

Results:

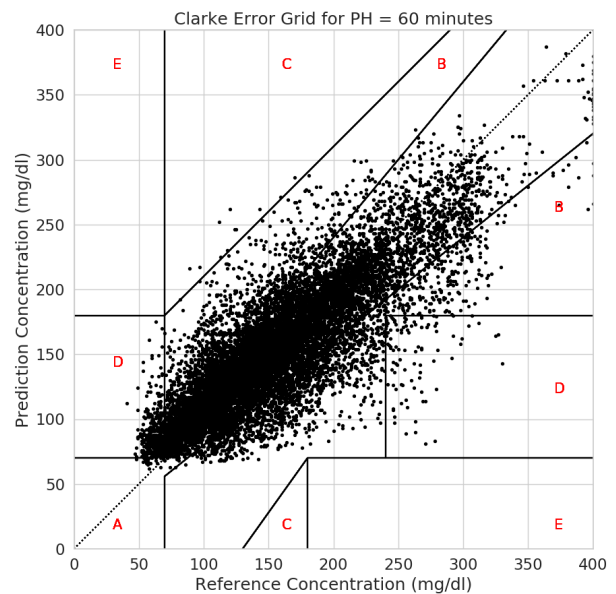
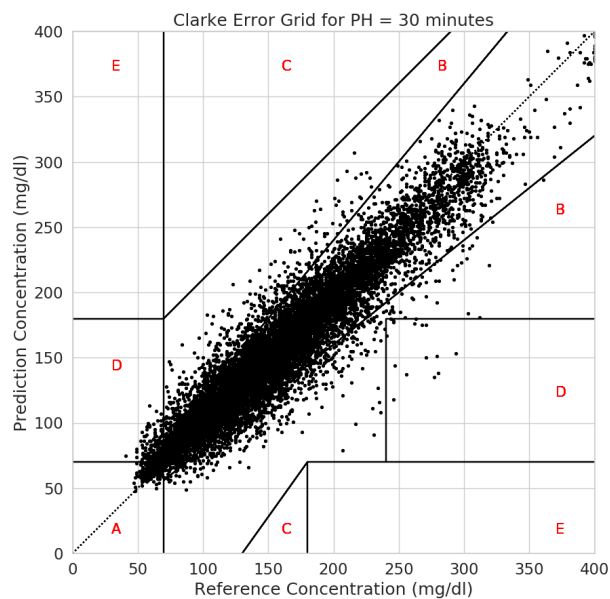
Model = RNN

