

1. НАУЧНЫЕ СТАТЬИ

1.1. СРАВНЕНИЕ ВОЗМОЖНОСТЕЙ НЕЙРОННЫХ СЕТЕЙ И МЕТОДА ШЕПАРДА ДЛЯ ПРОГНОЗИРОВАНИЯ НАВОДНЕНИЙ НА ОСНОВЕ ОБРАБОТКИ СТАТИСТИЧЕСКИХ ДАННЫХ

Карандеев А.А.¹, Осипов В.П.²

¹РЭУ им. Г.В. Плеханова

²ИПМ им. М.В.Келдыша РАН

Задача прогнозирования того или иного стихийного бедствия и сопутствующих ему проблем является ключевой во многих странах и регионах. Своевременное принятие соответствующих предупредительных мер помогает свести к минимуму повреждения и траты на последующее восстановление инфраструктуры, а также избежать человеческих жертв. Одним из ключевых подходов для решения подобного рода задач является математическое моделирование и анализ временных рядов. В статье рассмотрена задача прогнозирования наводнений в прибрежном городе Туапсе (Краснодарский край, Российской Федерации), при этом за основу для исследований взяты данные, предоставленные МЧС России. Рассмотрены алгоритмы первичной обработки данных для нивелирования помех и их корректировки для последующего анализа и использования. Представлены результаты сравнительного анализа возможностей нейросетевого моделирования, а также метода Шепарда для решения задач прогнозирования наводнений. Рассмотрены преимущества и недостатки каждого из подходов. Приведены результаты численных экспериментов, демонстрирующие особенности применения каждого из подходов, а также сделан ряд выводов на основе полученных данных.

Введение

Нейронные сети нашли свое применение во многих сферах нашей жизни. Однако ряд задач, которые на первый взгляд могут быть решены методами нейросетевого моделирования, по-прежнему остаются нерешенными. К таким задачам можно отнести задачу предсказания развития тех или иных комплексных процессов. Этому способствует целый ряд причин, которые будут разобраны ниже на примере решения задачи предсказания наводнений.

Данная задача является типичным примером задачи прогнозирования на основе обработки временных рядов, что предполагает анализ большого количества данных, формирование модели и ее калибровку [Chollet, 2017].

Могут быть использованы традиционные подходы на основе методов статистики, методы машинного обучения нейронных сетей, а также метод Шепарда (метод обратных взвешенных расстояний IDW Inverse Distance Weighted) [Shepard, 1968], который является методом многомерной интерполяции.

Основной областью применения данного метода Шепарда является обработка картографических, геологических и других экспериментальных данных [Shepard, 1968, Korusov, 2013]. Под интерполяцией понимается метод нахождения неизвестных промежуточных значений некоторой функции по имеющемуся дискретному набору ее известных значений. Типичным примером такой функции является временной ряд, значения которого — результаты наблюдения, зафиксированные через определенный интервал времени.

За исключением некоторых нюансов, метод нейронных сетей и метод Шепарда в данном конкретном случае решают одни и те же задачи аппроксимации и интерполяции. В этой связи практически полезно исследовать особенности методов и осуществить сравнительный анализ результатов их применения для решения одной и той же задачи.

1. Исходные данные задачи

В качестве такой задачи была рассмотрена проблема прогнозирования наводнений в городе Туапсе (Краснодарский край, РФ) на основе метеорологических данных (количество осадков, направление ветра, уровень воды и облачности в регионе), за период с 2014 по 2018 год. Актуальные данные были предоставлены МЧС России в виде Excel-таблицы, изображение которой представлено на рис. 1. Были даны различные показатели и результаты осуществления замеров с частотой в 3 часа в период с ноября 2018 года по сентябрь 2020 года. Так, среди основных параметров, которые использовались для проведения расчетов, можно выделить показатель уровня воды в регионе, атмосферное давление, показатель облачности в регионе, направление ветра и его скорость.

