

7. Kadantsev M. M., Negrobov S. O. *Rekreacionnyj monitoring jekosistem botanicheskogo sada Voronezhskogo gosuniversiteta* [Attendance of hard-winged introduced plants of the Voronezh Botanical garden] // Bulletin of the Voronezh state University. 2017. No. 4. P. 35-37.
8. Lepeshkina L. A. *Rekreacionnyj monitoring jekosistem botanicheskogo sada Voronezhskogo gosuniversiteta* [Recreational monitoring of the ecosystems of the Botanical garden of the Voronezh state University] // Collection of scientific works SWORLD. 2012. Vol. 35. No. 3. S. 3-6.
9. Medvedev L. N., Roginsky N. *Katalog kormovyh rastenij listoedov SSSR*. [Catalogue of host plants of leaf beetles of the USSR]. M., 1988. 191 p.
10. Pavlovsky E. N., Bay-Bienko G. Ya. *Opredelitel' nasekomyh, povrezhdajushhih derev'ja i kustarniki polezashhitnyh polos* [Keys to the insects that damage trees and shrubs shelterbelts bands]. M., 1950. P. 207.
11. Bieńkowski, A. O. Leaf-beetles (Coleoptera: Chrysomelidae) of the Eastern Europe. New key to subfamilies, genera and species. M., 2004. 278 p.
12. Salberg, J. Till kännedomen om Haltica Engströmi och dess biologi // Entomologisk Tidskritt. – 1913. – Arg. 34. P. 262-266.

### Сведения об авторе

*Каданцев Михаил Михайлович* – аспирант кафедры экологии и систематики беспозвоночных животных ФГБОУ ВО «Воронежский государственный университет», г. Воронеж, Российская Федерация; email: Mitredel@inbox.ru

### Information about the author

*Kadantsev Mikhail Mikhailovich* – postgraduate student of the Department of ecology and systematics of invertebrate animals of the "Voronezh state University", Voronezh, Russian Federation; email: Mitredel@inbox.ru

DOI: 10.12737/article\_5ab0dfbb946859.24647128

УДК 630\*161:614.842

### ИСПОЛЬЗОВАНИЕ ЗАКОНА БЕНФОРДА ДЛЯ ОЦЕНКИ ДОСТОВЕРНОСТИ СВЕДЕНИЙ О ЛЕСНЫХ ПОЖАРАХ

кандидат технических наук **Р.В. Котельников**<sup>1</sup>  
доктор сельскохозяйственных наук **А.А. Маргынюк**<sup>2</sup>

1– Филиал ФБУ ВНИИЛМ «Центр лесной пирологии, развития технологий охраны лесных экосистем, защиты и воспроизводства лесов», г. Красноярск, Российская Федерация

2– ФБУ «Всероссийский научно-исследовательский институт лесоводства и механизации лесного хозяйства», г. Пушкино, Российская Федерация

Важнейшим управленческим аспектом при организации системы охраны лесов от пожаров является своевременное получение достоверных сведений о площади, пройденной огнем. В настоящее время для контроля за точностью соответствующих сведений создана многоуровневая система специальных мероприятий, включая выездные проверки площадей горельников. Вместе с тем большие объемы информации из разных источников, накопленные в лесопожарных базах данных, позволяют проводить статистическую оценку достоверности (или точности) сведений, значительно сокращая временные и финансовые затраты на выполнение проверочных мероприятий. Математически доказано, что множество чисел, характеризующих реальные природные

объекты, которые могут расти экспоненциально, подчиняются закону Бенфорда. В работе доказана применимость положений закона Бенфорда к оценке достоверности сведений о площадях лесных пожаров на примере анализа встречаемости первых цифр в числах, характеризующих пройденную огнем покрытую лесом площадь на территории Российской Федерации в 2016 году, а также проведена оценка минимального количества значений, необходимых для получения адекватного результата. Кроме того, в работе показана возможность проведения сравнительной оценки достоверности сведений, полученных из разных источников. Учитывая, что отклонения частоты появления отдельных цифр в анализируемых совокупностях значений могут иметь разный знак для разных цифр, для сравнительных оценок предлагается использовать показатель, представляющий среднее значение модулей отклонений вероятности появления соответствующей цифры. Предложенный метод основан на использовании закона Бенфорда и может стать одним из элементов риск-ориентированного подхода для планирования контрольно-надзорных мероприятий в области лесных отношений.

