In the paper "Nonsmooth optimization based hyperparameter-free neural networks for large-scale regression", a new nonsmooth optimization-based algorithm for solving large-scale regression problems is introduced. The regression problem is modelled as fully-connected feedforward neural networks with one hidden layer, the piecewise linear activation, and the L_1-loss functions. A modified version of the limited memory bundle method is applied to minimize this nonsmooth objective. In addition, a novel constructive approach for automated determination of the proper number of hidden nodes is developed. Finally, large real-world data sets are used to evaluate the proposed algorithm and to compare it with some state-of-the-art neural network algorithms for regression. The results demonstrate the superiority of the proposed algorithm as a predictive tool in most data sets used in numerical experiments. Here, we recall the used performance measures and give results with respect to mean absolute error, coefficient of determination, and Pearson's correlation coefficient.

1. Performance measures

We now describe the performance measures used in our evaluations and comparisons in the paper "Nonsmooth optimization based hyperparameter-free neural networks for large-scale regression". Let $e_1, \ldots, e_n$ for $n \geq 1$ be actual observed values for some parameter e and $\bar{e}_1, \ldots, \bar{e}_n$ be their predicted values. We use the following performance measures:

- *root mean square error*:

$$\text{RMSE} = \Big(\frac{1}{n}\sum_{i=1}^{n}(\bar{e}_i - e_i)^2\Big)^{1/2};$$

- *mean absolute error*:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\bar{e}_i - e_i|;$$

- *coefficient of determination*:

$$R^2 = 1 - \Big(\frac{\sum_{i=1}^{n}(e_i - \bar{e}_i)^2}{\sum_{i=1}^{n}(e_i - e_0)^2}\Big);$$

- *Pearson's correlation coefficient*:

$$r = \frac{\sum_{i=1}^{n}(e_i - e_0)(\bar{e}_i - \bar{e_0})}{\Big(\sum_{i=1}^{n}(e_i - e_0)^2 \sum_{i=1}^{n}(\bar{e}_i - \bar{e_0})^2\Big)^{1/2}}.$$

In the R^2 and r measures, e_0 is the mean of observed values while $\bar{e}_0$ is the mean of predicted values. The small values of RMSE and MAE measures indicate small deviations of the predictions from actual observations. The R^2 measure ranges from $-\infty$ to 1: $R^2 = 1$ means a perfect prediction, $R^2 = 0$ indicates that the model predictions are as accurate as the mean of the observed data and an efficiency $-\infty < R^2 < 0$ occurs when the observed mean is a better predictor than the model. The range of r is from -1 to 1: $r = 1$ implies that a

linear equation describes the relationship between observed and predicted values perfectly, $r = -1$ means that all the samples lie on a line for which a predicted value decreases as an observed value increases and $r = 0$ happens when there is no linear correlation between these values.

2. Results

The results with respect to the RMSE values of the test set and the elapsed CPU times are given in the paper "Nonsmooth optimization based hyperparameter-free neural networks for large-scale regression". Here we give results with respect to MAE, R^2, and r. Note that in the following tables "Best" corresponds to the best solution obtained with LMBNNR with respect to RMSE.

Table S1. MAE, R^2 and r for test set in Combined cycle power plant data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	3.487	6.703	15.910	3.634	16.660	5.276
5	3.342	4.807	16.650	3.470	4.125	5.172
10	3.176	4.563	14.706	3.332	4.125	5.140
50	3.079	3.727	16.093	3.226	4.125	7.614
100	3.079	3.638	14.870	3.182	4.125	14.777
200	3.079	3.605	15.118	3.221	4.125	t-limit
500		3.582	14.315	3.198	4.125	t-limit
5000		3.571	14.529	3.249	4.125	t-limit
Best:	3.087	H=29				
ASP:	3.079	H=61				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.934	0.750	–0.256	0.930	–0.246	0.840
5	0.939	0.874	–0.618	0.935	0.907	0.845
10	0.941	0.887	–0.094	0.939	0.907	0.846
50	0.943	0.926	–0.213	0.941	0.907	0.693
100	0.943	0.929	0.004	0.942	0.907	0.000
200	0.943	0.930	–0.046	0.941	0.907	t-limit
500		0.931	0.085	0.942	0.907	t-limit
5000		0.931	0.066	0.941	0.907	t-limit
Best:	0.944	H=29				
ASP:	0.943	H=61				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.967	0.867	0.108	0.967	–0.288	0.916
5	0.969	0.935	0.169	0.969	0.953	0.919
10	0.970	0.942	0.087	0.970	0.953	0.920
50	0.971	0.962	–0.204	0.972	0.953	0.910
100	0.971	0.964	0.110	0.972	0.953	0.787
200	0.971	0.965	–0.042	0.972	0.953	t-limit
500		0.965	0.491	0.972	0.953	t-limit
5000		0.965	0.738	0.972	0.953	t-limit
Best:	0.971	H=29				
ASP:	0.971	H=61				

