

Научная статья

УДК 629.7.051.83

URL: <https://trudymai.ru/published.php?ID=182680>

ФОРМИРОВАНИЕ ВЫБОРКИ ПОЛЕТОВ И МАССИВОВ ДАННЫХ ДЛЯ ИССЛЕДОВАНИЯ ПОСАДКИ САМОЛЕТА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

**Иван Александрович Петров¹, Надежда Алексеевна Прошкина²,
Владимир Викторович Стрелков³✉**

^{1,3}Центральный аэрогидродинамический институт имени профессора
Н.Е. Жуковского (ЦАГИ), Жуковский, Московская область, Россия

²Московский физико-технический институт (МФТИ), Долгопрудный, Московская
область, Россия

¹peredreyhus1@mail.ru

²proshkina.na@phystech.edu

³v_strelkov@tsagi.ru ✉

Аннотация. Потенциально доступные объемы информации о полете и тенденция их многократного увеличения в ближайшие годы говорят об актуальности и важности развития новых подходов к исследованию проблем обеспечения безопасности полетов, которые базируются на анализе информации с использованием методов машинного обучения (МО). Достоверность результатов такого анализа в значительной степени определяется качеством предобработки первичной полетной информации и подготовленных массивов для непосредственного применения

методов МО. Массивы данных формируются под конкретную прикладную задачу, которую предполагается решать.

В настоящей работе рассматриваются вопросы формирования выборки полетов и массивов данных для их последующего анализа с использованием методов МО применительно к задаче прогнозирования координаты точки приземления самолета на взлетно-посадочной полосе (ВПП). Весь процесс подготовки данных разделен на отдельные этапы, каждый из которых подробно обсуждается и иллюстрируется результатами обработки реальных полетных данных.

В работе использованы записи параметров полетов пассажирского самолета в условиях эксплуатации на маршрутной сети одной из авиакомпаний, архивы METAR, справочные данные об аэродромах посадки и номинальных технических характеристиках воздушного судна.

Показано, что работа по формированию выборки полетов и массивов данных для исследования посадки самолета с использованием методов МО является очень трудоемкой, но необходимой частью в рамках реализации перспективного подхода к исследованию проблем безопасности полетов.

Ключевые слова: безопасность полетов, машинное обучение, прогнозирование, полетная информация, выборка полетов, массив данных, посадка, координата точки приземления

Для цитирования: Петров И.А., Прошкина Н.А., Стрелков В.В. Формирование выборки полетов и массивов данных для исследования посадки самолета с использованием методов машинного обучения // Труды МАИ. 2024. № 138. URL:

<https://trudymai.ru/published.php?ID=182680>

Original article

GENERATION OF FLIGHT SAMPLES AND DATA SETS FOR AIRCRAFT LANDING STUDIES USING MACHINE LEARNING TECHNIQUES

Ivan A. Petrov¹, Nadezhda A. Proshkina², Vladimir V. Strelkov³✉

^{1,3}Central Aerohydrodynamic Institute named after Professor N.E. Zhukovsky (TsAGI), Zhukovsky, Moscow Region, Russia

²Moscow Institute of Physics and Technology (MIPT), Dolgoprudny, Moscow Region, Russia

¹peredreyhus1@mail.ru

²proshkina.na@phystech.edu

³v_strelkov@tsagi.ru✉

Abstract. Potentially available volumes of flight data and the trend of their multiple increase in the coming years indicate the relevance and importance of the development of new approaches to the study of flight safety problems, which are based on the analysis of information using machine learning (ML) methods. The reliability of the results of such analysis is largely determined by the quality of preprocessing of primary flight data and prepared arrays (datasets) for the direct application of ML methods. Data arrays are formed for a specific applied problem to be solved.

This paper deals with the formation of flight samples and data arrays for their subsequent analysis using ML methods as applied to the task of predicting the coordinate of an aircraft touchdown point on a runway.

The whole process of data preparation is divided into the following separate stages:

- collection, transcription and annotation of flight parametric recorder data;

- formation of flight sample;
- calculation of non-registered parameters;
- alignment of records of different flights;
- validation of recorded flight data;
- data consolidation and synchronization;
- preliminary (exploratory) data analysis.

Each of the steps is discussed in detail and illustrated by the results of real flight data processing.

The paper uses records of passenger aircraft flight parameters under operating conditions on the route network of one of the airlines, METAR archives, reference data on landing airfields and nominal technical characteristics of the aircraft.

It is shown that the work on the formation of a sample of flights and data arrays for the study of aircraft landing using machine learning methods is a very labor-intensive, but necessary part of the implementation of a promising approach to the study of flight safety problems.

Keywords: flight safety, machine learning, prediction, flight data, flight sampling, dataset landing, touchdown point coordinate

For citation: Petrov I.A., Proshkina N.A., Strelkov V.V. Generation of flight samples and data sets for aircraft landing studies using machine learning techniques. *Trudy MAI*, 2024, no. 138. URL: <https://trudymai.ru/eng/published.php?ID=182680>

Введение

Значительное расширение потоков разнородной информации о полете и доступа к ней открывает новые перспективы при решении прикладных задач с использованием методов машинного обучения (МО), в частности в области обеспечения безопасности полетов на посадке. Решению проблемы грубых посадок и выкатываний самолета с ВПП посвящено огромное количество публикаций отечественных и зарубежных авторов [1÷7].

Существуют различные вычислительные методы (методы МО), которые позволяют на основании обработки информации о ранее выполненных полетах построить алгоритмы выявления аномалий в текущей конкретной полетной ситуации, связать их с вероятными последствиями и спрогнозировать развитие ситуации, что делает их весьма актуальными для решения задач обеспечения безопасности полетов, [8÷12]. Речь идет не просто о доступной оцифрованной информации, имеющей отношение к исследуемой проблеме (например, безопасность посадки самолета), но о специально предобработанных и структурированных массивах данных – датасет (dataset). Подготовка массивов данных для реализации обозначенного подхода к решению задач безопасности полетов с применением методов МО является обязательной и очень значимой частью работы. Датасет формируется под решение конкретной задачи (или круга задач) и с учетом методов МО, которые предполагается использовать. Например, в работе [13] описывается датасет траекторий для исследования уходов на 2-ой круг.

В настоящей работе рассматривается процесс подготовки массивов данных на примере задачи прогнозирования координаты точки касания самолетом поверхности

ВПП на посадке. От координаты точки приземления зависит и координата точки остановки самолета после пробега, и, следовательно, необходимая длина ВПП в конкретных условиях посадки.

