

# Vorhersage von E/A-Leistung im Hochleistungsrechnen unter der Verwendung von neuronalen Netzen

Jan Fabian Schmid  
Betreuer: Dr. Julian Kunkel

Universität Hamburg  
*2schmid@informatik.uni-hamburg.de*

12.02.2016

# Übersicht

## Einleitung

- Motivation und Ziele der Bachelorarbeit
- Modellierungsstrategien

## Künstliche neuronale Netze

## Modell des Ein-/Ausgabe-Pfads

## Messdaten

- Attribute der Messdaten
- Exploration der Messdaten

## Modelle zur Leistungsvorhersage

- Ergebnisse

## Einbeziehung des E/A-Pfads

- Fehlerklassen
- Ergebnisse von Modellen mit Fehlerklassen

## Zusammenfassung

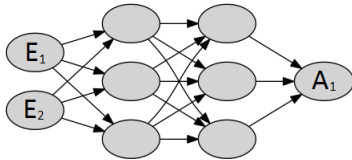
## Motivation und Ziele der Bachelorarbeit

- ▶ Hochleistungsrechnen ist wichtig für die Wissenschaft
  - ▶ Verwendung massiv paralleler Systeme
- ▶ Effiziente Nutzung der Ressourcen ist schwierig
  - ▶ Hilfreich wäre ein Analysewerkzeug zur Beurteilung der Effizienz der Ein-/Ausgabe (E/A) eines Programms
- ▶ Ziel: Erstellung eines Modells für das E/A-System

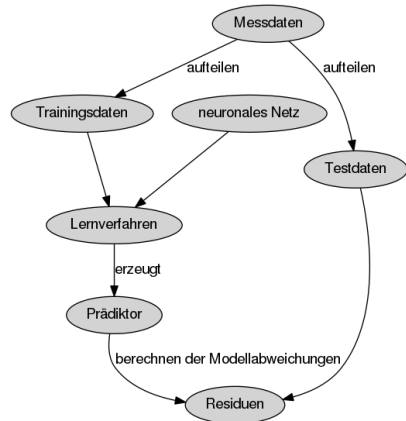
## Modellierungsstrategien

- ▶ Meist nur E/A-Leistung auf einzelnen Festplatten betrachtet
- ▶ Zwei mögliche Ansätze:
  - ▶ White-Box-Modellierung
  - ▶ Black-Box-Modellierung

# Künstliche neuronale Netze

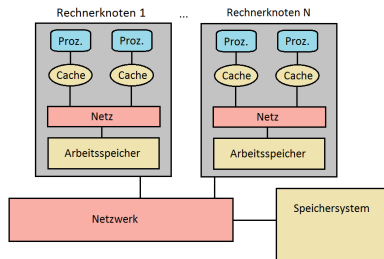


- ▶ Verfahren des maschinellen Lernens
- ▶ Zur Regressionsanalyse verwendet
  - ▶ Überwachtes Lernen
- ▶ Trainieren durch Fehlerrückführung
- ▶ Mächtiger als lineare Regression



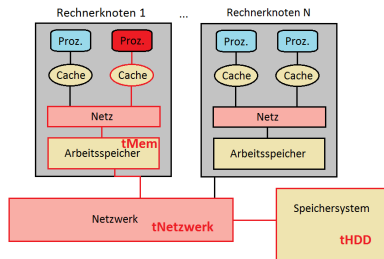
## Modell des Ein-/Ausgabe-Pfads

- ▶ Verarbeitung eines Zugriffs entlang des E/A-Pfades
- ▶ Zugriffszeit wesentlich durch tiefe des Pfades bestimmt



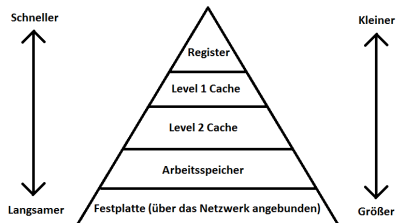
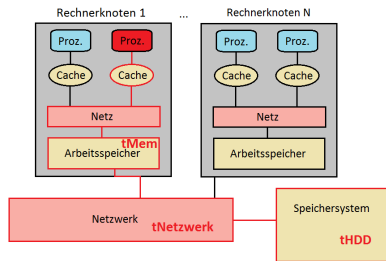
## Modell des Ein-/Ausgabe-Pfads

