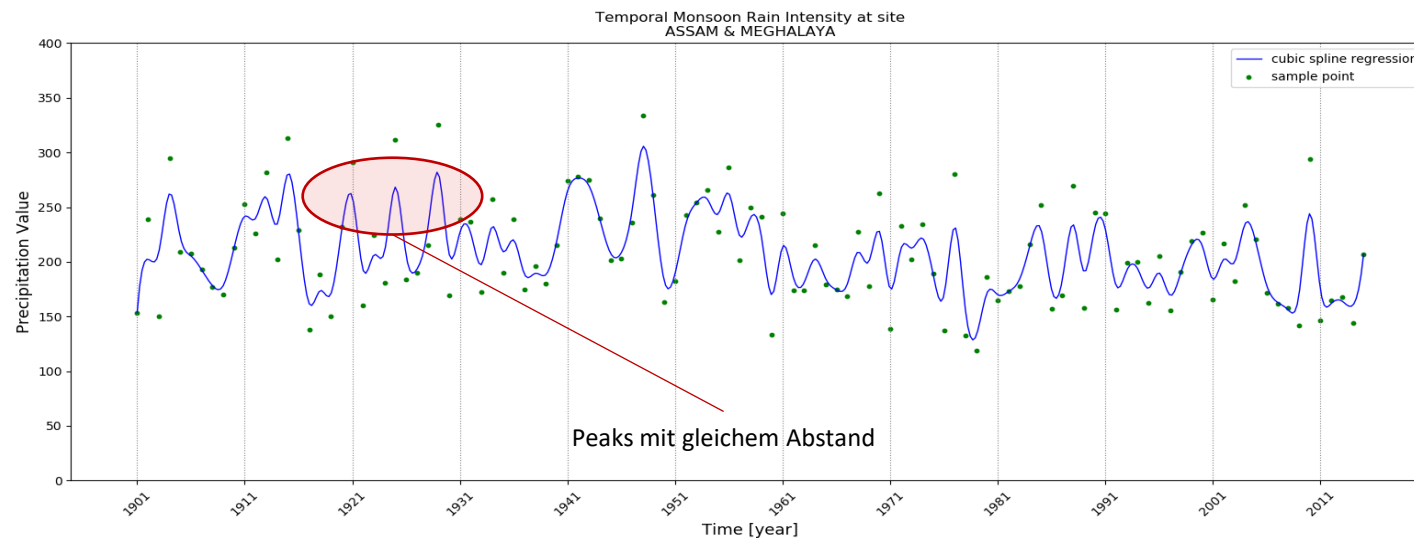


Vorgehen

Nachprozessieren

Niederschlagsmengenzeitreihen:

-> Glättung der Daten durch kubische Spline Regression visualisiert das angenommene periodische Verhalten besser.



Nur ein qualitativer Hinweis auf Bestand der Hypothese. Gibt es eine quantitative Methode?

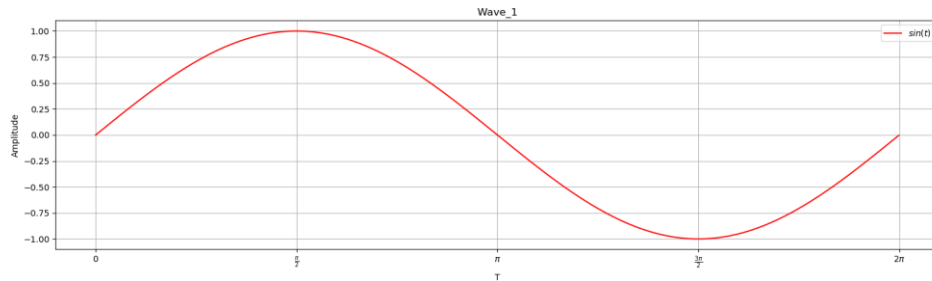
-> FOURIER-Transformation



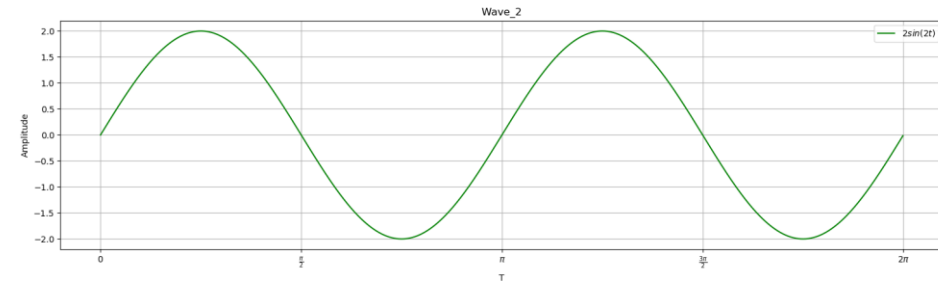
Vorgehen

FOURIER-Transformation zum Untersuchen auf Periodizität im Signal

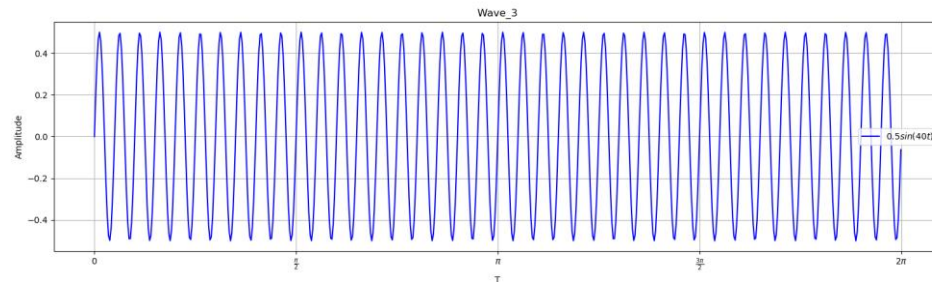
Überlagerung verschiedener periodischer Signale zu einem Summensignal.



