CDF/PHYS/EXOTIC/CDFR/XXXX
June 25, 2007
V1.00

Integrating the Matrix Elements into the Neural Nets for the H→ W^+W^- Analysis[1]

D. Benjamin[1], P. Bussey[2], D. Hidas[1], S. Hsu[3],M. Kruse[1]
E. Lipeless[3],M. Neubauer[3], A. Robson[2],
R. St. Denis[2], A.S. Thompson[2], F. Wuerthwein[3],W. Yao[5]
[1]Duke University
[2]Glasgow University
[3]University of California- San Diego
[4]University of Illinois- Urbana-Champaign
[5]Lawrence Berekely Laboratories

Abstract

Training of an artifical neural net using TMVA for H→ W^+W^- based on Monte Carlo prepared for 1 fb^{-1} of integrated luminosity is performed. Training using the twelve input variables used in the standard NN analysis is compared to training with the same twelve variables plus the matrix element compuation. No significant improvement in signal vs purity is observed as is seen in the low ranking (11 of 13) of the matrix element liklihood ratio relative to the dilepton mass, the leading lepton energy, scalar sum of the lepton transeverse momenta and the missing eneryg as well as the missing energy magnitude.

1 Introduction

Two approaches to assessing the contribution of H→ W^+W^- to dilepton events have been considered at CDF. The Matrix Element analysis uses likelihood ratios as described in [1] and the neural net analysis described in [2]

[1]Hww/doc/CombineMENN/MENNTraining.tex

trains artificial neural nets to distinguish the Higgs contribution. A statistical combination of the analyses [3] has been performed. This note describes that results of directly adding the ME value to the NN input.

The analysis is performed on the 1 fb^{-1} Monte Carlo dataset used in the ME analysis and in the NN analysis. Events come from the small ntuples made from the lepton selection done by the ME analysis group. A loose set of cuts against Drell-Yan is applied and the Matrix Element computed for a random portion of the events passing the cuts. The twelve input quantities used in the NN anlaysis are computed and written together with the Matrix Element liklihood ratio to a tree used as input to the TMVA analysis package.

The training has been done with Higgs having a mass of 160 $\mathrm{GeV/c^2}$ as signal and either a Pythia or MC@NLO WW sample as background. The neural net is trained both with no discrimination against lepton type and then three more times for the ee, eμ and $\mu\mu$ cases. For comparison, the training is also done without the Matrix Element. The net is trained with 200 epochs.

2 Results

The ranked importance of the input variables is listed in Table 3 for the neural net training with pythia but without the Matrix Element. The same ranked lists are provided in tables 3 for the training with pythia and using the Matrix Element Likelihood Ratio as input. The results of the same exercise for MC@NLO are shown in Tables **??** and **??**. It can be seen that the matrix element ranks 11 of 13 and two orders of magnitude weaker than the most discriminant variable, the dilepton mass.

The plots of the signal and background discrimnant values are given in Figures 17 to 32. Plots of rejection vs efficiency are given in Figures 1 to 16. There is slight improvement of the rejection vs efficiency when the matrix element is included. There is no strong dendency on lepton type.

Plots of the correlations amoungst the variables for the signal and background are given in Figures 33 to 64. For the purpose of discussion, the correlations for all leptons and with Pythia WW background are examined first. For the signal, the correlation matrix is shown in Figure 41 and for the Pythia WW background, the correlation matrix is shown in Figure 42. The correlations of all variables with the matrix element liklihood ratio, LRHWW, is summarized in Table 5.

For the signal, LRHWW is most strongly anticorrelated with “dildphi”, the azimuthal separation of the leptons, and “lepR”, the separation of the leptons in η–ϕ. A correlation of simliar strength is observed for the missing E_t significance, metsig, the magnitude of the missing E_t, metsig, and the scalar sum of the lepton transverse momenta with the missing E_t. There is a weaker anticorrelation with “dilmass” the dilepton mass with the leading and subleading lepton energies. Couriously there is no correlation with the leading lepton energy. There is also no correlation with the jet transverse momenta.

For the background, LRHWW, has strongest (anti)correlation with dildphi, dilmass and lepR. There is a weaker anticorrelation with the leading and subleading lepton energies. There is now a weak correlation with the two jet E_t values and the number of jets. The missing E_t magnitude and significance have a weaker correlation than for the signal.

The behaviour does not change significantly when one looks at individual leptons or at the MC@NLO results.

3 Conclusions and Futher Work

The analysis seems to indicate that there is little discrimination to be gained by simply adding the matrix element liklihood ratio computation to the neural net. There is potentially gains in understanding and hence in the determination of the systematic uncertainties and a study of the likelihood ratio and jet varaibles might be quite useful to undertake. This amounts to releasing or modifying some of the input variables so as to place more reliance on the ME computation.

A tighter preselection should be explored, matching both that in the ME and in the standard NN analysis.

References

[1] CDF 8719

[2] CDF 8700

[3]

Leptons	All		ee	
Rank	Variable	Separation	Variable	Separation
1	dilmass	6.137e+04	dilmass	1.176e+04
2	LepAEn	2.953e+04	addEt	3.393e+03
3	addEt	2.868e+04	LepAEn	3.058e+03
4	metmag	1.441e+04	metmag	1.620e+03
5	LepBEn	5.675e+03	LepBEn	1.176e+03
6	jet1_Et	3.875e+03	jet1_Et	6.952e+02
7	lepR	9.973e+02	lepR	1.672e+02
8	metsig	8.596e+02	metsig	1.419e+02
9	dildphi	5.901e+02	dildphi	1.029e+02
10	jet2_Et	3.321e+02	jet2_Et	6.399e+01
11	ntightjets	1.017e+02	ntightjets	1.779e+01
12	dPhiMetLJ	7.248e+01	dPhiMetLJ	1.327e+01

Leptons	eμ		$\mu\mu$	
Rank	Variable	Separation	Variable	Separation
1	dilmass	3.400e+04	dilmass	1.436e+04
2	addEt	1.232e+04	LepAEn	1.196e+04
3	LepAEn	1.149e+04	addEt	1.098e+04
4	metmag	6.243e+03	metmag	5.353e+03
5	LepBEn	3.003e+03	LepBEn	1.634e+03
6	jet1_Et	2.092e+03	jet1_Et	1.075e+03
7	lepR	5.498e+02	metsig	2.260e+02
8	metsig	4.810e+02	lepR	1.993e+02
9	dildphi	3.215e+02	dildphi	1.554e+02
10	jet2_Et	1.811e+02	jet2_Et	9.032e+01
11	ntightjets	5.422e+01	ntightjets	2.749e+01
12	dPhiMetLJ	4.206e+01	dPhiMetLJ	1.936e+01

Table 1: Ranked impact of parameters used in NN anlaysis with Pythia for WW

Leptons	All		ee	
Rank	Variable	Separation	Variable	Separationn
1	dilmass	6.137e+04	dilmass	1.176e+04
2	LepAEn	2.953e+04	addEt	3.393e+03
3	addEt	2.868e+04	LepAEn	3.058e+03
4	metmag	1.441e+04	metmag	1.620e+03
5	LepBEn	5.675e+03	LepBEn	1.176e+03
6	jet1_Et	3.875e+03	jet1_Et	6.952e+02
7	lepR	9.973e+02	lepR	1.672e+02
8	metsig	8.596e+02	metsig	1.419e+02
9	dildphi	5.901e+02	dildphi	1.029e+02
10	jet2_Et	3.321e+02	jet2_Et	6.399e+01
11	LRHWW	2.279e+02	LRHWW	3.945e+01
12	ntightjets	1.017e+02	ntightjets	1.779e+01
13	dPhiMetLJ	7.248e+01	dPhiMetLJ	1.327e+01

Leptons	eμ		$\mu\mu$	
Rank	Variable	Separation	Variable	Separation
1	dilmass	3.400e+04	dilmass	1.436e+04
2	addEt	1.232e+04	LepAEn	1.196e+04
3	LepAEn	1.149e+04	addEt	1.098e+04
4	metmag	6.243e+03	metmag	5.353e+03
5	LepBEn	3.003e+03	LepBEn	1.634e+03
6	jet1_Et	2.092e+03	jet1_Et	1.075e+03
7	lepR	5.498e+02	metsig	2.260e+02
8	metsig	4.810e+02	lepR	1.993e+02
9	dildphi	3.215e+02	dildphi	1.554e+02
10	jet2_Et	1.811e+02	jet2_Et	9.032e+01
11	LRHWW	1.238e+02	LRHWW	6.172e+01
12	ntightjets	5.422e+01	ntightjets	2.749e+01
13	dPhiMetLJ	4.206e+01	dPhiMetLJ	1.936e+01

Table 2: Ranked impact of parameters used in NN+ME(HWW) anlaysis with Pythia for WW

Leptons	All		ee	
Rank	Variable	Separation	Variable	Separation
1	dilmass	5.490e+04	dilmass	1.046e+04
2	addEt	1.827e+04	addEt	2.842e+03
3	LepAEn	1.507e+04	LepAEn	2.815e+03
4	metmag	1.051e+04	metmag	1.603e+03
5	LepBEn	4.655e+03	LepBEn	8.556e+02
6	jet1_Et	1.947e+03	jet1_Et	3.545e+02
7	lepR	9.860e+02	lepR	1.674e+02
8	metsig	7.726e+02	metsig	1.235e+02
9	dildphi	5.600e+02	dildphi	9.795e+01
10	jet2_Et	2.679e+02	jet2_Et	5.927e+01
11	dPhiMetLJ	8.782e+01	dPhiMetLJ	1.736e+01
12	ntightjets	5.523e+01	ntightjets	9.723e+00
Leptons	eμ		$\mu\mu$	
Rank	Variable	Separation	Variable	Separation
1	dilmass	3.040e+04	dilmass	1.280e+04
2	addEt	9.889e+03	addEt	5.269e+03
3	LepAEn	8.402e+03	LepAEn	3.682e+03
4	metmag	6.095e+03	metmag	2.774e+03
5	LepBEn	2.453e+03	LepBEn	1.437e+03
6	jet1_Et	1.075e+03	jet1_Et	5.723e+02
7	lepR	5.290e+02	metsig	2.098e+02
8	metsig	4.328e+02	lepR	2.035e+02
9	dildphi	3.027e+02	dildphi	1.487e+02
10	jet2_Et	1.457e+02	jet2_Et	7.896e+01
11	dPhiMetLJ	4.879e+01	dPhiMetLJ	2.273e+01
12	ntightjets	2.939e+01	ntightjets	1.469e+01

Table 3: Ranked impact of parameters used in NN anlaysis with MC@NLO for WW

Leptons	All		ee	
Rank	Variable	Separation	Variable	Separation
1	dilmass	5.490e+04	dilmass	1.046e+04
2	addEt	1.827e+04	addEt	2.842e+03
3	LepAEn	1.507e+04	LepAEn	2.815e+03
4	metmag	1.051e+04	metmag	1.603e+03
5	LepBEn	4.655e+03	LepBEn	8.556e+02
6	jet1_Et	1.947e+03	jet1_Et	3.545e+02
7	lepR	9.860e+02	lepR	1.674e+02
8	metsig	7.726e+02	metsig	1.235e+02
9	dildphi	5.600e+02	dildphi	9.795e+01
10	jet2_Et	2.679e+02	jet2_Et	5.927e+01
11	LRHWW	2.077e+02	LRHWW	3.568e+01
12	dPhiMetLJ	8.782e+01	dPhiMetLJ	1.736e+01
13	ntightjets	5.523e+01	ntightjets	9.723e+00
Leptons	eμ		$\mu\mu$	
Rank	Variable	Separation	Variable	Separation
1	dilmass	3.040e+04	dilmass	1.280e+04
2	addEt	9.889e+03	addEt	5.269e+03
3	LepAEn	8.402e+03	LepAEn	3.682e+03
4	metmag	6.095e+03	metmag	2.774e+03
5	LepBEn	2.453e+03	LepBEn	1.437e+03
6	jet1_Et	1.075e+03	jet1_Et	5.723e+02
7	lepR	5.290e+02	metsig	2.098e+02
8	metsig	4.328e+02	lepR	2.035e+02
9	dildphi	3.027e+02	dildphi	1.487e+02
10	jet2_Et	1.457e+02	jet2_Et	7.896e+01
11	LRHWW	1.127e+02	LRHWW	5.592e+01
12	dPhiMetLJ	4.879e+01	dPhiMetLJ	2.273e+01
13	ntightjets	2.939e+01	ntightjets	1.469e+01

Table 4: Ranked impact of parameters used in NN + ME (HWW) anlaysis with MC@NLO for WW

Variable	Correlation	
	Signal	Background
LepAEn	+0	−17
LepBEn	+26	−23
addEt	+44	−12
dPhiMEtLJ	+14	+18
dildphi	−43	−49
dilmass	−23	−45
jet1_Et	+0	+11
jet2_Et	+0	+5
lepR	−46	−55
metmag	+43	+10
metsig	+43	+26
ntightjets	+7	+12

Table 5: Correlations of variables with the Matrix Element liklihood ratio for all leptons trained using Pythia WW background.

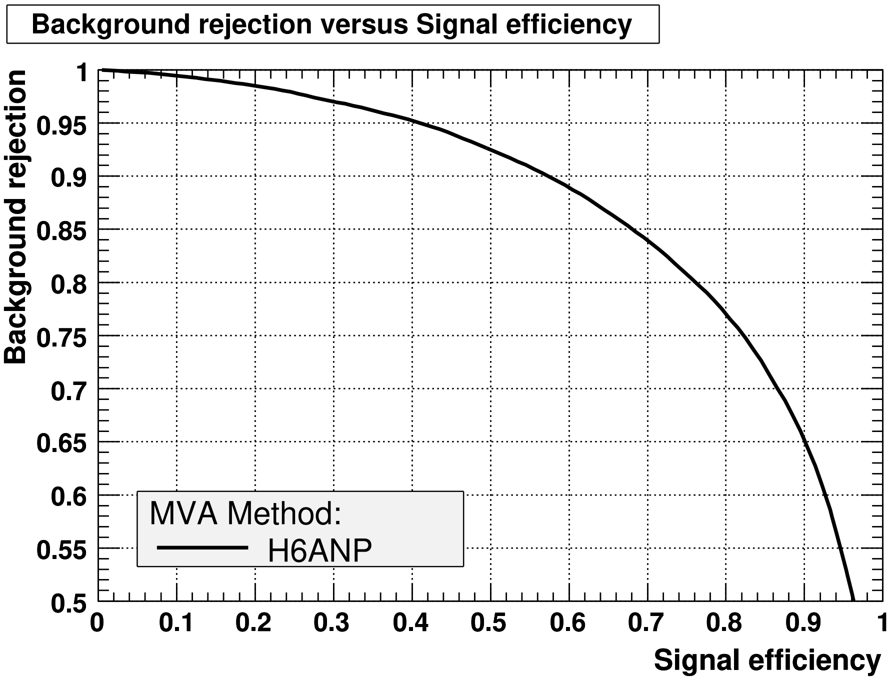

Figure 1: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, All leptons, NN variables, Pythia WW.

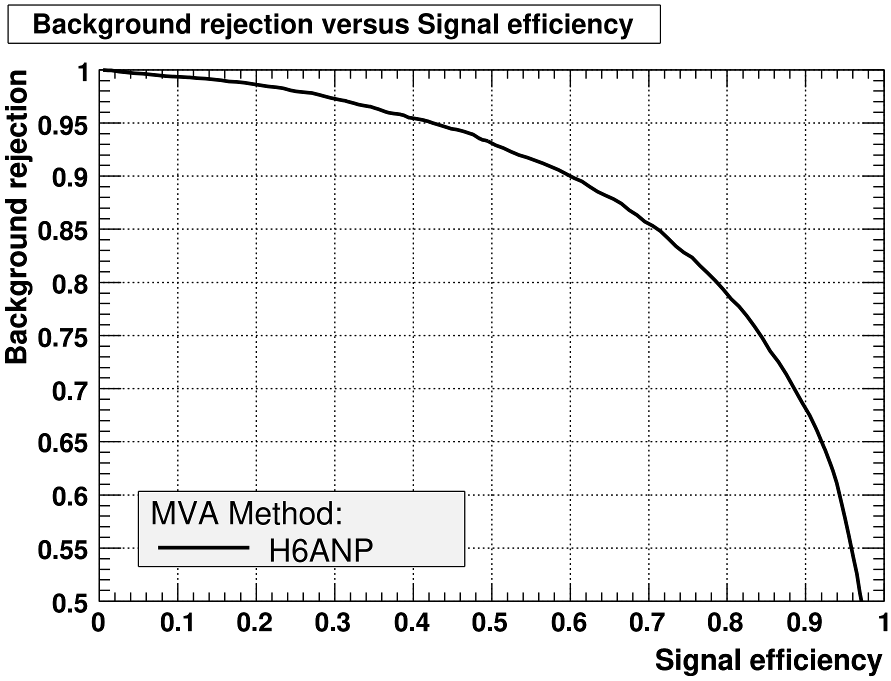

Figure 2: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, ee, NN variables, Pythia WW.

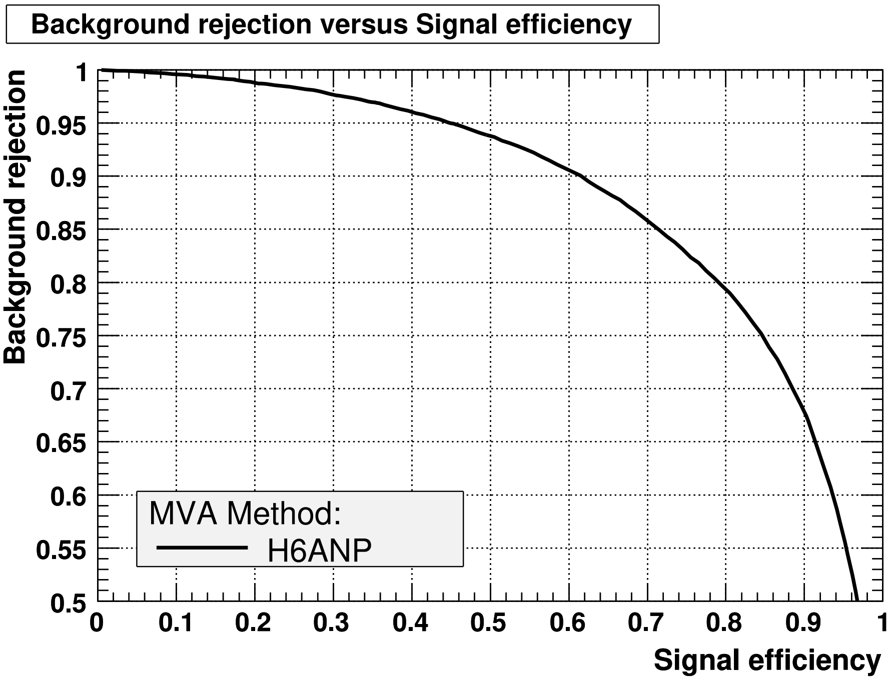

Figure 3: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, eμ, NN variables, Pythia WW.

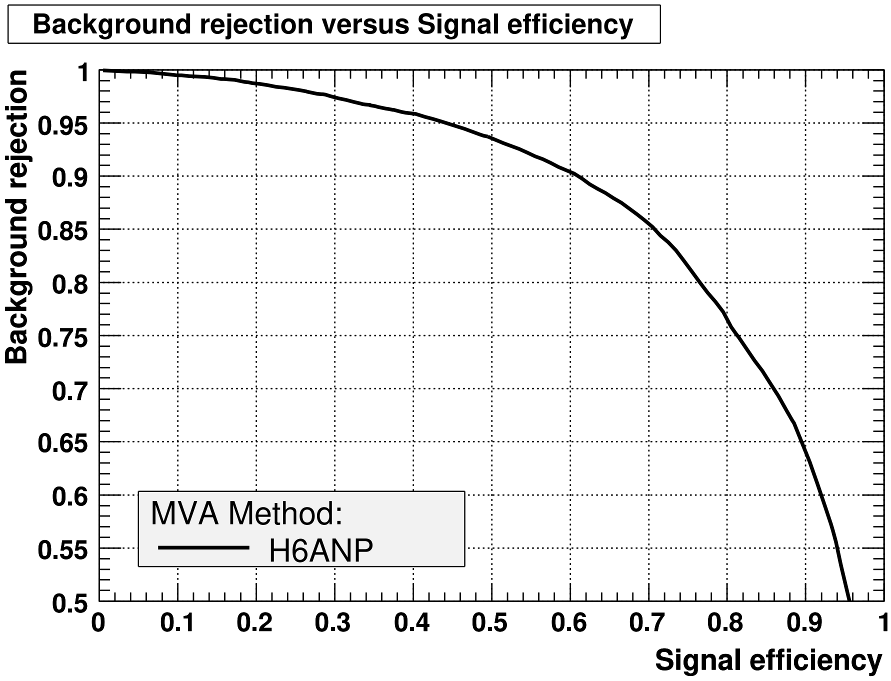

Figure 4: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, $\mu\mu$, NN variables, Pythia WW.

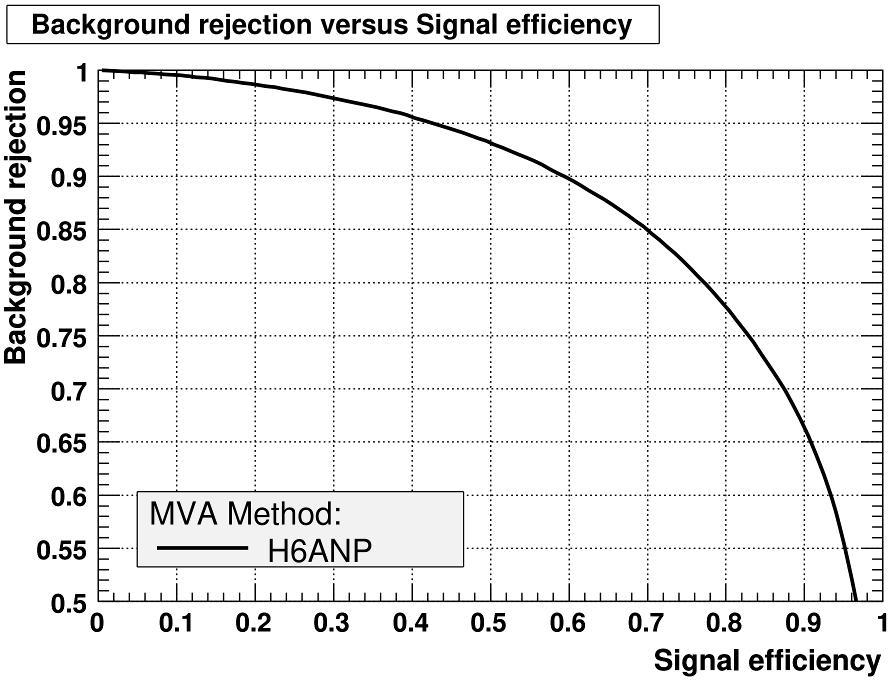

Figure 5: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, All leptons, NN + LRHWW variables, Pythia WW.

