



УДК 519.216.3

ОПРЕДЕЛЕНИЕ НАИЛУЧШИХ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ПАССАЖИРОПОТОКОВ В АЭРОПОРТУ ПУЛКОВО В УСЛОВИЯХ НОРМАЛЬНОЙ РАБОТЫ И КРИЗИСА

С. А. Андронов

Санкт-Петербургский университет аэрокосмического приборостроения

Рассмотрена задача нахождения лучших по точности прогнозных моделей пассажиропотоков в аэропорту Пулково в условиях его штатной работы и действия неблагоприятных факторов. В качестве средства моделирования использовалась программа Statistica, а также были задействованы системы Дедуктор и Loginom. Моделирование позволило получить адекватные модели и прогнозы как в условиях нормальной работы аэропорта так и в условиях пандемии, вызванной коронавирусной инфекцией. Средние относительные ошибки прогнозирования на разных горизонтах упреждения незначительно различаются и располагаются в пределах от 0,5 до 5%. Описан вариант автоматизации экспертных процедур прогнозирования пассажиропотоков в аэропорту с учетом влияния внешних факторов. Кризисный сценарий в рассматриваемом периоде оценивается потерями порядка 40%.

Ключевые слова: аэропорт, пассажиропоток, модели прогнозирования, временные ряды, экспертный прогноз, коронавирус.

Для цитирования:

Андронов С. А. Определение наилучших моделей прогнозирования пассажиропотоков в аэропорту Пулково в условиях нормальной работы и кризиса // Системный анализ и логистика: журнал.: выпуск №2(24), ISSN 2007-5687. – СПб.: ГУАП., 2020 – с. 13-29. РИНЦ.

DETERMINING THE BEST MODELS FOR PREDICTING PASSENGER FLOWS AT PULKOVO AIRPORT IN NORMAL OPERATION AND CRISIS CONDITIONS

S. A. Andronov

Saint-Petersburg state University of Aerospace Instrumentation

The problem of finding the best accuracy forecast passenger traffic models at Pulkovo Airport under the conditions of its regular operation and the action of adverse factors is considered. The Statistica program was used as a modeling tool, and the systems Deductor and Loginom were also used. Modeling made it possible to obtain adequate models and forecasts both in the normal operation of the airport and in the context of a pandemic caused by coronavirus infection. The relative relative forecast errors at different lead horizons differ slightly and range from 0,5 to 5%. A variant of automation of expert procedures for forecasting passenger flows at the airport, taking into account the influence of external factors, is described. The crisis scenario in the period under review is estimated at a loss of about 40%.

Keywords: airport, passenger flow, forecasting models, time series, expert forecast, coronavirus.

For citation:

Andronov S. A. Determining the best models for predicting passenger flows at Pulkovo airport in normal operation and crisis conditions // System analysis and logistics.: №2(24), ISSN 2007-5687. – Russia, Saint-Petersburg.: SUAI., 2020 – p. 13-29.

Введение

По количеству принимаемых пассажиров Пулково занимает четвертое место в России. Рост пассажиропотока аэропорта обусловлен рядом факторов. Один из ключевых среди них — активная работа оператора аэропорта совместно с администрацией Санкт-Петербурга по расширению маршрутной сети Пулково и привлечению новых авиакомпаний, в том числе низкобюджетных перевозчиков, для открытия новых регулярных полетов как по российским, так и зарубежным направлениям [7]. Положительное влияние оказывают транспортная доступность, модернизация аэропорта для удобства сервиса и увеличения пропускной способности. Кроме того, с 01.01.2020 действует режим «открытого неба» по седьмой степени «свободы воздуха», позволяющий



иностранным авиаперевозчикам не совершать посадку в стране своей регистрации и использовать электронные визы. По оценкам пассажиропоток в «Пулково» к 2039 году достигнет 35 млн пассажиров, а уже в 2025 году составит 25 млн пассажиров [8].

Возможность эффективного управления деятельностью аэропорта зависит от качества прогнозирования пассажиропотоков. Перспективы, которые необходимо предсказывать, чаще всего описываются временными рядами, то есть последовательностью значений некоторых величин, полученных в определенные моменты времени. Математические формальные методы прогнозирования временных рядов, основанные на идее экстраполяции, исходят из предпосылки, что в краткосрочной перспективе существующая тенденция сохранится и в будущем. Если устойчивая тенденция нарушается, возникает точка перелома, которую можно предсказать только с помощью экспертных процедур. Подобная ситуация наблюдается в условиях высокой неопределенности внешней среды. Этот вариант развития событий также будет затронут в настоящей работе.

При прогнозировании временного ряда требуется определить модель прогнозирования, т.е. функциональную зависимость, адекватно описывающую временной ряд. Целью при этом является нахождение модели, ошибка которой, например, средняя относительная, является наименьшей для времени упреждения. После того, как модель прогнозирования временного ряда определена, требуется вычислить будущие значения временного ряда, а также их доверительный интервал.

Постановка задачи

Целью исследования является определение наилучших регрессионных моделей для прогнозирования пассажиропотоков аэропорта Пулково Санкт-Петербург и прогноза с их помощью пассажиропотоков в 2020 году. Объектами прогнозирования являются внутренний, международный и суммарный пассажиропотоки.

В результате развития коронавирусной инфекции COVID-19 международные перелеты оказались в кризисе. Большинство наиболее привлекательных стран было закрыто на въезд [9,10], а по возвращении из них россияне должны отсидеть дома две недели карантина. Отмена рейсов, режим карантина, настрой населения на нежелание перемещаться сказался не только на международных направлениях, но и на внутренних направлениях. Помимо названных, присутствуют и другие негативные факторы, такие как, снижения курса рубля из-за падения цен на нефть, реальных доходов населения и др., которые однако, на фоне пандемии, представляются не настолько значительными. Таким образом, интерес представляет также прогнозирование степени падения пассажиропотоков в условиях объявленной пандемии.

Стоит подчеркнуть, что в условиях распространения вируса аналогичный процесс коснулся не только авиационного, но также и других видов транспорта, в частности, внутренних пассажирских перевозок на железнодорожном транспорте, автомобильного дальнего следования, а также городского пассажирского транспорта, включая метро. Соответственно, все названные виды транспорта и транспортные предприятия нуждаются в предиктивном управлении, а значит, и в прогнозировании.

Особенности регрессионных моделей

К настоящему времени разработано множество моделей для решения задачи прогнозирования временного ряда, среди которых наибольшую применимость имеют: статистические (регрессионные, авторегрессионные, скользящего среднего, модели экспоненциального сглаживания) и структурные, в частности, нейросетевые, где функциональная зависимость между будущими и фактическими значениями временного ряда, а также внешними факторами задана структурно. В данной работе для построения регрессионных моделей главным влияющим фактором является время и поэтому далее регрессия понимается относительно только этого фактора. По этой причине были выбраны модели в классах авторегрессии, скользящего среднего, экспоненциального сглаживания и нейросетевые.



Перечисленные модели обладают определенными достоинствами и недостатками, которые описаны во многих источниках, например [2, 3].

Составить наиболее точный прогноз можно, не ограничиваясь каким-то одним методом, а используя ряд методик. Это позволит наиболее объективно оценить ситуацию и обеспечить разработку качественного прогноза. Точность прогноза считается оптимальной, если результаты, полученные в результате использования различных методов отличаются друг от друга не более, чем на 10% [2].

Кроме того, при выборе наиболее подходящей модели требуется учитывать величину периода упреждения. В задачах краткосрочного, среднесрочного и долгосрочного прогнозирования вышеперечисленные модели имеют различную применимость.

