

УДК 004.94

Д.В. Рагозин

## МЕТОД МОНИТОРИНГА АНОМАЛИЙ ЛОКАЛЬНОЙ ПОГОДЫ НА ОСНОВЕ ТЕХНОЛОГИЙ СЕНСОРНЫХ СЕТЕЙ

В статье предложен метод уточнения метеорологического прогноза на основании поиска аналогий погоды в истории местных метеорологических наблюдений, полученных с помощью сенсорной сети, и истории прогнозов погоды. Предложена методика предупреждения местных опасных метеорологических явлений на основе поиска аналогий погоды в истории метеорологических наблюдений. Также продемонстрированы и проанализированы результаты обработки массива погодных данных с целью поиска погодных аномалий.

Существующие модели расчёта погоды (яркий пример – GFS) – имеют известный недостаток, связанный с достаточно большими размерами ячейки расчётной сетки – от 20 до 50 км. Это означает, что прогноз погоды для конкретного участка в рамках большого квадрата неточен, поэтому для современных видов деятельности, зависящих от погоды (точное земледелие, обработка городских насаждений и специальных угодий), необходимы методы, позволяющие уточнять прогноз внутри ячейки сетки модели прогноза.

### Экономические предпосылки уточнения прогноза

Мониторинг и анализ погоды являются критичным для ряда отраслей экономики. Если бы глобальный (всемирный) прогноз погоды в точности исполнялся, это бы позволило резко уменьшить потери, возникающие в связи с неожиданным наступлением аномальных явлений. Аномальные явления в нашем случае трактуются широко: от месячной засухи и более низких средних температур в течении недели до ураганных порывов ветра в течении нескольких минут или осадков в виде града.

Для оценки аномальных явлений необходимо в статистических целях охарактеризовать климат территории. Под статистической характеристикой климата будем понимать усреднённое значение погоды на данной территории, принимая во внимание, что это усреднённое значение – временные ряды за время от трёх десятилетий, и это значение определяет хозяй-

ственную деятельность территории. Усредняемых погодных факторов может быть очень много, тем не менее метеорологи выделяют ряд параметров (температура, давление, влажность, осадки, скорость и направление ветра), которые оказывают основное влияние на погоду (климат) территории. Рельеф местности и наличие водоёмов могут очень сильно влиять на статистику местных наблюдений за погодой, но это единственная доступная статистика о климате территории для задачи уточнения прогноза погоды.

Хорошим примером попытки уточнения прогноза погоды является старый метод предсказания погоды – по Замбретти, в основу которого положены долговременные наблюдения за среднеевропейским климатом. Статистика изменения европейской погоды в целом подтверждает на экспериментальных данных метод Замбретти [1] (который сам по себе не даёт точных данных, а лишь подсказывает тенденцию изменения погоды). Однако, горные местности (Карпаты, Крымские горы) имеют другую статистику, которая плохо укладывается в прогнозирование по Замбретти.

Неоднократно отмечалось [2], что повысить точность прогнозирования позволяет расширение сети метеостанций. Для точного земледелия (как зерновых культур, так и садоводства и овощеводства) или транспортных задач плотность сети станций может быть «одна станция на поле» или «одна станция на теплицу». Основной проблемой расширения и модернизации сети метеостанций является минимальное участие человека в съёме показаний датчиков и минимальная стоимость

обслуговування такої системи. Предпочтительна необслуговувана мережа метеостанцій, постійно аналізує параметри оточуючої середовища і дає аналіз розвитку подій.

Далі будемо позначати місцевий (або місцевий) клімат як  $L_i$ , глобальний клімат як  $G_i$ , індекс  $i$  вказує на територію прогнозу.

### **Застосування сенсорної мережі для аналізу погоди**

Одним з економічно вигідних способів збору даних про місцеву погоду є застосування сенсорних мереж з погодними датчиками. Для отримання погодних даних в базі сенсорних мереж важливі декілька параметрів: 1) бездротовий зв'язок датчика з мережею Інтернет, можливо використовуючи інші датчики своєї мережі як ретранслятори сигналу для забезпечення мінімальної кількості точок доступу в мережу Інтернет; 2) економічне споживання енергії для збільшення інтервалу між вимірами датчиків; 3) місцевий аналіз даних з сегмента сенсорної мережі з доступом до даних з мережі – «географічних сусідів»; 4) простота установки і обслуговування. Економічні і споживачі параметри при розгляді рішення на базі сенсорної мережі переважають.

Окремо розташована сенсорна мережа, пов'язана з мережею Інтернет, отримує прогноз погоди, місцеві дані про погоду і повинна дозволити користувачеві прогнозувати зміну місцевого клімату, враховуючи вимірювану поточну погоду і історичні погодні дані в цій точці. Якщо сенсорна мережа отримує дані з сусідніх мереж, то це забезпечує прорив з точки зору прогнозування погоди, так як аналіз даних може охоплювати не тільки місцеві дані з ділянок, «накритих» сенсорною мережею, але і з найближчих і дальніх сусідніх мереж. Отримання і аналіз даних з мереж-сусідів дозволяє в режимі реального часу (15 хвилин) помічати відмінності показань близько і далеко розташованих сусідніх мереж від гло-

бального прогнозу погоди і попередити користувача про непередбачену ситуацію – непередбаченому зміні погоди, що дозволяє заздалегідь підготуватися до захисту від погодніх явищ.