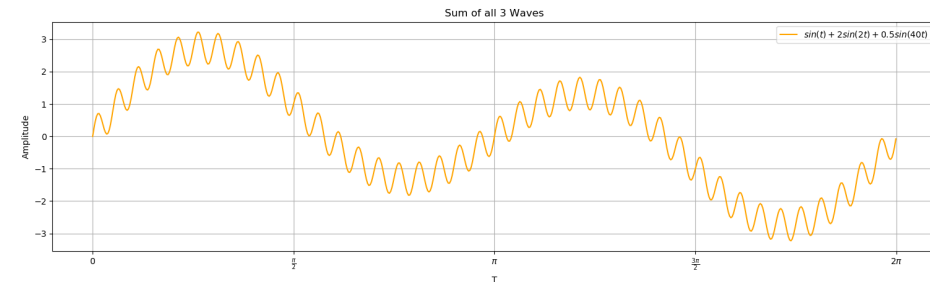
+



+



=



Vorgehen

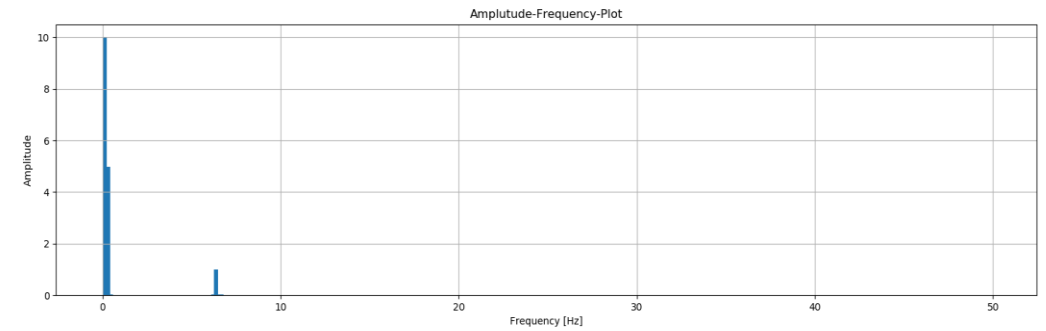
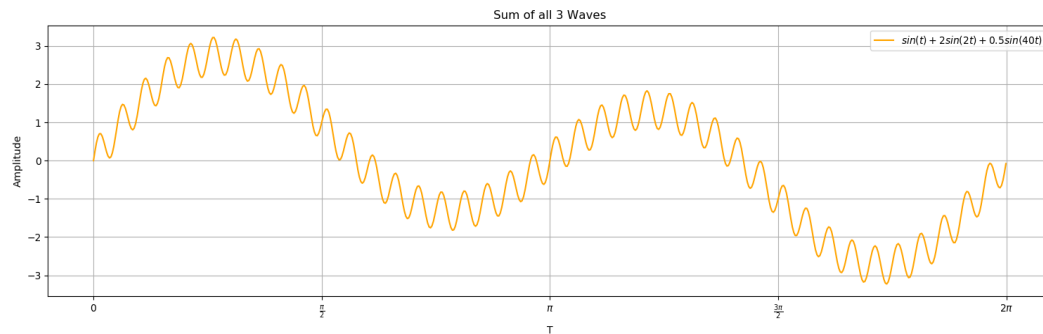
FOURIER-Transformation

Transformation vom...

