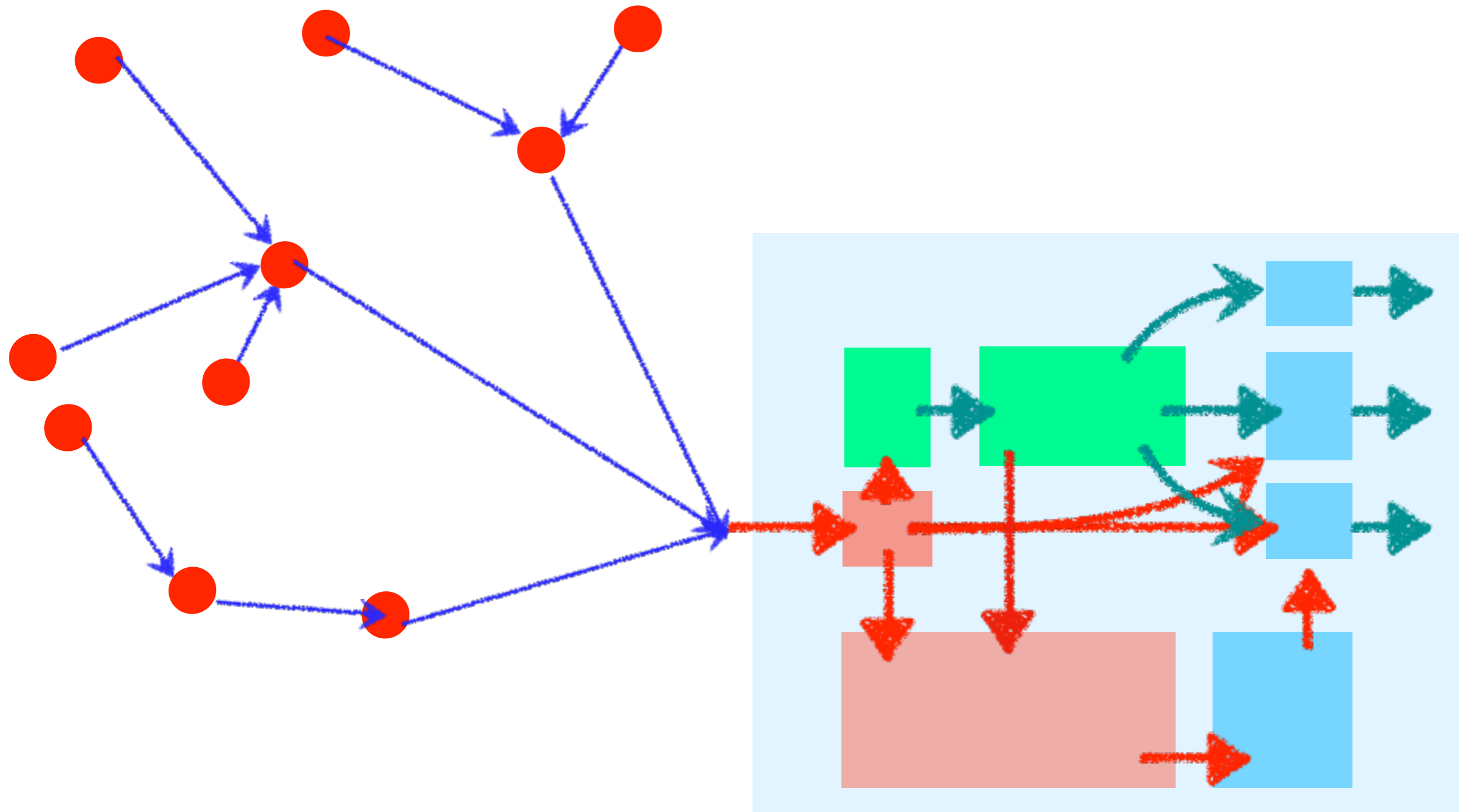


**Установочный семинар для
специалистов ИТ и синоптиков
по вопросам использования
COSMO-ЦА**

Ташкент 4-9 июня 2018 г.
Группа Всемирного банка

Широкое видение: от автоматической погодной сети к пост процессингу (последующей обработке) COSMO

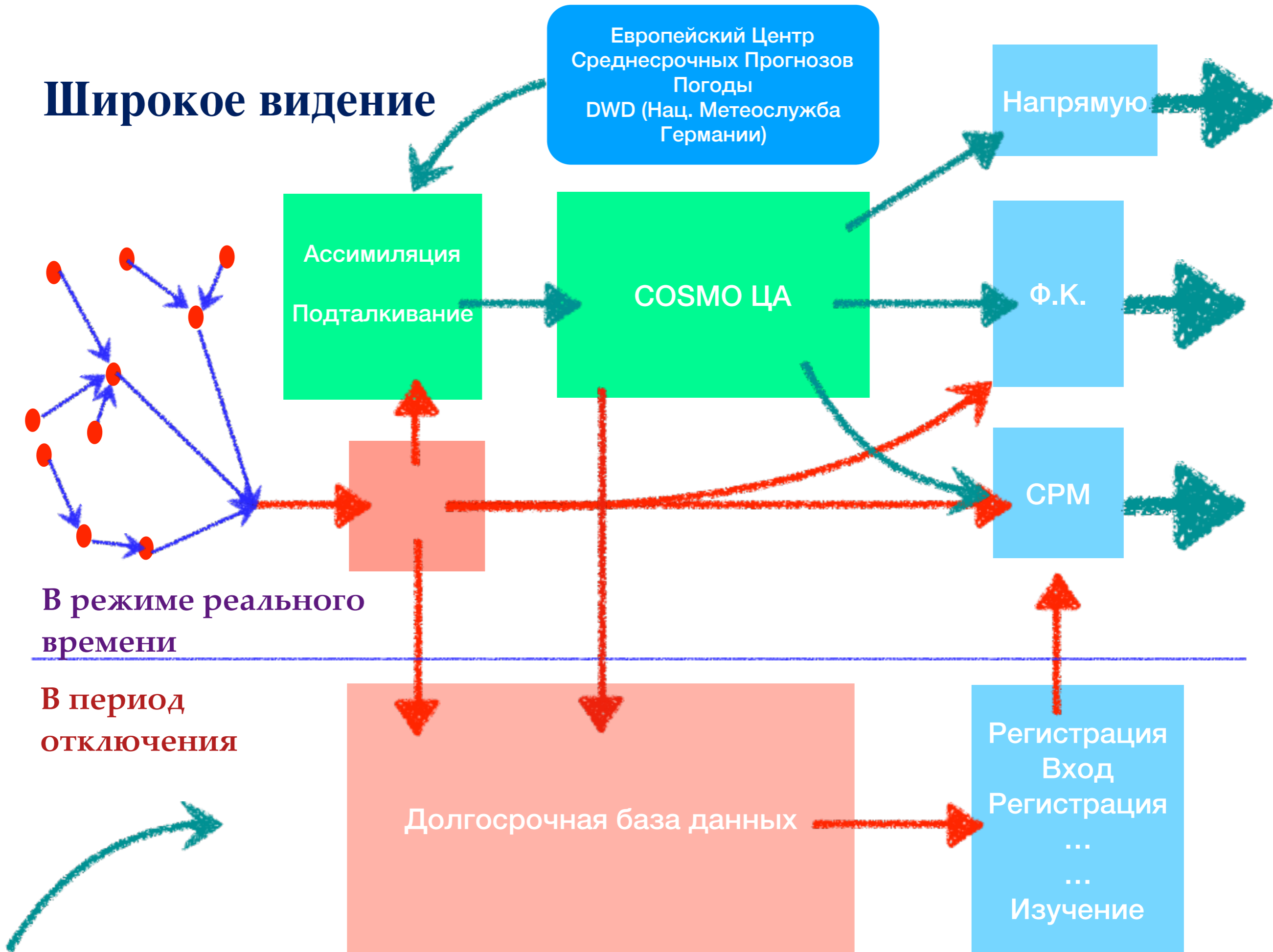


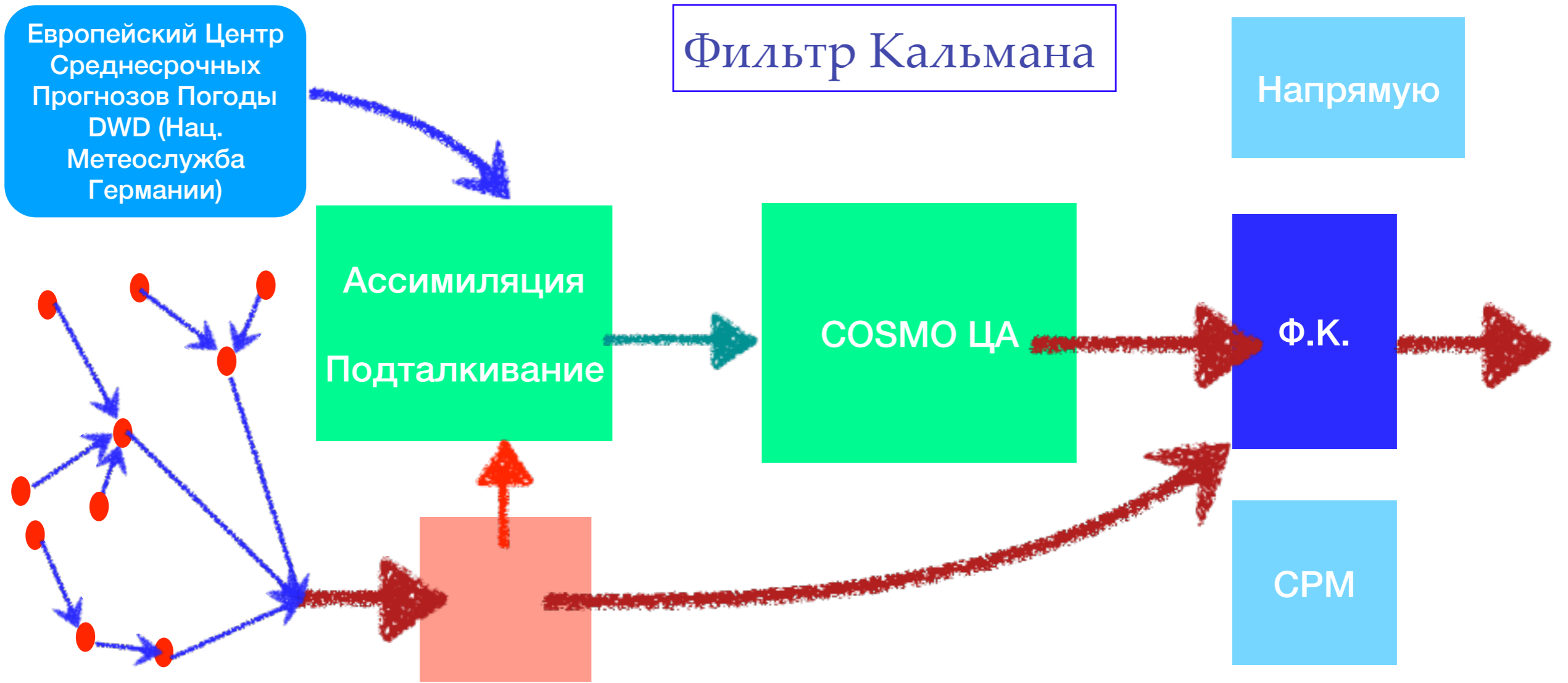
Группа Всемирного банка. Жак Амбюль. Консультант

Широкое видение: от автоматической погодной сети к последующей обработке COSMO

- Широкое видение**
- Фильтр Кальмана**
- Статистика результатов моделирования**
- Логистическая регрессия**
- Индексы экстремального прогноза**
- Графические оповещения/предупреждения**
- Относительная характеристика операций**
- Решение о посылке оповещения**
- Решение о последовательности**

Широкое видение





В режиме реального времени

В период отключения

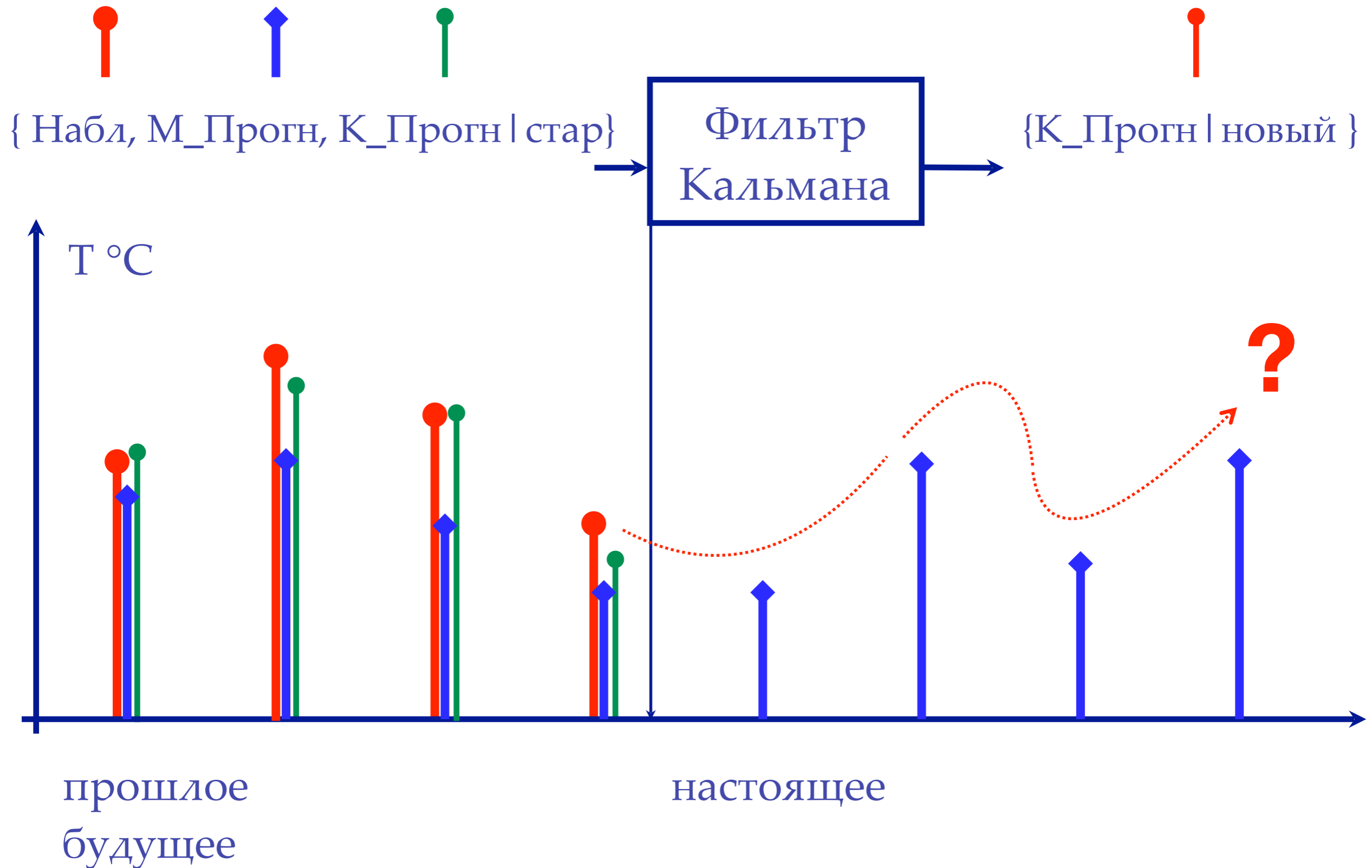


Долгосрочная база данных

Регистрация
Вход
Регистрация...
...
Изучение

Фильтр Кальмана: установка

Фильтр Кальмана



Фильтр Кальмана



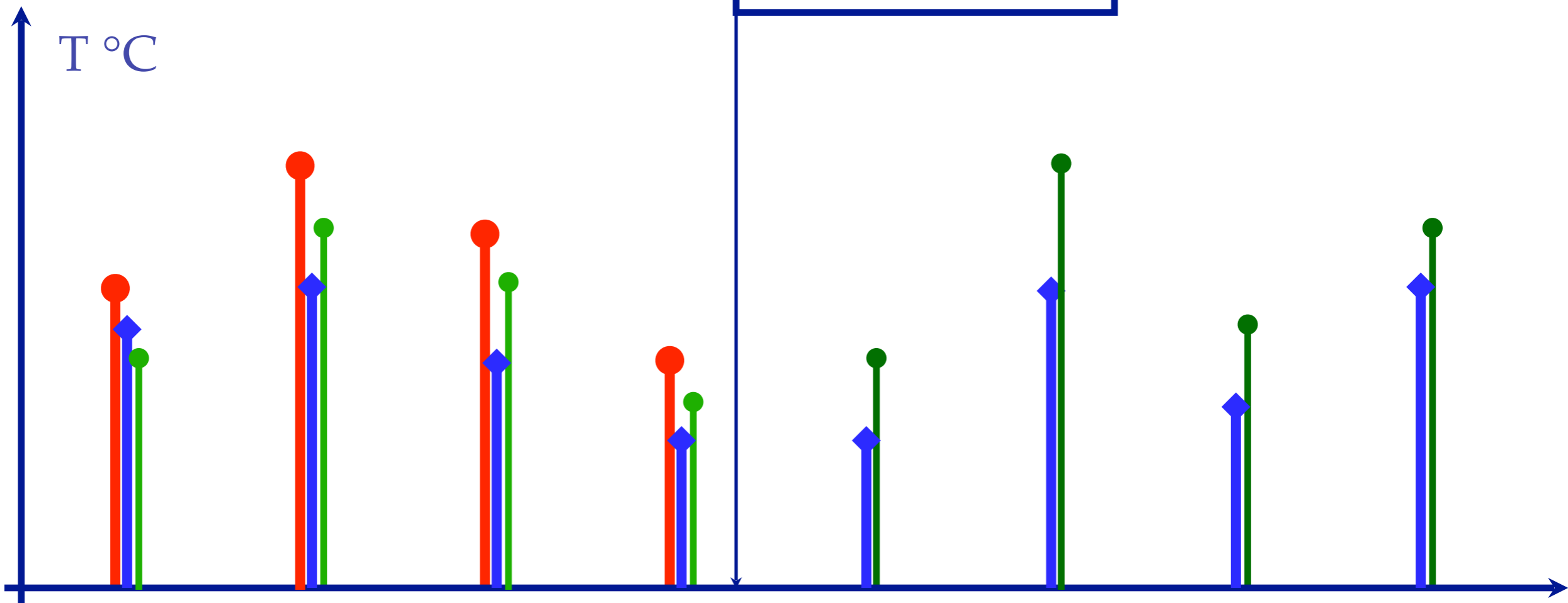
{ Набл, М_Прогн, К_Прогн | стар }



Фильтр
Кальмана



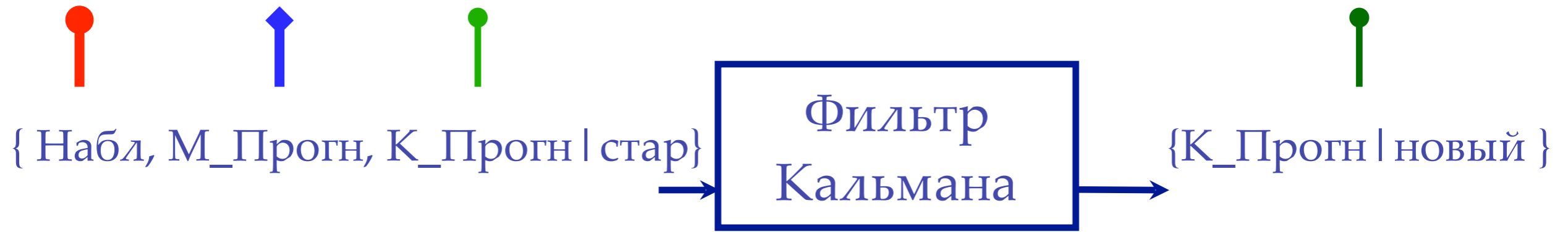
{ К_Прогн | новый }



прошлое
будущее

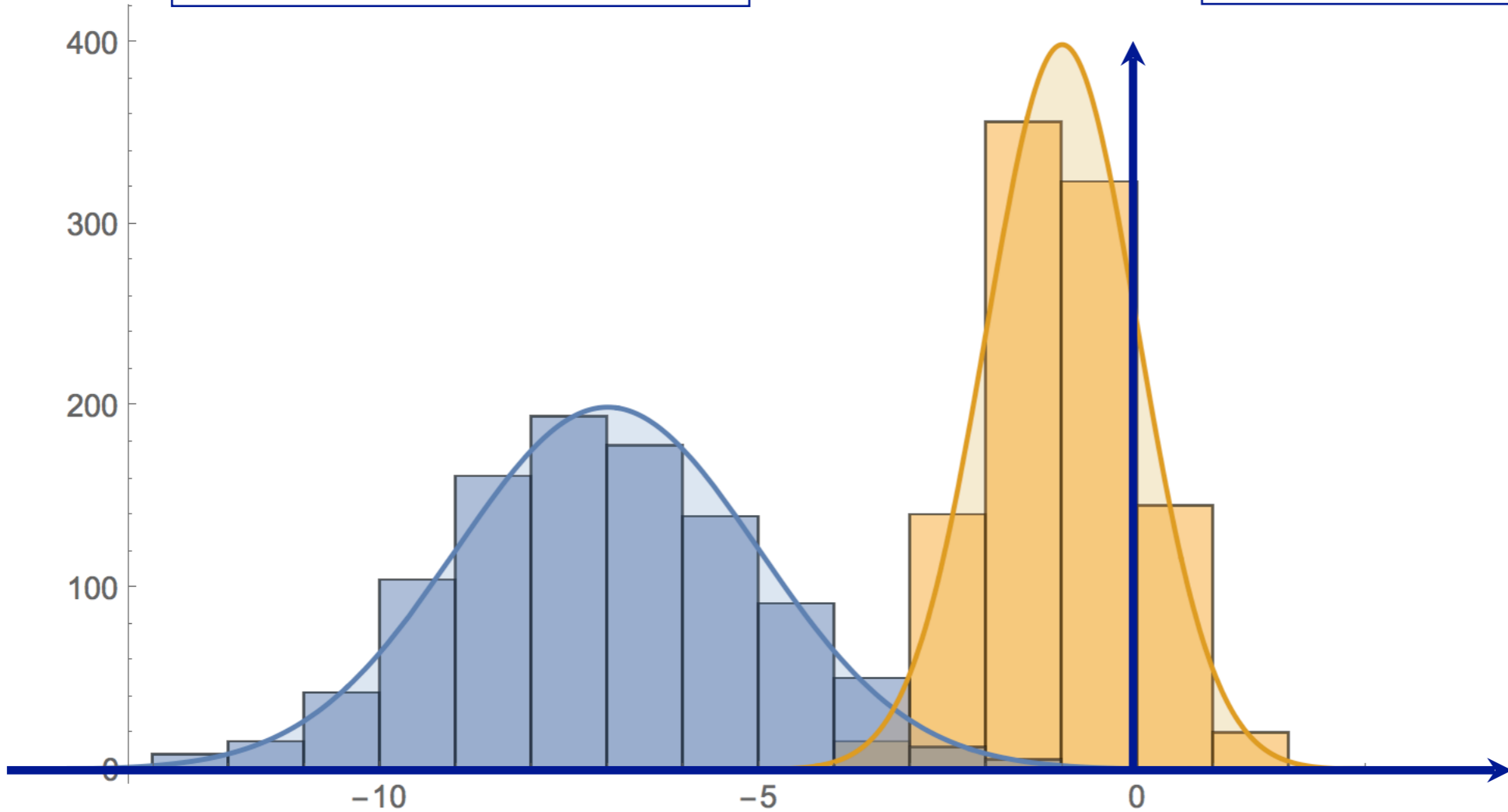
настоящее

Фильтр Кальмана



$|\text{M_Прогн} - \text{Набл}|$

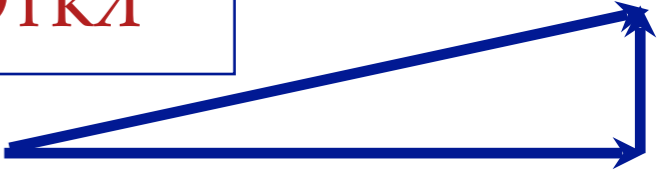
$|\text{K_Прогн} \mid_{\text{новый}} - \text{Набл}|$



