

10-Я МЕЖДУНАРОДНАЯ КОНФЕРЕНЦИЯ  
ПО ГЛУБОКОМУ ОБУЧЕНИЮ  
В ВЫЧИСЛИТЕЛЬНОЙ ФИЗИКЕ (DLCP2026)

Использование алгоритмов машинного обучения  
для доменной адаптации при переходе  
между данными космических аппаратов

**Пятикоп А.В.<sup>1</sup>, Гаджиев И.М.<sup>1</sup>, Баринов О.Г.<sup>2</sup>, Доленко С.А.<sup>2</sup>**

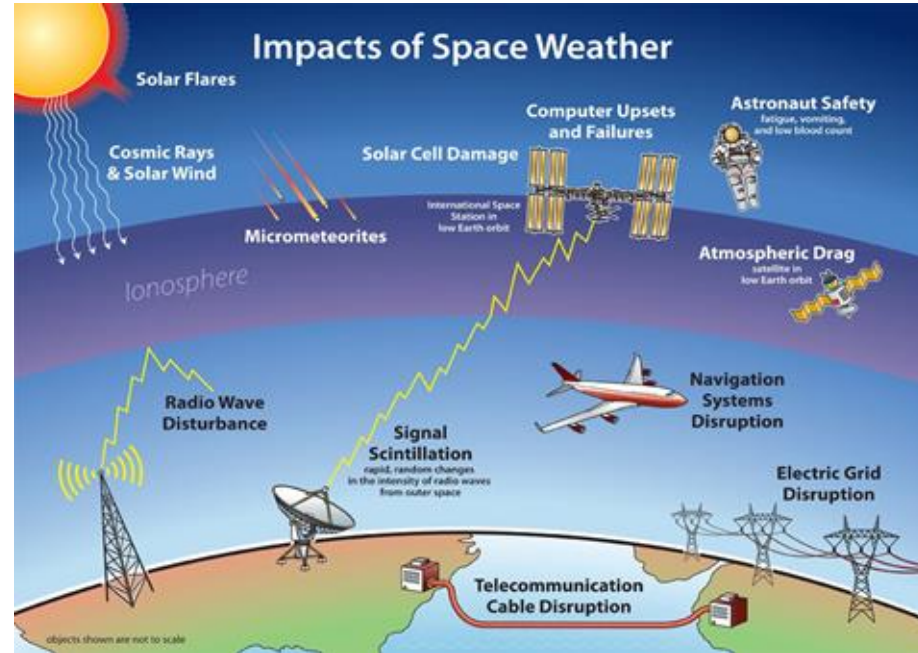
*<sup>1</sup>Физический факультет МГУ имени М.В. Ломоносова*

*<sup>2</sup>Научно-исследовательский институт ядерной физики имени Д.В. Скобельцына  
МГУ имени М.В. Ломоносова (НИИЯФ МГУ)*

# Космическая погода

Сильные магнитные бури  
могут вызвать  
нарушение работы:

- космических аппаратов
- радиосвязи
- линий электропередач
- электрических сетей
- навигационных систем

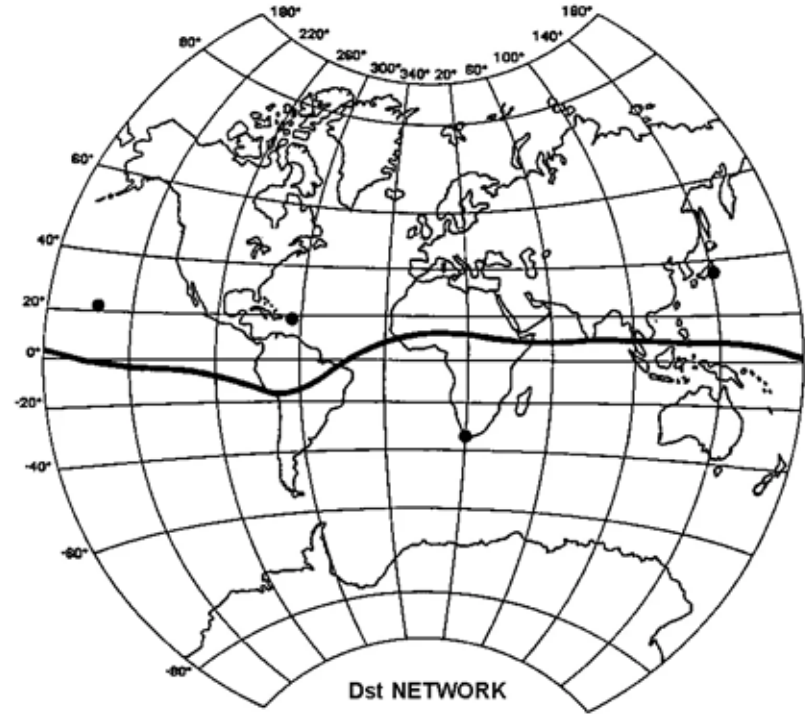


# Оценка интенсивности геомагнитного возмущения

## Dst-индекс

(Disturbance Storm Time index) – оценка осесимметричной относительно геомагнитного диполя компоненты возмущенного геомагнитного поля.

Определяется на основе измерений магнитного поля на приэкваториальных станциях: Сан-Хуан, Херманус, Какиока, Гонолулу.



# Источники данных

Измерения параметров солнечного ветра и межпланетного магнитного поля, необходимых для прогнозирования геомагнитной обстановки, проводятся космическими аппаратами (КА) в точке Лагранжа L1 между Солнцем и Землей ( $\approx 1,5$  млн км от Земли), где гравитационное равновесие позволяет минимизировать расход топлива.

## Основные КА:

- **ACE** (Advanced Composition Explorer) – запущен в 1997 г.  
(планируется вывод из эксплуатации в 2027–2028 гг.)
- **DSCOVR** (Deep Space Climate Observatory) – запущен в 2016 г.

# Актуальность работы

КА ACE скоро **будет выведен из эксплуатации** из-за истощения запаса топлива.

**Наблюдаются проблемы** с передачей данных:

- фиксируются пропуски данных в реальном времени, **ухудшающие качество оперативного прогноза.**

Необходим переход на данные КА DSCOVР, однако:

- его **объём данных меньше**
- качество данных - **ниже**

В ближайшие годы предполагается появление новых КА на замену существующих.

## Цель работы

Разработать и исследовать методику доменной адаптации, позволяющую согласовать данные ACE и DSCOV, снизить ошибку прогнозирования индекса Dst и обеспечить сохранение качества прогноза после вывода из эксплуатации КА ACE.

# Доменная адаптация

**Доменная адаптация** – подход в рамках машинного обучения, направленный на **перенос обучения модели с одного источника данных (source domain) на другой (target domain)**, имеющий отличающиеся распределения признаков или целевых меток.

**Цель адаптации** – сохранить высокую точность модели при работе на данных, которые **статистически отличаются** от обучающего набора, без необходимости собирать и размечать большой набор данных в новой области.

В данной задаче **домен** – это данные, полученные с конкретного КА (ACE или DSCOVN).

**Данные имеют различия** из-за конструктивных особенностей приборов, разного положения в пространстве и шума.

# Доменная адаптация

**Проблема:** модель, обученная на данных одного КА, демонстрирует ухудшение качества при применении к данным другого КА.

**Идея подхода:** построить отображение между доменами, позволяющее привести данные одного КА к статистике другого и обеспечить совместное использование большого массива данных.

# Этапы работы

1. Предварительная **обработка данных** KA ACE и DSCOVr:
  - **интерполяция** пропусков, имеющих длину меньше 12 часов;
  - **подбор глубины погружения** для каждого признака **по автокорреляционной функции**: в качестве критерия использовалась задержка, при которой автокорреляция падала ниже  **$1/e$** , также рассматривался вариант с фиксированной глубиной **24 часа** для всех параметров;
  - нормировка данных и удаление пропусков;
2. Реализация и сравнение методов обучения **прогностических моделей без применения адаптации**
3. Реализация и сравнение различных **сценариев использования** и **методов доменной адаптации**

# Используемые данные

- **Параметры межпланетного магнитного поля (ММП) в системе GSM** (измеряются в нТл):
  - Амплитуда  $|B|$  (модуль ММП);
  - $B_x, B_y, B_z$  (компоненты ММП);
- **Параметры солнечного ветра:**
  - Скорость протонов (измеряется в км/с)
  - Плотность протонов СВ (измеряется в  $\text{см}^{-3}$ )
  - Температура плазмы СВ (измеряется в Кельвинах)
- **Циклическая компонента** для учета вращения Земли вокруг Солнца (sin/cos)

# Глубины погружения переменных

1) Построенные по автокорреляционной функции:

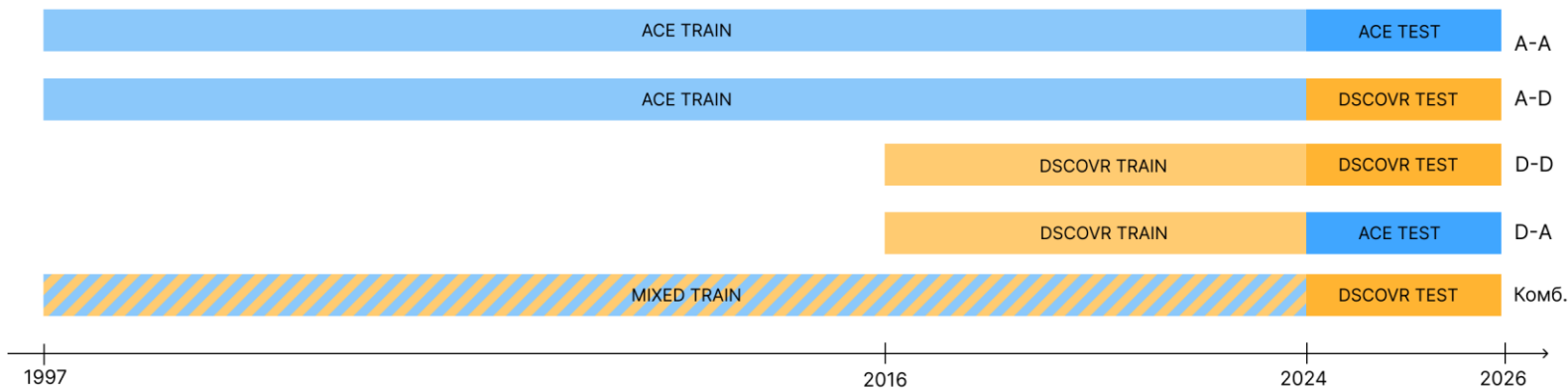
Dst – 43ч, Bx – 26ч, B\_GSM\_y – 12ч, B\_GSM\_z – 3ч, B\_magn – 19ч,  
SW\_spd – 56ч, H\_den -16ч

