Robust Gaussian Process Regression with Huber Likelihood

To comprehensively evaluate the performance of HuberLA and HuberMCMC, we expanded our experiments to include various outlier scenarios. For comparative analysis, we also incorporated the robust and conjugate Gaussian process (RCGP) and its sparse variant, RCSVGP, models proposed by Altamirano et al. (2024). To further validate our findings, experiments were extended to the Energy and Yacht datasets characterized by asymmetric and focused outlier distributions from the study of Altamirano et al.

1 Neal

We experimented on Neal dataset by introducing bad data points into the training dataset for the following cases

- **Case 1: significant vertical outliers with bad leverage.** Bad data points were added to both input and output dimensions, with the latter being significantly distant from the main data cluster.
- **Case 2: significant vertical outliers without bad leverage.** Only output dimensions were contaminated with bad data points.
- **Case 3: near vertical outliers with bad leverage.** Bad data points were added to both input and output dimensions, with the latter being relatively close to the main data cluster compared to Case 1.
- **Case 4: near vertical outliers without bad leverage.** Only output dimensions were contaminated with data points relatively close to the data cloud compared to Case 1.

For each of the four considered error distributions—$\mathcal{N}(0.01, 0.08)$, Student-t(10),Laplace$(0, 0.1)$, Student-t(1)—experiments were conducted across all bad data cases.

1.1 When is HuberMCMC better?

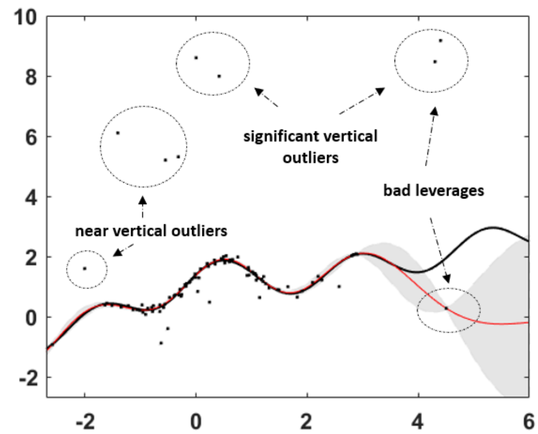
In scenarios with significant vertical outliers with bad leverage (Case 1), HuberMCMC performed better than HuberLA (see, Tables 1 and 2). HuberMCMC also outperformed tLA in terms of predictive accuracy, demonstrating a more robust fit that is less influenced by bad leverage points (Figure 1) HuberLA generally provided more reliable uncertainty quantification compared to HuberMCMC (Figures 1 and 2), while maintaining competitive predictive performance. In scenarios with significant vertical outliers but without influential points, HuberMCMC exhibited superior performance across Student's-t, Laplace, and Cauchy error distributions (see, Table 2). This suggests that HuberMCMC is a robust choice for datasets containing extreme outliers in the response variable.

1.2 When is HuberLA better?

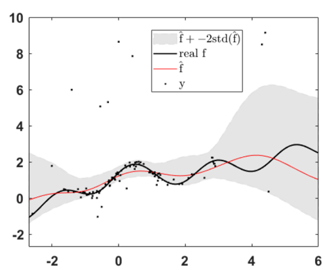
HuberLA exhibited superior performance in handling near vertical outliers compared to HuberMCMC (Tables 3 and 4). Figure 3 highlights HuberLA's robustness to bad leverage points, in contrast to tLA which is influenced by such points. While HuberLA generally provided more accurate predictions and reliable uncertainty quantification than both HuberMCMC and tLA, HuberMCMC demonstrated competitive performance in terms of RMSE, MAE, and NLP (Table 3).

Preprint. Under review (Rebuttal phase).

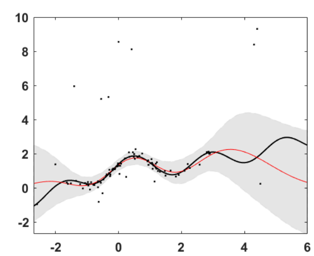
When we did not add leverage points (Cases 2 and 4), HuberLA and HuberMCMC exhibited performance comparable to other models, indicating their robustness to exclusively vertical outliers. Compared to the RCGP model, both HuberLA and HuberMCMC consistently produced better predictive performance across all outlier conditions and error distributions when applied to the Neal dataset.