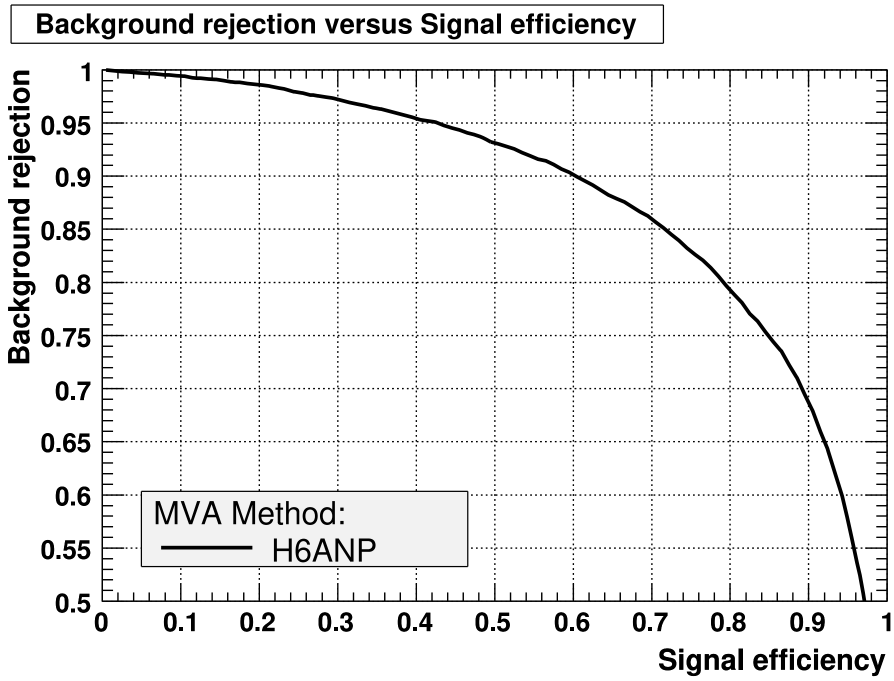

Figure 6: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, ee, NN + LRHWW variables, Pythia WW.

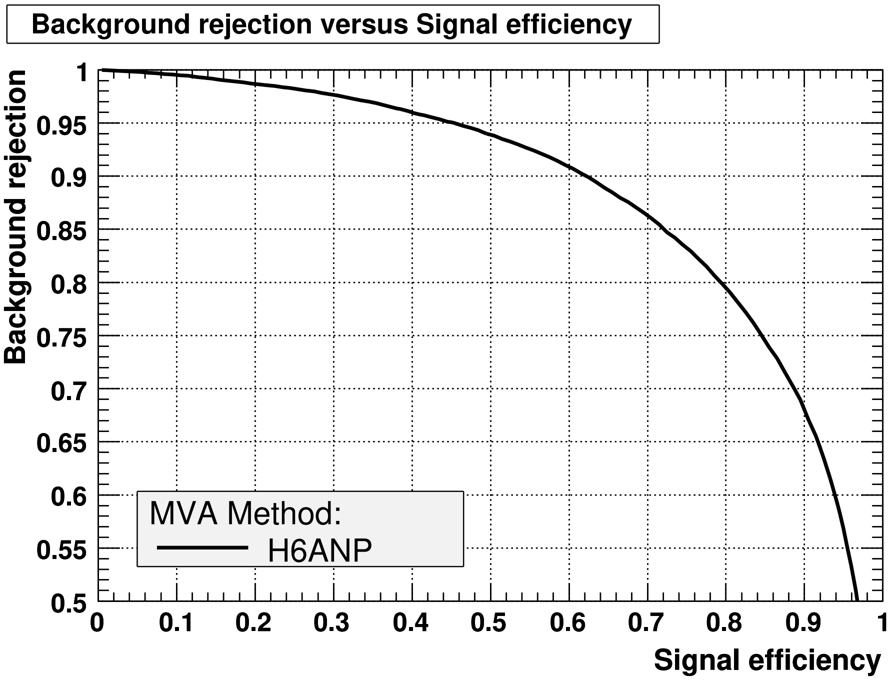

Figure 7: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, eμ, NN + LRHWW variables, Pythia WW.

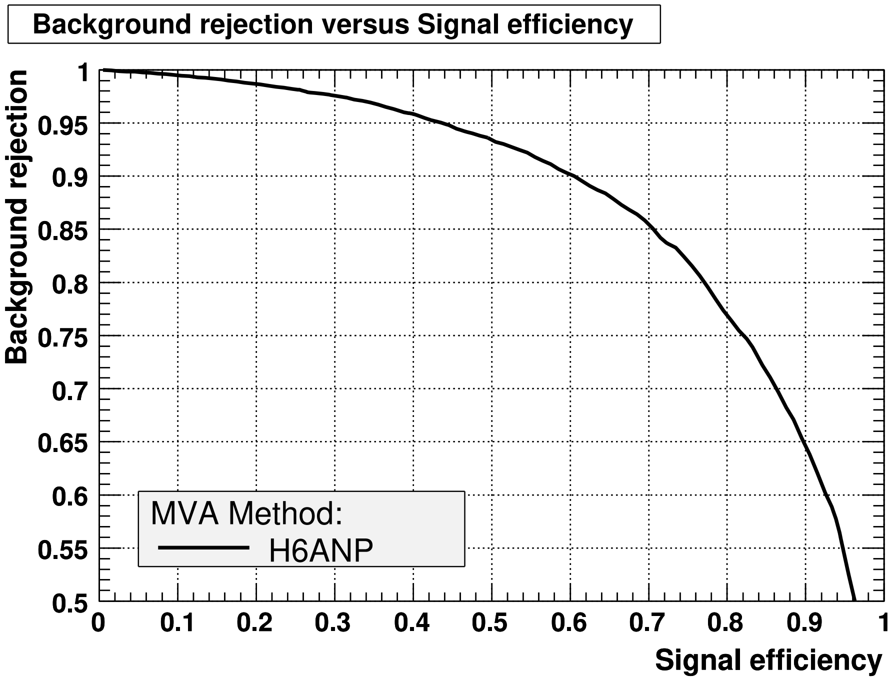

Figure 8: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, $\mu\mu$, NN + LRHWW variables, Pythia WW.

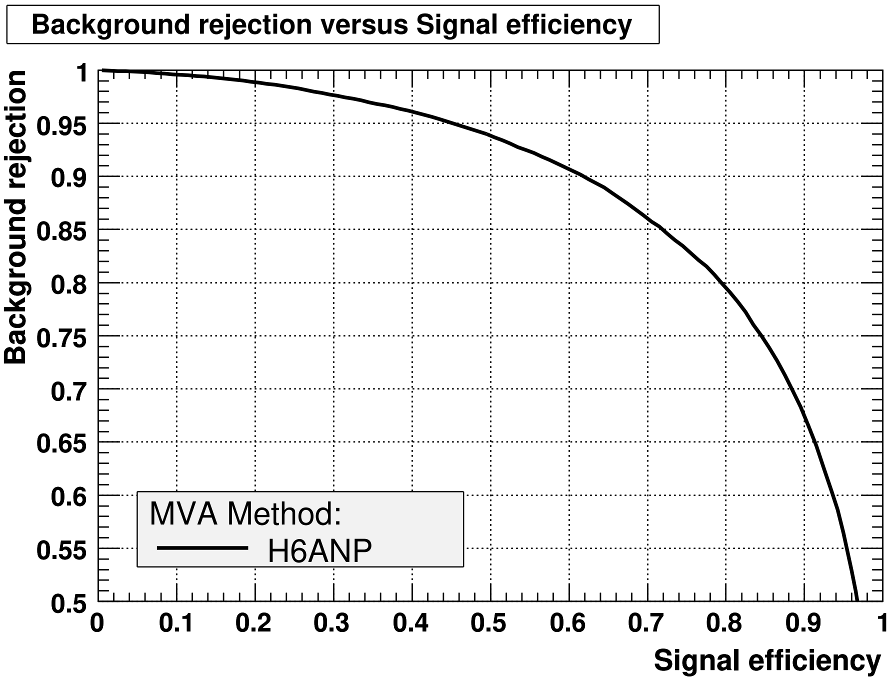

Figure 9: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, All leptons, NN variables, MC at NLOWW.

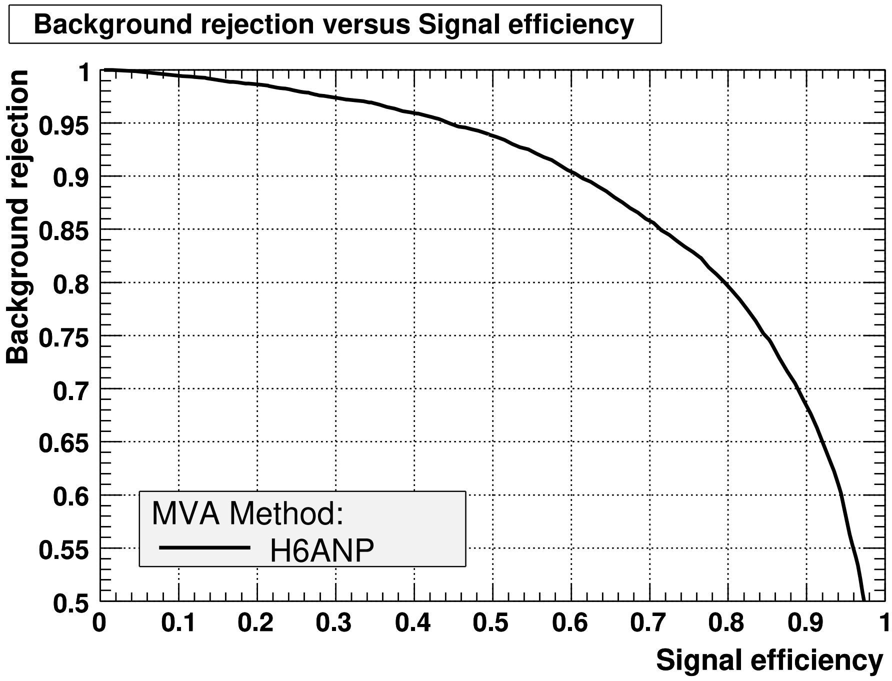

Figure 10: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, ee, NN variables, MC at NLOWW.

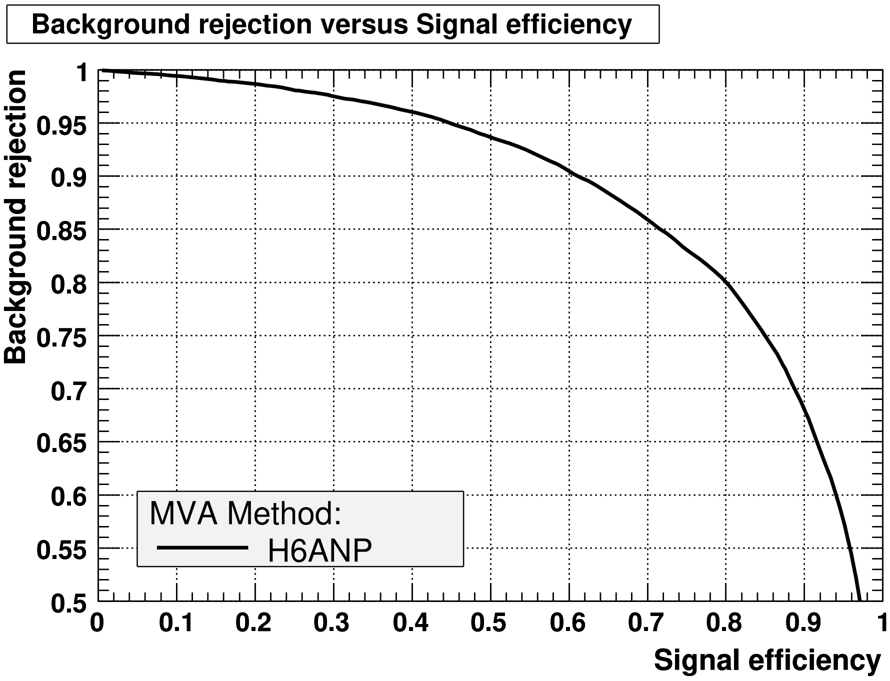

Figure 11: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, eμ, NN variables, MC at NLOWW.

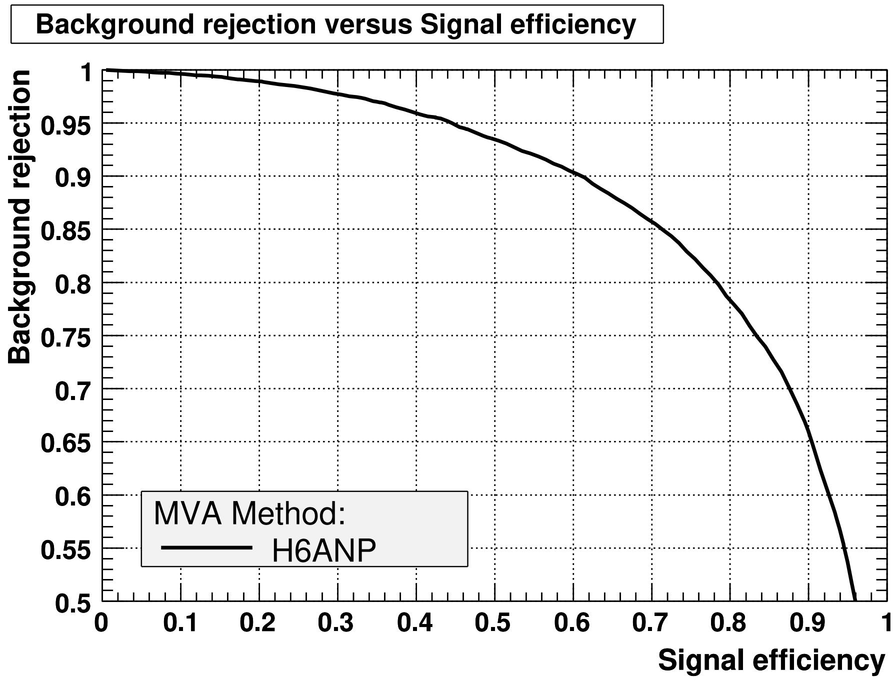

Figure 12: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, $\mu\mu$, NN variables, MC at NLOWW.

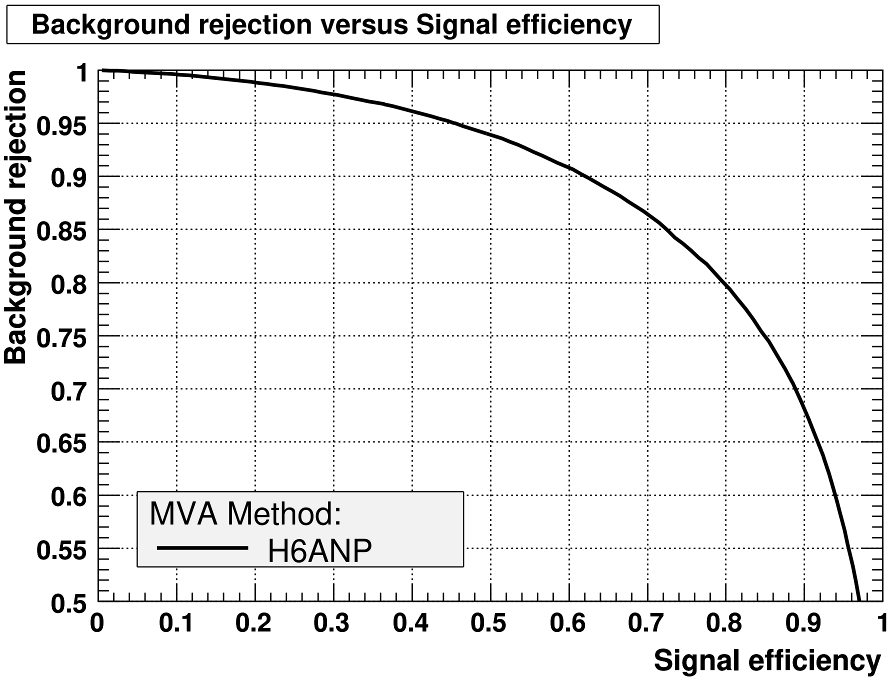

Figure 13: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, All leptons, NN + LRHWW variables, MC at NLOWW.

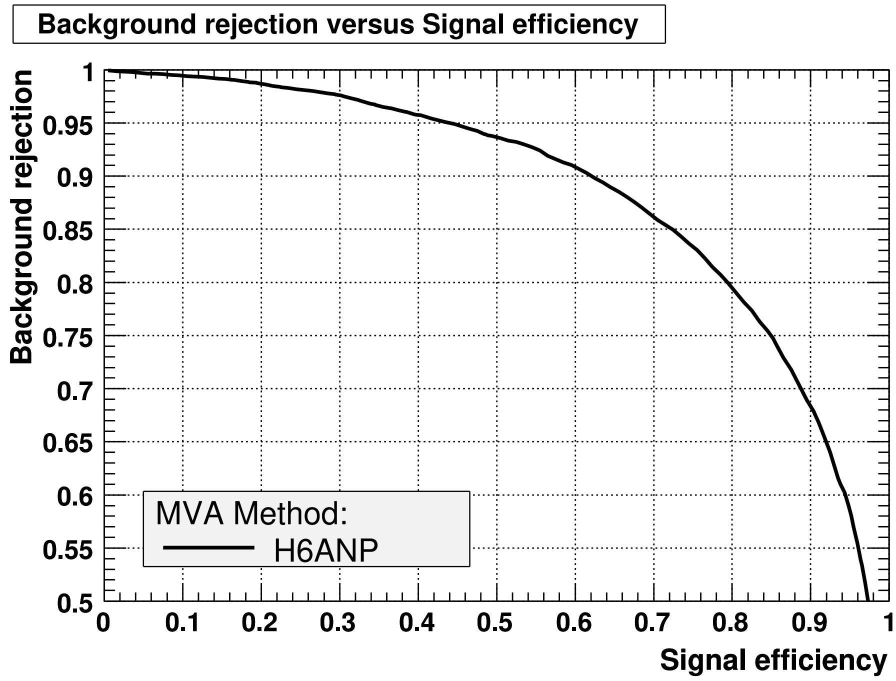

Figure 14: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, ee, NN + LRHWW variables, MC at NLOWW.

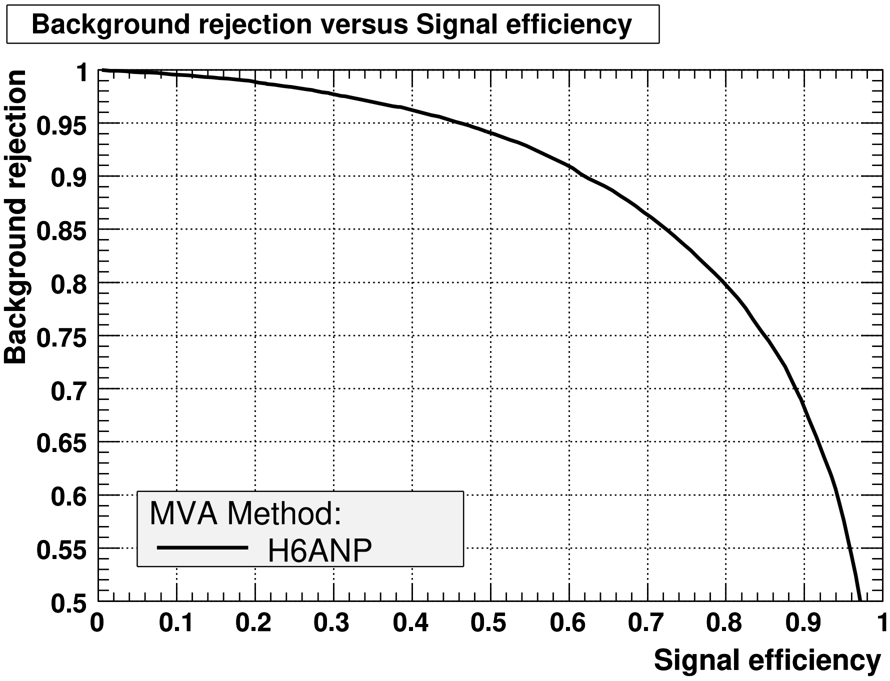

Figure 15: $M_h = 160 GeV/c^2$, Background Rejection vs Signal Efficiency, eμ, NN + LRHWW variables, MC at NLOWW.

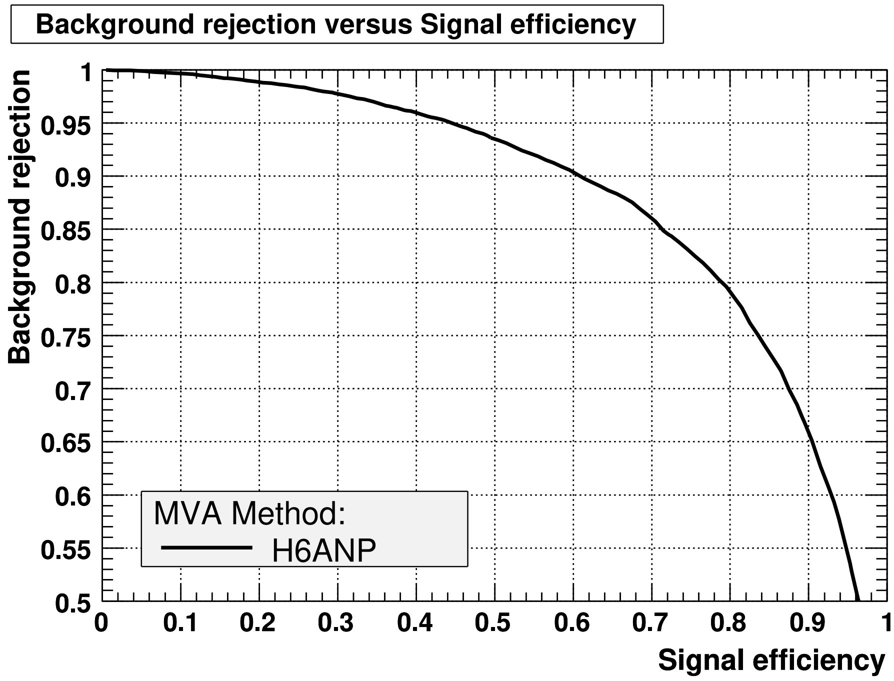

Figure 16: $M_h = 160GeV/c^2$, Background Rejection vs Signal Efficiency, $\mu\mu$, NN + LRHWW variables, MC at NLOWW.