Краткосрочный прогноз - прогноз на несколько шагов вперед. Для него подходят все вышеназванные модели. Среднесрочный прогноз – это прогноз на один или на половину сезонного цикла. Для него используют авторегрессионную модель со скользящим средним и экспоненциальное сглаживание, которые позволяют отслеживать качество прогноза в зависимости от горизонта упреждения.

При построении долгосрочного прогноза стандартные статистические методы прогнозирования практически не используют, и применяют комбинированный подход, который позволяет объединить два и более метода, компенсировать недостатки одних методов достоинствами других, особенно в случае, если они построены на различной информационной базе. При этом строится средневзвешенная оценка прогноза, где весовые коэффициенты тем меньше, чем менее точным является прогноз.

В отсутствии статистических данных качество прогноза на различных горизонтах можно повысить в результате использования комбинаций экспертных и статистических методов. В общем случае при такой комбинации необходимо учитывать являются ли такие прогнозы противоречивыми, перекрытие их доверительных интервалов и т.д.

Выбор лучшей модели в условиях штатной работы аэропорта

При построения прогнозных модели для нахождения сезонного лага, а также дальнейшей стационаризации производят исследование автокорреляционной и частной автокорреляционной функции, применяют спектральный анализ. Важно проверять, насколько построенная модель является адекватной. Для этого можно, во-первых, провести сравнение результатов прогнозирования и фактических данных со сдвигом прогноза на несколько шагов назад, а во-вторых, воспользоваться анализом остатков.

В качестве критерия точности в настоящем исследовании используется средняя относительная ошибка в виде

$$\delta = \frac{1}{p} \sum_{i=1}^p \left| \frac{(\hat{y}_i - y_i)}{\hat{y}_i} \right| 100\% ,$$

где p - период упреждения, $\hat{y}_i$, y_i - прогнозные и фактические значения соответственно.

Задача определения лучшей прогнозных модели сводится к тому, чтобы средняя относительная ошибка отклонения истинного значения от прогнозируемого стремилась к минимальному для заданного p .

В качестве основного средства исследований использовалась программа Statistica [6]. В дополнение к ней применялись возможности прогнозирования в программах Дедуктор 5.3 [4] и его наследнике – программе Logipom 6.3 [5], дополняющие основную программу. Данные системы



позиционированы как аналитические платформы интеллектуального анализа «сырых» данных произвольного вида из разнородных источников, в том числе промышленных.

На рисунке 1 представлены графики пассажиропотоков, взятых из открытого источника за период с января 2007 по декабрь 2019 г. [11]. Общее число пассажиров – сумма внутреннего и международного потока.

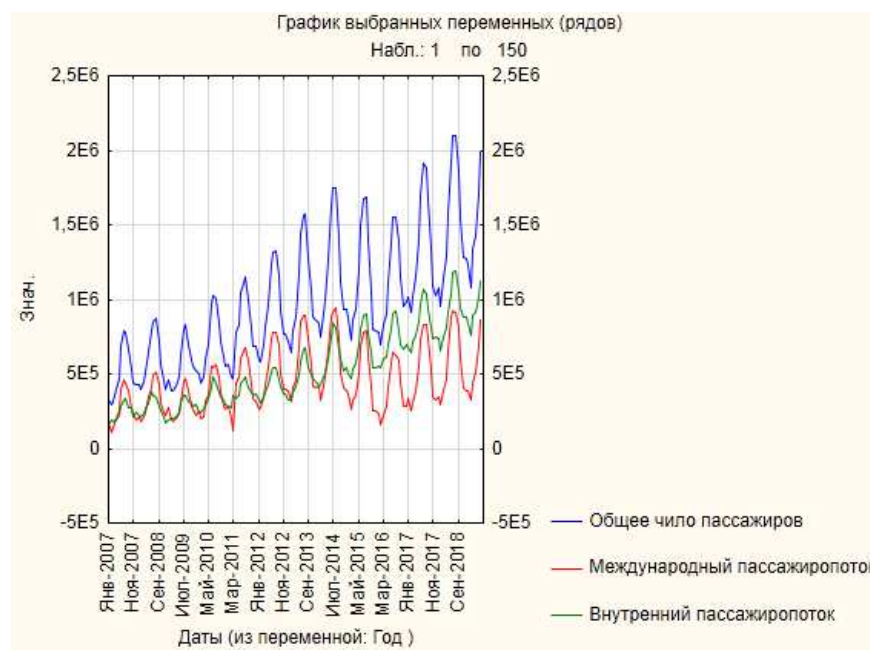


Рис.1. Динамика пассажиропотоков

Можно заметить, что динамика носит сезонный характер с пиками в середине лета когда число отпусков достигает пиковых значений.

Рассмотрим требуемый порядок действий для подготовки к прогнозированию. Вначале необходимо оценить параметры модели, если модель содержит параметры, или структуру, если модель относится к категории структурных моделей.

В результате выполнения операции Автокорреляция (рис. 2) на всех потоках пассажиров выделяется сезонный всплеск на 12 лаге, т.е. присутствует очевидная годовичная сезонность.

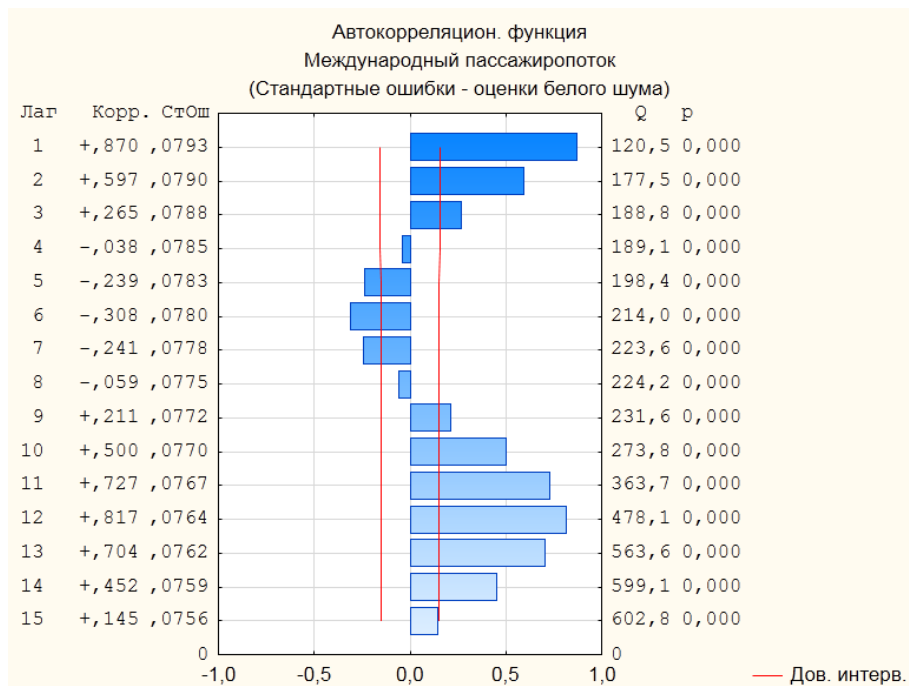


Рис. 2. Вид автокорреляционной функции

Для оценки компонент ряда выполним классическую сезонную декомпозицию Census I, предполагая мультипликативную модель исходя из вида процессов. В результате выделяются нерегулярные компоненты (рис.3), а также сезонная и восходящая тренд-циклическая (рис.4). Можно заметить, что разброс значений со временем меняется, что свидетельствует о непостоянстве среднего и дисперсии. Также на основе построения частной автокорреляционной функции можно сделать вывод о ее достаточно медленном затухании, что свидетельствует о нестационарности исследуемых временных рядов. Нерегулярные компоненты рядов в основном являются стационарными процессами и только международный поток содержит единственный выброс.

