

МАТЕМАТИКО-СТАТИСТИЧЕСКИЙ АНАЛИЗ ВОЗНИКНОВЕНИЯ И ДИНАМИКИ ЛЕСНЫХ ПОЖАРОВ В РОССИЙСКОЙ ФЕДЕРАЦИИ

Кандидат физ.-мат. наук *Л.Р. Борисова*
Финансовый университет при Правительстве РФ,
Московский физико-технический институт (государственный университет)

Доктор сельхоз. наук, кандидат техн. наук *Ю.В. Подрезов*
ФБГУ ВНИИ ГОЧС (ФЦ)
Московский физико-технический институт

Я.С. Гурьев
Московский физико-технический институт (государственный университет)

Целью исследования в работе является анализ факторов, влияющих на возникновение лесных пожаров и поиск устойчивых закономерностей для улучшения прогнозирования и, следовательно, увеличение скорости реакции органов управления МЧС России на лесные пожары и вызываемые ими чрезвычайные лесопожарные ситуации.

Ключевые слова: анализ данных; лесные пожары; машинное обучение; метеорологические данные; регрессия, статистические факторы.

MATHEMATICAL AND STATISTICAL ANALYSIS OF THE OCCURRENCE AND DYNAMICS OF FOREST FIRES IN THE RUSSIAN FEDERATION

Ph.D. (Phys.-Mat.) *L.R. Borisova*
Financial University under the Government of RF, Moscow Institute of physics
and technology (state University)

Dr. of agricultural sciences, Ph.D (Tech) *J.V. Podrezov*
FC VNII GOCHS EMERCOM of Russia
Moscow Institute of physics and technology (state University)

J.S. Guriev
Moscow Institute of physics and technology (state University)

The purpose of the research is to analyze the factors that affect the occurrence of forest fires and search for established patterns to improve forecasting and, consequently, increase the speed of response of the Russian Ministry of emergency situations to forest fires and forest fire emergencies caused by them.

Keywords: data analysis; forest fires; machine learning; meteorological data; regression, statistical factors.

Масштабные лесные пожары и вызываемые ими чрезвычайные лесопожарные ситуации (далее - ЧЛС) представляют серьезную опасность для многих стран, в том числе и Российской Федерации, которая имеет наибольшую площадь лесной территории.

При выполнении исследований мы использовали методы факторного анализа, который часто определяется в литературе как раздел многомерного статистического анализа, объединяющий методы оценки размерности множества наблюдаемых переменных посредством исследования структуры корреляционных и ковариационных матриц. Целями факторного анализа является определение взаимосвязей между переменными, их классификация и сокращение количества переменных, оставляя только самые значимые. Величина, называемая корреляцией, показывает насколько физические величины статистически зависят друг от друга. Иллюстрация корреляции приведена на рис. 1.