Фильтр Кальмана

Среднее квадратичное
Квадратичная ошибка

$$RMSE^2 = \text{Сист.ош.}^2 + \text{СтдОткл}^2$$



Сист.ошибка М_Прогн

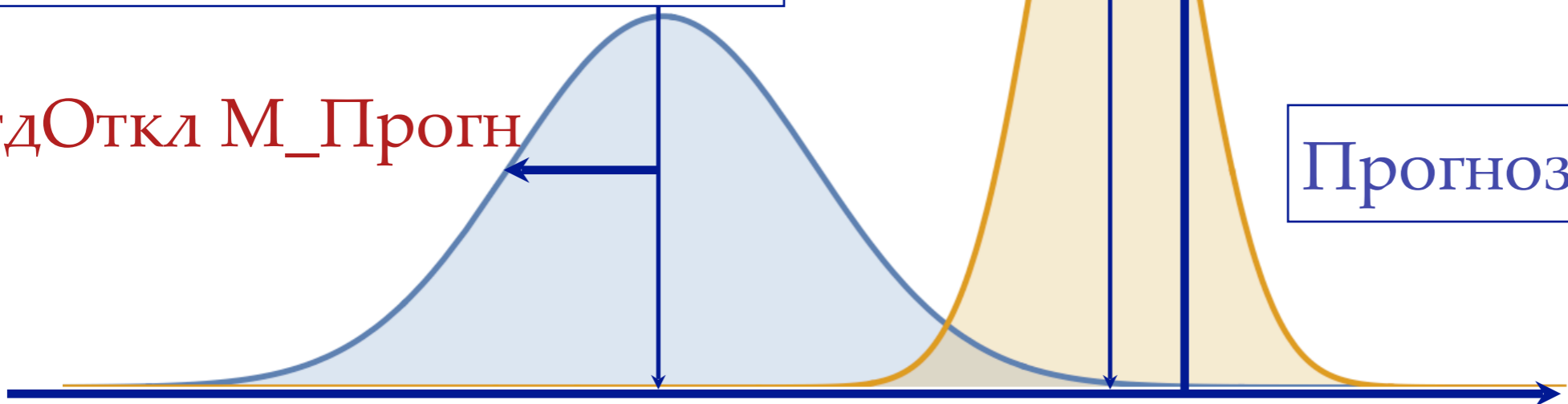
Сист.ош. К_Прогн

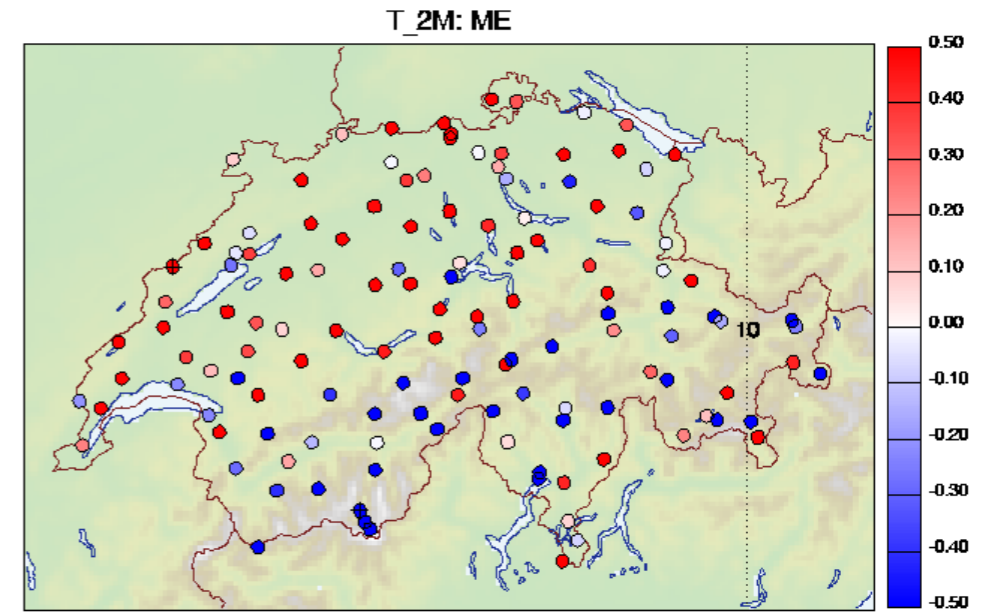
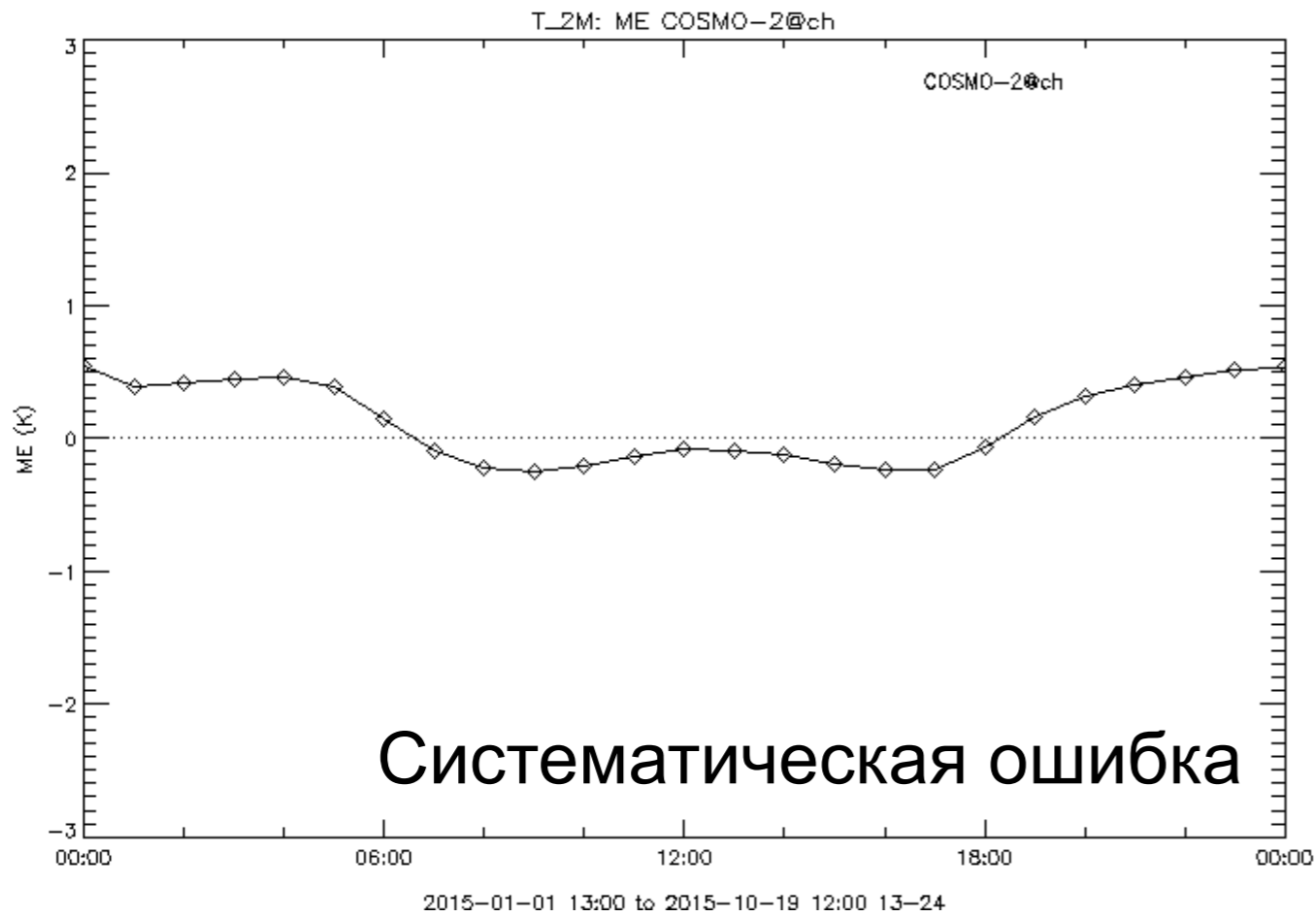
СтдОткл К_Прогн

Прогноз_моделирования

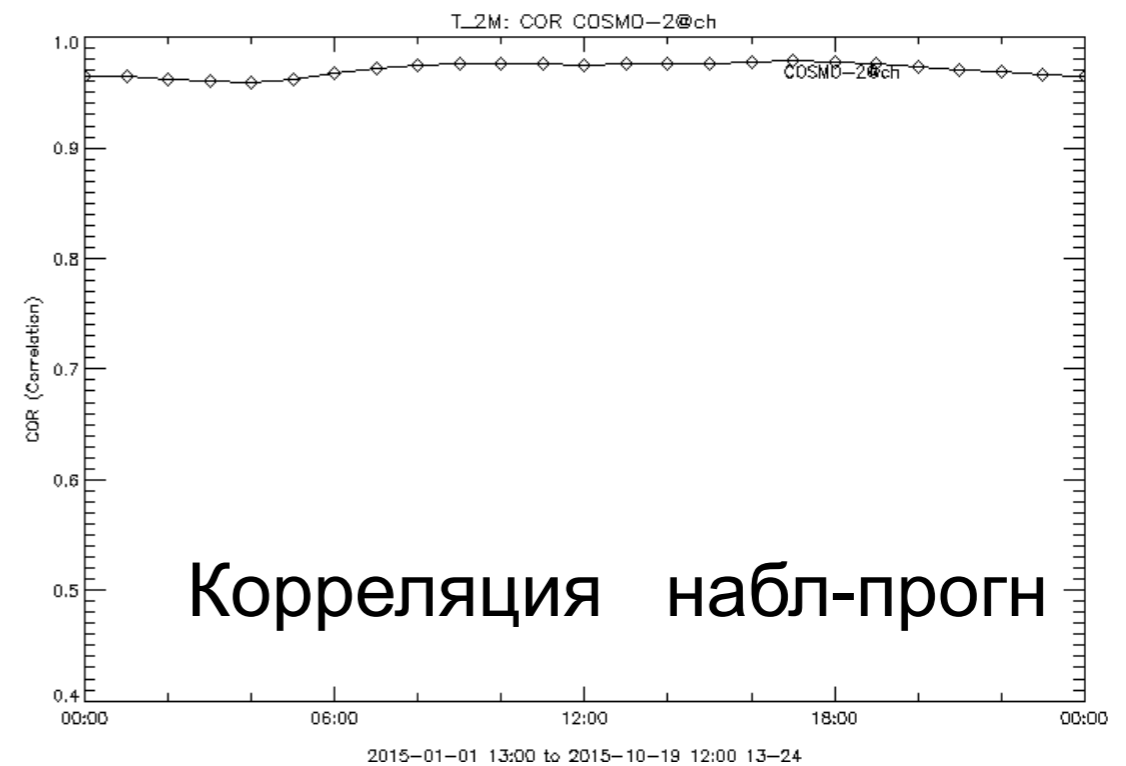
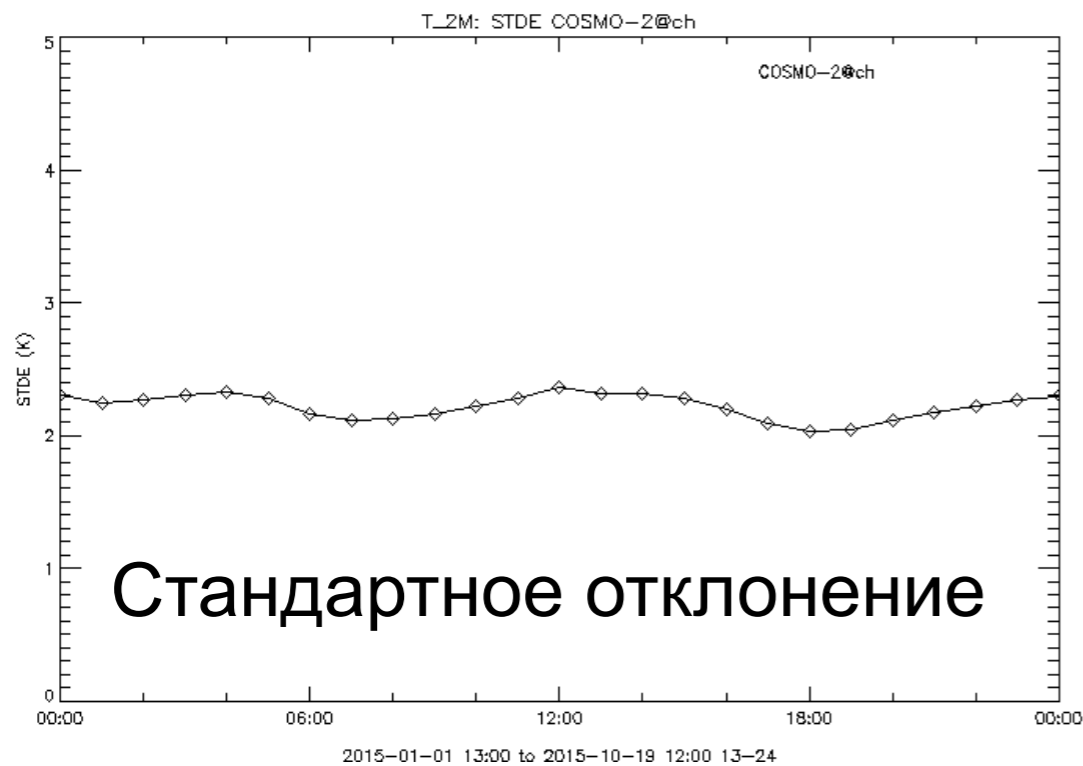
СтдОткл М_Прогн

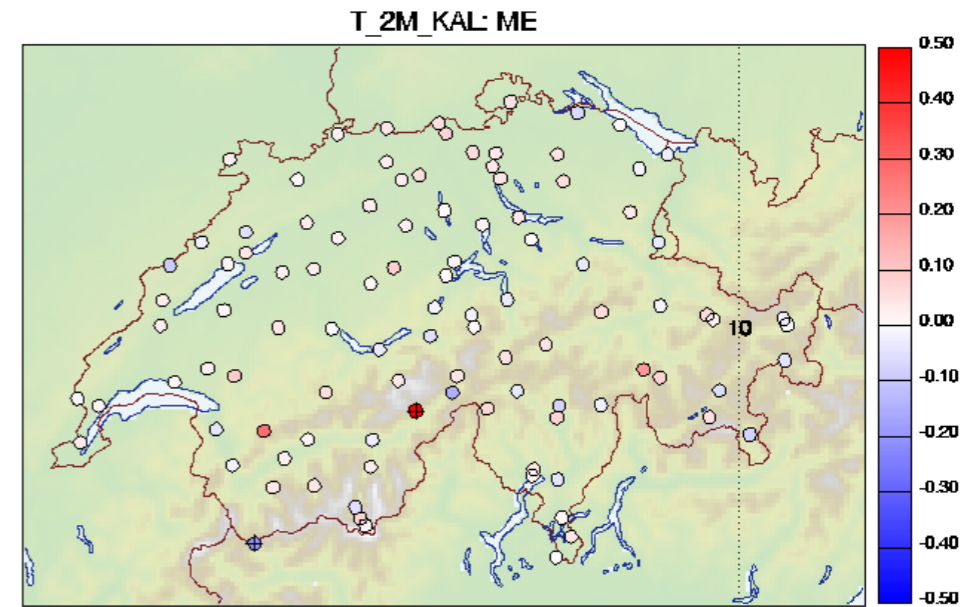
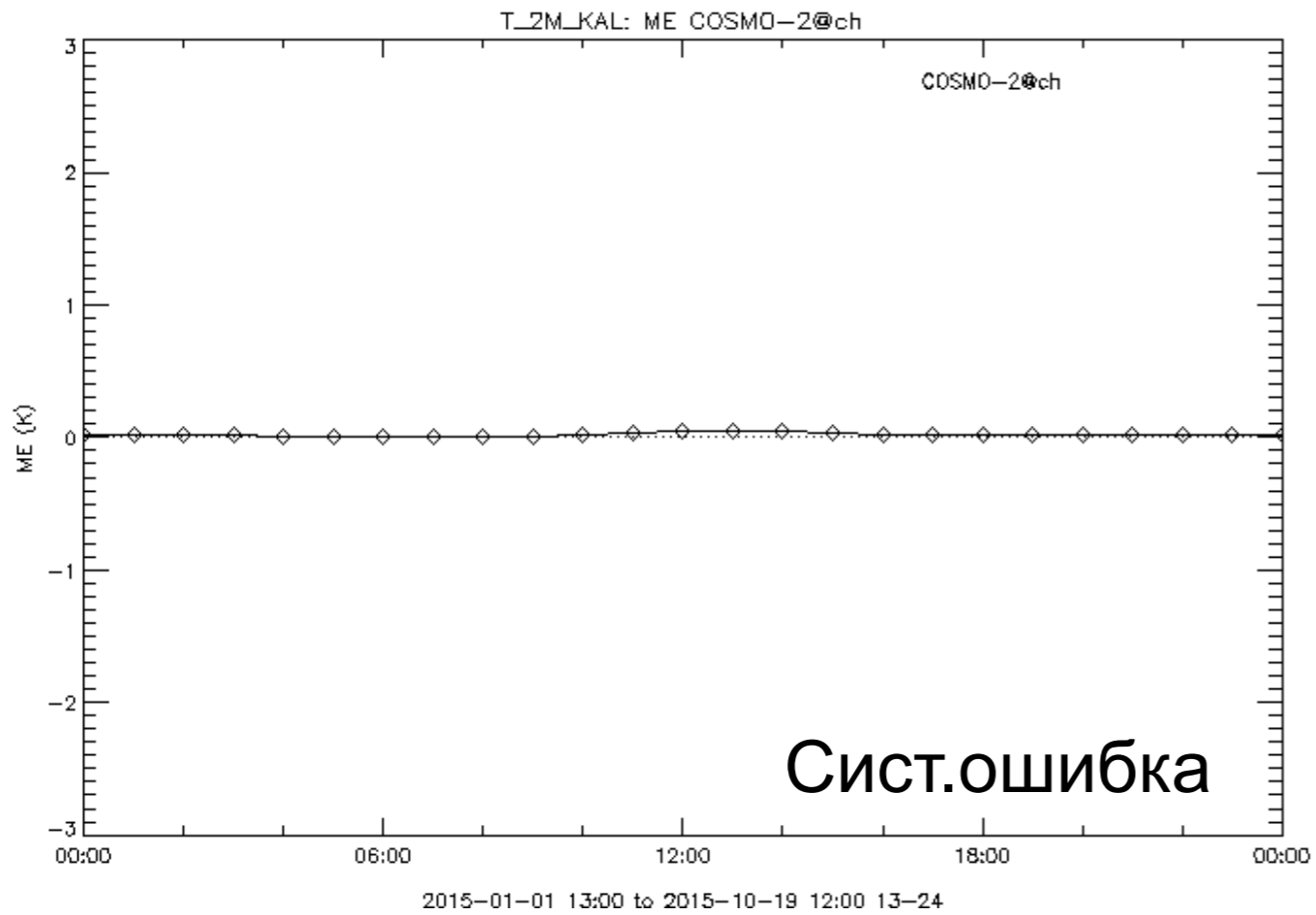
Прогноз_Кальмана



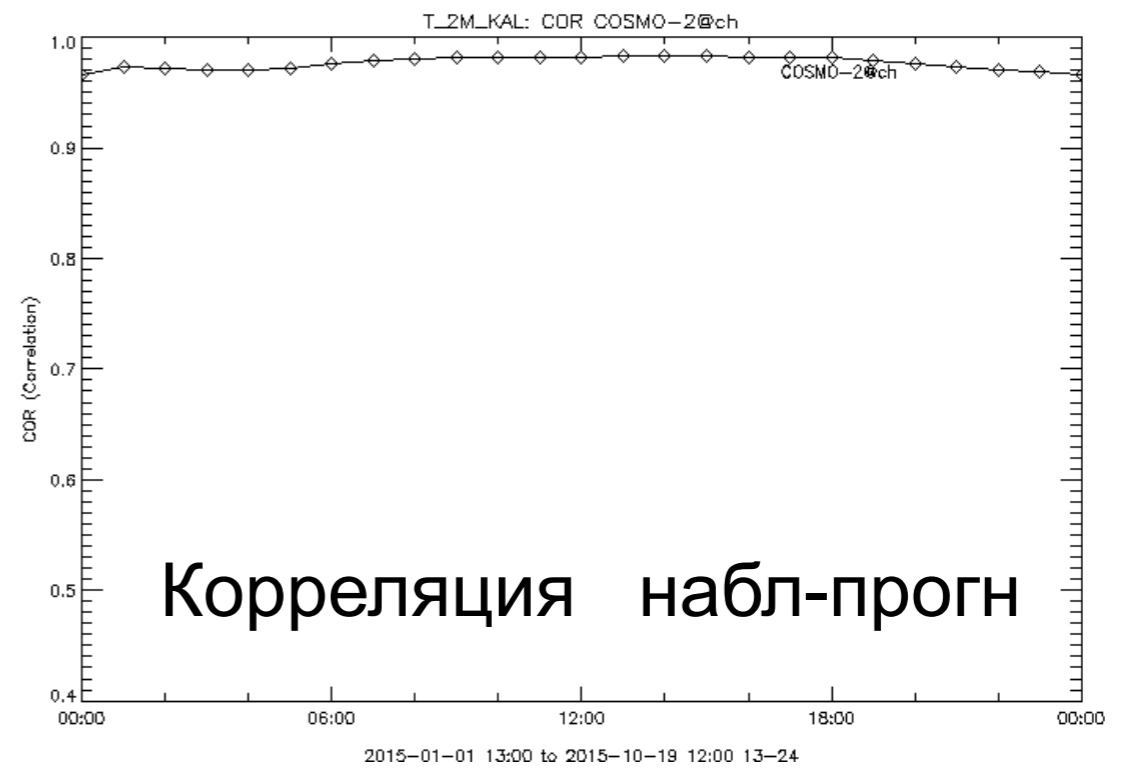
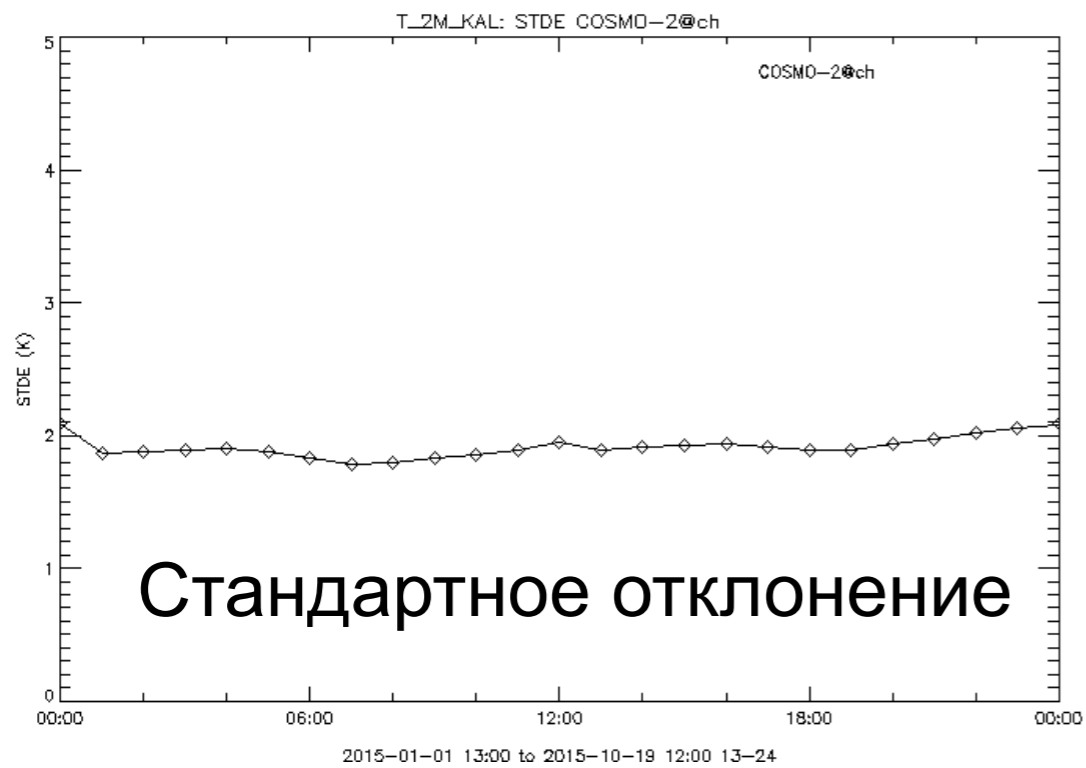


T_2M Прямой результат моделирования

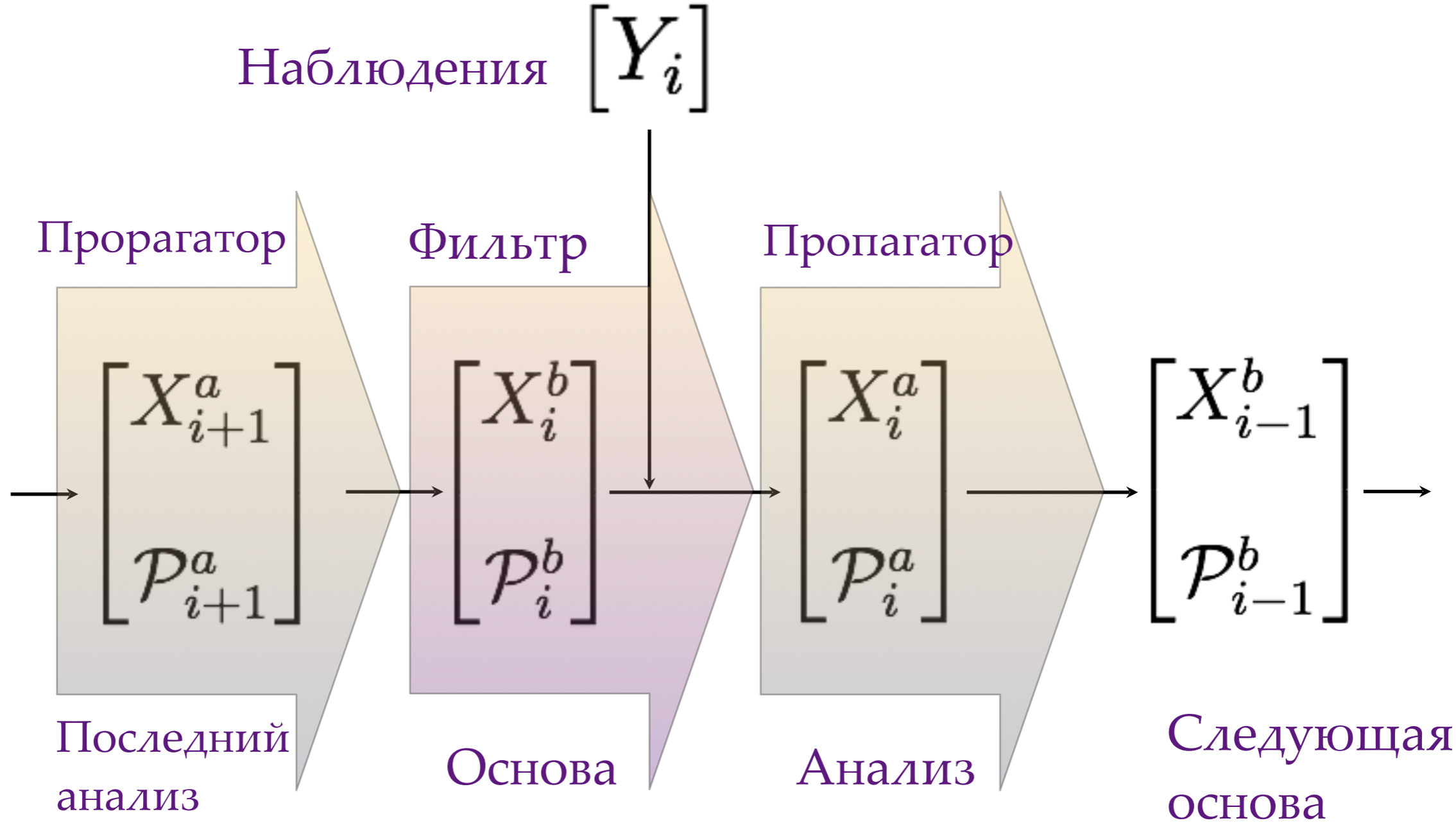




T_2M Кальман



Фильтр Кальмана: Последовательность операций



Фильтр Кальмана: последовательность операций

$$\begin{bmatrix} X_i^a \\ P_i^a \end{bmatrix}$$

- представляет среднее значение и ковариацию гауссового распределения вероятностей, описывающего вероятность состояния системы, учитывая наблюдения от времени t_n до времени t_i .

$$P_i^a$$

- является пропагатором неопределенности, как и ожидалось из уравнения Фоккера-Планка-Колмогорова (позволяет вычислять матрицы $R(\tau)$ в схеме ассимиляции 4D-Var).

Фильтр Кальмана: последовательность операций

$$X_i^b = \mathcal{F}_{i,i+1} X_{i+1}^a$$

$$\mathcal{P}_i^b = \mathcal{F}_{i,i+1} \mathcal{P}_{i+1}^a \mathcal{F}_{i,i+1}^T$$

Пропагатор

$$\mathcal{P}_i^a = (\mathcal{I} + \mathcal{P}_i^b \mathcal{K}_i^T \mathcal{R}_i^{-1} \mathcal{K}_i)^{-1} \mathcal{P}_i^b$$

$$X_i^a = X_i^b + \underbrace{\mathcal{P}_i^a \mathcal{K}_i^T \mathcal{R}_i^{-1}}_{\text{Преимущество}} (Y - \mathcal{K}_i X_i^b)$$

ФИЛЬТ
Р

Преимущество

Фильтр Кальмана



Journal of Statistical Software

March 2011, Volume 39, Issue 2.

<http://www.jstatsoft.org/>

Kalman Filtering in R

Fernando Tusell

University of the Basque Country

Abstract

Support in R for state space estimation via Kalman filtering was limited to one package, until fairly recently. In the last five years, the situation has changed with no less than four additional packages offering general implementations of the Kalman filter, including in some cases smoothing, simulation smoothing and other functionality. This paper reviews some of the offerings in R to help the prospective user to make an informed choice.

Keywords: state space models, Kalman filter, time series, R.