1	Местное время в	T	Po	P	Pa	U	DD	Ff	ff10	ff3	N	WW
2	09.11.2018 18:00	14,3	760,4	766,0	0,1	72	Ветер, дующий с севера	1			Облаков нет.	
3	09.11.2018 15:00	19,8	760,3	765,7	-0,5	53	Ветер, дующий с западо-юго-запада	2			Облаков нет.	
4	09.11.2018 12:00	18,5	760,8	766,3	0,2	48	Ветер, дующий с юго-запада	1			10% или менее, но не 0	
5	09.11.2018 09:00	11,8	760,6	766,2	0,8	57	Ветер, дующий с востоко-северо-востока	2			Облаков нет.	
6	09.11.2018 06:00	11,4	759,8	765,4	0,1	59	Ветер, дующий с востоко-северо-востока	3			Облаков нет.	
7	09.11.2018 03:00	11,2	759,7	765,4	-0,1	62	Ветер, дующий с востоко-северо-востока	2			Облаков нет.	
8	09.11.2018 00:00	13,1	759,8	765,4	0,1	54	Ветер, дующий с северо-северо-востока	2			Облаков нет.	
9	08.11.2018 21:00	13,8	759,7	765,2	-0,2	50	Ветер, дующий с востоко-северо-востока	1			Облаков нет.	
10	08.11.2018 18:00	14,8	759,9	765,4	-0,2	64	Ветер, дующий с северо-северо-востока	2			Облаков нет.	
11	08.11.2018 15:00	20,2	760,1	765,5	-0,4	46	Ветер, дующий с юго-запада	2			10% или менее, но не 0	
12	08.11.2018 12:00	20,1	760,5	765,9	-0,7	45	Ветер, дующий с юго-запада	2			20–30%.	
13	08.11.2018 09:00	11,9	761,2	766,9	0,3	62	Ветер, дующий с востоко-северо-востока	2			20–30%.	
14	08.11.2018 06:00	9,1	760,9	766,5	0,0	74	Ветер, дующий с северо-востока	2			Облаков нет.	
15	08.11.2018 03:00	11,3	760,9	766,5	-0,3	64	Ветер, дующий с северо-востока	2			Облаков нет.	
16	08.11.2018 00:00	12,9	761,2	766,7	0,0	56	Ветер, дующий с северо-северо-востока	2			Облаков нет.	
17	07.11.2018 21:00	13,5	761,2	766,8	-0,2	56	Ветер, дующий с северо-северо-востока	2			90 или более, но не 100%	
18	07.11.2018 18:00	15,2	761,4	766,9	0,1	62	Ветер, дующий с северо-северо-запада	1			70 – 80%.	
19	07.11.2018 15:00	20,8	761,3	766,7	-0,7	46	Ветер, дующий с юго-запада	2			40%.	
20	07.11.2018 12:00	18,6	762,0	767,5	-0,1	51	Ветер, дующий с юго-юго-востока	2			90 или более, но не 100%	
21	07.11.2018 09:00	11,1	762,1	767,7	0,6	64	Ветер, дующий с востоко-северо-востока	3			100%.	
22	07.11.2018 06:00	8,5	761,5	767,2	0,2	76	Ветер, дующий с северо-востока	2			Облаков нет.	Состояние неба в обще
23	07.11.2018 03:00	10,2	761,3	766,9	0,2	71	Ветер, дующий с северо-востока	1			Облаков нет.	Состояние неба в обще
24	07.11.2018 00:00	11,2	761,1	766,7	0,4	64	Ветер, дующий с северо-северо-востока	2			Облаков нет.	Состояние неба в обще
25	06.11.2018 21:00	12,1	760,7	766,3	0,6	54	Ветер, дующий с северо-северо-востока	3			Облаков нет.	Состояние неба в обще
26	06.11.2018 18:00	13,4	760,1	765,7	0,4	49	Ветер, дующий с северо-востока	3			Облаков нет.	Облака в целом рассеи
27	06.11.2018 15:00	16,3	759,7	765,2	-0,4	43	Ветер, дующий с северо-северо-востока	4			10% или менее, но не 0	Облака в целом образ
28	06.11.2018 12:00	16,1	760,1	765,6	-0,2	45	Ветер, дующий с северо-востока	4			Облаков нет.	Состояние неба в обще
29	06.11.2018 09:00	12,0	760,3	765,8	0,4	59	Ветер, дующий с северо-востока	3			Облаков нет.	Состояние неба в обще
30	06.11.2018 06:00	10,7	759,9	765,5	0,0	67	Ветер, дующий с северо-северо-востока	3		10	Облаков нет.	

Рис. 1. Пример данных в формате Excel-таблицы

Для решения поставленной задачи данные были систематизированы и предварительно обработаны для машинного обучения и использования метода Шепарда. Основной проблемой на данном этапе является неоднородность и неполнота данных, отсутствие различных показателей на конкретных отрезках времени, а также возможные ошибки и погрешности измерительных приборов.

Для уточнения и корректировки данных были использованы дополнительные алгоритмы предобработки для сглаживания и фильтрации на основе spline-интерполяции и медианных фильтров [Ferguson, 1964; Volshakov, 1978]. Используемые алгоритмы достаточно часто рекомендуют применять для первоначальной обработки сигналов в интересах машинного обучения. Особенностью данных алгоритмов является избирательность по отношению к выпадающим элементам массива, резко выделяющихся на фоне соседних отсчетов. Благодаря этой особенности исключаются или сглаживаются некоррелированные или слабо коррелированные помехи.

