INFN Istituto Nazionale di Fisica Nucleare
LABORATORI NAZIONALI DI FRASCATI
SIS-Pubblicazioni

LNF–10/04 (P)
February 16, 2010
CMS Note 2009/69

A NEW APPROACH IN MODELING THE RESPONSE OF RPC DETECTORS

L. Benussi[1], S. Bianco[1], S. Colafranceschi[1,2,3,*], F. L. Fabbri[1], D. Piccolo[1], G. Saviano[2], A. K. Bhattacharyya[3], A. Sharma[3]

[1)] *INFN Laboratori Nazionali di Frascati, Via E. Fermi 40, I-00044 Frascati, Italy*
[2)] *INFN Laboratori Nazionali di Frascati, Via E. Fermi 40, I-00044 Frascati, Italy and Università degli Studi di Roma "La Sapienza", Piazzale A. Moro*
[3)] *CERN CH-1211 Genéve 23 F-01631 Switzerlan*

[*)] *Corresponding Author*

Abstract

The response of RPC detectors is highly sensitive to environmental parameters. A novel approach is presented to model the response of RPC detectors in a variety of experimental conditions. The algorithm, based on Artificial Neural Networks, has been developed and tested on the CMS RPC gas gain monitoring system during commissioning.

PACS: 29.00.00

Submitted to
Journal of Instrumentation

1 Introduction

Resistive Plate Counters (RPC) detectors [1] are widely used in HEP experiments for muon detection and triggering at high-energy, high-luminosity hadron colliders, in astroparticle physics experiments for the detection of extended air showers, as well as in medical and imaging applications. At the LHC, the muon system of the CMS experiment [2]relies on Drift Tubes (DT), Cathode Strip Chambers (CSC) and RPCs [3] for the muon trigger system, with a total gas volume of about 50 m^3.

The constituent elements of RPCs are two parallel electrode, high-resistivity bakelite plates between which, in a 2 mm gap, a gas mixture at ordinary pressure is circulated. A 9.0-10.0 kV voltage drop is applied between plates. RPC use fluorine-based gas mixture whose main component is Freon. Due to the large gas volume and high costs, the RPC detector in CMS uses a recirculation gas system ("Closed Loop") with gas purifiers.

The absence of gas contaminants is of paramount importance in all gas detectors, and especially in RPC detectors due to the high reactivity of the F-based gas mixture used [4] [5]. The operation of the CMS RPC system is strictly correlated to the ratio of the gas components, and to the presence of pollutants that can be produced inside the gaps during discharges (i.e. HF produced by SF_6 or $C_2H_2F_4$ molecular break-up and further fluorine recombination), accumulated in the closed-loop, or by pollution that can be present in the gas piping system (tubes, valves, filters, bubblers, etc.) and flushed into the gaps by the gas flow. The monitoring of the presence of these contaminants, as well as the gas mixture stability, is therefore mandatory to avoid RPC damage and to ensure their correct functionality.

The response of RPC's is, however, strongly dependent on environmental parameters as temperature, pressure and relative humidity, as well as on other operational parameters typical of the application chosen, such as radiation dose. The dependence of RPC response from environmental parameters has been studied in the past [6] and several parameterizations have been proposed.

In this paper a new approach is proposed to model the response of RPC detector via a multivariate strategy. The algorithm, based on Artificial Neural Networks (ANN), allows one to predict the response of RPC's as a function of a set of parameters, once enough data is available to provide a training to the ANN. As initial stage, environmental parameters (temperature T, atmosferic pressure p and relative humidity H) have been considered. Further studies including radiation dose are underway and will be subject of a forthcoming paper.

The data for this study have been collected utilizing the gas gain monitoring (GGM) system [7][8][9] of the CMS RPC muon detector, during commissioning with cosmic

rays.

This paper is organized as follows. In Sect.2 the GGM is described, while the main features of the ANN-based algorithm are described in Sect.4. Datasets used are discussed in Sect.3, while in Sect.5 results are presented.

2 The Gas Gain Monitoring system as online calibration system

The GGM system monitors efficiency and signal charge continuously by means of a cosmic ray telescope based on RPC detectors. The GGM is described in details elsewhere [7][8][9].

The GGM system is composed by the same type of RPC used in the CMS detector but of smaller size (2 mm-thick Bakelite gaps, 50×50 cm^2).Twelve gaps are arranged in a stack. The trigger is provided by four out of twelve gaps of the stack, while the remaining eight gaps are used to monitor the working point stability. A 3/4 majority coincidence is required to acquire the cosmic ray event.

In standard CMS operation the eight gaps are arranged in three sub-system: one subsystem (two gaps) is operated with the fresh CMS mixture. The second sub-system (three gaps) is operated with CMS gas coming from the closed-loop gas system and extracted before the gas purifiers, while the third sub-system (three gaps) is operated with CMS gas extracted from the closed-loop after the gas purifiers. A fraction of the eight gaps work in pure avalanche mode, while the remaining will be operated in avalanche+streamer mode. Comparison of signal charge distributions and the ratio of the avalanche to streamer components of the ADC provides a monitoring of the stability of working point for changes due to gas mixture variations.

In this study, the GGM was operated in open loop mode with freon 95.5%, isobutane 4.2%, SF_6 0.3% gas mixture. Six gaps out of eight were used. The monitoring is performed by measuring the charge distributions of each chamber. The six gaps are operated at different high voltages, fixed for each chamber, in order to monitor the total range of operating modes of the gaps (Tab. 1). The operation mode of the RPC changes as a function of the voltage applied.

Table 1: Applied high voltage power supplies for GGM RPC detectors used in this study.

