

Введение Известно, что атмосферные процессы инерционны. Сохранение высоких концентраций примесей после дней с высоким уровнем загрязнения атмосферы отмечается в [1-3]. Поэтому, при прогнозировании высокого уровня загрязнения следует учитывать инерционный фактор, в качестве которого обычно используется значение параметра P в предшествующий прогнозу день. Нами ставилась задача рассчитать инерционный фактор, т.е. значения параметра P_{ϕ} предшествующего дня. Но учет инерционного фактора оставляет открытым вопрос о временном интервале возникновения высокого уровня загрязнения атмосферы после формирования групп неблагоприятных метеорологических условий (НМУ). Однако, при наличии длительного ряда измерений уровня концентраций загрязняющих веществ в атмосферном воздухе города в совокупности с значениями метеорологических параметров с использованием современных информационных технологий можно определить за какой период до возникновения высокого уровня загрязнения воздуха сформировались НМУ. Экспериментальная часть Как следует из табл. 1, которая составлена по материалам [4] при исследовании качества атмосферного воздуха в группе городов, с преобладающими выбросами источников полимерных производств, параметр P зависит от значения параметра P_{ϕ} в предшествующий день. Теснота связей между показателями загрязнения воздуха в различные периоды снижается с увеличением промежутка времени между этими периодами. Таблица 1 - Повторяемость (%) повышенного уровня загрязнения воздуха в зонах действия полимерных производств P' В общем случае При застое воздуха 0,1 11 24 ³ 0,3 92 95 Анализ результатов наблюдений на стационарных постах наблюдений за качеством атмосферного воздуха в г.Нижнекамске показал, что если в предшествующий день значение параметра P велико, то и в текущий день загрязнение воздуха, как правило, повышено. Обратная картина имеет место, когда значение обобщенного по городу показателя загрязнения в предшествующий день мало (P_{ϕ} 0,1). В этом случае в последующем загрязнение воздуха чаще всего понижено, в том числе и в такой неблагоприятной ситуации, как застой воздуха. Коэффициент корреляции между значениями параметра P в соседние дни составляет 0,6—0,7. Положительная корреляция имеет место между значениями параметра P , рассчитанными по данным наблюдений в 1-й и 2-й половине одного и того же дня. Даже в летний сезон при большом суточном ходе метеорологических параметров коэффициент корреляции между загрязнением воздуха в 1-й и во 2-й половине дня составил $r = 0,61 \pm 0,10$. Действие инерционного фактора в значительной степени определяется метеорологической инерцией, которая означает тенденцию к сохранению атмосферных процессов, определяющих уровень концентраций. Некоторые из метеорологических факторов, влияющих на концентрации примесей в воздухе, могут быть неизвестны, и при учете установившегося уровня загрязнения воздуха они в какой-то степени учитываются автоматически [5]. Существенную