Розглянемо докладніше дані, які потенційно можуть збиратися сенсорною мережею. Для складання цього списку було взято список параметрів, які фіксуються в записках офіційних метеостанцій Всесвітньої метеорологічної організації ВМО. Опущено параметри, які необхідно заповнювати з участю людини (наприклад, дані про висоту хмарності, хоча зняття цього параметра теж можна автоматизувати):

- температура повітря;
- атмосферний тиск на рівні станції і приведений до рівня моря;
- відносна вологість на висоті 2 м над поверхнею Землі;
- напрямок вітру (румби) і швидкість вітру на висоті 10–12 метрів над поверхнею, усереднені за 10-хвилинний період, що передує часу спостережень;
- максимальне значення пориву вітру на висоті 10–12 метрів над поверхнею, усереднені за 10-хвилинний період, що передує часу спостережень;
- максимальне значення пориву вітру на висоті 10–12 метрів над поверхнею, за період між датами спостережень;
- мінімальна і максимальна температура повітря, за період між датами спостережень;
- кількість всіх спостережуваних хмар;
- кількість опадів, висота снігового покриву.

Приведений перелік даних не вичерпуючий. Розташування сенсорного блоку на буй в воді дозволяє мати дані про температуру води і хвилювання моря, солоність і т.д. Можливі виміри температури і вологості ґрунту на різних глибинах (20 см, 50 см, 1 м), що дозволяє мати інформацію про стан ґрунту і використовувати цю інформацію в сільському господарстві і комунальних службах.

Для изучения локальных особенностей погоды необходим анализ собранных погодных данных, в частности постоянных расхождений прогноза глобальной погоды и локальной погоды. Если расхождения глобального прогноза и локальных данных имеют постоянную (в рамках некоторого классификатора) компоненту, то возможно (в рамках классификатора) предсказание разности глобального прогноза и локальной погоды в определённых условиях. Далее опишем постановку задачи. Постановка задачи опирается на широко известные исследования [3, 4], в рамках которых предпринято сравнение погоды в течении длительного времени (от 20 лет) в рамках больших участков территории (до нескольких сотен километров в поперечнике) с целью нахождения корреляций развития погоды. При этом сеть метеостанций была достаточно редкой, и снятие данных до десятка раз в сутки, и не учитывались особенности частей территории. Наша задача – масштабирование погодных данных до конкретного участка территории, покрываемой сенсорной сетью, с целью прогнозирования изменения локальной погоды. Для этого используются три источника данных: 1) исторические данные о локальной погоде на территории (в регионе); 2) история глобальных прогнозов для ретроспективного сравнения глобального прогноза и реальных местных погодных данных; 3) глобальный прогноз для территории на ближайшее время.

### Метод поиска аналогий для локальной (местной) погоды

В качестве исходных данных для прогнозирования используются данные от датчиков сенсорной сети и глобальный прогноз на участок территории  $T$ . Обозначим ансамбль параметров  $S^T = \{s_1^T, s_2^T, \dots, s_k^T\}$  – параметры, измеряемых датчиком сенсорной сети для локального участка территории. Ансамбль необязательно охватывает все возможные датчики устройства, для устройства можно определить множество ансамблей  $\{S^T\}$ , так как не по всем ансамблям возможно

предсказание. Значение  $s_i^T$  – значение параметра, измеренного датчиком узла сети; параметр, измеренный в момент времени  $t$  обозначим как  $s_i^T(t)$ , значение ансамбля параметров в момент  $t$  обозначим как  $S^T(t)$ . Соответственно,  $N$  последовательных измерений погоды начиная с момента  $t$  для территории будут являться последовательностью ансамблей  $S^T(t), S^T(t+1), \dots, S^T(N)$ . Обозначим эту последовательность как  $S^T(t:N)$ .

В рамках задачи нам необходимо найти «подобные» последовательности измеренных значений ансамблей параметров для одной и той же точки в разное время, предварительно определив критерий подобия. После чего можно сравнить соответствующие ансамбли с целью выявить корректность глобального прогноза. «Подобие» локальных данных не означает, что до или после последовательности данных не присутствовало действие определённых неучитываемых нами глобальных факторов, из-за которых прогноз на будущее (относительно взятой временной последовательности) будет разный. На момент исследования анализ всех возможных прогнозов для некоторой последовательности не ставится из-за высокой вычислительной сложности задачи.