(a)

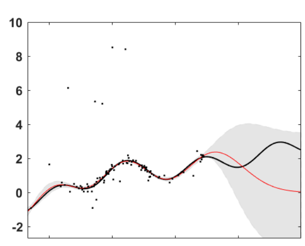


(b)

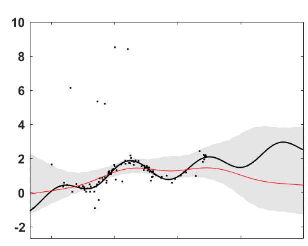


(c)

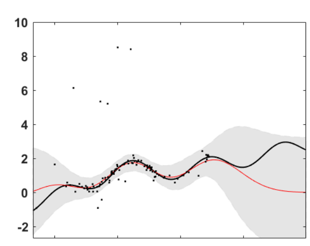
Figure 1: Predicted values for (a) tLA; (b) HuberMCMC; (c) HuberLA with standard deviations for the case 1 with error following Student's t distribution on Neal dataset.



(a)



(b)



(c)

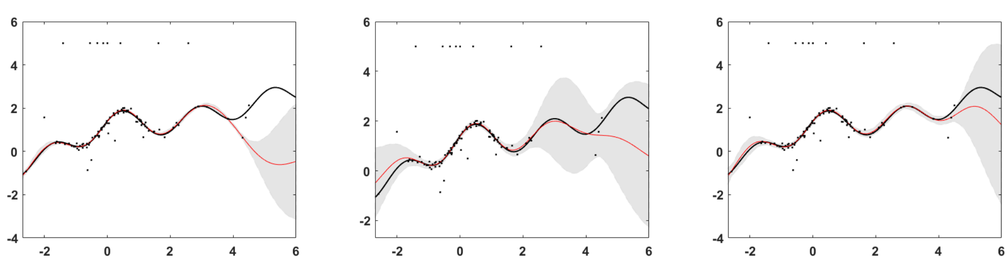
Figure 2: Predicted values for (a) tLA; (b) HuberMCMC; (c) HuberLA with standard deviations for the case 2 with error following Student's t distribution on Neal dataset.

Table 1: Case 1: Significant vertical outliers with bad leverage.

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP
				$e \sim \mathcal{N}(0.01, 0.08)$					
RMSE	0.74	0.61	0.75	**0.37**	0.73	1.44	0.43	0.59	1.84
MAE	0.47	0.36	0.48	**0.31**	0.52	1.24	0.33	0.43	1.48
NLP	0.80	**0.34**	7.38	0.66	0.63	2.03	0.52	0.36	2.94
				$e \sim$ Student-t(10)					
RMSE	4.86	1.192	1.22	**0.50**	1.17	1.52	0.59	1.01	1.89
MAE	1.67	0.728	0.77	**0.41**	0.79	1.34	0.43	0.67	1.71
NLP	134.90	0.76	6.80	0.82	1.296	2.09	**0.71**	0.73	4.13
				$e \sim$ Laplace$(0, 0.1)$					
RMSE	4.76	1.12	1.23	**0.58**	1.17	1.51	1.06	1.07	1.95
MAE	1.64	0.69	0.76	**0.44**	0.81	1.33	0.75	0.75	1.72
NLP	6.49	**0.76**	6.73	0.85	1.36	2.08	0.92	0.93	4.18
				$e \sim$ Student-t(1)(Cauchy)					
RMSE	4.75	1.19	1.25	**0.61**	1.20	1.50	0.42	1.08	1.97
MAE	1.65	0.74	0.78	**0.47**	0.83	1.30	0.32	0.76	1.78
NLP	4.94	0.86	4.52	0.84	1.45	2.05	**0.81**	0.95	4.91

Table 2: Case:2 significant vertical outliers without bad leverage.

