



Российский  
научный  
фонд



II всероссийская школа молодых учёных  
Системный анализ динамики природных процессов в российской Арктике

**Методы машинного обучения в задачах диагностирования  
и прогнозирования геофизических параметров  
(на примере высокоширотных геомагнитных данных)**

Воробьев Андрей Владимирович  
[e. mail: geomagnet@list.ru](mailto:geomagnet@list.ru)

ФГБОУ ВО Уфимский государственный авиационный технический университет  
ФГБУН Геофизический центр РАН

6–9 июня 2022 г.  
Московская область

Интеллект - способность системы создавать в ходе самообучения программы (в первую очередь эвристические) для решения задач определённого класса сложности и решать эти задачи»

Машинное обучение — класс методов искусственного интеллекта, характерной чертой которых является не прямое решение задачи, а обучение за счёт применения решений множества сходных задач. Для построения таких методов используются средства математической статистики, численных методов, математического анализа, методов оптимизации, теории вероятностей, теории графов, различные техники работы с данными в цифровой форме.

Различают два типа обучения:

Обучение по прецедентам, или индуктивное обучение, основано на выявлении эмпирических закономерностей в данных.

Дедуктивное обучение предполагает формализацию знаний экспертов и их перенос в компьютер в виде базы знаний.

Дедуктивное обучение принято относить к области экспертных систем, поэтому термины машинное обучение и обучение по прецедентам можно считать синонимами.



# ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И МАШИННОЕ ОБУЧЕНИЕ: НЕМНОГО ИСТОРИИ

- 1943 г. У. Маккалок и У. Питтс формализуют понятие нейронной сети в фундаментальной статье о логическом исчислении идей и нервной активности. В начале своего сотрудничества с Питтсом Н. Винер предлагает ему вакуумные лампы в качестве средства для реализации эквивалентов нейронных сетей
- 1949 г. Д. Хебб предлагает первый алгоритм обучения.
- 1958 г. Ф. Розенблатт изобретает однослойный перцептрон и демонстрирует его способность решать задачи классификации[6]. Перцептрон использовали для распознавания образов, прогнозирования погоды
- 1982 г. Дж. Хопфилд показал, что нейронная сеть с обратными связями может представлять собой систему, минимизирующую энергию (сеть Хопфилда). Кохоненом представлены модели сети, обучающейся без учителя (нейронная сеть Кохонена).
- 1986 г. Дэвидом И. Румельхартом, Дж. Е. Хинтоном и Рональдом Дж. Вильямсом, а так же независимо и одновременно С. И. Барцевым и В. А. Охониным переоткрыт и развит метод обратного распространения ошибки.
- 1997 г. Г. Каспаров проигрывает искусственному интеллекту Deep Blue в шахматы
- 2007 г. Джеффри Хинтоном в университете Торонто созданы алгоритмы глубокого обучения многослойных нейронных сетей.
- В декабре 2017 вышел AlphaZero, еще более мощный вариант AlphaGo Zero. Он смог стать лучше предшественника за 8 часов, одновременно достигнув уровня гроссмейстера в шахматах. Так го стала второй игрой, в которой люди больше не могут выиграть.
- В марте 2018 года небольшой робот собрал кубик Рубика за 0,38 секунды. Рекорд среди людей — 4,69 секунды.
- В мае 2018 искусственный интеллект стал лучше людей распознавать рак кожи.



Матч Deep Blue — Каспаров (Нью-Йорк, май 1997)



# МАШИННОЕ ОБУЧЕНИЕ: ОБЛАСТИ ПРИМЕНЕНИЯ

## Распознавание образов и классификация

В качестве образов могут выступать различные по своей природе объекты: символы текста, изображения, образцы звуков и т. д. При обучении сети предлагаются различные образцы образов с указанием того, к какому классу они относятся. Образец, как правило, представляется как вектор значений признаков. При этом совокупность всех признаков должна однозначно определять класс, к которому относится образец. В случае, если признаков недостаточно, сеть может соотнести один и тот же образец с несколькими классами, что неверно. По окончании обучения сети ей можно предъявлять неизвестные ранее образы и получать ответ о принадлежности к определённому классу.

## Поддержка принятия решений и управление

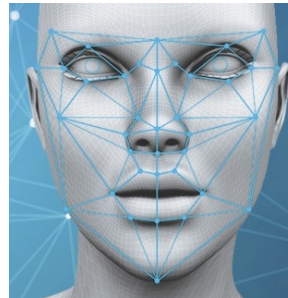
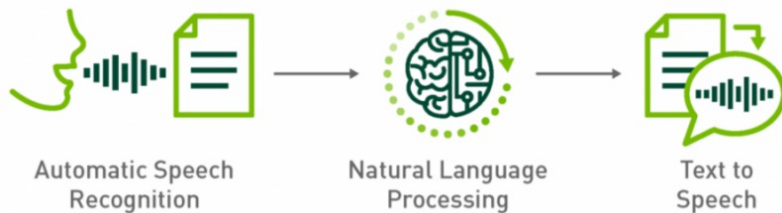
Эта задача близка к задаче классификации. Классификации подлежат ситуации, характеристики которых поступают на вход нейронной сети. На выходе сети при этом должен появиться признак решения, которое она приняла. При этом в качестве входных сигналов используются различные критерии описания состояния управляемой системы

## Кластеризация

Под кластеризацией понимается разбиение множества входных сигналов на классы, при том, что ни количество, ни признаки классов заранее не известны. После обучения такая сеть способна определять, к какому классу относится входной сигнал. Сеть также может сигнализировать о том, что входной сигнал не относится ни к одному из выделенных классов — это является признаком новых, отсутствующих в обучающей выборке, данных. Таким образом, подобная сеть может выявлять новые, неизвестные ранее классы сигналов. Соответствие между классами, выделенными сетью, и классами, существующими в предметной области, устанавливается человеком.

## Прогнозирование

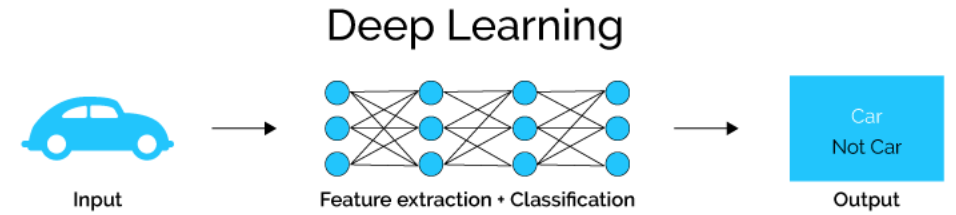
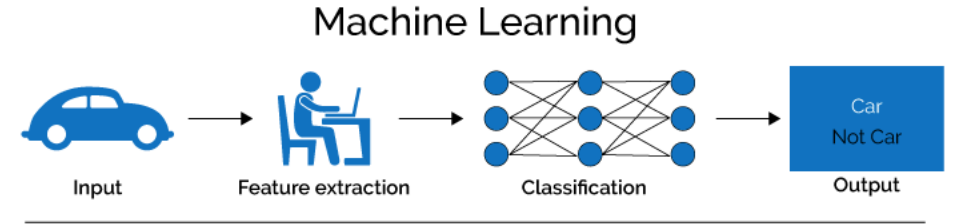
Способности нейронной сети к прогнозированию напрямую следуют из её способности к обобщению и выделению скрытых зависимостей между входными и выходными данными. После обучения сеть способна предсказать будущее значение некой последовательности на основе нескольких предыдущих значений и (или) каких-то существующих в настоящий момент факторов. Прогнозирование возможно только тогда, когда предыдущие изменения действительно в какой-то степени определяют будущее.





# СОВРЕМЕННОЕ ПОНИМАНИЕ МАШИННОГО ОБУЧЕНИЯ

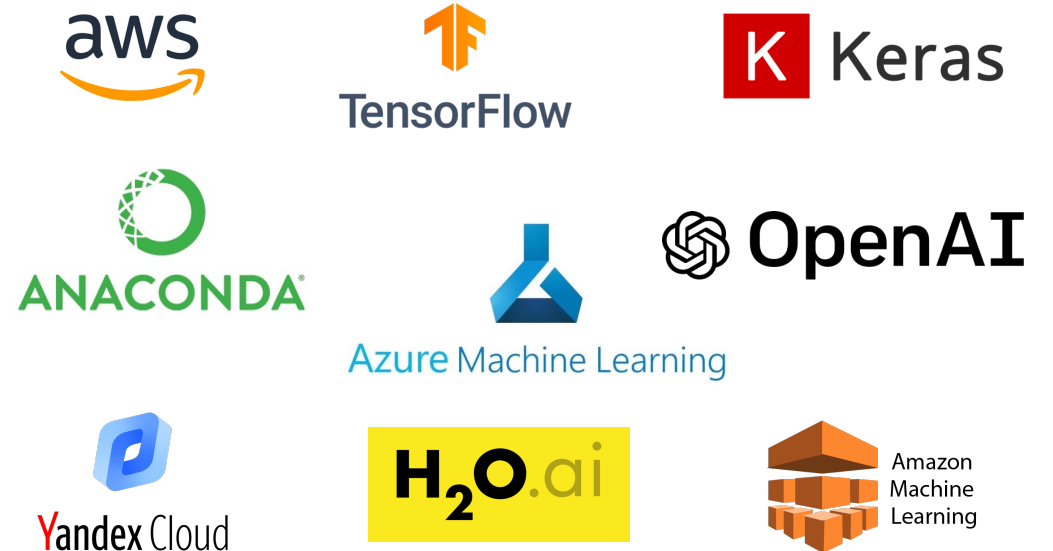
Машинное обучение заключается в извлечении знаний из данных. Это научная область, находящаяся на пересечении статистики, искусственного интеллекта и компьютерных наук и также известная как прогнозная аналитика или статистическое обучение.



## Классическое Обучение



## ОСНОВНЫЕ ТЕХНОЛОГИИ



# ГЕОМАГНИТНОЕ ПОЛЕ И ЕГО ВАРИАЦИИ, КАК ФАКТОР НЕОПРЕДЕЛЕННОСТИ

Геомагнитное поле (ГМП), нТл:

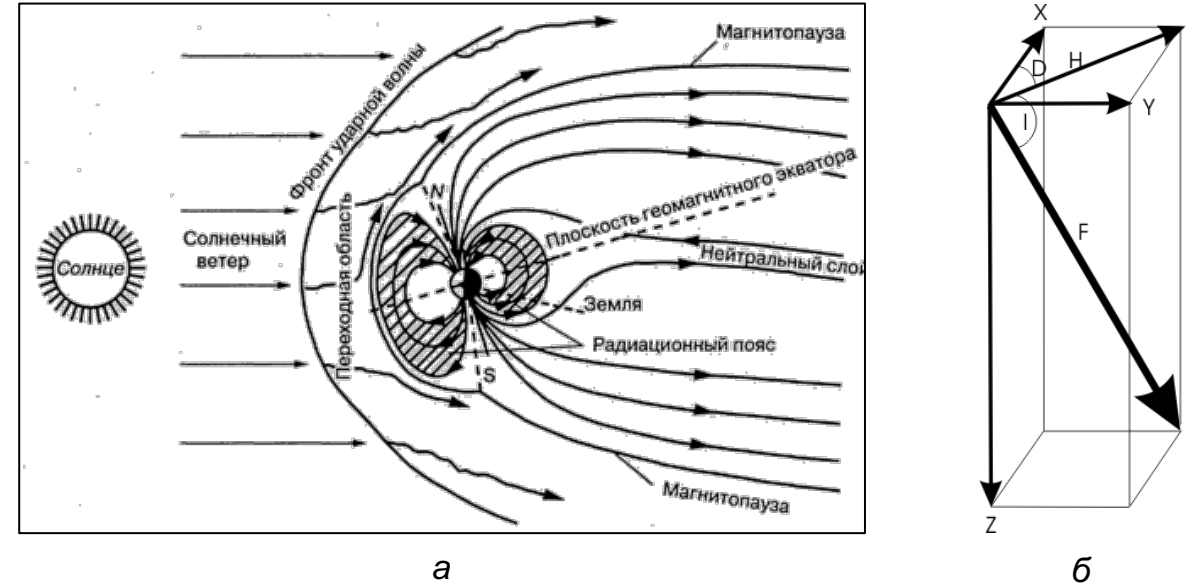
$$B_{\text{ГМП}} = B_1 + \delta B = (B_0 + B_a) + \delta B, \quad (1)$$

где  $B_1$  – ГМП внутриземных источников;  $B_0$  – главное поле (ГОСТ 25645.126-85);  $B_a$  – поле земной коры, являющееся полем магнетизма горных пород;  $\delta B$  – поле геомагнитных вариаций.

$$\delta B = S + L + DP + DR + DCF + DT, \quad (2)$$

где  $S$  – регулярная часть поля, возникающая из-за волнового излучения Солнца (ее статистическая аналогия –  $S_q$ -вариация: токовая система, развивающаяся на освещенной стороне Земли);  $L$  – регулярная часть поля, возникающая из-за лунных приливов в верхней атмосфере, статистически выделяемая методом наложения этих эпох по возможно большему периоду наблюдения;  $DP$  – нерегулярная часть поля, возникающая из-за корпускулярного излучения Солнца (ее статистическая аналогия –  $S_q$ -вариация: токи развиваются в виде электроструй в зоне полярных сияний);  $DR$  – поле кольцевого тока, существенно усиливающегося в период магнитных бурь планетарного масштаба, его аналогия – главная фаза в  $D_{st}$ -вариации: токи развиваются во внешней части радиационных поясов Земли;  $DCF$  – нерегулярная часть поля, возникающая из-за токов на поверхности магнитосферы при обтекании ее солнечным ветром (ее статистическая аналогия – начальная фаза  $D_{st}$ -вариации);  $DT$  – поле токов хвоста магнитосферы, определяемое движением плазмы, статистической аналогии не имеет и является наименее изученной составляющей вариаций ГМП [1].

Вклад каждого из перечисленных в выражении (1) слагаемых существенно зависит от времени суток, сезона, географической и геомагнитной широт наблюдения, а также от состояния околоземного космического пространства. Однако можно сказать, что составляющие  $S$ ,  $DP$  и  $DR$  являются основными.



**Рис. 1.** Схематическое изображение структуры магнитосферы Земли (а) и составляющие вектора ГМП, регистрируемые на поверхности Земли (б)

Регистрируемые на поверхности Земли параметры ГМП:

$$F^2 = X^2 + Y^2 + Z^2; H^2 = X^2 + Y^2; D = \arctg\left(\frac{Y}{X}\right); I = \arcsin\left(\frac{Z}{F}\right) = \arctg\left(\frac{Z}{H}\right), \quad (3)$$

где  $X$ ,  $Y$  и  $Z$  – северная, восточная и вертикальная составляющие вектора ГМП, нТл;  $H$  – горизонтальная составляющая, нТл;  $F$  – полный вектор, нТл;  $D$  – магнитное склонение, определяемое как угол между истинным севером (географическим севером) и магнитным севером (горизонтальной составляющей поля), град.  $I$  – магнитное наклонение, определяемое как угол, измеренный от горизонтальной плоскости к вектору магнитного поля, град.