**Ключевые слова:** лесные пожары, статистический анализ, информационные технологии, риск-ориентированный подход.

### APPLICATION OF THE BENFORD LAW IN ASSESSMENT OF WILDFIRE DATA ACCURACY

Candidate of technical sciences **R.V. Kotelnikov**<sup>1</sup>

Doctor of agricultural sciences **A.A. Martynuik**<sup>2</sup>

1 – VNIILM subsidiary “Center for forest pyrology, forest ecosystem conservation technology development, forest protection and regeneration”, Krasnoyarsk, Russian Federation

2 – FBE “Russian Research Institute for Silviculture and Mechanization of Forestry”, Pushkino, Moscow region, Russian Federation

#### Abstract

Timely availability of accurate burned out area data is a key management aspect in forest protection arrangements. Special operation multilevel net-work including field surveys of burned out areas has been established now to verify appropriate data accuracy. In the mean time extensive levels of information from various sources accumulated in wildfire databases enable statistical assessment of the data accuracy drastically reducing time and financial costs of verification operations. Mathematically proven that amount of numbers that specify real natural facilities may grow exponentially due to the Benford law. The paper proves applicability of the Benford law provisions in assessment of wildfire area data accuracy through analysis of first figure occurrence in numbers specifying forest covered burned out area in the Russian Federation territory in 2016 and assessed a minimum set of values needed for an adequate result. In addition the paper highlights an opportunity of variously outsourced data accuracy comparative analysis. Taking into consideration that variation of individual figure occurrence frequency in analyzed value packages may have a different sign for various figures it is offered to apply an indicator representing a mean value of appropriate figure occurrence probability variation modules. The offered procedure based on the Benford law application may be a part of a risk-targeted approach to plan control supervisory operations in forest relations.

**Keywords:** wildfires, statistical analysis, information technologies, risk targeted approach.

#### Введение

Организация эффективной системы охраны лесов от пожаров во многом определяется оперативностью принятия адекватных управленческих решений, а также качеством планирования распределения ресурсов для осуществления плановых и

внеплановых лесопожарных мероприятий. В этой связи важнейшим управленческим аспектом является своевременное получение достоверной информации о лесных пожарах.

Известно, что в настоящее время для контроля за достоверностью сведений о пожарной

опасности в лесах и лесных пожарах, поступающих от региональных лесопожарных служб в Российской Федерации, на федеральном уровне создана, в соответствии с законодательством, многоуровневая система специальных мероприятий. В частности, оперативный контроль осуществляется путем автоматического сравнения региональных данных с данными Информационной системы дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) [4, 6], полученными методом детектирования тепловых аномалий. В случае, если расхождения не устранены, проводится детальная проверка крупных лесных пожаров по данным снимков высокого пространственного разрешения. В отдельных спорных случаях может назначаться выездная проверка для натурного замера площади пожаров с использованием приборов спутниковой навигации в процессе облета на воздушном судне или пешего обхода контура пожара.

Вместе с тем большие объемы информации из разных источников, накопленные в лесопожарных базах данных, позволяют проводить статистическую оценку точности сведений, в том числе полученных из космических источников. Это значительно сокращает временные и финансовые затраты на выполнение проверочных мероприятий.

## Методология

Практический опыт показывает, что количество малых объектов (и явлений) в природе существенно превышает количество больших. Это относится и к лесным пожарам. Математически доказано, что множество чисел, характеризующих реальные природные объекты, которые могут расти экспоненциально (другими словами, темп роста величины пропорционален её текущему значению), подчиняются закону Бенфорда [1, 5, 7-10]. В соответствии с этим законом, в десятичной системе исчисления вероятность первой значащей цифры  $d$  ( $d \in \{1, \dots, 9\}$ ) составляет

$$P(d) = \log_{10} \left( 1 + \frac{1}{d} \right). \quad (1)$$

Расчетные значения вероятности появления в численных значениях первой цифры  $d$  приведены в табл. 1.

Используя представленный математический аппарат, а также классические методы статистики,

покажем возможность использования закона Бенфорда для анализа сведений о площади лесных пожаров.