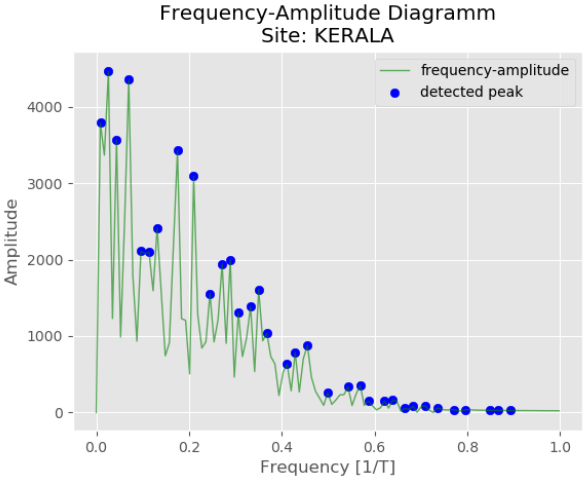
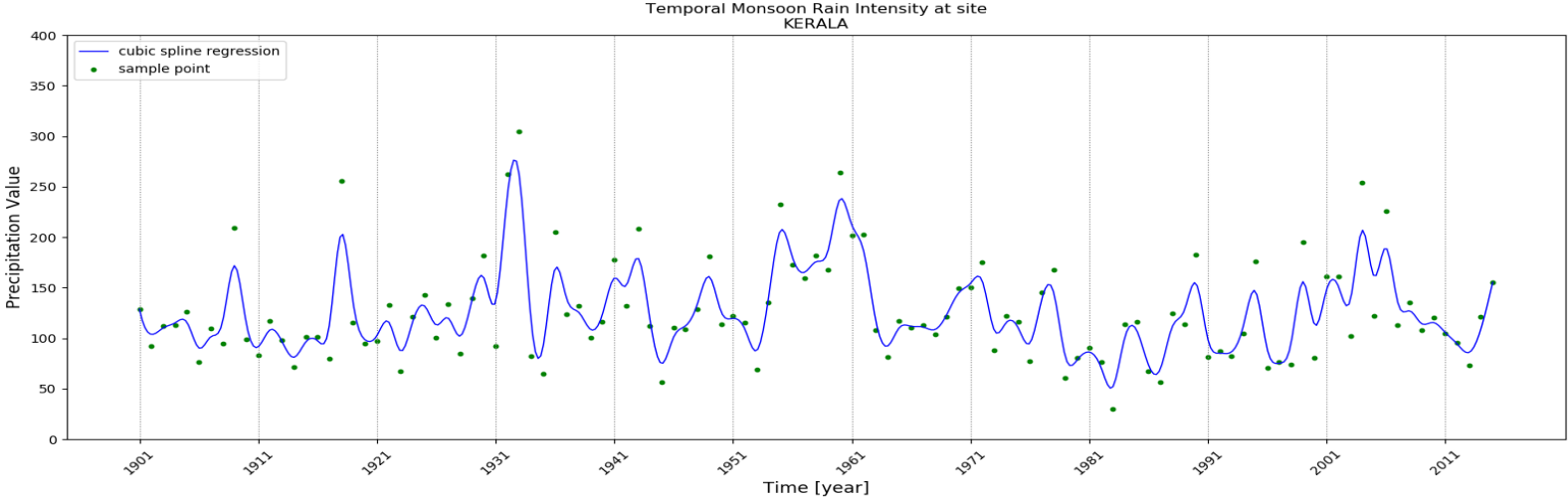
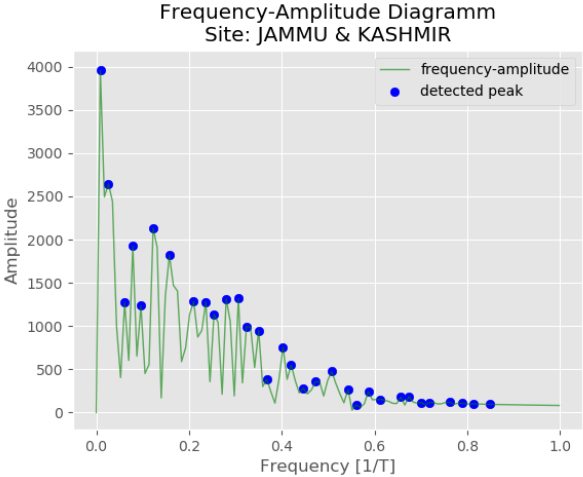
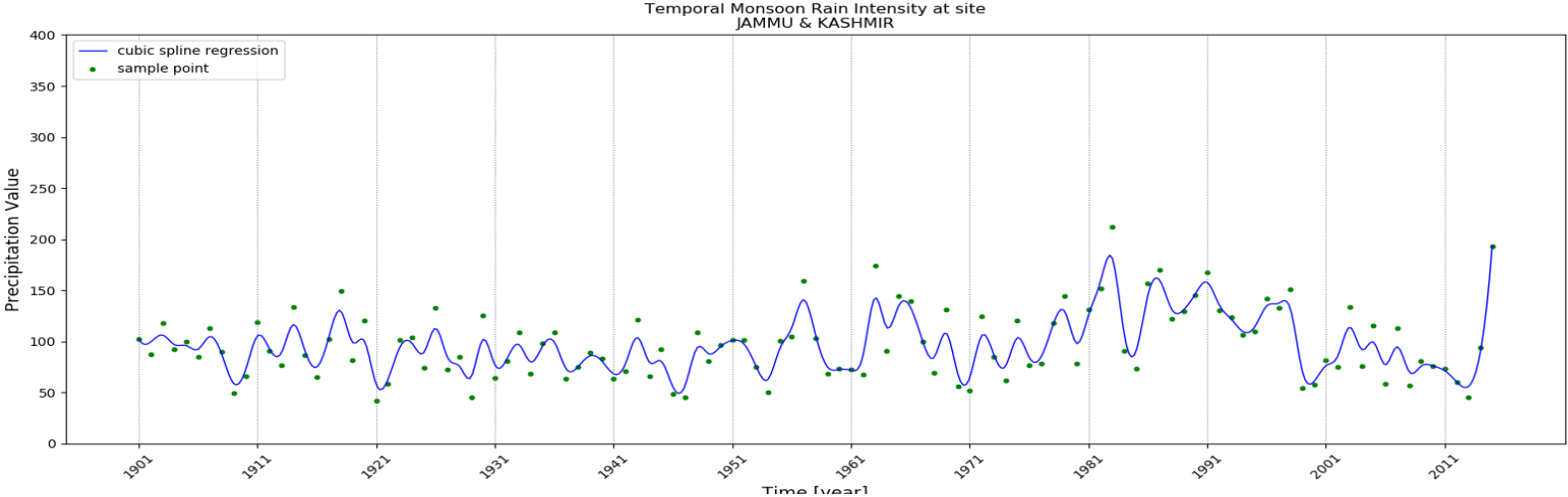
Amplituden-Zeit-Raum

-> in den ->

Amplituden-Frequenzraum



Ergebnisse



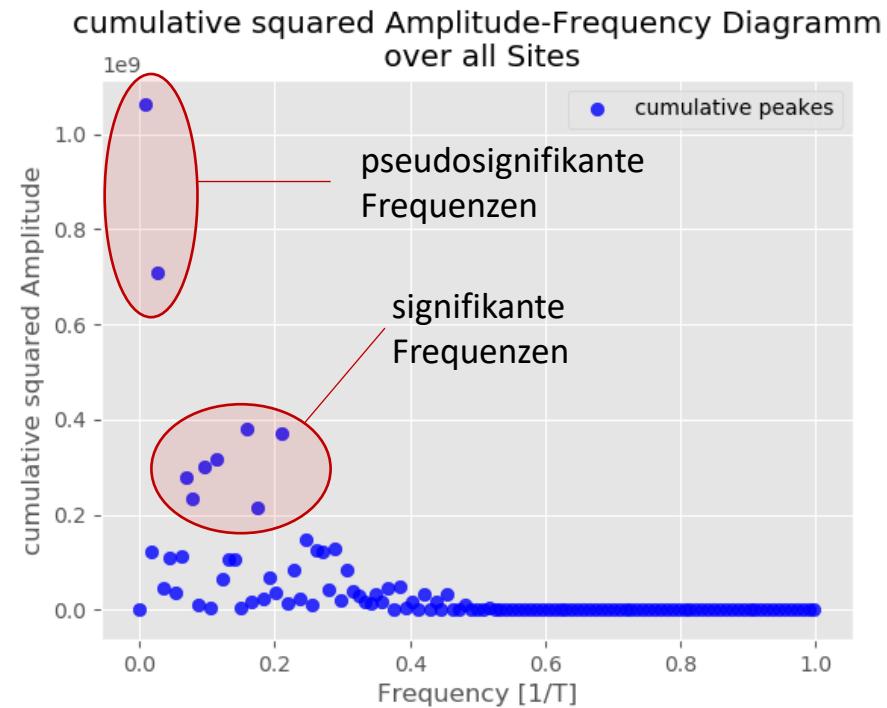
Ergebnisse

Gibt es also signifikante Periodizitäten?

-> sehr wahrscheinlich ja

Worauf basieren sie?

