Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Neural network cloud top pressure and height for MODIS

Nina Håkansson[1], Claudia Adok[2], Anke Thoss[1], Ronald Scheirer[1], and Sara Hörnquist[1]

[1]Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden
[2]Regional Cancer Center Western Sweden, Gothenburg, Sweden

Correspondence: Nina Håkansson (nina.hakansson@smhi.se)

Abstract.

Cloud top height retrieval from imager instruments is important for Nowcasting and for satellite climate data records. A neural network approach for cloud top height retrieval from the imager instrument MODIS is presented. The neural networks are trained using cloud top layer pressure data from the CALIOP dataset.

Results are compared with two operational reference algorithms for cloud top height: the MODIS Collection 6 level 2 height product and the cloud top temperature and height algorithm (CTTH) in the 2014 version of the NWCSAF Polar Platform System (PPS-v2014). All three techniques are evaluated using both CALIOP and CPR (CloudSat) height.

Instruments like AVHRR and VIIRS contain fewer channels useful for cloud top height retrievals than MODIS, therefore several different neural networks are investigated to test how infrared channel selection influences retrieval performance.

Also a network with only channels available for the AVHRR1 instrument is trained and evaluated. To examine the contribution of different variables, networks with fewer variables are trained. It is shown that variables containing imager information for neighbouring pixels are very important.

Overall results for the neural network height retrievals are very promising. The neural networks using the brightness temperatures at 11μm and 12μm show at least 33% (or 627m) lower mean absolute error (MAE) compared to the two operational reference algorithms when validating with CALIOP height. Validation with CPR (CloudSat) height gives at least 25% (or 433m) reduction of MAE. For the network trained with a channel combination available for AVHRR1, the MAE is at least 542m better when validated with CALIOP and 414m when validated with CPR (CloudSat) compared to the two operational reference algorithms. The NWCSAF PPS-2018 release will contain a neural network based cloud height algorithm.

1 Introduction

The retrieval of cloud top temperature, pressure and height from imager data from polar orbiting satellites is used both as a vital product in global cloud climatologies (Stubenrauch et al., 2013) and for nowcasting at high latitudes where data from geostationary satellites are either not available or not available in sufficient quality and spatial resolution. Cloud top height products from VIS/IR imagers are used in the analysis and early warning of thunderstorm development and for height assignment in aviation forecasts. The cloud height can serve as input to mesoscale analysis and models for use in nowcasting in general, or as input to other satellite retrievals used in nowcasting (e.g. cloud micro physical properties retrieval, or cloud type retrieval). It

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

is important that climatologists and forecasters have reliable and accurate cloud height products from recent and past satellite measurements.

Several algorithms to retrieve cloud top height from polar orbiting satellites are available and used operationally for nowcasting purposes or in cloud climatologies. These include the CTTH (cloud top temperature and height) from the PPS (Polar Platform System) package (Dybbroe et al., 2005), which is also used in the CLARA-A2 climate data record of CMSAF (EUMETSAT Satellite Application Facility for Climate Monitoring) (Karlsson et al., 2017), ACHA (cloud height algorithm) used in PATMOS-x (Heidinger et al., 2014), CC4CL used in ESA Cloud_CCI (Stengel et al., 2017), MODIS Collection-6 algorithm (Ackerman et al., 2015) and the ISCCP algorithm (Rossow and Schiffer, 1999).

We will use both the MODIS Collection-6 (MODIS-C6) and the version 2014 CTTH from PPS (PPS-v2014) as references to evaluate the performance of neural network based cloud height retrieval. The MODIS-C6 algorithm is developed for the MODIS instrument. The PPS, delivered by the NWC SAF (EUMETSAT Satellite Application Facility for Nowcasting and very shortrange forecasting), is adapted to handle data from instruments AVHRR, VIIRS and MODIS.

Artificial neural networks are widely used for non-linear regression problems, see for example Gardner and Dorling (1998), Meng et al. (2007) or Milstein and Blackwell (2016) for neural network applications in atmospheric science. In CC4CL a neural network is used for the cloud detection (Stengel et al., 2017). Artificial neural networks have also been used on MODIS data to retrieve cloud optical depth (Minnis et al., 2016). The COCS algorithm uses artificial neural networks to retrieve cirrus cloud optical thickness and cloud top height for the SEVIRI instrument (Kox et al., 2014).

One type of neural network is the multilayer perceptron described in (Gardner and Dorling, 1998) which is a supervised learning technique. If the output for a certain input, when training the multilayer perceptron, is not equal to the target output an error signal is propagated back in the network and the weights of the network are adjusted resulting in a reduced overall error. This algorithm is called the back-propagation algorithm.

In this study we will compare the performance of back-propagation neural network algorithms for retrieving cloud top height (NN-CTTH) with the CTTH algorihm from PPS version 2014 (PPS-v2014) and MODIS Collection 6 (MODIS-C6) algorithm. Several networks will be trained to estimate the contribution of different training variables to the overall result. The networks will be validated using both CALIOP and CloudSat height data.

In section 2 the different datasets used are briefly described and in section 3 the three algorithms are described. Results are presented in section 4, discussed in section 5 and final conclusions are found in section 6.

2 Instruments and data

For this study we used data from the MODIS instrument on the polar orbiting satellite Aqua in the A-Train, as it is co-located with both CALIPSO and CloudSat at most latitudes and has multiple channels useful for cloud top height retrieval.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

2.1 Aqua - MODIS

The MODIS (Moderate-resolution Imaging Spectro-radiometer) is a spectro-radiometer with 36 channels covering the solar and thermal spectra. We are using level 1 data from the MODIS instrument on the polar orbiter Aqua. For this study the MYD021km (MODIS Science Data Support Team, 2015a) and MYD03 (MODIS Science Data Support Team, 2015b) for all orbits from 24 dates were used (1^{st} and 14^{th} of every month of 2010). The data were divided into four parts which were used for training, validation during training (used to decide when to quit training), testing under development (used to test different combinations of variables during prototyping) and final validation. See Table 1 for distribution of data.

2.1.1 Aqua - MODIS Collection 6 cloud products

The MODIS Collection-6 climate data records produced by the National Aeronautics and Space Administration (NASA) Earth Observation System are using data from the MODIS sensor. The MODIS Collection 6 cloud products were used as an independent algorithm with which to compare the performance of the NN-CTTH. The 1-km cloud top height and cloud top pressure from the MYD06_L2-product (Ackerman et al., 2015) for the dates in Table 1 were used.

2.2 CALIPSO - CALIOP

The CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarisation) lidar on the polar orbiting satellite CALIPSO is an active sensor and therefore more sensitive to particle conglomerates with low density than typical imagers. The horizontal pixel resolution is 0.07km x 0.333km, this means that when co-locating with MODIS one should remember that CALIOP samples only a small part of each MODIS pixel. The CALIOP 1km Cloud Layer product (version 3) data were used (for the dates, see Table 1) as the truth to train the networks against, and for validation of the networks.

