

ОПРЕДЕЛЕНИЕ ОСТРОВА ТЕПЛА Г. МИНСКА ПО ИНСТРУМЕНТАЛЬНЫМ ДАННЫМ С ПОМОЩЬЮ ГИС-МЕТОДОВ И СТАТИСТИКИ

Е. А. Ярош¹⁾, Т. В. Шлендер^{1,2)}, П. А. Силков²⁾

¹⁾ Факультет географии и геоинформатики БГУ

*²⁾ Национальный научно-исследовательский центр мониторинга озоносферы БГУ,
г. Минск, Республика Беларусь, email: zenua2004.6@gmail.com*

Проведена верификация массива данных различных датчиков г. Минска по температуре воздуха с помощью ГИС методов и методов статистики. Особенности распространения городского острова тепла представлены посредством анализа карто-схемы пространственного распределения медианных значений нормированных температур, построенной при помощи статистических данных, ГИС платформы QGIS и языка программирования Python.

Ключевые слова: остров тепла; AirMQ; АМС; rp5; метод IQR; распределение нормированных температур.

DETERMINATION OF THE HEAT ISLAND OF MINSK BY INSTRUMENTAL DATA USING GIS METHODS AND STATISTICS

E. A. Yarosh¹⁾, T. V. Shlender^{1,2)}, P. A. Silkov²⁾

¹⁾ Faculty of Geography and Geoinformatics BSU

*²⁾ National Research Center for Monitoring of the Ozonosphere BSU,
Minsk, Republic of Belarus, email: zenua2004.6@gmail.com*

Verification of the air temperature data set using GIS methods and statistical methods has been carried out. The peculiarities of urban heat island distribution are presented by analyzing the map scheme of spatial distribution of median values of normalized temperatures, constructed with the help of statistical data, GIS platform QGIS and Python programming language.

Keywords: heat island; AirMQ; AMS; rp5; IQR method; distribution of normalized temperatures; land surface temperature.

В настоящее время в связи с осознанием антропогенного влияния на окружающую среду и урбанизацией, ростом городов, всё более важной становится проблема городского острова тепла (ГОТ) [1]. Основными причинами ГОТ являются огромное количество тепла, вырабатываемого городскими сооружениями, транспортом, промышленными предприятиями. Крупные города и представители разнообразной экономической деятельности сталкиваются с еще более серьезными трудностями, вследствие

чего возникла необходимость изучения данного явления для дальнейшего его контроля и предотвращения негативного влияния на здоровье и жизнедеятельность людей в целом [2,3].

В настоящей работе для получения информации о температуре воздуха использовались данные трех источников информации: автоматические метеостанции (АМС) Белгидромета (9 точек), данные о температуре воздуха с погодного сайта *gr5* (23 точки) [4], данные датчиков общественного проекта *AirMQ* (63 точки) [5], что в совокупности даёт 95 пунктов измерений в г. Минске.

С целью обработки полученных массивов данных температуры воздуха, использовалось фильтрование наблюдений при помощи проверки статистики: исключались значения, которые значительно отличаются от средней или медианной температуры. Кроме того, проводилась фильтрация на основании проверки согласованности данных: исключались любые наблюдения, значения которых не были согласованы с физическими процессами, лежащими в основе изменений температуры. Таким образом, для верификации массива данных по температуре применялись следующие методы:

- проверка статистики с ограничениями по верхней и нижней границам массива данных;
- проверка статистики распределения температур воздуха на нормальность;
- идентификация выбросов методом IQR (межквартильный диапазон).

Поскольку данные АМС Белгидромета предположительно имеют наилучшую точность и надёжность измерений температуры, то они были использованы в качестве референсных измерений для валидации и фильтрования данных датчиков *gr5* [4] и *AirMQ* [5]. На рисунке 1 представлена картосхема расположения датчиков, которые были использованы в обработке данных.

