

Сравнительный анализ методов машинного обучения для прогнозирования суточного хода температуры воздуха степной зоны Северной Евразии

Васильев Д.Ю.^{1,2}, Семенов В.А.^{3,4}, Христодуло О.И.¹, Ковальчук В.А.¹

¹ Уфимский университет науки и технологий, Уфа

² Институт степи ОФИЦ РАН, Оренбург

³ Институт физики атмосферы им. А.М. Обухова РАН, Москва

⁴ Институт географии РАН, Москва



Международная ежегодная конференция «Современные проблемы дистанционного зондирования Земли из космоса - 2025»

АКТУАЛЬНОСТЬ

За последние десятилетия в гидрометеорологической практике стали интенсивно внедрять методы машинного обучения [1-3], которые могут эффективно использоваться вместе с традиционными методами статистического анализа и численных методов прогноза погоды и изменений климата [4-7]. Необходимо отметить, что применение методов машинного обучения обусловлено также постоянно увеличивающимся объемом баз данных гидрометеорологических наблюдений станционных сетей, реанализов, а также дистанционных измерений. Несмотря на большое количество опубликованных работ и многолетний опыт использования методов машинного обучения в гидрометеорологической практике, не существует строго определенных моделей искусственного интеллекта для прогноза погоды. Цель работы – прогноз и пространственно-временной анализ данных приземной температуры воздуха в степной зоне Северной Евразии.

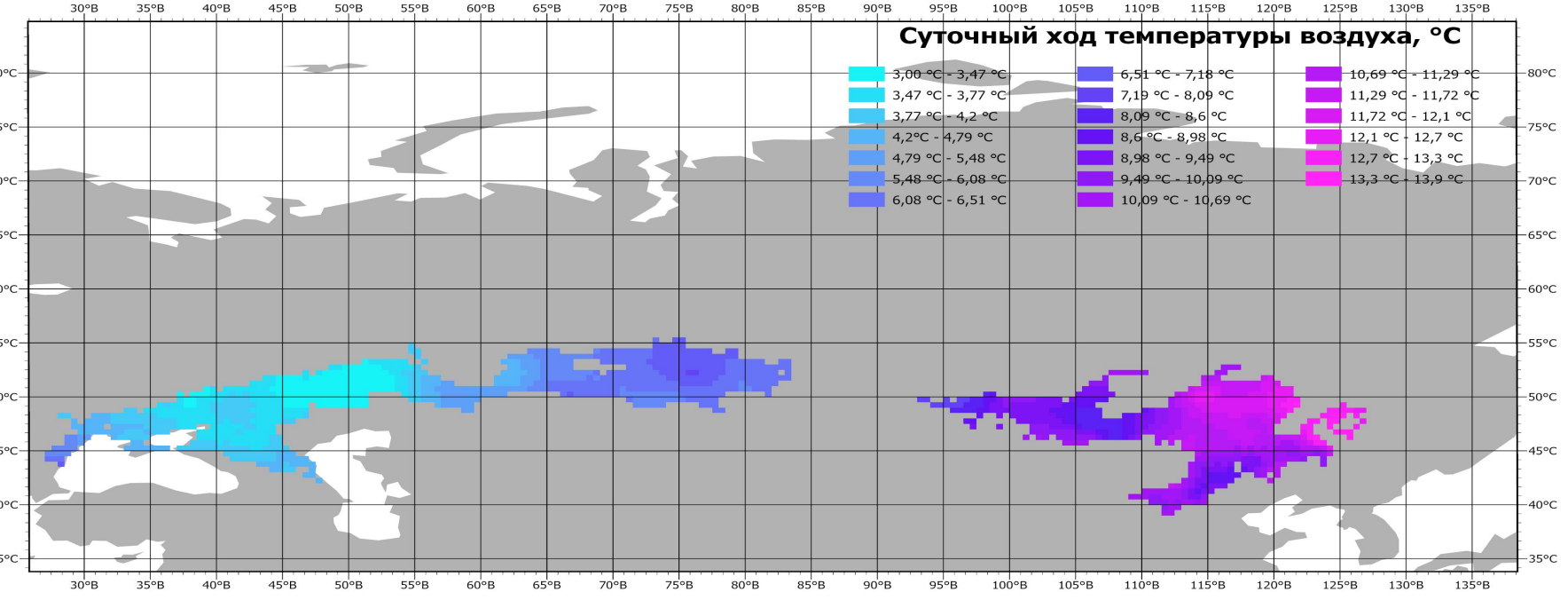


Рис. 1. Суточный ход температуры воздуха степной зоны Северной Евразии зимнего сезона