2) 24ч по всем переменным

| A |    |   |    |      |      |      |      |      |   | A_+0 A_+1 A_+2 A_+3 A_+4 A_+5 |   |      |      |      |      |      |
|---|----|---|----|------|------|------|------|------|---|-------------------------------|---|------|------|------|------|------|
| 0 | 2  | 0 | 2  | NaN  | NaN  | NaN  | NaN  | NaN  |   | 5                             | 3 | -4.0 | 10.0 | 6.0  | -5.0 | 2.0  |
| 1 | -5 | 1 | -5 | 2.0  | NaN  | NaN  | NaN  | NaN  |   | 6                             | 1 | 3.0  | -4.0 | 10.0 | 6.0  | -5.0 |
| 2 | 6  | 2 | 6  | -5.0 | 2.0  | NaN  | NaN  | NaN  | ➡ | 7                             | 0 | 1.0  | 3.0  | -4.0 | 10.0 | 6.0  |
| 3 | 10 | 3 | 10 | 6.0  | -5.0 | 2.0  | NaN  | NaN  |   | 8                             | 5 | 0.0  | 1.0  | 3.0  | -4.0 | 10.0 |
| 4 | -4 | 4 | -4 | 10.0 | 6.0  | -5.0 | 2.0  | NaN  |   | 9                             | 8 | 5.0  | 0.0  | 1.0  | 3.0  | -4.0 |
| 5 | 3  | 5 | 3  | -4.0 | 10.0 | 6.0  | -5.0 | 2.0  |   |                               |   |      |      |      |      |      |
| 6 | 1  | 6 | 1  | 3.0  | -4.0 | 10.0 | 6.0  | -5.0 |   |                               |   |      |      |      |      |      |
| 7 | 0  | 7 | 0  | 1.0  | 3.0  | -4.0 | 10.0 | 6.0  |   |                               |   |      |      |      |      |      |
| 8 | 5  | 8 | 5  | 0.0  | 1.0  | 3.0  | -4.0 | 10.0 |   |                               |   |      |      |      |      |      |
| 9 | 8  | 9 | 8  | 5.0  | 0.0  | 1.0  | 3.0  | -4.0 |   |                               |   |      |      |      |      |      |

# Сценарии экспериментов без адаптации

1. **ACE**→**ACE** (A-A): обучение и тестирование на данных ACE;
2. **ACE**→**DSCOV**R (A-D): обучение на ACE, тестирование на DSCOVR;
3. **DSCOV**R→**DSCOV**R (D-D): обучение и тестирование на данных DSCOVR.
4. **DSCOV**R→**ACE** (D-A): обучение на DSCOVR, тестирование на ACE;
5. **Комбинированные данные**: обучение на объединённом наборе (**ACE** + **DSCOV**R) с дополнительным бинарным признаком **is\_ace**, тестирование на DSCOVR.



# Методология адаптации доменов

Рассматриваются два сценария ДА между данными КА ACE и DSCOVР.

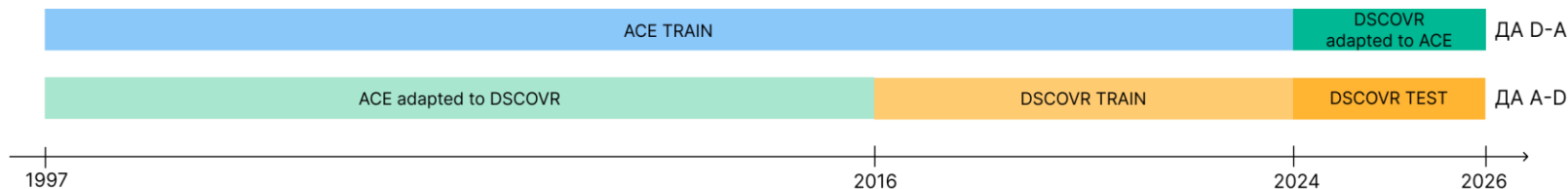
## 1. Сценарий *DSCOVР* → *ACE* (D-A)

Строится модель адаптации  $L : DSCOVР \rightarrow ACE$

Модель адаптации обучается на интервале совместных наблюдений (28.07.2016 – 01.01.2024).

- модель прогнозирования  $M\_A$  обучается на исходных данных ACE;
- **при тестировании** используются **данные DSCOVР, приведённые к домену ACE.**

Новые данные DSCOVР преобразуются к статистике ACE, что позволяет использовать уже обученную модель без потери качества.



# Методология адаптации доменов

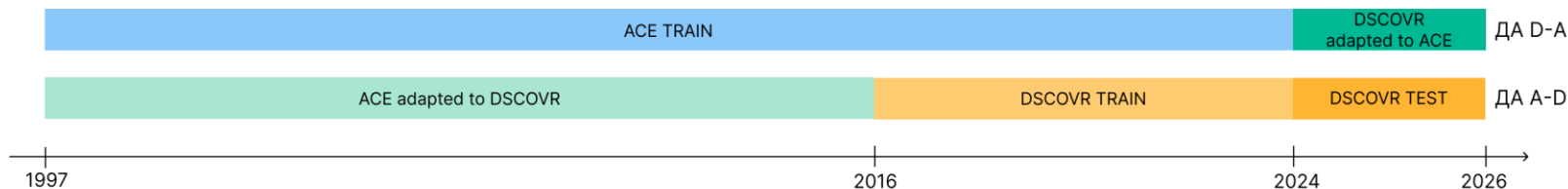
Рассматриваются два сценария ДА между данными КА ACE и DSCOVER.

## 2. Сценарий *ACE* → *DSCOVER* (A-D)

Строится модель адаптации K: *ACE* → *DSCOVER*

- обучающая выборка модели прогнозирования  $M_D$  расширяется за счёт адаптированных данных **ACE** (1997–2016) и исходных данных **DSCOVER**;
- тестирование проводится на исходных данных DSCOVER.

При недоступности ACE во время тестирования его исторические данные переносятся в домен DSCOVER, что увеличивает объём обучающей выборки и улучшает качество прогноза.



# Модели адаптации

## 1) Линейная регрессия (ЛР)

## 2) Градиентный бустинг (ГБ - LGBM)

- количество деревьев – 1000
- число листьев – 31
- максимальная глубина – 3
- темп обучения – 0.1
- L1-регуляризация с коэффициентом  $\alpha=0.7$
- L2-регуляризация с коэффициентом  $\lambda=0.7$

## 3) Многослойный персептрон (МСП)

- 2 скрытых слоя (8 и 4 нейрона)
- L2-регуляризация с коэффициентом  $\alpha \times 10^{-4}$
- начальная скорость обучения –  $5 \times 10^{-4}$
- максимальное число итераций – 1500 с ранней остановкой

## 4) Метод проекций на латентные структуры (PLS)

- число компонент – 120

# Модели адаптации

## 5) Доменно-состязательная нейронная сеть (DANN)

### а) Адаптация $DSCOV\!R \rightarrow ACE$

- размер скрытого слоя – 2048
- скорость обучения –  $2 \times 10^{-4}$
- коэффициент адаптации – 6
- число эпох обучения – 40

### б) Адаптация $ACE \rightarrow DSCOV\!R$

- размер скрытого слоя – 256
- скорость обучения –  $4 \times 10^{-4}$
- коэффициент адаптации – 2
- число эпох обучения – 30

**6) ЛР+ГБ, ЛР+МСП** – комбинированное использование методов (ЛР для полей, др. модель для ветра) с аналогичными параметрами соответствующих моделей

**7) Фильтр Калмана** – итеративный процесс, где скрытое состояние – это сигнал целевого домена (например, ACE), а наблюдения – данные исходного домена (DSCOV\!R). На этапе предсказания фильтр экстраполирует, каким должен быть сигнал ACE в следующий момент, а на этапе обновления корректирует эту оценку, получив новые данные DSCOV\!R через матрицу связи  $H$ , обученную на периоде совместной работы аппаратов.

