

Исследование статистической значимости результатов обработки метеорологических архивных данных

И. И. Ульшин, e-mail: ulshin@rambler.ru,
А. С. Фатеев

Военный учебно-научный центр Военно-воздушных сил
«Военно-воздушная академия имени профессора Н.Е. Жуковского и
Ю.А. Гагарина» (г. Воронеж)

***Аннотация.** Рассматривается вопрос об определении климатических характеристик различных регионов при наличии в распоряжении исследователя только коротких рядов метеорологических наблюдений. Для решения поставленной в работе задачи использованы положения трендового анализа и теории статистических гипотез. Установлено сходство результатов, полученных на основе обоих подходов, что может дополнительно свидетельствовать об их достоверности.*

***Ключевые слова:** климатические характеристики, ряд наблюдений, тренд, статистическая гипотеза, минимальный объем выборки.*

Введение

В современных условиях все более высокие требования предъявляются к оперативности и обоснованности принимаемых в экономической сфере метеозависимых решений. Эффективное функционирование транспорта, энергетики, сельского хозяйства может быть достигнуто только при широком внедрении современных технологий получения и использования различной информации, в том числе, о погодно-климатических условиях проведения тех или иных мероприятий. Необходимы всестороннее изучение и учет закономерностей формирования погоды и климата, отражающих особенности циркуляции атмосферы, температурного режима и других гидрометеорологических характеристик региона, представление комплекса состояний климатической системы в форме, соответствующей решаемым задачам [1–3].

1. Постановка задачи

Одной из проблем при проведении климатологических исследований является отсутствие достаточного объема исходных данных, которыми являются результаты метеорологических наблюдений в рассматриваемом районе. Характеристики глобальных погодообразующих процессов могут быть рассчитаны и по коротким рядам наблюдений, однако, достоверность полученных результатов не может не вызывать сомнений. В связи с этим целью работы является повышение надежности погодно-климатической информации о регионах, в которых не проводились многолетние метеорологические наблюдения.

Достижение поставленной цели требует решения следующих задач:

исследование изменчивости статистических характеристик атмосферных параметров в зависимости от объема используемого архивного материала;

выбор методов, позволяющих определить минимально необходимое число случаев в выборке для получения статистически значимых климатологических параметров;

определение указанных значений для нескольких географических районов и метеорологических величин и анализ полученных результатов.

Для решения поставленной задачи в работе использовались методы трендового анализа и теории статистических гипотез.

2. Исследование изменчивости статистических характеристик атмосферных параметров в зависимости от объема используемого архивного материала

Одним из значимых климатических характеристик является температурный режим, от которого во многом зависят иные составляющие климатической системы. В работе рассматривались значения температуры воздуха на стандартном изобарическом уровне 1000 гПа на метеорологических станциях, находящихся в городах Брянск, Курск и Воронеж. Для каждого города была организована архивная выборка объемом 1000 случаев по результатам наблюдений в период с 2020 по 2022 г. Из состава этих выборок для каждого из городов последовательно формировались выборки объемом 10, 20, 30... и т.д. значений. По ним проводились расчеты оценок математического ожидания температуры воздуха.

В целях последующего сравнительного анализа были исследованы значения еще одного погодного параметра – атмосферного давления.

Для проведения вычислений использовались электронные таблицы, входящие в состав офисного пакета программ. Пример полученных результатов для пункта Воронеж показан на рис. 1.

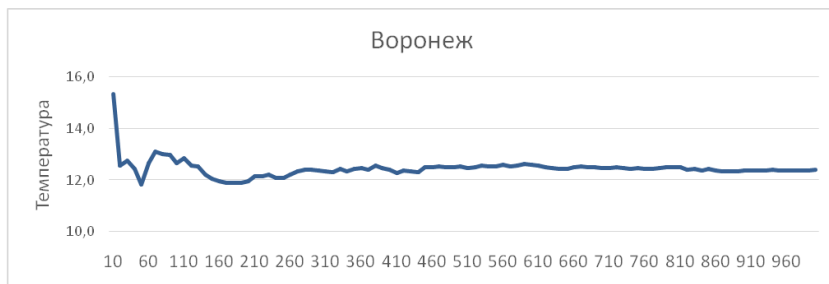


Рис. 1. Значения оценок математического ожидания температуры при расчетах по выборкам различного объема (пункт Воронеж)

Видно, что при малых объемах выборок наблюдаются существенные колебания оценок математического ожидания, Примерно после 200 случаев оценки стабилизируются и с увеличением объема выборки имеют тенденцию к приближению к истинному значению. Точное значение требуемого объема выборки визуально определить практически невозможно.

