**Модель**

Гипер-параметрами модели являются WIN\_SIZE – ширина окна, STRIDE – шаг на который будет сдвигаться окно, in\_channels – число каналов в модели данных, kernel\_size – размер сверточного ядра, dropout и scale\_factor – число слоев кодировщика (причем каждый следующий слой, число событий уменьшается в 2 раза, число каналов не изменяется), декодер симметричен кодировщику (то есть количество TCN слоев в автокодировщике равно scale\_factor\*2 + 1, т.к. после последнего upsampling слоя, используется ещё один TCN слой).

В тестах использовал 2 модели TCN автокодировщика общими параметрами у которых были in\_channels = 30, kernel\_size = 3, dropout = 0, STRIDE = 4. Отличные параметры моделей:

1. WIN\_SIZE = 64, scale\_factor = 4
2. WIN\_SIZE = 256, scale\_factor = 5

При обучении в среднем использовал, 50 эпох и скорость обучения 1e-3. Т.к. используется сдвиг окна, в некоторых ситуациях ВР нельзя будет целиком разбить на соответственные семплы, поэтому будем брать максимальное число разбиений, после чего передвинем окно в самый конец, для того, чтобы создать семпл с остаточными данными.

!!! В качестве ошибки обучения использовал MSE Loss. Но в качестве IRE решил брать максимум из MSE Loss по каждому из каналов (по умолчанию берется среднее значение, но предполагая, что аномалии одновременно будут находиться в малом числе ВР, для лучшего детектирования стоит смотреть на максимальную ошибку по всем каналам).

**Данные**

Синтетическая выборка состоит из 30-ти временных рядов по 20 000 событий. События ряда генерируются по формуле event = tfunc((t – t0)/w) + scale\*bias (1), где

tfunc – cos или sin (выбирается случайно)

t – текущий момент времени

t0, w, scale – задаются случайно для каждого ВР (t0 ∈ [50;100], w ∈ [40,50], scale = 0.3)

**Тесты**

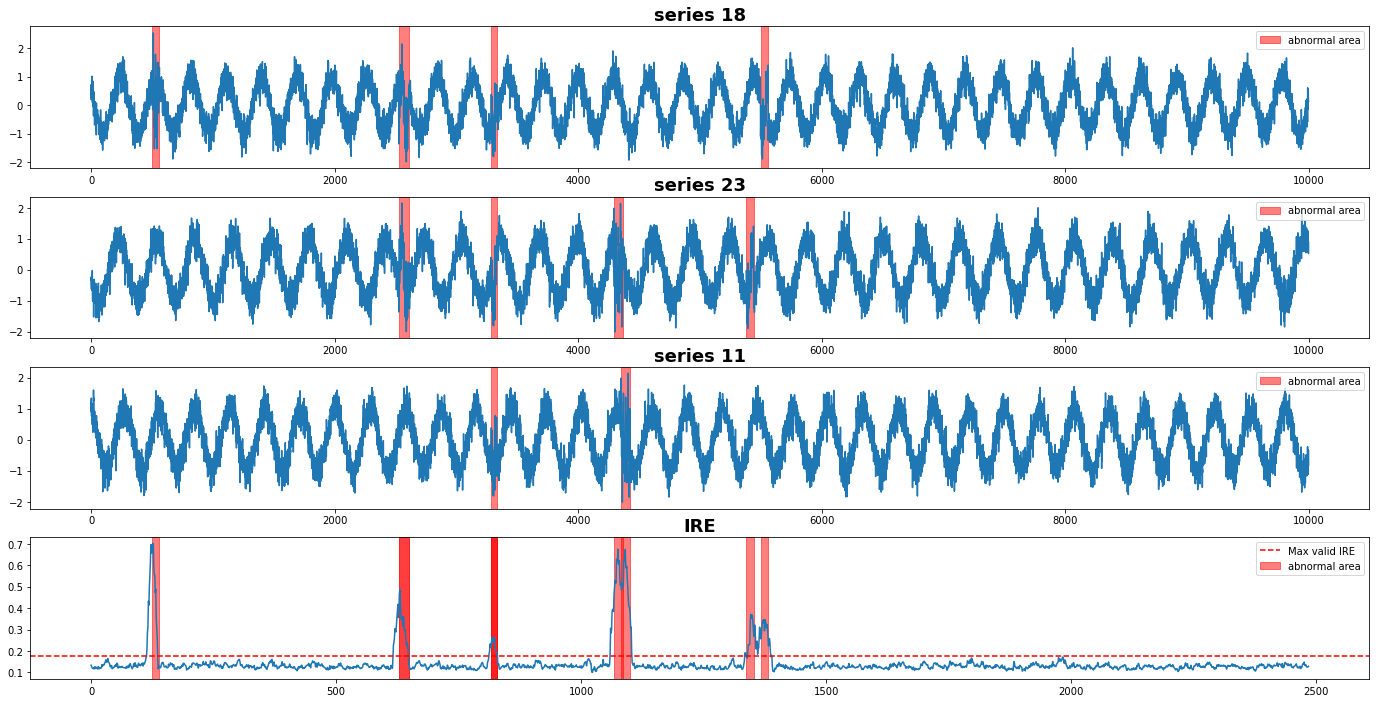
**Shift test**

Генерируем 5 временных рядов ошибок по формуле 1, при этом параметры t0, w выбираются из промежутков [30;50] и [50;80] соответственно, а scale = 0.5. Длинна каждого ряда выбирается случайным образом из промежутка [30, 90].