# НИЗКАЯ СИТУАЦИОННАЯ ОСВЕДОМЛЕННОСТЬ И НЕОПРЕДЕЛЕННОСТЬ, КАК ПРОБЛЕМА ОБРАБОТКИ ВЫСОКОШИРОТНОЙ ГЕОФИЗИЧЕСКОЙ ИНФОРМАЦИИ ПРИ ПОДДЕРЖКЕ ПРИНЯТИЯ РЕШЕНИЙ

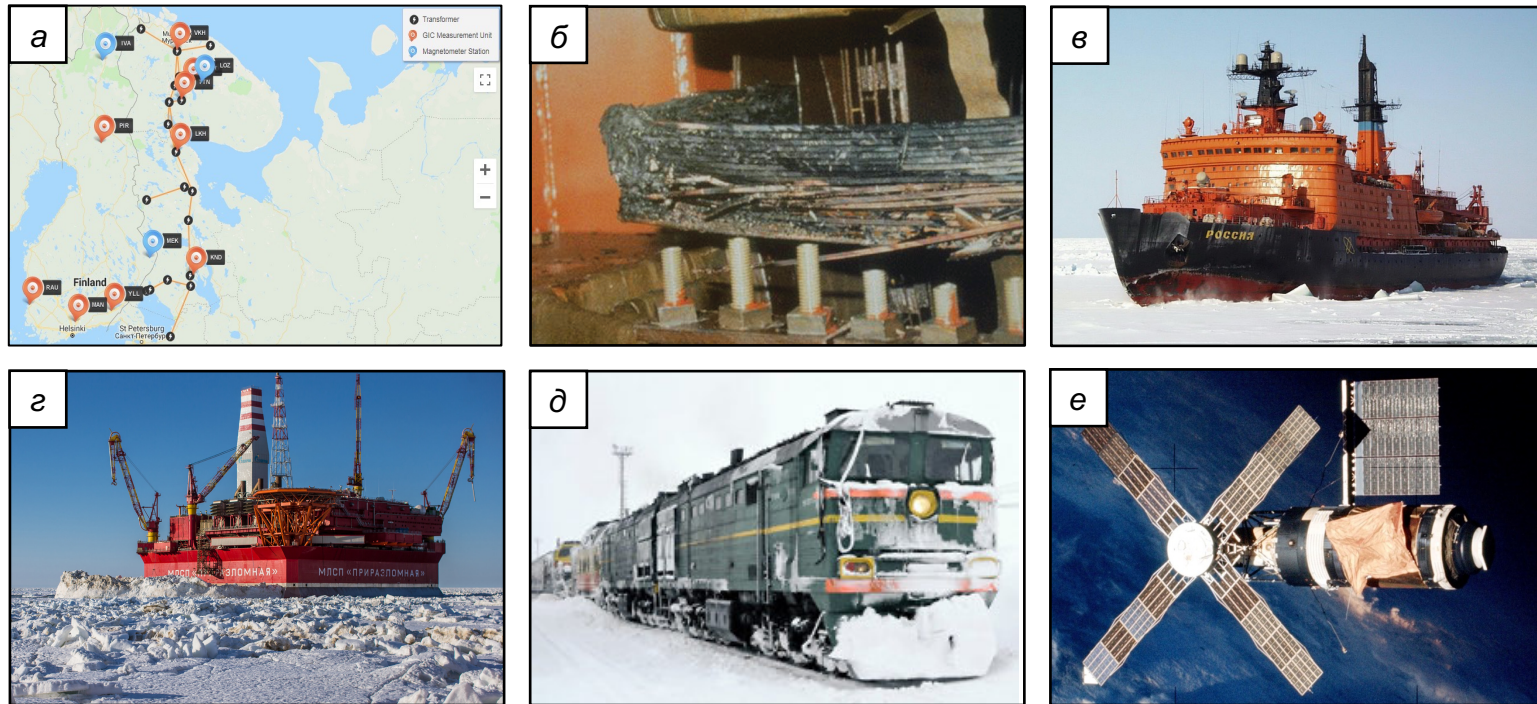
Основные типы неопределенности при поддержке принятия решений:

- неопределенности, связанные с человеческим фактором;
- неопределенности, связанные с социальным и административным фактором;
- неопределенности, связанные с экономическим фактором;
- неопределенности, связанные с внешне- и внутривнутриполитическим фактором;
- неопределенности, связанные с отсутствием / недостатком информации (информационная неопределенность);
- неопределенности, связанные с недостаточными знаниями о природе, природных явлениях, механизмах их возникновения и воздействия на биологические и технические системы.



# ИДЕНТИФИКАЦИЯ ПРОБЛЕМЫ И АКТУАЛЬНОСТЬ ИССЛЕДОВАНИЙ

- Государственная программа «Социально-экономическое развитие Арктической зоны Российской Федерации» Утверждена постановлением Правительства от 21 апреля 2014 года №366
- Указ Президента РФ от 26 октября 2020 г. № 645 «О Стратегии развития Арктической зоны Российской Федерации и обеспечения национальной безопасности на период до 2035 года»



**Рис. 2.** Объекты техносферы – потенциальные цели ГМА: а) высокоширотные ЛЭП: на рисунке показан проект «EURISGIC» (European Risk from Geomagnetically Induced Currents) для исследования ГИТ (<http://eurisgic.org/map>); б) разрушение силового трансформатора компании «Electric and Gas Company» (США) вследствие ГИТ, вызванных ГМА 13-14 марта 1989 г; в) системы навигации и КВ-радиосвязи в АЗРФ г) морская ледостойкая стационарная платформа (МЛСП) «Приразломная»; д) подвижные составы и системы сигнальной автоматики высокоширотных железных дорог; е) искусственные спутники Земли (на рисунке показана национальная орбитальная станция США SKYLAB, потерпевшая крушение в результате возросшей солнечной активности в июле 1979 г.)

- Магнитная буря 13 марта 1989 г. явилась причиной выхода из строя силовых трансформаторов и каскадного отключения (блэкаута) ЛЭП более, чем на 9 ч в провинции Квебек (Канада). В объединенной энергосистеме северо-запада России в ноябре 2001 г. вследствие ГМА дважды имело место одностороннее отключение воздушной ЛЭП (330 кВ) «Оленегорск-Мончегорск» со стороны подстанции «Оленегорск». В октябре 2003 г. аналогичная причина привела к перерыву электроснабжения в энергосистеме в Мальме на юге Швеции. [1] (Согласно отчету «Zurich Insurance Group» страховые выплаты в США в результате сбоев электрооборудования в следствии магнитных бурь с 2005 по 2015 гг превысили \$ 1.9 млрд. [2])

- Ухудшение условий КВ-радиосвязи на трассах Кольского п-ва.

- Синхронные аномалий в работах сигнальной автоматики северных веток Октябрьской и Северной железных дорог РФ.

- Демпфирование угловой скорости ИСЗ и спутниковые аномалии, а также дополнительные погрешности в системах спутниковой и подземной навигации [3]. (Так, с 3-го по 7-е февраля 2022 г магнитная буря вывела из строя 40 из 49 спутников Starlink, запущенные 3 февраля компанией SpaceX.) и пр.

1. Данилов Г.А. Повышение качества функционирования линий электропередачи: монография. Москва-Берлин, 2015

2. Dobbins R.W., Schriever K. Electrical Claims and Space Weather Measuring the visible effects of an invisible force June 2015

3. Гвишиани А.Д., Лукьянова Р.Ю., Оценка влияния геомагнитных возмущений на траекторию наклонно-направленного бурения глубоких скважин в арктическом регионе, Физика Земли, 2018, № 4, с. 19–30, doi: 10.1134/S0002333718040051.



# ИСТОЧНИКИ ГЕОМАГНИТНЫХ ДАННЫХ



а



б



в



г

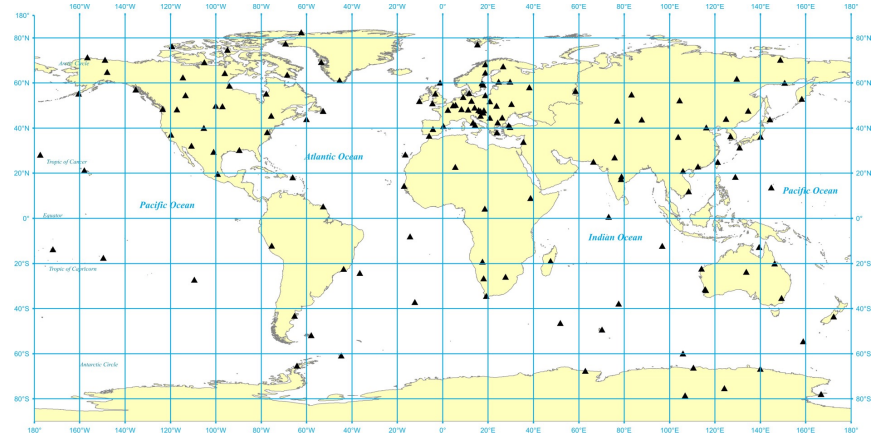


д

**Рис. 3.** а – магнитная станция Хорнсунн (архипелага Шпицберген); б – адаптированный для магнитных измерений самолет Ан-12; в – морской магнитометр Geometrics G-882; г – протонный магнитометр Geometrics G-857; д – исследовательский спутник «Браво» миссии SWARM



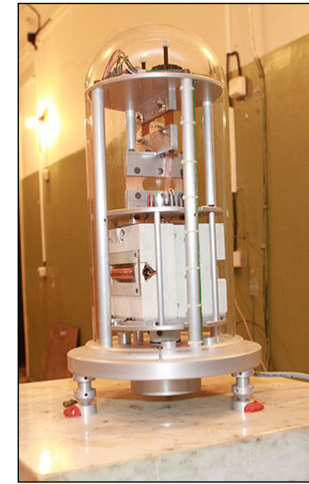
# СЕТИ МАГНИТНЫХ СТАНЦИЙ



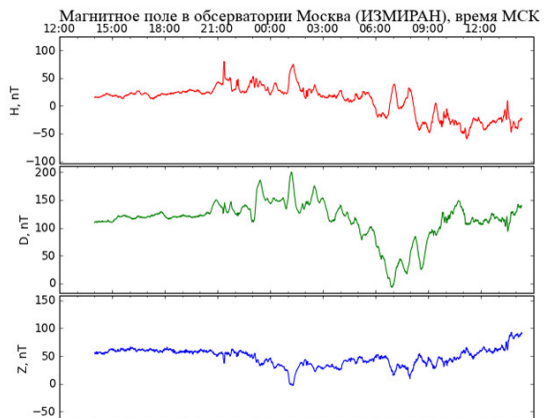
а



в



а



б



б

**Рис. 4.** Характер пространственного распределения магнитных обсерваторий сети INTERMAGNET (<https://www.intermagnet.org/>), включающей более 130 геодезических пунктов – источников данных (а), агрегатора данных SuperMag, нацеливающего более 500 магнитных станций (б) и сети высокоширотных магнитометров IMAGE, включающей 36 магнитных станций (в)

**Рис. 5.** а – магнитный теодолит; б – временные диаграммы регистрация ГМВ в обсерватории «Москва» (ИЗМИРАН)

DATE	TIME	DOY	AAAX	AAAY	AAAZ	AAAL	AAAM
2015-07-29	00:00:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:01:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:02:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:03:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:04:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:05:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:06:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:07:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:08:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:09:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:10:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:11:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:12:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:13:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:14:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:15:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:16:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:17:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:18:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:19:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:20:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:21:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:22:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:23:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:24:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:25:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:26:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:27:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:28:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-29	00:29:00.000	210	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:05:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:06:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:07:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:08:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:09:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:10:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:11:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:12:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:13:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:14:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00
2015-07-28	00:15:00.000	209	99999.00	99999.00	99999.00	99999.00	99999.00

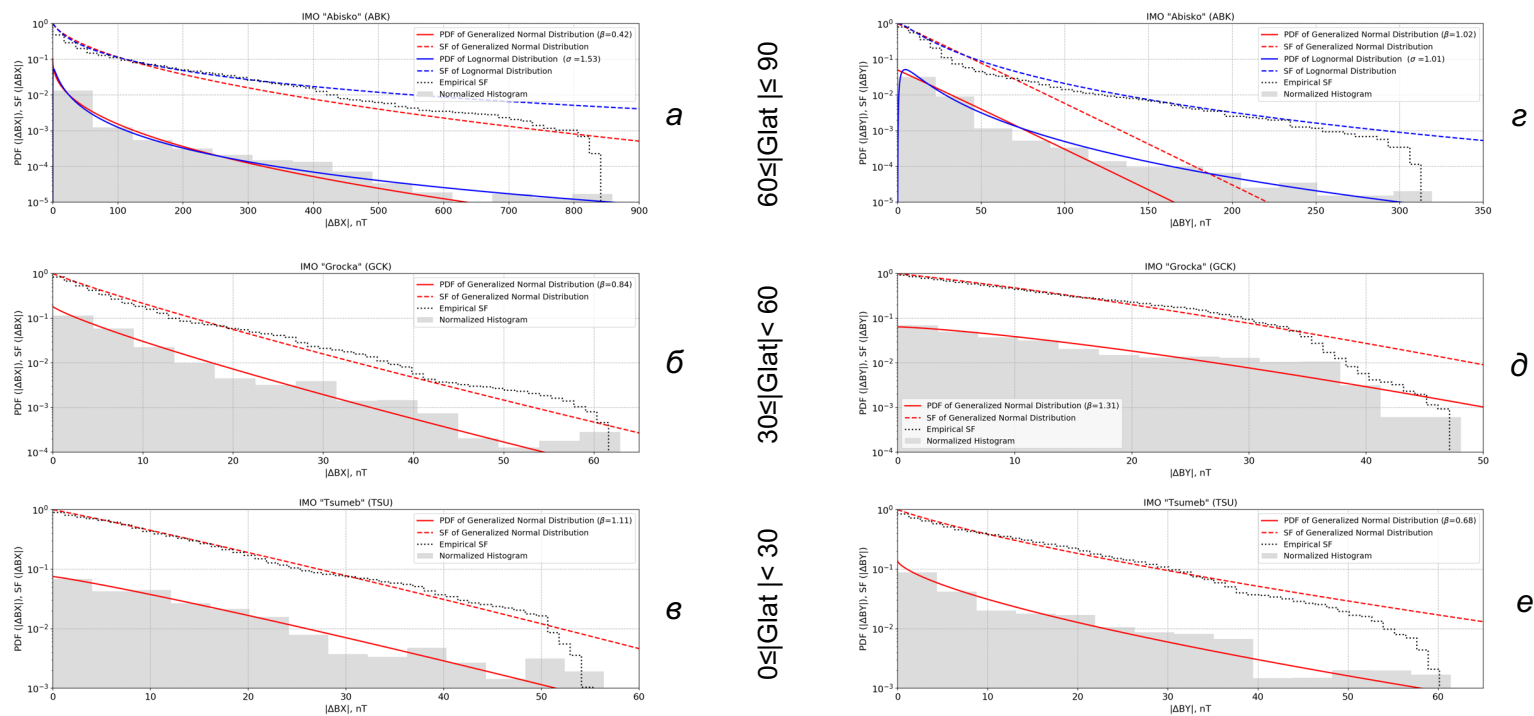


# ГЛОБАЛЬНАЯ КЛАСТЕРИЗАЦИЯ ИСТОЧНИКОВ ГЕОМАГНИТНЫХ ДАННЫХ ПО ПРИЗНАКУ СТАТИСТИЧЕСКОЙ ОДНОРОДНОСТИ РЕГИСТРИРУЕМОЙ ИНФОРМАЦИИ