## Результаты

Для доказательства применимости положений закона Бенфорда к оценке лесопожарной ситуации проведем анализ встречаемости первых цифр в числах, характеризующих пройденную огнем покрытую лесом площадь на территории Российской Федерации в 2016 году, и значений расчётных вероятностей появления в значениях величин показателей первой цифры, предсказанной законом Бенфорда. Данные ИСДМ-Рослесхоз получены по методу детектирования тепловых аномалий и доступны всем желающим через авторизацию сайта государственных услуг. Как видно из рис. 1, вероятность появления соответствующей первой цифры, характеризующей площади лесных пожаров, почти полностью совпадает со значением, предсказываемым законом Бенфорда.

Для доказательства соответствия распределения значений площадей лесных пожаров (и, соответственно, вероятности появления первой цифры в значениях численной величины площади) в исследуемой выборке закона Бенфорда воспользуемся критерием корреляции Пирсона (2)

$$r = \frac{\sum_{i=1}^9 (Pb_i - Pf_i)}{\sqrt{\sum_{i=1}^9 Pb_i^2 \times \sum_{i=1}^9 Pf_i^2}}, \quad (2)$$

где  $r$  – коэффициент корреляции;

$Pf_i$  – фактическая частота появления соответствующей цифры в качестве первой;

$Pb_i$  – расчетная вероятность появления соответствующей цифры по закону Бенфорда. Вычисленное значение коэффициента корреляции составляет  $r = 0.991430$ .

Найдем значение  $t$ -критерия для оценки статистической значимости корреляционной связи:

$$t = \frac{r \times \sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.991430 \times \sqrt{7}}{\sqrt{1-0.991430^2}} = 20.08. \quad (3)$$

Значение расчетной величины  $t$ -критерия превышает его критическую величину  $t_{кр} = 2.365$  по соответствующим таблицам критических значений  $t$ -критерия Стьюдента для числа степеней свободы  $f = 9 - 2 = 7$ , уровне значимости  $\alpha = 0.05$ , следовательно, связь является статистически значимой.

Вероятности появления на первом месте числа соответствующей цифры, рассчитанные на основе закона Бенфорда

D	1	2	3	4	5	6	7	8	9
$P(d)$	30.1 %	17.6 %	12.5 %	9.7 %	7.9 %	6.7 %	5.8 %	5.1 %	4.6 %

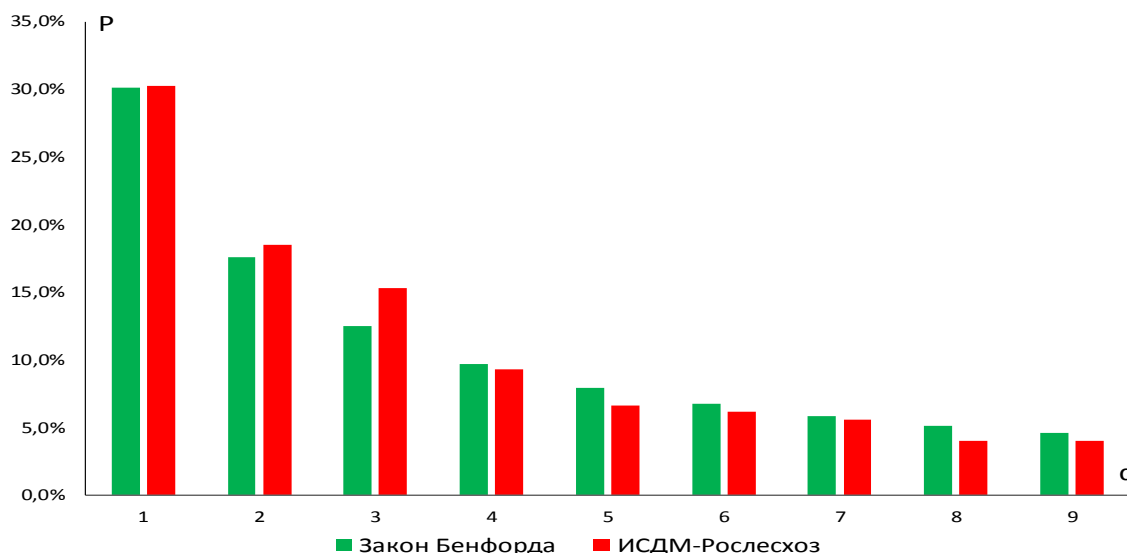


Рис. 1. Зависимость распределения первых цифр множества значений по закону Бенфорда и первых цифр значений площади лесных пожаров (по форме 1-ИСДМ)

