Atmos. Meas. Tech. Discuss., 7, C1721–C1734, 2014
www.atmos-meas-tech-discuss.net/7/C1721/2014/


# *Interactive comment on* "A neural network approach for the simultaneous retrieval of volcanic ash parameters and SO<sub>2</sub> using MODIS data" *by* A. Piscini et al.

**A. Piscini et al.**

alessandro.piscini@ingv.it

Received and published: 15 July 2014

REVIEW: "A neural network approach for the simultaneous retrieval of volcanic ash parameters

and SO2 using MODIS data", by A. Piscini, M. Picchiani, M. Chini, S. Corradini, L. Merucci, F.

Del Frate, and S. Stramondo.

Roger Denlinger (Referee) Received and published: 29 April 2014



General Comments:

This paper presents a back-propagation method on a feed-forward multi-layer perceptron neural network, using data from MODIS satellite measurements to constrain mass of volcanic gas and ash in eruption-derived ash clouds produced in May 2010 by Eyjafjallajokull volcano in Iceland. The ash clouds were measured and simulated over sea surface alone and over clouds overlying a sea surface, and the results between the two situations are compared and analyzed. The results are very good, but the method raises questions regarding over-fitting and the use of prior distributions. Both concerns are related to the complexity of the neural network.

Specific comments:

1) Over-fitting:

a) The back propagation algorithm in complex neural networks is notorious for over-fitting (an output highly tuned to a particular set of inputs), producing overly-confident results. That is why procedures such as cross-validation (which are used in the method the authors present here) were developed. As I am concerned about this, I spent some time analyzing the authors work and previous publications.

What concerns me most is the choice of data sets to train the network and the choice of data to test it.

The network is trained using brightness temperature differences, or BTD, for ash mass, effective radius, and aerosol optical depth, and training the network to estimate sulfur dioxide using the differences between simulated and observed radiances in the band around 8.7 microns. The test data set, or validation set, is bracketed by the data sets used to calibrate the model in time. As a consequence, this is an in-data-set or (IS) model versus an out-of-data-set or (OOS) model. Models tuned and tested with IS data do not test the veracity of a back-propagation algorithm, even in this method where care has been taken to limit over-fitting. The results appear to be very good, but are only compelling to me because the same method was used by Picchiani et al, 2011. What I would have liked to see was the use of a data set outside the time frame used to train the model, rather than having the model trained on data sets that bracket the validation set in time, or tests of the model using independent data from other instruments. For example, CALIPSO data exist for May 8 2010 for cloud temperature and emissivity (Pavolonis et al., 2013). These results could be compared with the weighted estimates of cloud temperature and emissivity determined from their training algorithm using MODIS data alone. Such a comparison would be a better measure of the tendency towards overfitting with their algorithm.

Answer: We have considered MODIS acquisitions outside the timeframe used to train the networks, in order to better evaluate their generalization capabilities. A MODIS image of last phase of 2010 Eyja eruption (16 May 2010, 12:30 UTC) has been used. The results confirm that the network is suitable for retrieving the ash parameters and SO2 estimates also in OOS scenario (see figures from 1 to 4 representing scatterplot NN vs model for each parameters distinguishing also between "above sea" and "above cloud" cases and containing main statistics for each case). As already clarified, the aim of this work is to verify that the neural network approach is able to replicate the 'standard' volcanic ash and SO2 retrievals using MODIS data. The products validation, using CALIPSO measurements for example, is an extremely interesting and important topic but is out of the scope of this study.

Figures captions follow

Figure 1. Scatter plots for ash mass networks, 3 (top row), 28 (middle row) and 28 inputs pruned (bottom row), applied to the validation set of 16 May 2010, 12:30 UTC, divided in total pattern (left column), pattern over sea (middle column), patterns over meteorological cloud (right column).

Figure 2. Scatter plots for reff networks, 3 (top row), 28 (middle row) and 28 inputs pruned (bottom row), applied to the validation set of 16 May 2010, 12:30 UTC, divided



in total pattern (left column), pattern over sea (middle column), patterns over meteorological cloud (right column).

Figure 3. Scatter plots for AOD networks, 3 (top row), 28 (middle row) and 28 inputs pruned (bottom row), applied to the validation set of 16 May 2010, 12:30 UTC, divided in total pattern (left column), pattern over sea (middle column), patterns over meteorological cloud (right column).

