First look at data

Valentin Pestel
Nikhef KM3NeT group meeting
09/07/2020

What am I doing ? Playing with ORCA4 data !

- Work strongly based on Jannik's analysis : wiki, git
- Mainly for educational purposes

Work around 2 points mainly :

- Data and MC comparison and quality over time
- Understanding and improvements of neutrino selection

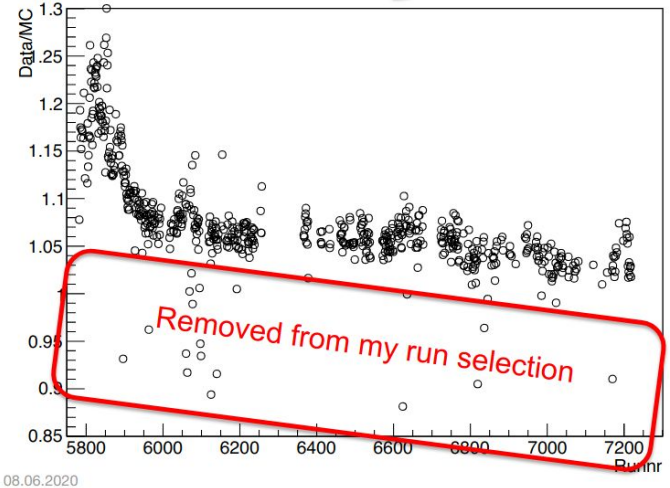
Data/MC(muons) ratio evolves over long time period :

data/MC ratio

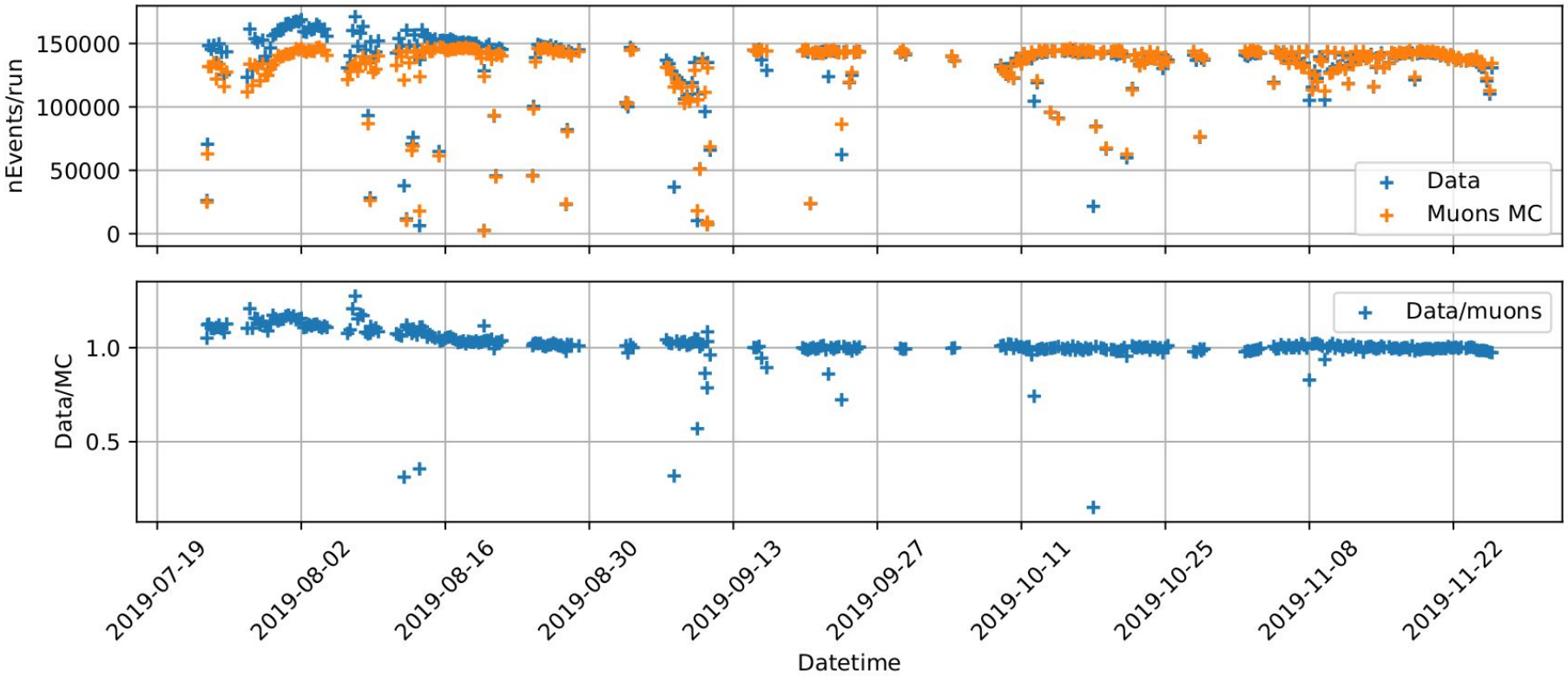
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- Number of events as a function of runnr, gives some observations:
- 1. there is a downward trend
- 2. there are outliers (data<mc)
- 3. there was a biolum burst ~5850

