Supplementary Material for Semi-Supervised Conditional Density Estimation with Wasserstein Laplacian Regularisation

Experimental Procedure

In this section, we give additional details to the experimental procedures carried out in the main paper.

Selection of Test and Training Sets

Random selection of these sets was achieved using the Python sklearn.utils.shuffle function (Buitinck et al. 2013).

The test set was firstly separated from the training set. We shuffled the dataset, allocated the first t data points to the test set and the remainder to the training set. The same shuffle seed (0) was used across all runs of all experiments to provide the same test set for each run.

The training set was then separated into Labelled, Unlabelled and Validation sets. We shuffled the training set, allocated the first l data points to the labelled set, the following u data points to the unlabelled set, and the following v data points to the validation set. Seeds $0, 1, 2, \ldots, 49$ were used to shuffle the training set, to provide different and independent labelled, unlabelled and validation sets for each run.

Stopping Criteria

During training, MDN, WMDN, UMAL and Mean Teacher calculated validation performance (NLL for MDN, WMDN, UMAL, or RMSE for Mean Teacher) every 10 epochs. If validation performance worsened over 20 epochs, training was stopped.

Parameter Tuning

Some algorithm parameters were tuned for each dataset and number of labelled data points. For example, experiments on Electric with 100 labelled data points, Electric with 300 labelled data points, and Protein with 100 labelled data points, had independently tuned parameters. Tuning was carried out by running 50 seeds for each parameter combination, and selecting the parameter combination which had the highest average performance on the validation set.

Parameters which were not tuned for each dataset were set to values which performed well across datasets. All parameters used in the algorithms are shown in Table 1.

Deep Learning Architecture

All algorithms with a deep learning architecture used three fully connected layers of 32 neurons with ReLU activation, however, different algorithms have different output layers. As typical with MDNs (Bishop 1994), our MDN and WMDN have three output layers with c neurons each; one with a linear activation function to define μ, one with an exponential activation function to define σ, and one with a softmax activation function to define π. The UMAL model used a modified ELU activation function to define b and no activation function to define μ (Brando et al. 2019). The output layer of Mean Teacher is a single neuron with no activation function.

Algorithms

CDE for Regression

When using CDE models for regression, the distribution mean was used as models' point prediction. Other aggregation methods include using the distribution mode, or the centre of the largest Gaussian component (Pahlevan et al. 2020). We used the distribution mean because from our experiments, this produced the lowest RMSE for all datasets.

Mean Teacher for Regression

In our experiments, we adapted Mean Teacher (Tarvainen and Valpola 2017) for regression. This required some changes to be made to the original code. Firstly, the output to the deep learning architecture was modified to produce a single continuous output. The "classification" loss and consistency loss were changed from cross entropy loss to Mean Squared Error. Finally, the noise added to input data was replaced with Gaussian noise with standard deviation σ_c (Table 1).

Datasets

In this section, we give additional details to the datasets used in the main paper.

Chlorophyll-a Datasets

Both the Real and Synthetic chlorophyll-a datasets had to be processed from hyperspectral data to multispectral data, as unlabelled satellite data would be multispectral. The data

Algorithm	Fixed Parameters	Tuned Parameters
MDN	$s = 10, c = 5$	$\eta \in \{10^n \mid n = -5, -4.5, \ldots, -2\}$
WMDN	$s = 10, c = 5, d = 1$ $q = 2, k = 5$	$\eta = \eta_{best}$, $\gamma_u \in \{10^n \mid n = -2, -1.5, \ldots, 2\}$, $b \in \{10, 20\}$
UMAL	$s = 10, n_\tau = 100$	$\eta \in \{10^n \mid n = -5, -4.5, \ldots, -2\}$
CoREG	$n_i = 100$	$d_1, d_2 \in \{2, 3, 5\}$ $(k_1, k_2) \in \{(5, 5), (10, 10)\}$
Mean Teacher	$s = 10, \alpha = 0.995$ $\gamma = 100, \gamma_{ramp} = 5$	$\eta \in \{10^n \mid n = -4, -3.5, \ldots, -2\}$, $\sigma_c \in \{0.1, 0.2, 0.4\}$