- ▶ Verarbeitung eines Zugriffs entlang des E/A-Pfades
- ▶ Zugriffszeit wesentlich durch tiefe des Pfades bestimmt



## Modell des Ein-/Ausgabe-Pfads

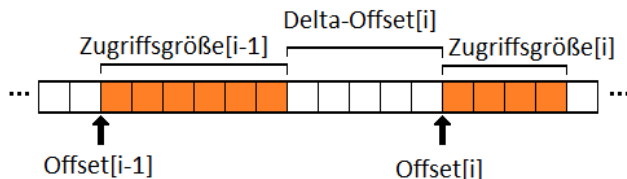
- ▶ Verarbeitung eines Zugriffs entlang des E/A-Pfades
- ▶ Zugriffszeit wesentlich durch tiefe des Pfades bestimmt





## Attribute der Messdaten

- ▶ Parameter eines E/A-Aufrufs:
  - ▶ Datei ID
  - ▶ **Zugriffsgröße**
  - ▶ **Operationstyp**
  - ▶ Offset
  - ▶ **Delta-Offset**
- ▶ Messbare Größen:
  - ▶ Zeitpunkt der Anfrage
  - ▶ **Zugriffszeit**
- ▶ Keine Informationen über den E/A-Pfad



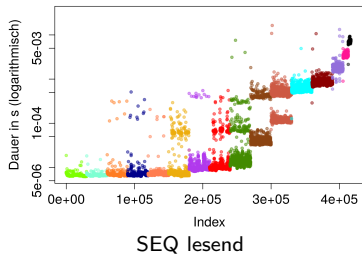
## Benchmark-Tests

- ▶ Daten wurden zur Verfügung gestellt
- ▶ Zwei untersuchte Anwendungsfälle:
  - ▶ **SEQ**: Sequentieller Zugriff auf hintereinanderliegende Daten
  - ▶ **RND**: Zufälliger Zugriffsort
- ▶ Zugriffsgrößen von 1 B bis 16 MiB
- ▶ Jeweils lesende und schreibende Messreihen
- ▶ Ausführung mit einem Thread

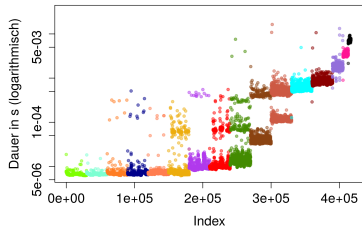
## Hardwarekonfiguration des Testsystems Mistral

- ▶ Über 1500 Knoten
- ▶ 30 Petabyte Speicherkapazität
- ▶ Speichersystem Lustre

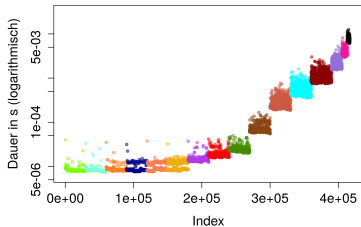
## Messungen sortiert nach Zugriffsgröße