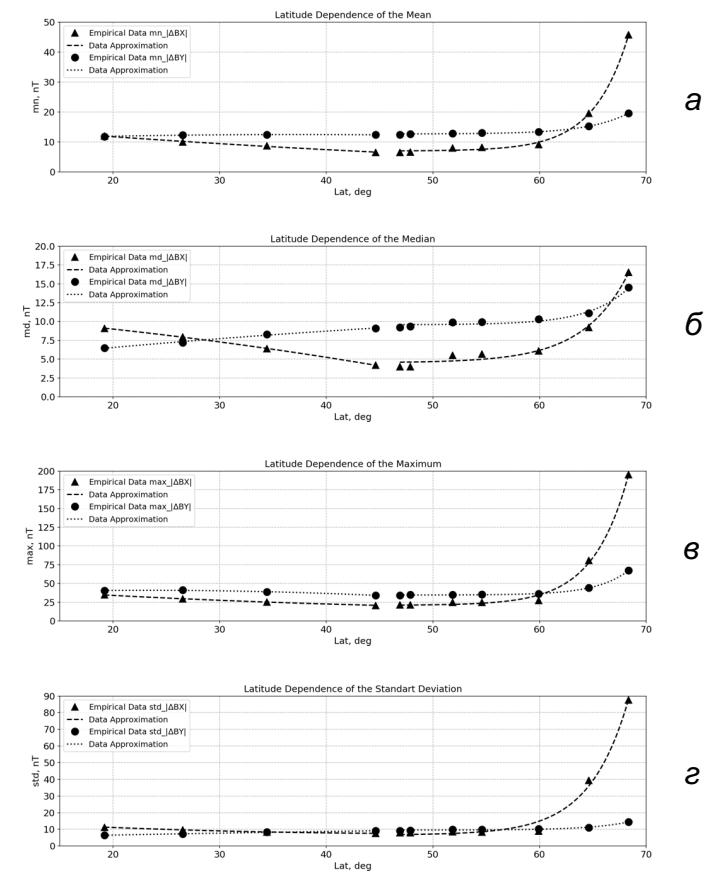
Семейство распределений, задаваемых выражением (4) демонстрирует наличие хвостов, которые тяжелее нормальных хвостов при  $\beta < 2$  и легче нормальных хвостов при  $\beta > 2$ ; при  $\beta = 2$  хвосты распределения соответствуют нормальному закону (5) Рис. 6 (в, д). При  $\beta = 1$  хвосты распределения соответствуют хвостам лапласовского распределения. При  $\beta < 1$  хвосты распределения тяжелее лапласовских хвостов и соответствуют логнормальному закону (6), Рис. 6(а, з).

$$f(x, \beta) = \frac{\beta}{\Gamma(1/\beta)} \exp(-|x|^\beta); \quad (4) \quad f(x) = \sqrt{\frac{2}{\pi}} \exp\left(-\frac{x^2}{2}\right); \quad (5) \quad f(x, s) = \frac{1}{sx\sqrt{2\pi}} \exp\left(-\frac{\log^2(x)}{2s^2}\right), \quad (6)$$

где  $\Gamma$  – гамма-функция;  $\beta$  и  $s$  – коэффициенты формы



**Рис. 6.** Характер изменения статистики северной (а, б, в) и восточной (з, д, е) составляющей вариаций ГМП для высокоширотных (а, з), среднеширотных (б, д) и приэкваториальных магнитных станций (в, е)



**Рис. 7.** Характер широтного распределения среднего, медианного и максимального значений ЦД/ФП (а), (б) и (в) соответственно и стандартного отклонения (з)

Время неработоспособного состояния системы ( $T_F$ ):

$$T_F = T - T_W, \quad (7)$$

где  $T$  – наработка;  $T_W$  – число информативных значений (общее время работоспособного состояния) за период времени  $T$ .

Среднее время до восстановления рабочего состояния  $\langle T2R \rangle$  (эквивалент математического ожидания размера отсутствующего фрагмента) и среднее время наработки между отказами системы  $\langle T2F \rangle$  (эквивалент среднего фрагмента данных без пропусков):

$$\langle T2R \rangle = \frac{1}{N_F} \sum_{i=1}^{N_F} T2R_i = \frac{T_F}{N_F}; \quad (8) \quad \langle T2F \rangle = \frac{1}{N_W + k} \sum_{i=1}^{N_W + k} T2F_i = \frac{T_W}{N_W + k}, \quad (9)$$

где  $T2R_i$  – время до  $i$ -го восстановления системы после отказа,  $T2F_i$  – время до  $i$ -го отказа системы;  $N_F$  и  $N_W$  – число отказов системы и число восстановлений после отказа соответственно;  $k = 1$  или  $0$ , если в момент начала наблюдения система находилась работоспособном или неработоспособном состоянии соответственно.

Время безотказной работы системы:

$$T_W^{\sim} = \left| \bigcap_N K \right|, \text{ где } K = \{k_j\}_{j=1}^{T-T_F}: S_{k_j} \neq \emptyset, k_j \in [1, T]; S = \{s\}_{i=1}^T. \quad (10)$$

Для магнитных обсерваторий сети INTERMAGNET [1] (более 120 станций), поддерживающей наивысший стандарт качества, размеры пропущенных фрагментов занимают достаточно широкий диапазон и варьируются как во времени, так и от станции к станции. Так, например, для станции Alma Ata (AAA) в 2015 г. число пропущенных значений составило 36 % от годовой наработки, для станции Dalat (DLT) – более 12 %, для станции Sodankyla (SOD) – 0.4 %, и т. д.

Табл. 1. Результаты оценки показателей надежности МВС (на примере магнитометров сети IMAGE, 2015 г)

№ п/п	IAGA код	Геогр. координаты		T <sub>w</sub>		T <sub>F</sub>		F	<T2R> [мин]	<T2F> [мин]
		Ф N, [град]	Λ E, [град]	[мин]	[%]	[мин]	[%]			
1	NAL	78.92	11.95	509551	96.947	16049	3.053	20	802.45	25477.55
2	LYR	78.20	15.82	506314	96.331	19286	3.669	11	1753.27	46028.55
3	HOR	77.00	15.60	466554	88.766	59046	11.234	4	14761.5	116638.5
4	HOP	76.51	25.01	492524	93.707	33076	6.293	49	675.02	10051.51
5	BJN	74.50	19.20	525523	99.985	77	0.015	7	11	75074.71
6	NOR	71.09	25.79	519087	98.761	6513	1.239	144	45.23	3604.77
7	SOR	70.54	22.22	523740	99.646	1860	0.354	43	43.26	12180.0
8	KEV	69.76	27.01	525569	99.994	31	0.006	11	2.82	47779.0
9	TRO	69.66	18.94	524713	99.831	887	0.169	15	59.13	34980.87
10	MAS	69.46	23.70	524144	99.723	1456	0.277	73	19.95	7180.05
11	AND	69.30	16.03	525284	99.94	316	0.06	6	52.67	87547.33
12	KIL	69.06	20.77	523732	99.645	1868	0.355	33	56.61	15870.67
13	IVA	68.56	27.29	486940	92.645	38660	7.355	6	6443.33	81156.67
...	...	...	...	...	...	...	...	...	...	...
31	KAR	59.21	5.24	524637	99.817	963	0.183	41	23.49	12796.02
32	TAR	58.26	26.46	525137	99.912	463	0.088	12	38.58	43761.42
33	BRZ	56.17	24.86	523584	99.616	2016	0.384	3	672.0	174528.0
34	SUW	54.01	23.18	487904	92.828	37696	7.172	20	1884.8	24395.2
35	WNG	53.74	9.07	525577	99.996	23	0.004	19	1.21	27661.95
36	NGK	52.07	12.68	525600	100	0	0	0	–	–

В 50 % случаях магнитных станций сети IMAGE [2] (более 40 станций) математическое ожидание размера пропущенного фрагмента превышает 58.5 мин. Усредненный (по всем источникам данных) размер пропущенного фрагмента составляет 1066 мин. Математическое ожидание числа отказов с восстановлением по всем станциям превышает 45/год. При этом 50 % станций испытывают за год более 17 отказов. В крайних случаях общий объем отсутствующих фрагментов одной станции может превышать 11.2% (более 41 суток) от общего размера годовой выборки, при этом среднее время восстановления может превышать 10 суток.

**Т.о., хотя сети IMAGE и INETERMAGNET являются одними из наиболее надежных источников геомагнитных данных, пропущенные и аномальные значения для них также представляют существенную проблему, которые из-за специфики регистрируемой информации, в большинстве случаев не могут быть достоверно восстановлены известными методами**

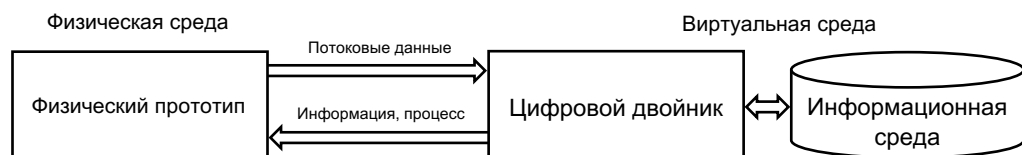
1. Love J. An International Network of Magnetic Observatories // EOS, transactions, American Geophysical Union. V. 94. No 42. P. 373–384. 2013.

2. Tanskanen, E.I. (2009): A comprehensive high-throughput analysis of substorms observed by IMAGE magnetometer network: Years 1993-2003 examined. J. Geophys. Res., 114, A05204, doi:10.1029/2008JA013682



# ПОДХОД К РЕШЕНИЮ ПРОБЛЕМЫ ОТСУТСТВУЮЩИХ ЗНАЧЕНИЙ НА БАЗЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ И КОНЦЕПЦИИ ЦИФРОВЫХ ДВОЙНИКОВ

*Цифровой двойник (ЦД)* – динамическое виртуальное представление физического объекта (процесса, системы или продукта) в течение всего жизненного цикла с использованием данных в режиме реального времени для понимания, изучения и рассуждения [1-2].



**Рис. 8.** Концепция цифрового двойника согласно ISO/DIS 23247 [3-5]

## Классификация цифровых двойников:

- *Цифровые двойники-прототипы* (Digital Twin Prototype, DTP) содержат информацию, необходимую для описания и создания физических версий экземпляров объекта.
- *Цифровые двойники-экземпляры* (Digital Twin Instance, DTI) описывают конкретный физический экземпляр объекта, с которым двойник остается связанным на протяжении всего срока эксплуатации.
- *Агрегированные двойники* (Digital Twin Aggregate, DTA) определяются как информационная система управления физическими экземплярами семейства объектов или изделий, которая имеет доступ ко всем их цифровым двойникам.



**Рис. 9.** Концептуальная модель цифрового двойника-экземпляра магнитной станции (МС – магнитная станция; ЦОС – цифровая обработка сигнала; ЦФ – цифровая фильтрация; МО – машинное обучение; ИНС – искусственная нейронная сеть;  $\Delta t_1$  – время между реальным установлением параметра и фактической передачей его значения в виде потоковых данных;  $\Delta t_2$  – время, затраченное ЦД на синтез соответствующего параметра и принятие решения о его регистрации)

1. Bolton R. N. et al. Customer experience challenges: bringing together digital, physical and social realms //Journal of Service Management. – 2018. – Т. 29. – №. 5. – С. 776-808.
2. Söderberg R. et al. Toward a Digital Twin for real-time geometry assurance in individualized production //CIRP Annals. – 2017. – Т. 66. – №. 1. – С. 137-140.
3. ISO/DIS 23247-1 Automation systems and integration — Digital Twin framework for manufacturing — Part 1: Overview and general principles
4. ISO/DIS 23247-2 Automation systems and integration — Digital Twin framework for manufacturing — Part 2: Reference architecture
5. ISO/DIS 23247-3 Automation systems and integration — Digital Twin framework for manufacturing — Part 3: Digital representation of manufacturing elements

# СИНТЕЗ МАТЕМАТИЧЕСКОЙ МОДЕЛИ ДЛЯ DTI-ДВОЙНИКА МАГНИТНОЙ ВАРИАЦИОННОЙ СТАНЦИИ НА ОСНОВЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

IDW-метод (Inverse Distance Weighting) пространственной интерполяции в модификации Шеппарда (45) и обобщенная регрессионная модель (46):

$$X_{KIL}^* = \sum_{i=1}^m \frac{1}{d_i^p} X_i / \sum_{i=1}^m \frac{1}{d_i^p}, \quad (11)$$

где  $m$  – число станций аврорального кластера,  $X_i$  – значение  $X$ -составляющей  $i$ -й станции,  $d$  – расстояние между станцией KIL и  $i$ -й станцией аврорального кластера;  $p$  – весовой коэффициент.

$$f(x, \beta) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k = \sum_{j=1}^k \beta_j x_j = x^T \beta, \quad (12)$$

где  $x^T = (x_1, x_2, \dots, x_k)$  – вектор регрессоров;  $\hat{\beta} = (\beta_1, \beta_2, \dots, \beta_k)^T$  – вектор столбец коэффициентов;  $k$  – число признаков модели.

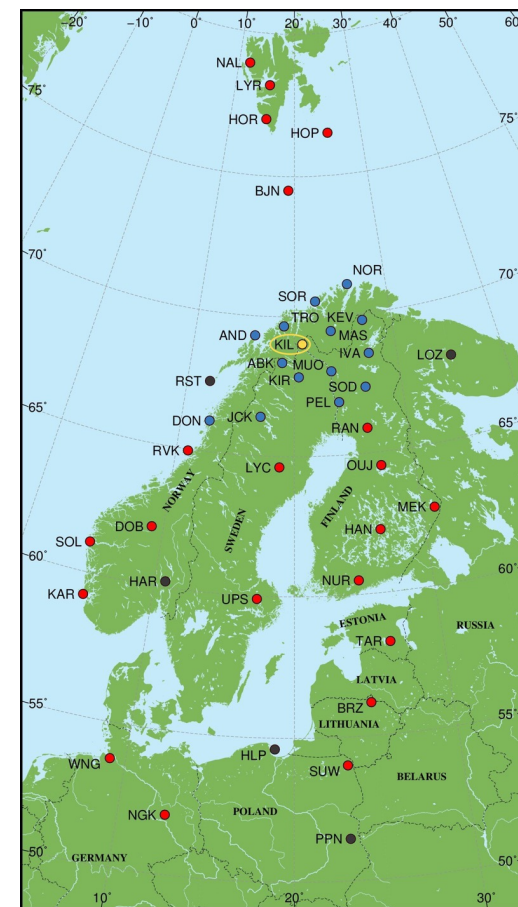
**Табл. 2.** Оценка коэффициента корреляции Пирсона между  $X$ -составляющей ГМП, регистрируемой станцией KIL, и аналогичным параметром, регистрируемым прочими магнитометрами сети IMAGE

Магнитные станции, входящие в авроральный кластер													
NOR	SOR	KEV	TRO	MAS	AND	IVA	ABK	MUO	KIR	SOD	PEL	JCK	DON
0.872	0.933	0.978	0.985	0.99	0.987	0.975	0.986	0.957	0.958	0.909	0.875	0.845	0.820
Магнитные станции, не входящие в авроральный кластер													
NAL	LYR	HOR	HOP	BJN	RAN	RVK	LYC	OUJ	MEK	HAN	DOB	SOL	NUR
-0.164	-0.129	0.015	0.015	0.427	0.053	0.694	0.642	0.617	0.432	0.384	0.363	0.262	0.274
UPS	KAR	TAR	BRZ	SUW	WNG	NGK							
0.218	0.142	0.176	0.098	-0.045	-0.017	-0.044							

$$X_{KIL}^* = \alpha + \beta_1 X_{NOR} + \beta_2 X_{SOR} + \beta_3 X_{KEV} + \beta_4 X_{TRO} + \beta_5 X_{MAS} + \beta_6 X_{AND} + \beta_7 X_{IVA} + \beta_8 X_{ABK} + \beta_9 X_{MUO} + \beta_{10} X_{KIR} + \beta_{11} X_{SOD} + \beta_{12} X_{PEL} + \beta_{13} X_{JCK} + \beta_{14} X_{DON}, \quad (13)$$

где  $\alpha = 418$  нТл – коэффициент вертикального смещения;  $\beta_1, \beta_2, \dots, \beta_{14}$  – коэффициенты, рассчитанные методом наименьших квадратов (МНК):  $\beta_1 = -0.05119918$ ;  $\beta_2 = -0.07917934$ ;  $\beta_3 = 0.0119317$ ;  $\beta_4 = 0.58589785$ ;  $\beta_5 = -0.21993329$ ;  $\beta_6 = -0.2039247$ ;  $\beta_7 = 0.11381287$ ;  $\beta_8 = 0.68734227$ ;  $\beta_9 = 0.00202142$ ;  $\beta_{10} = -0.28453331$ ;  $\beta_{11} = 0.01707591$ ;  $\beta_{12} = 0.01524056$ ;  $\beta_{13} = 0.00379651$ ;  $\beta_{14} = -0.02637731$ .