Table 1: List of parameters for each algorithm used in experiments. For all experiments, η represents learning rate, s represents the number of batches per epoch. WMDN: η_{best} is the best η as tuned by MDN, as described in the main paper. CoREG: n_i represents the number of iterations undertaken during each run. Mean Teacher: γ_u represents the consistency cost weight, γ_{ramp} represents the ramp up epochs of the consistency cost weight, σ_c represents the magnitude of consistency noise.

was converted to multispectral data corresponding to the OLCI sensor of Sentinel-3, as this data allows for greatest sensitivity in wavelengths of chlorophyll-a reflection (Syariz et al. 2020). The conversion to multispectral was done using publicly available OLCI Spectral Response Function Data. We note that the Real dataset contained spectral data from 400nm to 800nm, whereas the Synthetic dataset containted spectral data from 400nm to 750nm. Hence, the Synthetic dataset had fewer dimensions ($d = 12$ compared to $d = 16$). Normally distributed noise was added to features and labels of the Synthetic data, with a standard deviation $\sigma_{noise} = \frac{\sigma}{20}$, where σ is the standard deviation of that feature.

Supplementary Results

Experimental results for Interval Cover, Interval Width, over all datasets and CDE algorithms, are shown in Tables 2 and 3.

References

Bishop, C. M. 1994. Mixture Density Networks. Technical report, Aston University.

Brando, A.; Rodríguez-Serrano, J. A.; Vitrià, J.; and Rubio, A. 2019. Modelling heterogeneous distributions with an Uncountable Mixture of Asymmetric Laplacians. In Wallach, H. M.; Larochelle, H.; Beygelzimer, A.; d'Alché-Buc, F.; Fox, E. B.; and Garnett, R., eds., *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, 8836–8846.

Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; Layton, R.; VanderPlas, J.; Joly, A.; Holt, B.; and Varoquaux, G. 2013. API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, 108–122.

Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. 2020. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sensing of Environment*, 240: 111604.

Syariz, M. A.; Lin, C.-H.; Nguyen, M. V.; Jaelani, L. M.; and Blanco, A. C. 2020. WaterNet: A convolutional neural network for chlorophyll-a concentration retrieval. *Remote Sensing*, 12(12): 1966.

Tarvainen, A.; and Valpola, H. 2017. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 1195–1204.

