# Status of OPERA open data at CERN

*S.Dmitrievsky, JINR, Dubna*

**Second International Workshop on Nuclear Emulsions
for Neutrino Studies and WIMP Search,**

*Anacapri, Italy, 2018/05/31*

# *CERN Open Data Portal*

## CERN Open Data Portal

**Documentation**  **About**

The CERN Open Data portal is the access point to a growing range of data produced through the research performed at CERN. It disseminates the preserved output from various research activities, including accompanying software and documentation which is needed to understand and analyse the data being shared.

The portal adheres to established global standards in data preservation and Open Science: the products are shared under open licenses; they are issued with a digital object identifier (DOI) to make them citable objects in the scientific discourse (see details below on how to do this).



**22nd of May 2018**: release of the first set of OPERA data samples.

OPERA became the first non-LHC experiment presented in the portal.

# *Preparation of the first data sample: $\nu_\mu$ sample*

The first data sample initially contained **818 $\nu_\mu$ events** used in the multiplicity analysis. It was prepared in May 2017 with help of **Ç.Kamışcıoğlu**. In Feb 2018 it was realized that one of the tau-candidate events (12123032048) was included in the sample by mistake. The event (with updated information) has been moved to the $\nu_\tau$ sample, so the total number of events in the multiplicity sample (before applying the cut for **$W^2$**) is now **817**!

From the multiplicity paper

## 3 Analysis

During the physics runs between 2008 and 2012, OPERA collected data corresponding to $1.8 \times 10^{20}$ protons on target. The electronic detectors recorded 19,505 neutrino interactions in the target fiducial volume. The search of the neutrino vertex in the first and second most probable bricks plus some additional selections (see [18] for more details) resulted in a sample of 5603 located events out of which 4406 have an identified muon. For the present measurement an unbiased sub-sample of 818 events occurring in the lead with a negatively charged muon identified by the muon spectrometer was selected in order to measure the track and vertex parameters in the target including a detailed check of the nuclear break-up and evaporation processes.

3

# Preparation of the second data sample: $\nu_\tau$ sample

General information about the 10 $\nu_\tau$-candidate events (*event Id, vertex brick Id, ECC-track types, etc.*) was kindly provided by **G.Galati**.

Feedback files with parameters of vertices and tracks found in ECC were kindly provided by **A.Ariga** and **T.Ariga**.

Some missing information (EM shower track, etc.) was kindly recovered by **A.Di Crescenzo**.

Also I thank **S.Vasina** who helped me to find information about energy reconstructed in the events.

Electronic detectors data (*hit positions, amplitudes, etc.*) for both data sets were prepared easily by using the **OpBrickFinder** package.

Preparation of the ECC data for the $\nu_\tau$ sample required a lot of manual handling. To obtain nice pictures of the tau-candidate events all tracks in the feedback files were processed one by one in order to align consecutive track segments to each other and to recover some missing segments.

# *OPERA ED and ECC data for the two data samples*

**_opendata.cern.ch/search?page=1&size=20&experiment=OPERA&file_type=zip_**

**Filter by experiment**
- ☑ OPERA     4

**Filter by year**
- ☐ 2009-2012     2
- ☐ 2010-2012     2

**Filter by file type**
- ☐ csv     831
- ☐ pdf     2
- ☑ zip     4

## Electronic detector data for multiplicity studies

The dataset was extracted from the official OPERA data repository. It contains 817 muon neutrino interactions with the lead target where a muon was reconstructed in the final state. This happens in th...

`Dataset` `Derived` `OPERA`

## Electronic detector data for tau neutrino appearance studies

This dataset was extracted from the official OPERA data repository and it contains all the data of the electronic detectors for the ten tau neutrino candidates, identified after an extensive analys...

`Dataset` `Derived` `OPERA`

## Emulsion data for neutrino tau appearance studies

This dataset was extracted from the official OPERA data repository and it contains all the emulsion data information for the ten tau neutrino candidates, identified after an extensive analysis that...

`Dataset` `Derived` `OPERA`

## Emulsion data for track multiplicity

The dataset was extracted from the official OPERA data repository. It contains 817 muon neutrino interactions with the lead target where a muon was reconstructed in the final state. This happens in th...

