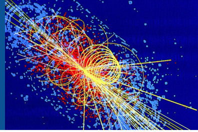
Machine Learning Based Anomaly Detection for Online Data Quality Monitoring of the CMS Electromagnetic Calorimeter

Kyungmin Park, Abhirami Harilal, Manfred Paulini, Michael Andrews

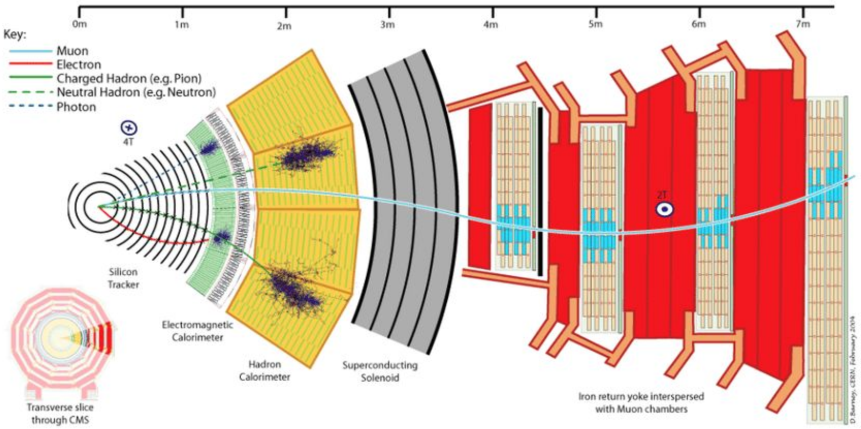
On Behalf of the CMS Collaboration

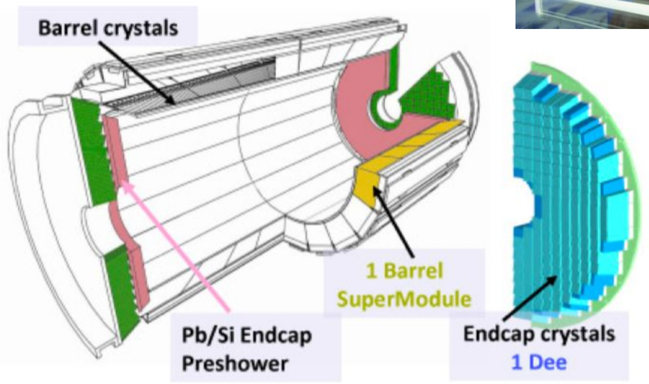
August 25, 2023 @ AI4DQM

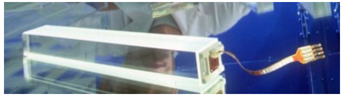
CMS Electromagnetic Calorimeter



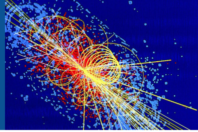
- <u>Electromagnetic calorimeter (ECAL)</u>
  - Measures energy of **photons and electrons**.
  - Consists of 75,848 lead tungstate ($PbWO_4$) crystals, arranged in central *barrel* **(EB)** section and two *endcaps* **(EE+, EE–)**.
    - Electrons and photons produce scintillation light when passing through the crystal, in proportion to the particle’s energy.
    - Scintillation light is detected by the photo-detectors, converted to electrical signal.



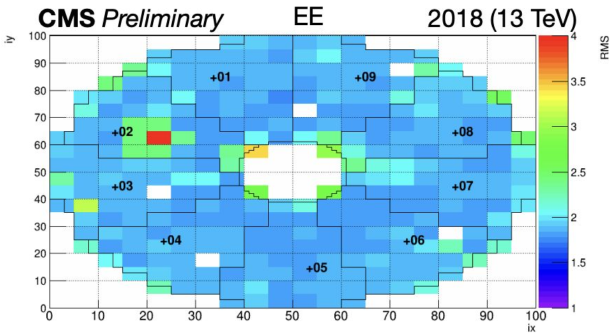


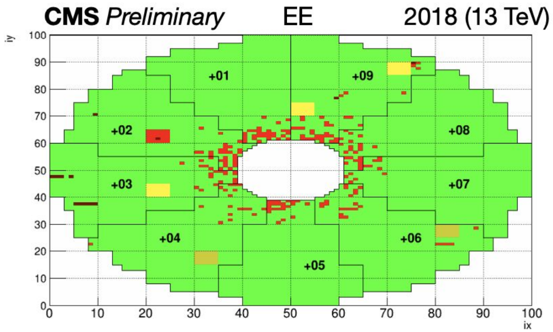


Online Data Quality Monitoring for ECAL



- In the operation of the particle detector, it is important to monitor the quality of data and **catch anomalies** in the detector.
- Online **Data Quality Monitoring (DQM)** of CMS ECAL offers a real-time snapshot of a subset of raw data, followed by a quality interpretation.



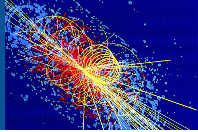


<u>Quality Histogram</u>

Green: GOOD
Red: BAD
Brown: KNOWN ISSUE
Yellow: NO DATA

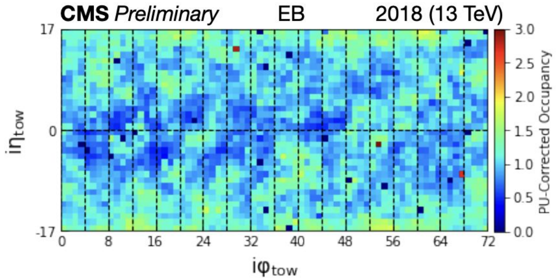
  - Highly time-sensitive operational task
  - <u>Challenge</u>: Anomalies can come in all shapes and sizes. → impossible to anticipate all failure modes.
    - ➔ Utilize machine learning as a possible solution.