Для сравнения последовательностей ансамблей введём метрику подобия  $M(S^T(t:N), S^T(t_x:N))$ . Это метрика сравнивает последовательности и результатом сравнения является значение подобия в интервале  $[0; 1]$ . При значении метрики 0 последовательности полностью различаются, при значении метрики 1 последовательности считаются подобными (одинаковыми). Значение метрики 0 не означает, что не существует корреляции между  $S^T(t:N)$  и  $S^T(t_x:N)$ , но последовательности считаются различными для данной местности, так же как и для значения метрики 1 не означают, что значения  $S^T(t:N)$  и  $S^T(t_x:N)$  будут идентичными. Введём пороговое значение метрики (близкое к 1), при достижении которого

считается, что последовательности равны. В нашем исследовании для метрики используется формула:

$$M(S^T(t:N), S^T(t_x:N)) = K^T \sum_{i=0}^N \left( \sum_k \left| s_k^T(t+i) - s_k^T(t_x+i) \right| \right) * w_k + D^T \sum_{i=0}^N \left( \sum_k \left| \frac{s_k^T(t+i)}{dt} - \frac{s_k^T(t_x+i)}{dt} \right| \right) * r_k,$$

где  $K^T$ ,  $D^T$  – масштабирующие коэффициенты для значений параметров и их изменений между последовательными измерениями,  $w_k$ ,  $r_k$  – масштабирующие коэффициенты для параметров определённого типа (температура, давление, влажность и т. д.). Фактически метрика определяет, насколько велика разница временного ряда параметров ансамбля и производных от параметров ансамбля (например, изменение температуры, влажности и т. д.). Важно, что множество коэффициентов метрики уникально для территории  $T$ , но методика их точной настройки – предмет отдельного исследования, на данный момент мы пользуемся упрощёнными методиками. При практическом применении данной формулы важно задание интервала для каждого типа параметров, например, 4 градуса при сравнении значений температуры и 2 градуса для сравнения изменений температуры во времени. Если разница параметров меньше введённого значения, считаем, что значения «соответствуют» друг другу – так как понятие «равенства» здесь не подходит. В зависимости от особенностей местного климата эти значения могут увеличиваться или уменьшаться. Также может устанавливаться минимальное количество «соответствующих» значений (и производных), при которых значение метрики достигает порогового, например «85 из 100» или «90 из 100». Это существенно упрощает подбор параметров в формуле (1) для реализации в программе. Например, в случае сравнение трёхсуточных интервалов с измерениями каждые три часа при шести параметрах в ансамбле производится 144 сравнения параметров с помощью метрики подобия. 90 % (или

85 %) этих сравнений должны давать результат «подобны», чтобы две последовательности были признаны в итоге «подобными».

Для поиска аналогов в локальных измеренных данных мы производим поиск корреляций внутри последовательности ансамбля параметров, полученных с метеостанции за некоторый промежуток времени. Для этого используется полный перебор с рядом ограничений, позволяющих экономить вычислительные ресурсы, например, сравнение идёт в рамках сезонов. В результате для некоторой последовательности  $S^T(t:N)$  получается множество соответствий  $\{S^T(t_i:N)\}$ , при этом при поиске корреляций определяется как  $N$ , так и мощность множества  $\{S^T(t_i:N)\}$  (обозначим далее как  $\|\{S^T(t_i:N)\}\|$ ). В случае, если  $S^T(t)$  принадлежит нескольким  $S^T(t_i:N)$ , то возможно выполнение эвристической процедуры, которая удаляет «лишние»  $S^T(t_i:N)$ , так чтобы выполнялось условие, что для ряда  $S^T(t_i)$  значение  $\|\{S^T(t_i:N)\}\|$  максимально – для упрощения дальнейшего анализа.

Далее можно приступить к определению наиболее вероятного прогноза из найденных в рамках процедуры поиска аналогов. Для каждого множества соответствий  $\{S^T(t_i:N)\}$  выбираем множество прогнозных значений  $\{S^T(t_{i+N}:M)\}$ , непосредственно во времени следующих за соответствиями, обозначим его  $R$  (опустим далее индекс территории  $T$ , подразумевая его). В рамках множества  $R$  тоже проводится процедура поиска аналогов, но в упрощённом виде: сравниваются только ансамбли полностью методом полного перебора, чтобы разделить  $R$  на подмножества, для которых поиск аналогов дал соответствие. Обозначим полученные подмножества как  $r_i$ ,  $R = \{r_i\}$ . Для каждого  $r_i$  вычисляется вероятность его появления  $P_i$ , исходя из мощности подмножества и

мощности  $R$ . Каждому  $r_i$  сопоставляется необходимый ансамбль из глобального прогноза путём простой выборки из базы данных глобальных прогнозов.

Архив глобальных прогнозов для всей поверхности Земли доступен с помощью сети Интернет, желающие могут ознакомиться с возможностями и получить доступ, например, через программу *zyGRIB* [5]. Таким образом, каждому  $r_i$  сопоставляется  $f_i$  – исторический глобальный прогноз для территории  $T$ . Дополнительно как исходные данные используется глобальный прогноз на текущее время для локальной территории (на самом деле для большой территории, включающей в себя и целевую)  $f^*$ , который и необходимо откорректировать в рамках решения задачи о прогнозе локальной погоды.

Для корректировки глобального прогноза и уточнения до локального вводится матричная  $\Delta$ -функция, служащая для определения корректировки глобального прогноза на основе истории глобальных прогнозов. Скорректированный прогноз рассчитывается по формуле:

$$S^T(t_{i+N} : M) = \Delta(R, F, f^*),$$

где функция использует множества значений  $R$  и  $F$  и текущий глобальный прогноз  $f^*$ . Вычислительная реализация  $\Delta$ -функции в базовом случае достаточно проста: например, 1) подсчёт минимальной разности векторов исторического глобального прогноза и текущего глобального прогноза и нахождения наиболее близких исторических прогнозов; 2) нахождение разности наиболее близких исторических прогнозов и измеренного ансамбля параметров из истории наблюдений; 3) применение этой разности для корректировки текущего глобального прогноза для территории  $T$ . Если текущий прогноз будет неточен, то в рамках следующего предсказания погоды в множество  $\{S^T(t_{i+N} : M)\}$  будет добавлен как текущий ансамбль, так и «неудачный» исторический прогноз, что позволит уточнить статистическую выборку в будущем.

