

УДК 667.6
DOI: 10.15827/0236-235X.117.143-147

Дата подачи статьи: 19.09.16
2017. Т. 30. № 1. С. 143–147

РАЗРАБОТКА И ВНЕДРЕНИЕ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ЦВЕТОВОГО ОТКЛОНЕНИЯ ПОЛИМЕРНОГО ПОКРЫТИЯ ОЦИНКОВАННОЙ ПОЛОСЫ

В.М. Осколков, аспирант, vasilij_oskolkov@outlook.com;
И.А. Варфоломеев, к.т.н., доцент, igor.varf@gmail.com;
Л.Н. Виноградова, к.т.н., доцент, lvinogradova@bk.ru;
Е.В. Ершов, д.т.н., профессор, eve@chs.u.ru
*(Череповецкий государственный университет,
просп. Луначарского, 5, г. Череповец, 162600, Россия)*

Статья посвящена результатам исследования метода повышения качества оцинкованной полосы с полимерным покрытием за счет уменьшения цветового отклонения, используя методы моделирования.

Предложенная модель прогнозирования цветового отклонения состоит из трех подмоделей, каждая из которых прогнозирует одну из координат цветового пространства CIE Lab. Основой каждой из подмоделей является алгоритм машинного обучения Random Forest. По спрогнозированным значениям цветовых координат вычисляется полное цветовое отклонение.

Рассмотрен алгоритм построения дерева принятия решений. Описаны основные параметры, влияющие на цветовое отклонение, получаемые из трех источников: данные из сертификата о краске, характеристика входного рулона для покраски, технологические параметры.

Разработан подход быстрой и эффективной интеграции модели прогнозирования в существующую ИТ-инфраструктуру предприятия с помощью трансляции модели в БД. Разработан скрипт для трансляции модели на используемые при разработке ПО на промышленных предприятиях языки программирования SQL и .NET.

Описаны этапы трансляции модели прогнозирования с языка R на язык SQL: генерация кода, заполнение таблиц. Прогнозирование в БД осуществляется за 0,3 секунды, этого достаточно для применения в производстве в режиме реального времени.

Использование разработанной модели делает возможным прогнозирование полного цветового отклонения полимерного покрытия со средней относительной ошибкой модели 6,1 %.

Ключевые слова: интеграция моделей, генерация кода, Random Forest, регрессия, дерево решений, прогнозирование, цветовое отклонение, предварительно окрашенный прокат, полимерное покрытие, coil coating.

В настоящее время на рынках России, США и Западной Европы по-прежнему сохраняются относительно высокие цены на листовую прокат. Одним из главных направлений металлургической промышленности является производство предварительно окрашенного проката – продукции с высокой добавленной стоимостью. В связи с этим предприятия ищут новые пути для увеличения конкурентных преимуществ своей продукции. Например, «Северсталь» расширяет производство и готовит к запуску третью линию по покрытию металла, которая включает агрегат полимерных покрытий.

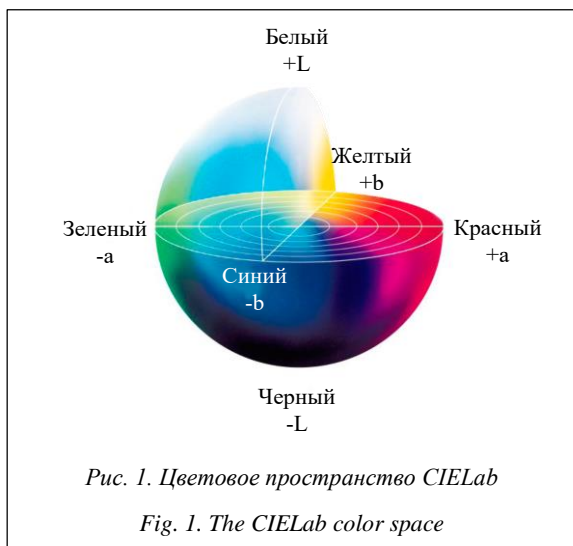
Процесс производства предварительно окрашенного проката оцинкованной полосы по технологии coil coating осуществляется на агрегате полимерных покрытий и состоит из следующих этапов: обработка поверхности полосы хроматирующим раствором, нанесение грунтовочного слоя, сушка грунтовочного слоя, нанесение эмали (отделочного слоя) окрасочными валами, сушка отделочного слоя, смотка полосы в рулоны.