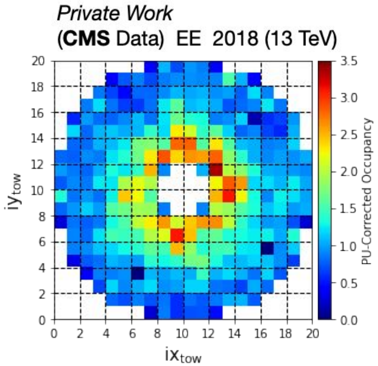
Dataset and Pre-processing



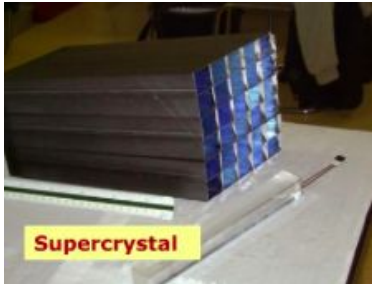
- Dataset taken from 2018 LHC collisions, manually certified as “good”.
    - Use ***occupancy*** histograms processed as 2D images for the quality plot.
        - Crystal with energy deposit above a set threshold → occupancy of 1 for the crystal.
        - Each image represents each *Lumi-Section (LS)* of ~23 seconds.

- Pre-processing
    - Normalize the occupancy data by ***PileUp (PU)***, additional proton-proton interactions within the same proton bunch crossing.
      → Make occupancy images consistent across different LHC run conditions.



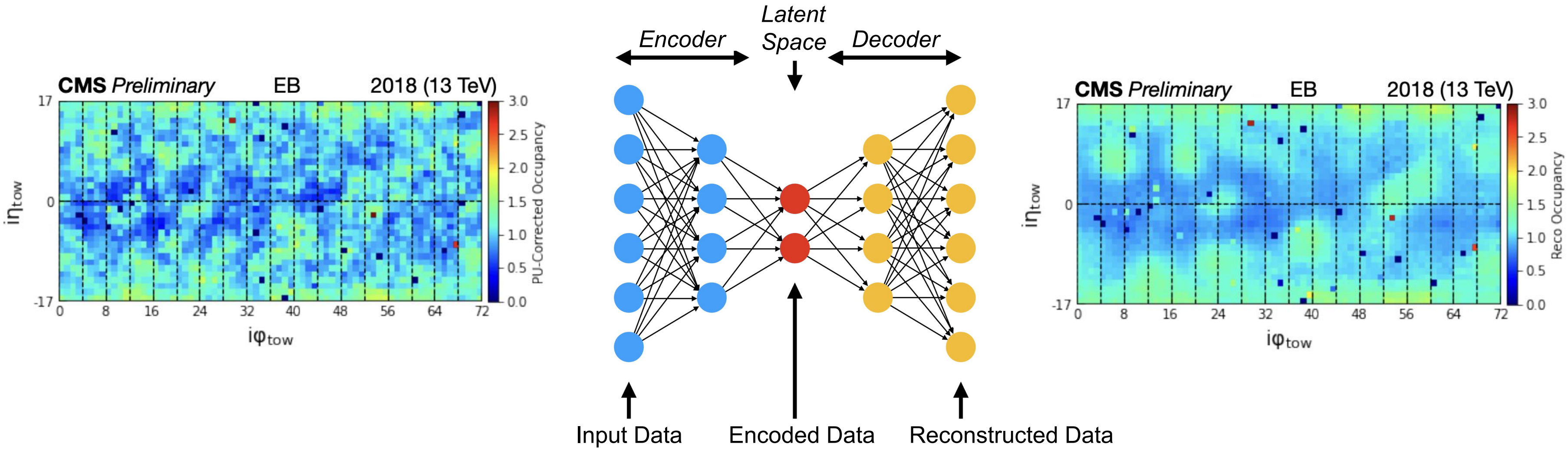


Each pixel = *tower (supercrystal)* of 5x5 channels.



Machine Learning for Anomaly Detection

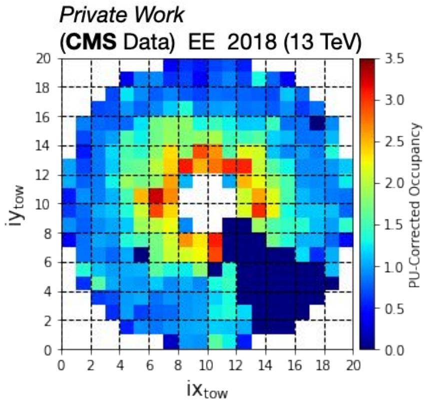
- **Unsupervised machine learning** for a more robust anomaly detection
  - **Autoencoder (AE)** model can learn the “good” pattern of detector data.
    - Encoding layers: learn a representation of data and compress the feature.
    - Decoding layers: reconstruct the input from the compressed features.



➔ AE can spot the anomalies that deviate from the learned norm, eliminating the need for hand-coded rules for every failure mode.

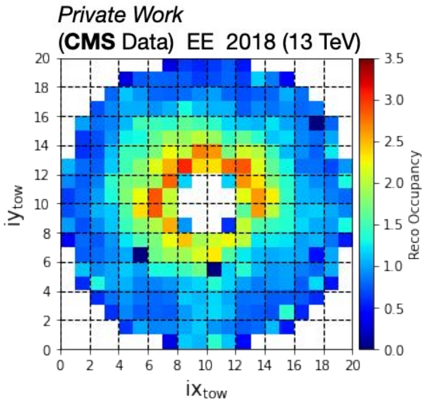
Anomaly Detection Strategy using AutoEncoder

- AE-based anomaly detection and localization



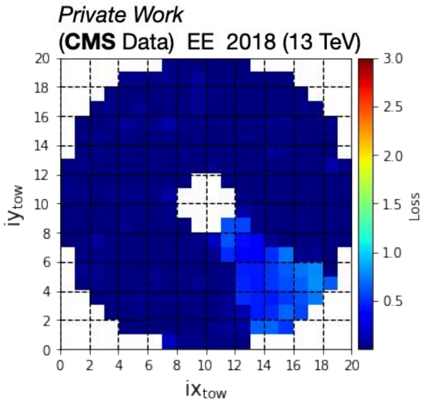
[1] **Input map**

Anomaly present



[2] **AE-reconstructed map**

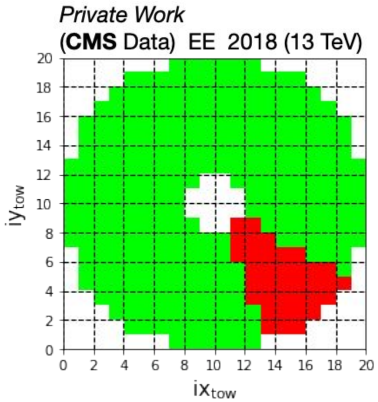
Anomaly not reconstructed



[3] **Loss map**

Mean Squared Error loss

High loss in the anomaly

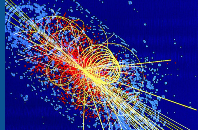


[4] **Quality plot**

Using anomaly tagging threshold

- **Anomaly tagging threshold** is obtained from **validation** sets with “fake anomalies”.
  - Based on loss values of anomalous towers from the anomaly validation sets, choose a **loss threshold** that AE can catch 99% of all anomalies.