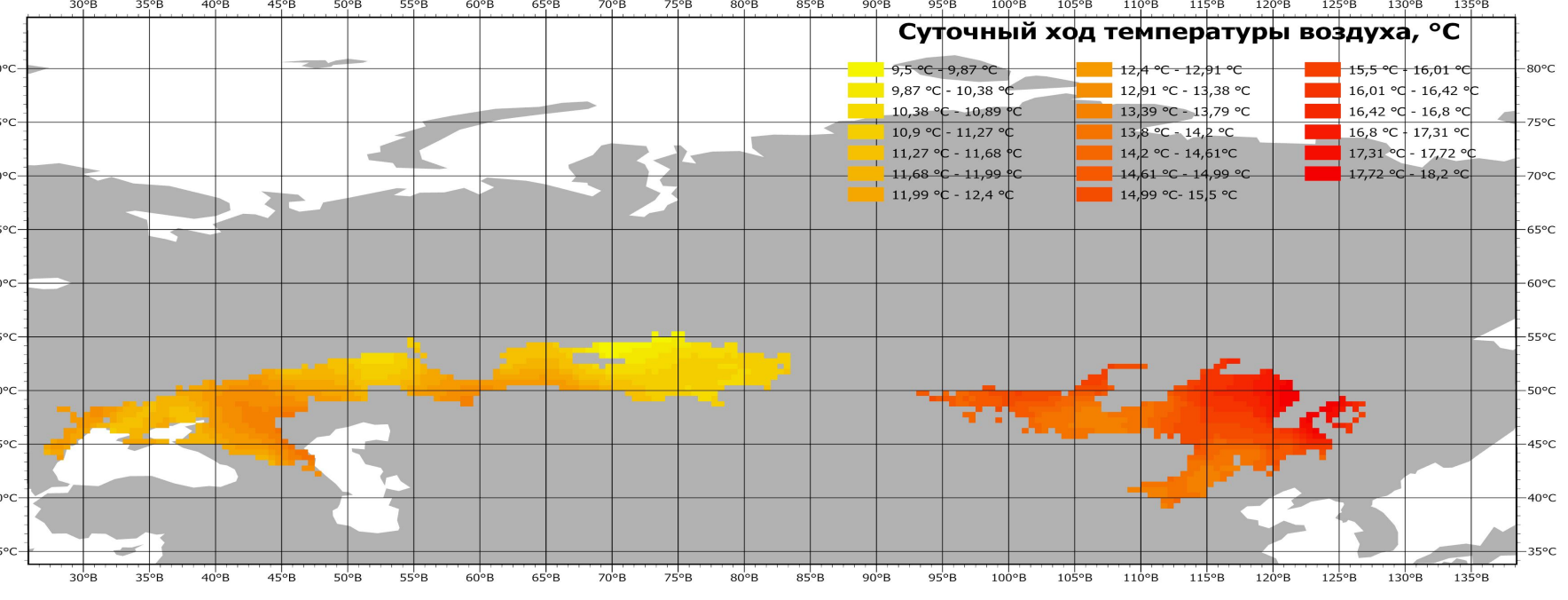


Рис. 2. Суточный ход температуры воздуха степной зоны Северной Евразии летнего сезона