# Модели прогнозирования

## 1) Многослойный персептрон (МСП)

- один скрытый слой с 12 нейронами
- L2-регуляризация с коэффициентом  $\alpha=10^{-4}$
- начальная скорость обучения –  $10^{-3}$
- максимальное число итераций – 1500 с ранней остановкой

## 2) Линейная регрессия (ЛР)

### 3) ЛР с Lasso-регуляризацией (ЛР L1)

- регуляризация с коэффициентом  $\alpha=5 \times 10^{-4}$

### 4) ЛР с Ridge-регуляризацией (ЛР L2)

- регуляризация с коэффициентом  $\alpha=1.5$

## 5) Градиентный бустинг (ГБ – LGBM)

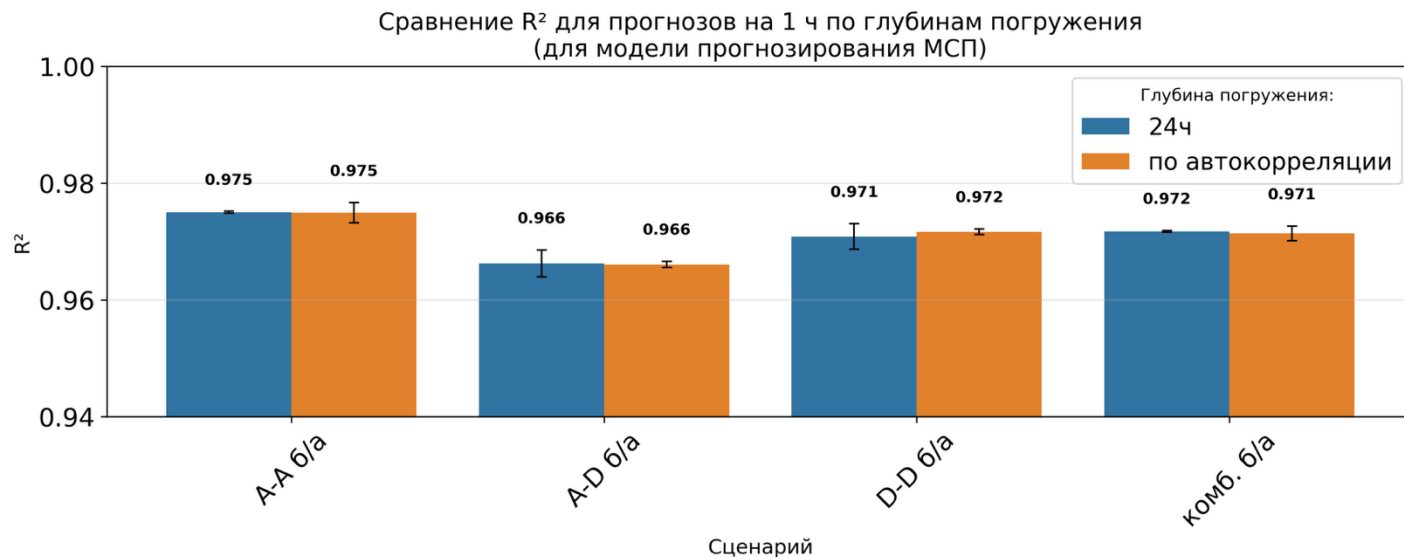
- количество деревьев – 1160
- число листьев – 93
- максимальная глубина – 5
- темп обучения – 0.0144
- L1- и L2-регуляризации с коэффициентами  $\alpha=0.023$  и  $\lambda=0.006$

# Выбор глубины погружения временных рядов

Использование **фиксированной глубины (24 ч)**

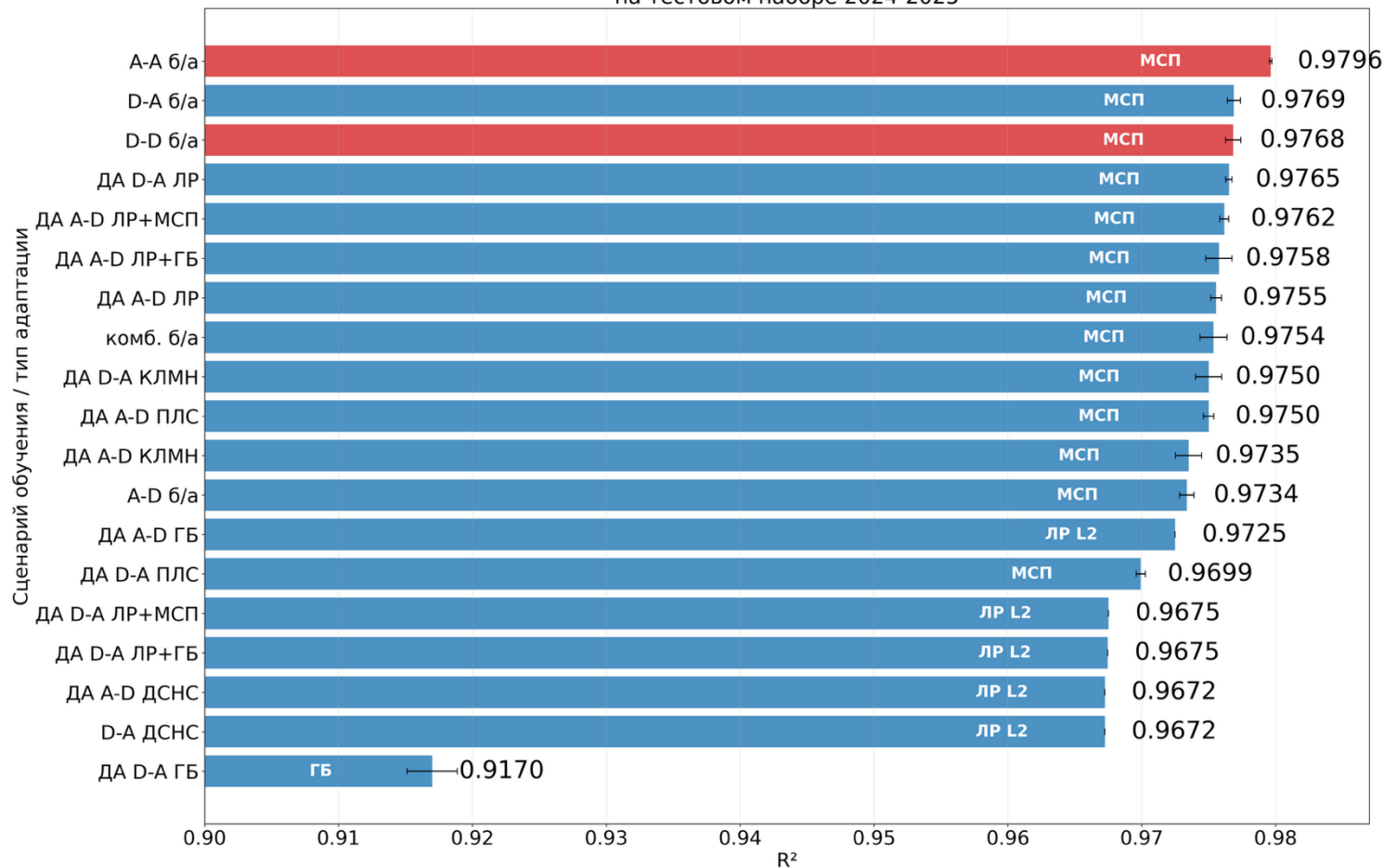
и **глубины, определённой по автокорреляционной функции**, приводит к **сопоставимому качеству** прогноза.

В дальнейшем анализе используются результаты для глубины, определённой по автокорреляционной функции, как более обоснованной с физической точки зрения.



# Выбор оптимальных сценариев адаптации

$R^2$  моделей для прогнозов на 1 ч при погружении по автокорреляционной ф-ции  
на тестовом наборе 2024-2025

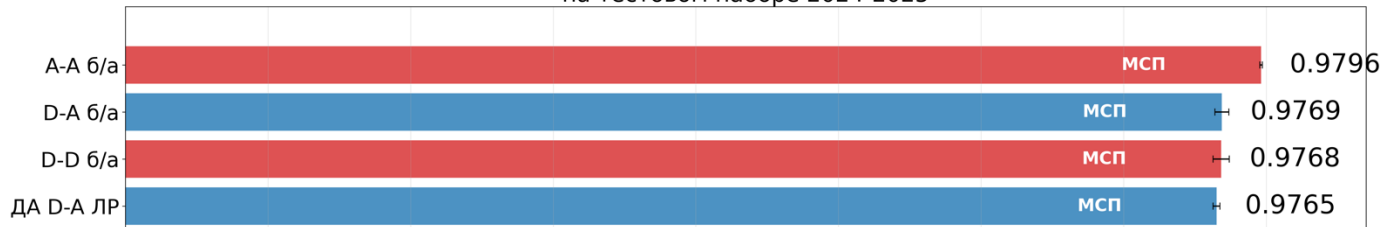


Модель внутри столбца гистограммы – наилучшая прогностическая модель для данного сценария

Рассматривается тестовый набор 2024-2025

# Выбор оптимальных сценариев адаптации

$R^2$  моделей для прогнозов на 1 ч при погружении по автокорреляционной ф-ции  
на тестовом наборе 2024-2025



Модель внутри  
столбца  
гистограммы –  
наилучшая  
прогностическая  
модель для  
данного сценария