Возможно использование способа повышения чувствительности прогноза. Множество  $\{S^T(t_i : N)\}$  содержит в себе последовательность ансамблей длиной  $N$  и анализирует последующие варианты длиной  $M$  (итого  $N+M$ ) ансамблей. Каждый ансамбль длины  $N$  можно рассмотреть как ансамбль длины  $N-m$  и  $m$  значений ансамблей перенести в последующий вариант развития погоды, уменьшив прежнее значение  $M$  на  $m$ . Соответствующим образом изменяется (упрощается)  $\Delta$ -функция, но при этом прогноз вычисляется на меньшее время.

Отдельно остановимся на мониторинге аномалий локальной погоды. Существует как минимум два типа аномалий: 1) аномально большие расхождения с прогнозом погоды; 2) аномальное отклонение величин ряда параметров от многолетних наблюдаемых средних, например, температура выше 35 и ниже 25 градусов, сверхвысокие и сверхнизкие давление, сверхнизкая влажность, сверхвысокие осадки, максимальные порывы ветра более 20 м/с и иные параметры (промерзание грунта, и т. д.). Для практического изучения аномальных отклонений были выбраны такие параметры как максимальная температура и порывы ветра (фактически, буря и ураганный ветер). Минимальный вариант мониторинга аномалий погоды состоит в поиске аномальных погодных состояний в рамках множеств «подобий»  $\{S^T(t_i : N)\}$  и в дальнейшем учёте количества повторяющихся случаев аномальной погоды на выбранное множество подобий. Поскольку проводить прямой поиск аналогий не имеет смысла (события эти точечные – как ураган – или слабоповторяемые – как высокая температура), в рамках методики мы довольствуемся констатацией вероятности повторения аномалии в рамках множества подобий и расчёта вероятности наступления события при развитии местной погоды.

Отметим, что важнейшим свойством сенсорной сети для предупреждения ущерба от аномальных погодных явлений есть возможность получать информацию о текущих погодных аномалиях от развёрнутых соседних сенсорных сетей. В случае,

если сенсорная сеть с некоторой плотностью покрывает десяток районов области, предупреждения о погодных аномалиях могут быть разосланы по абонентам сенсорной сети в течении 10 минут. При этом для мониторинга сенсорная сеть не нуждается в глобальном прогнозе погоды, но может использовать информацию о глобальном прогнозе для повышения готовности системы информирования пользователей.

### Применение метода

В качестве тестового набора данных использовалось несколько наборов данных от 3 месяцев до 7 лет, полученных в центральной части России (специально выбирался континентальный климат). В качестве базового ансамбля был выбран ансамбль параметров (температура воздуха, давление, влажность), максимальный интервал для изучения локальных параметров климата брался 7 лет и более. Это позволяет иметь большое количество сравнений между сезонами различных лет и хорошо оценить вычислительную сложность метода, так как количество сравнений в базовых вариантах поиска аналогов растёт пропорционально  $N^2$ , где  $N$  – длина последовательности ансамблей. Опыт оптимизации программного подобного кода позволяет судить о том, что производительность можно улучшить до значений  $CN$ , где  $C$  – некоторая константа, умеренно увеличивающаяся при росте  $N$ . Американские исследователи [4] использовали 20-летние выборки. Ансамбль содержал 6 параметров – кроме вышеперечисленных, использовались их производные. Метрика подобия сравнивала параметры в небольших пределах (несколько градусов, миллиметров ртутного столба, процентов влажности), ограничение изменения временной производной параметра было как правило 50–70 % от ограничения значения параметра. Метрика считала условие подобия выполненным, если 85–90 % сравнений ансамбля были выполнены успешно. Система была запрограммирована таким образом, что выяснялась максимальная последовательность «подобных» данных, при этом минимальная

длина последовательности была ограничена тремя сутками.

Производилось множество прогнозов анализа данных с различными параметрами с целью нахождения наиболее длительных «подобий» прогнозов для погодной станции. Для части входных данных использовались данные сайта RP5.UA, предоставляющего информацию о погоде в табличном виде. Программа использует для своей работы данные в табличном формате .csv, что позволяет оперативно генерировать таблицы или выгружать их из других приложений.

В этих условиях был проведен анализ данных. Наименьшей длиной подобия были выбраны трое суток, далее анализировались интервалы 3–3.5 суток, 3.5–4 суток, 4–4.5 суток, 4.5–5 суток, 5–5.5 суток, 5.5–6 суток и более 6 суток. С учётом особенностей параметров ансамбля были выбраны два порога для метрик – 86.7 % и 91.7 %. Далее мы рассмотрим один (мотивирующий) пример с одной из станций в России, интервал анализа составляет 7.5 лет.