Предварительная обработка рядов измерений различных датчиков состояла в реализации пространственной и временной интерполяции. Для анализа взаимосвязи метаданных и данных по температурам, матрица температур связывалась с таблицей с метаданными. В качестве главного инструмента связывания всех данных вместе и приведения их к одной временной шкале с временным шагом один час использовался язык программирования Python.

Для фильтрования данных датчиков *AirMQ* [4] и *gr5* [5] использовались два варианта проверки:

- проверка на наличие данных на выбранную дату: данных по температуре воздуха должно быть не менее 97 % (23 часа из 24);
- проверка MAE (англ. mean absolute error - средняя абсолютная ошибка).

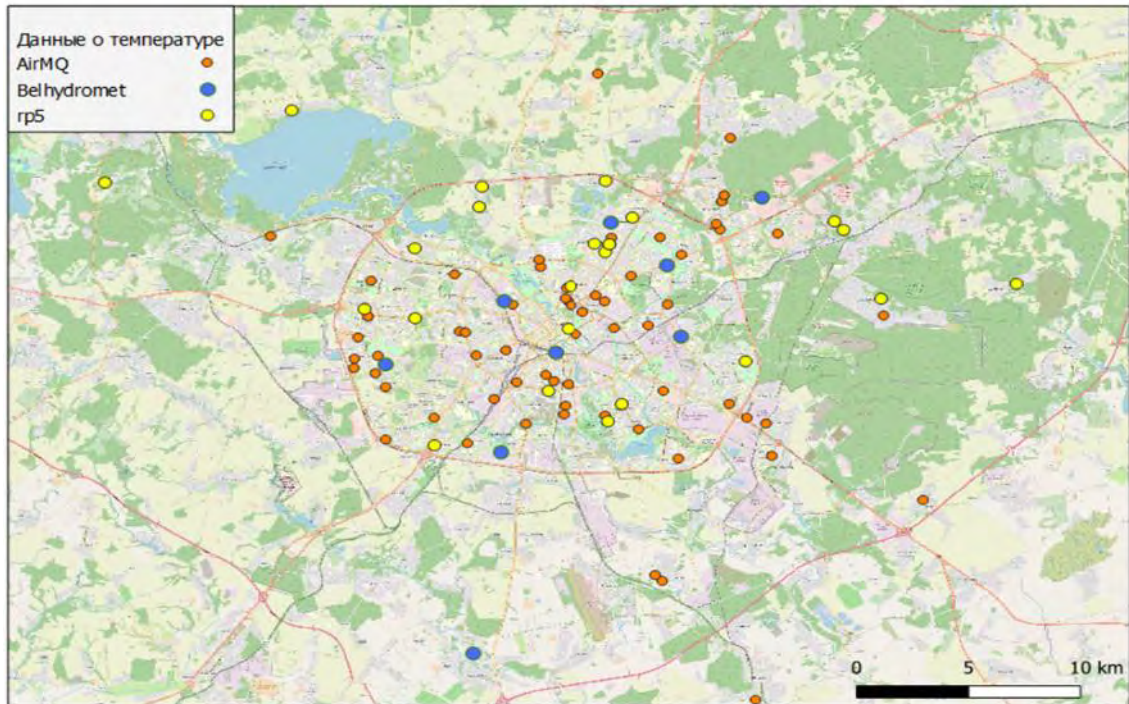


Рис. 1. Картограмма локализации используемых метеодатчиков для г. Минска

МАЕ показывает на сколько градусов в среднем отличаются одно измерение проверяемого датчика относительно одного значения эталонного ряда.

На рисунке 2 представлены измерения температуры выборки группы датчиков за 05.04.2021, откуда можно заметить, что для выбранного дня проверку прошли суммарно только 4 датчика AirMQ и rp5.