роль может играть и инерция самого загрязнения воздуха. Нами спроектирована нейронная сеть для прогнозирования параметра P в зависимости от входных данных – метеоусловий. Если рассчитанное нейросетью значение P (P' при прогнозировании) ≤ 0.1 то при любом сочетании метеовеличин параметр P следующего дня не будет превышать $P = P' + (P' \cdot 0.89)$, где 0.89 это приведенное значение 89% сохранения низкого P (11% повышения P), если же $P' > 0.3$, то отмечаем, что параметр P следующего дня не будет ниже $P = P' - (P' \cdot 0.08)$, где 0.08 это 8% вероятность снижения (92% сохранения высокого P). В случае, когда рассчитанные нейросетью значения P' находятся в интервале 0,1-0,3 мы можем прогнозировать только интервал колебаний, взяв за основу коэффициент корреляции 0.7, то есть P будет в пределах от $P' - (P' \cdot r)$ до $P' + (P' \cdot r)$, где $r = 0.7$. Таким образом, инерционный фактор четко проявляется только на крайних значениях P' . Рассмотрим пример: был отмечен низкий уровень параметра P' , равный 0,1. Согласно условию $P' \leq 0,1$, параметр P следующего дня тоже будет небольшим, и не будет превышать $P = 0,1 + (0,1 \times 0,89) = 0,189$. Таким образом, при наличии длительного ряда измерений уровня концентраций загрязняющих веществ, в атмосферном воздухе города в совокупности с данными метеорологических параметров за этот же период времени, можно составить достаточно точный прогноз вероятности повышения приземных концентраций примесей в конкретном городе. Но определение периода до формирования высокого уровня загрязнения вследствие НМУ (так называемого критического временного интервала) не описано в нормативных и руководящих документах. Нейронные сети, с успехом применяемые для прогнозирования значений концентраций токсикантов в приземном слое атмосферного воздуха [6], могут также использоваться для определения критического временного интервала. В этом случае необходимо сформировать несколько нейронных сетей-экспертов, каждая из которых может давать прогноз относительно наступления превышения ПДК в своем временном интервале [7]. Тот эксперт, который покажет лучшие результаты прогнозирования на обучающем и тестовом множествах, является победителем и временной интервал, охватываемый им, и является искомым критическим интервалом. Превышение предельно допустимой концентрации хотя бы по одному из загрязняющих веществ характеризуется как высокий уровень загрязнения. Отдаленность формирования высокого уровня загрязнения в результате возникновения НМУ не может превышать четырех дней. Исходя из этого, для определения критического временного интервала, т.е. истинного временного промежутка между моментом формирования НМУ и днем превышения ПДК по любому из загрязняющих веществ, можно выдвинуть три взаимоисключающие гипотезы: 1) Превышение ПДК произойдет через четыре дня после возникновения НМУ. 2) Превышение ПДК произойдет через два дня после возникновения НМУ. 3) Превышение ПДК произойдет на тот же день, в который сложились НМУ. Данные гипотезы необходимо сравнить по степени

достоверности. Для этого необходимо построить нейронные сети, соответствующие каждой выдвинутой гипотезе. Для обучения каждой на входы нейронных сетей подаются метеоданные за охватываемый сетью временной период, а на выход – факт превышения (или не превышения) ПДК в конце периода, охватываемого сетью. Проиллюстрируем данную идею на примере. Первая сеть проверяет гипотезу, что превышение ПДК произойдет через четыре дня после формирования НМУ, вторая - что превышение ПДК произойдет через два дня после формирования НМУ, а третья – что превышение ПДК произойдет в тот же самый день (рис.1). На вход первой сети подаются метеоданные за текущий день измерений, на вход второй – метеоданные за второй день, на вход третьей сети – данные за четвертый день. Так как превышения ПДК не произошло ни через четыре, ни через два, ни в течение текущего дня, ответ всех нейросетей должен быть равен 0. Таким образом, для каждой сети-эксперта формируется первый обучающий набор данных. После этого интервал поиска сдвигается на два дня вправо. На вход первой сети, таким образом, подаются метеоданные второго дня измерений, второй сети – четвертого дня, третьей сети – шестого дня. Рис. 1 - Условия для формирования первого обучающего набора данных. Так как вновь превышения ПДК не произошло в конце ни одного из периодов, все сети должны иметь нулевой выход (рис.2). Рис. 2 - Условия для формирования второго обучающего набора данных. Таким образом, формируется второй набор обучающих данных для всех трех нейросетевых экспертов. Наконец, сдвинем интервал поиска еще на два дня вправо. Теперь на входы первой, второй и третьей нейросетей подаются метеоданные четвертого, шестого и восьмого дней соответственно. Для первой нейронной сети гипотеза подтвердилась: через четыре дня действительно произошло превышение ПДК. Так же, как и для второй и для третьей сети-эксперта. Таким образом, в качестве выхода всех трех сетей необходимо задать 1 (рис.3). Рис. 3 - Условия для формирования третьего обучающего набора данных. Продолжая подобным образом сдвигать интервал поиска, будут сформированы обучающие наборы для всех трех сетей. Каждый нейросетевой эксперт обучается по одному и тому же алгоритму одинаковое количество эпох. Для данной методики тип и топология построенных сетей-экспертов являются несущественными моментами, главное, чтобы для всех гипотез использовались сети с единой топологией. Выигрывает тот эксперт, ошибка обучения которого оказалась минимальной. Эксперт-победитель в дальнейшем не может использоваться для непосредственного прогнозирования. Он лишь определяет критический временной интервал, который впоследствии и должен использоваться для прогнозирования, например, многоуровневым нейросетевым экспертом. Проведем численные эксперименты. Эксперимент №1. Построение нейросетей-экспертов для прогнозирования единичного события – превышения ПДК – в зависимости от метеорологических условий при помощи сети MLP. Как уже рассматривалось,