Задача формулируется следующим образом: необходимо на режиме снижения самолета по глиссаде и при выравнивании перед приземлением в каждый момент времени (для каждой дальности до торца ВПП - X) определить ожидаемое значение координаты точки касания ВПП стойками шасси ($X_{кас}$) в зависимости от текущих параметров полета:

$$X_{кас}(X) = F(\Delta H, \Delta Z, \Delta V_{пр}, СКО-\gamma, СКО-\vartheta, \dots), \text{ где:}$$

$\Delta H, \Delta Z$ – отклонение самолета от номинальной траектории посадки в вертикальной и горизонтальной плоскостях;

$\Delta V_{пр}$ – отклонение текущей приборной скорости от заданной скорости захода на посадку;

$СКО-\gamma, СКО-\vartheta$ – среднеквадратичное отклонение углов крена и тангажа на определенном временном интервале, отнесенное к текущему моменту времени;

Первоначально перечень параметров, от которых может зависеть координата точки приземления, вообще говоря, априори неизвестен и формируется на основании предварительной экспертной оценки, а затем уточняется по результатам анализа. Также можно было бы включить в рассмотрение и другие параметры, характеризующие условия захода на посадку, которые потенциально оказывают влияние на результат, если они доступны. Например, представляется логичным включить в первоначальный перечень такие параметры как продольная и боковая

составляющие скорости ветра и др. В контексте данной работы перечень параметров, определяющих координату точки приземления, детально не обсуждается, поскольку целью работы является рассмотрение порядка формирования выборки полетов и массивов данных для применения методов МО.

В качестве метода прогнозирования на первом этапе предполагается использовать линейную регрессию (или более сложные ее варианты), нейронную сеть или деревья решений. Таким образом, массив данных для изложенной выше постановки задачи должен включать в себя значения исследуемых параметров полета в выбранных сечениях (при заданных дальностях до торца ВПП), координаты точки приземления для каждого полета, вес самолета и конфигурацию закрылков, угол наклона посадочной глиссады, метеоусловия посадки. Доступными источниками информации, имеющей отношение к решаемой задаче, являются записи бортового параметрического регистратора, архивы погоды, справочник аэронавигационной информации.

Подготовку массивов данных можно разделить на следующие этапы:

- Сбор, расшифровка и аннотирование записей бортового параметрического регистратора;
- Формирование выборки полетов;
- Вычисление нерегистрируемых параметров;
- Выравнивание записей разных полетов;
- Валидация зарегистрированных полетных данных;
- Консолидация и синхронизация данных;

- Предварительный (разведочный) анализ данных.

Ниже детально рассматриваются все перечисленные этапы подготовки данных применительно к решаемой задаче.

Сбор, расшифровка и аннотирование записей бортового регистратора

Сбор и расшифровка полетных данных включает в себя считывание с бортового параметрического регистратора файла первичных полетных данных, в котором содержатся записи параметров, как правило, нескольких полетов, наземных отработок самолетных систем (если такие были в период между моментами считывания информации с борта) и др. При расшифровке файла бортового регистратора двоичный код преобразуется в физические величины, исходный файл разделяется на файлы записей отдельных полетов и из каждого файла извлекается информация для формирования его уникального имени, под которым он сохраняется в хранилище данных. Расшифровка выполняется с помощью специального программного обеспечения (ПО), в данном случае было использовано ПО «ARINC-конвертор», [14].

Имя файла записи отдельного полета фактически является аннотацией содержимого файла. Аннотирование (или разметка) данных позволяет однозначно идентифицировать полет среди множества всех имеющихся в хранилище данных. В нашем случае аннотация файла в доступной базе полетов содержит в себе следующую информацию:

- Номер полета в выборке,
- Тип бортового регистратора,

- Дата взлета,
- Время взлета,
- Дата посадки,
- Время посадки,
- Номер борта,
- Номер рейса,
- Код аэропорта взлета;
- Код аэропорта посадки.

Формирование выборки полетов

Применительно к рассматриваемой задаче (прогнозирование координаты точки приземления) формируется выборка из первичных полетных данных – файлов бортового параметрического регистратора. Выборка должна быть релевантной (соответствовать решаемой задаче), полной (репрезентативной) и однородной (все записи имеют одинаковую структуру), [15]. В настоящей работе доступное для проведения исследований хранилище данных содержит более 10 тысяч полетов пассажирского самолета одного типа с посадками в разных аэропортах. На первом этапе были отобраны только полеты с посадками в одном аэропорту (более 3300 полетов), чтобы исключить из рассмотрения такой параметр как угол наклона посадочной глиссады, который оказывает существенное влияние на продольную координату точки приземления. Репрезентативность выборки можно было бы повысить за счет объединения всех заходов на посадку в разных аэропортах с

одинаковым углом наклона посадочной глissады, но это потребовало бы дополнительного анализа условий захода на посадку в каждом из аэропортов.

Вычисление нерегистрируемых параметров

Не все параметры полета, которые являются предметом исследования, регистрируются на борту летательного аппарата. К числу таких параметров могут относиться как дополнительные временные ряды, например, положение ЛА в координатах ВПП, так и значения некоторых показателей в определенные моменты времени, например:

- координата точки приземления;
- отклонение самолета от номинальной траектории посадки в вертикальной и горизонтальной плоскостях на разных расстояниях до торца ВПП;
- отклонение скорости захода на посадку от заданной для фактического посадочного веса и конфигурации самолета;
- СКО- γ , СКО- ϑ и др. (зависит от детализации задачи).

Кроме того, некоторые новые параметры формируются как результат объединения двух однотипных параметров, регистрируемых со сдвигом по времени. Например, на бортовом регистраторе сигналы от двух датчиков отклонения самолета от равносигнальной зоны курсового радиомаяка (LocDev1 и LocDev2) фиксируются с частотой 1 Гц со сдвигом по времени друг относительно друга на 0.5 секунды. Объединение этих сигналов в один параметр позволяет повысить информативность нового параметра, который будет иметь вдвое большую частоту.

Перечень вычисляемых параметров определяется постановкой и детализацией задачи. Однако практически для всех задач, связанных с безопасностью посадки, требуется знать траекторию движения самолета относительно ВПП посадки. Восстановление гладкой траектории движения самолета в системе координат, привязанной к ВПП, является сложной задачей, решение которой зависит от имеющихся систем позиционирования самолета в пространстве, погрешностей измерения координат и характеристик бортового регистратора.

Среди регистрируемых на борту воздушного судна параметров полета присутствуют сигналы барометрической и радиометрической высот. При этом ни один из этих параметров нельзя рассматривать как координату самолета в вертикальной плоскости относительно ВПП.

Если с использованием сигнала радиовысоты при послеполетной обработке информации, как правило, не возникает заметных сложностей, то с баровысотой в отдельных случаях могут быть проблемы. Во-первых, регистрируемая баровысота может отсчитываться либо от уровня моря (QNH), либо относительно порога ВПП (QFE), либо от стандартного давления на уровне моря (QNE). Во-вторых, в перечне параметров может присутствовать несколько сигналов баровысоты, отличия между которыми не всегда описаны в документации на воздушное судно. В итоге оказывается, что регистрируемая на борту и доступная для последующей обработки барометрическая высота полета требует определенной коррекции. В настоящей работе речь идет именно о такой ситуации, которая сложилась при расшифровке доступной информации с бортовых регистраторов пассажирского самолета зарубежного производства.

На рисунке 1 приведены зарегистрированные значения баровысоты $H_{bar}(X)$ в зависимости от дальности до торца ВПП для 24-х посадок в одном из аэропортов Российской Федерации. Видно, что кривые заметно «расслоились» по высоте. Для привязки баровысоты к высоте аэродрома требуется соответствующая коррекция значений зарегистрированного параметра. Но даже приведенная к высоте аэродрома баровысота не может рассматриваться как координата самолета в вертикальной плоскости относительно земли из-за заметного искажения баровысоты вблизи подстилающей поверхности.