Результаты анализа подобия данных показаны на следующих рисунках (рис. 1 и 2). По горизонтальной оси отложены длины интервалов подобия в сутках, по вертикальной – количества подобных интервалов. При анализе рассмотрено два случая анализа: в одном случае («обычный» на рисунках) минимальная разница интервала подобия была 3 часа, что позволяет находить «самоподобные» плавные изменения ансамбля в течении длительного интервала, что соответствует длительному плавному изменению погоды, во втором случае («прорежен» на рисунках) минимальная разница интервала подобия была 1 сутки – таким образом сравнение погодных данных начинало делаться не менее чем через сутки. Такое изменение интервала достаточно сильно влияет на изменение характера результатов (это видно при сравнении двух колонок). Увеличение количества «подобий» связано с тем, что при сравнении одной последовательности ансамблей со всеми возможными ансамблями множество сравнений не перекрывается, и при изменении минимально

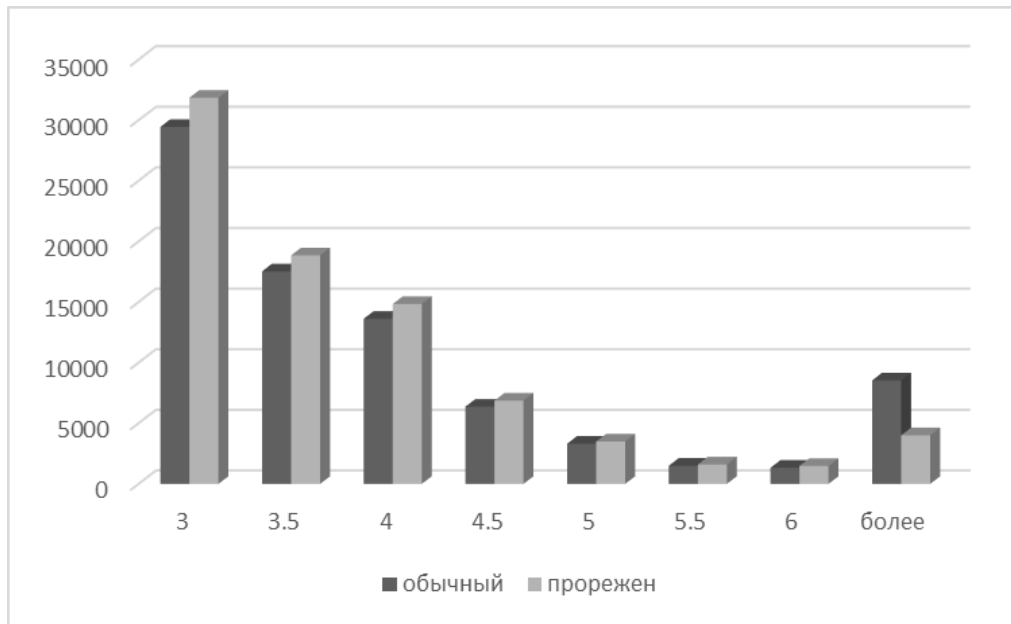


Рис. 1. Результаты анализа для порогового значения метрики 86.7 %

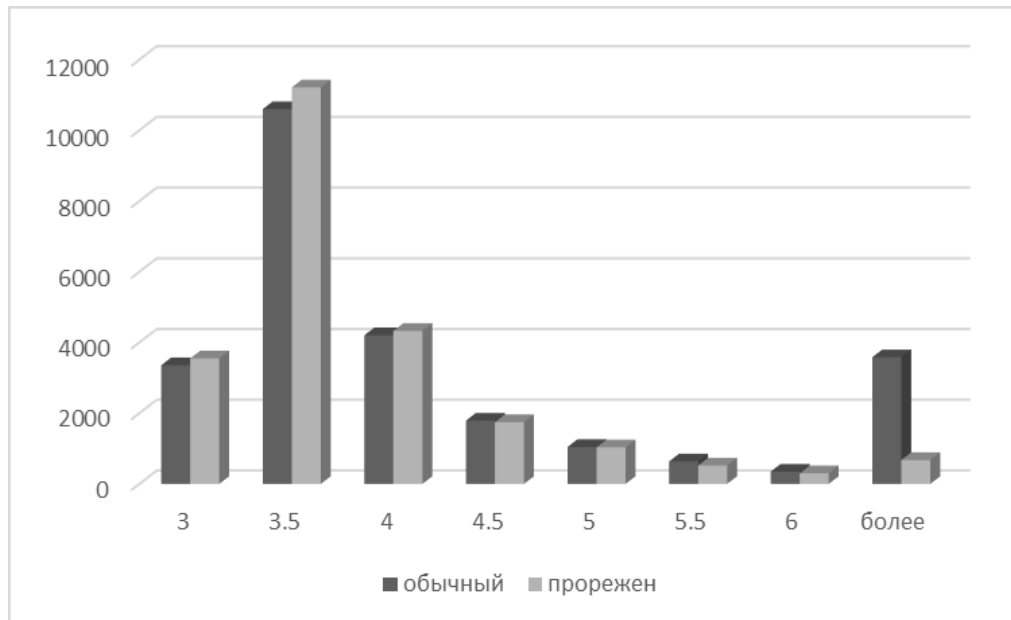


Рис. 2. Результаты анализа для порогового значения метрики 91.7 %

интервала подобия поиск ведётся по разному. Особенности подсчёта для интервала в трое суток связаны с тем, что метрика подобия рассчитывается при сравнении трёхсуточных интервалов, далее анализ пытается увеличить длину интервалов. Длина увеличивается до тех пор, пока метрика превышает порог. При превышении порога фиксируется достигнутая длина.

