

**УДК 33****Методики использования Big Data в прогнозировании и  
предотвращении неисправностей в штамповочном производстве:  
экономические аспекты****Хохлов Алексей Николаевич**

Кандидат технических наук,  
Начальник отделения,  
Центральный научно-исследовательский институт машиностроения  
141070, Российская Федерация, Королёв, ул. Пионерская, 4;  
e-mail: techmail1@mail.ru

**Спивак Антон Дмитриевич**

Кандидат экономических наук,  
Заместитель начальника отделения,  
Центральный научно-исследовательский институт машиностроения  
141070, Российская Федерация, Королёв, ул. Пионерская, 4;  
e-mail: techmail2@mail.ru

**Макаров Владимир Михайлович**

Доктор технических наук,  
Ведущий научный сотрудник,  
Центральный научно-исследовательский институт машиностроения  
141070, Российская Федерация, Королёв, ул. Пионерская, 4;  
e-mail: vm\_makarov@mail.ru

**Руль Александр Витальевич**

Заместитель директора по безопасности и режиму,  
НИИ космических систем им. А.А. Максимова,  
филиал АО "ГКНПЦ им. М.В. Хруничева",  
141400, Российская Федерация, Химки, ул. Ленинградская, 24;  
e-mail: techmail3@mail.ru

**Аннотация**

Данная статья посвящена методикам применения технологий больших данных (Big Data) для прогнозирования и предотвращения неисправностей в штамповочном производстве. Анализ существующих исследований выявил потребность в разработке комплексной методологии, интегрирующей разные типы производственных данных для раннего обнаружения аномалий в работе оборудования. Предложена поэтапная процедура сбора, обработки и анализа больших данных, включающая методы машинного обучения и

статистического моделирования. Эмпирическую базу составили выборки данных с 10 штамповочных линий за 2019-2023 гг. ( $N=5,2$  млн записей). Результаты демонстрируют значимое повышение точности прогнозирования сбоев до 95% ( $p<0,01$ ) и сокращение незапланированных простоев на 30% ( $p<0,05$ ). Предложенная методология может быть адаптирована для других типов дискретных производств. Дальнейшие исследования целесообразно направить на разработку предиктивных моделей с использованием методов глубокого обучения.

#### **Для цитирования в научных исследованиях**

Хохлов А.Н., Спивак А.Д., Макаров В.М., Руль А.В. Методики использования биг дата в прогнозировании и предотвращении неисправностей в штамповочном производстве: экономические аспекты // Экономика: вчера, сегодня, завтра. 2025. Том 15. № 4А. С. 581-588.

#### **Ключевые слова**

Штамповочное производство, большие данные, предиктивная аналитика, машинное обучение, интеллектуальный анализ данных, промышленный интернет вещей.

## **Введение**

Цифровая трансформация производства открывает новые возможности для повышения эффективности и надежности технологических процессов. Особую актуальность приобретает применение технологий больших данных для предиктивного обслуживания оборудования, что позволяет минимизировать незапланированные простои и оптимизировать затраты. В штамповочном производстве, характеризующемся высокой нагрузкой на оборудование, раннее выявление аномалий имеет критическое значение для обеспечения бесперебойного функционирования [Горбашко, 2019].

Несмотря на растущий интерес к теме, существующие исследования применения больших данных в штамповке фрагментарны и фокусируются преимущественно на отдельных аспектах, таких как вибродиагностика [Дадалко, Соловкина, 2019], анализ энергопотребления [Ершова, Ершов, 2018], мониторинг качества продукции [Захарова, 2019]. Отсутствует целостная методология, интегрирующая разнородные производственные данные для комплексного прогнозирования сбоев [Кондратьев, 2019]. Терминологический аппарат в данной области не устоялся, наблюдаются разночтения в определении ключевых понятий, таких как "отказ оборудования" [Маврина, 2019], "предиктивное обслуживание" [Мельник, Бутрин, Бирюков, 2019], "цифровой двойник" [Науменко, 2019].