2.3 CloudSat - CPR

The CPR (Cloud Profiling Radar for CloudSat) is a radar on CloudSat which derives a vertical profile of cloud water. Its horizontal resolution is 1.4km x 3.5km, and its vertical resolution is 0.5km. The CPR product 2B-GEOPROF-R04 (Marchand et al., 2008) was used as an additional source for independent validation of the networks. See Table 1 for selected dates. The validation with CloudSat will have a lower percentage of low clouds compared to CALIOP because ground clutter is a problem for space bourne radar instruments.

2.4 Other data

Numerical weather prediction (NWP) data are needed as input for the PPS-v2014 and the Neural network algorithm. In this study NWP data from ECMWF (European Centre for Medium-range Weather Forecasting) were used. Also ice maps from OSISAF (Satellite Application Facility on Ocean and Sea Ice) were used as input for the PPS cloud mask algorithm.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

3 Algorithms

3.1 PPS-v2014 cloud top temperature and height

The cloud top height algorithm in PPS-v2014, uses two different algorithms for cloud height retrieval, one for pixels classified as opaque and another for semi-transparent clouds. The reason for having two different algorithms is that the straight forward opaque algorithm can not be used for pixels with optically thin clouds like cirrus or broken cloud fields like cumulus. The signals for these pixels are a mixture of contributions from the cloud itself and underlying clouds and/or the surface.

The retrieval for opaque clouds matches the observed brightness temperatures at 11µm against a temperature profile derived from a short term forecast or (re)analysis of a NWP model, adjusted for atmospheric absorption. The first match, going along the profile from the ground and upwards, gives the cloud top height and pressure. Temperatures colder or warmer than the profile are fitted to, respectively, the coldest or warmest temperature of the profile below tropopause.

The algorithm uses a split-window technique to decide whether to apply the opaque or semi-transparent retrieval. All pixels with a difference between the 11µm and 12µm brightness temperatures of more than 1.0 K are treated as semi-transparent. This is a slight modification of the PPS version 2014 algorithm where also the clouds classified as non-opaque by cloud type product are considered semi-transparent.

The algorithm for semi-transparent pixels uses a histogram method, based on the work of Inoue (1985) and Derrien et al. (1988), which fits a curve to the brightness temperature difference between the 11µm and 12µm bands as a function of 11µm brightness temperatures for all pixels in a segment (32x32 pixels). One parameter of this fitting is the cloud top temperature. The solution is checked for quality (low root mean square error) and sanity (inside physically meaningful interval and not predicted too far from data). The solution is accepted if both tests are passed. The height and pressure are then retrieved from the temperature, in the same way as for opaque clouds. For more detail about the algorithms see SMHI (2015).

PPS height uses the unit altitude above ground. For all comparisons this is transformed to height above mean sea level, using elevations given in the CloudSat or CALIOP datasets.

3.2 MODIS Collection 6 Aqua Cloud Top Properties product

In MODIS Collection 6 the CO_2-slicing method (described in Menzel et al., 2008) is used to retrieve cloud top pressure using the 15µm channels for ice clouds (as determined from MODIS phase algorithm). For low level clouds the 11µm channel and the IR-window approach (IRW) with a latitude dependent lapse rate is used over ocean (Baum et al., 2012). Over land the 11µm temperature is fitted against a 11µm temperature profile calculated from GDAS temperature, water vapour and ozone profiles and PFAAST radiative transfer model are used for low clouds (Menzel et al., 2008). For more details about the updates in Collection 6 see Baum et al. (2012). Cloud pressure is converted to temperature and height using the National Centers for Environmental Prediction Global Data Assimilation System (Baum et al., 2012).

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

3.3 Neural network cloud top temperature and height NN-CTTH

Neural networks are trained using MODIS data co-located with CALIOP data. The Aqua and CALIPSO satellites are both part of the A-Train and the matched FOV are close in time (only 75s apart). The top layer pressure variable from CALIOP data was used as training truth. Temperature and height for the retrieved cloud top pressure are extracted using NWP-data. Pressure predicted higher than surface pressure are set to surface pressure. For pressures lower than 70hPa neither height nor temperature values are extracted.

3.3.1 Neural network variables

To reduce sun-zenith angle dependence and to have the same algorithm for all illumination conditions it was decided to use only infra-red channels to train the neural networks. Several different types of variables were used to train the network. The most basic ones were the NWP temperatures at pressure levels (surface, 950, 850, 700, 500 and 250hPa). This together with the 11µm or 12µm brightness temperature gives the network much of what is needed to predict cloud top pressure for opaque clouds. For opaque clouds that are geometrically thin, with little or no water vapour above the cloud, the 11µm and 12µm brightness temperatures will be the same as the cloud top temperature. If the predicted NWP temperatures are correct the neural network could fit the 11µm brightness temperature to the NWP temperatures and receive the cloud pressure (similar to what is done in PPS-v2014 and MODIS-C6). For cases without inversions in the temperature profile, the retrieved cloud top pressure should be accurate. The cases with inversions are more difficult to fit correctly, since multiple solutions exist and the temperature inversion might not be accurately captured regarding its strength and height in the NWP data. For semi-transparent clouds the network needs more variables to make a correct retrieval.

To give the network information on opacity of the pixel, brightness temperature difference variables were included ($B_{11}-B_{12}, B_{11}-B_{3.7}, B_{8.5}-B_{11}$). Texture variables with the standard deviation of brightness temperature, or brightness temperature difference, for 5x5 pixels were included. These contain information about whether pixels with large $B_{11}-B_{12}$ are more likely to be semi-transparent or more likely to be fractional or cloud edges.

As described in section 3.1, PPS-v2014 uses $B_{11}-B_{12}$ and B_{11} for neighbouring pixels to retrieve temperatures for semi-transparent clouds. In order to feed the network with some of this information the neighbouring warmest and coldest pixels (in brightness temperature 11µm in a 5x5 pixel neighbourhood) were identified. Variables using the brightness temperature at these warmest and coldest pixels were calculated, for example the 12µm brightness temperature for the coldest pixel minus the same for the current pixel: $B_{12}^{C}-B_{12}$, see Table 2 for more information about what variables were calculated.

The surface pressure was also included, which provides the network with a value for the maximum reasonable pressure. Also the brightness temperature for the CO_2 channel at 13.3µm and the water vapour channels at 6.7µm and 7.3µm were included as variables. The CO_2 channel at 13.3µm is used in the CO_2-slicing method of MODIS-C6 and should improve the cloud height retrieval for high clouds.

The instruments AVHRR, VIIRS, MERSI-2, MetImage and MODIS all have different selections of IR channels. Most of them have the 11µm and 12µm channels. The first AVHRR instrument AVHRR1 had only two IR channels at 11µm and 3.7µm

and no channel at 12μm. Networks were trained using combinations of MODIS IR-channels corresponding to the channels available for the other instruments. See Table 3 for specifications of the networks trained. Table 4 gives an overview of what imager channels were used for which network.

To see how much the different variable types contribute to the result, some basic networks were trained using less or no imager data. These are also described in Table 3. Also one network using only NWP data was included as a sanity check. For this network we expect bad results. However good results for this network would indicate that height information retrieved was already available in the NWP-data.