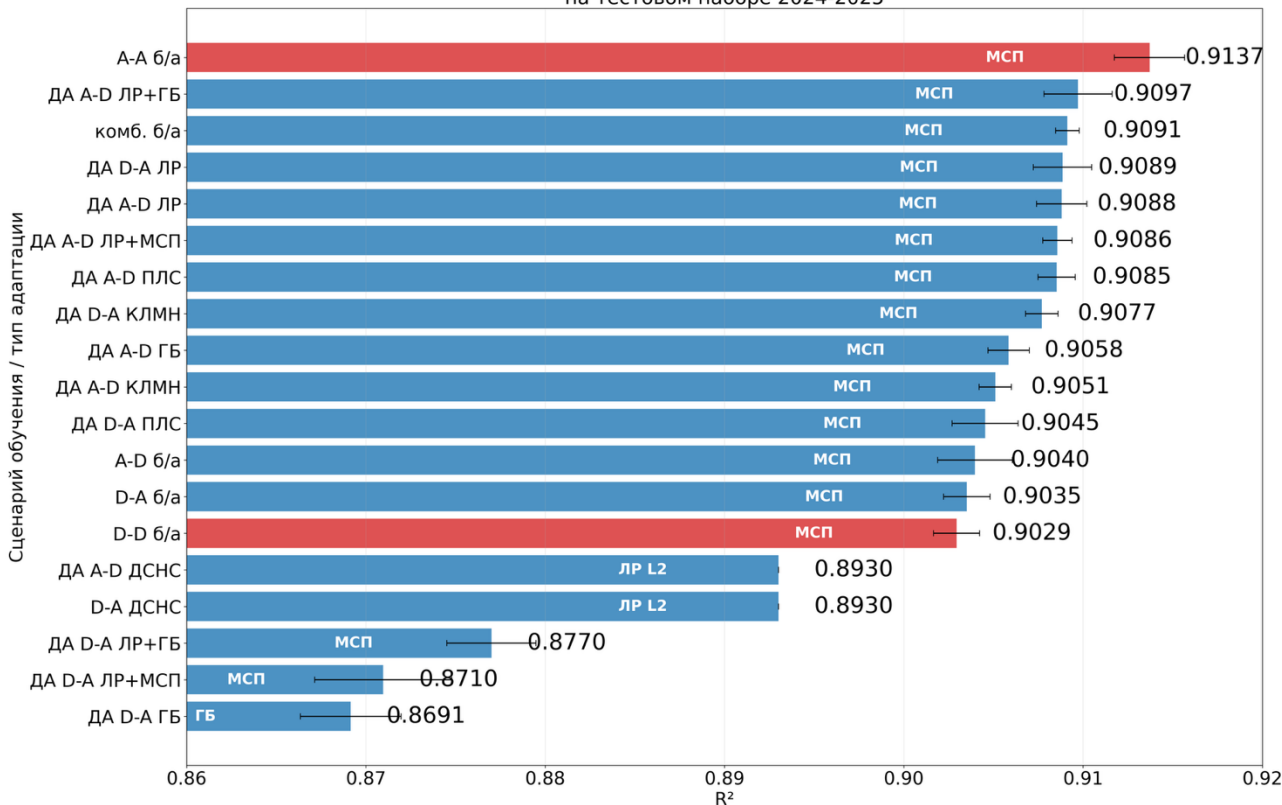
Рассматривается  
тестовый набор  
2024-2025

**Наилучшее качество краткосрочного прогноза**  
достигается в сценарии **ACE-ACE без адаптации**,  
т.е. при обучении и применении модели на данных  
одного и того же аппарата без необходимости перехода  
на другой источник данных.

**Для прогноза на 1 час** применение доменной  
адаптации **не даёт существенного выигрыша.**

# Выбор оптимальных сценариев адаптации

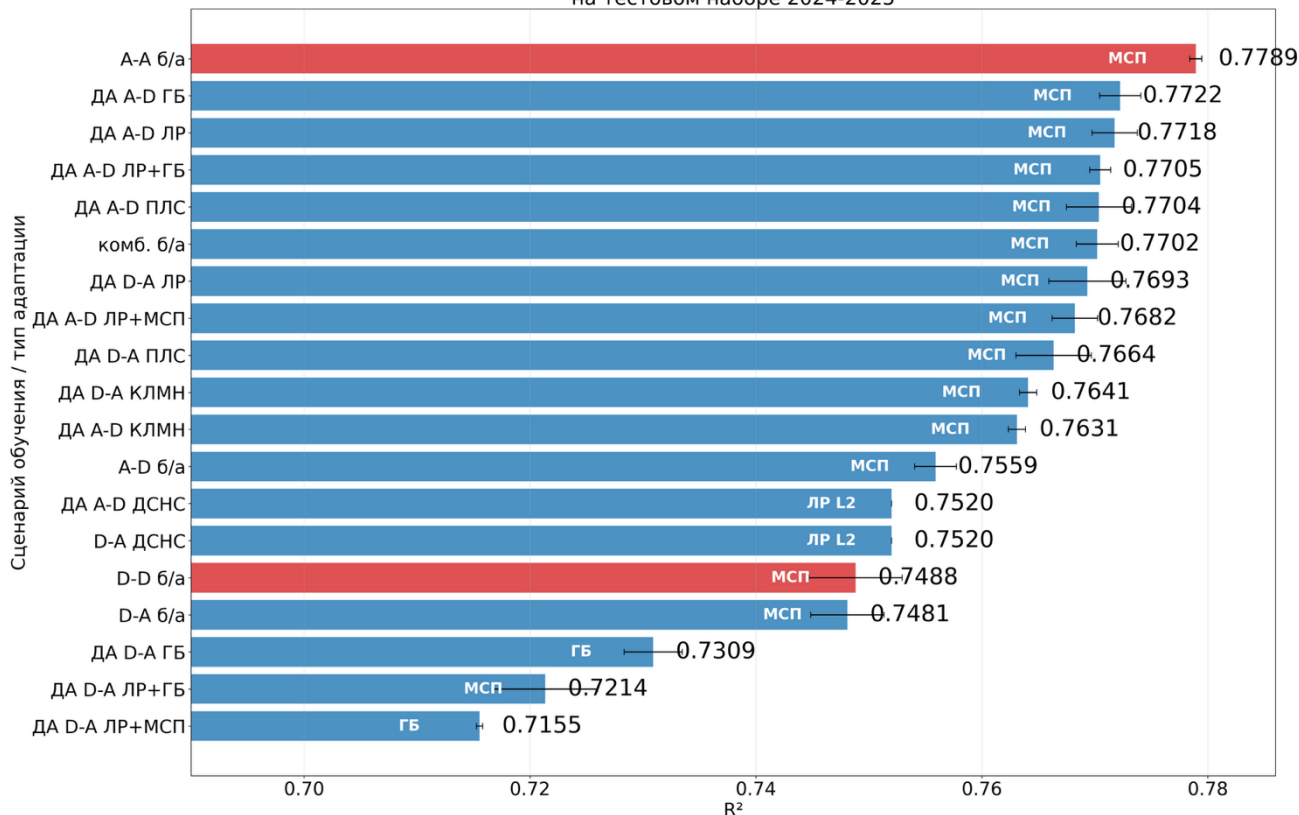
$R^2$  моделей для прогнозов на 3 ч при погружении по автокорреляционной ф-ции  
на тестовом наборе 2024-2025



На горизонте предсказания  
**3 часа** сценарии  
**комбинированной**  
**адаптации (ЛР для полей +**  
**ГБ для параметров ветра) и**  
**обучением**  
**на комбинированных**  
**данных** показывают  
результаты, близкие к A-A

# Выбор оптимальных сценариев адаптации

$R^2$  моделей для прогнозов на 6 ч при погружении по автокорреляционной ф-ции  
на тестовом наборе 2024-2025

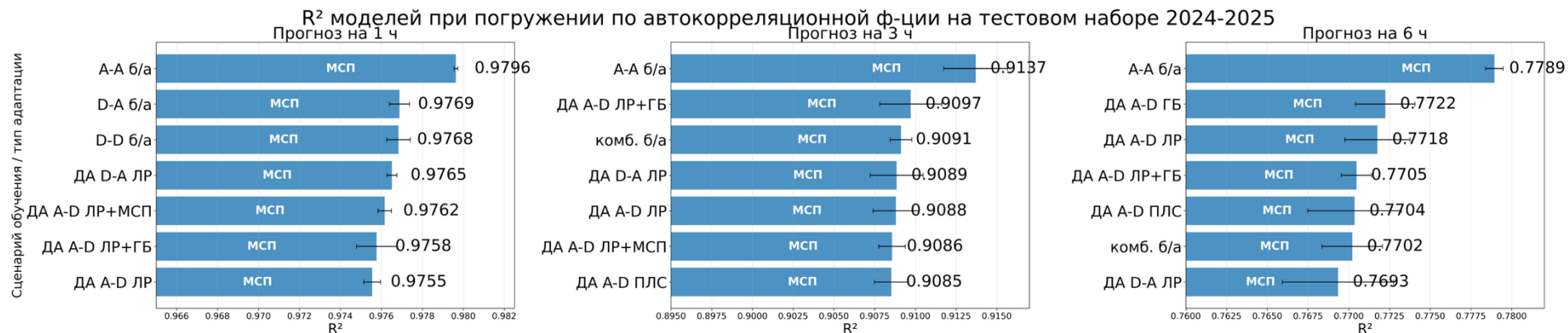


На горизонте **6 часов** результаты, близкие к сценарию A-A, достигаются для сценариев адаптации **ACE-DSCOV** с помощью моделей ЛР, ГБ и их комбинаций, а также при обучении на комбинированных данных

# Моделирование совместной работы КА

**Идея эксперимента:** смоделировать реальную эксплуатацию системы прогнозирования Dst на протяжении всего периода наблюдений (1997–2026).

- 1) Были выбраны 7 моделей и сценариев адаптации, показавших себя лучше всего на тестовом периоде 2024-2025;
- 2) Прогностическая модель – многослойный персептрон (МСП).
- 3) Рассматриваются 5 временных диапазонов, отражающих последовательные этапы работы системы.

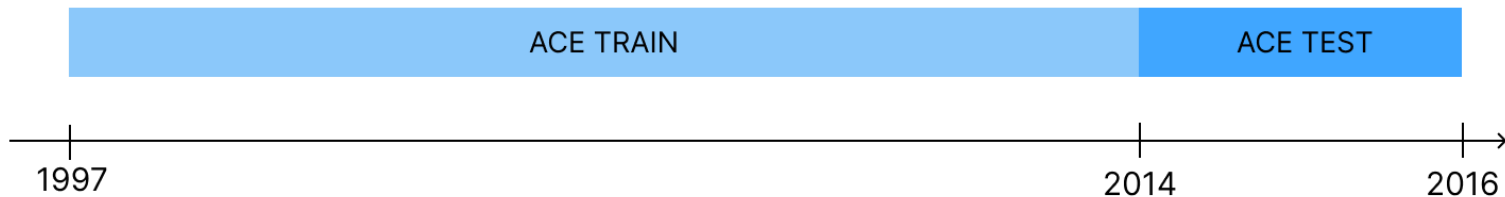


# Моделирование совместной работы КА

**Этап 1:** временной диапазон 23.10.1997-29.07.2016

Зафиксируем момент времени, когда КА DSCOVР еще не был введен в эксплуатацию. В данном случае рассматриваем прогнозирование индекса Dst только на данных КА ACE.