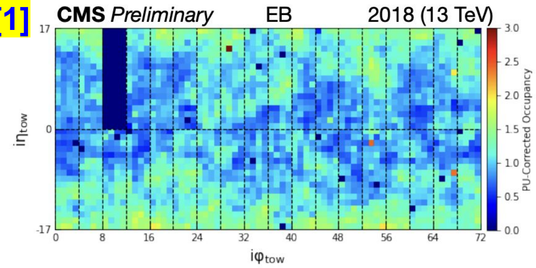
Training and Validation Strategy



- **Training**
  - <u>Network</u>: Convolutional Neural Network (**CNN**) with **ResNet** architecture
    - Train AE models each for the barrel and endcaps.
  - <u>Dataset</u>: 90k occupancy images from “good” dataset

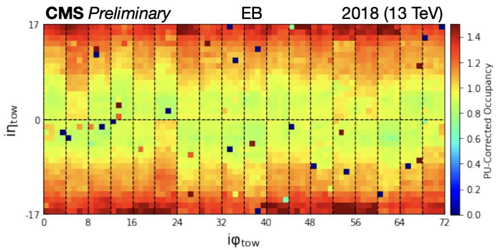
- **Validation**
  - Validation sets using 10k “good” occupancy images
    - Nominal validation: using the “good” occupancy images
    - **Fake anomaly** validation: same images with anomalies manually introduced
      - Zero occupancy tower
      - Hot tower
      - Missing supermodule (barrel) [1] / sector (endcaps)
  - Test using real anomalous data

[1]

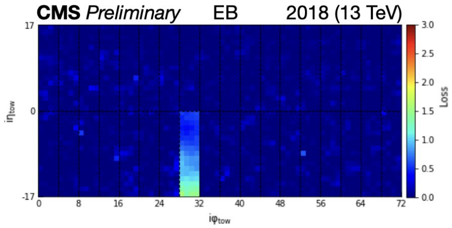


Post-processing: Spatial and Time Correction

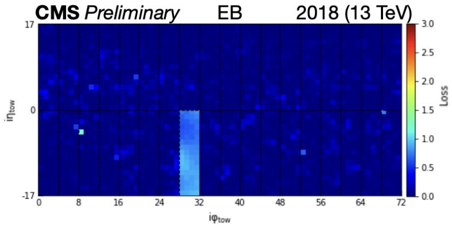
- **Spatial correction**: account for **spatial variations in the ECAL response**
  - Crystals at high $|\eta|$ exhibit higher occupancy than those at low $|\eta|$.
  - Normalize loss maps by the **average occupancy map**.



Average occupancy map

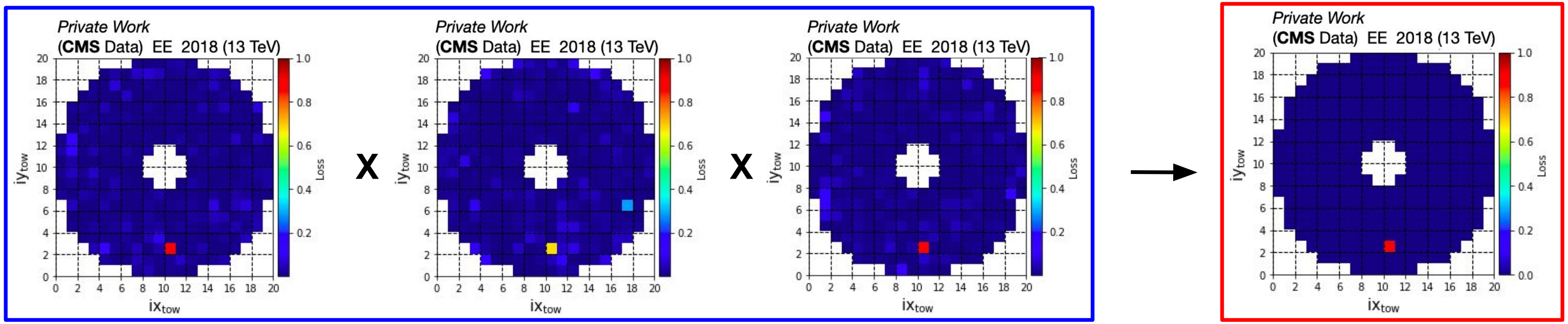


Before spatial correction

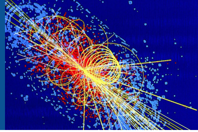


After spatial correction

- **Time correction**: exploit **time-dependent nature of anomalies**.
  - Anomalies in towers persist throughout several LSs, while fluctuations do not.

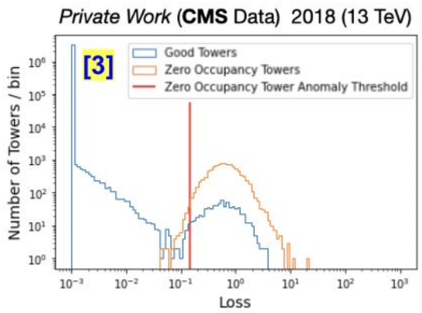


Validating with Fake Anomalies



- <u>Performance metric</u>: **False discovery rate (FDR)**

$$\textbf{FDR} = \frac{\text{Number of good towers above anomaly threshold}}{\text{Number of good and bad towers above anomaly threshold}}$$