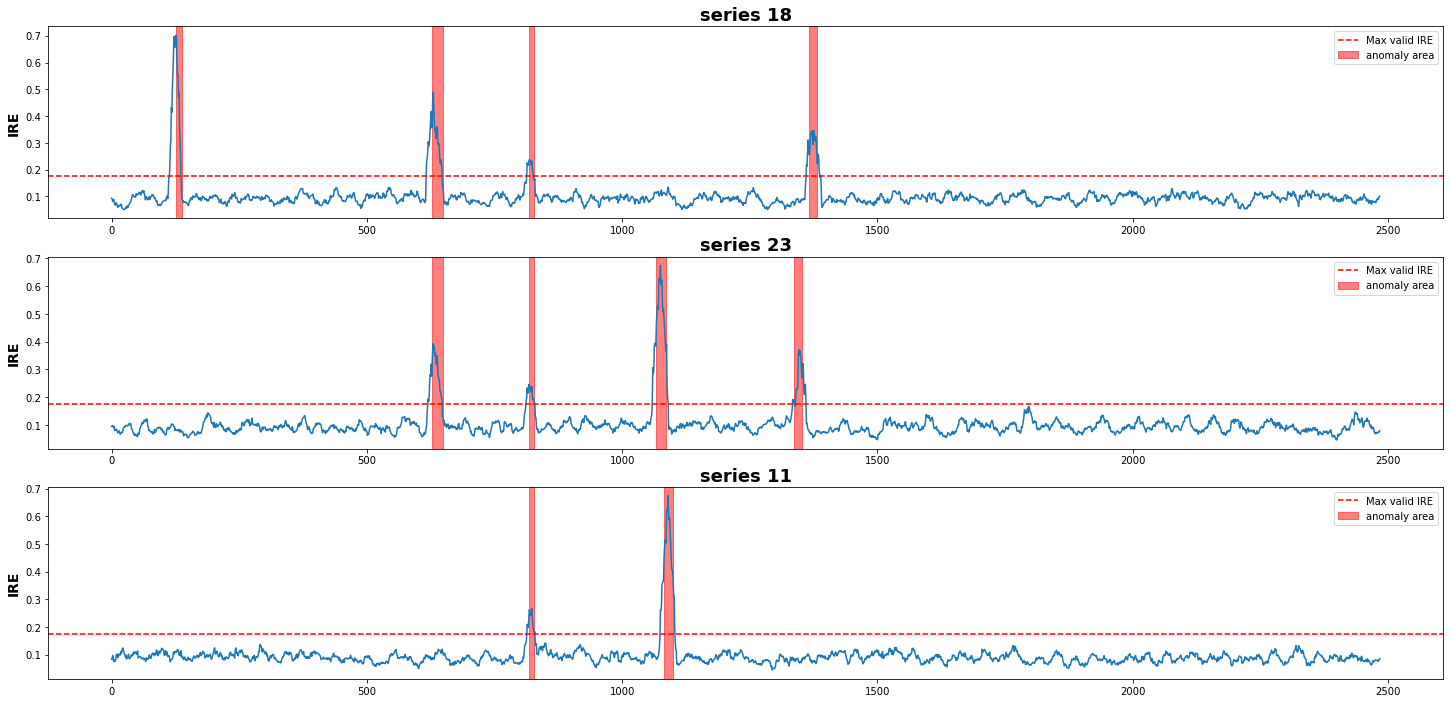
Выбираем 3 временных ряда из всей выборки, в которые будем интегрировать аномалии. Первую ошибку вставляем в один из ВР, вторую в два ВР, третью в три (предполагается, что аномалия проявляется в одном, двух или трех ВР одновременно). Четвертую ошибку вставляем в два ВР с некоторым небольшим сдвигом (предполагается, что аномалия проявляется в двух ВР с некоторой задержкой). Аналогично, пятую ошибку вставляем в два ряда так, но так, чтобы ошибки не пересекались.

*Результаты*

Модель 64\_4



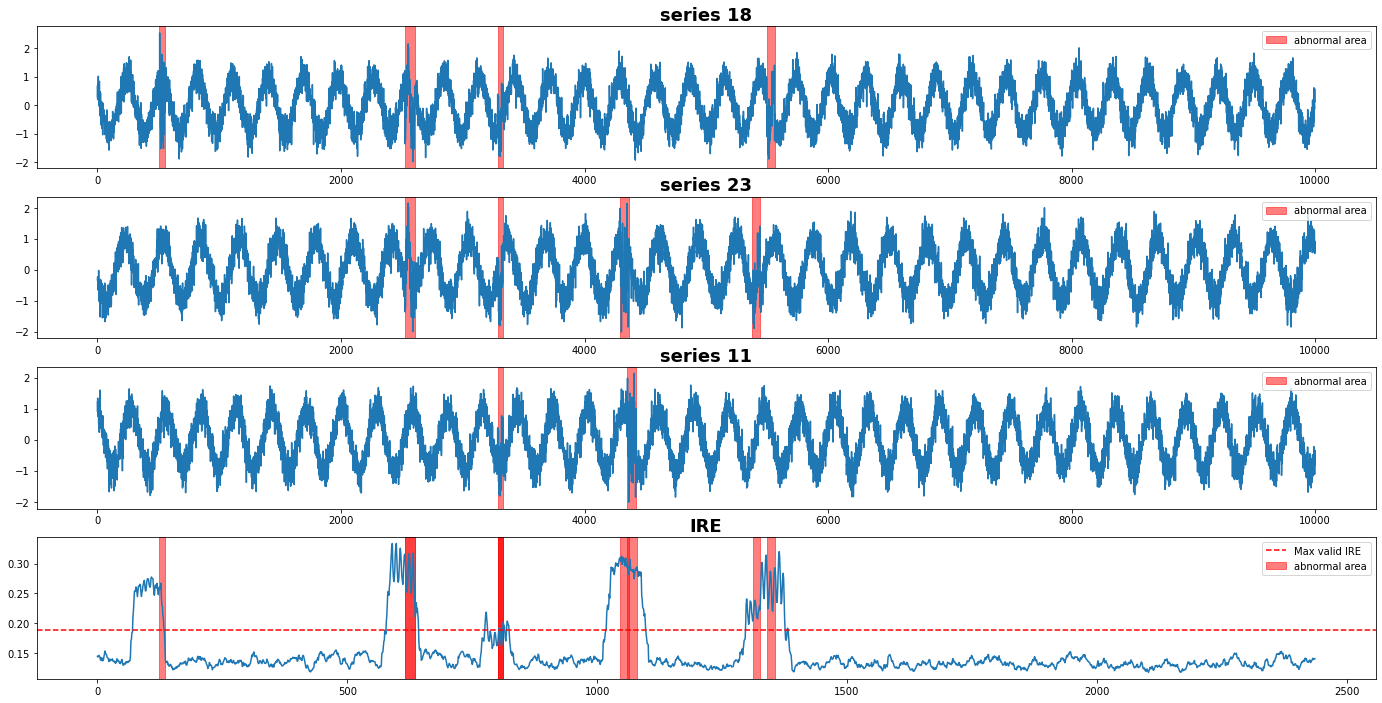
(рис.1 График аномальных ВР и результирующей IRE)



