УДК 621.396.96 DOI: 10.26467/2079-0619-2025-28-1-20-38

# Применение Seq2seq-моделей для прогнозирования развития грозовой деятельности с целью повышения уровня ситуационной осведомленности пилота в полете

## Г.В. Коваленко<sup>1</sup>, И.А. Ядров<sup>1</sup>

<sup>1</sup>Санкт-Петербургский государственный университет гражданской авиации имени Главного маршала авиации А.А. Новикова, г. Санкт-Петербург, Россия

Аннотация: В статье представлены результаты применения моделей Seq2seq на основе нейронных сетей для наукастинга – прогнозирования с заблаговременностью до 2 часов грозовой активности с целью повышения ситуационной осведомленности экипажей воздушных судов. На основе данных радиолокационных метеорологических наблюдений за грозовыми очагами были созданы и обучены различные рекуррентные и сверточные рекуррентные модели. Результаты исследования показали, что сверточные рекуррентные нейронные сети (ConvRNN, ConvLSTM, ConvGRU) превосходят классические рекуррентные модели, и при этом позволяют улучшить прогноз развития грозы на 25-30 % по метрике RMSE (корень среднеквадратической погрешности) по сравнению с базовой моделью, каждый раз в качестве предсказания выбирающей последнее доступное на момент предсказания радиолокационное изображение. Тем не менее, несмотря на то что сверточные рекуррентные сети позволяют достаточно точно передать общую тенденцию изменения формы грозового облака, точность предсказания интенсивности элементов грозового очага оказывается, как правило, завышенной. Применение предложенной технологии прогнозирования грозовой активности может способствовать повышению уровня ситуационной осведомленности летного экипажа, улучшая проекцию текущей обстановки на ближайшее будущее и оптимизируя процесс принятия решений по обходу грозы за счет предоставления членам экипажа прогностической информации о развитии грозы на экране навигационного дисплея. В рамках будущих исследований предполагается дальнейшая оптимизация архитектуры моделей, а также интеграция прогностической технологии в системы поддержки принятия решений экипажем.

Ключевые слова: Seq2seq, прогнозирование, рекуррентные нейронные сети, сверточные рекуррентные нейронные сети, ситуационная осведомленность, обход грозы.

Для цитирования: Коваленко Г.В., Ядров И.А. Применение Seq2seq-моделей для прогнозирования развития грозовой деятельности с целью повышения уровня ситуационной осведомленности пилота в полете // Научный вестник МГТУ ГА. 2025. Т. 28, № 1. С. 20–38. DOI: 10.26467/2079-0619-2025-28-1-20-38

# Application of Seq2Seq models for predicting the development of thunderstorm activity to enhance the pilot's situational awareness in flight

# G.V. Kovalenko<sup>1</sup>, I.A. Yadrov<sup>1</sup>

<sup>1</sup> Saint Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A.A. Novikov, St. Petersburg, Russia

Abstract: The paper presents the results of application of Seq2seq models based on neural networks for nowcasting-forecasting with a lead time of up to 2 hours – of thunderstorm activity in order to increase situational awareness of aircraft crews. Various recurrent and convolutional recurrent models were created and trained on the basis of radar meteorological observations of thunderstorm cells. The results showed that convolutional recurrent neural networks (ConvRNN, ConvLSTM, ConvGRU) outperform classical recurrent models and improve the thunderstorm forecast by 25–30% in terms of RMSE (root mean square

#### <u>Том 28, № 01, 2025</u> Vol. 28, No. 01, 2025

error) metric compared to the baseline model, which always selects the most recent radar image available at the time of prediction. Nevertheless, despite the fact that the convolution recurrence models can accurately represent the general trend of thunderstorm cloud shape changes, the accuracy of predicting the intensity of thunderstorm cells is usually overestimated. Application of the proposed thunderstorm activity forecasting technology can enhance the situational awareness of the flight crew improving the projection of the current situation into the near future and optimizing the decision-making process for thunderstorm avoidance by providing crew members with predictive information about thunderstorm development on the navigation display screen. Future research is expected to further optimize the model architecture and integrate the predictive technology into flight crew decision support systems.

Key words: Seq2seq, prediction, recurrent neural networks, convolutional recurrent neural networks, situational awareness, thunderstorm avoidance.

For citation: Kovalenko, G.V., Yadrov, I.A. (2025). Application of Seq2Seq models for predicting the development of thunderstorm activity to enhance the pilot's situational awareness in flight. Civil Aviation High Technologies, vol. 28, no. 1, pp. 20–38. DOI: 10.26467/2079-0619-2025-28-1-20-38

### Введение

Ситуационную осведомленность члена летного экипажа воздушного судна (ВС) при выполнении полета можно определить как совокупность: 1) восприятия элементов окружающей среды; 2) понимания их значения; 3) проекции их состояния на ближайшее будущее [1]. Уровень безопасности и надежности, с которым экипажи ВС гражданской авиации выполняют полет, во многом обусловлен уровнем их ситуационной осведомленности [2, 3]: этот факт подтверждает актуальность работы по поддержанию и повышению ситуационной осведомленности пилота на различных этапах полета.

Сложности в обеспечении необходимого уровня ситуационной осведомленности могут возникать при выполнении полета в условиях неопределенности, например при полетах в зонах грозовой деятельности и сильных ливневых осадков. Нетривиальность процесса принятия решения по выбору оптимального маршрута обхода грозы подчеркивается в работе [4]: находясь в одних и тех же условиях, пилоты могут принять кардинально различающиеся решения по обходу зоны грозовой активности, причем сам процесс принятия решения усложняется необходимостью учитывать тенденцию развития грозовых очаговых движений относительно ВС. Предполагается, что предоставление членам экипажа на экране навигационного дисплея прогностической информации о развитии грозы (а также информации о местоположении ВС

относительно грозовых очагов в соответствующие моменты времени в соответствии с текущим планом полета) может способствовать повышению ситуационной осведомленности пилотов, оптимизации процесса работы экипажа при выполнении полетов в сложных метеорологических условиях и, как следствие, принятию наиболее рациональных решений по обходу грозы.

Стоит отметить, что в настоящей работе бортовые радиолокаторы рассматриваются в качестве единственного источника информации, на основе которой предполагается осуществлять прогнозирование развития очагов грозовой деятельности. Достоинство такого подхода заключается в простоте и автономности предлагаемого метода прогнозирования, который, таким образом, не будет зависеть от наличия информации от прочих источников спутниковых систем, станций наземной сети, грозопеленгаторов и т. д. В то же время объединение данных из различных источников потенциально могло бы позволить учесть большее количество факторов, влияющих на развитие гроз, и обеспечить более полную картину атмосферных процессов: такой подход имеет большой практический интерес, поскольку может способствовать значительному улучшению качества и точности прогнозов, тем не менее его реализация требует разработки новых алгоритмов обработки и интеграции данных о развитии гроз и выходит за пределы настоящего исследования.