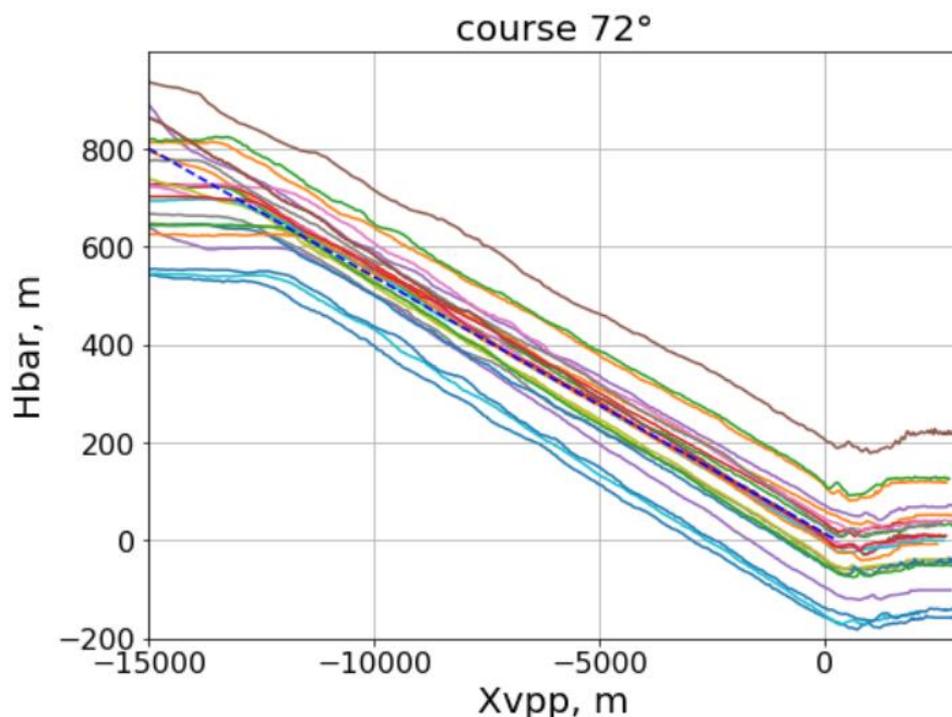


Рис. 1 – Зарегистрированные значения баровысоты в зависимости от дальности до торца ВПП (24 посадки)

Радиовысота успешно используется для анализа процесса выравнивания самолета перед приземлением. Но данный сигнал не пригоден для исследования режима снижения самолета по глиссаде, поскольку значение радиометрической высоты зависит от рельефа земной поверхности в зоне подхода к ВПП. На рисунке 2

приведены зависимости радиовысоты от дальности до ВПП для 24-х посадок в том же аэропорту. Характерное изменение всех кривых на удалении 5÷8 км от торца ВПП отражает рельеф земной поверхности на подходе к ВПП и никак не связано с отклонением самолета от глиссады.

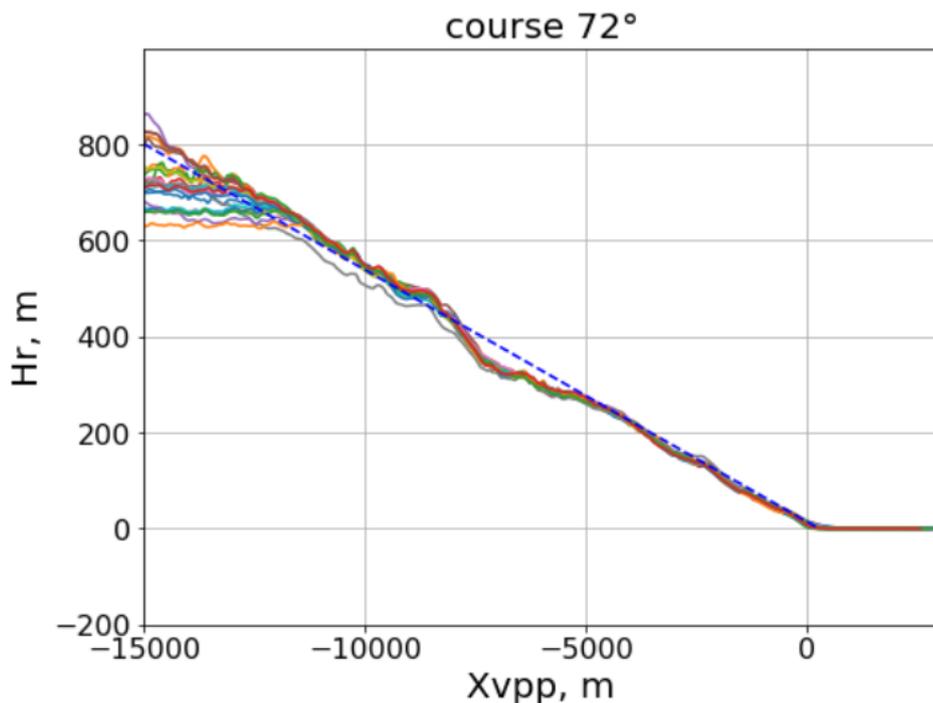


Рис. 2 – Зарегистрированные значения радиовысоты в зависимости от дальности до торца ВПП (24 посадки)

Для исследования режимов захода на посадку, включая выравнивание самолета перед приземлением, был сформирован новый параметр $H(X)$ (синтезированный сигнал высоты) из регистрируемой радиовысоты и скорректированной баровысоты, который определяется следующим образом:

$$H(X) = H_{\text{bar}_{\text{cor}}}(X), \text{ при } H_r(X) > 15 \text{ м};$$

$$H(X) = H_r(X), \text{ при } H_r(X) \leq 15 \text{ м}.$$

Коррекция баровысоты осуществляется таким образом, чтобы при $H_r = 15 \text{ м}$

$$H_{\text{bar}_{\text{cor}}} = H_r.$$

Данный параметр можно трактовать как траекторию движения колес шасси самолета в выпущенном положении и необжатом состоянии. На рисунке 3 приведены графики изменения барометрической, радиометрической и синтезированной высот.

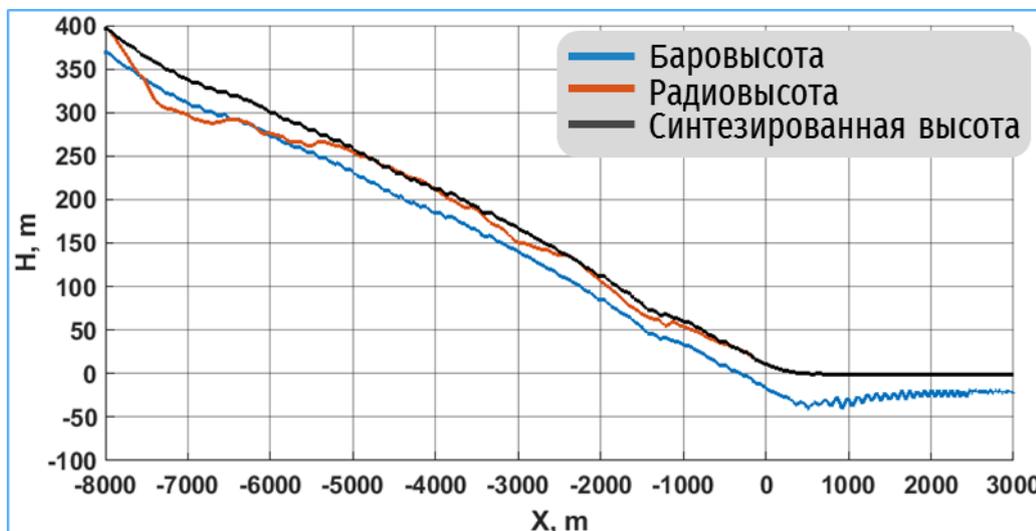


Рис. 3 – Баровысота, радиовысота и синтезированная траектория движения самолета в вертикальной плоскости