70% Confidence Interval							
		Interval Coverage (%)			Interval Width		
Data	Labels	MDN	UMAL	WMDN	MDN	UMAL	WMDN
Electric	100	55.8 ± 1.3	56.6 ± 1.4	**68.3 ± 1.1**	**0.22 ± 0.01**	**0.21 ± 0.01**	**0.23 ± 0.01**
($d = 6$)	300	61.5 ± 0.9	52.1 ± 1.4	**75.5 ± 0.8**	0.15 ± 0.00	**0.14 ± 0.00**	0.20 ± 0.00
	1000	61.3 ± 0.7	60.5 ± 1.4	**74.7 ± 0.9**	**0.12 ± 0.00**	0.15 ± 0.01	0.16 ± 0.00
Protein	100	53.2 ± 0.5	**67.4 ± 0.9**	53.7 ± 0.7	**1.23 ± 0.02**	2.15 ± 0.04	1.35 ± 0.02
($d = 9$)	300	61.1 ± 0.5	61.9 ± 1.0	**63.7 ± 0.5**	**1.28 ± 0.01**	1.32 ± 0.03	1.43 ± 0.01
	1000	66.4 ± 0.2	62.9 ± 0.7	**66.8 ± 0.2**	**1.33 ± 0.01**	1.52 ± 0.05	1.36 ± 0.01
Air Quality	100	65.9 ± 0.8	55.1 ± 1.7	**72.0 ± 0.5**	**0.85 ± 0.01**	1.21 ± 0.13	0.89 ± 0.01
($d = 12$)	300	64.4 ± 0.4	54.8 ± 1.0	**72.2 ± 0.4**	0.69 ± 0.01	**0.54 ± 0.01**	0.72 ± 0.01
	1000	65.8 ± 0.2	58.9 ± 0.7	**71.1 ± 0.2**	0.61 ± 0.00	**0.52 ± 0.01**	0.64 ± 0.00
Elevators	100	61.0 ± 0.6	**76.1 ± 1.8**	61.9 ± 0.7	**0.27 ± 0.01**	0.73 ± 0.07	**0.27 ± 0.01**
($d = 18$)	300	63.2 ± 0.4	**74.8 ± 1.9**	63.3 ± 0.4	**0.25 ± 0.00**	0.67 ± 0.07	**0.24 ± 0.00**
	1000	65.0 ± 0.3	**75.5 ± 1.6**	65.0 ± 0.2	**0.20 ± 0.00**	0.64 ± 0.07	**0.19 ± 0.00**
Parkinsons	100	55.0 ± 0.8	**77.8 ± 1.0**	60.8 ± 0.6	**5.75 ± 0.17**	31.04 ± 1.18	6.39 ± 0.15
($d = 20$)	300	60.3 ± 0.5	61.3 ± 1.2	**66.0 ± 0.4**	**5.08 ± 0.08**	7.00 ± 1.21	5.52 ± 0.10
	1000	64.6 ± 0.4	57.6 ± 1.1	**70.2 ± 0.3**	4.59 ± 0.06	**3.55 ± 0.15**	4.30 ± 0.06
Appliances	100	60.3 ± 0.7	53.0 ± 1.8	**71.5 ± 0.6**	**65.53 ± 1.98**	73.27 ± 5.43	79.44 ± 3.09
($d = 27$)	300	64.1 ± 0.4	57.4 ± 1.3	**71.1 ± 0.4**	**65.00 ± 1.23**	83.70 ± 8.29	71.72 ± 1.66
	1000	66.6 ± 0.3	54.8 ± 0.6	**72.3 ± 0.3**	**63.75 ± 0.62**	**62.09 ± 1.60**	71.09 ± 1.15
Song Year	100	62.1 ± 0.6	**79.6 ± 0.7**	61.1 ± 0.4	**15.04 ± 0.23**	40.37 ± 1.46	16.47 ± 0.20
($d = 90$)	300	60.2 ± 0.5	**78.7 ± 0.7**	63.0 ± 0.4	**14.17 ± 0.23**	36.98 ± 1.44	17.03 ± 0.19
	1000	61.5 ± 0.3	**77.7 ± 0.8**	66.0 ± 0.3	**14.76 ± 0.16**	33.77 ± 1.43	17.65 ± 0.12
Synthetic	100	59.2 ± 0.5	49.9 ± 0.7	**62.4 ± 0.5**	1.78 ± 0.03	**1.57 ± 0.03**	1.93 ± 0.03
($d = 12$)	300	**64.9 ± 0.4**	61.4 ± 1.0	**65.9 ± 0.4**	**1.93 ± 0.02**	3.88 ± 1.23	**1.97 ± 0.02**
	1000	**66.6 ± 0.3**	64.9 ± 0.5	**67.0 ± 0.2**	1.91 ± 0.01	**1.80 ± 0.02**	1.93 ± 0.01
Real	100	60.5 ± 0.7	**78.9 ± 1.3**	68.2 ± 0.6	**1.08 ± 0.02**	3.73 ± 0.22	1.17 ± 0.02
($d = 16$)	300	62.8 ± 0.6	55.1 ± 1.4	**70.1 ± 0.6**	**0.94 ± 0.02**	**0.91 ± 0.06**	0.98 ± 0.01
	1000	64.6 ± 0.5	58.1 ± 0.7	**68.1 ± 0.3**	0.77 ± 0.01	**0.64 ± 0.01**	0.78 ± 0.01