Прогнозирование развития грозового очага может быть рассмотрено как решение за-



Рис. 1. Seq2seq-модель [12]: RNN – Recurrent Neural Network, рекуррентная нейронная сеть; C – контекстный вектор;  $X_i$  – элемент входной последовательности;  $y_i$  – элемент выходной последовательности Fig. 1. Seq2seq architecture [12]: RNN – Recurrent Neural Network; C – context vector;  $X_i$  – input sequence element;  $y_i$  – output sequence element

дачи предсказания временных рядов, при которой параметры используемой для этого модели машинного обучения подбираются таким образом, чтобы наиболее точно уловить закономерности во временных данных, а также выделить наиболее правдоподобный сценарий развития ситуации [5]. К классическим методам предсказания временных рядов можно отнести как разработанные специально для решения данной задачи модели, такие как, например, интегрированная модель авторегрессии – скользящего среднего ARIMA (Autoregressive Integrated Moving Average) и ее модификации [6] или модель экспоненциального сглаживания (Exponential Smoothing) [7], так и более общие классические подходы, как, например, модели градиентного бустинга [8]. Главными достоинствами таких моделей являются их стабильность при работе с «табличными» данными и относительная легкость интерпретации параметров моделей [9]. Тем не менее современные исследования показывают, что в случае применения таких моделей для прогнозирования в задачах, где данные имеют более сложную, «нетабличную» структуру (изображения, текст, звук) или где необходимо учитывать более сложные нелинейные взаимосвязи между элементами входной последовательности, классические модели показывают заметно более слабый результат по сравнению с нейросетевыми подходами: так, модель LSTM (Long Short-Term Memory, длинная цепь элементов краткосрочной памяти) про-

Энкодер

демонстрировала превосходство над классической ARIMA на 5–15 % по метрикам MAPE (Mean Absolute Percentage Error, средняя абсолютная ошибка) и sMAPE (Symmetric MAPE, симметричная средняя абсолютная ошибка) [10], а более сложные архитектурные решения, как, например, модели Seq2seq (Sequence-to-sequence, генерация последовательности на основе последовательности), превзошли показатели классических моделей на 20–30 % [11].

Декодер

Архитектура Seq2seq [12] представляет собой нейросетевую модель, предназначенную для прогнозирования выходной последовательности  $Y = [y_1, y_2, ..., y_m]^{\mathsf{T}}$  некоторой заданной длины *m* на основе входной последовательности  $X = [x_1, x_2, ..., x_n]^{\top}$  длины n таким образом, чтобы максимизировать условную вероятность  $\Pr(Y|X) = \prod_{t=1}^{m} \Pr(y_t|y_{1:(t-1)}, X)$ , при этом выходное значение  $y_t$  в момент времени t определяется как матрицей входных векторов Х, так и всеми предшествующими вы-Модель ходными значениями  $y_{1:(t-1)}$ . Seq2seq состоит из энкодера и декодера: на основе набора обучающих примеров Х энкодер формирует контекстный вектор С, который в свою очередь используется в декодере для генерации результирующей матрицы У и в дальнейшем обновляется после получения на вход декодера очередного предыдущего предсказания  $y_{t-1}$  (рис. 1). В основе функционирования описанной модели лежит принцип формирования и обновления вектора скрыто-



 Рис. 2. Схемы ячеек рекуррентных нейронных сетей [16]: RNN – рекуррентная нейронная сеть; LSTM – длинная цепь элементов краткосрочной памяти; GRU – управляемый рекуррентный блок
 Fig. 2. Recurrent neural networks cells [16]: RNN – Recurrent Neural Network; LSTM – Long Short-Term Memory; GRU – Gated Recurrent Unit

го состояния, реализованный в рекуррентных нейронных сетях (Recurrent Neural Networks, RNN).

К рекуррентным относят класс нейронных сетей, в которых связи между элементами образуют направленную последовательность [13]: за счет формирования и обновления вектора скрытого состояния  $h_t$  нейронная сеть в каждый момент времени несет в себе информацию об элементах входной последовательности, поступивших на вход сети ранее. Структуру классической RNN [14] можно описать с помощью уравнений обновления скрытого состояния  $h_t$  и получения предсказания  $y_t$  в момент времени t (рис. 2):

$$h_t = \sigma(W_{hh}h_{t-1} + W_{xh}x_t + b_h),$$
 (1)

$$y_t = W_{hy}h_t + b_y, \tag{2}$$

где  $h_t$  – вектор скрытого состояния в момент времени t;  $x_t$  – входной вектор в момент времени t;  $\sigma$  – некоторая нелинейная функция активации, как правило гиперболический тангенс или ReLU (Rectified Linear Unit, линейный блок с выпрямлением) [15];  $W_{hh}$ ,  $W_{xh}$ и  $W_{hy}$  – матрицы весов (обучаемые параметры сети);  $b_h$  и  $b_y$  – свободные члены (обучаемые параметры сети).

Усовершенствованный способ хранения и обработки информации, попавшей на вход сети, реализован в модели LSTM [17]: в ней в дополнение к вектору скрытого состояния, в котором хранится краткосрочная информация, вводятся отвечающая за долгосрочное хранение информации ячейка памяти  $c_t$ , а также

коэффициенты  $i_t$ ,  $f_t$  и  $o_t$ , известные как фильтр входа (input gate), фильтр забывания (forget gate) и фильтр выхода (output gate) соответственно: эти коэффициенты принимают значения на отрезке [0, 1] и позволяют изменять значения соответствующих векторов путем их поэлементного перемножения на коэффициенты  $i_t$ ,  $f_t$  и  $o_t$  при прохождении входного вектора через LSTM-ячейку (рис. 2):

$$\dot{u}_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i),$$
 (3)

$$f_t = \sigma \Big( W_{xf} x_t + W_{hf} h_{t-1} + b_f \Big), \qquad (4)$$

$$p_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o),$$
 (5)

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh (W_{xc} x_t + W_{hc} h_{t-1} + b_c),$$
(6)

$$h_t = o_t \odot \tanh(c_t), \tag{7}$$

где  $W_{xi}$ ,  $W_{hi}$ ,  $W_{xf}$ ,  $W_{hf}$ ,  $W_{xo}$ ,  $W_{ho}$ ,  $W_{xc}$  и  $W_{hc}$  – матрицы весов (обучаемые параметры сети);  $b_i$ ,  $b_f$ ,  $b_o$  и  $b_c$  – свободные члены (обучаемые параметры сети);  $\bigcirc$  – оператор поэлементно-го умножения.

Другой разновидностью рекуррентных нейронных сетей является более легковесная по сравнению с LSTM (с меньшим числом обучаемых параметров) модель GRU (Gated Recurrent Unit, управляемый рекуррентный блок): в ней не используется ячейка памяти  $c_t$ , фильтры входа и забывания объединены в единый фильтр обновления  $z_t$  (update gate), отвечающий за то, какая доля предыдущего скрытого состояния  $h_{t-1}$  сохраняется в оче-



**Рис. 3.** Механизм свертки [20]: *X* – входная матрица; *K* – матрица-фильтр; \* – оператор свертки **Fig. 3.** Convolution [20]: *X* – input matrix; *K* – kernel matrix; \* – convolution operator

редном состоянии  $h_t$ , а также добавлен фильтр сброса  $r_t$  (reset gate), определяющий долю информации  $h_{t-1}$ , которая на очередной итерации будет проигнорирована (рис. 2):

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z),$$
 (8)

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1} + b_r), \qquad (9)$$

$$\tilde{h}_t = \tanh(W_{xh}x_t + r_t \odot (W_{hh}h_{t-1}) + b_h), (10)$$

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t, \qquad (11)$$

где  $W_{xz}$ ,  $W_{hz}$ ,  $W_{xr}$ ,  $W_{hr}$ ,  $W_{xh}$  и  $W_{hh}$  – матрицы весов (обучаемые параметры сети);  $b_z$ ,  $b_r$  и  $b_h$  – свободные члены (обучаемые параметры сети).