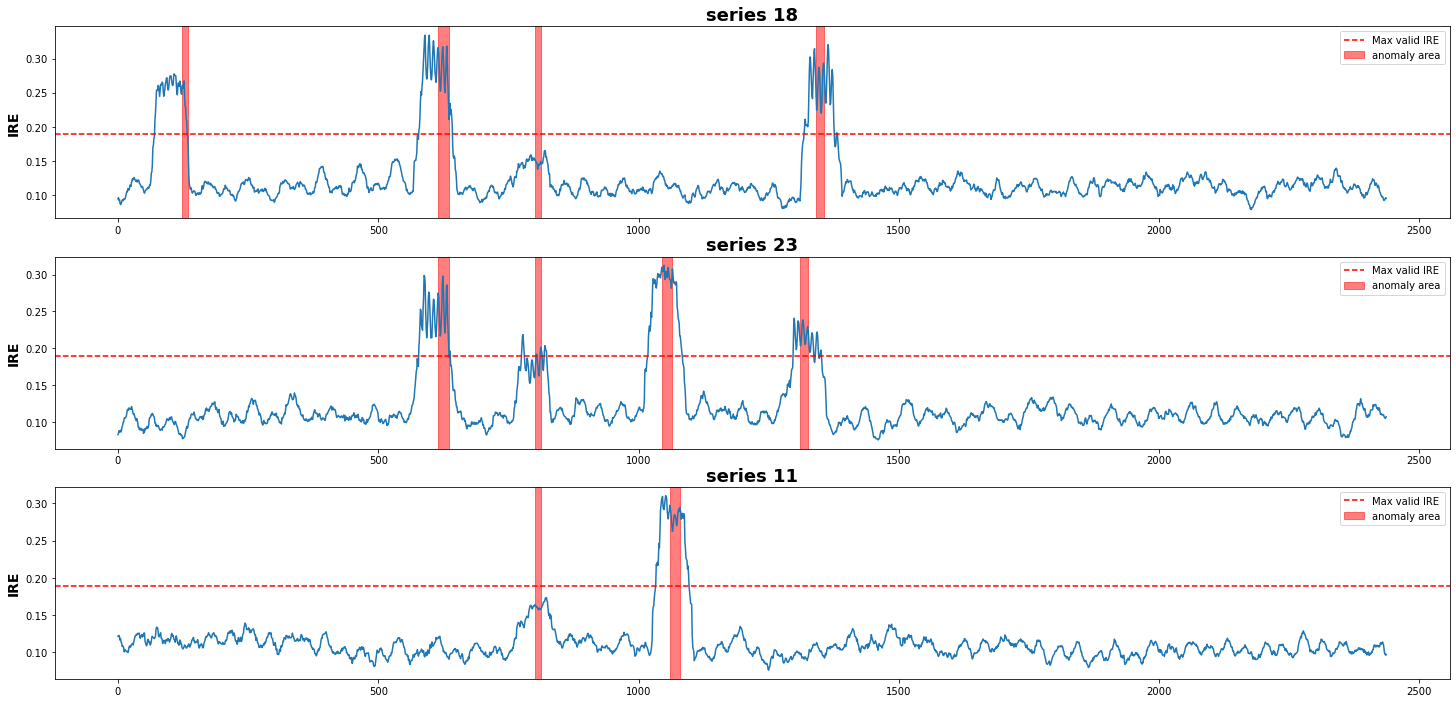
(рис.2 График ошибок аномальных ВР)

Все аномалии выявлены корректно, при этом можно отследить источник аномалий (рис.2).

Модель 256\_4



(рис.3 График аномальных ВР и результирующей IRE)



(рис.4 График ошибок аномальных ВР)

Аномалии на графике IRE присутствуют в большом диапазоне, так получается потому, что достаточно широкое окно с небольшим шагом будет содержать аномалию во многих соседних семплах.