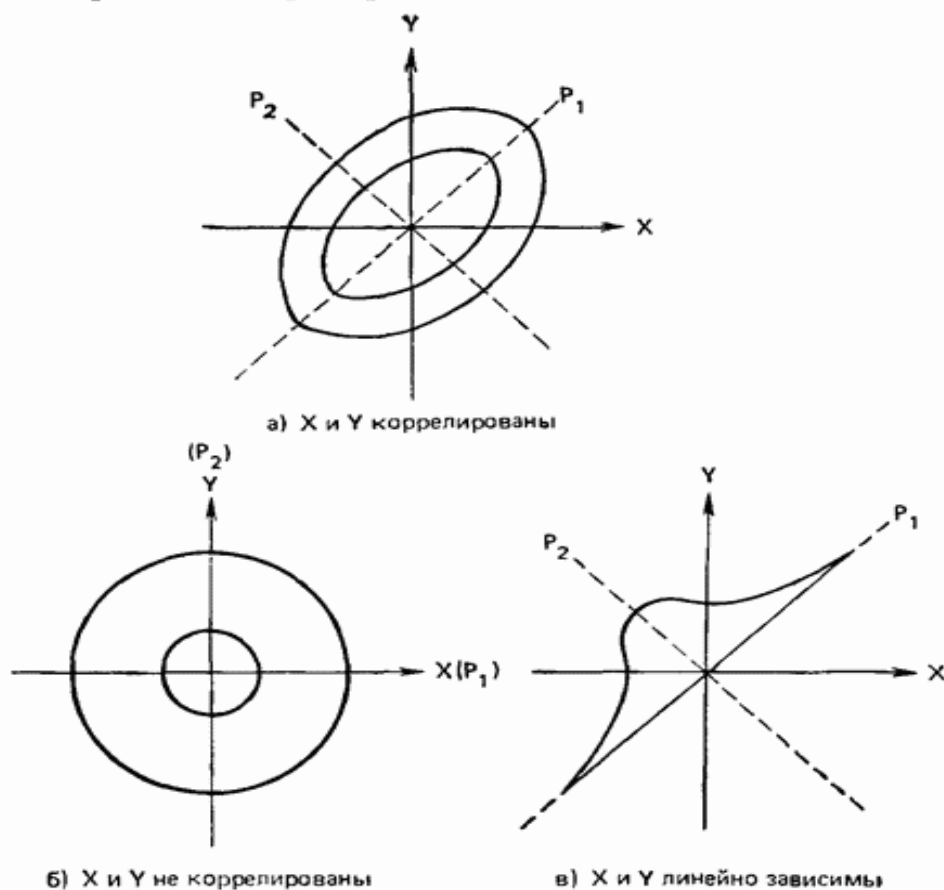


Рис. 1. Иллюстрация корреляции

Важная величина, используемая при тестировании достоверности статистических гипотез – это p -значение или p -value. Эта величина определяется как вероятность того, что случайная величина с данным распределением примет значение не меньшее, чем фактическое значение тестовой статистики. Иными словами, это наименьшее значение уровня значимости, для которого вычисленная проверочная статистика ведет к отказу от нулевой гипотезы. Если p -значение больше определенного значения (0,005 или 0,01) то гипотезу считают недостоверной.

Еще один критерий проверки гипотез при построении регрессионных моделей - это критерий Фишера или F-тест. В общем случае критерий Фишера оценивает отношение дисперсий двух выборок. Чем ближе F-значение к 1, тем больше вероятность признать гипотезу справедливой. Другой критерий проверки гипотез - это t-критерий Стьюдента, в котором проверяется отношение разницы выборочного среднего и известным значением к стандартной ошибке выборочного среднего.

При проведении наших исследований используется дискриминантный анализ, позволяющий предсказать к какой из нескольких групп будет относиться данный объект, и по каким переменным выполняется разделение.

В статье авторов "Анализ изменения статистических факторов возникновения лесных пожаров в Российской Федерации" было проведено исследование данных с помощью методов машинного обучения, таких как метод дерева решений, метод случайного леса и метод k-ближайших соседей. Для представления результатов были построены ROC-кривые для каждого из этих методов и получены оценки точности [1].

Точность каждой из трёх моделей, описанных в вышеуказанной статье, оказалась невысока, но можно сделать вывод о том, что лучшей моделью является простое дерево решений, показавшую наибольшую точность и скорость расчёта. При проведении этого анализа использовались метеоданные, а также публичные данные о пожарах на территории Российской Федерации [1-3].

Проверим эти данные о числе случаев возникновения пожаров на достоверность. На факт возникновения ЛП влияет множество независимых факторов, таких как влажность, количество топлива, температура и т.д. В этом случае, согласно центральной предельной теореме, случайная величина должна иметь нормальное распределение. Для проверки этого утверждения для данных из двух областей Российской Федерации (0), взятых из статистических сборников, выполним QQ-анализ [4].

На рис. 2 показан QQ-анализ числа пожаров в Московской области, а на рис. 3 - QQ-анализ числа пожаров в Новосибирской области.