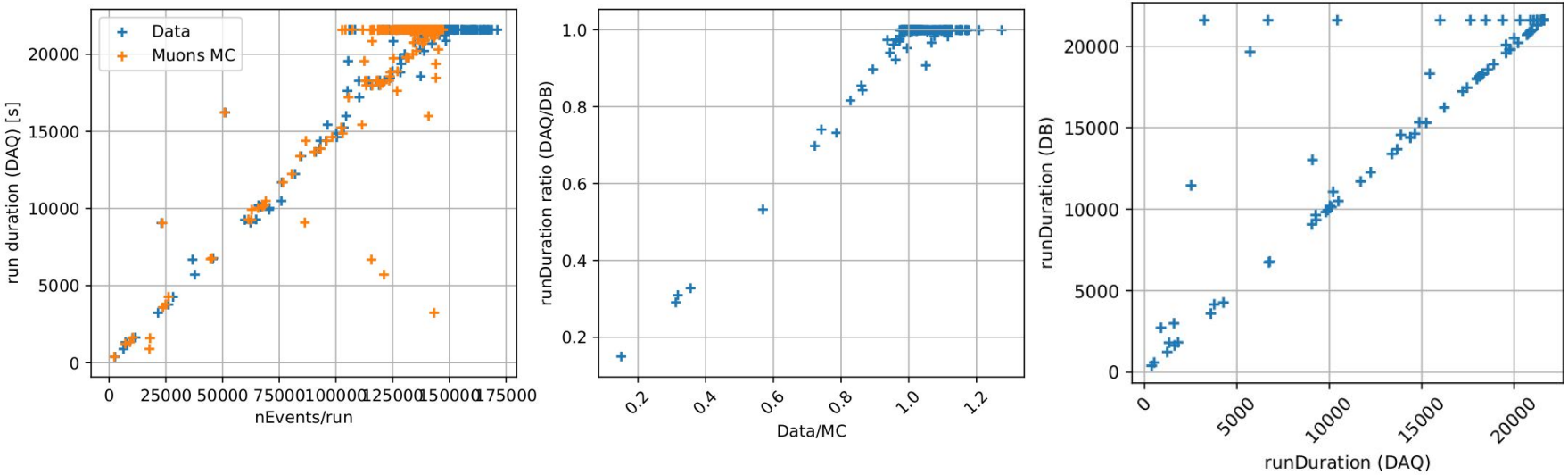
Should be followed up!



Time evolution



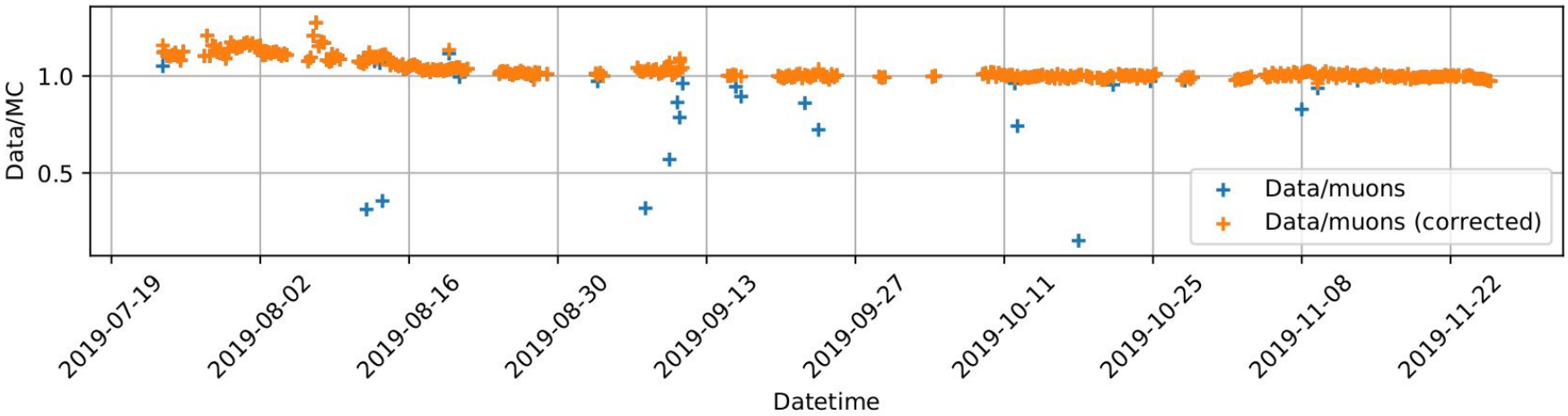
Outliers : wrong MC run duration ?