Table S2. MAE, R^2 and r for test set in Airfoil self-noise data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	3.519	5.469	7.868	3.308	77.016	4.535
5	3.010	5.317	7.552	2.819	25.095	4.503
10	3.011	4.755	7.082	2.187	25.095	4.499
50	2.921	4.490	5.712	1.833	25.095	5.179
100	1.364	4.440	5.567	1.549	25.095	5.522
200	1.364	4.318	5.630	1.580	25.095	5.513
500		4.297	5.454	1.683	25.095	t-limit
5000		4.294	5.417	1.532	25.095	t-limit
Best:	1.364	H=83				
ASP:	1.364	H=105				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.467	–0.013	–1.414	0.570	–169.584	0.276
5	0.570	0.034	–1.026	0.696	–19.363	0.295
10	0.569	0.228	–0.851	0.822	–19.363	0.291
50	0.582	0.303	–0.125	0.871	–19.363	0.116
100	0.920	0.325	–0.040	0.909	–19.363	–0.005
200	0.920	0.366	–0.080	0.907	–19.363	–0.004
500		0.373	–0.008	0.890	–19.363	t-limit
5000		0.378	0.009	0.908	–19.363	t-limit
Best:	0.920	H=83				
ASP:	0.920	H=105				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.699	0.033	0.063	0.769	–0.221	0.528
5	0.768	0.245	0.041	0.849	–0.155	0.544
10	0.768	0.478	–0.035	0.911	–0.155	0.541
50	0.784	0.593	0.096	0.940	–0.155	0.537
100	0.960	0.663	0.108	0.956	–0.155	–0.080
200	0.960	0.711	–0.012	0.957	–0.155	–0.042
500		0.723	0.054	0.949	–0.155	t-limit
5000		0.726	0.164	0.956	–0.155	t-limit
Best:	0.960	H=83				
ASP:	0.960	H=105				

Table S3. MAE, R^2 and r for test set in Concrete compressive strength data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	8.761	14.271	22.520	6.099	12.873	4.619
5	4.951	12.909	18.357	4.991	11.255	3.713
10	4.482	12.493	18.999	4.286	7.819	3.238
50	3.795	11.401	14.427	3.811	7.819	11.810
100	3.780	11.090	14.663	3.866	7.819	12.464
200	3.780	10.364	14.168	3.680	7.819	12.465
500		10.315	13.593	3.622	7.819	t-limit
5000		10.332	13.536	3.425	7.819	t-limit
Best:	3.780	H=96				
ASP:	3.781	H=74				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.521	–0.118	–2.544	0.787	–0.079	0.841
5	0.836	0.102	–1.011	0.854	0.205	0.900
10	0.869	0.170	–1.101	0.888	0.615	0.926
50	0.903	0.323	–0.153	0.912	0.615	0.104
100	0.904	0.367	–0.181	0.908	0.615	–0.001
200	0.904	0.442	–0.089	0.915	0.615	0.000
500		0.451	0.021	0.918	0.615	t-limit
5000		0.452	0.025	0.928	0.615	t-limit
Best:	0.904	H=96				
ASP:	0.904	H=74				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.762	0.116	–0.017	0.889	0.082	0.917
5	0.917	0.351	–0.048	0.930	0.465	0.949
10	0.933	0.441	0.021	0.946	0.785	0.962
50	0.952	0.594	0.141	0.959	0.785	0.886
100	0.952	0.655	–0.008	0.959	0.785	–0.252
200	0.952	0.722	0.012	0.962	0.785	–0.241
500		0.743	0.181	0.963	0.785	t-limit
5000		0.751	0.237	0.967	0.785	t-limit
Best:	0.952	H=96				
ASP:	0.952	H=74				