Из рис. 2 и рис. 3 следует, что коэффициенты корреляции для Московской и Новосибирской областей равны 0,97 и 0,98 соответственно. Это говорит нам о нормальности распределений этих величин и действительности статистических данных.

Выполним анализ случаев гибели людей при пожарах. Данные для анализа гибели на пожарах на территории РФ, взятые с сайта представлены в 0 [4].

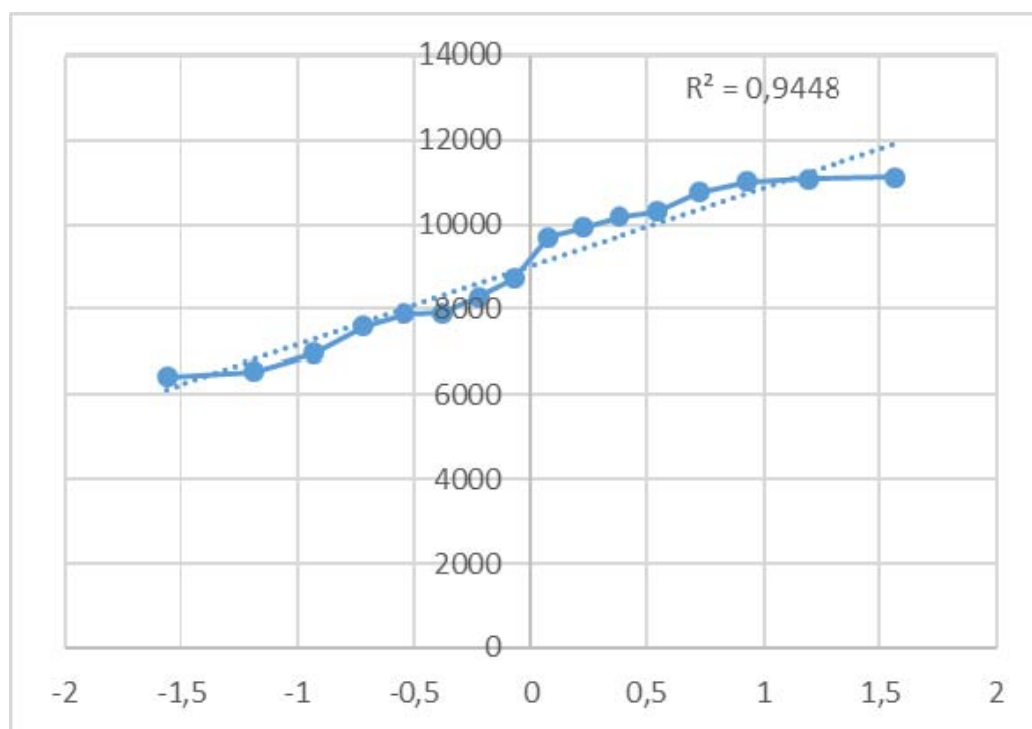


Рис. 2. QQ-анализ числа пожаров в Московской области

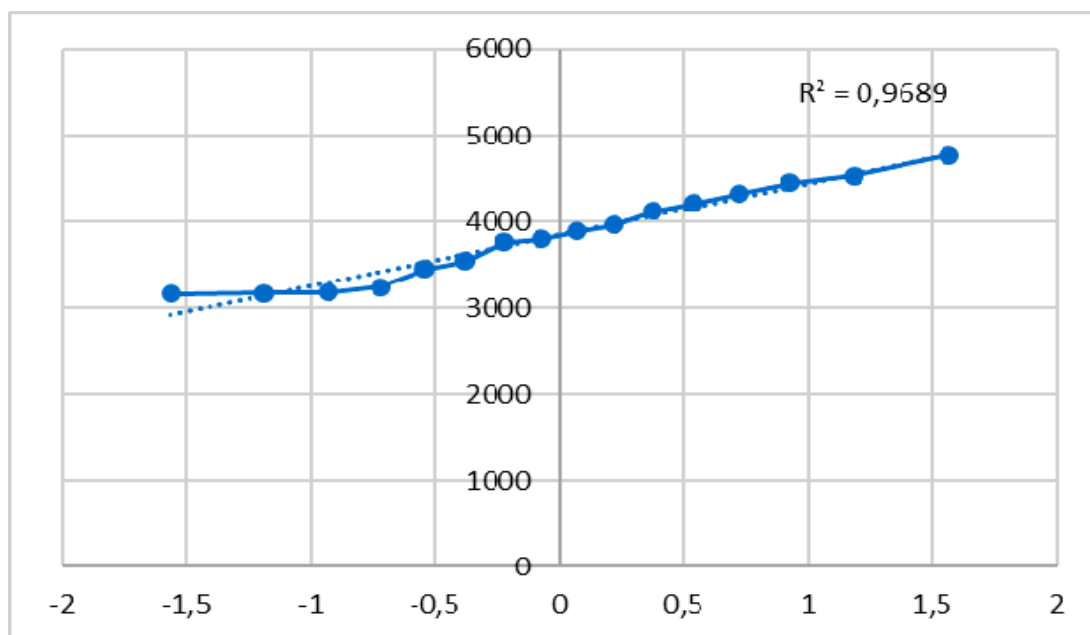


Рис. 3. QQ-анализ числа пожаров в Новосибирской области

Таблица 1

Динамика случаев возникновения ЛП в Московской и Новосибирской областях

Московская область								
Год	2003	2004	2005	2006	2007	2008	2009	2010
Число пожаров	11076	11133	11011	10767	10179	9710	10302	9933

Год	2011	2012	2013	2014	2015	2016	2017	2018
Число пожаров	8725	8286	7902	7878	7585	6959	6552	6426

Новосибирская область								
Год	2003	2004	2005	2006	2007	2008	2009	2010
Число пожаров	4765	4534	4449	4326	4217	4121	3958	3891

Год	2011	2012	2013	2014	2015	2016	2017	2018
Число пожаров	3796	3758	3541	3456	3245	3181	3162	3175

Динамика числа погибших при лесных пожарах

Проверим данные на нормальность, построив экспериментальную и теоретическую квантили (QQ-анализ) (0. 4). Так как коэффициент детерминации R^2 равен 0,96, то только 4% данных по гибели на пожарах не могут быть объяснены нормальным распределением. Таким образом, исходя из нормального распределения исходных данных (что вполне согласуется с центральной предельной теоремой, так как независимых причин возникновения пожаров очень много), можно использовать регрессионные модели.

Динамика количества погибших, линейная регрессия и прогноз с помощью алгоритма экспоненциального сглаживания в пределах доверительного интервала представлены на

рис. 5. Проверка модели линейной регрессии по критериям Фишера и Стьюдента показала, что модель статистически адекватная: F-статистика равна 1095, значимость критерия равна $1,075 \cdot 10^{-14}$; t - статистика равна 33.09, значимость критерия равна $4,785 \cdot 10^{-12}$. Проверка по критерию Дарбина - Уотсона показала, что в данных отсутствуют автокорреляции, что вполне ожидаемо, ведь количество пожаров в один год не зависит от того, сколько их было в предыдущем году.