3.3.2 Training

For the training 1.5 million pixels were used, with the distribution 50% low clouds, 25% medium level clouds and 25% high clouds. A higher percentage of low clouds was included because the mean square error (MSE) is often much higher for high clouds. Previous tests showed that less low clouds caused the network to focus too much on predicting the high clouds correctly and showed degraded results for low clouds. For training validation 375000 pixels were randomly selected with the same low/medium/high distribution as for the training data.

The machine learning module Scikit-learn (Pedregosa et al., 2011), the Keras package (Chollet et al., 2015), the Theano (Theano Development Team, 2016) backend and the language Python were used for training the network.

3.3.3 Parameters and configurations

During training of the network the MSE was used as the loss function that is minimized during training. The data were standardized by subtracting the mean and dividing with the standard deviation before training.

Choosing the number of hidden neurons and hidden layers for the neural network is also important for the training to be effective. Too few hidden neurons will result in under-fitting. We used two hidden layers with 30 neurons in the first layer and 15 neurons in the second.

The initialization of weights before training the network is important for the neural network to learn faster. There are many different weight initialization methods, for training the networks the glorot uniform weight initialization was used.

The activation function used for the hidden layers was the tangent hyperbolic (see Karlik and Olgac, 2011) and for the output layer a linear activation function was used.

To determine the changes in the weights an optimization method is used during the back-propagation algorithm. The optimization method used for the multilayer perceptron is mini-batch stochastic gradient descent which performs mini-batch training. A mini-batch is a sample of observations in the data. Several observations are used to update weights and biases, which is different from the traditional stochastic gradient descent where one observation at a time is used for the updates (Cotter et al., 2011). Having an optimal mini-batch size is important for the training of a neural network because overly large batches can cause the network to take a long time to converge. We used a mini-batch size of 250.

When training the neural network there are different learning parameters that need to be tuned to ensure an effective trainig procedure. During prototyping several different combinations were tested. The learning rate is a parameter that determines the

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

size of change in the weights. A too large learning rate will result in large weight changes and can result in an unstable model (Hu and Weng, 2009). If a learning rate on the other hand is too small the training time of the network will be long. We used a learning rate of 0.01.

The momentum is a parameter which adds a part of the weight change to the current weight change, using momentum can help avoid the network getting trapped in local minima (Gardner and Dorling, 1998). A high value of momentum speeds up the training of the network. We had a momentum of 0.9. The parameter *learning rate decay*, set to 10^{-6}, in Keras, is used to decrease the learning rate after each update as the training progresses.

To avoid the neural network from over-fitting (which makes the network extra sensitive to unseen data), a method called *early stopping* was used. In early stopping the validation error is monitored during training to prevent the network from over-fitting. If the validation error is not improved for some (we used 10) epochs training is stopped; this helps to reduce risk of over-fitting. The network for which the validation error was at its lowest is then used. The neural networks were trained for a maximum of 2650 epochs, but the early stopping method caused the training to stop much earlier.

4 Results

First the performance of all the trained networks were validated with CALIOP in terms of mean absolute error (MAE). Results in Table 5 show that both PPS-v2014 and MODIS-C6 have a MAE close to 120hPA. Even the NN-OPAQUE network using only B_{12} and the basic NWP-data has a 10hPa improvement in MAE. By including the variable $B_{11}-B_{12}$, the MAE improves by an additional 19hPa because $B_{11}-B_{12}$ contains information about the semi-transparency of the pixel. Adding the NWP variable Ciwv, which allows the network to attempt to predict the expected values of $B_{11}-B_{12}$, has a smaller effect of 2hPa on MAE. However adding all variables containing information on neighboring pixels improves the result by additional 20hPa. The NN-AVHRR network using 11µm and 12µm from MODIS provides an MAE which is reduced by about 50hPA compared to both from MODIS-C6 and PPS-v2014. Notice also that the scores improve for all categories (low, medium and high) when compared with both PPS-v2014 and MODIS-C6.

Adding more IR channels improves the results further. Adding channel 8.5µm ($B_{8.5}-B_{11}$, NN-VIIRS) improves MAE by 7hPa and adding 7.3µm ($B_{7.3}$, NN-MERSI-2) improves MAE by 5hPa. Including the other watervapor channel at 6.7µm ($B_{6.7}$, NN-MetImage-$NoCO_2$) improves MAE only by 1hPa. The CO_2 channel at 13.3µm ($B_{13.3}$, NN-MetImage) improves the MAE by an additional 6hPa.

The NN-AVHRR1 network trained using 3.7µm and 11µm (MAE 76.1hPa) is a little worse compared to NN-AVHRR (MAE 72.2hPa). Also NN-AVHRR1 shows better scores for all categories (low, medium, and high) compared to PPS-v2014 and MODIS-C6. Notice that the network using only the NWP information and no imager channels (NN-NWP) shows high MAE. This was included as a sanity check to see that the predicted height is using mainly the satellite data, and the high MAE for NN-NWP is supporting this.

The training with CALIOP using only MODIS from Aqua includes only near NADIR observations. Figure 1 shows that NN-AVHRR and NN-AVHRR1 networks perform robustly also for higher satellite zenith angles. The N-VIIRS and NN-

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

MetImage-$NoCO_2$ results deviate for satellite zenith angles larger than 60 degrees. The NN-MERSI-2 results deviate for satellite zenith angles larger than 40 degrees. The NN-MetImage retrieval shows deviations already above 20 degree satellite zenith angles and for satellite zenith angles larger than 40 the retrieval has no predictive skill. Notice that the distribution for MODIS-C6 also depend on the satellite zenith angle (with less high clouds at higher angles). For PPS-v2014 instead there are less low clouds at higher satellite zenith angles. PPS-v2014 also has the least reasonable pressure distribution with one peak for mid-level clouds instead of capturing the two peaks for low and high clouds observed in the CALIOP data.

Results for MAE in meters compared to CALIOP top layer height are provided for the best performing networks in Table 6. Notice that the MAE for high clouds is 1.5km better for the NN-AVHRR than for PPS-v2014. Compared to MODIS-C6 the NN-AVHRR is 0.8km better for both medium level clouds and high clouds. The NN-AVHRR MAE is 627m lower (corresponding to 33% reduction of MAE) than MODIS-C6 and 797m (corresponding to 38% reduction of MAE) lower than PPS-v2014. The NN-MetImage-$NoCO_2$ has the best result while performing well at all satellite zenith angles, with a 44% reduction in MAE when compared to MODIS-C6 and a 48% reduction when compared to PPS-v2014.

Co-located comparisons with CloudSat were also perfomed. In Table 7 the MAE in meters compared to height from CPR (CloudSat) are presented. All neural networks show at least 400m better MAE compared to PPS-v2014 and MODIS-C6. Note that when using CPR (CloudSat) as reference, the neural networks show better results for all categories (low, medium and high) compared to both PPS-v2014 and MODIS-C6. The NN-AVHRR shows 433m lower MAE (cooresponding to 25% reduction of MAE) compared to MODIS-C6 and 483m (corresponding to 27% reduction of MAE) compared to PPS-v2014. The NN-MetImage-$NoCO_2$ shows 550m lower MAE (corresponding to 32% reduction of MAE) compared to MODIS-C6 and 600m lower MAE (correspoding to 34% reduction of MAE) compared to PPS-v2014.