Существенно более сложной задачей является восстановление траектории движения самолета в горизонтальной плоскости. На борту рассматриваемого самолета с помощью системы спутниковой навигации с периодичностью 1 раз в секунду определяются геодезические координаты самолета: широта ψ_c и долгота λ_c . Для определения положения самолета в координатах ВПП выполняется пересчет зарегистрированных значений широты и долготы в дальность до входного порога и боковое отклонение относительно оси ВПП, [16]. При этом используется информация о геодезических координатах соответствующего торца ВПП $\psi_{впп}$, $\lambda_{впп}$ и курс (ориентация) полосы $\Psi_{впп}$.

$$X = F1(\psi_c, \lambda_c, \psi_{впп}, \lambda_{впп}, \Psi_{впп})$$

$$Z = F2(\psi_c, \lambda_c, \psi_{впп}, \lambda_{впп}, \Psi_{впп})$$

Реальная погрешность измерения координат на поверхности земли с помощью спутниковой системы позиционирования зависит от многих факторов, но она не лучше $\pm 10 \div 25$ метров. То есть фактическая траектория самолета в координатах $Z(X)$ лежит в коридоре шириной от 20 до 50 метров. В ряде случаев ситуация еще более осложняется из-за характеристик регистратора, а именно – количества бит, отведенных для регистрации сигналов широты и долготы. Например, если в 32-х битном слове разрешение по широте и долготе (цена младшего разряда) в пересчете на метры оценивается величиной 1 см, то в 18-и битном слове дискретность сигнала будет равна 152 метрам. Очевидно, что при таком высоком уровне неопределенности координат невозможно их использовать для анализа траектории без специальной обработки.

Существуют технологии восстановления гладкой траектории движения самолета в горизонтальной плоскости [17, 18]. В настоящей работе были использованы методика и программное обеспечение, описанные в источниках [18, 19]. Окончательная коррекция траектории выполняется по показаниям курсового радиомаяка (параметр `LocDev`), если посадка осуществляется на аэродром, оборудованный соответствующей наземной радиотехнической системой посадки.

На рисунке 4 показаны результаты всех трех этапов построения траектории самолета в горизонтальной плоскости: пересчет широты и долготы в координаты относительно ВПП, восстановление гладкой траектории, коррекция траектории.

Наличие скорректированной гладкой траектории самолета в горизонтальной плоскости позволяет определить координаты точки приземления самолета на ВПП.

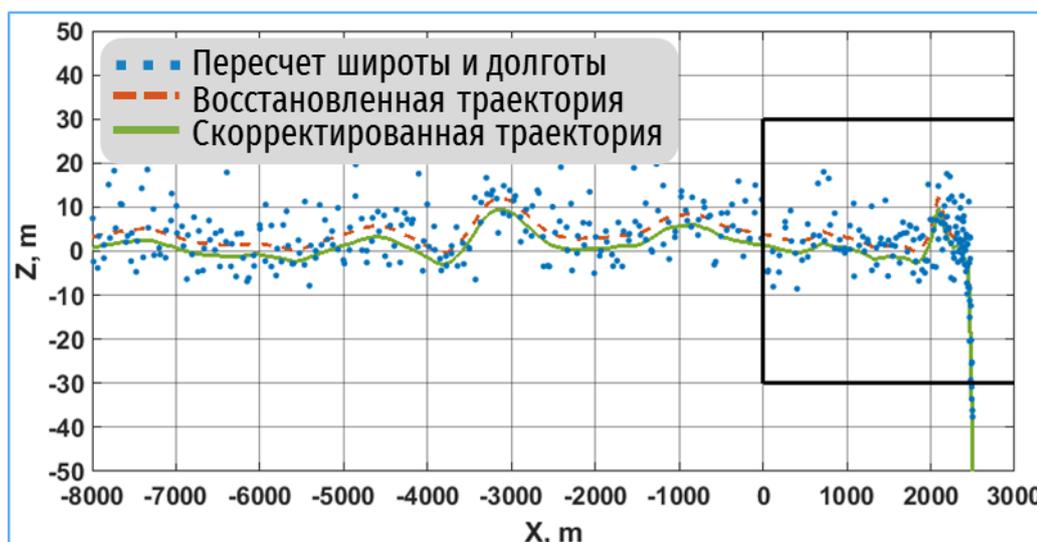
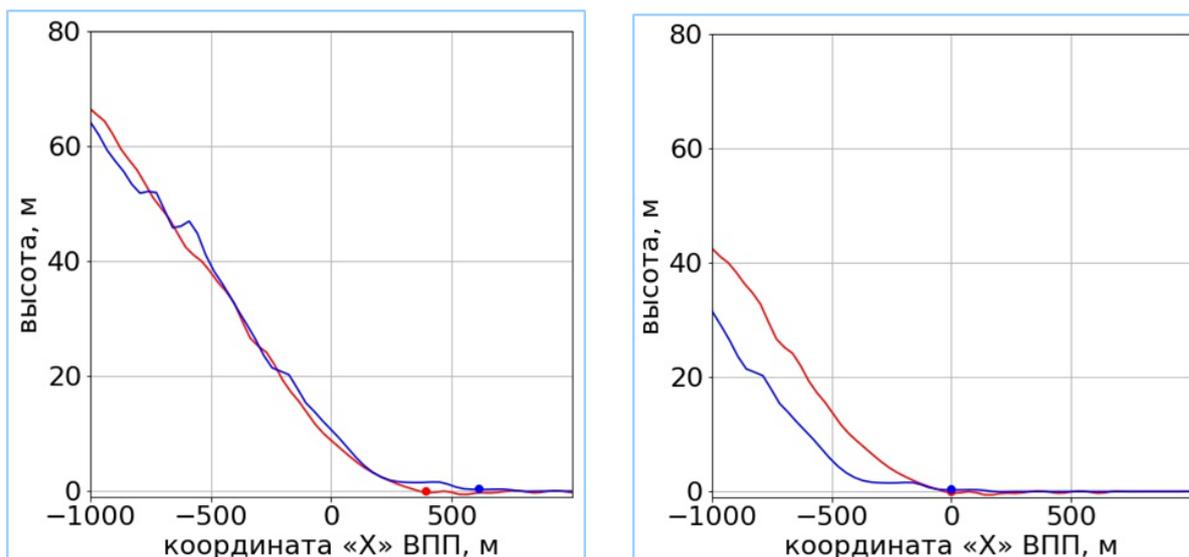


Рис. 4 – Зарегистрированная, восстановленная и скорректированная траектории движения самолета в горизонтальной плоскости

Выравнивание записей параметров полета

Все полеты имеют разную продолжительность и запись каждого из полетов осуществляется в «своем» времени. Как правило, время отсчитывается от момента включения бортового регистратора или от начала файла, если речь идет о файлах отдельных полетов. Чтобы сформировать массивы данных для применения методов машинного обучения, необходимо предварительно записи всех полетов выровнять друг относительно друга.