Table S4. MAE, R^2 and r for test set in Physicochemical properties of protein data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	3.941	4.399	6.651	4.265(1)	5.665	3.998
5	3.756	4.263	6.723	NaN	5.295	3.849
10	3.743	4.233	6.587	3.800(2)	4.347	3.994
50	3.521	4.118	5.732	NaN	4.347	t-limit
100	3.520	4.073	5.710	NaN	4.347	t-limit
200	3.520	4.084	5.424	NaN	4.347	t-limit
500		4.122	5.471	NaN	4.347	t-limit
5000		4.092	5.423	NaN	4.347	t-limit
Best:	3.520	H=56				
ASP:	3.520	H=74				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.238	0.260	–0.777	0.287(1)	–0.192	0.347
5	0.297	0.289	–0.979	NaN	–0.106	0.389
10	0.305	0.300	–0.844	0.361(2)	0.278	0.351
50	0.377	0.329	–0.263	NaN	0.278	t-limit
100	0.377	0.335	–0.137	NaN	0.278	t-limit
200	0.377	0.338	–0.027	NaN	0.278	t-limit
500		0.330	–0.019	NaN	0.278	t-limit
5000		0.338	0.010	NaN	0.278	t-limit
Best:	0.377	H=56				
ASP:	0.377	H=74				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.539	0.511	–0.045	0.551(1)	0.068	0.589
5	0.583	0.539	–0.008	NaN	0.245	0.625
10	0.587	0.550	–0.001	0.611(2)	0.528	0.593
50	0.632	0.577	–0.027	NaN	0.528	t-limit
100	0.632	0.581	–0.010	NaN	0.528	t-limit
200	0.632	0.583	0.055	NaN	0.528	t-limit
500		0.579	0.031	NaN	0.528	t-limit
5000		0.585	0.168	NaN	0.528	t-limit
Best:	0.632	H=56				
ASP:	0.632	H=74				

(1) 2/10 runs led to NaN loss function value.
(2) 8/10 runs led to NaN loss function value.

Table S5. MAE, R^2 and r for test set in Boston housing data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	3.298	7.388	9.626	3.265	7.480	2.099
5	3.296	6.662	9.265	2.750	6.255	1.731
10	2.650	6.229	9.371	2.693	4.560	1.821
50	2.594	5.519	7.685	2.622	3.332	2.880
100	2.594	5.163	7.412	2.619	3.332	4.183
200	2.594	4.904	6.955	2.528	3.332	5.423
500		4.966	7.086	2.532	3.332	t-limit
5000		4.917	6.681	2.494	3.332	t-limit
Best:	2.594	H=30				
ASP:	2.594	H=39				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.703	–0.195	–0.783	0.785	–0.113	0.912
5	0.704	0.141	–0.513	0.841	0.093	0.933
10	0.800	0.239	–0.608	0.849	0.529	0.924
50	0.827	0.366	–0.145	0.861	0.738	0.821
100	0.827	0.425	–0.120	0.861	0.738	0.611
200	0.827	0.455	–0.085	0.874	0.738	0.344
500		0.425	–0.051	0.870	0.738	t-limit
5000		0.437	0.040	0.869	0.738	t-limit
Best:	0.827	H=30				
ASP:	0.827	H=39				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.853	0.127	0.067	0.901	0.533	0.955
5	0.853	0.470	0.298	0.922	0.414	0.966
10	0.895	0.477	0.042	0.923	0.733	0.961
50	0.910	0.640	0.099	0.931	0.862	0.909
100	0.910	0.685	0.169	0.931	0.862	0.815
200	0.910	0.720	0.138	0.937	0.862	0.687
500		0.698	0.132	0.936	0.862	t-limit
5000		0.712	0.542	0.934	0.862	t-limit
Best:	0.910	H=30				
ASP:	0.910	H=39				