Наиболее широкое применение рекуррентные нейронные сети нашли в задачах обработки естественного языка [18], когда входная последовательность нейронной сети представляет собой последовательность векторных представлений слов – эмбеддингов. Тем не менее в виде вектора может быть представлено и изображение окружающей метеорологической обстановки, предоставляемое пилоту на экране бортового радиолокатора: для этого матрицу  $\mathbb{R}^{m \times n}$ , каждый элемент которой соответствует интенсивности пикселя навигационного дисплея, необходимо преобразовать в вектор  $\mathbb{R}^{mn}$ , после чего последовательность таких векторов может быть передана на вход рекуррентной нейронной сети – элемента Seq2seq-модели – с целью получения прогноза развития грозы на заданный промежуток времени. Как отмечается в [19], очевидным недостатком такого подхода является неизбежная потеря полезной информации о взаимном расположении участков грозового очага в пространстве. Один из потенциальных путей решения описанной проблемы заключается в дополнительном использовании механизма свертки и переходе сверточным рекуррентным К нейронным сетям (ConvRNN, Convolutional RNN) [19].

Свертка представляет собой один из основных нейросетевых механизмов, предназначенный для сохранения заложенной в некоторой входной матрице X пространственной информации при прохождении X через нейронную сеть. Процесс свертки может быть представлен как создание новой матрицы U = X \* K, получаемой в результате матричного умножения элементов исходной матрицы X на матрицу-фильтр K при последовательном «скольжении» фильтра по Xслева направо, сверху вниз [20] (рис. 3).

Математически механизм свертки для входной матрицы  $X \in \mathbb{R}^{H \times W}$  и фильтра  $K \in \mathbb{R}^{h \times w}$  может быть описан следующим образом:

$$U[i,j] = (X * K)[i,j] = \sum_{m=0}^{h-1} \sum_{n=0}^{w-1} X[i+m,j+n] \cdot K[m,n],$$
(12)

где \* – оператор свертки; запись U[i, j] соответствует элементу матрицы U, записанному в *i*-й строчке *j*-го столбца.

Таким образом, применение механизма свертки в рекуррентных нейронных сетях позволяет перейти к архитектуре ConvRNN, на вход которой в момент времени t подается матрица  $X_t$ , а матрица скрытого состояния  $H_t$ (в отличие от вектора скрытого состояния RNN $h_t$ ) обновляется в соответствии с законом

$$H_t = \sigma(W_{hh} * H_{t-1} + W_{xh} * X_t + b_h).$$
(13)

По аналогии с ConvRNN были разработаны сверточная разновидность модели LSTM – ConvLSTM – с дополнительным хранением долгосрочной информации в матрице памяти C<sub>t</sub> [20]:

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i), \quad (14)$$

$$f_t = \sigma \big( W_{xf} * X_t + W_{hf} * H_{t-1} + b_f \big), \quad (15)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o), \quad (16)$$

$$C_{t} = f_{t} \odot C_{t-1} + i_{t} \odot \tanh (W_{xc} * X_{t} + W_{hc} * H_{t-1} + b_{c}), \quad (17)$$

$$H_t = o_t \odot \tanh(\mathcal{C}_t), \tag{18}$$

а также сверточная разновидность модели GRU – ConvGRU [21]:

$$z_t = \sigma(W_{xz} * X_t + W_{hz} * H_{t-1} + b_z), \quad (19)$$

$$r_t = \sigma(W_{xr} * X_t + W_{hr} * H_{t-1} + b_r), \quad (20)$$

$$\widetilde{H}_t = \tanh (W_{xh} * X_t + r_t \odot (W_{hh} * H_{t-1}) + b_h), \quad (21)$$

$$H_t = z_t \odot H_{t-1} + (1 - z_t) \odot \widetilde{H}_t.$$
 (22)

В настоящей работе предпринята попытка оценки потенциала прогнозирования развития грозового очага с заблаговременностью менее 2 часов (наукастинга) как инструмента повышения ситуационной осведомленности пилота при выполнении полета в условиях сложной метеорологической обстановки, а также реализовано сравнение рекуррентных и сверточных рекуррентных моделей прогнозирования по таким характеристикам, как точность прогноза и скорость обучения, применительно к решению задачи предсказания развития грозы.

#### Методы исследования

Создание моделей рекуррентных и сверточных рекуррентных нейронных сетей, их обучение и оценка качества, а также предшествующие этапы получения и подготовки данных для обучения осуществлялись с использованием языка объектно ориентированного программирования Python, а также библиотек: NumPy с поддержкой операций над многоуровневыми массивами, PyTorch для создания архитектуры нейронных сетей, их обучения и оценки качества, а также OpenCV, Rasterio и Matplotlib для предобработки и визуализации растровых изображений. Код находится в открытом доступе<sup>1</sup>.

Для обучения моделей были использованы данные<sup>2</sup> радиолокационных метеорологических наблюдений за грозовыми очагами и сильными ливневыми осадками Финского метеорологического института (Finnish Meteorological Institute, FMI) за период с января 2021 года по июнь 2024 года включительно. Достоинства указанного источника заключаются в том, что данные находятся в открытом доступе, представлены за достаточно большой промежуток времени, а также обновляются регулярно с периодичностью в 5 мин, то есть достаточно часто для того, чтобы на их основе реализовать качественное прогнозирование развития грозовых облаков.

<sup>&</sup>lt;sup>1</sup> Код для предобработки и визуализации растровых изображений [Электронный ресурс] //github.com. URL: https://github.com/yaaadrov/WxRADNet-PyTorch (дата обращения: 08.10.2024).

<sup>&</sup>lt;sup>2</sup> Схема данных [Электронный ресурс] // XML-фал. URL: http://s3-eu-west-1.amazonaws.com/fmiopendata-radar-geotiff (дата обращения: 08.10.2024).



Рис. 4. Покрытие сети радаров FMI; пример полученного радиолокационного изображения; отображение полученной конфигурации грозы на экране самолетного метеолокатора
 Fig. 4. FMI radar network coverage; example of the obtained radar image; display of the received thunderstorm configuration on the aircraft weather radar screen

Метеорологические наблюдения проводятся на базе сети FMI из 12 импульснодоплеровских двухполяризационных радаров С-диапазона, покрывающих территорию Финляндии; полученные с их помощью значения радиолокационной отражаемости Z (dBZ) переводятся в значения интенсивности пикселя I (число от 0 до 255), после чего картина радиолокационной обстановки сохраняется в виде растрового черно-белого изображения формата tiff. Получить значение Z на основе интенсивности пикселя I можно по формуле  $Z = 0,5 \cdot I - 32$  (рис. 4).

Современные наземные доплеровские грозовые радары позволяют обеспечить высокую достоверность получаемых с их помощью данных, а также хорошее пространственное покрытие и детализацию в радиусе до 200-250 км [22]; тем не менее недостаток таких радаров заключается в их подверженности влиянию помех, возникающих при наличии в воздухе отражающих объектов, не связанных с осадками (например, птиц). Чтобы избавиться от помех подобного рода, на этапе предобработки данных пиксели со значениями интенсивности I менее 100 (с отражаемостью Z менее 20 dBZ, что соответствует слабому дождю) отфильтровывались: значения их интенсивности заменялись на 0, однако это не приводило к потере информации о сильных ливневых осалках.