Алгоритмы выравнивания записей полетов друг относительно друга зависят от решаемой задачи. В задаче прогнозирования координаты точки приземления записи разных полетов целесообразно выравнивать по моменту времени пересечения самолетом торца ВПП (по дальности до торца ВПП), рис. 5а, а для исследования послепосадочного пробега выравнивание должно производиться по моменту приземления, рис. 5б.



а) Выравнивание по моменту
пролета торца ВПП

б) Выравнивание по моменту
касания ВПП

Рис. 5 – Выравнивание записей полетов

При выравнивании записей полетов встает и вопрос вырезания из всех полетов однотипных участков для формирования массива данных. Необходимо определить единые для всех полетов критерии начала и конца вырезаемого участка полета. В зависимости от решаемой задачи это может быть временной интервал относительно некоторого момента времени (например, момента обжатия стоек шасси), или диапазон рассматриваемых дальностей полета, или какой-то другой критерий. Увеличение длины исследуемого участка полета приводит к сокращению количества полетов, которые могут быть включены в выборку. На рисунке 6 приведены траектории захода на посадку в горизонтальной плоскости. Очевидно, что траектории посадки с заходом «по коробочке» должны быть исключены из выборки, если исследуемый диапазон начинается с дальности до торца ВПП $> 4000 \div 5000$ метров. В противном случае эти полеты будут вносить ошибку, например, в распределение по боковому отклонению.

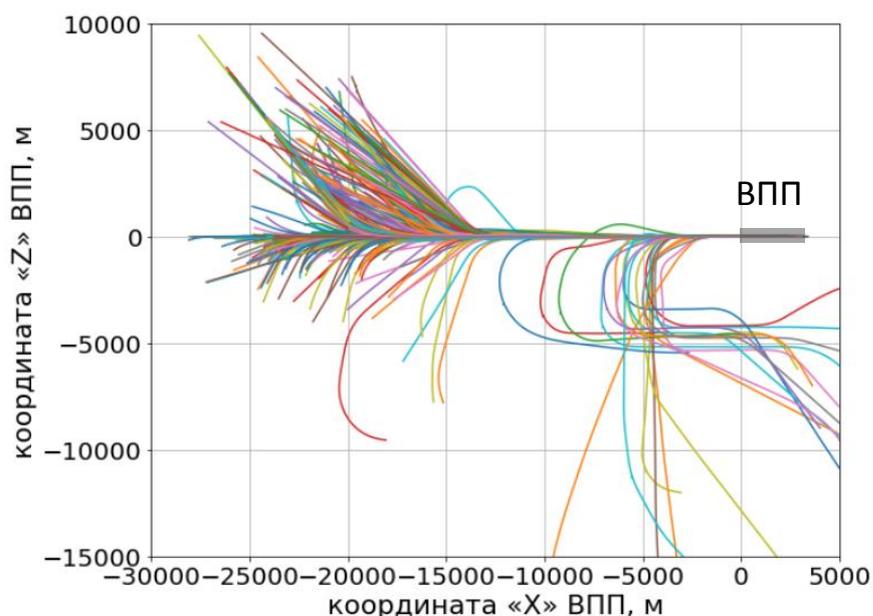
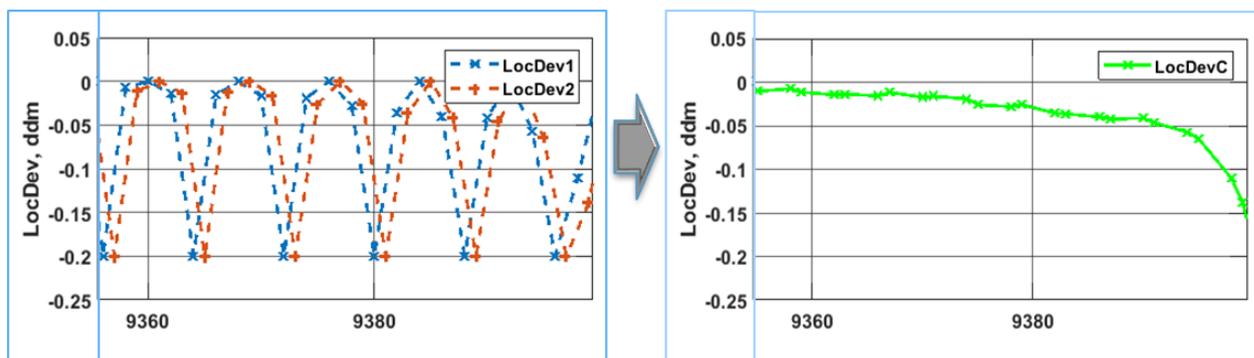


Рис. 6 – Траектории захода на посадку в горизонтальной плоскости

Валидация записей параметров полета

На следующем этапе данные проверяются на предмет наличия в выборке некорректных значений параметров, в том числе из-за сбоев записи при регистрации (валидация данных). Полеты с такими признаками или отбраковываются, или поврежденные данные исправляются. На рисунке 8 приведен пример «ремонта» зарегистрированных значений параметра $LocDev$, характеризующего отклонение самолета от равносигнальной зоны курсового радиомаяка. Параметр $LocDev$ сформирован из сигналов двух датчиков ($LocDev1$ и $LocDev2$), которые опрашиваются в разное время. В данном случае было обнаружено, что все невалидные точки имеют значения $LocDev=0$ или $LocDev=-0.2$, рис. 7а. Эти точки были отбракованы, в результате был восстановлен требуемый для анализа сигнал, рис. 7б.



a) Неважные сигналы двух датчиков LocDev

b) Восстановленный сигнал LocDev

Рис. 7 – Исправление неважного сигнала отклонения самолета от равносигнальной зоны курсового радиомаяка (LocDev)

Консолидация и синхронизация данных

Далее информация с бортового регистратора дополняется данными об аэродроме посадки (длина ВПП, угол наклона посадочной глиссады и др.), информацией о погоде и состоянии ВПП и характеристиками воздушного судна (конфигурации, скорость захода на посадку и др.). Источниками дополнительной информации являются справочники аэронавигационной информации, руководство по летной эксплуатации, архивы METAR, из которых извлекаются метеоданные, соответствующие времени посадки. Таким образом осуществляется консолидация и синхронизация данных.

Предварительный (разведочный) анализ данных

Завершающим шагом в подготовке данных для их последующего глубокого анализа с использованием методов МО является их «визуальный» и статистический анализ. Основная цель такого анализа – дополнительно проверить валидность данных

в выборке и оценить характерный диапазон и тип распределения данных.

Пучки траекторий всех отобранных полетов дают наглядное представление относительно вида коридора снижения самолета по глиссаде. На рисунках 8 и 9 приведены отобранные для дальнейшего анализа траектории движения самолета в вертикальной и горизонтальной плоскостях на этапе снижения по глиссаде и пробега по ВПП.

Кроме того, отрисовка на одном графике однотипных параметров из разных полетов позволяет дополнительно оценить валидность используемых данных, поэтому целесообразно проверять таким образом все исследуемые параметры. Следует отметить, что построение и анализ пучков траекторий используется и другими авторами при исследовании вопросов безопасности посадки, [20].

