Chapitre X
Reconstruction de la masse d'une résonance grâce au *Machine Learning*

Sommaire

1 **Introduction** . . . 1
2 **HTT events dataset** . . . 2
2.1 generation with FastSim . . . 2
2.2 selection . . . 3
3 **XGBoost** . . . 4
3.1 Principle . . . 4
3.2 Training . . . 4
4 **DNN** . . . 5
4.1 Neurons . . . 5
4.2 Neural networks . . . 5
4.3 Training . . . 5
5 **Models and hyperparameters tuning** . . . 6
5.1 Selection using classic hyperparams . . . 6
6 **Discussions** . . . 10
6.1 Mass range . . . 10
6.2 PU effect . . . 10
6.3 Reco effects . . . 10
6.4 Fakes . . . 10
6.5 channel splitting . . . 10
6.6 Boundaries effects . . . 11
6.7 Final model . . . 11
7 **Use in the MSSM HTT analysis** . . . 11
8 **Conclusion** . . . 11

1 Introduction

AI = exploding for last years
Siri, autonomus cars, AlphaGO, ...
AI > Machine Learning > Deep Learning
principle of ML : get a program mapping inputs to outputs
classification vs regression tasks
in HEP : [1] and the one used for the SM categories in the analysis !
Also cite Gaël thesis :

G. Touquet. « Search for an additional neutral MSSM Higgs boson decaying to tau leptons with the CMS experiment ». Thèse de doct. Université Claude Bernard Lyon 1, oct. 2019. URL : https://hal.archives-ouvertes.fr/tel-02526393

and Mortiz's :
M. SCHAM. « Standard Model $H \to \tau\tau$ Analysis with a Neural Network Trained on a Mix of Simulation and Data Samples ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), juin 2020. URL : https://publish.etp.kit.edu/record/21993
and Tanja's :
T. KOPF. « Recoil Calibration as a Neural Network Task ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), fév. 2019. URL : https://publish.etp.kit.edu/record/21500

The Elements of statistical learning : Trees have one aspect that prevents them from bering the ideal tool for predictive learning, namely inaccuracy. –> they work great with the data used to create them, but they are not flexible when it comes to classifying new samples...

- **Aim :** reconstruction of di-τ mass.
  - Already done in :

P. BÄRTSCHI & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ». *Nuclear Instruments and Methods in Physics Research* **A929** (2019), p. 29-33. DOI : 10.1016/j.nima.2019.03.029. URL : http://www.sciencedirect.com/science/article/pii/S0168900219303377

P. BALDI, P. SADOWSKI & D. WHITESON. « Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning ». *Physical Review Letters* **114**.11 (mar. 2015). DOI : 10.1103/physrevlett.114.111801

- **BUT**
  - CMS simulated at phase-0 with Delphes $\to$ we used CMS Fast Simulation (FASTSIM).
  - No Pile-Up $\to$ we added it using the 2017 PU profile.
  - $m_h \in [80;300]$GeV per steps of 5 GeV $\to$ we do from 50 to 800 per steps of 1.
  - 270 000 training events $\to$ we have 2 180 992 ($\times 8$).
  - 100 000 testing events $\to$ we have 311 504 ($\times 3$).

Try both XGBOOST and DNN and compare, then apply to MSSM HTT analysis

✎ *To cite :*

- DELPHES *3.4.2 [7, 8] ?*
- *CMS Fast Simulation (*FASTSIM*) [9-12]*
- PYTHIA *8.235 [13]*
- FASTJET *[14, 15]*
- KERAS *[16]*
- TENSORFLOW *[17]*
- XGBOOST *[18]*
- *for an example of nn use in HEP*
  D. GUEST & coll. « Jet flavor classification in high-energy physics with deep neural networks ». Physical Review ***D94**.11 (déc. 2016).* DOI *: 10.1103/physrevd.94.112002*
- *W. SARLE. « Neural Networks and Statistical Models ». 1994.* URL *: https://people.orie.cornell.edu/davidr/or474/nn_sas.pdf*
- *P. BÄRTSCHI & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ».* Nuclear Instruments and Methods in Physics Research ***A929** (2019), p. 29-33.* DOI *: 10.1016/j.nima.2019.03.029.* URL *: http://www.sciencedirect.com/science/article/pii/S0168900219303377*
- SVFIT *[20]*

2 HTT events dataset

2.1 generation with FASTSIM

Give the settings used.
Also cite the followings :