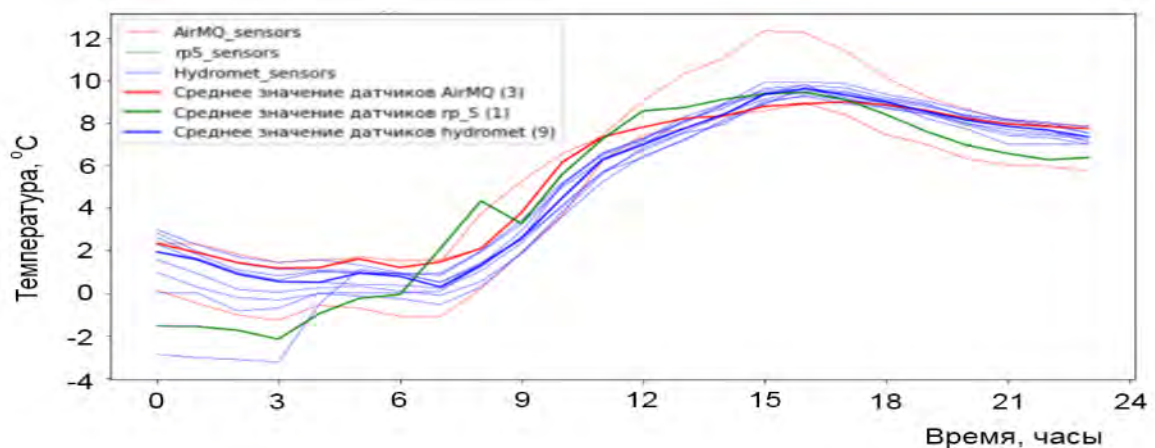


Рис. 2. Временной ряд температуры воздуха приземного слоя г. Минска по отфильтрованной выборке датчиков для 05.04.2021

Для демонстрации того, как именно применение статистических фильтров к датчикам AirMQ и gp5 меняет медианные значения MAE на масштабе времени один день, были отфильтрованы данные датчиков за период 01.03.2021 по 15.05.2021.

Такой диапазон был выбран в целях удобства визуализации изменений в рядах наблюдений, так как при выборе более широкого диапазона сложнее проводить анализ графической информации. Для анализа применялись медианные значения температуры, а не среднеарифметические, поскольку оказалось, что наличие шума в ряду данных значительно влияет на оценки математического ожидания, однако практически не влияет на медианные значения MAE, что показано на рисунке 3.



Рис. 3. Медианные значения MAE для данных о температуре воздуха приземного слоя г. Минска за период с 01.03.2021 по 15.05.2021:
а) с применением фильтра по MAE; б) без применения фильтра по MAE;
в) MAE для АМС Белгидромета

На рисунке 3 жёлтым цветом отображено медианное значение MAE датчиков AirMQ и gp5 до фильтрации для каждого дня, синим цветом - медианное значение MAE датчиков AirMQ и gp5 после фильтрации, зелёным - медианное значение MAE датчиков станций Белгидромета. Эти результаты показывают, что проведённая фильтрация приблизила среднюю абсолютную ошибку AirMQ и gp5 к средней абсолютной ошибке станций Белгидромета. Это говорит о том, что после фильтрации по признаку MAE остались измерения, приближенные к измерениям Белгидромета. Изменения статистических показателей датчиков после применения алгоритмов фильтрации за период 01.03.2021 – 15.05.2021 представлены в таблице.

По результатам, представленным в таблице, можно заметить, что медианное значение MAE для датчиков AirMQ и gp5 до применения процедуры фильтрации превышает медианные значения MAE датчиков станций Белгидромета более чем в 10 раз.