MSE модели (8), рассчитанная с применением процедуры кросс-валидации, составила 11.5 нТл, что соответствует 0.51% от размаха значений параметра  $X_{KIL}$  за 2015 г. Коэффициент корреляции Пирсона ( $R=0.999$ ) и результаты  $t$ -теста Стьюдента (статистический критерий  $\approx 0$ ,  $p$ -значение  $\approx 1$ ) указывают на то, что исходные и синтезированные данные статистически неразличимы и принадлежат к одной и той же выборке. Однако, вероятность срабатывания модели (47) ограничивается вероятностью отказа хотя бы одной из станций аврорального кластера (Рис. 5) и по имеющимся данным составляет 77.4%.



**Рис. 10.** Кластеризация IMAGE-магнитометров относительно параметра  $X$ , регистрируемого станцией Kilpisjärvi (KIL)

## ПОДХОД К МАШИННОМУ СОКРАЩЕНИЮ РЕГРЕССИОННЫХ ПРИЗНАКОВ DTI-ДВОЙНИКА МАГНИТНОЙ ВАРИАЦИОННОЙ СТАНЦИИ

Повысить надежность срабатывания DTI-двойника МС возможно путем модификации модели (7), например, за счет применения при оценке ее коэффициентов подхода, заключающегося во введении ограничения на норму вектора коэффициентов модели  $\hat{\beta}$ , что, например в случае метода LASSO (15) приведет к обращению в ноль некоторых ее коэффициентов, т. е. фактически исключению из выражения (7) одной или нескольких станций.

$$\hat{\beta}_{\text{ridge}} = \underset{\beta}{\operatorname{argmin}} \left( \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^k \beta_j^2 \right), \text{ где } \beta_0 = \frac{1}{n} \sum_{i=1}^n y_i; \quad (14)$$

$$\hat{\beta}_{\text{LASSO}} = \underset{\beta}{\operatorname{argmin}} \left( \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^k |\beta_j| \right), \quad (15)$$

$y$  – ожидаемый отклик модели;  $\lambda \geq 0$  – параметр регуляризации, имеющий смысл штрафа за сложность модели (при  $\lambda=0$  модели (8) и (9) вырождаются в МНК. Выражение (10), помимо повышения устойчивости повышает интерпретируемость модели, поскольку в итоге отбираются признаки, оказывающие наибольшее влияние на вектор ответов.

Так, например, в случае LASSO-регрессии (9), при  $\lambda = 1$  возможно добиться сокращения выражения (47) до 3 слагаемых ( $\beta_3, \beta_9, \beta_{12} = 0$ ), повысив тем самым вероятность срабатывания модели с 77.4 до 86.3%, при этом практически не теряя в точности (MSE  $\sim 12$  нТл) и сохраняя параметры корреляции и статистической однородности оригинальной и синтезируемой выборок на уровне модели (47). Обобщив выражений (8) и (9) позволяет рассматривать их как байесовские оценки (50) или модель типа elastic-net (51):

$$\tilde{\beta} = \underset{\beta}{\operatorname{argmin}} \left( \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^k |\beta_j|^q \right), \quad (16)$$

$$\hat{\beta}_{\text{elastic-net}} = \underset{\beta}{\operatorname{argmin}} \left( \sum_{i=1}^n \left( y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^k (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|) \right), \quad (17)$$

где при  $q=1$  – выражение (16) вырождается в (15),  
а при  $q=2$  – в (14).

где  $\lambda \sum_{j=1}^k (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|)$  – штраф за сложность модели.

Однако, как показал эксперимент, дальнейшее сокращение модели (47) путем вариаций параметров регуляризации в выражениях (14)-(17) малоэффективна и приводит к значительному росту MSE при относительно небольшом снижении числа ее слагаемых. Т.е. дальнейшее применение методов машинного сокращения нецелесообразно и последующую минимизацию числа признаков следует производить вручную, например, путем попарного сравнительного анализа статистик доступных предикатов.

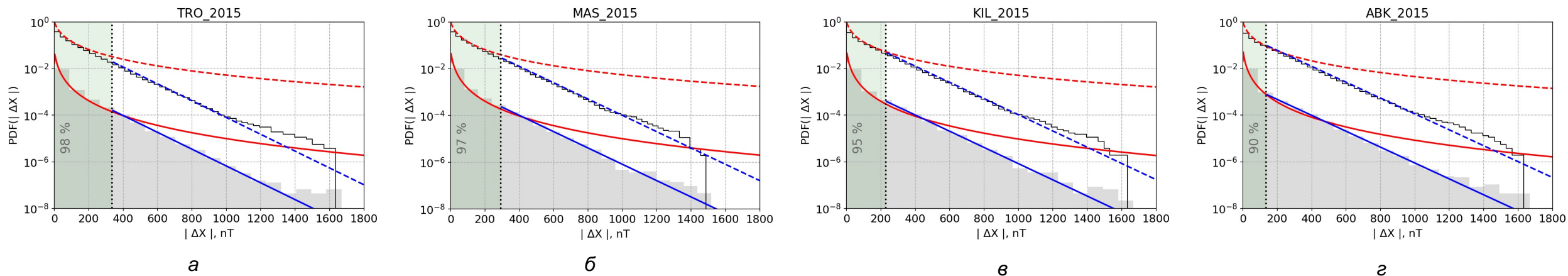


# ПОДХОД К СОКРАЩЕНИЮ ЧИСЛА РЕГРЕССОРОВО МОДЕЛИ ДТИ-ДВОЙНИКА МВС ПУТЕМ ВЫЯВЛЕНИЯ И АНАЛИЗА СТАТИСТИЧЕСКОЙ ОДНОРОДНОСТИ ВРЕМЕННЫХ РЯДОВ

Исключим из временных рядов каждой МС базовую линию в соответствии с выражением (18), нормализуем гистограмму и, руководствуясь критериями Колмогорова-Смирнова для полученных выборок  $|\Delta X|$ , подберем функцию, наилучшим образом аппроксимирующую характер распределения ее значений.

$$|\Delta X_{ij}| = |X_{ij} - \text{Me}(X_j)|, \quad (18)$$

где  $X_{ij}$  –  $i$ -е значение за  $j$ -е сутки  $X$ -составляющей на данной станции;  $X_j$  выборка за  $j$ -е сутки;  $i$  и  $j$  соответствуют порядковым номерам минуты в сутках (от 1 до 1440) и дня в году (от 1 до 365) соответственно.



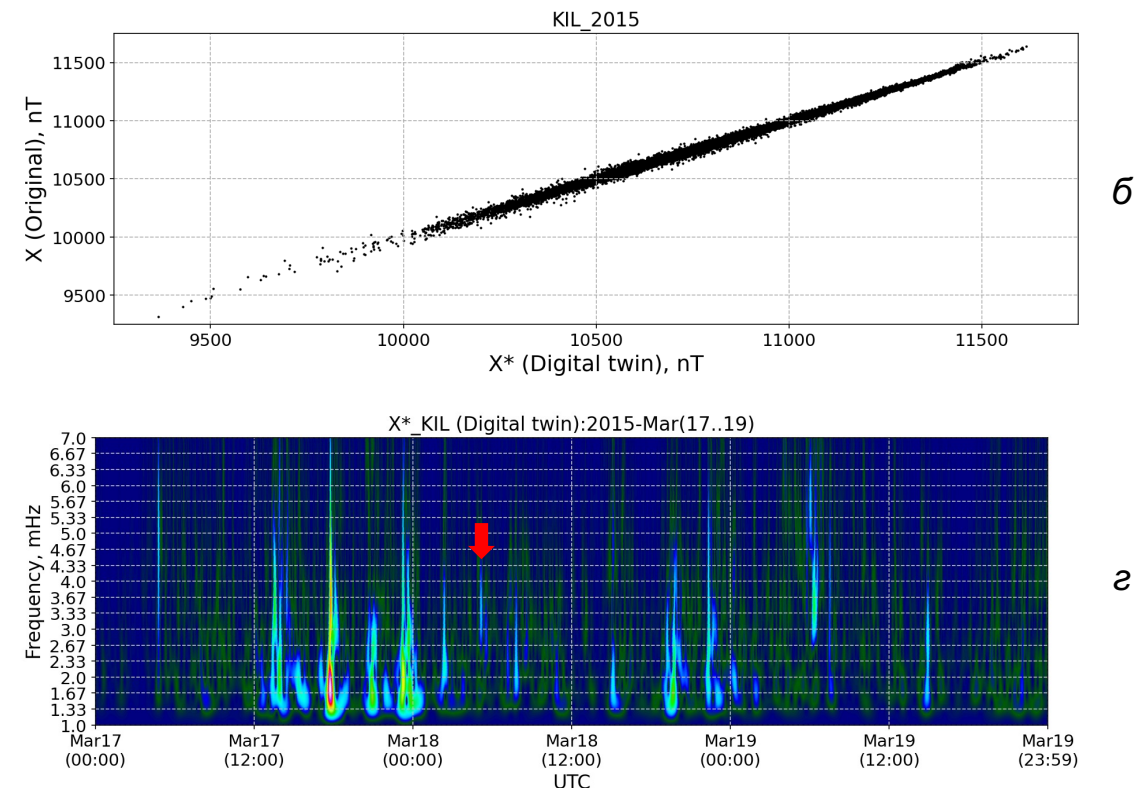
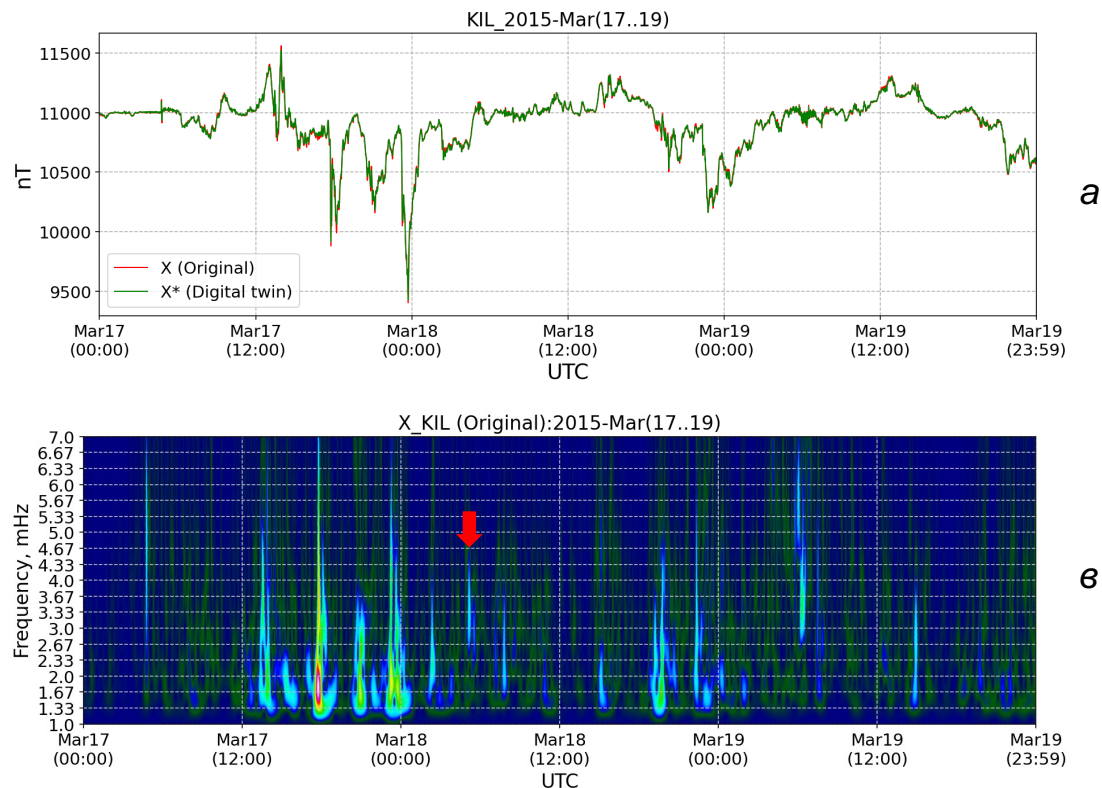
**Рис. 11** – Статистика возмущенных ГМВ: красные и синие сплошные (пунктирные) линии соответствуют функциям плотности вероятности (выживаемости) логнормального и экспоненциального законов распределения соответственно; черная сплошная линия – эмпирической выживаемости

$$\{KIL, TRO, MAS, ABK\} \in \{GEO_{Aurora} \cap GEO_{IMAGE}\} \quad (19)$$

$$X_{KIL}^* = \alpha + \beta_4 X_{TRO} + \beta_5 X_{MAS} + \beta_8 X_{ABK}, \quad \text{где } \alpha = 248.719 \text{ нТл}; \beta_4 = 0.29147953; \beta_5 = 0.28620399; \beta_8 = 0.44050465. \quad (20)$$

Параметры валидации:  $MSE \leq 28.9$  нТл;  $R = 0.999$ ;  $p\text{-value} \approx 1$ ;  $statistic \approx 0$ ;  $P_w = 99.954\%$

# ВАЛИДАЦИЯ МОДЕЛЕЙ В АМПЛИТУДНО-ЧАСТОТНОМ ДИАПАЗОНЕ



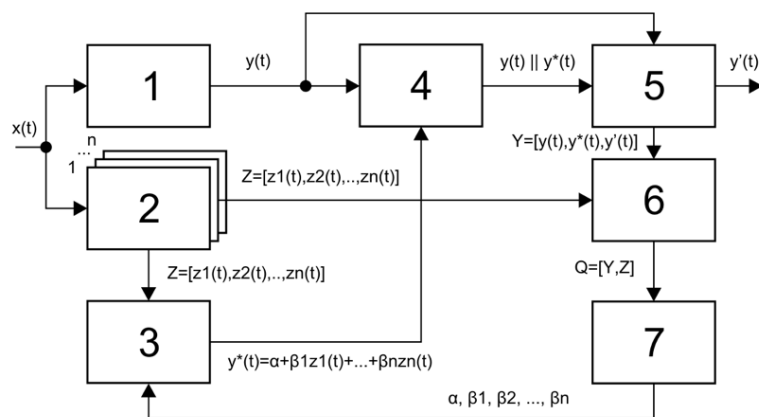
**Рис. 12.** Временные ряды, полученные на базе DTI-двойника магнитной станции и его физического прототипа (а); диаграмма рассеяния оригинальных и синтезированных цифровым двойником данных (б); вейвлет-спектрограммы верхней частотной области, сигнала, регистрируемого станцией Kilpisjärvi (в) и получаемые с выхода его ЦД-экземпляра (г)

**Табл. 3.** Параметры валидации и вероятности безотказной работы моделей DTI станции Kilpisjärvi (KIL)

Модель	Параметр	MSE, [нТл]	MSE, [%]	R	T-test Стьюдента		T <sub>W</sub> , [мин]	T <sub>F</sub> , [мин]	P <sub>W</sub> , [%]
					стат. критерий	р-знач.			
Выр. (8), МНК		11.5	0.51	0.89	~0	~1	406936	118664	77.423
Выр. (8), LASSO		12.0	0.54	0.87	~0	~1	453819	71781	86.343
Выр. (10), LASSO		28.9	1.25	0.87	~0	~1	523257	2343	99.554
Выр. (11), IDW (ρ=3)		114.1	4.94	0.58	~0	~1	406936	118664	77.423

Примечание: P<sub>W</sub> – ожидаемая вероятность срабатывания модели

# МОДЕЛЬ И АЛГОРИТМ ИНТЕГРАЦИИ ЦИФРОВОГО ДВОЙНИКА МВС В ПРОЦЕССЫ СБОРА, ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ И РЕГИСТРАЦИИ ГЕОМАГНИТНЫХ ДАННЫХ