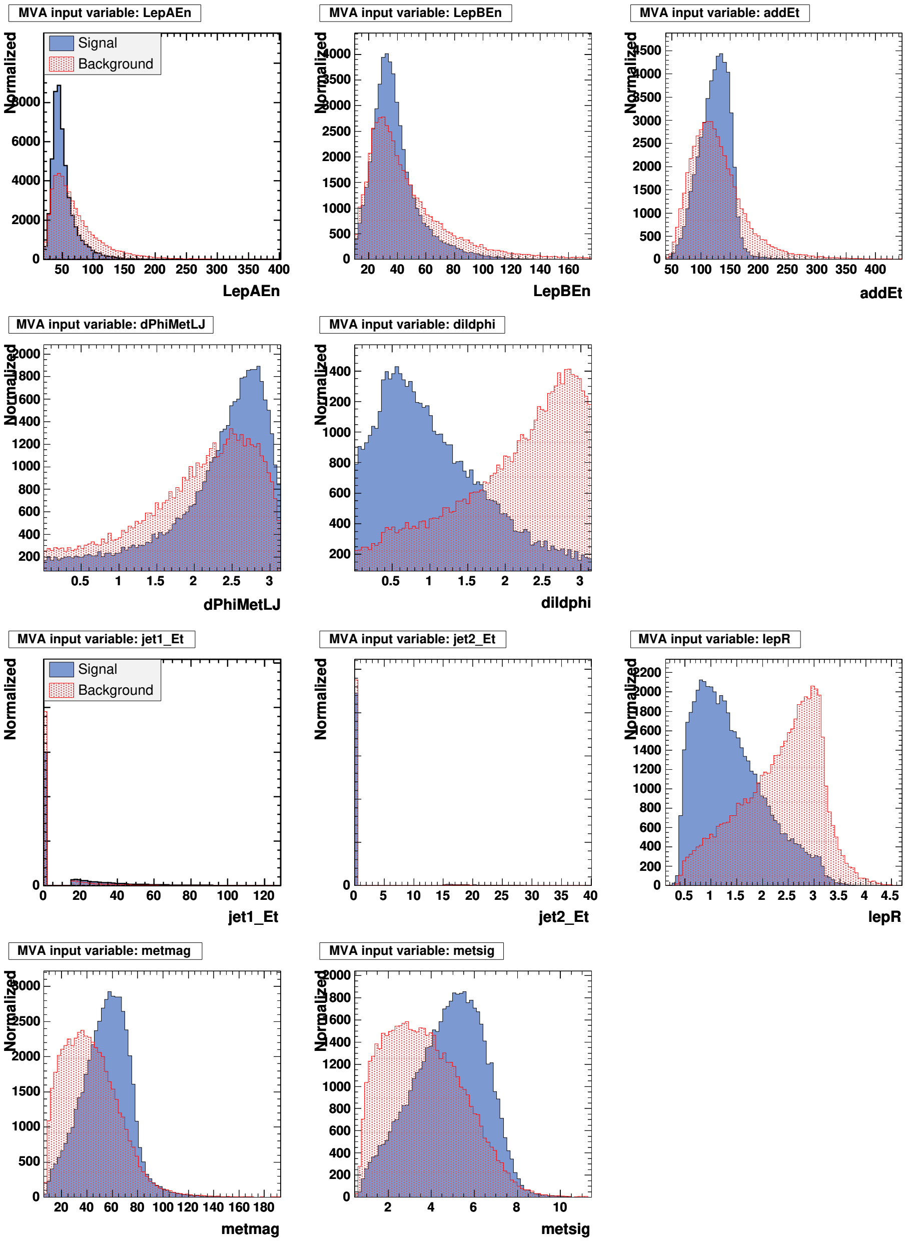

Figure 17: $M_h = 160 GeV/c^2$, Discriminant variables, All leptons, NN variables, Pythia WW.

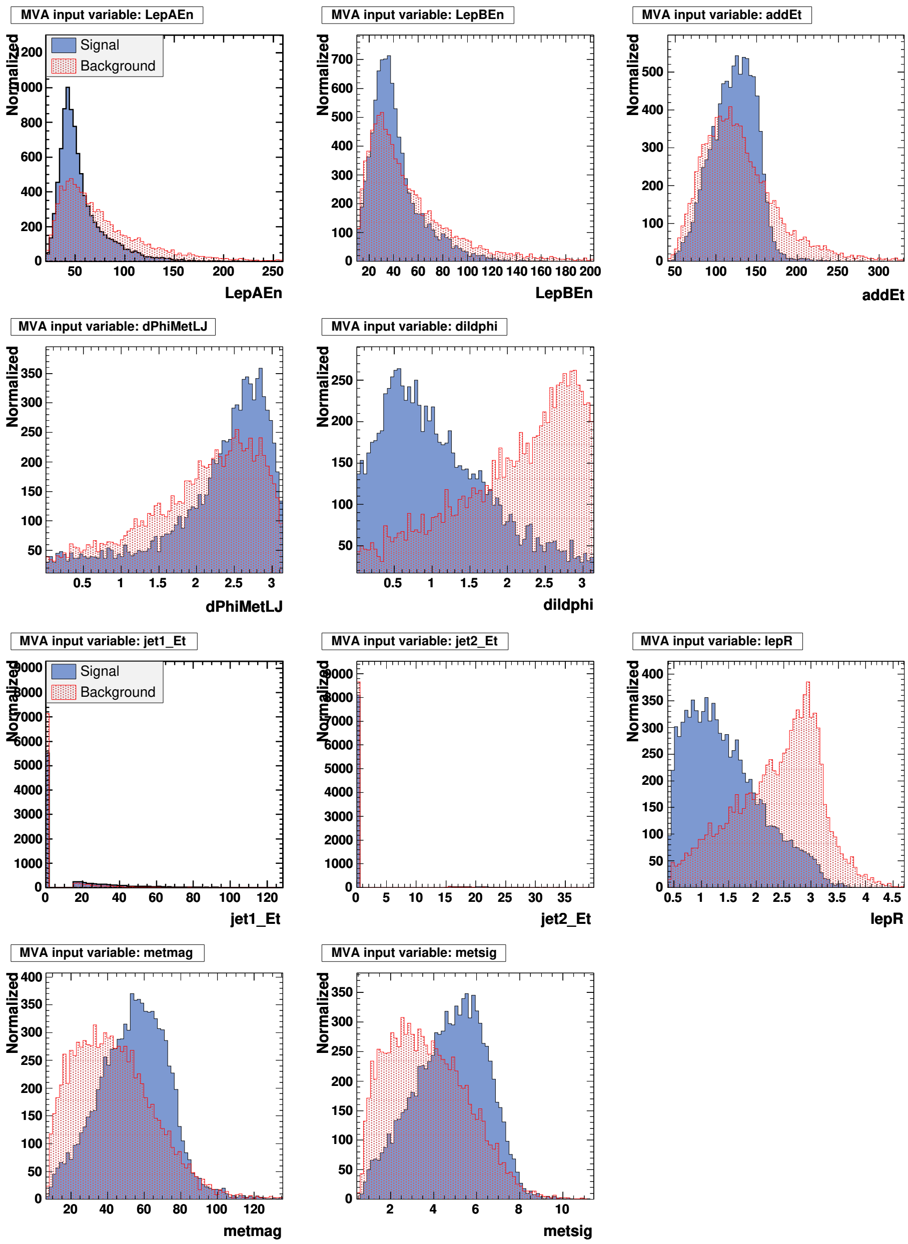

Figure 18: $M_h = 160GeV/c^2$, Discriminant variables, ee, NN variables, Pythia WW.

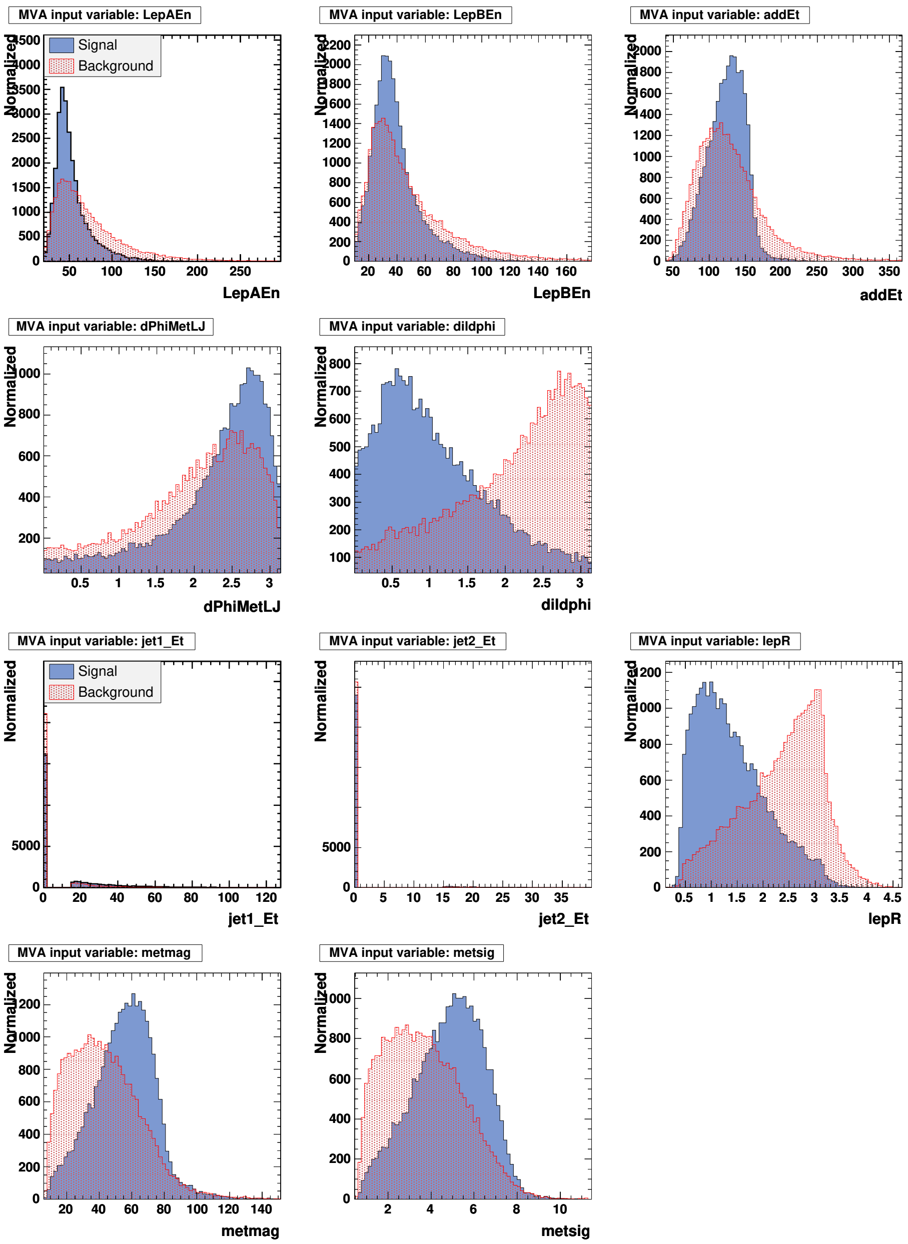

Figure 19: $M_h = 160GeV/c^2$, Discriminant variables, eμ, NN variables, Pythia WW.

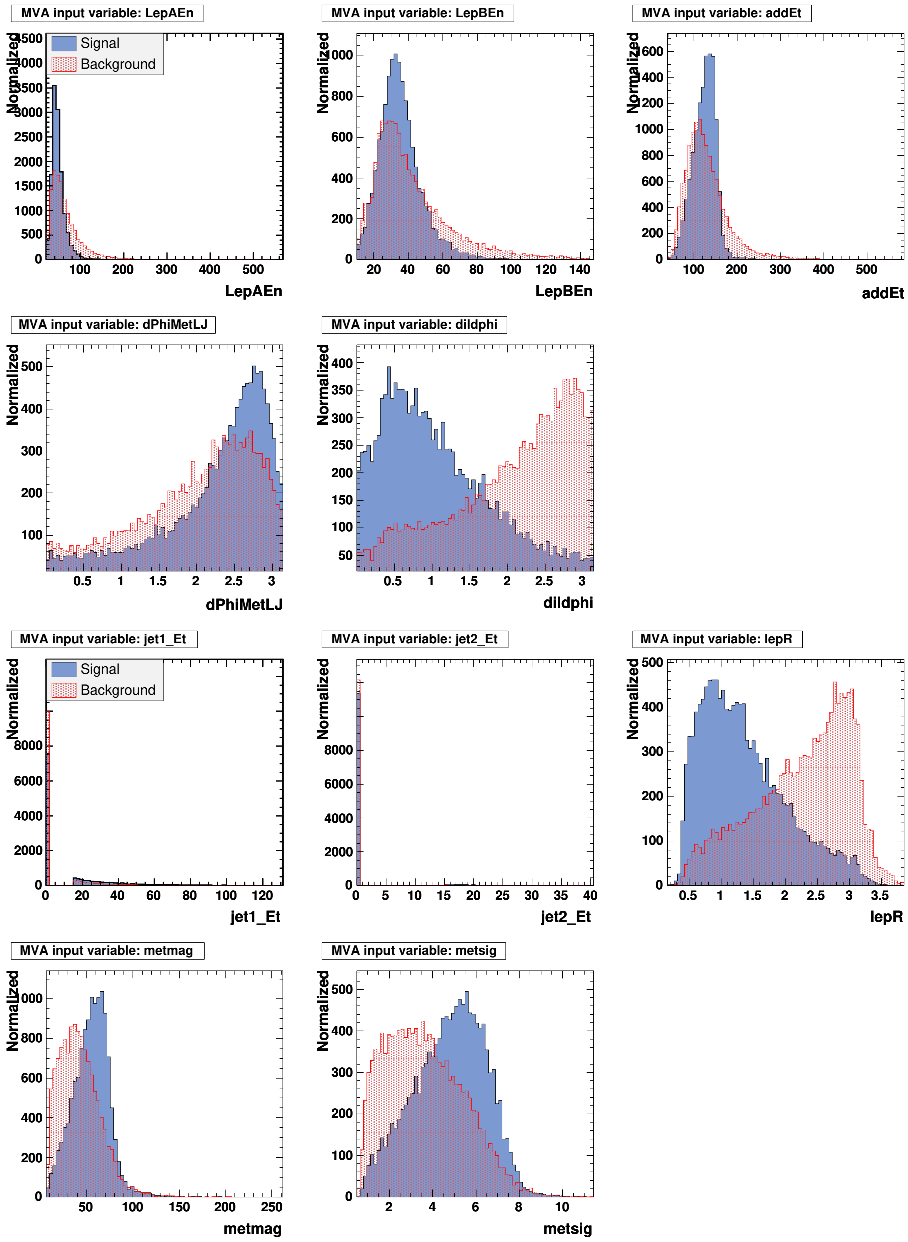

Figure 20: $M_h = 160GeV/c^2$, Discriminant variables, $\mu\mu$, NN variables, Pythia WW.

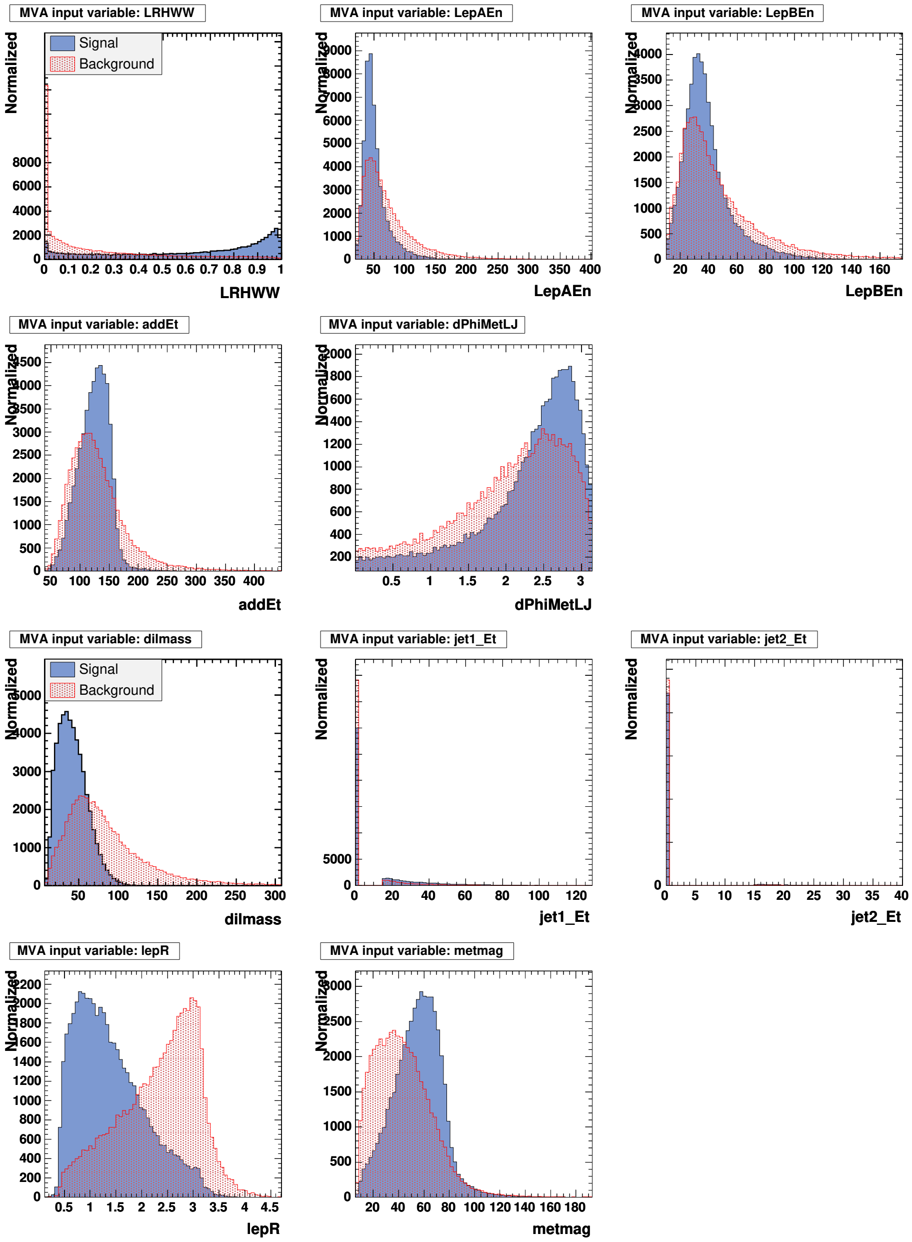

Figure 21: $M_h = 160 GeV/c^2$, Discriminant variables, All leptons, NN + LRHWW variables, Pythia WW.

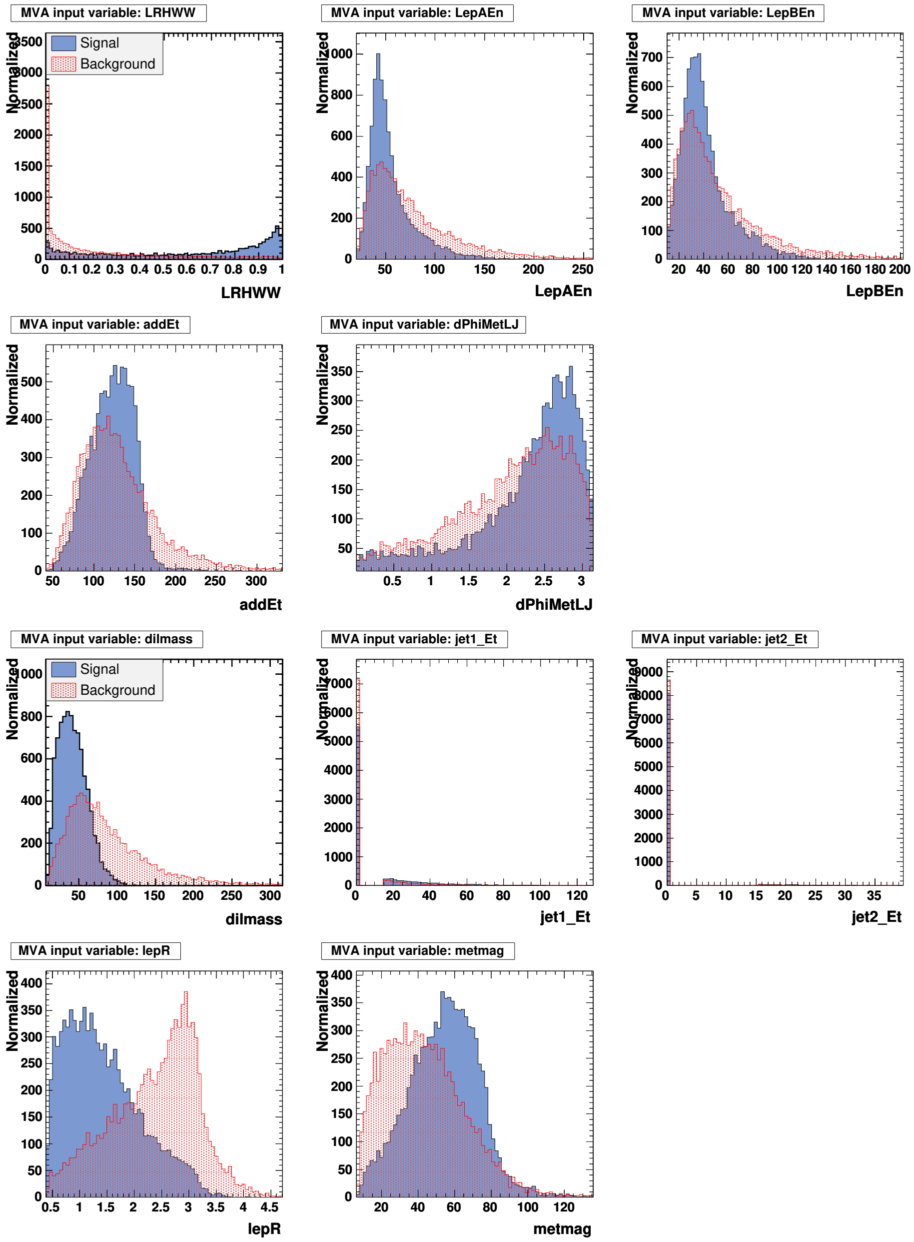

Figure 22: $M_h = 160 GeV/c^2$, Discriminant variables, ee, NN + LRHWW variables, Pythia WW.

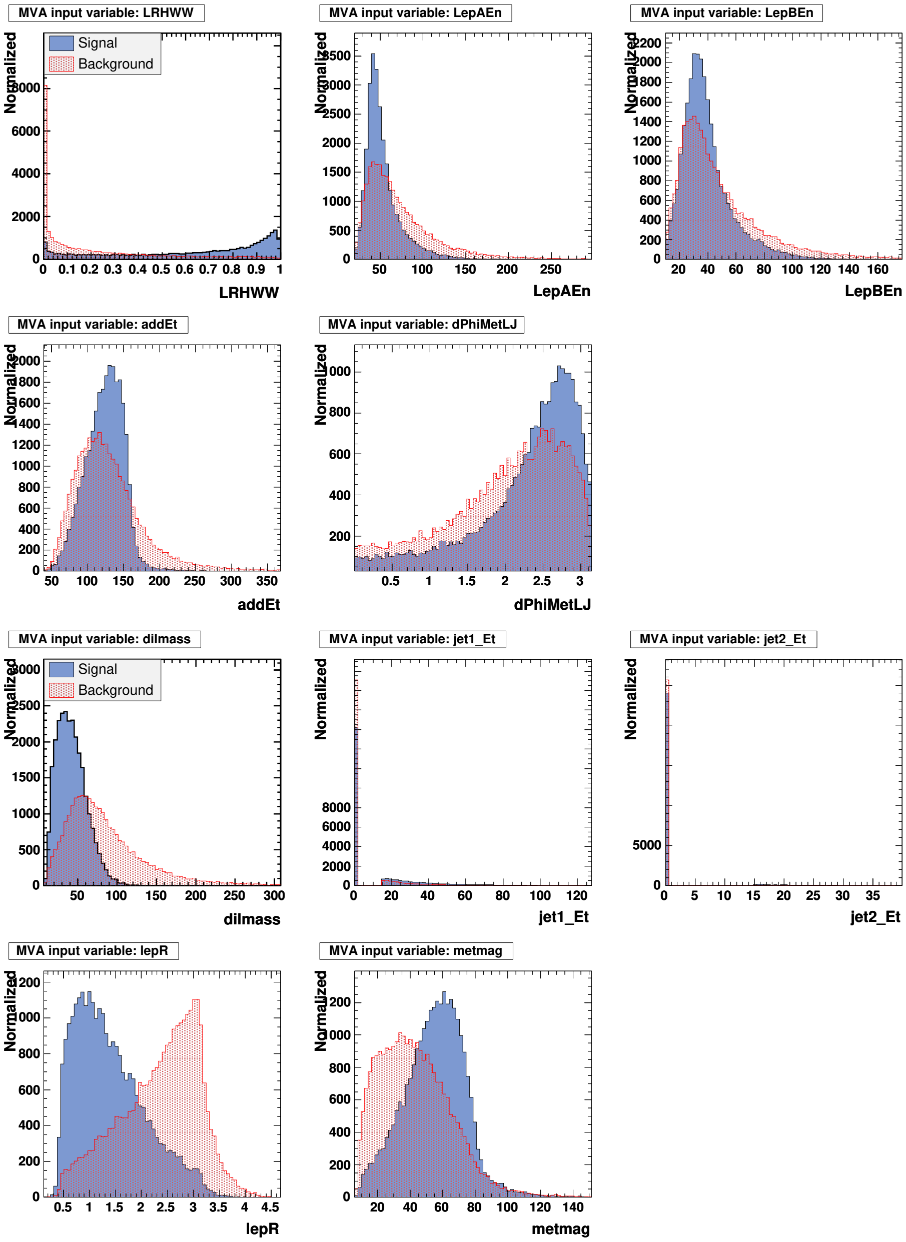

Figure 23: $M_h = 160 GeV/c^2$, Discriminant variables, eμ, NN + LRHWW variables, Pythia WW.