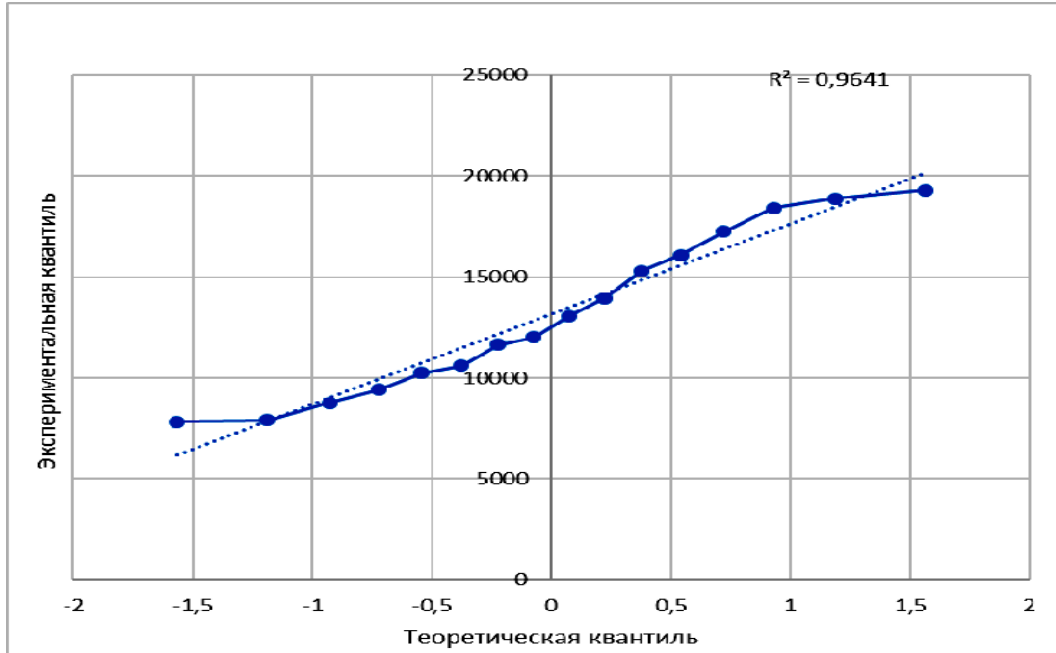


Рис. 4. QQ-анализ динамики количества погибших при ЛП

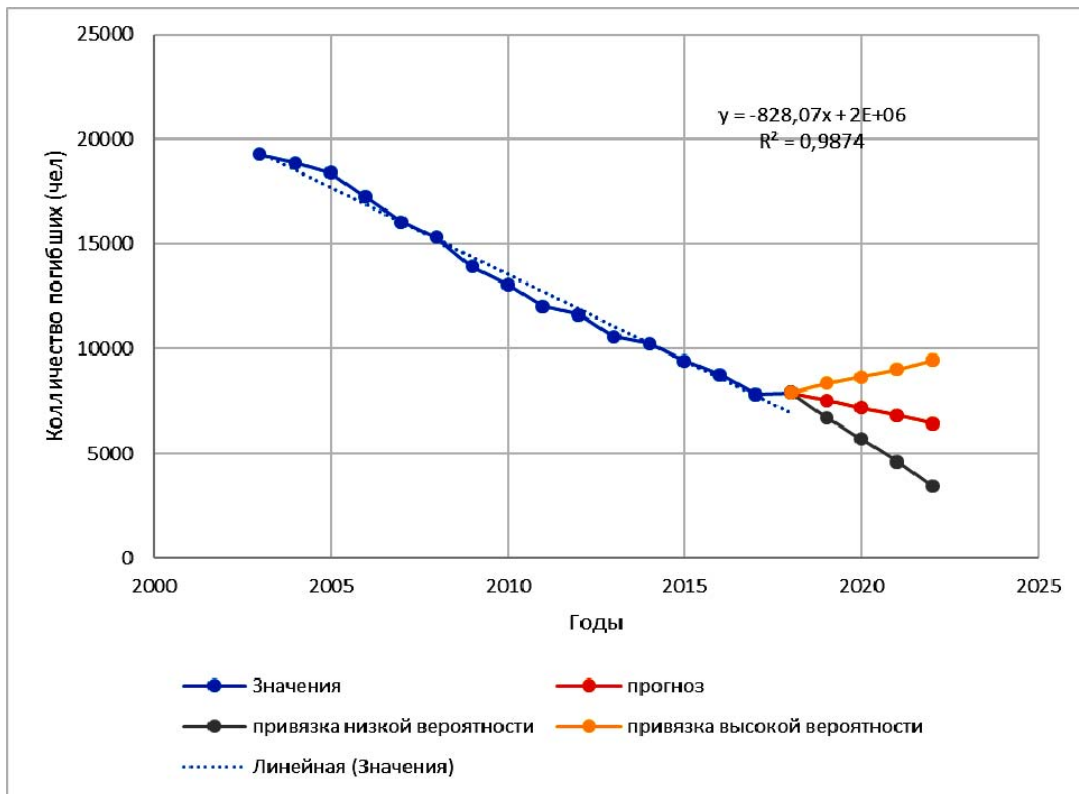


Рис. 5. Динамика количества погибших, линейная регрессия и прогноз

Также в открытом доступе на сайте имеется статистика причин гибели на ЛП [4]. Данные удобно представить в виде гистограммы на рис. 6. На этой гистограмме наглядно видно, что самая распространенная причина смерти при пожаре - это «отравление токсичными продуктами горения при пожаре».