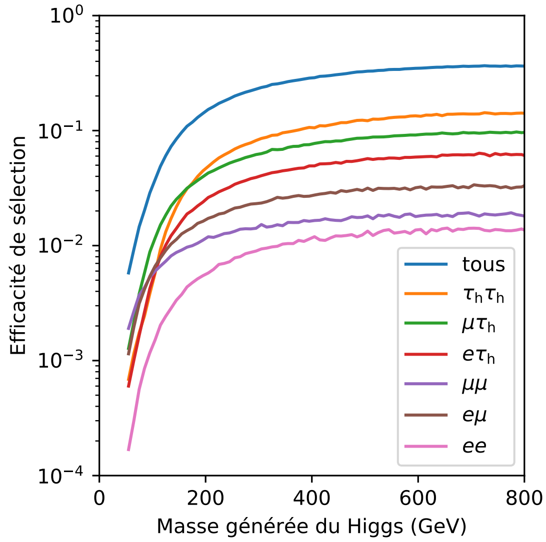
S. ABDULLIN & coll. « The Fast Simulation of the CMS Detector at LHC ». *Journal of Physics : Conference Series* **331**.3 (déc. 2011). DOI : 10.1088/1742-6596/331/3/032049

A. GIAMMANCO. « The Fast Simulation of the CMS Experiment ». *Journal of Physics : Conference Series* **513**.2 (juin 2014). DOI : 10.1088/1742-6596/513/2/022012

M. KOMM. « Fast emulation of track reconstruction in the CMS simulation ». *Journal of Physics : Conference Series* **898** (oct. 2017). DOI : 10.1088/1742-6596/898/4/042034

S. SEKMEN. *Recent Developments in CMS Fast Simulation*. 2017. arXiv : 1701.03850

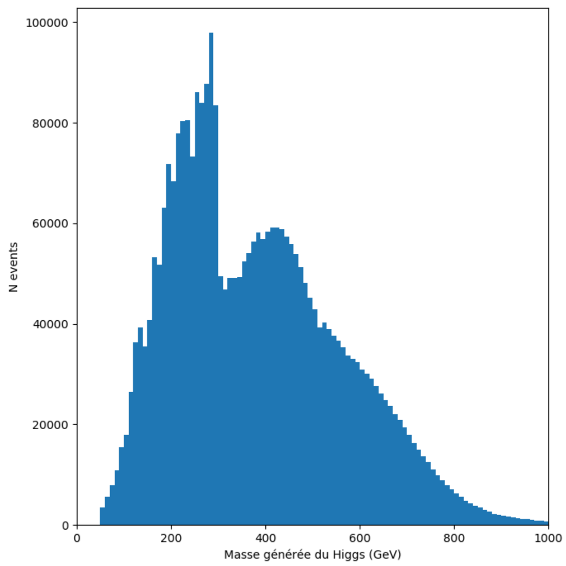
from 50 to 800 GeV
justify by showing plots when trained on smaller mass ranges
above problem not solved yet
below the selections (detailed after) drops everything (plot below)



amount of events : 60/20/10 k due to the cut efficiencies

2.2 selection

about the same as in the analysis (p_T cuts, DEEPTAU, ...)
+ ee and $\mu\mu$ channels as well !
mass distribution obtained



weights per steps of 2 GeV
70/20/10 % fraction for train/valid/test

3 XGBoost

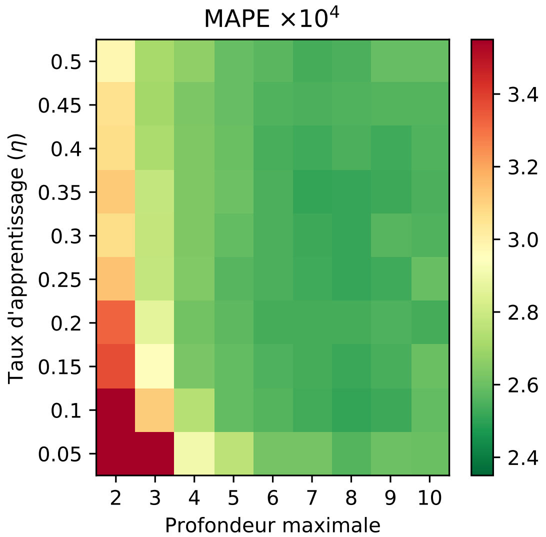
why ? faster a good at challenges (see Colin's slides from *somewhere in the past*)