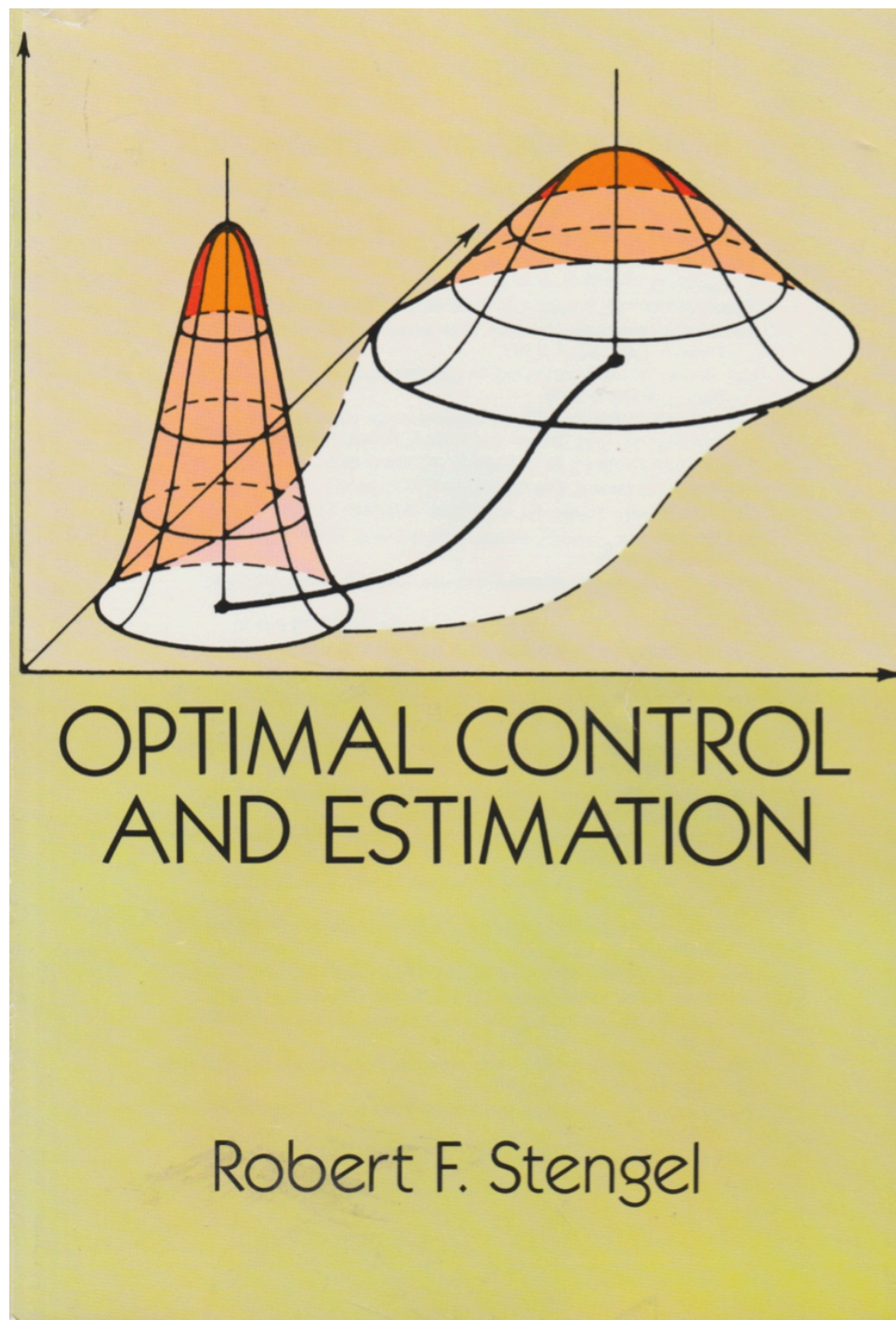
Визуальная статистика:
мы используем R!

Заключения по Фильтру Кальмана

- учитывает местную климатологию
- подходит для непрерывных параметров:
 - температура, влажность, интенсивность ветра, давление, ...
- фильтр не учится, но динамически подгоняет модельные прогнозы к наблюдениям
 - темп -> темп, влажность -> влажность, ... -> ...
- включен в пакет прогнозирования
- довольно чувствителен
- не очень хорошо справляется с резкими изменениями погоды
- Однако, ...

... Фильтр Кальмана используется во многих областях

Фильтр Кальмана

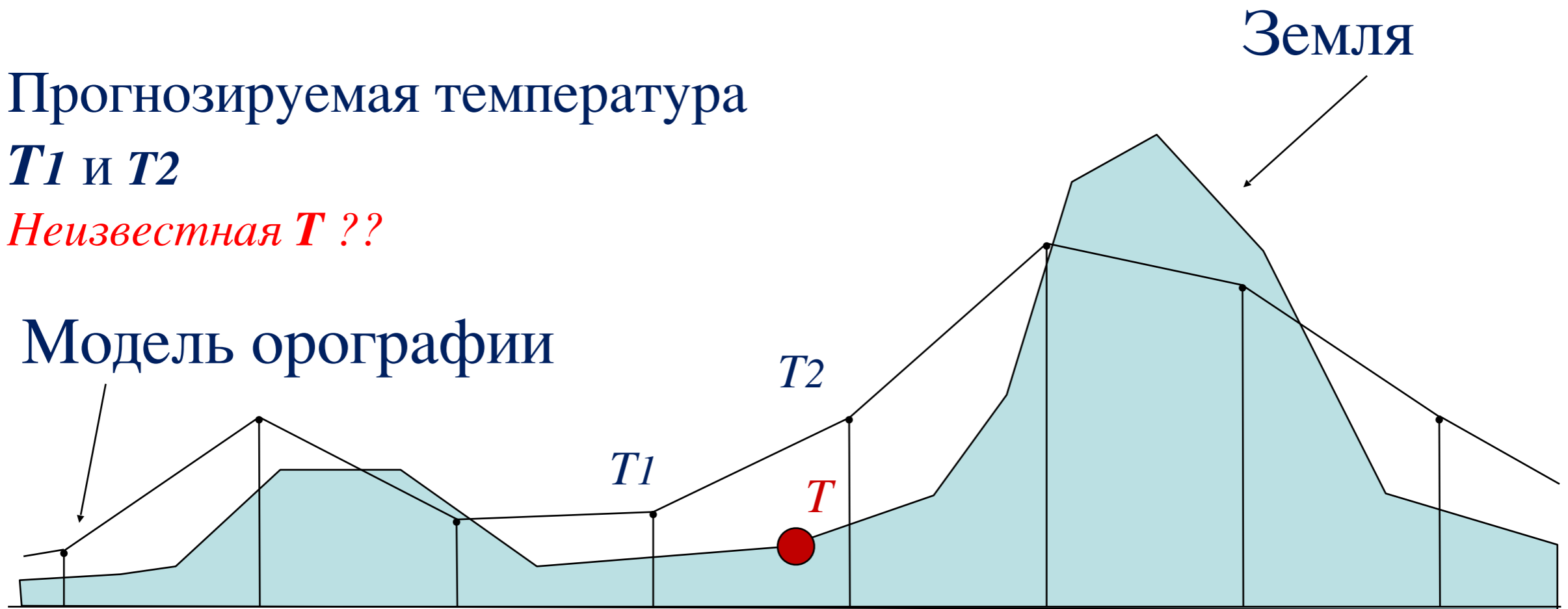


Прогнозируемая температура

T_1 и T_2

Неизвестная T ??

Модель орографии



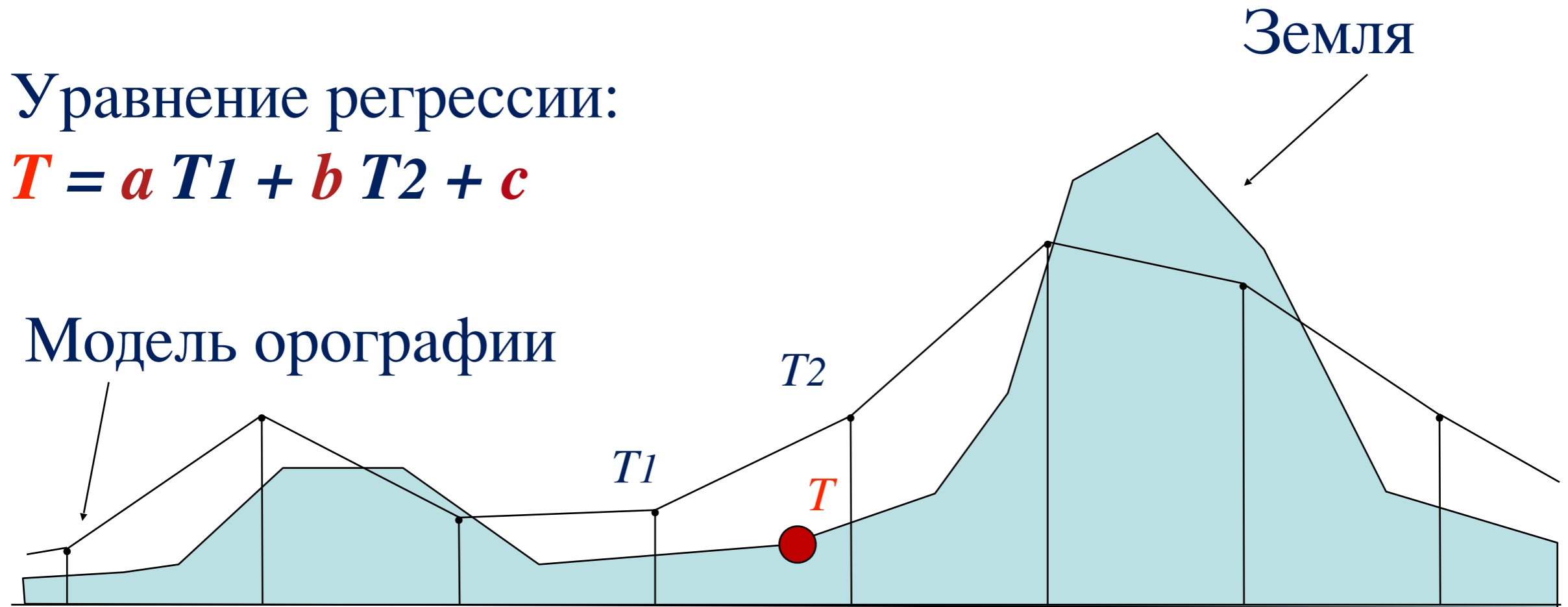
Основа для обучения: данные климатологии и прошлые модели прогнозирования:

$\{\{T_1, T_2, T\}_1, \dots, \{T_1, T_2, T\}_2, \dots, \{T_1, T_2, T\}_N\}$, с $N \sim 100 - 1000$

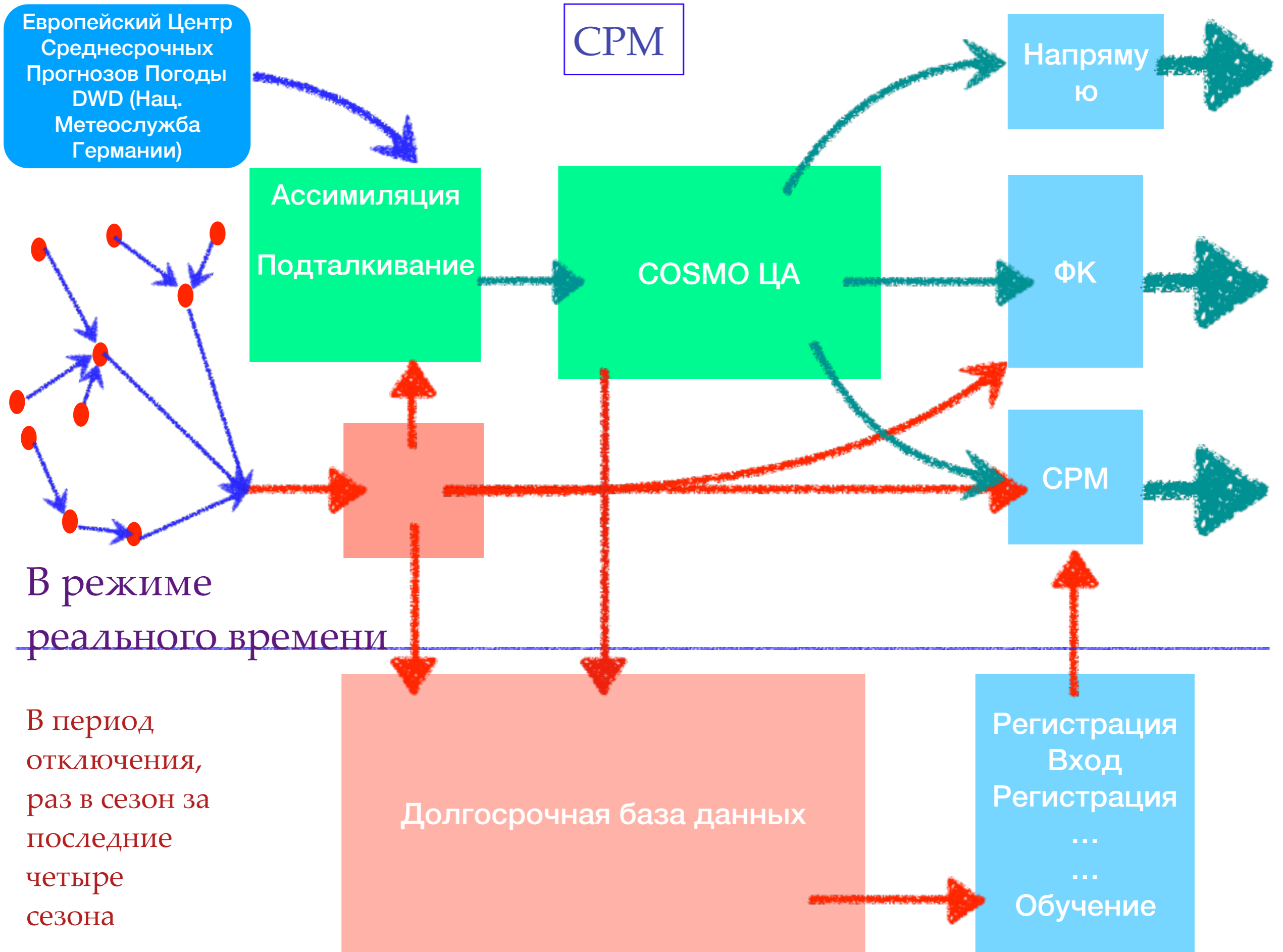
Уравнение регрессии:

$$T = a T1 + b T2 + c$$

Модель орографии



- ❑ Обучение машины: a, b, c будут вычисляться статистически
- ❑ В использовании: Модель предоставляет $T1$ & $T2$ на +48 часов.
- ❑ И в вычислениях: $T+48 = a T1 + b T2 + c$



CRM

Напрямую

Ассимиляция
Подталкивание

COSMO ЦА

ФК

CRM

[Red box]

Долгосрочная база данных

Регистрация
Вход
Регистрация
...
...
Обучение

В режиме реального времени

В период отключения, раз в сезон за последние четыре сезона

Европейский Центр
Среднесрочных
Прогнозов Погоды
DWD (Нац.
Метеослужба
Германии)

СРМ

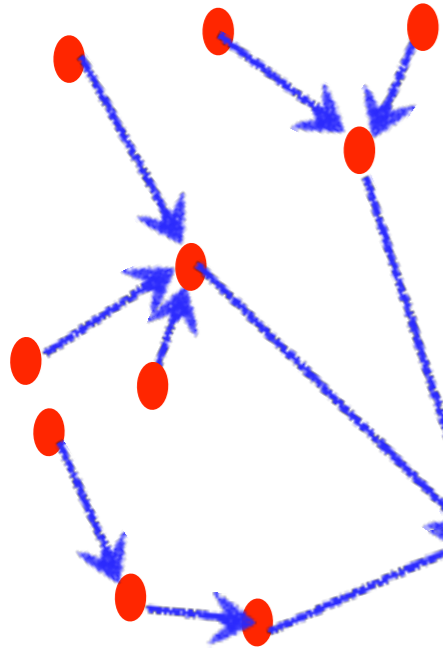
Напрямую

Ассимиляция
Подталкивание

COSMO ЦА

Ф.К.

СРМ



В режиме реального
времени

В период
отключения,
раз в сезон за
последние
четыре
сезона

СРМ: запись реляционных данных

$\{T, T_1, T_2\}_1,$
 $\{T, T_1, T_2\}_2,$
.... ,
 $\{T, T_1, T_2\}_N, N \sim 300$

Регрессия
для
параметров.
 a, b, c, \dots

Фабрика CPM

$$\text{Станция \#1: } X_{\#1} = a_{11} P_1 + a_{12} P_2 + a_{13} P_3 + \dots + a_{1N} P_N$$

$$\text{Станция \#2: } X_{\#2} = a_{21} P_1 + a_{22} P_2 + a_{23} P_3 + \dots + a_{2N} P_N$$

$$\text{Станция \#3: } X_{\#3} = a_{31} P_1 + a_{32} P_2 + a_{33} P_3 + \dots + a_{3N} P_N$$

.....

$$\text{Станция \#M: } X_{\#M} = a_{M1} P_1 + a_{M2} P_2 + a_{M3} P_3 + \dots + a_{MN} P_N$$

Прогнозируемый параметр (предоставляемый моделью):

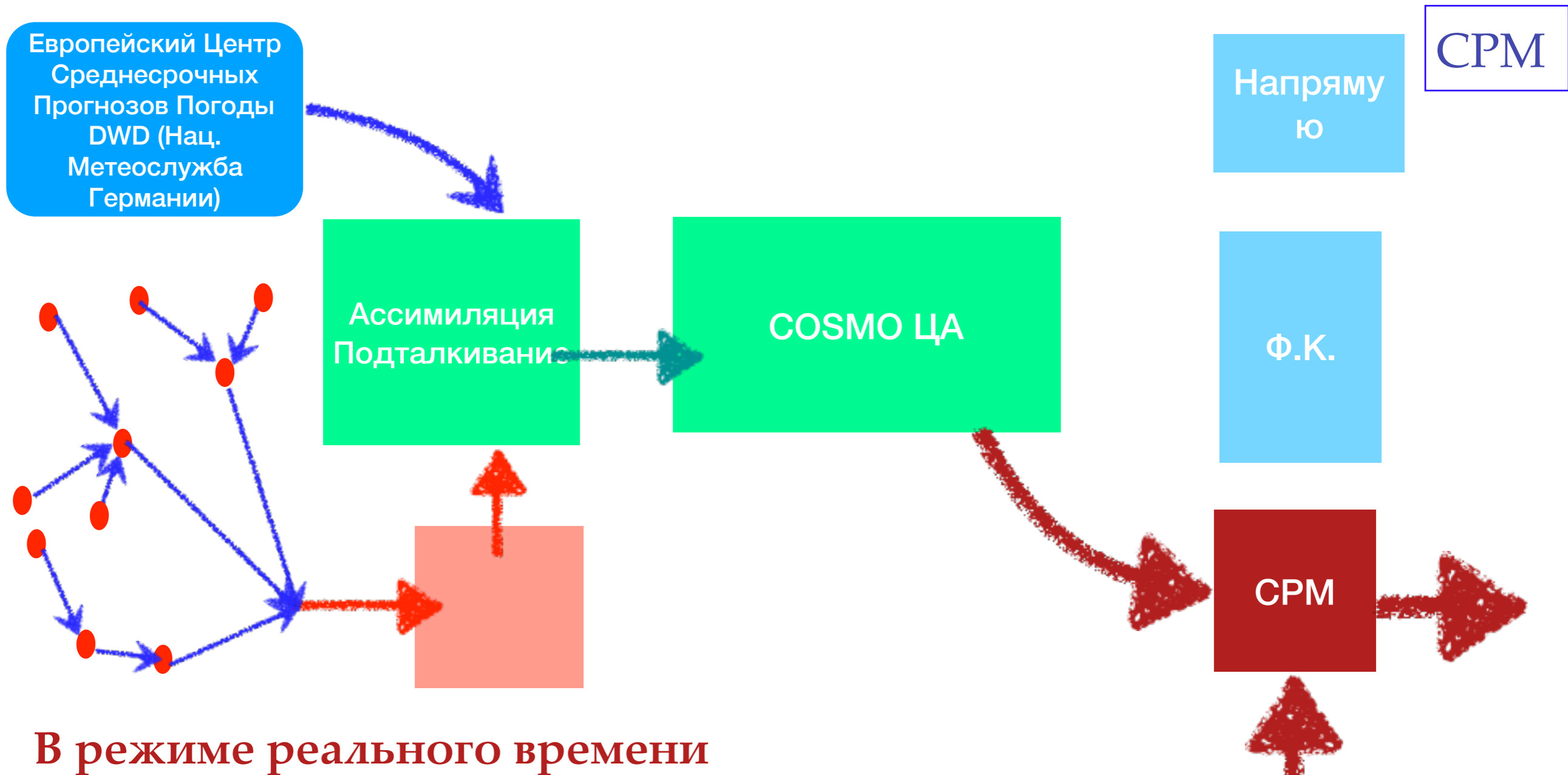
$P_j \rightarrow \{ \text{геопотенциал, адвекция вихря, модель осадков, ...} \}$

Прогнозируемый элемент (для каждой станции):

$X_{\#i} \rightarrow \{ \text{осадки, солнечное излучение, сильные порывы ветра, ...} \}$

Алгоритм вычисляет для каждой станции $\#i$

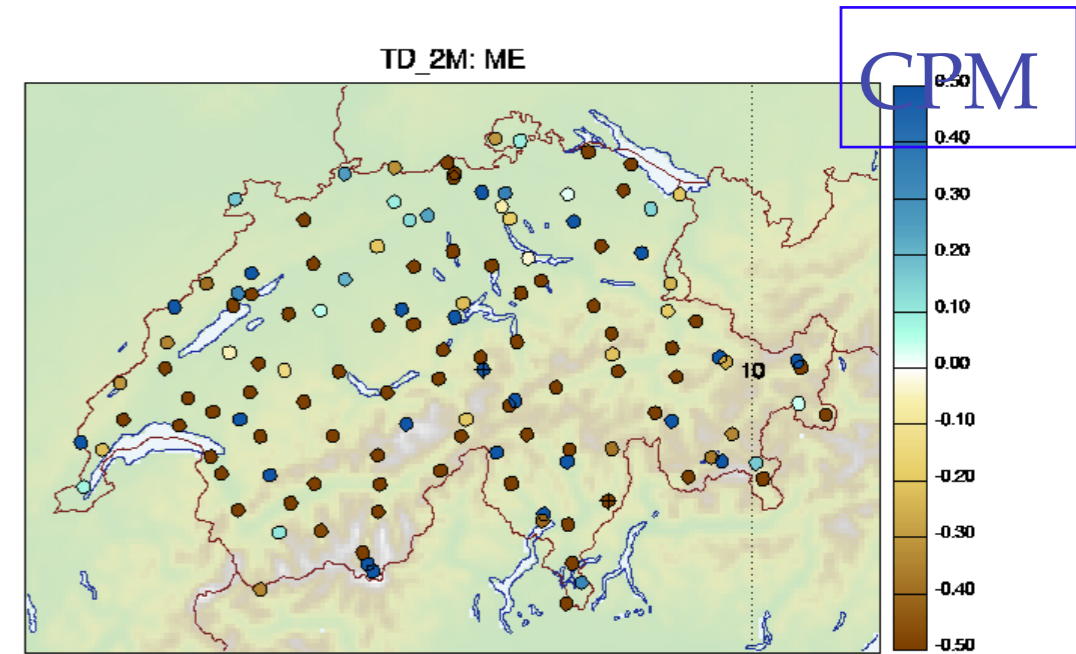
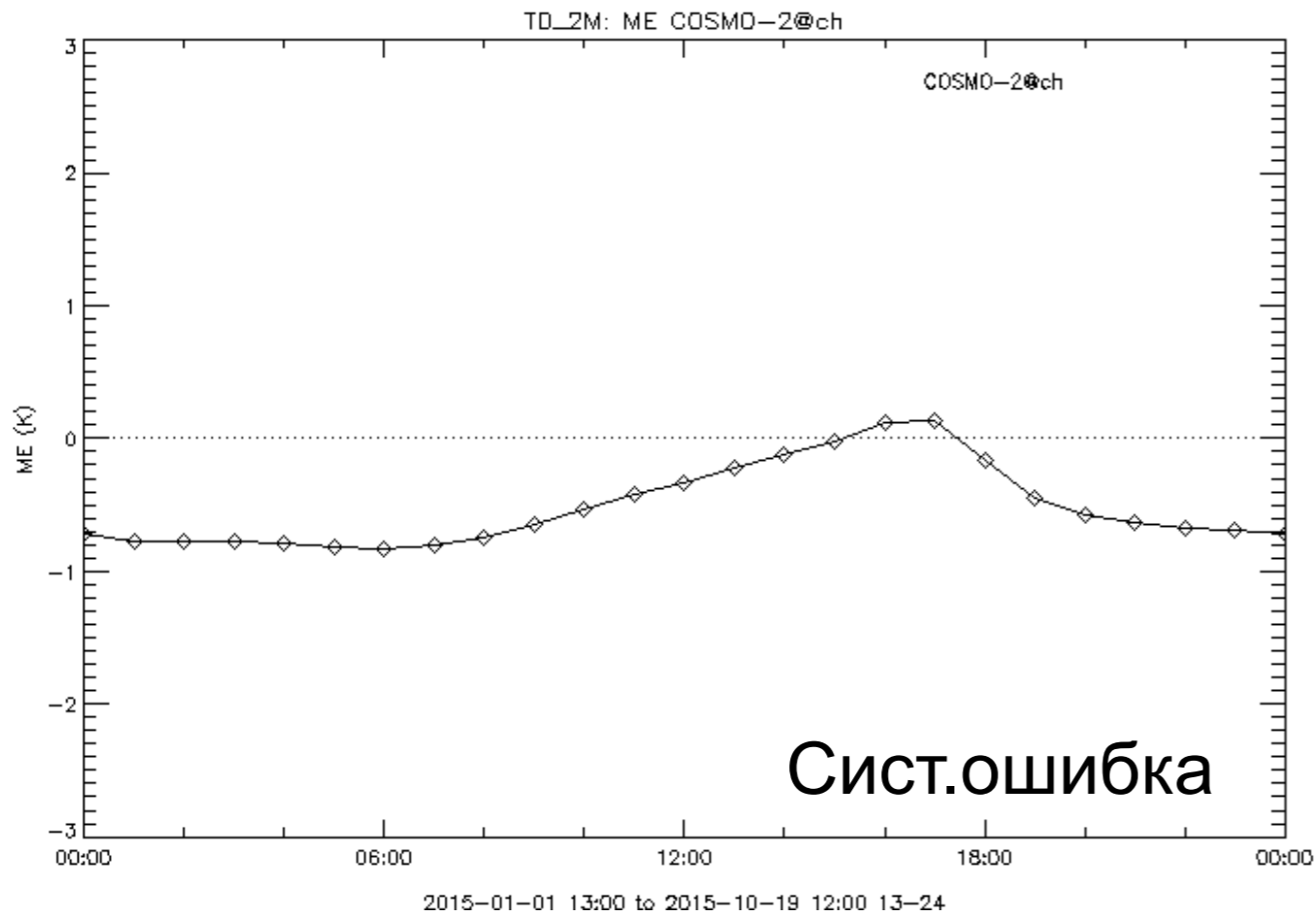
коэффициенты CPM a_{ij} и отказывается от неактуальных