Time-series = multi-variate

Forecasting = single-step

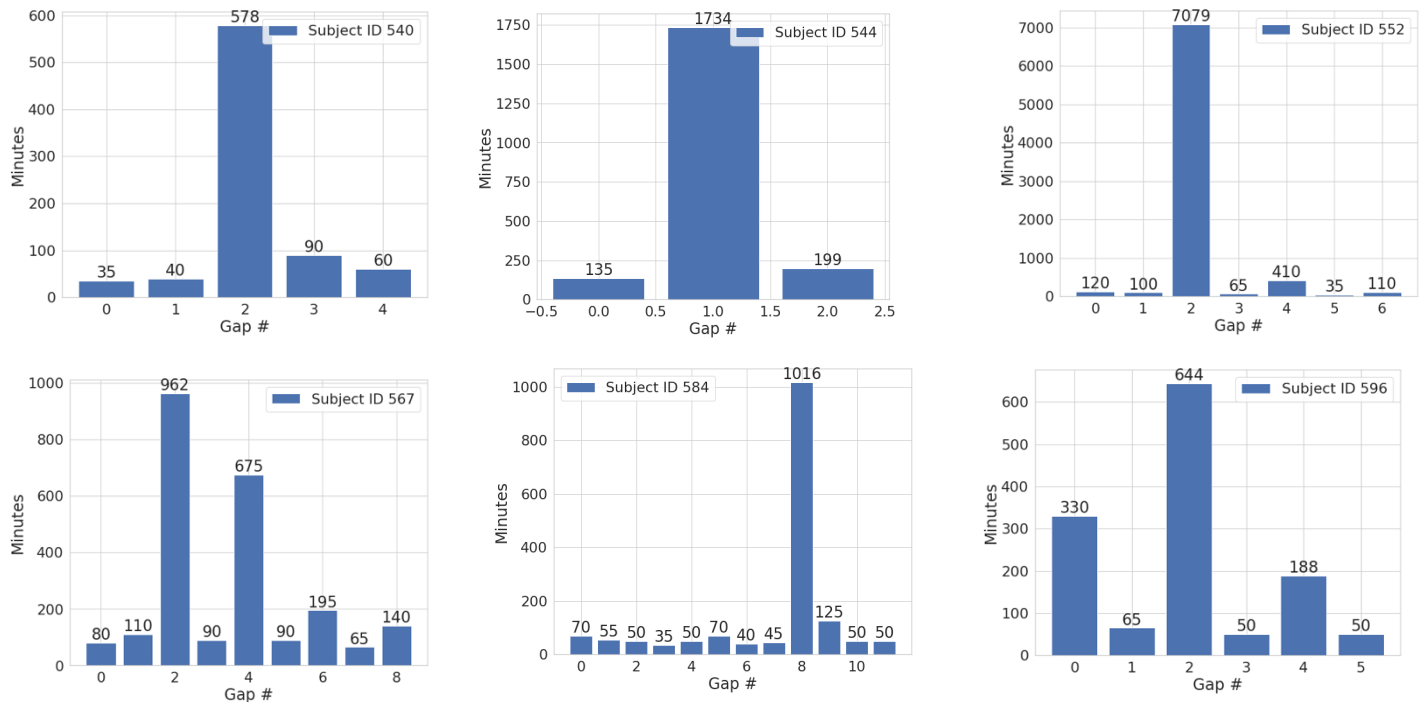
Extrapolation = Yes

	PH = 30 minutes		PH = 60 minutes	
Subject ID	RMSE	MAE	RMSE	MAE
540	19.56471436	14.01250434	33.44946507	25.13679411
544	16.58024153	11.52241571	27.8058579	20.30795073
552	15.06236149	11.151683	26.8045382	20.2190352
567	23.0919235	14.69224283	36.53305452	25.72762646
584	25.17940162	16.14048338	36.27235085	26.35526316
596	15.8717577	11.00220103	26.64410591	19.08752328
Mean:	19.2250667	13.08692172	31.25156208	22.80569882



Note on the number of test datapoints and extrapolating across large gaps:

The plots below show the gaps present in test data for each of the 6 subjects. For example, in the first plot for subject 540, we can see that there are 5 instances where there were gaps of 35, 40, 578, 90 and 60 minutes between adjacent recordings, respectively. Extrapolating across these large gaps that sometimes last for hours, even a day, can deteriorate the performance of the model.



If no extrapolation is done, the results are as follows:

	PH = 30 minutes		PH = 60 minutes	
Subject ID	RMSE	MAE	RMSE	MAE
540	18.62839239	13.63407722	33.01446038	24.84054344
544	16.56833717	11.46533534	27.81626299	20.26159696
552	14.92218959	10.99776086	26.66894909	19.96294602
567	19.34821554	13.32687978	34.71675566	24.73854697
584	21.5495254	15.17512077	34.93184666	25.51691419
596	15.68574569	10.90106545	26.72131388	19.12374323
Mean:	17.7837343	12.58337324	30.64493144	22.4073818

Summary of the methods used (detailed in systems description paper)

Four machine learning pipelines were tried and tested. A bigger open-source dataset OpenAPS was used to pre-train a model which was tested on OhioT1DM with and without additional training (approach II and III respectively). Their performance was comparable to training a model using OhioT1DM only (approach I)