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8
Applied high voltage (kV)	10.2	9.8	10.0	N/A	N/A	10.4	10.2	10.4

3 Environmental parameters and Datasets

The experimental setup is shown in Fig. 1 and Fig. 2. The environmental parameters are monitored by an Oregon Scientific weather station WMR100. The WMR100 has relative humidity, pressure and temperature built-in sensors in the main station and the possibility to add remote wireless sensors for both temperature and relative humidity. The DAQ has been modified in order to acquire via USB the environmental informations and merge environmental parameters with performance detector parameters such as efficiency, average anodic charge and avalanche and streamer area. The accuracy of the temperature sensor $\pm 1^o$C in the range $0 - 40^o$C and the resolution is 0.1^oC. The relative humidity sensor has an operating range from 2% to 98% with a 1% resolution, $\pm 7\%$ absolute accuracy from 25% to 40%, and $\pm 5\%$ from 40% to 80%. The barometer operational range is between 700 mbar and 1050 mbar with a 1 mbar resolution and a ± 10 mbar accuracy. The online monitoring system records the ambient temperature, pressure and humidity of the GGM box, as well as the gas mixture temperature before and after each RPC gap, also the pressure and the relative humidity are monitored and recorded both inside the box that contains the RPC stack and in the gas mixture before and after each gap. The dataset used is composed of four periods, each period composed of runs. Each run contains 10^4 cosmic ray events where environmental parameters and GGM anodic output charges are collected. The acquisition rate is typically 9.5 Hz. Tab.2 shows features of the three periods.

Table 2: Periods summary table.

Periods	Number of runs	Timeline		Environmental range		
		Start	Finish	temperature (°C)	relative humidity (%)	pressure (mbar)
1	396	01/09/08	26/09/08	16.5-20.5	35-70	953-970
2	86	27/09/08	01/10/08	16.5-20.5	35-55	962-973
3	460	01/10/08	17/10/08	16.5-22.0	32-60	958-980
4	92	27/10/08	1/11/08	18.5-20.0	35-45	944-960

4 The Artificial Neural Network simulation code

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information [10]. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of input units is connected to a layer of hidden units, which is connected to a layer of output unit as shown in Fig. 3.

Figure 1: Gas Gain Monitoring system setup.

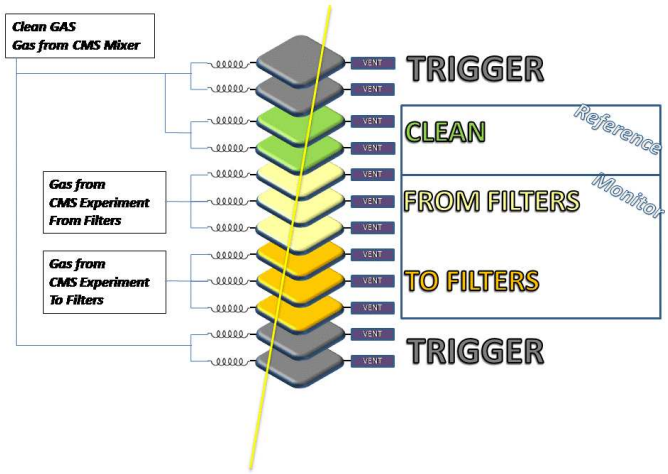

Figure 2: Layout of GGM experimental setup.

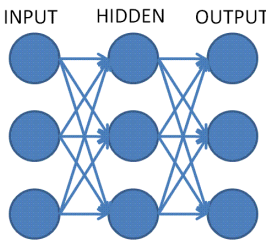

Figure 3: Example of a simple Neural Network configuration.

The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. The behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units. For this study temperature, humidity and pressure have been selected as inputs and anodic charge as output. For the ANN an error back-propagation pattern with 3 hidden layers. It was demonstrated [11] that the number of layers is not critical for the network performance, so we decided to go with 3 layers and give to the neural network a sufficient number of hidden units automatically optimized by a genetic algorithm that can take into account several configuration.

For each configuration, in each layer there are a number of neurons between 2 and 12, the genetic algorithm performs the training process with an estimation of the global error; then the configuration is stored and the genetic algorithm continues to evaluate a slightly different configuration. Once the algorithm has taken into account all the possible configurations the best one in terms of global error is chosen. The error is calculated point by point just with the comparison between the neural network forecast and the experimental data.

During the training phase the network is taught with environmental data as input, the output depends on the neuronal weights, that at the very beginning are initialized with random numbers. The network output is compared with the experimental data we want to model, and in this phase the network has an estimation of the error, the error itself is back-propagated into the network in order to modify the weights to minimize the error.

Once the training is complete the network's weights are optimized to have the minimum error for the chosen network pattern, the genetic algorithm goes on considering several configuration in an automatic way and the really optimal network along with its structure is returned. Such a network is ready to be executed in a non taught period, with different input data.

5 Results

Typical simulation outputs show generally good agreement between data and prediction Fig. 4. In periods where prediction is not accurate, the discrepancy is typically concentrated in narrow regions (”spikes”) as shown in Fig. 5.

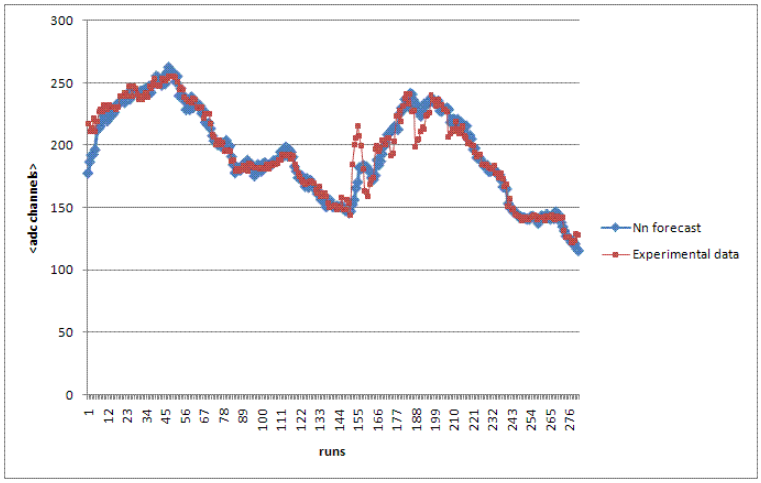

Figure 4: Gap 7 trained on period3 - prevision on period3.