Характерной особенностью медианных фильтров является их нелинейность. Во многих случаях применение медианных фильтров необходимо для обработки реальных сигналов в случаях, когда перепады значений сигналов велики по сравнению с дисперсией аддитивного белого шума, что дает меньшее значение среднеквадратической ошибки по сравнению с оптимальными линейными фильтрами. Особенно эффективный медианный фильтр оказывается при очистке сигналов от импульсных шумов при обработке изображений, акустических сигналов, передаче кодовых сигналов и т. п.

К сожалению, часть данных за определенные периоды времени отсутствовало, а потому было принято решение заполнить отсутствующие значения случайными данными в диапазоне от последнего известного значения и вновь полученного значения спустя определенный период времени.

2. Метод Шепарда

Метод аппроксимации Шепарда – один из малоизвестных способов построения статических моделей, основанных на “нерегулярных данных”. Под нерегулярными данными в данном случае подразумевается набор обучающих выборок, расположение узлов в которых носит произвольный характер. Функция Шепарда представляет собой отношение двух дробно-рациональных функций, для её построения, в отличие от других способов аппроксимации, не требуется решения оптимизационных задач. Также используются обратные расстояния при вычислении весовых коэффициентов, с помощью которых взвешиваются значения экспериментальных значений в точках наблюдений. Таким образом, аппроксимация методом Шепарда довольно эффективно применяется при решении задач анализа данных, невыпуклой оптимизации и оптимального управления. Именно это и послужило основной причиной для выбора этого метода при решении задачи аппроксимации на основе больших данных.

Для реализации метода Шепарда средствами QT и Visual Studio 2019 был разработан программный комплекс на языке C++. В качестве входных данных программный комплекс принимал на вход предварительно обработанные параметры за определенный промежуток времени, из представленной выше на рис. 1 таблицы. Затем на основе известных данных предсказывались значения уровня воды в течение последующих 24 часов с 3-х часовым интервалом. На рис. 2 синей линией представлен график реального уровня воды, а красной – результат предсказания, сделанный с помощью метода Шепарда.

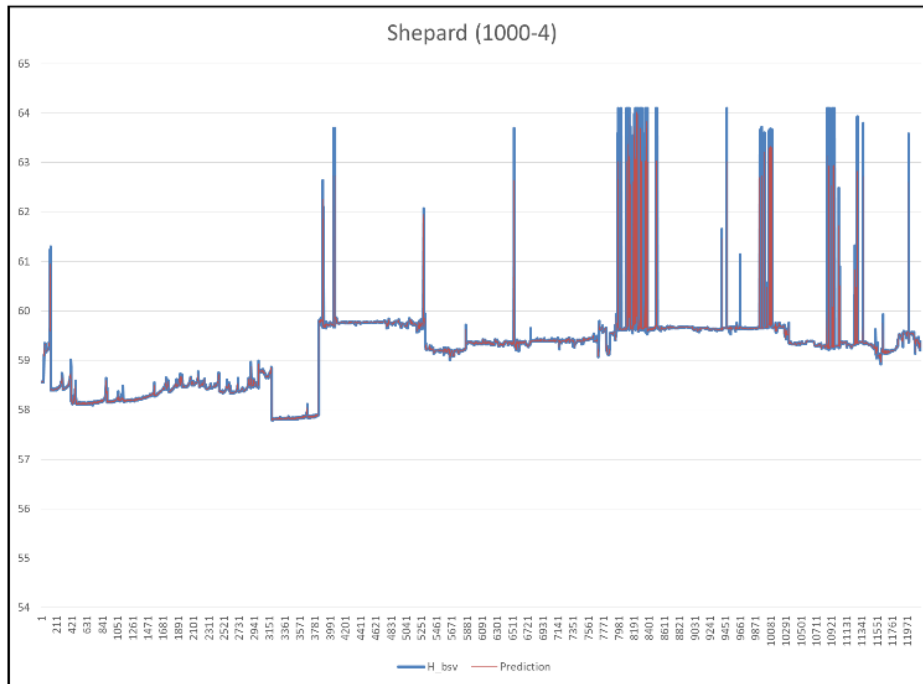


Рис. 2. График предсказания событий методом Шепарда

При этом на рис. 3 можно видеть Excel-таблицу, на основе которой был построен график, представленный на рис. 2, на котором левый столбец отображает реальный уровень воды в регионе, а правые столбцы, выделенные красным, отображают попытку предсказать данный уровень с помощью метода Шепарда с частотой в 3 часа. Интенсивность цвета отображает высоту уровня воды. Соответственно, чем интенсивнее цвет, тем выше уровень прогнозируемого либо реального значения.