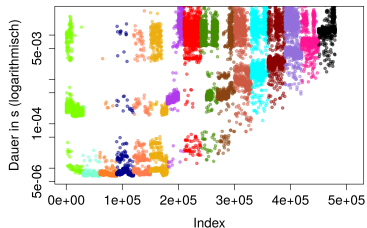
# Messungen sortiert nach Zugriffsgröße



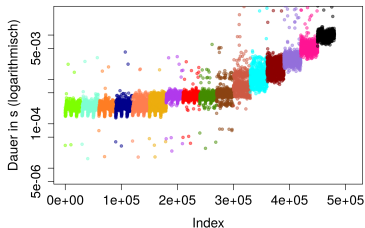
SEQ lesend



SEQ schreibend

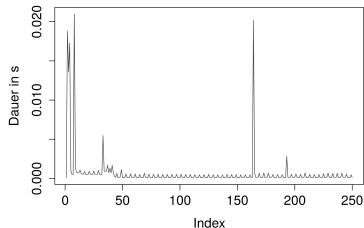


RND lesend

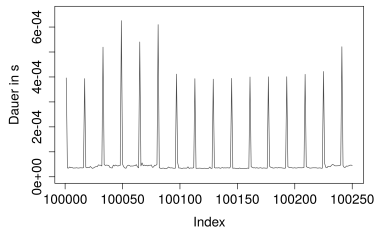


RND schreibend

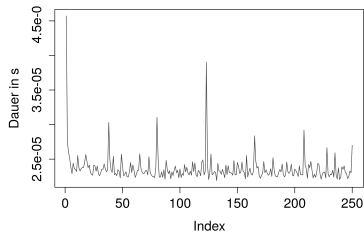
## Detailbetrachtung: 250 Messungen auf SEQ



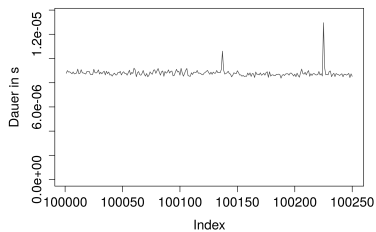
SEQ lesend



SEQ lesend

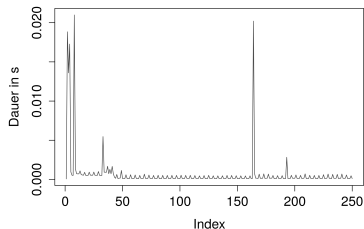


SEQ schreibend

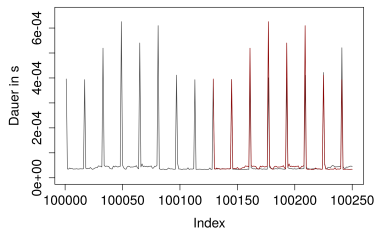


SEQ schreibend

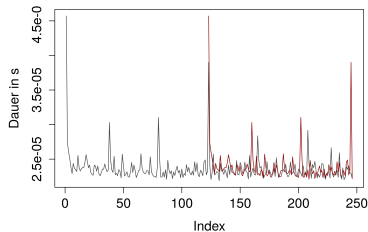
## Detailbetrachtung: 250 Messungen auf SEQ



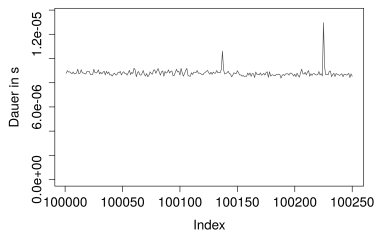
SEQ lesend



SEQ lesend



SEQ schreibend



SEQ schreibend

## Modelle zur Leistungsvorhersage

- ▶ Durchschnittliche Zugriffszeit  
(**Durchschnitt**)

## Modelle zur Leistungsvorhersage

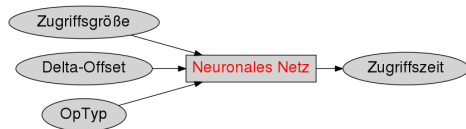
- ▶ Durchschnittliche Zugriffszeit (**Durchschnitt**)
- ▶ Lineare Regression nach der Zugriffsgröße (**LinReg G**)





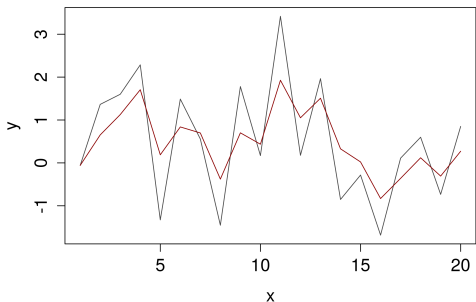
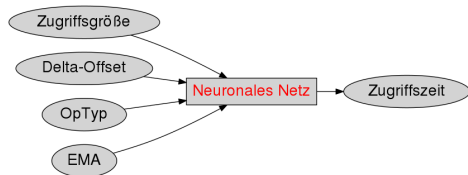
## Modelle zur Leistungsvorhersage