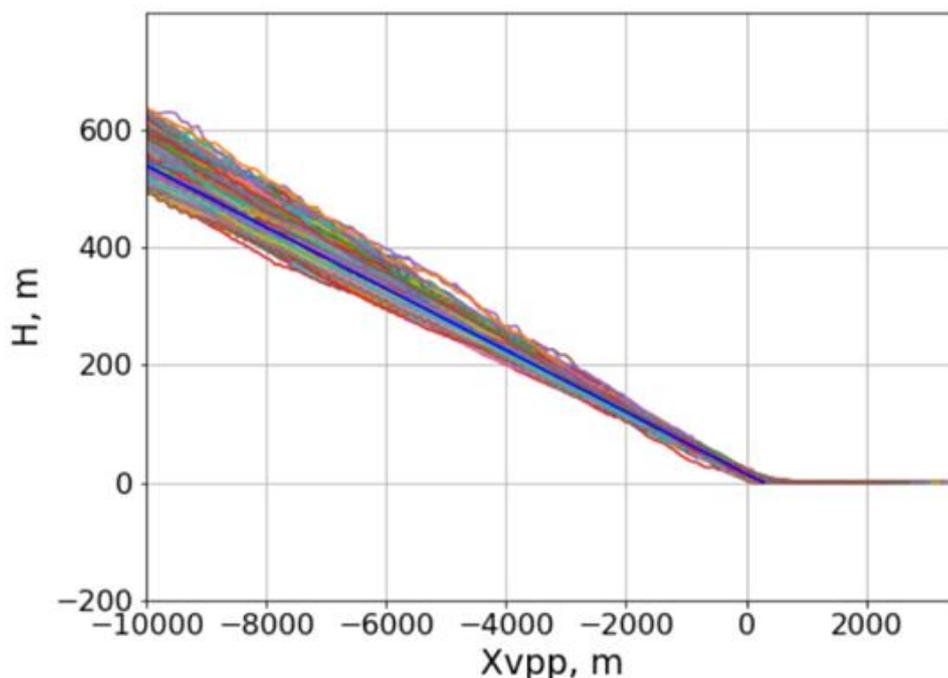


Рис. 8 – Траектории посадки самолетов в вертикальной плоскости

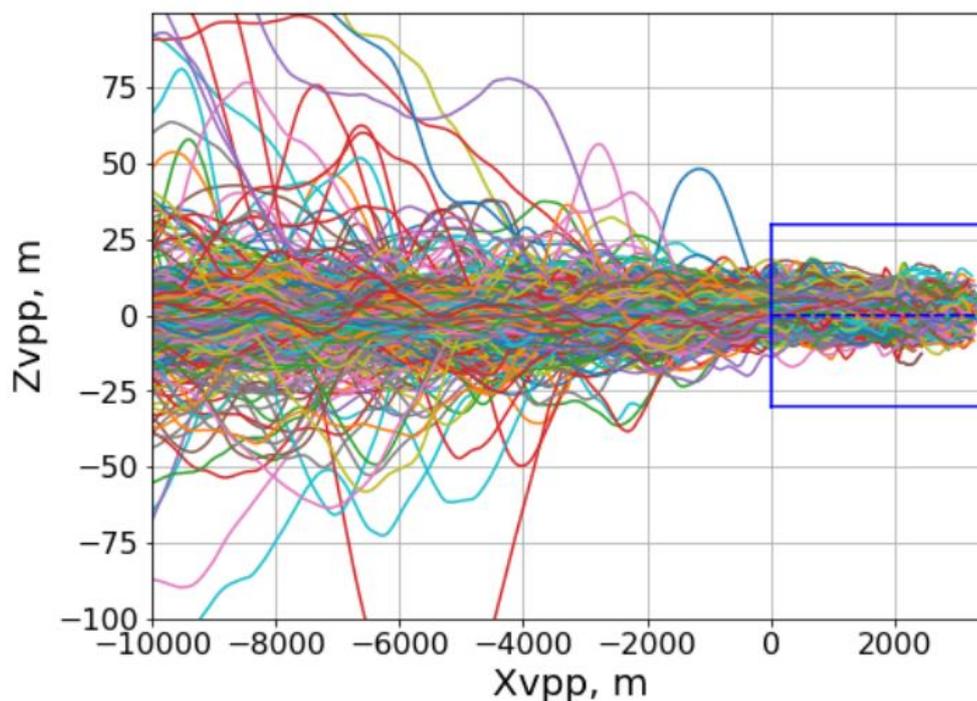


Рис. 9 – Траектории посадки самолетов в горизонтальной плоскости

Строятся также диаграммы рассеяния и гистограммы плотности распределения исследуемых параметров. В качестве примера на рисунке 10 приведена диаграмма рассеяния точек приземления самолета на ВПП примерно для 3000 посадок.

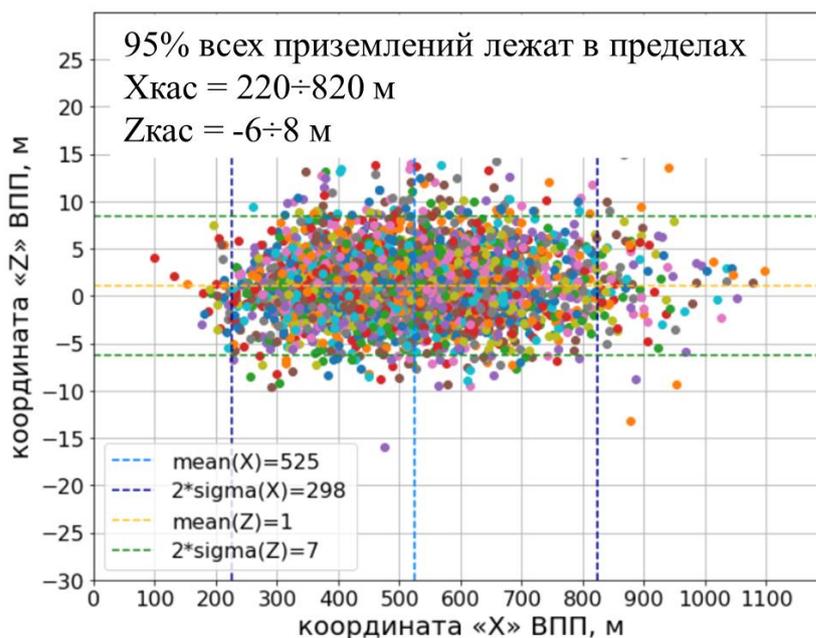


Рис. 10 – Диаграмма рассеяния точек приземления самолета на ВПП

Гистограммы распределения значений бокового отклонения на различных удалениях от торца ВПП (рис. 11) дают возможность увидеть некоторые закономерности, важные для понимания процесса пилотирования самолета на заключительном этапе посадки и пробеге по ВПП.