В период отключения, раз в сезон за последние четыре сезона

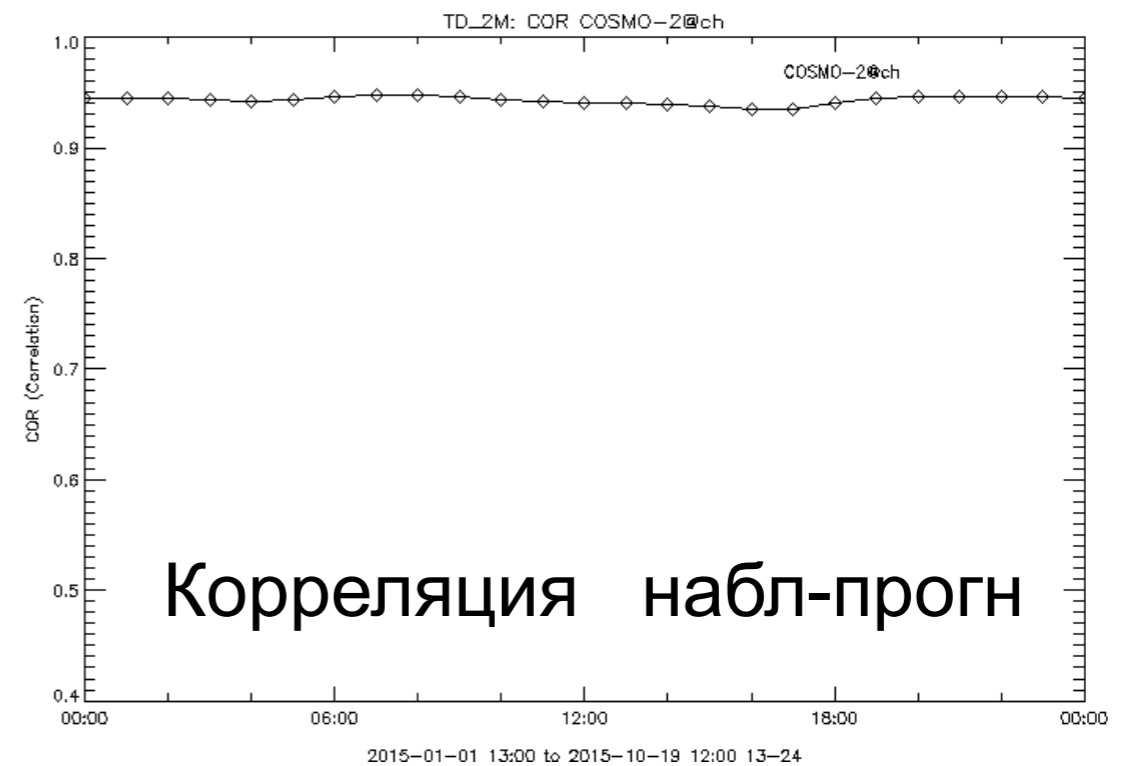
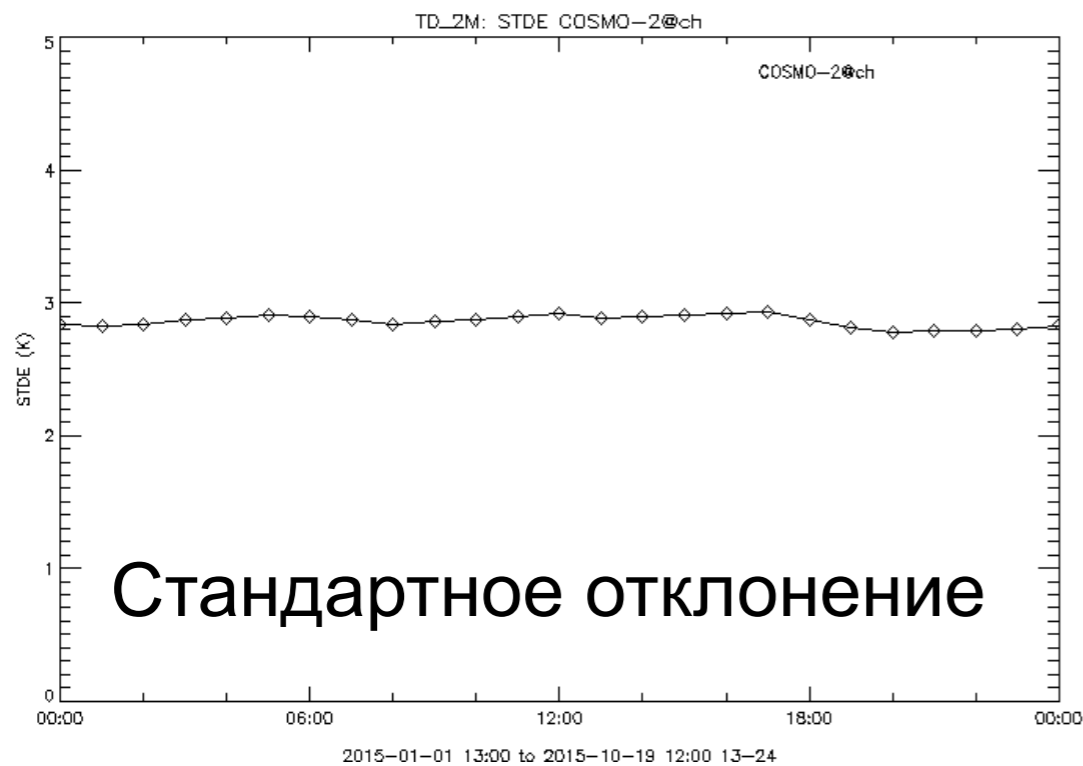
Долгосрочная база данных

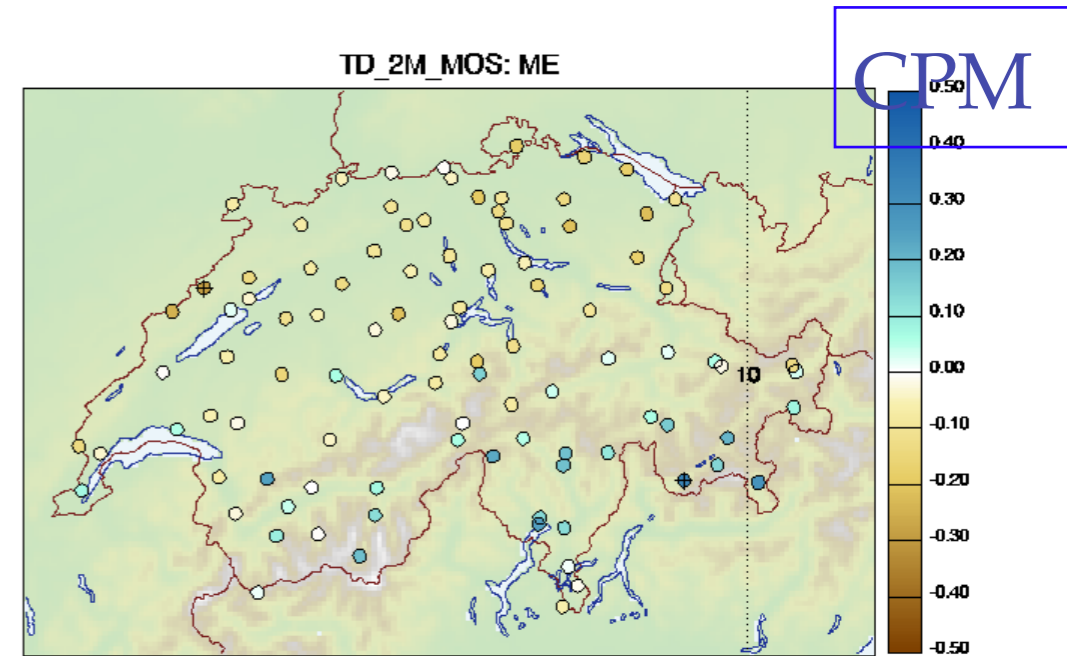
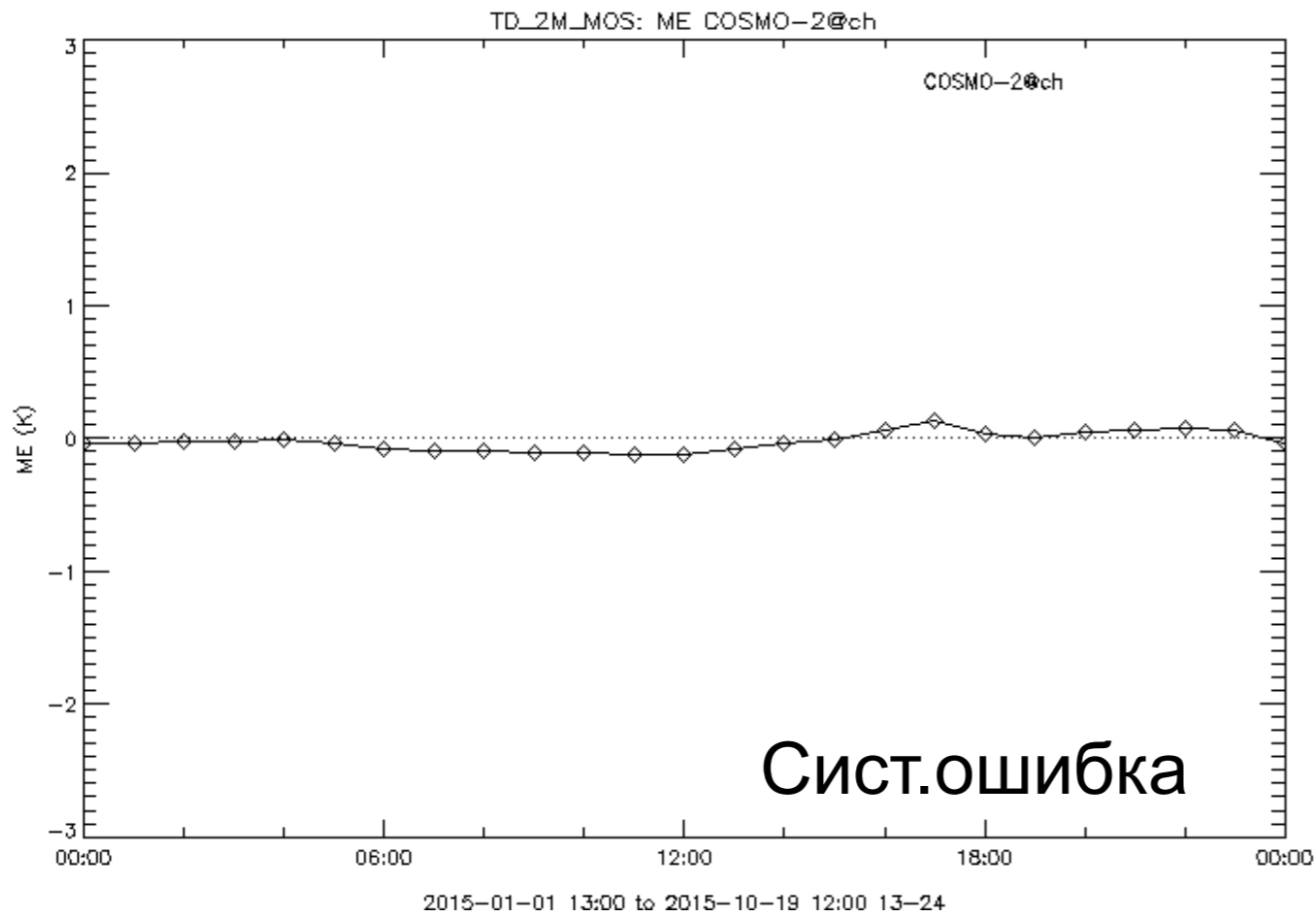
Регистрация
Вход
Регистрация
...
...
Обучение



COSMO-2@ch 2015-01-01 13:00 to 2015-10-19 12:00 13-24
 +Min: -3.446 K at station 06758 +Max: 4.692 K at station 06740

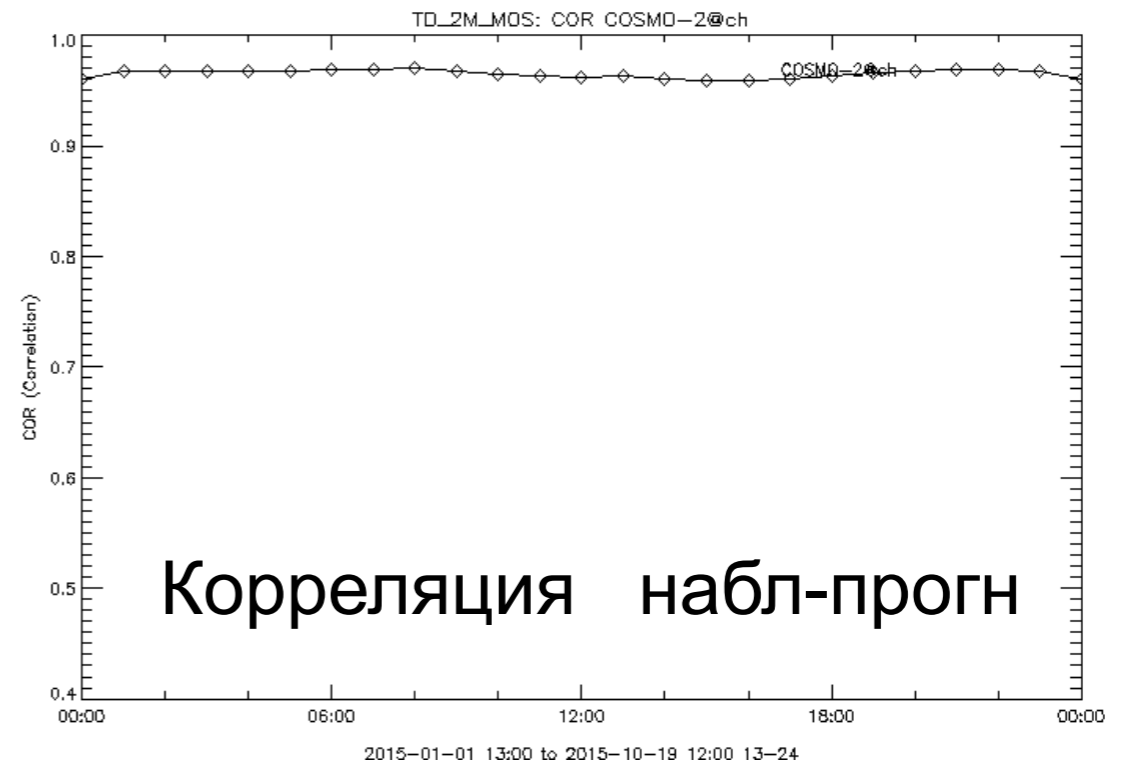
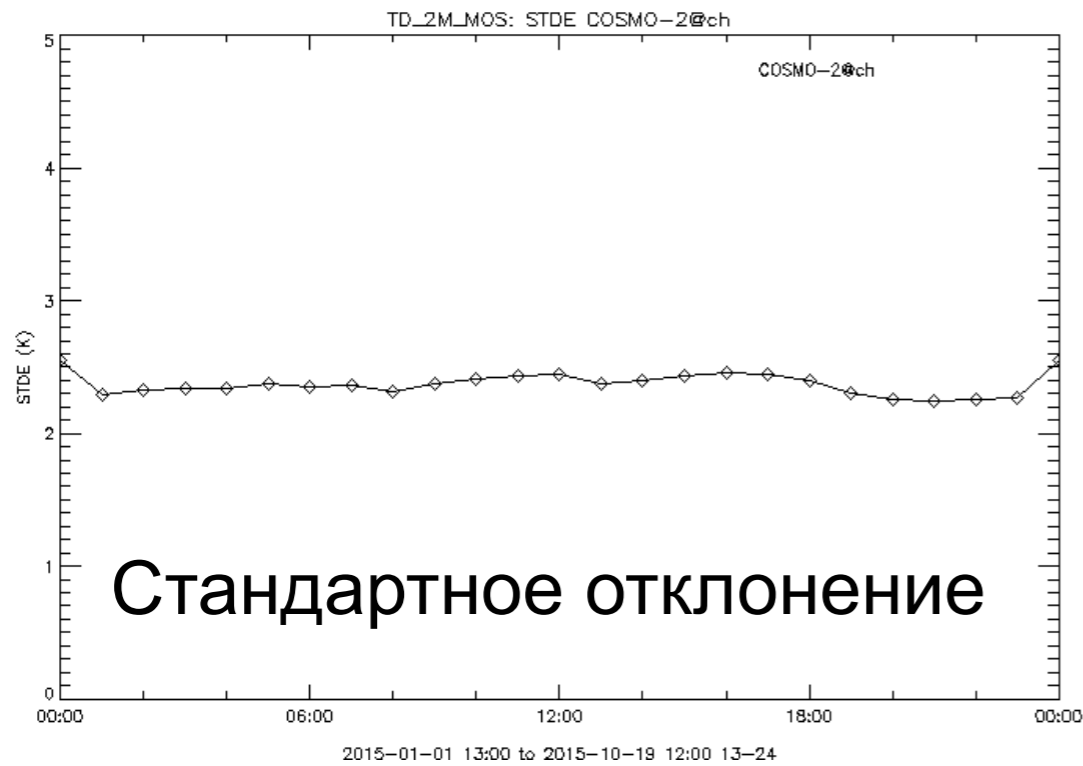
TD_2M Прямые результаты моделирования

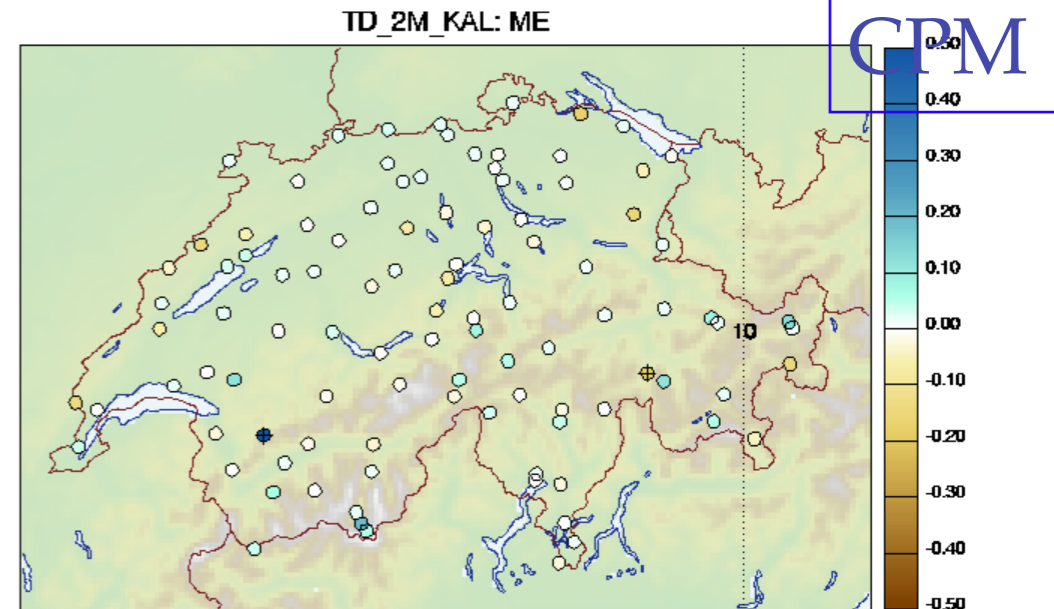
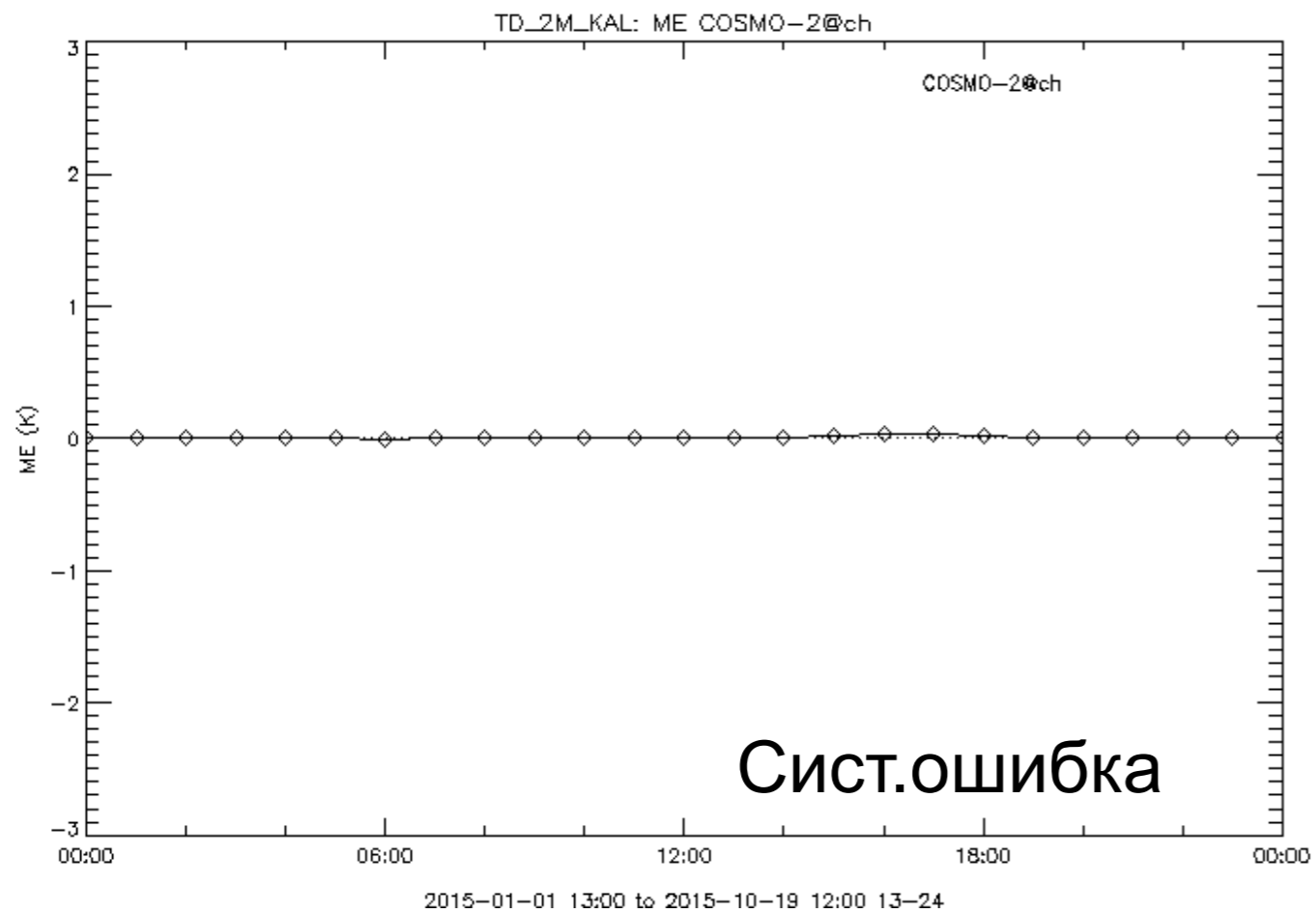




COSMO-2@ch 2015-01-01 13:00 to 2015-10-19 12:00 13-24
 +Min: -0.2912 K at station 06612 +Max: 0.3457 K at station 06788

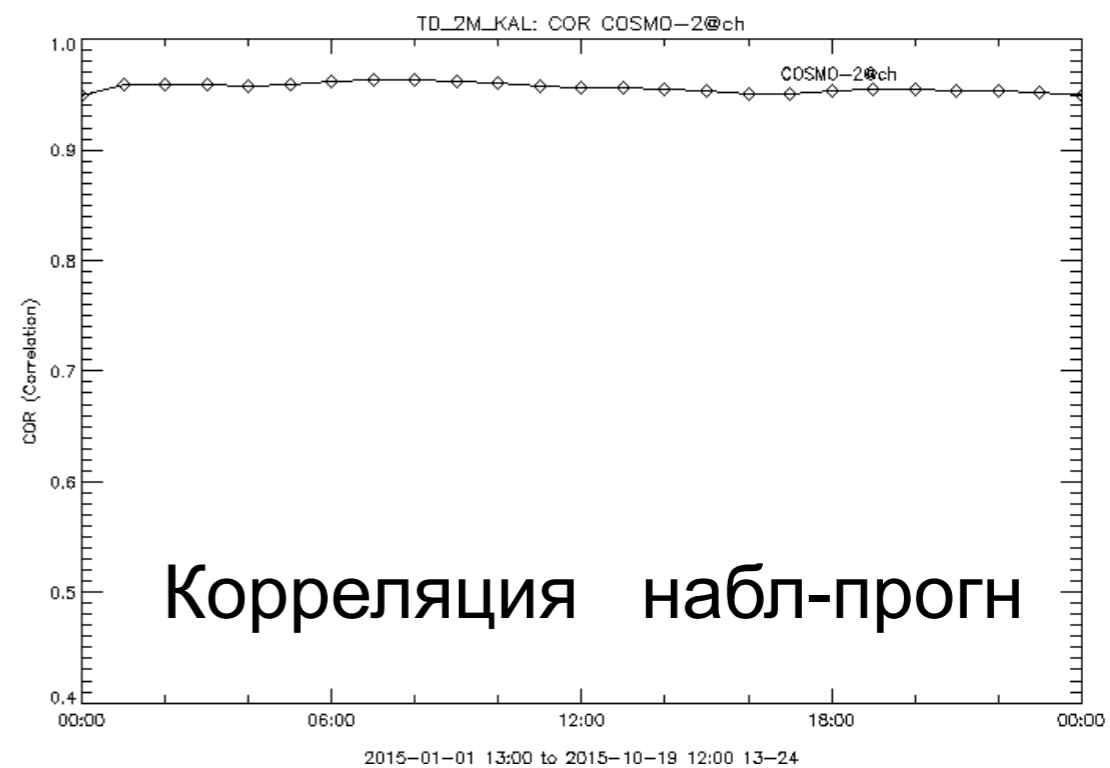
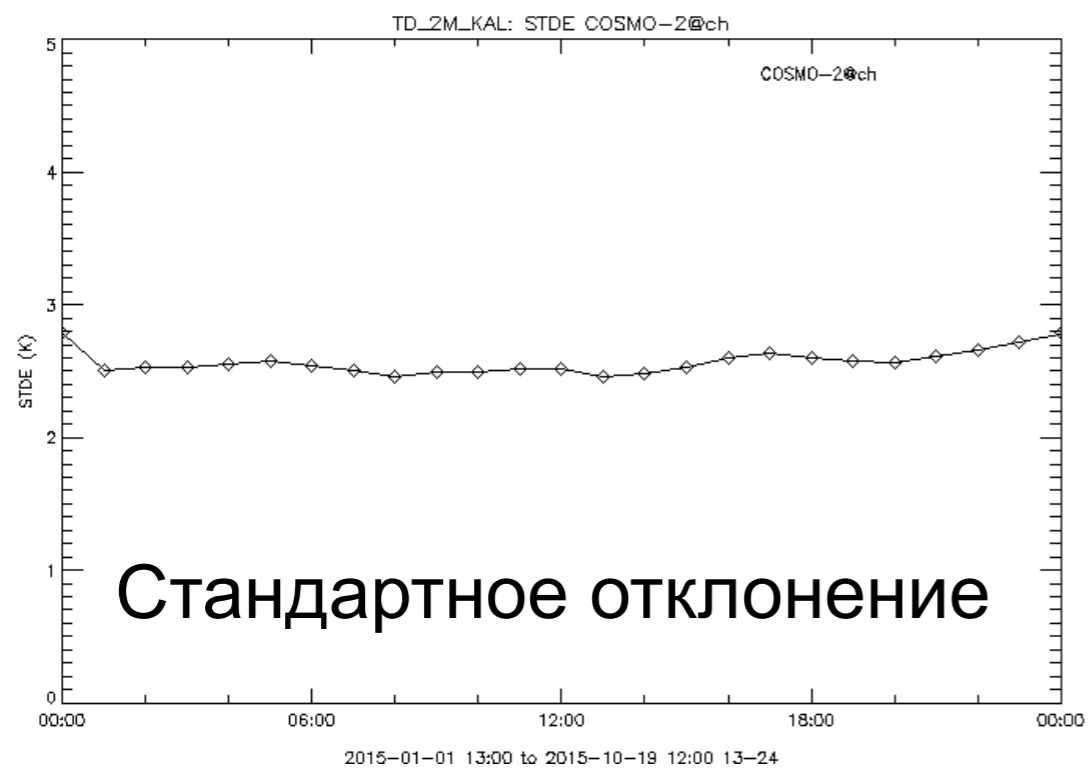
TD_2M CPM





COSMO-2@ch 2015-01-01 13:00 to 2015-10-19 12:00 13-24
 + Min: -0.1762 K at station 06787 + Max: 0.5977 K at station 06714