Notice that for MODIS-C6 the MAE for low clouds is high, 1206m, in Table 7, and when compared to CALIOP (Table 6) MODIS-C6 has the highest MAE for low clouds. When checking the MAE per month we found that scores for low clouds were worst for December (at the same time the scores for high clouds were best in December). There turned out to be a bug in the algorithm for low marine cloud height (Richard Frey, MODIS Team, 2017 pers. comm.) which likely affected the results and the bug has been corrected in Version 6.1. However overall validation scores for MODIS-C6 were not affected by the bug (Steve Ackerman, MODIS Team, 2017 pers. comm.).

In Figure 2 the height bias distributions comparing MODIS-C6, PPS-v2014 and NN-AVHRR to CALIOP and CPR (CloudSat) are shown. The NN-AVHRR has the highest and best centered distributions especially for low clouds. For medium level clouds both the NN-AVHRR and the PPS-v2014 show good results. For high clouds we can see that the NN-AVHRR have less clouds predicted too low, especially compared to PPS-v2014. For the high clouds comparison to CloudSat, MODIS-C6 has the highest peak, however MODIS-C6 also has a large amount of clouds with bias close to -6km, which explains why the overall MAE (Table 7) for high clouds is better for the NN-AVHRR.

The scatter plots in Figure 3 show how the height data of the neural networks and the reference methods are distributed compared to CloudSat. Figure 4 show the same type of scatter plots for pressure with CALIOP as truth. These scatter plots show that the NN-CTTH all have similar appearance with most of the data retrieved close to the truth. All methods (NN-CTTH, PPS-v2014 and MODIS-C6) retrieve some heights and pressures that are very far from the true values of CloudSat or CALIOP.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

It is important to remember that some of these seemingly bad results are due to the different FOV for the MODIS and the CALIOP or CloudSat sensors.

To show how performance varies between surfaces and different parts of the globe, the MAE in meters compared to CALIOP are calculated on a Fibonacci grid (constructed using the method described in González, 2009) with a grid evenly spread out on the globe approximately 250km apart and plotted in Figure 5. We can see that all algorithms have problems with clouds around the equator in areas where very thin high cirrus is common. The MAE-difference (Figure 6) shows that the NN-AVHRR is better than MODIS-C6 in most parts of the globe, with the greatest benefit observed closer to the poles. At a few isolated locations MODIS-C6 is better than NN-AVHRR. Figure 7 compares the NN-AVHRR and PPS-v2014 for one scene. Note that the NN-AVHRR has a less noisy appearance and has less nodata.

5 Discussion

The neural networks show great potential for retrieveing cloud height. The NN-CTTH is better in terms of MAE than both PPS-v2014 and the MODIS Collection 6. The neural network algorithms are also useful for instruments with fewer channels than MODIS, including the channels available for AVHRR1. This is important for climate data records which include AVHRR1 data to produce a long, continuous time series.

Neural networks can behave unexpectedly for unseen data. By using a large training dataset and early stopping the risk for unexpected behaviour is decreased. Also the risk for unexpected results in a neural network algorithm can be a fair price to pay given the significant improvements when compared to the current algorithms.

Only near nadir satellite zenith angles were used for training. This might limit the performance for the neural networks at other satellite zenith angles. The NN-MetImage network using the CO_2 channel at 13.3µm shows strong satellite zenith angle dependence and is not useful for higher satellite zenith angles. A solution to train networks to perform better at higher satellite zenith angles could be to include MODIS data from satellite Terra co-located with CALIPSO in the training data, as they will get matches at any satellite zenith angle. Radiative transfer calculations for the CO_2-channels for different satellite-zenith angles could be another way to improve the performance for higher satellite-zenith angles.

Several technical parameters influence the performance of the neural network, for example: learning rate, learning rate decay, momentum, number of layers, number of neurons, weight initialization function and early stopping criteria. For several combinations tested, the differences were in the order of a few hPa. Networks tested using two hidden layers were found to perform better than those using only one hidden layer.

The NN-CTTH algorithm currently has no pixel specific error estimate. The MAE provides a constant error estimate (the same for all pixels). However for some clouds the height retrieval is more difficult, e.g. thin clouds and sub-pixel clouds. Further work to include pixel specific error estimates could be valuable.

The training of neural networks requires reference data (truth). For optimal performance a neural network approach for upcoming new sensors (e.g. MERSI-2, MetImage) being launched, when data from CALIPSO or CloudSat are no longer available, either another truth is needed or a method to robustly transform network trained for one sensor to other sensors is

needed. A way forward could be to include variables with radiative transfer calculations of cloud free brightness temperatures and brightness temperature differences. Further work is needed to test how the networks trained for the MODIS sensor perform for AVHRR, AVHRR1, VIIRS and other sensors. Our results show that networks can be trained using only the channels available on AVHRR, but they might need to be retrained with actual AVHRR data as the spectral response functions of the channels differ. The spectral response functions also differ between different AVHRR instruments, and more investigations are needed to see how networks trained for one AVHRR instrument will perform for other AVHRR instruments.

6 Conclusions

The neural network approach shows high potential to improve cloud height retrievals. Including variables with information on neighbouring pixel values was very important to get good results. Compared to two existing algorithms (MODIS Collection 6 and PPS-v2014) the neural networks show an improvement of mean absolute error (MAE) of at least 400m. This is valid both for validation with CloudSat and CALIOP height products. Low, medium and high level clouds all show better results for the NN-CTTH compared with both PPS-v2014 and MODIS-C6.

The networks trained using only two IR-channels at 11µm and 12µm or 3.7µm showed the most robust performance at higher satellite zenith angles. Including more IR channels does improve results for nadir observations, but introduces some differences between satellite zenith angles. The NN-CTTH could run for the AVHRR1 instrument, in contrast to the other two algorithms neither of which could be applied for the AVHRR1 instrument.

A neural network cloud top pressure, temperature and height algorithm will be be part of the PPS-v2018 release. The PPS software package is accessible via the NWCSAF site nwc-saf.eumetsat.int.

Author contributions. All authors contributed to designing the study. Claudia Adok and Nina Håkansson wrote the code and carried out the experiments. Nina Håkansson drafted the manuscript and prepared the figures and tables. All authors discussed results and revised the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge that the work was mainly funded by EUMETSAT. The CALIOP data were obtained from the NASA Langley Research Center Atmospheric Science Data Center. The CloudSat Data Processing Center (DPC) and Science Teams are further acknowledged for providing CloudSat datasets. NWP data were downloaded from ECMWF. The MODIS/Aqua dataset was acquired from the Level-1 & Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC), located in the Goddard Space Flight Center in Greenbelt, Maryland (https://ladsweb.nascom.nasa.gov/). The authors thank Thomas Heinemann (EUMETSAT) for suggesting adding CloudSat as an independent validation truth.

References

Ackerman, S., Menzel, P., and Frey, R.: MODIS Atmosphere L2 Cloud Product (06_L2), https://doi.org/http://dx.doi.org/10.5067/MODIS/MYD06_L2.006, 2015.

Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and Yang, P.: MODIS Cloud-Top Property Refinements for Collection 6, Journal of Applied Meteorology and Climatology, 51, 1145–1163, https://doi.org/10.1175/JAMC-D-11-0203.1, http://dx.doi.org/10.1175/JAMC-D-11-0203.1, 2012.

Chollet, F. et al.: Keras, https://github.com/fchollet/keras, 2015.

Cotter, A., Shamir, O., Srebro, N., and Sridharan, K.: Better Mini-Batch Algorithms via Accelerated Gradient Methods, in: Advances in Neural Information Processing Systems 24, edited by Shawe-Taylor, J., Zemel, R. S., Bartlett, P. L., Pereira, F., and Weinberger, K. Q., pp. 1647–1655, Curran Associates, Inc., http://papers.nips.cc/paper/4432-better-mini-batch-algorithms-via-accelerated-gradient-methods.pdf, 2011.

Derrien, M., Lavanant, L., and Le Gleau, H.: Retrieval of the cloud top temperature of semi-transparent clouds with AVHRR, in: Proceedings of the IRS'88, pp. 199–202, Deepak Publ., Hampton, Lille, France, 1988.

Dybbroe, A., Karlsson, K.-G., and Thoss, A.: AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modelling - part one: Algorithm description, Journal of Applied Meteorology, 41, 39–54, https://doi.org/http://dx.doi.org/10.1175/JAM-2188.1, http://journals.ametsoc.org/doi/pdf/10.1175/JAM-2188.1, 2005.

Gardner, M. and Dorling, S.: Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences, Atmospheric Environment, 32, 2627 – 2636, https://doi.org/https://doi.org/10.1016/S1352-2310(97)00447-0, http://www.sciencedirect.com/science/article/pii/S1352231097004470, 1998.

González, Á.: Measurement of Areas on a Sphere Using Fibonacci and Latitude–Longitude Lattices, Mathematical Geosciences, 42, 49, https://doi.org/10.1007/s11004-009-9257-x, http://dx.doi.org/10.1007/s11004-009-9257-x, 2009.

Heidinger, A. K., Foster, M. J., Walther, A., and Zhao, X. T.: The Pathfinder Atmospheres–Extended AVHRR Climate Dataset, Bulletin of the American Meteorological Society, 95, 909–922, https://doi.org/10.1175/BAMS-D-12-00246.1, https://doi.org/10.1175/BAMS-D-12-00246.1, 2014.

Hu, X. and Weng, Q.: Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks, Remote Sensing of Environment, 113, 2089 – 2102, https://doi.org/https://doi.org/10.1016/j.rse.2009.05.014, http://www.sciencedirect.com/science/article/pii/S0034425709001655, 2009.

Inoue, T.: On the Temperature and Effective Emissivity Determination of Semi-Transparent Cirrus Clouds by Bi-Spectral Measurements in the 10μm Window Region, Journal of the Meteorological Society of Japan. Ser. II, 63, 88–99, 1985.

Karlik, B. and Olgac, A. V.: Performance analysis of various activation functions in generalized mlp architectures of neural networks, International Journal of Artificial Intelligence and Expert Systems, 1, 111–122, 2011.

Karlsson, K.-G., Anttila, K., Trentmann, J., Stengel, M., Fokke Meirink, J., Devasthale, A., Hanschmann, T., Kothe, S., Jääskeläinen, E., Sedlar, J., Benas, N., van Zadelhoff, G.-J., Schlundt, C., Stein, D., Finkensieper, S., Håkansson, N., and Hollmann, R.: CLARA-A2: the second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, Atmospheric Chemistry and Physics, 17, 5809–5828, https://doi.org/10.5194/acp-17-5809-2017, https://www.atmos-chem-phys.net/17/5809/2017/, 2017.

Kox, S., Bugliaro, L., and Ostler, A.: Retrieval of cirrus cloud optical thickness and top altitude from geostationary remote sensing, Atmospheric Measurement Techniques, 7, 3233–3246, https://doi.org/10.5194/amt-7-3233-2014, https://www.atmos-meas-tech.net/7/3233/2014/, 2014.

Marchand, R., Mace, G. G., Ackerman, T., and Stephens, G.: Hydrometeor Detection Using Cloudsat—An Earth-Orbiting 94-GHz Cloud Radar, Journal of Atmospheric and Oceanic Technology, 25, 519–533, https://doi.org/10.1175/2007JTECHA1006.1, https://doi.org/10.1175/2007JTECHA1006.1, 2008.

Meng, L., He, Y., Chen, J., and Wu, Y.: Neural Network Retrieval of Ocean Surface Parameters from SSM/I Data, Monthly Weather Review, 135, 586–597, https://doi.org/10.1175/MWR3292.1, https://doi.org/10.1175/MWR3292.1, 2007.

Menzel, W. P., Frey, R. A., Zhang, H., Wylie, D. P., Moeller, C. C., Holz, R. E., Maddux, B., Baum, B. A., Strabala, K. I., and Gumley, L. E.: MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results, Journal of Applied Meteorology and Climatology, 47, 1175–1198, https://doi.org/10.1175/2007JAMC1705.1, http://dx.doi.org/10.1175/2007JAMC1705.1, 2008.

Milstein, A. B. and Blackwell, W. J.: Neural network temperature and moisture retrieval algorithm validation for AIRS/AMSU and CrIS/ATMS, Journal of Geophysical Research: Atmospheres, 121, 1414–1430, https://doi.org/10.1002/2015JD024008, http://dx.doi.org/10.1002/2015JD024008, 2015JD024008, 2016.

Minnis, P., Hong, G., Sun-Mack, S., Smith, W. L., Chen, Y., and Miller, S. D.: Estimating nocturnal opaque ice cloud optical depth from MODIS multispectral infrared radiances using a neural network method, Journal of Geophysical Research: Atmospheres, 121, 4907–4932, https://doi.org/10.1002/2015JD024456, http://dx.doi.org/10.1002/2015JD024456, 2015JD024456, 2016.

MODIS Science Data Support Team: MYD021KM, https://doi.org/http://dx.doi.org/10.5067/MODIS/MYD021KM.006, 2015a.

MODIS Science Data Support Team: MYD03, https://doi.org/http://dx.doi.org/10.5067/MODIS/MYD03.006, 2015b.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Rossow, W. B. and Schiffer, R. A.: Advances in Understanding Clouds from ISCCP, Bulletin of the American Meteorological Society, 80, 2261–2287, https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2, 1999.

SMHI: Algorithm Theoretical Basis Document for Cloud Top Temperature, Pressure and Height of the NWC/PPS, NWCSAF, 4.0 edn., http://www.nwcsaf.org/AemetWebContents/ScientificDocumentation/Documentation/PPS/v2014/NWC-CDOP2-PPS-SMHI-SCI-ATBD-3_v1_0.pdf, 2015.

Stengel, M., Stapelberg, S., Sus, O., Schlundt, C., Poulsen, C., Thomas, G., Christensen, M., Carbajal Henken, C., Preusker, R., Fischer, J., Devasthale, A., Willén, U., Karlsson, K.-G., McGarragh, G. R., Proud, S., Povey, A. C., Grainger, D. G., Meirink, J. F., Feofilov, A., Bennartz, R., Bojanowski, J., and Hollmann, R.: Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud_cci project, Earth System Science Data Discussions, 2017, 1–34, https://doi.org/10.5194/essd-2017-48, https://www.earth-syst-sci-data-discuss.net/essd-2017-48/, 2017.

Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Girolamo, L. D., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B. C., Menzel, W. P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, Bulletin of the American Meteorological Society, 94, 1031–1049, https://doi.org/10.1175/BAMS-D-12-00117.1, https://doi.org/10.1175/BAMS-D-12-00117.1, 2013.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Theano Development Team: Theano: A Python framework for fast computation of mathematical expressions, arXiv e-prints, abs/1605.02688, http://arxiv.org/abs/1605.02688, 2016.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 1. MODIS data from 2010 used for training and validation of the neural networks.

Dataset	Days used
Training	1^{st} January March July September
	14^{th} February April May
	14^{th} August October December
Validation during training	1^{st} May
	14^{th} March July November
Testing under development	1^{st} November
	14^{th} January June September
Final validation	1^{st} February April June
	1^{st} August October December

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 2. Description of variable types used to train neural networks.

Variable type	Variable names	Note
Surface pressure	P_S	Max pressure for pixel
NWP temperatures at surface, 950, 850, 700, 500, 250hPa	$T_S, T_{950}, T_{850}, T_{700}, T_{500}, T_{250}$	BT to pressure conversion
NWP column integrated water vapour	$Ciwv$	Expected BT differences
Brightness Temperature (BT) for 11µm or 12µm	B_{11} or B_{12}	Opaque temperature
BT for water vapour channels at 6.7µm 7.3µm	$B_{6.7}$, $B_{7.3}$	High or low
BT for CO_2 channel at 13.3µm	$B_{13.3}$	High or low
BT differences.	$B_{11} - B_{12}$, $B_{11} - B_{3.7}$, $B_{8.5} - B_{11}$	Opacity or phase
BT differences *to* warmest/coldest neighbour.	$B^W_{12} - B_{12}$, $B^C_{12} - B_{12}$ or	Edge or thin
	$B^W_{11} - B_{11}$, $B^C_{11} - B_{11}$	Edge or thin
BT differences *for* warmest/coldest neighbour.	$B^W_{11} - B^W_{12}$, $B^C_{11} - B^C_{12}$ or	Opacity
	$B^W_{11} - B^W_{3.7}$, $B^C_{11} - B^C_{3.7}$	Opacity
Texture: standard deviation of variable for 5x5 pixels	$S_{B_{11}-B_{12}}, S_{B_{11}}, S_{B_{3.7}}$	Edge or thin

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 3. Description of the different networks. See Table 2 for explanation of the variables. The NWP variables: P_S, T_S, T_{950}, T_{850}, T_{700}, T_{500}, T_{250} are used in all networks.

Network name	Network specific variables
NN-NWP	$Ciwv$
NN-OPAQUE	B_{12}
NN-BASIC	B_{12}, $B_{11} - B_{12}$,
NN-BASIC-CIWV	B_{12}, $B_{11} - B_{12}$, $Ciwv$,
NN-AVHRR	B_{12}, $B_{11} - B_{12}$, $Ciwv$, $B_{11}^W - B_{12}^W$, $B_{11}^C - B_{12}^C$, $B_{12}^W - B_{12}$, $B_{12}^C - B_{12}$, $S_{B_{11}-B_{12}}$, $S_{B_{11}}$,
NN-VIIRS	B_{12}, $B_{11} - B_{12}$, $Ciwv$, $B_{11}^W - B_{12}^W$, $B_{11}^C - B_{12}^C$, $B_{12}^W - B_{12}$, $B_{12}^C - B_{12}$, $S_{B_{11}-B_{12}}$, $S_{B_{11}}$, $B_{8.5} - B_{11}$
NN-MERSI-2	B_{12}, $B_{11} - B_{12}$, $Ciwv$, $B_{11}^W - B_{12}^W$, $B_{11}^C - B_{12}^C$, $B_{12}^W - B_{12}$, $B_{12}^C - B_{12}$, $S_{B_{11}-B_{12}}$, $S_{B_{11}}$, $B_{8.5} - B_{11}$, $B_{7.3}$
NN-MetImage-$NoCO_2$	B_{12}, $B_{11} - B_{12}$, $Ciwv$, $B_{11}^W - B_{12}^W$, $B_{11}^C - B_{12}^C$, $B_{12}^W - B_{12}$, $B_{12}^C - B_{12}$, $S_{B_{11}-B_{12}}$, $S_{B_{11}}$, $B_{8.5} - B_{11}$, $B_{7.3}$, $B_{6.7}$
NN-MetImage	B_{12}, $B_{11} - B_{12}$, $Ciwv$, $B_{11}^W - B_{12}^W$, $B_{11}^C - B_{12}^C$, $B_{12}^W - B_{12}$, $B_{12}^C - B_{12}$, $S_{B_{11}-B_{12}}$, $S_{B_{11}}$, $B_{8.5} - B_{11}$, $B_{7.3}$, $B_{6.7}$, $B_{13.3}$
NN-AVHRR1	B_{11}, $B_{11} - B_{3.7}$, $Ciwv$, $B_{11}^W - B_{3.7}^W$, $B_{11}^C - B_{3.7}^C$ $B_{11}^W - B_{11}$, $B_{11}^C - B_{11}$ $S_{B_{3.7}}$, $S_{B_{11}}$

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 4. Description of the imager channels used for different networks.

Imager channel: Network name	B_{11}	B_{12}	$B_{8.5}$	$B_{7.3}$	$B_{6.7}$	$B_{13.3}$	$B_{3.7}$
NN-NWP							
NN-OPAQUE	x						
NN-BASIC	x	x					
NN-BASIC-CIWV	x	x					
NN-AVHRR	x	x					
NN-VIIRS	x	x	x				
NN-MERSI-2	x	x	x	x			
NN-MetImage-$NoCO_2$	x	x	x	x	x		
NN-MetImage	x	x	x	x	x	x	
NN-AVHRR1	x						x

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 5. Mean absolute error (MAE) for different algorithms compared to CALIOP top layer pressure. The final validation dataset (see Table 1), containing 1796428 pixels (45% high, 39% low and 16% medium level clouds) is used. Pixels with valid pressure for PPS-v2014, MODIS-C6, and CALIOP are considered. The low, medium and high classes are from CALIOP feature classification flag.

	MAE [hPa]			
	all	low	medium	high
PPS-v2014	122.9	79.4	88.6	173.5
MODIS-C6	124.3	90.4	140.0	148.4
NN-NWP	191.6	140.8	110.5	266.0
NN-OPAQUE	113.3	81.3	105.1	144.5
NN-BASIC	93.9	66.7	92.8	118.3
NN-BASIC-CIWV	92.1	66.4	91.4	115.0
NN-AVHRR	72.2	54.1	67.4	89.9
NN-VIIRS	65.7	49.1	59.2	82.7
NN-MERSI-2	61.2	46.7	52.0	77.3
NN-MetImage-$NoCO_2$	60.0	45.5	54.3	74.8
NN-MetImage	53.6	42.7	51.3	64.1
NN-AVHRR1	76.1	53.6	70.0	98.1

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 6. Mean absolute error (MAE) in meters for different algorithms compared to CALIOP top layer altitude. The final validation dataset (see Table 1), containing 1793142 pixels (45% high, 39% low and 16% medium level clouds), where all algorithms had a cloud height is used. The low, medium and high classes are from CALIOP feature classification flag. A small amount 0.2% of the pixels were excluded because of missing height or pressure below 70hPa for any of the algorithms.