3. Методы определения минимально необходимого числа случаев в выборке для получения статистически значимых климатологических результатов

Минимально необходимый для получения надежных оценок объем архивного материала определялся методами трендового анализа и теории статистических гипотез.

Для каждого рассматриваемого пункта на соответствующих графиках были построены линии тренда [4]. Поскольку графики имеют ярко выраженный нелинейный вид, использовались полиномиальные функции:

$$y = a_0 + \sum_{i=1}^k a_i x_i, \quad (1)$$

где a_i – коэффициенты полинома; k – степень полинома (в работе использовались полиномы шестой степени).

Приближение линии тренда к истинному значению статистического параметра и достижение разности между ними,

соответствующей точности измерений рассматриваемой метеорологической величины, очевидно, свидетельствует о достаточности архивного материала. Подобная точка на рис. 2 показана вертикальным отрезком. Представленная на графике горизонтальная пунктирная линия, проведена через значение температуры на оси ординат, соответствующее математическому ожиданию, принимаемому за истинное значение, определенному по всей выборке, рассматриваемой в качестве генеральной совокупности.

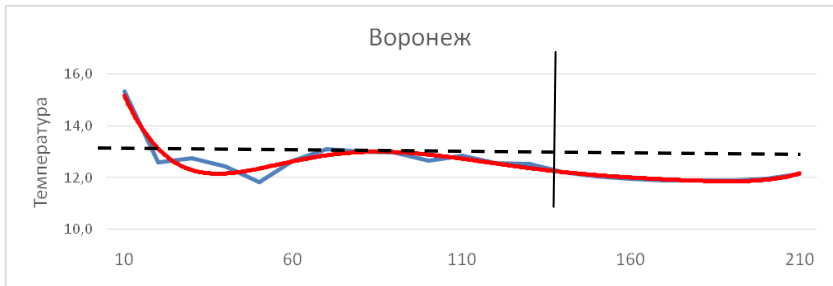


Рис. 2. Линия тренда на графике изменения оценок математического ожидания (пункт Воронеж)

Аналогичные действия были проведены на основе данных об атмосферном давлении.

На следующем этапе работы с использованием методов теории статистических гипотез было проведено сравнение оценок математического ожидания, рассчитанных по всей выборке и по выборке ограниченного объема.

Если считать истинное значение дисперсии неизвестным, аналитически данную задачу можно описать следующим образом. На основе выборки значений метеорологической величины X требуется при выбранном уровне значимости α проверить нулевую гипотезу о равенстве математических ожиданий, рассчитанных по ограниченной и по полной выборкам $H_0 : M_{огр}(X) = M_{полн}(X)$ при конкурирующей гипотезе $H_1 : M_{огр}(X) \neq M_{полн}(X)$.

Для решения этой задачи вводится новая случайная величина δ — разность между рассчитанными по ограниченной и полной выборкам значениями математического ожидания, а также ее среднее значение.

Для проверки гипотезы вычисляется наблюдаемое (фактическое) значение критерия согласия статистической гипотезы [5, 6]:

$$T^* = \frac{\sqrt{n}}{\sigma_\delta}, \quad (2)$$

где n – объем выборки, а σ_δ определяется по формуле (3):

$$\sigma_\delta = \sqrt{\frac{\sum_{i=1}^n \delta_i^2 - \frac{1}{n} (\sum_{i=1}^n \delta_i)^2}{n-1}}. \quad (3)$$

Далее по таблице критических точек распределения Стьюдента [6] по заданному уровню значимости и числу степеней свободы $g = n - 1$ определяется критическая точка (значение). Если рассчитанное T^* превышает критическое значение, нулевая гипотеза принимается, если не превышает – нулевая гипотеза отвергается.

Результаты проверок статистических гипотез были сведены в таблицу, фрагменты которой представлены ниже.