90% Confidence Interval							
Electric	100	77.7 ± 1.0	79.3 ± 1.1	**85.6 ± 0.6**	**0.37 ± 0.02**	**0.40 ± 0.03**	**0.38 ± 0.01**
($d = 6$)	300	82.8 ± 0.7	76.0 ± 1.3	**90.5 ± 0.4**	**0.26 ± 0.01**	**0.25 ± 0.01**	0.32 ± 0.01
	1000	83.9 ± 0.5	83.8 ± 1.0	**91.7 ± 0.5**	**0.20 ± 0.00**	0.26 ± 0.01	0.26 ± 0.00
Protein	100	84.7 ± 0.5	**93.4 ± 0.3**	85.7 ± 0.5	**1.93 ± 0.02**	3.29 ± 0.05	2.01 ± 0.02
($d = 9$)	300	84.9 ± 0.3	84.6 ± 1.0	**86.2 ± 0.2**	**1.85 ± 0.01**	2.04 ± 0.07	1.95 ± 0.01
	1000	**87.4 ± 0.2**	**88.0 ± 0.6**	**87.6 ± 0.2**	**1.88 ± 0.01**	2.59 ± 0.07	**1.90 ± 0.01**
Air Quality	100	85.0 ± 0.4	79.2 ± 1.5	**87.9 ± 0.3**	**1.51 ± 0.03**	2.75 ± 0.29	**1.52 ± 0.02**
($d = 12$)	300	85.6 ± 0.3	79.2 ± 0.7	**88.3 ± 0.2**	1.24 ± 0.02	**1.11 ± 0.04**	1.21 ± 0.01
	1000	87.3 ± 0.1	81.7 ± 0.6	**89.1 ± 0.1**	1.10 ± 0.01	**1.02 ± 0.02**	1.11 ± 0.01
Elevators	100	83.1 ± 0.5	**92.6 ± 0.7**	82.9 ± 0.5	**0.49 ± 0.01**	1.15 ± 0.07	**0.47 ± 0.01**
($d = 18$)	300	85.0 ± 0.3	**92.9 ± 0.6**	85.0 ± 0.3	**0.45 ± 0.01**	1.11 ± 0.07	**0.44 ± 0.01**
	1000	86.0 ± 0.2	**93.9 ± 0.4**	86.1 ± 0.1	0.35 ± 0.00	1.10 ± 0.07	**0.33 ± 0.00**
Parkinsons	100	76.6 ± 0.7	**94.1 ± 0.5**	80.7 ± 0.5	**9.70 ± 0.24**	48.53 ± 1.36	10.75 ± 0.21
($d = 20$)	300	82.1 ± 0.4	83.5 ± 1.0	**86.3 ± 0.3**	**8.65 ± 0.14**	15.48 ± 3.43	9.97 ± 0.20
	1000	85.9 ± 0.3	80.9 ± 1.0	**89.4 ± 0.2**	7.85 ± 0.09	6.61 ± 0.36	**7.36 ± 0.09**
Appliances	100	83.6 ± 0.5	77.0 ± 1.2	**90.6 ± 0.4**	**183.12 ± 7.38**	**169.65 ± 12.92**	246.82 ± 7.92
($d = 27$)	300	85.8 ± 0.3	79.9 ± 1.1	**91.6 ± 0.2**	**206.93 ± 5.59**	**195.15 ± 18.32**	272.54 ± 5.25
	1000	88.1 ± 0.2	78.4 ± 0.6	**92.0 ± 0.2**	258.70 ± 3.65	**143.72 ± 3.94**	272.36 ± 4.30
Song Year	100	87.4 ± 0.3	**92.3 ± 0.5**	87.1 ± 0.3	**26.18 ± 0.40**	69.43 ± 3.32	27.74 ± 0.27
($d = 90$)	300	87.6 ± 0.3	**91.6 ± 0.4**	85.9 ± 0.3	**27.50 ± 0.42**	66.14 ± 3.38	28.80 ± 0.26
	1000	87.6 ± 0.2	**90.8 ± 0.5**	88.2 ± 0.2	**29.26 ± 0.31**	61.79 ± 3.26	31.14 ± 0.20
Synthetic	100	80.6 ± 0.4	79.4 ± 0.7	**83.6 ± 0.4**	**2.83 ± 0.03**	4.69 ± 0.15	3.04 ± 0.03
($d = 12$)	300	85.6 ± 0.3	83.1 ± 0.9	**86.7 ± 0.3**	**3.06 ± 0.03**	7.60 ± 2.57	3.14 ± 0.02
	1000	87.2 ± 0.2	83.5 ± 0.4	**87.8 ± 0.2**	3.04 ± 0.02	**2.80 ± 0.04**	3.09 ± 0.01
Real	100	81.9 ± 0.5	**96.6 ± 0.5**	86.4 ± 0.5	**1.84 ± 0.03**	5.85 ± 0.21	1.94 ± 0.03
($d = 16$)	300	85.3 ± 0.4	79.9 ± 1.4	**88.3 ± 0.4**	**1.70 ± 0.03**	**1.69 ± 0.12**	**1.68 ± 0.03**
	1000	86.1 ± 0.3	81.3 ± 0.5	**87.7 ± 0.2**	1.38 ± 0.02	**1.16 ± 0.02**	1.38 ± 0.02