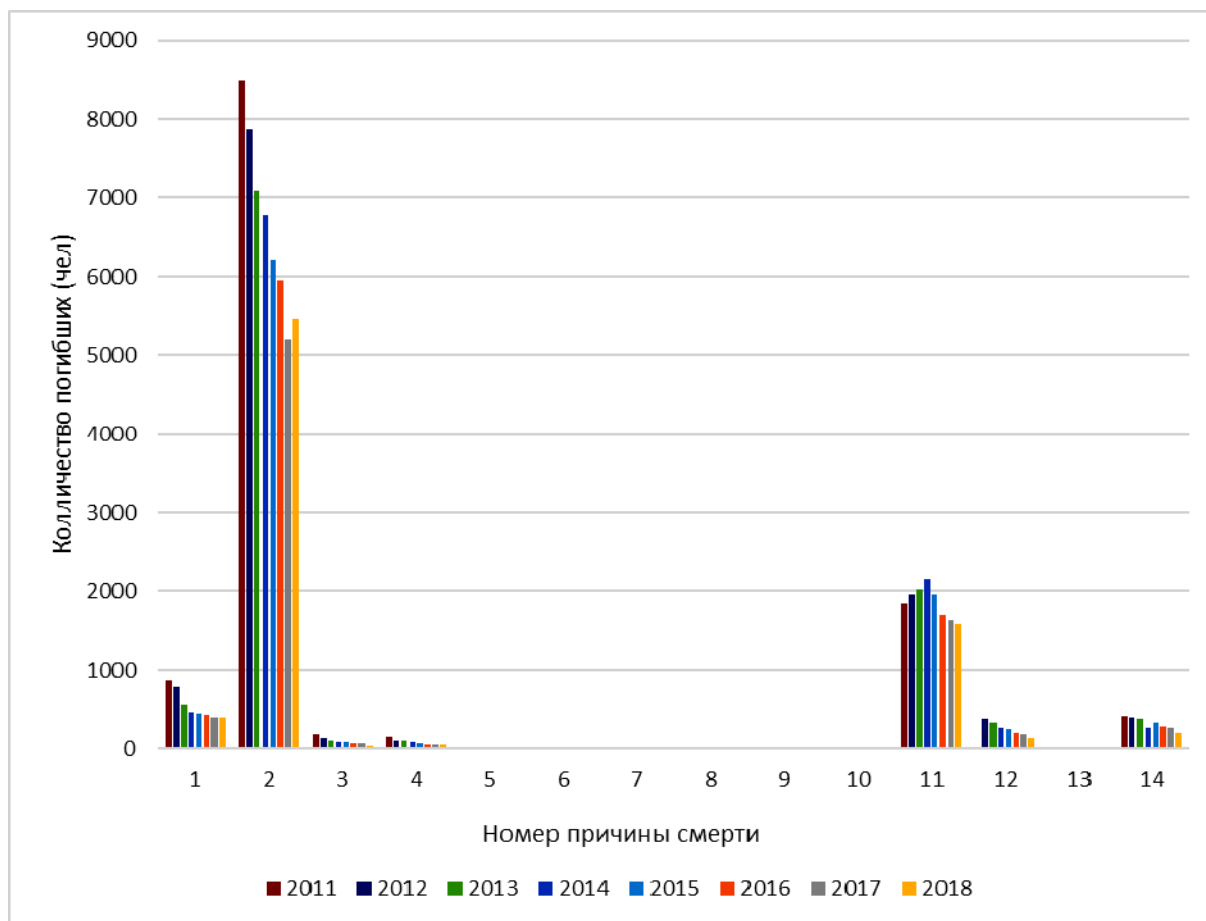


Рис. 6. Гистограмма причин смерти при лесных пожарах

Таким образом, в результате исследований осуществлено прогнозирование количества погибших при лесных пожарах и выявлена основная причина гибели при ЧЛС, вызванных такими пожарами.

В ходе исследований решена новая исследовательская задача по усовершенствованию математической модели прогноза возникновения крупных отслеживаемых лесных пожаров, выявлены значимые факторы, которые в большей степени влияют на возникновение ЛП и их последствия, найдены прочие полезные закономерности.

При этом решены следующие частные задачи:

- произведено преобразование разнотипных входных данных из разных источников к виду, пригодному для исследования и анализа,
- выполнен анализ существующих математических моделей предсказания возникновения лесных пожаров;
- осуществлена проверка основных переменных на пригодность для использования в прогнозе, их проверка на корреляцию между собой для исключения переобучения модели.

Литература

1. Борисова Л.Р., Подрезов Ю.В., Анисимов А.А. Анализ изменения статистических факторов возникновения лесных пожаров в Российской Федерации. //Проблемы безопасности и чрезвычайных ситуаций. – М. – 2019, № 6. – 33-38.
2. <http://aisori-m.meteo.ru/waisori/select.xhtml>.
3. http://rosleshoz.gov.ru/information_systems.
4. fireman.club/literature/.

Сведения об авторах

Борисова Людмила Робертовна, доцент департамента анализа данных, принятия решений и финансовых технологий Финансового университета при Правительстве РФ; доцент кафедры «Высоких технологий в обеспечении безопасности жизнедеятельности» МФТИ (ГУ). Тел. 8-916-245-71-27 E-mail: borisovalr@mail.ru