Графическое отображение анализа корреляций на протяжении всех 7.5 лет затруднительно, поэтому приведём крат-

кие итоги: 1) расстояние по временной оси между подобными последовательностями может быть как 2 дня, так и 6 лет. Это зависит от методов анализа, например «обычный» или «прорежен»; 2) существуют «неотождествимые» последовательности ансамблей вплоть до нескольких суток, это связано с довольно коротким интервалом анализа, с другой стороны многие последовательности имеют полдюжины подобий с разными сдвигами во времени; 3) на рассмотренных временных интервалах количество подобий минимум в пол-

тора раза більше количества ансамблей в 7.5 летней выборке, то есть результаты не являются исключительно стечением обстоятельств или особенностями выборки, что доказано на тестировании метода анализа на многих выборках. Это количество подобий опосредованно позволяет понять минимальную длину погодных данных для станции для осуществления прогноза по подобиям.

Отметим, что небольшое увеличение порогового значения метрики до 91.7 % значительно уменьшает количество подобий, но они есть и все довольно точные, существенная часть подобий по длине достигает 5 суток и более.

На рис. 3 показано распределение мощности множеств подобий – более всего однократных повторений (цифра «1» на горизонтальной оси), но количество многократных повторений в 3–4 раза на временном промежутке в 7.5 лет превышает количество однократных подобий. Длины временных промежутков, на которых не зарегистрировано никаких подобий, могут достигать недели и более, что в целом найденные подобия распределены по оси времени более-менее равномерно. Отсутствие подобия обозначает лишь то, что в этом случае лучше воспользоваться информацией глобального прогноза погоды.