Table 2: Interval Cover and Interval Width for 70%, 90% confidence intervals.

95% Confidence Interval							
		Interval Coverage (%)			Interval Width		
Data	Labels	MDN	UMAL	WMDN	MDN	UMAL	WMDN
Electric	100	85.2 ± 0.8	86.3 ± 0.8	**90.5 ± 0.5**	**0.47 ± 0.03**	0.58 ± 0.04	**0.47 ± 0.01**
($d = 6$)	300	89.2 ± 0.5	85.5 ± 1.1	**94.0 ± 0.3**	**0.32 ± 0.01**	**0.34 ± 0.01**	0.38 ± 0.01
	1000	90.9 ± 0.4	91.2 ± 0.8	**95.3 ± 0.3**	**0.26 ± 0.01**	0.34 ± 0.02	0.33 ± 0.01
Protein	100	93.0 ± 0.4	**97.5 ± 0.1**	93.8 ± 0.3	**2.26 ± 0.03**	3.70 ± 0.05	2.32 ± 0.03
($d = 9$)	300	92.3 ± 0.2	91.5 ± 0.7	**93.0 ± 0.2**	**2.12 ± 0.01**	2.40 ± 0.07	2.18 ± 0.01
	1000	93.5 ± 0.1	**94.7 ± 0.4**	93.5 ± 0.1	**2.11 ± 0.01**	2.96 ± 0.07	**2.13 ± 0.01**
Air Quality	100	90.1 ± 0.3	86.6 ± 1.2	**91.6 ± 0.2**	**1.95 ± 0.03**	4.32 ± 0.55	**1.92 ± 0.03**
($d = 12$)	300	91.4 ± 0.2	87.1 ± 0.4	**92.4 ± 0.2**	1.66 ± 0.03	**1.60 ± 0.07**	**1.58 ± 0.02**
	1000	93.2 ± 0.1	88.8 ± 0.4	**93.9 ± 0.1**	**1.50 ± 0.01**	**1.48 ± 0.03**	**1.50 ± 0.01**
Elevators	100	89.5 ± 0.4	**95.4 ± 0.4**	89.1 ± 0.4	**0.61 ± 0.01**	1.32 ± 0.07	**0.59 ± 0.01**
($d = 18$)	300	91.3 ± 0.2	**95.7 ± 0.4**	91.2 ± 0.2	**0.58 ± 0.01**	1.30 ± 0.07	**0.56 ± 0.01**
	1000	91.9 ± 0.2	**96.6 ± 0.3**	92.0 ± 0.1	0.45 ± 0.01	1.30 ± 0.07	**0.41 ± 0.01**
Parkinsons	100	83.9 ± 0.6	**97.0 ± 0.3**	86.9 ± 0.4	**12.09 ± 0.28**	54.76 ± 1.41	13.33 ± 0.27
($d = 20$)	300	88.6 ± 0.4	**90.9 ± 0.7**	**91.8 ± 0.2**	**10.70 ± 0.19**	22.93 ± 5.43	12.46 ± 0.32
	1000	91.7 ± 0.2	89.1 ± 0.8	**94.3 ± 0.1**	9.63 ± 0.12	9.08 ± 0.54	**8.92 ± 0.12**
Appliances	100	90.5 ± 0.5	84.2 ± 0.9	**94.9 ± 0.3**	274.39 ± 9.41	**239.81 ± 18.74**	**335.68 ± 8.07**
($d = 27$)	300	92.3 ± 0.2	86.8 ± 0.8	**95.9 ± 0.2**	320.10 ± 5.76	**276.43 ± 25.89**	380.24 ± 5.29