`Dataset` `Derived` `OPERA`

# *csv-files*: Electronic detector data for $\nu_\mu$ and $\nu_\tau$ samples

11 csv-files[*] for each event: EventInfo.csv, RawTTHitsXZ.csv, RawTTHitsYZ.csv, FilteredTTHitsXZ.csv, FilteredTTHitsYZ.csv, RawRPCHitsXZ.csv, RawRPCHitsYZ.csv, FilteredRPCHitsXZ.csv, FilteredRPCHitsYZ.csv, RawDTHitsXZ.csv, FilteredDTHitsXZ.csv

---

## Dataset Semantics    (Description of the variables in the csv-files):

**amplL:** PMT amplitude measured from the "left" side of a scintillator strip (in photo-electrons)

**amplR:** PMT amplitude measured from the "right" side of a scintillator strip (in photo-electrons)

**amplRec:** PMT amplitude reconstructed from the "left" and "right" side amplitudes of a scintillator strip taking into account light attenuation in a WLS fiber (in photo-electrons)

**clLength:** cluster length (in cm)

**driftDist:** drift distance (in cm)

**enHad:** energy of a hadron jet (in GeV)

**enNeu:** energy of a neutrino (in GeV)

**enVis:** visible energy (in MeV)

**evID:** event Id (11-digit number)

**muMom:** momentum of a muon (in GeV/c)

**posX:** X position of a drift tube, RPC, Target Tracker hit in the OPERA detector system of reference (in cm)

**posY:** Y position of an RPC hit in the OPERA detector system of reference (in cm)

**posZ :** Z position of a drift tube, RPC, Target Tracker hit in the OPERA detector system of reference (in cm)

**timestamp:** event time in milliseconds since 01/01/1970

---

[*] csv-file is a delimited text file that uses a comma to separate values.

# *csv-files: Emulsion data for the $\nu_\mu$ sample*

2 csv-files for each of 817 events: Tracks.csv, Vertex.csv

## Dataset Semantics (Description of the variables in the csv-files):

**evID:** event Id (11-digit number)

**globPosX:** X position of a vertex in the OPERA detector system of reference (in cm)

**globPosY:** Y position of a vertex in the OPERA detector system of reference (in cm)

**globPosZ:** Z position of a vertex in the OPERA detector system of reference (in cm)

**mult:** number of ECC tracks attached to the vertex

**posX:** X position of a track/vertex in the OPERA brick system of reference (in micrometers)

**posY:** Y position of a track/vertex in the OPERA brick system of reference (in micrometers)

**posZ:** Z position of a track/vertex in the OPERA brick system of reference (in micrometers)

**slopeXZ:** tangent of a track angle in XZ view

**slopeYZ:** tangent of a track angle in YZ view

**timestamp:** time in milliseconds since 01/01/1970

**trType:** type of a track: 1 - muon; 2 - hadron; 3 - electron; 4 - black; 5 - back black; 6 - gray; 7 - back gray

# *csv-files: Emulsion data for the $\nu_\tau$ sample*

### 3 csv-files for each of 10 events: Tracks.csv, Vertices.csv, Lines.csv

**Dataset Semantics**   (Description of the variables in the csv-files):

**amplL:** PMT amplitude measured from the "left" side of a scintillator strip (in photo-electrons)

**amplR:** PMT amplitude measured from the "right" side of a scintillator strip (in photo-electrons)

**amplRec:** PMT amplitude reconstructed from the "left" and "right" side amplitudes of a scintillator strip taking into account light attenuation in a WLS fiber (in photo-electrons)

**clLength:** cluster length (in cm)

**driftDist:** drift distance (in cm)

**enHad:** energy of a hadron jet (in GeV)

**enNeu:** energy of a neutrino (in GeV)

**enVis:** visible energy (in MeV)

**evID:** event Id (10- or 11-digit number)

**globPosX:** X position of a vertex in the OPERA detector system of reference (in cm)

**globPosY:** Y position of a vertex in the OPERA detector system of reference (in cm)

**globPosZ:** Z position of a vertex in the OPERA detector system of reference (in cm)

**muMom:** momentum of a muon (in GeV/c)

**posX:** For Electronic Detector events, X position of a drift tube, RPC, Target Tracker hit in the OPERA detector system of reference (in cm). For Emulsion Detector events, X position of a track/vertex in the OPERA brick system of reference (in micrometers).