Figure 4. Scatter plots for SO2 networks, 3 (top row), 28 (middle row) and 28 inputs pruned (bottom row), applied to the validation set of 16 May 2010, 12:30 UTC, divided in total pattern (left column), pattern over sea (middle column), patterns over meteorological cloud (right column).

I urge the authors to more completely test their model, as they did using the same approach for ash detection and mass retrieval in Picchiani et al., 2011, and to use other measurements of ash mass obtained from different equipment to test their method, as in Corradini et al., 2010. As reported here, their model is trained to minimize differences with BTD measurements of the same clouds observed just before and just after the validation data set, and can be expected to produce a very good fit to the clouds within this sequence.

Answer: The main object of this study is to replicate the results of inversion method developed by Corradini et al. (2009,2010) with an accurate and fast implementation based on NNs. Different to Picchiani et al. 2011, this goal is here addressed considered the full set of ash parameters and the SO2 estimation. The expected accuracy of the NN based implementation will not exceed the MODTRAN based approach. This explanation is better formulated through the text in order to better highlights the specific goals of the work. To this reason the results here shown have not been compared to the outcomes of other methods. Anyway we would thanks the reviewer for this specific suggestion, that would be considered in our future work.

b) In the optimal brain surgeon approach to obtaining a good balance between the data and the degrees of freedom in their model, they use a Hessian matrix. Is this Hessian determined from the output relative to the inner hidden layer output, or determined from the output of their neural network relative to the input to the neural network? The latter may be more appropriate since the results are used to prune the number of inputs.

Answer: The Optimal Brain surgeon implementation of SNNS simulator computes the Hessian matrix considering all the weights, so including also the input links (SNNS manual pg. 2018). Moreover we have adopted a further constrain to prune only the input weights in order to focus on the effects due to inputs selection instead of topology selection.

c) The scatter plots suggest that using only 3 channels, as Picchiani et al., 2011 did, would provide a better Occam factor than using all 28 channels. This would apply both to gas as well as to ash estimates. The Occam factor can be estimated from the curvature of the posterior distribution of weights output from the network, evaluated at the maximum likelihood. Defining this matrix as A, and Taylor-expanding the log posterior probability around the value of maximum likelihood for the weights, the posterior can be locally approximated as a Gaussian and with covariance matrix A-1. This approximation provides an additional criteria for pruning network inputs, and may provide a better comparison of results obtained with 3 channels versus 28 channels.

Answer: From the suggestion of the reviewer seems that the Occam factor of the 3 input network and of 28 could be computed in the case of multi-output neural network. After a trial and error phase, not reported in this work, we have chosen to split the several ash parameters in to different neural network. As a metric to evaluate the complexity vs. the generalization performance of the NN topologies we have adopted the saliency of the different weights, included the output one. Anyway if the reviewer can provides some further information to apply its suggested procedure we would to implement it.

2) The use of prior data:



This concern is secondary to over-fitting, but may still affect the outcomes if informative priors are used. Placing a prior constraint on one parameter, such as water vapor, may influence the determination of other parameters in complex ways that affect the learning. A single layer network avoids this since the prior distributions are simple and are directly used to calculate output, but in a multilevel network prior distributions are developed at each level. If biases are excluded and unconstrained, then the priors across multiple levels form improper (unable to normalize) distributions.

On page 10, line 1 it is stated that a 2 layer network with sigmoidal activation functions are independent from a priori assumptions. Unless the prior distributions are constructed properly, this is not strictly true. Sensitivity to prior distributions was a point raised by Picchiani et al., 2011.

The meaning of the sentence wanted to be different. Indeed a trained NN is certainly dependent by the distribution of the data used in the training phase. We would like just to clarify that the algorithms based on multilayer perceptron and back propagation can be trained by data with any distribution and it is not strictly necessary to know it before to train the network, since the training procedure is applied always in the same manner. We would like to thank the reviewer to have raised this point and to give us the chance to review the text accordingly.