Компания «Тримо-ВСК» отмечает, что 5–18 % продукции проката с полимерным покрытием, поставляемой отечественными заводами, имеет отклонения от требований заказчика и норм ГОСТ, тогда как у зарубежных компаний доля такой продукции составляет в среднем 0,28 % [1]. Также от-

мечается, что партии поставляемой большинством российских компаний продукции одного цвета отличаются оттенками [1]. Поэтому так актуальна для российских металлургических предприятий задача повышения качества данной продукции.

Нанесение покрытий осуществляется с помощью грунтовочного и отделочного коутеров, которые состоят из наносящего валька, заборного валька и ванны с краской. Цвет *лакокрасочного материала* (ЛКМ) определяется согласно каталогу цветовых гамм RAL в соответствии со стандартными образцами. Измерение цветовых координат полимерного покрытия образца выполняется с помощью спектрофотометра во время лабораторных испытаний готовой продукции.

Геометрический смысл цветового отклонения заключается в том, что оно описывает разницу между двумя цветовыми воздействиями, определяемую как евклидово расстояние между точками, описывающими эти цвета в пространстве CIE Lab (рис. 1) с системой координат L^* , a^* , b^* . Данное пространство разработано Международной комиссией по свету с целью упрощения обмена информацией о цвете. Оно образуется в результате представления основных цветов (белого, черного, синего, желтого, зеленого, красного) в системе координат.



Значение координаты L характеризует цветовое различие в светлую (белую) и темную (черную) стороны относительно стандарта. Положительное значение координаты a говорит о том, что цвет образца более красный, чем стандарт, отрицательное – цвет образца более зеленый. Координата b показывает отклонение по желтой и синей составляющим цвета.

Полное цветовое различие рассчитывается по формуле $\Delta E = \sqrt{\Delta L^2 + \Delta a^2 + \Delta b^2}$, где $\Delta L = L - L_0$, L – координата образца, L_0 – координата стандарта; $\Delta a = a - a_0$, a – координата образца, a_0 – координата стандарта; $\Delta b = b - b_0$, b – координата образца, b_0 – координата стандарта.

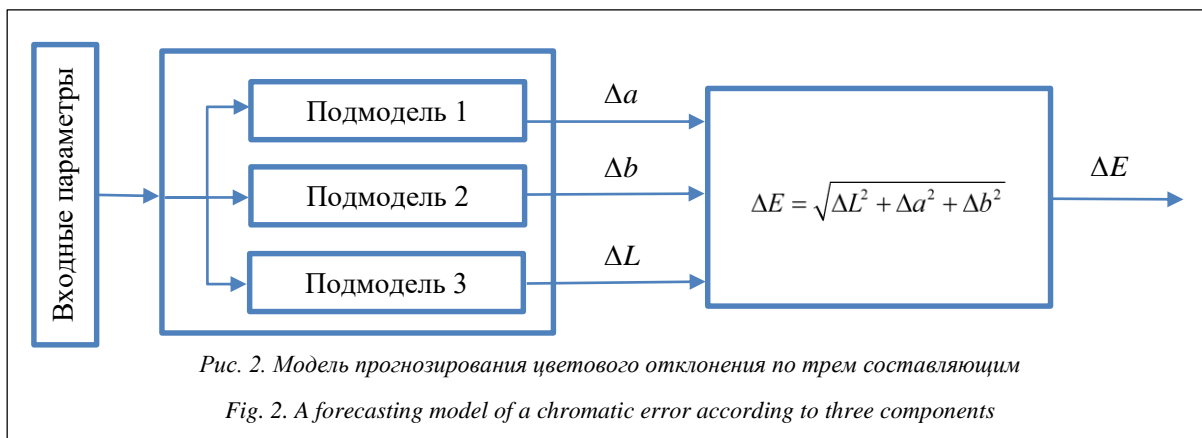
Построение модели прогнозирования

Применение моделирования сложного технологического процесса позволяет решать различные задачи по повышению качества продукции в металлургии [2]. Результаты прогнозирования можно использовать для подбора оптимальных значений технологических параметров для повышения качества продукции [3]. Для управления технологическими параметрами с целью уменьшения значения

цветового отклонения отделочного покрытия готовой продукции с помощью методов машинного обучения была построена прогнозная модель, состоящая из трех подмоделей, каждая из которых осуществляет прогнозирование одной из цветовых координат (рис. 2). По полученным координатам по вышеприведенной формуле вычисляется полное цветовое отклонение.