- ▶ Durchschnittliche Zugriffszeit (**Durchschnitt**)
- ▶ Lineare Regression nach der Zugriffsgröße (**LinReg G**)
- ▶ Einfaches Modell mit neuronalem Netz (**NN-Tupel1**)



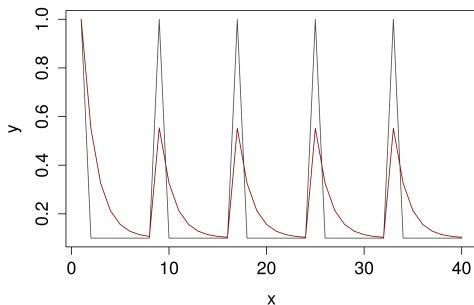
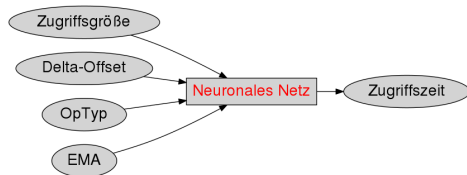
## Modelle zur Leistungsvorhersage

- ▶ Durchschnittliche Zugriffszeit (**Durchschnitt**)
- ▶ Lineare Regression nach der Zugriffsgröße (**LinReg G**)
- ▶ Einfaches Modell mit neuronalem Netz (**NN-Tupel1**)
- ▶ Ausnutzen zeitlicher Abhängigkeit mit dem *exponential moving average* (**NN-EMA**)

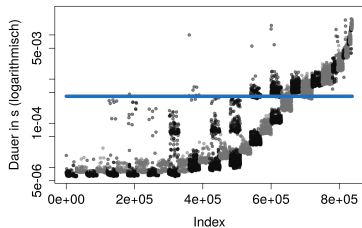


## Modelle zur Leistungsvorhersage

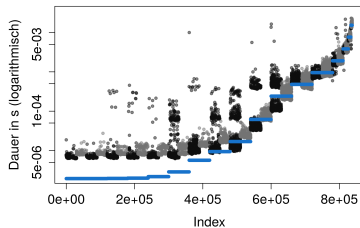
- ▶ Durchschnittliche Zugriffszeit (**Durchschnitt**)
- ▶ Lineare Regression nach der Zugriffsgröße (**LinReg G**)
- ▶ Einfaches Modell mit neuronalem Netz (**NN-Tupel1**)
- ▶ Ausnutzen zeitlicher Abhängigkeit mit dem *exponential moving average* (**NN-EMA**)