**posX1:** X position of the beginning of a line in the OPERA brick system of reference (in micrometers)

**posX2:** X position of the end of a line in the OPERA brick system of reference (in micrometers)

**posY:** For Electronic Detector events, Y position of an RPC hit in the OPERA detector system of reference (in cm). For Emulsion Detector events, Y position of a track/vertex in the OPERA brick system of reference (in micrometers).

**posY1:** Y position of the beginning of a line in the OPERA brick system of reference (in micrometers)

**posY2:** Y position of the end of a line in the OPERA brick system of reference (in micrometers)

**posZ :** For Electronic Detector events, Z position of a drift tube, RPC, Target Tracker hit in the OPERA detector system of reference (in cm). For Emulsion Detector events, Z position of a track/vertex in the OPERA brick system of reference (in micrometers).

**posZ1:** Z position of the beginning of a line in the OPERA brick system of reference (in micrometers)

**posZ2:** Z position of the end of a line in the OPERA brick system of reference (in micrometers)

**primary:** flag of a vertex: 1 - primary vertex; 0 - not primary vertex

**slopeXZ:** tangent of a track angle in XZ view

**slopeYZ:** tangent of a track angle in YZ view

**timestamp:** event time in milliseconds since 01/01/1970

**trType:** type of a track: 1 - muon; 2 - hadron; 3 - electron/positron; 8 - tau lepton

# *Disclaimer for the OPERA open data*

(from opendata.cern.ch/docs/about-opera)

• **The open data are released under the Creative Commons CC0 waiver. Neither OPERA nor CERN endorse any works, scientific or otherwise, produced using these data.**

• **All released data samples will have a unique DOI that you are requested to cite in any applications or publications.**

# *Browser-based event display*

**OPERA browser-based event display** was implemented by analogy with an existing web-interface for visualization of open data of the CMS experiment:

### *opendata.cern.ch/visualise/events/CMS*

The OPERA display includes two 2D-views of an event in ED + one simplified 3D-view of tracks found in ECC near the primary vertex.

CERN open data **GitHub** repository with the source code of the display web-application:

### *https://github.com/cernopendata/demobbed-viewer*

The working web-application can be run here:

### *opendata.cern.ch/visualise/events/opera*

*Please note that a modern browser (with the ES6 support) is needed to run the application!*

# *Basic functionality of the ED event display*

• browsing through the event list

• moving of a camera views (up, down, left, and right)