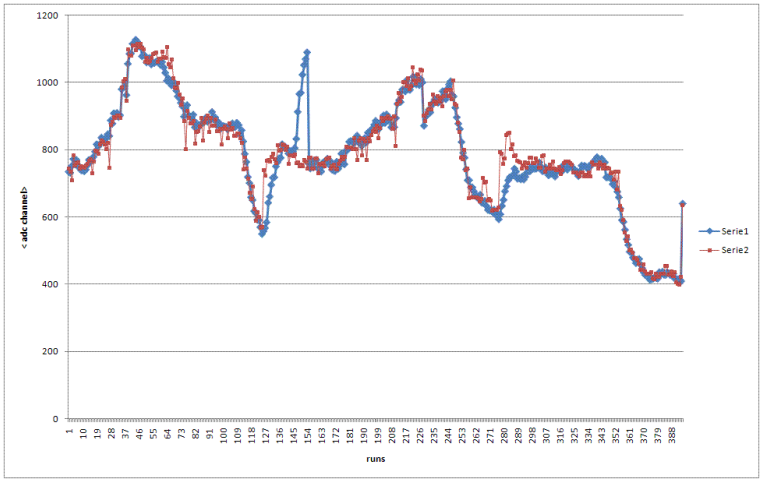

Figure 5: Gap 8 trained on period1 - prevision on period1.

The overall agreement between data and prediction is shown in Figs. 6,7 where the quantity

$$\frac{\Delta Q}{Q} \equiv \frac{Q_{EXP} - Q_{PRED}}{Q_{EXP}} \tag{1}$$

is plotted as a function of the experimental points for all four periods, divided for training and prediction respectively. The error distribution for the predictions is much wider than for the training as expected (Fig. 8).

The distribution of the error for the predictions shows a $\sigma_{fwhm} \sim 7\%$ where $\sigma_{fwhm} \equiv \Gamma_{fwhm}/2.36$ width with very long tails, due to points with very large discrepancy between data and prediction. The cases with very large discrepancy were studied in detail, and found to be characterized by a (p, T, H) value at the edges of the parameter space. Fig. 9 shows the correlation between (p,T,H) and error.

To quantify the position of each point in the (p, T, H) parameter space, the centroid of the distribution of runs in the (p, T, H) parameter space

$$C_i \equiv \frac{\sum_{i=1,N} x_i}{N} \quad ; \quad \mathbf{x} \equiv (p, T, H) \tag{2}$$

and the norm $\|\mathbf{x}\|$ the distance of each run to the centroid

$$\|\mathbf{x}\| \equiv \sqrt{\sum_{j=1,3} (x_j - C_j)^2} \tag{3}$$

were computed. The distribution of the $\frac{\Delta Q}{Q}$ error as a function of the norm $\|\mathbf{x}\|$ (Fig. 10) shows three distinct structures. The satellite bands with very large error were studied in detail. All data point in such bands belong to period four and channel six for which problems were detected. Period four and channel six therefore were excluded by the analysis. The distribution of the error as a function of r_b after this selection is shown in Fig. 11, with a $\sigma fwhm \sim 4\%$ width and nongaussian tails extending up to $\frac{\Delta Q}{Q} = 200\%$.

A selection on the fiducial volume in the **x** parameter space (Tab. 3) was applied in order to avoid runs on the boundaries of the (p, T, H) space. After the selection cuts, predictions on two periods based on training on the third period were performed. Results are shown in Fig. 13 to Fig. 21. The selection cuts provide $\sigma_{fwhm} \sim 5\%$ error, as summarized in Tab.4.

Table 3: Synopsis of selection cuts for fiducial volume.

$(958 < p < 968)$mbar	$(19.4 < T < 20.4)^o$C	$(34 < H < 44)\%$

6 Conclusions

A new approach based on ANN in modeling the response of RPC detectors was presented, and preliminary results obtained with data from the CMS RPC GGM system were

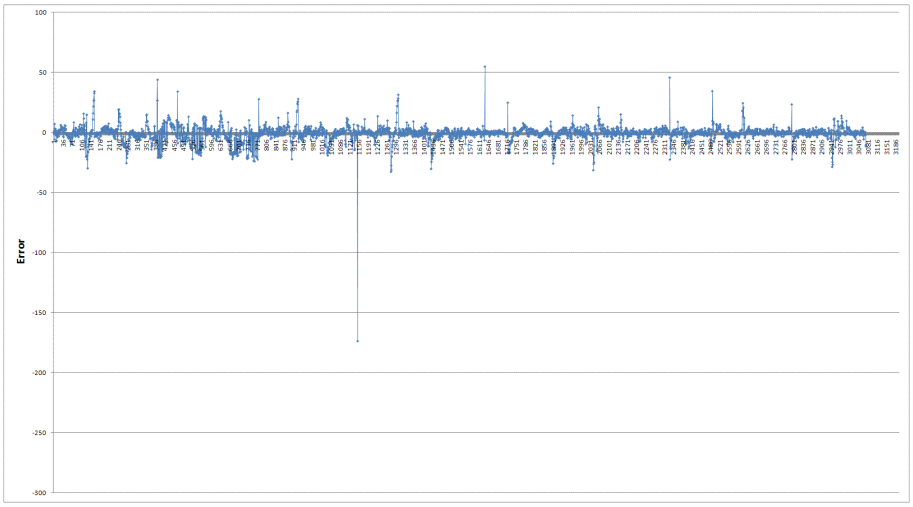

Figure 6: Distribution of error for training for all runs.