Table S6. MAE, R^2 and r for test set in SGEMM GPU kernel performance data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	84.187	106.274	332.182	97.193	196.175	57.215
5	56.191	75.854	403.268	63.499	181.867	41.097
10	47.621	63.318	360.189	45.560	176.276	48.089
50	47.010	61.959	286.637	28.370	176.153	t-limit
100	47.010	61.185	267.953	28.346	176.153	t-limit
200	47.010	59.193	246.155	24.309	176.153	t-limit
500		57.267	221.834	23.878	176.153	t-limit
5000		56.482	211.507	19.731	t-limit	t-limit
Best:	47.010	H=19				
ASP:	47.410	H=18				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.634	0.768	–1.038	0.838	0.223	0.940
5	0.897	0.878	–1.273	0.916	0.303	0.966
10	0.934	0.900	–0.892	0.954	0.395	0.952
50	0.936	0.913	–0.366	0.986	0.407	t-limit
100	0.936	0.917	–0.195	0.987	0.407	t-limit
200	0.936	0.920	–0.092	0.990	0.407	t-limit
500		0.924	–0.079	0.991	0.407	t-limit
5000		0.924	0.011	0.994	t-limit	t-limit
Best:	0.936	H=19				
ASP:	0.935	H=18				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.825	0.870	0.004	0.924	0.541	0.969
5	0.948	0.938	–0.016	0.960	0.551	0.983
10	0.967	0.949	0.005	0.981	0.628	0.976
50	0.968	0.956	0.007	0.993	0.638	t-limit
100	0.968	0.958	0.007	0.994	0.638	t-limit
200	0.968	0.960	0.093	0.996	0.638	t-limit
500		0.962	–0.091	0.996	0.638	t-limit
5000		0.962	0.125	0.997	t-limit	t-limit
Best:	0.968	H=19				
ASP:	0.967	H=18				

Table S7. MAE, R^2 and r for test set in Online news popularity data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	2460.334	3333.375	7630.585	3800.756(1)	3225.340	3061.343
5	2450.368	3249.126	10810.528	NaN	3205.684	3023.191
10	2449.567	3464.855	10630.898	NaN	3096.292	3000.213
50	2449.517	3642.130	9047.129	NaN	3089.455	t-limit
100	2449.517	3632.464	7357.324	NaN	3087.765	t-limit
200	2449.517	4408.260	6285.794	NaN	3087.765	t-limit
500		3609.625	4917.437	NaN	3087.765	t-limit
5000		3484.741	3420.317	NaN	t-limit	t-limit
Best:	2449.517	H=19				
ASP:	2449.517	H=19				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	–0.016	–0.002	–0.690	–0.031(1)	–0.009	0.097
5	–0.014	–0.002	–1.331	NaN	–0.002	0.278
10	–0.012	–0.010	–1.209	NaN	0.014	0.271
50	–0.012	–0.065	–0.806	NaN	0.024	t-limit
100	–0.012	–0.073	–0.440	NaN	0.024	t-limit
200	–0.012	–0.112	–0.298	NaN	0.024	t-limit
500		–0.038	–0.134	NaN	0.024	t-limit
5000		–0.005	–0.012	NaN	t-limit	t-limit
Best:	–0.012	H=19				
ASP:	–0.012	H=19				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.110	0.042	0.003	0.092(1)	0.005	0.344
5	0.119	0.066	0.004	NaN	0.055	0.643
10	0.127	0.074	–0.014	NaN	0.123	0.644
50	0.127	0.068	0.005	NaN	0.154	t-limit
100	0.127	0.066	–0.001	NaN	0.154	t-limit
200	0.127	0.063	–0.004	NaN	0.154	t-limit
500		0.081	–0.006	NaN	0.154	t-limit
5000		0.114	0.011	NaN	t-limit	t-limit
Best:	0.127	H=19				
ASP:	0.127	H=19				

(1) 7/10 runs led to NaN loss function value.

Table S8. MAE, R^2 and r for test set in Residential building data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	119.297	787.701	1161.594	149.070(1)	1391.853	41.736
5	120.137	745.838	998.809	126.355(2)	1377.299	308.112
10	90.924	746.882	1179.940	105.905(3)	742.278	910.196
50	91.620	695.736	850.053	94.044(3)	130.341	918.273
100	91.620	656.278	782.288	NaN	81.144	t-limit
200	91.620	627.105	708.014	NaN	81.144	t-limit
500		607.388	655.423	NaN	81.144	t-limit
5000		492.349	568.573	NaN	81.144	t-limit
Best:	90.393	H=11				
ASP:	91.620	H=27				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.967	0.169	–0.714	0.964(1)	–1.348	0.997
5	0.967	0.330	–0.107	0.977(2)	–1.332	0.889
10	0.981	0.330	–0.427	0.985(3)	0.337	0.000
50	0.981	0.375	0.112	0.987(3)	0.980	0.000
100	0.981	0.392	0.241	NaN	0.989	t-limit
200	0.981	0.440	0.376	NaN	0.989	t-limit
500		0.368	0.404	NaN	0.989	t-limit
5000		0.636	0.499	NaN	0.989	t-limit
Best:	0.981	H=11				
ASP:	0.981	H=27				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.986	0.456	0.320	0.985(1)	–0.005	0.999
5	0.986	0.593	0.280	0.990(2)	–0.107	1.000
10	0.992	0.597	0.160	0.993(3)	0.583	–0.558
50	0.992	0.615	0.466	0.994(3)	0.990	–0.560
100	0.992	0.690	0.547	NaN	0.994	t-limit
200	0.992	0.704	0.632	NaN	0.994	t-limit
500		0.673	0.641	NaN	0.994	t-limit
5000		0.852	0.708	NaN	0.994	t-limit
Best:	0.992	H=11				
ASP:	0.992	H=27				