TD_2M Кальман

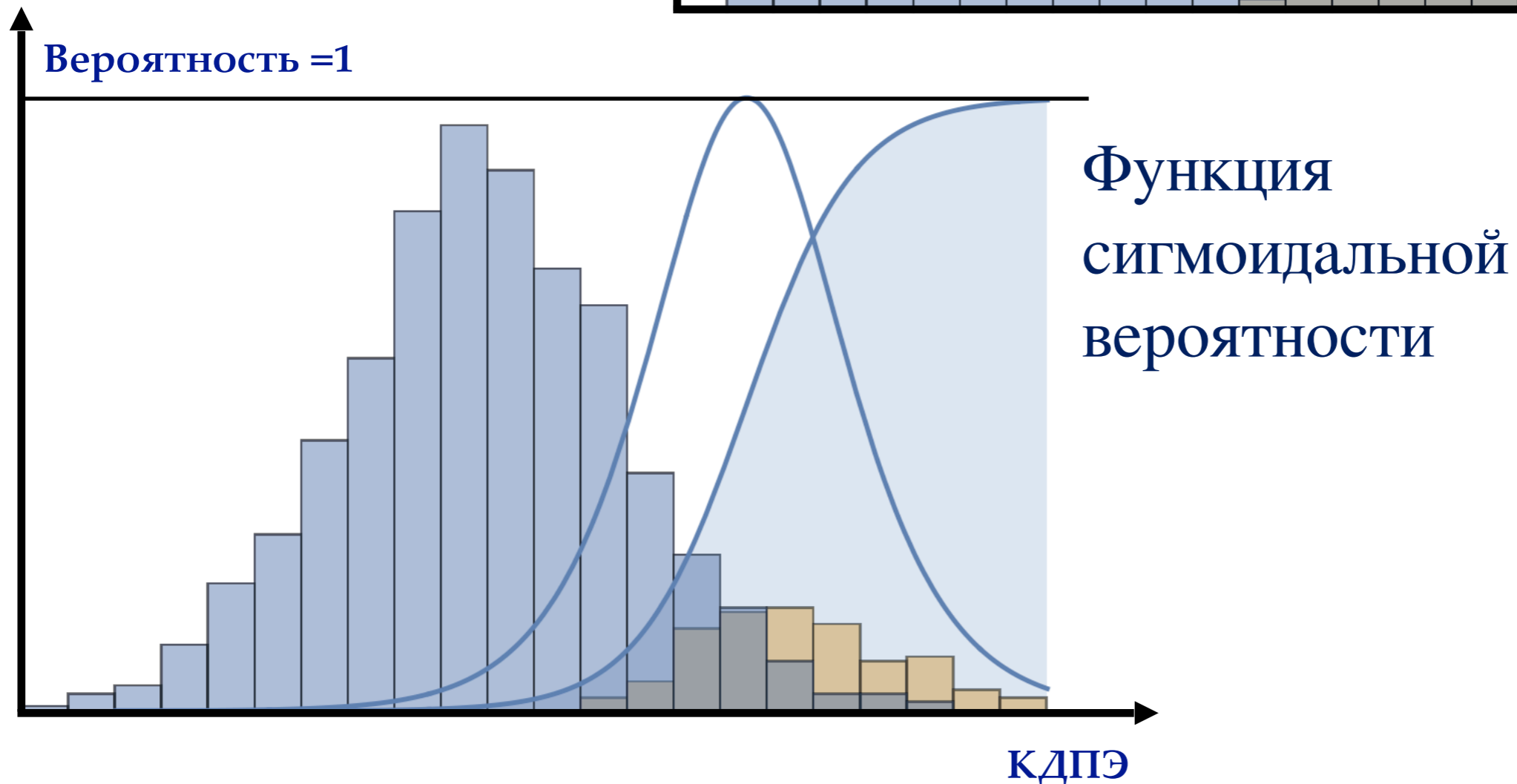
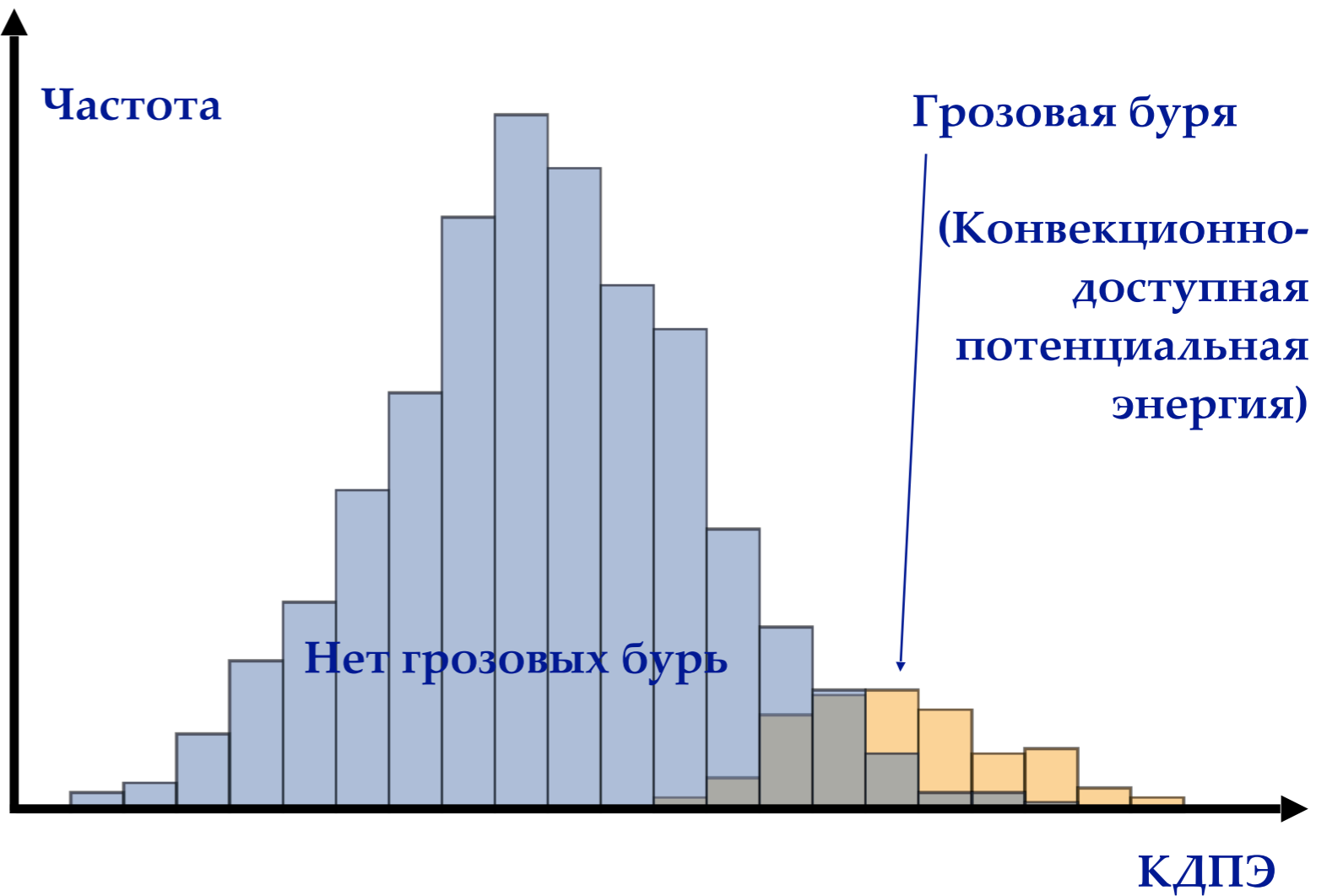


Заключения по Регрессии CRM:

- принимает во внимание местную климатологию
- подходит для спорадических событий:
 - осадки, ветровые штормы, загрязнение воздуха, ..
- изучает и записывает характеристики прошлых событий
- процесс обучения происходит «вне режима реального времени»,
- CRM связывает внешние события, не включенные в модель, с параметрами модели. Пример: загрязнение воздуха в городе с {геопотенциалом, осадками, интенсивностью ветра, ...}
- менее чувствителен, чем Кальман
- вполне может справиться с резкими изменениями

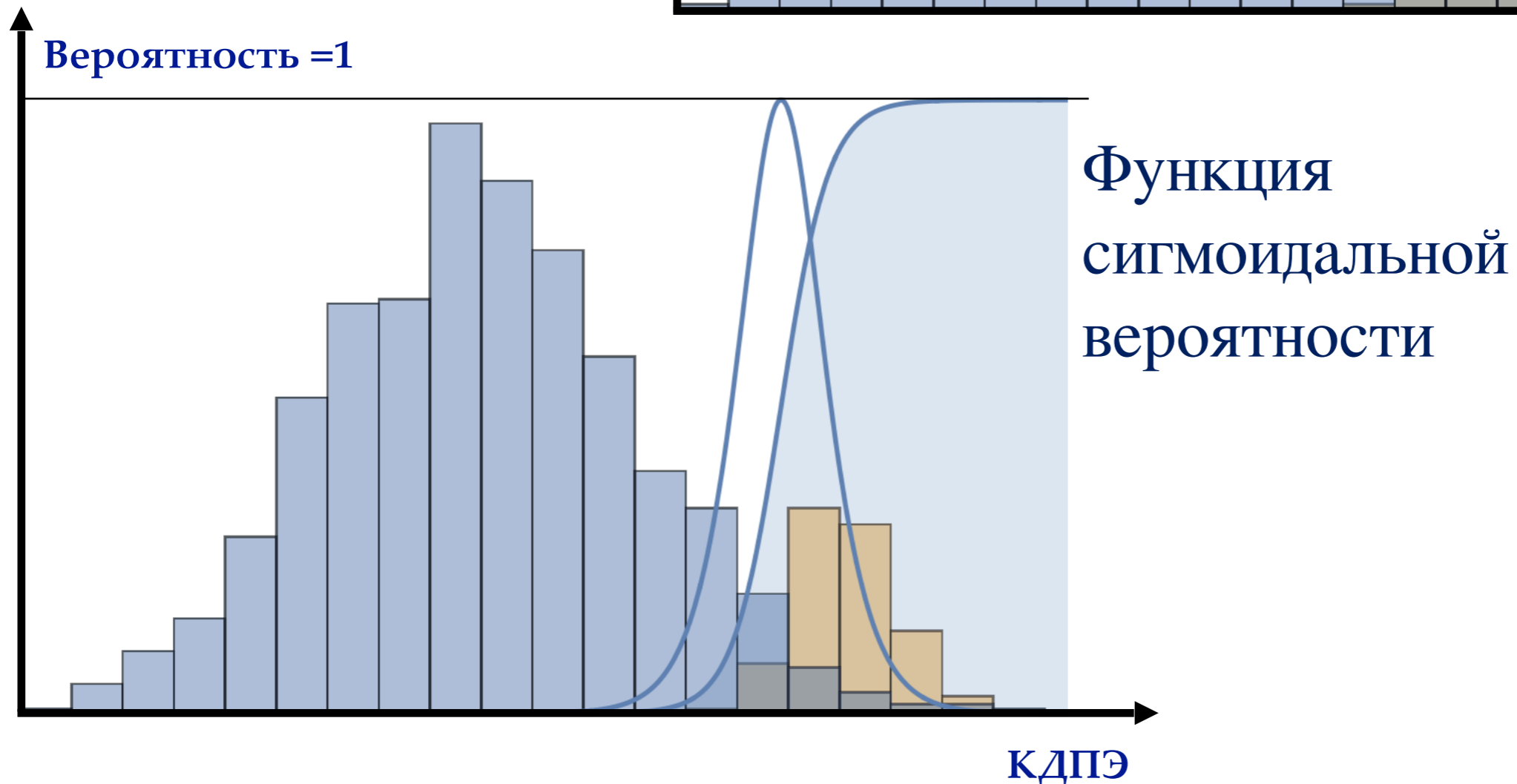
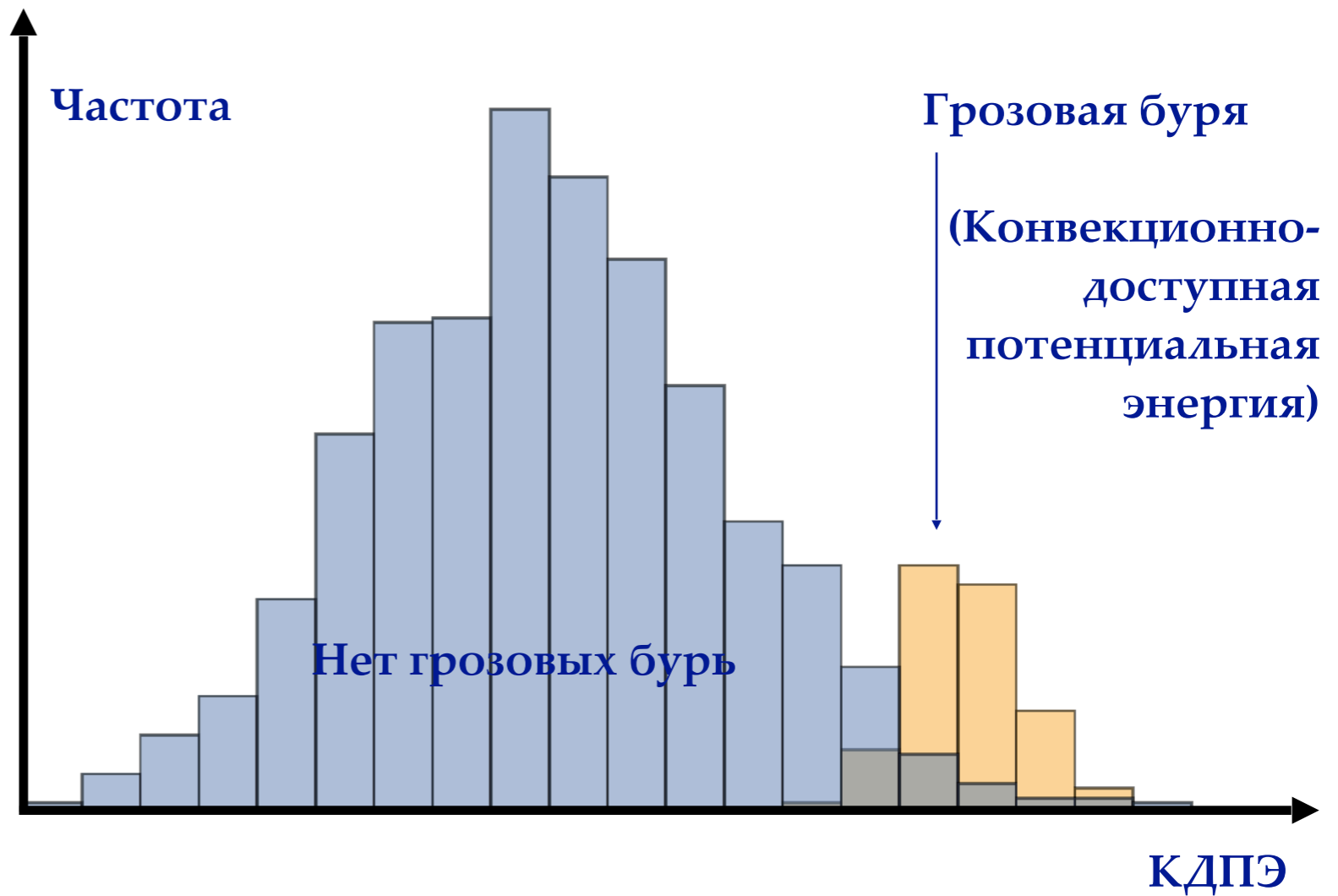
Базовые понятия логистической регрессии

Основа
неправильного
решения



ОСНОВЫ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ

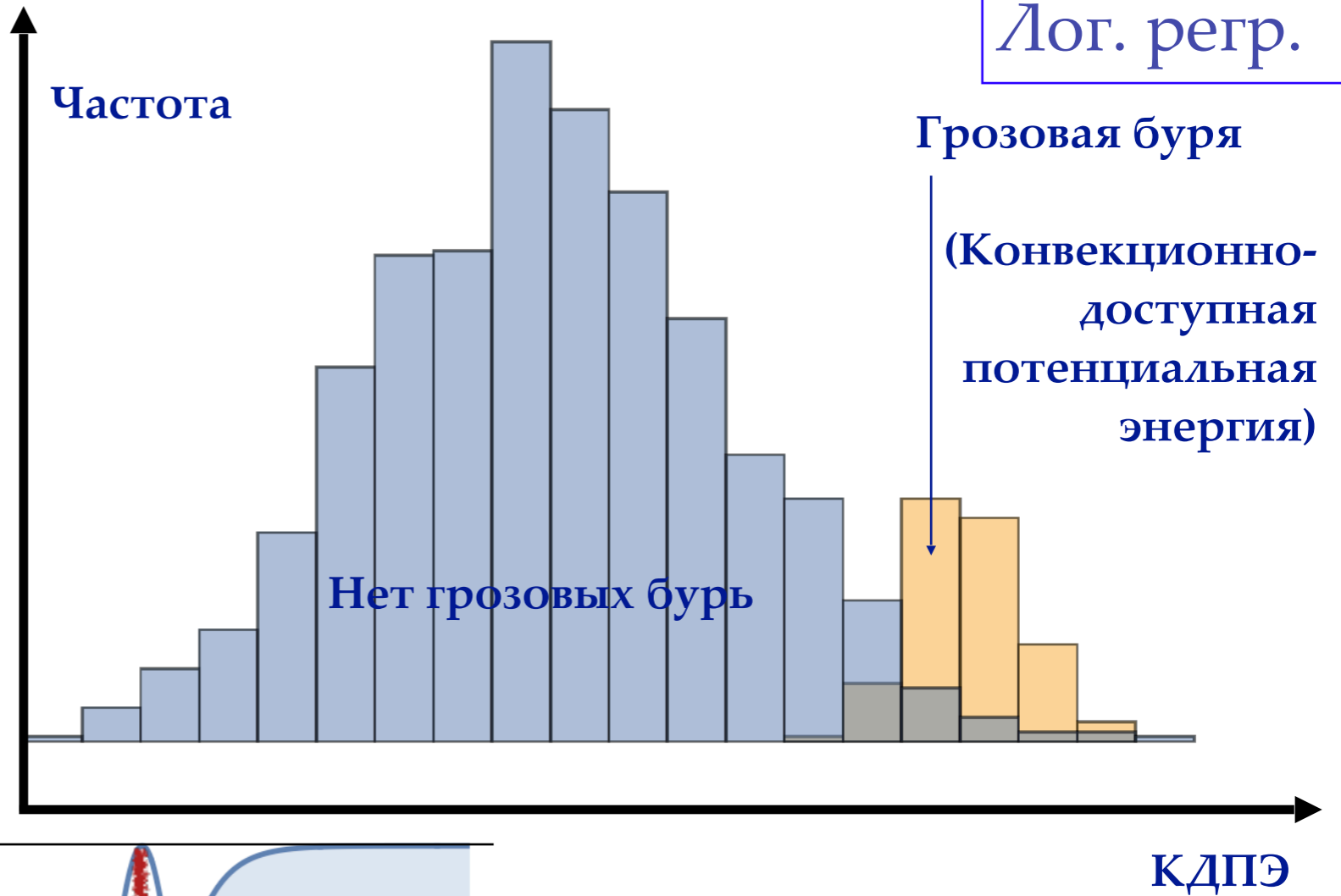
Основа
правильного
решения



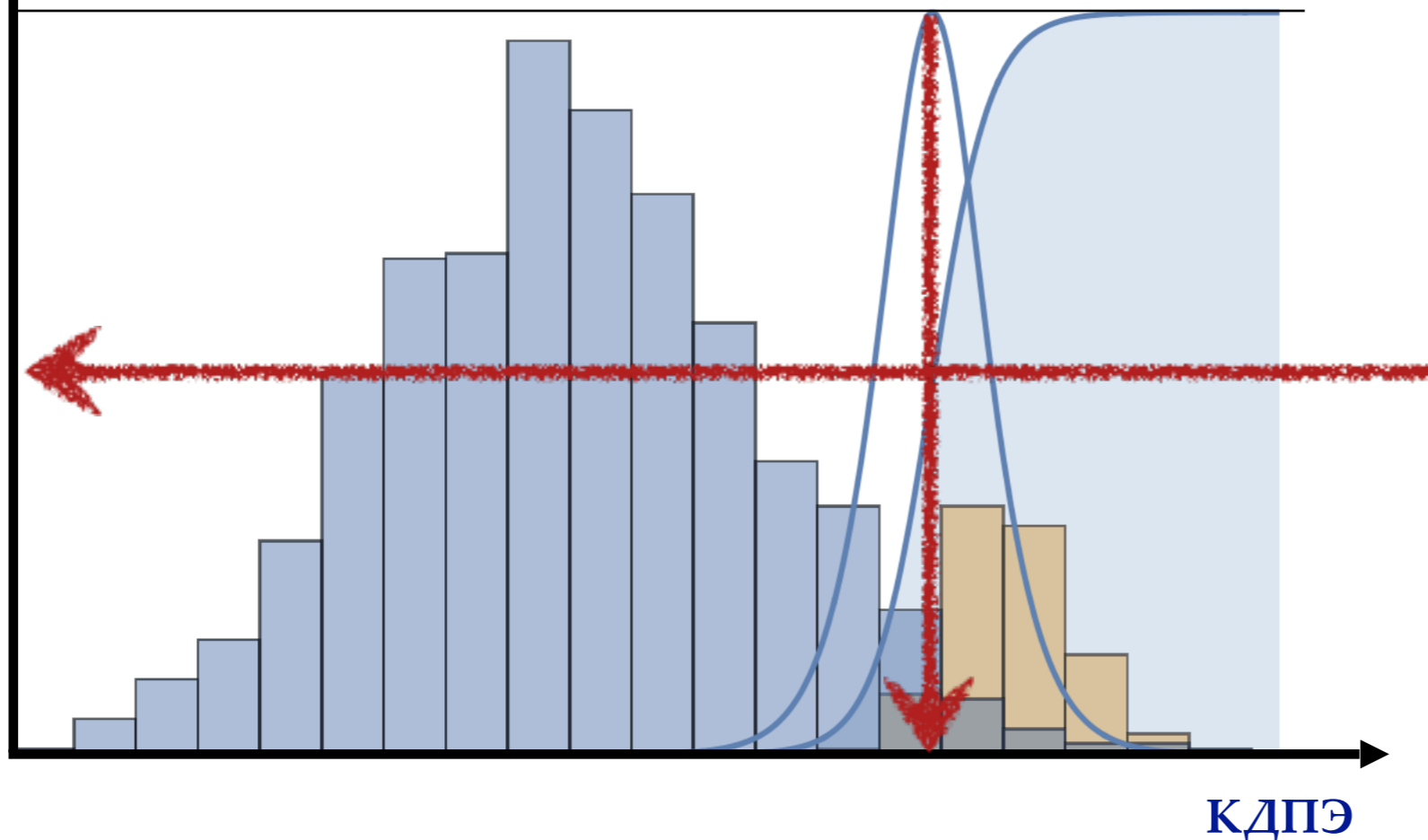
ОСНОВЫ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ

Основа
правильного
решения

Лог. регр.



Вероятность = 1



Пороги
вероятности и
интенсивности

Логистическая регрессия: байесовское предположение

Лог. регр.

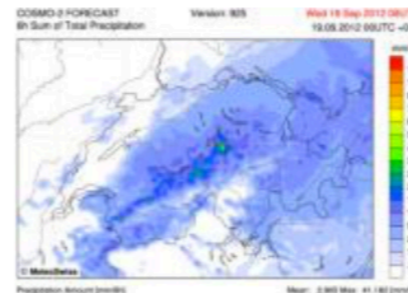
$$\ln \left[\frac{\mathbb{P}}{1 - \mathbb{P}} \right] = \mathbf{z} = \sum_i \alpha_i \mathbf{x}_i = \alpha_0 + \alpha_1 \mathbf{x}_1 + \dots + \alpha_N \mathbf{x}_N$$

$$\mathbb{P}[\mathbb{E} \leq \mathbf{z}] = \frac{1}{1 + e^{-\sum_i \alpha_i \mathbf{x}_i}} = \frac{1}{1 + e^{-\mathbf{z}}}$$

Функция
сигмоидальной
вероятности

Факторы
прогнозирования
отбираются в
соответствие с
методами, схожими с
СРМ

Archived COSMO-2 Forecast records



Historical observations



Predictors X

Predictand Y

Подбор подходящих
факторов

Оценка коэффициентов
логистической регрессии

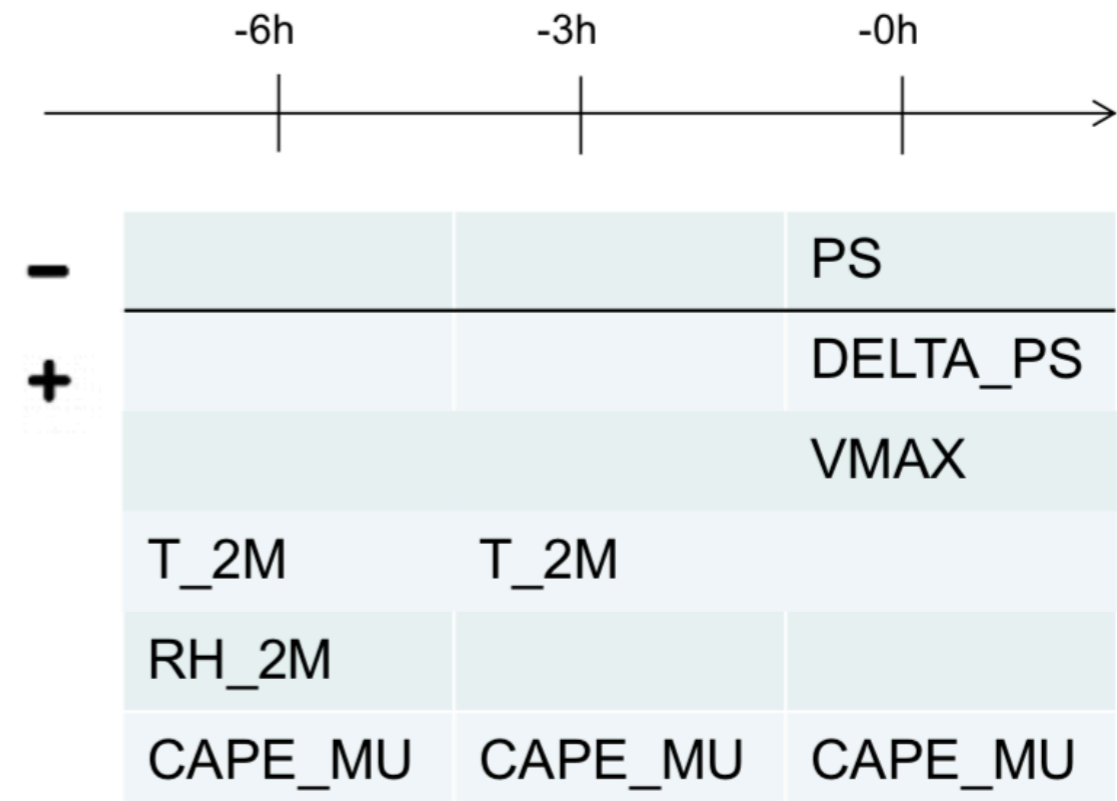
Прогноз грозовой бури в аэропортах Швейцарии

Лог. регр.