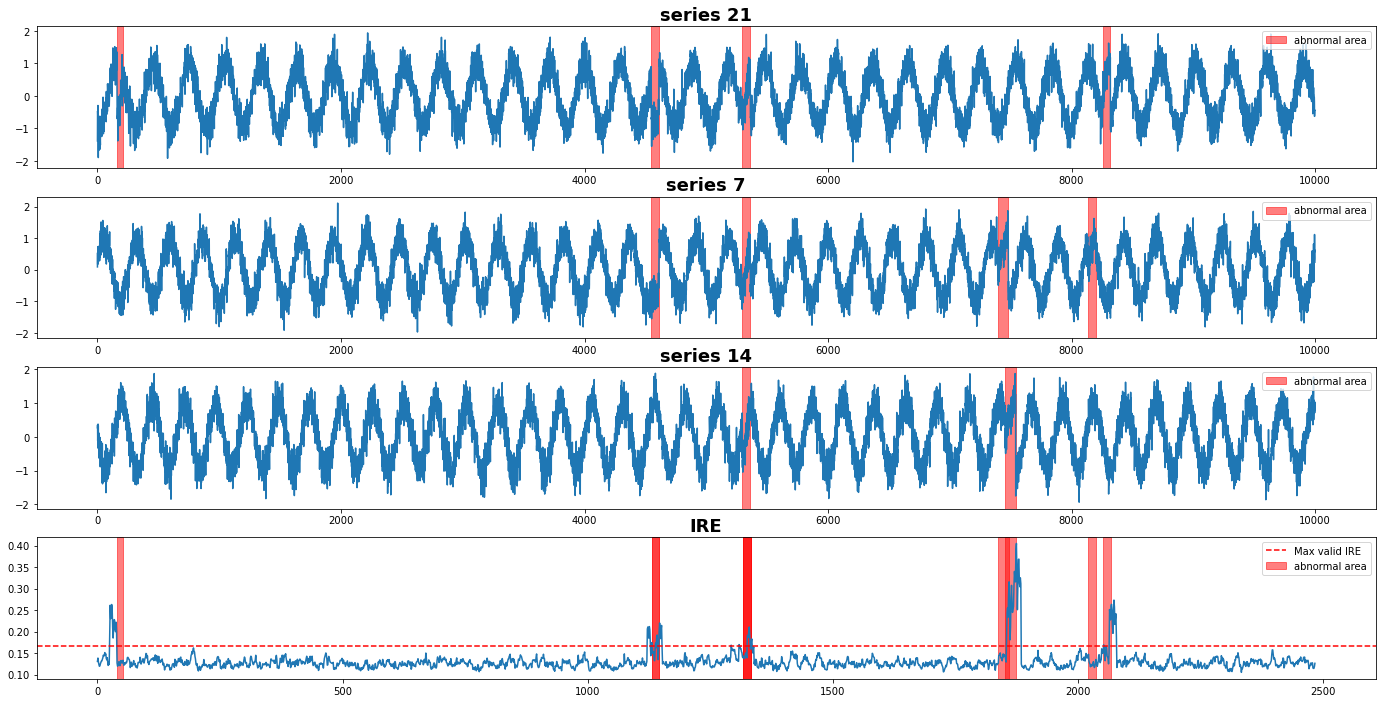
Можно сделать вывод – если окно сильно больше длины аномалии, появляются проблемы с детектированием конкретного аномального набора событий (т.к. аномалия будет присутствовать в большом количестве семплов), значит следует либо уменьшать окно, либо увеличивать шаг STRIDE.

**Shift similar test**

Тест аналогичный предыдущему, только теперь параметры ошибок выбираются из тех же промежутков, что и параметры для генерации обучающей выборки.