Technical corrections: Answer: all technical correction have been made

Page 4:

Line 9: delete ‘in’ Answer: done

Line 10: replace ‘has’ with ‘have’ Answer: done

Line 12: replace ‘have been also’ with ‘were’ Answer: done

Line 12: insert ‘a scenario for’ between ‘to’ and ‘the detection’ Answer: done

Page 6:

Line 17: replace ‘to’ with ‘of’ Answer: done

Page 9:

Line 2: replace ‘insist’ with ‘exist’ Answer: done

Line 12: replace ‘the achieved’ with ‘can be achieved’ Answer: done

Line 24: delete ‘such as’ Answer: done

Line 25: replace ‘as a’ with ‘that an’ Answer: done

Page 10:

Line 2: replace ‘type’ with ‘types’ Answer: done

Line 6: replace ‘a MLP’ with ‘an MLP’ Answer: done

Line 6: replace ‘neuron’ with ‘a neuron’ Answer: done

Line 7: replace ‘input,’ with ‘input’ Answer: done

Line 13: delete ‘in the form’ Answer: done

Line 14: replace ‘Digital Numbers (DNs), e.g.’ with ‘data such as’ Answer: done Line 14: replace ‘temperatures,’ with ‘temperatures’ Answer: done

Line 15: replace ‘through a’ with ‘where the’ Answer: done

Line 15: replace ‘equals’ with ‘is equal’ Answer: done

Line 22: replace ‘vanish’ with ‘eliminate’ Answer: done

Line 29: delete ‘of’ Answer: done

Page 11:

Line 2: replace ‘aiming’ with ‘aimed’ Answer: done

Line 3: replace ‘are,’ with ‘are’ Answer: done



Line 5: delete ‘which are’ Answer: done

Line 7: delete ‘the’ Answer: done

Line 20: replace ‘is capable to give’ with ‘gives’ Answer: done

Line 21: delete ‘its’ Answer: done

Page 12:

Line 10: replace ‘the all’ with ‘all’ Answer: done

Line 11: insert ‘from the’ after ‘parameters’ Answer: done

Line 11: delete ‘have been used’ Answer: done

Line 16: delete ‘an error step’ Answer: done

Line 16: replace ‘The’ with ‘This’ Answer: done

Page 13:

Line 1: replace ‘has’ with ‘have’ Answer: done

Lines 11 and 23: replace ‘plume’ with ‘cloud’ Answer: done

Line 14: replace ‘parameters’ with ‘parameter’ Answer: done

Line 23: replace ‘sea and clouds’ with ‘sea and meteorological clouds’ Answer: done

Line 24: replace ‘were lying’ with ‘were located’ Answer: done

Page 14:

Line 21: delete ‘also’ Answer: done

Line 21: replace ‘for the two’ with ‘show’ Answer: done

Line 21: insert ‘that’ between ‘images’ and ‘have’ Answer: done

Line 22: replace ‘in’ with ‘into’ Answer: done

Page 15:

Line 6: replace ‘on’ with ‘at the’ Answer: done

Line 15: replace ‘allowing to train the network’ with ‘allowing the network to train’ Answer: done Line 28: replace ‘As regard to the’ with ‘With regard to’ Answer: done

Page 16:

Line 13: replace ‘of ’ with ‘for’ in both cases. Answer: done

Line 14: replace ‘retrieval’ with ‘retrieval,’ Answer: done

Line 14: replace ‘considered’ with ‘considered,’ Answer: done

Page 17:

Line 20: replace ‘plume and clouds’ with ‘the ash cloud and meteorological clouds’ Answer: done

Page 27:

Line 1: replace ‘divided’ with ‘in Table 2 and 3 by separately’ Answer: ?!?

Page 28:

Line 1: replace ‘for divided’ with ‘for parameters in Tables 2 and 3 divided’ Answer: ?!?

Line 1: replace ‘plume on’ with ‘ash clouds over’ Answer: done

Line 2: replace ‘plume on’ with ‘ash clouds over meteorological’ Answer: done

Page 34:

Line 4: replace ‘laying on sea’ with ‘clouds over sea alone’ Answer: done

Line 5: replace ‘on’ with ‘and ash clouds over’ Answer: done

Pages 34-47:



The axes on all figures are labeled with miniscule labels that are unreadable without significant