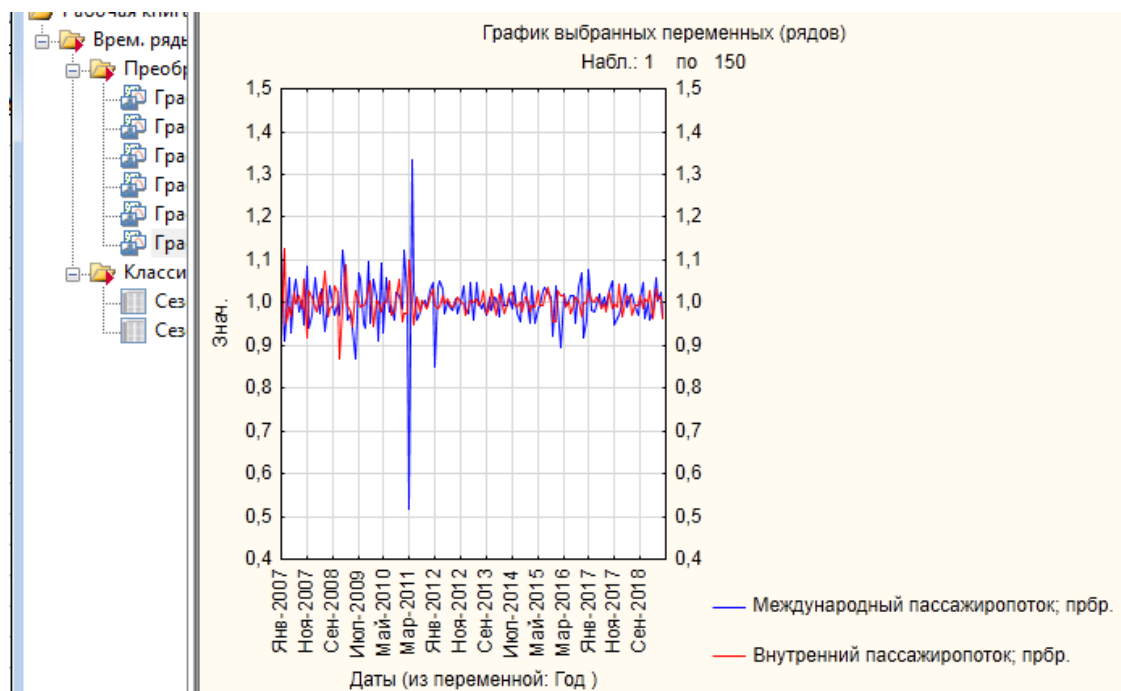


Рис. 3. Нерегулярные компоненты внутреннего и международного пассажиропотоков



С целью исключения аномальных выбросов и сглаживания шумовых составляющих для дальнейшего анализа (за исключением применения метода экспоненциального сглаживания) в программе Statistica предварительно выполнялась фильтрация шумов фильтром 4253Н (рис. 5). В программах Дедуктор и Loginom для этой цели, применялось исключение аномальных значений (ограничение $\pm 3\sigma$) и вычитание шума. В частности, в Loginom использовался фильтр Ходрика-Прескотта с параметром $\lambda=1$.



Рис. 4. Результат выделения трен-циклических компонент внутреннего и международного пассажиропотока

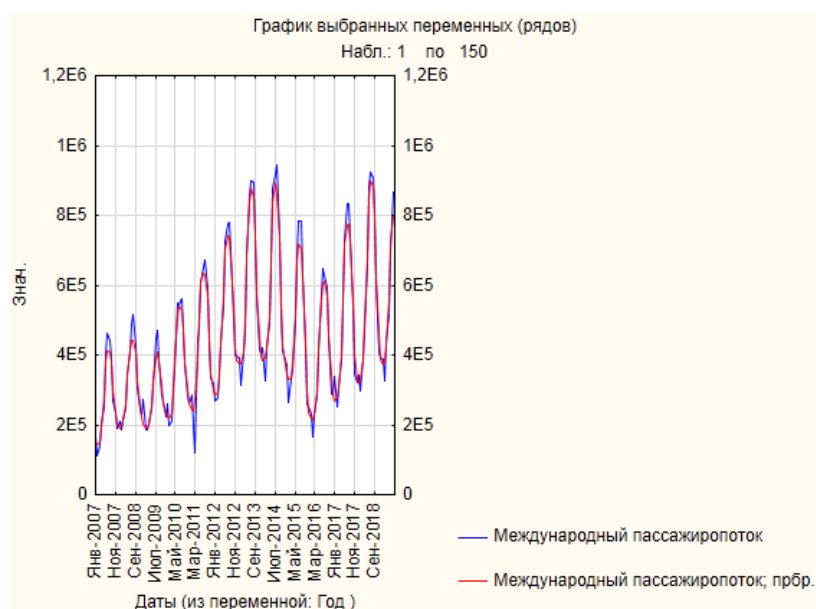


Рис. 5. Результат сглаживания международного пассажиропотока

Для сравнения моделей на интервале нормального протекания процессов был выделен набор исторических данных без последних 6 месяцев 2019 года, на которых планировалось сравнивать фактические пассажиропотоки с прогнозными.



1. Линейная авторегрессионная модель

Для интервалов прогнозов на один – два отсчета, т.е. в нашем случае порядка одного месяца, эффективными могут оказаться линейные авторегрессионные модели, которые, однако вблизи сезонных пиков работают плохо.

В основу авторегрессионных моделей p (AR(p)-процесс) заложено предположение о том, что значение прогноза процесса линейно зависит от некоторого количества предыдущих значений того же процесса. Данная имеет вид

$$\hat{y}_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \varepsilon_t,$$

где $\hat{y}_t$ – прогноз в момент времени t , y_{t-i} – исторические данные, ε_t – случайная составляющая (белый шум), a_i – коэффициенты авторегрессии.

В результате корреляционного анализа было выявлено, что наибольшим весом на интервале прогнозирования для рассматриваемой модели имеют 3 предшествующих прогнозу отсчета, поэтому они использовались в качестве входных данных для настройки модели. Результат по международному потоку, полученный в программе Дедуктор и справедливый при прогнозировании как для первого так и второго полугодия 2019 г. имеет вид

$$\hat{y}_t = 10752,95 + 0,89y_{t-1} - 2,5y_{t-2} + 2,59y_{t-3} + \varepsilon_t.$$

Следует отметить, что ошибки данной модели с ростом горизонта линейно нарастают (июль – 6%, август – 17%, август 24%).

2. Модель скользящего среднего

Данная модель применима для стационарных процессов. Формальная запись модели q -го порядка $MA(q)$ имеет вид

$$\hat{y}_t = \sum_{j=1}^q b_j \varepsilon_{t-j},$$

где b_j – параметры модели, ε_t – белый шум.

Построение модели выполнялось в программе Statistica. При этом для стационаризации исследуемых рядов требовалось исключить трендовую и сезонную составляющие.

Чтобы определить порядок q были взяты разности по D(-1) первому лагу, который имел существенное значение и D(-12) – для исключения годичной сезонности. Повторное построение автокорреляционной функции показало, что на ее графике сохранился единственный пик, а частная автокорреляционная функция при этом экспоненциально убывала. Данная ситуация характерна для наличия единственного значимого параметра $MA(1)$. Таким образом, было определено, что для внутреннего пассажиропотока оцениваемый параметр имеет значение 0.44, для международного пассажиропотока 0.624, для общего пассажиропотока 0.39. Все параметры определены по t -критерию Стьюдента как статистически значимые с доверительной вероятностью 95%. Данная модель показала достаточно высокую точность.