То есть можно утверждать, что распределение частоты появления первых цифр в реальных значениях площадей лесных пожаров полностью подчиняются закону Бенфорда, что создает возможность оценивать достоверность сведений, полученных из разных источников, в том числе детектированных разными методами, через сравнение отклонений от закона Бенфорда. При этом следует отметить, что отклонения частоты появления отдельных цифр в анализируемой совокупности цифр связаны с ошибками методов, которые планируется сравнить (или мошенническими действиями, которые необходимо выявить). Они могут иметь разный знак для разных цифр. Однако критерий корреляции Пирсона оценивает корреляционную связь зависимости в целом. Это приводит к тому, что отклонения в отдельных точках могут частично компенсироваться (иметь разный знак).

Таким образом, для выявления отклонений от закона Бенфорда в исследуемой выборке можно использовать среднее значение модулей отклонений вероятности появления соответствующей цифры (4)

$$\delta = \frac{1}{9} \times \sum_{i=1}^9 \left( \left| \frac{P_{fi}}{P_{bi}} - 1 \right| \right) \times 100\%. \quad (4)$$

Эта формула позволяет сравнить две выборки значений (два источника информации, два разных метода замера одного процесса и т. д.) с точки зрения того, какая больше соответствует закону Бенфорда, т. е. реальным природным объектам или явлениям.

Для примера воспользуемся той же общедоступной информацией в ИСДМ-Рослесхоз по площадям крупных лесных пожаров за 2010 год (площадь, покрытая лесом), по которым есть данные региональных диспетчерских служб и ИСДМ-Рослесхоз [3] (соответственно графа 8 и 9 формы 7 ИСДМ). Рис. 2 показывает, что вероятности первых цифр значений площади крупных лесных пожаров, представленных региональными диспетчерскими службами, в доминирующем большинстве случаев отличаются на большую величину от предсказанных по закону Бенфорда, чем полученные по данным ИСДМ-Рослесхоз. Соответствующее среднее значение отклонений для региональных данных составило 14,7 %, для данных ИСДМ-Рослесхоз –

8,3 %. Эти значения еще раз доказывают, что формируемые автоматически данные ИСДМ-Рослесхоз, не зависящие от человеческого фактора, более достоверно характеризуют региональную ситуацию с лесными пожарами.

Для оценки условий, в которых метод применим в зависимости от числа анализируемых случаев лесных пожаров, проведем расчеты частоты появления первой цифры для разных значений количества произвольно выбранных лесных пожаров (в данном случае исследование проводилось на примере данных ИСДМ-Рослесхоз за 2016 год).

Как видно из графика (рис. 3), зависимость для исследуемой выборки соблюдается уже начиная со 100 случаев. В дальнейшем, при увеличении чисел в выборке, значение отклонения от закона Бенфорда уменьшается, а величина t-критерия, а соответственно, и статистическая значимость результата, – увеличивается.

### Выводы

Таким образом, анализ данных о лесных пожарах на основе использования закона Бенфорда позволяет также статистически сравнить различные алгоритмы автоматизированных систем детектирования лесных пожаров. Одной из положительных

сторон метода является возможность автоматизации процесса расчетов и интеграции соответствующих алгоритмов в федеральные информационные системы, используемые в лесном хозяйстве. При этом следует учитывать, что метод достаточно точно работает только на большой выборке данных (более 100).

Следует иметь в виду, что подобный подход применим не только к лесным пожарам. Например, проверка распределения данных по закону Бенфорда уже используется для выявления злонамеренных манипуляций с данными, в том числе для выявления подлогов в финансовых документах [2, 3, 7-8]. Открываются возможности его использования в сфере лесного хозяйства не только для оценки качества учета лесных пожаров в субъектах Российской Федерации, но и для повышения эффективности контрольных мероприятий по использованию, охране, защите и воспроизводству лесов путем сравнения соответствующего комплекса межгодовых показателей. Кроме того, предложенный метод, основанный на использовании закона Бенфорда, может стать одним из элементов риск-ориентированного подхода для планирования контрольно-надзорных мероприятий в области лесных отношений.