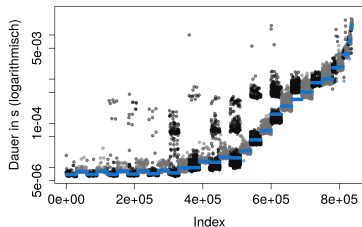
## Modelle angewendet auf SEQ



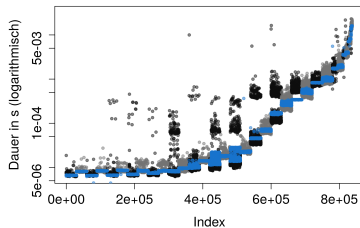
Durchschnitt



LinReg G

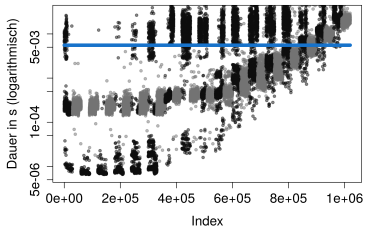


NN-Tupel1

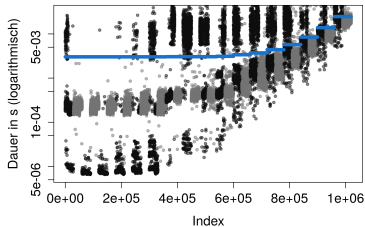


NN-EMA

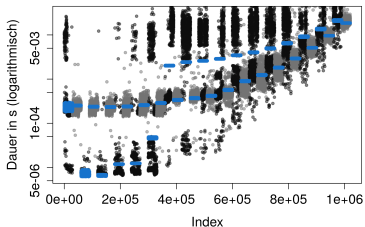
## Modelle angewendet auf RND



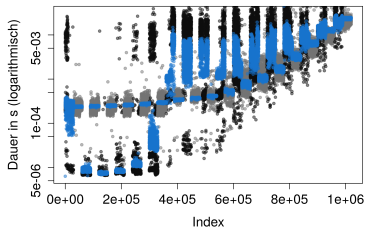
Durchschnitt



LinReg G



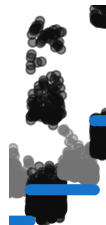
NN-Tupel1



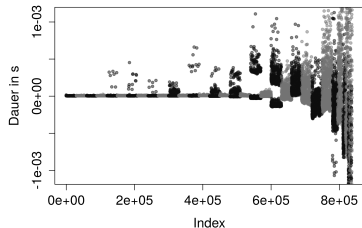
NN-EMA

## Einbeziehung des E/A-Pfads

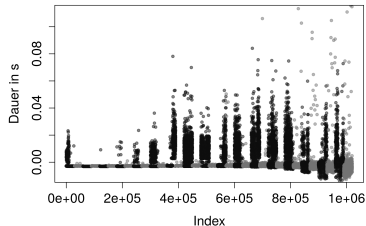
- ▶ Bisher keine Unterscheidung von E/A-Pfaden
- ▶ **Idee:** Residuen eines einfachen Modells zur Bestimmung einer Näherung des E/A-Pfades eines Dateizugriffs nutzen
- ▶ Clustering der Residuen mit *k-Means* in 10 Gruppen
  - ▶ Diese bezeichnen wir als **Fehlerklassen**
- ▶ Fehlerklassen sind nicht für die Vorhersage von Zugriffszeiten geeignet
  - ▶ Dafür kann eine Näherung des E/A-Pfades bestimmt werden



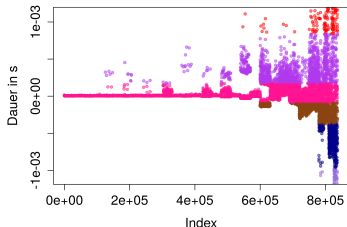
## Erstellung der Fehlerklassen



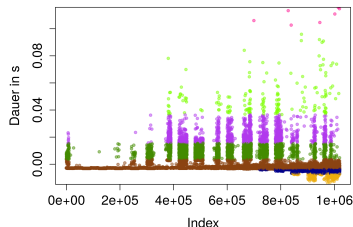
Residuen von LinReg G auf SEQ



Residuen von LinReg G auf RND



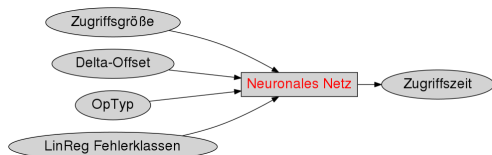
Farblich markierte Fehlerklassen



Farblich markierte Fehlerklassen

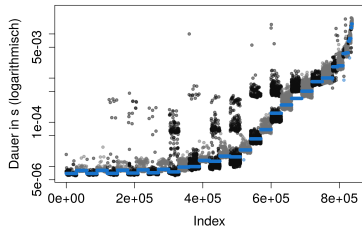
## Modell mit Fehlerklassen

- ▶ **Durchschnitt**
- ▶ **LinReg G**
- ▶ **NN-Tupel1**
- ▶ **NN-EMA**
- ▶ Einfaches Modell mit neuronalem Netz und Fehlerklassen (**NN-LinRegFK**)