3. Модель авторегрессии (AR) проинтегрированного скользящего среднего (МА) - АРПСС или AutoRegressive Integrated Moving Average (ARIMA)

Рассмотренная выше модель AR относится к категории стационарных моделей, которые описывают процессы, для которых математическое ожидание и дисперсия являются константами. Однако, поскольку нестационарность часто присутствует, как в нашем случае, такая модель не может быть адекватной.

Стационарной может быть некоторая разность наблюдаемого процесса [1]. Вводится оператор Δ^d разности порядка d . В частности, $\Delta X_t = X_{t-1} - X_t$ - оператор разности временного порядка 1; $\Delta^2 X_t = \Delta^2 X_{t-1} - \Delta X_t$ - оператор разности 2-го порядка и т.д. Формальная запись модели ARIMA (p, d, q) имеет вид

$$\Delta^d X_t = c + \sum_{i=1}^p a_i \Delta^d X_{t-i} + \sum_{j=1}^q b_j \Delta^d \varepsilon_{t-j} + \varepsilon_t,$$

где ε_t — стационарный временной ряд; c, a_i, b_j — параметры модели.

Таким образом, модель ARIMA позволяет с приемлемым качеством прогнозировать как стационарные временные ряды, так и нестационарные, временные ряды с трендом, а также при небольшой модификации и ряды с сезонной компонентой (SARIMA). Расширением модели является учет внешних данных (ARIMAX), которые могут быть той же структуры или данными вообще из другого источника. Это может быть влияющий на прогноз временной процесс будущего периода (например, влияние температуры на энергопотребление).

В нашей задаче в качестве внешних были использованы предыдущие отсчеты значений ряда (-1, -2, -3, -12). На данный момент прогнозные значения внешних данных будущих периодов находятся в стадии реализации. Структура правой части математической модели ARIMAX подбиралась автоматически в программе Loginom с помощью минимизации Akaike Information Criterion (AIC).

На рисунках 6,7 представлен сценарии решения задачи в программе Loginom и Дедуктор соответственно. На рисунке 8 приведены прогнозы и остатки для модели ARIMA. Найденные коэффициенты для международного пассажиропотока представлены в таблице 1.

Таблица 1 – Коэффициенты модели ARIMAX

Параметр	Коэффициент
AR[1]	0,386851627
AR[2]	-0,27645517
AR[3]	0,164040615
AR[4]	0,172706303
AR[5]	-0,026304568
MA[1]	0,411309599
MA[2]	-0,086701789
MA[3]	-0,407997287
MA[4]	-0,879521337
MA[5]	-0,61059801
Сезонная AR[12]	-0,187054154
Сезонная AR[24]	-0,13010997
Сезонная MA[12]	-0,410756694
Сезонная MA[24]	-0,232059488
Внешний 12	8864,23279