3.1 Principle

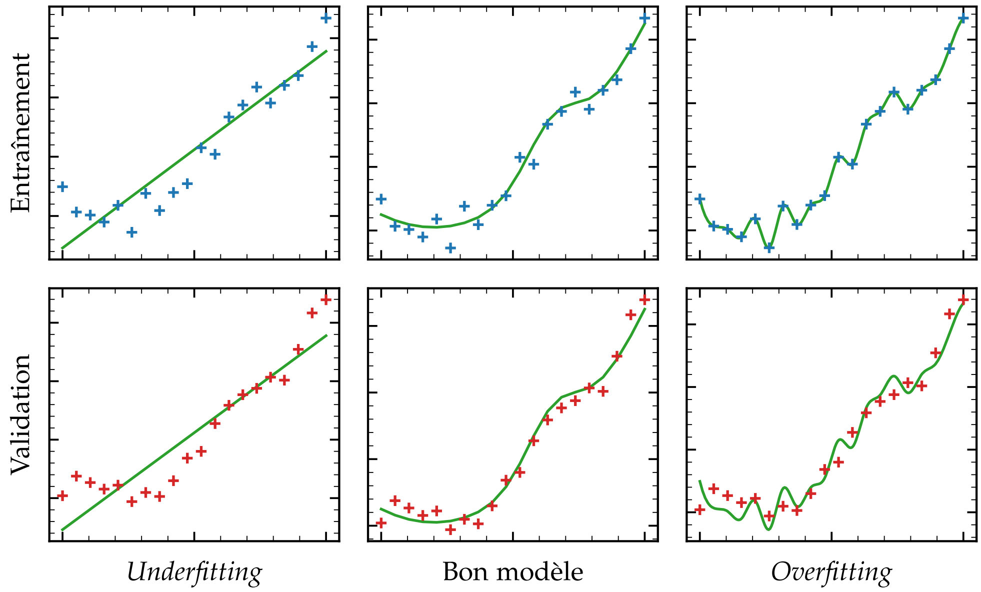
tree
max depth, n estimators

3.2 Training

objective
learning rate
early stopping



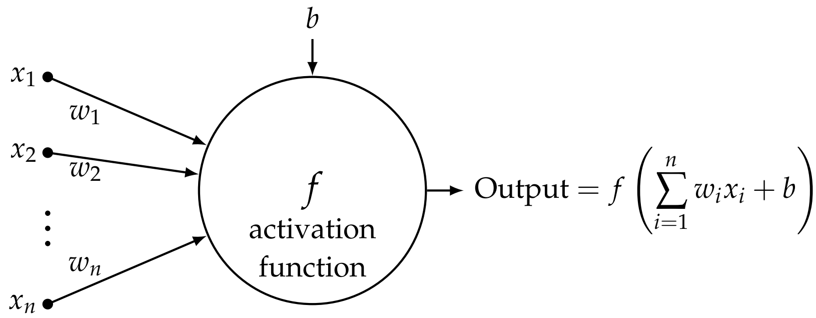
3.2.1 Overfitting and early stopping



4 DNN

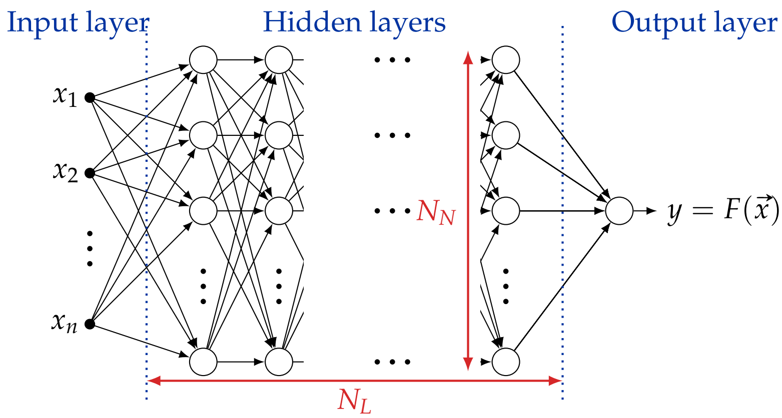
I. GOODFELLOW, Y. BENGIO & A. COURVILLE. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016

4.1 Neurons



Activation functions : tanh, sigmoïd mostly for classification, linear, relu, elu, selu, softmax, softplus ...

4.2 Neural networks



4.3 Training

4.3.1 Loss function

loss == objective
= 0 when prediction == truth
minimize it !

4.3.2 Optimizer and weights init

Adam, Adadelta, SGD
parameters to optimize = weights and biais
need to init : (Glorot) uniform/normal
Glorot : X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks", in Proceedings of the thirteenth international conference on artificial intelligence and statistics, p. 249. 2010.