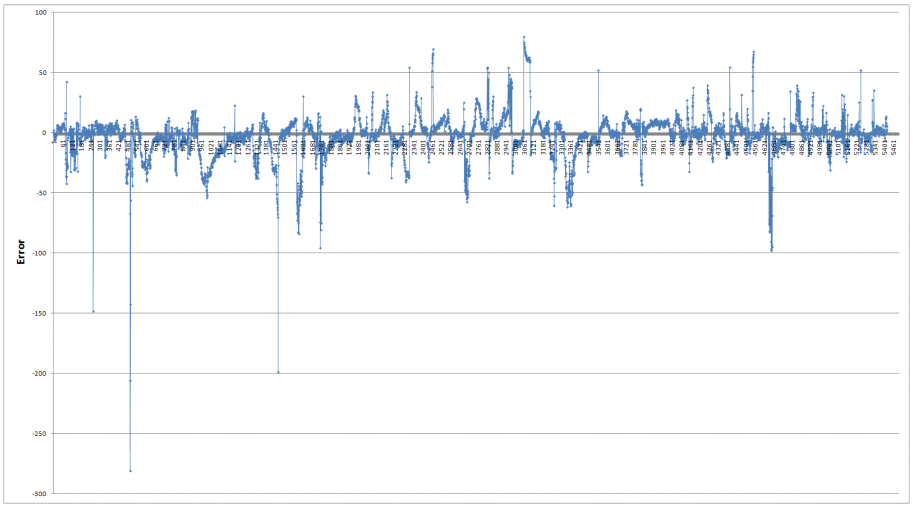

Figure 7: Distribution of error for predictions for all runs.

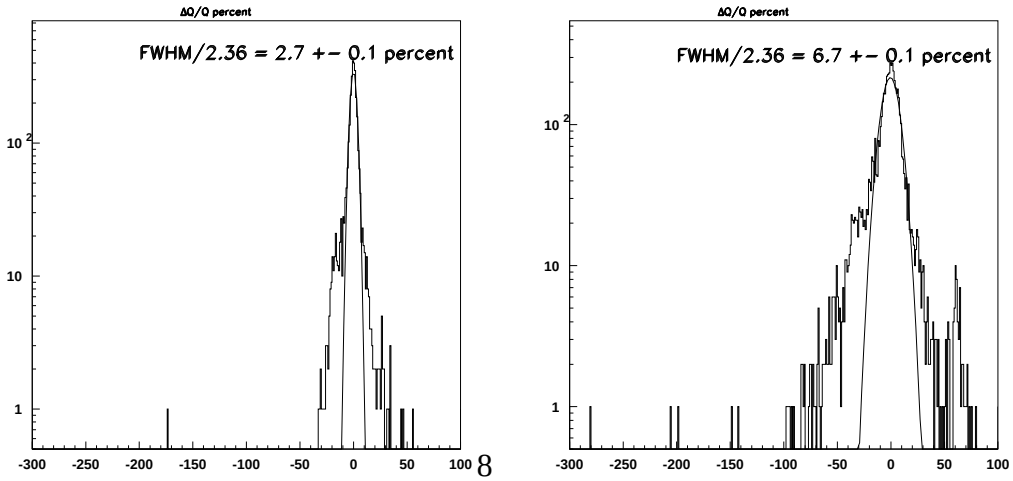

Figure 8: Error for training (left) and prediction (right) for all runs. Gaussian fit superimposed.

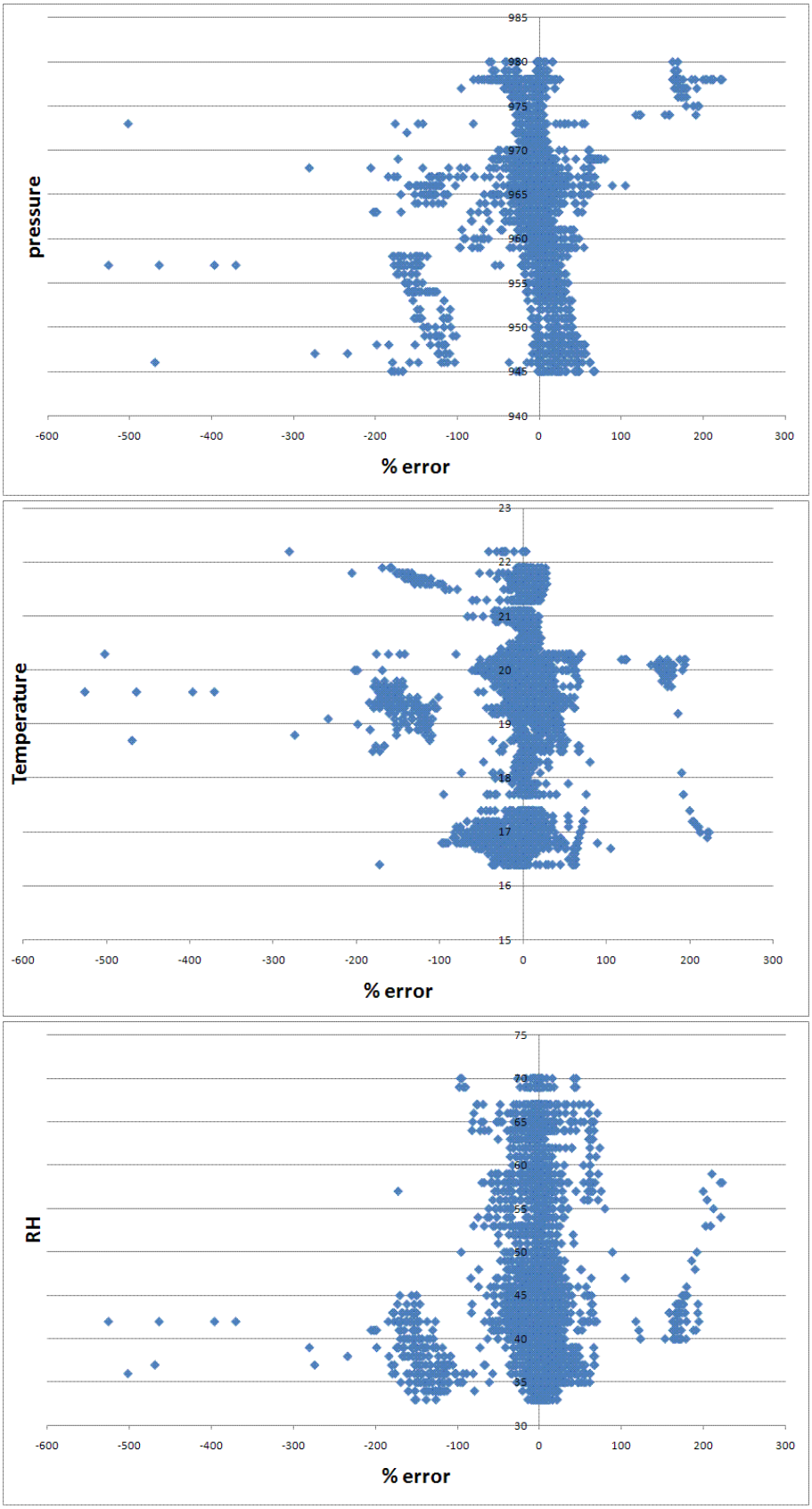

Figure 9: (Top) pressure (center) temperature (bottom) humidity versus error. $\frac{\Delta Q}{Q}$