63.67	34	9	59.655	59.654	59.655	59.677	59.694	59.695	59.698	59.699	59.695	59.698	59.697	59.654	59.685	59.683	59.697	59.699	59.683	59.688	59.699
63.67	34	12	63.094	63.178	63.631	60.3	59.811	63.963	63.669	63.863	63.657	63.638	63.626	63.603	63.683	63.814	63.559	63.697	63.694	62.437	60.02
59.71	34	15	63.67	63.669	60.5	59.811	63.662	63.669	61.67	63.657	63.641	63.649	63.645	63.649	62.514	63.649	63.669	63.665	63.666	63.585	62.904
63.58	34	18	59.977	60.5	59.811	63.662	63.667	62.551	63.628	61.02	60.308	63.623	62.649	62.911	63.655	63.114	60.839	60.611	60.756	62.541	62.218
59.7	34	21	60.5	59.811	63.662	63.667	62.696	63.626	63.139	61.691	63.636	62.671	62.912	63.678	63.458	63.631	62.062	62.631	62.396	61.252	62.29
59.7	35	0	59.811	63.662	63.668	62.598	62.898	59.79	62.623	63.628	62.862	62.96	63.631	63.631	61.272	63.721	62.281	62.499	62.184	62.238	61.996
59.7	35	3	61.454	63.66	62.598	62.898	59.79	62.623	63.628	62.868	60.533	63.631	63.631	61.272	63.721	62.281	62.499	60.498	61.277	61.655	62.733
59.7	35	6	63.658	62.598	61.838	59.79	62.623	63.655	62.857	59.995	61.82	63.631	63.631	61.272	63.721	62.281	60.766	60.514	60.996	63.681	61.72
59.7	35	9	62.502	59.842	59.776	62.623	63.662	62.881	59.705	61.397	61.395	61.202	63.631	62.279	59.775	59.756	59.83	59.974	60.942	60.233	60.309
59.69	35	12	59.701	59.706	62.622	60.249	60.272	59.703	59.921	60.538	60.055	61.876	62.266	59.725	59.724	59.735	59.756	59.955	59.834	59.887	59.907
59.71	35	15	59.698	62.236	59.801	59.722	59.703	59.692	59.695	59.695	60.437	61.892	59.718	59.722	59.727	59.732	59.764	59.778	59.798	59.862	59.731
59.69	35	18	62.235	59.801	59.722	59.703	59.692	59.695	59.695	60.431	61.888	59.718	59.72	59.724	59.725	59.755	59.728	59.728	59.729	59.729	59.738
62.68	35	21	59.811	59.722	59.705	59.692	59.695	59.695	60.424	61.882	59.718	59.719	59.72	59.728	59.727	60.541	59.725	59.728	59.713	59.729	59.732
59.7	36	0	59.705	60.354	59.541	59.775	63.635	63.638	61.793	60.096	59.93	59.886	60.333	59.736	59.73	59.714	59.749	59.750	59.757	59.773	59.907
59.7	36	3	59.707	59.743	59.772	63.638	63.638	60.554	59.957	59.878	59.842	60.002	59.735	59.73	59.714	59.749	59.762	59.756	59.77	59.874	60.205
59.69	36	6	59.743	59.772	63.638	63.638	60.422	59.704	59.713	59.739	59.982	59.717	59.723	59.71	59.702	59.703	59.703	59.705	59.759	60.098	59.73
59.69	36	9	59.69	63.504	63.632	60.235	59.704	59.71	59.709	59.708	59.708	59.715	59.709	59.702	59.703	59.704	59.705	59.704	59.704	59.704	59.704
59.68	36	12	60.711	63.598	60.223	59.704	59.709	59.704	59.699	59.703	59.706	59.709	59.702	59.703	59.703	59.704	59.715	59.725	59.722	59.723	59.743
62.68	36	15	63.63	60.216	59.703	59.704	59.686	59.681	59.682	59.682	59.698	59.703	59.703	59.703	59.703	59.69	59.687	59.699	59.703	59.703	59.703
63.66	36	18	60.216	59.703	59.704	59.686	59.681	59.682	59.682	59.698	59.703	59.703	59.703	59.703	59.69	59.687	59.699	59.703	59.703	59.703	59.703
63.66	37	0	59.703	59.704	59.686	59.682	59.682	59.699	59.702	59.703	59.703	59.703	59.703	59.693	59.702	59.703	59.711	59.703	59.705	59.745	59.724
59.69	37	3	63.649	63.649	59.698	60.037	59.756	59.822	61.698	63.122	61.417	62.554	62.611	60.386	60.44	60.19	63.673	63.653	63.551	62.598	60.482
59.69	37	6	59.706	59.913	59.741	59.723	59.75	59.738	59.708	59.726	59.722	59.888	60.085	59.822	59.868	60.14	60.303	59.955	59.972	59.928	59.906
