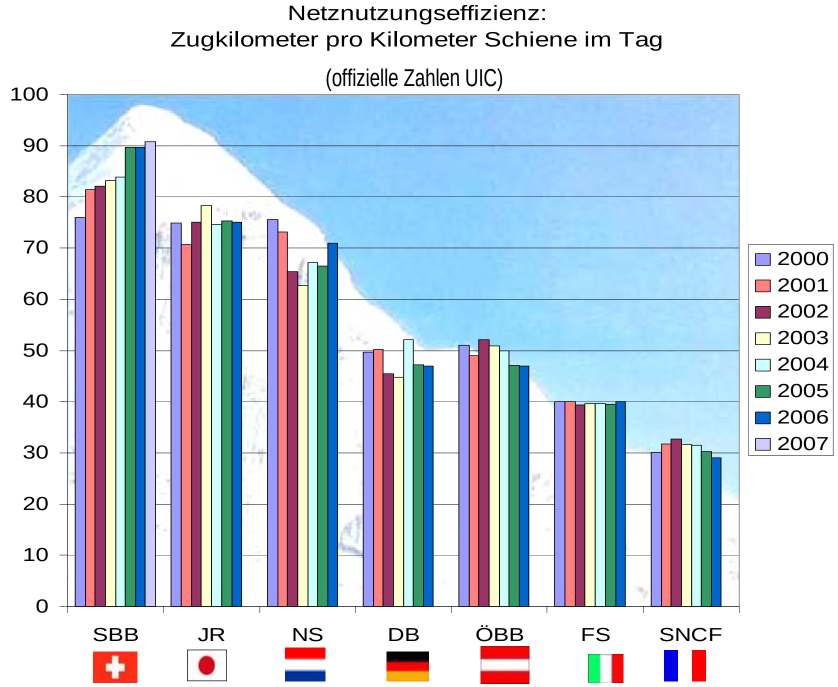
Finding Robust Train Paths in Dense Corridors

Holger Flier (ETH)
Thomas Graffagnino (SBB)
Marc Nunkesser (ETH)

Motivation and Problem Statement

- Demand for passenger train transportation increases faster than infrastructure can be extended
- Traffic gets denser
- Risk of delay increases



Problem: Find a schedule for an additional train on a dense corridor that has a low risk of being delayed upon arrival at the final station

Basic Idea

- Historical delay data contains implicit information
  - Dependencies between trains
  - Resource bottlenecks
  - Dispatching decisions
  - …

- Learn linear regression models from the data
  - to predict the risk of delay of a train
- Combine statistical and combinatorial models
  - to find optimal suggestions for train paths

LINEAR REGRESSION MODELS

Linear Regression Models

- Goal: “predict” the delay of a train for *a day in the past*
- Linear regression models use predictors from the data

$$\hat{\delta} = \alpha + \sum_i^n \beta_i \, \mathrm{predictor}_i$$

Linear Regression Models:
Predictors from the Data

- Train density around planned arrival/departure time
- Timetable measures
  - Scheduled time difference to previous / next trains in the same / opposite direction
  - Slack time
- Delays of neighbor trains
- Track properties
  - Average increase/decrease of delay of all trains between two operating points, scheduled one hour around the predicted train
- Previous delay