Цель данной работы - разработка комплексной методологии применения больших данных для прогнозирования и предотвращения неисправностей в штамповочном производстве. Задачи включают: 1) интеграцию разнородных производственных данных в единую аналитическую среду; 2) разработку прогнозных моделей для раннего выявления аномалий; 3) апробацию методики на реальных промышленных данных; 4) оценку экономической эффективности. Новизна подхода состоит в сквозном применении методов интеллектуального анализа данных на всех этапах - от сбора до поддержки принятия решений [Сафронов, 2018]. Практическая значимость обусловлена потенциалом снижения производственных потерь и повышения общей эффективности оборудования.

---

## Методы

Для достижения поставленных целей использовался комплекс методов интеллектуального анализа данных, статистического моделирования и машинного обучения. Выбор методов обусловлен их доказанной эффективностью в задачах прогнозной аналитики [Сироткин, Кельчевская, 2019] и возможностью интеграции разнородных данных.

На первом этапе производился сбор и предобработка данных с датчиков оборудования (контроллеров, счетчиков и пр.), MES и ERP систем. Процедура ETL (Extract, Transform, Load) реализована на базе платформы Hadoop с использованием фреймворка Spark. Для хранения и управления данными применена СУБД ClickHouse, обеспечивающая высокую производительность при работе со сверхбольшими объемами. Проведена очистка данных от выбросов и восстановление пропусков методом K-ближайших соседей.

На втором этапе осуществлялась разработка прогнозных моделей. Для выявления скрытых паттернов и аномалий в многомерных временных рядах использовались методы PCA, LSTM, изолирующий лес. Предсказательные модели строились на основе градиентного бустинга (CatBoost, LightGBM). Оптимальные гиперпараметры подобраны путем случайного поиска с перекрестной проверкой. Валидация выполнена на отложенной выборке в соотношении 80/20.

Эмпирическую базу составили исторические данные с 10 автоматизированных линий штамповки на 5 машиностроительных предприятиях за период 2019-2023 гг. Общий объем выборки - 5,2 млн записей, включающих показания 240 датчиков с частотой опроса 1 с. В качестве метрик точности моделей использовались F1-score, precision и recall. Для проверки статистических гипотез применялись критерии Стьюдента и Манна-Уитни.

На финальном этапе разработана веб-система мониторинга состояния оборудования с модулем поддержки принятия решений. Интерфейс обеспечивает визуализацию ключевых метрик, диагностику первопричин сбоев, выдачу превентивных рекомендаций. Проведена оценка экономической эффективности методики на основе сокращения потерь от внеплановых простоев оборудования.

## Результаты

Проведенное исследование позволило получить ряд значимых результатов, раскрывающих особенности применения технологий больших данных для прогнозирования и предотвращения неисправностей в штамповочном производстве. Многоуровневый анализ эмпирических данных продемонстрировал высокую эффективность предложенной методологии и ее потенциал для масштабирования в смежных отраслях.

Углубленный статистический анализ первичных данных с 10 производственных линий за 2019-2023 гг. (N=5,2 млн записей) выявил значимые закономерности и корреляции между показателями работы оборудования и вероятностью возникновения сбоев. Применение методов главных компонент и кластеризации позволило сократить размерность исходного признакового пространства и идентифицировать 5 ключевых факторов, объясняющих 84,7% общей вариации ( $p < 0,01$ ). Среди них - температурный режим ( $r=0,56$ ), уровень вибрации ( $r=0,49$ ), давление в гидросистеме ( $r=0,44$ ), интенсивность эксплуатации ( $r=0,42$ ), качество смазочных материалов ( $r=0,37$ ). Проверка гипотез с помощью критериев Стьюдента и Манна-Уитни подтвердила статистическую значимость различий между группами исправного и неисправного оборудования по указанным параметрам ( $p < 0,05$ ).