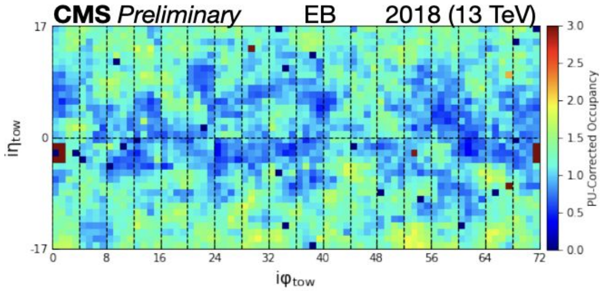
- <u>Summary of FDR</u> for 99% anomaly detection

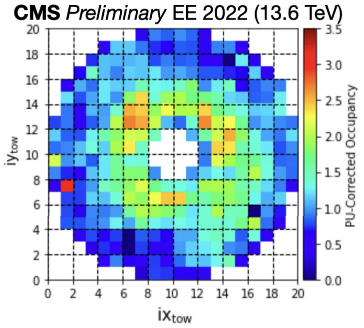
Scenario	Missing Supermodule (EB) / Missing Sector (EE)			Zero Occupancy Tower			Hot Tower		
	Barrel	**EE+**	**EE–**	**Barrel**	**EE+**	**EE–**	**Barrel**	**EE+**	**EE–**
<u>No correction</u>	3.6%	28%	27%	51%	86%	87%	2.8%	0.07%	< 0.01%
<u>After spatial correction</u>	3.1%	1.6%	1.9%	49%	14%	14%	2.9%	0.11%	0.07%
<u>After spatial and time correction</u>	0.13%	0.19%	0.23%	4.1%	5.6%	6.4%	< 0.01%	< 0.01%	< 0.01%

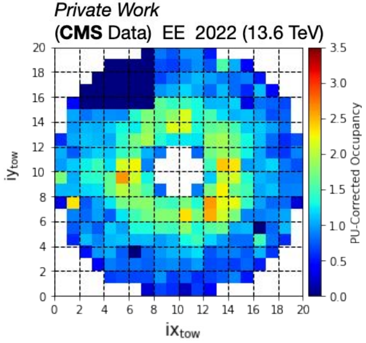
Testing with Real Anomalies

- Test on data with real anomalies, using the loss threshold from fake anomaly validation.

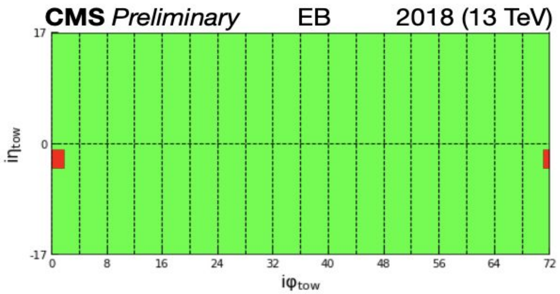
Input plot

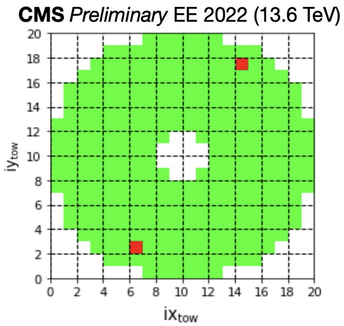


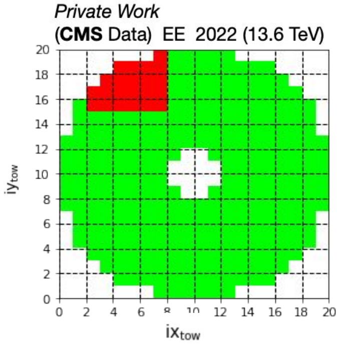




Quality plot



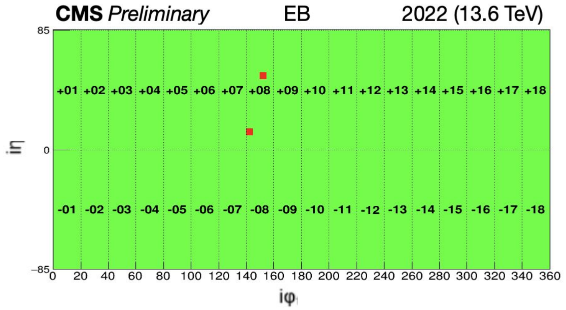




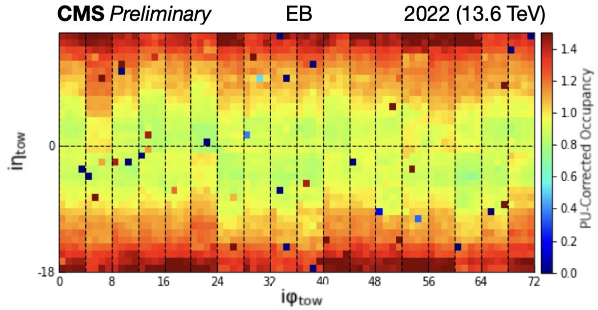
➔ Catches anomalies well with various shapes and sizes.

Deployment in Online DQM for ECAL

- AE-based DQM is **deployed** in the ECAL Online DQM workflow, as a new ML quality plot.
- Detecting potential bad towers



[1] Quality plot from AE

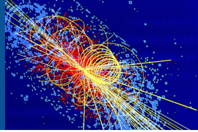


[2] Average occupancy map

  - ML quality plot [1] shows a tower that had very low occupancy in several LSs – “**semi-transient anomaly”** – not shown in other plots in current DQM.
  - Its low occupancy shows up in the average occupancy produced offline [2].

➔ AE can also spot potentially degrading towers.