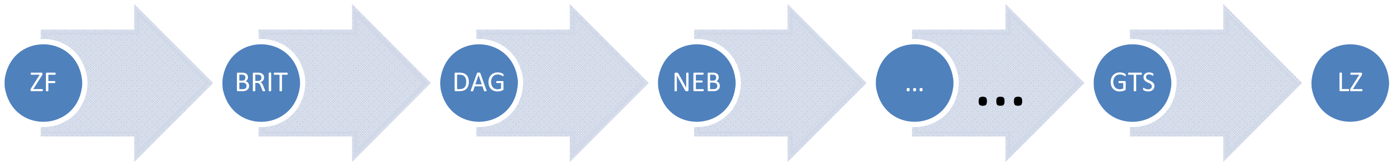
Linear Regression Models

- Goal: "predict" the delay of a train for *a day in the past*

- Linear regression models use predictors from the data

$$\hat{\delta} = \alpha + \sum_i^n \beta_i \,\mathrm{predictor}_i$$

- Set up a series of in-station and between-station models

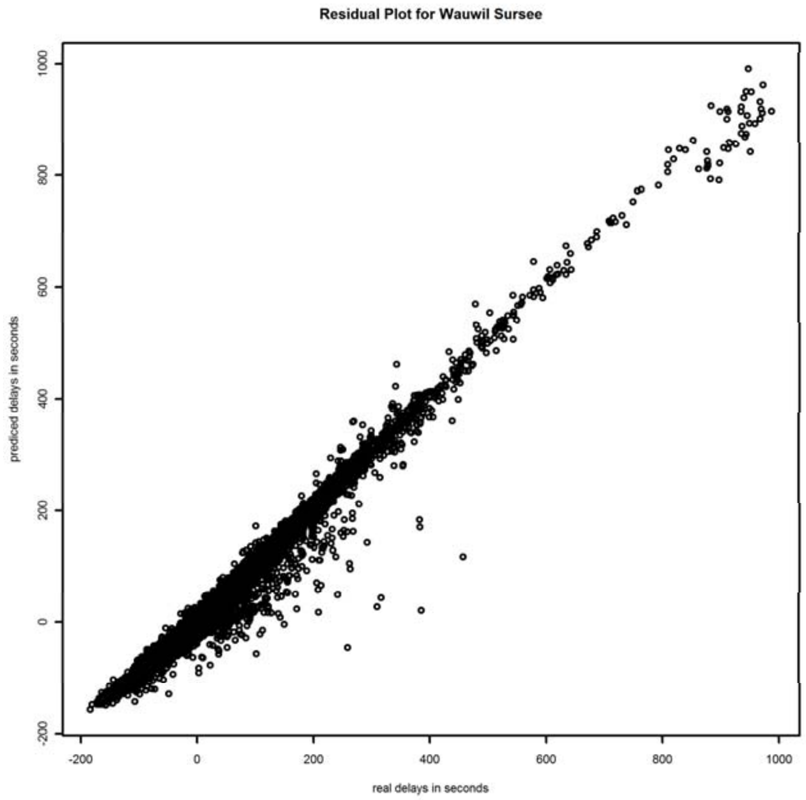


- Special predictor *previous delay* concatenates models

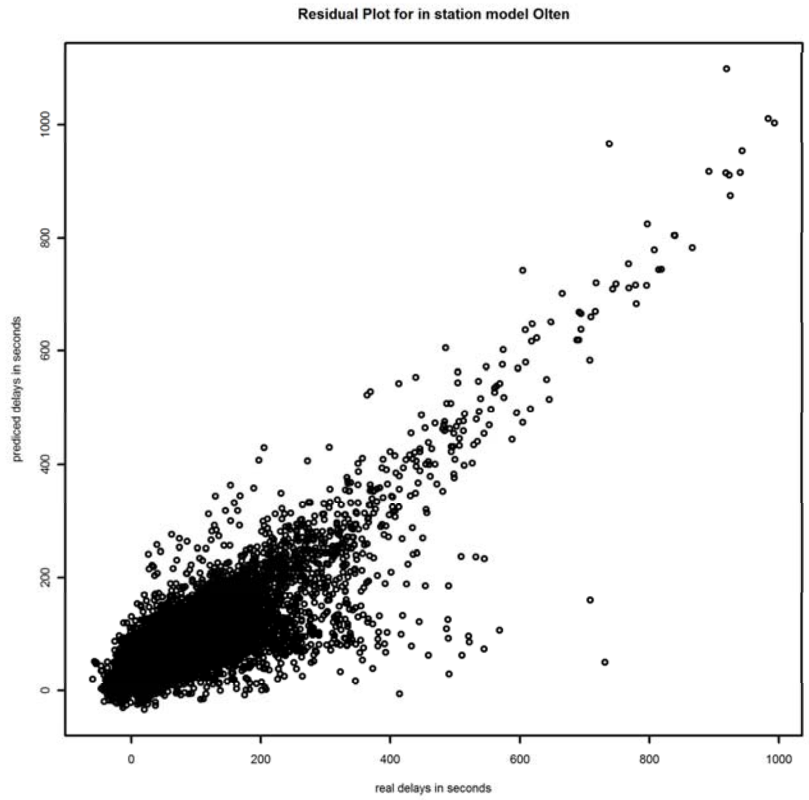
$$\hat{\delta}_i = \mathrm{model}(\Theta, d, \pi, \hat{\delta}_{i-1})$$