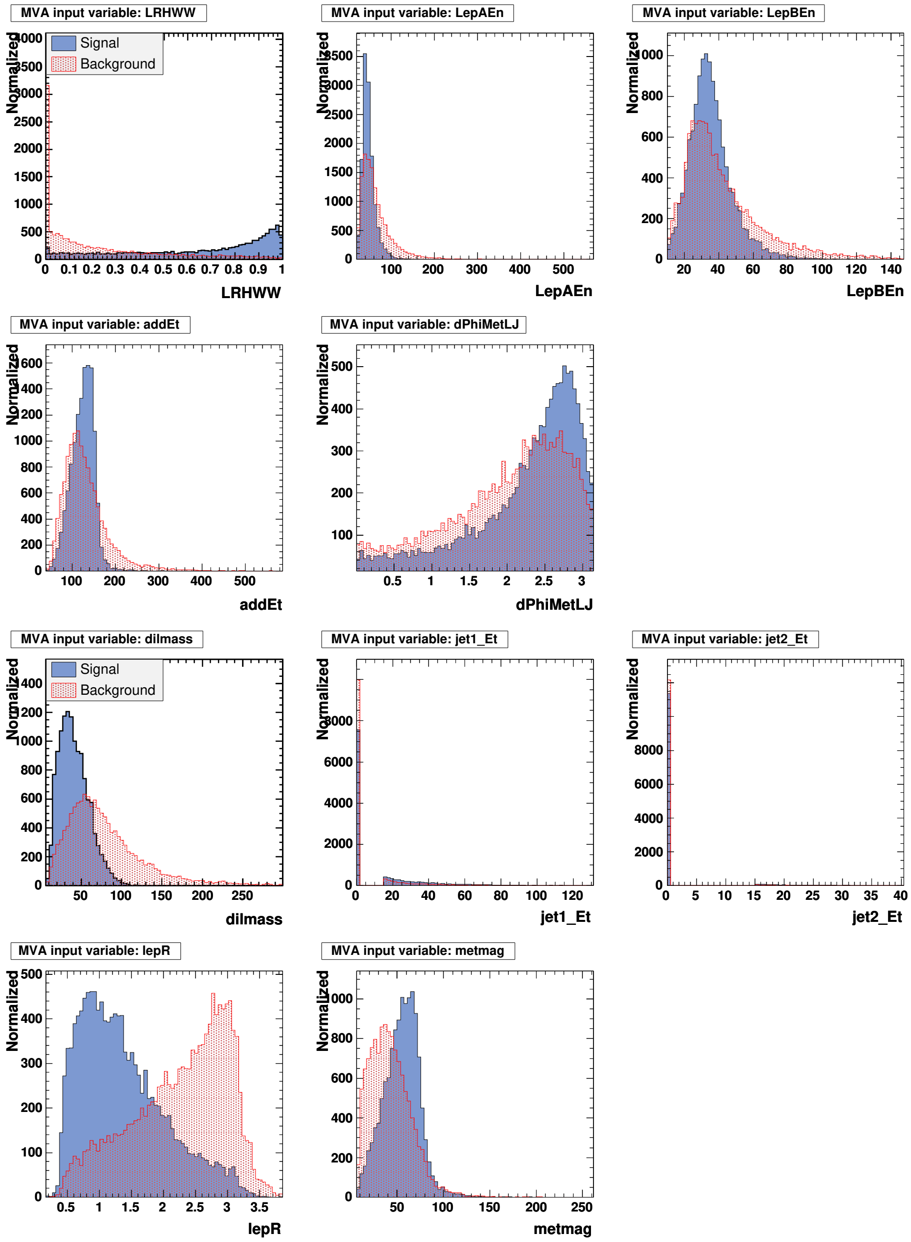

Figure 24: $M_h = 160GeV/c^2$, Discriminant variables, $\mu\mu$, NN + LRHWW variables, Pythia WW.

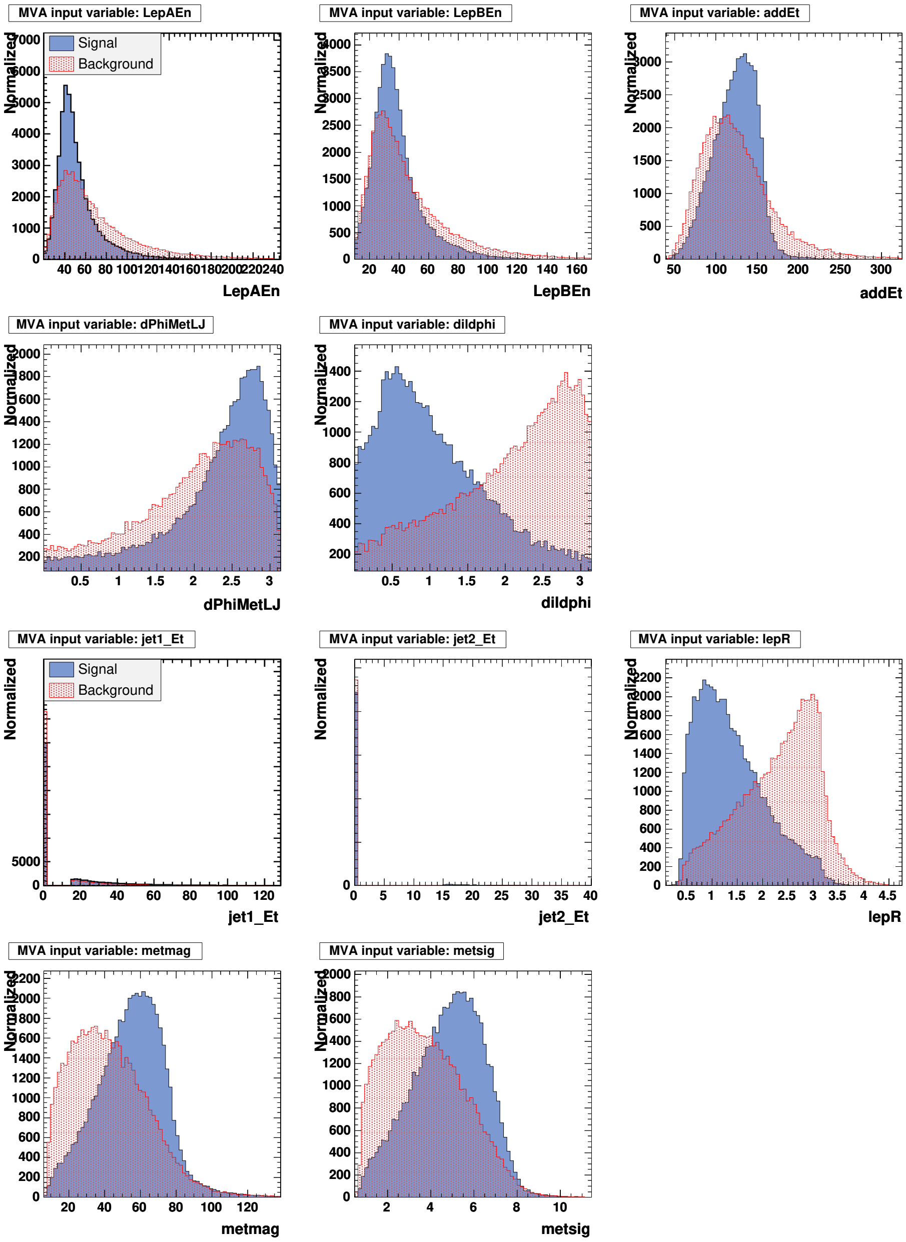

Figure 25: $M_h = 160 GeV/c^2$, Discriminant variables, All leptons, NN variables, MC at NLOWW.

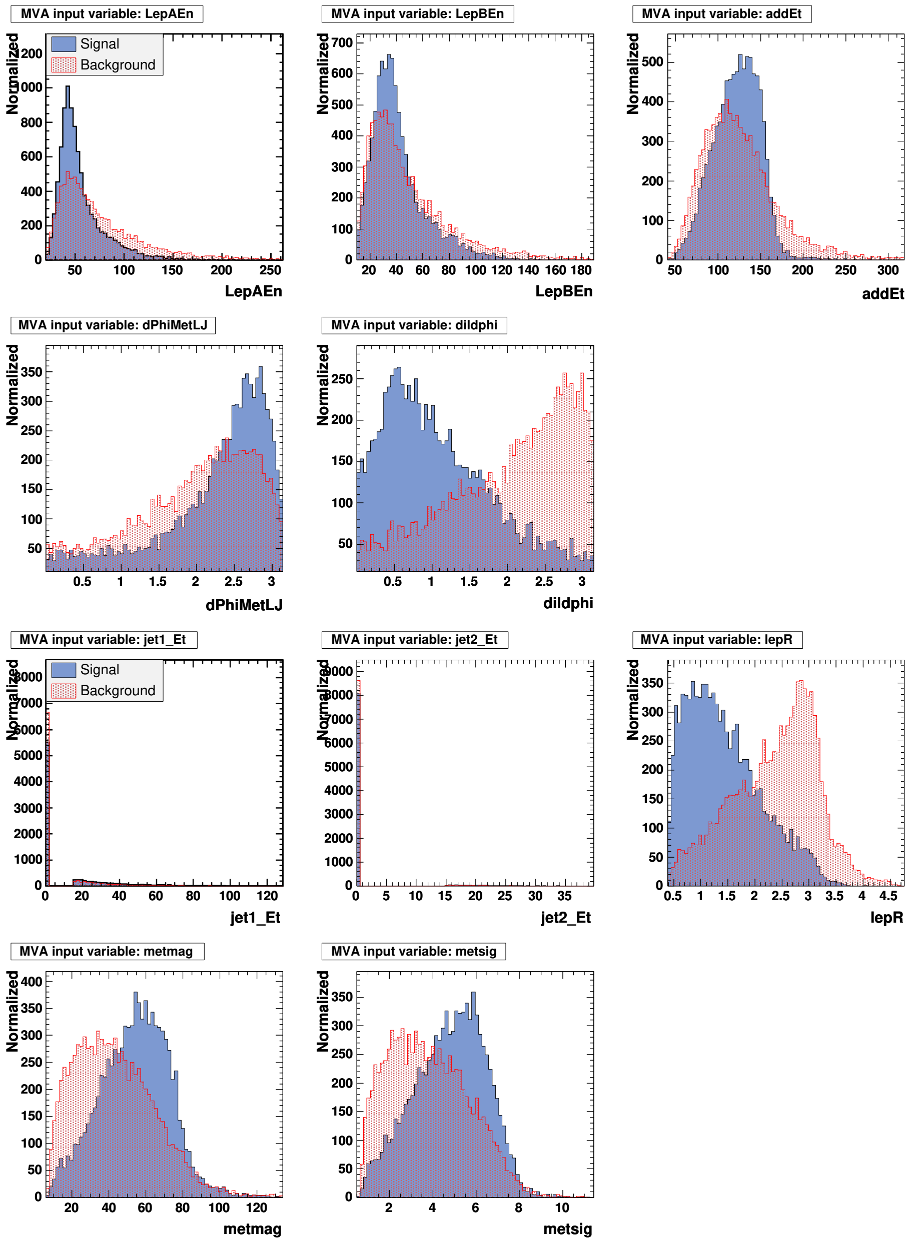

Figure 26: $M_h = 160 GeV/c^2$, Discriminant variables, ee, NN variables, MC at NLOWW.

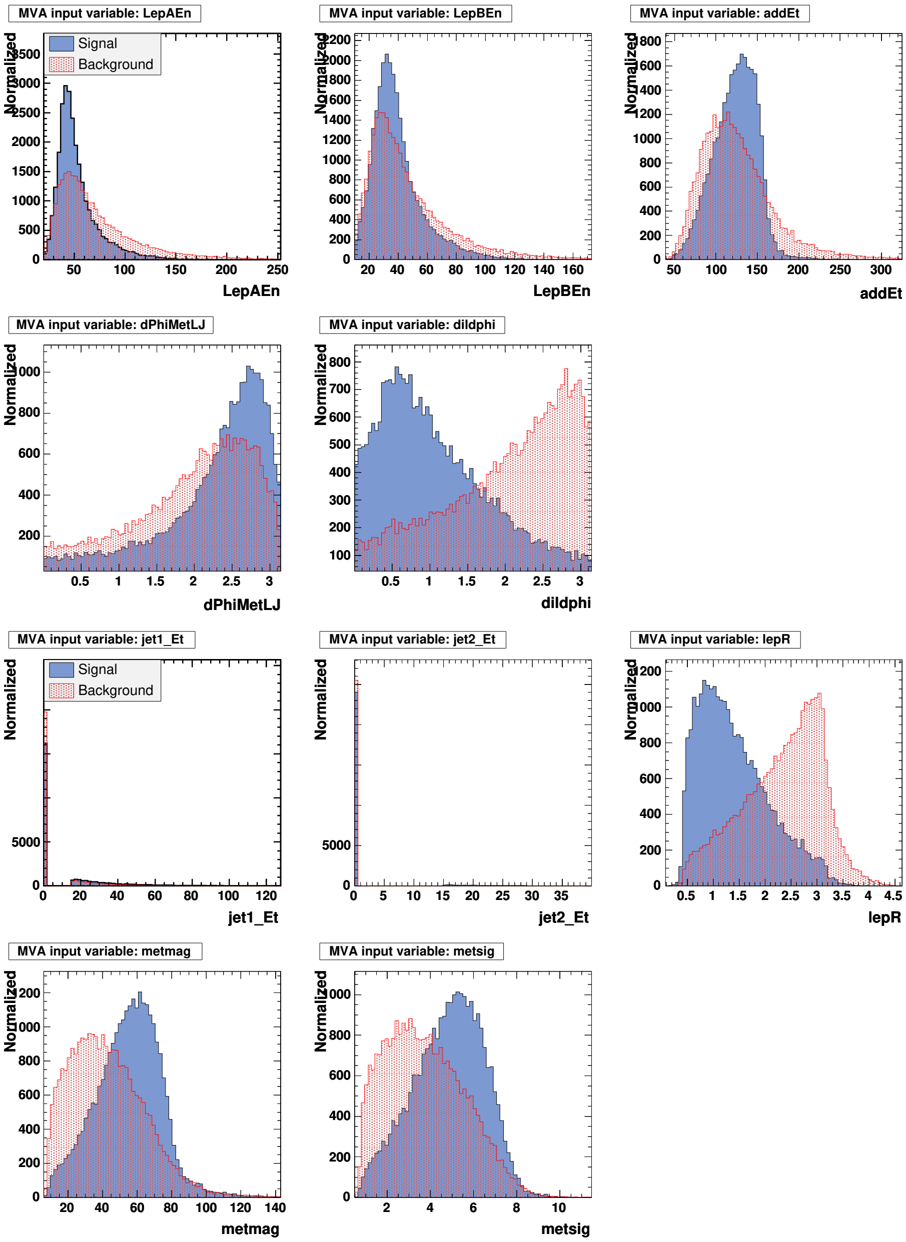

Figure 27: $M_h = 160GeV/c^2$, Discriminant variables, eμ, NN variables, MC at NLOWW.

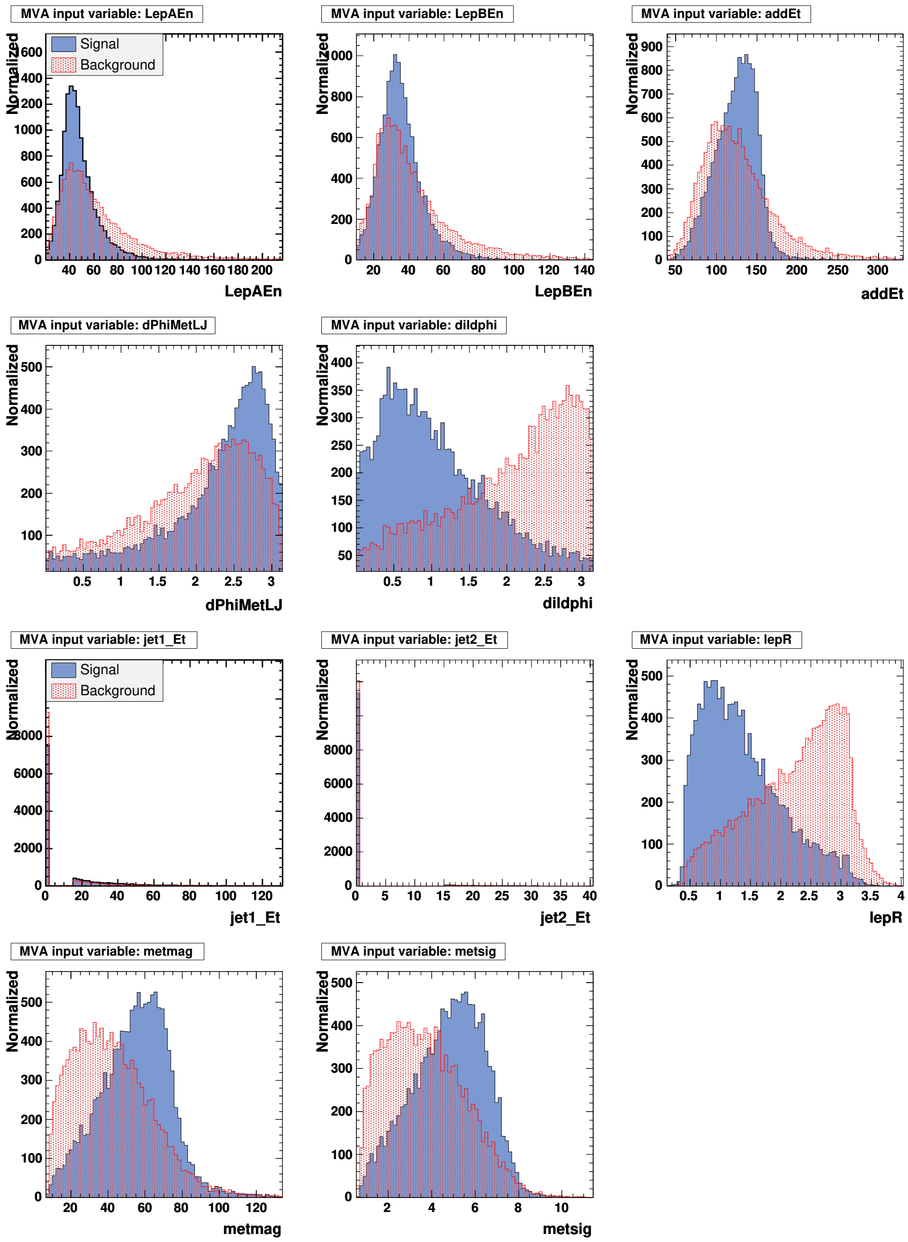

Figure 28: $M_h = 160GeV/c^2$, Discriminant variables, $\mu\mu$, NN variables, MC at NLOWW.

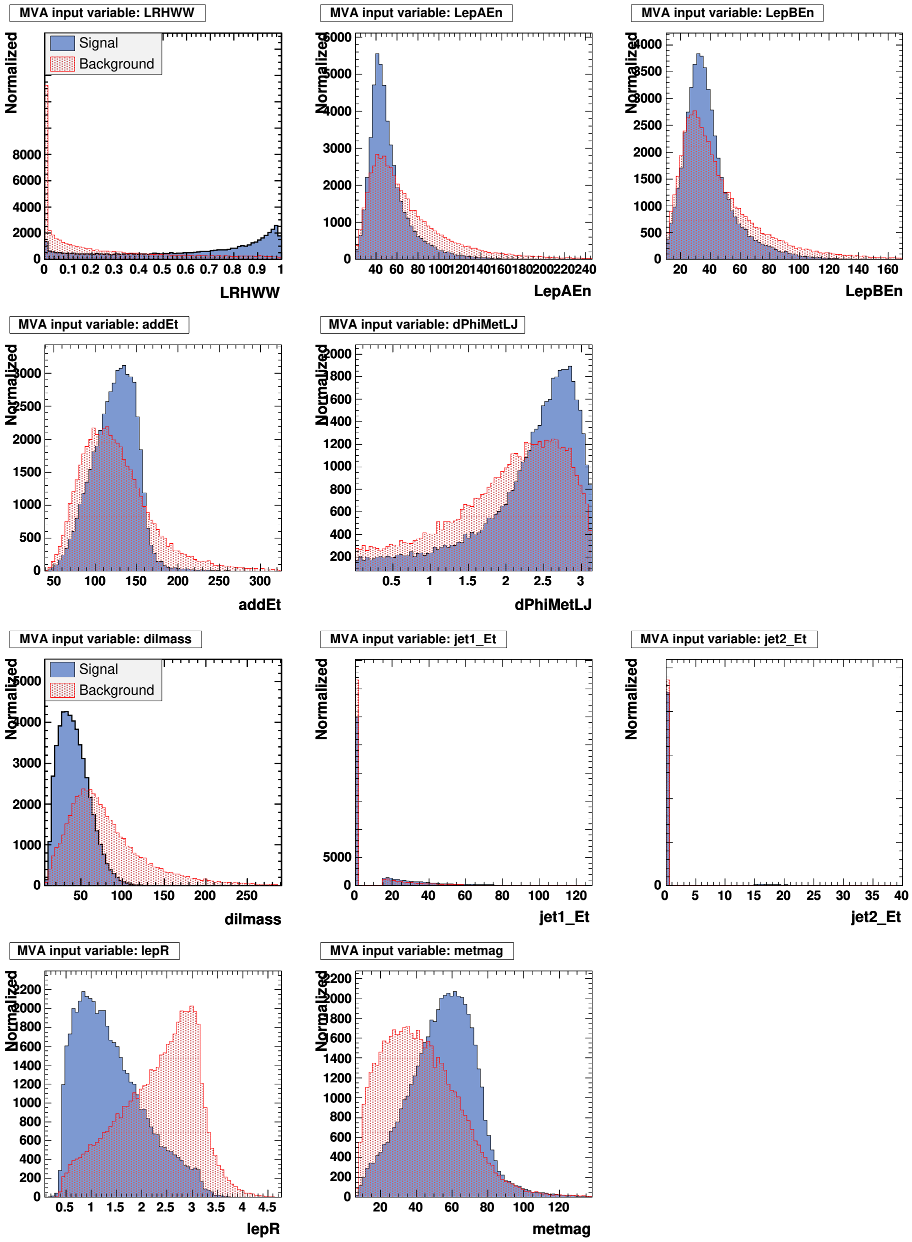

Figure 29: $M_h = 160GeV/c^2$, Discriminant variables, All leptons, NN + LRHWW variables, MC at NLOWW.