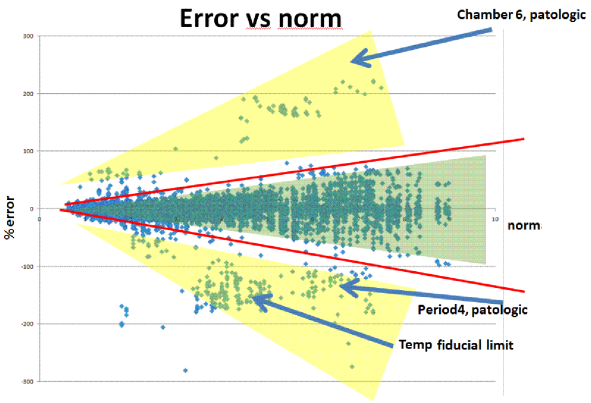

Figure 10: Distribution of error as a function of the $||\mathbf{x}||$ norm for all runs, six chambers and both training and prediction.

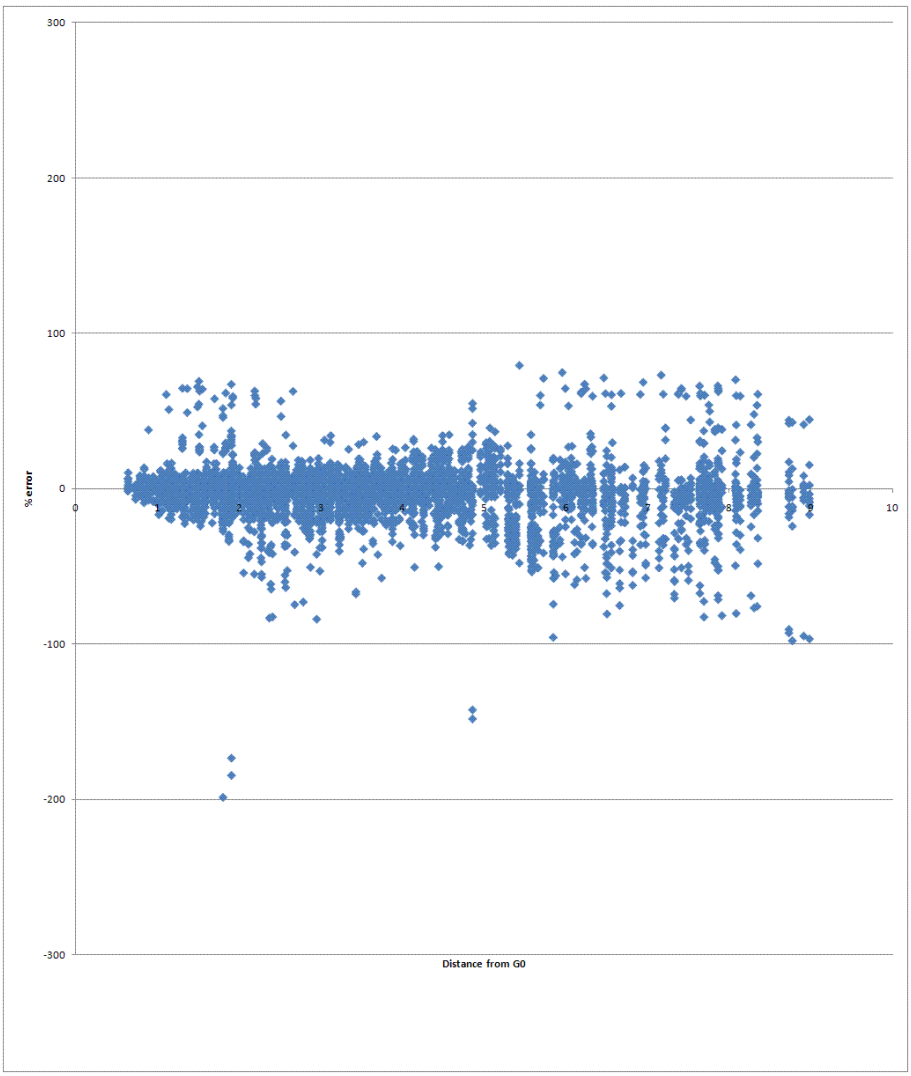

Figure 11: Error $\frac{\Delta Q}{Q}$ vs norm $||\mathbf{x}||$ after selection cuts.

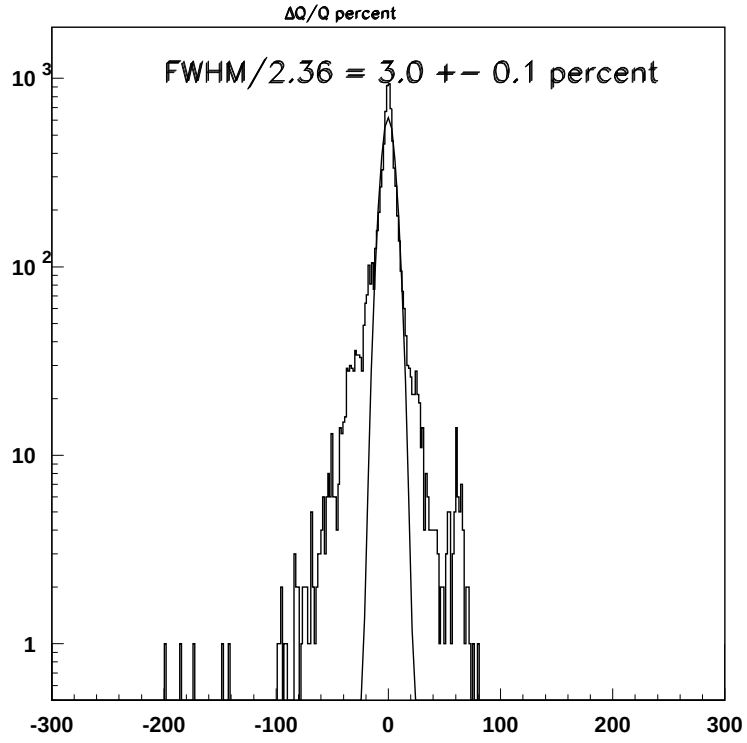

Figure 12: Distribution of errors for all norms $\|\mathbf{x}\|$. Gaussian fit superimposed, showing large nongaussian tails.