Linear Regression Models: Residual Plots

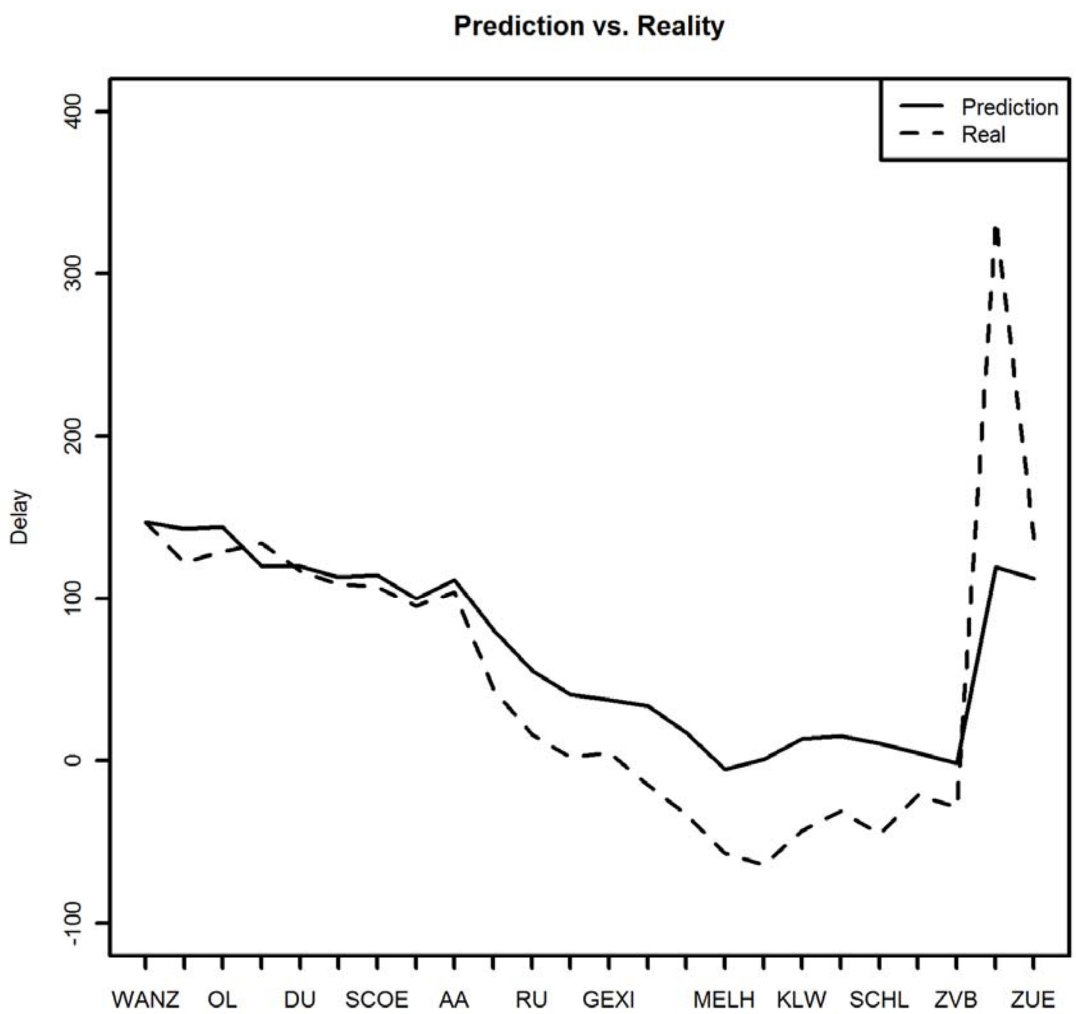
Wauwil – Sursee



Olten

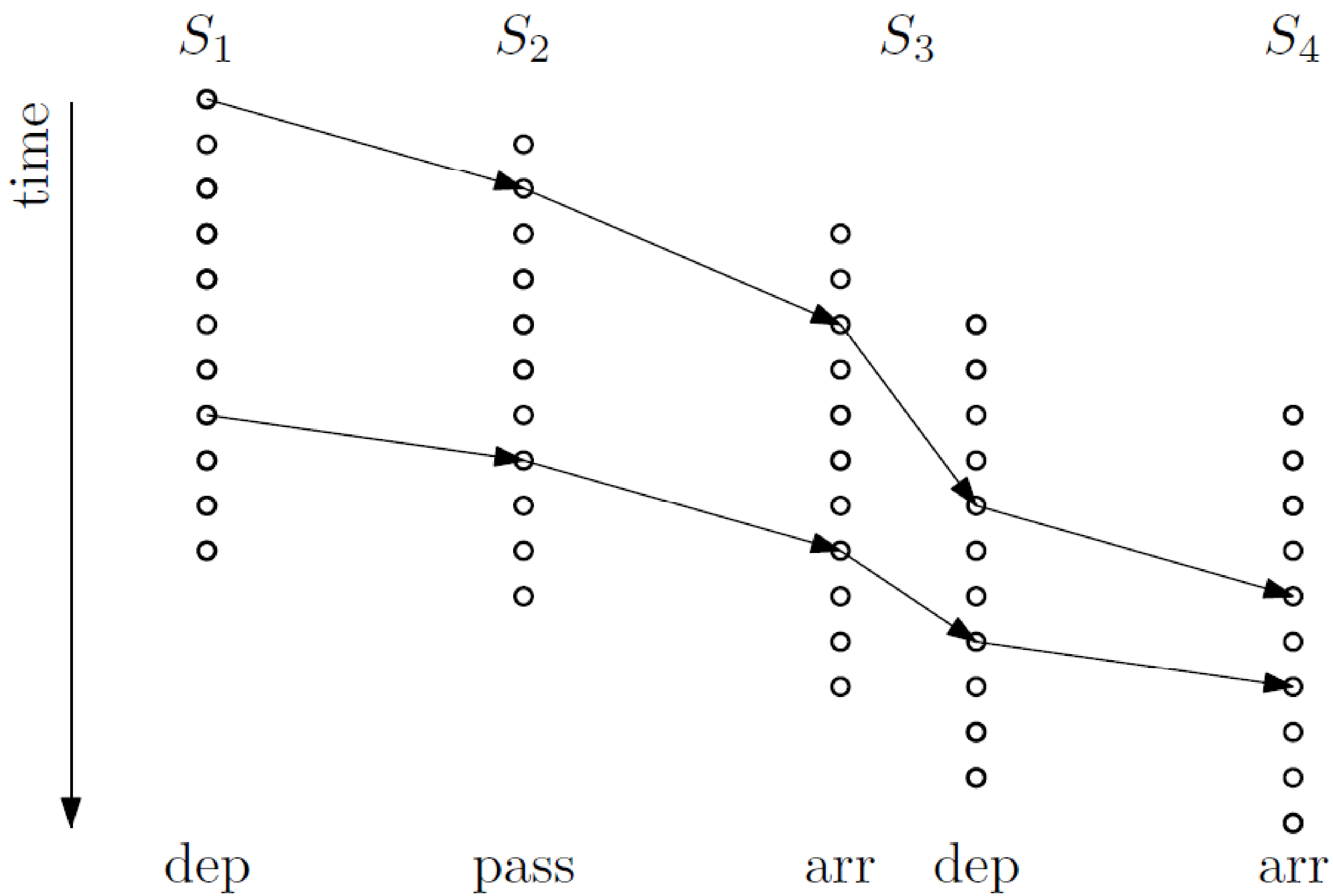


Sequence of Linear Regression Models: Quality of Prediction