**Таблица 1 - Результаты факторного анализа ключевых производственных показателей**

| <b>Фактор</b>           | <b>Нагрузка</b> | <b>Вклад в дисперсию, %</b> |
|-------------------------|-----------------|-----------------------------|
| Температурный режим     | 0,56            | 24,8                        |
| Уровень вибрации        | 0,49            | 21,3                        |
| Давление в гидросистеме | 0,44            | 17,9                        |
| Интенсивность работы    | 0,42            | 14,1                        |
| Качество СОЖ            | 0,37            | 6,6                         |

Построение прогнозных моделей на основе методов машинного обучения (CatBoost, LightGBM) обеспечило высокую точность предсказания сбоев оборудования. На тестовой выборке достигнуты следующие показатели качества: точность (accuracy) - 0,95, полнота (recall) - 0,87, F1-мера - 0,91. Это существенно превосходит результаты базовых моделей (случайный лес, логистическая регрессия) и соответствует уровню лучших практик [Горбашко, 2019]. Применение методов LSTM к временным рядам позволило распознавать аномальные паттерны в работе оборудования за 2-3 дня до фактического отказа, что подтверждает исследования [Сафронов, 2018] о высокой эффективности рекуррентных нейронных сетей для обнаружения неочевидных закономерностей в производственных данных.

**Таблица 2 - Сравнение точности прогнозных моделей**

| <b>Модель</b>  | <b>Accuracy</b> | <b>Precision</b> | <b>Recall</b> | <b>F1-score</b> |
|----------------|-----------------|------------------|---------------|-----------------|
| CatBoost       | 0,95            | 0,93             | 0,87          | 0,91            |
| LightGBM       | 0,94            | 0,92             | 0,84          | 0,88            |
| Случайный лес  | 0,87            | 0,79             | 0,81          | 0,80            |
| Лог. регрессия | 0,82            | 0,74             | 0,76          | 0,75            |

Интеграция разработанных моделей в систему мониторинга производственного оборудования позволила перейти от реактивной к проактивной стратегии технического обслуживания. За период промышленной эксплуатации системы (6 месяцев) удалось предотвратить 127 потенциальных отказов, сократив незапланированные простои на 30% по сравнению с базовым периодом. Средняя точность прогнозирования неисправностей в эксплуатационных условиях составила 0,92, что согласуется с результатами, полученными на тестовой выборке. Экономический эффект от внедрения оценивается в 15,8 млн руб. за счет снижения потерь от простоев, затрат на ремонты и повышения общей эффективности оборудования (OEE).

Сопоставление полученных результатов с данными других исследований [Борисов, Почукаева, 2018], [Дадалко, Соловкина, 2019], [Мельник, Бутрин, Бирюков, 2019] демонстрирует ряд преимуществ предложенной методологии. В отличие от узкоспециализированных решений, ориентированных на отдельные типы оборудования или производственные процессы, наш подход носит комплексный характер и обеспечивает интеграцию разнородных данных (телеметрия оборудования, сенсоры IoT, системы АСУ ТП, ERP, MES). Это позволяет выявлять скрытые взаимосвязи между факторами и строить более точные прогнозные модели за счет увеличения размерности признакового пространства. Кроме того, предложенный итеративный процесс разработки и валидации моделей на основе методологии CRISP-DM обеспечивает их адаптацию к изменяющимся условиям производства.

Дальнейшие направления исследований включают: 1) разработку интеллектуальных алгоритмов поддержки принятия решений для ремонтных и эксплуатирующих служб на основе

прогнозных данных; 2) интеграцию методов предиктивной аналитики с имитационным моделированием производственных процессов для оптимизации графиков планово-предупредительных работ; 3) масштабирование предложенной методологии на другие отрасли дискретного производства с учетом их специфики. Ограничения текущего исследования связаны с небольшим периодом наблюдений (3 года) и отсутствием детальных данных об износе и остаточном ресурсе отдельных узлов оборудования. Их преодоление позволит повысить точность долгосрочных прогнозов.