Стоит отметить, что полученные данные использовались для обучения модели лишь в исследовательских целях – для сравнения различных подходов к прогнозированию – и предобученные на этих данных модели не могут быть использованы сразу же для дальнейшего практического применения, так как при их обучении не учитывались индивидуальные характеристики конкретных самолетных радиолокаторов – их чувствительность, пороговые уровни интенсивности, разрешение, частота обновления и другие параметры, которые могут существенно повлиять на точность и актуальность прогнозов. Для реального применения предложенных в настоящей работе моделей требуется дополнительно реализовать их дообучение на данных, собранных непосредственно с бортового радиолокатора конкретного типа ВС.

Получение и предварительная обработка данных производились следующим образом.

1. Соответствующее началу очередного часа изображение в формате tiff загружалось, преобразовывалось в массив и разделялось на четыре непересекающиеся квадратные области размером  $200 \times 200$  км каждая, образуя массивы размерности  $256 \times 256$  пикселей (далее – [h, w]).

2. В качестве информативных выбирались те области, в которых пиксели с интенсивностью более 100 составляли не менее



Рис. 5. Примеры часовых последовательностей из 12 изображений, включая аугментированные данные Fig. 5. Examples of hourly sequences from 12 images including augmented data

15 % от общего числа; неинформативные области из дальнейшей обработки исключались.

3. Для каждой из полученных информативных областей загружались следующие доступные за текущий час радиолокационные изображения, образуя массив размерности [a, h, w], где a = 12 – число изображений, полученных для информативной области за 1 час.

4. Для каждой из полученных последовательностей изображений реализовывалась аугментация данных [23]: набор обучающих данных расширялся путем создания трех синтетических последовательностей размерности [a, h, w], каждая за счет поворота изображений на случайные углы, лежащие в диапазонах (60°, 120°), (150°, 210°) и (240°, 300°).

5. Все информативные области за текущий час объединялись в массив размерности [i, a, h, w], где i - число информативных сегментов за текущий час, после чего данный массив расширялся аналогично за счет данных за последующие часы.

В результате общее число полученных часовых последовательностей из 12 изображений с добавлением аугментированных данных составило 34 240 (рис. 5), 32 800 из которых использовались для обучения моделей и 1 440 (за июнь 2024 года) – для тестирования и сравнения моделей. Данные были использованы для обучения двух групп моделей – рекуррентных (RNN, LSTM и GRU) и сверточных рекуррентных (ConvRNN, ConvLSTM и ConvGRU) нейронных сетей. Каждая из часовых последовательностей была разделена на две равные части по шесть последовательных изображений: первая из них (то есть первые полчаса соответствующего часа) использовалась в качестве известных входных векторов  $x_t$ , а вторая (то есть вторые полчаса каждого часа) – в качестве выходных векторов  $y_t$ , значение которых требовалось предсказать.

Для всех видов рекуррентных нейронных сетей использовалась схожая архитектура, состоящая из следующих элементов:

1) энкодера, представляющего собой две последовательные ячейки (RNN, LSTM или GRU), реализующие процесс создания и обновления вектора скрытого состояния  $h_t$ (и дополнительно  $c_t$  для LSTM) согласно уравнениям (1), (3–7) и (8–11) соответственно;

2) декодера, представляющего собой две последовательные ячейки (RNN, LSTM или GRU), за которыми следует полносвязный слой (Fully-connected, FC), реализующий преобразование вектора скрытого состояния  $h_t$  в вектор-предсказание  $y_t$  в соответствии с уравнением (2).

Особенность рекуррентных нейронных сетей заключается в том, что в качестве входной последовательности  $x_t$  они способны получать лишь вектор, поэтому каждый из двумерных входных массивов преобразовывался в вектор размерности  $[h \cdot w]$ ; модель обучалась параллельно на батчах [24] размера 32: таким образом, в каждый момент времени t на вход модели подавался массив размерности  $[b, h \cdot w]$ , где b – размер батча.

#### Научный Вестник МГТУ ГА Civil Aviation High Technologies



Рис. 6. Архитектура использованных сетей: RNN – рекуррентный слой; ConvRNN – сверточный рекуррентный слой; FC – полносвязный слой; Conv – сверточный слой Fig. 6. Architecture of neural networks used: RNN – Recurrent Layer; ConvRNN – Convolutional Recurrent Layer; FC – Fully-connected Layer; Conv – Convolutional Layer

Принцип обучения рекуррентной модели заключается в следующем: энкодер последовательно получает каждый из шести входных векторов  $x_1 \dots x_6$ , на основе которых формируется вектор скрытого состояния  $h_6$ ; вектор  $h_6$  поступает на вход декодера, который в свою очередь формирует первое предсказание  $y_1$  и новый вектор скрытого состояния  $h_7$  (рис. 6).

Схожая архитектура использовалась и для всех видов сверточных рекуррентных сетей, за исключением следующих особенностей (рис. 6):

1) на вход сверточной рекуррентной сети поступали массивы размерности [b, h, w], так как механизм свертки предназначен для обработки матриц [h, w], причем в силу ограниченности вычислительных ресурсов использовался размер батча b = 1;

2) вместо классических рекуррентных ячеек RNN, LSTM или GRU использовались их сверточные аналоги – ConvRNN, ConvLSTM или ConvGRU, – реализующие обновление процесса создания и обновления матрицы скрытого состояния  $H_t$  (и дополнительно  $C_t$ для LSTM) в соответствии с уравнениями (13), (14)–(18) и (19)–(22) соответственно;

3) вместо полносвязного слоя в декодере в качестве последнего использован сверточный слой (Conv) с единичным фильтром таким образом, чтобы размер выходной матрицы соответствовал требуемому значению [h, w].

Обучение моделей (настройка обучаемых параметров сетей) реализовывалось с помощью градиентного спуска – методом обратного распространения ошибки во времени [25] с использованием алгоритма AdamW (Adaptive Momentum Estimation с неявным добавлением  $\ell_2$ -регуляризации) [26]: коэффициенты обновлялись на основе антиградиента среднеквадратичной функции ошибки (MSE, Mean Squared Error) с учетом скорости обучения (Learning

#### Таблица 1 Table 1

Параметры обучения рекуррентных и сверточных рекуррентных нейронных сетей Used RNN and ConvRNN parameters

	RNN	ConvRNN
Число нейронов скрытого слоя	128	64
Число слоев	2	2
Размер батча	32	1
Скорость обучения α	0,001	0,001
Число эпох обучения	30	20
Размер фильтра свертки	_	3

Rate) α, для чего на каждой итерации алгоритма рассчитывалось значение MSE на обучающей выборке (а также дополнительно и на тестовой выборке для отслеживания динамики обучения моделей и предотвращения их переобучения [27]) по формуле

MSE = 
$$\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
, (23)

где n – общее число наблюдений, участвующих в расчете;  $y_i$  – реальное значение массива;  $\hat{y}_i$  – значение, предсказанное моделью.

Для ускорения обучения моделей использовалась видеокарта GPUNVIDIAA100 с тензорными ядрами. Кроме того, во время обубыли реализованы: 1) методика чения уменьшения скорости обучения α на плато замедлении уменьшения значений (при MSE) [28]; 2) ранняя остановка обучения в случае, если в течение пяти последовательных эпох обучения значение MSE на тестовых данных не опускается ниже предыдущего лучшего значения; 3) обрезка градиентов (Gradient Clipping) во время обратного распространения ошибки [29] для предотвращения негативного влияния взрывных градиентов. Параметры обучения указаны в табл. 1.