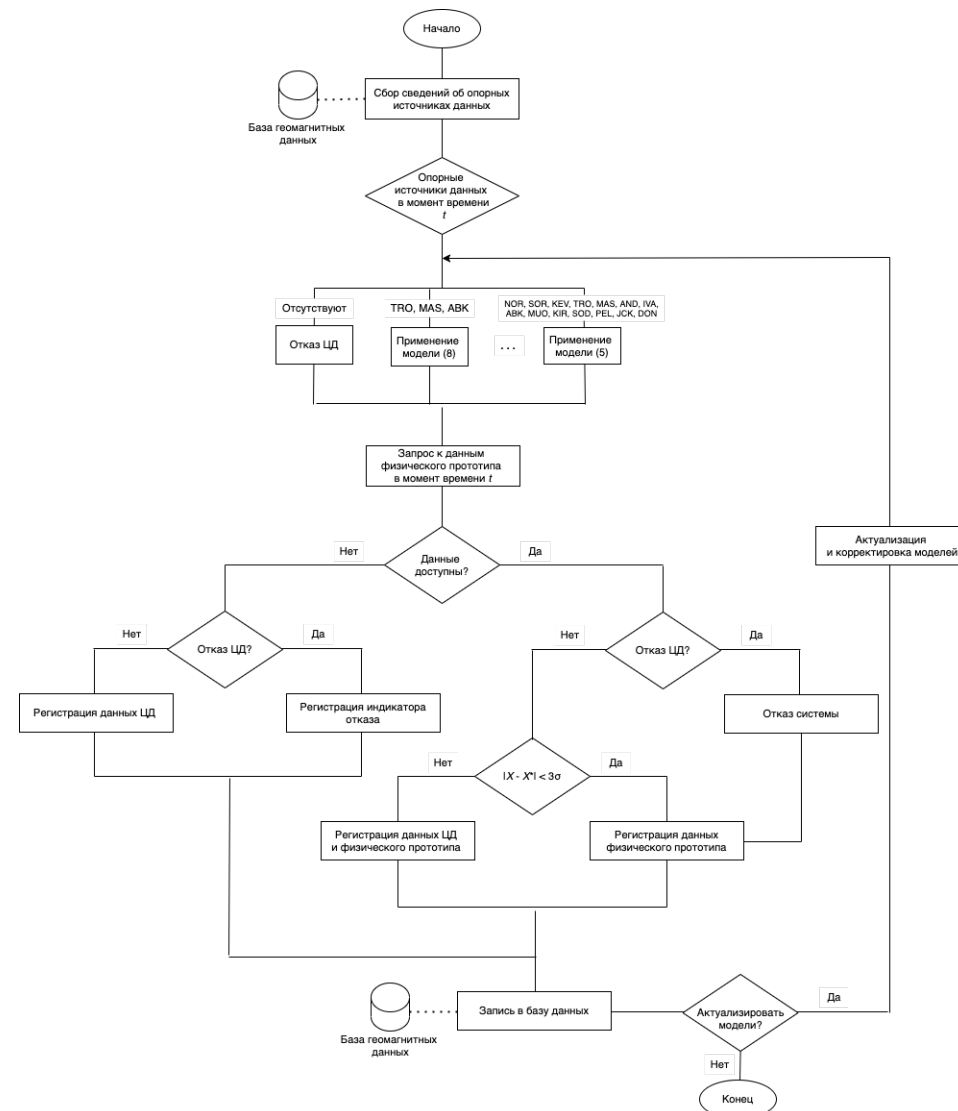
**Рис. 13.** Модель интеграции базового модуля цифрового двойника в процессы сбора и регистрации геомагнитных данных: 1 – МС-прототип; 2 – опорные источники данных (МС); 3 – математическое и алгоритмическое обеспечение DTI МС; 4 – устройство сравнения; 5 – выходной буфер; 6 – база геомагнитных данных; 7 – система корректировки весовых коэффициентов

Критерий регистрации данных физического прототипа DPI-двойника:

$$|X - X^*| < 3\sigma, \text{ где } \sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n ((x_i - x_i^*) - \bar{x})^2}; \quad (15)$$

$x_i^*$  и  $x_i$  – значения цифрового двойника и его физического прототипа соответственно в  $i$ -й момент времени

Предложенные модели и алгоритмы способны восстановить (на примере станции станции «Kilpisjärvi») 99.55% годовой информации, при этом 86.73% с ошибкой, не превышающей 12 нТл или 0.54 % от динамического диапазона



**Рис. 14.** Алгоритм (одного цикла) регистрации геомагнитных данных с применением технологии цифровых двойников



# ПРОГНОЗИРОВАНИЕ ГЕОМАГНИТНЫХ ДАННЫХ НА БАЗЕ МОДЕЛЕЙ СЕМЕЙСТВА ARIMA

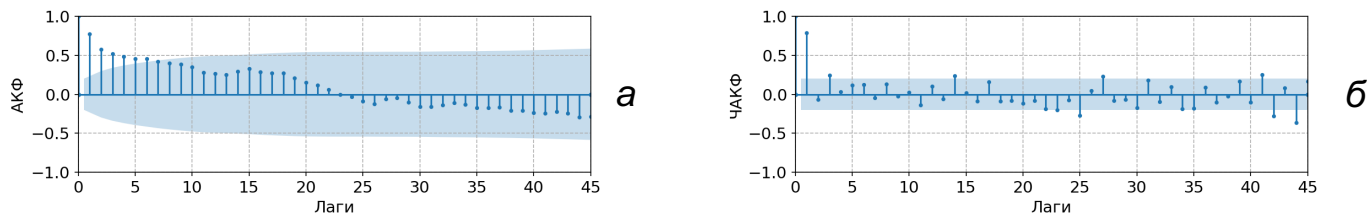
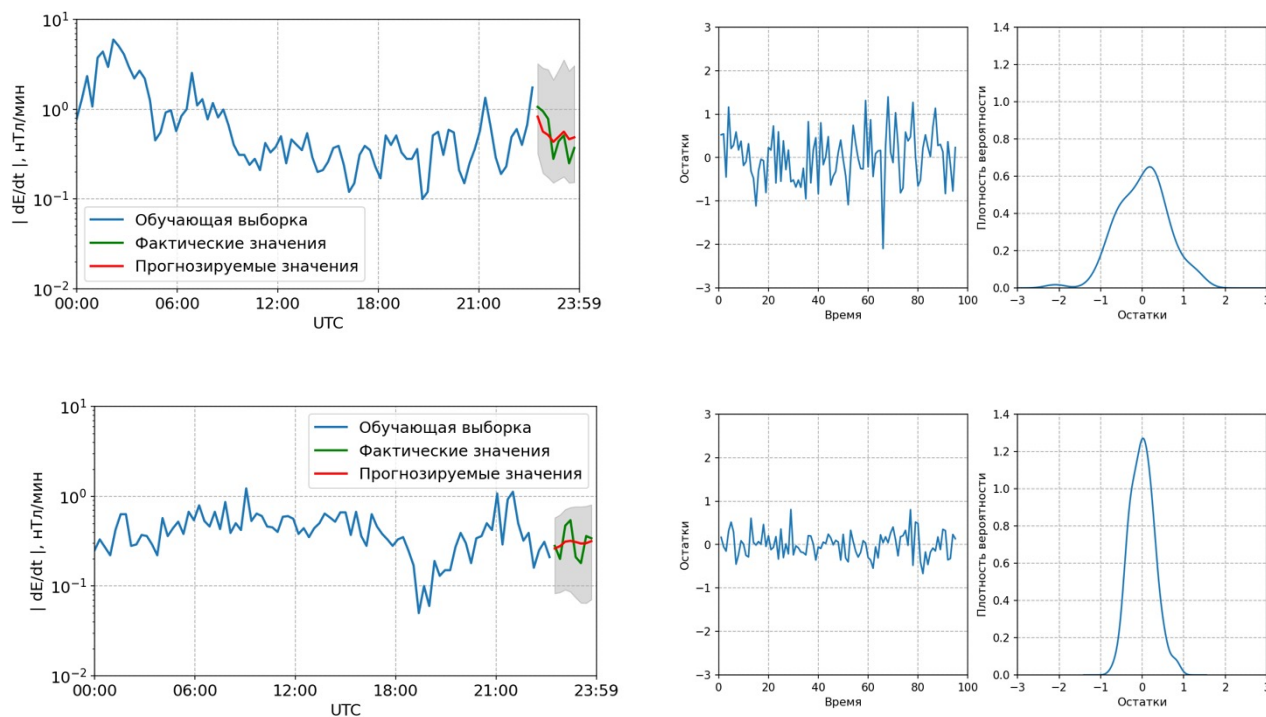


Рис. 12. – Характерная АКФ (а) и ЧАКФ (б) выборки  $|dE_{LOZ}/dt|$  за 9 марта 2015 г. (голубая область соответствует доверительному интервалу 95 %)

$$\Delta^d X_t = c + \sum_{i=1}^p a_i \Delta^d X_{t-1} + \sum_{j=1}^q b_j \varepsilon_{t-j} + \varepsilon_t, \quad (21)$$

где  $X_t$  – нестационарный временной ряд;  $c, a_i, b_j$  – параметры модели;  $\Delta^d$  – оператор разности временного ряда порядка  $d$ .



$$x_i(\lambda) = \begin{cases} \frac{(x_i + \alpha)^\lambda - 1}{\lambda}, & \text{при } \lambda \neq 0, \\ \ln(x_i + \alpha), & \text{при } \lambda = 0 \end{cases}, \quad (22)$$

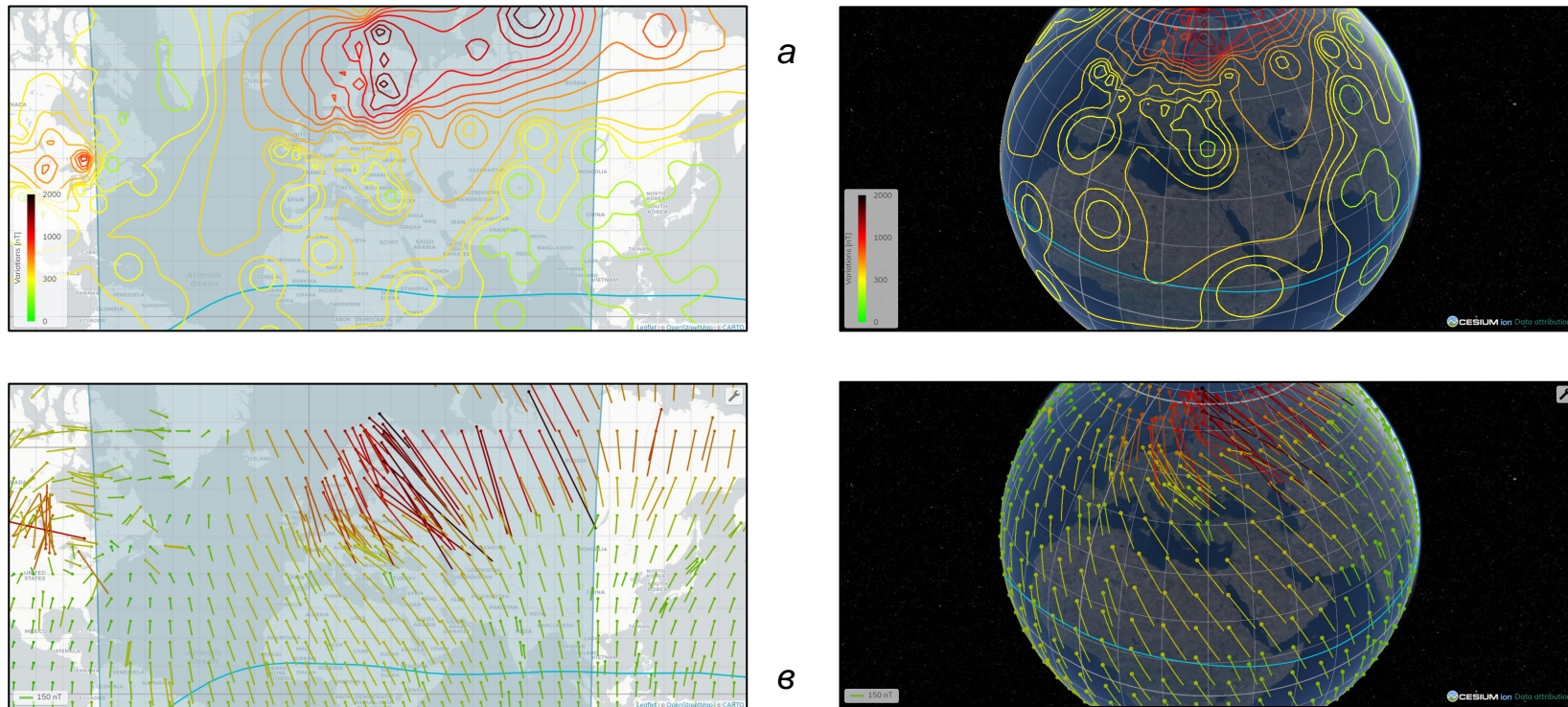
где  $\lambda$  – параметр, оптимальное значение которого можно определить путем максимизации функции правдоподобия:

$$f(x, \lambda) = -\frac{N}{2} \ln \left[ \sum_{i=0}^{N-1} \frac{(x_i(\lambda) - \bar{x}(\lambda))^2}{N} + (\lambda - 1) \sum_{i=0}^{N-1} \ln(x_i) \right], \quad (23)$$

где  $\bar{x}(\lambda) = \frac{1}{N} \sum_{i=0}^{N-1} x_i(\lambda)$ .

Рис. 15. – Краткосрочный (двухчасовой) прогноз  $|dE_{LOZ}/dt|$  на базе модели ARIMA(4,1,3)

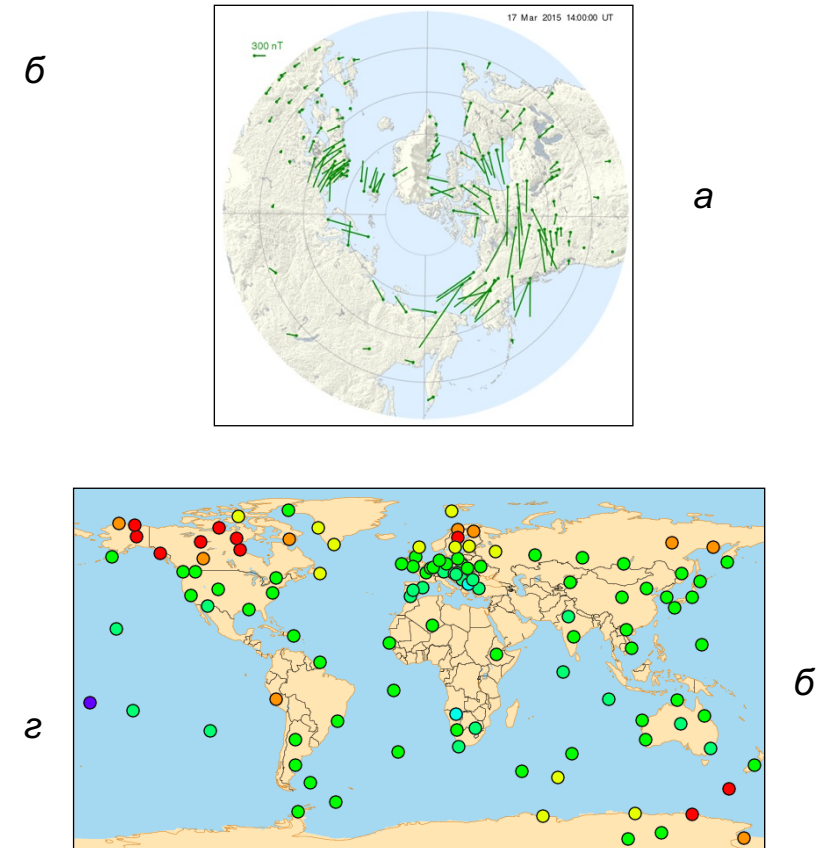
# РЕЗУЛЬТАТ РАБОТЫ ПОДСИСТЕМЫ ИНТЕРАКТИВНОЙ ДИНАМИЧЕСКОЙ ВИЗУАЛИЗАЦИИ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ГМП