Two different run duration available :

- runDuration (DAQ) : available in the data header
- runDuration (DB) : available in the database

Very good correlation between Data/MC discrepancies and runDuration discrepancies ...



Data/Muons (corrected) = data / MC * runDuration(DB)/runDuration(DAQ)

Remove the outliers.
Can it be a wrong duration in MC production ?

Still to explain/study : Bio-luminescence event and long time decreasing trends.

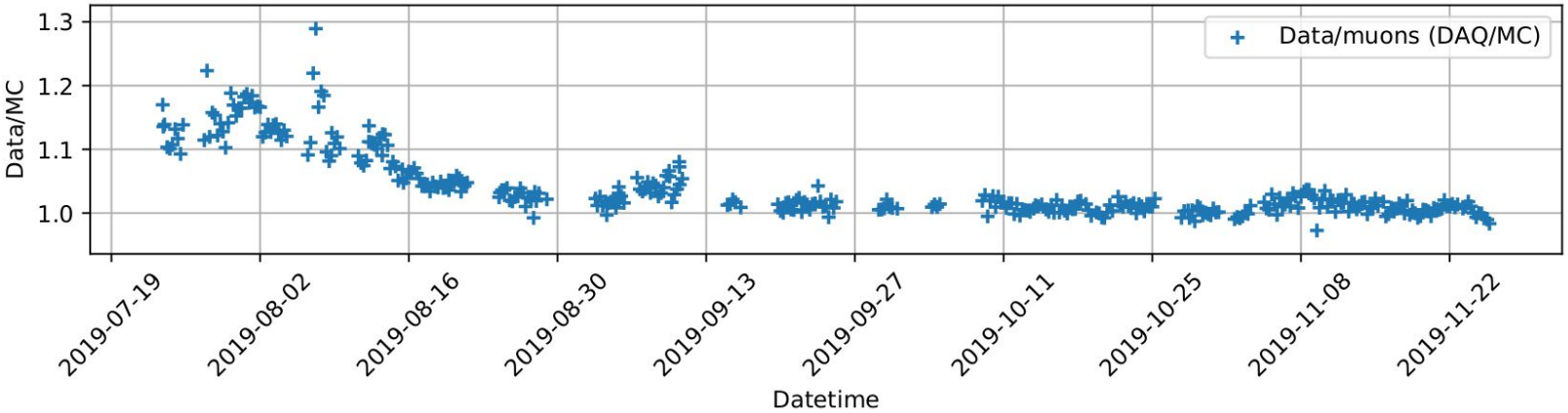
Correction with event weighting

Previous situation :

- MC muons weight = livetime (DB) / livetime(MC header)

New situation :

- MC muons weight = livetime (DAQ) / livetime(MC header)



Very few outliers now

Jannik's cuts summary

n_doms_cher_cond >= 5

Number of DOMs meeting Cherenkow condition (see "Cherenkov condition" - by Anna Sinopoulou at Caserta meeting, modified)

n_trig_doms_cher_cond >= 3

Same as previous but also triggered DOMs.

ntrighits_chercond_downFacing - ntrighits_chercond_upFacing >= 0

ntrighits_chercond_downFacing >= 5

nhits_chercond_downFacing - nhits_chercond_upFacing >= 10

nhits >= 20

Number of hits used in track reco (fitinf[3]). Remove pure noise events in data.

tot_max_trighit < 240

Qup/nhits>5-0.045*Qup

Qup ->best likelihood for upgoing track.

Qup/nhits>2.0

Qup-Qdn > 40

Qdn -> best likelihood for down going tracks

dz > 0

rtrk < 35 (m)

Track vertex distance to detector center

1.*(ztrk-lowest_dom_z+10)>rtrk || 1.*(ztrk-lowest_dom_z+10)>35.)