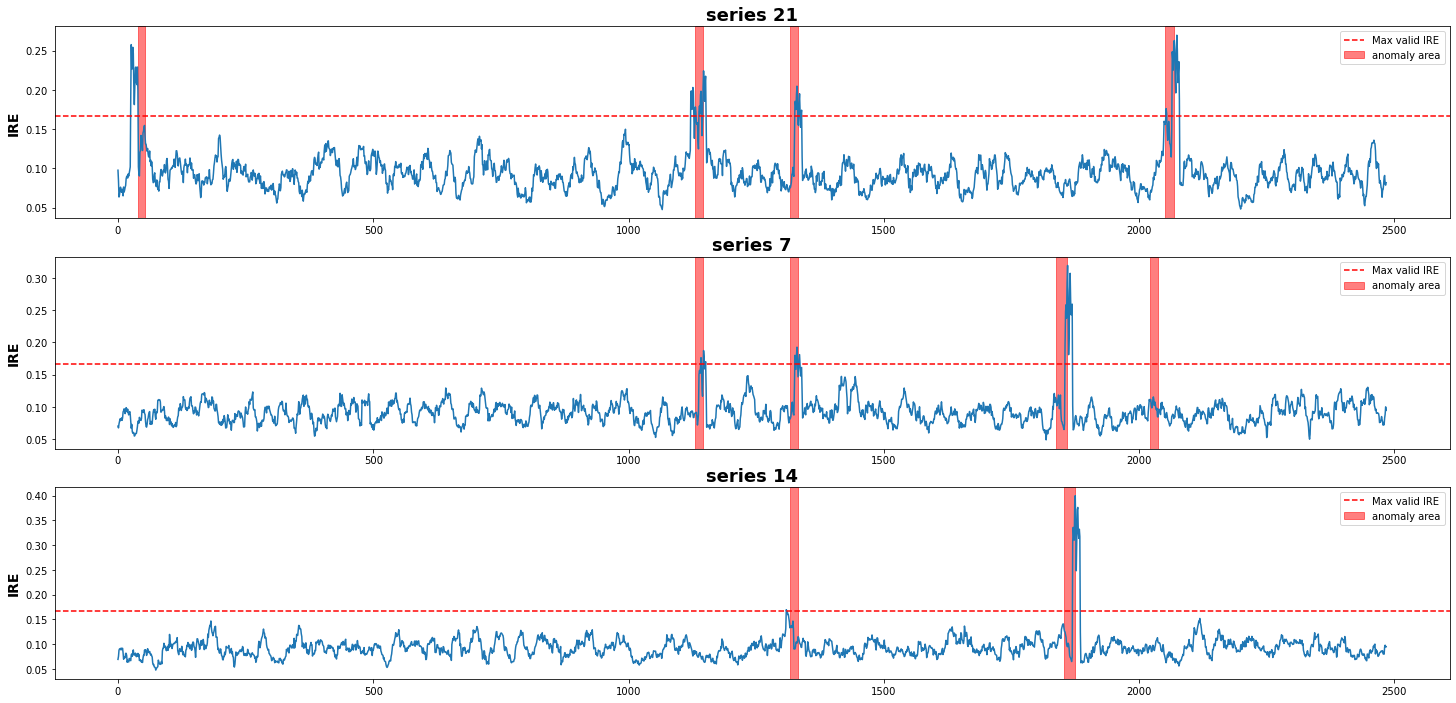
*Результаты*

Модель 64\_4

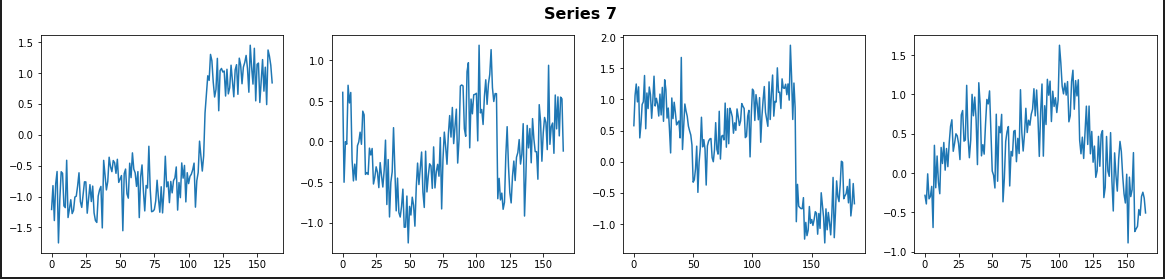


(рис.5 График аномальных ВР и результирующей IRE)

(рис.6 График ошибок аномальных ВР)