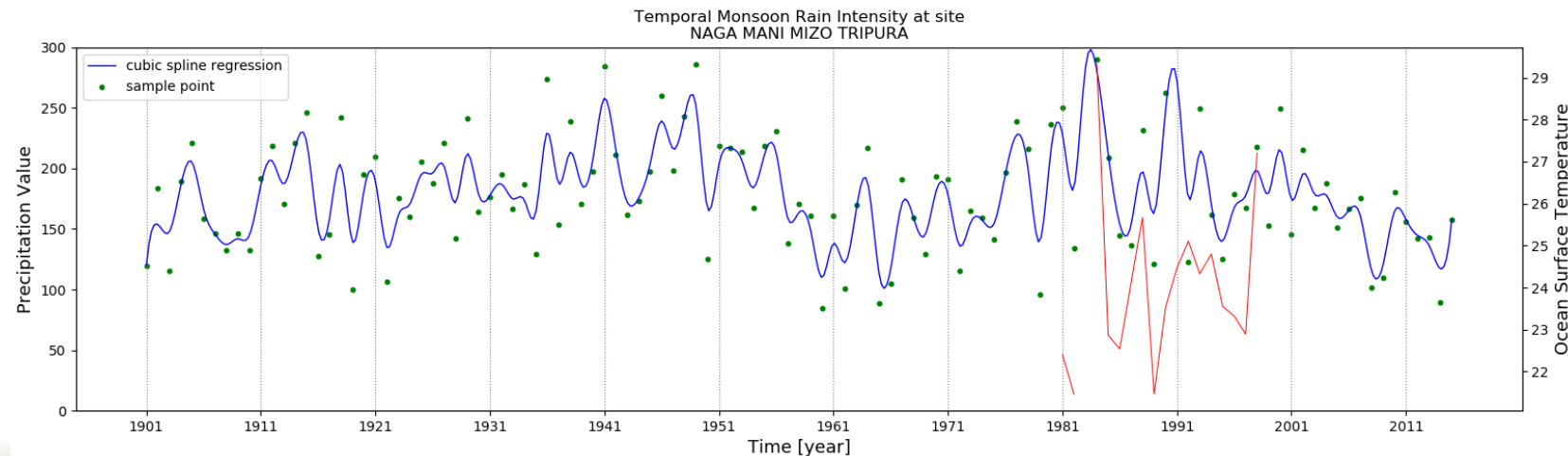
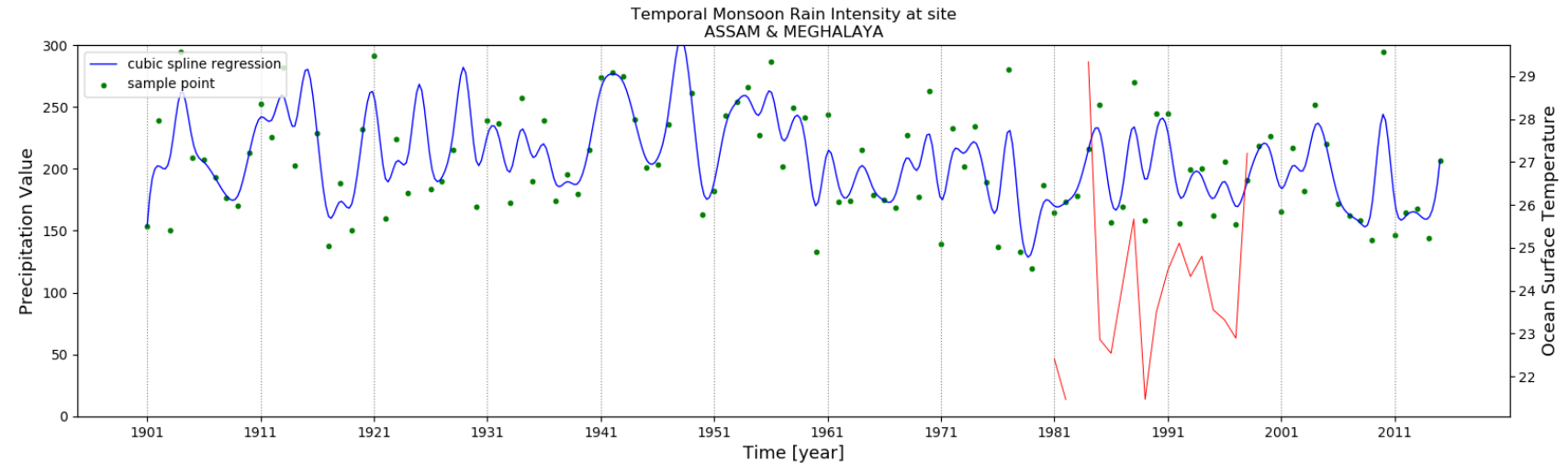
-> kinematische Erde-Sonne-Beziehungen (Melancović-Zyklen)



Ergebnisse

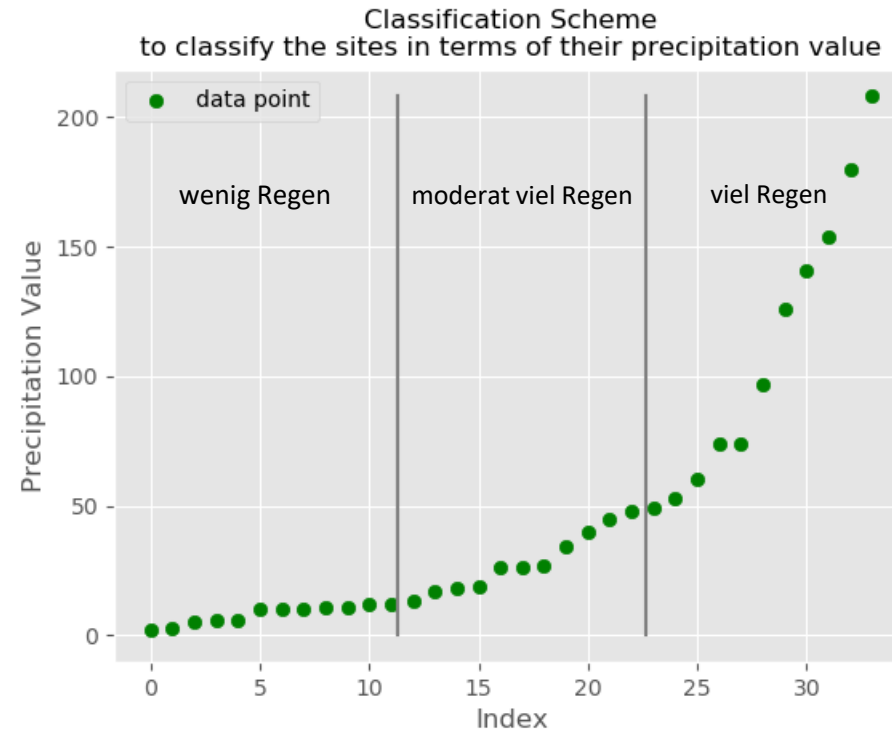
Gibt es eine Beziehung zwischen Ozeanoberflächentemperatur und Monsunintensität?