Quite empirical cut

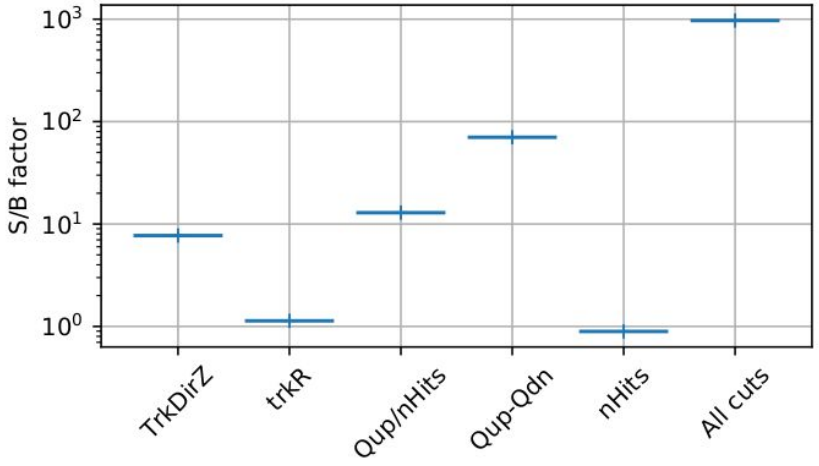
For now, focus on the green one

Cuts effects on data : (84 days of data)

	TrkDirZ	trkR	Qup/nHits	Qup-Qdn	nHits	All cuts
data	6.943	44.126	3.476	0.374	87.077	0.012
muons	4.532	41.981	1.810	0.410	92.782	0.013
nu mu CC	50.782	42.696	39.209	27.888	85.375	10.683
a nu mu CC	50.495	42.311	38.443	28.230	85.860	10.524
nu mu NC	55.451	56.709	51.140	22.568	69.392	11.248
nu e CC	62.870	74.642	63.580	20.924	52.616	13.900

% of events after cut

Cut impact on S/B ratio



Next step : look at the effect when all other cuts are applied.

Cuts effects on data : (84 days of data)

With these cuts :

- ~ 4.3 nu/day
- Full Jannik's cuts, 3 nu\day expected
  - eff = ~10-12% efficiency

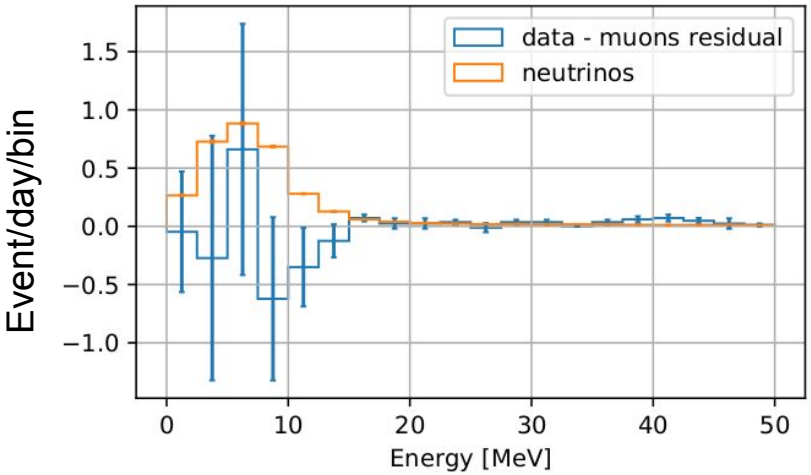
Events per day after cuts

data	70.964 +- 0.9188
muons	71.986 +- 1.5838
nu mu CC	1.713 +- 0.0086
a nu mu CC	1.158 +- 0.0091
nu mu NC	0.209 +- 0.0028
nu e CC	1.085 +- 0.0146
data - muons residual	-1.022 +- 1.831
neutrinos	4.165 +- 0.019

Only stat uncertainty

Too much muons ? :

- data - muons =/= neutrinos
- Still a problem in muon weighting ?



Perspectives

Data/MC time evolution :

- Try to correlate environnemental data and maybe calibration data to explore the long term decreasing trend
- Is there something in particular to look in bio-luminescence period ?

Neutrino signal extraction :

- Finish to reproduce Jannik's analysis (cuts not include yet, plots etc …)
- Try new variables (multivariate PID etc … e.g. Jerzy talk from collab meeting)

And wait for ORCA6 data ...

Data used :

Processing v5.40

- **data :** /sps/km3net/repo/data/KM3NeT_00000044/v5.40/reco/
- **mc Muons :** /sps/km3net/repo/mc/atm_muon/KM3NeT_00000044/v5.40/reco/
- **mc Nu :** /sps/km3net/repo/mc/atm_neutrino/KM3NeT_00000044/v5.40/reco/
    - mcv5.40.genhen_numuCC_10GeV.sirene.jterbr
    - mcv5.40.genhen_anumuCC_10GeV.sirene.jterbr
    - mcv5.40.gsg_muonNC-NC_1-500GeV.km3sim.jterbr
    - mcv5.40.gsg_elecCC-CC_1-500GeV.km3sim.jterbr

In this presentation, ~ 84 days of data