

Введение Как известно [1,2], к резкому возрастанию концентраций вредных веществ в приземном слое атмосферы приводят определенные сочетания метеорологических условий, которые называют неблагоприятными метеорологическими условиями (НМУ). Следовательно, актуальнейшей задачей является выделение и прогноз формирования таких групп НМУ за счет использования инновационных методов математического анализа и IT-технологий [3]. Первым этапом выделения групп НМУ, формирующих высокие приземные концентрации примесей является определение значащих, малозначащих и практически незначащих метеопараметров для прогноза высокого уровня загрязнения. Зная промежуток времени между возникновением неблагоприятных метеоусловий и проявлением высокого уровня загрязнения, следующим важным шагом становится адекватная интерпретация конкретного сочетания метеопараметров (групп НМУ), определение которых, согласно [1], должно осуществляться в каждом населенном пункте. Следовательно, актуальной задачей является разработка и апробация методики выделения групп НМУ для городов с особо опасными источниками загрязнения атмосферы – источниками полимерных производств [4]. Экспериментальная часть

Наиболее простым, и вместе с тем, надежным способом проведения такого анализа является исследование значений весов межнейронных связей сети-эксперта, которая оказалась победителем – сети, прогнозирующей наступление события (превышение ПДК) за два дня до формирования высокого уровня загрязнения. Если веса, соединяющие входные нейроны со скрытым слоем, оказываются близки к нулю по модулю (порядка 10^{-5}), то соответствующие этим входным нейронам параметры можно считать мало- или незначащими. Исследование весов сети-эксперта победителя дало следующие результаты (табл.1):

Таблица 1
Значения весов нейронов входной слой - скрытый слой

Номера нейронов скрытого слоя	1	2	3	4	Номера нейронов вход. слоя	1	5.63	-2.09	9.30	-8.06	2	1.33
1	1.21	3.98	-4.22	3	-10.65	-6.29	-6.85	-10.33	4	3.74	-2.72	2.57
5	-2.50	16.97	0.68	-3.40	Из таблицы 1 следует, что модули всех весов значительно отличаются от нуля, то есть, что ни один из входных параметров нельзя назвать малозначащим. Все измеряемые метеорологические параметры (температура, давление, скорость ветра, направление ветра, атмосферное давление) в значительной степени влияют на формирование высокого уровня загрязнения приземного слоя воздуха. Анализ особенностей сочетания метеопараметров, проявляющихся в критическом временном интервале, проводился методом кластерного многомерного анализа. Попытки разделить множества неблагоприятных и благоприятных сочетаний метеопараметров по априори неопределенным значениям факторов проводились четырьмя методами [5-8]: агломеративным методом минимального локального расстояния, дивизимным методом МакНотон-Смита, методом k-средних Мак-Кина и методом самоорганизующихся карт Кохонена. В качестве меры расстояния выбрана							

Евклидова метрика. Во всех методах предпринималась попытка разделить все множество метеоусловий на три возможных кластера («благоприятные», «неблагоприятные», «неопределенные» метеоусловия). Эксперимент показал следующие результаты: 1. Агломеративный метод. В качестве минимального межгруппового расстояния в серии экспериментов подобран параметр, равный 0,5. В результате получено разбиение всего множества метеоусловий на три кластера. Качественный состав каждого кластера по принадлежности благоприятных и неблагоприятных сочетаний факторов составляет: . 1) Процент благоприятных сочетаний метеоусловий 1.1) Кластер №1- 52% 1.2) Кластер №2 - 80% 1.3) Кластер №3 - 50% 2) Процент неблагоприятных сочетаний метеоусловий 2.1) Кластер №1- 48% 2.2) Кластер №2 - 20% 2.3) Кластер №3 - 50%

2. Дивизимный метод МакНотон-Смита. Максимальное межгрупповое расстояние в результате серии экспериментов выбрано равным 0,6. Множество метеопараметров в результате было разбито на два кластера с параметрами: 1) Процент благоприятных сочетаний метеоусловий 1.1) Кластер №1- 51% 1.2) Кластер №2 - 80% 2) Процент неблагоприятных сочетаний метеоусловий 2.1) Кластер №1- 49% 2.2) Кластер №2 - 20%

3. Метод k-средних Мак-Кина. Задано число групп разделения, равное трем. Получены следующие результаты: 1) Процент благоприятных сочетаний метеоусловий 1.1) Кластер №1- 50% 1.2) Кластер №2 - 50% 1.3) Кластер №3 - 70% 2) Процент неблагоприятных сочетаний метеоусловий 2.1) Кластер №1- 50% 2.2) Кластер №2 - 50% 2.3) Кластер №3 - 30%

4. Метод самоорганизующихся карт Кохонена. Выделено три кластера. Распределение сочетаний благоприятных и неблагоприятных метеоусловий рассеивания примесей в атмосфере составили: 1) Процент благоприятных сочетаний метеоусловий 1.1) Кластер №1- 60% 1.2) Кластер №2 - 38% 1.3) Кластер №3 - 58% 2) Процент неблагоприятных сочетаний метеоусловий 2.1) Кластер №1- 40% 2.2) Кластер №2 - 62% 2.3) Кластер №3 - 42%