	MAE [m]			
	all	low	medium	high
PPS-v2014	2087	837	1124	3542
MODIS-C6	1917	944	1759	2833
NN-AVHRR	1290	567	962	2049
NN-VIIRS	1177	514	828	1891
NN-MERSI-2	1110	488	727	1797
NN-MetImage-$NoCO_2$	1081	478	757	1732
NN-Metimage	964	451	707	1510
NN-AVHRR1	1375	560	978	2239

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

Table 7. Mean absolute error (MAE) in meters for different algorithms compared to CPR (CloudSat) Height. The final validation dataset (see Table 1), containing 1121199 pixels (53% high, 27% low and 21% medium level clouds) is used. The low, medium and high classes are derived comparing the CloudSat height to the NWP height at 440hPa and 680hPa. A cloudy threshold of 30% is used for CloudSat.

	MAE [m]			
	all	low	medium	high
PPS-v2014	1761	977	1365	2315
MODIS-C6	1711	1206	1912	1888
NN-AVHRR	1278	771	1218	1559
NN-VIIRS	1223	766	1144	1486
NN-MERSI-2	1135	748	1061	1362
NN-MetImage-$NoCO_2$	1161	768	1095	1386
NN-Metimage	1186	802	1119	1407
NN-AVHRR1	1297	858	1225	1548

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

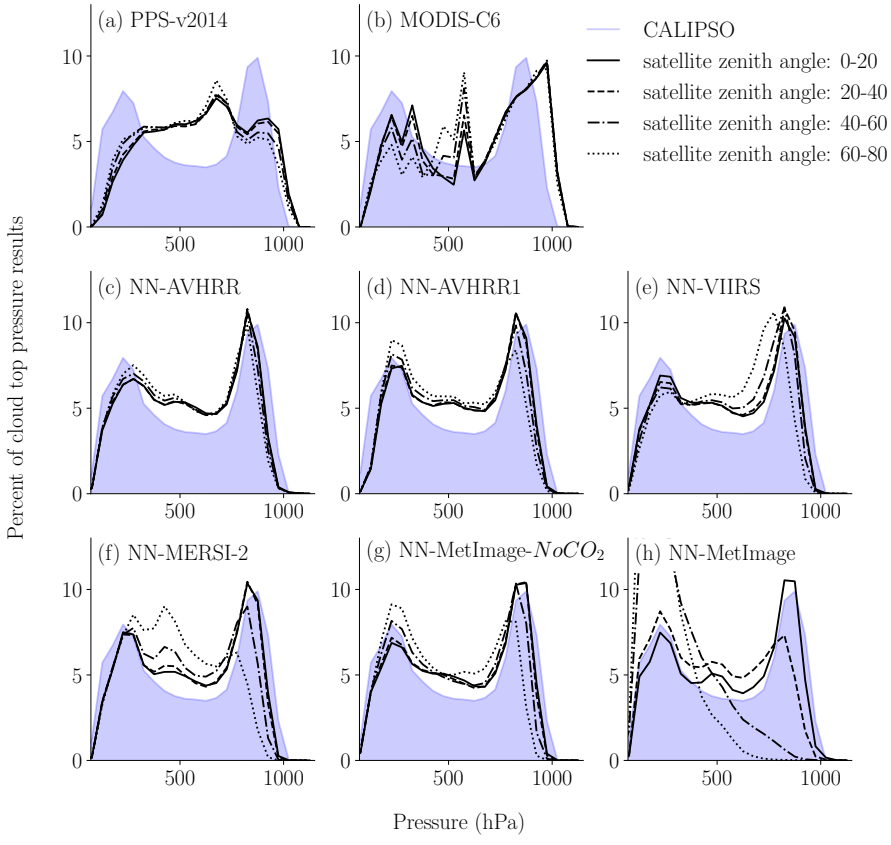

Figure 1. Retrieved pressure dependence on satellite zenih angle. CALIOP pressure distribution is shown in light blue. The percent of results are calculated in 50hPa bins. The final validation dataset is used (see Table 1).