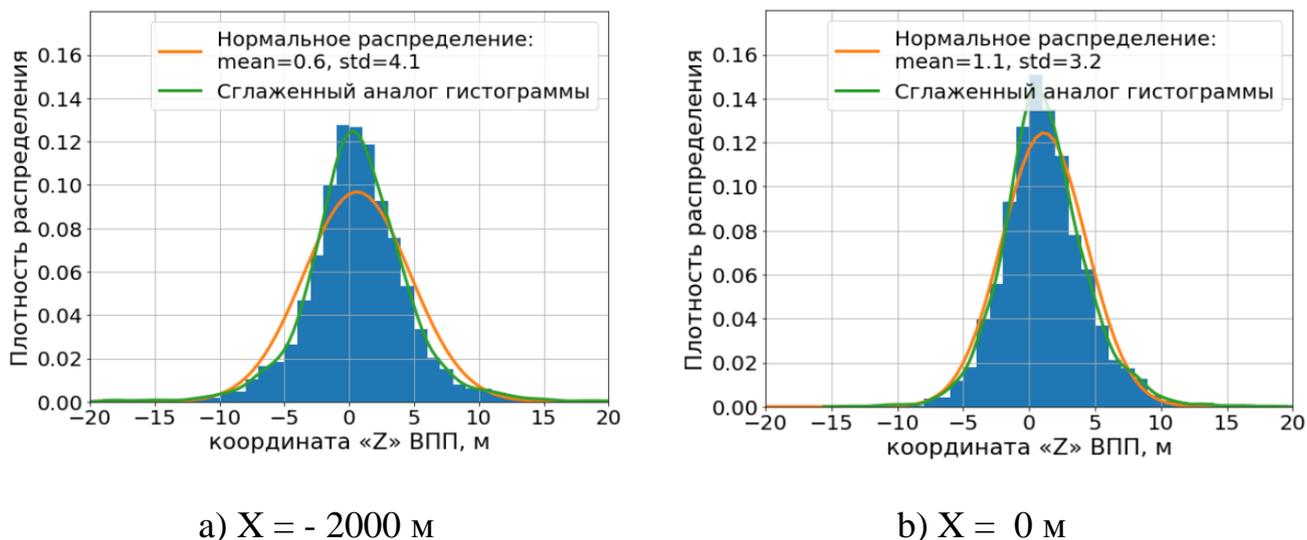


Рис. 11 – Гистограммы плотности распределения величины бокового отклонения траектории самолета на удалении 2000 метров и над торцом ВПП

Приведенные на рисунке 11 диаграммы для выборки примерно из 3000 траекторий посадки позволяют сделать следующие выводы:

- 1) вид распределения значений боковых отклонений близок к нормальному;
- 2) все траектории посадки лежат в достаточно узком диапазоне отклонений самолета от оси ВПП, который незначительно сужается по мере приближения к торцу ВПП.

Результаты обработки пучка траекторий самолета в горизонтальной плоскости приведены в таблице 1.

Таблица 1 – Характеристики распределения бокового отклонения в зависимости от дальности торца ВПП

Дальность до торца ВПП X, м	Среднее значение бокового отклонения μ, м	Среднеквадратичное боковое отклонение σ, м
-5000	0	5.6
-4000	0.3	5
-2000	0.6	4.1
0	1.1	3.2
300	1.2	3.3

Заключение

Применение методов МО в задачах безопасности полетов – это новый подход, который сегодня начинает активно разрабатываться и обещает не только уточнение результатов, полученных традиционными методами исследования, но в перспективе должен помочь вскрывать новые неявные факторы риска авиационного происшествия. Ключевым звеном для реализации данного подхода является информация, доступная для анализа, и технологии ее подготовки, которые сегодня практически не освещаются в публикациях.

Приведенное в настоящей работе описание выполненного авторами процесса поэтапной подготовки массивов данных, начиная с формирования выборки полетов и заканчивая разведочным анализом, дает наглядное представление о трудоемкости этого процесса. Вся работа базируется исключительно на реальных полетных данных

(файлы бортового параметрического регистратора пассажирского самолета, архивы погоды, необходимая информация о воздушном судне и аэропортах посадки из соответствующих справочников) и ориентирована на решение актуальных задач безопасности полетов. Важно отметить, что подготовка данных должна выполняться применительно к решению конкретной задачи.

Показано, что работа по формированию выборки полетов и массивов данных является очень трудоемкой, но необходимой, поскольку достоверность результатов применения методов МО в прикладных задачах определяется качеством подготовки массивов данных.

Список источников

1. G. van Es. A Study of Runway Excursions from a European Perspective. NLR-CR-2010-259. URL: <https://skybrary.aero/sites/default/files/bookshelf/2069.pdf>
2. Jenkins M., Aaron R.F. Reducing Runway Landing Overruns. Boeing Aero Magazine. URL: https://www.boeing.com/commercial/aeromagazine/articles/2012_q3/3/
3. Шаров В.Д. Прогнозирование и предотвращение выкатываний самолета за пределы взлетно-посадочной полосы. - Германия: Lambert Academic Publishing, 2013. С. 112.
4. Рыбин А.В. Разработка методики детектирования и анализа грубых посадок самолёта на основе численного моделирования происшествия // Труды МАИ. 2015. № 81. URL: <https://trudymai.ru/published.php?ID=57761>
5. Еремин А.И., Сельвесюк Н.И. Уточнённая оценка коэффициента опасности посадки при снижении по глиссаде с учётом действия вертикального ветра // Труды

МАИ. 2018. № 100. URL: <https://trudymai.ru/published.php?ID=93451>

6. Семаков С.Л. Об одном подходе к вероятностной оценке безопасности посадки гражданского самолета // Труды МАИ. 2019. № 104. URL:

<https://trudymai.ru/published.php?ID=102188>

7. Barry D., Greeves M., Solis T., Angel M. Advanced methods for analysis flight data for runway excursions risk factors. Cranfield University, Airbus Defence&Space. Future Sky Safety. Project: Solutions for Runway Excursions, 2017. URL: [https://www.futuresky-](https://www.futuresky-safety.eu/wp-content/uploads/2018/03/FSS_P3_CU_D3.5_v2.0.pdf/)

[safety.eu/wp-content/uploads/2018/03/FSS_P3_CU_D3.5_v2.0.pdf/](https://www.futuresky-safety.eu/wp-content/uploads/2018/03/FSS_P3_CU_D3.5_v2.0.pdf/)

8. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. – М.: Физматлит, 2006. - 813 с.

9. Онлайн-учебник по машинному обучению от ШАД. URL:

<https://education.yandex.ru>

10. Lishuai Li R., John Hansman, Rafael Palacios, Roy Welsch. Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring // Transportation Research Part C Emerging Technologies, 2016, vol. 64, pp, 45–57. DOI:

[10.1016/j.trc.2016.01.007](https://doi.org/10.1016/j.trc.2016.01.007)

11. Janakiraman V., Matthews B., Oza N. Discovery of Precursors to Adverse Events Using Time Series Data // SIAM International Conference on Data Mining, May 5-7, 2016, Miami, FL. DOI: [10.1137/1.9781611974348.72](https://doi.org/10.1137/1.9781611974348.72)

[10.1137/1.9781611974348.72](https://doi.org/10.1137/1.9781611974348.72)

12. Melnyk I., Banerjee A., Matthews B., Oza N. Semi-Markov Switching Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems // ACM International Conference on Knowledge Discovery and Data Mining (KDD), 2016. DOI:

[10.1145/2939672.2939789](https://doi.org/10.1145/2939672.2939789)