Подрезов Юрий Викторович, доцент, главный научный сотрудник научно-исследовательского центра ФГБУ ВНИИ ГЧС (ФЦ); заместитель заведующего кафедрой Московского физико-технического института (государственного университета). Тел.: 8-903-573-44-84

Гурьев Яков Сергеевич, студент 4 курса Московского физико-технического института (государственного университета).: Тел.:8-985-988-79-79, e-mail: gurev.yas@phystech.edu

УДК 623.094

DOI: 10.36535/0869-4176-2020-06-4

ОЦЕНКА ЖИВУЧЕСТИ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ ПРИ ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЯХ: МЕТОДИЧЕСКИЕ ОСНОВЫ

Доктор техн. наук **Р.А. Дурнев**, **А.С. Гусева**, кандидат техн. наук **Е.В. Свиридок**
ФГБУ РАРАН

И.В. Жданенко
ФГБУВНИИГОЧС (ФЦ)

В первой статье данной серии рассмотрены методические основы оценки живучести беспилотных летательных аппаратов при воздействии поражающих факторов источников чрезвычайных ситуаций. Приведена вербальная модель сохранения живучести, описан логико-вероятностный метод оценки рассматриваемого свойства. В последующих статьях будут рассмотрены предложения по адаптации и усовершенствованию данного метода применительно к решаемой задаче оценке живучести беспилотных летательных аппаратов.

Ключевые слова: беспилотный летательный аппарат, чрезвычайная ситуация, источник ЧС, поражающие факторы источника ЧС, живучесть, логико-вероятностный метод.