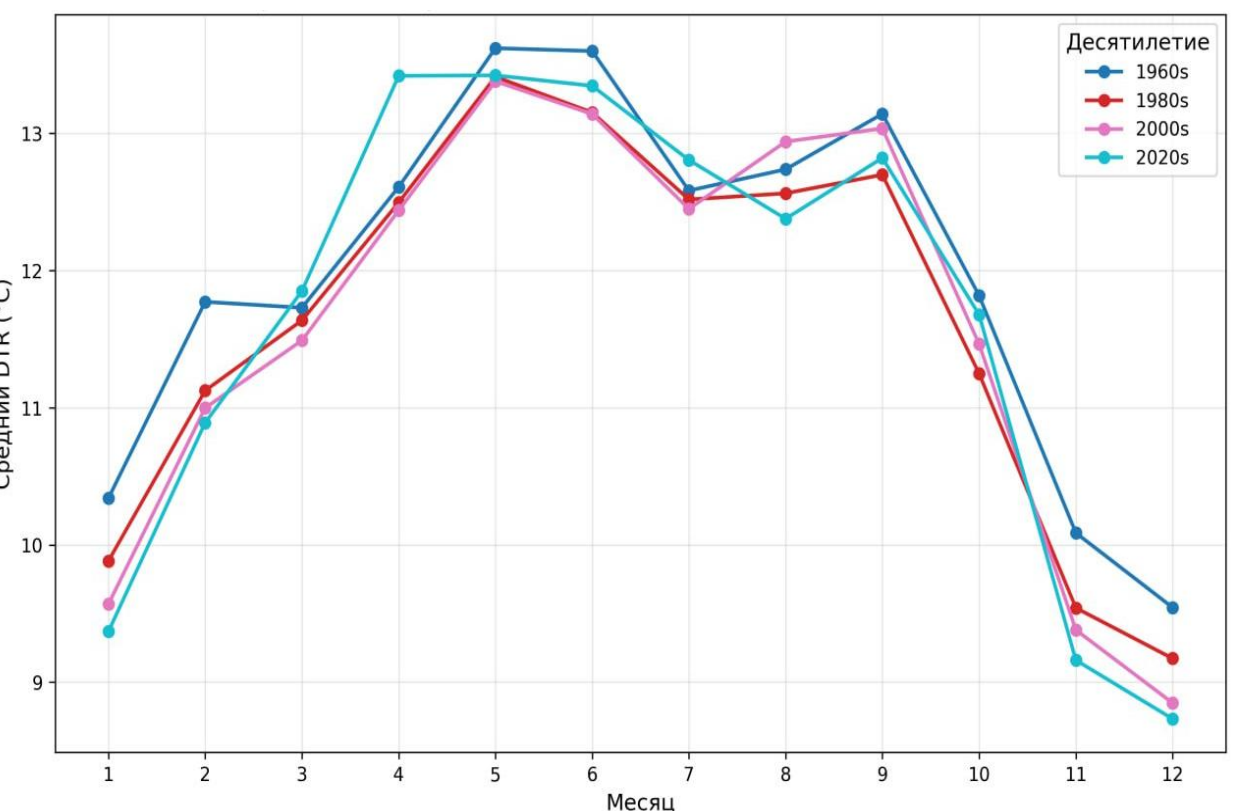


Рис. 3. Внутригодовое распределение суточного хода температуры воздуха степной зоны Северной Евразии по отдельным десятилетиям