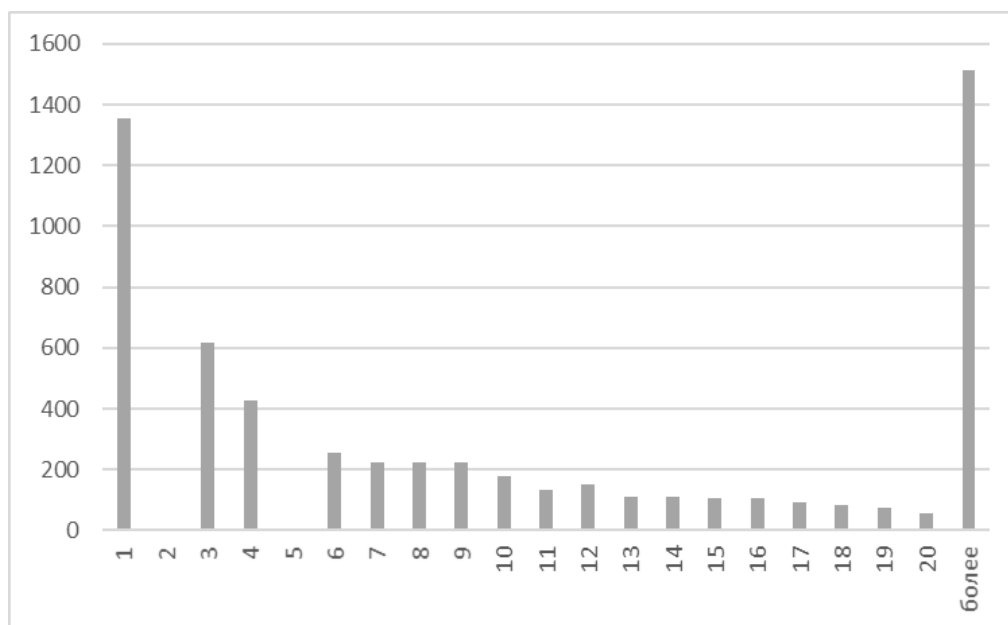


Рис. 3. Распределение количества подобий в группах

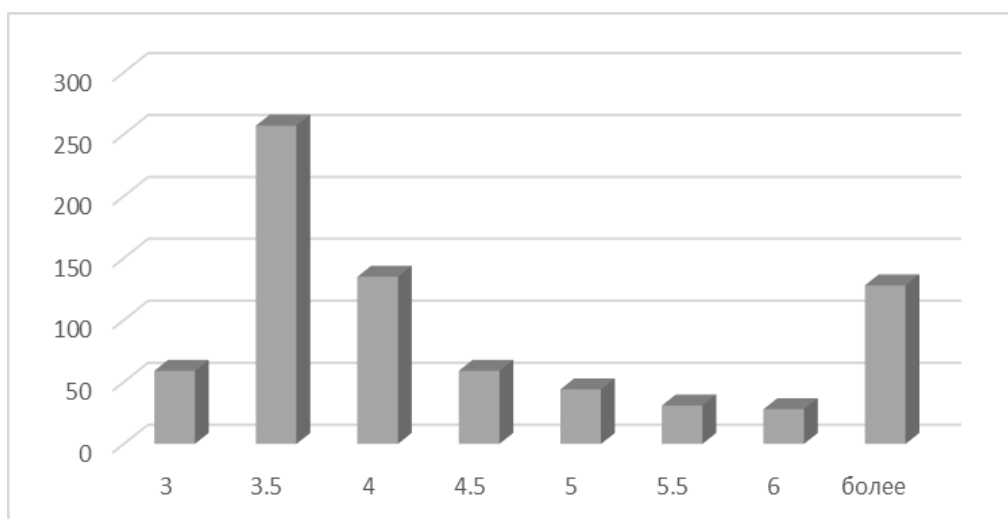


Рис. 4. Количество подобий погоды с учётом регистрации порывов ветра



Для анализа аномальных погодных данных в качестве параметра анализа была взята максимальная скорость ветра за период между снятиями показаний – этот параметр легко снимается на метеостанции с минимальным количеством электроники. Для задач анализа были выбраны подобию, в каждом их которых регистрировались порывы ветра более чем 12 м/с (43 км/час). Распределение подобию в режиме анализа «прорежен» показано на рис. 4.

В результате анализа было установлено, что для анализируемых данных порядка 2 % данных содержат порывы ветра, превышающие выбранный порог скорости. При этом для большого количества «аномальных событий» количество подобию было 2 и более, что позволяет давать прогноз возможности аномального явления – метод может использоваться для предупреждения пользователя о ненулевой вероятности такого явления.

На рис. 5 показано распределение мощности множеств подобию для порывов ветра. Как и следует ожидать, более всего однократных повторений (цифра «1» на горизонтальной оси), но множеств с мощностью более 11 обнаружить не удалось, что согласуется с постановкой задачи, так как аномальные явления довольно редки для 7.5-летней последовательности.

График ясно показывает, что предсказание аномальных явлений предлагаемым методом является именно результатом обработки статистической информа-

ции, а не случайным стечением обстоятельств.

Рассмотрим два примера прогнозирования аномальных ситуаций. На рис. 6 и 7 показаны трёхсуточные последовательности ансамблей (каждый рисунок – одно подобие) для двух разных по времени состояний погоды.

Для рис. 6 и 7 в квадрате показана последовательность ансамблей, с которой необходимо проверять последние измеренные данные (1.5 суток на рис. 6 и 2 суток на рис. 7) для того, чтобы сигнализировать о возможных порывах ветра. (Примечание: если на графике не отображены значения порывов ветра, то они не были зарегистрированы станцией, т.е. значение порыва ненамного отличалось от средней измеренной скорости ветра.)

Анализ данных о локальной погоде не ограничивается представленными методами перебора. Перспективны методы классификации на основе нейросетей, например, на основе многослойного персептрона или сети радиально-базисных функций, и векторных машин. Современные программные средства позволяют с помощью интерактивных построителей оценить работу классификатора, например, на основе сенсорной сети, что является дальнейшими целями исследования. Что касается методов на основе деревьев решений, то эти методы могут лишь собирать статистику по историческим погодным данным. Применение сингулярного