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP
				$e \sim \mathcal{N}(0.01, 0.08)$					
RMSE	1.41	1.42	**1.30**	1.40	1.36	1.74	1.38	1.34	2.04
MAE	0.90	0.95	**0.81**	0.99	0.94	1.51	0.98	0.93	1.78
NLP	**0.77**	1.04	1.89	1.22	1.50	2.16	1.31	1.21	4.32
				$e \sim$ Student-t(10)					
RMSE	1.22	1.22	1.14	**0.91**	1.12	1.66	1.01	1.11	2.04
MAE	0.63	0.63	0.56	0.62	0.62	1.45	**0.58**	0.63	1.76
NLP	-0.48	-0.48	-0.49	0.78	0.64	2.07	0.58	0.56	4.59
				$e \sim$ Laplace$(0, 0.1)$					
RMSE	1.38	1.41	2.73	**1.33**	1.37	1.73	1.33	1.34	2.06
MAE	**0.88**	0.93	1.82	0.97	0.96	1.51	0.95	0.95	1.80
NLP	**0.57**	0.89	78.73	1.24	1.68	2.16	1.12	1.32	4.25
				$e \sim$ Student-t(1)(Cauchy)					
RMSE	4.74	1.67	2.11	**1.33**	1.38	1.75	**1.33**	1.37	2.11
MAE	1.67	1.11	1.36	0.96	0.98	1.51	**0.95**	0.97	1.84
NLP	3.92	**1.12**	1.37	1.17	1.73	2.17	1.21	1.36	3.94



(a) (b) (c)

Figure 3: Predicted values for (a) tLA; (b) HuberMCMC; (c) HuberLA with standard deviations for the case 3 with error following Student's t distribution on Neal dataset.

Table 3: Case 3: Near vertical outliers with bad leverage.

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP
				$e \sim \mathcal{N}(0.01, 0.08)$					
RMSE	0.52	0.52	1.31	0.42	**0.25**	0.90	0.46	0.67	0.82
MAE	0.25	0.25	0.61	0.25	**0.14**	0.61	0.25	0.34	0.57
NLP	-1.14	-1.14	7.8	-0.41	-0.65	**0.84**	-0.53	-0.45	0.88
				$e \sim$ Student-t(10)					
RMSE	11.56	1.56	1.31	0.81	**0.37**	0.9	0.38	0.67	0.85
MAE	1.25	1.25	0.61	0.39	**0.18**	0.6	0.22	0.35	0.61
NLP	32.92	2.92	7.79	-0.35	-1.02	**0.84**	-0.52	-0.45	0.87
				$e \sim$ Laplace$(0, 0.1)$					
RMSE	0.48	0.48	1.31	0.42	**0.35**	0.89	0.82	0.67	0.86
MAE	0.23	0.23	0.61	0.24	**0.18**	0.6	0.41	0.34	0.61
NLP	-1.21	-1.21	7.76	-0.47	-0.77	**0.84**	-0.31	-0.45	0.88
				$e \sim$ Student-t(1)(Cauchy)					
RMSE	0.57	0.57	1.32	0.49	**0.17**	0.89	0.75	0.67	0.86
MAE	0.27	0.27	0.61	0.27	**0.11**	0.6	0.38	0.34	0.62
NLP	-1.06	-1.06	7.9	-0.46	-0.84	**0.84**	-0.19	-0.45	0.89

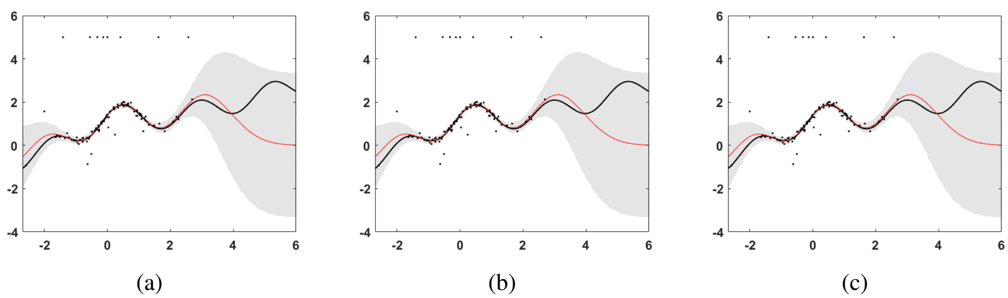


Figure 4: Predicted values for (a) tLA; (b) HuberMCMC; (c) HuberLA with standard deviations for the case 4 with error following Student's t distribution on Neal dataset.