5 Models and hyperparameters tuning

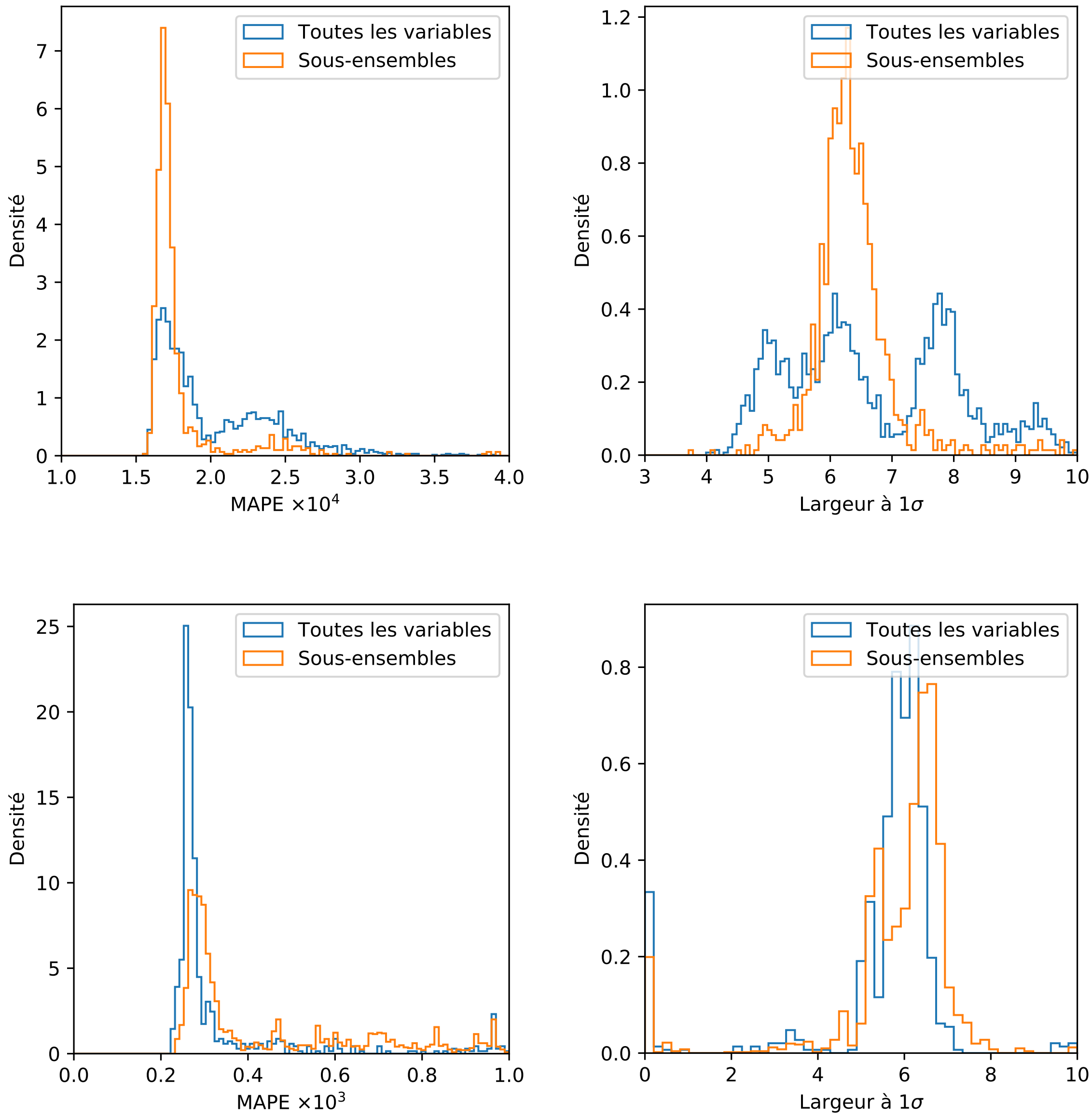
5.1 Selection using classic hyperparams

no customized loss yet.

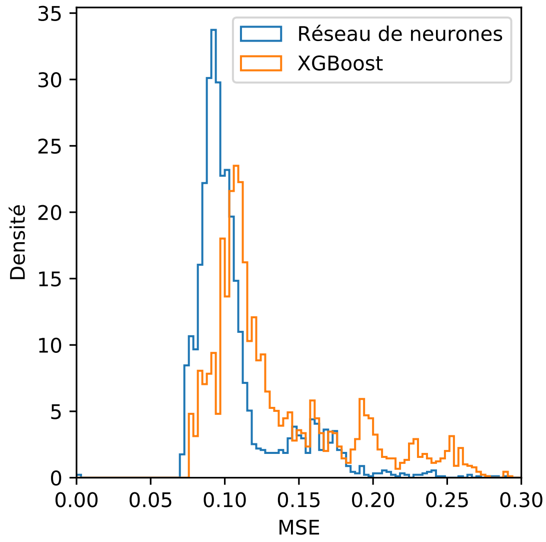
hard to get one single score to determine which model is the best : use mse, mae, mape, median diff, $\pm$ 1 or 2 σ width ... low, medium, high and full mass regions as well.