Mit den vorhandenen Daten
nicht sicher zu sagen.
Aber es gibt einige Daten
welche die Hypothese stützen.



Ergebnisse

Gibt es langzeitliche Trends bzgl. des Verhaltens der Monsunintensität?

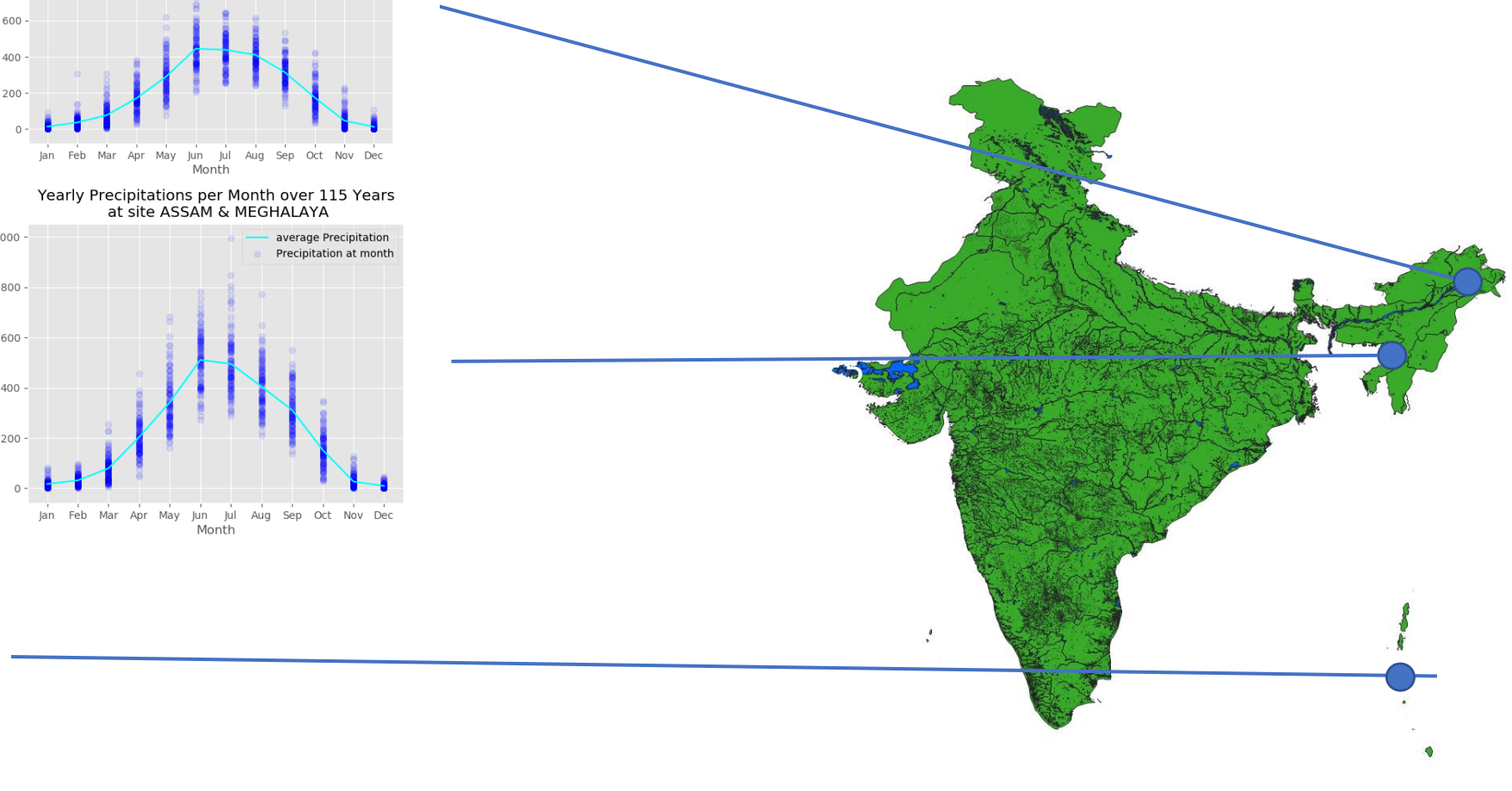
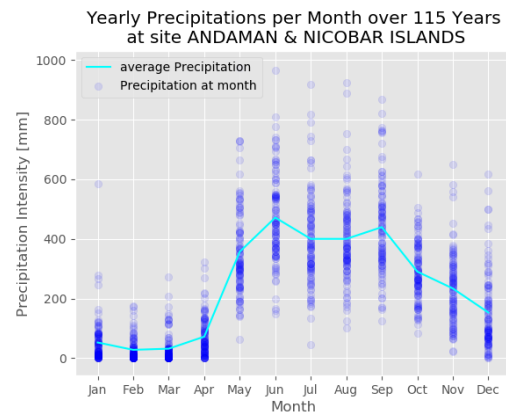
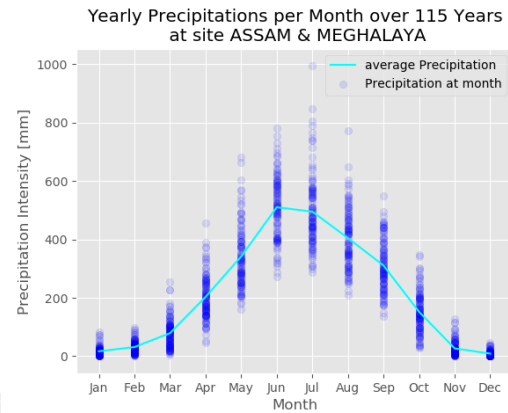
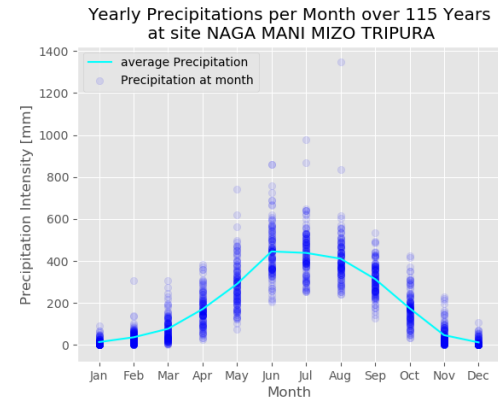