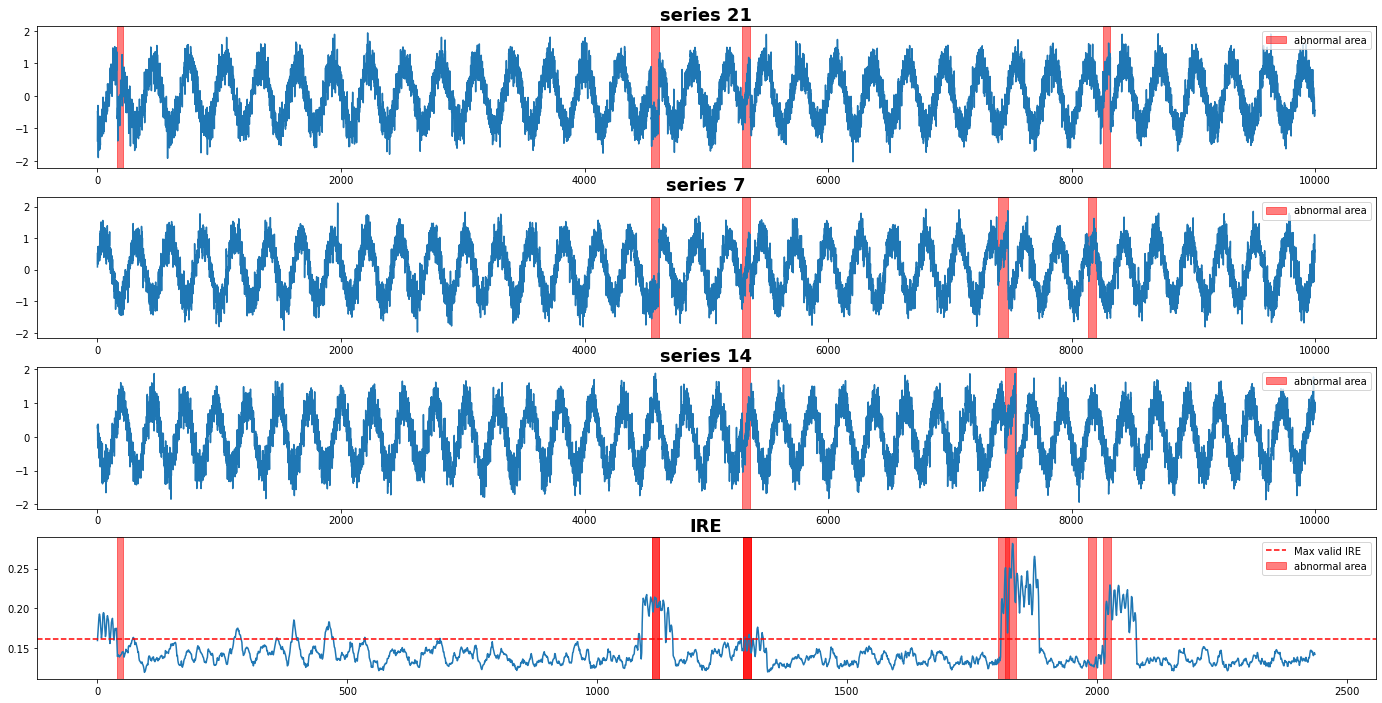


Аномалии стали менее выраженными, скорее всего на такую разницу повлиял параметр генерации ошибок – scale. Также ошибка 5 (которая присутствует в ВР 7 и 21 с большим сдвигом), не найдена в ВР 7, скорее всего из-за сходства сегрегированной ошибки с временным рядом. Построим график ВР 7 в местах его ошибок с отступами от аномалии по 50 событий с обеих сторон (рис.7). Заметим, что последняя ошибка, которая является ошибкой 5 для всего набора данных, действительно достаточно схожа с временным рядом и возможно модель не смогла обнаружить такую аномалию из-за устойчивости к шуму различного масштаба.

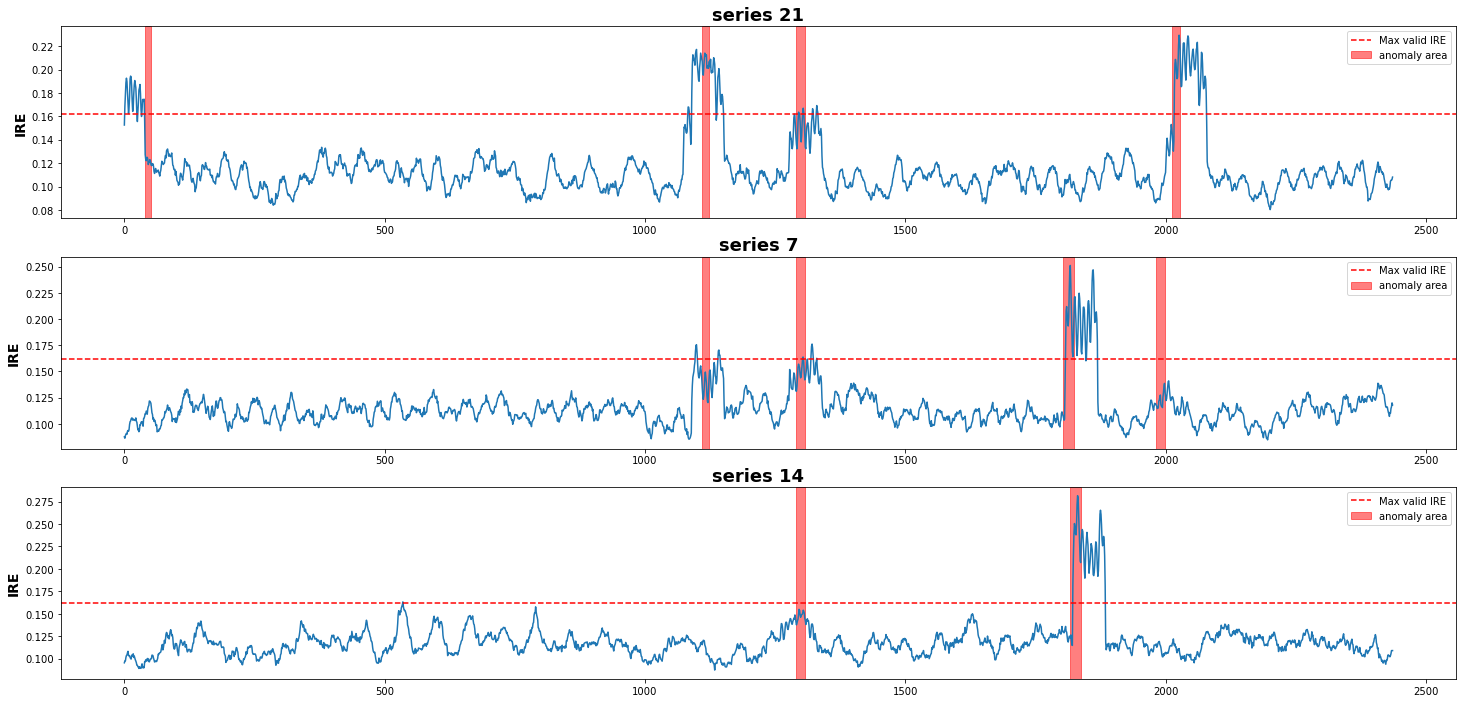


(рис.7 График ВР 7 в окрестностях его аномалий)

Модель 256\_4



(рис.8 График аномальных ВР и результирующей IRE)



(рис.9 График ошибок аномальных ВР)

В случае модели с широким окном, аналогично предыдущему тесту, появляется проблема поиска конкретного аномального промежутка, при этом наличие аномальные сегменты стали более явными за счет большой ошибки реконструкции в нескольких соседних семплах.

(Была идея вычесть из IRE среднее и прибавить 1, чтобы средняя ошибка была равна 1. После чего пройти окном небольшой ширины (3-5 элементов) и пересчитывать ошибку как произведение элементов в этом окне. Метод вроде сглаживания, только вместо среднего считать произведение, таким образом если несколько соседних событий будут иметь ошибку значительно больше одного, то на графике будет явный пик).