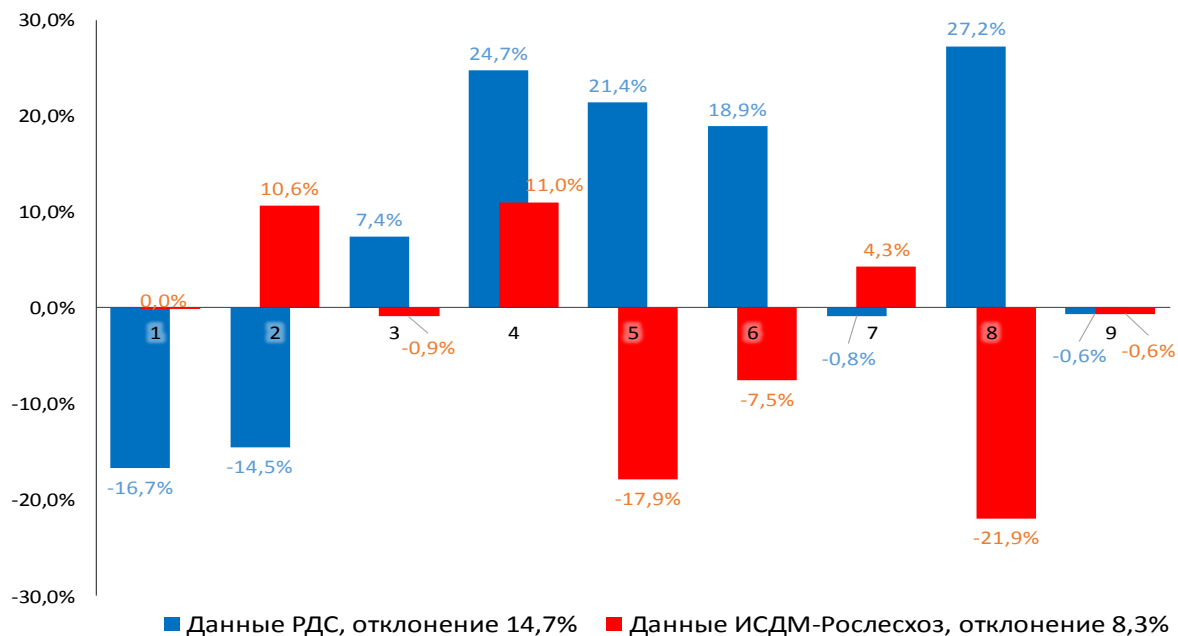


Рис. 2. Сравнение отклонения распределения первых цифр значений площади сопоставленных крупных лесных пожаров от закона Бенфорда для данных региональных диспетчерских служб (РДС) и данных ИСДМ-Рослесхоз