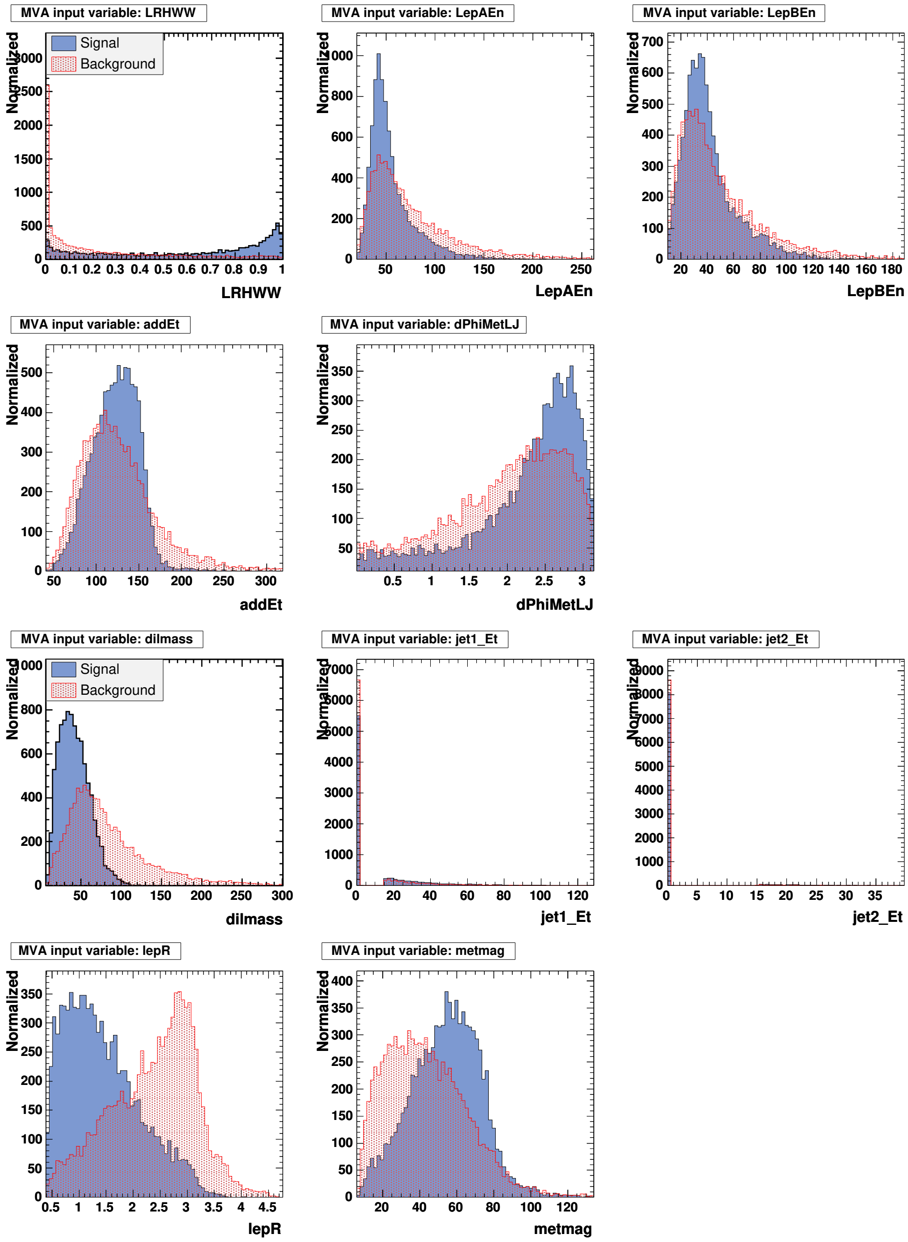

Figure 30: $M_h = 160 GeV/c^2$, Discriminant variables, ee, NN + LRHWW variables, MC at NLOWW.

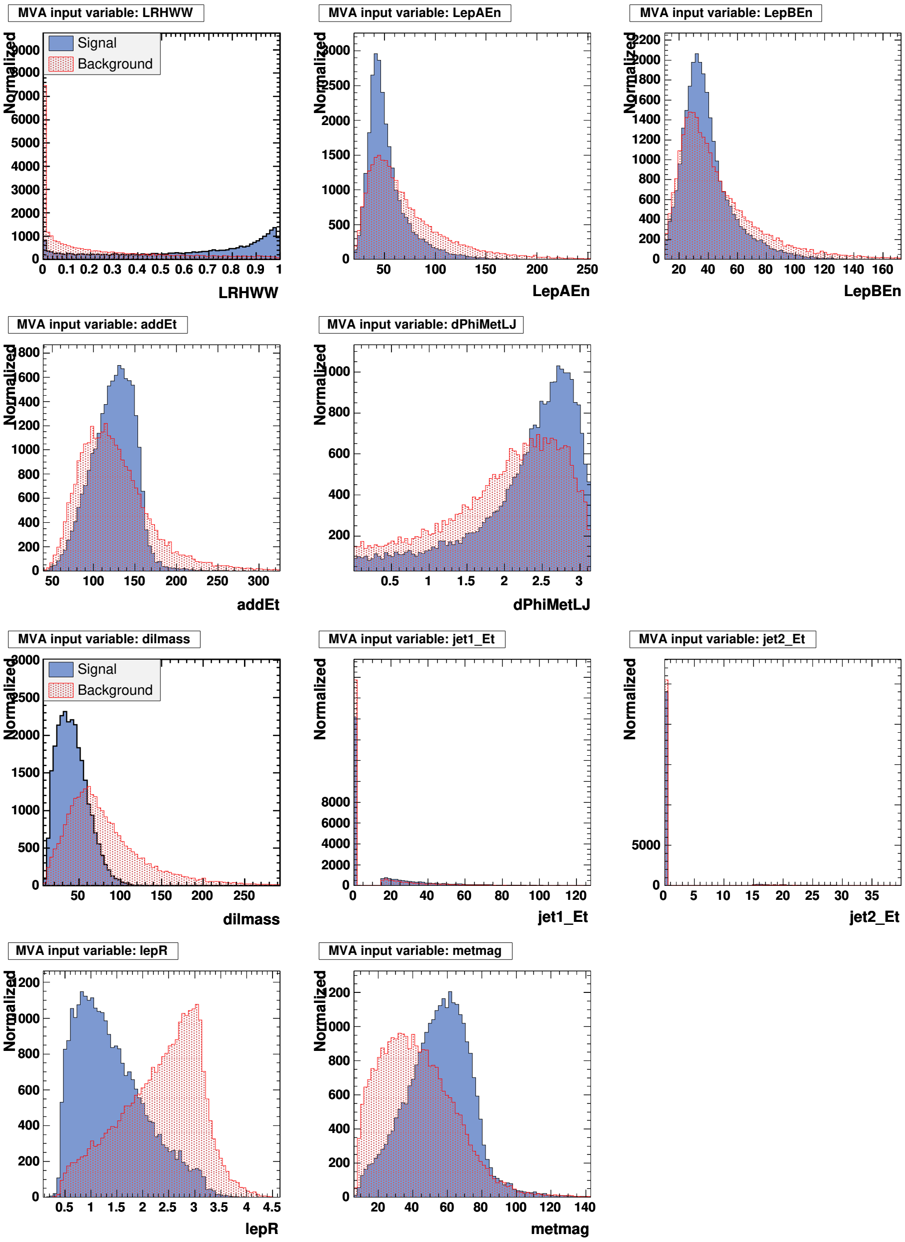

Figure 31: $M_h = 160GeV/c^2$, Discriminant variables, eμ, NN + LRHWW variables, MC at NLOWW.

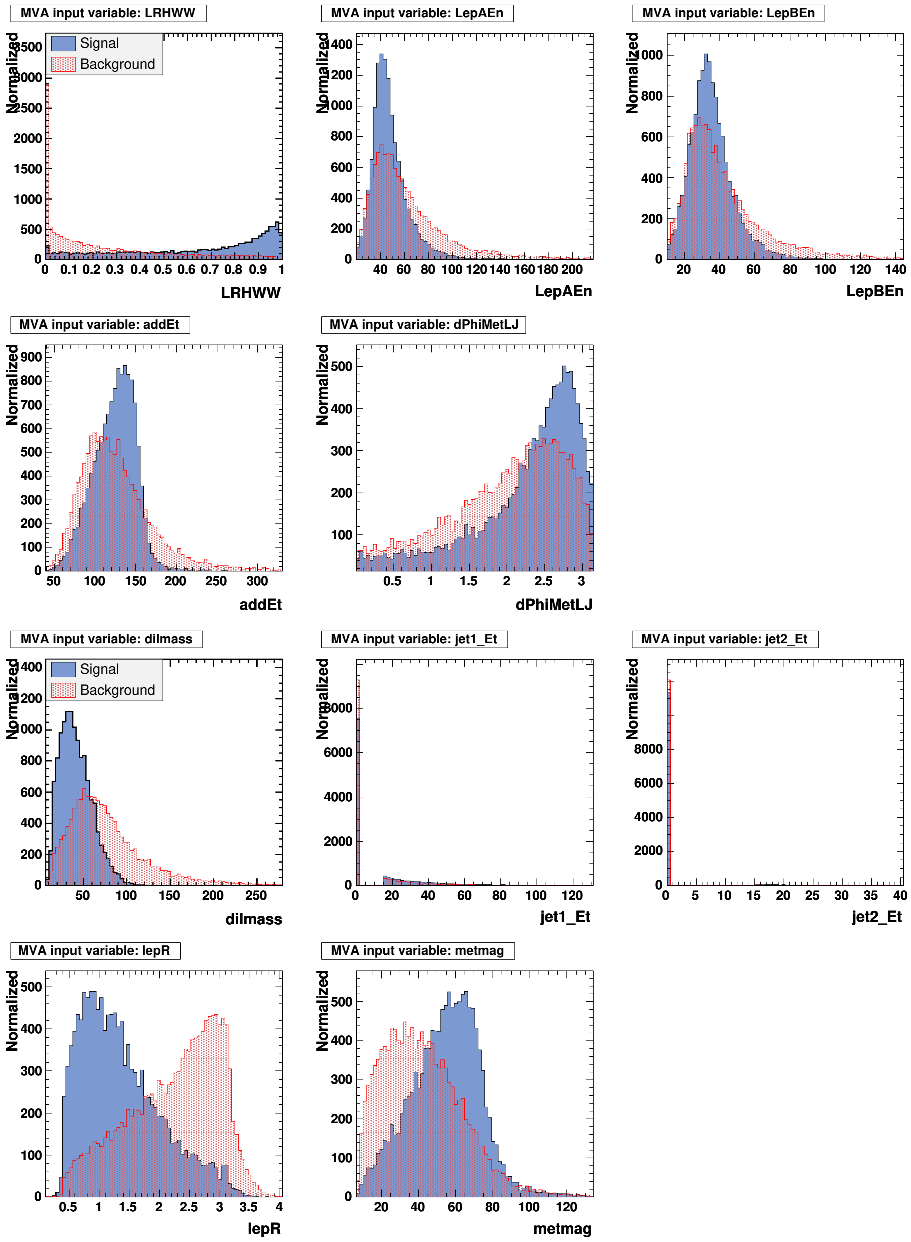

Figure 32: $M_h = 160GeV/c^2$, Discriminant variables, $\mu\mu$, NN + LRHWW variables, MC at NLOWW.

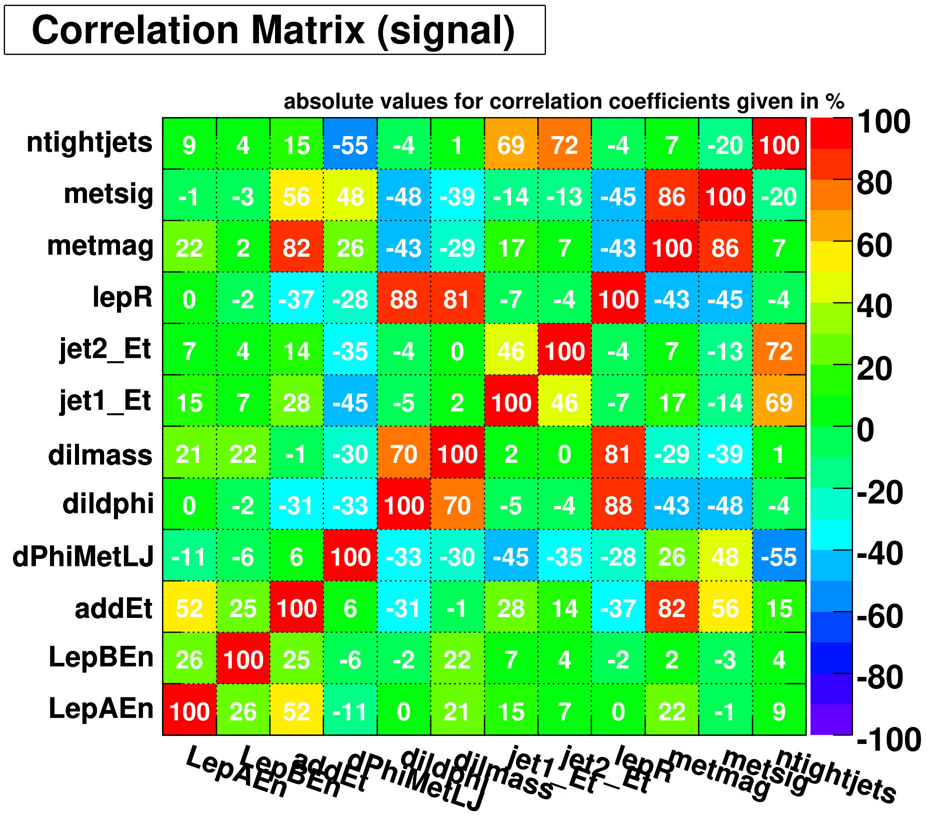

Figure 33: $M_h = 160 GeV/c^2$, Signal correlation matrix, All leptons, NN variables, Pythia WW.

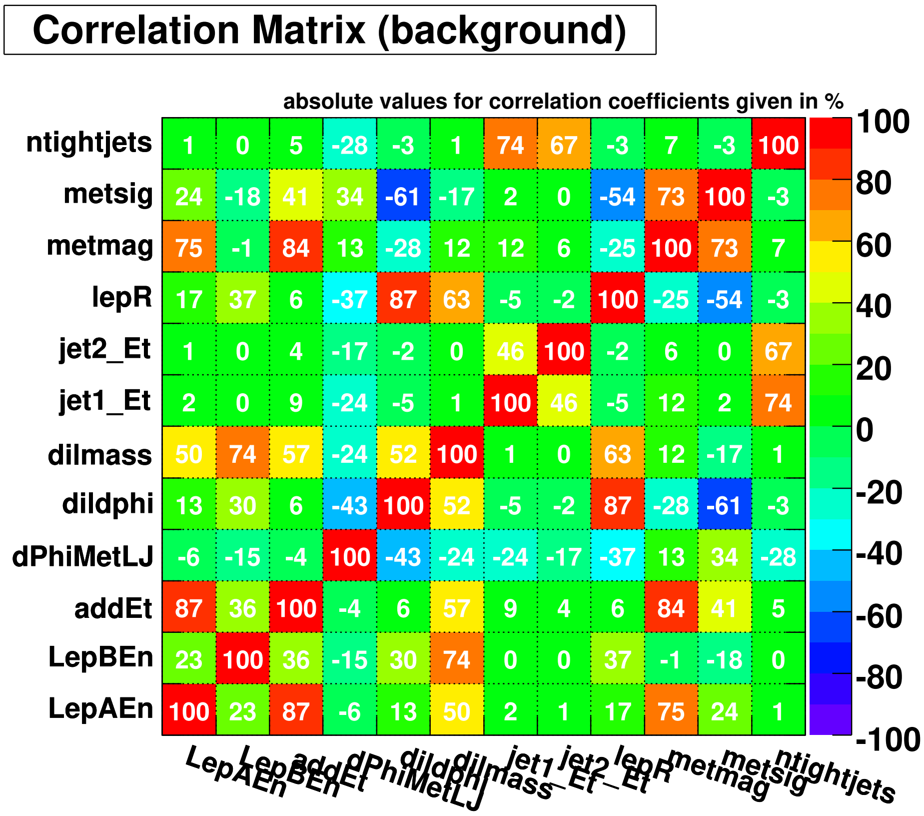

Figure 34: $M_h = 160GeV/c^2$, Background correlation matrix, All leptons, NN variables, Pythia WW.

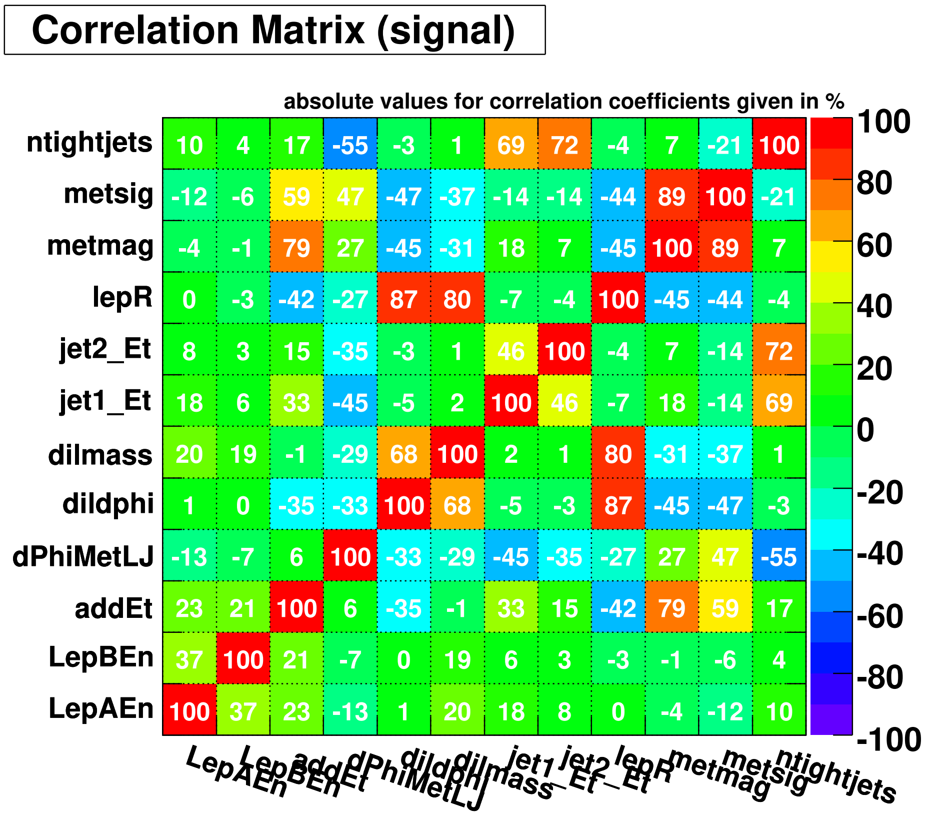

Figure 35: $M_h = 160 GeV/c^2$, Signal correlation matrix, ee, NN variables, Pythia WW.

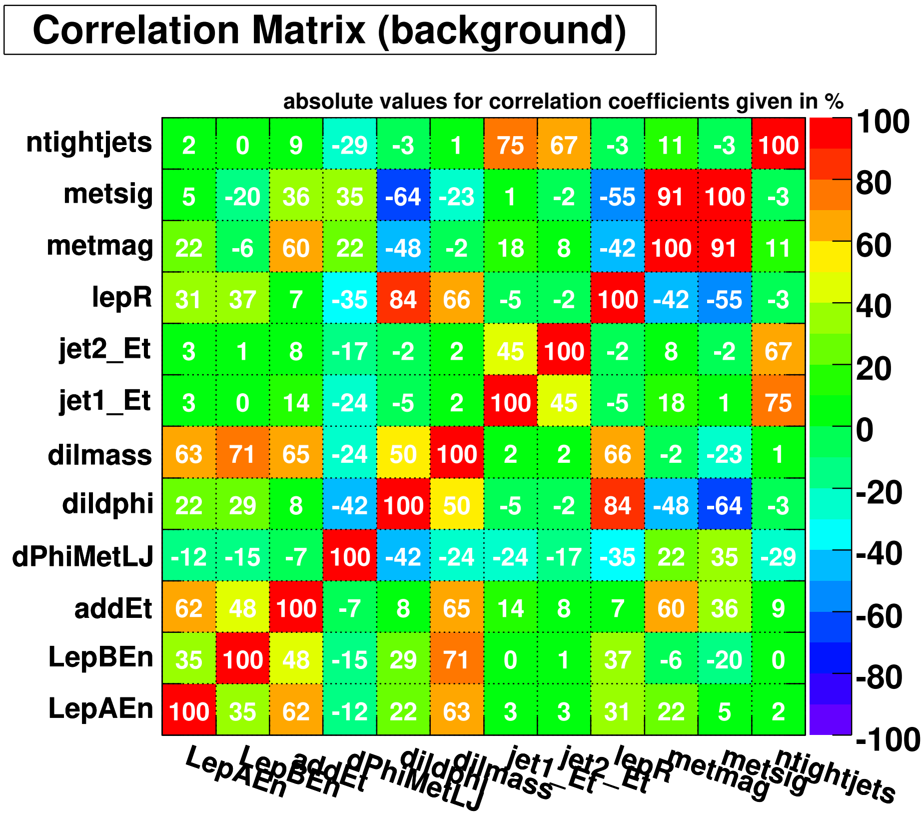

Figure 36: $M_h = 160 GeV/c^2$, Background correlation matrix, ee, NN variables, Pythia WW.

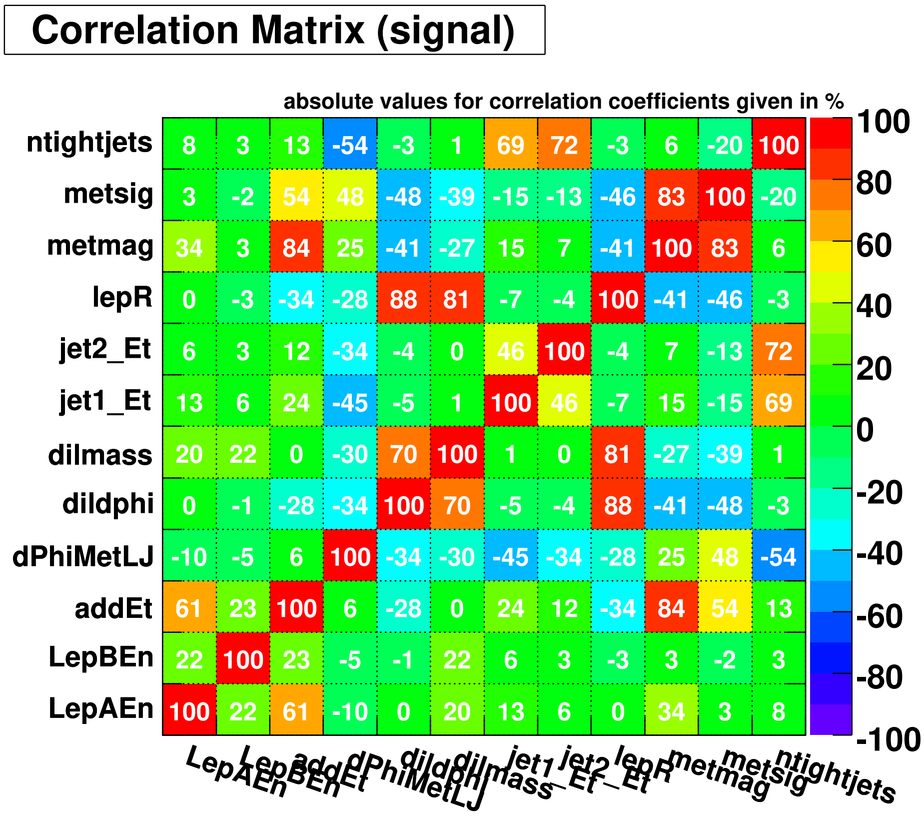

Figure 37: $M_h = 160GeV/c^2$, Signal correlation matrix, eμ, NN variables, Pythia WW.