Набор первоначальных факторов прогнозирования

Predictors	
GLOB	Global solar radiation
CAPE_MU	Conv. avail. potential energy
VMAX_10M	Maximum wind speed at 10m
T_2M	Temperature at 2m
RH_2M	Relative humidity at 2m
PS	Surface pressure
DELTA_PS	Pressure difference

Выбранные факторы



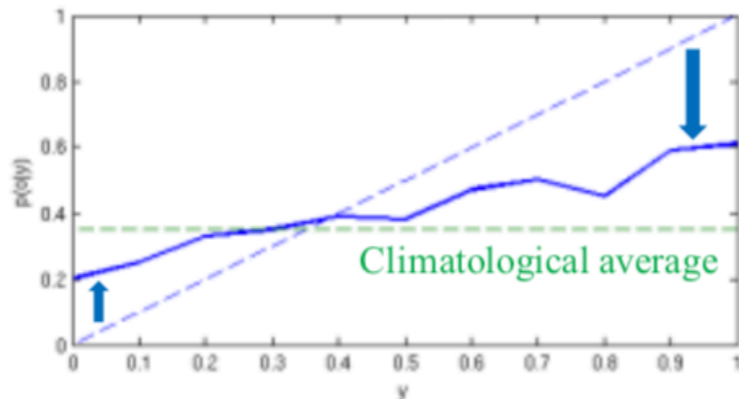
Initial set : $6 * 6 = 36$ predictors
 Selected subsets : 3-4 predictors

Прогноз грозовой бури в аэропортах Швейцарии

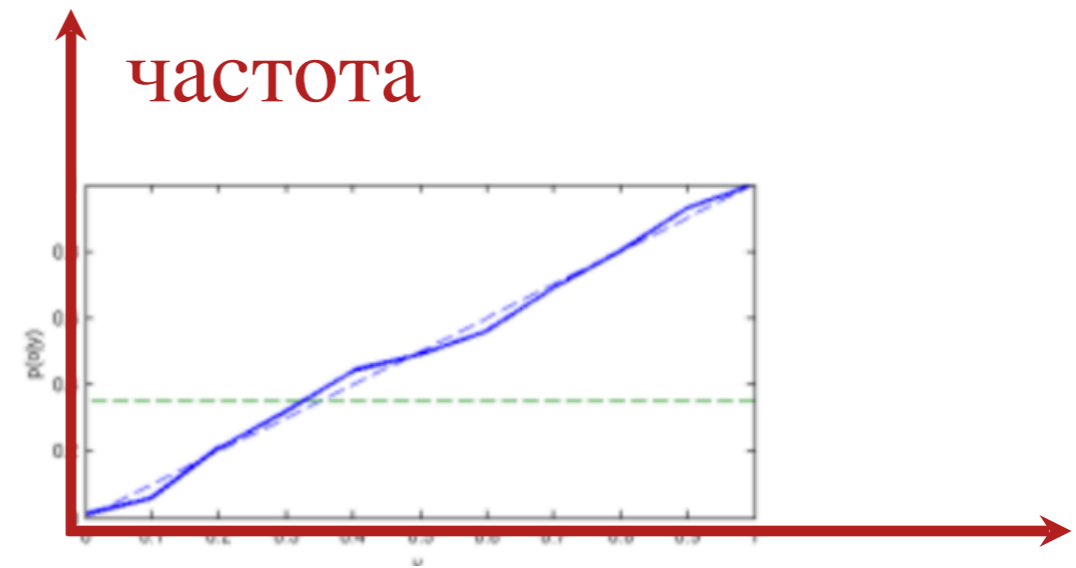
Оценка качества

Лог. регр.

Диаграмма надежности



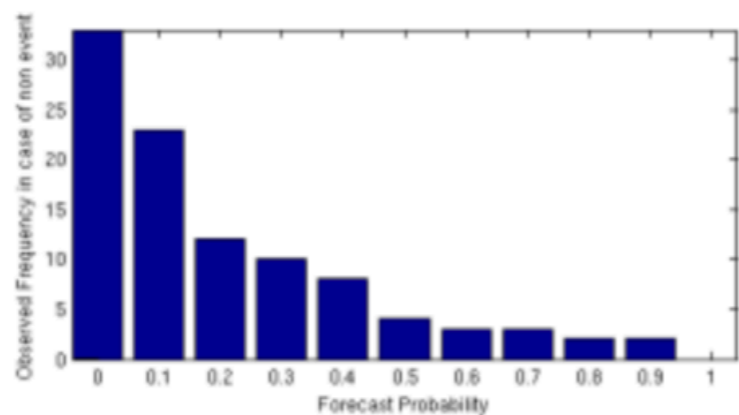
Плохое разрешение



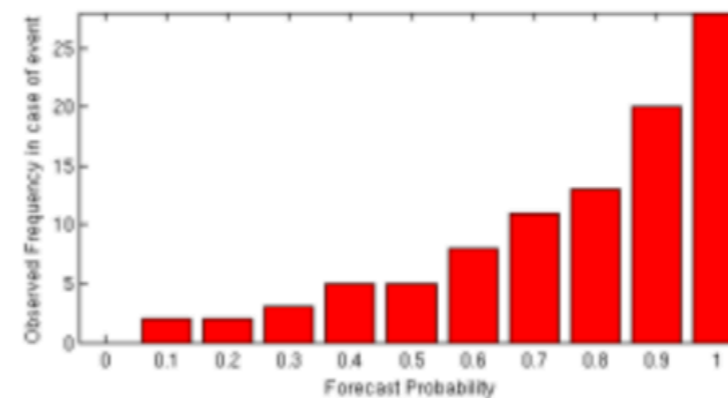
Хорошая калибровка

вероятн.

Доработанное распределение



Явления не происходят



Явления происходят

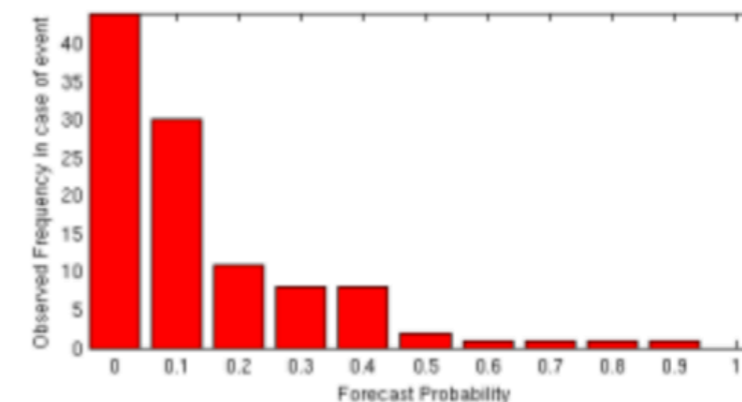
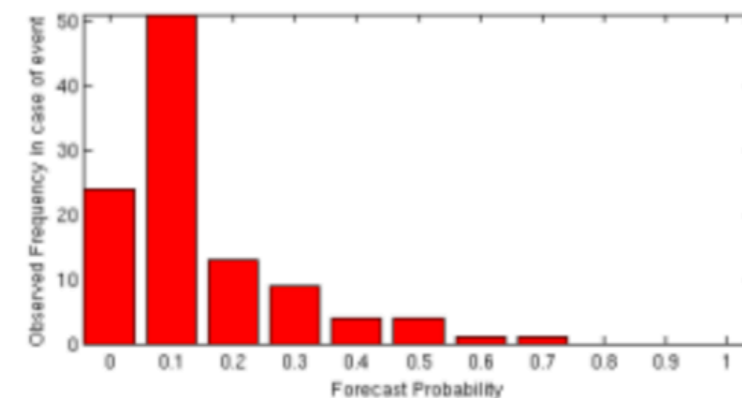
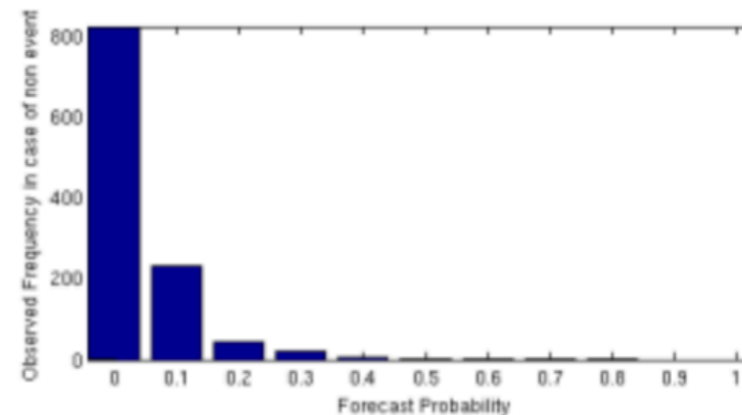
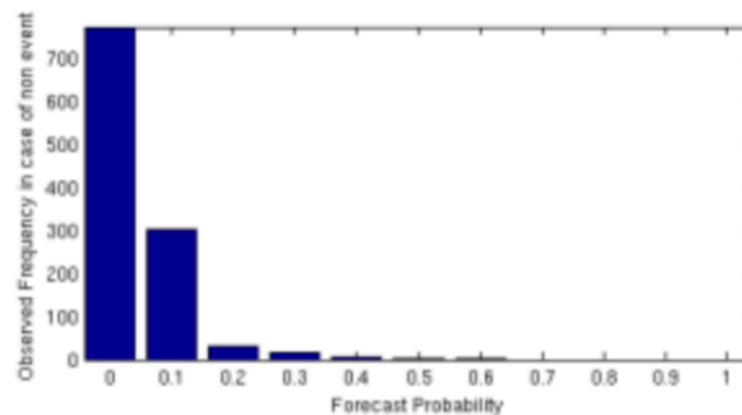
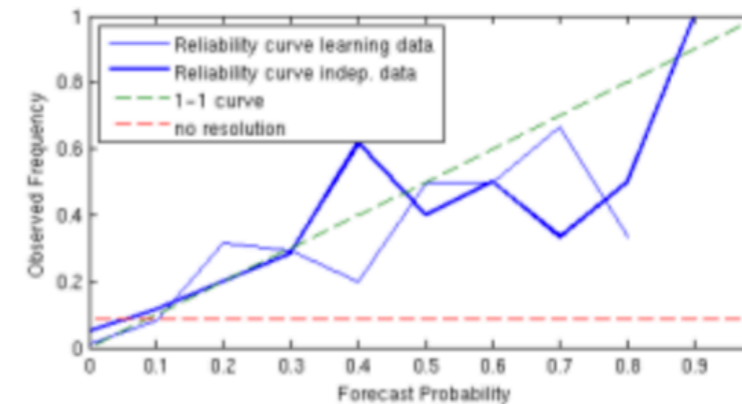
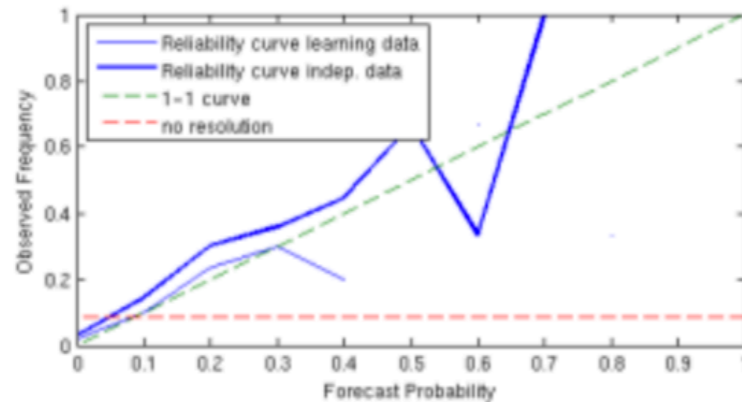
Прогноз грозовой бури в аэропортах Швейцарии

Оценка качества

Лог.
регр.

Послеобеденное время 14-17 UTC

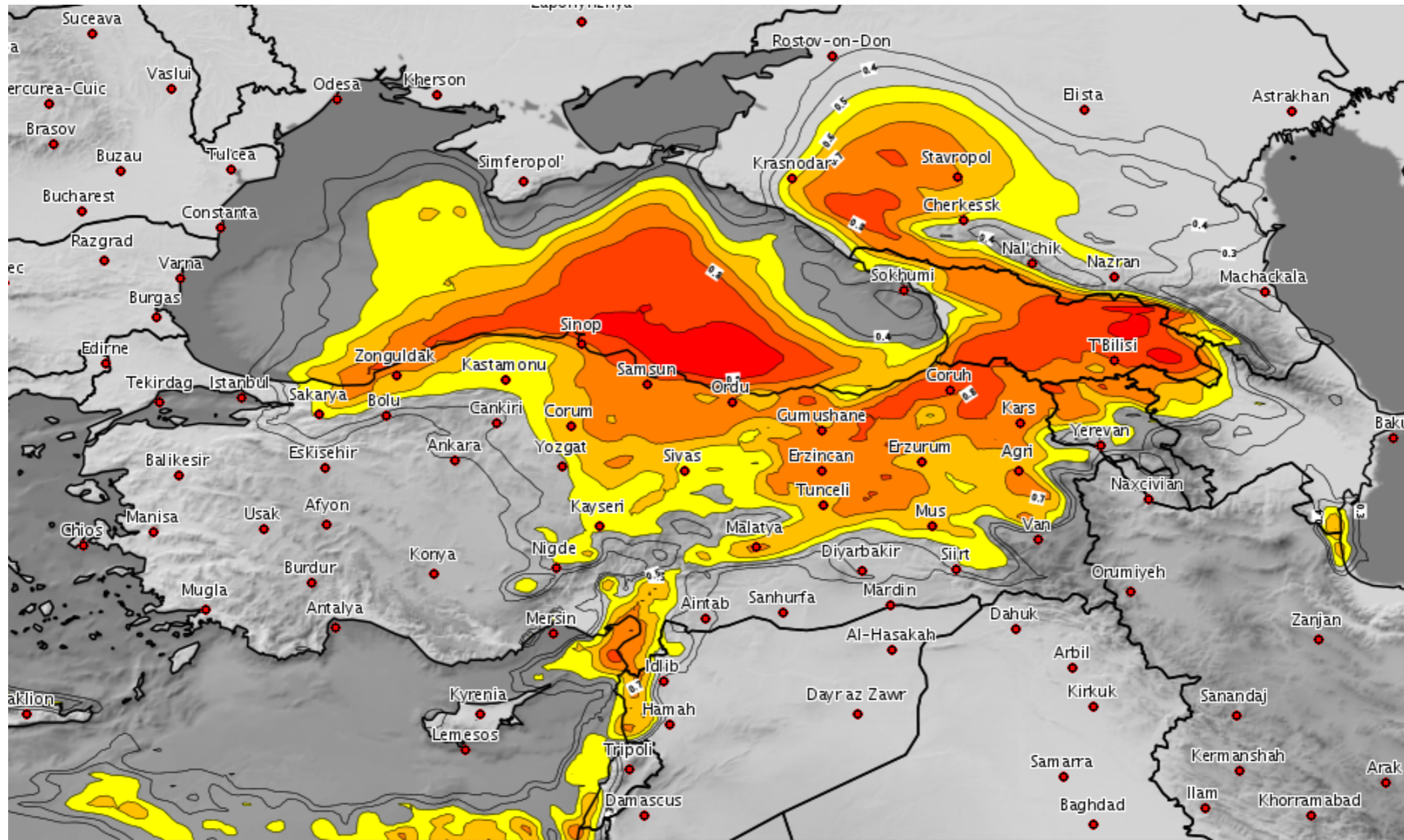
Вечер 18-21 UTC



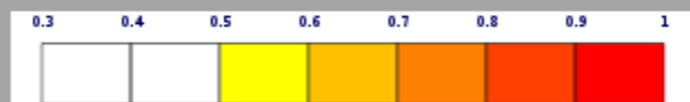
Заключения по Логистической регрессии

- Принимает во внимание местную климатологию
- Подходит для спорадических и/или редких явлений:
 - Тяжелые осадки, ветровые штормы, грозовые бури, ...
- Извлекает уроки и записывает характеристики прошлых событий
- Процесс обучения происходит «вне режима реального времени»
- Лог. регр. связывает внешние события, не включенные в модель, с параметрами модели.
Пример: Диапазон видимости на взлетно-посадочной полосе в аэропортах предоставляет вероятность происхождения редких/экстремальных явлений

Европейский Центр Среднесрочных Прогнозов Погоды (ECMWF) - EFI (Индекс экстремального прогноза)



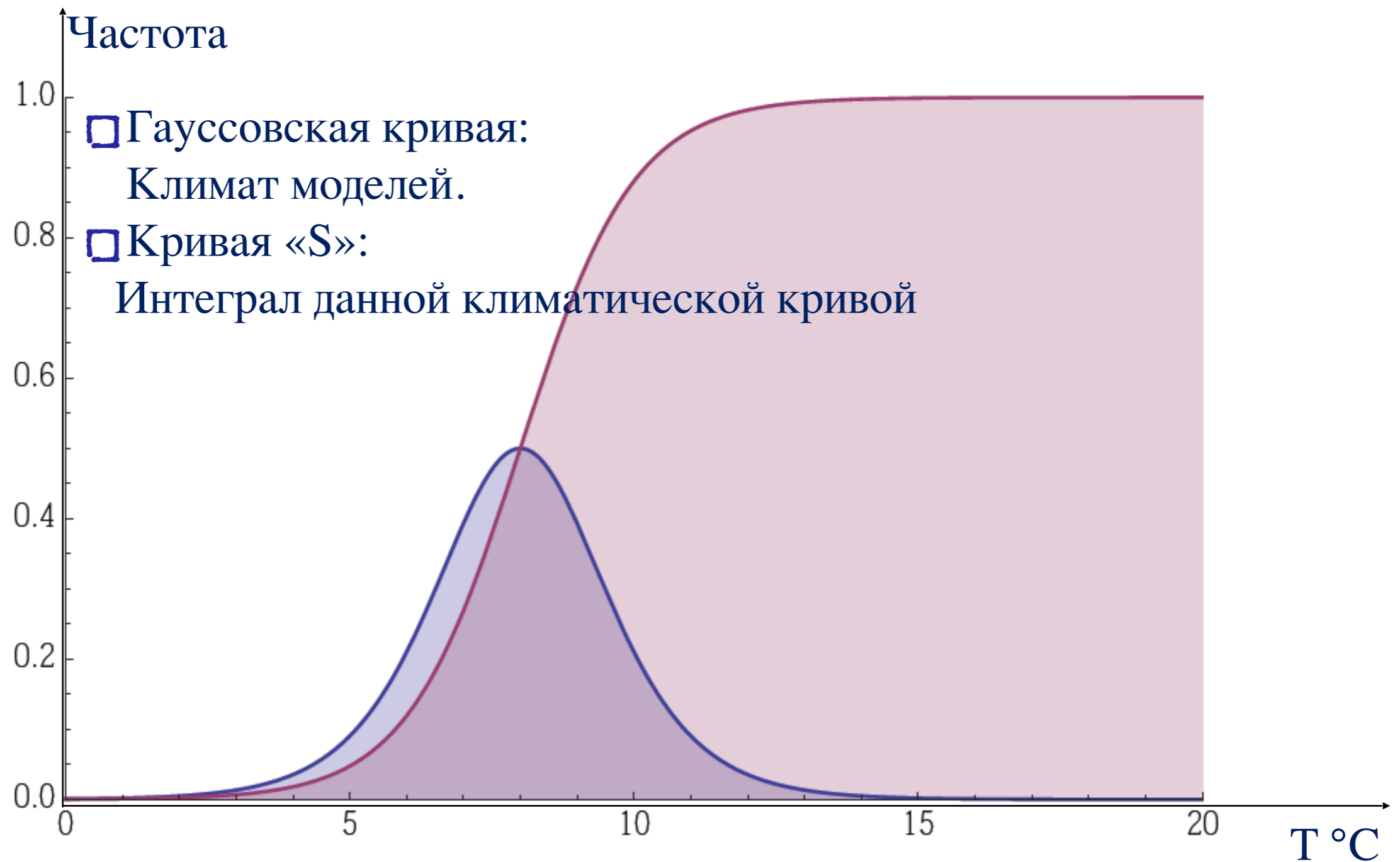
ECMWF 12Z 22 SEP 2016 - EPS FORECAST : Total Precipitation Extreme Forecast Index EFI -
24h During 23-24 SEP 2016 - BY: KHALID AL-OTAIBI



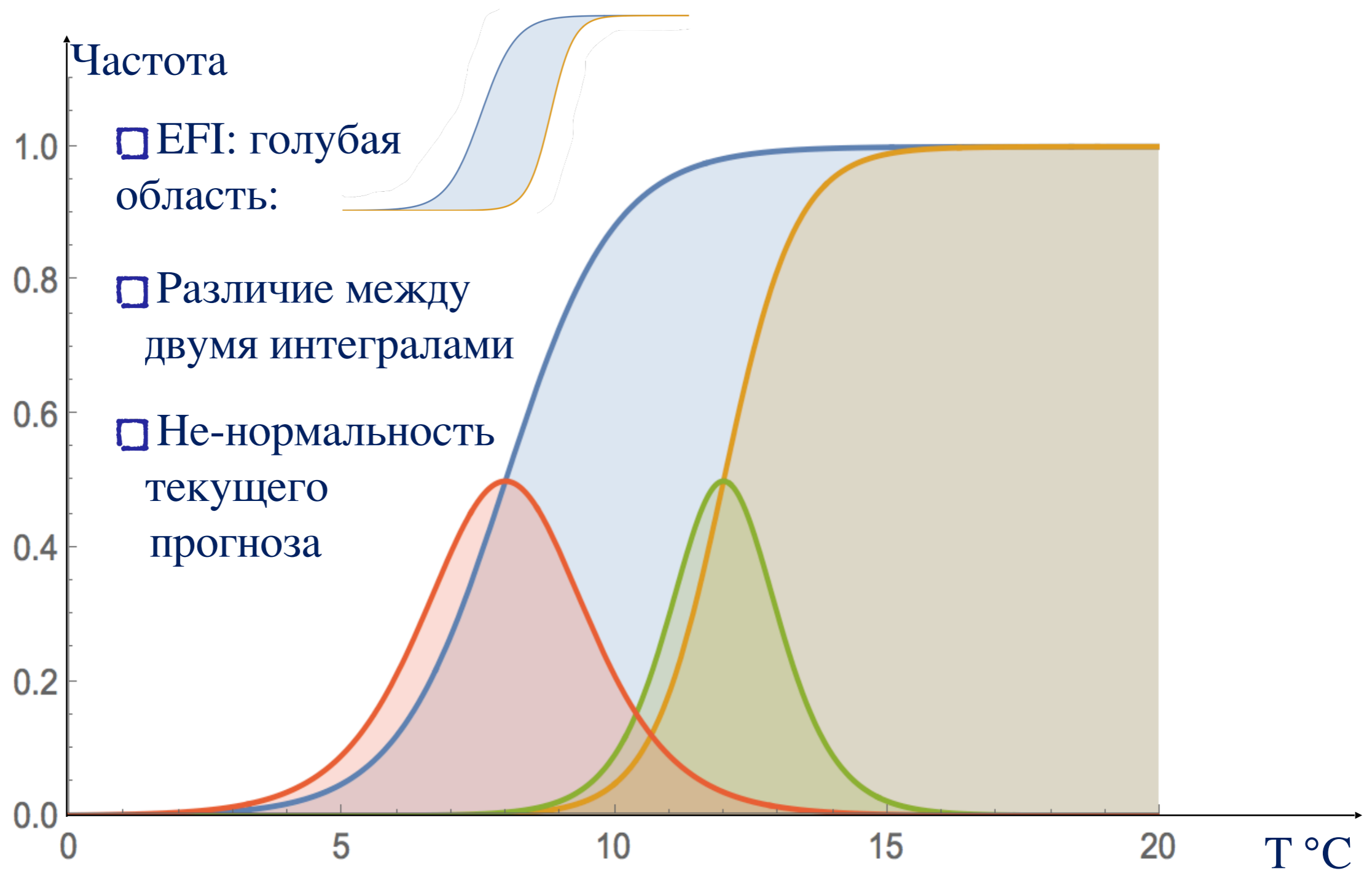
Европейский Центр Среднесрочных Прогнозов Погоды (ECMWF) - EFI (Индекс экстремального прогноза)



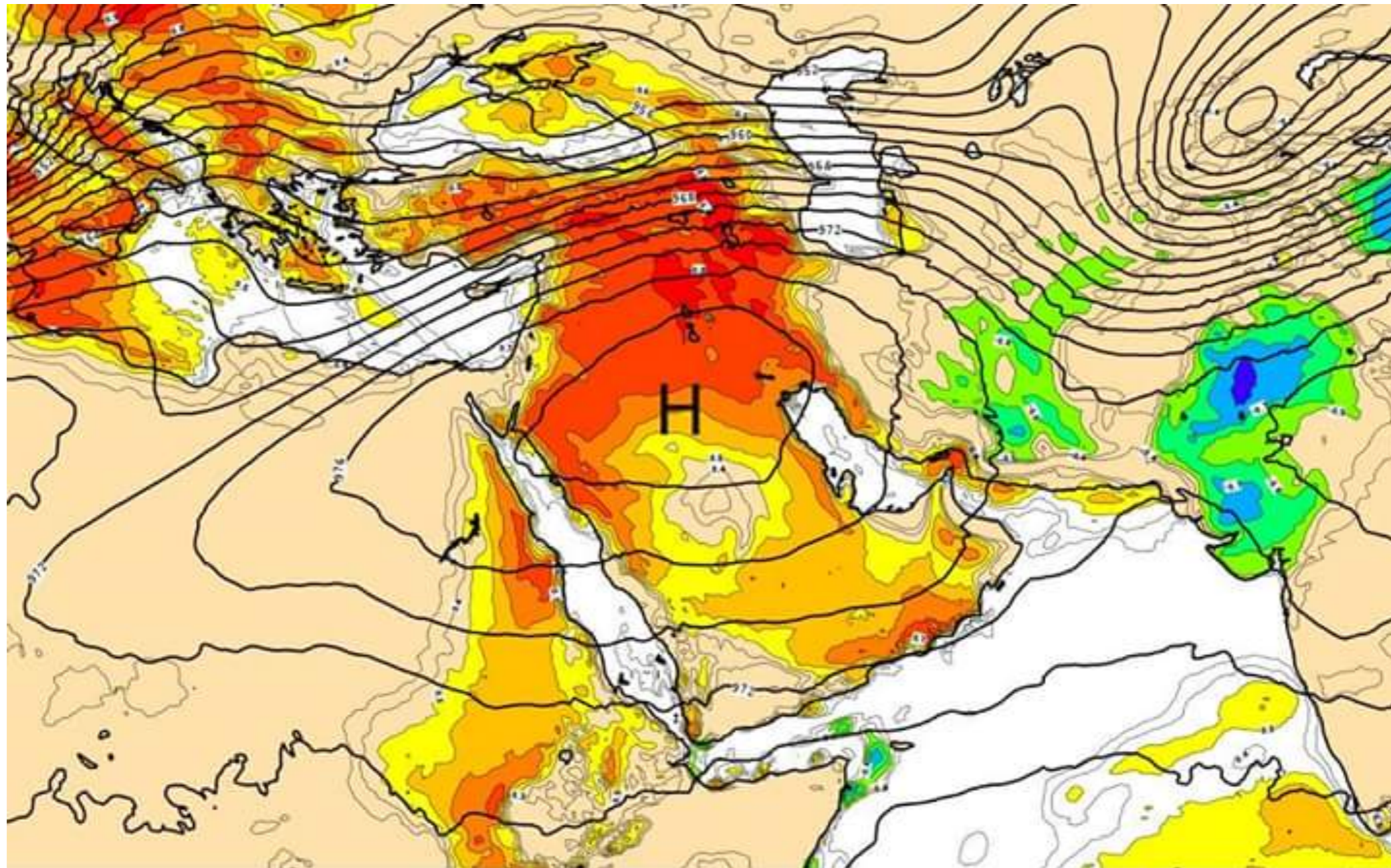
Европейский Центр Среднесрочных Прогнозов Погоды (ECMWF) - EFI (Индекс экстремального прогноза)



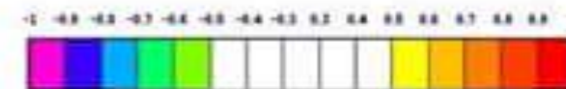
Европейский Центр Среднесрочных Прогнозов Погоды (ECMWF) - EFI (Индекс экстремального прогноза)



Европейский Центр Среднесрочных Прогнозов Погоды (ECMWF) - EFI (Индекс экстремального прогноза)