Unterteilung in 3 Klassen (wenig Regenintensität, moderate Regenintensität, viel Regenintensität)



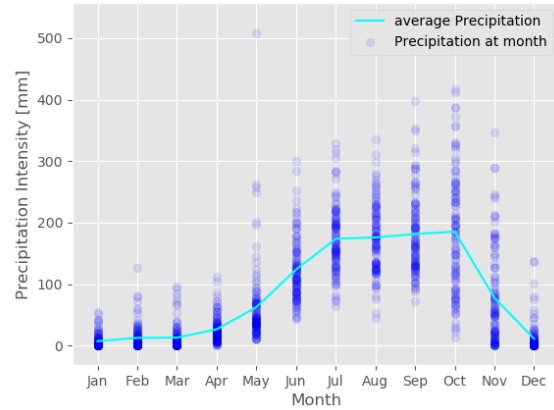
Ergebnisse

Gebiete mit viel Niederschlag in der Monsunzeit

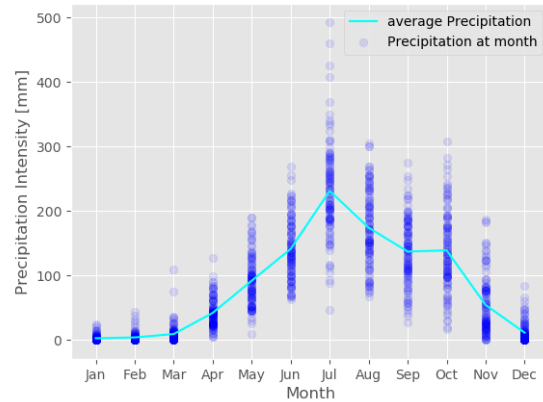


Ergebnisse

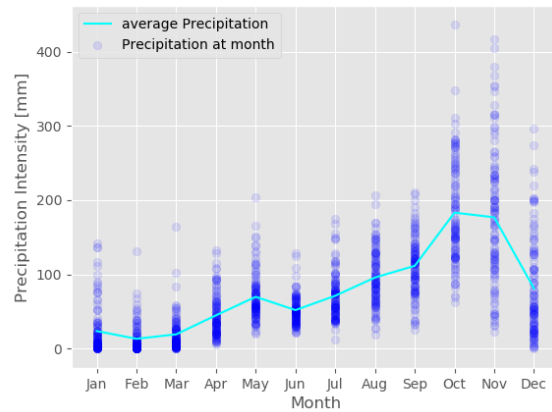
Yearly Precipitations per Month over 115 Years
at site COASTAL ANDHRA PRADESH



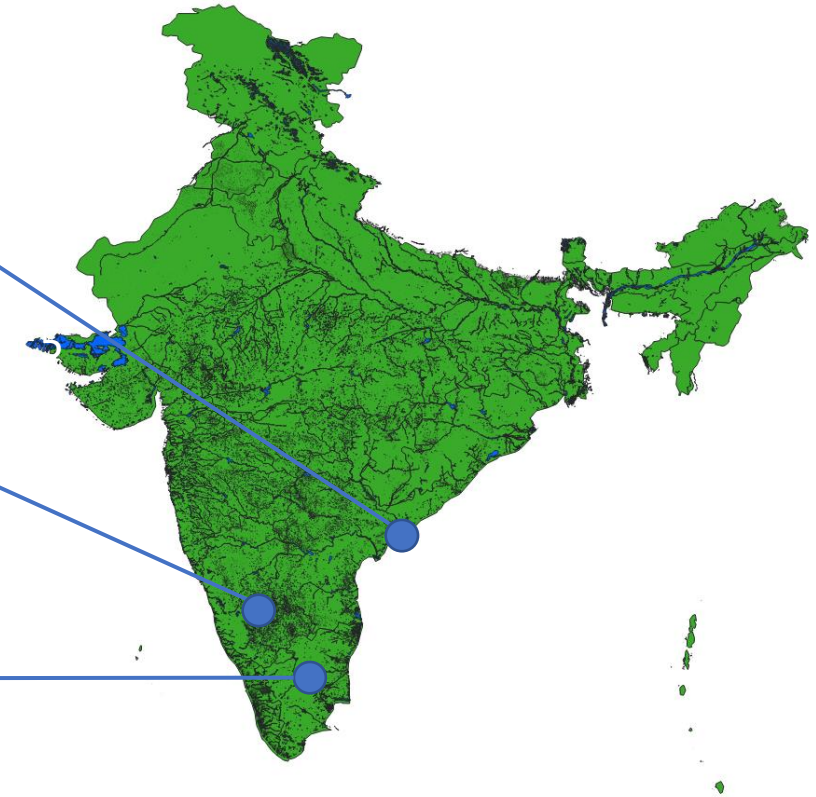
Yearly Precipitations per Month over 115 Years
at site SOUTH INTERIOR KARNATAKA