• general (un)zooming

• zooming to the area of the beginning of an event

• zooming to the area of a whole event

• zooming to the area of the whole detector (default)
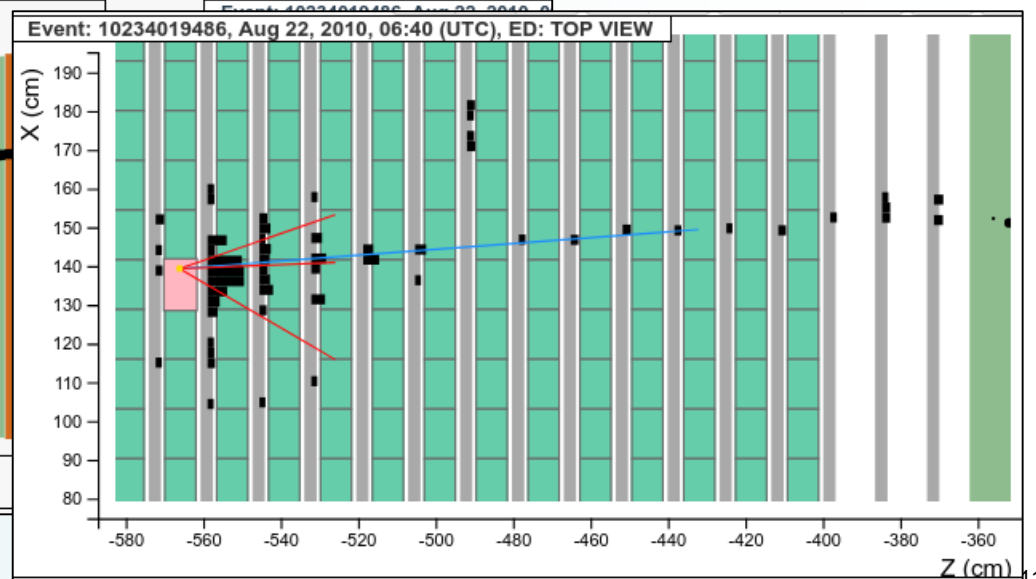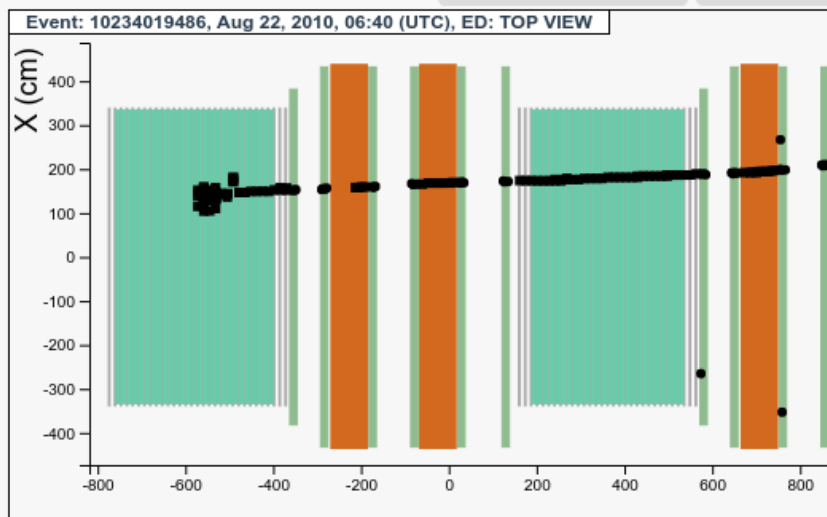
# *Basic functionality of the ECC event display*

(based on the **three.js** graphics library)

The **ECC-display** can be opened below the ED-display with the [⇐] button.

- browsing through the event list

- moving of a camera views (up, down, left, and right)