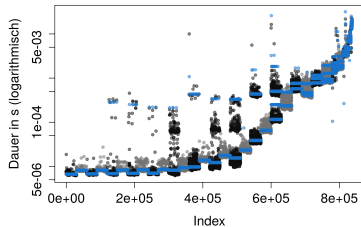




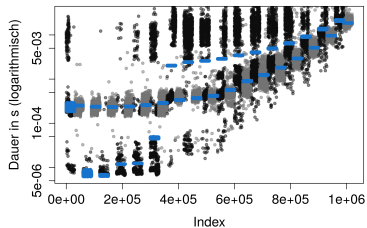
## Mit und ohne Fehlerklassen



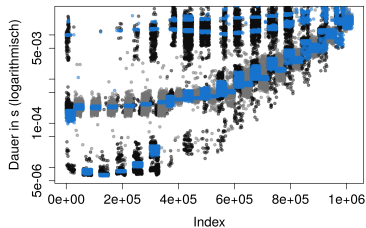
NN-Tupel1 auf SEQ



NN-LinRegFK auf SEQ



NN-Tupel1 auf RND



NN-LinRegFK auf RND

## Zusammenfassung

- ▶ Vorhersagen einfacher Modelle mit neuronalen Netzen gut für schreibende Zugriffe
- ▶ Lineare Modelle unzureichend
- ▶ E/A-Pfad entscheidender Faktor für die Zugriffszeiten
  - ▶ Keine Kenntnisse über den E/A-Pfad einer Messung gegeben
- ▶ Periodische Abhängigkeiten konnten nicht auf einfache Weise genutzt werden
- ▶ Unterscheidung von E/A-Pfaden durch Fehlerklassen möglich

Ende der Präsentation

## Modell mit Fehlerklassen auf SEQ

Modell	MAE (s)	MAPE (%)	MSPE (%)	RMax (%)
NN-LinRegFK	2.0e-05	8.6	14	275
NN-EMA	5.7e-05	13.7	22	2336
NN-Tupel1	6.0e-05	14.1	22	295
LinReg G	7.6e-05	50.8	59	326
Durchschnitt	5.9e-04	2939.6	3757	6537

- ▶ MAE: *mean absolute error*
- ▶ MAPE: *mean absolute percentage error*
- ▶ MSPE: *mean square percentage error*
- ▶ RMax: maximale relative Modellabweichung

## Modell mit Fehlerklassen auf RND

Modell	MAE (s)	MAPE (%)	MSPE (%)	RMax (%)
NN-LinRegFK	0.00103	31	119	4272
NN-Tupel1	0.00313	103	530	21786
NN-EMA	0.00305	86	619	45320
LinReg G	0.00476	5578.4	14185	46941
Durchschnitt	0.00692	10243.5	26035	86139

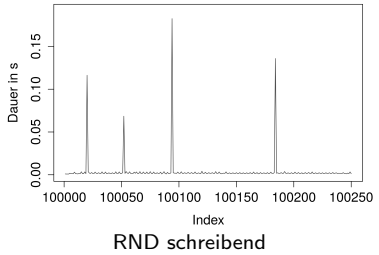
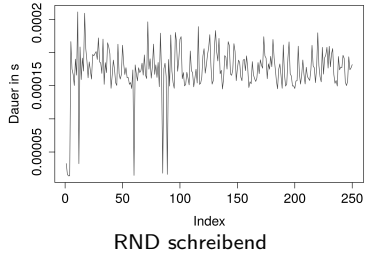
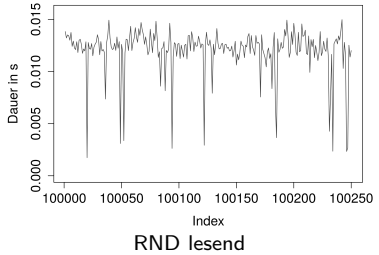
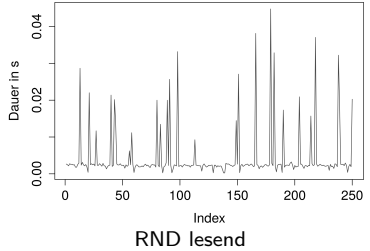
- ▶ MAE: *mean absolute error*
- ▶ MAPE: *mean absolute percentage error*
- ▶ MSPE: *mean square percentage error*
- ▶ RMax: maximale relative Modellabweichung