63.64	37	9	59.724	59.693	59.694	59.704	59.73	59.703	59.711	59.7	59.706	59.74	59.708	59.718	59.915	59.894	59.766	59.749	59.799	59.81	59.906
63.64	37	12	59.69	59.69	59.693	59.729	59.703	59.709	59.697	59.691	59.692	59.693	59.695	59.673	59.794	59.722	59.706	59.703	59.709	59.739	59.893
63.68	37	15	60.618	60.63	59.79	59.708	59.711	59.708	60.217	63.8	60.142	60.79	59.908	60.693	59.961	60.315	59.929	59.922	60.072	59.949	59.937
59.7	37	18	63.632	59.747	59.724	60.005	61.234	61.279	61.555	63.662	63.637	59.985	60.382	61.209	63.129	62.349	62.325	62.983	60.511	60.828	61.346
59.7	37	21	59.747	59.724	59.396	60.995	61.872	61.942	60.209	60.035	59.984	60.37	61.081	62.06	61.252	61.335	61.033	60.44	60.667	60.962	61.207
59.71	38	0	59.724	59.952	60.703	63.669	63.941	59.714	59.722	59.944	60.369	61.062	63.775	59.813	59.771	59.737	60.252	60.386	60.412	60.89	61.347
59.7	38	3	59.71	60.181	61.669	63.942	59.714	59.722	59.873	60.333	60.962	61.787	59.758	59.757	59.736	59.71	59.711	59.727	60.686	61.336	61.338
59.7	38	6	59.7	61.668	61.941	59.714	59.722	59.71	60.029	59.974	61.837	59.744	59.749	59.735	59.709	59.706	59.705	59.911	61.183	61.295	60.411
59.69	38	9	61.667	63.941	59.714	59.722	59.71	60.038	59.936	61.828	59.712	59.718	59.724	59.706	59.702	59.701	59.806	61.078	61.209	60.988	60.203
59.69	38	12	61.808	59.714	59.722	59.707	59.662	59.717	60.301	59.708	59.718	59.718	59.706	59.703	59.697	59.702	60.351	60.746	59.944	60.057	59.896
63.64	38	15	59.713	59.722	59.706	59.719	59.704	59.803	59.698	59.713	59.717	59.706	59.703	59.697	59.694	59.702	59.71	59.713	59.713	59.715	59.745
63.67	38	18	59.752	59.713	59.968	59.838	60.238	59.71	59.717	59.725	59.706	59.701	59.698	59.979	62.439	63.243	62.323	62.117	60.501	60.836	62.218
59.7	38	21	61.394	60.469	59.805	59.703	60.114	61.278	60.194	60.279	59.691	59.693	59.726	60.083	61.14	61.311	60.85	60.744	59.727	59.767	59.827
59.69	39	0	59.982	59.995	63.134	61.262	63.596	61.937	59.71	59.719	59.849	60.775	62.413	63.327	60.905	60.561	60.667	61.05	62.178	62.827	63.081
59.69	39	3	59.69	63.092	61.238	61.596	63.938	59.709	59.713	59.821	60.257	62.324	63.091	60.904	60.561	59.694	59.718	60.238	62.078	62.852	62.855
59.69	39	6	62.236	61.255	61.594	63.938	59.709	59.715	59.802	59.703	61.203	62.825	60.894	60.56	59.691	59.696	59.705	60.402	61.708	62.183	61.792
59.69	39	9	61.225	61.275	63.938	59.709	59.712	59.717	59.709	60.424	62.997	60.803	60.545	59.691	59.693	59.711	60.156	61.391	61.748	61.543	61.849
63.61	39	12	61.173	61.809	59.709	59.71	59.708	59.703	60.114	61.208	60.194	60.279	59.691	59.693	59.726	60.083	61.14	61.311	60.85	60.739	59.725
63.64	39	15	61.63	61.496	60.393	59.79	59.703	60.114	61.278	60.194	60.279	59.691	59.693	59.726	60.083	61.14	61.311	60.85	60.744	59.727	59.767
59.7	39	18	61.394	60.469	59.805	59.703	60.114	61.278	60.194	60.279	59.691	59.693	59.726	60.083	61.14	61.311	60.85	60.744	59.727	59.767	59.827
59.71	40	0	60.462	59.804	59.703	60.114	61.278	60.194	60.279	59.691	59.693	59.726	60.083	61.14	61.311	60.85	60.744	59.727	59.767	59.827	59.861
59.7	40	3	59.705	59.703	60.114	61.278	60.194	60.279	59.691	59.693	59.726	60.083	61.139	61.31	60.85	60.739	59.727	59.767	59.823	59.846	60.363
59.7	40	6	59.703	60.114	61.278	60.193	60.278	59.691	59.693	59.726	60.083	61.138	61.308	60.846	60.733	59.723	59.751	59.775	59.823	59.851	60.287
59.69	40	9	60.111	61.238	60.377	60.267	59.694	59.708	60.077	61.125	61.286	60.804	60.498	59.711	59.717	59.714	59.718	59.722	59.91	59.91	