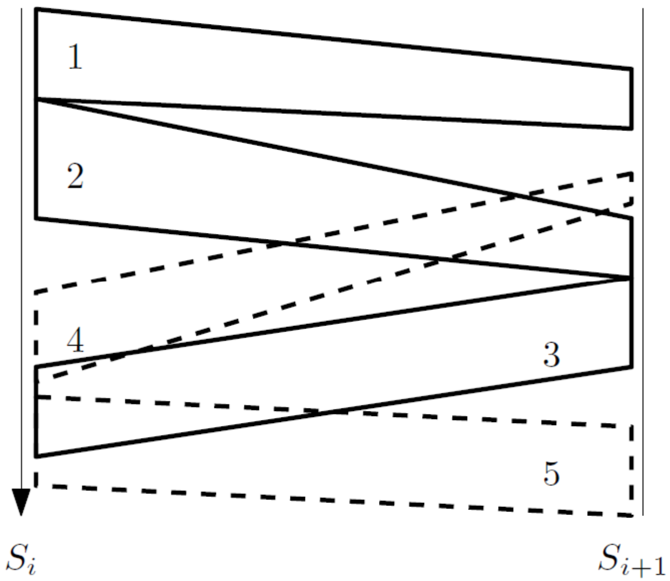


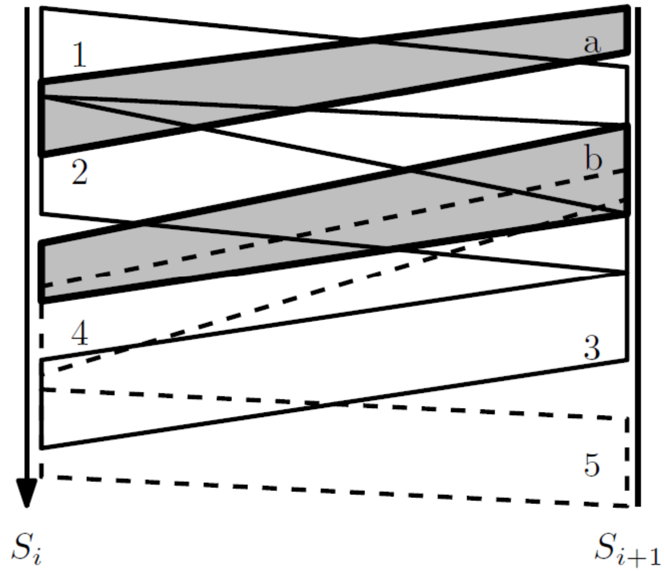
COMBINATORIAL MODEL

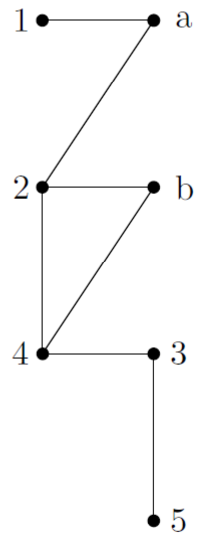
Shortest Path Model



Shortest Path Model: Tracks and Trapezoids







Minimum Risk Problem

Minimize the predicted delay of the additional train upon arrival at the final station for a

a) single day (P)