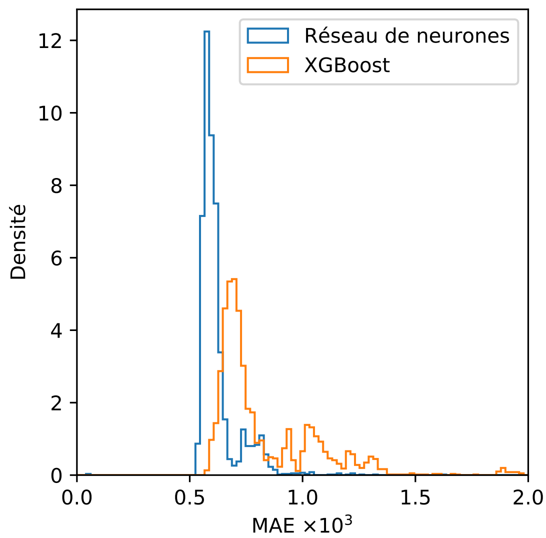
Model inputs : DNN not that sensible but XGB is better when having all of them, then use all inputs (give list) and not a subset of them.

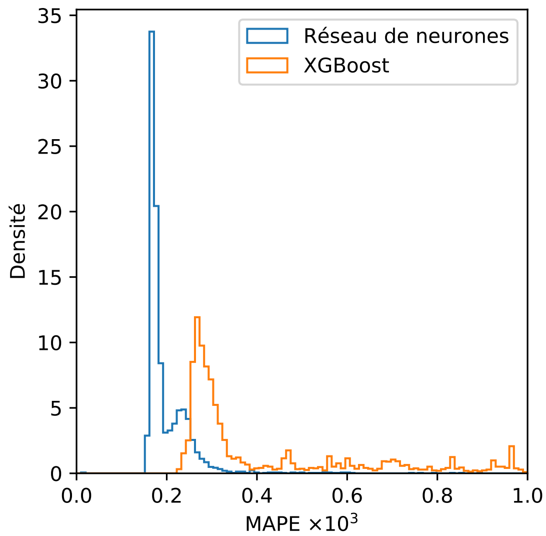
use plots with ref when relevant



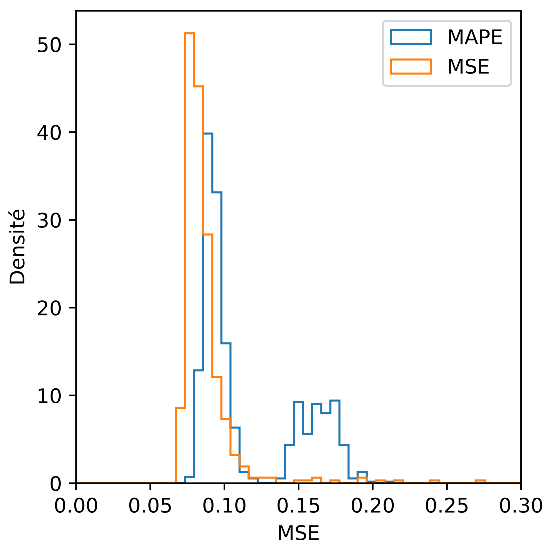
DNN vs XGB : use DNN!