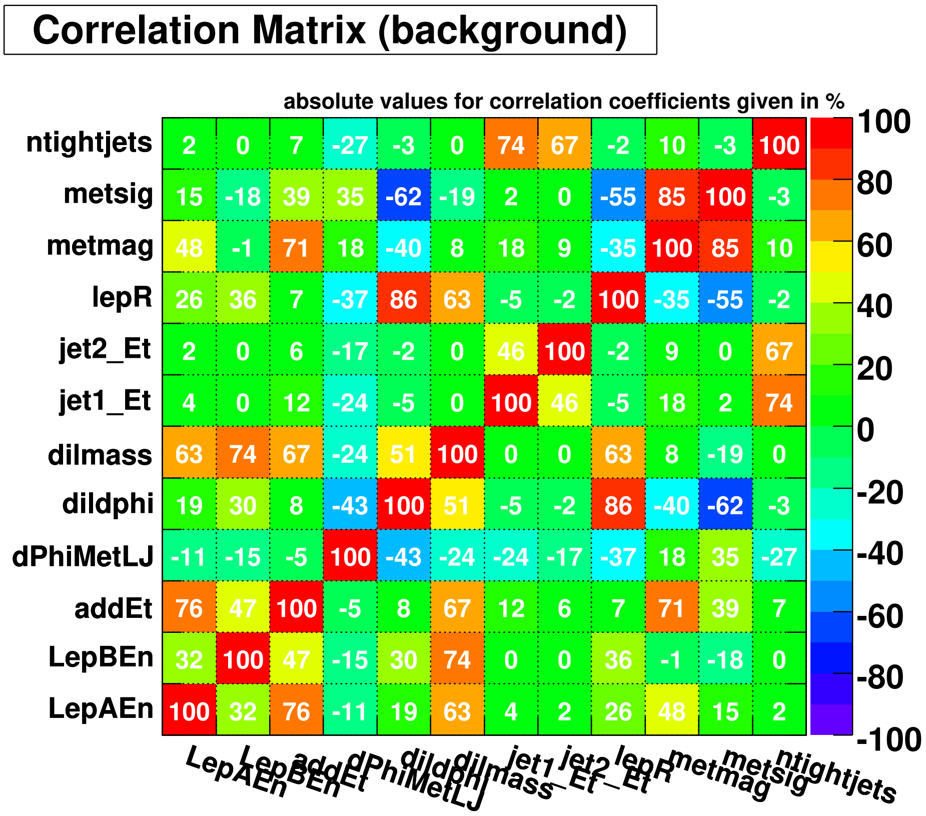

Figure 38: $M_h = 160 GeV/c^2$, Background correlation matrix, eμ, NN variables, Pythia WW.

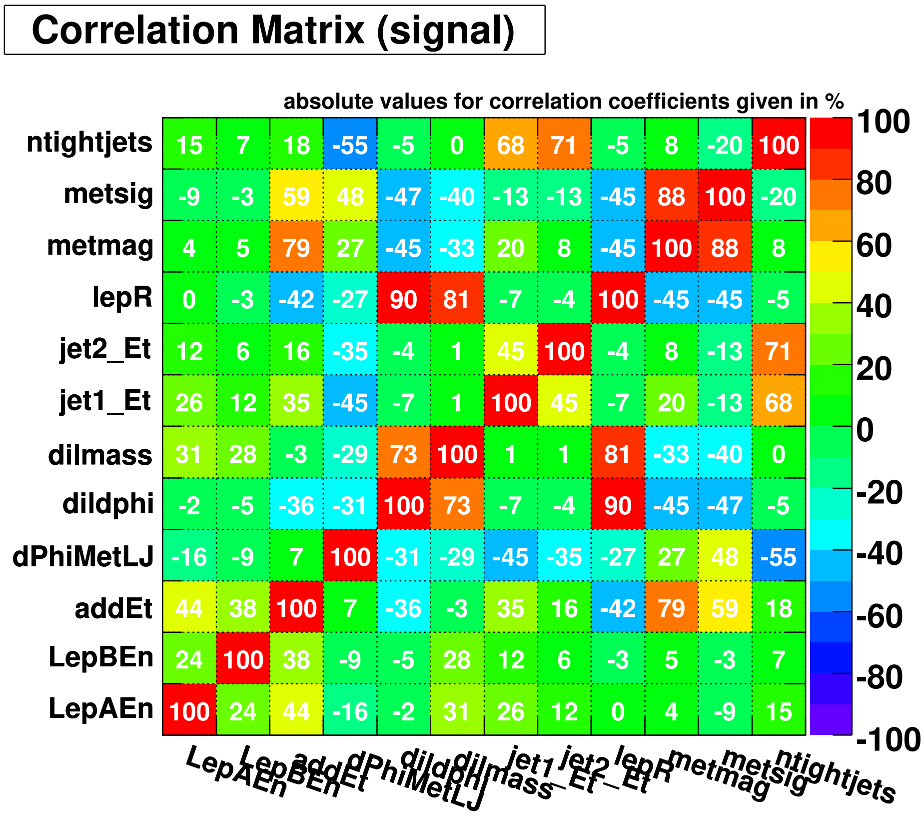

Figure 39: $M_h = 160GeV/c^2$, Signal correlation matrix, $\mu\mu$, NN variables, Pythia WW.

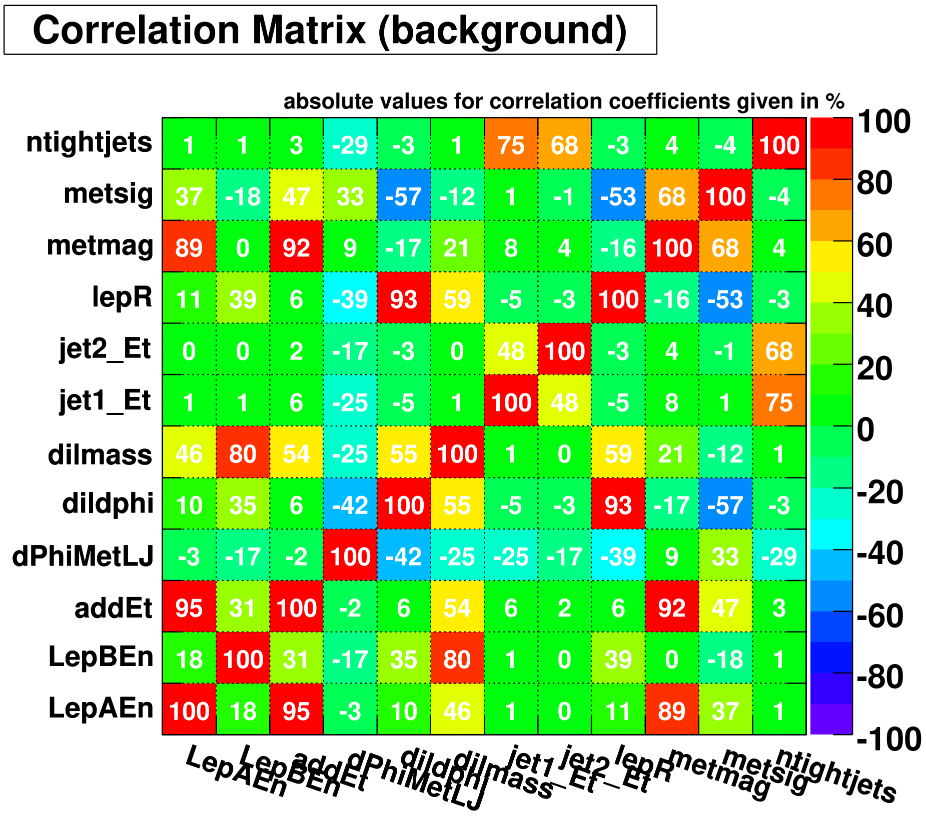

Figure 40: $M_h = 160GeV/c^2$, Background correlation matrix, $\mu\mu$, NN variables, Pythia WW.

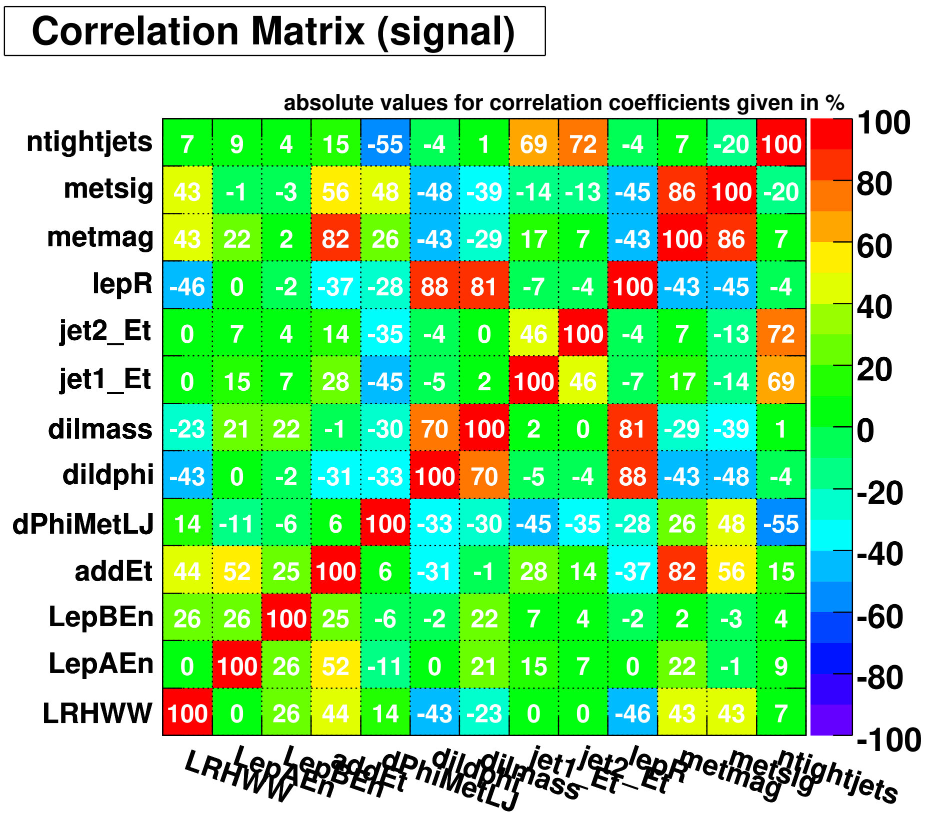

Figure 41: $M_h = 160GeV/c^2$, Signal correlation matrix, All leptons, NN + LRHWW variables, Pythia WW.

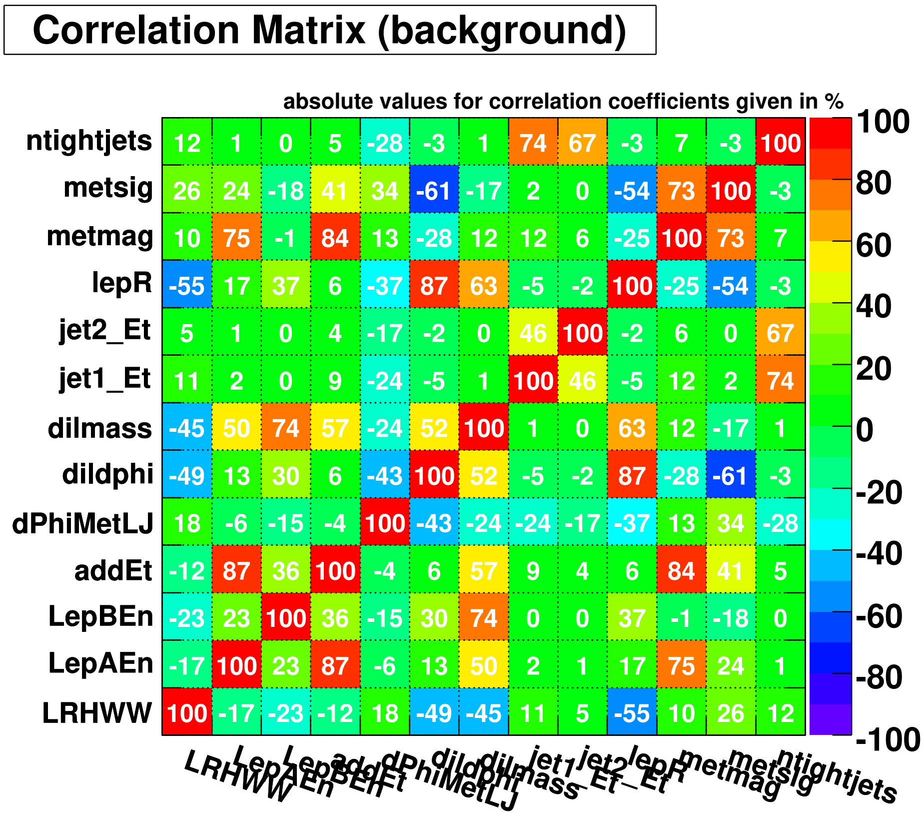

Figure 42: $M_h = 160 GeV/c^2$, Background correlation matrix, All leptons, NN + LRHWW variables, Pythia WW.

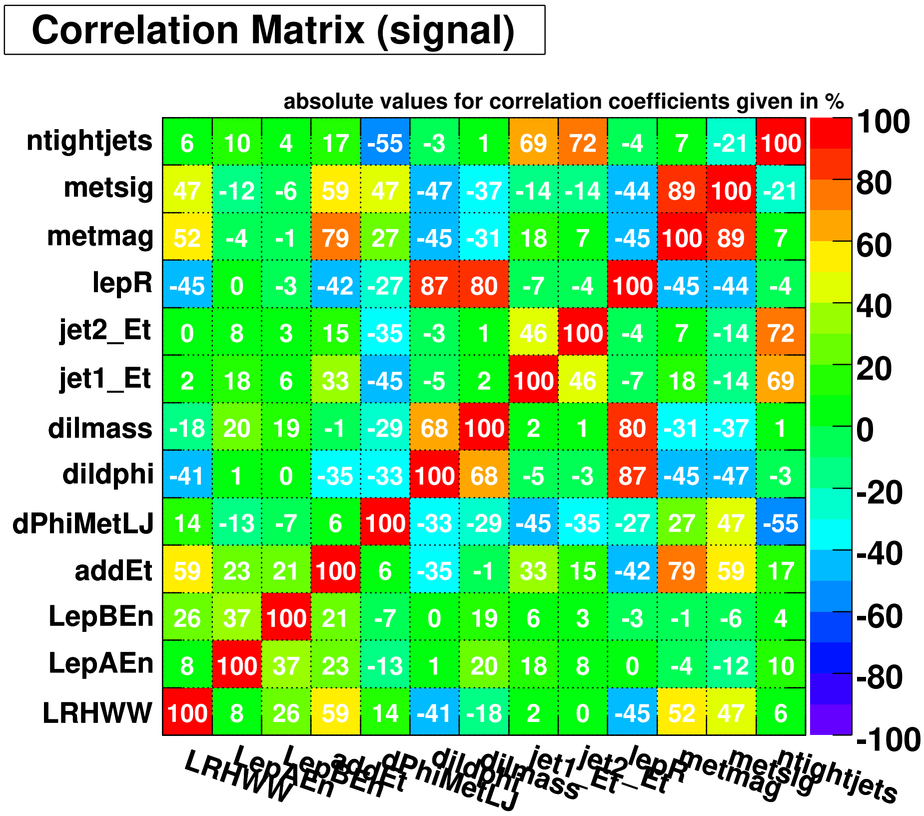

Figure 43: $M_h = 160 GeV/c^2$, Signal correlation matrix, ee, NN + LRHWW variables, Pythia WW.

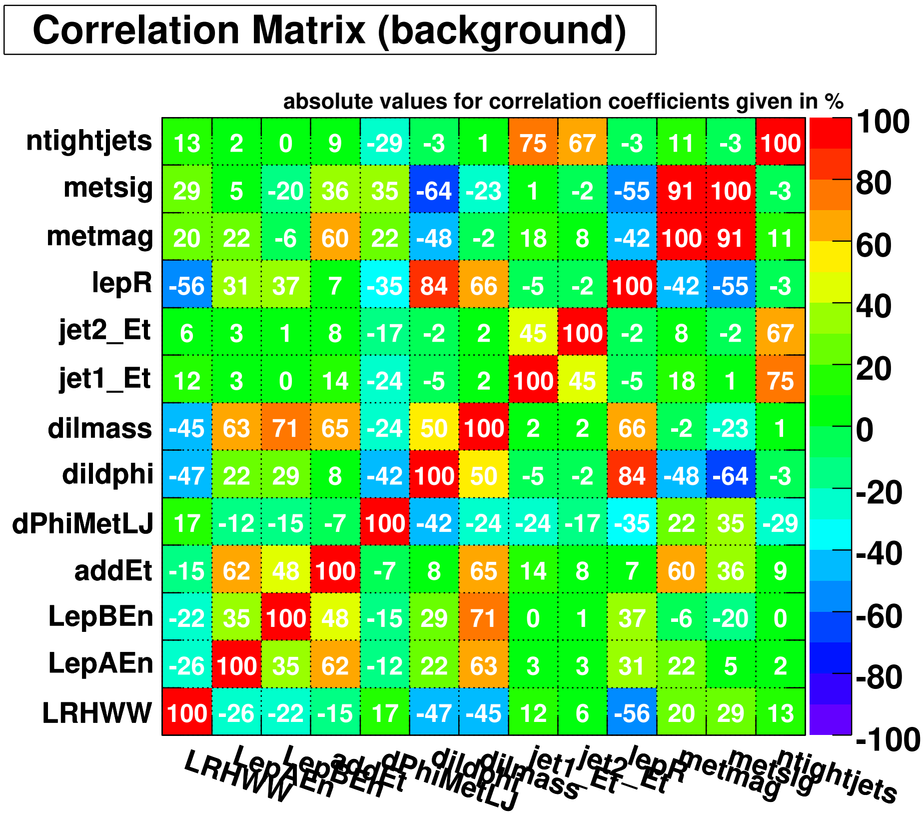

Figure 44: $M_h = 160 GeV/c^2$, Background correlation matrix, ee, NN + LRHWW variables, Pythia WW.

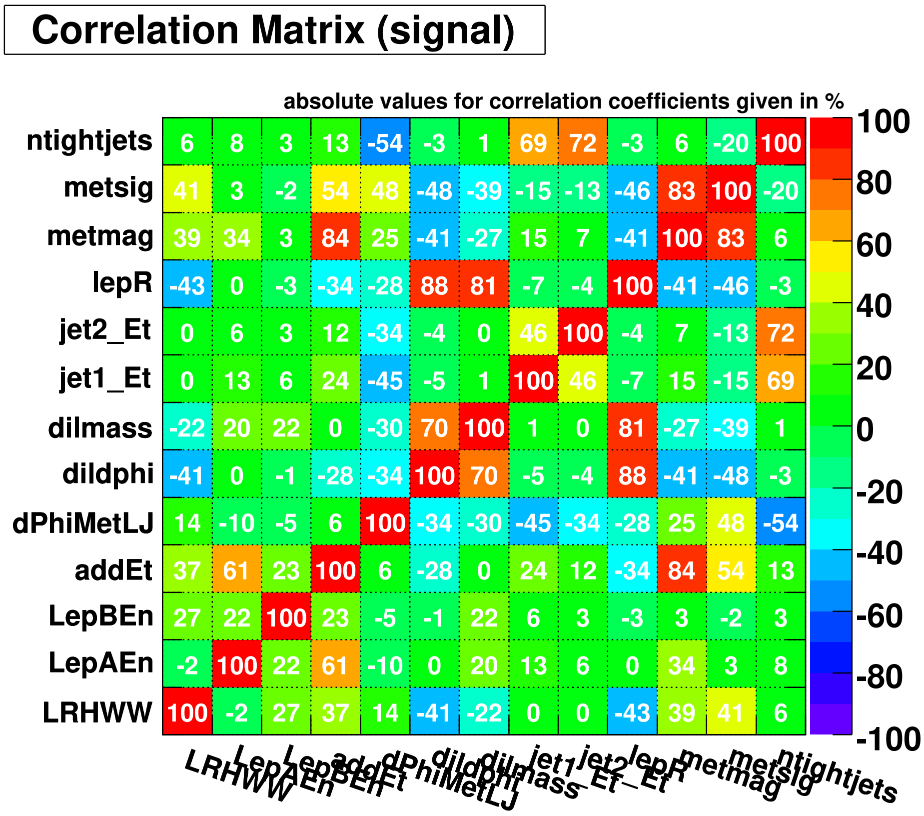

Figure 45: $M_h = 160 GeV/c^2$, Signal correlation matrix, eμ, NN + LRHWW variables, Pythia WW.

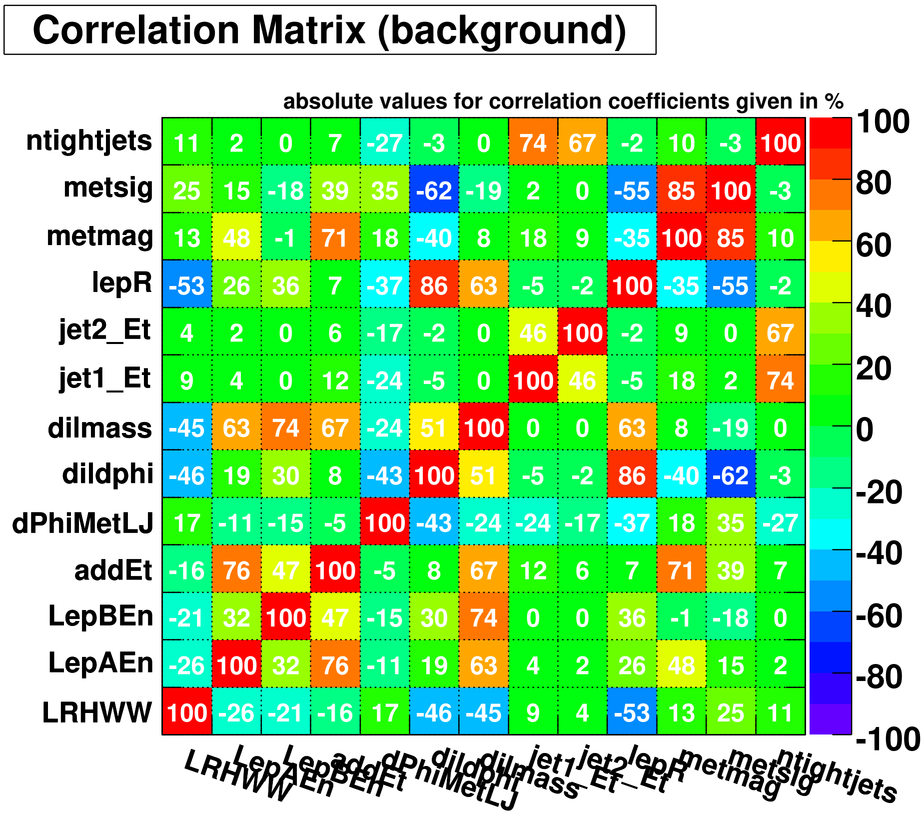

Figure 46: $M_h = 160GeV/c^2$, Background correlation matrix, eμ, NN + LRHWW variables, Pythia WW.