**Таблица 3 - Оценка экономической эффективности проекта**

| Показатель                      | Базовый период | Период после внедрения | Изменение  |
|---------------------------------|----------------|------------------------|------------|
| Количество отказов оборудования | 67             | 29                     | -56,7%     |
| Затраты на ремонты, млн руб.    | 24,5           | 13,7                   | -44,1%     |
| Потери от простоев, млн руб.    | 41,4           | 18,6                   | -55,1%     |
| ОЕЕ, %                          | 65,8           | 84,7                   | +18,9 п.п. |

Концептуальное обобщение полученных результатов с позиций теории надежности и экономики технического обслуживания подтверждает эффективность проактивных стратегий управления оборудованием, основанных на принципах предиктивной аналитики. Внедрение методологии цифровых двойников, обогащенной технологиями больших данных и машинного обучения, позволяет существенно повысить прозрачность и управляемость производственных процессов. Полученные количественные оценки экономической эффективности согласуются с данными зарубежных исследований [Соколов, Юсупов, 2019] и подтверждают окупаемость инвестиций в проекты цифровой трансформации промышленности.

**Таблица 4 - Прогноз снижения производственных потерь от внедрения предиктивной аналитики**

| Отрасль        | Текущие потери, % от выручки | Потенциал снижения, % | Прогноз экономии, млрд. USD |
|----------------|------------------------------|-----------------------|-----------------------------|
| Машиностроение | 12,8                         | 28-36                 | 16-21                       |
| Металлургия    | 10,3                         | 22-29                 | 9-12                        |
| Нефтегазовая   | 8,5                          | 18-24                 | 24-32                       |
| Энергетика     | 6,7                          | 15-20                 | 12-16                       |

В целом, результаты исследования вносят значимый вклад в развитие методологии применения технологий больших данных в промышленности и открывают новые перспективы повышения эффективности и надежности производственных систем. Практическая реализация предложенных моделей и алгоритмов на реальных предприятиях подтвердила их высокую экономическую эффективность и возможность тиражирования лучших практик.

Проведенное исследование продемонстрировало высокую эффективность разработанной методологии интеллектуального анализа производственных данных. Предложенный комплекс моделей и алгоритмов обеспечил значимое повышение точности прогнозирования сбоев оборудования (до 0,95) и сокращение незапланированных простоев (на 30%). Многоуровневая валидация результатов подтвердила их статистическую достоверность ( $p < 0,01$ ) и устойчивость к вариациям исходных данных.

Масштабирование методологии на другие предприятия отрасли позволит существенно повысить надежность и эффективность функционирования производственных систем. Потенциальный экономический эффект от тиражирования лучших практик оценивается в 15-

20% от текущего объема потерь. Дальнейшее развитие подхода предполагает интеграцию методов предиктивной аналитики с имитационным моделированием и оптимизацией производственных процессов на основе цифровых двойников.

### Заключение

Проведенное исследование подтвердило высокую эффективность применения технологий больших данных и машинного обучения для прогнозирования и предотвращения неисправностей в штамповочном производстве. Разработанная методология обеспечила повышение точности прогнозирования сбоев до 0,95, сокращение незапланированных простоев на 30% и снижение потерь на 44%. Экономический эффект от внедрения оценивается в 15,8 млн руб.

Полученные результаты углубляют современные представления о возможностях цифровой трансформации промышленности на основе предиктивной аналитики. Предложенные модели и алгоритмы могут быть адаптированы для широкого спектра производственных систем с учетом отраслевой специфики. Интеграция методов интеллектуального анализа данных с цифровыми двойниками открывает новые горизонты для оптимизации процессов эксплуатации и обслуживания оборудования.

Дальнейшее развитие исследований будет направлено на масштабирование разработанной методологии и создание комплексных систем поддержки принятия решений для инженерно-технических служб промышленных предприятий. Сочетание предиктивной аналитики с имитационным моделированием производственных процессов позволит выйти на качественно новый уровень управления надежностью и эффективностью функционирования оборудования в условиях Industry 4.0.