### Результаты исследования

В ходе обучения моделей после прохождения очередной эпохи (то есть после того, как все обучающие данные очередной раз поступили на вход модели) рассчитывались значения MSE на обучающей и тестовой выборках; результаты обучения представлены на рис. 7.

Для большинства моделей (RNN, GRU, ConvLSTM и ConvGRU) используемого числа эпох оказалось достаточно для выхода функции потерь на плато, в то время как графики для сетей типа LSTM и ConvRNN свидетельствуют о возможности потенциального уменьшения MSE при условии дальнейшего обучения моделей. Для классических рекуррентных сетей свойственно плавное уменьшение метрики в начале обучения, за которым следуют колебания, сопровождающиеся переменным ростом ошибки на тестовой выборке, что свидетельствует о возможном выходе моделей на стадию переобучения. В то же время графики обучения сверточных рекуррентных моделей характеризуются скачкообразными изменениями функции потерь на начальных эпохах обучения, за которыми следует плавное уменьшение метрики, показывающее, что в течение рассматриваемого числа эпох модели не успевают переобучиться. Отдельного внимания заслуживает тот факт, что в отличие от традиционных RNN, LSTM и GRU для их сверточных аналогов в течение всего периода обучения значения функции потерь на тестовой выборке оказываются ниже, чем на обучающей.

Метрики качества моделей – корень среднеквадратической ошибки (RootMSE, RMSE) на тестовой выборке после обучения, – а также среднее время, затрачиваемое на одну эпоху обучения, представлены в табл. 2.



Рис. 7. Результаты обучения моделей Fig. 7. Models' training results

Таблица 2 Table 2

метрики качества моделеи
Models' metrics

	Среднее время обучения эпохи	RMSE
Базовая модель	—	39,5244
RNN	3 мин 9 с	43,3217
LSTM	3 мин 7 с	41,8138
GRU	2 мин 53 с	42,3127
ConvRNN	17 мин 21 с	29,3511
ConvLSTM	47 мин 36 с	27,5801
ConvGRU	39 мин 32 с	26,2939

### Обсуждение

В контексте решаемой задачи значение метрики RMSE можно рассматривать как среднюю ошибку определения интенсивности пикселя при прогнозировании развития грозового очага на ближайшие 30 мин на основе данных о предыдущих 30 мин с дискретностью 5 мин. В качестве базы для интерпретации точности моделей было использовано значение RMSE константной модели, каждый раз предсказывающей последнюю доступную на момент предсказания (шестую) конфигурацию грозовых очагов: для такой модели получено значение RMSE, равное 39,5244, что составляет около 15,5 % максимальной интенсивности пикселя.

Среди рассмотренных моделей наихудшая точность соответствует классической RNN: значение RMSE, равное 43,3217, составляет приблизительно 17 % максимальной интенсивности пикселя, что на 9,6 % больше (хуже), чем базовое значение точности 39,5244. Лучшая из классических рекуррентных моделей (LSTM) со значением средней ошибки 41,8138 также не смогла преодолеть базовый порог RMSE. Наилучший результат продемонстрировала сверточная рекуррентная модель ConvGRU со значением RMSE, равным 26,2939, то есть около 10 % максимальной интенсивности пикселя. Здесь важно отметить, что всем моделям, использующим механизм свертки, удалось преодолеть базовое значение RMSE константной модели и показать значения на 25-30 % меньше (лучше).

Полученный результат свидетельствует о непригодности классических рекуррентных моделей для прогнозирования развития грозы. Более точное представление о причинах установленного факта, а также особенностях предсказания развития грозы каждой из моделей можно получить, рассмотрев результаты прогнозирования развития грозовых очагов для конкретного обучающего примера (рис. 8).

Рис. 8 показывает, что обычные рекуррентные модели оказываются неспособны уловить особенности распределения отдельных участков грозового очага в пространстве и, несмотря на то что определяют общую тенденцию перемещения грозы, не могут сохранить пространственную структуру данных, в результате чего особенности формы грозы теряются и участки грозового облака имеют нехарактерные для такого метеоявления пространственные очаги. Данный факт можно объяснить особенностью работы классических RNN: они были созданы для работы с одномерными массивами данных, а разворачивание двумерного изображения в вектор приводит к потере пространственных взаимосвязей пикселей, в результате чего теряется и возможность спрогнозировать точную форму грозового облака.

В то же время рис. 8 показывает, что сверточные рекуррентные нейронные сети верно уловили тенденцию к уменьшению интенсивности рассматриваемого грозового очага, а также в общих чертах передали его перемещение налево. Данный факт можно объяснить тем, что использование в моделях ConvRNN, ConvLSTM и ConvGRU механизма свертки позволяет реализовать обработку изображений в виде двумерных массивов, извлекая из данных пространственные признаки и учитывая зависимости между соседними пикселями на всех уровнях свертки:

1) в процессе обучения параметры (веса) фильтров свертки подбираются таким образом, чтобы фильтры были способны выявлять важные локальные признаки изображения – вертикальные и горизонтальные линии, углы, криволинейные структуры и т. п.;

2) в то время как один сверточный слой захватывает простые пространственные признаки, добавление дополнительных сверток позволяет нейронной сети выявлять более высокоуровневые признаки, например контуры грозовых облаков;

3) поскольку свертка выполняется на небольших локальных областях изображения, сверточные слои учитывают главным образом локальные пространственные зависимости: это свойство позволяет сети быть инвариантной к небольшим сдвигам, то есть после смещения грозового очага сверточный слой сможет найти и идентифицировать каждый его участок на следующем изображении.

Тем не менее, несмотря на достаточно высокое качество предсказания формы грозового облака с использованием сверточных рекуррентных сетей, значения интенсивности предсказанных пикселей, очевидно, являются завышенными по сравнению с реальной картиной развития грозы; поиск решения указанного недостатка ConvRNN, ConvLSTM и ConvGRU планируется развить в дальнейших исследованиях.

Предполагаемый положительный результат применения предлагаемой технологии прогнозирования развития грозовых очагов заключается в потенциальном росте ситуационной осведомленности пилотов, выполняющих полеты в условиях грозовой деятельности и ливневых осадков, за счет представле-

#### Научный Вестник МГТУ ГА Civil Aviation High Technologies

#### <u>Том 28, № 01, 2025</u> Vol. 28, No. 01, 2025



**Рис. 8.** Первые и последующие шесть радиолокационных изображений одной из конфигураций тестовой выборки, а также предсказания развития грозы для рассматриваемых моделей

Fig. 8. First and next six radar images of one of the test sample configurations, as well as predictions of thunderstorm development for the considered models

ния на экране навигационного дисплея прогностической информации о развитии грозового фронта.