- general (un)zooming and also stretching along the vertical (Y) axis

- animated rotation of the ECC tracks around the neutrino interaction vertex
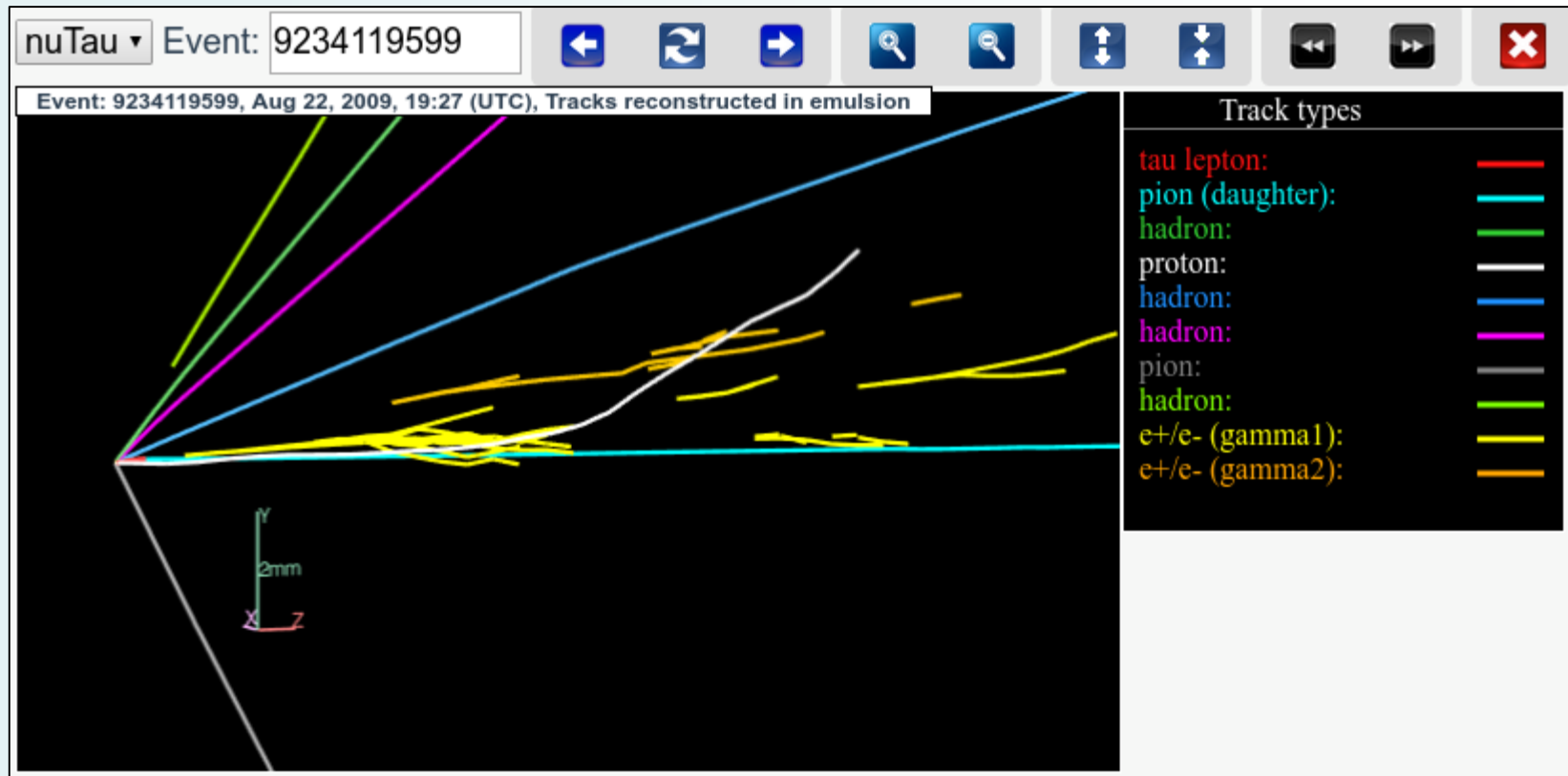
# *User manual for the event display*

A brief description of the event display info and all available button actions can be opened with the help [?] button.

[←] / [→] - go to the previous/next event (if any) in the selected sub-sample. Also the event Id (10- or 11-digit number) can be specified directly in the nearest edit box.

[↻] - reload the currently displayed event.

[🔍] / [🔍] - zoom in/out. Please note, that the default (whole detector) views of the ED display can't be zoomed out.

[■] - move the XZ and YZ camera views to the vertex brick. ECC-brick that contains the primary neutrino interaction vertex will be shown (an appropriate zooming will be adjusted automatically). The vertex position as well as projections of muon and some long hadron tracks found in the emulsion will be displayed as well.

[⊞] - move the XZ and YZ camera views and adjust zooming in order to display the whole neutrino event region.

[▥] - show the whole detector region (default option).

[↑] , [↓] , [←] , and [→] - move the camera views, correspondingly, up, down, left, and right.

[↰] / [✗] - open/close the ECC-event display.

[↕] / [↕] - stretch/compress the ECC-view in the vertical (Y) direction.

[▶▶] - start/accelerate animated rotation of the ECC-event around the Y (vertical) axis.

[◀◀] - decelerate/stop animated rotation of the ECC-event.

# *Summary*

- *Two first OPERA data samples* have been prepared and released in the CERN **Open Data Portal**:
    - the "multiplicity sample" (**817 $\nu_\mu$ events**),
    - the "tau appearance sample" (**10 $\nu_\tau$ candidate events**).

- *Browser-based event display* has been developed and integrated to the portal for visualization of the OPERA events.