**Рис. 16.** Результат визуализации возмущенной составляющей полного вектора ГМП на базе плоской карты (а) и виртуального глобуса (б) и горизонтальной составляющей ( $H$ -компонента вектора ГМП) в виде векторного поля на базе плоской карты (в) и виртуального глобуса (г). (Голубой линией показан магнитный экватор)

$$W = (dS/dt)_{\text{при } B > B_0} \quad (24)$$

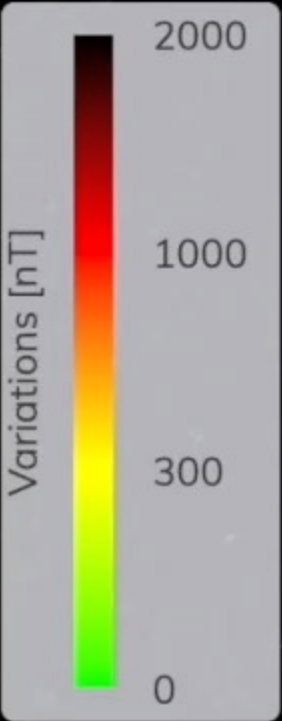
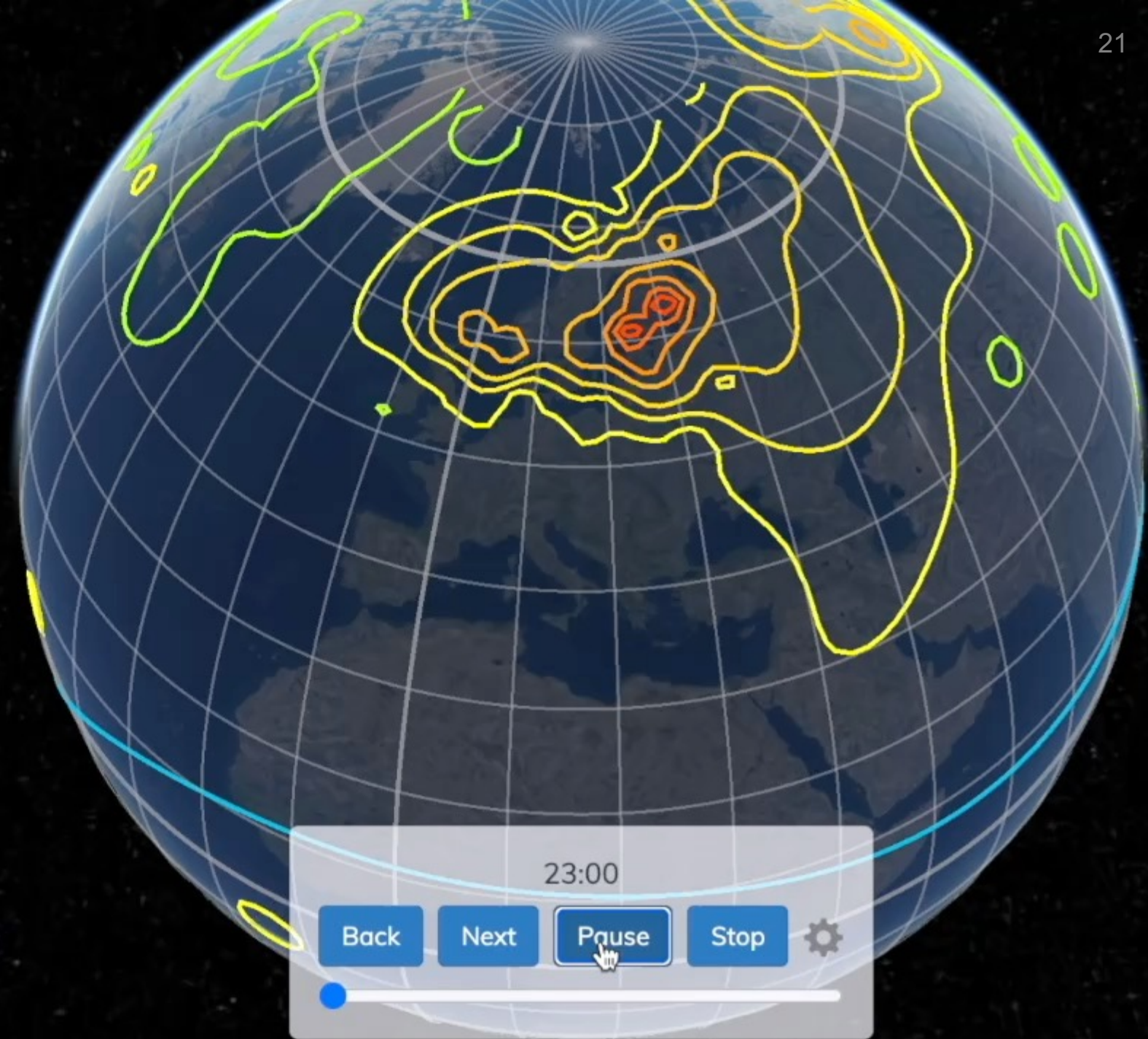
где  $S$  – площадь земной поверхности, внутри которой параметр ГМП удовлетворяет условию  $B > B_0$ ;  $t$  – время.



**Рис. 17.** Результат веб-визуализации геомагнитных возмущений (2015-03-17 14:00 UT) посредством порталов (а) и INTERMAGNET (б)



17 марта 2015 г



23:00

Back Next Pause Stop

A control panel for the visualization. It features four blue buttons labeled "Back", "Next", "Pause", and "Stop". A mouse cursor is hovering over the "Pause" button. To the right of the buttons is a gear icon for settings. Below the buttons is a white progress bar with a blue dot on the left side, indicating the current position in the visualization sequence.



# ПОДХОД К ДИАГНОСТИРОВАНИЮ ГЕОИНДУЦИРОВАННЫХ ТОКОВ НА БАЗЕ РЕГРЕССИОННЫХ МОДЕЛЕЙ

Табл. 4 Корреляция значений |JVКН| со значениями геомагнитных вариаций (данные усреднены по 15-минутным интервалам)

	$ dY_{LOZ}/dt $	$ dY_{IVA}/dt $	$ dY_{SOD}/dt $	$ dY_{KEV}/dt $	IE-index
r	0.882	0.878	0.847	0.841	0.772

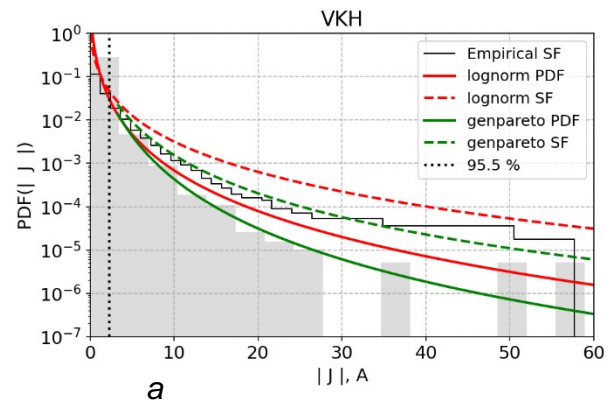
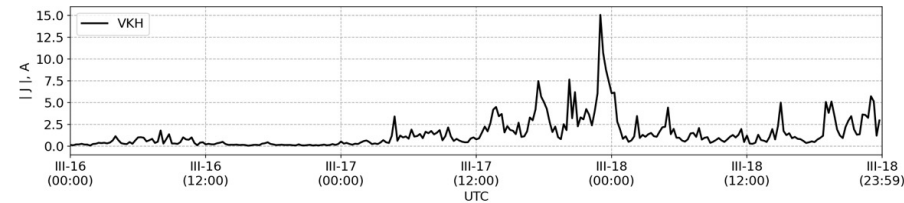
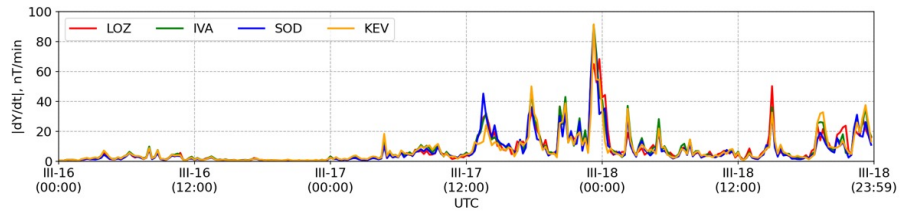


Рис. 14 Вариации значений целевой функции и признаков переменных

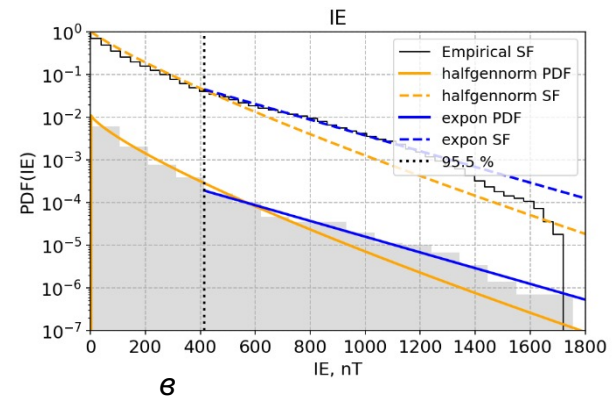
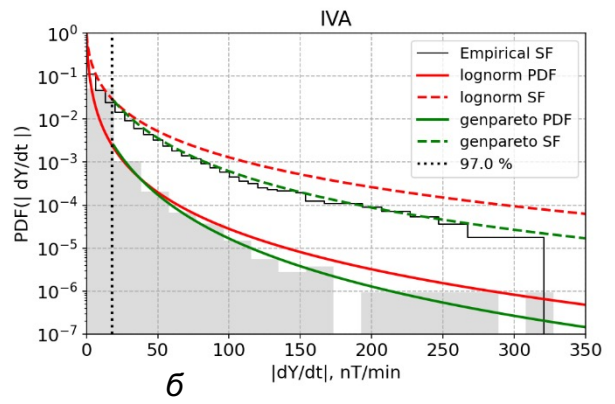


Рис. 19. Анализ однородности информации в процессе формирования информационной среды многоуровневой системы ЦД: а – геоиндуцированные токи на трансформаторной станция VKH; б и в – вариабельность восточной составляющей и проксителлурическое поле на станции IVA; IE-индекс ГМА

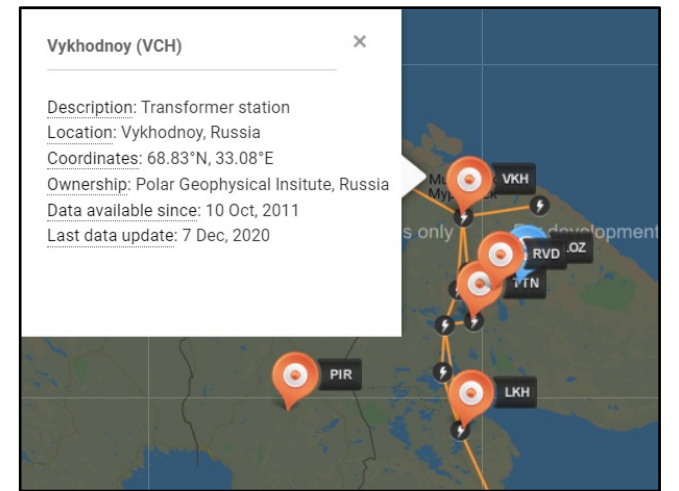


Рис. 18. Структура многоуровневой системы цифровых двойников (а) и география ее физических прототипов (б)

# ВАЛИДАЦИЯ ДИАГНОСТИЧЕСКОЙ МОДЕЛИ ГИТ НА ОСНОВЕ РЕГРЕССИОННОГО ПОДХОДА И ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Оценка коэффициента детерминации  $R^2$  показала, что в рамках решаемой задачи подходы, основанные на методе множественной линейной регрессии и искусственной нейронной сети с функцией активации на базе линейного выпрямителя (ReLU), являются наилучшими ( $R^2=0.816$  и  $R^2=0.814$  соответственно). Так, например в случае применения диагностики ГИТ посредством множественной линейной регрессии (45), искомая модель будет иметь вид выражения (57).

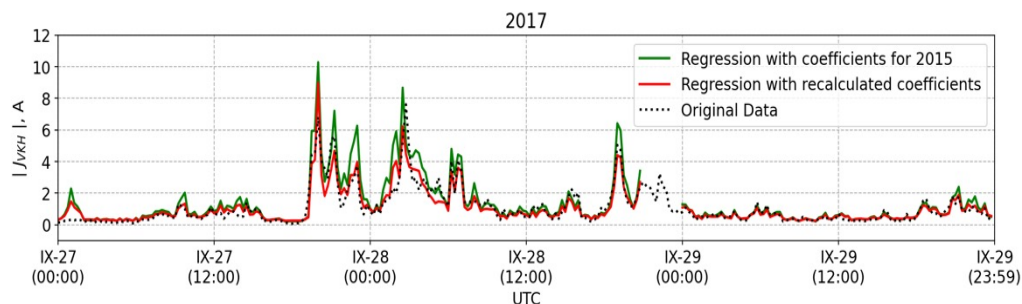
$$f(x, \beta) = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k = \sum_{j=1}^k \beta_j x_j = x^T \beta, \quad (24)$$

где  $x^T = (x_1, x_2, \dots, x_k)$  – вектор регрессоров;  $\hat{\beta} = (\beta_1, \beta_2, \dots, \beta_k)^T$  – вектор столбец коэффициентов;  $k$  – число признаков модели.

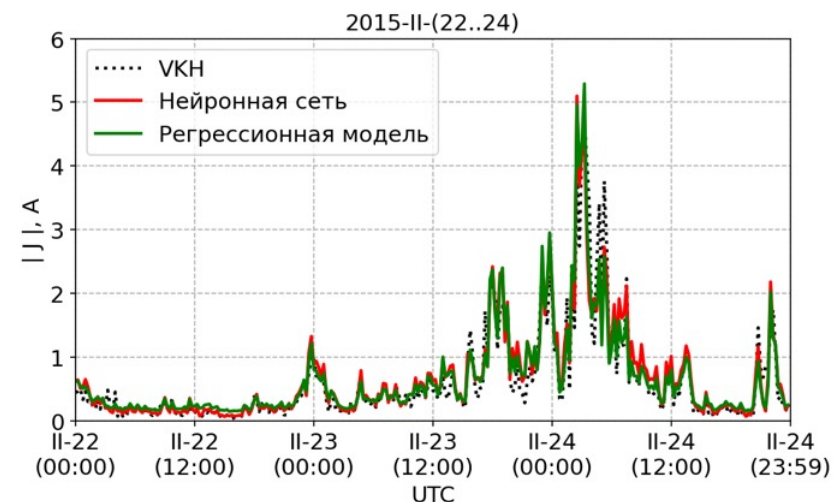
$$|J_{VKH}| = \beta_0 + \beta_1 \left| \frac{dY_{LOZ}}{dt} \right| + \beta_2 \left| \frac{dY_{IVA}}{dt} \right| + \beta_3 \left| \frac{dY_{KEV}}{dt} \right| + \beta_4 IE, \quad (25)$$

где  $\beta_0 = 0.1$ ;  $\beta_1 = 90.56 \cdot 10^{-3}$ ;  $\beta_2 = 32.25 \cdot 10^{-3}$ ;  $\beta_3 = 32.36 \cdot 10^{-3}$ ;  $\beta_4 = 0.37 \cdot 10^{-3}$ .

**RMSE ~ 0.34 A; R ~ 0.9; Т-тест Стьюдента ( $p$ -значение = 0.87)**



а



б

**Рис. 20.** Верификация данных с выхода ЦД ТС «VYKODNOY» (VKH) и его физического прототипа

Pc5 пульсации (периоды порядка нескольких мин) являются одним из наиболее мощных проявлений волновой активности в околоземном космическом пространстве. Pc5 пульсации способны заметно модулировать потоки захваченных энергичных электронов и протонов в магнитосфере [Sarris et al., 2007], [Белаховский, Пилипенко, Самсонов 2013].