Внешний 3	-23799,19777
Внешний 2	47151,13734
Внешний1	0,386851627

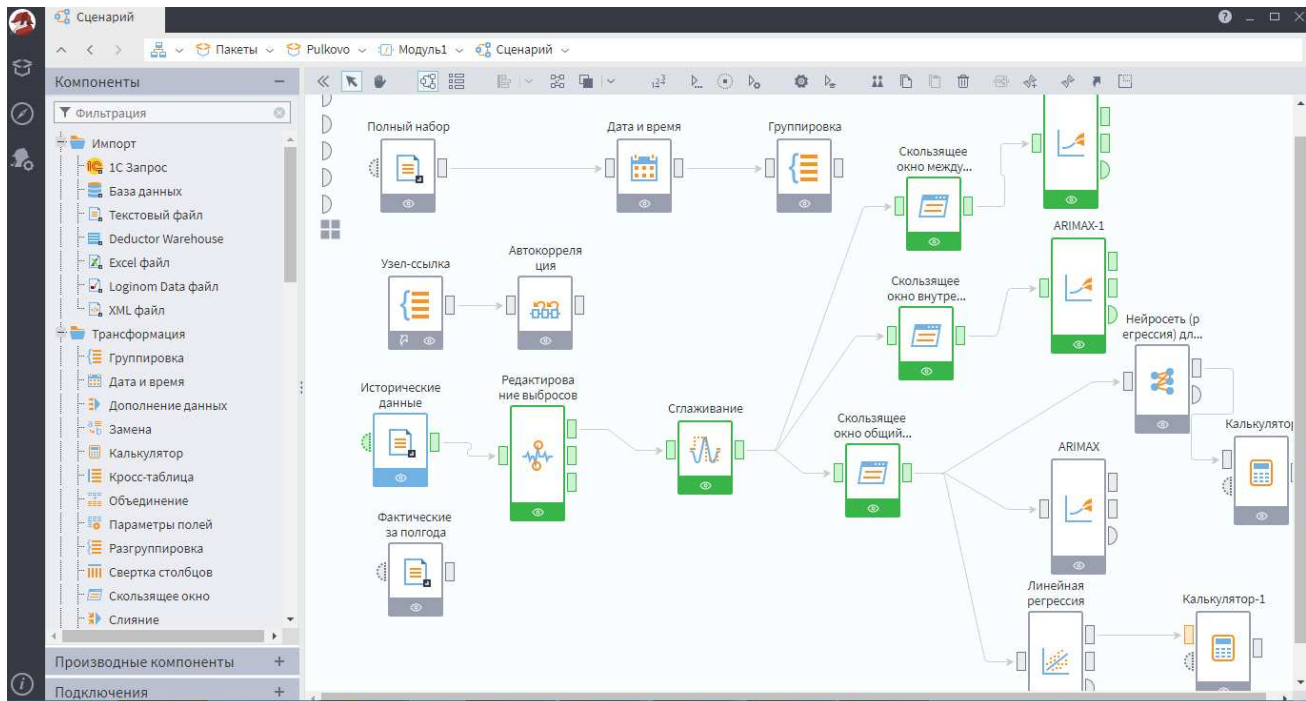


Рис.6. Вид сценария в программе Loginom

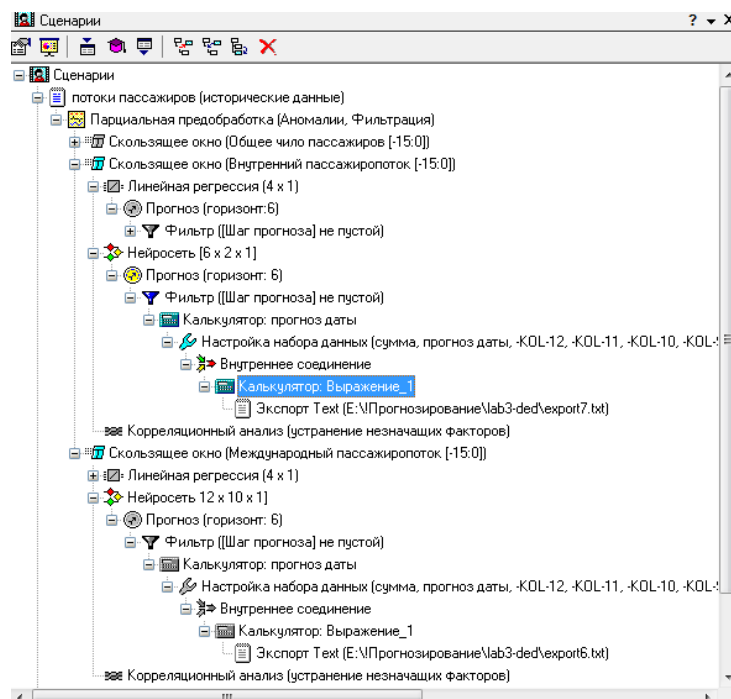


Рис. 7. Вид сценария в программе Дедуктор

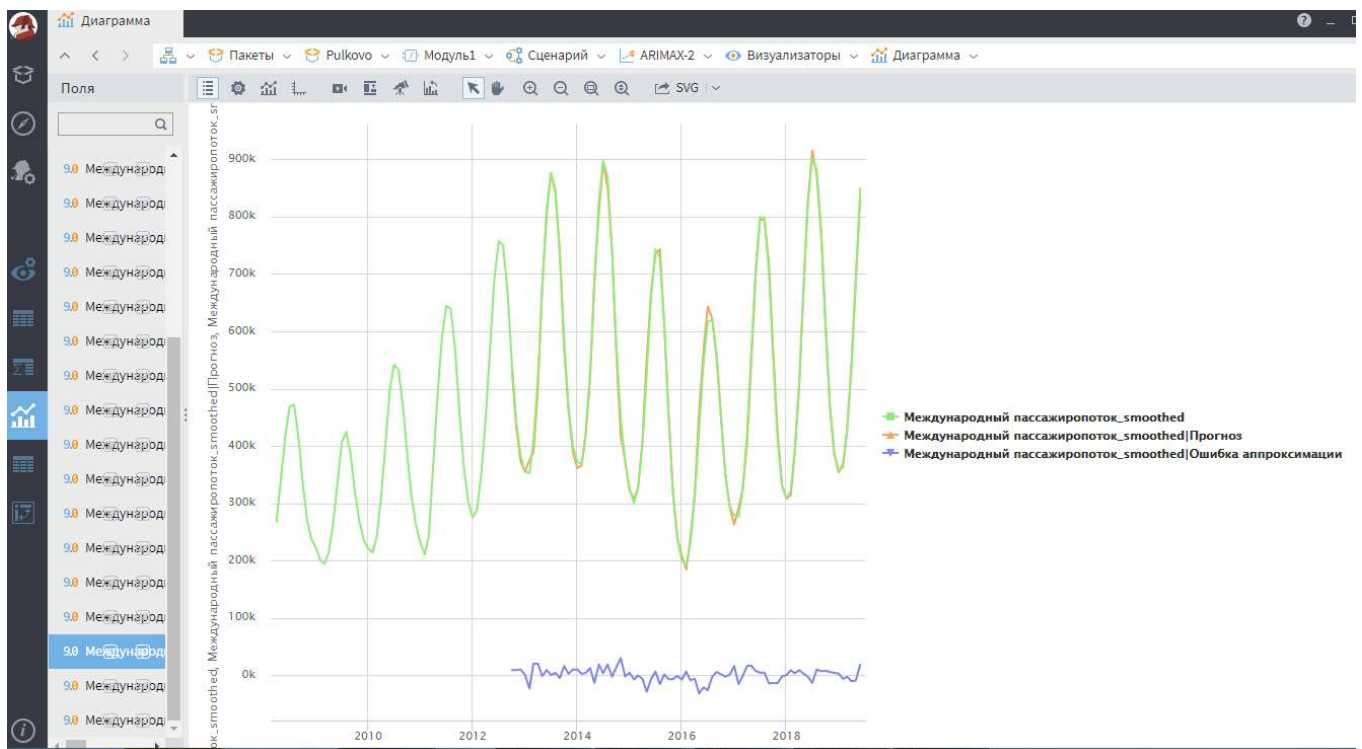


Рис. 8. Прогноз международного потока и остатки для модели ARIMA

4. Нейросетевые модели

Были рассмотрены различные структуры нейросети. В программе Statistica на всех 3-х потоках в первом во втором полугодии 2019 г. хорошо зарекомендовала себя модель с 12-тью входами из данных предыдущих месяцев и 8 нейронами на скрытом слое, т.е. вида 12 x 8 x 1, полученная обработчиком «автоматизированные нейронные сети» (АНС). Вид структуры нейросети показан на рисунке 9.

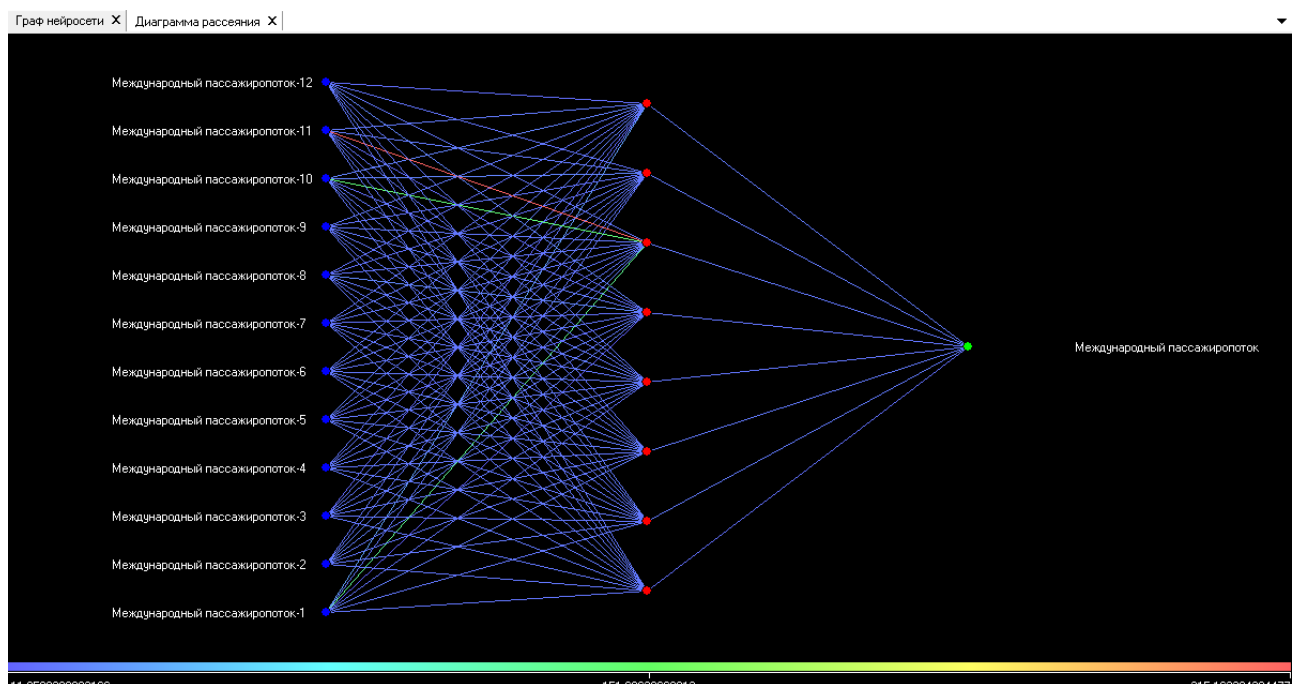


Рис.9. Вид структуры нейросети в программе Дедуктор



5. Экспоненциальное сглаживание

Данный вариант прогноза, как известно [2], может применяться для прогнозов на различных горизонтах. Программа Statistica поддерживает 3-х параметрическое сглаживание, позволяющее учесть тренд и сезонность в аддитивном и мультипликативном случае.