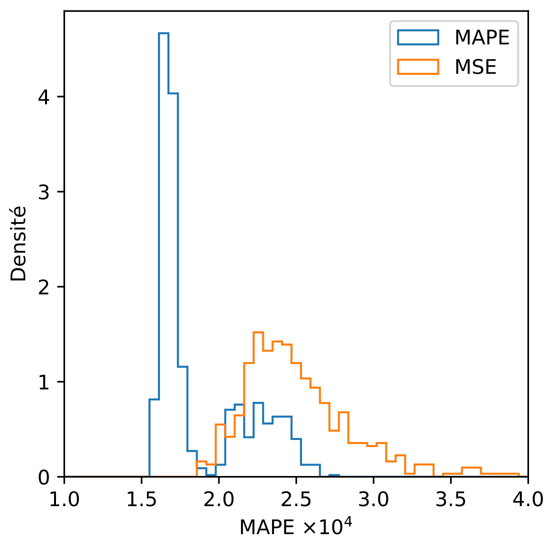


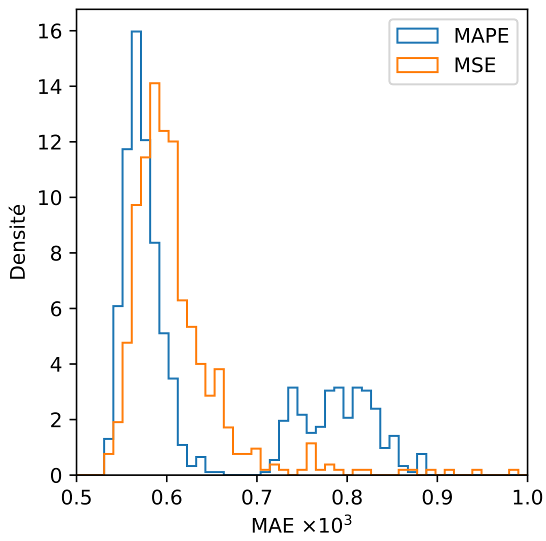




loss : when using a given loss, the corresponding models are of course better when using the loss as score.