(1) 2/10 runs led to NaN loss function value.
(2) 1/10 runs led to NaN loss function value.
(3) 4/10 runs led to NaN loss function value.

Table S9. MAE, R^2 and r for test set in BlogFeedback data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	6.009	NaN	NaN	NaN	10.379	6.454
5	5.770	NaN	NaN	NaN	10.059	t-limit
10	5.574	NaN	NaN	NaN	8.843	t-limit
50	5.574	NaN	NaN	NaN	8.718	t-limit
100	5.574	NaN	NaN	NaN	8.829	t-limit
200	5.574	NaN	NaN	NaN	9.124	t-limit
500		NaN	NaN	NaN	9.123	t-limit
5000		NaN	NaN	NaN	9.123	t-limit
Best:	5.574	H=10				
ASP:	5.574	H=12				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.352	NaN	NaN	NaN	0.126	0.323
5	0.352	NaN	NaN	NaN	0.152	t-limit
10	0.443	NaN	NaN	NaN	0.340	t-limit
50	0.443	NaN	NaN	NaN	0.374	t-limit
100	0.443	NaN	NaN	NaN	0.375	t-limit
200	0.443	NaN	NaN	NaN	0.380	t-limit
500		NaN	NaN	NaN	0.381	t-limit
5000		NaN	NaN	NaN	0.381	t-limit
Best:	0.443	H=10				
ASP:	0.443	H=12				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.629	NaN	NaN	NaN	0.358	0.676
5	0.634	NaN	NaN	NaN	0.394	t-limit
10	0.685	NaN	NaN	NaN	0.585	t-limit
50	0.685	NaN	NaN	NaN	0.613	t-limit
100	0.685	NaN	NaN	NaN	0.614	t-limit
200	0.685	NaN	NaN	NaN	0.618	t-limit
500		NaN	NaN	NaN	0.618	t-limit
5000		NaN	NaN	NaN	0.618	t-limit
Best:	0.685	H=10				
ASP:	0.685	H=12				

Table S10. MAE, R^2 and r for test set in ISOLET data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	3.621	4.270	7.494	NaN	10.648	2.102
5	2.954	3.831	8.759	NaN	8.320	t-limit
10	2.749	3.665	7.982	NaN	7.518	t-limit
50	2.748	3.642	7.888	NaN	5.327	t-limit
100	2.748	3.607	7.815	NaN	4.873	t-limit
200	2.748	3.605	7.301	NaN	3.932	t-limit
500		3.784	7.502	NaN	3.414	t-limit
5000		NaN	6.648	NaN	3.250	t-limit
Best:	2.749	H=10				
ASP:	2.748	H=14				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.563	0.463	–0.682	NaN	–1.972	0.859
5	0.669	0.547	–1.444	NaN	–0.949	t-limit
10	0.717	0.574	–0.850	NaN	–0.448	t-limit
50	0.716	0.587	–0.791	NaN	0.248	t-limit
100	0.716	0.591	–0.754	NaN	0.367	t-limit
200	0.716	0.592	–0.539	NaN	0.569	t-limit
500		0.558	–0.666	NaN	0.669	t-limit
5000		NaN	–0.315	NaN	0.706	t-limit
Best:	0.717	H=10				
ASP:	0.716	H=14				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.761	0.676	0.050	NaN	0.004	0.934
5	0.825	0.743	0.118	NaN	0.024	t-limit
10	0.852	0.763	0.086	NaN	–0.042	t-limit
50	0.852	0.774	0.147	NaN	0.518	t-limit
100	0.852	0.779	0.142	NaN	0.612	t-limit
200	0.852	0.782	0.210	NaN	0.756	t-limit
500		0.800	0.242	NaN	0.818	t-limit
5000		NaN	0.392	NaN	0.840	t-limit
Best:	0.852	H=10				
ASP:	0.852	H=14				

Table S11. MAE, R^2 and r for test set in CIFAR-10 data.