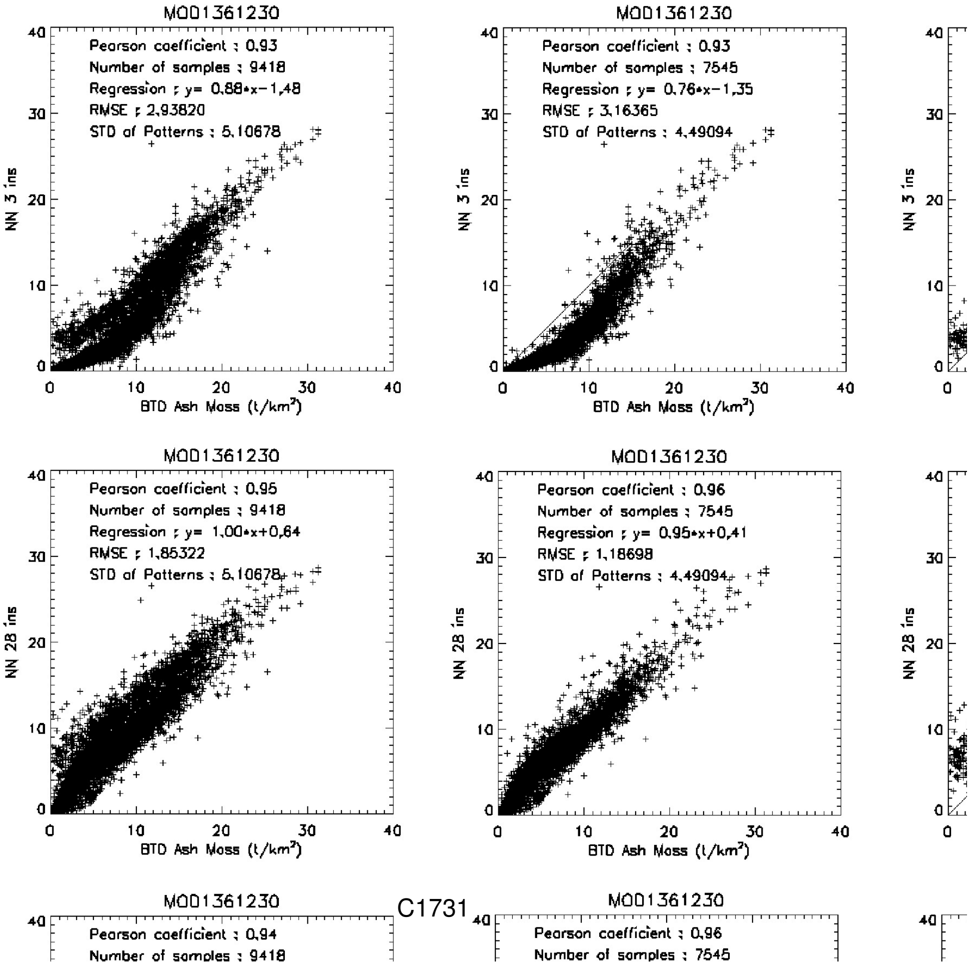
magnification. Please increase these font sizes to be readable. Answer: it will be done in the final manuscript revision

References Cited:

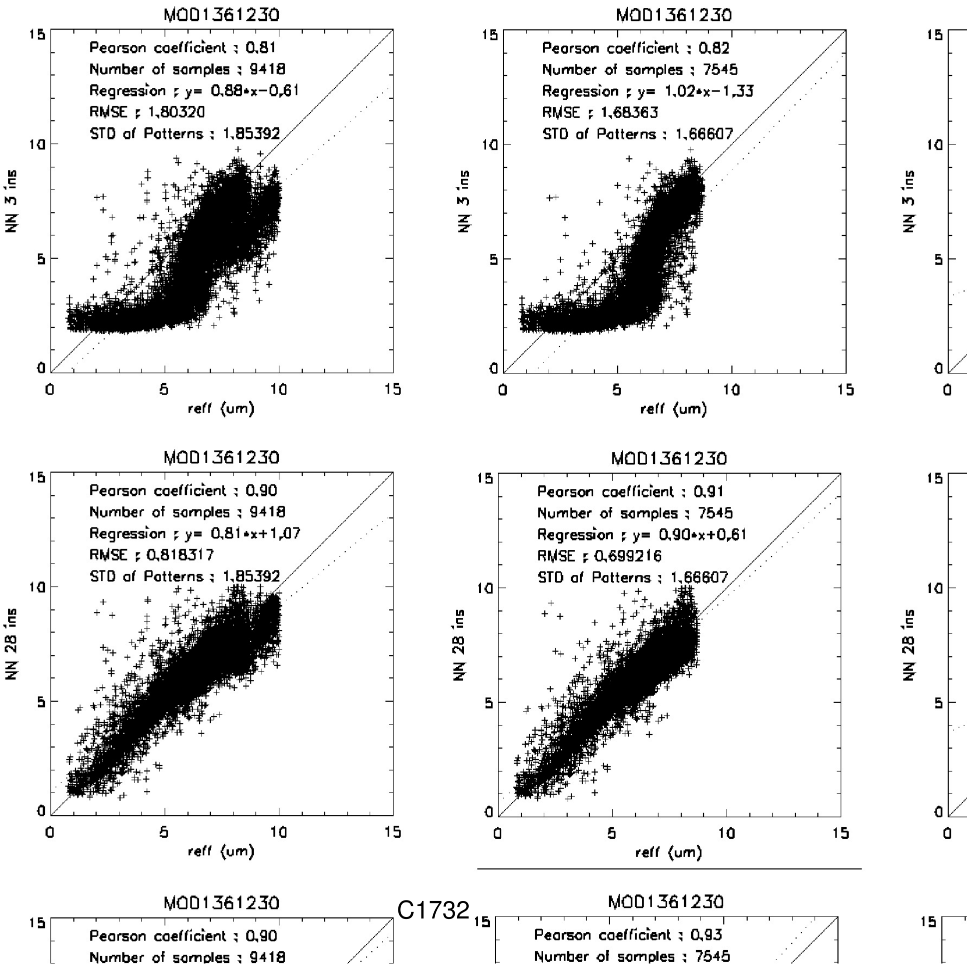
Corradini, S., L. Merucci, A.J. Prata, and A. Piscini, 2010, Volcanic ash and SO2 in the 2008 Kasatochi eruption: Retrievals comparison from different IR satellite sensors, JGR, doi: 10.1029/2009/JD013634.