## Fehlerklassen auf SEQ

Klasse	gemittelte Angaben			Angaben zu den Residuen			zugeordnete Messungen
	Durchsatz (B/s)	Größe (B)	Dauer (s)	Min (s)	Durchschnitt (s)	Max (s)	auf SEQ
1	3.8e+09	1.7e+07	0.00452	-1.2e-02	-1.1e-02	-6.7e-03	101
2	1.3e+09	1.2e+07	0.00948	-6.0e-03	-1.4e-03	-1.0e-03	2128
3	1.3e+09	6.1e+06	0.00490	-1.0e-03	-6.6e-04	-4.7e-04	8689
4	1.7e+09	1.8e+06	0.00137	-4.7e-04	-2.8e-04	-2.0e-04	38820
5	1.6e+09	9.4e+05	0.00073	-2.0e-04	-1.3e-04	-5.9e-05	71237
6	4.3e+08	1.1e+05	0.00011	-5.9e-05	9.2e-06	2.2e-04	682997
7	6.8e+08	1.6e+06	0.00193	2.2e-04	4.3e-04	1.1e-03	29496
8	7.8e+08	7.5e+06	0.00869	1.1e-03	1.9e-03	3.7e-03	3527
9	6.5e+08	1.2e+07	0.01653	3.8e-03	5.7e-03	1.1e-02	527
10	1.6e+08	4.1e+06	0.02094	1.1e-02	1.7e-02	4.0e-02	78

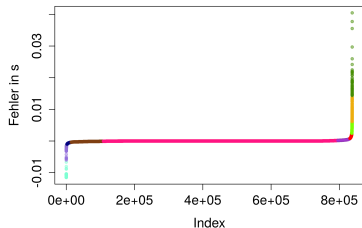
# Fehlerklassen auf RND

Klasse	gemittelte Angaben			Angaben zu den Residuen			zugeordnete Messungen auf RND
	Durchsatz (B/s)	Größe (B)	Dauer (s)	Min (s)	Durchschnitt (s)	Max (s)	
1	1.4e+09	1.5e+07	0.0130	-0.0210	-0.0090	-0.0069	9467
2	9.9e+08	9.1e+06	0.0101	-0.0069	-0.0047	-0.0036	54371
3	2.3e+08	1.4e+06	0.0024	-0.0036	-0.0025	0.0036	825974
4	5.5e+07	1.2e+06	0.0143	0.0036	0.0096	0.0156	85462
5	6.1e+07	2.3e+06	0.0276	0.0156	0.0216	0.0366	37862
6	5.8e+07	4.0e+06	0.0598	0.0366	0.0516	0.0976	4695
7	3.2e+07	4.7e+06	0.1528	0.0977	0.1438	0.2066	1443
8	4.5e+06	1.2e+06	0.2741	0.2067	0.2696	0.4728	567
9	3.0e+05	1.6e+05	0.6956	0.4822	0.6923	1.0063	123
10	9.4e+02	1.0e+03	1.3627	1.0396	1.3597	2.1216	36

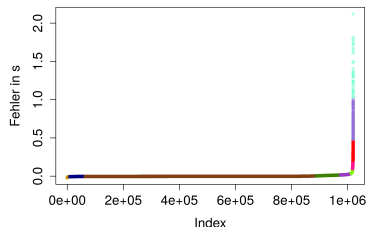




## Darstellung der Fehlerklassen nach sortierten Residuen



sortierte Residuen auf SEQ



sortierte Residuen auf RND

## Ausblick

- ▶ Periodische Zusammenhänge im E/A-System besser ausnutzen
- ▶ Zusammenhang zwischen Fehlerklassen und E/A-Pfaden genauer untersuchen
- ▶ Fehlerklassen aus Residuen verschiedener Modelle vergleichen
- ▶ Problem beheben, dass die Residuen größerer Laufzeiten die Fehlerklassen dominieren