	1000	93.5 ± 0.1	85.8 ± 0.4	**96.1 ± 0.1**	378.58 ± 3.89	**206.62 ± 5.86**	388.39 ± 4.17
Song Year	100	91.5 ± 0.3	**95.4 ± 0.3**	92.8 ± 0.2	**32.27 ± 0.52**	79.35 ± 3.91	33.66 ± 0.35
($d = 90$)	300	93.1 ± 0.3	**95.0 ± 0.3**	93.1 ± 0.2	**35.09 ± 0.53**	76.10 ± 3.97	**35.04 ± 0.29**
	1000	**94.2 ± 0.2**	**94.5 ± 0.4**	**94.5 ± 0.1**	**37.09 ± 0.35**	71.74 ± 3.87	37.81 ± 0.20
Synthetic	100	87.5 ± 0.4	88.2 ± 0.5	**90.1 ± 0.3**	**3.37 ± 0.04**	6.11 ± 0.15	3.61 ± 0.04
($d = 12$)	300	91.8 ± 0.2	90.0 ± 0.7	**92.6 ± 0.2**	**3.65 ± 0.03**	10.23 ± 3.42	3.73 ± 0.03
	1000	93.0 ± 0.1	90.3 ± 0.4	**93.5 ± 0.1**	**3.64 ± 0.02**	3.90 ± 0.12	3.69 ± 0.02
Real	100	88.3 ± 0.5	**98.3 ± 0.3**	91.6 ± 0.4	**2.28 ± 0.04**	6.68 ± 0.20	2.39 ± 0.04
($d = 16$)	300	91.3 ± 0.3	89.0 ± 1.0	**92.9 ± 0.3**	**2.20 ± 0.04**	2.31 ± 0.16	**2.13 ± 0.03**
	1000	92.2 ± 0.3	90.1 ± 0.5	**92.9 ± 0.2**	1.81 ± 0.03	**1.60 ± 0.03**	1.78 ± 0.03
99% Confidence Interval							
Electric	100	92.9 ± 0.5	93.5 ± 0.4	**95.4 ± 0.3**	**0.68 ± 0.04**	1.16 ± 0.09	0.66 ± 0.02
($d = 6$)	300	95.8 ± 0.3	95.7 ± 0.4	**97.4 ± 0.2**	**0.50 ± 0.01**	0.71 ± 0.03	0.55 ± 0.01
	1000	97.5 ± 0.1	97.9 ± 0.3	**98.5 ± 0.1**	**0.40 ± 0.01**	0.62 ± 0.05	0.48 ± 0.01
Protein	100	98.5 ± 0.1	**99.5 ± 0.1**	98.6 ± 0.1	**2.81 ± 0.03**	4.20 ± 0.06	**2.83 ± 0.03**
($d = 9$)	300	**98.2 ± 0.1**	97.2 ± 0.2	**98.2 ± 0.1**	**2.61 ± 0.02**	2.93 ± 0.08	**2.62 ± 0.02**
	1000	98.5 ± 0.0	**98.8 ± 0.1**	98.5 ± 0.0	**2.55 ± 0.01**	3.42 ± 0.07	**2.56 ± 0.01**
Air Quality	100	**95.5 ± 0.2**	94.3 ± 0.7	**95.6 ± 0.2**	2.99 ± 0.05	8.19 ± 1.19	**2.79 ± 0.05**
($d = 12$)	300	**96.9 ± 0.1**	94.9 ± 0.2	**96.8 ± 0.1**	2.79 ± 0.05	3.11 ± 0.13	**2.57 ± 0.04**
	1000	**98.3 ± 0.1**	96.2 ± 0.3	**98.4 ± 0.1**	**2.76 ± 0.02**	3.53 ± 0.11	**2.79 ± 0.02**