Таблица 1

Изменение метрик датчиков после фильтрации

Метрика	Значение
Медианное количество активных датчиков gr5 и AirMQ за день	38,0
Медианное количество датчиков gr5 и AirMQ после фильтрации за день	8,25
Медианный процент датчиков gr5 и AirMQ, не прошедших фильтрацию, %	80,7
Медианное MAE датчиков AirMQ и gr5 до фильтрации	3,88
Медианное MAE датчиков AirMQ и gr5 после фильтрации	0,86
Медианное MAE датчиков Белгидромета	0,36
Медианная дисперсия датчиков AirMQ и gr5 до фильтрации	6,22
Медианная дисперсия датчиков AirMQ и gr5 после фильтрации	4,36
Медианная дисперсия датчиков Белгидромета	3,97

Сравнивая статистические параметры соответствующих рядов температур воздуха датчиков до и после применения процедуры фильтрации, можно отметить разницу в медианных значениях MAE в 4.4 раза. Медианные изменения дисперсии оказались не столь значительны: в 1.6 раз между датчиками до фильтрации и станциями Белгидромета и в 1.4 раза между датчиками до фильтрации и датчиками после фильтрации.

Влияние процедуры фильтрации датчиков AirMQ и gr5 на дисперсию анализируемых данных показано на рисунке 4, где жёлтым цветом отображена медианная дисперсия датчиков AirMQ и gr5 до фильтрации для каждого дня, синим цветом – медианная дисперсия датчиков AirMQ и gr5 после фильтрации, зелёным – медианная дисперсия станций Белгидромета.

Информация о распределении тепла в г. Минске собиралась на основе средних значений датчиков за период в один год. Для анализа выбирался период с 01.02.2021 по 01.02.2022, так как именно в нём были представлены датчики всех имеющихся измерительных сетей: AirMQ, gr5 и Белгидромета.



Рис. 4. Медианные значения дисперсии у датчиков после фильтрации, датчиков до фильтрации и датчиков станций Белгидромета за период с 01.03.2021 по 15.05.2021

Для получения среднего распределения температурного поля по данным отфильтрованных датчиков за один год, проводилась нормировка данных с использованием стандартных отклонений по формуле:

$$x_{i, \text{станд}} = \frac{x_i - \mu}{\sigma},$$

где $x_{i, \text{станд}}$ – нормированное значение температуры; x_i – исходное значение температуры; μ – среднее значение температуры выборки всех датчиков за конкретный момент времени; σ – стандартное отклонение выборки всех датчиков за конкретный момент времени.

В матрице данных всех доступных для анализа датчиков за один год содержится 8760 строк, представляющие собой все моменты измерений с дискретностью раз в 10 минут, каждая из которых образует выборку отфильтрованных рабочих датчиков в конкретный момент времени. Для каждой отфильтрованной выборки рассчитывались параметры измерений μ и σ , на базе которых с использованием формулы вычислялись нормированные значения температуры $x_{i, \text{станд}}$, что позволило, не привязываясь к абсолютным значениям температур воздуха, анализировать статистические параметры распределения нормированных значений температур воздуха для различных сезонов и времени суток.

В результате расчета среднегодовых нормированных значений температур отфильтрованных датчиков, из изначальной выборки в 95 датчиков была извлечена выборка из 62 датчиков. В итоговой выборке, которая использовалась для проведения дальнейшего анализа, вышло: 35 датчиков сети AirMQ (28 датчиков ни разу не прошли проверку в течении года), 18 датчиков сети gr5 (5 датчиков ни разу не прошли проверку в течении года) и 9 станций сети Белгидромета.

Таким образом, предложенный метод фильтрации разнообразных температурных данных, основанный на оценке средней абсолютной ошибки MAE и дисперсии относительно референсного набора данных, позволил выделить из массива доступных датчиков те, которые предположительно характеризуются наличием полезной информации, поскольку заданные рамки отклонений для регистрируемых значений позволяют надеяться на принадлежность референсных и выбранных датчиков к одной и той же генеральной совокупности.

Нормированные температуры воздуха отображают то, насколько измерения датчика в рассматриваемый момент отличаются от среднеарифметического значения всех датчиков.

Картосхема среднегодовых значений отклонений стандартизированных температур воздуха от среднего для каждого датчика за 2021 год представлена на рисунке 5.