Входные параметры собираются из нескольких источников: данные из сертификата о краске, характеристики входного рулона для покраски, технологические данные. Краска обладает следующими свойствами: вязкость, плотность жидкой краски, плотность сухой краски, сухой остаток по объему, сухой остаток по весу, толщина сухой пленки, толщина жидкого слоя эмали, укрывистость, субстрат, приведенная толщина металла, блеск, адгезия, Т-изгиб, прочность при растяжении по Эриксену, прочность при обратном ударе, стойкость к истиранию растворителем, твердость по карандашу, начальное цветовое отклонение ЛКМ ($\Delta L'_{st}, \Delta a'_{st}, \Delta b'_{st}, \Delta E'_{st}$). Характеристики входного рулона для покраски: марка стали, толщина листа, температура перед входом в агрегат, толщина цинкового покрытия, теплоемкость металла.

Технологические данные собираются в процессе производства с датчиков основных узлов агрегата полимерных покрытий: разматыватель, накопитель полосы на входе, печь грунтовочного покрытия, печь отделочного покрытия, секция очистки, химический коутер, грунтовочный коутер, отделочный коутер, выходной накопитель, моталка. Технологический процесс характеризуется следующими основными параметрами: натяжение полосы, скорость полосы на каждом технологическом участке, проводимость чистящего раствора, температура чистящего раствора, температура сушильного воздуха, пиковая температура металла в печи грунтовочного покрытия, пиковая температура металла в печи отделочного покрытия, температура охлаждающей воды после печи грунтовочного слоя, температура воды после печи отделочного слоя, зазор, усилие и скорость вращения на каждом ролике, температура и скорость вра-



щения вентиляторов в каждой из семи печных зон, толщина покрытия ЛКМ. Все эти данные накапливаются по каждому метру производимой продукции.

При реализации каждой из подмоделей использовался алгоритм машинного обучения Random Forest, который применяется в решении задач прогнозирования в различных отраслях [4, 5]. Данный алгоритм использует ансамбль решающих деревьев и сочетает в себе два метода – беггинга (bagging) и случайных подпространств. Дерево принятия решений представляет собой бинарное дерево. На ребрах (ветках) дерева – атрибуты, от которых зависит целевая функция, в узлах – атрибуты, по которым различаются случаи, в терминальных вершинах (листьях) – значения целевой функции. Общий принцип построения дерева решений – выбор атрибута из входных атрибутов и разбиение значений по этому атрибуту. Дальнейшее разбиение производится аналогично для каждого. Важным этапом построения является выбор критерия атрибута таким образом, чтобы в результате разбиения получаемые подмножества принадлежали к одному классу [6]. При разработке модели прогнозирования для каждой составляющей цветового отклонения автоматически строятся 300 деревьев. Расчет интервала, в котором находится результат, производится путем голосования каждого дерева из ансамбля. Пример дерева решений для прогнозирования Δa представлен на рисунке 3.

Использование описанной модели делает возможным прогнозирование полного цветового отклонения полимерного покрытия со средней относительной ошибкой модели 6,1 %. Это позволяет более гибко реагировать на требования клиента, повысить качество продукции, снизить себестоимость и сроки производства.

Внедрение модели прогнозирования цветового отклонения в систему управления технологическим процессом

Построенная модель прогнозирования решает задачу статистической обработки большого набора данных. Для реализации был выбран язык программирования R, который широко используется для статистического анализа, предсказания и визуализации данных [7]. По данным исследования, наиболее эффективной реализацией алгоритма Random Forest с точки зрения точности является его параллельная реализация parRF_t, включенная в пакет caret для языка R [8].

После разработки модели необходимо интегрировать ее в существующие ИТ-системы компании. Простота интеграции для крупных промышленных предприятий является важным фактором при использовании нового программного продукта. Технологические данные с агрегата записываются в БД. Исходя из этого можно сделать вывод о том, что максимально эффективным с точки зрения