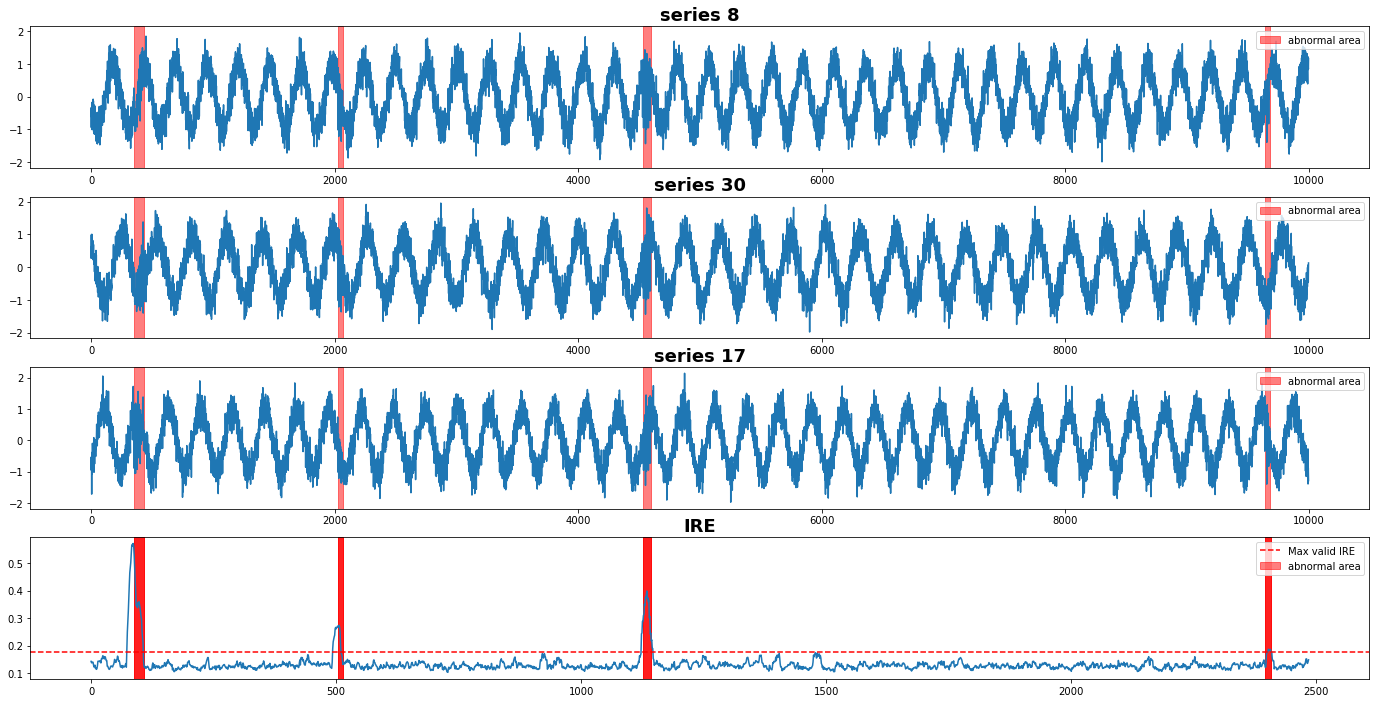
Вывод – модель с широким окном (большим чем максимально возможная аномалия), позволяет лучше детектировать факт возникновения аномалий, при этом модель с более узким окном позволяет точнее находить конкретные аномальные промежутки. Обе модели не смогли обнаружить слабо отличимую аномалию в ВР 7.

**Dilation similar test**

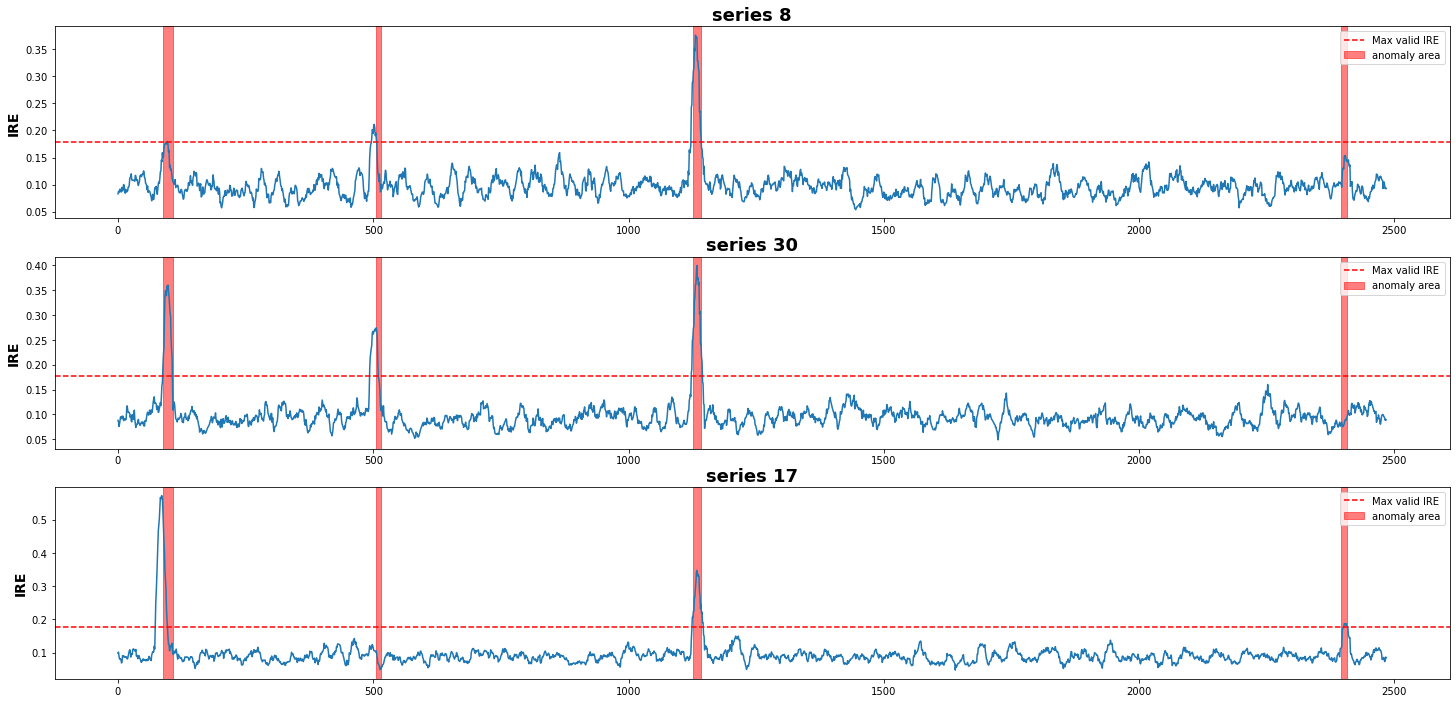
В этом тесте разбавлял аномальные события обычными. То есть при интеграции ряда аномалии, заменял только каждый 2-ой, 4-тый, 8-ой и 16-тый элемент, соответственно для четырех сгенерированных ошибок.