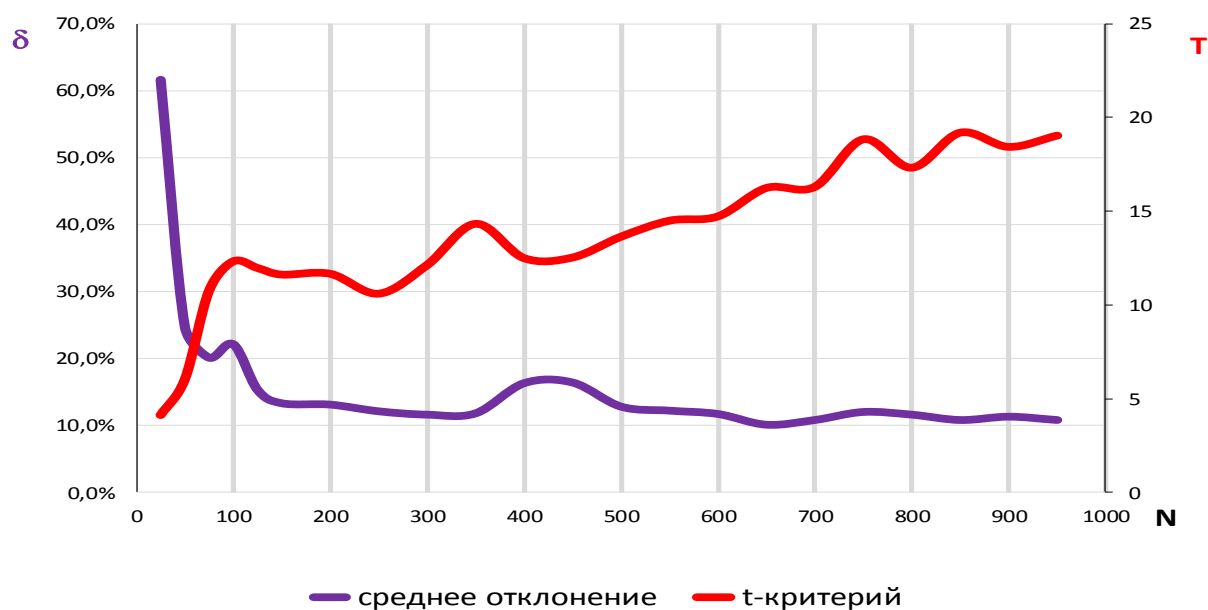


Рис. 3. Среднее значение ( $\delta$ ) отклонения реальной зависимости от вероятности, предсказываемой законом Бенфорда для разных значений количества лесных пожаров в выборке (N), а также величина t-критерия статистической значимости значений корреляции

### Библиографический список

1. Алексеев, М. А. Применимость закона Бенфорда для определения достоверности финансовой отчетности [Текст] / М. А. Алексеев // Вестник НГУЭУ. – 2016. – № 4. – С. 114-128.
2. Воробьев, А. Аномальные цифры финансовых махинаций [Электронный ресурс] / А. Воробьев // Бухгалтерия.ру. – Режим доступа: <http://www.buhgalteria.ru/article/n28684>.
3. Кечкова, И. В. Закон Бенфорда как метод выявления мошеннических действий [Текст] / И. В. Кечкова // Молодой ученый. – 2017. – № 11. – Ч.2. – С. 219-221.
4. Котельников, Р. В. Технология идентификации спутниковых данных о лесных пожарах с данными наземного и авиационного мониторинга с использованием ИСДМ-Рослесхоз [Текст] / Р. В. Котельников, Е. В. Флитман // Современные проблемы дистанционного зондирования Земли из космоса. – 2007. – Т. I. – № 4. – С. 162-166.
5. Кувакина, Л. В. Закон Бенфорда: сущность и применение [Текст] / Л. В. Кувакина, А. Ф. Долгополова // Современные проблемы науки и образования. – 2013. – № 6. – С. 74-76.
6. Организация работы со спутниковыми данными в информационной системе дистанционного мониторинга лесных пожаров Федерального агентства лесного хозяйства (ИСДМ-Рослесхоз) [Текст] / Е. А. Лупян [и др.] // Современные проблемы дистанционного зондирования Земли из космоса. – 2015. – Т. 12. – № 5. – С. 222-250.
7. Badal-Valero, E. *Combining Benford's Law and machine learning to detect money laundering. An actual Spanish court case* [Text] / E. Badal-Valero, J. A. Alvarez-Jareno, J. M. Pavia // FORENSIC SCIENCE INTERNATIONAL. – Vol. 282 – P. 24-34.
8. Hullemann, S. *Application of Benford's law: a valuable tool for detecting scientific papers with fabricated data? A case study using proven falsified articles against a comparison group* [Text] / S. Hullemann, G. Schupfer, J. Mauch // ANAESTHESIST. – Vol. 66. – No. 10. – P. 795-802.
9. Kruger, P. S. *The power of one: Benford's law* [Text] / P. S. Kruger, V. S. S. Yadavalli // SOUTH AFRICAN JOURNAL OF INDUSTRIAL ENGINEERING. – Vol. 28. – No. 2. – P. 1-13.

10. Shiu, P. *Benford's law: Theory and application* [Text] / P. Shiu // *Mathematical Gazette*. – Vol. 100. – No. 549. – P. 564-567.