топология нейросетевых экспертов, проверяющих каждую гипотезу, не играет существенной роли в определении критического временного интервала. Поэтому для прогнозирования была выбрана простейшая топология нейросети – многослойный персептрон (MLP). Общие параметры и характеристики сетей:

Структура: число нейронов входного слоя – 5, количество скрытых слоев – 1, число нейронов в скрытом слое – 4, количество нейронов выходного слоя – 1 (рис.4). Параметры обучения: двухэтапное обучение. · Первый этап – метод обратного распространения ошибки. Кол-во эпох – 1000. · Второй этап: метод сопряженных градиентов. Кол-во эпох – 5000.

Сеть №2. Гипотеза: Превышение ПДК произойдет через два дня. Характеристики: Производительность: 0,59 Ошибка обучения: 0,27 Средняя ошибка: 0,00015 Корреляция: 0,80

Сеть №3. Гипотеза: Превышение ПДК произойдет через четыре дня. Характеристики: Производительность: 0,78 Ошибка обучения: 0,34 Средняя ошибка: 0,000015 Корреляция: 0,63

Рис. 4 - Общая структура сетей типа MLP для нейросетевых экспертов

Сеть №1. Гипотеза: Превышение ПДК произойдет в этот же день. Характеристики: Производительность: 0,66 Ошибка обучения: 0,30 Средняя ошибка: 0,000026 Корреляция: 0,75

Эксперимент №2. Построение нейросетей-экспертов для прогнозирования единичного события – превышения ПДК – в зависимости от метеорологических условий при помощи сети для классификации. Так как в качестве выхода нейросетевых экспертов выступают булевы переменные, прогнозирование превышения ПДК можно рассматривать как классификацию входных векторов измерений метеопараметров. При этом определяется два класса – 0 и 1. Поэтому, для подтверждения результатов, полученных при помощи сети MLP, целесообразно провести тот же эксперимент для экспертов построенных по принципу нейронной сети для классификации.

Общие параметры и характеристики сетей: Структура: число нейронов входного слоя – 5, количество скрытых слоев – 1, число нейронов в скрытом слое – 6. Параметры обучения: двухэтапное обучение. · Первый этап – метод обратного распространения ошибки. Кол-во эпох – 1000. · Второй этап: метод сопряженных градиентов. Кол-во эпох – 5000.

Сеть №1. Гипотеза: Превышение ПДК произойдет в этот же день. Характеристики: Производительность: 0,80 Ошибка обучения: 0,59 Процент правильно распознанных примеров: 79%

Сеть №2. Гипотеза: Превышение ПДК произойдет через два дня. Характеристики: Производительность: 0,82 Ошибка обучения: 0,48 Процент правильно распознанных примеров: 82%

Сеть №3. Гипотеза: Превышение ПДК произойдет через четыре дня. Характеристики: Производительность: 0,65 Ошибка обучения: 0,72 Процент правильно распознанных примеров: 62%

Таким образом, разработанная методика определения критического временного интервала является достоверной и эффективной, что подтверждается численными экспериментами. В практическом плане для прогнозирования высокого уровня загрязнения атмосферного воздуха в городе - наилучшие результаты показал

нейросетевой эксперт, прогнозирующий превышение ПДК за два дня до самого события. Следовательно, критическим временным интервалом является промежуток времени от 24 до 48 часов до превышения концентраций загрязняющих веществ.