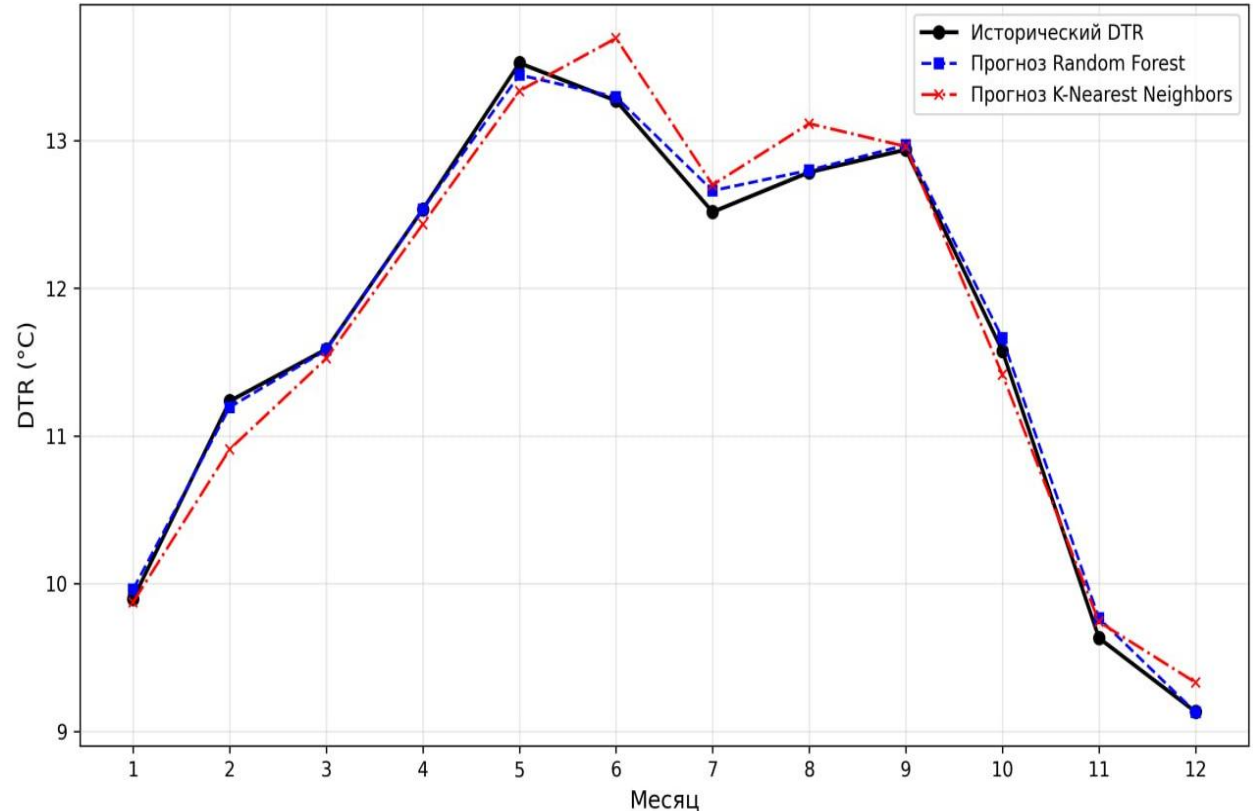


Рис. 4. Сравнительный анализ фактического суточного хода температуры воздуха степной зоны Северной Евразии и прогнозного с использованием методов машинного обучения

ИСПОЛЗУЕМЫЕ ДАННЫЕ, ТЕРРИТОРИЯ

В работе были использованы архивные данные суточного хода приземной температуры воздуха (Climatic Research Unit–CRU), находящиеся в свободном доступе (<https://archive.ceda.ac.uk>), за период с 1960 по 2023 гг. На аналогичном временном интервале был произведен корреляционный анализ данных архива CRU с фактическими данными наблюдательной сети Росгидромет (<http://meteo.ru/>). Для суточного, декадного и месячного прогноза были использованы методы k-ближайших соседей (k-nearest neighbors) и случайного леса (random forests).

Степная зона Северной Евразии — это большой по площади регион, простирающийся от Восточной Европы до Тихого океана, характеризующийся преимущественно безлесными пространствами с травянистой растительностью. Климат здесь резко континентальный, с холодной зимой, жарким и сухим летом, средняя годовая температура 5°C, счастливыми суховеями и пыльными бурями, а количество осадков уменьшается с севера на юг (200-350 мм).

ВЫВОДЫ:

Установлено, что суточный ход температуры воздуха по сезонам в отдельные десятилетия отличается, наблюдается тенденция к небольшому снижению в летние месяцы и увеличению в зимние, что приводит к уменьшению амплитуды сезонных колебаний в последние два десятилетия по сравнению с предыдущими годами. Прогноз методом случайного леса по сравнению с методом k-ближайших соседей более адекватные результаты. При этом, анализ гистограмм ошибок показал, что распределение ошибок для обеих моделей центрировано и близкую к нормальному (распределение Гаусса), что является хорошим признаком. В целом, обе модели показывают сопоставимую производительность, но random forests, как правило, имеет более высокие метрики R² (коэффициент детерминации), RMSE (среднеквадратическая ошибка) и MAE (средняя абсолютная ошибка), что делает его предпочтительной моделью для данной выборки данных.

ЛИТЕРАТУРА:

- Krasnopolsky V. M., Lin Y. A neural network nonlinear multimodel ensemble to improve precipitation forecasts over continental US // Advances in Meteorology. 2012. Article ID 649450. P.11. DOI: 10.1155/2012/649450
- Marzban C. A neural network for post-processing model output: ARPS // Monthly Weather Review. 2003. V. 131. № 4. PP. 1103-1111.
- Varentsov M.I., Krinitskiy M.A., Stepanenko V.M. Approximation of spatial and temporal variability of the urban heat island in Moscow using machine learning // Moscow university physics bulletin. 2024. V. 79. S2. PP. 784-797.
- Алферов Ю.В., Климова Е.Г. Опыт использования фильтра Калмана для коррекции численного прогноза приземной температуры воздуха // Гидрометеорологические исследования и прогнозы. 2020. № 4. С. 28-42.
- Багров А.Н., Быков Ф.Л., Гордин В.А. Схема оперативного краткосрочного комплексного прогноза приземной температуры и влажности воздуха // Метеорология и гидрология. 2018. № 8. С. 5-18.
- Вильфанд Р.М., Емелина С.В., Тищенко В.А., Толстых М.А., Хан В.М. Статистическая коррекция долгосрочных прогнозов приземной температуры воздуха по модели ПЛАВ для территории Северной Евразии // Метеорология и гидрология. 2024. № 5. С. 5-16.
- Зароченцев Г.А., Рубинштейн К.Г., Бычкова В.И., Игнатов Р.Ю., Юсупов Ю.И. Сравнение нескольких численных методов прогноза тумана // Оптика атмосферы и океана. 2018. Т. 31. № 12. С. 981-987.

*Работа выполнена в рамках госзадания № AAAA-A21-121011190016-1, при поддержке гранта РНФ № 24-17-00138