С математической точки зрения ситуационная осведомленность *SA* (Situational Awareness) может быть представлена как некоторая функция f переменных: 1)  $P_{env}$  – восприятия пилотом текущей обстановки (Perceptionof Current Environment); 2)  $C_{env}$  – комплексного понимания текущей обстановки (Comprehensionof Current Environment); 3)  $P_{fut}$  – проекции текущей обстановки на ближайшее будущее; 4) H – комплексного параметра, отвечающего за вклад человеческого фактора (внимание, опыт членов экипажа, текущий уровень рабочей нагрузки и т. д.) в обеспечение ситуационной осведомленности [30]:

$$SA = f(P_{env}, C_{env}, P_{fut}, H).$$
(24)

Предоставление пилоту прогностической информации о развитии грозового очага мо-

#### Научный Вестник МГТУ ГА Civil Aviation High Technologies

жет способствовать повышению качества проекции текущего состояния окружающей среды на ближайшее будущее за счет обновления вероятностного распределения будущей метеорологической обстановки  $W_{fut}$  (Future Weather), которое теперь будет вычисляться не только на основе параметров текущей метеорологической обстановки  $I_{cur}$  (Current Information, текущая информация), но также на основе прогностической информации  $I_{pred}$  (Predictive Information, прогностическая информация), таким образом снижая уровень неопределенности:

$$\Pr(W_{fut} \mid I_{cur}, I_{pred}) = \frac{\Pr(I_{cur} \mid W_{fut}, I_{pred}) \cdot \Pr(W_{fut})}{\Pr(I_{cur})},$$
(25)

где  $\Pr(W_{fut} | I_{cur}, I_{pred})$  – апостериорная вероятность получить некоторую радиолокационную картину в будущем с учетом дополнительного использования прогностической информации  $I_{pred}$ ,  $\Pr(I_{cur} | W_{fut}, I_{pred})$  – условная вероятность наблюдать текущую радиолокационную картину,  $\Pr(W_{fut})$  – предварительная вероятность получить некоторую радиолокационную картину в будущем (до получения предсказания),  $\Pr(I_{cur})$  – вероятность наблюдения текущей радиолокационной картины.

Кроме того, предполагается, что предложенная прогностическая модель не только позволит оптимизировать процесс принятия пилотами решений по обходу грозовых очагов и снизить когнитивную нагрузку членов летного экипажа, но также может применяться в Байесовских моделях оценки риска [31] и сможет быть интегрирована в самолетные системы для обеспечения автоматизированного обхода грозы [32].

### Заключение

Таким образом, в настоящей работе приведены результаты применения Seq2seq-моделей (нейронных сетей) для наукастинга развития грозовой деятельности с целью повышения уровня ситуационной осведомленности членов летного экипажа в полете.

1. На этапе подготовки к исследованию были получены данные радиолокационных метеорологических наблюдений за грозовыми очагами и сильными ливневыми осадками Финского метеорологического института за период с января 2021 года по июнь 2024 года. и подготовлено Всего было получено 32 800 обучающих и 1 440 тестовых последовательностей по 12 радиолокационных изображений, которые могут использоваться и в дальнейших исследованиях для объективного сравнения метрик качества различных подходов к прогнозированию развития грозы.

2. Созданы и обучены рекуррентные (RNN, LSTM и GRU) и сверточные рекуррентные (ConvRNN, ConvLSTM и ConvGRU) нейронные сети, реализующие предсказание шести последовательных радиолокационных изображений на основании первых шести изображений.

3. Интерпретация предсказаний моделей, а также анализ графиков изменения функции потерь на различных эпохах обучения и сравнение моделей на основе метрики RMSE (корень среднеквадратической ошибки на тестовой выборке) показали, что классические рекуррентные модели оказались непригодны для прогнозирования развития грозы: разворачивание двумерного радиолокационного изображения в вектор приводит к потере пространственных взаимосвязей, в результате чего теряется возможность спрогнозировать точную форму грозового облака.

4. В то же время сверточные рекуррентные модели продемонстрировали результат на 25–30 % лучше (по RMSE), чем базовая модель, в качестве предсказаний использующая последнее доступное на момент предсказания радиолокационное изображение: применение в моделях ConvRNN, ConvLSTM и ConvGRU механизма свертки позволяет реализовать обработку изображений в виде двумерных массивов, извлекая из данных пространственные признаки и учитывая зависимости между соседними пикселями на всех уровнях свертки.

5. Потенциальный положительный эффект применения предлагаемой технологии прогнозирования развития грозы заключается в возможном повышении уровня ситуационной осведомленности членов летного экипажа за счет предоставления им прогностической информации о развитии грозы на экране навигационного дисплея: предполагается, что наличие указанной информации будет способствовать более качественной реализации проекции текущей обстановки на ближайшее будущее, а также позволит оптимизировать процесс принятия пилотами решений по обходу грозовых очагов и снизить когнитивную нагрузку членов летного экипажа при выполнении полетов в сложных метеорологических условиях.

Полученные результаты предполагается развить в будущих исследованиях.

Использование в архитектурах ConvRNN, ConvLSTM и ConvGRU нескольких уровней свертки разного размера может позволить захватывать пространственные признаки на различных масштабах и, как результат, способствовать улучшению качества прогнозов.

Другим потенциальным подходом к прогнозированию развития грозы могут быть гибридные модели, объединяющие в себе подходы ConvRNN и ARIMA.

Оптимизация процесса принятия решений при обходе грозы может быть достигнута за счет интеграции предлагаемой технологии в

интеллектуальную адаптивную систему поддержки принятия экипажем ВС решений по обходу очагов грозовой деятельности, для чего требуется дополнительно разработать алгоритм определения оптимального маршрута обхода грозы на основе прогностической информации о ее развитии в ближайшее время.

### Список литературы

**1. Bolstad C.A., Riley J.M.** Using goal directed task analysis with Army brigade officer teams // Proceedings of the Human Factors and Ergonomics Society Annual Meeting. SAGE Publications. 2002. Vol. 46, no.3. Pp. 472–476. DOI:10.1177/154193120204600354

2. Stanton N.A., Chambers P.R.G., Piggott J. Situational awareness and safety // Safety science. 2001. Vol. 39, no. 3. Pp. 189–204. DOI: 10.1016/S0925-7535(01)00010-8

**3.** Sarter N.B., Woods D.D. Situation awareness: A critical but Ill-defined phenomenon // The International Journal of Aviation Psychology. 1991. Vol. 1, no. 1. Pp. 45–57. DOI: 10.1207/s15327108ijap0101 4

4. Стрелков Ю.К. Инженерная и профессиональная психология: учеб. пособие. М.: Академия; Высшая школа, 2001. 360 с.

**5. De Gooijer J.G., Hyndman R.J.** 25 years of time series forecasting // International Journal of Forecasting. 2006. Vol. 22, no. 3. Pp. 443–473. DOI: 10.1016/j.ijforecast.2006.01.001

6. Stevenson S. A comparison of the forecasting ability of ARIMA models // Journal of Property Investment & Finance. 2007. Vol. 25, no. 3. Pp. 223–240. DOI: 10.1108/14635780 710746902

7. Chatfield C. A new look at models for exponential smoothing / C. Chatfield, A.B. Koehler, J.K. Ord, R.D. Snyder // Journal of the Royal Statistical Society: Series D (The Statistician). 2001. Vol. 50, no. 2. Pp. 147–159. DOI: 10.1111/1467-9884.00267

8. Bentéjac C., Csörgő A., Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms // Artificial Intelligence Review. 2021. Vol. 54. Pp. 1937–1967. DOI: 10.1007/s10462-020-09896-5 9. Faloutsos C. Classical and contemporary approaches to big time series forecasting / C. Faloutsos, J. Gasthaus, T. Januschowski, Y. Wang // SIGMOD '19: Proceedings of the 2019 International Conference on Management of Data, 2019. Pp. 2042–2047. DOI: 10.1145/ 3299869.3314033