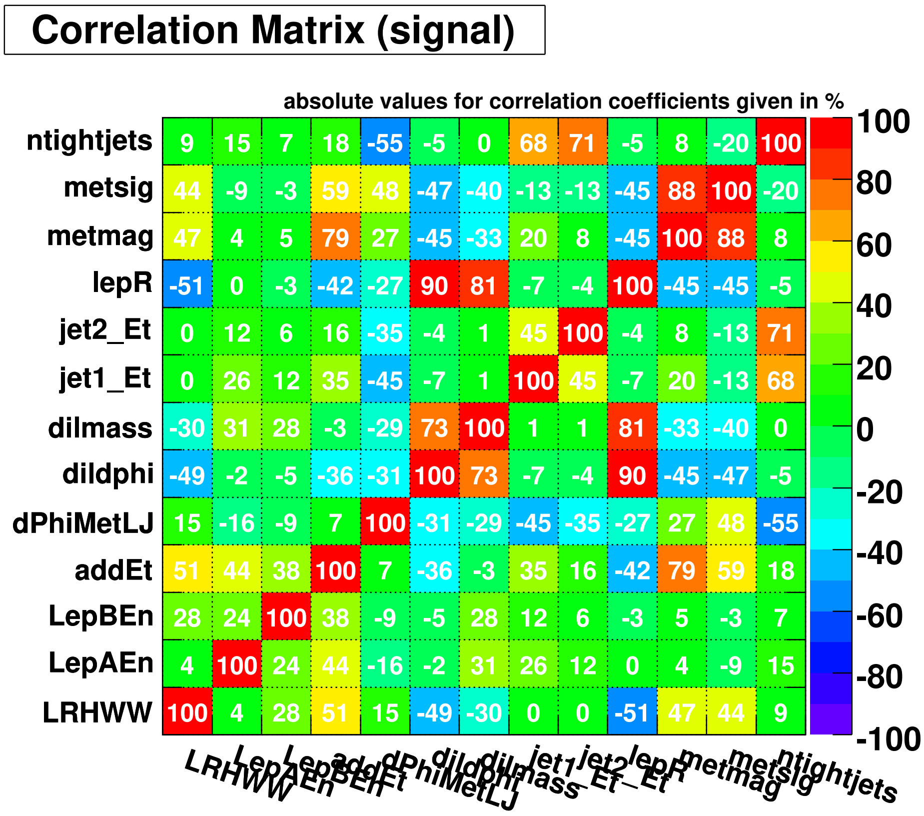

Figure 47: $M_h = 160GeV/c^2$, Signal correlation matrix, $\mu\mu$, NN + LRHWW variables, Pythia WW.

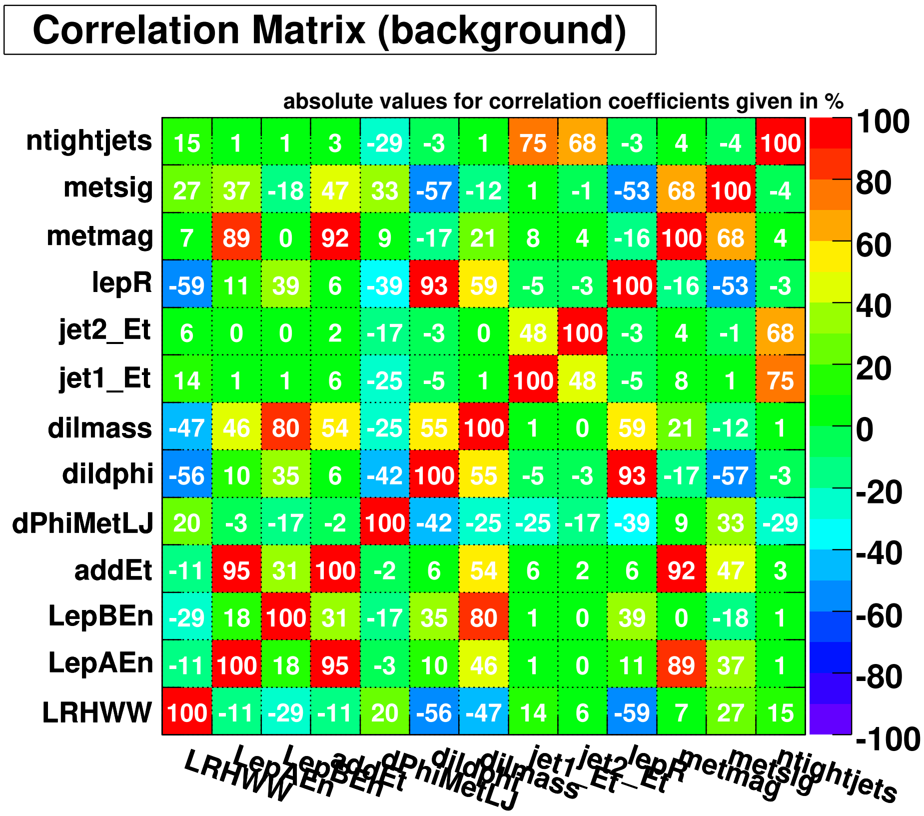

Figure 48: $M_h = 160 GeV/c^2$, Background correlation matrix, $\mu\mu$, NN + LRHWW variables, Pythia WW.

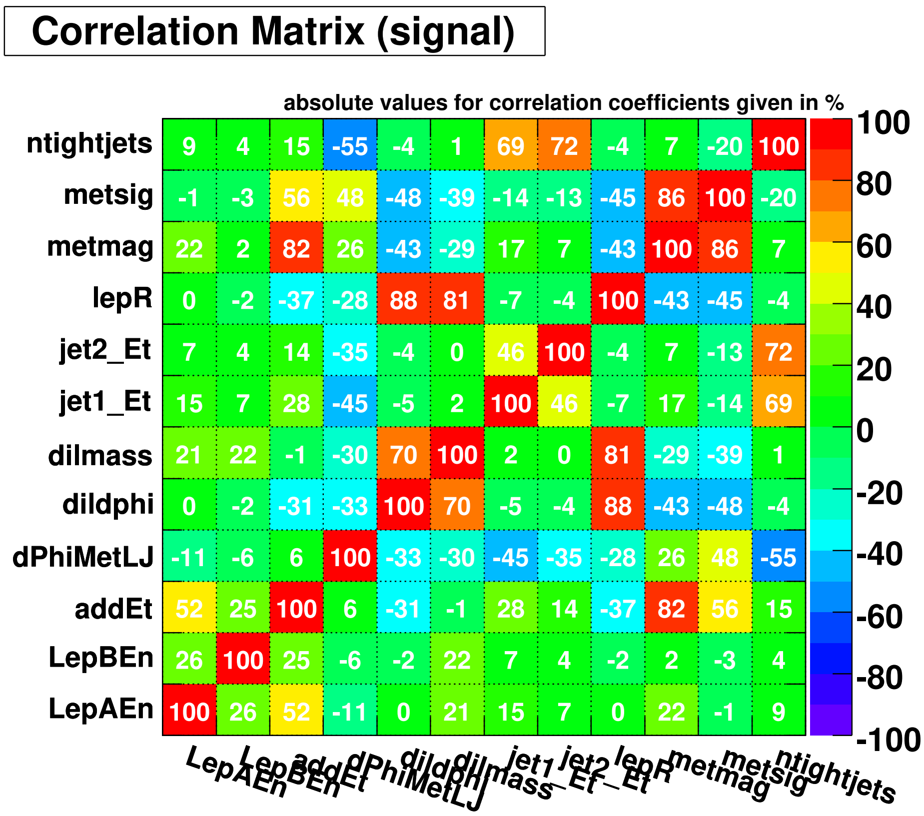

Figure 49: $M_h = 160 GeV/c^2$, Signal correlation matrix, All leptons, NN variables, MC at NLOWW.

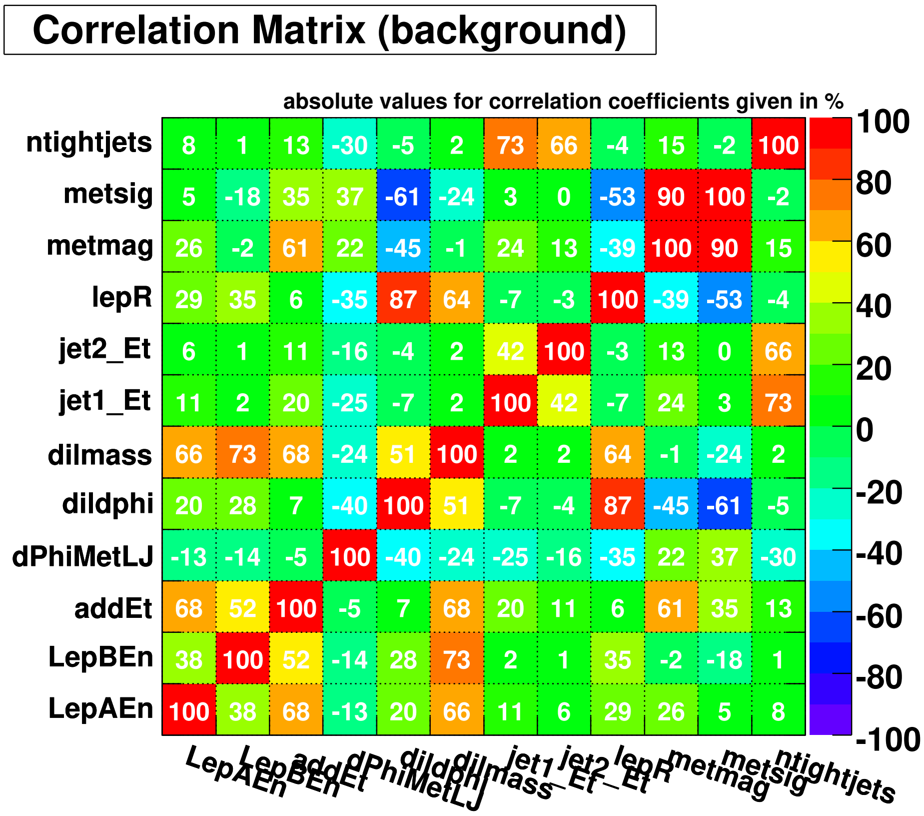

Figure 50: $M_h = 160 GeV/c^2$, Background correlation matrix, All leptons, NN variables, MC at NLOWW.

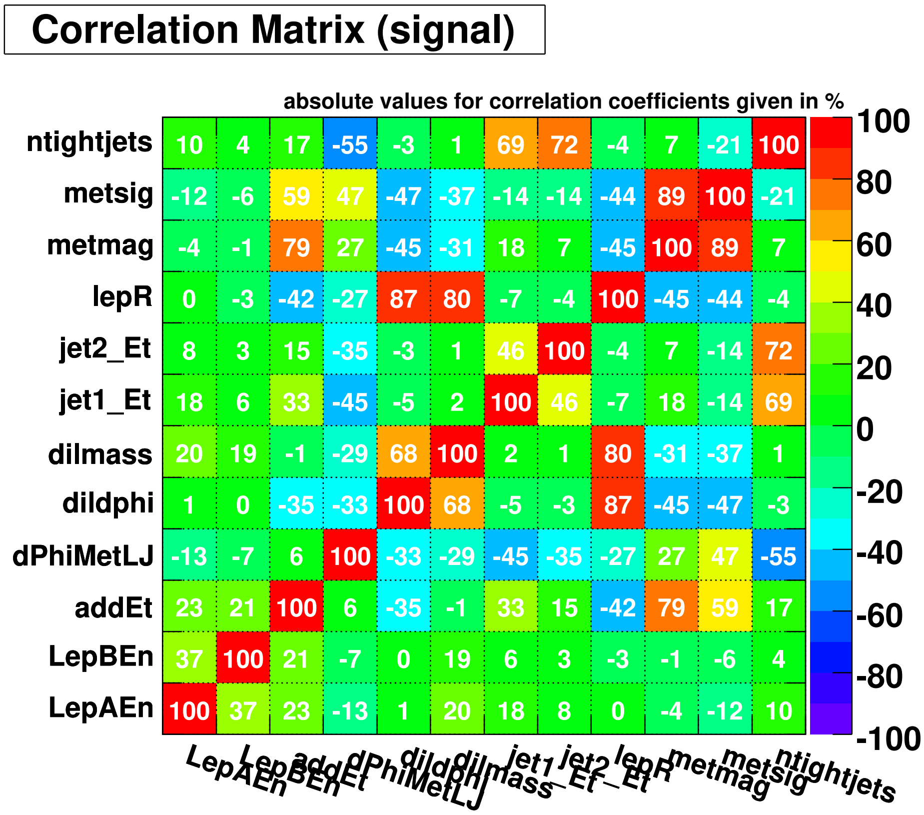

Figure 51: $M_h = 160 GeV/c^2$, Signal correlation matrix, ee, NN variables, MC at NLOWW.

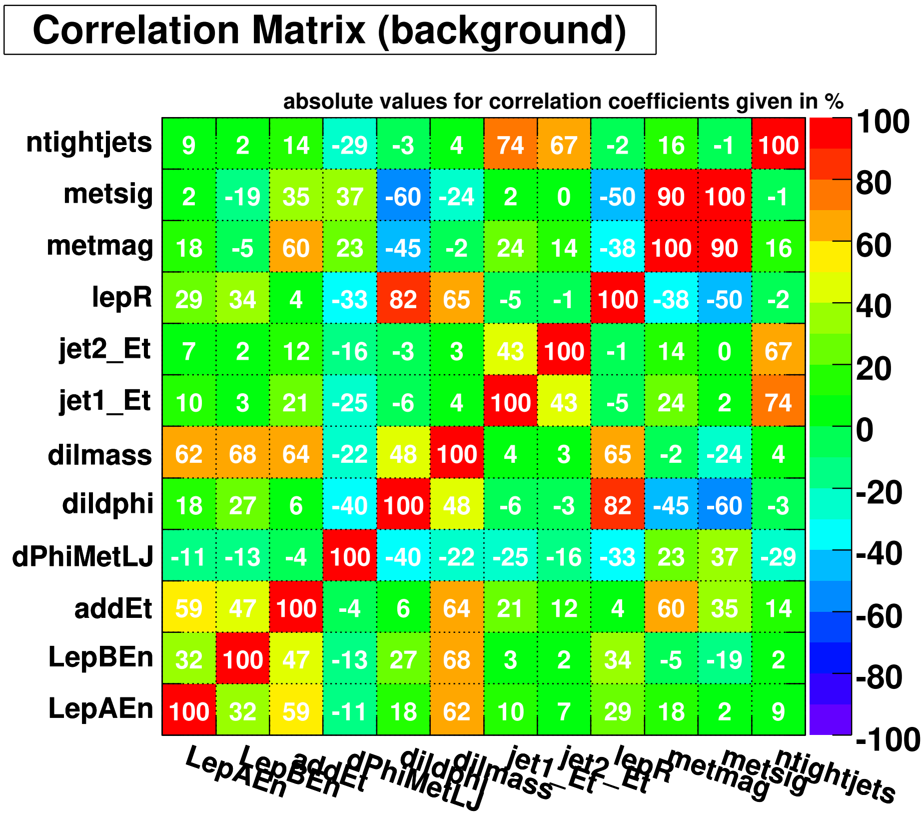

Figure 52: $M_h = 160 GeV/c^2$, Background correlation matrix, ee, NN variables, MC at NLOWW.

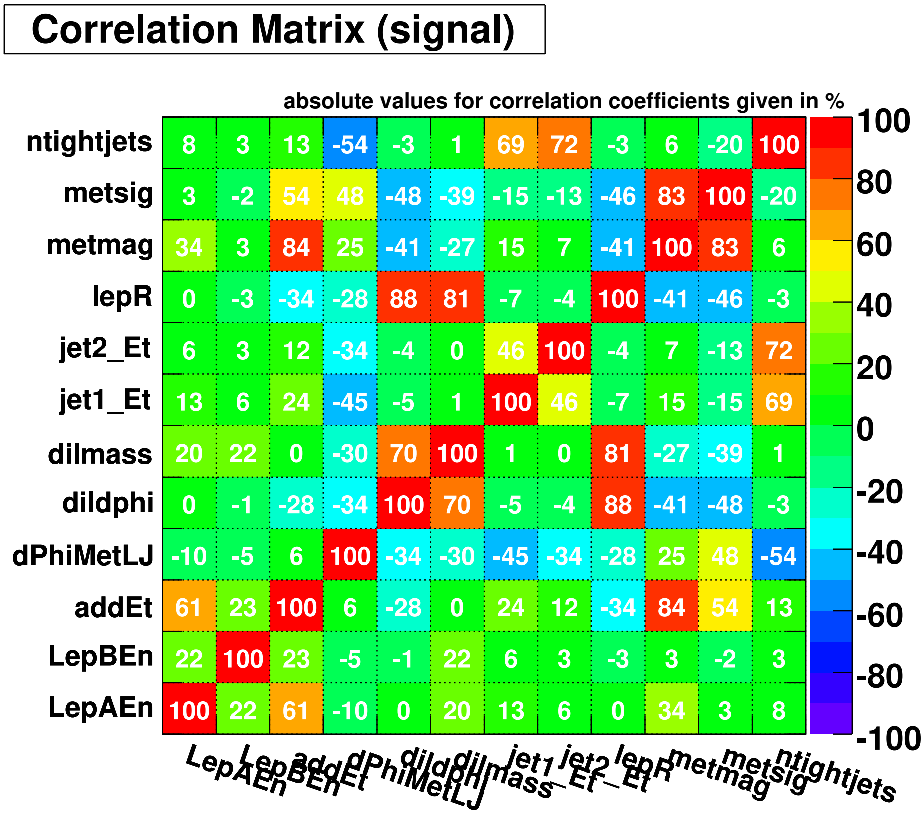

Figure 53: $M_h = 160GeV/c^2$, Signal correlation matrix, eμ, NN variables, MC at NLOWW.

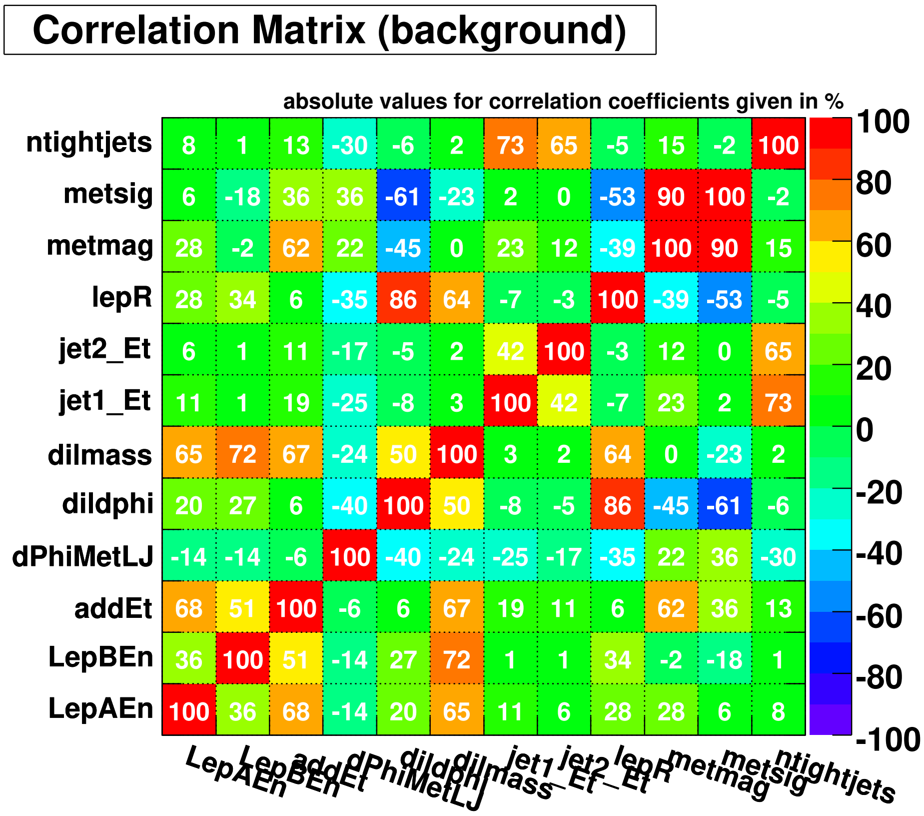

Figure 54: $M_h = 160 GeV/c^2$, Background correlation matrix, eμ, NN variables, MC at NLOWW.

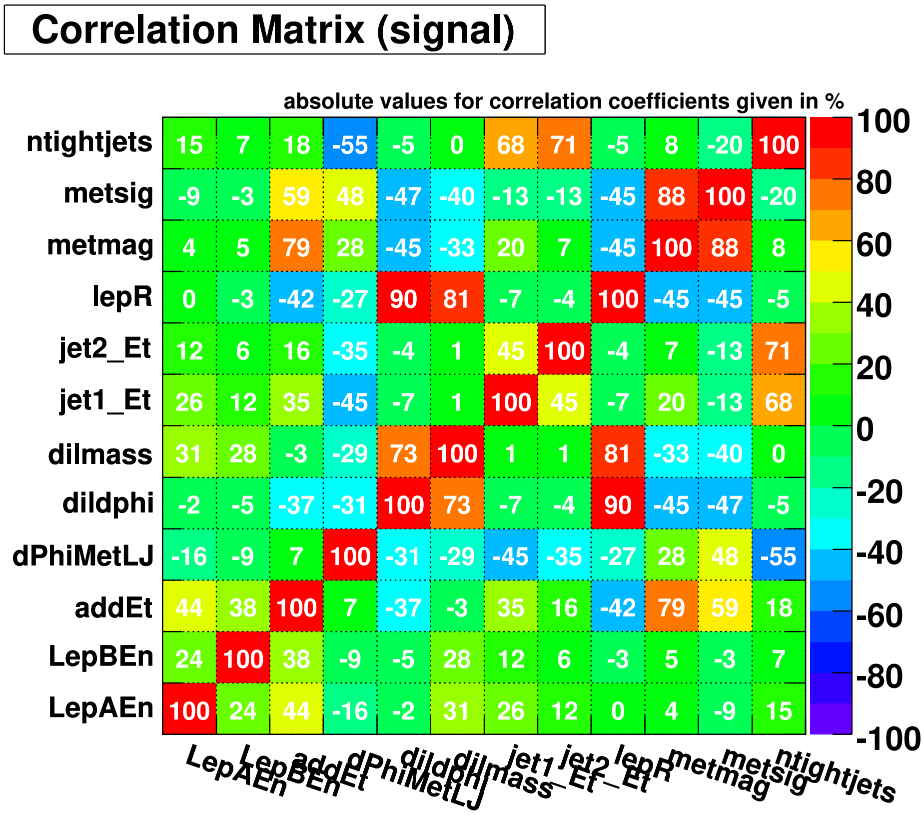

Figure 55: $M_h = 160GeV/c^2$, Signal correlation matrix, $\mu\mu$, NN variables, MC at NLOWW.