13. Raphael Monstein, Benoit Figuet, Timothé Krauth, Manuel Waltert and Marcel Dettling. Large Landing Trajectory Dataset for Go-Around Analysis // Engineering Proceedings, 2022, vol. 28 (2). DOI: [10.3390/engproc2022028002](https://doi.org/10.3390/engproc2022028002)
14. Буров Я.В., Стрелков В.В. Программа расшифровки полетной информации с бортовых параметрических регистраторов пассажирских самолетов «ARINC-конвертор». Свидетельство о регистрации программы для ЭВМ № 2017613196, 13.03.2017.
15. Alberto Rizzoli. Training Data Quality: Why It Matters in Machine Learning // An Introductory Guide to Quality Training Data for Machine Learning. URL: <https://www.v7labs.com/blog/quality-training-data-for-machine-learning-guide>
16. Побединский Г.Г. Системы координат в геодезии и их связи. - Нижний Новгород: Нижегородский государственный архитектурно-строительный университет, 2023. – 144 с.
17. Geraud de Rivals. Landing Trajectory Computation // 3rd conference of the European Operators Flight Data Monitoring forum (EOFDM), Cologne, 06 February 2014. URL: <https://www.pressreleasepoint.com/european-operators-flight-data-monitoring-eofdm>
18. Емельянов Д.С., Стрелков В.В., Горский Е.Б. Восстановление и кодировка траектории движения самолета на заключительном этапе снижения по глиссаде, выравнивании и пробеге по ВПП // XXX Научно-техническая конференция по аэродинамике (пос. Володарского, 25÷26 апреля 2019): тезисы докладов. – Жуковский: ЦАГИ, 2019. С. 111-112.
19. Емельянов Д.С. Программа восстановления траектории движения самолета на заключительном этапе снижения по глиссаде, выравнивании и движении по взлетно-

посадочной полосе (TrajectoryScan). Свидетельство № 2018664260, 2018.

20. G. van Es, P. der Geest, Cheng A., Hackler L., Dillard A.E. A Study of Normal Operational Landing Performance on Subsonic, Narrow-Body Jet Aircraft during Instrument Landing System Approaches. Report No. DOT/FAA/AR-07/72007.

References

1. G. van Es. *A Study of Runway Excursions from a European Perspective*. NLR-CR-2010-259. URL: <https://skybrary.aero/sites/default/files/bookshelf/2069.pdf>

2. Jenkins M., Aaron R.F. *Reducing Runway Landing Overruns*. *Boeing Aero Magazine*. URL: https://www.boeing.com/commercial/aeromagazine/articles/2012_q3/3/

3. Sharov V.D. *Prognozirovanie i predotvrashchenie vykatyvanii samoleta za predely vzletno-posadochnoi polosy* (Forecasting and Prevention of Aircraft Rolling Out of the Runway), Germaniya, Lambert Academic Publishing, 2013, pp. 112.

4. Rybin A.V. *Trudy MAI*, 2015, no. 81. URL: <https://trudymai.ru/eng/published.php?ID=57761>

5. Eremin A.I., Sel'vesyuk N.I. *Trudy MAI*, 2018, no. 100. URL: <https://trudymai.ru/eng/published.php?ID=93451>

6. Semakov S.L. *Trudy MAI*. 2019. № 104. URL: <https://trudymai.ru/eng/published.php?ID=102188>

7. Barry D., Greeves M., Solis T., Angel M. *Advanced methods for analysis flight data for runway excursions risk factors*. Cranfield University, Airbus Defence&Space. Future Sky Safety. Project: Solutions for Runway Excursions, 2017. URL: https://www.futuresky-safety.eu/wp-content/uploads/2018/03/FSS_P3_CU_D3.5_v2.0.pdf/

8. Kobzar' A.I. *Prikladnaya matematicheskaya statistika* (Applied mathematical statistics), Moscow, Fizmatlit, 2006, 813 p.
9. *Onlain-uchebnik po mashinnomu obucheniyu ot ShAD*. URL: <https://education.yandex.ru>
10. Lishuai Li R., John Hansman, Rafael Palacios, Roy Welsch. Anomaly detection via a Gaussian Mixture Model for flight operation and safety monitoring, *Transportation Research Part C Emerging Technologies*, 2016, vol. 64, pp. 45–57. DOI: [10.1016/j.trc.2016.01.007](https://doi.org/10.1016/j.trc.2016.01.007)
11. Janakiraman V., Matthews B., Oza N. Discovery of Precursors to Adverse Events Using Time Series Data, *SIAM International Conference on Data Mining*, May 5-7, 2016, Miami, FL. DOI: [10.1137/1.9781611974348.72](https://doi.org/10.1137/1.9781611974348.72)
12. Melnyk I., Banerjee A., Matthews B., Oza N. Semi-Markov Switching Vector Autoregressive Model-Based Anomaly Detection in Aviation Systems, *ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016. DOI: [10.1145/2939672.2939789](https://doi.org/10.1145/2939672.2939789)
13. Raphael Monstein, Benoit Figuet, Timothé Krauth, Manuel Waltert and Marcel Dettling. Large Landing Trajectory Dataset for Go-Around Analysis, *Engineering Proceedings*, 2022, vol. 28 (2). DOI: [10.3390/engproc2022028002](https://doi.org/10.3390/engproc2022028002)
14. Burov Ya.V., Strelkov V.V. *Svidetel'stvo o registratsii programmy dlya EVM № 2017613196*, 13.03.2017.
15. Alberto Rizzoli. Training Data Quality: Why It Matters in Machine Learning, *An Introductory Guide to Quality Training Data for Machine Learning*. URL: <https://www.v7labs.com/blog/quality-training-data-for-machine-learning-guide>

16. Pobedinskii G.G. *Sistemy koordinat v geodezii i ikh svyazi* (Coordinate systems in geodesy and their relationships), Nizhnii Novgorod, Nizhegorodskii gosudarstvennyi arkhitekturno-stroitel'nyi universitet, 2023, 144 p.
17. Geraud de Rivals. Landing Trajectory Computation, *3rd conference of the European Operators Flight Data Monitoring forum (EOFDM)*, Cologne, 06 February 2014. URL: <https://www.pressreleasepoint.com/european-operators-flight-data-monitoring-eofdm>
18. Emel'yanov D.S., Strelkov V.V., Gorskii E.B. *XXX Nauchno-tehnicheskaya konferentsiya po aerodinamike: tezisy dokladov*. Zhukovskii, TsAGI, 2019, pp. 111-112.
19. Emel'yanov D.S. *Programma vosstanovleniya traektorii dvizheniya samoleta na zaklyuchitel'nom etape snizheniya po glissade, vyravnivanii i dvizhenii po vzletno-posadochnoi polose (TrajectoryScan)*. *Svidetel'stvo № 2018664260*, 2018.
20. G. van Es, P. der Geest, Cheng A., Hackler L., Dillard A.E. *A Study of Normal Operational Landing Performance on Subsonic, Narrow-Body Jet Aircraft during Instrument Landing System Approaches*. Report No. DOT/FAA/AR-07/72007.

Статья поступила в редакцию 05.08.2024

Одобрена после рецензирования 13.08.2024

Принята к публикации 25.10.2024

The article was submitted on 05.08.2024; approved after reviewing on 13.08.2024; accepted for publication on 25.10.2024