**10.** Makridakis S., Spiliotis E., Assimakopoulos V. The M4 Competition: Results, findings, conclusion and way forward // International Journal of Forecasting. 2018. Vol. 34, iss. 4. Pp. 802–808. DOI: 10.1016/j.ijforecast.2018. 06.001

11. Taieb S.B., Sorjamaa A., Bontempi G. Multiple-output modeling for multi-step-ahead time series forecasting // Neurocomputing. 2010. Vol. 73, iss. 10–12. Pp. 1950–1957. DOI: 10.1016/ j.neucom.2009.11.030

**12.** Sutskever I., Vinyals O., Quoc V.L. Sequence to sequence learning with neural networks // Proceedings of the 27th International Conference on Neural Information Processing Systems. 2014. No. 2. Pp. 3104–3112. DOI: 10.48550/arXiv.1409.3215

**13. Caterini A.L.** Recurrent neural networks / A.L. Caterini, D.E. Chang, A.L. Caterini, D.E. Chang. In book: Deep Neural Networks in a Mathematical Framework. Springer Briefs in Computer Science. Springer, Cham, 2018. Pp. 59–79. DOI: 10.1007/978-3-319-75304-1\_5

**14. Rumelhart D.E., Hinton G.E., Williams R.J.** Learning representations by backpropagating errors // Nature. 1986. No. 323. Pp. 533–536. DOI: 10.1038/323533a0

**15. Rasamoelina A.D., Adjailia F., Sin**čák P. A review of activation function for artificial neural network // 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI). Slovakia, Herlany, 2020. Pp. 281–286. DOI: 10.1109/SAMI48414. 2020.9108717

**16. Toharudin T.** Employing long shortterm memory and Facebook prophet model in air temperature forecasting / T. Toharudin, R.S. Pontoh, R.E. Caraka, S. Zahroh, Y. Lee Y, R.C. Chen // Communications in Statistics-Simulation and Computation. 2023. Vol. 52, iss. 2. Pp. 279–290. DOI: 10.1080/03610918.2020. 1854302 **17. Schmidhuber J., Hochreiter S.** Long short-term memory // Neural Computation. 1997. Vol. 9, iss. 8. Pp. 1735–1780. DOI: 10.1162/ neco.1997.9.8.1735

**18. De Mulder W., Bethard S., Moens M.F.** A survey on the application of recurrent neural networks to statistical language modeling / Computer Speech & Language. 2015. Vol. 30, iss. 1. Pp. 61–98. DOI: 10.1016/j.csl. 2014.09.005

**19. Shi X.** Convolutional LSTM network: A machine learning approach for precipitation nowcasting / X. Shi, Z. Chen, H. Wang, D.Y. Yeung, W.K. Wong, W.C. Woo // NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems. 2015. Vol. 1. Pp. 802–810. DOI: 10.48550/ arXiv.1506.04214

**20.** Li Z. A survey of convolutional neural networks: analysis, applications, and prospects / Z. Li, F. Liu, W. Yang, S. Peng, J. Zhou // IEEE transactions on neural networks and learning systems. 2021. Vol. 33, no. 12. Pp. 6999–7019. DOI: 10.1109/TNNLS.2021.3084827

**21. Ballas N.** Delving deeper into convolutional networks for learning video representations / N. Ballas, L. Yao, C.J. Pal, A. Courville [Электронный ресурс] // 4th International Conference on Learning Representations (ICLR 2016), 2016. 2 р. DOI: 10.48550/arXiv.1511. 06432 (дата обращения: 08.10.2024).

**22. Mahafza B.R.** Radar systems analysis and design using MATLAB. 2nd ed. Chapman and Hall, CRC, 2005. 638 p. DOI: 10.1201/978 1420057072

**23.** Van Dyk D.A., Meng X.L. The art of data augmentation // Journal of Computational and Graphical Statistics. 2001. Vol. 10, no. 1. Pp. 1–50. DOI: 10.1198/10618600152418584

**24.** Masters D., Luschi C. Revisiting small batch training for deep neural networks [Электронный ресурс] // Computer Science and Machine Learning. 2018. Pp. 1–18. DOI: 10.48550/arXiv.1804.07612 (дата обращения: 08.10.2024).

**25.** Werbos P.J. Backpropagation through time: what it does and how to do it // Proceedings of the IEEE. 1990. Vol.78, no. 10. Pp. 1550–1560. DOI: 10.1109/5.58337

**26.** Llugsi R. Comparison between Adam, AdaMax and Adam W optimizers to implement a weather forecast based on neural networks for the Andean city of Quito / R. Llugsi, S.E. Yacoubi, A. Fontaine, P. Lupera // 2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM), 2021. Pp. 1–6. DOI: 10.1109/ETCM 53643.2021.9590681

**27. Bejani M.M., Ghatee M.** A systematic review on overfitting control in shallow and deep neural networks // Artificial Intelligence Review. 2021. Vol. 54. Pp. 6391–6438. DOI: 10.1007/s10462-021-09975-1

**28.** Passos D., Mishra P. A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks [Электронный ресурс] // Chemometrics and Intelligent Laboratory Systems. 2022. Vol. 223. ID: 104520. DOI: 10.1016/j.chemolab.2022.10 4520 (дата обращения: 08.10.2024).

**29.** Koloskova A., Hendrikx H., Stich S.U. Revisiting gradient clipping: Stochastic bias and tight convergence guarantees // International Conference on Machine Learning. 2023. Pp. 17343–17363. DOI: 10.48550/arXiv.2305. 01588

**30.** Endsley M.R. Toward a theory of situation awareness in dynamic systems // Human factors. 1995. Vol. 37, no. 1. Pp. 32–64. DOI: 10.1518/001872095779049543

**31. Kaikkonen L.** Bayesian networks in environmental risk assessment: A review / L. Kaikkonen, T. Parviainen, M. Rahikainen, L. Uusitalo, A. Lehikoinen // Integrated environmental assessment and management. 2021. Vol. 17, no. 1. Pp. 62–78. DOI: 10.1002/ieam.4332

**32. Kovalenko G.V., Yadrov I.A., Kuts K.A.** Intelligent adaptive flight crew decision support system for thunderstorm avoidance // Russian Aeronautics. 2023. Vol. 66. Pp. 552–559. DOI: 10.3103/S1068799823030170

## References

**1.** Bolstad, C.A., Riley, J.M. (2002). Using goal directed task analysis with Army brigade officer teams. *In: Proceedings of the Human Factors and Ergonomics Society Annual* 

*Meeting. SAGE Publications*, vol. 46, no. 3, pp. 472–476. DOI: 10.1177/154193120204600354

**2.** Stanton, N.A., Chambers, P.R.G., Piggott, J. (2001). Situational awareness and safety. *Safety science*, vol. 39, no. 3, pp. 189–204. DOI: 10.1016/S0925-7535(01)00010-8

**3.** Sarter, N.B., Woods, D.D. (1991). Situation awareness: A critical but Ill-defined phenomenon. *The International Journal of Aviation Psychology*, vol. 1, no. 1, pp. 45–57. DOI: 10.1207/s15327108ijap0101 4

**4. Strelkov**, **Yu.K.** (2001). Engineering and professional psychology: textbook. Moscow: Vysshaya shkola; Akademiya, 360 p. (in Russian)

**5.** De Gooijer, J.G., Hyndman, R.J. (2006). 25 years of time series forecasting. *International Journal of Forecasting*, vol. 22, no. 3, pp. 443–473. DOI: 10.1016/j.ijforecast. 2006.01.001