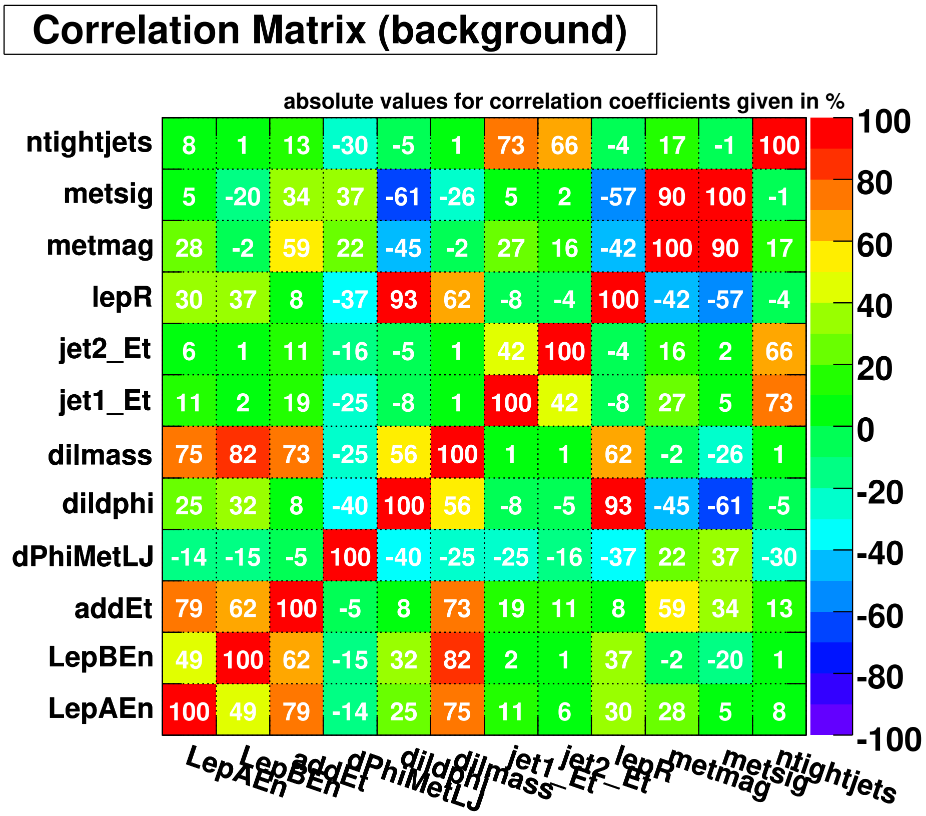

Figure 56: $M_h = 160GeV/c^2$, Background correlation matrix, $\mu\mu$, NN variables, MC at NLOWW.

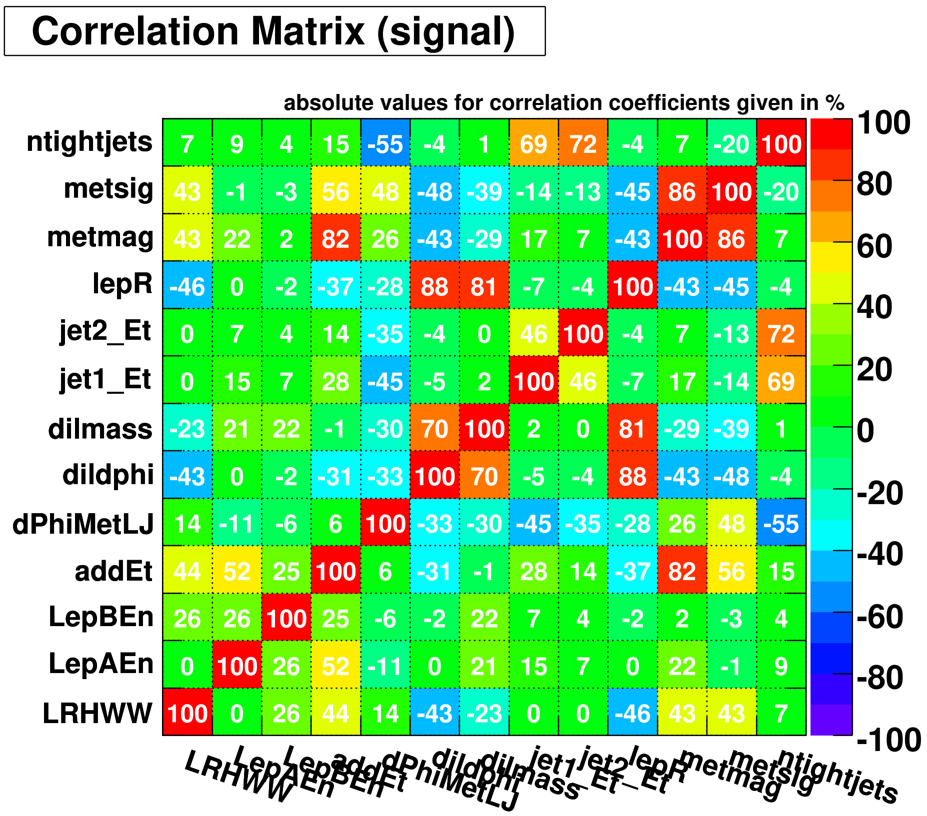

Figure 57: $M_h = 160GeV/c^2$, Signal correlation matrix, All leptons, NN + LRHWW variables, MC at NLOWW.

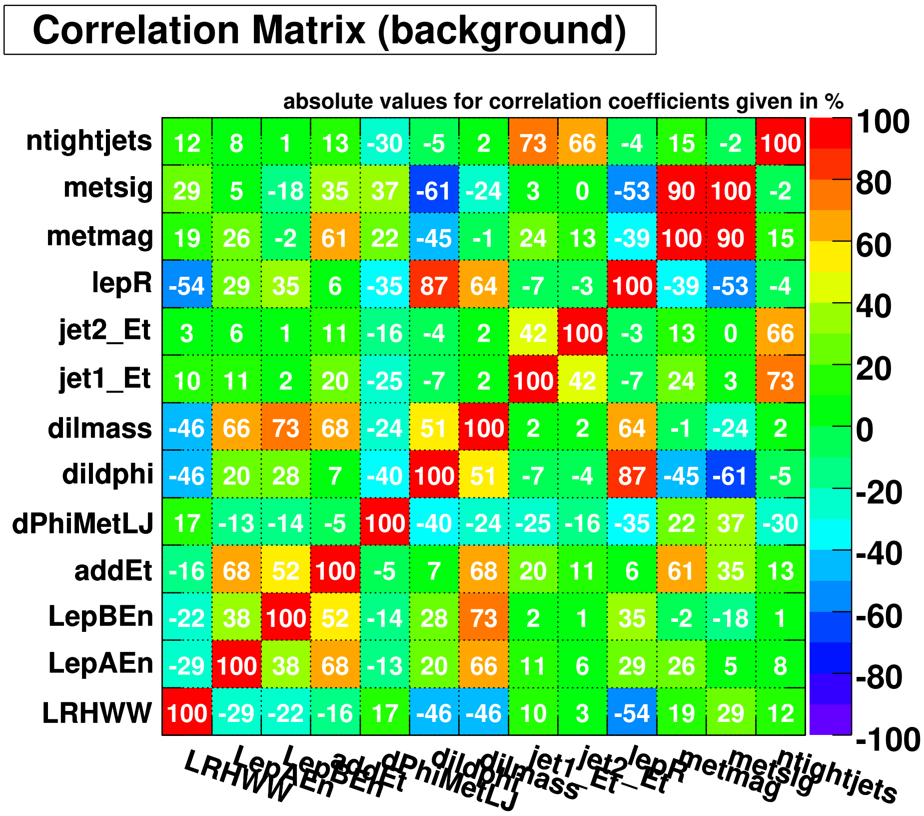

Figure 58: $M_h = 160GeV/c^2$, Background correlation matrix, All leptons, NN + LRHWW variables, MC at NLOWW.

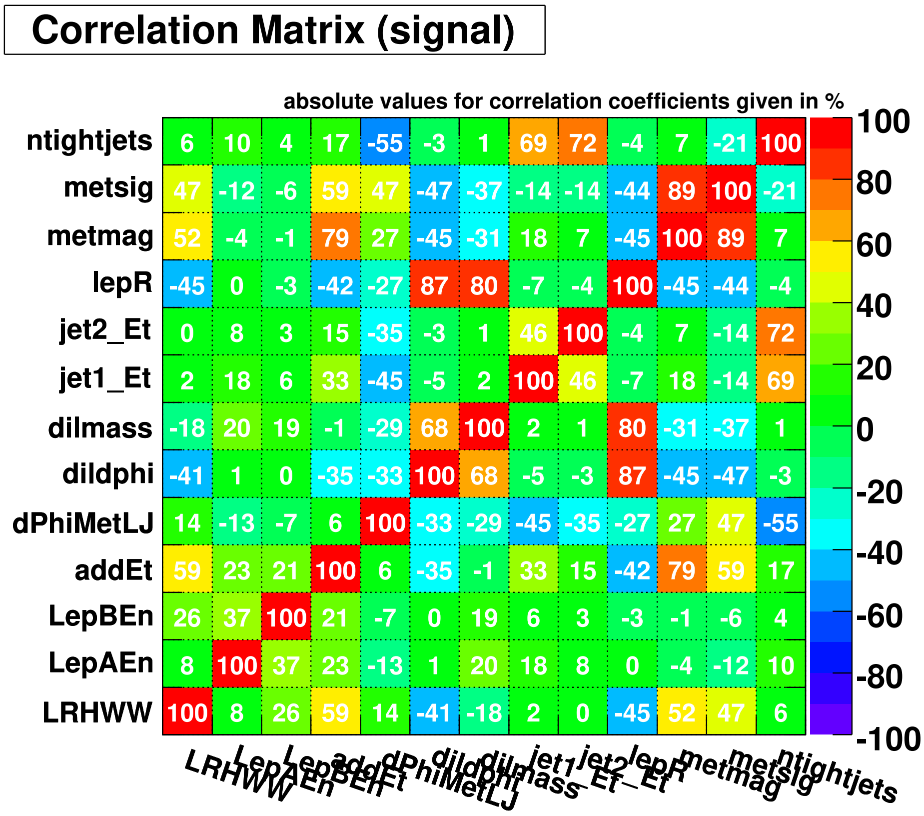

Figure 59: $M_h = 160 GeV/c^2$, Signal correlation matrix, ee, NN + LRHWW variables, MC at NLOWW.

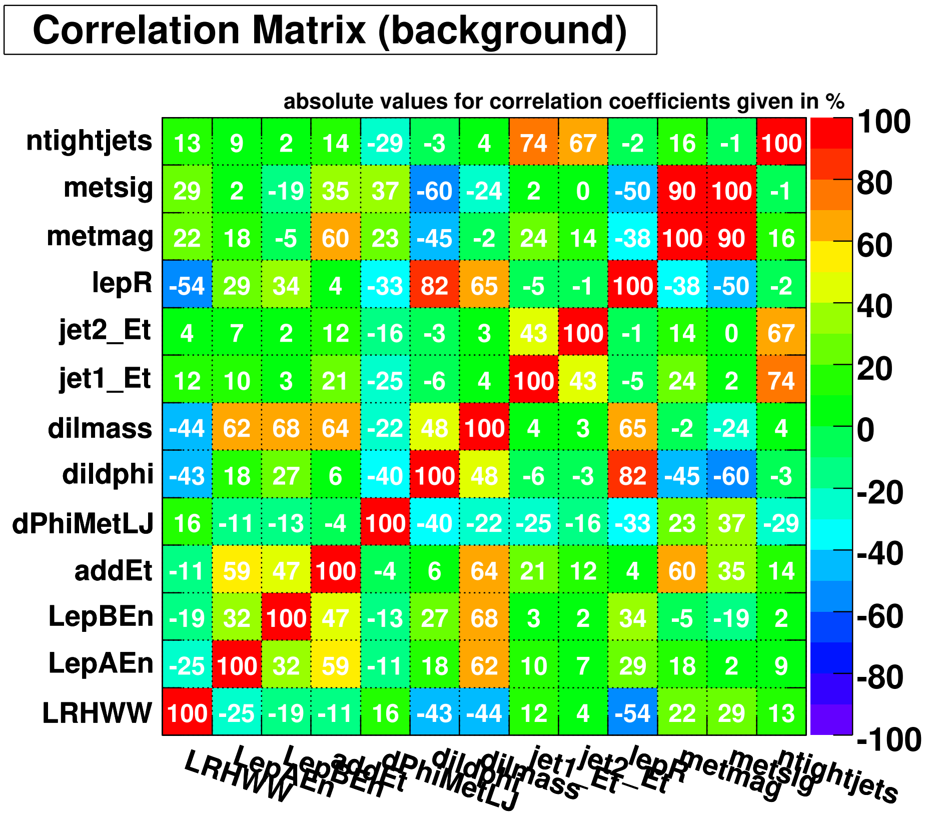

Figure 60: $M_h = 160 GeV/c^2$, Background correlation matrix, ee, NN + LRHWW variables, MC at NLOWW.

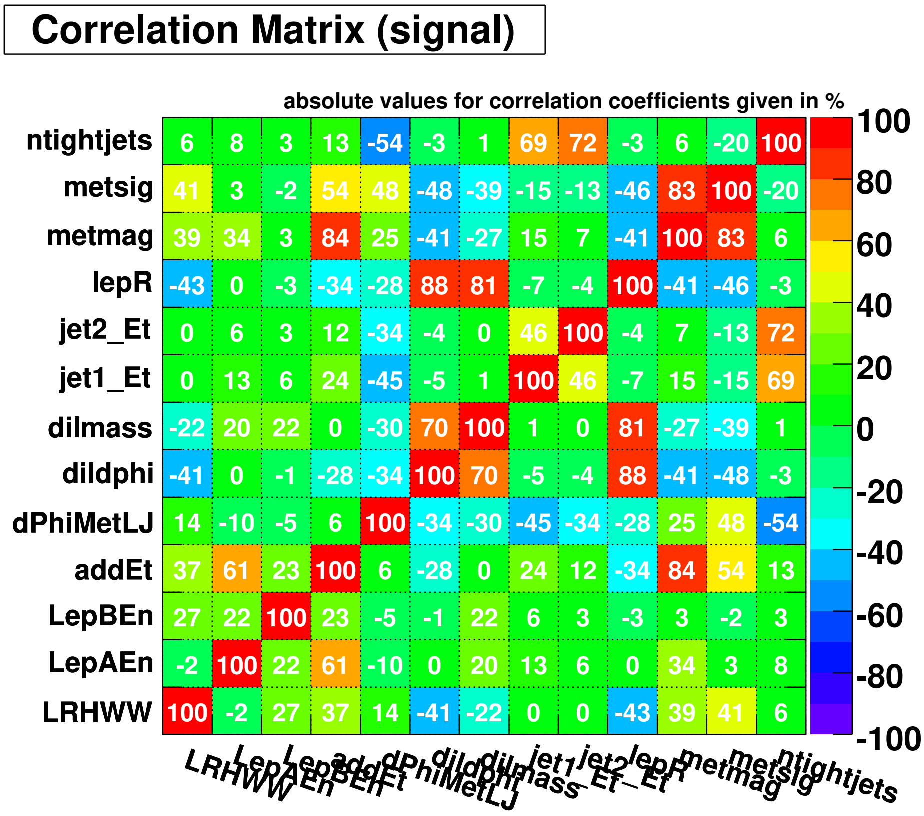

Figure 61: $M_h = 160 GeV/c^2$, Signal correlation matrix, eμ, NN + LRHWW variables, MC at NLOWW.

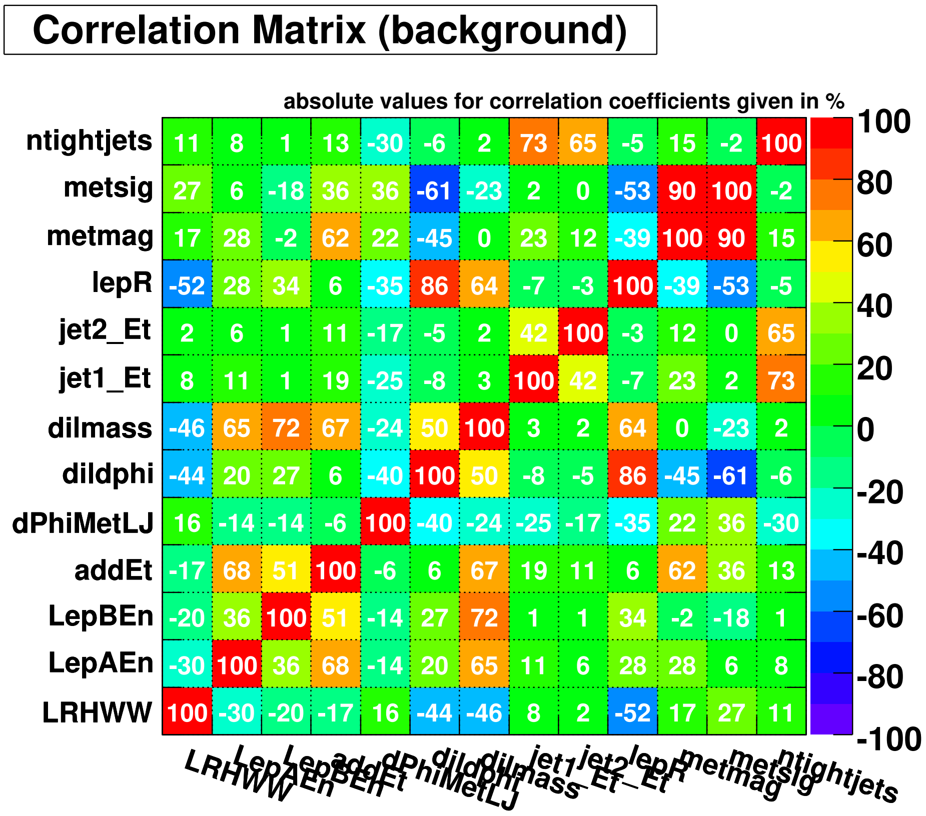

Figure 62: $M_h = 160 GeV/c^2$, Background correlation matrix, eμ, NN + LRHWW variables, MC at NLOWW.

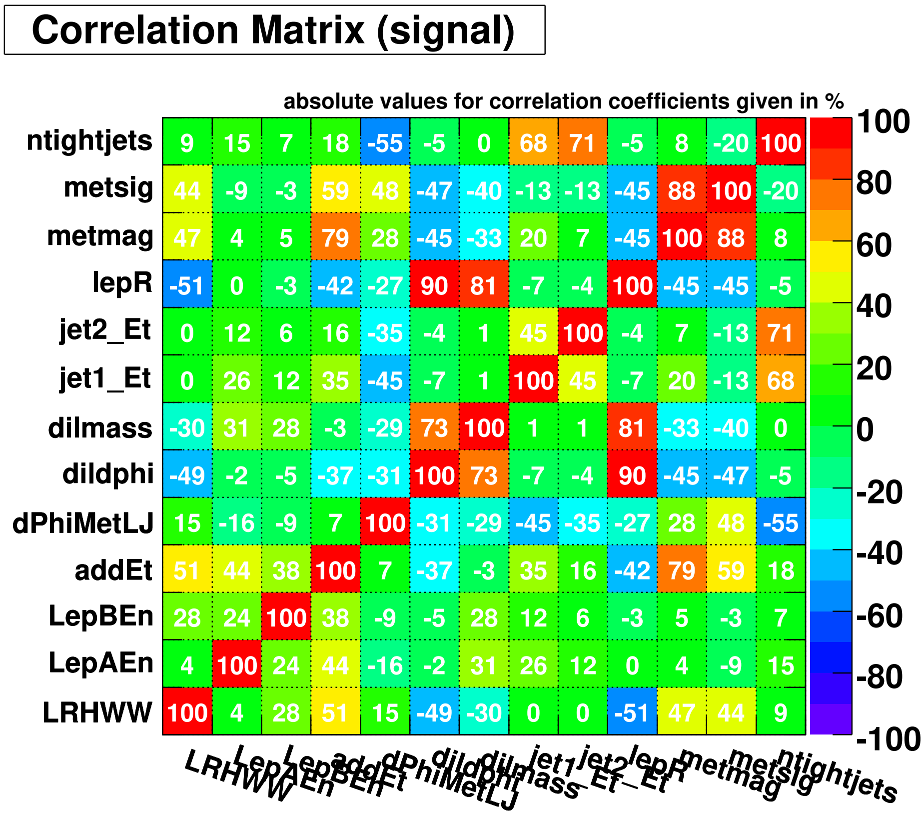

Figure 63: $M_h = 160GeV/c^2$, Signal correlation matrix, $\mu\mu$, NN + LRHWW variables, MC at NLOWW.

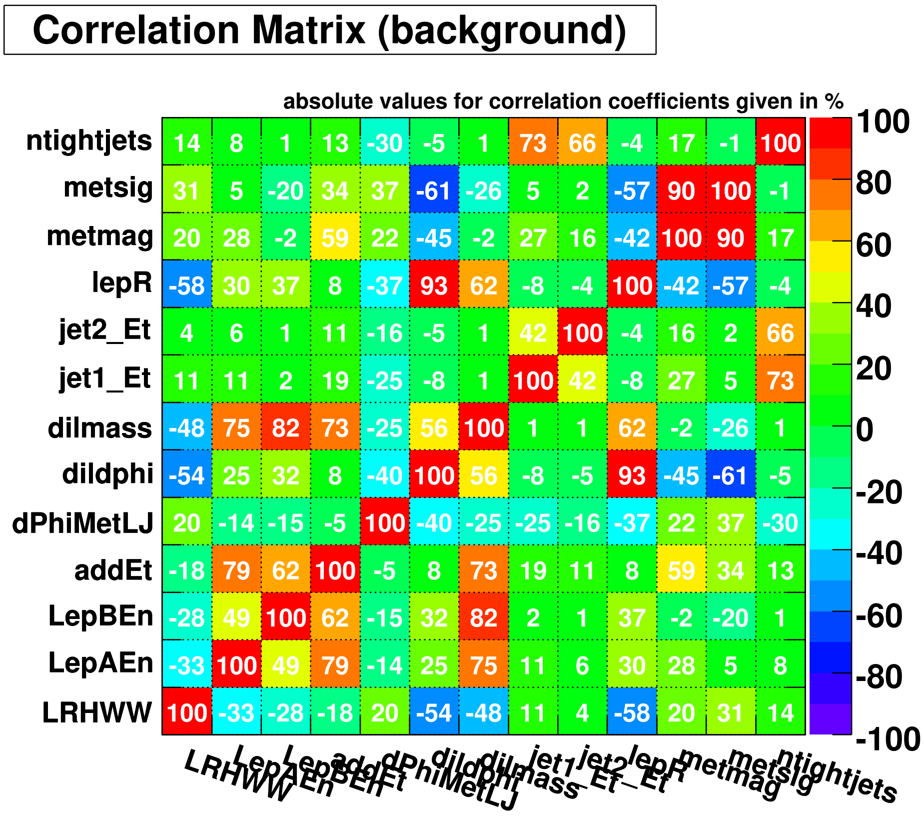

Figure 64: $M_h = 160GeV/c^2$, Background correlation matrix, $\mu\mu$, NN + LRHWW variables, MC at NLOWW.