Yearly Precipitations per Month over 115 Years
at site TAMIL NADU

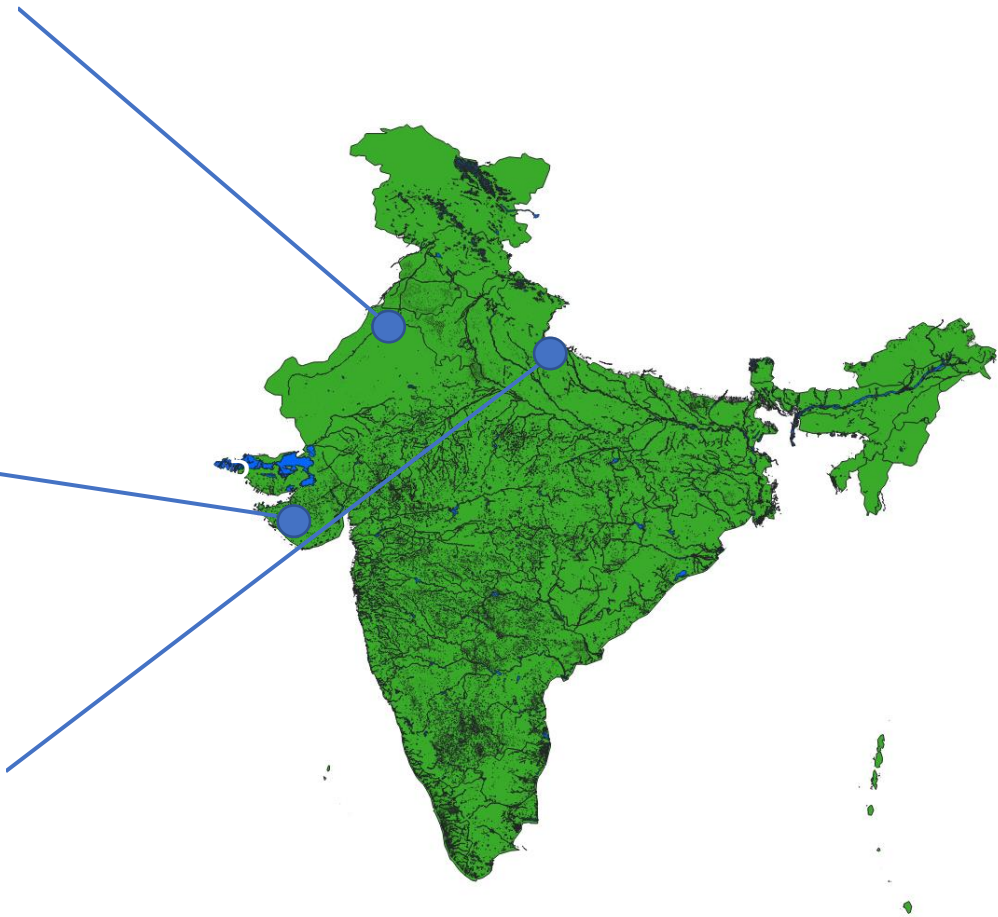
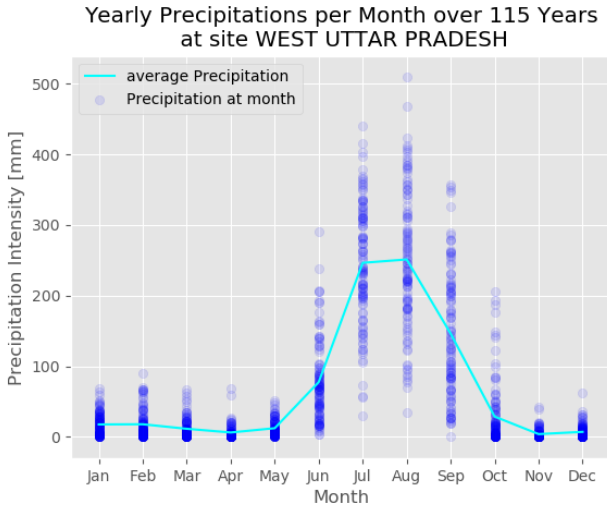
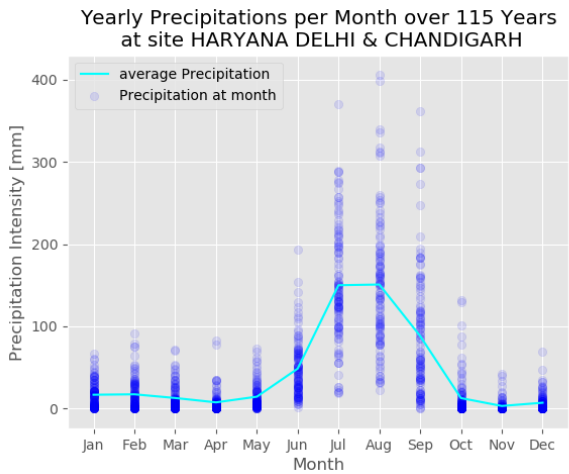
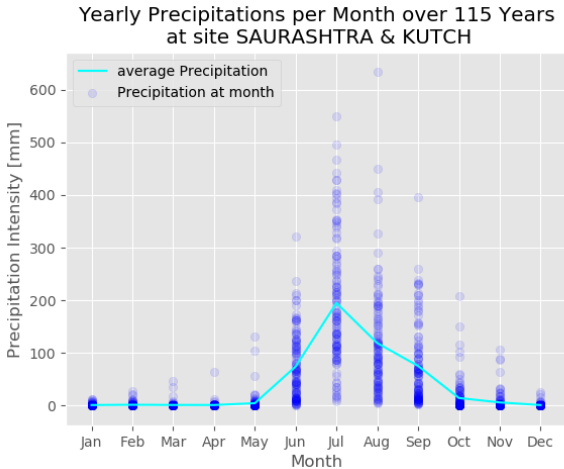