Длительное существование (несколько часов) ГИТ умеренной интенсивности, индуцированных этими пульсациями, может быть более опасным для долговременной эксплуатации сетей, чем кратковременные всплески ГИТ во время начал суббурь и бурь. [Сахаров, Ягова, Пилипенко 2021]

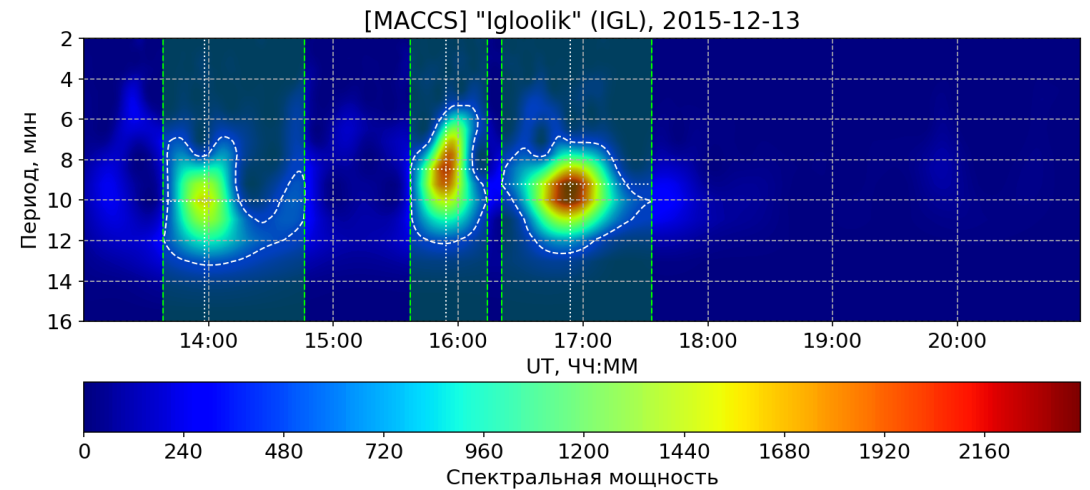
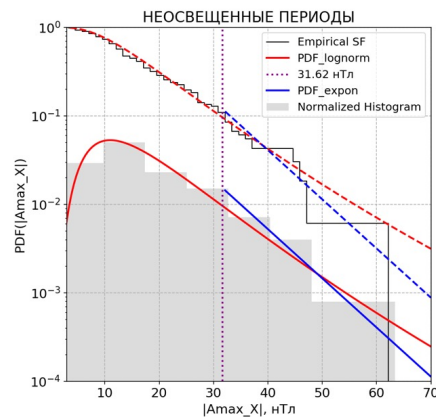
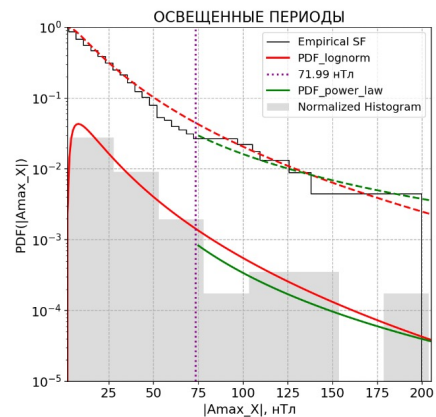
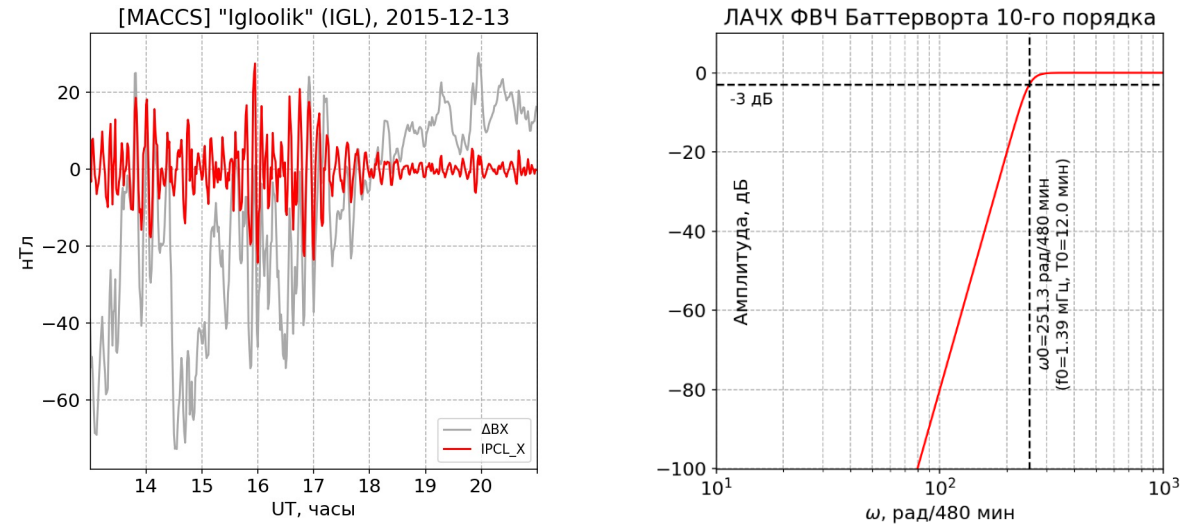
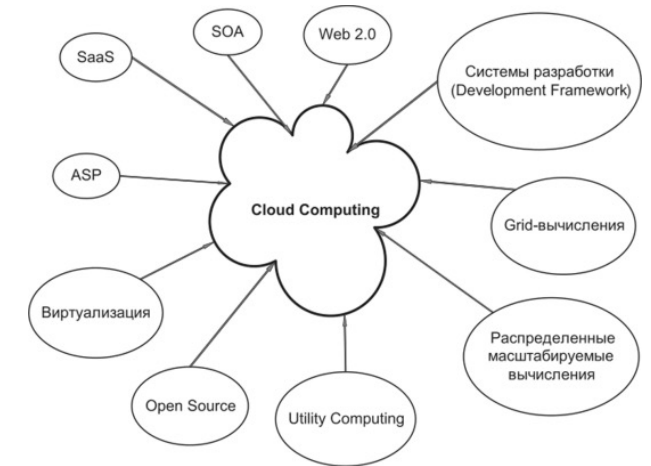


Рис. 21. Результаты автоматизированного статистического анализа параметров IPCL пульсаций



# ОБЛАЧНЫЕ СЕРВИСЫ И СЕРВИСЫ ОБЛАЧНЫХ ВЫЧИСЛЕНИЙ

- Amazon Web Services (AWS) – 2006 г.
  - Microsoft Azure – 2010 г.
  - Google Cloud Platform – 2011 г.
  - Yandex Cloud
  - VK Cloud Solutions
- и др.



**Инфраструктура как услуга (IaaS)** Основная группа облачных вычислительных служб. В схеме IaaS вы арендуете ИТ-инфраструктуру (серверы, виртуальные машины, хранилище, сети и операционные системы) у облачного поставщика с системой оплаты по мере использования.

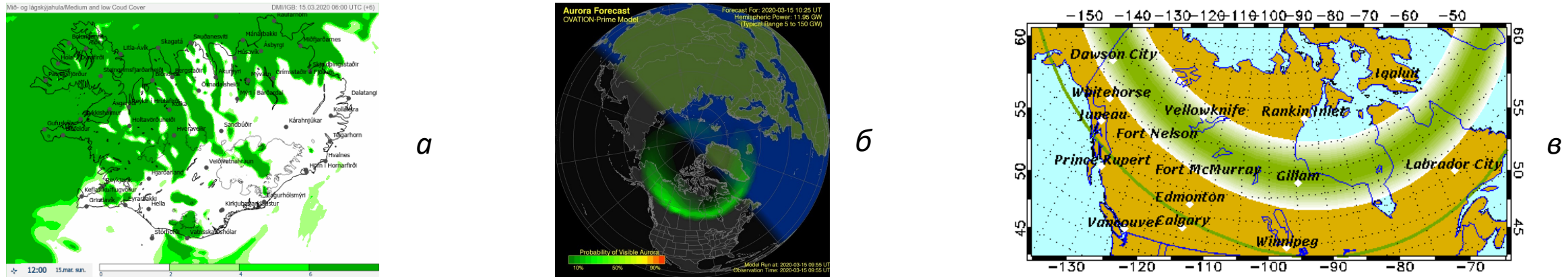
## Платформа как услуга (PaaS)

Платформа как услуга относится к облачным вычислительным службам, которые поставляют среду, доступную по запросу, для разработки, тестирования, доставки приложений программного обеспечения и управления ими. PaaS упрощает разработчикам задачу быстрого создания веб-приложений или мобильных приложений без необходимости иметь дело с базовой инфраструктурой серверов, хранилища, сети и баз данных, необходимых для разработки.

## Программное обеспечение как услуга (SaaS)

Программное обеспечение как услуга (SaaS) — это метод доставки программного обеспечения для приложений через Интернет по запросу и обычно на основе подписки. В схеме SaaS облачные поставщики размещают программное обеспечение и базовую инфраструктуру и управляют ими, а также занимаются всем обслуживанием, включая обновление программного обеспечения и установку исправлений безопасности. Пользователи подключаются к приложению по Интернету, обычно с помощью веб-браузера на своем телефоне, планшете или ПК.

## СЕРВИСЫ ПРОГНОЗИРОВАНИЯ ПОЛЯРНЫХ СИЯНИЙ



**Рис. 22** Пример визуализации краткосрочного прогноза вероятности видимости полярных сияний метеорологической службой Исландии (а), Аляскинским университетом в Фэрбенксе (б), США и сервисом NOAA (в)

В качестве базового математического обеспечения используется эмпирическая модель OP [Newell, 2009], принимающая на вход параметры солнечного ветра и ММП. Основу модели составляют регрессионные коэффициенты, определяемые сезоном, типом полярных сияний и набором координат, всего включающие 245760 регрессионных соответствий (4 сезона × 4 типа полярных сияний × 96 отсчетов местного времени × 160 отсчетов геомагнитной широты). Локальная мощность сияний  $p_A$  рассчитывается по соотношению

$$\begin{cases} p_A(\text{MLat}, \text{MLT}, Q, G) = a + b \left( \frac{d\Phi_{\text{МП}}}{dt} \right), \\ \frac{d\Phi_{\text{МП}}}{dt} = V^{4/3} B_T^{2/3} \sin^{8/3} \left( \frac{\theta}{2} \right), \end{cases} \quad (26)$$

где  $\text{MLat} \in \{[50...90] \cup [-50...-90]\}$  – магнитная широта с шагом  $0.25^\circ$ ;  $\text{MLT} \in [0...24)$  – магнитное локальное время с шагом дискретизации  $0.25$  ч;  $a$  и  $b$  – регрессионные коэффициенты, определяемые сезоном  $Q$  и типом полярного сияния  $G$  для каждого значения  $\text{MLat}$  и  $\text{MLT}$ ;  $d\Phi_{\text{МП}}/dt$  – скорость изменения магнитного потока на магнитопаузе, Вб/с;  $V$  – скорость солнечного ветра, км/с;  $B_Y$  и  $B_Z$  – компоненты ММП, нТл;  $\theta = \arctan(B_Y/B_Z)$ ;  $B_T = (B_Y^2 + B_Z^2)^{1/2}$ .

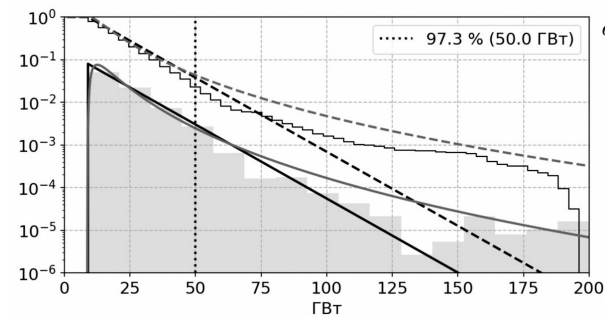
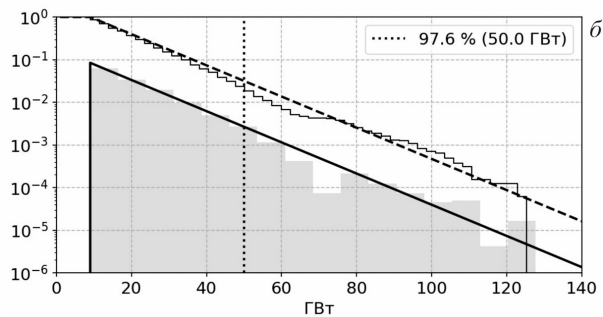
Объединив 4 основных типа полярных сияний можно оценить характер пространственного распределения суммарной мощности светимости:

$$p_A(\text{MLat}, \text{MLT}, Q) = \sum_{g=1}^G \left[ a_g + b_g \left( \frac{d\Phi_{\text{МП}}}{dt} \right) \right], \quad (27)$$

Просуммировав выражение (2) по долготе ( $M = 186$ ) и широте ( $N=96$ ) отдельно для каждого полушария можно получить интегральную оценку мощности полярных сияний в северном (N) и южном (S) полушариях [Newell, 2009]:

$$P_A^{(N,S)}(Q) = \sum_{g=1}^G \sum_{m=1}^M \sum_{n=1}^N \left[ a_{g,m,n} + b_{g,m,n} \left( \frac{d\Phi_{\text{МП}}}{dt} \right) \right]; \quad (28)$$

Для  $P_A \leq 20$  ГВт имеет место слабое или визуально неразличимое сияние; при  $20 \leq P_A \leq 50$  наблюдение сияния возможно лишь на небольшом от него расстоянии; для  $50 < P_A \leq 100$ , сияние должно быть видимым невооруженным глазом;  $P_A > 100$  соответствует экстремальной авроральной активности и значимому расширению аврорального овала к низким широтам.



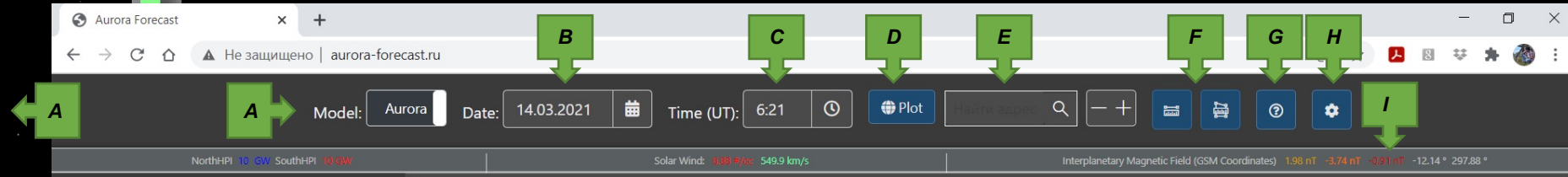
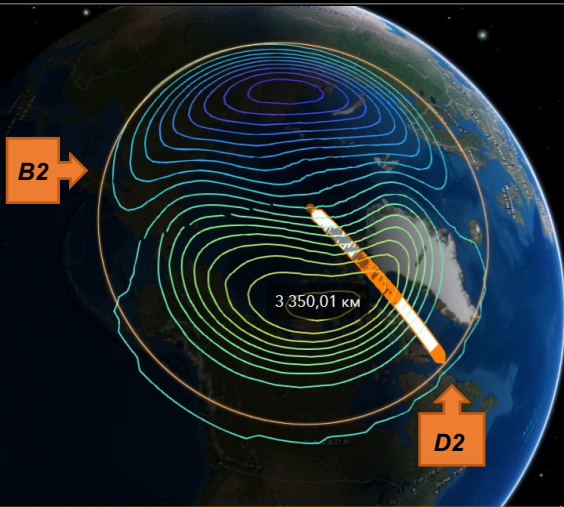
Для минимума солнечного цикла среднегодовую вероятность наблюдения полярных сияний можно оценить как  $\sim 2.4\%$  ( $P_A^N \geq 50$  ГВт), а вероятность наблюдения экстремальных сияний составит  $\sim 0.1\%$  ( $P_A^N \geq 100$  ГВт).