- Тренировочный набор – 1997-2014
- Тестовый набор – 2014-2016 (2 года)
- Адаптация не применяется

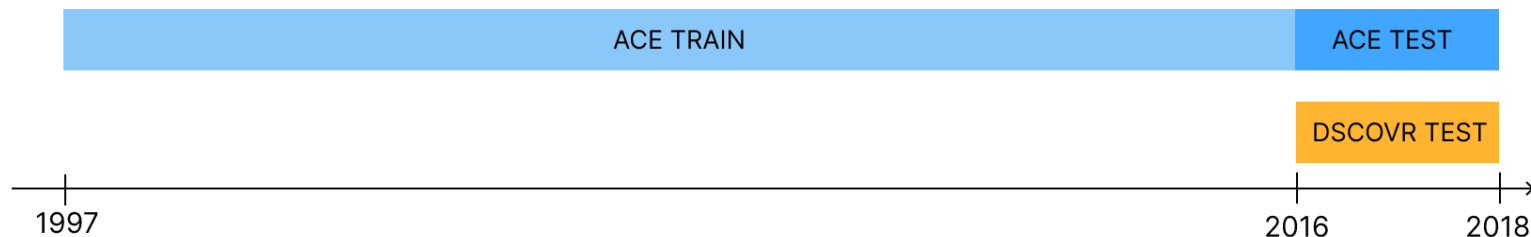


# Моделирование совместной работы КА

**Этап 2:** временной диапазон 23.10.1997-01.01.2018

Также рассматриваем обучение и тест прогностической модели на данных КА ACE, дополнительно используем первые накопившиеся данные КА DSCOVР для прямого переноса обучения (б/а А-D).

- Тренировочный набор – 1997-2016
- Тестовый набор – 2016-2017 (1 год 5 месяцев)
- Адаптация не применяется



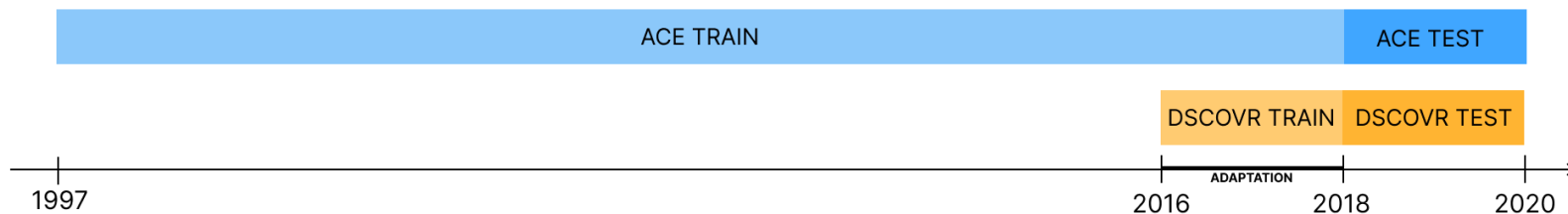
# Моделирование совместной работы КА

**Этап 3:** временной диапазон 23.10.1997-01.01.2020

Набор данных КА DSCOVР увеличивается на 2 года.

С этого момента начинаем обучать модели адаптации на диапазоне пересечения данных обоих КА перед тестовым набором.

- Тренировочный набор – 1997-2017
- Тестовый набор – 2018-2019 (2 года)
- Модели адаптации обучаются на диапазоне 2016-2017

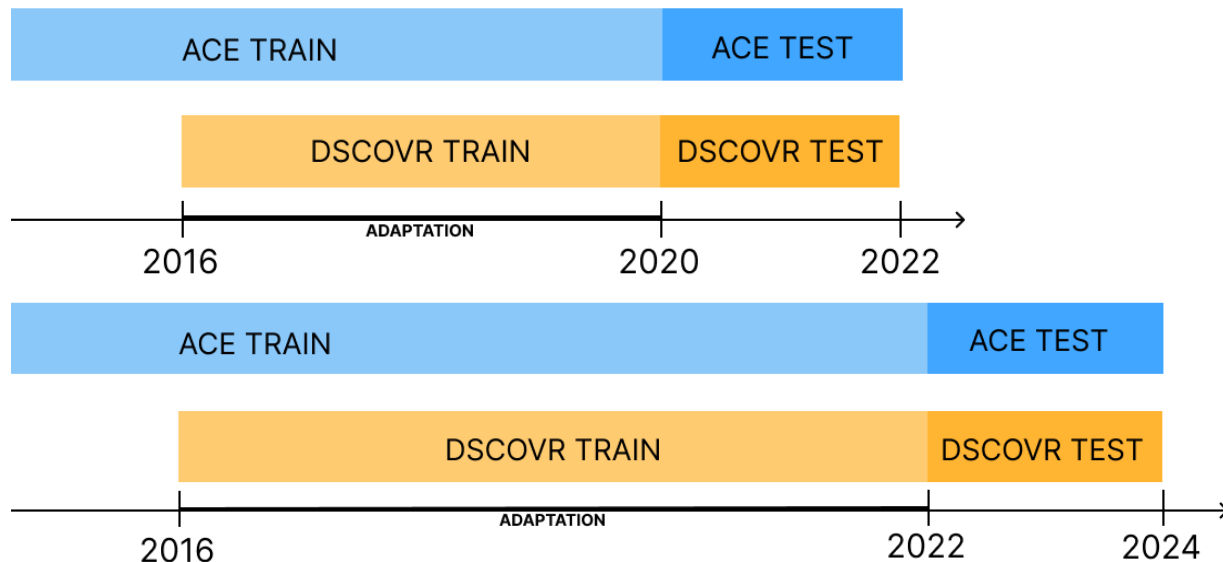


# Моделирование совместной работы КА

**Этап 4:** временной диапазон 23.10.1997-01.01.2022 (тест на – 2020-2021)

**Этап 5:** временной диапазон 23.10.1997-01.01.2024 (тест на – 2022-2023)

Аналогично этапу 3 последовательно расширяем наборы данных на 2 года, тем самым увеличивая наборы для обучения прогностической и адаптационных моделей.



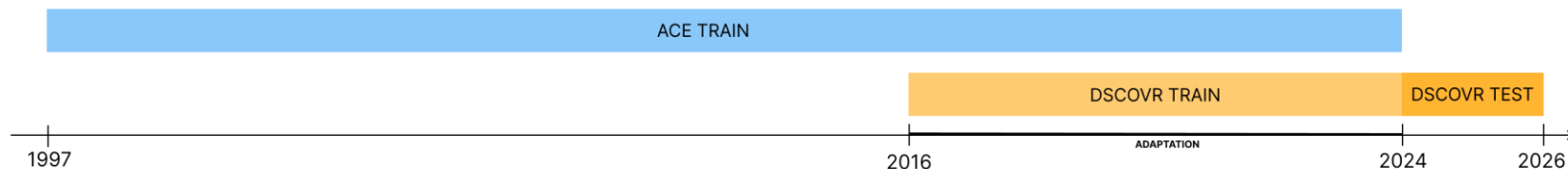
# Моделирование совместной работы КА

**Этап 6:** временной диапазон 23.10.1997-01.01.2026

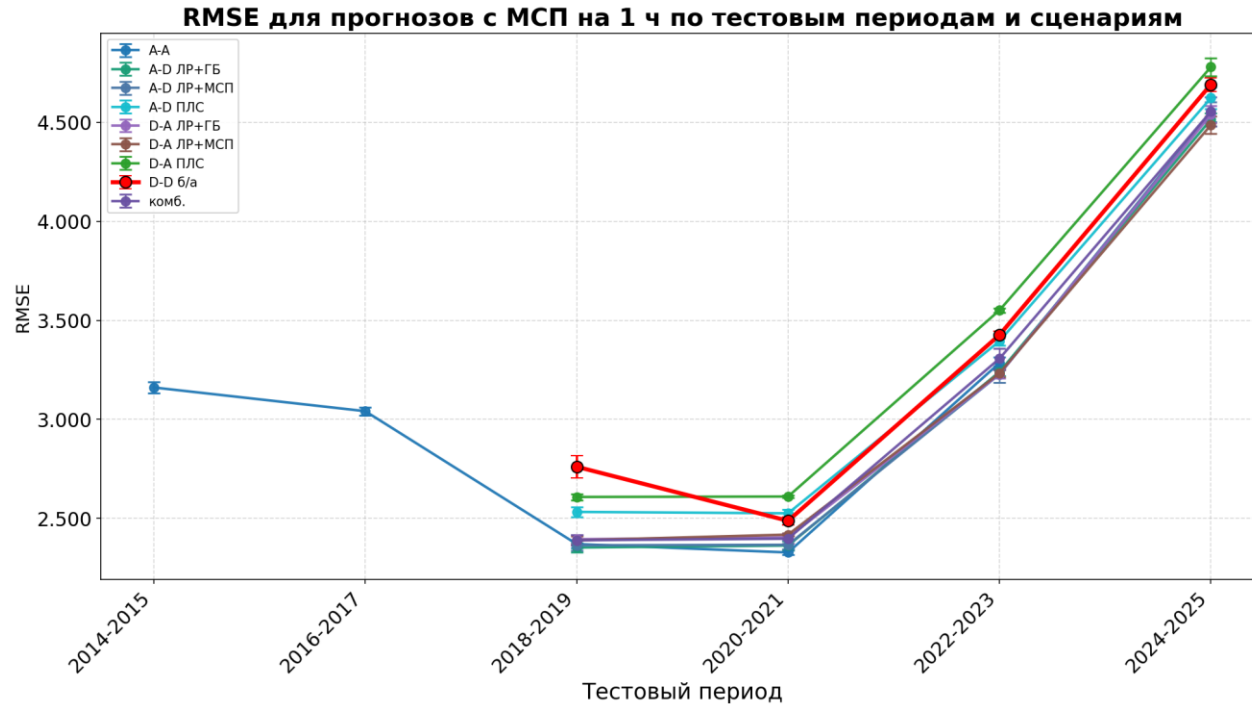
Рассматриваем случай, в котором КА ACE был отключен 01.01.2024 и теперь поступают только данные КА DSCOVER.