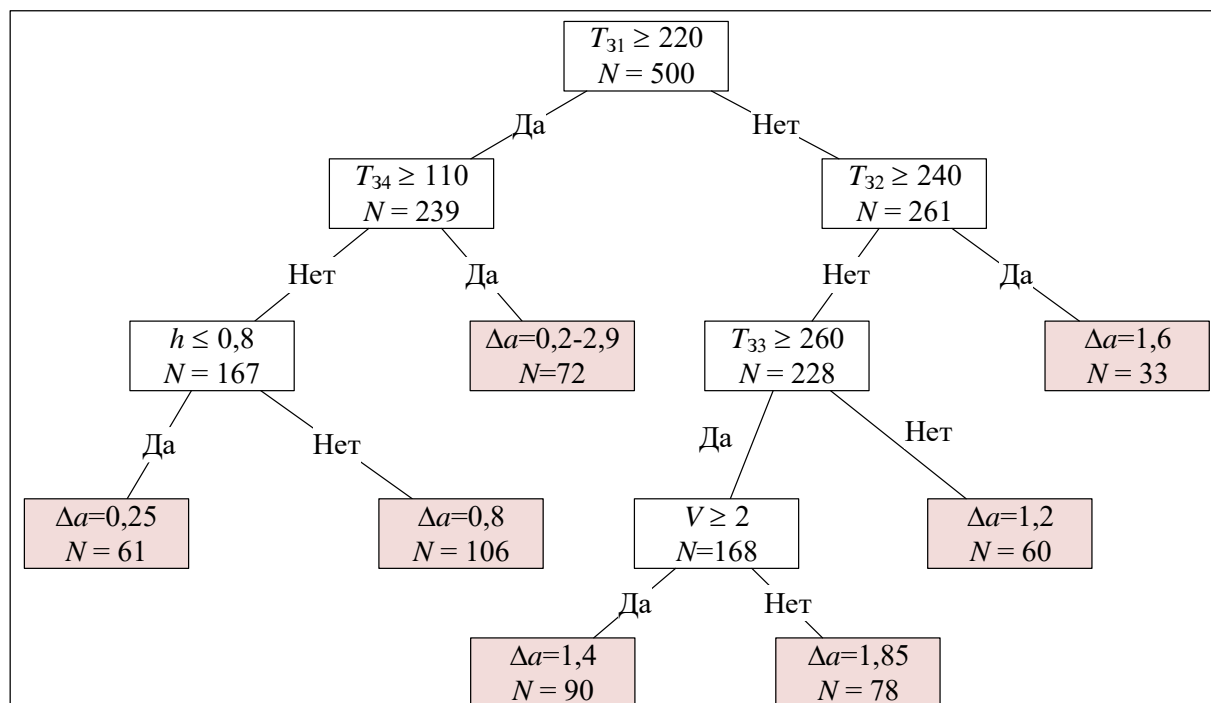


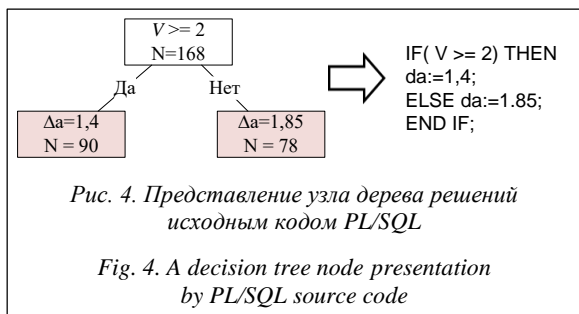
Рис. 3. Дерево принятия решений для величины Δa :
 T_{31}, \dots, T_{34} – температура сушки в зоне печи; h – толщина отделочного слоя ЛКМ; V – скорость движения полосы в печи сушки отделочного покрытия; N – количество объектов в узле

Fig. 3. A decision tree for Δa value

IT-архитектуры будет внедрение модели прогнозирования на уровне БД.

Для упрощения интеграции модели был разработан скрипт на языке R, который в автоматическом режиме транслирует модель, построенную с помощью алгоритма Random Forest с языка R в программный код на языке SQL.

Каждый узел полученных деревьев может быть преобразован в конструкцию вида if-then-else. Таким образом, итоговая модель представляет собой набор соответствующих вложенных операторов языка PL/SQL (рис. 4).



Генерация SQL-кода

Генерация if-then-else конструкций приводит к большому количеству строк кода (1 дерево – 1 200 строк кода). В данном случае БД долго исполняет запрос, поэтому эффективнее транслировать деревья в записи.

На первом этапе генерируется код для создания таблицы (табл. 1) в БД со следующими полями: номер дерева, имя параметра, значение условия перехода, ссылка на узел дерева при истинном результате выполнения условия, ссылка на узел дерева при ложном результате выполнения условия, тип узла, значение для терминального узла.