Аддитивная модель:

$$\hat{y}_t = S_t + I_{t-r}.$$

Мультипликативная модель:

$$\hat{y}_t = S_t \cdot I_{t-r}.$$

В этой формуле S_t обозначает (простое) экспоненциально сглаженное значение ряда в момент t ; I_{t-r} обозначает сглаженный сезонный фактор в момент t минус r (r длина сезона). Таким образом, в сравнении с простым экспоненциальным сглаживанием, прогноз улучшается добавлением или умножением сезонной компоненты. Эта компонента оценивается независимо (от тренда) с помощью простого экспоненциального сглаживания по формуле:

Аддитивная модель:

$$I_t = I_{t-r} + \delta \cdot (1 - \alpha) \cdot \varepsilon_t.$$

Мультипликативная модель:

$$I_t = I_{t-r} + \delta \cdot (1 - \alpha) \cdot \varepsilon_t / S_t.$$

Предсказанная сезонная компонента в момент t вычисляется через значения соответствующей компоненты на последнем сезонном цикле плюс ошибка (ε_t , наблюдаемое минус прогнозируемое значение в момент t). Параметры сглаживания должны соответствовать условиям: $0 \leq \alpha \leq 1$; $0 \leq \gamma \leq 1$; $0 \leq \delta \leq 1$.

Следует отметить, что внешний вид исходных данных на рассматриваемом интервале с точки зрения типа модели Винтера отличается незначительно. На больших интервалах для общего и международного пассажиропотока вид процессов предполагает использование мультипликативной модели, для внутреннего - аддитивной. Также, как следует из рис.4, тренд изменяется по линейной зависимости, что имеет значение при использовании модели Винтера. Для международного потока в результате поиска «на сетке» для мультипликативной модели с линейным трендом на горизонте 6 месяцев наименьшие значения средней относительной ошибки были получены при: $\alpha=0.9$; $\gamma=0.1$; $\delta=0.1$. На рисунке 10 приведен вид прогноза для международного пассажиропотока и остатков.

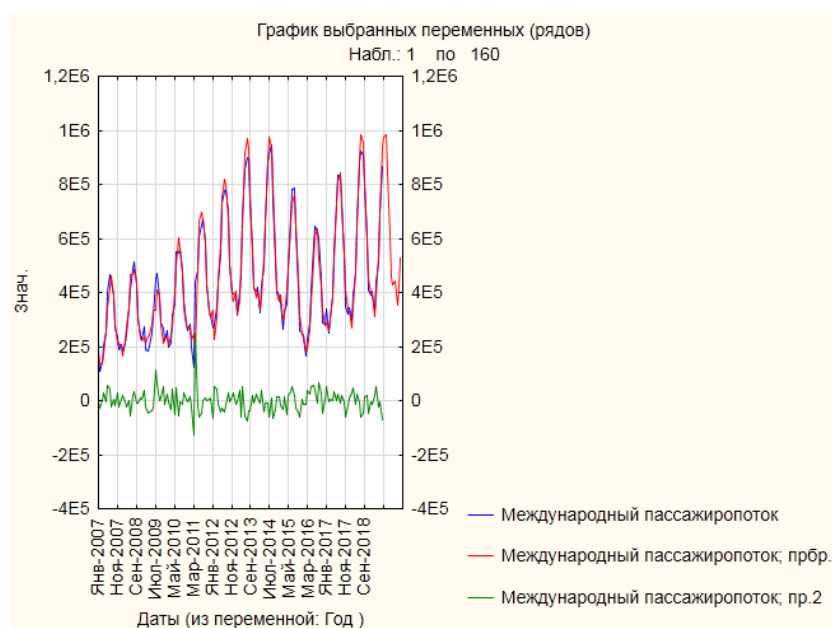


Рис.10. Иллюстрация экспоненциального сглаживания

6. Экспертный метод

Данный подход применим в случае когда методы экстраполяции не работают и необходимо определить точки перелома тенденции на основе учета влияния внешних факторов. В краткосрочной перспективе эксперт может определить точку перелома тенденции и скорректировать прогноз при заданном наборе внешних факторов, иначе говоря, ввести поправочный коэффициент к регрессионной модели прогноза.

Точки начала снижения пассажиропотоков вовсе не обязательно могут быть обусловлены такими негативными событиями как карантин, запрет на авиаперелеты и пр. Подобные сбои уже встречались в экономические кризисы 2008г., 2014г. Очевидно, возможны, наоборот, точки восстановления тенденции к росту, которые также могут быть предсказаны экспертно.

Внешние факторы могут иметь долгосрочный характер, как например: состояние инфраструктуры аэропорта, транспортная доступность, уровень авиационной подвижности населения, активность аэропортов – конкурентов, визовые барьеры, курс национальной валюты, присутствие авиакомпаний-дискаунтеров и лоукостеров, открытие новых направлений, наличие прямых рейсов, решения по «открытому небу» и т.д.; среднесрочный характер: эффективность работы персонала, скорость выдачи багажа, размеры очередей у стоек, качество питания, организация досуга; краткосрочный характер: погодные условия, сезонные скидки на перелеты, масштабы развития пандемии, введение карантина, мероприятия по обеззараживанию самолетов и мест общего пользования, наличие защитных масок у персонала и т.д.

Полученные результаты

1. Регрессионные модели

В таблицах 2-4 представлены помесечные и средние относительные ошибки, полученные в результате сравнения моделей прогноза для второго полугодия 2019. На рисунках 11-13 показаны соответствующие диаграммы. Для первого полугодия значения ошибок имеют примерно тот же уровень.



Таблица 2 – Внутренний пассажиропоток

Месяц	Скольз. ср.	ARIMA	Нейросеть 12x8x1	Винтер
01.07.2019	2	1	2	0
01.08.2019	1	1	1	1
01.09.2019	2	2	2	3
01.10.2019	3	3	0	5
01.11.2019	0	0	0	3
01.12.2019	2	2	2	1
средняя относительная ошибка	1,7	1,5	1,2	2,2

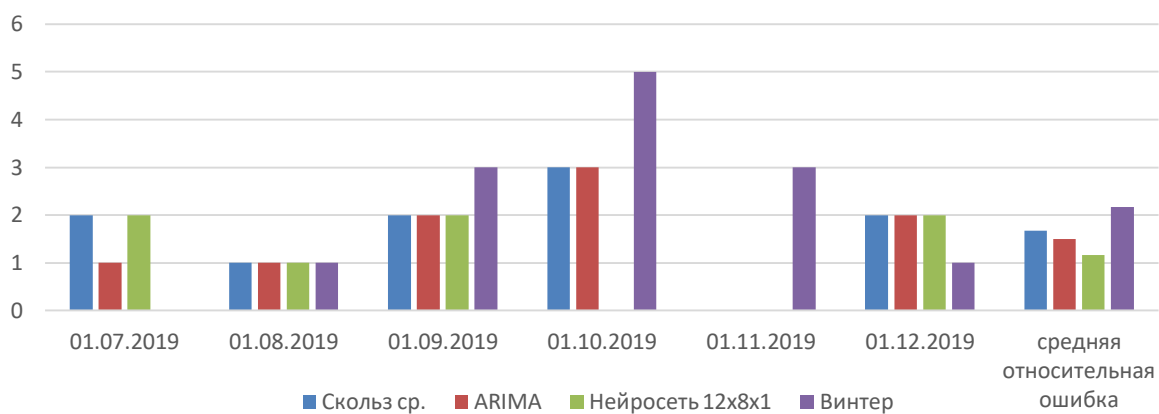


Рис.11. Диаграмма ошибок внутреннего пассажиропотока

Таблица 3 – Международный пассажиропоток

Месяц	Скольз. ср.	ARIMA	Нейросеть 12x8x1	Винтер
01.07.2019	1	0	6	4
01.08.2019	3	3	1	3
01.09.2019	3	3	3	1
01.10.2019	6	6	8	9
01.11.2019	8	6	6	8
01.12.2019	9	7	2	5
средняя относительная ошибка	5,0	4,2	4,3	5