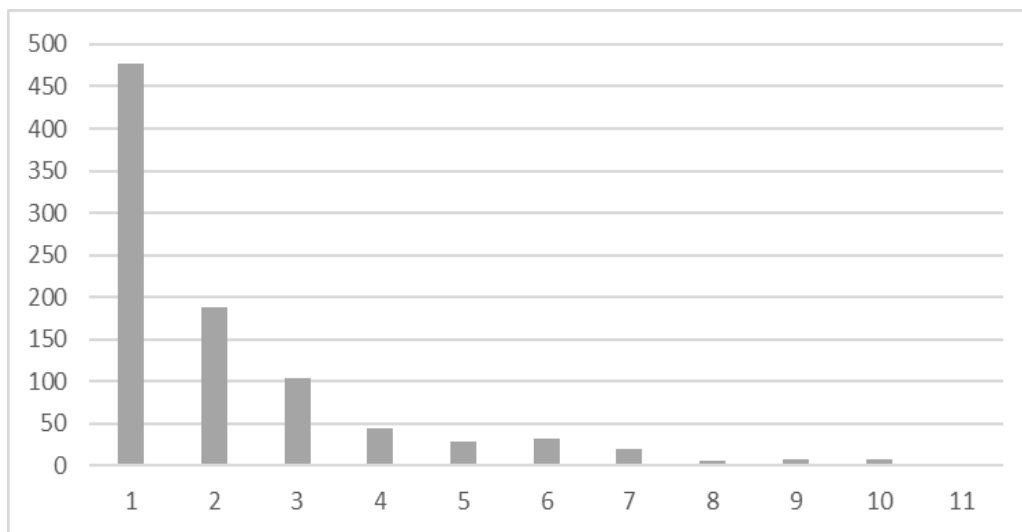


Рис. 5. Распределение количества подобию в группах для аномалию

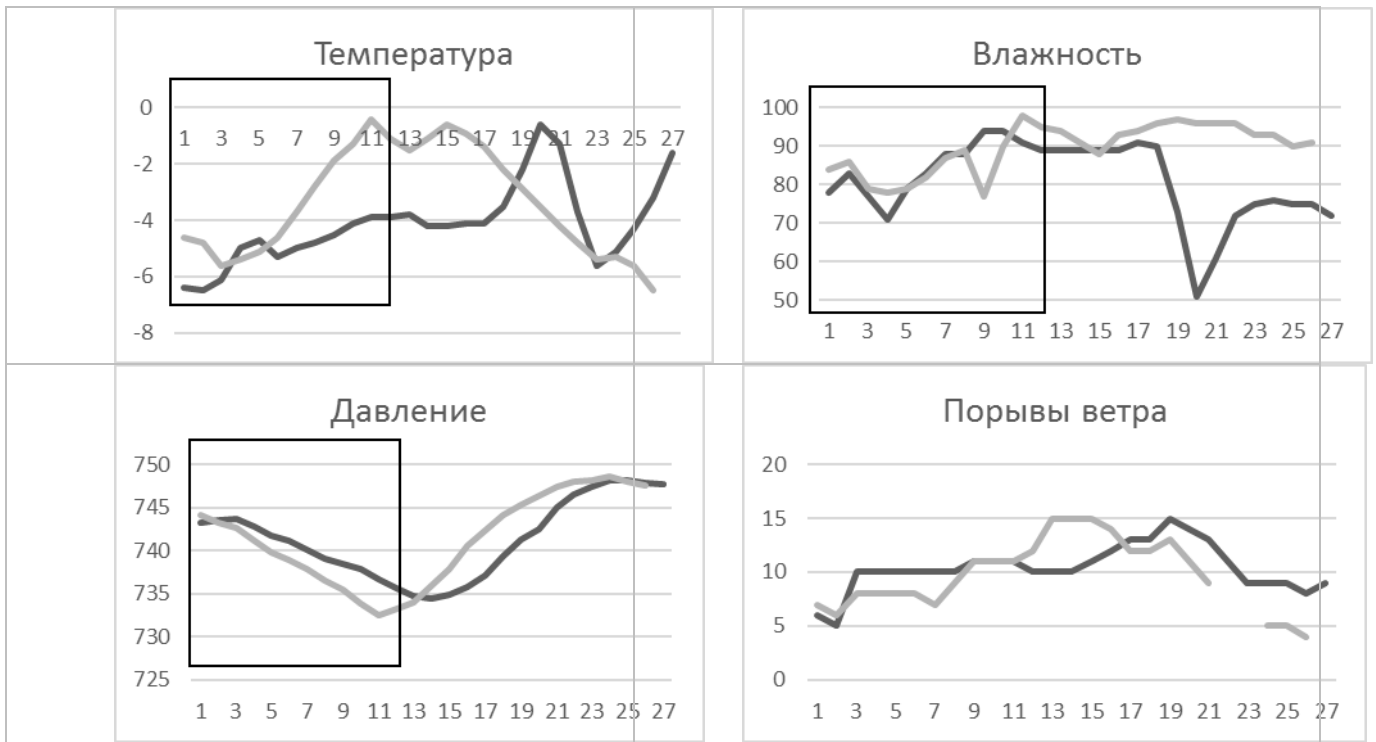


Рис. 6. Последовательность ансамблей 1



Рис. 7. Последовательность ансамблей 2

спектрального анализа (SSA), также известного как «гусеница», позволяет найти долгопериодические (сезонные, месячные) ряды, или же ряды, описывающие изменения погоды типа день-ночь. Иные флукту-

ации погоды метод «гусеницы» относит к неперiodическому шуму, поэтому его использование для анализа погодных данных крайне ограничено.

## Основные результаты работы

Разработана методика анализа временных рядов ансамблей погодных данных, позволяющая находить зависимости в локальных погодных данных для конкретной территории.

Разработана методика применения беспроводных сенсорных сетей для снятия необходимых для анализа погодных данных.

Разработано программное обеспечение для анализа погодных данных, описаны и продемонстрированы результаты анализа.

Описаны и проиллюстрированы результаты анализа аномальных погодных явлений в применении к разработанному методу анализа погодных данных. Анализ может давать предупреждения о вероятности наступления аномального погодного явления – на примере сильных порывов ветра.

1. *Zambretti Forecaster*, 1915. Доступно по ссылке <http://weather.vouthead.gr/wxforecast-zambretti.php?lang=en>
2. *National Geophysical Data Center*. <http://www.ngdc.noaa.gov/>

3. *Thomas M. Hamill and Jeffrey S. Whitaker*, 2006: Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application. *Mon. Wea. Rev.*, **134**, 3209–3229.
4. *H.M. van den Dool*, 1989: A New Look at Weather Forecasting through Analogues. *Mon. Wea. Rev.*, **117**, 2230–2247.
5. Программный пакет zyGRIB. Доступно по ссылке: <http://www.zygrib.org/>

Получено 26.03.2015

### Об авторе:

*Рагозин Дмитрий Васильевич*,  
кандидат технических наук,  
директор ТОВ «МОБИОН».

### Место работы автора:

ТОВ «МОБИОН»,  
603105, РФ,  
г. Нижний Новгород,  
ул. Ижорская 40, кв. 64,  
Тел.: +7(920) 010 1738.  
E-mail: [Dmytro.Rahozin@gmail.com](mailto:Dmytro.Rahozin@gmail.com)