Рис. 3. Данные, полученные с использованием метода Шепарда

Метод Шепарда показывает приемлемые результаты аппроксимации на множестве повторяющихся событий, однако он неудовлетворительно решает задачу экстраполяции параметров еще не состоявшихся событий. Это связано с тем, что данный метод вносит сглаживающий эффект и прогнозируемые значения рассчитываются с помощью взвешенного среднего, полученного из уже измеренных значений исследуемых параметров [Shepard, 1968; Łukaszuk, 2004].

За счет того, что большую часть времени приведенные данные имеют стабильные значения, прогнозируемые параметры искомого уровня воды сглаживаются и далеки от реальных значений наводнений. По этой причине на большом количестве однородных временных данных предсказание на основе метода Шепарда неспособно предсказывать те или иные резкие всплески и отклонения от усредненных значений.

3. Машинное обучение нейронных сетей

Нейронная сеть — это математическая модель и ее программная реализация, построенная по принципу сетей нервных клеток живого организма. В последнее время использование подобного рода моделей приобрело массовый характер, за счет их возможностей решается довольно широкий спектр задач классификации и аппроксимации на основе большого объема данных.

При решении задачи с помощью нейронных сетей в первую очередь следует определиться с архитектурой, принципом построения и типом нейронной сети, так как все эти и многие другие параметры наравне с методом построения обучающей выборки напрямую влияют на качество обучения нейронной сети. Поскольку основной задачей было предсказание уровня воды в регионе, то есть выходное пространство решений сети известно, было принято решение использовать стандартный метод обучения с учителем.

При данном подходе входные данные делятся на несколько блоков, которые формируют обучающие выборки. Основной задачей при этом является выявление целевых зависимостей между множеством объектов описания и множеством допустимых ответов. То есть данный алгоритм должен вывести общие закономерности и зависимости из набора частных факторов или особенностей, что в данном случае является обобщением классических задач аппроксимации функций.

Для работы с нейронными сетями использовались открытые нейросетевые библиотеки Keras, которые являются надстройкой над фреймворком Tensorflow и нацелены на работу с нейронными сетями методами машинного и глубокого обучения [Chollet, 2017, Gulli, 2017]. Данный выбор обусловлен открытым исходным кодом проекта, а также поддержкой технологии CUDA и Cudnn, что значительно оптимизирует работу библиотеки и существенно ускоряет скорость обучения. Еще одной отличительной особенностью библиотеки Keras является ее направленность на работу с сетями глубокого обучения, что позволяет моделировать сложные нелинейные отношения и в конечном счете находить комплексные зависимости, которые невозможно найти при первичном анализе имеющихся данных.

Для реализации разработанных алгоритмов был применен компилятор PyCharm, который применялся при написании программы на языке Python версии 3.6. Данная программа осуществляла построение нейронной сети в зависимости от выбранной архитектуры, с последующим обучением разработанной модели на основе размеченных данных, которые предварительно были поделены на обучающую, тестовую и валидационную выборки. Стоит отметить, что для того чтобы составить корректные обучающие данные, а именно, учесть предыдущие показатели и долгосрочные зависимости, обучающие выборки были сформированы вручную, без автоматизированных алгоритмов сегрегации данных. Были проведены вычислительные эксперименты с нейронными сетями различной архитектуры и размерности. Наиболее интересные результаты представлены на рис. 4-6.

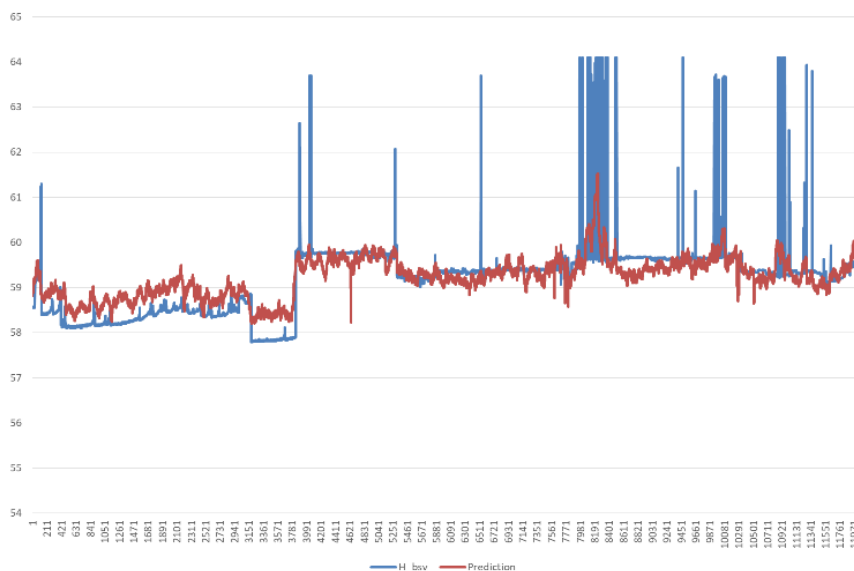


Рис. 4. График предсказания события методом нейронных сетей (30000 эпох, глубина теневого слоя 120)