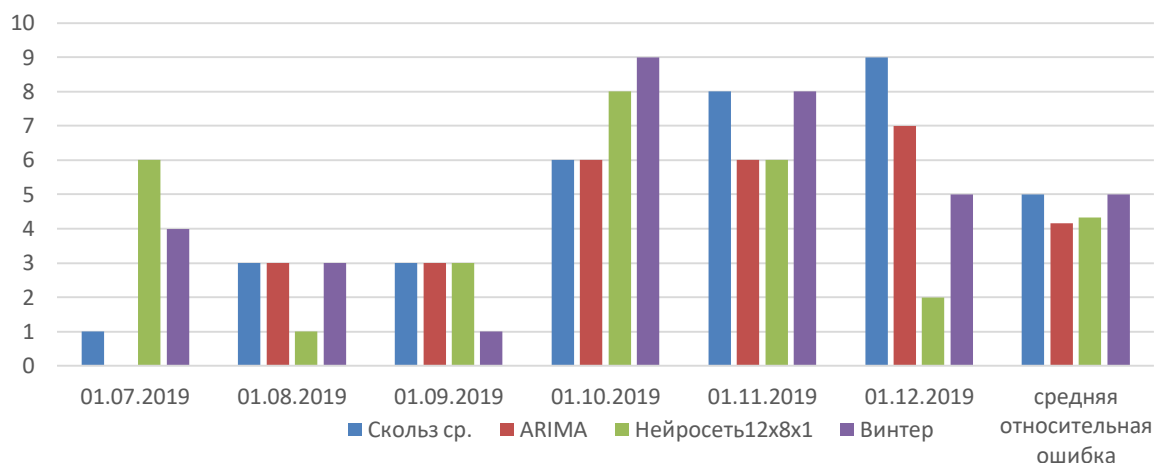


Рис. 12. Диаграмма ошибок международного пассажиропотока

Таблица 4 – Общий пассажиропоток

Месяц	Скользя. ср.	ARIMA	Нейросеть 12x8x1	Винтер
01.07.2019	1	1	4	1
01.08.2019	2	2	6	0
01.09.2019	1	1	7	4
01.10.2019	5	5	3	9
01.11.2019	1	0	6	2
01.12.2019	1	1	3	1
средняя относительная ошибка	1,8	1,7	4,8	2,8

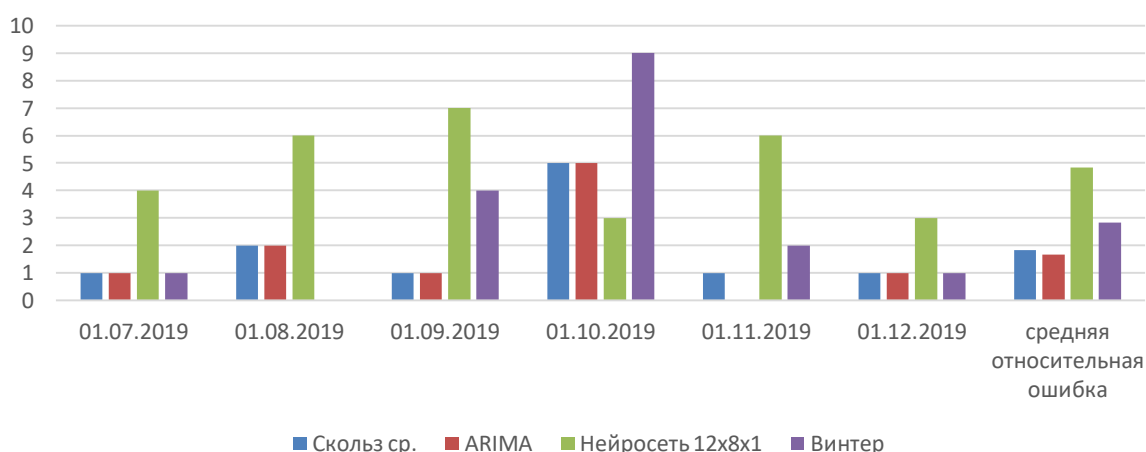


Рис. 13. Диаграмма общего пассажиропотока

Исходя из приведенных результатов можно заключить, что ошибки для всех рассмотренных моделей за исключением линейной регрессии, показавшей свою неудовлетворительность, незначительно отличаются на разных горизонтах и пассажиропотоках. Модель скользящего среднего с параметрами MA(1) (внутренний 0.44, международный - 0.624, общий 0.39) показала себя весьма достойно. Точность модели ARIMA оказалась несколько выше. Например, на горизонте 2 месяца наименьшее значение ошибок для международного пассажиропотока у модели ARIMA (1.5%). Для



внутреннего потока пассажиров лидирует модель Винтера (0.5%). На горизонте 6 месяцев ошибки всех рассмотренных моделей для внутреннего пассажиропотока примерно одинаковы и не превышают 2%. Наименьшие значения ошибок при прогнозе международного потока на горизонте 6 месяцев у модели ARIMA (4.2%).

2. Экспертный прогноз

Если говорить конкретно об аэропорте Пулково, главными источниками пассажиропотока являются деловые полеты, а также туристический поток из Китая, поэтому, вместе с этими источниками в условиях пандемии деградируют и потоки пассажиров.

Первые отмены международных рейсов в аэропорту Пулково начались с 13.03.2020 (10), 16.03(35), 18.03 (40), 19.03 (40). Затем, с 23.03.2020 синхронно с резким нарастанием заболеваемости в России, последовала полная отмена всех международных сообщений.

По причине отсутствия официальной информации о пассажиропотоках в феврале, марте 2020г. на момент написания данной статьи автор опирался на прогнозные данные за первый квартал 2020г. Ошибка прогноза для января 2020г составляла 0.1%. Потенциальное значение для марта составило 372533 пассажира или в день среднем $372533 / 31 = 12017$ чел. Потенциальный прирост относительно прогноза на февраль мог быть 40936 чел. или в день 1320 чел. Потенциальные потери только из за отмены рейсов в Европу (учитывалась средняя загрузка среднемагистральных самолетов A320 ка 150 чел.) составили $12017 * (31 - 23) + 150 * (10 + 35 + 40 + 40) = 108154 + 12750$ чел. или примерно 32%.

Автоматизировать процесс рассуждений эксперта для получения промежуточных оценок можно составив таблицу в виде комбинаций действующих факторов. В графе «Прогноз» эксперт проставляет процент, выражающий ожидаемую оценку изменения (роста или падения) пассажиропотока от 100%, соответствующих прогнозных значений в нормальной ситуации. Далее выполняется генерация обучающей выборки на основе таблицы искусственных прецедентов из комбинаций возможных факторов и строится модель нейронной сети с входными параметрами внешних факторов. С помощью визуализатора «Что если» (Дедуктор/ Loginom) можно выполнять эксперименты вводя новые значения внешних факторов и получая процент коррекции прогноза (рис. 15).