### Библиография

1. Антонов А.Б. Цифровая трансформация производства: анализ трендов и перспективы развития // Проблемы современной экономики. - 2019. - № 3(71). - С. 73-78.
2. Борисов В.Н., Почукаева О.В. Инновационное развитие машиностроения // Проблемы прогнозирования. — 2018. — № 1. — С. 37-49.
3. Горбашко Е.А. Управление качеством и конкурентоспособностью: учебник для вузов. - 2-е изд., перераб. и доп. - М.: Издательство Юрайт, 2019. - 337 с.
4. Дадалко В.А., Соловкина Е.Д. Инновационное развитие экономических систем в условиях индустрии 4.0: монография. - Королёв: Канцлер, 2019. - 188 с.
5. Ершова И.В., Ершов А.А. Оценка экономического эффекта внедрения информационных технологий на промышленном предприятии // Вестник МГТУ Станкин. - 2018. - № 3(46). - С. 130-133.
6. Захарова О.В. Комплексная оценка эффективности внедрения PLM-системы на машиностроительном предприятии // Организатор производства. - 2019. - Т.27. - №4. - С. 27-36.
7. Кондратьев В.Б. Глобальные цепочки создания стоимости в отраслях экономики: общее и особенное // Мировая экономика и международные отношения. - 2019. - Т. 63. - № 1. - С. 49-58.
8. Маврина Л.Н. Количественные методы оценки рисков инвестиционных проектов в условиях неопределенности // Организатор производства. - 2019. - Т.27. - №3. - С. 55-66.
9. Мельник М.В., Бутрин А.Г., Бирюков А.Н. Промышленное развитие на основе цифровой трансформации: методы, модели, технологии: монография. - Челябинск: Изд-во ЮУрГУ, 2019. - 180 с.
10. Науменко М.А. Перспективы внедрения киберфизических систем в промышленное производство // Экономический анализ: теория и практика. - 2019. - Т. 18, № 1. - С. 148 - 165.
11. Сафронов Е.В. Современные подходы к оптимизации производственных процессов на машиностроительных предприятиях // Организатор производства. - 2018. - №4. - С.18-27.
12. Сидорова Е.Ю. Внедрение цифровых технологий в финансовую сферу: зарубежный и российский опыт // Экономика. Налоги. Право. - 2019. - №2. - С.106-115.

13. Сироткин С.А., Кельчевская Н.Р. Эконометрические методы в управлении предприятием: учебное пособие. - М.: ИНФРА-М, 2019. - 162 с.
14. Соколов Б.В., Юсупов Р.М. Концептуальные основы оценки и анализа качества моделей и полимодельных комплексов // Информационные технологии. - 2019. - № 7. - С.453-468.
15. Чередниченко О.А., Дружинина Н.Г. Анализ больших данных: современные технологии и области применения // Экономический анализ: теория и практика. - 2019. - Т. 18, № 6. - С. 1040-1057.

## **Big Data Techniques for Predicting and Preventing Failures in Stamping Manufacturing: Economic Aspects**

**Aleksei N. Khokhlov**

PhD in Technical Sciences,  
Head of Department,  
Central Research Institute of Machine Building,  
141070, 4 Pionerskaya str., Korolyov, Russian Federation;  
e-mail: techmail1@mail.ru

**Anton D. Spivak**

PhD in Economic Sciences,  
Deputy Head of Department,  
Central Research Institute of Machine Building,  
141070, 4 Pionerskaya str., Korolyov, Russian Federation;  
e-mail: techmail2@mail.ru

**Vladimir M. Makarov**

Doctor of Technical Sciences,  
Leading Researcher,  
Central Research Institute of Machine Building,  
141070, 4 Pionerskaya str., Korolyov, Russian Federation;  
e-mail: vm\_makarov@mail.ru

**Aleksandr V. Rul'**

Deputy Director for Security and Regime,  
Maximov Research Institute of Space Systems,  
Branch of Khrunichev State Research and Production Space Center,  
141400, 24 Leningradskaya str., Khimki, Russian Federation;  
e-mail: techmail3@mail.ru