Summary

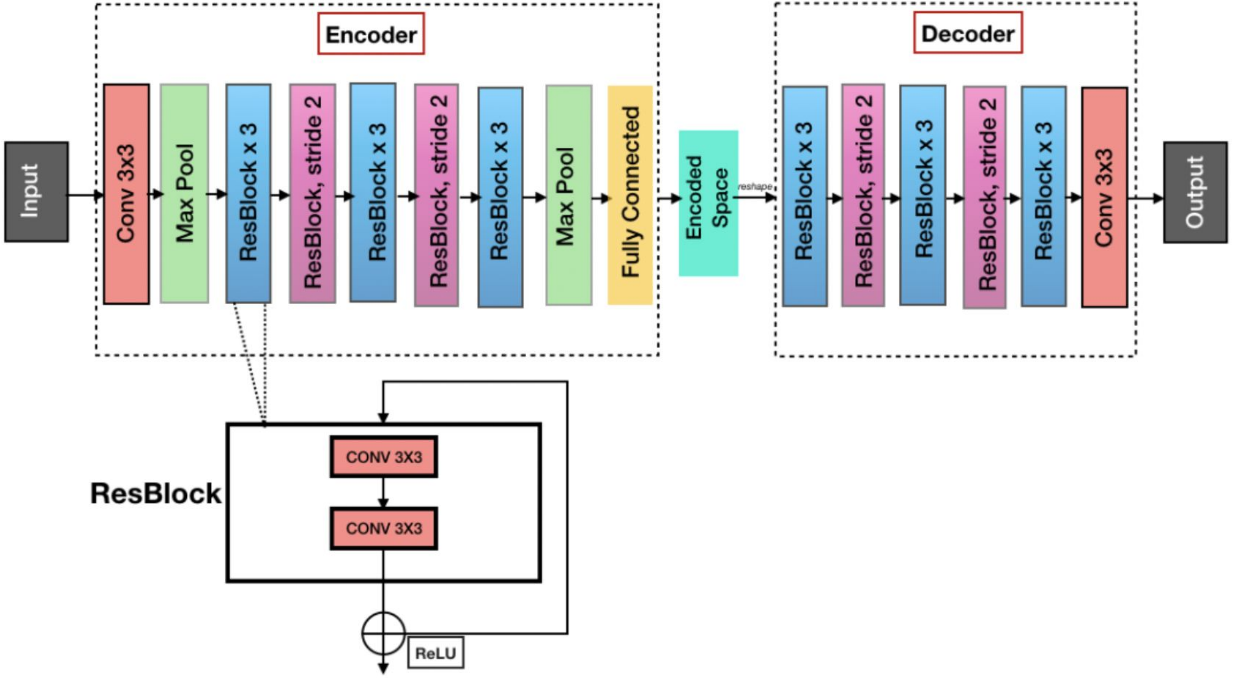


- Developed a robust ML-based anomaly detection & localization system for CMS ECAL.
  - After accounting for spatial and time variations in the detector response, AE-based system is able to detect various anomalies in fake anomaly with an estimated FDR of ~6% at 99% detection rate.
  - AE-based system is able to detect real anomalies of arbitrary shapes.

- The system has been deployed in the ECAL Online DQM workflow for barrel, detecting bad and potentially degrading towers.
  - This method can be generalized to anomalies of arbitrary shapes and extended to other experiments requiring data quality monitoring.

BACK UP

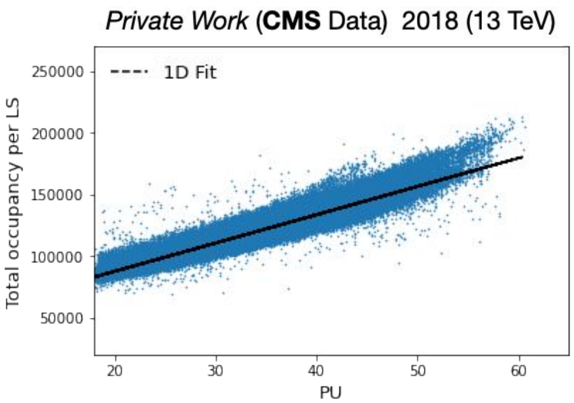
AutoEncoder Model



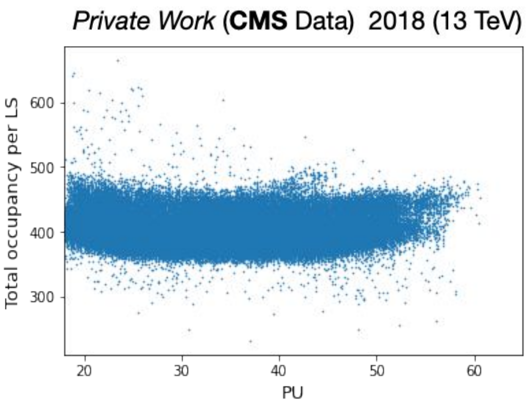
Pre-processing: PU Correction

- Occupancy is linearly related to PU.
  - In order to make our data consistent across different runs and LHC conditions, remove PU dependency from our occupancy data.

<u>EE+ Total occupancy</u> <u>*(sum of occupancies in all towers)* vs.</u> **<u>PU</u>**



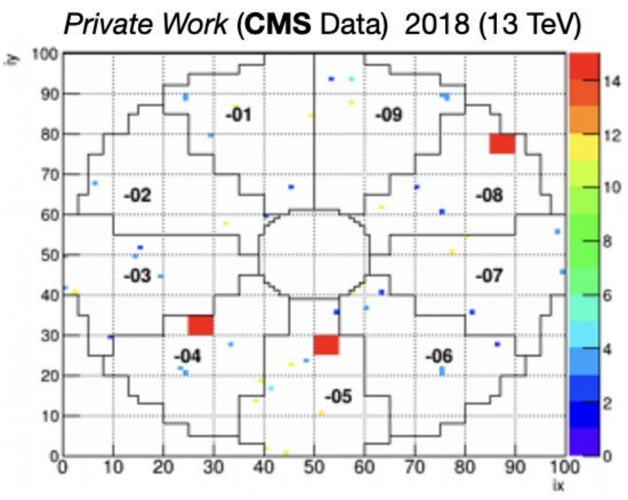
<u>Before PU correction</u>



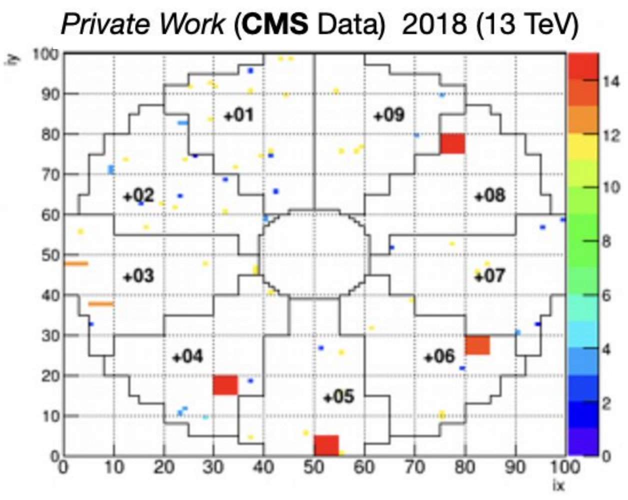
<u>After PU correction</u>

Pre-processing: Masking Problematic Towers

- Masking towers with known issues
  - Some towers are known to have issues throughout the 2018 runs.
  - Mask those towers from the dataset, and not include them in the list of “valid towers” of our interest for training and validation.