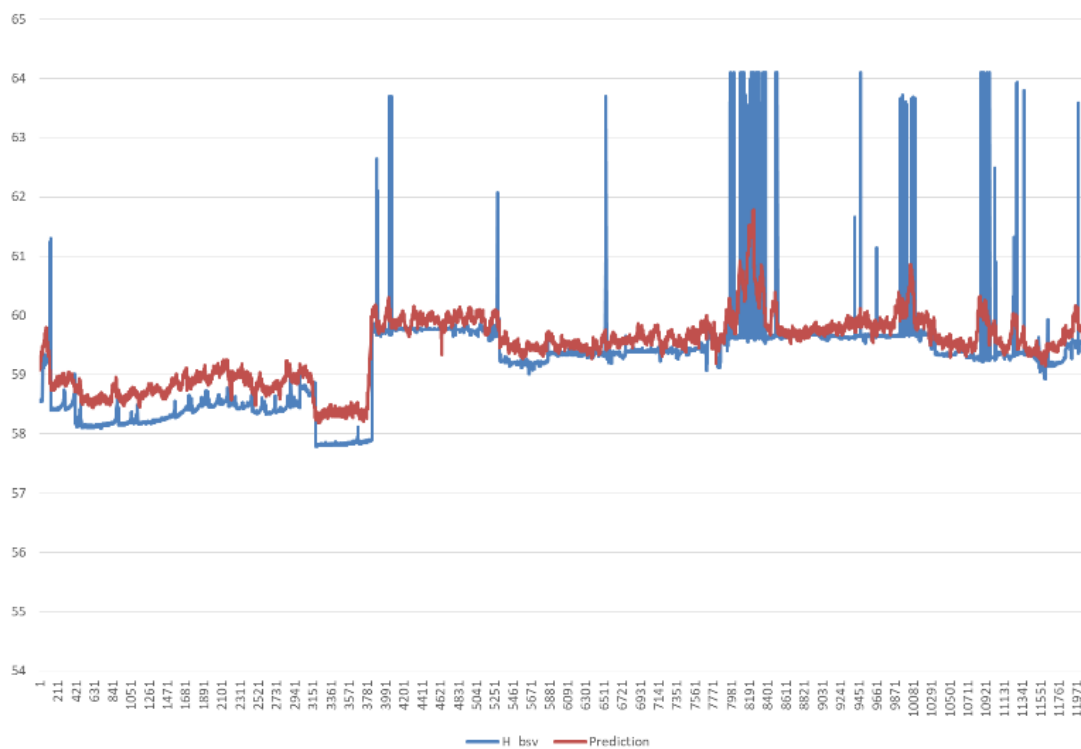


Рис. 5. График предсказания событий методом нейронных сетей (10000 эпох, глубина теневого слоя 200)