Pavolonis, M.J., A.K. Heidinger, and Justin Sieglaff, Automated retrievals of volcanic ash and dust cloud properties from upwelling infrared measurements, JGR: Atmospheres, doi: 10.1002/jgrd.50173 Picchiani, M., M. Chini, S. Corradini, L. Merucci, P. Sellitto, F. Del Frate, and S. Stramondo, 2011, Volcanic ash detection and retrievals using MODIS data by means of neural networks, Atmos. Meas. Tech., doi: 10.5194/amt-4-2619-2011.

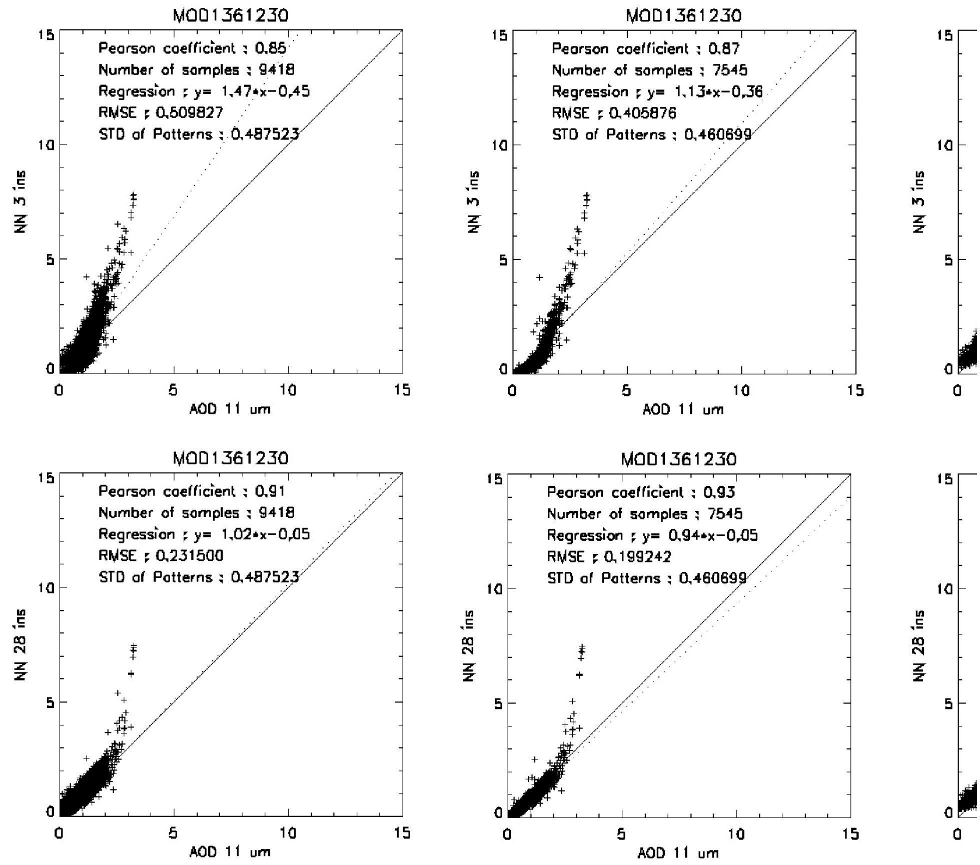
Interactive comment on Atmos. Meas. Tech. Discuss., 7, 3349, 2014.