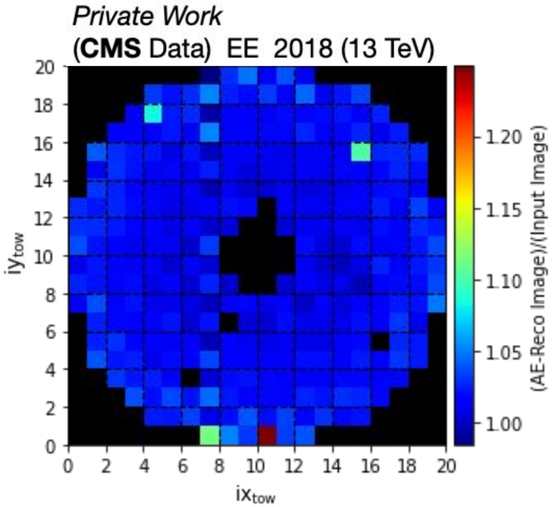
EE– channel status map



EE+ channel status map

Padding

- Ratio of AE-reconstructed image and original input image for **Endcap**

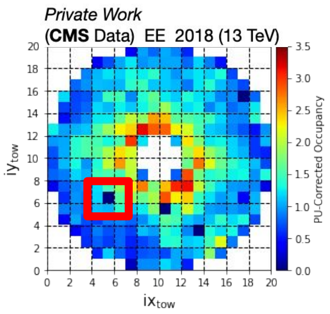


→ *Pattern around the edges of EndCap*
→ *Similar "edge effect" observed for barrel*

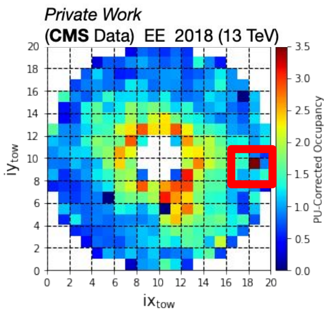
- Padding for training and validation dataset
  - Add padding around the occupancy map image for training and validation, copying the adjacent edges.
  - When calculating anomaly threshold from loss map, do not consider the loss from the padded pixels.

Validation with Fake Anomalous Towers

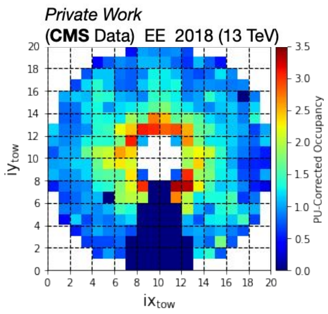
- Fake anomaly validation sets for each AE model (EB and EE):
  - 1) “**Zero occupancy tower**” validation: *fake zero occupancy tower* manually introduced to one of valid towers from each occupancy map.
  - 2) “**Hot tower**” validation [*]: *fake hot tower (high occupancy)* manually introduced to one of valid towers from each occupancy map.
  - 3) “**Missing supermodule (EB) / sector (EE)**” validation: missing supermodule/sector manually introduced.



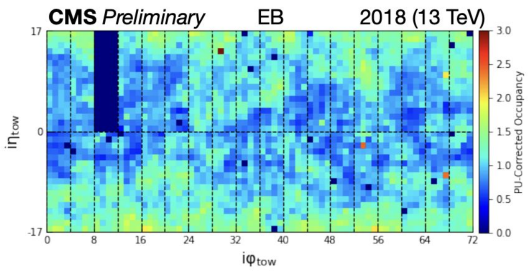
Zero occupancy tower



Hot tower



Missing sector



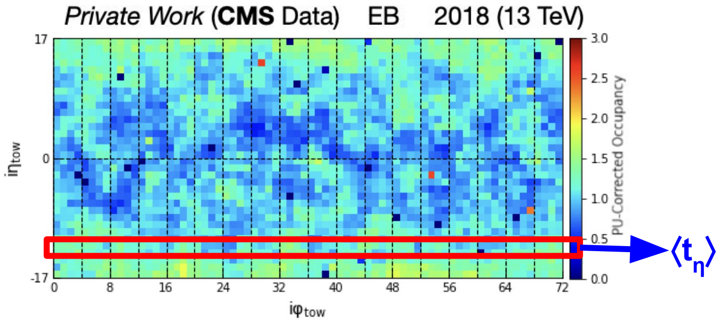
Missing supermodule

[*] Hot tower value (before PU normalization): ***25 × 500 × f***, f = (0,1]

- 25 crystals per tower, 500 events per LS, f can be 1 at max. Target *f > 0.1 for barrel, f > 0.2 for endcap*.

Baseline for Comparison

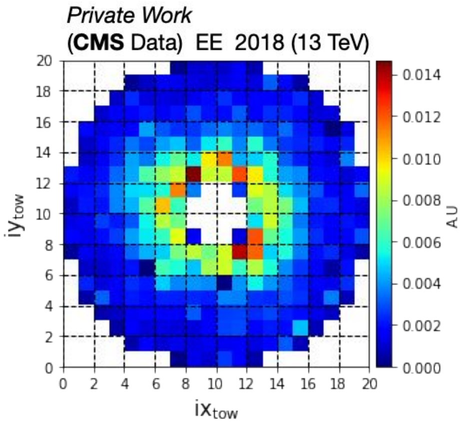
- Baseline studies for comparison performed for barrel.
- <u>Baseline loss per tower:</u> compare each tower occupancy $t_{\varphi,\eta}$ to η-ring average occupancy $\langle t_\eta \rangle$.
  - Define baseline tower loss$_{\varphi,\eta}$ = $| t_{\varphi,\eta} - \langle t_\eta \rangle |$