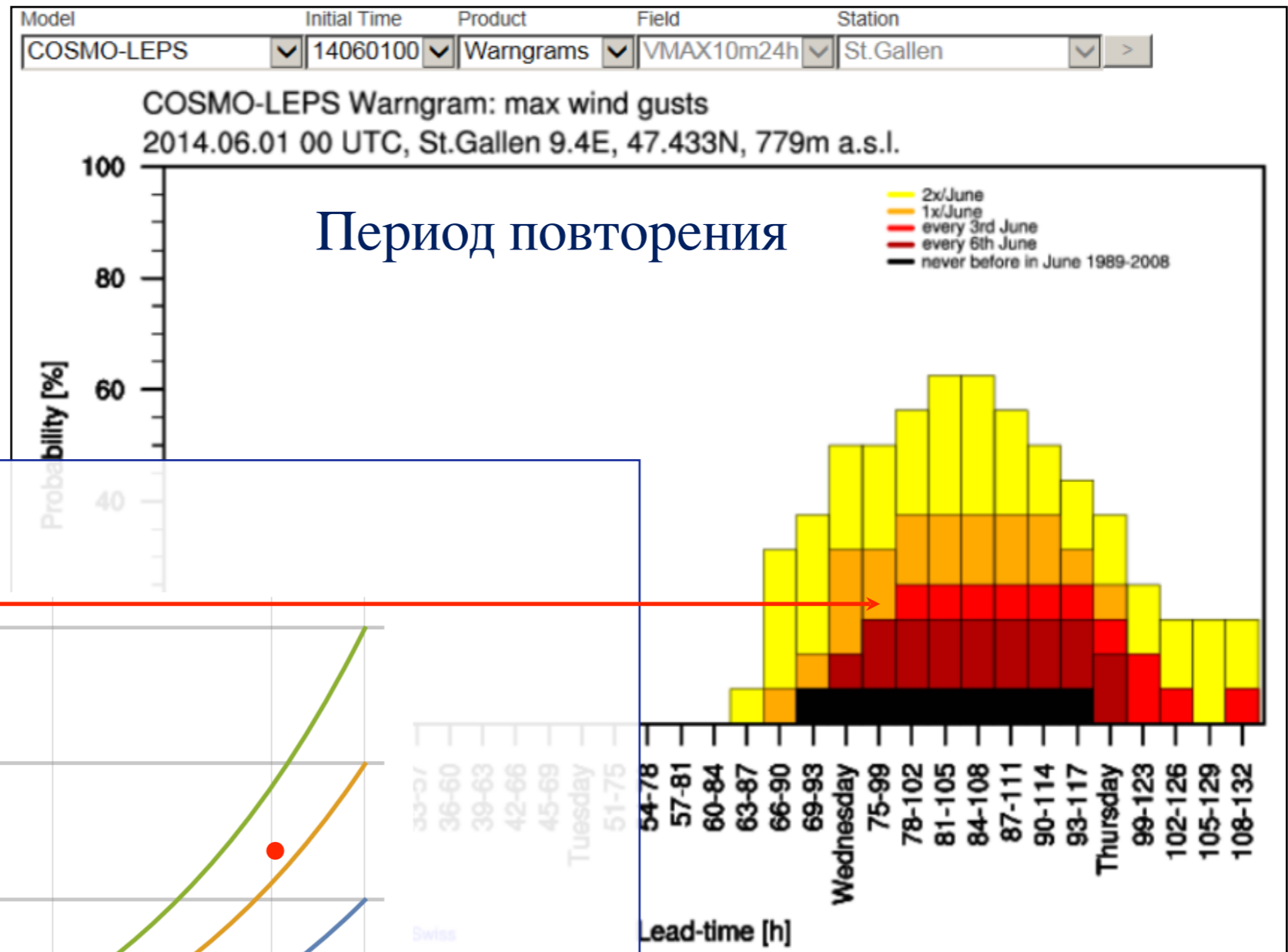
ECMWF 00Z 25 Jun 2017 - HRES+EPS EFI FORECAST :
300-hPa geopotential height , During 12Z 28 Jun 2017
Maximum 2m Temperature Extreme Forecast Index EFI - 24h During 28-29 Jun 2017
BY: KHALID AL-OTAIBI



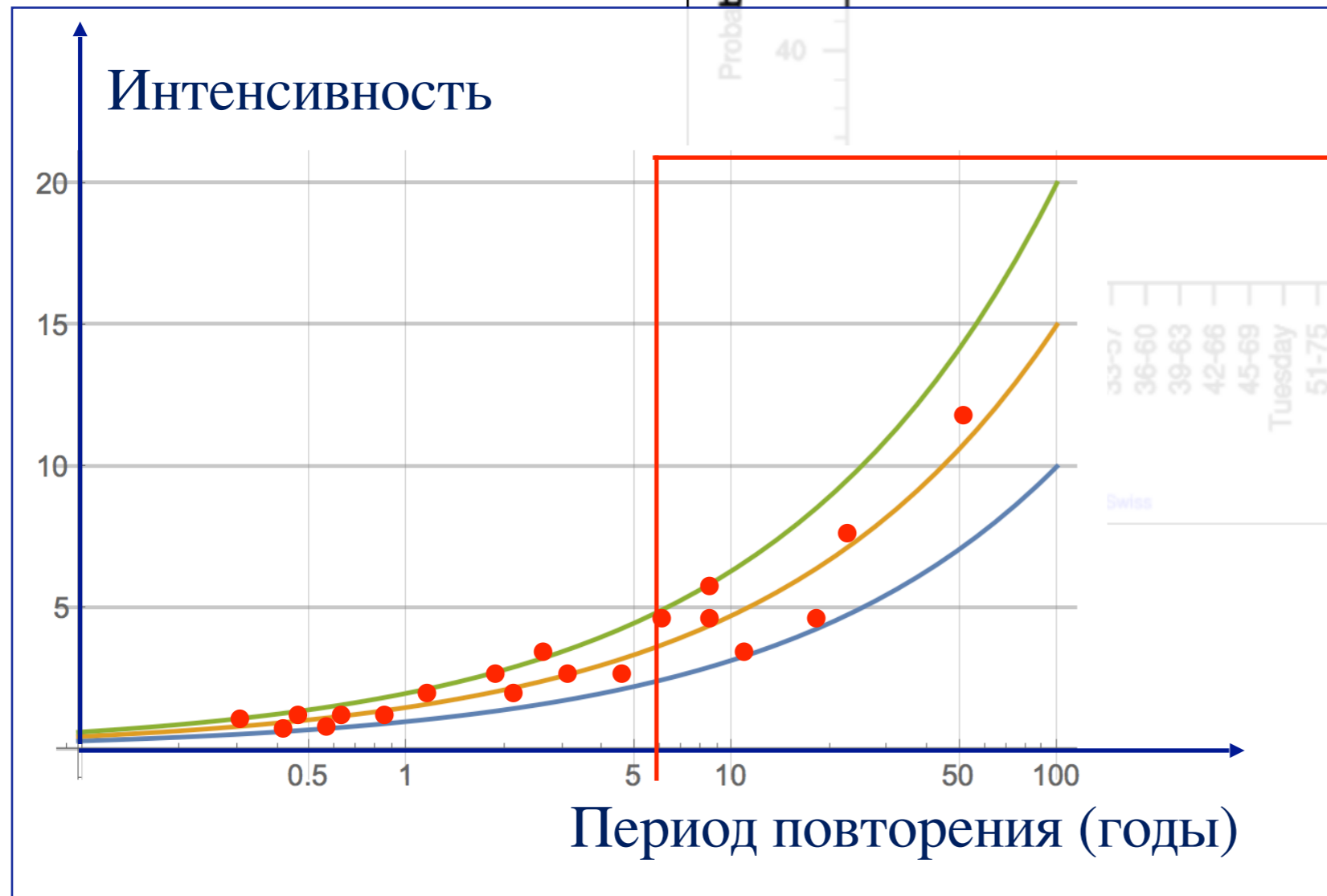
Заключения по EFI (Индекс экстремального прогноза)

- EFI оценивает не-нормальность ситуации прогноза погоды.
- Является интегральной единицей измерения разницы между распределением ансамблевого прогноза (ENS) и распределением модельного климата (M-climate). EFI принимает значения от -1 до +1.
- Опыт предполагает, что масштаб EFI
 - 0,5 – 0,8 (независимо от знака) может рассматриваться как «необычная» погода
 - свыше 0,8 означает «очень необычную» погоду.
- Несмотря на то, что более высокие значения EFI указывают на то, что экстремальное явление более вероятно, EFI, как таковой, не представляет вероятности.

Период повторения явлений и графические оповещения (COSMO)



Период повторения



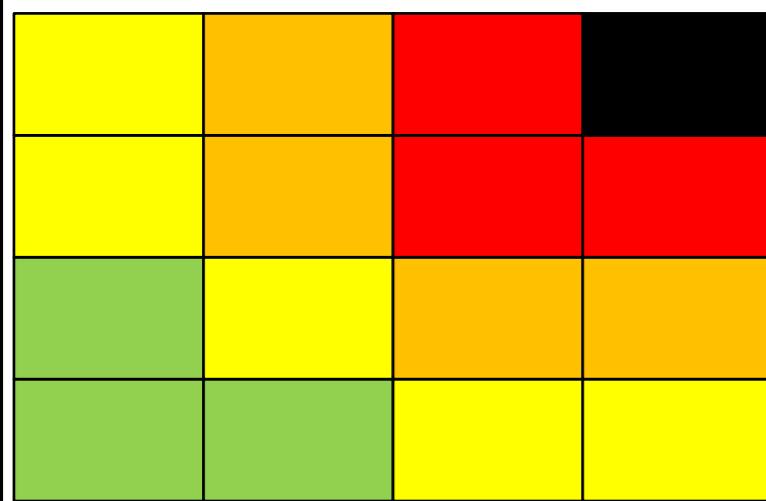
Период повторения явлений и графические оповещения (COSMO)



решение

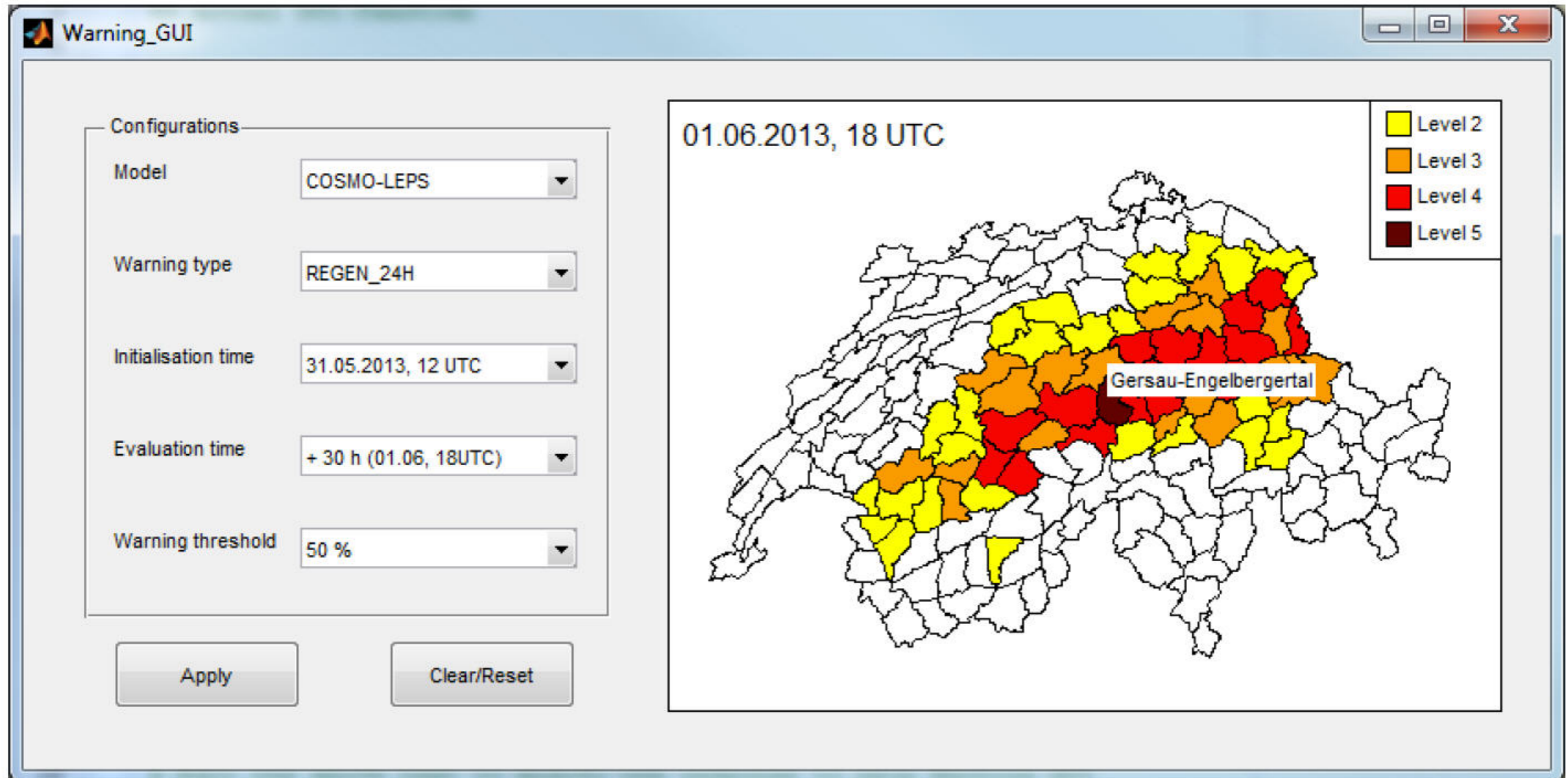
вероятность

Воздействие




Интенсивность

Решение о предупреждении/оповещении (на основе COSMO-LEPS)




Информирование населения


Wetterbericht




Gefahren




Naturgefahrenbulletin




Line Prognosen 


Niederschlag (Radar)



Temperatur



Wind

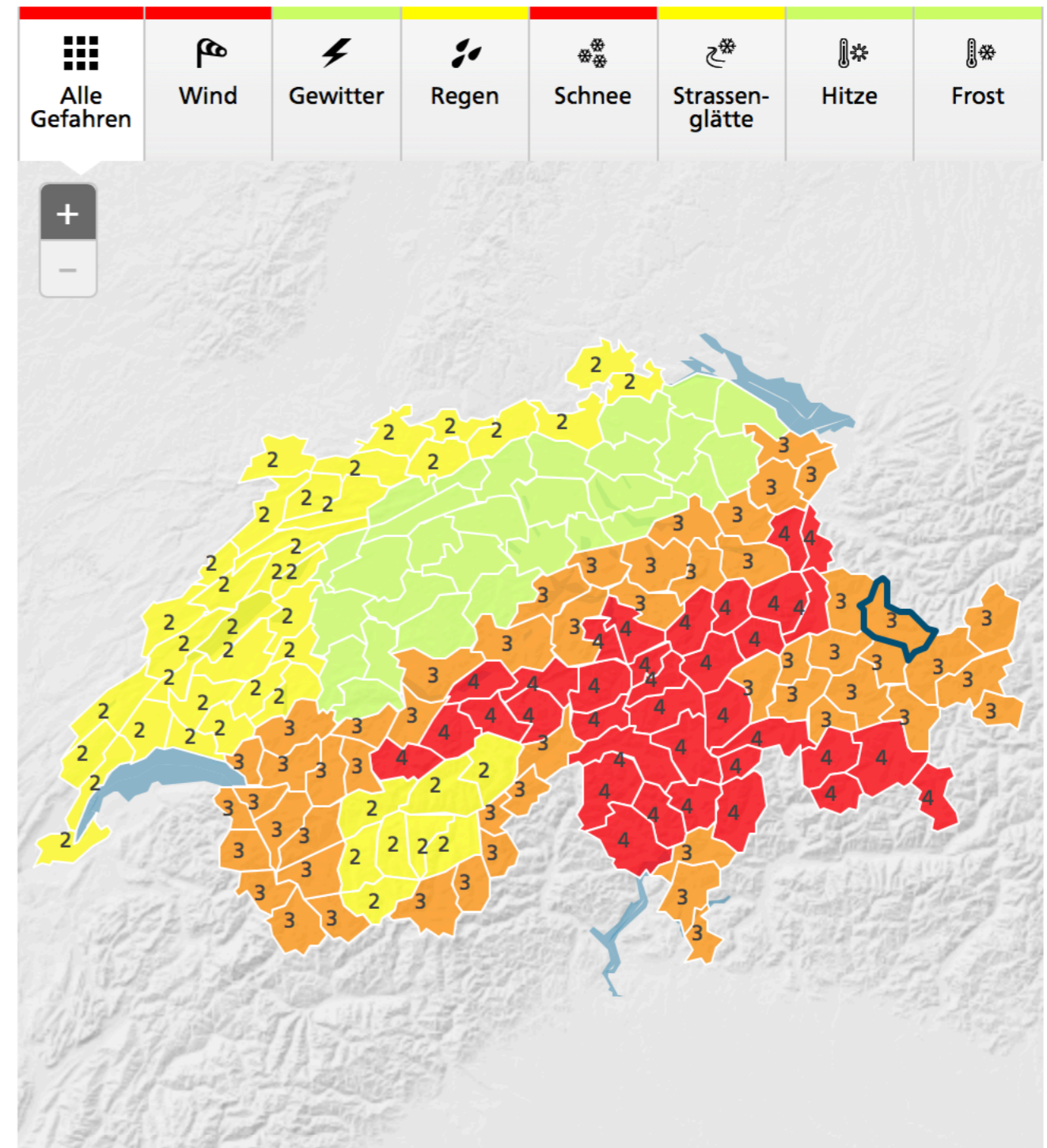


Bewölkung

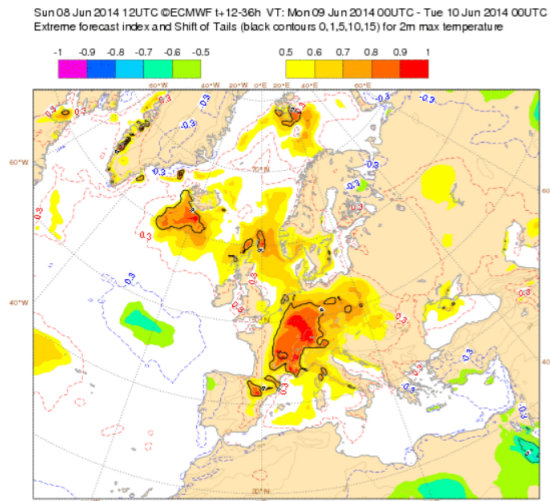
Gefahren

In den nächsten 24 Stunden

► **Legende**



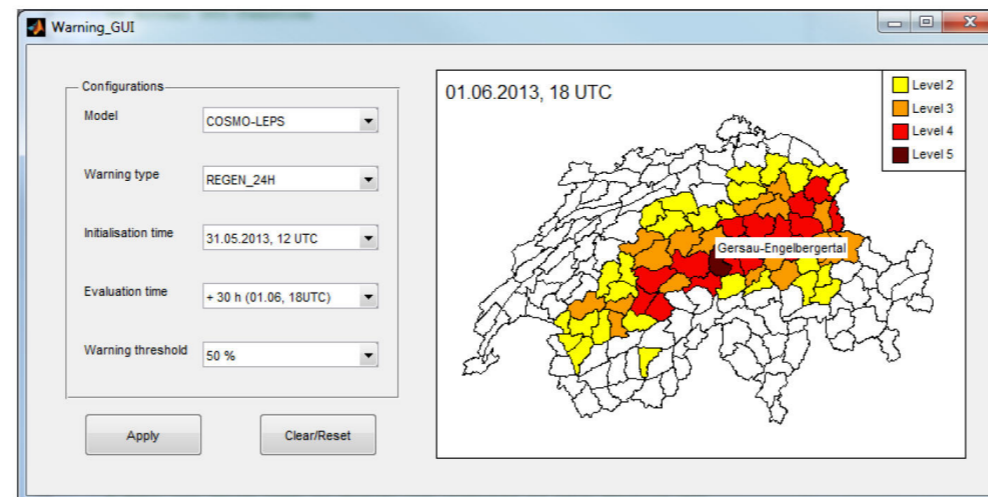
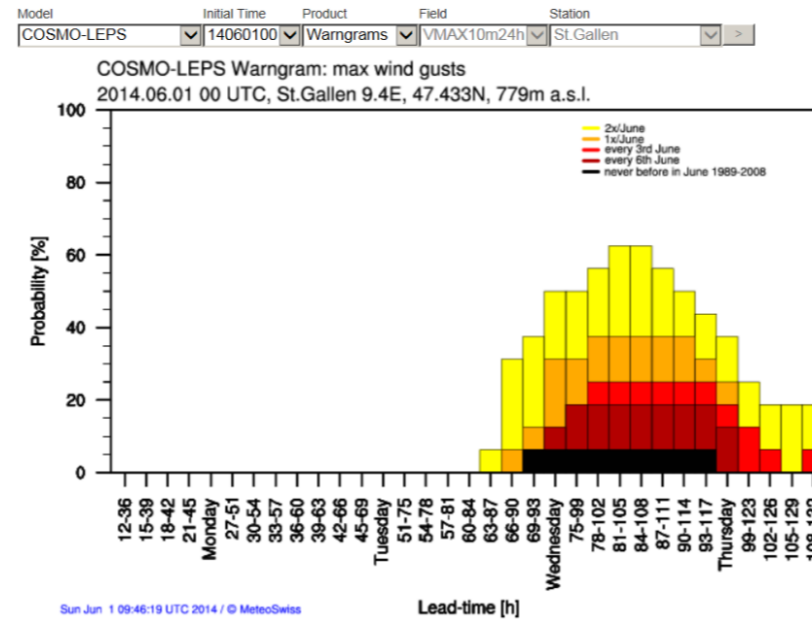
Заключение: Последовательность решений



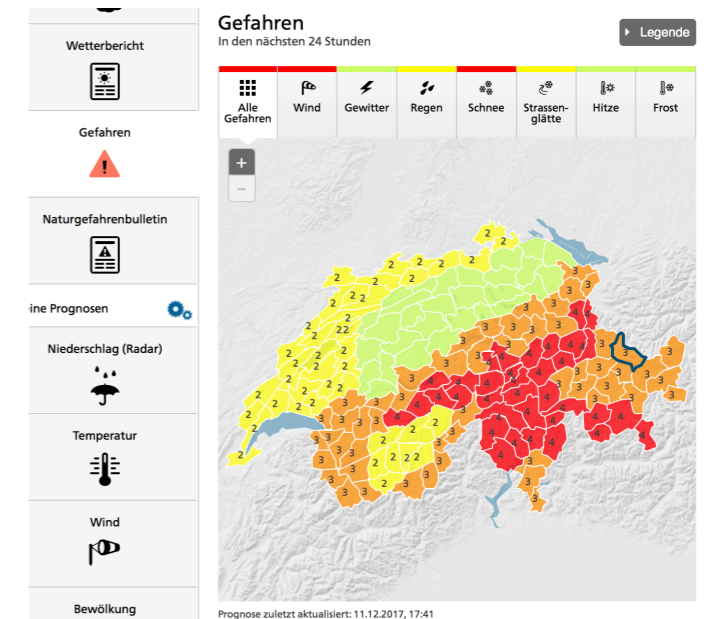
EFI



Служба
погоды



Органы власти



СМИ / Население

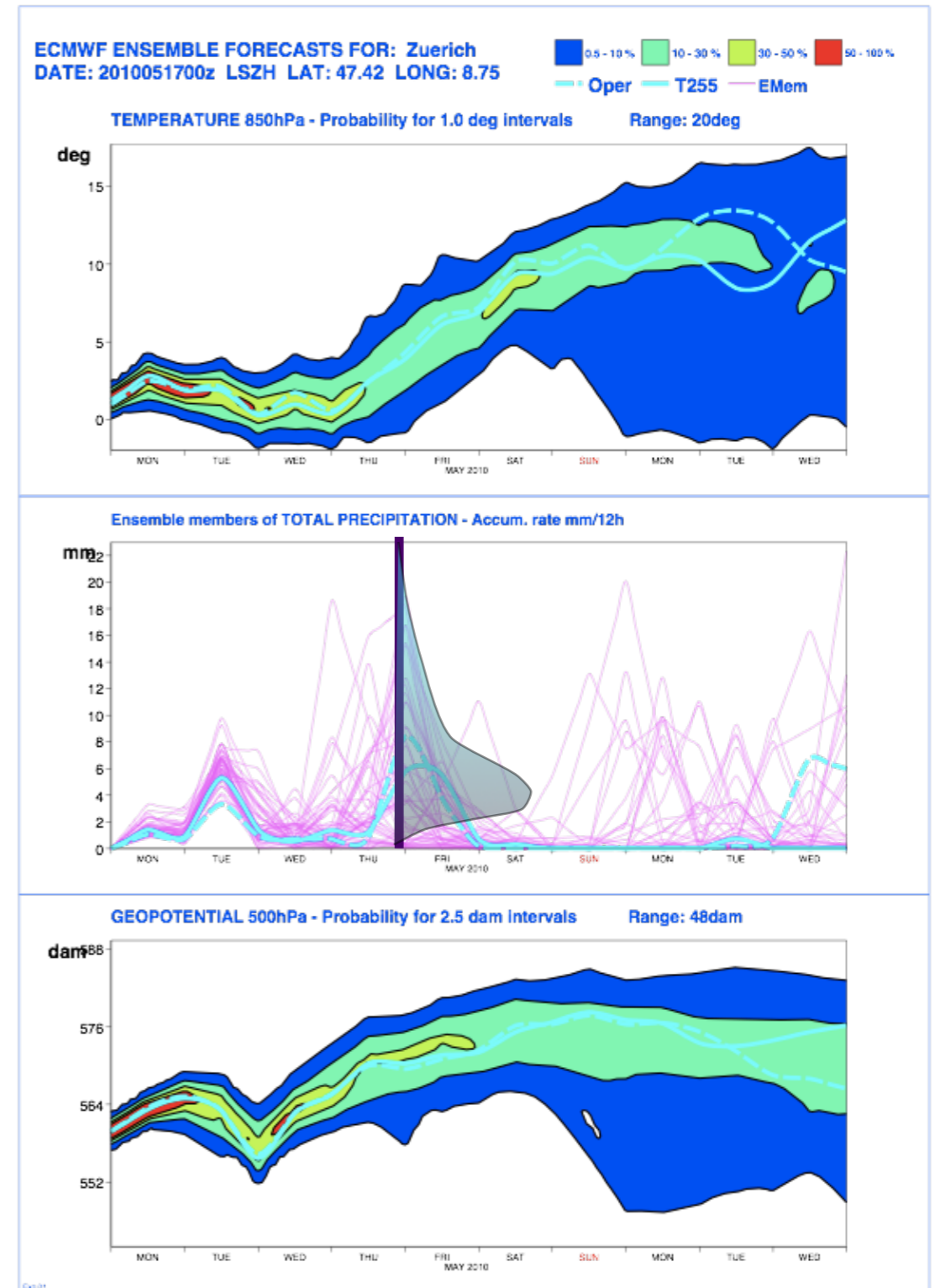
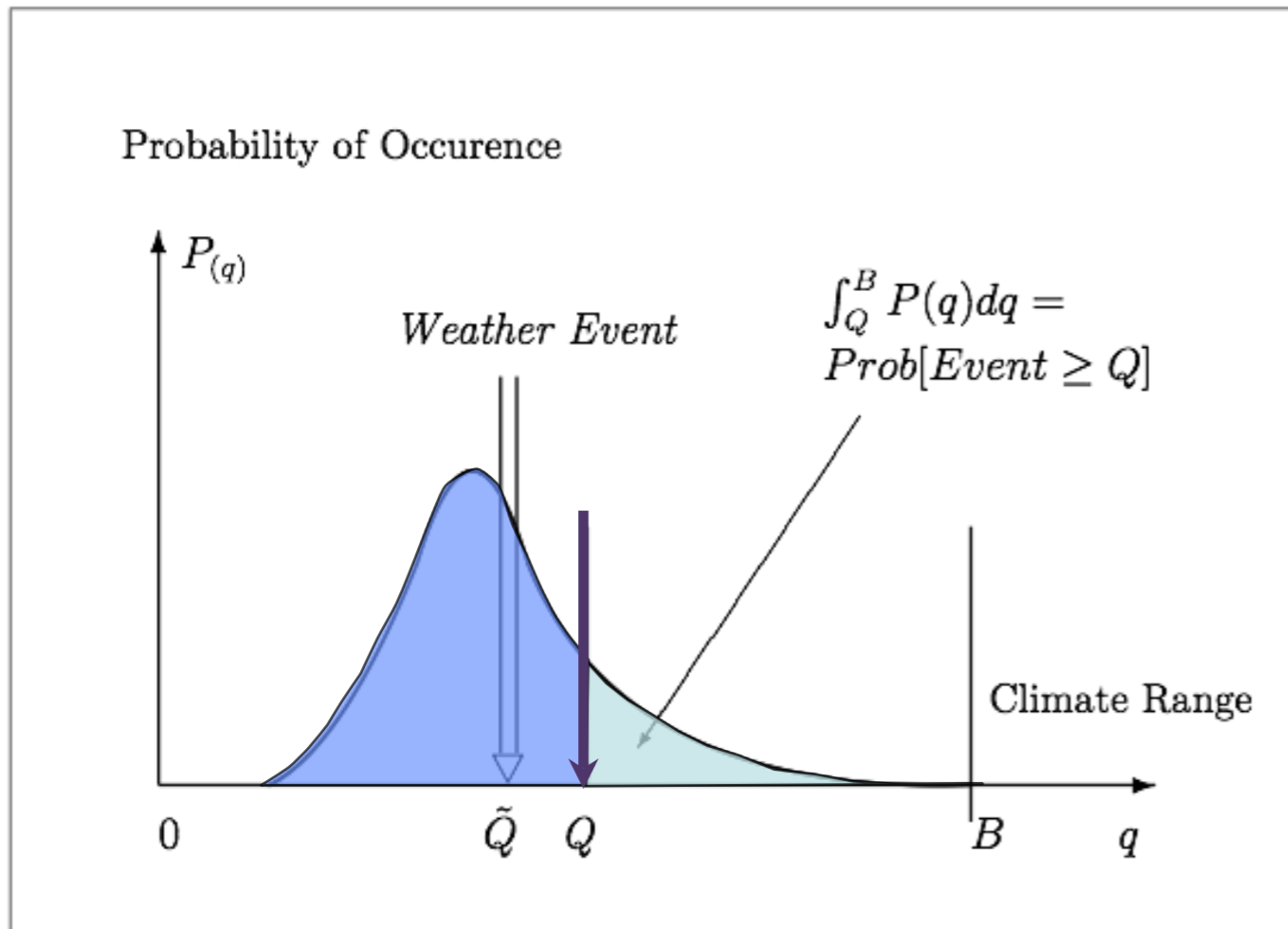