Таблица 3

значения критерия согласия при проверке статистических гипотез

Объем выборки	Воронеж		Курск		Брянск	
	Темпера тура	Давление	Температура	Давление	Температура	Давление
100	-9,91	-2,05	-3,51	-2,32	-17,28	-2,05
110	-10,41	-2,20	-11,90	-2,54	-22,16	-2,08
120	-9,34	-1,89	-3,21	-2,24	-19,81	-1,82
130	-8,84	-1,61	-1,13	-2,23	-20,55	-1,60
140	12,35	-1,39	14,55	-1,79	-14,39	-1,73
150	17,82	-1,26	19,29	-1,67	8,45	-1,87
160	21,63	-1,12	23,92	-1,71	15,45	-1,59
...
420	17,00	1,10	15,41	-1,49	-23,13	-3,35
430	21,70	2,93	18,96	2,35	-19,50	-1,65
440	24,84	3,77	26,09	3,48	-14,54	2,37

Значения критерия согласия, при которых гипотеза принимается, выделены полужирным шрифтом.

4. Анализ полученных результатов

Полученные результаты позволили провести сравнение и сделать следующие выводы.

Характер изменений оценок математического ожидания является аналогичным для обеих метеорологических величин. Вместе с тем, Атмосферному давлению свойственны большие колебания значений оценок математического ожидания на начальной части графика. Минимально необходимый объем выборки при этом увеличивается.

В результате реализации обоих методов установлено, что для получения статистически значимых результатов обработки наблюдений за температурой воздуха необходимо располагать выборкой объемом не менее 170 случаев, за атмосферным давлением – не менее 430 случаев.

При проведении аналогичных исследований необходимо учитывать, что требования к архивным материалам существенно зависят от изменчивости рассматриваемых метеорологических величин. Универсальных рекомендаций, по-видимому, не существует.

Заключение

Анализ полученных в работе результатов позволяет сделать следующие выводы.

1. Для эффективного функционирования экономической организации государства необходимо широкое внедрение современных технологий получения и использования различной информации, в том числе, о погодно-климатических условиях проведения тех или иных мероприятий. Особенно актуальным это является с учетом происходящих климатических изменений, о которых свидетельствуют смещение «реперных» климатических дат, рост частоты и интенсивности погодных аномалий и экстремальных погодных условий.

2. Одной из основных сложностей, возникающих при изучении климатических процессов и перемен, является проблема коротких рядов метеорологических наблюдений и, следовательно, недостаточной надежности результатов их статистической обработки.

3. Как показали исследования, основанные на трендовом анализе и направленные на определение минимально необходимого объема исходных данных, для получения статистически значимых оценок математического ожидания температуры воздуха достаточно 170 случаев. Атмосферному давлению свойственны большие колебания оценок при малом числе наблюдений, и минимально необходимый объем выборки при этом увеличивается до 400 случаев.

4. На основе теории статистических гипотез получены схожие результаты. Для оценки математического ожидания температуры

воздуха необходим объем выборки 150 случаев. Для получения надежных оценок математического ожидания атмосферного давления необходима выборка объемом 430 наблюдений.

Результаты, полученные в работе, могут служить основой для решения задач использования климатической информации в системах поддержки принятия решений по обеспечению гидрометеорологической безопасности.

Список литературы

1. Климатическая доктрина Российской Федерации / Утв. распоряжением Президента РФ от 17 декабря 2009 г. № 861-рп. – М., 2009. – 20 с.

2. Катцов В.М. Современные приоритеты фундаментальных исследований климата / В.М. Катцов, В.П. Мелешко // Труды ГГО, вып. 557. 2008. С. 3-19.

3. Израэль Ю.А. Изменения глобального климата. Роль антропогенных воздействий / Г.В. Груза, В.М. Катцов, В.П. Мелешко, Ю.А. Израэль // Метеорология и гидрология. – № 5. – 2001. – С. 5-21.

4. Трендовый анализ: цели, сущность, основные приемы [Электронный ресурс]. – Режим доступа: https://studopedia.ru/5_152750_vopros--trendoviy-analiz-tseli-sushchnost-osnovnie-priemi.html.

5. Ульшин И.И. Методы статистической обработки гидрометеорологической информации / И.И. Ульшин – Воронеж: ВУНЦ ВВС «ВВА». 2016. – 187 с.

6. Гмурман В.Е. Теория вероятностей и математическая статистика: Учеб. пособие для бакалавров / В.Е. Гмурман – М.: Юрайт, 2013. – 489 с.