На втором этапе для каждого узла всех деревьев в ансамбле генерируется конструкция INSERT для вставки записи в созданную таблицу. Обход дерева производится рекурсивно в глубину, начиная с корня дерева. Каждый узел характеризуется следующими параметрами: имя, тип узла, для нетерминальных узлов – логическое условие перехода к дочерним узлам, ссылки на дочерние узлы, для тер-

минальных узлов – результирующее значение. Поле тип узла (node_type) содержит определенное значение для каждого вида узла: 0 – корневой, 1 – внутренний, 2 – терминальный.

Таблица 1

Структура таблицы для хранения деревьев

Table 1

The structure of a storing tree table

№	Имя	Тип данных	Описание
	Id	int	ID узла
	tree_num	int	Номер дерева
	param_name	nvarchar(10)	Наименование параметра
	cond_value	float	Значение условия перехода
	true_id	int	Ссылка, если ИСТИНА
	false_id	int	Ссылка, если ЛОЖЬ
	node_type	int	Тип узла
	value	int	Значение

После завершения работы программы сгенерированный код исполняется в системе управления БД Oracle для создания в ней таблицы с записями (табл. 2), соответствующими модели прогнозирования.

Прогнозирование новой записи осуществляется с помощью вызова специальной процедуры, хранимой в БД, которая вычисляет значение для каждого дерева и усредняет его. Вычисление выполняется за 0,3 секунды. Этого достаточно для использования в производстве в режиме реального времени.

Заключение

Таким образом, собрана вся информация, накапливаемая в процессе производства и доступная в сертификатах о краске. Используя эти параметры, авторы разработали модель, которая позволяет прогнозировать полное цветное отклонение полимерного покрытия со средней относительной ошибкой модели 6,1 %. В основе модели прогнозирования лежит алгоритм Random Forest, а именно – его параллельная реализация в пакете parRF_t языка R. Кроме того, разработан подход, который

Таблица 2

Фрагмент таблицы БД, содержащей узлы дерева

Table 2

A fragment of a database table containing tree nodes

	ID	TREE_NUM	PARAM_NAME	COND_VALUE	TRUE_ID	FALSE_ID	NODE_TYPE	VALUE
1	1101	3	V178	2,8515	1102	1103	1	
2	1104	3					2	1,85
3	1100	3	V207	0,024	1101	1104	1	
4	1098	3	V212	0,00605	1099	1100	1	
5	1106	3					2	0,35
6	1107	3					2	1,4
7	1105	3	V210	0,1085	1106	1107	1	
8	1097	3	V215	0,0037	1098	1105	1	

позволяет быстро и эффективно внедрять разработанные модели прогнозирования на предприятии.

ПО «Генератор программного кода моделей Random Forest» было зарегистрировано в Роспатенте, свидетельство № 2016616932 от 22.06.2016.

Литература

1. Игнатенко Т. Оцинкованный и окрашенный прокат: нерадужные перспективы // *Металлоснабжение и сбыт*. 2015. № 4. С. 36–45.
2. Виноградова Л.Н., Шумилова Е.С. Алгоритмическое обеспечение системы прогнозирования параметров макроструктуры и качества агломерата на основе нейронной сети и фрактальной аппроксимации // *Инновационные тенденции развития российской науки: сб. матер. III Междунар. науч.-практич. конф.* 2010. С. 282–284.
3. Варфоломеев И.А., Ершов Е.В., Богачев Д.В. Оптико-

электронный контроль цветового отклонения полимерного покрытия оцинкованной полосы // *Дефектоскопия*. 2015. № 5. С. 44–51.

4. Картиев С.Б., Курейчик В.М. Алгоритм классификации, основанный на принципах случайного леса, для решения задачи прогнозирования // *Программные продукты и системы*. 2016. № 2. С. 11–15.
5. Иванова Е.И. Интеллектуальная информационная система для решения задач прогнозирования неисправностей вагонного оборудования на железнодорожном транспорте // *Программные продукты и системы*. 2015. № 4. С. 231–236.
6. Чистяков С.П. Случайные леса: обзор // *Тр. Карельского НЦ РАН*. 2013. № 2. С. 117–136.
7. Кабаков Р.И. *R в действии. Анализ и визуализация данных на языке R*. М.: ДМК Пресс, 2014. С. 33–36.
8. Fernández-Delgado M., Cernadas E., Barro S., Amorim D. Do we need hundreds of classifiers to solve real world classification problems? *Jour. of Machine Learning Research*. 2014, no. 15, pp. 3133–3181.