6. Stevenson, S. (2007). A comparison of the forecasting ability of ARIMA models. *Journal of Property Investment & Finance*, vol. 25, no. 3, pp. 223–240. DOI: 10.1108/14635780 710746902

7. Chatfield, C., Koehler, A.B., Ord, J.K., Snyder, R.D. (2001). A new look at models for exponential smoothing. *Journal of the Royal Statistical Society: Series D (The Statistician)*, vol. 50, no. 2, pp. 147–159. DOI: 10.1111/1467-9884.00267

8. Bentéjac, C., Csörgő, A., Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, vol. 54, pp. 1937–1967. DOI: 10.1007/s10462-020-09896-5

9. Faloutsos, C., Gasthaus, J., Januschowski, T., Wang, Y. (2019). Classical and contemporary approaches to big time series forecasting. *In: SIGMOD '19: Proceedings of the* 2019 International Conference on Management of Data, pp. 2042–2047. DOI: 10.1145/3299 869.3314033

**10.** Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, vol. 34, issue 4, pp. 802–808. DOI: 10.1016/j.ijfore cast.2018.06.001 11. Taieb, S.B., Sorjamaa, A., Bontempi, G. (2010). Multiple-output modeling for multi-step-ahead time series forecasting. *Neurocomputing*, vol. 73, issue 10–12, pp. 1950–1957. DOI: 10.1016/j.neucom.2009.11.030

**12.** Sutskever, I., Vinyals, O., Quoc, V.L. (2014). Sequence to sequence learning with neural networks. *In: Proceedings of the 27th International Conference on Neural Information Processing Systems*, no. 2, pp. 3104–3112. DOI: 10.48550/arXiv.1409.3215

**13.** Caterini, A.L., Chang, D.E., Caterini, A.L., Chang, D.E. (2018). Recurrent neural networks. *In book: Deep Neural Networks in a Mathematical Framework*, Springer Briefs in Computer Science. Springer, Cham, pp. 59–79. DOI: 10.1007/978-3-319-75304-1 5

14. Rumelhart, D.E., Hinton, G.E., Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, no. 323, pp. 533–536. DOI: 10.1038/323533a0

**15.** Rasamoelina, A.D., Adjailia, F., Sinčák, P. (2020). A review of activation function for artificial neural network. *In: 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Herlany, Slovakia, pp. 281–286. DOI: 10.1109/SAMI48 414.2020.9108717

16. Toharudin, T., Pontoh, R.S., Caraka, R.E., Zahroh, S., Lee, Y., Chen, R.C. (2023). Employing long short-term memory and Facebook prophet model in air temperature forecasting. *Communications in Statistics-Simulation and Computation*, vol. 52, issue 2, pp. 279–290. DOI: 10.1080/03610918.2020. 1854302

**17. Schmidhuber, J., Hochreiter, S.** (1997). Long short-term memory. *Neural Computation*, vol. 9, issue 8, pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735

18. De Mulder, W., Bethard, S., Moens, M. F. (2015). A survey on the application of recurrent neural networks to statistical language modeling. *Computer Speech & Language*, vol. 30, issue 1, pp. 61–98. DOI: 10.1016/ j.csl.2014.09.005

**19.** Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K., Woo, W.C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *In:*  NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems, vol. 1, pp. 802–810. DOI: 10.48550/ arXiv.1506.04214

**20.** Li, Z., Liu, F., Yang, W., Peng, S., Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *In: IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019. DOI: 10.1109/TNNLS.2021. 3084827

**21.** Ballas, N., Yao, L., Pal, C.J., Courville, A. (2016). Delving deeper into convolutional networks for learning video representations. *In: 4th International Conference on Learning Representations*, 2 p. DOI: 10.48550/arXiv.1511.06432 (accessed: 08.10.2024).

**22. Mahafza, B.R.** (2005). Radar systems analysis and design using MATLAB, 2nd ed. Chapman and Hall, CRC, 638 p. DOI: 10.1201/9781420057072

**23.** Van Dyk, D.A., Meng, X.L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1–50. DOI: 10.1198/10618600152 418584

**24.** Masters, D., Luschi, C. (2018). Revisiting small batch training for deep neural networks. *Computer Science and Machine Learning*, pp. 1–18. DOI: 10.48550/arXiv. 1804.07612 (accessed: 08.10.2024).

**25.** Werbos, P.J. (1990). Backpropagation through time: what it does and how to do it. *In: Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560. DOI: 10.1109/5.58337

**26.** Llugsi, R., Yacoubi, S.E., Fontaine, A., Lupera, P. (2021). Comparison between Adam, AdaMax and Adam W optimizers to implement a weather forecast based on neural networks for the Andean city of Quito. *In: 2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, pp. 1–6. DOI: 10.1109/ETCM53643. 2021.9590681

**27. Bejani, M.M., Ghatee, M.** (2021). A systematic review on overfitting control in shallow and deep neural networks. *Artificial Intelligence Review*, vol. 54. Pp. 6391–6438. DOI: 10.1007/s10462-021-09975-1

28. Passos, D., Mishra, P. (2022). A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemometrics and Intelligent Laboratory Systems*, no. 223. ID: 104520. DOI: 10.1016/j.chemolab.2022.104520 (accessed: 08.10.2024).

**29.** Koloskova, A., Hendrikx, H., Stich, S.U. (2023). Revisiting gradient clipping: Stochastic bias and tight convergence guarantees. *International Conference on Machine Learning*, pp. 17343–17363. DOI: 10.48550/arXiv.2305. 01588

**30.** Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Hu*-

*man factors*, vol. 37, no. 1, pp. 32–64. DOI: 10.1518/001872095779049543

**31. Kaikkonen, L., Parviainen, T., Rahikainen, M., Uusitalo, L., Lehikoinen, A.** (2021). Bayesian networks in environmental risk assessment: A review. *Integrated environmental assessment and management*, vol. 17, no. 1, pp. 62–78. DOI: 10.1002/ieam.4332

**32. Kovalenko, G.V., Yadrov, I.A., Kuts, K.A.** (2023). Intelligent adaptive flight crew decision support system for thunderstorm avoid-ance. *Russian Aeronautics*, vol. 66. Pp. 552–559. DOI: 10.3103/S1068799823030170

### Сведения об авторах

Коваленко Геннадий Владимирович, доктор технических наук, профессор, профессор кафедры летной эксплуатации и безопасности полетов в гражданской авиации СПбГУ ГА, kgvf@inbox.ru.

**Ядров Илья Александрович,** аспирант кафедры летной эксплуатации и безопасности полетов в гражданской авиации СПбГУ ГА, yadrov.ilya@gmail.com.

### Information about the authors

**Gennadiy V. Kovalenko**, Doctor of Technical Sciences, Professor, Professor of the Chair of Flight Operations and Flight Safety in Civil Aviation, Saint Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A.A. Novikov, kgvf@inbox.ru.

**Ilya A. Yadrov,** Postgraduate Student of the Chair of Flight Operations and Flight Safety in Civil Aviation, Saint Petersburg State University of Civil Aviation named after Chief Marshal of Aviation A.A. Novikov, yadrov.ilya@gmail.com.

Поступила в редакцию	14.10.2024	Received	14.10.2024
Одобрена после рецензирования	06.11.2024	Approved after reviewing	06.11.2024
Принята в печать	23.01.2025	Accepted for publication	23.01.2025