### References

1. Alekseev M. A. *Primenimost' zakona Benforda dlya opredeleniya dostovernosti finansovoy otchetnosti* [Applicability of the Benford law to verify financial reporting accuracy]. *Vestnik NGUEU*. 2016. № 4. P. 114-128.
2. Vorob'ev A. *Anomal'nye tsifry finansovykh makhinatsiy. Bukhgalteriya.ru* [Financial fraud abnormal figures]. Available at: <http://www.buhgalteriya.ru/article/n28684>.
3. Kechkova I. V. *Zakon Benforda kak metod vyyavleniya moshennicheskikh deystviy* [Financial fraud abnormal figures]. *Molodoy uchenyy*. [Young scientist]. 2017. № 11. Ch. 2. P. 219-221.
4. Kotelnikov R. V., Flitman E. V. *Tekhnologiya identifikatsii sputnikovykh dannykh o lesnykh pozharakh s dannymi nazemnogo i aviatsionnogo monitoringa s ispol'zovaniem ISDM-Rosleskhoz* [Application of the ISDM-Rosleskhoz to identify wildfire remote sensing data with ground and aerial monitoring data]. *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. 2007. T. I. № 4. P. 162-166.
5. Kuvakina L. V., Dolgoplova A. F. *Zakon Benforda: sushchnost' i primeneniye* [The Benford law its implications and applications]. *Sovremennyye problemy nauki i obrazovaniya*. 2013. № 6. P. 74-76.
6. Lupyan E. A. [et al.] *Organizatsiya raboty so sputnikovymi dannymi v informatsionnoy sisteme distantsionnogo monitoringa lesnykh pozharov Federal'nogo agentstva lesnogo khozyaystva (ISDM-Rosleskhoz)* [Arrangements of satellite data operation in wildfire remote monitoring information system of the federal Forestry Agency (ISDM-Rosleskhoz)]. *Sovremennyye problemy distantsionnogo zondirovaniya Zemli iz kosmosa*. [Modern problems of remote sensing of the Earth from space]. 2015. T. 12. № 5. P. 222-250.
7. Badal-Valero E., Alvarez-Jareno J. A., Pavia, J. M. *Combining Benford's Law and machine learning to detect money laundering. An actual Spanish court case*. *FORENSIC SCIENCE INTERNATIONAL*. Vol. 282. P. 24-34.
8. Hulleman S., Schupfer G., Mauch J. *Application of Benford's law: a valuable tool for detecting scientific papers with fabricated data? A case study using proven falsified articles against a comparison group*. *ANAESTHESIST*. Vol. 66. No. 10, P. 795-802.
9. Kruger P. S., Yadavalli V. S. S. *The power of one: Benford's law*. *SOUTH AFRICAN JOURNAL OF INDUSTRIAL ENGINEERING*. Vol. 28. No. 2. P. 1-13.
10. Shiu P. *Benford's law: Theory and application*. *Mathematical Gazette*. Vol. 100. No. 549. P. 564-567.

### Сведения об авторах

*Котельников Роман Владимирович* – директор филиала ФБУ ВНИИЛМ «Центр лесной пирологии, развития технологий охраны лесных экосистем, защиты и воспроизводства лесов», г. Красноярск, Российская Федерация, кандидат технических наук; e-mail: [kotelnikovrv@mail.ru](mailto:kotelnikovrv@mail.ru)

*Мартынюк Александр Александрович* – директор ФБУ «Всероссийский научно-исследовательский институт лесоводства и механизации лесного хозяйства», г. Пушкино, Российская Федерация, доктор сельскохозяйственных наук; e-mail: [vniilm\\_martinuk@mail.ru](mailto:vniilm_martinuk@mail.ru)

### Information about authors

*Kotelnikov Roman* – director of the VNIILM subsidiary “Center for forest pyrology, forest ecosystem conservation technology development, forest protection and regeneration”, Krasnoyarsk, Russian Federation, candidate of technical sciences; e-mail: [kotelnikovrv@mail.ru](mailto:kotelnikovrv@mail.ru)

*Martyniuk Alexander* – director of the Russian Research Institute for Silviculture and Mechanization of Forestry, Pushkino, Moscow region, Russian Federation, doctor of agricultural sciences; e-mail: [vniilm\\_martinuk@mail.ru](mailto:vniilm_martinuk@mail.ru)