Ошибки генерируются аналогично обычным ВР.

Модель 64\_4

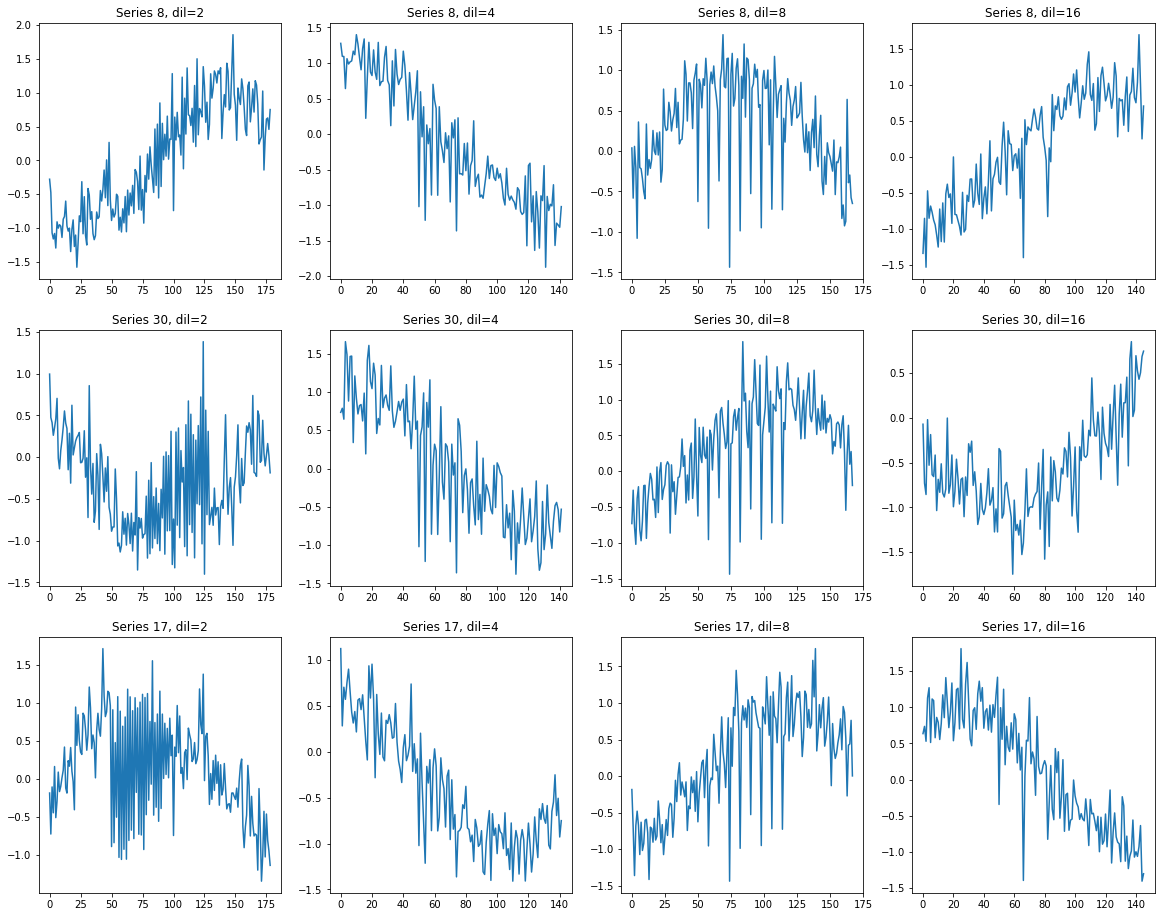


(рис.10 График аномальных ВР и результирующей IRE)



(рис.11 График ошибок аномальных ВР)

Аномалия с пропуском в 16-ть событий практически не выделяется на графике IRE. Стоит заметить, что последней аномалии (с dil = 16) равен 46, то есть аномальных событий с таким расширением (пропуском) в аномальном сегменте будет всего 3.



(рис.12 Графики всех ВР в окрестностях их аномалий)

Результат даже улучшился, относительно предыдущего теста с этой моделью. Скорее всего резкий рост дисперсии (из-за того, что аномальные события чередуются с обычными) помогает модели детектировать аномалии.