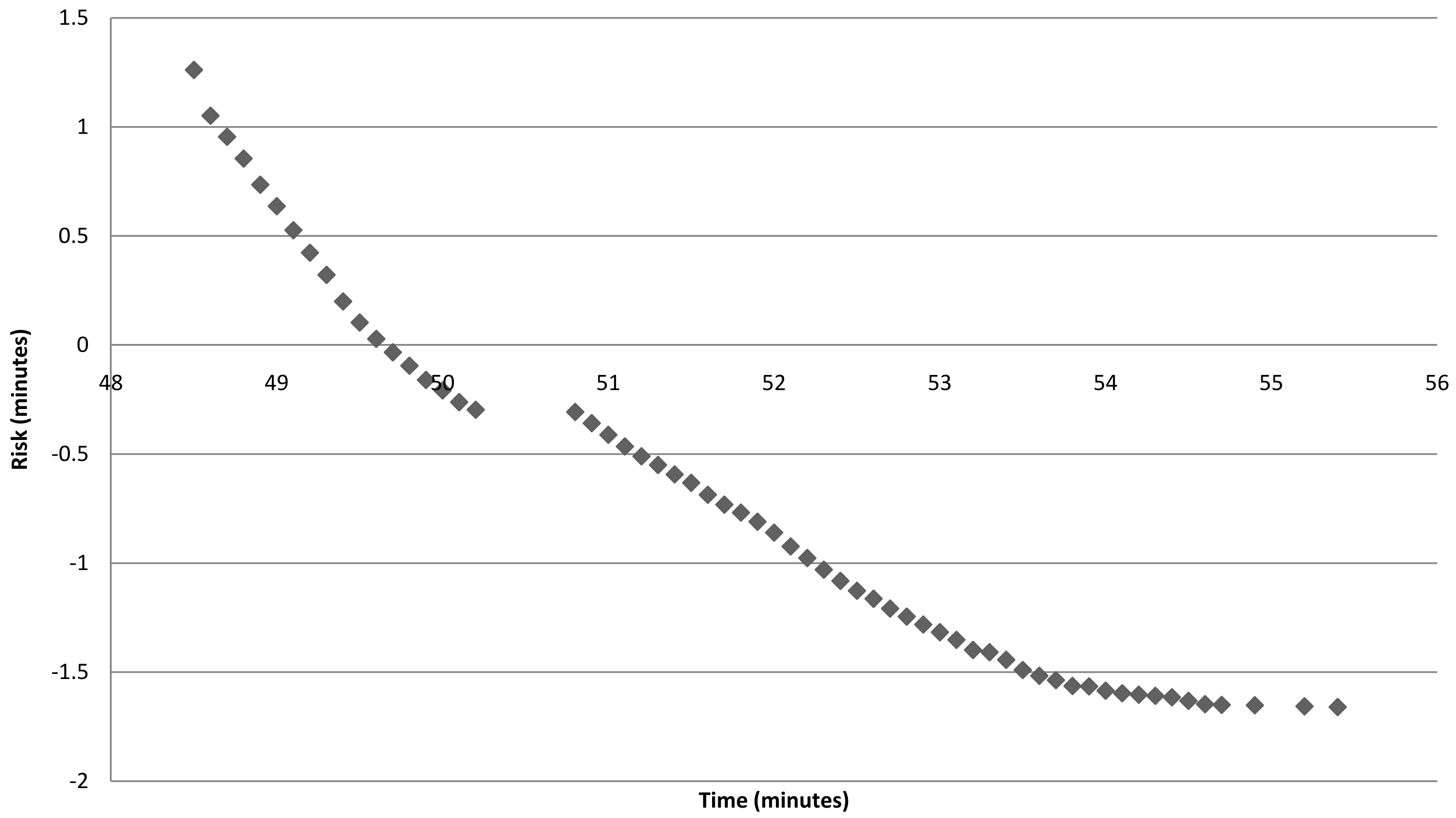
b) set of days:

   i. Average of the predicted delays (P)