Сравниваем только различные сценарии адаптации и прогноз только на данных КА DSCOVER.

- Тестовый набор – 2024-2025 (2 года)
- Модели адаптации обучаются на диапазоне пересечения данных обоих КА 2016-2023

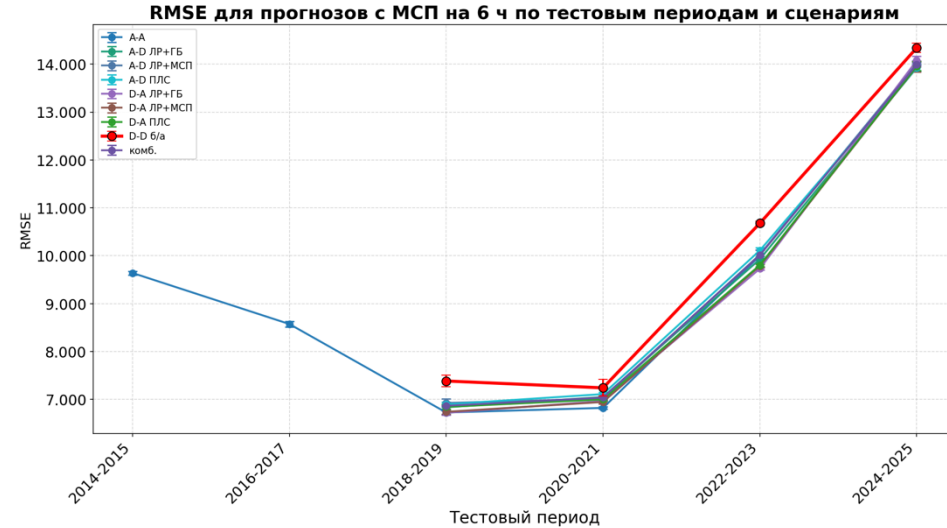
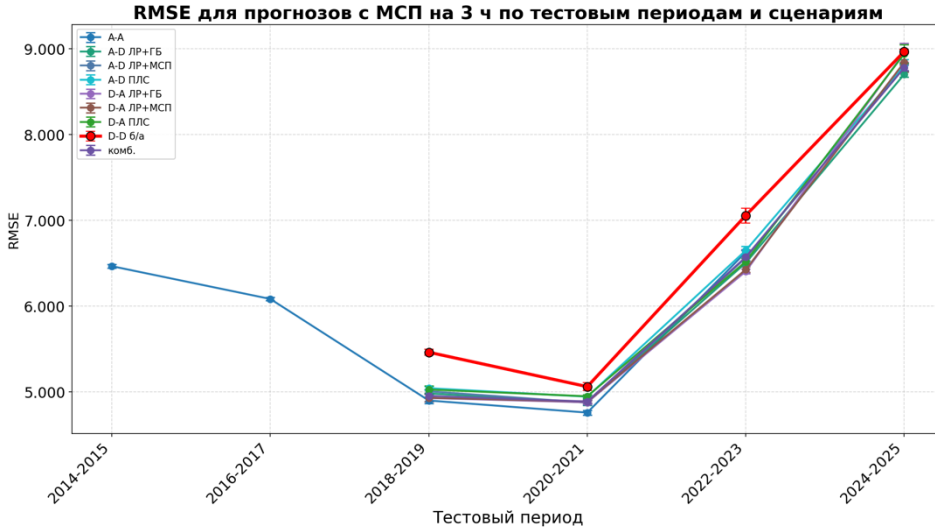


# Результаты



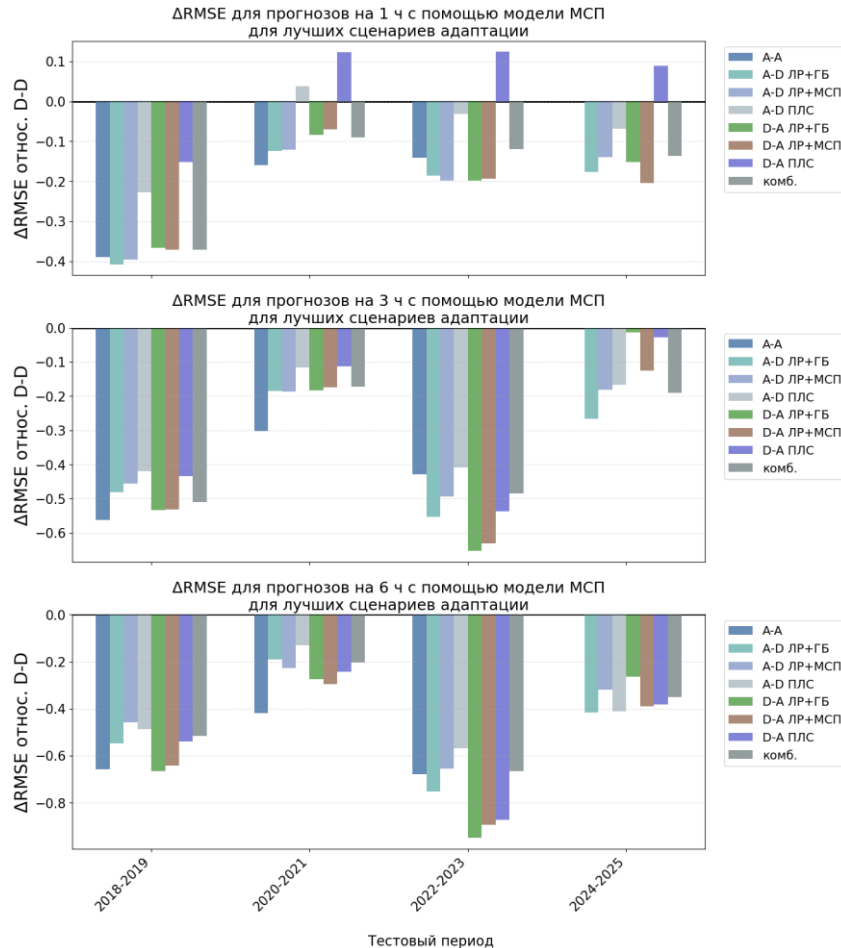
- Полный переход на данные КА DSCOVN ухудшает качество прогнозирования;
- Некоторые сценарии ДА позволяют приблизиться к значению СКО при прогнозировании на данных КА ACE.

# Результаты



На горизонтах прогнозирования 3 и 6 часов наблюдается аналогичная ситуация.

# Результаты



На графике показана разница СКО между каждым сценарием адаптации и базовым сценарием D-D без адаптации.

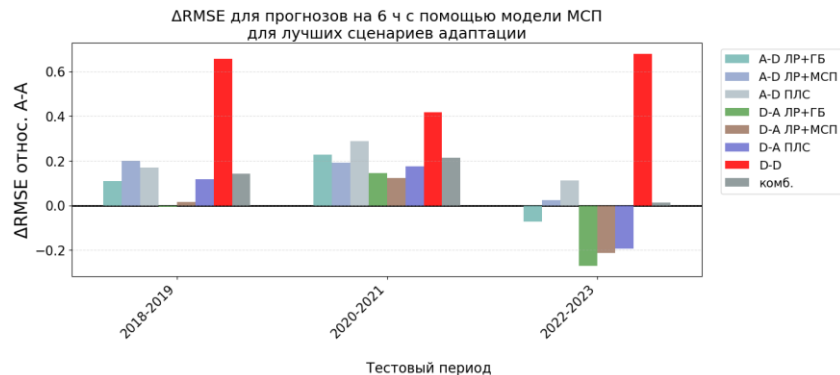
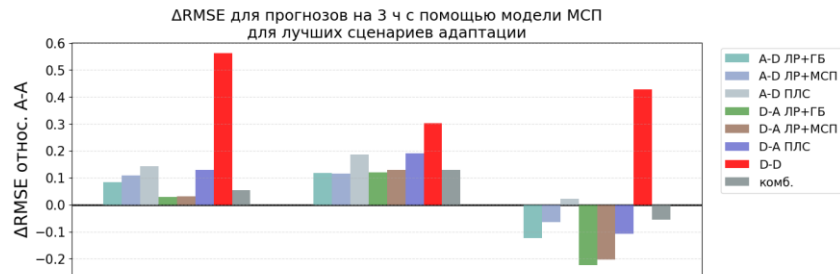
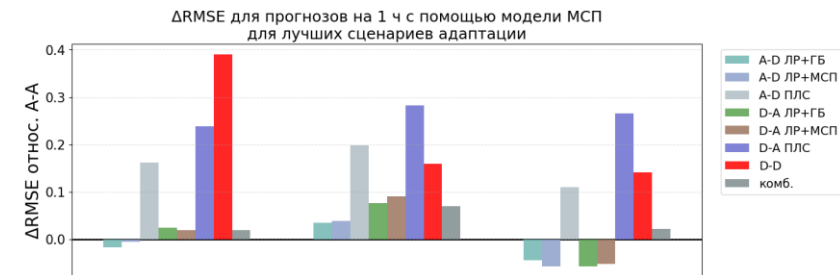
$$\Delta RMSE_f = RMSE_i - RMSE_{D-D} \text{ б/а}$$

Если  $\Delta RMSE < 0$ , значит сценарий работает лучше базового, а если  $\Delta RMSE > 0$ , то хуже.