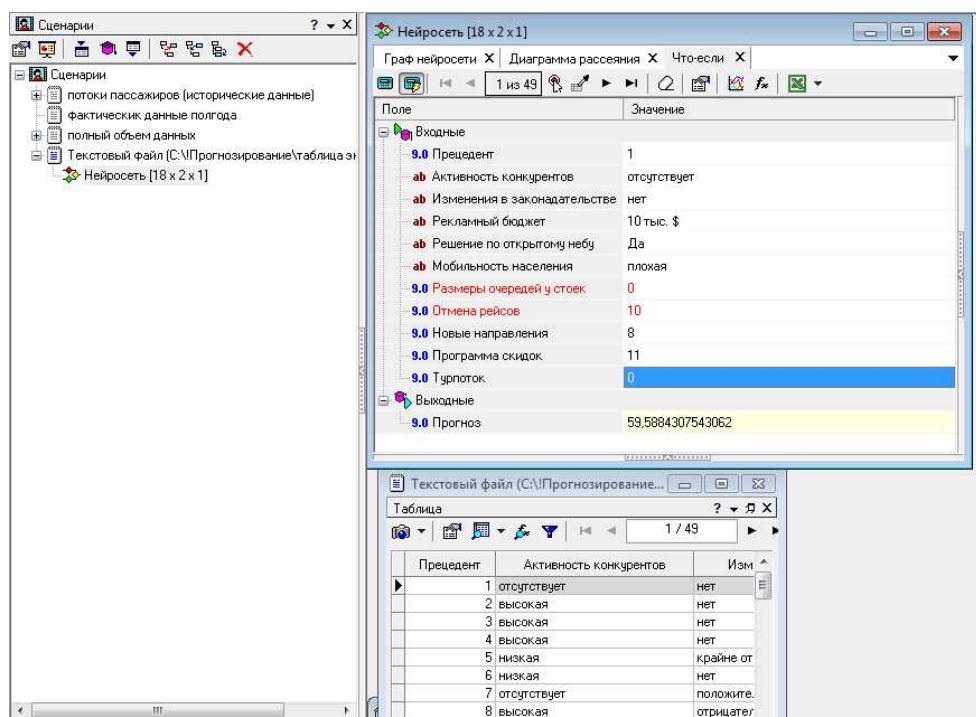


Рис.14. Визуализатор «Что если» в программе Дедуктор



Если учитывать закрытие европейских аэропортов, которое случилось на неделю раньше Пулково, низкую мобильность населения из-за запретов и карантина, отсутствие турпотока из Китая и пр., потери будут выше. Экспертная оценка падения международного пассажиропотока за март составила порядка 40% от значения 100% для прогноза на март. Данная оценка согласуется с прогнозами суммарных данных российских авиакомпаний для международных перевозок как минус 44% [12].

На рисунке 15 приведена утрированная оценка картины изменения международного потока в марте 2020 г.

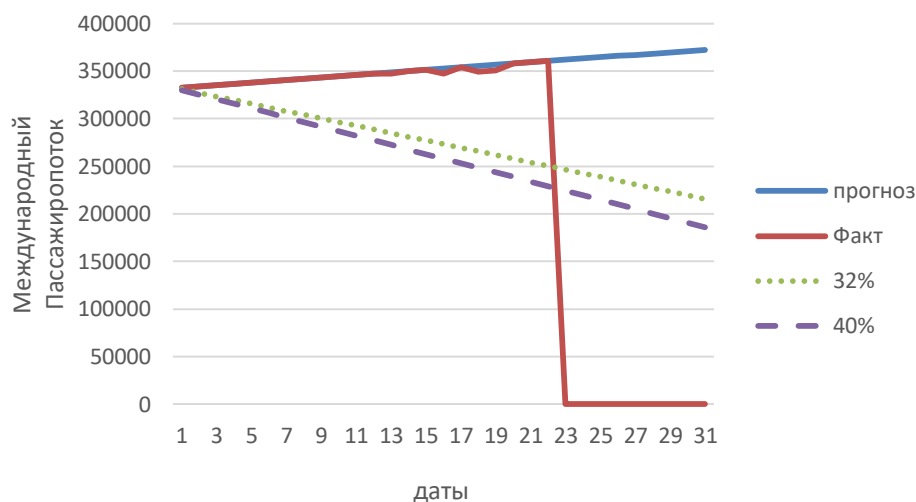


Рис. 15. Оценки международного пассажиропотока в марте 2020 г.

Заключение

От точности и надежности прогноза зависит эффективность реализации планов компаний. В настоящей статье была предпринята попытка определения лучших моделей прогноза пассажиропотоков в аэропорту Пулково в классе регрессионных, как в условиях нормального функционирования, так и в условиях кризиса, вызванного пандемией. Все рассмотренные модели, за исключением линейной авторегрессионной, продемонстрировали достаточно низкий уровень ошибок, которые незначительно отличаются. Лидеры моделей для разных горизонтов и пассажиропотоках отмечены в полученных результатах.

В отсутствии данных статистики приведены экспертные оценки снижения международного пассажиропотока при наличии вирусной инфекции, которые согласуются с выводами российских авиакомпаний. Описан подход автоматизации получения экспертных оценок при воздействии комбинации факторов.

СПИСОК ЛИТЕРАТУРЫ

1. Бокс Дж., Дженкинс Г.М. Анализ временных рядов, прогноз и управление / Дж. Боксб, Г.М. Дженкинс. - М.: Мир, 1974. - 406 с.
2. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобию : дис. ... канд. техн. наук : 05.13.18: защищена 12.03.2012 / Чучуева Ирина Александровна. — М., 2012. — 153 с.
3. Андронов С.А. Прогнозирование и планирование в сервисе : текст лекций / С. А. Андронов. - Санкт-Петербург : ГУАП, 2008. - 191 с.
4. Технологии анализа данных [Электронный ресурс]. — Режим доступа: <https://basegroup.ru/> (дата обращения: 02.04.2020).



5. Аналитическая платформа Loginom [Электронный ресурс]. — Режим доступа: <https://loginom.ru> (дата обращения: 02.04.2020).
6. Statsoft [Электронный ресурс]. — Режим доступа: <http://statsoft.ru/> (дата обращения: 02.04.2020).
7. Пулково установил исторический рекорд по пассажиропотоку [Электронный ресурс]. — Режим доступа: https://www.dp.ru/a/2017/11/09/Pulkovo_ustanovil_istorich (дата обращения: 02.04.2020).
8. Пассажиропоток «Пулково» к 2025 году достигнет 25 млн пассажиров [Электронный ресурс]. — Режим доступа: <https://regnum.ru/news/economy/2580685.html> (дата обращения: 02.04.2020).
9. Потери авиакомпаний России из-за коронавируса превысят 70 млрд. Как компании могут сократить убытки [Электронный ресурс]. — Режим доступа: <https://www.rbc.ru/business/13/03/2020/5e6a5f039a7947302d6e5ae4> (дата обращения: 02.04.2020).
10. Авиация России [Электронный ресурс]. — Режим доступа: <https://aviation21.ru/passazhiropotok-pulkovo-v-yanvare-2020-goda-uvelichilsya-na-8/> (дата обращения: 02.04.2020).
11. Аэропорт Пулково Санкт-Петербург [Электронный ресурс]. — Режим доступа: <https://pulkovoairport.ru/about/performance/> (дата обращения: 02.04.2020).
12. Пассажиропоток авиакомпаний РФ в марте на внутренних рейсах снизился почти на 17% [Электронный ресурс]. — Режим доступа: <https://tass.ru/ekonomika/8102203> (дата обращения: 02.04.2020).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Андронов Сергей Александрович –

к.т.н., доцент

Санкт-Петербургский университет аэрокосмического приборостроения

190000, Санкт-Петербург, ул. Большая Морская, д. 67, лит. А

E-mail: andronov_00@mail.ru

INFORMATION ABOUT THE AUTHOR

Andronov Sergej Aleksandrovich –

PhD in Technical Sciences

Saint-Petersburg state University of Aerospace Instrumentation

SUAI, 67, BolshayaMorskaia str., Saint-Petersburg, 190000, Russia

E-mail: andronov_00@mail.ru