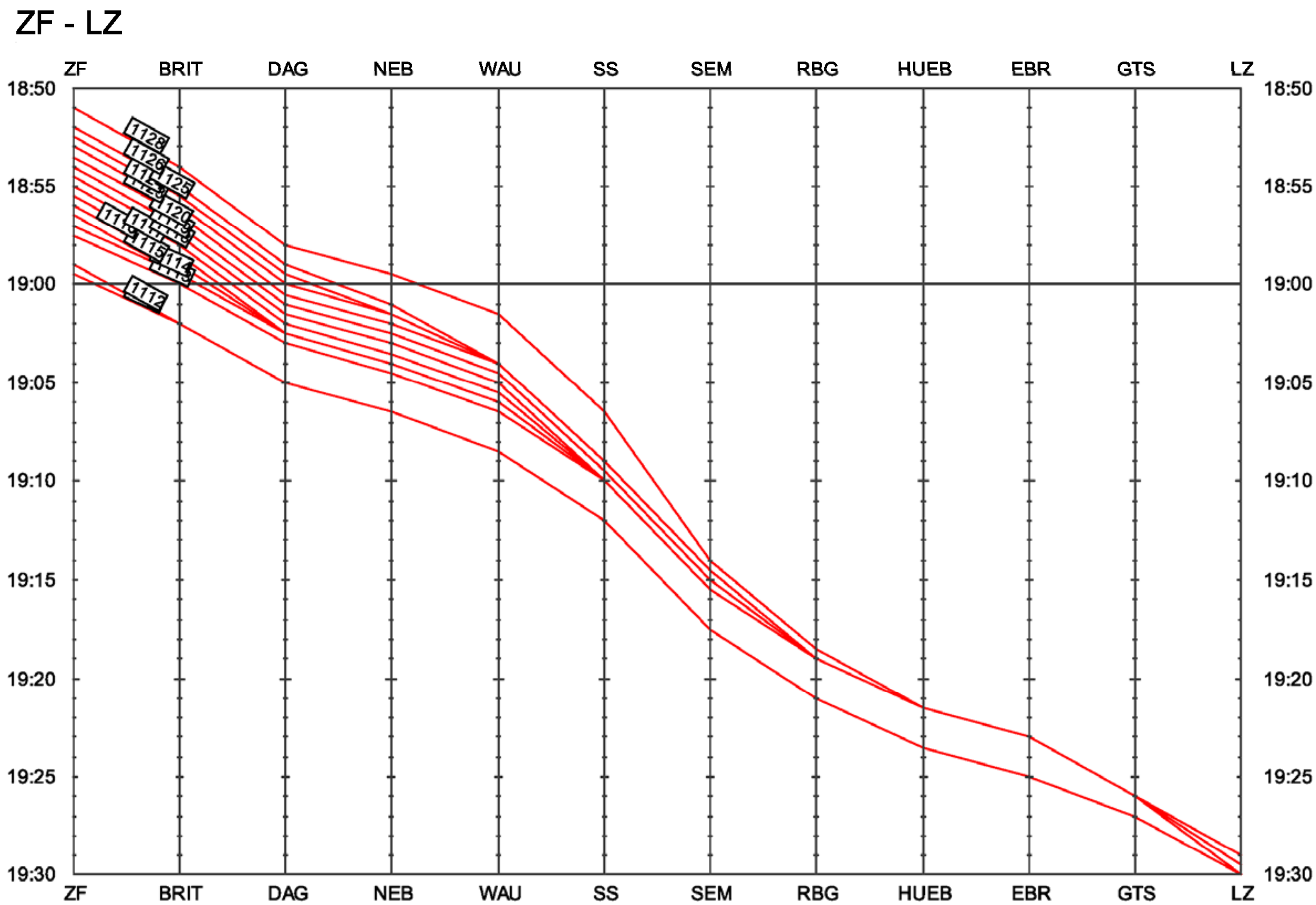
   ii. Median of the predicted delays (NP-hard)

(using basic prediciton models)

Travel Time vs. Risk:
Pareto Optimal Solutions



Travel Time vs. Risk: Pareto Optimal Solutions



Conclusion and Outlook

- Novel approach to minimize risk of delay
  - Combination of linear regression and a shortest path algorithm yielding Pareto optimal solutions w.r.t. travel time vs. risk of delay
- Profit from existing delay data
- Keep infrastructure modeling efforts to a minimum
- Reasonable risk estimates instead of simulation
  - Detailed simulation only for most promising suggestions

- Outlook and future research:
  - Integrate approach into existing planning tools at SBB
  - Modified risk measure, including delay of follow-up trains
  - Option to remove trains of less priority


Thank you!

Linear Regression Models: Achievable Quality

From	To	SE	DoF	Min	1Q	2Q	3Q	Max	R^2
ZF	BRIT	9.2	11903	-27.8	-5.1	-1.2	3.7	126	0.9931
BRIT	DAG	12.6	11903	-61.5	-4.8	-1.3	3.5	295.9	0.9874
DAG	NEB	8.2	11902	-57.5	-3.5	-0.8	2.3	219.2	0.9948
NEB	WAU	6.2	11902	-33.8	-3.4	-0.5	2.7	177.2	0.9971
WAU	SS	18.3	11898	-105.1	-8.5	-2.3	5.4	370.5	0.9778
SS	SEM	27.8	14273	-108.0	-14.6	-2.3	11.4	708.9	0.9403
SEM	RBG	21.1	14274	-141.3	-7.9	-2.2	3.5	1032	0.9686
RBG	HUEB	25.2	14274	-102.1	-11.4	-0.8	8.3	382.9	0.9582
HUEB	EBR	16.3	19964	-58.5	-7.7	-1.7	4.6	663.7	0.9820
EBR	GTS	49.0	19966	-223.1	-27.4	-5.9	18.9	587.5	0.8342
GTS	LZ	41.2	37761	-202.4	-22.5	-7.3	14	760.5	0.8769