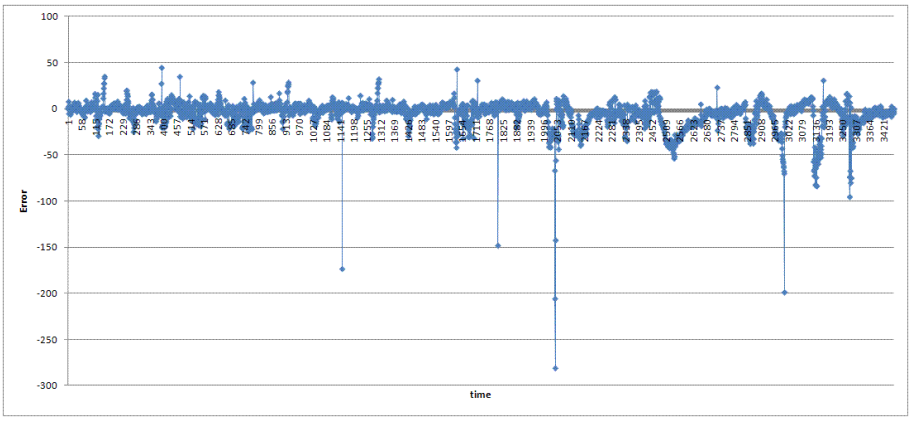

Figure 13: Period 1 training, prediction on periods 2 and 3.

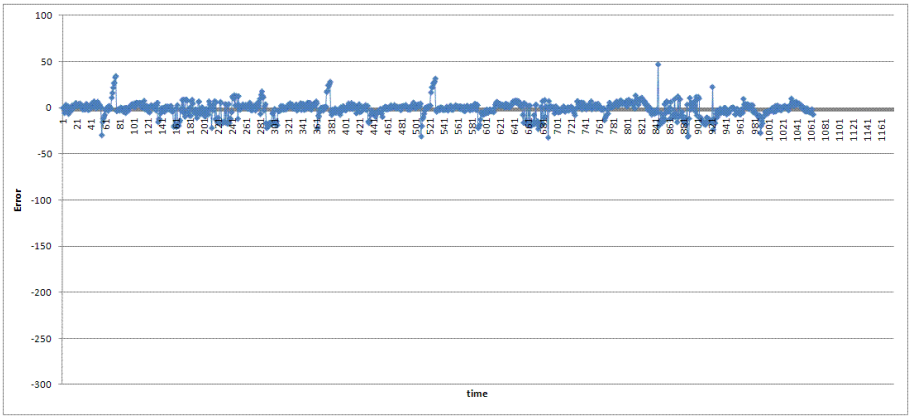

Figure 14: Period 1 training, prediction on periods 2 and 3, fiducial selection.

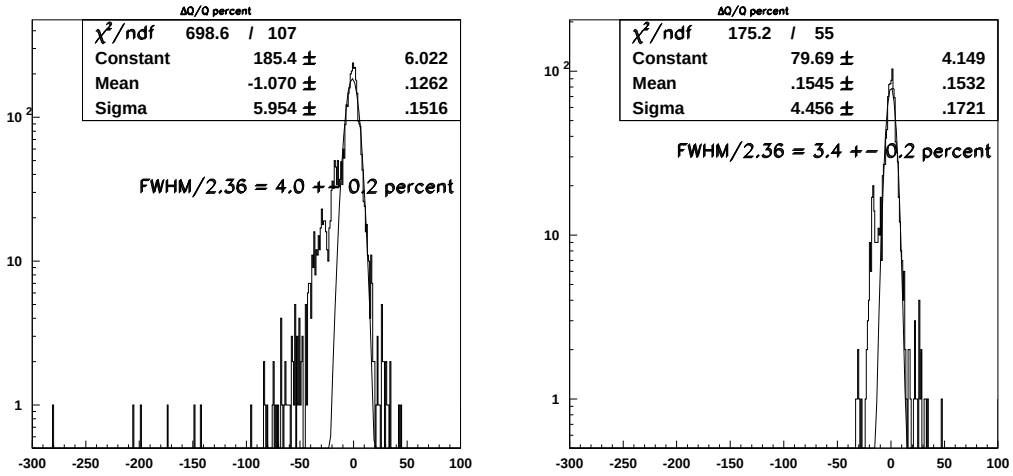

Figure 15: Period 1 training, prediction on periods 2 and 3, (left) no cuts, (right) fiducial cut, gaussian fit superimposed.

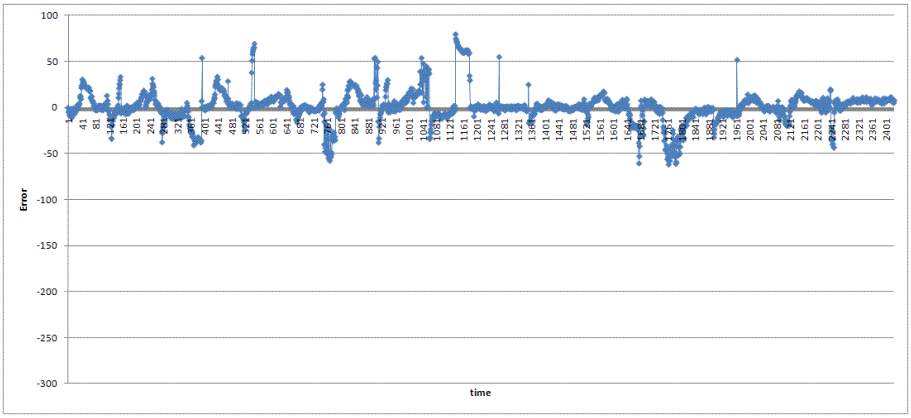

Figure 16: Period 2 training, prediction on periods 3 and 1.