- The feedback-files + recovered track info for the tau candidate events have been sent to **C.Bozza** for data preservation.

- Next steps to be discussed. E.g., preparation of the *$\nu_e$ data sample* for the Open Data project (see **Svetlana**'s talk).
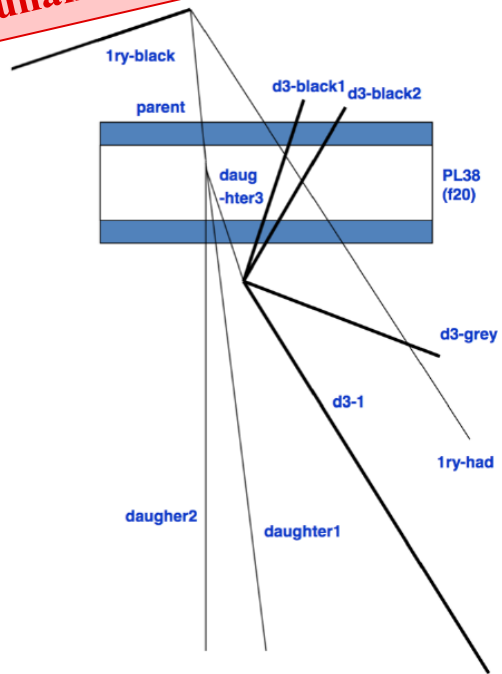
# *Backup slides*

# *Initial design of the event display*
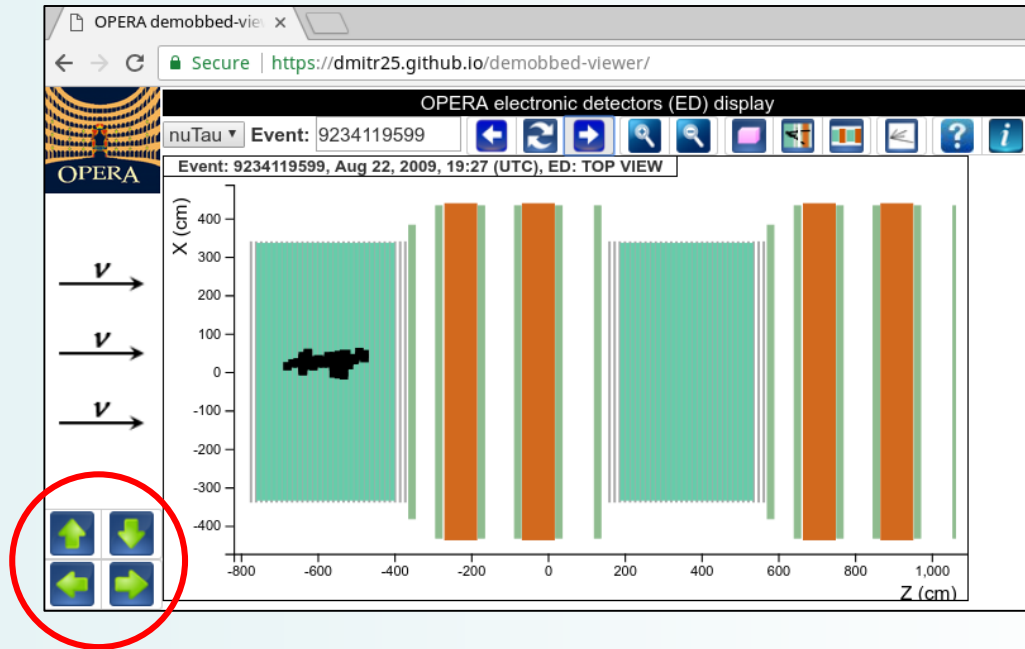
# *Example of missed ECC information*

*Event 11213015702:*



**From Giuliana's PhD thesis**

**Figure 5.7.** Event display for the Event 11213015702

3 tracks specified as "d3-black1", "d3-black2", and "d3-grey" in Giuliana's PhD thesis were not present in the provided feedback file.

**RECOVERED:** *missing track slopes were found in a report given by **Nakatsuka Yuji** in 2013.*