Table 4: Case 4: Near vertical outliers without bad leverage.

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP
				$e \sim \mathcal{N}(0.01, 0.08)$					
RMSE	1.02	1.02	**1.01**	1.48	1.06	1.17	0.98	1.08	1.10
MAE	**0.51**	0.51	0.52	0.79	0.54	0.78	0.53	0.56	0.76
NLP	-0.48	-0.48	-0.44	**0.14**	-0.33	1.03	-0.02	1.05	1.01
				$e \sim$ Student-t(10)					
RMSE	1.58	1.58	1.02	1.17	1.13	1.17	**0.61**	1.07	1.11
MAE	1.28	1.28	0.52	0.67	0.57	0.78	**0.35**	0.56	0.76
NLP	4.11	4.11	-0.43	**0.13**	-0.14	1.03	-0.11	-0.06	1.03
				$e \sim$ Laplace$(0, 0.1)$					
RMSE	1.04	1.04	**1.01**	1.06	1.05	1.17	1.18	1.07	1.16
MAE	0.51	**0.51**	0.52	0.59	0.53	0.78	0.66	0.56	0.76
NLP	-0.47	-0.47	-0.44	**0.13**	-0.34	1.03	0.16	-0.06	**1.02**
				$e \sim$ Student-t(1)(Cauchy)					
RMSE	1.58	1.58	**1.02**	1.18	**1.02**	1.17	1.04	1.07	1.10
MAE	1.28	1.28	0.52	0.63	0.53	0.78	**0.52**	0.56	0.76
NLP	5.04	5.04	-0.41	**0.06**	-0.18	1.03	0.07	-0.05	1.02

2 Additional results on Energy and Yacht dataset

2.1 Energy

Table 5: RMSE, MAE, and NLP values for Energy dataset

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP	RCSVGP
					Asymmetric Outliers					
RMSE	0.03	3.92	0.95	0.08	0.28	0.94	6.91	0.23	1.43	**0.00029**
MAE	0.02	3	0.85	4.84	0.07	0.71	5.07	0.14	1.23	**0.00029**
NLP	-1.84	74.22	240.26	226.45	**0.84**	1.71	547.64	2.21	1.19	1691.56
					Focused Outliers					
RMSE	0.03	0.12	0.05	0.15	0.05	0.03	0.14	7.76	0.95	**0.0038**
MAE	0.02	0.09	0.03	0.12	0.04	0.03	0.11	4.58	0.86	**0.0038**
NLP	-1.94	-0.66	-1.65	-0.18	-1.47	-1.80	**0.11**	30.82	0.32	128.35

2.2 Yacht

Table 6: RMSE, MAE, and NLP values for Yacht

	SCtMCMC	SCt4MCMC	tLA	HuberMCMC	HuberLA	GP	LaplaceMCMC	LaplaceEP	RCGP	RCSVGP
					Asymmetric Outliers					
RMSE	0.01	0.26	0.03	0.59	0.57	0.8	0.39	0.39	0.43	**0.03**
MAE	0.01	0.16	0.02	0.42	0.29	0.67	0.22	0.24	0.31	**0.01**
NLP	-2.87	0.65	-2.13	1.05	0.73	3.6	0.45	**0.44**	0.74	41.0
					Focused Outliers					
RMSE	0.55	0.43	**0.13**	0.66	0.72	0.56	0.58	2.16	0.38	0.62
MAE	0.13	0.18	**0.07**	0.37	0.28	0.26	0.31	0.82	0.22	0.52
NLP	137.25	9.76	**0.81**	4.46	1.61	7.12	2.16	3.1	1.19	3.72

2.3 Discussion

As expected, HuberLA performs demonstrates to be more robust than HuberLA since the asymmetrical and focused outliers cases broadly fall under the category of near outliers in our study. On the Energy dataset, HuberLA outperformed both tLA and RCGP, achieving performance comparable to LaplaceEP in the presence of asymmetric outliers (Table 5). While HuberMCMC and HuberLA demonstrated comparable efficacy on the Yacht dataset (Table 6), HuberLA and tLA yielded similar results on the Energy dataset under focused outlier conditions.