Дни -8 .. -6

День-3

День-1

Проверка ансамблевых (вероятностных) прогнозов



Проверка вероятностных прогнозов

Относительная характеристика операций (ОХО)

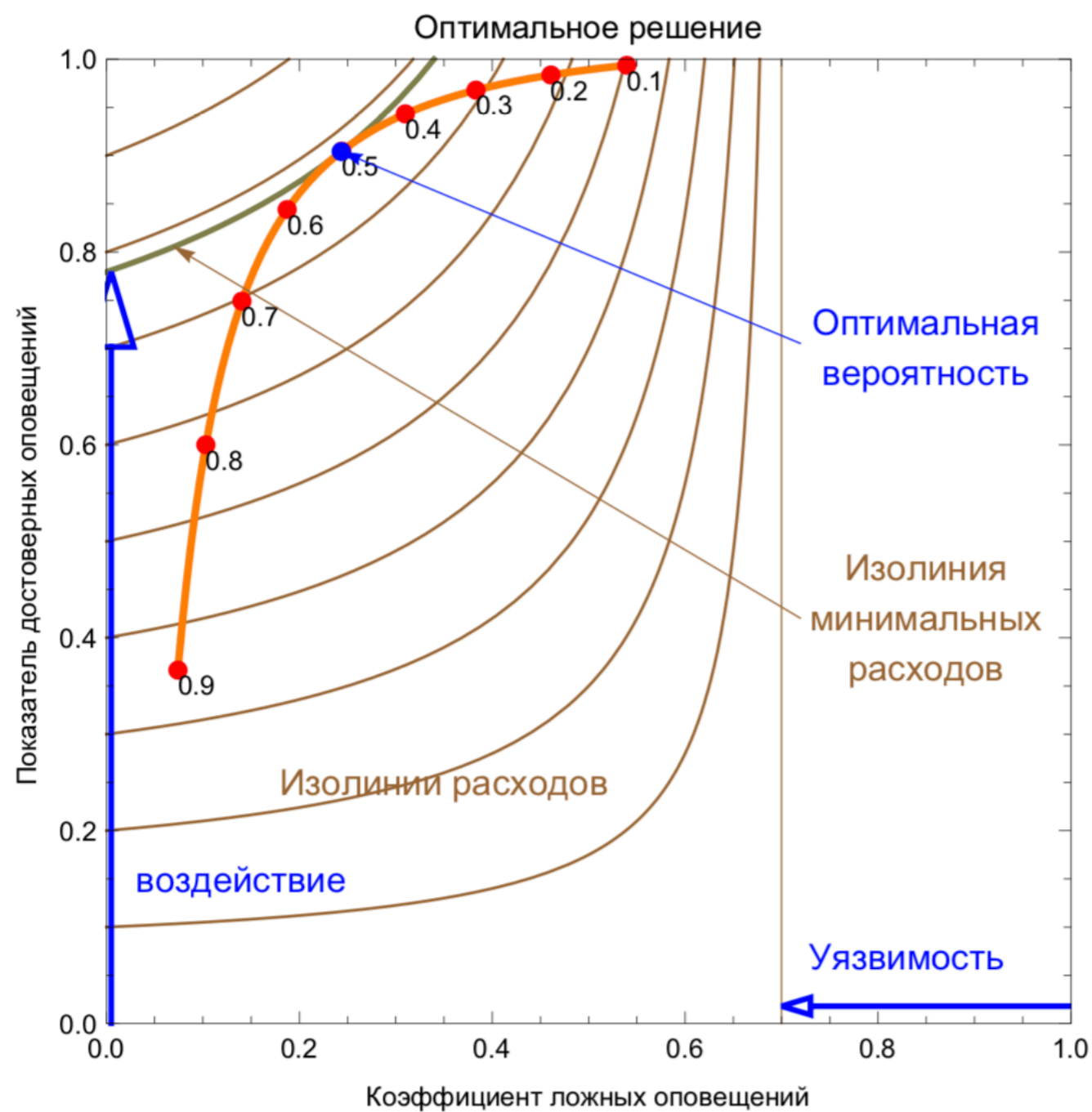
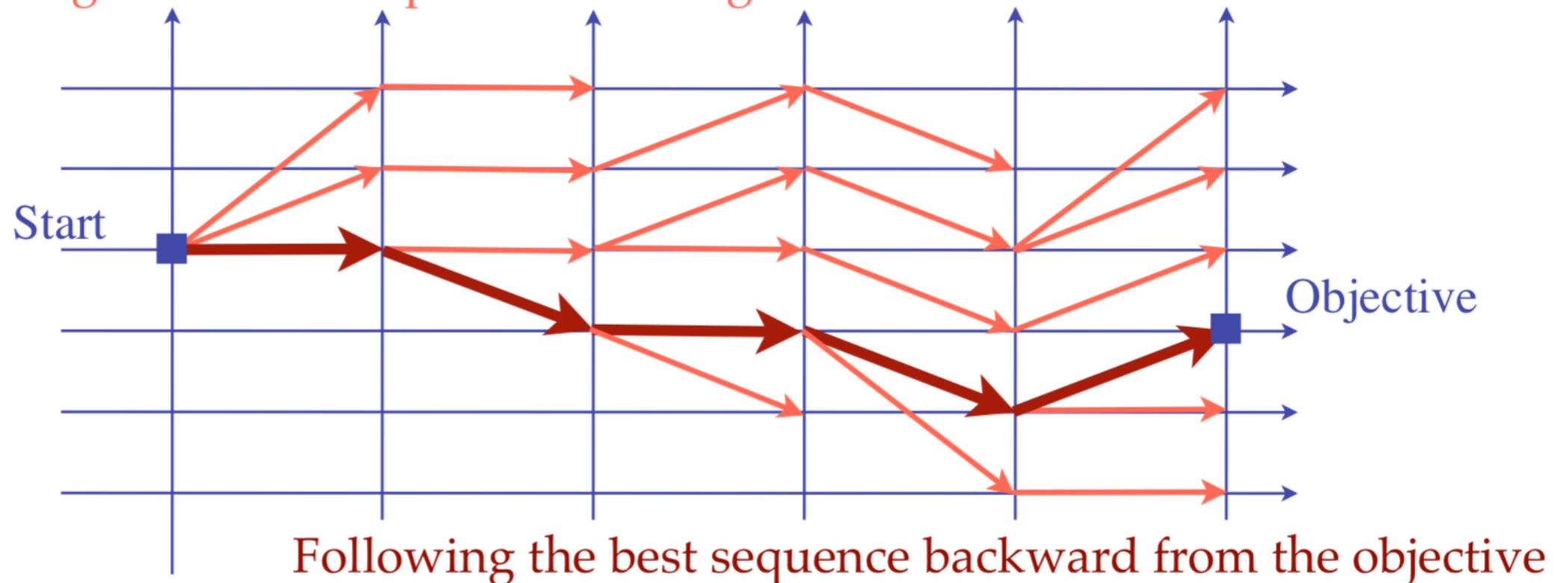


Рисунок 3. Оптимальное решение и соответствующий выбор предельного значения для вероятного оповещения: 51%. Оранжевая кривая: профиль эффективности системы метеорологического оповещения. Коричневая кривая: изолиния расходов в связи с уязвимостью пользователя к погодным явлениям. Все критические параметры находятся в оптимуме и просматриваются на данной диаграмме.

Решение о последовательности действий

Sequential Decision: Dynamical Programming

Seeking forward for options: building the tree

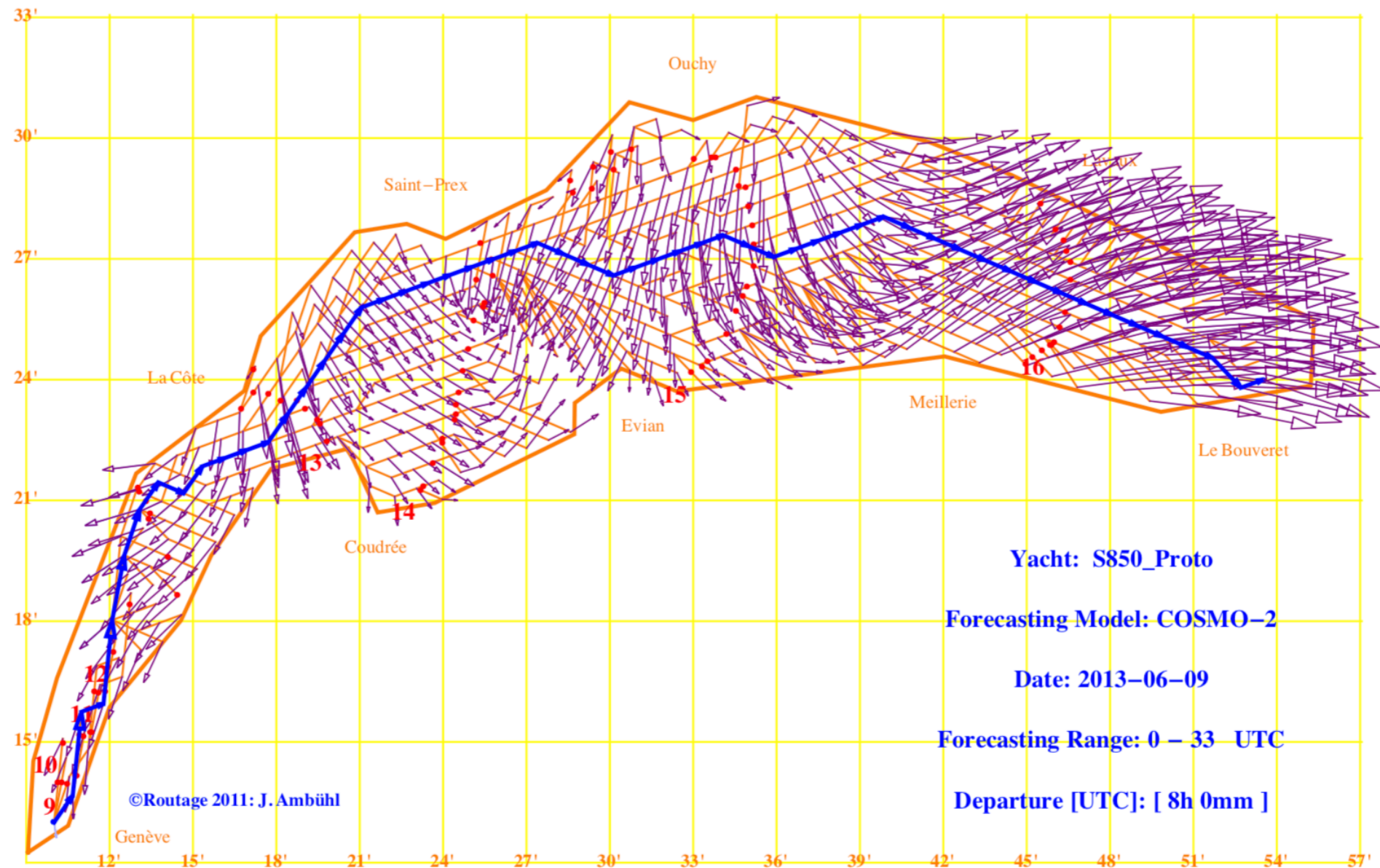


Classical technique in (financial) options trading

Alinghi, America Cup (EPFL, Prof. Dalang)

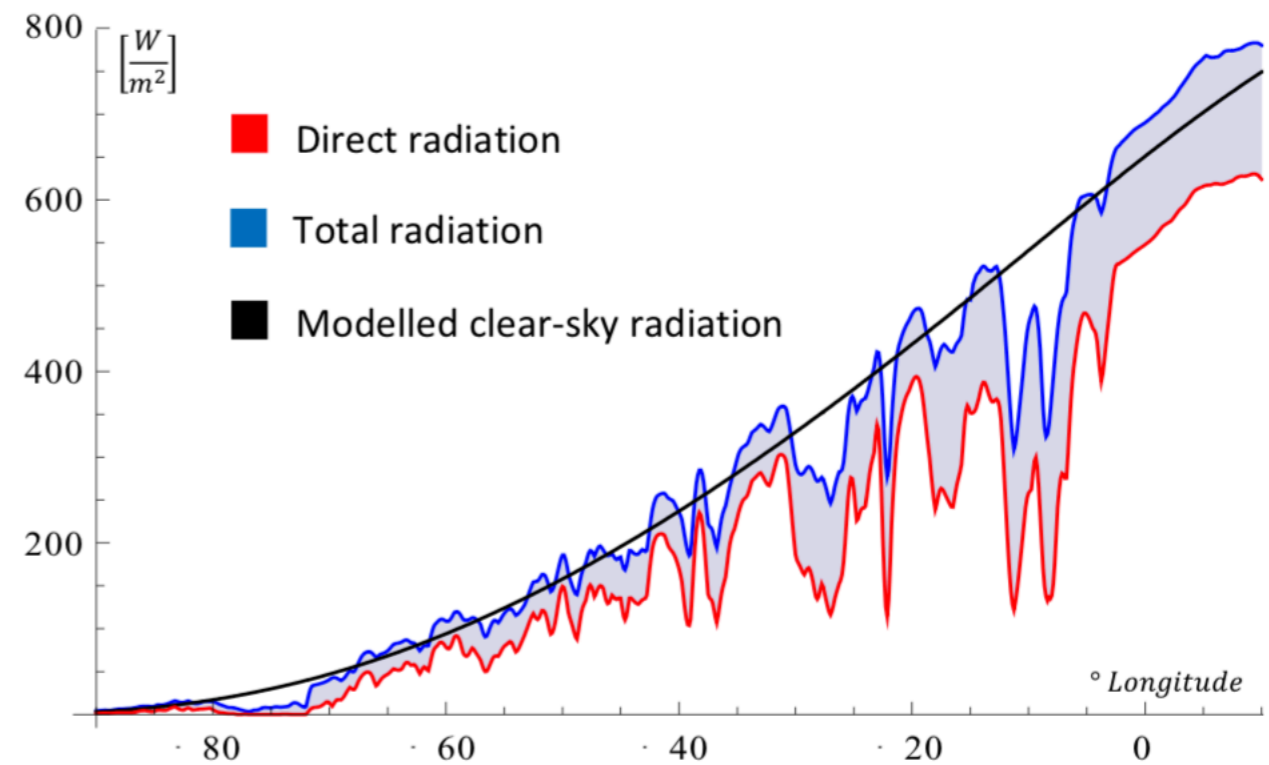
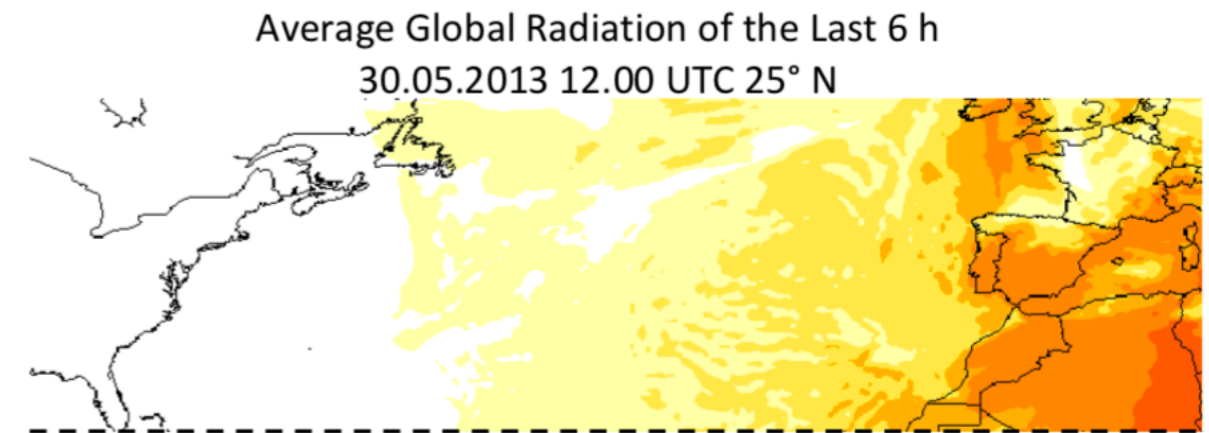
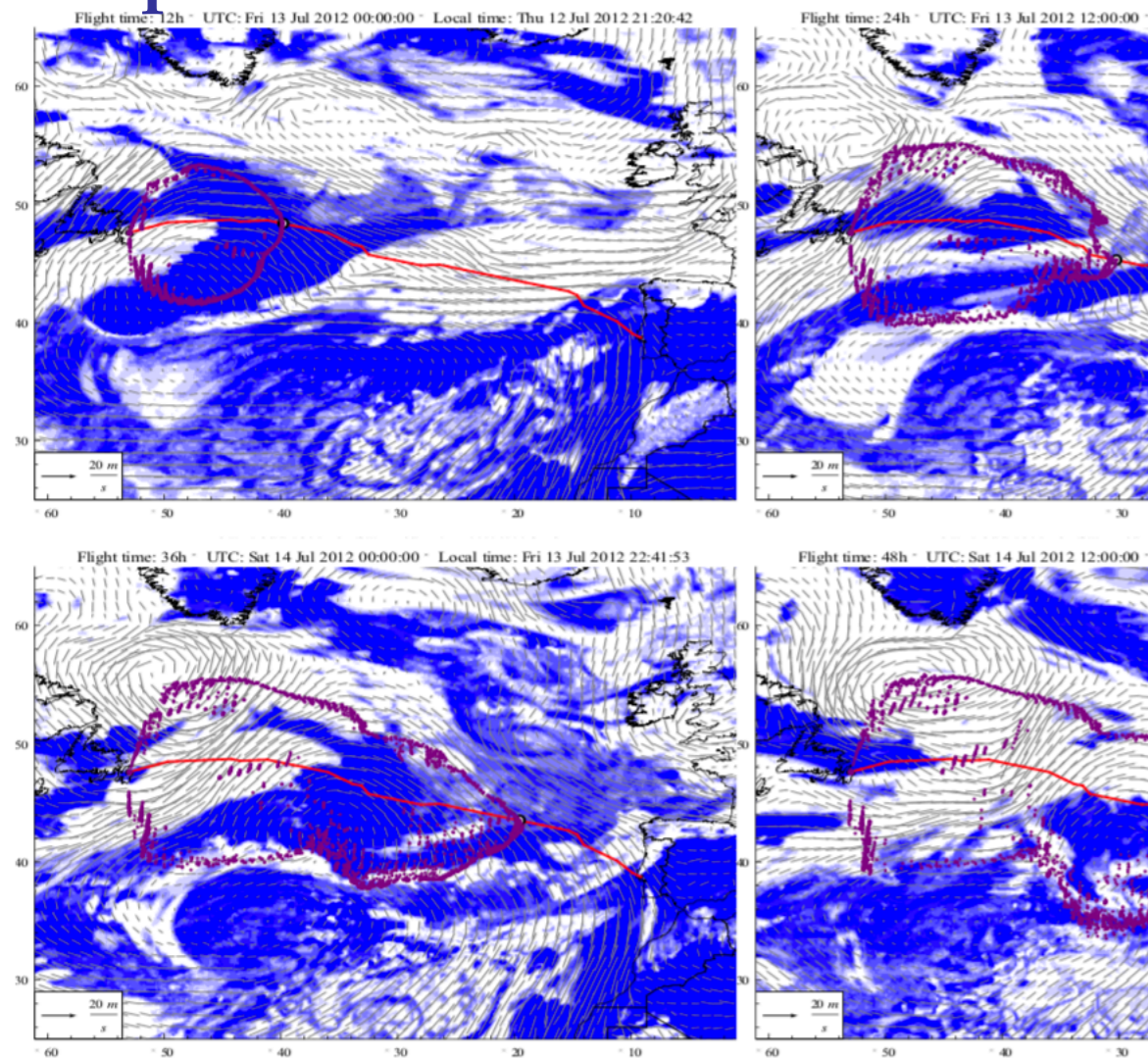
Решение о последовательности действий

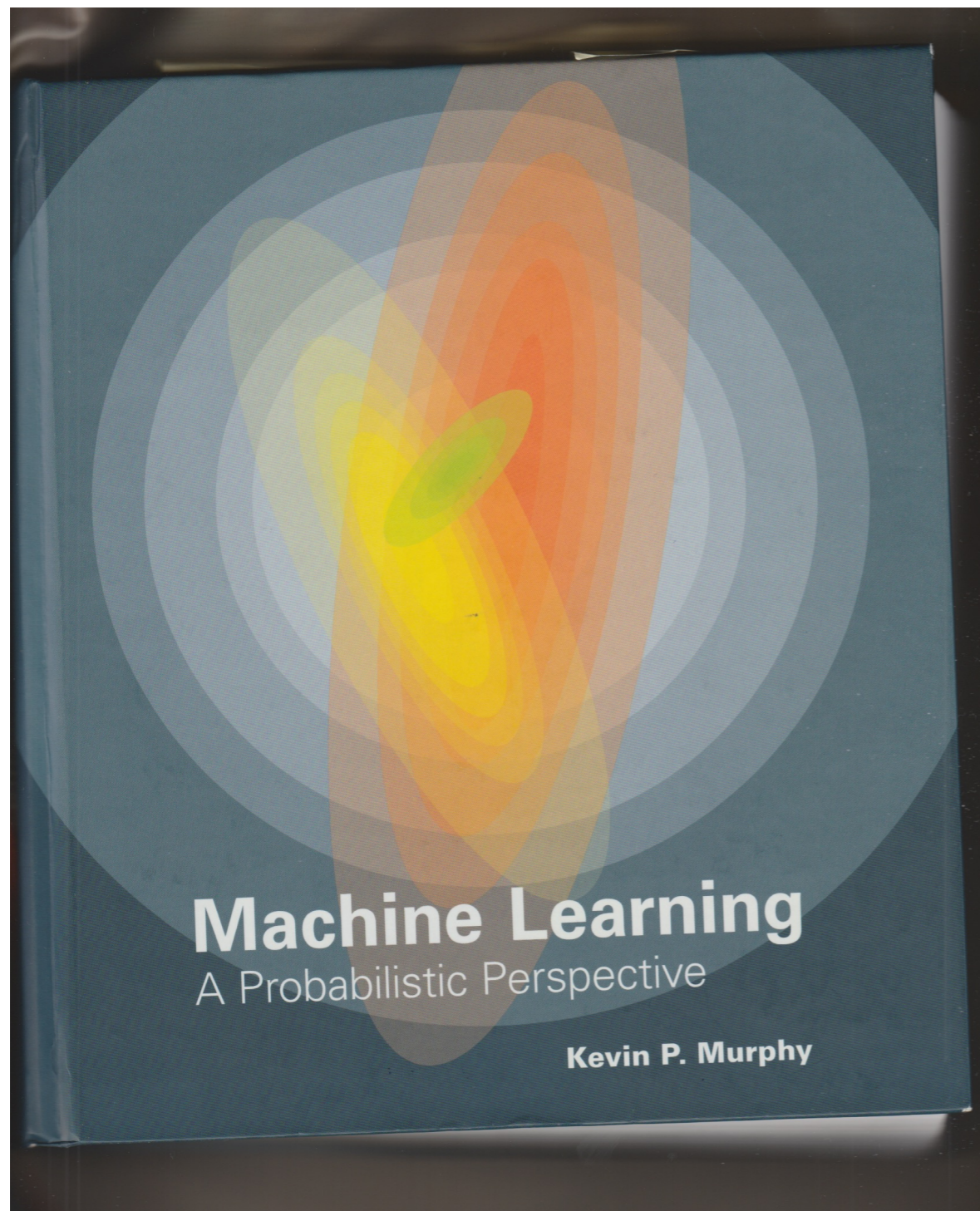
Sequential Decision (example: sailing boat routing)



Решение о последовательности действий

Автоматическая навигация дронов, подпитываемых солнечной энергией





ОБУЧЕНИЕ МАШИНЫ, ВЕРОЯТНОСТНАЯ ПЕРСПЕКТИВА

**Хорошая книга обо всем этом,
и многое другое ...**

Широкое видение



Нейронные сети в метеорологии, пример

Говоря о *бионике*, мы говорим о подходе инженера, вдохновлённого живой природой на планирование и создание машин и технических приборов. Леонардо да Винчи наблюдает за летучими мышами, представляя себе летательные аппараты; первые конструкторы самолётов пытаются повторить взмах птичьего крыла. Многочисленные современные механические и архитектурные конструкции вдохновлены мотивами из флоры и фауны. Но как работает это в структурах мозга? Это постепенно изучается в течении 20-

Дилемма Кассандры

Жак Амбюль (Jacques Ambühl) 2013 г.

Опубликовано в 2013 году на французском языке в «Ревю Метемагазин» (Revue Météo-Magazine) и в 2015 – на немецком языке в «Инженёр Биологи Ревю» (Ingenieur Biologie Revue).

Кассандра, дочь троянского царя, предсказывает будущее. Гектор, вождь троянского войска, защищает осаждённый афинянами город. Достоверность предсказаний Кассандры носит ограниченный характер, их можно описать как предположения. Гектор знает слабые места в расположении своих войск. Говоря современным языком, он осведомлён о своём профиле уязвимости. При принятии решений он советуется, в том числе, и с Кассандрой. Как же протекает их совместная работа?

оявляются первые модели ютера. Однако применение ель этого эссе - представить eoSchweiz.

**Благодарю за
внимание!**