Данные результаты говорят о том, что множества неблагоприятных и благоприятных сочетаний метеоусловий практически неразделимы. Так как в явном виде определить интервальные оценки неблагоприятных значений или сочетаний метеопараметров не удалось, для отнесения некоторого произвольного метеонаблюдения, произведенного в критическом временном интервале, к группе НМУ необходимо разработать специальный алгоритм. Наиболее простым, и вместе с тем эффективным способом такой идентификации является использование в качестве эксперта уже спроектированной и обученной ранее нейронной сети для классификации, которая оказалась победителем в процессе оценки критического временного интервала. Действительно, данная сеть для группы НМУ выдаст ответ «1», а если параметры являются благоприятными – «0». Подобный ответ сети-эксперта для введенного вектора метеонаблюдений даст ответ на вопрос, стоит ли в течение следующих двух дней ожидать высокого уровня загрязнения атмосферного воздуха. В случае положительного

ответа системы принятия решений могут инициировать мероприятий по оперативному регулированию качества атмосферного воздуха, согласно действующему руководящему документу [9]. Обобщая результаты проведенных исследований, разработана следующая методика создания и использования нейронных сетей для прогнозирования превышения концентраций токсикантов в атмосферном воздухе над ПДК на основе ряда измерений концентрации загрязняющих веществ: 1. Сформировать временной ряд наблюдений по метеопараметрам в сочетании с фактами превышения ПДК концентрациями загрязняющих веществ в приземном слое атмосферы. 2. Выдвинуть гипотезы относительно критического временного интервала – периода времени между появлением НМУ и проявлением высокого уровня загрязнения. 3. Для каждой гипотезы построить сеть типа «сеть для классификации» с парадигмой MLP и функцией классификации «кросс-энтропия», обучить сети. 4. Сеть, показавшую минимальную ошибку при обучении, считать экспертом-победителем.

Соответствующую гипотезу о критическом временном интервале считать достоверной. 5. В дальнейшем при получении нового набора метеопараметров подвергнуть его обработке экспертом-победителем. В случае положительного ответа эксперта считать, что в определенном им критическом временном интервале следует ожидать превышения ПДК. Разработанная методика была апробирована для прогнозирования высокого уровня загрязнения атмосферного воздуха в г. Нижнекамске в течение 335 дней в период с 2008-2009 гг. Обучение и проверка адекватности построенной нейронной сети проводились на экспериментальных данных. Сопоставление экспериментальных и расчетных значений параметра P показано на рис.1. Основные параметры полученных результатов: Минимальное значение ошибки – 0,0002 (0,4%) Среднее значение ошибки – 0,005 (0,9%) Максимальное значение ошибки – 0,06 (12%) Корреляция между экспериментальными и рассчитанными данными – 0,99. Рис. 1 -

Сопоставление расчетных и экспериментально измеренных значений параметра P . Для определения критического временного интервала выдвигалось пять гипотез: превышение ПДК произойдет в тот же день, в который сформировались НМУ (гипотеза №1), через день (гипотеза №2), через два дня (гипотеза №3), через три дня (гипотеза №4) и через четыре дня (гипотеза №5). В качестве сетей-экспертов использовались сети для классификации с функцией классификации кросс-энтропия. Эксперт-победитель подтвердил гипотезу № 3: задержка между формированием НМУ и превышением ПДК составляет от 24 до 48 часов. Ошибка обучения эксперта-победителя (0,48) отличалась от максимальной ошибки экспертов других гипотез почти в два раза (максимальную ошибку показал эксперт, подтверждающий гипотезу №5: 0,72). Эксперт – победитель показал следующие результаты: производительность – 0,82, ошибка обучения- 0,476349. Результаты распознавания групп НМУ в эксперименте указаны в табл.2: Таблица 2 - Результаты распознавания групп

НМУ экспертом – победителем Характеристика групп метеоусловий
Характеристика групп НМУ Благоприятные (ответ эксперта -0) Неблагоприятные
(ответ эксперта - 1) Всего исследовано групп МУ 97 38 Правильно распознано 80
31 Ошибочно распознано 17 7 Не распознано 0 0 % правильно распознанных
82,47423 81,57895 % ошибочно распознанных 17,52577 18,42105 %
нераспознанных 0,00000 0,00000 Таким образом, точность идентификации
набора метеопараметров составила в среднем 82%. Результаты предсказания
критического временного интервала экспертом-победителем: · Количество дней
наблюдений - 335 · Количество зафиксированных превышений ПДК токсикантов
в атмосфере (дней) - 92 · Кол-во правильно предсказанных временных задержек
- 72 · Кол-во ошибочно предсказанных временных задержек - 20 · Кол-во
правильно предсказанных временных задержек в % -78 · Кол-во ошибочно
предсказанных временных задержек в %- 22 Таким образом, точность прогноза
критического временного интервала составила 78%, Разработанная методика
является оптимальным вариантом обеспечения быстрого реагирования для
систем принятия решений в периоды НМУ. Данная методика универсальна и
может использоваться для проведения работ на территории любых населенных
пунктов, при условии обучения спроектированных нами нейросетей на
результатах экспериментальных наблюдений за содержанием загрязняющих
веществ в приземном слое атмосферы и на измеряемых величинах
метеопараметров.