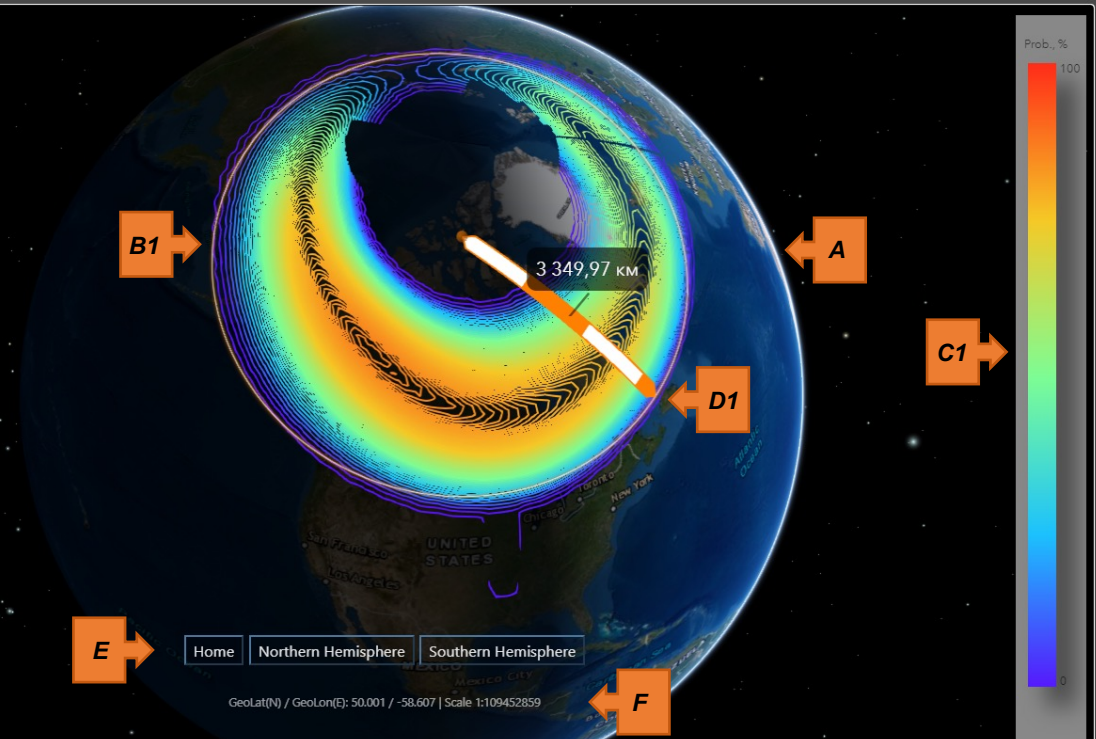
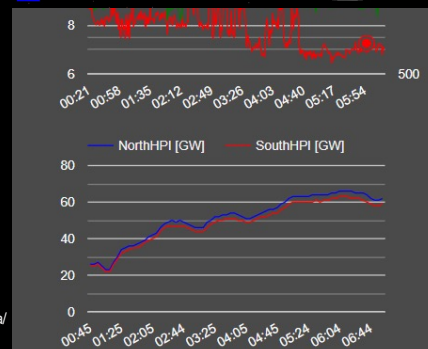
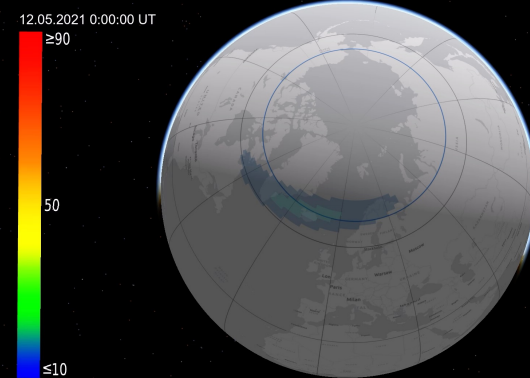
# СИСТЕМА ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА БАЗЕ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ПАРАМЕТРОВ В АВРОРАЛЬНОЙ ЗОНЕ «AURORA-FORECAST»

<http://aurora-forecast.ru/>

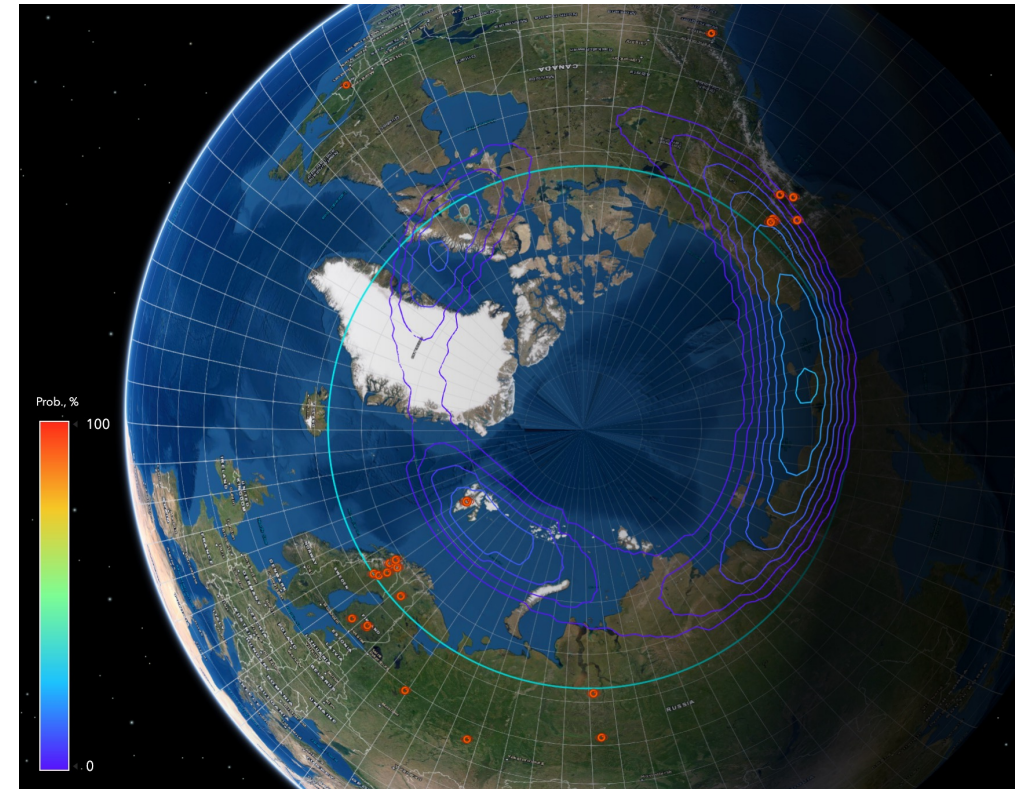
**ПАНЕЛЬ УПРАВЛЕНИЯ:** **A** – выбор параметра визуализации (пространственное распределение вероятности наблюдения полярных сияний или электрического потенциала в авроральной зоне); **B** и **C** – выбор даты и времени визуализации соответственно; **D** – построение пространственного распределения для выбранных даты, времени и модели; **E** – функционал обратного геокодирования (фокусировка виртуального глобуса, согласно полнотекстовому запросу, например, г. Мурманск, Россия); **F** – функционал для геометрического анализа картины пространственного распределения выбранного параметра; **G** – информация о проекте; **H** – информация о проекте; **I** – текущее состояние параметров солнечного ветра (СВ) и межпланетного магнитного поля (ММП)



**ПАНЕЛЬ ОТОБРАЖЕНИЯ:** **A** – виртуальный (интерактивный) глобус; **B1** и **B2** – характер пространственного распределения вероятности наблюдения полярных сияний и электрического потенциала соответственно (по состоянию на 2021-03-14, 06:21 UT); **C1** и **C2** – интерпретация цветowych схем визуализации; **D1** и **D2** – результат оценки диаметра аврорального овала (~3350 км); **E** – управление камерой наблюдения (вид на северное полушарие, вид на южное полушарие, вид по умолчанию); **F** – отображение вспомогательных параметров (географические координаты текущего местонахождения курсора и текущий масштаб); **G** – характер изменения параметров СВ и ММП за последние 6 часов; **H** – оценка интегральной мощности NHP и SHPI, ГВт.



№ п/п	Координаты, [град.]		Название	URL
	Сев. широта	Вост. долгота		
1	78.15	16.04	Svalbard, Norway	<a href="https://aurorainfo.eu/aurora-live-cameras/svalbard-norway-all-sky-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/svalbard-norway-all-sky-aurora-live-camera.jpg</a>
2	69.35	19.13	Ramfjordmoen, Norway	<a href="https://aurorainfo.eu/aurora-live-cameras/ramfjordmoen-norway-all-sky-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/ramfjordmoen-norway-all-sky-aurora-live-camera.jpg</a>
3	69.02	20.85	Kilpisjarvi, Finland	<a href="https://aurorainfo.eu/aurora-live-cameras/kilpissafarit-all-sky-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/kilpissafarit-all-sky-aurora-live-camera.jpg</a>
4	67.85	20.41	Kiruna, Sweden	<a href="https://aurorainfo.eu/aurora-live-cameras/kiruna-sweden-all-sky-aurora-live-camera.jpg?1634790312">https://aurorainfo.eu/aurora-live-cameras/kiruna-sweden-all-sky-aurora-live-camera.jpg?1634790312</a>
5	67.41	26.60	Sodankyla, Finland	<a href="https://aurorainfo.eu/aurora-live-cameras/sodankyla-finland-all-sky-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/sodankyla-finland-all-sky-aurora-live-camera.jpg</a>
6	66.95	19.82	Porjus, Sweden	<a href="https://aurorainfo.eu/aurora-live-cameras/porjus-sweden-west-view-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/porjus-sweden-west-view-aurora-live-camera.jpg</a>
7	66.58	18.85	Jokkmokk, Sweden	<a href="https://jokkmokk.jp/photo/nr3/latest.jpg">https://jokkmokk.jp/photo/nr3/latest.jpg</a>
8	66.00	76.00	Novy Urengoy, Russia	<a href="https://starvisor.ru/wp-content/uploads/webcam/capture_nur.jpg">https://starvisor.ru/wp-content/uploads/webcam/capture_nur.jpg</a>
9	64.75	-147.3	North Pole, Alaska	<a href="https://auroranotify.com/image10.jpg">https://auroranotify.com/image10.jpg</a>
10	63.07	-151.00	Denali, Alaska	<a href="http://denaliview1.ddns.net:8080/nph-jpeg.cgi">http://denaliview1.ddns.net:8080/nph-jpeg.cgi</a>
11	62.39	26.43	Hankasalmi, Finland	<a href="https://aurorasnow.fmi.fi/public_service/images/latest_SIR_AllSky.jpg">https://aurorasnow.fmi.fi/public_service/images/latest_SIR_AllSky.jpg</a>
12	62.30	-145.27	Gakona, Alaska	<a href="http://optics.gi.alaska.edu/realtime/latest/gak_latest.jpg">http://optics.gi.alaska.edu/realtime/latest/gak_latest.jpg</a>
13	62.25	26.59	Hankasalmi, Finland	<a href="https://aurorainfo.eu/aurora-live-cameras/hankasalmi-finland-all-sky-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/hankasalmi-finland-all-sky-aurora-live-camera.jpg</a>
14	61.58	-147.45	Wasilla, Alaska	<a href="https://auroranotify.com/kickaxcamimage.jpg">https://auroranotify.com/kickaxcamimage.jpg</a>
15	61.52	23.50	Tampere, Finland	<a href="https://aurorainfo.eu/aurora-live-cameras/tampere-finland-aurora-live-camera.jpg">https://aurorainfo.eu/aurora-live-cameras/tampere-finland-aurora-live-camera.jpg</a>
16	61.00	77.00	Strezhevoy, Russia	<a href="https://starvisor.ru/wp-content/uploads/webcam/185/capture.jpg">https://starvisor.ru/wp-content/uploads/webcam/185/capture.jpg</a>



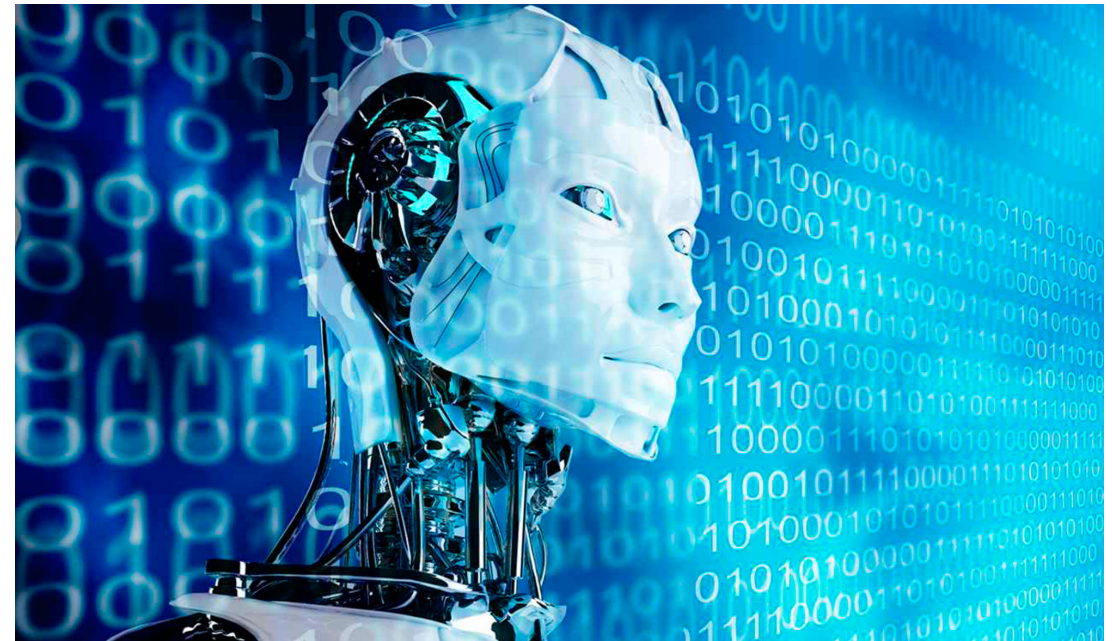
Доли ошибок первого ( $B/(A+B)$ ) и второго ( $C/(C+D)$ ) рода модели составляют 14% и 26 % соответственно. При этом доля совпадения предсказанных и наблюдаемых полярных сияний ( $A/(A+B)$ ), составляет 86 %. Общую точность компьютерной модели можно оценить как  $(A+D)/(A+B+C+D) \sim 77\%$ .

A – число случаев, когда прогноз был положительный, при этом сияние наблюдалось  
 B – число случаев, когда прогноз был положительный, при этом сияние не наблюдалось  
 C – число случаев, когда прогноз был отрицательный, при этом сияние наблюдалось  
 D – число случаев, когда прогноз был отрицательный, при этом сияние не наблюдалось



## ОБЛАСТИ ПРИМЕНЕНИЯ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В ЗАДАЧАХ АНАЛИЗА И ОБРАБОТКИ ГЕОФИЗИЧЕСКОЙ ИНФОРМАЦИИ

- Прогнозирование и диагностирование геофизических параметров;
- Реконструкция информационного сигнала в пространственно-временной области;
- Выделение признаков, связанных с вариациями наблюдаемого параметра;
- Обнаружение и классификация событий по выделенным признакам;
- Разработка цифровых двойников геофизических станций.







**Воробьев Андрей Владимирович**

Доктор технических наук

Доцент кафедры геоинформационных систем Уфимского государственного  
авиационного технического университета

Старший научный сотрудник Геофизического центра РАН

geomagnet@list.ru