MOD1361230
Pearson coefficient : 0.93
Number of samples : 9418
Regression : y= 0.88*x-1.48
RMSE : 2.93820
STD of Patterns : 5.10678
NN 3 ins
BTD Ash Mass (t/km$^2$)
MOD1361230
Pearson coefficient : 0.93
Number of samples : 7545
Regression : y= 0.76*x-1.35
RMSE : 3.16365
STD of Patterns : 4.49094
NN 3 ins
BTD Ash Mass (t/km$^2$)
NN 3 ins
MOD1361230
Pearson coefficient : 0.95
Number of samples : 9418
Regression : y= 1.00*x+0.64
RMSE : 1.85322
STD of Patterns : 5.10678
NN 28 ins
BTD Ash Mass (t/km$^2$)
MOD1361230
Pearson coefficient : 0.96
Number of samples : 7545
Regression : y= 0.95*x+0.41
RMSE : 1.18698
STD of Patterns : 4.49094
NN 28 ins
BTD Ash Mass (t/km$^2$)
NN 28 ins
MOD1361230
Pearson coefficient : 0.94
Number of samples : 9418
MOD1361230
Pearson coefficient : 0.96
Number of samples : 7545

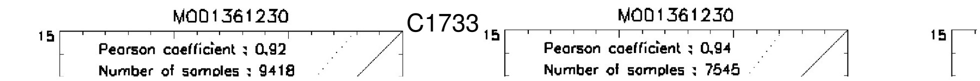
C1731

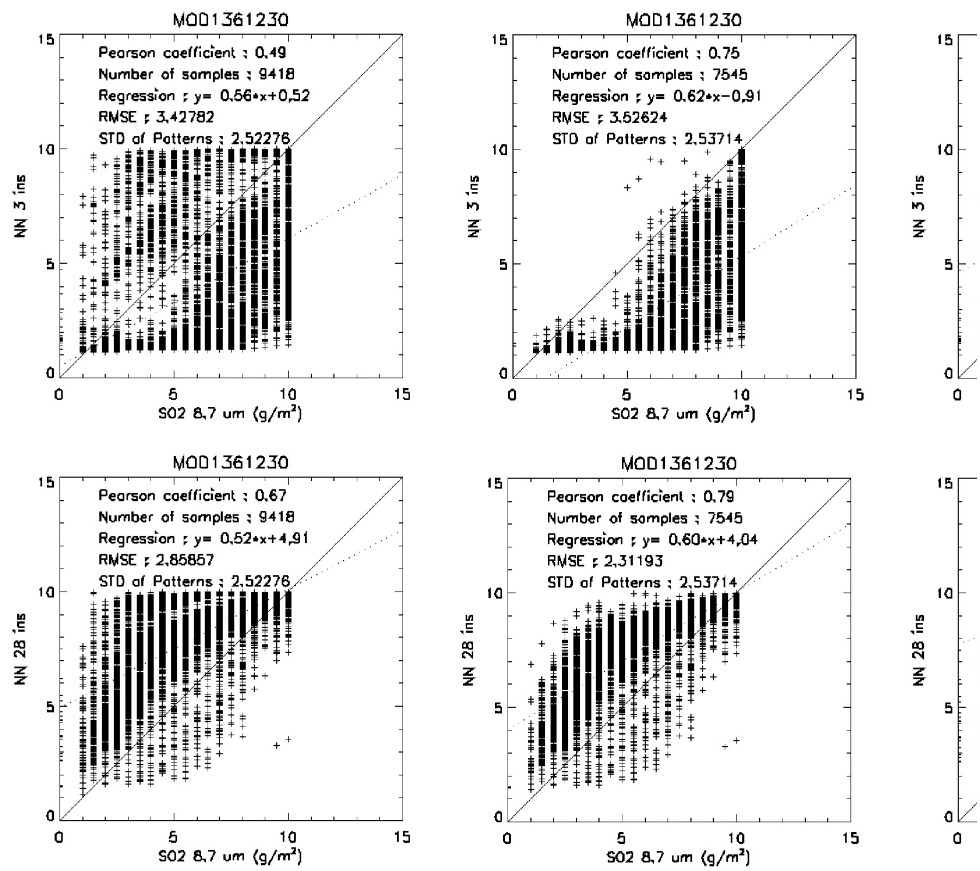

MOD1361230
Pearson coefficient : 0.81
Number of samples : 9418
Regression : y= 0.88*x-0.61
RMSE : 1.80320
STD of Patterns : 1.85392
NN 3 ins
reff (um)
MOD1361230
Pearson coefficient : 0.82
Number of samples : 7545
Regression : y= 1.02*x-1.33
RMSE : 1.68363
STD of Patterns : 1.66607
NN 3 ins
reff (um)
NN 3 ins
MOD1361230
Pearson coefficient : 0.90
Number of samples : 9418
Regression : y= 0.81*x+1.07
RMSE : 0.818317
STD of Patterns : 1.85392
NN 28 ins
reff (um)
MOD1361230
Pearson coefficient : 0.91
Number of samples : 7545
Regression : y= 0.90*x+0.61
RMSE : 0.699216
STD of Patterns : 1.66607
NN 28 ins
reff (um)
NN 28 ins
MOD1361230
Pearson coefficient : 0.90
Number of samples : 9418
MOD1361230
Pearson coefficient : 0.93