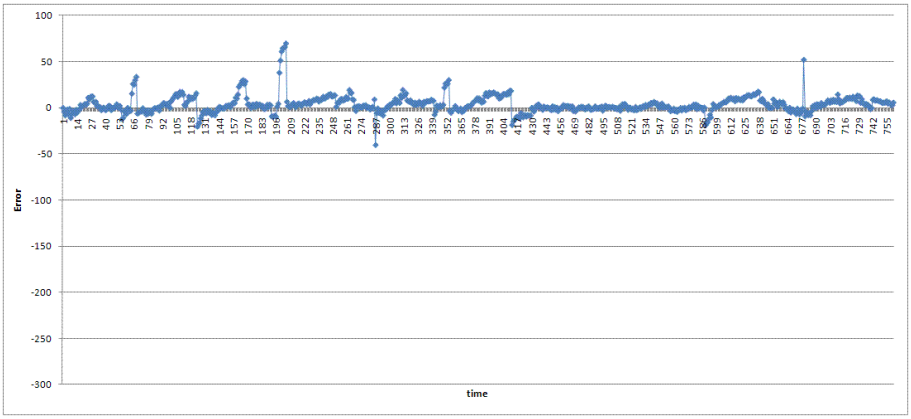

Figure 17: Period 2 training, prediction on periods 3 and 1, fiducial selection.

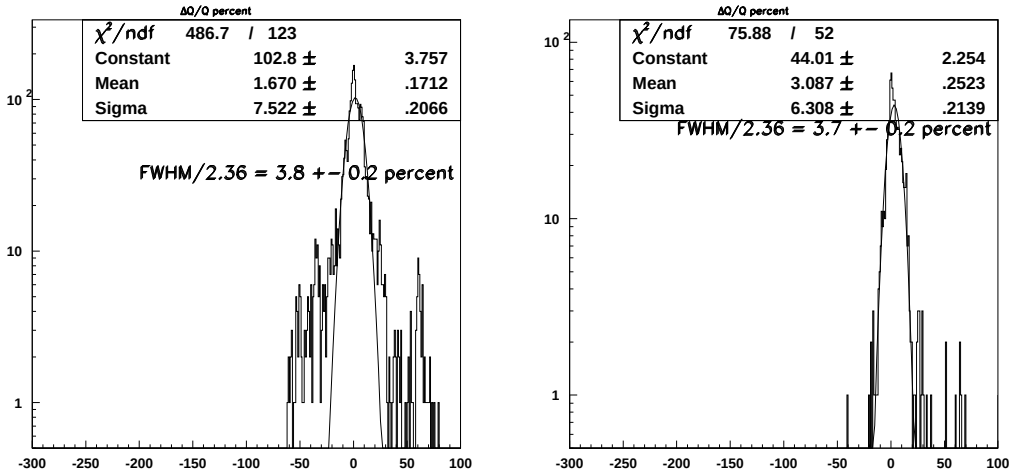

Figure 18: Period 2 training, prediction on periods 3 and 1, (left) no cuts, (right) fiducial cut, gaussian fit superimposed.

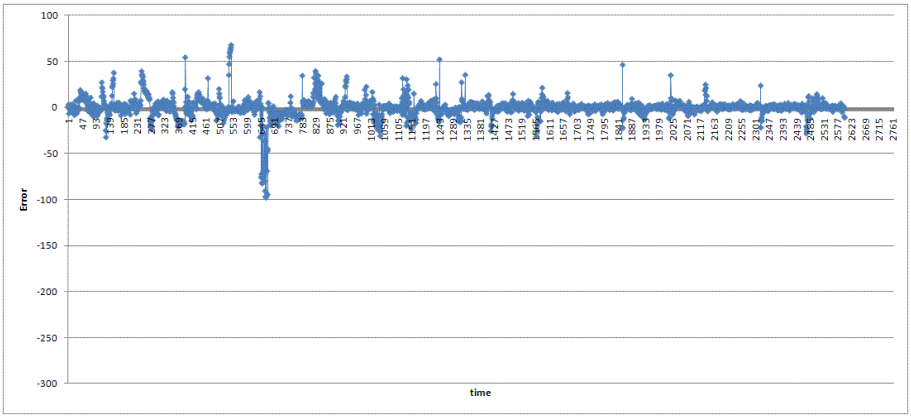

Figure 19: Period 3 training, prediction on periods 1 and 2.

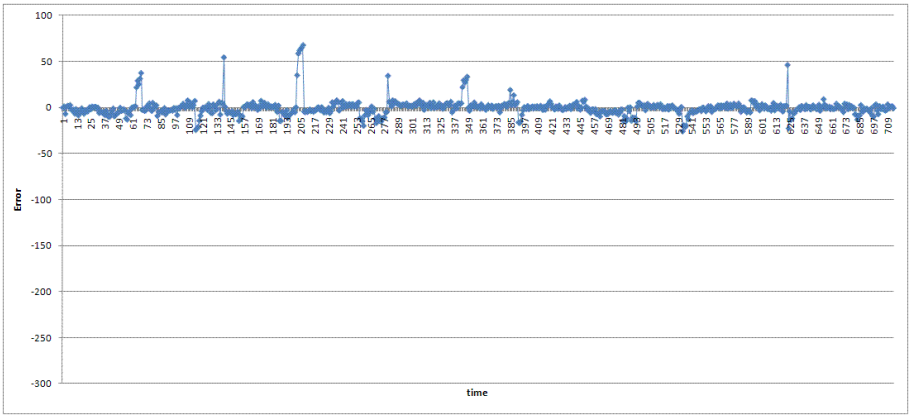

Figure 20: Period 3 training, prediction on periods 1 and 2, fiducial selection.