Software & Systems

DOI: 10.15827/0236-235X.117.143-147

Received 19.09.16

2017, vol. 30, no. 1, pp. 143–147

DEVELOPMENT AND IMPLEMENTATION OF A FORECASTING MODEL FOR A CHROMATIC ERROR OF GALVANIZED STRIP POLYMERIC COATING

V.M. Oskolkov¹, Postgraduate Student, vasilij_oskolkov@outlook.com

I.A. Varfolomeev¹, Ph.D. (Engineering), Associate Professor, igor.varf@gmail.com

L.N. Vinogradova¹, Ph.D. (Engineering), Associate Professor, lvinogradova@bk.ru

E.V. Ershov¹, Dr.Sc. (Engineering), Professor, eve@chs.u

¹ Cherepovets State University, Lunocharskogo Ave. 5, Cherepovets, 162600, Russian Federation

Abstract. The article presents the results of the study on quality improvement method for galvanized strip polymeric coating using modelling methods for chromatic deviation reduction. A predictive model of the chromatic deviation consisting of 3 sub-models is proposed; each sub-model predicts one CIELab color space model coordinate. Each sub-model is based on Random Forest machine learning algorithm. Full chromatic deviation output value is calculated from predictive coordinates.. Each sub-model is based on Random Forest machine learning algorithm.

The paper considers a decision tree algorithm. It also describes the main parameters affecting chromatic deviation. Those parameters are received from 3 sources: paint certificate values, characteristics of an incoming strip coil for further painting, process parameters.

The authors have developed an approach for prompt and efficient integration of the mentioned forecasting model into existing IT infrastructure by model translation into a database. The developed script allows translating the model into programming languages used for industrial control systems (SQL, .NET).

The paper describes the following stages of forecasting model translation from R language into SQL language: code generation, filling the tables. Forecasting in a database takes 0,3 seconds which is enough for real time mode production. Application of the developed model allows forecasting chromatic deviation of a polymeric coating with a mean error of 6,1 %.

Keywords: model integration, code generation, random forest, regression, decision tree, prediction, chromatic aberration, pre coated rolling, polymer coating, coil coating.

References

1. Ignatenko T. Galvanized and coloured roll stock: doesn't look good. *Metallsnabzhenie i sbyt* [Metal Supply and Sales]. 2015, no. 4, pp. 36–45.
2. Vinogradova L.N., Shumilova E.S. Algorithmic support of a forecasting system for sinter macrostructure and quality parameters based on a neural network and fractal approximation. *Innovatsionnye tendentsii razvitiya rossiyskoy nauki: sb. mater. 3 Mezhdunar. nauch.-praktich. konf.* [Proc. 3rd Int. Science and Practice Conf. of Young Scientists “Innovative trends of Russian Science”]. 2010, pp. 282–284 (in Russ.).
3. Varfolomeev I.A., Ershov E.V., Bogachev D.V. Electro-optical monitoring of the chromatic error of the polymeric coating of a galvanized strip. *Defektoskopiya* [Russian Jour. of Nondestructive Testing]. 2015, no. 5, pp. 44–51 (in Russ.).
4. Kartiev S.B., Kureychick V.M. Classification algorithm based on random forest principles for a forecasting problem. *Programmnye produkty i sistemy* [Software & Systems]. 2016. no. 2, pp. 11–15 (in Russ.).
5. Ivanova E.I. Intelligent information system for rolling stock carriage equipment failures forecasting on railways. *Programmnye produkty i sistemy* [Software & Systems]. 2015. no. 4, pp. 231–236 (in Russ.).
6. Chistyakov S.P. Random forests: an overview. *Trudy KarNTs RAN* [Trans. of the Karelian Research Centre of RAS]. 2013, no. 2, pp. 117–136 (in Russ.).
7. Kabakov R.I. *R v deystvii. Analiz i vizualizatsiya dannykh na yazyke R* [R in Action. Data Analysis and Graphics with R]. Moscow, DMK Press, 2014, pp. 33–36 (in Russ.).
8. Fernández-Delgado M., Cernadas E., Barro S., Amorim D. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *Jour. of Machine Learning Research*. 2014, no. 15, pp. 3133–3181.