Gebiete mit moderat viel Niederschlag
in der Monsunzeit



Ergebnisse

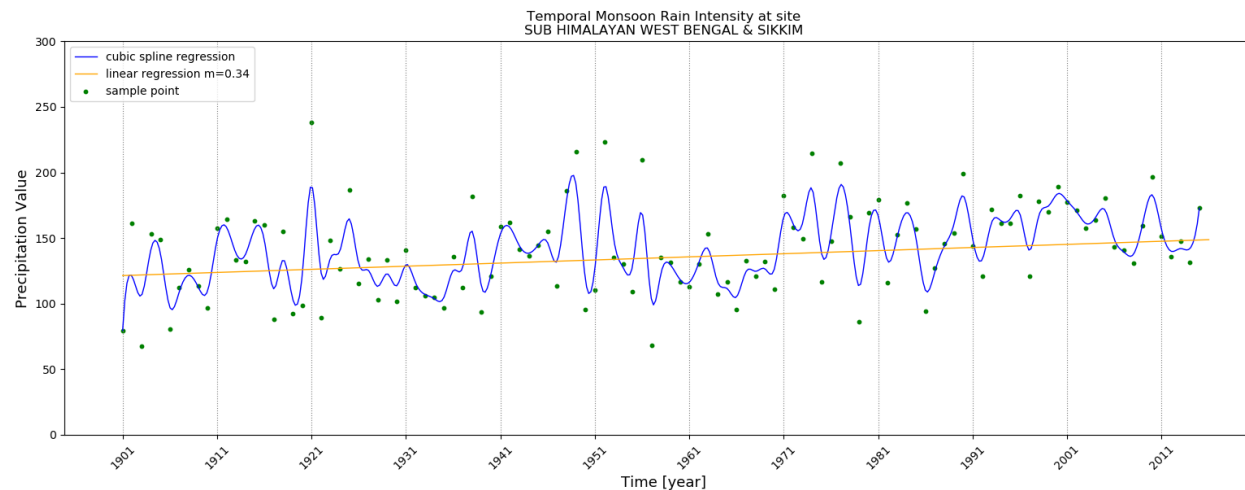
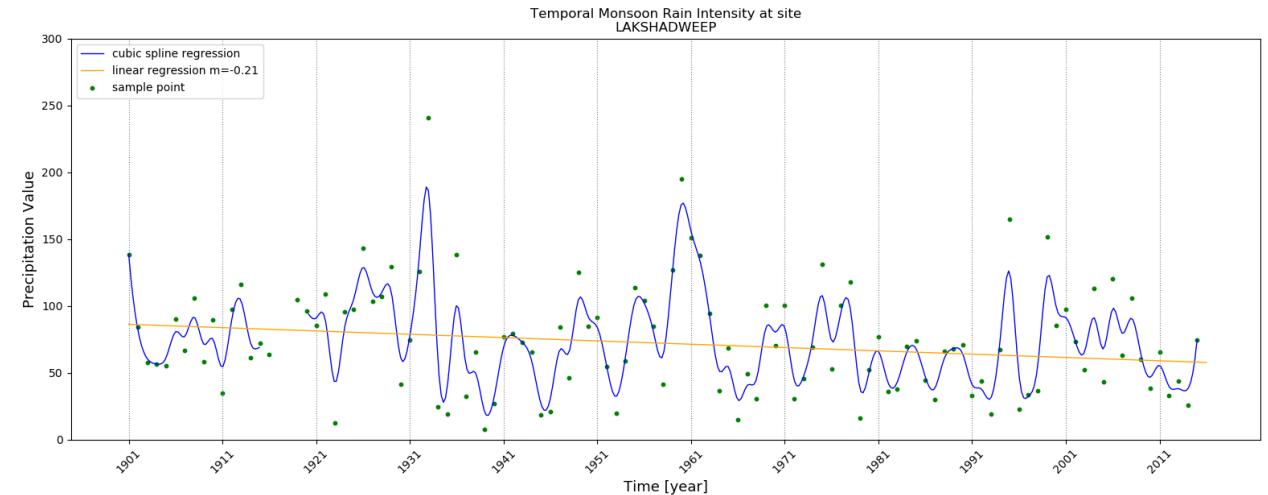
Gebiete mit wenig Niederschlag in der Monsunzeit



Ergebnisse

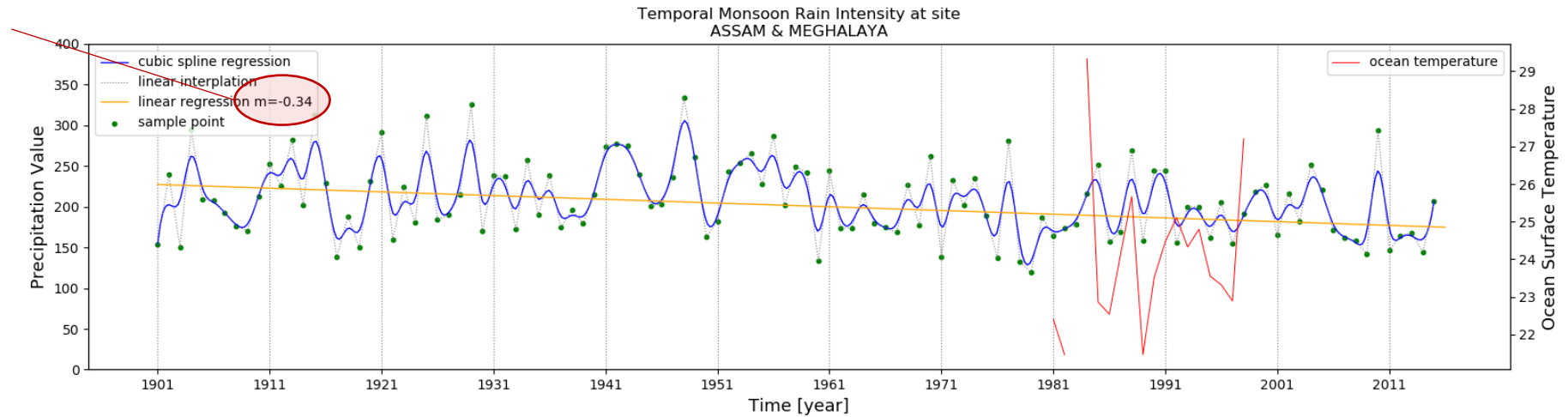
Gibt es langzeitliche Trends bzgl. des Verhaltens der Monsunintensität?

Es sind interessante Trends zu erkennen, wenn man die Daten klassifiziert betrachtet.



Ergebnisse

Analyse der Regressionsanstiege aller Messorte



gesamt :

44% positive Tendenz
56% negative Tendenz

viel Regen:

60% positive Tendenz
40% negative Tendenz

0.153 mean positive gradient
0.160 mean negative gradient

moderat viel Regen:

58% positive Tendenz
42% negative Tendenz

0.034 mean positiv gradient
0.040 mean negative gradient

wenig Regen;

25% positive Tendenz
75% negative Tendenz

0.026 mean positiv gradient
0.024 mean negativ gradient