MAE						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	2.302	NaN	NaN	NaN	4.192	t-limit
5	2.256	NaN	NaN	NaN	2.682	t-limit
10	2.244	NaN	NaN	NaN	2.672	t-limit
50	2.244	NaN	NaN	NaN	2.506	t-limit
100	t-limit	NaN	NaN	NaN	2.491	t-limit
200	t-limit	NaN	NaN	NaN	2.464	t-limit
500		NaN	NaN	NaN	2.445	t-limit
5000		NaN	NaN	NaN	t-limit	t-limit
Best:	2.244	H=12				
ASP:	2.244	H=17				

R^2						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.029	NaN	NaN	NaN	–2.095	t-limit
5	0.054	NaN	NaN	NaN	–0.227	t-limit
10	0.067	NaN	NaN	NaN	–0.222	t-limit
50	0.067	NaN	NaN	NaN	–0.073	t-limit
100	t-limit	NaN	NaN	NaN	–0.058	t-limit
200	t-limit	NaN	NaN	NaN	–0.043	t-limit
500		NaN	NaN	NaN	–0.032	t-limit
5000		NaN	NaN	NaN	t-limit	t-limit
Best:	0.067	H=12				
ASP:	0.067	H=17				

r						
H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.278	NaN	NaN	NaN	–0.029	t-limit
5	0.321	NaN	NaN	NaN	0.034	t-limit
10	0.331	NaN	NaN	NaN	0.002	t-limit
50	0.331	NaN	NaN	NaN	0.121	t-limit
100	t-limit	NaN	NaN	NaN	0.142	t-limit
200	t-limit	NaN	NaN	NaN	0.169	t-limit
500		NaN	NaN	NaN	0.192	t-limit
5000		NaN	NaN	NaN	t-limit	t-limit
Best:	0.331	H=12				
ASP:	0.331	H=17				

Table S12. MAE, R^2 and r for test set in Greenhouse gas observing network data.

MAE

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	18.547	32.923(1)	238.605	NaN	54.541	t-limit
5	17.477	NaN	204.267	NaN	32.349	t-limit
10	17.343	NaN	313.008	NaN	32.879	t-limit
50	17.342	NaN	294.423	NaN	20.904	t-limit
100	17.342	NaN	278.436	NaN	18.445	t-limit
200	17.342	NaN	300.219	NaN	16.227	t-limit
500		NaN	325.884	NaN	13.511	t-limit
5000	–	NaN	664.811	NaN	fail	t-limit
Best:	17.343	H=10				
ASP:	17.343	H=12				

R^2

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.942	0.509(1)	–20.987	NaN	0.210	t-limit
5	0.949	NaN	–11.939	NaN	0.617	t-limit
10	0.950	NaN	–27.330	NaN	0.677	t-limit
50	0.950	NaN	–24.579	NaN	0.891	t-limit
100	0.950	NaN	–19.414	NaN	0.920	t-limit
200	0.950	NaN	–25.177	NaN	0.940	t-limit
500		NaN	–33.636	NaN	0.958	t-limit
5000		NaN	–147.049	NaN	fail	t-limit
Best:	0.950	H=10				
ASP:	0.950	H=12				

r

H	LMBNNR	MBGD	BGD	SGD	ELM	MONMLP
2	0.971	0.700(1)	0.557	NaN	0.693	t-limit
5	0.974	NaN	0.667	NaN	0.806	t-limit
10	0.975	NaN	0.740	NaN	0.844	t-limit
50	0.975	NaN	0.832	NaN	0.933	t-limit
100	0.975	NaN	0.832	NaN	0.959	t-limit
200	0.975	NaN	0.753	NaN	0.970	t-limit
500		NaN	0.795	NaN	0.979	t-limit
5000		NaN	0.838	NaN	fail	t-limit
Best:	0.975	H=10				
ASP:	0.975	H=12				

(1) 4/10 runs led to NaN loss function value.