Elevators	100	96.0 ± 0.3	**97.1 ± 0.3**	95.7 ± 0.2	0.88 ± 0.02	1.53 ± 0.07	**0.83 ± 0.01**
($d = 18$)	300	**97.4 ± 0.1**	**97.6 ± 0.2**	97.1 ± 0.1	0.85 ± 0.01	1.52 ± 0.07	**0.81 ± 0.01**
	1000	97.4 ± 0.1	**98.0 ± 0.2**	97.4 ± 0.1	0.66 ± 0.01	1.53 ± 0.07	**0.60 ± 0.01**
Parkinsons	100	92.2 ± 0.4	**98.7 ± 0.1**	93.6 ± 0.3	**17.30 ± 0.43**	61.95 ± 1.46	18.67 ± 0.41
($d = 20$)	300	95.6 ± 0.2	**97.2 ± 0.3**	**97.0 ± 0.1**	**14.81 ± 0.28**	41.44 ± 10.43	16.61 ± 0.43
	1000	97.2 ± 0.1	97.2 ± 0.2	**98.3 ± 0.1**	12.61 ± 0.18	15.93 ± 0.98	**11.72 ± 0.21**
Appliances	100	97.1 ± 0.3	91.3 ± 0.6	**98.0 ± 0.1**	461.83 ± 13.37	**404.71 ± 33.63**	476.92 ± 10.33
($d = 27$)	300	98.1 ± 0.1	93.9 ± 0.5	**98.6 ± 0.1**	542.52 ± 9.49	**477.98 ± 48.34**	534.86 ± 9.15
	1000	98.0 ± 0.1	93.5 ± 0.3	**98.8 ± 0.0**	571.53 ± 5.45	**355.52 ± 12.30**	559.17 ± 5.28
Song Year	100	96.3 ± 0.3	**98.0 ± 0.2**	96.9 ± 0.2	**43.17 ± 0.80**	91.00 ± 4.56	**44.40 ± 0.53**
($d = 90$)	300	**98.1 ± 0.2**	**97.8 ± 0.2**	**98.1 ± 0.1**	48.89 ± 0.79	87.73 ± 4.57	**47.03 ± 0.37**
	1000	**99.0 ± 0.1**	97.6 ± 0.2	**99.0 ± 0.0**	51.54 ± 0.51	83.27 ± 4.50	**49.82 ± 0.24**
Synthetic	100	95.1 ± 0.3	95.3 ± 0.3	**96.6 ± 0.2**	**4.43 ± 0.06**	7.73 ± 0.10	4.71 ± 0.05
($d = 12$)	300	97.6 ± 0.1	96.1 ± 0.5	**98.0 ± 0.1**	**4.80 ± 0.05**	15.00 ± 4.44	**4.87 ± 0.04**
	1000	98.2 ± 0.1	96.3 ± 0.2	**98.5 ± 0.1**	**4.76 ± 0.03**	6.89 ± 0.15	4.82 ± 0.02
Real	100	94.7 ± 0.3	**99.5 ± 0.1**	96.3 ± 0.3	**3.15 ± 0.06**	7.62 ± 0.18	**3.25 ± 0.07**
($d = 16$)	300	**97.2 ± 0.2**	**97.0 ± 0.6**	**97.5 ± 0.2**	3.25 ± 0.05	4.21 ± 0.23	**3.12 ± 0.05**
	1000	**98.4 ± 0.1**	97.9 ± 0.2	**98.4 ± 0.1**	**2.94 ± 0.04**	3.02 ± 0.07	**2.87 ± 0.04**

Table 3: Interval Cover and Interval Width for 95%, 99% confidence intervals.