# Результаты

Аналогично сравниваются СКО со сценарием ACE-ACE без адаптации.

$$\Delta RMSE_f = RMSE_i - RMSE_{A-A} \text{ б/а}$$

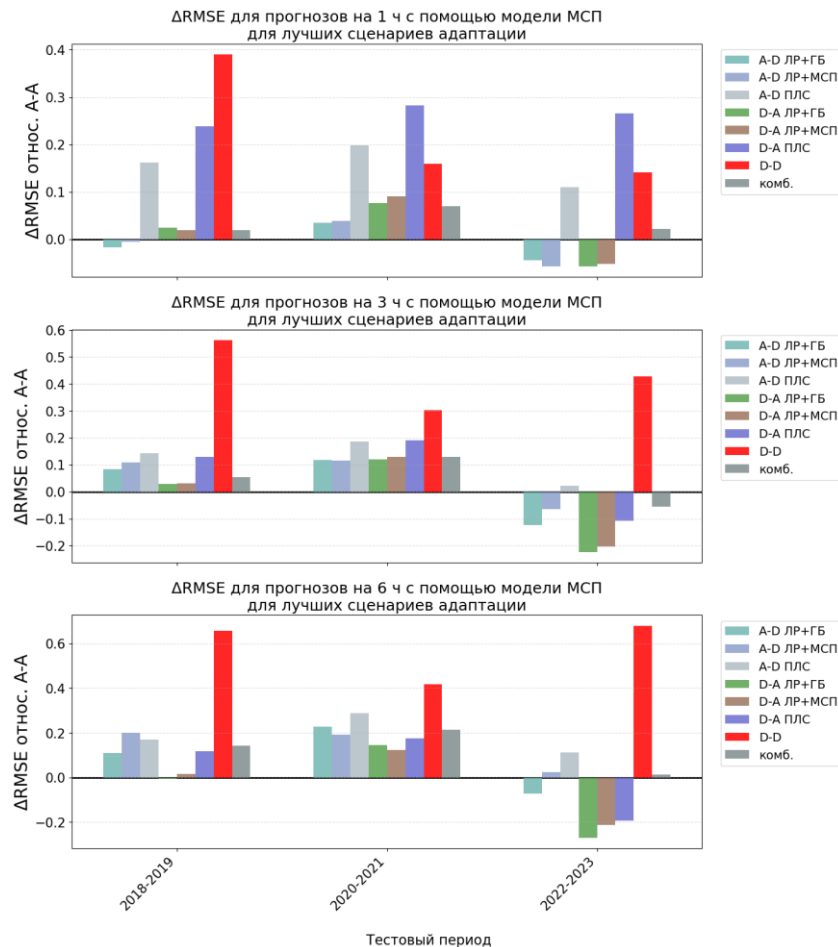


# Результаты

Аналогично сравниваются СКО со сценарием ACE-ACE без адаптации.

$$\Delta RMSE_f = RMSE_i - RMSE_{A-A \text{ б/а}}$$

Доменная адаптация помогает приблизить результаты модели к сценарию ACE-ACE б/а.



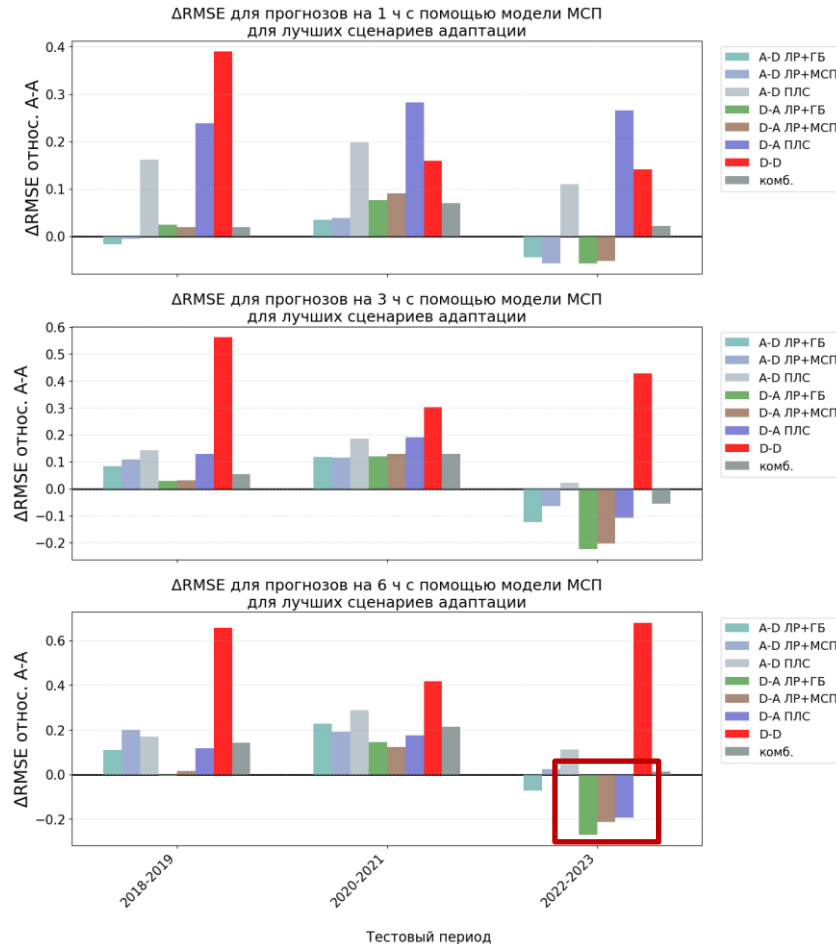
# Результаты

Аналогично сравниваются СКО со сценарием ACE-ACE без адаптации.

$$\Delta RMSE_f = RMSE_i - RMSE_{A-A \text{ б/а}}$$

Доменная адаптация помогает приблизить результаты модели к сценарию ACE-ACE б/а.

На горизонте в 6 часов модели адаптации D-A на моделях ЛР+МСП, ЛР+ГБ и ПЛС дают выигрыш по сравнению с базовым сценарием прогнозирования.



# Выводы и итоги

- **Различия между доменами** данных KA ACE и DSCOVР приводят к **снижению качества** прогнозирования **при прямом переносе обучения** между доменами.
- Модели прогнозирования, обученные исключительно на данных DSCOVР, уступают по точности моделям на ACE.
- **Методы доменной адаптации** позволяют частично компенсировать эти различия и обеспечить **сопоставимое качество прогноза** при использовании данных разных KA.

# Выводы и итоги

- **На коротких горизонтах (1 час) доменная адаптация не даёт существенного выигрыша**  
по сравнению с базовыми сценариями без адаптации.
- **С увеличением горизонта прогнозирования вклад доменной адаптации становится более заметным,**  
при этом наилучшие результаты достигаются при использовании:
  - **комбинированной адаптации (ЛР+ГБ/МСП)**  
в сценарии адаптации **DSCOV**R → **ACE**
  - **метода проекции на латентные структуры**  
в сценарии адаптации **DSCOV**R → **ACE**
  - **комбинированных данных** обоих КА.

# Выводы и итоги

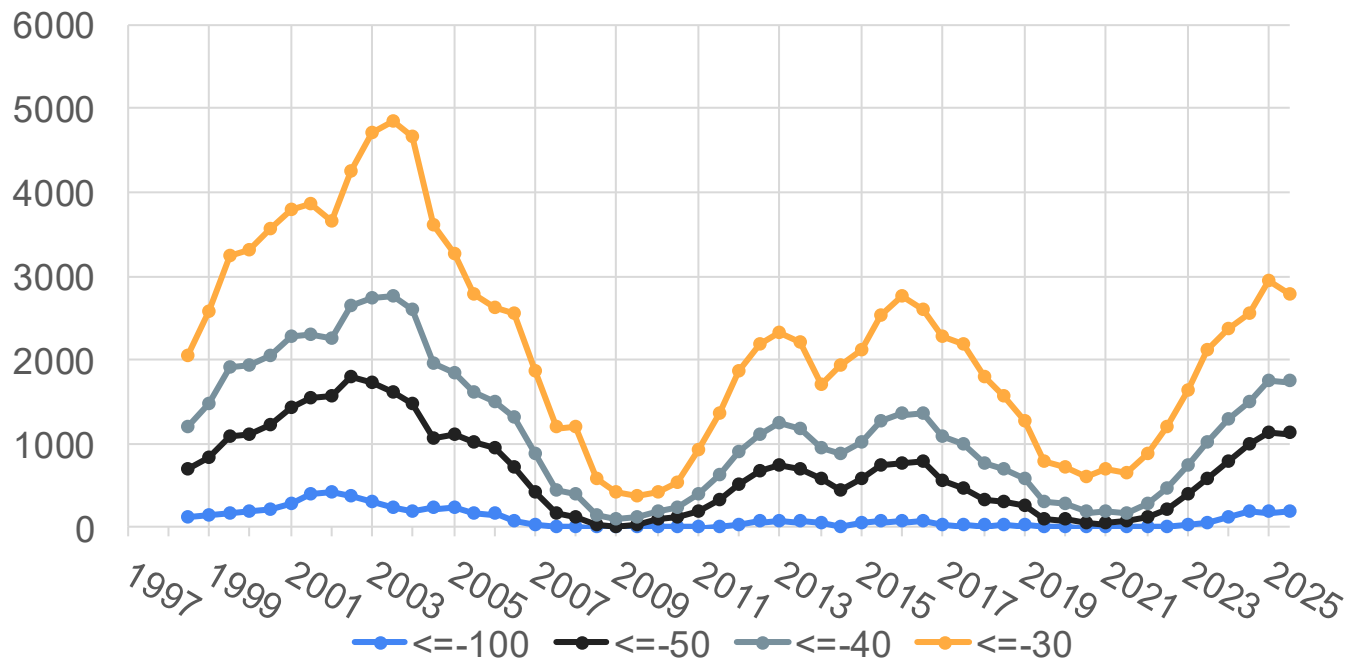
- Предложенный подход позволяет  
**сохранить качество** прогнозирования  
**после вывода из эксплуатации КА АСЕ**  
за счёт использования методов **доменной адаптации**.