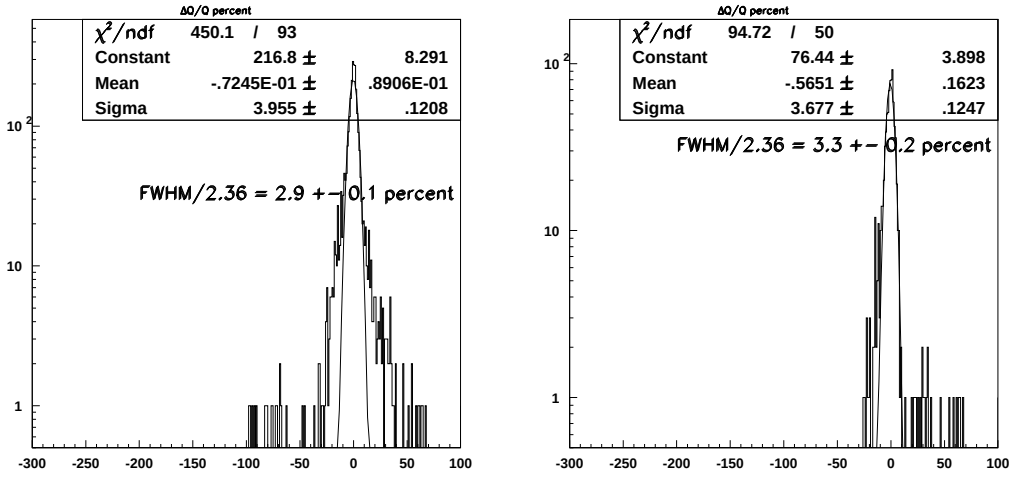

Figure 21: Period 3, prediction on periods 1 and 2, (left) no cuts, (right) fiducial cut, gaussian fit superimposed.

Table 4: Summary of errors σ_{fwhm} and nongaussian tails for various selection cuts and samples.

Data sets	$\frac{\Delta Q}{Q}$
All six chambers, all four periods training	2.7 ± 0.1
All six chambers, all four periods prediction	6.7 ± 0.1
Chamber six and period four excluded prediction	3.0 ± 0.1
Prediction on periods 2 and 3, training on period 1	4.0 ± 0.2
Prediction on periods 3 and 1, training on period 2	3.4 ± 0.2
Prediction on periods 1 and 2, training on period 3	3.8 ± 0.2
Prediction on periods 2 and 3, training on period 1, fiducial cuts	3.7 ± 0.3
Prediction on periods 3 and 1, training on period 2, fiducial cuts	2.9 ± 0.1
Prediction on periods 1 and 2, training on period 3, fiducial cuts	3.3 ± 0.2

described. The model, once trained on the response of a detector well within the parameter space (p, T, H), is able to predict the response in other periods with a better than $\sigma_{fwhm} \sim 10\%$ accuracy. Further studies are in progress to determine and cure the residual nongaussian tails of the $\frac{\Delta Q}{Q}$ errors distributions, to deal with training and prediction on detectors with different high voltage supply, to widen the sample of environmental conditions, and in adding new dimensions to the parameter space such as radiation levels.

7 Acknowledgements

The skills of M. Giardoni, L. Passamonti, D. Pierluigi, B. Ponzio, A. Russo (Frascati) in setting up the experimental setup are gratefully acknowledged. The technical support of the CERN gas group is gratefully acknowledged. Thanks are due to R. Guida (CERN Gas Group), Nadeesha M. Wickramage, Yasser Assran for discussions and help. This research was supported in part by the Italian Istituto Nazionale di Fisica Nucleare and Ministero dell' Istruzione, Università e Ricerca.

References

[1] R. Santonico and R. Cardarelli, "Development Of Resistive Plate Counters," Nucl. Instrum. Meth. **187** (1981) 377.

[2] R. Adolphi *et al.* [CMS Collaboration], "The CMS experiment at the CERN LHC", JINST **3** (2008) S08004.

[3] CMS Collaboration, The CMS muon project: Technical Design Report, CERN-LHCC-97-032 ; CMS-TDR-003. Geneva, CERN, 1997.

[4] J. Va'vra, "Physics and chemistry of aging: Early developments," ICFA Instrum. Bull. **24**, 1 (2002) [Nucl. Instrum. Meth. A **515**, 1 (2003)].

[5] J. Va'vra, "Summary of session 6: Aging effects in RPC detectors," Nucl. Instrum. Meth. A **515**, 354 (2003).

[6] K. Doroud et al., Nucl. Instrum. and Meth. A **602** (2009) 723-726. M. De Vincenzi et al., Nucl. Instrum. and Meth. A **508** (2003) 94-97. M. Bianco et al., Nucl. Instrum. and Meth. A **602** (2009) 700-704.

[7] M. Abbrescia *et al.*, "Gas analysis and monitoring systems for the RPC detector of CMS at LHC," LNF-06-34-P, LNF-04-25-P, Jan 2007. 9pp. Presented by S. Bianco on behalf of the CMS RPC Collaboration at the 2006 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and 15th International Room Temperature Semiconductor Detector Workshop, San Diego, California, 29 Oct - 4 Nov 2006. arXiv:physics/0701014.

[8] L. Benussi *et al.*, "The CMS RPC gas gain monitoring system: an overview and preliminary results," Nucl. Instrum. Meth. A **602** (2009) 805 [arXiv:0812.1108 [physics.ins-det]].

[9] L. Benussi *et al.*, JINST **4** (2009) P08006 [arXiv:0812.1710 [physics.ins-det]].

[10] W. S. Mc Culloch, W. Pitts, " A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics **5** (1943) 115-133.

[11] K. Hornik, M. Stinchcombe and H. White, "Multilayer Feedforward Networks are Universal Approximators ?", Neural Networks, vol. 2, pp. 359- 366, 1989.