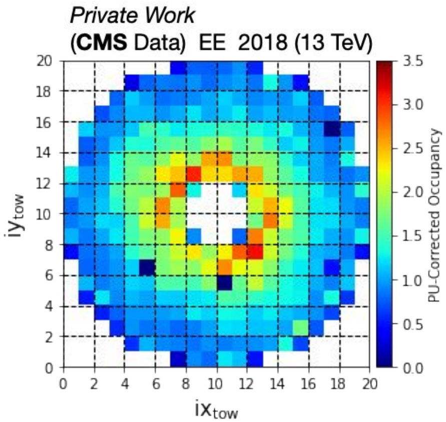
FDR for EB	Missing Supermodule		Zero Occupancy Tower		Hot Tower	
	AE	Baseline	AE	Baseline	AE	Baseline
<u>No correction</u>	3.6%	14%	51%	90%	2.8%	5.2%
<u>After time correction</u>	0.13%	5.9%	4.1%	80%	0.00%	0.00%

Spatial Correction

- Motivation: high FDR for zero occupancy towers
  - Low-loss towers are mostly around the *outer ring* of the endcap [1], related to the presence of *gradient* in the occupancy map [2].



[1] Average map for zero occupancy tower losses

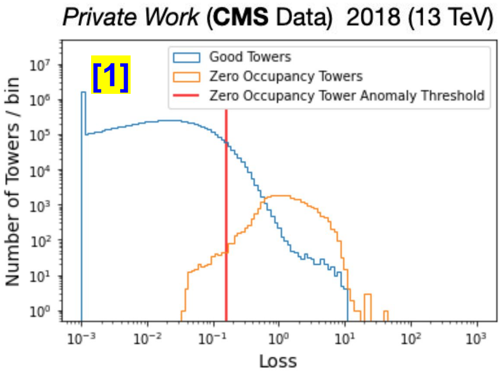


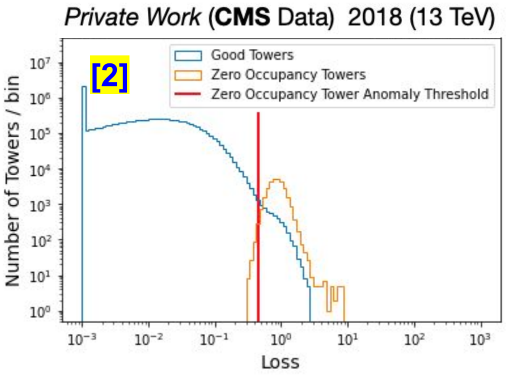
[2] Average occupancy map

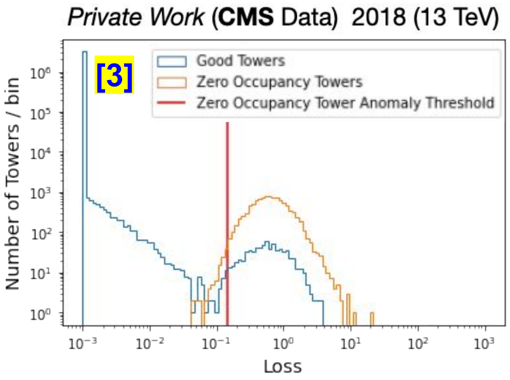
➔ **Spatial correction**: normalize the loss map with average occupancy map.

1D Loss Histograms after Each Corrections

- Loss histograms for EE–: before [1], after spatial [2], and time [3] corrections.





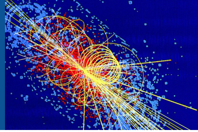


.

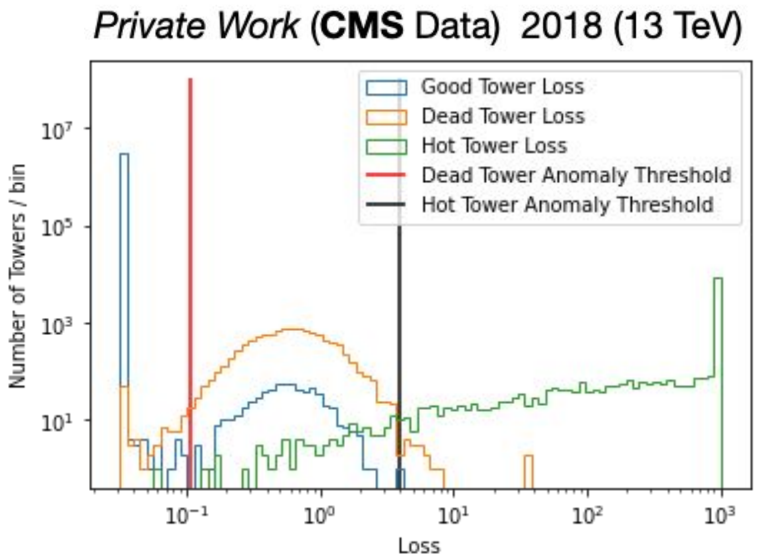
FDR for **EE–**	**Missing Sector**	**Zero Occupancy Tower**	**Hot Tower**
No correction	27%	87%	< 0.01%
After spatial correction	1.9%	14%	0.07%
After spatial and time correction	0.23%	6.4%	< 0.01%

Remaining false detections come from the actual anomalies that fell into the “good” dataset

Application for Actual Runs



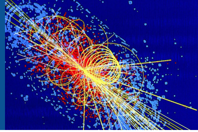
- Choose final anomaly threshold that can catch all anomalies considered



$$\textbf{FDR} = \frac{\text{Number of good towers above anomaly threshold}}{\text{Number of good and bad towers above anomaly threshold}}$$

→ Zero occupancy tower (dead tower) threshold can catch both dead and hot towers.

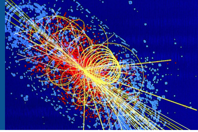
Further Studies with 2018 Data



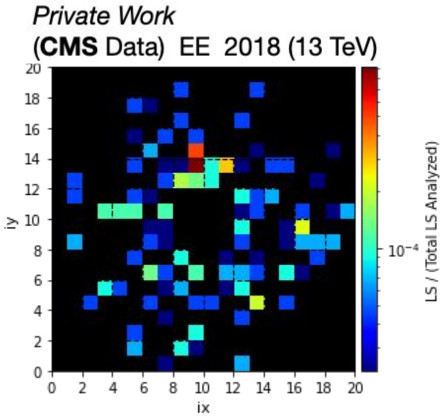
- Validation studies with the “fake” anomalies on 2018 dataset, but with the following changes:
  - Split dataset into each run era for 2018 (Run[A-D]). Average occupancy map used for spatial correction is also obtained from each corresponding run era.
  - Used the final anomaly threshold (more conservative one – dead tower anomaly threshold) to predict both dead and hot towers.