Спасибо за внимание!

# Влияние солнечных циклов на качества прогноза Dst

- Распределение Dst заметно меняется от периода к периоду из-за **11-летних циклов солнечной активности**, которые задают разную частоту и силу магнитных бурь.
- При прогнозировании метрики  $R^2$  и RMSE между разными тестовыми периодами меняются **не только из-за качества моделей**, но и из-за уровней солнечной активности и связанного с ними распределения Dst в каждом периоде.

Число часов по полугодиям с указанным и ниже значением Dst - сумма за предшествующие 2 года



# Линейная регрессия

Предполагаем, что будущее значение ряда  $y_t$  можно выразить как линейную комбинацию прошлых значений и входных параметров:

$$y_t = w_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{j=1}^q b_j z_{j,t},$$

где  $y_{t-i}$  — прошлые значения прогнозируемой величины,  
 $z_{j,t}$  — входные параметры (параметры ММП и СВ) в момент  $t$ ,  
 $a_i, b_j$  — веса (коэффициенты) линейной регрессии.

## Обучение модели

Коэффициенты  $w_i$  подбираются так, чтобы минимизировать среднеквадратичную ошибку на обучающей выборке:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

# Градиентный бустинг (LightGBM)

Градиентный бустинг — это ансамблевый метод, в котором итоговый прогноз строится как сумма многих простых моделей, обычно деревьев решений:

$$F_M(x) = F_0(x) + \sum_{m=1}^M \nu f_m(x),$$

где  $F_0(x)$  — начальное приближение,  $f_m(x)$  — очередное дерево,  $\nu$  — шаг обучения.

Каждое новое дерево обучается так, чтобы уменьшить ошибку предыдущего ансамбля.

Формально на каждом шаге минимизируется функция потерь  $L(y, F(x))$ :

$$F_m = \arg \min_F \sum_{i=1}^N L(y_i, F_{m-1}(x_i) + F(x_i))$$

Идея состоит в том, что модель движется в направлении антиградиента ошибки.

**LightGBM** — это быстрая реализация градиентного бустинга, рассчитанная на большие наборы данных. Для прогнозирования временных рядов LightGBM особенно удобен, когда признаки уже представлены в табличном виде, например как погружение ряда, параметры ММП и т.д.

# Метод проекции на латентные структуры

Метод проекций на латентные структуры (ПЛС) используется, когда входных признаков много и они сильно коррелированы между собой. Вместо того чтобы напрямую строить модель по всем исходным столбцам  $X$ , в ПЛС сначала **строятся новые признаки (латентные компоненты)** как линейные комбинации имеющихся признаков:

$$t_k = Xw_k, \quad k = 1, \dots, n_C,$$

где каждый  $t_k$  — новый признак (компонента), а  $w_k$  — вектор весов по исходным признакам.

Число таких компонент заранее задаётся параметром  $n_C$ . Затем целевая переменная моделируется уже по этим компактным и информативным компонентам:

$$y \approx c_0 + \sum_{k=1}^{n_C} c_k t_k.$$

Так мы одновременно уменьшаем размерность, боремся с мультиколлинеарностью и сохраняем именно ту часть вариации признаков, которая наиболее связана с таргетом.

# Многослойный персептрон

Многослойный персептрон – это нейронная сеть, в которой сигнал проходит от входного слоя через один или несколько скрытых слоёв к выходному слою. Каждый нейрон скрытого слоя вычисляет взвешенную сумму входов и применяет нелинейную функцию активации:

$$z = \sum_{i=1}^n w_i x_i + b, a = \varphi(z).$$

Для обучения сети задаётся функция потерь, которая показывает ошибку прогноза. В этой работе используется среднеквадратичная ошибка:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i^{source} - \hat{y}_i)^2.$$

Весовые коэффициенты подбираются методом стохастического градиентного спуска, который обновляет веса по мини-батчам данных. Чтобы снизить риск переобучения, применяется ранняя остановка: обучение прекращается, если ошибка на валидационной выборке не улучшается в течение заданного числа эпох.

# Комбинированные модели адаптации

В работе используется комбинированная схема адаптации: для одних групп признаков применяется **линейная регрессия**, а для других – более гибкая модель, например **градиентный бустинг** или **многослойный персептрон**.

Такой подход позволяет разделить задачу на две части:

$$\hat{y} = f_{\text{лин}}(X_{\text{ММП}}) + f_{\text{нелин}}(X_{\text{ветер}}),$$

где  $X_{\text{ММП}}$  – признаки межпланетного магнитного поля, а  $X_{\text{ветер}}$  – параметры солнечного ветра.

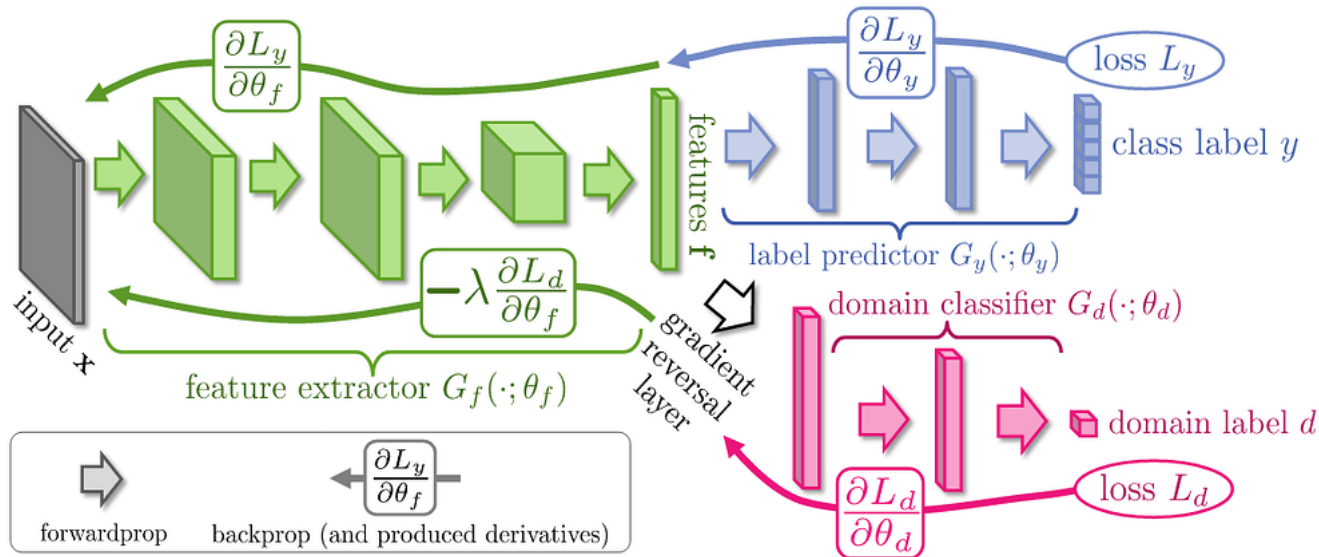
Линейная регрессия хорошо описывает более **плавные и устойчивые** зависимости, а ГБ или МСП лучше улавливают **нелинейные эффекты** и сложные зависимости между параметрами.

# Доменно-состязательная нейронная сеть

Доменно-состязательная сеть — это нейросеть для переноса обучения модели между разными наборами данных.

Сеть состоит из трёх частей: **экстрактора** признаков, **регрессора** и блока, который определяет принадлежность данных к домену — **дискриминатора**. Во время обучения модель делает признаки менее различимыми между доменами.

## Unsupervised Domain Adaptation by Backpropagation



# Фильтр Калмана для доменной адаптации

Цикл фильтра Калмана (два шага на каждый момент времени  $t$ ):

1. Предсказание – куда должно прийти состояние без новых данных:

$$\hat{s}_{t|t-1} = F \hat{s}_{t-1},$$

$$P_{t|t-1} = F P_{t-1} F^\top + Q$$

2. Обновление – коррекция по новому наблюдению  $z_t$ :

$$y_t = z_t - H \hat{s}_{t|t-1},$$

$$K_t = P_{t|t-1} H^\top (H P_{t|t-1} H^\top + R)^{-1},$$

$$\hat{s}_t = \hat{s}_{t|t-1} + K_t y_t,$$

$$P_t = (I - K_t H) P_{t|t-1}$$

Преимущество перед линейной регрессией: коэффициент  $K_t$  автоматически адаптируется – при высоком шуме (большое  $R$ ) фильтр доверяет модели, при стабильном сигнале – наблюдениям.

# Точки Лагранжа

Точки в пространстве, в которых тело малой массы может находиться в относительном равновесии по отношению к двум другим небесным телам.

Сигнал от L1 до Земли идёт ~5 с, а солнечный ветер достигает её за ~40 минут.