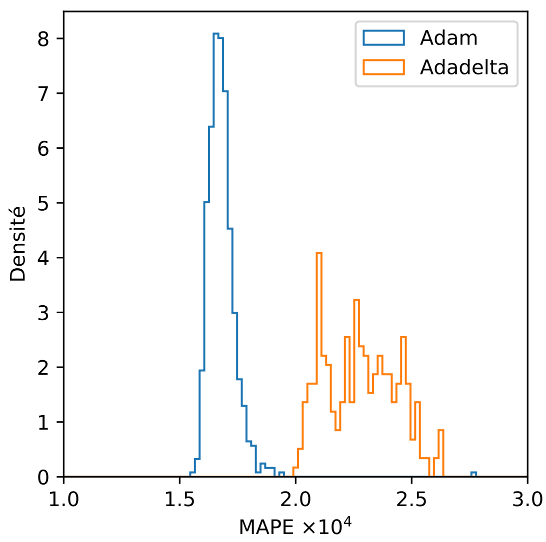




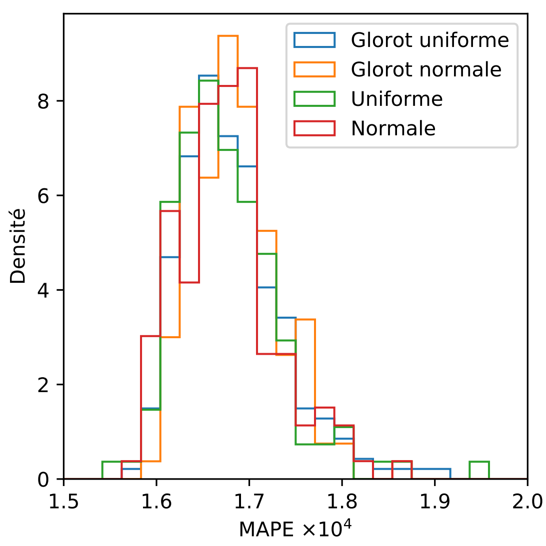


use mape loss, gives the better results

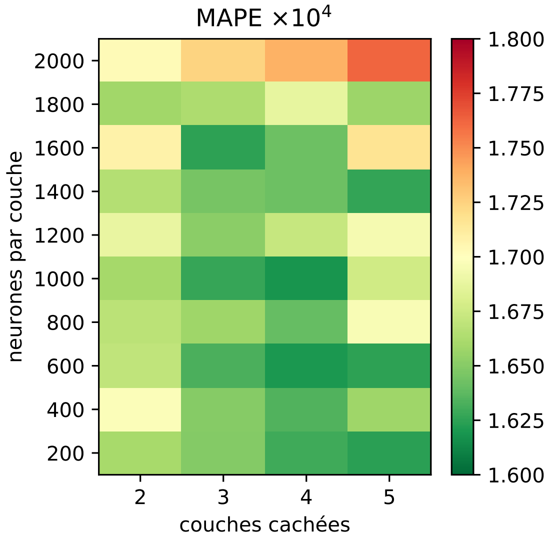
optimizer : Adam



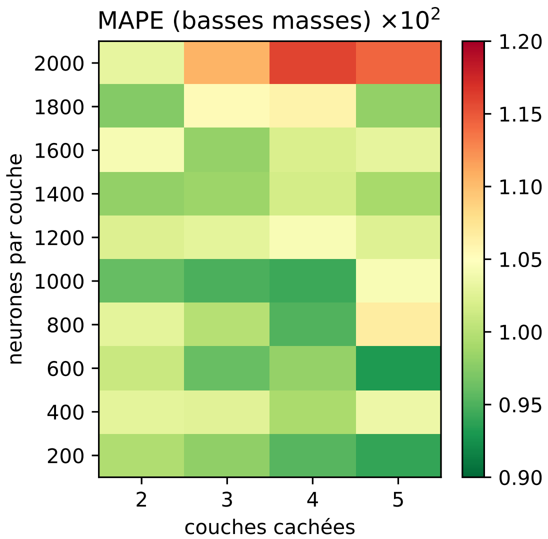
w init mode Glorot Uniform



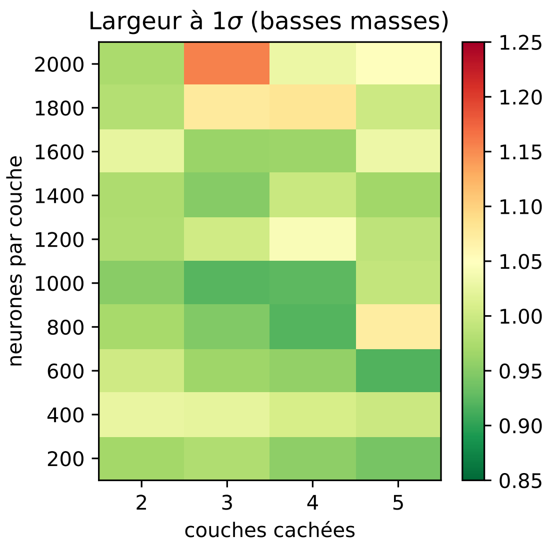
which structure ?



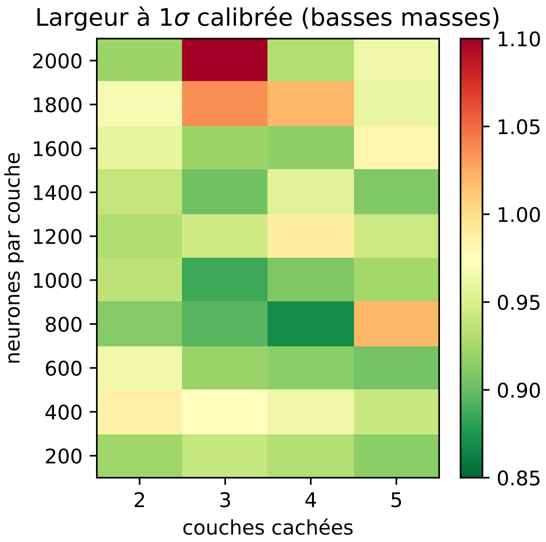
several possibilities, but the loss mass region contains the Z boson and is important



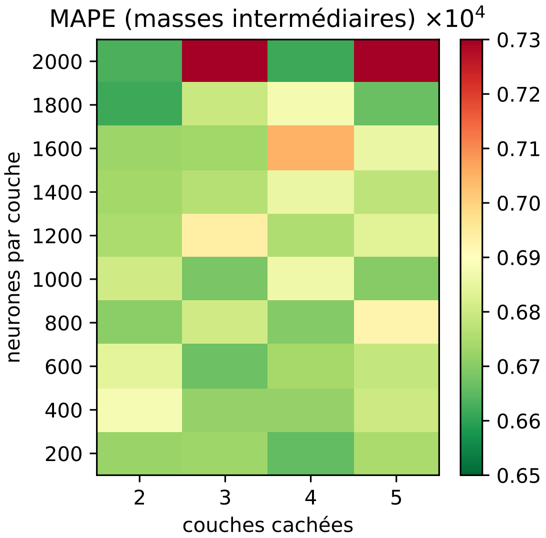
2x900 and 5x600 seem to be the best options, check the low mass resolution



5x600 seems good, check the low mass **calibrated** resolution



and in the medium mass region we have



3x1000 is the best compromise we found
activation = softplus

6 Discussions

6.1 Mass range

6.2 PU effect

also show PU effect (see fig 2 and 3 from report 2020-11-20, update with new models and samples)

6.3 Reco effects

show trained/tested on gen tau, gen tau decays, reco tau decays (=real), see fig 3 from report 2021-01-11
the model understand the physics, now it has to deal with the reco resolution and fakes.

6.4 Fakes

6.5 channel splitting

not relevant (fig3 report 2021-01-21)

6.6 Boundaries effects

use the custom loss with boundaries cuts (basically all the report 2021-02-04)

Follow report from 2021-02-04 but for section 3 : We saw that predictions come out too low, which already is a motivation to put larger weights on higher masses, i.e. to weight by truth. Choosing sqrt(truth) is of course just a guess then

extend up to 1TeV using the tails

6.7 Final model

DeepTau
1 TeV
all inputs
activation softplus
loss mapesqrt_b
opti Adam
glorot uniform
3 layers of 1000 neurons
show reponses and 2d histo

7 Use in the MSSM HTT analysis

show distributions of mTtot and ML predictions
discuss
show limits
discuss

8 Conclusion

Références

[1] D. Guest & coll. « Jet flavor classification in high-energy physics with deep neural networks ». *Physical Review* **D94**.11 (déc. 2016). doi : 10.1103/physrevd.94.112002.