Стоит отметить, что картосхема нормированных температур, представленная на рисунке 5, представляют собой карту безразмерных величин, которые характеризуют степень отклонения измеряемой температуры воздуха от ее математического ожидания в исследуемом временном ряду, что позволяет учитывать изменения температуры в связи с влиянием различных времен суток и сезонов года на городской остров тепла.

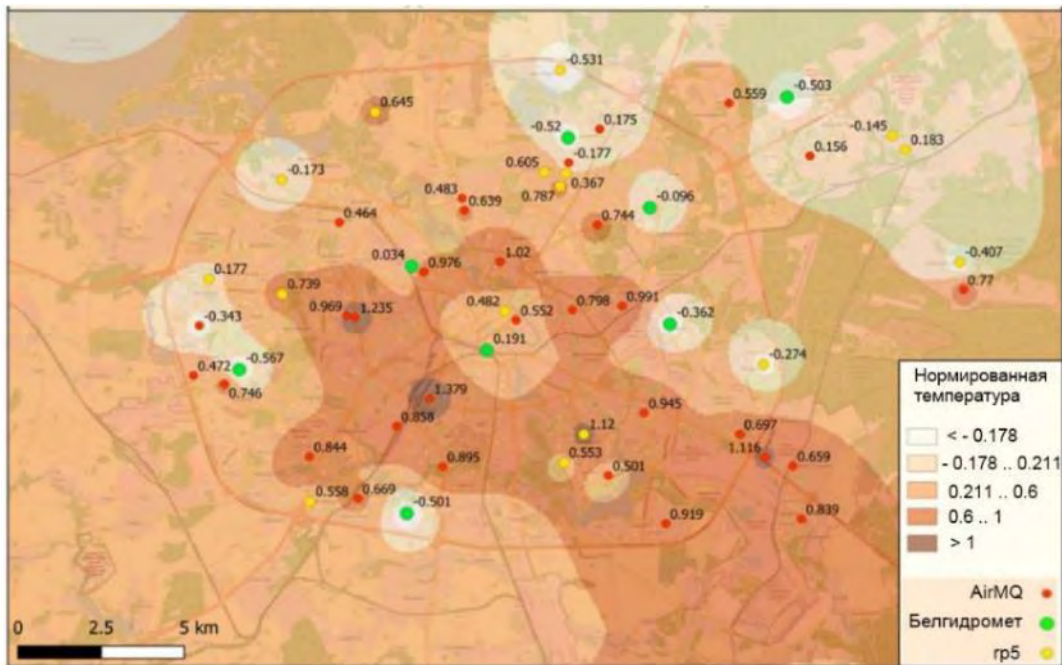


Рис. 5. Пространственное распределение медианных значений нормированных температур воздуха за период 01.02.2021 – 01.02.2022

Таким образом, из данных рисунка 5 можно заметить расположение городского острова тепла, локализованного в южной части города Минска по данным годовых наблюдений совокупностью в 62 датчика. При этом, отсутствие необходимого количества датчиков за пределами города не позволяет точно оценить границы распределения температурных отклонений на удалении более 5 км от внешних границ города в его южной и юго-западной частях. Наблюдаемое распределение нормированных температур может быть объяснено наличием различных локальных климатических зон внутри города, в рамках которых производилась регистрация температурных рядов, что напрямую связано с феноменом городского острова тепла.

Библиографические ссылки

1. *Oke T. R.* City size and the urban heat island // *Atmos. Environ.* 1973. № 7. P. 769–779.
2. *Basarin B., Lukić T., Matzarakis A.* Review of Biometeorology of Heatwaves and Warm Extremes in Europe // *Atmosphere.* 2020. Vol. 11. № 1276.
3. Конференция IOP 2023 г. Сер. Earth Environ / Х. Дж. Хаджер [и др.]
4. Архив погоды городских метеостанций г. Минска, [Электронный ресурс]. URL: <https://tp5.by/> (дата обращения: 25.01.2023).
5. Архив портативных датчиков проекта AirMQ г. Минска, [Электронный ресурс]. URL: <https://airmq.by/> (дата обращения: 09.01.2023).