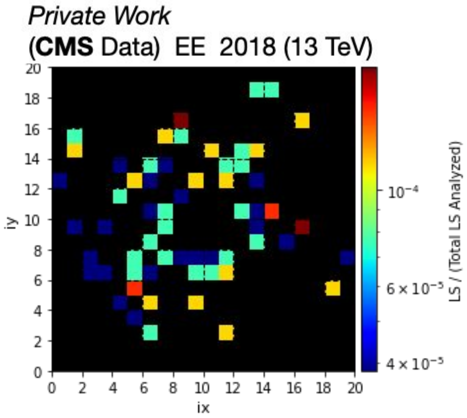
FDR after all corrections		**RunA**	**RunB**	**RunC**	**RunD**
Dead Tower	**EE+**	0.020	0.014	*0.087*	0.010
	EE–	0.020	0.013	*0.088*	0.022
Hot Tower	**EE+**	0.020	0.014	*0.086*	0.010
	EE–	0.020	0.013	*0.087*	0.021

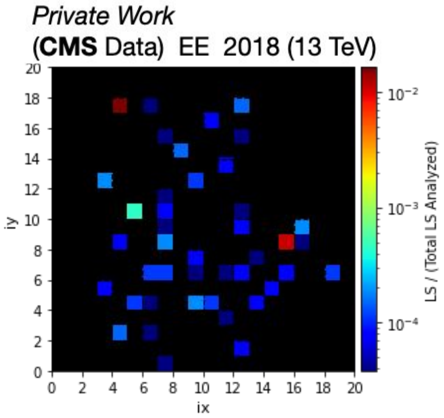
Further Studies with 2018 Data

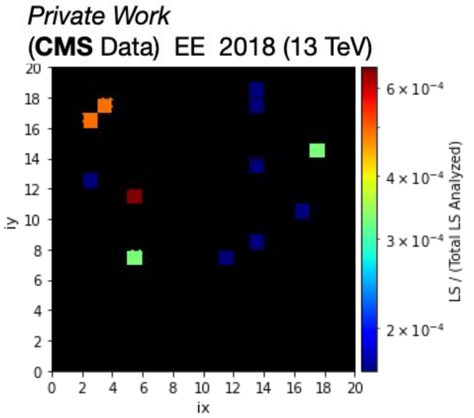


- Check which towers are contributing to the FDR for **EE+**.
  - Check <u>false positives</u> in FDR
    → "good" towers that are not "fake" anomaly towers, with occupancy above anomaly threshold.





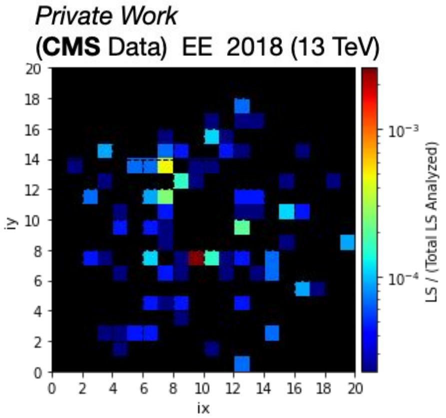


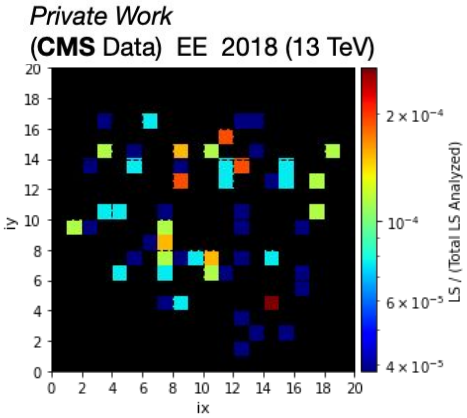


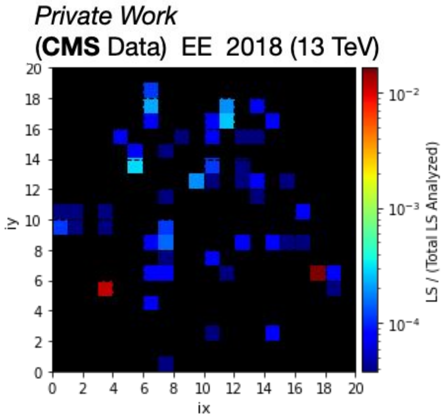
<u>FDR after all corrections</u>		**RunA**	**RunB**	**RunC**	**RunD**
Dead (Hot) Tower	**EE+**	0.020 (0.020)	0.014 (0.014)	*0.087 (0.086)*	0.010 (0.010)

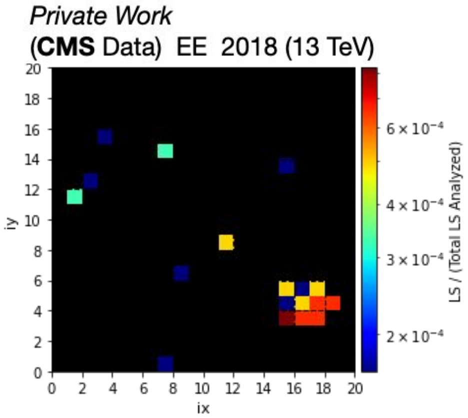
Further Studies with 2018 Data

- Check which towers are contributing to the FDR for **EE–**.
  - Check <u>false positives</u> in FDR
    → “good” towers that are not “fake” anomaly towers, with occupancy above anomaly threshold.









<u>FDR after all corrections</u>		**RunA**	**RunB**	**RunC**	**RunD**
Dead (Hot) Tower	**EE–**	0.020 (0.020)	0.013 (0.013)	*0.088 (0.087)*	0.022 (0.021)