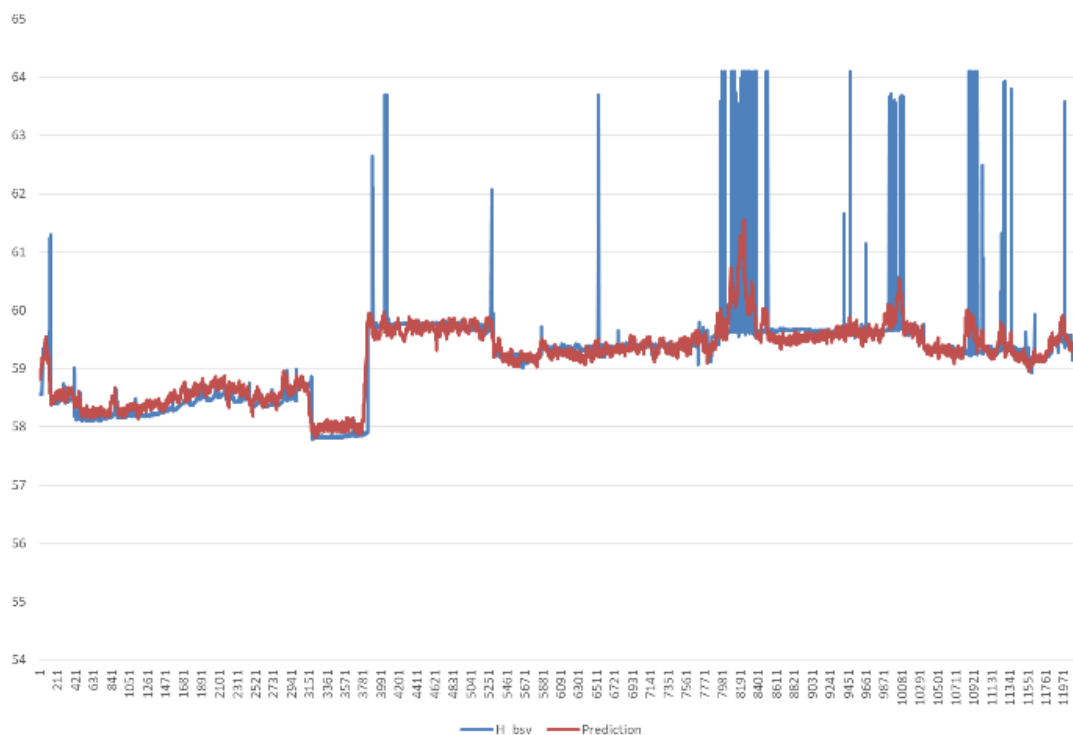


Рис. 6. График предсказания событий методом нейронных сетей (20000 эпох, глубина теневого слоя 300)

На приведенных выше графиках отчетливо видно, что в целом, несмотря на большой объем данных, на точность работы нейронной сети влияет непосредственно ее архитектура, но никак не количество итераций (эпох).

Заключение

В ходе данного исследования был проведен сравнительный анализ результатов прогнозирования наводнений методом нейронных сетей и методом Шепарда, который показал, что метод Шепарда неприемлем для экстраполяции предстоящих событий, а нейронные сети дают адекватный горизонт прогнозирования не более 3-х часов. В связи с этим можно констатировать, что нейросетевой подход может найти ограниченное применение для решения таких задач.

Очевидно, что для долгосрочного прогнозирования необходимо использовать гибридные методы нейросетевого и физического моделирования, например, комплекса гидродинамических (бассейновых) и нейросетевых моделей. Использование физических моделей позволяет обозначить наиболее приоритетные факторы, на которые следует обращать внимание при прогнозировании того или иного процесса. В то же время эта информация позволяет расставлять приоритеты в доступных данных при обучении нейронной сети.

Благодарности

Работа выполнена при финансовой поддержке РФФИ, грант № 20-31-90031.

Литература

1. F. Chollet. (2017) Deep Learning with Python. Manning Publications Co, CT, United, 2017. - 384 p.
2. D. Shepard. (1968) Two-dimensional interpolation function for irregularly-spaced data. // Proc. of the 1968 ACM National Conference. New York: ACM Press, 1968. – pp. 517-524.
3. S.P. Kopusov, I.M. Kuzmin, L.E. Tonkov. (2013) Methods for deforming meshes in conjugate problems. // Computational methods and programming, 2013. No 3. – pp. 269-278.
4. J.C. Ferguson. (1964) Multi-variable curve interpolation. J. ACM, Vol.11. April, 1964. – pp. 221-228.
5. I.A. Bolshakov, V.S. Rakoshiz. (1978) Applied theory of stochastic flows. Moscow: Sov. Radio, 1978 [in Russian].
6. S. Łukaszuk. (2004) A new concept of probability metric and its applications in approximation of scattered data sets. Computational Mechanics. 2004. 33 (4). – pp. 299-304.
7. Gulli, P. Sujit. (2017) Deep Learning with Keras: Implementing deep learning models and neural networks with the power of Python, 1st Edition, Kindle Edition. - M.: Packt Publishing. April, 2017. – 320 p.

Ключевые слова

Нейронные сети, нейросетевое моделирование, метод Шепарда, временные ряды, наводнение

*Карандеев Александр Андреевич, аспирант ИПМ им. М.В. Келдыша РАН,
млaд. науч. сотр. РЭУ им. Г.В. Плеханова,
KarAlex755@gmail.com*

*Осипов Владимир Петрович, к.т.н.,
вед. научн. сотр. ИПМ им. М.В. Келдыша РАН,
osipov@keldysh.ru*

Alexander Karandeev, Vladimir Osipov, Comparison of the capabilities of neural networks and the Shepard method for flood forecasting based on statistical data processing**Keywords**

Neural networks, neural network modeling, Shepard's method, time series, flooding
DOI: 10.34706/DE-2022-04-01

JEL classification C02 – Математические методы; M15 Управление информационными технологиями.

Abstract

The task of forecasting a particular natural disaster and related problems is a key one in many countries and regions. Since the timely adoption of appropriate preventive measures helps to minimize damage and expenses for the subsequent restoration of infrastructure, as well as to avoid human casualties. One of the key approaches to solving such problems is mathematical modeling and time series analysis. The article considers the problem of flood forecasting in the coastal city of Tuapse (Krasnodar Territory, Russian Federation), while the data provided by the Ministry of Emergency Situations of Russia is taken as the basis for research. Algorithms of primary data processing for noise leveling and their correction for subsequent analysis and use are considered. The results of a comparative analysis of the capabilities of neural network modeling, as well as the Shepard method for solving flood forecasting problems are presented. The advantages and disadvantages of each of the approaches are considered. The results of numerical experiments demonstrating the features of the application of each of the approaches are presented, and a number of conclusions are made based on the data obtained.