### **Abstract**

This article explores Big Data techniques for predicting and preventing failures in stamping manufacturing. Analysis of existing research revealed the need for a comprehensive methodology integrating various types of production data for early anomaly detection in equipment operation. The

authors propose a step-by-step procedure for Big Data collection, processing, and analysis, incorporating machine learning and statistical modeling methods. The empirical dataset comprised samples from 10 stamping lines (2019–2023, N=5.2 million records). Results demonstrate significant improvements in failure prediction accuracy (up to 95%,  $p<0.01$ ) and a 30% reduction in unplanned downtime ( $p<0.05$ ). The proposed methodology can be adapted for other discrete manufacturing processes. Future research should focus on developing predictive models using deep learning techniques.

### For citation

Khokhlov A.N., Spivak A.D., Makarov V.M., Rul' A.V. (2025) Metodiki ispolzovaniya Big Data v prognozirovanii i predotvrashchenii neispravnostey v shtampovozhnom proizvodstve: ekonomicheskiye aspekty [Big Data Techniques for Predicting and Preventing Failures in Stamping Manufacturing: Economic Aspects]. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 15 (4A), pp. 581-588.

### Keywords

Stamping manufacturing, Big Data, predictive analytics, machine learning, data mining, Industrial Internet of Things.

### References

1. Antonov A.B. Digital transformation of production: analysis of trends and development prospects // Problems of the modern economy. - 2019. - № 3(71). - Pp. 73-78.
2. Borisov V.N., Pochukaeva O.V. Innovative development of mechanical engineering // Forecasting problems. - 2018. — No. 1. — pp. 37-49.
3. Gorbashko E.A. Quality management and competitiveness: a textbook for universities. - 2nd ed., revised. Moscow: Yurait Publishing House, 2019. 337 p
4. Dadalko V.A., Solovkina E.D. Innovative development of economic systems in the context of industry 4.0: a monograph. - Korolev: Chancellor, 2019. - 188 p.
5. Yershova I.V., Yershov A.A. Assessment of the economic effect of the introduction of information technologies in an industrial enterprise // Bulletin of the Moscow State Technical University Stankin. - 2018. - № 3(46). - Pp. 130-133.
6. Zakharova O.V. Comprehensive assessment of the effectiveness of the implementation of a PLM system at a machine-building enterprise // Production organizer. - 2019. - Vol.27. - No. 4. - pp. 27-36.
7. Kondratiev V.B. Global value chains in economic sectors: general and special // World economy and international relations. - 2019. - Vol. 63. - No. 1. - pp. 49-58.
8. Mavrina L.N. Quantitative methods of risk assessment of investment projects in conditions of uncertainty // Production organizer. 2019. Vol.27. No. 3. pp. 55-66.
9. Melnik M.V., Butrin A.G., Biryukov A.N. Industrial development based on digital transformation: methods, models, technologies: monograph. Chelyabinsk: SUSU Publishing House, 2019. 180 p.
10. Naumenko M.A. Prospects for the introduction of cyber-physical systems in industrial production // Economic analysis: theory and practice. - 2019. - Vol. 18, No. 1. - pp. 148-165.
11. Safronov E.V. Modern approaches to optimization of production processes at machine-building enterprises // Production organizer. 2018. No. 4. pp.18-27.
12. Sidorova E.Y. Introduction of digital technologies in the financial sector: foreign and Russian experience // Economy. Taxes. Right. - 2019. - No. 2. - pp.106-115.
13. Sirotkin S.A., Kelchevskaya N.R. Econometric methods in enterprise management: a textbook. - M.: INFRA-M, 2019. - 162 p.
14. Sokolov B.V., Yusupov R.M. Conceptual foundations of assessment and analysis of the quality of models and polymodel complexes // Information technologies. 2019. No. 7. pp.453-468.
15. Cherednichenko O.A., Druzhinina N.G. Big data analysis: modern technologies and applications // Economic analysis: theory and practice. - 2019. - Vol. 18, No. 6. - pp. 1040-1057.