Number of samples : 7545

C1732

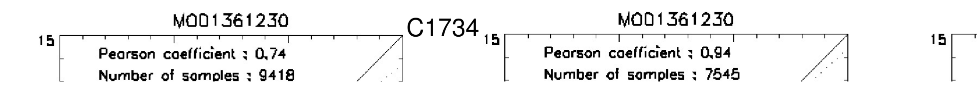

MOD1361230
Pearson coefficient : 0.85
Number of samples : 9418
Regression : y= 1.47*x-0.45
RMSE : 0.509827
STD of Patterns : 0.487523
NN 3 ins
AOD 11 um
MOD1361230
Pearson coefficient : 0.87
Number of samples : 7545
Regression : y= 1.13*x-0.36
RMSE : 0.405876
STD of Patterns : 0.460699
NN 3 ins
AOD 11 um
NN 3 ins
MOD1361230
Pearson coefficient : 0.91
Number of samples : 9418
Regression : y= 1.02*x-0.05
RMSE : 0.231500
STD of Patterns : 0.487523
NN 28 ins
AOD 11 um
MOD1361230
Pearson coefficient : 0.93
Number of samples : 7545
Regression : y= 0.94*x-0.05
RMSE : 0.199242
STD of Patterns : 0.460699
NN 28 ins
AOD 11 um
NN 28 ins

C1733


MOD1361230
Pearson coefficient : 0.92
Number of samples : 9418
MOD1361230
Pearson coefficient : 0.94
Number of samples : 7545


MOD1361230
Pearson coefficient : 0.49
Number of samples : 9418
Regression : y= 0.56*x+0.52
RMSE : 3.42782
STD of Patterns : 2.52276
NN 3 ins
SO2 8.7 um (g/m$^2$)
MOD1361230
Pearson coefficient : 0.75
Number of samples : 7545
Regression : y= 0.62*x-0.91
RMSE : 3.52624
STD of Patterns : 2.53714
NN 3 ins
SO2 8.7 um (g/m$^2$)
NN 3 ins
MOD1361230
Pearson coefficient : 0.67
Number of samples : 9418
Regression : y= 0.52*x+4.91
RMSE : 2.85857
STD of Patterns : 2.52276
NN 28 ins
SO2 8.7 um (g/m$^2$)
MOD1361230
Pearson coefficient : 0.79
Number of samples : 7545
Regression : y= 0.60*x+4.04
RMSE : 2.31193
STD of Patterns : 2.53714
NN 28 ins
SO2 8.7 um (g/m$^2$)
NN 28 ins

C1734


MOD1361230
Pearson coefficient : 0.74
Number of samples : 9418
MOD1361230
Pearson coefficient : 0.94
Number of samples : 7545