[2] G. Touquet. « Search for an additional neutral MSSM Higgs boson decaying to tau leptons with the CMS experiment ». Thèse de doct. Université Claude Bernard Lyon 1, oct. 2019. url : https://hal.archives-ouvertes.fr/tel-02526393.

[3] M. Scham. « Standard Model $H \to \tau\tau$ Analysis with a Neural Network Trained on a Mix of Simulation and Data Samples ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), juin 2020. url : https://publish.etp.kit.edu/record/21993.

[4] T. Kopf. « Recoil Calibration as a Neural Network Task ». Mém. de mast. Fakultät für Physik des Karlsruher Instituts für Technologie (KIT), fév. 2019. url : https://publish.etp.kit.edu/record/21500.

[5] P. Bärtschi & coll. « Reconstruction of τ lepton pair invariant mass using an artificial neural network ». *Nuclear Instruments and Methods in Physics Research* **A929** (2019), p. 29-33. doi : 10.1016/j.nima.2019.03.029. url : http://www.sciencedirect.com/science/article/pii/S0168900219303377.

[6] P. Baldi, P. Sadowski & D. Whiteson. « Enhanced Higgs Boson to $\tau^+\tau^-$ Search with Deep Learning ». *Physical Review Letters* **114**.11 (mar. 2015). doi : 10.1103/physrevlett.114.111801.

[7] J. de Favereau & coll. « Delphes 3 : a modular framework for fast simulation of a generic collider experiment ». *Journal of High Energy Physics* **2** (fév. 2014). doi : 10.1007/jhep02(2014)057.

[8] A. Mertens. « New features in Delphes 3 ». *Journal of Physics : Conference Series* **608**.1 (2015). Sous la dir. de L. Fiala, M. Lokajicek & N. Tumova. doi : 10.1088/1742-6596/608/1/012045.

[9] S. Abdullin & coll. « The Fast Simulation of the CMS Detector at LHC ». *Journal of Physics : Conference Series* **331**.3 (déc. 2011). doi : 10.1088/1742-6596/331/3/032049.

[10] A. Giammanco. « The Fast Simulation of the CMS Experiment ». *Journal of Physics : Conference Series* **513**.2 (juin 2014). doi : 10.1088/1742-6596/513/2/022012.

[11] M. Komm. « Fast emulation of track reconstruction in the CMS simulation ». *Journal of Physics : Conference Series* **898** (oct. 2017). doi : 10.1088/1742-6596/898/4/042034.

[12] S. Sekmen. *Recent Developments in CMS Fast Simulation*. 2017. arXiv : 1701.03850.

[13] T. Sjöstrand & coll. « An Introduction to pythia 8.2 ». *Computer Physics Communications* **191** (2015), p. 159-177. doi : 10.1016/j.cpc.2015.01.024. arXiv : 1410.3012 [hep-ph].

[14] M. Cacciari, G. P. Salam & G. Soyez. « fastjet user manual ». *European Physical Journal* **C72** (nov. 2012). doi : 10.1140/epjc/s10052-012-1896-2. arXiv : 1111.6097 [hep-ph].

[15] M. Cacciari & G. P. Salam. « Dispelling the N^3 myth for the k_T jet-finder ». *Physics Letters* **B641**.1 (sept. 2006), p. 57-61. doi : 10.1016/j.physletb.2006.08.037.

[16] F. Chollet & coll. Keras. https://keras.io. 2015.

[17] M. Abadi & coll. TensorFlow *: Large-scale machine learning on heterogeneous distributed systems*. Software available from tensorflow.org. 2015. url : https://www.tensorflow.org/.

[18] T. Chen & C. Guestrin. « XGBoost : A Scalable Tree Boosting System ». *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (août 2016). doi : 10.1145/2939672.2939785.

[19] W. Sarle. « Neural Networks and Statistical Models ». 1994. url : https://people.orie.cornell.edu/davidr/or474/nn_sas.pdf.

[20] L. Bianchini & coll. « Reconstruction of the Higgs mass in $H \to \tau\tau$ Events by Dynamical Likelihood techniques ». *Journal of Physics : Conference Series* **513**.2 (juin 2014). doi : 10.1088/1742-6596/513/2/022035.

[21] I. Goodfellow, Y. Bengio & A. Courville. *Deep Learning*. http://www.deeplearningbook.org. MIT Press, 2016.