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

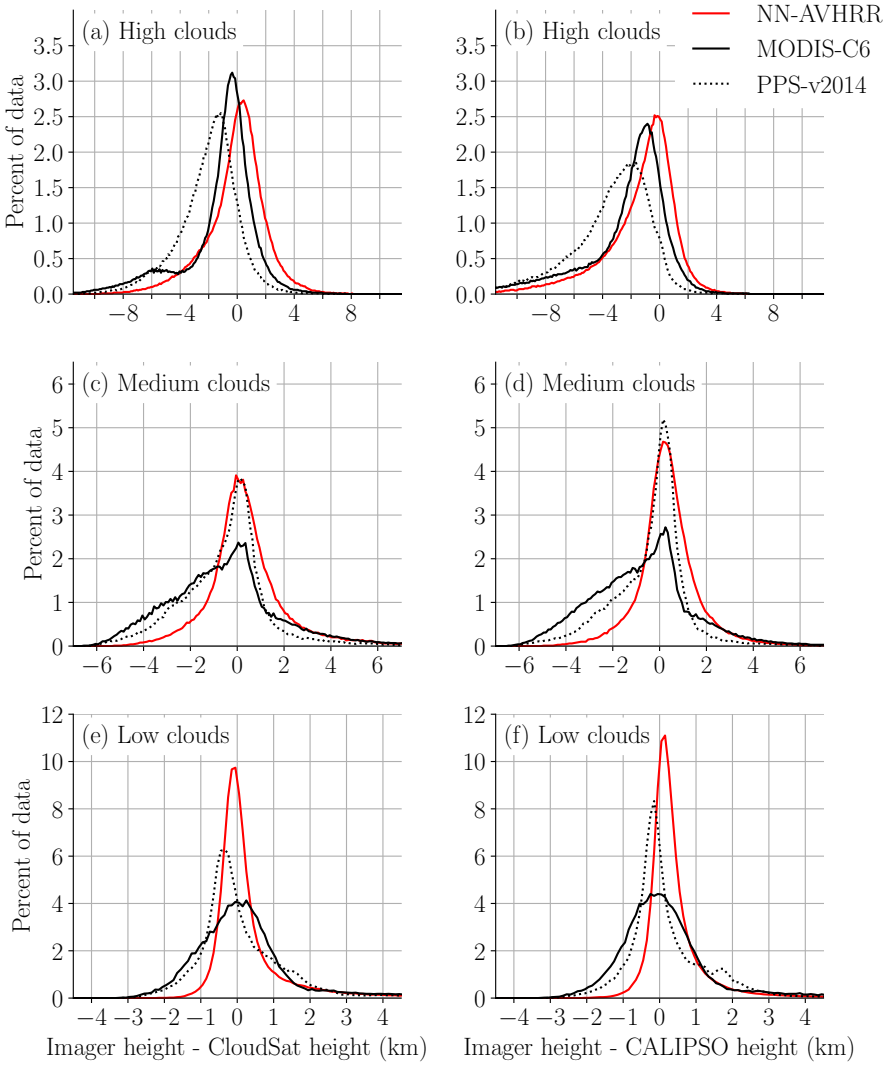

Figure 2. Bias distribution compared to CPR (CloudSat) (left) and CALIOP (right). The percent of data is calculated in 0.1km bins. For CALIOP the low, medium and high clouds are determined from CALIOP feature classification flag. For CloudSat the low, medium, high clouds are determined from CloudSat height compared to NWP geopotential height at 440hPa and 680hPa. The final validation dataset (see Table 1) where all algorithms had a height reported is used. Note that the values on the y-axis are dependent of the bin size. The peak at 11% for NN-AVHRR in subplot (f), means that 11% of the retrieved heights are between the CALIOP height and the CALIOP height + 0.1km.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

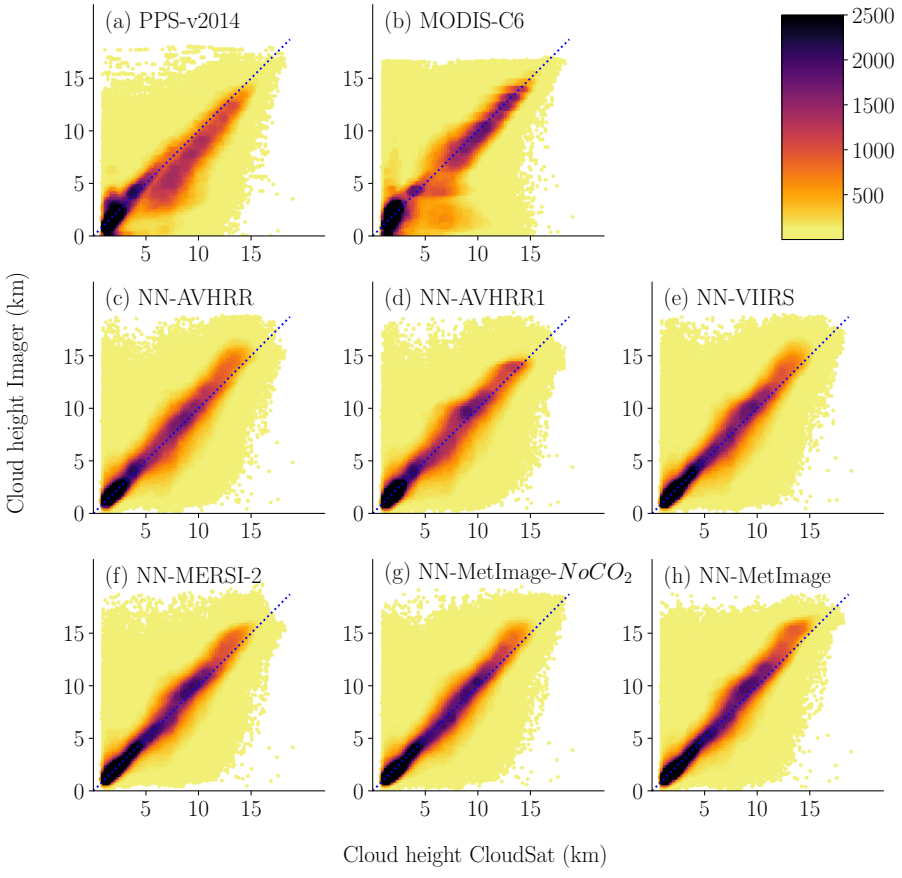

Figure 3. Scatters plot of the height for the neural networks and for the reference methods against CloudSat height. The data were divided in bins of size 0.25 x 0.25 (km) for colour coding. The number of points in each bin determines the colour of the point. The final validation dataset (see Table 1) where all algorithms had a height reported is used. Five points where CloudSat had a height above 22km where excluded. A cloudy threshold of 30% is used for CloudSat.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

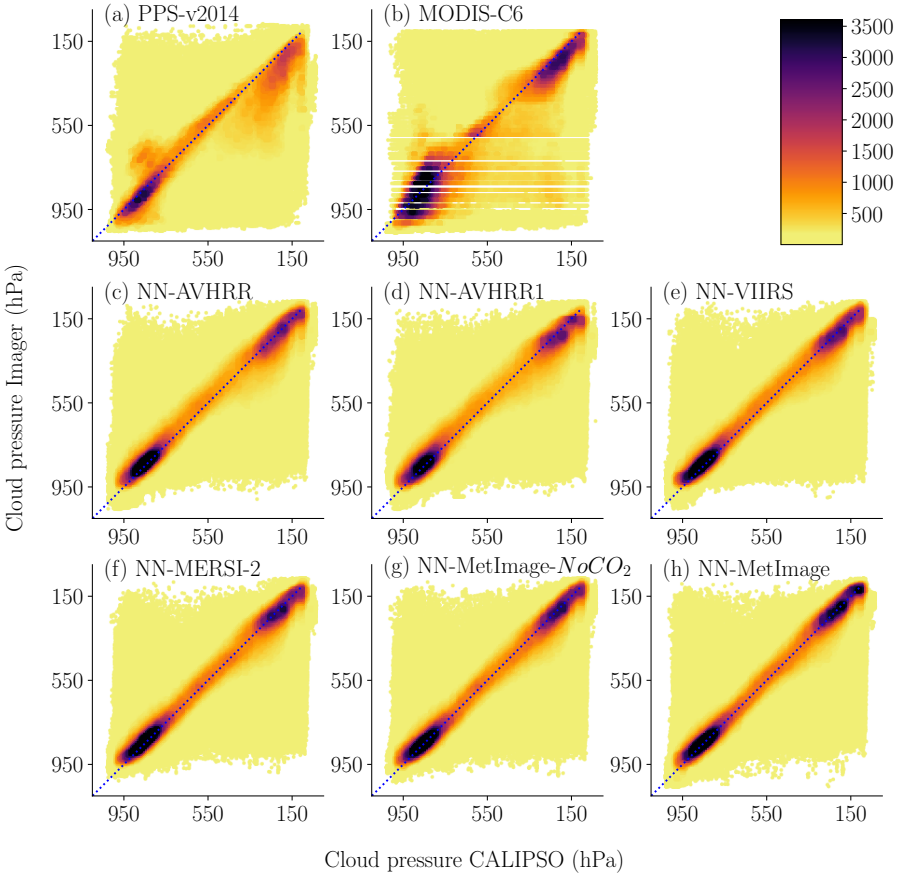

Figure 4. Scatter plots of the pressure for the neural networks and for the reference methods against CALIOP cloud top pressure. The data were divided in bins of size 10 x 10 (hPa) for colour coding. The number of points in each bin determines the colour of the point. The final validation dataset (see Table 1) where all algorithms had a height reported is used.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

Atmospheric Measurement Techniques Discussions

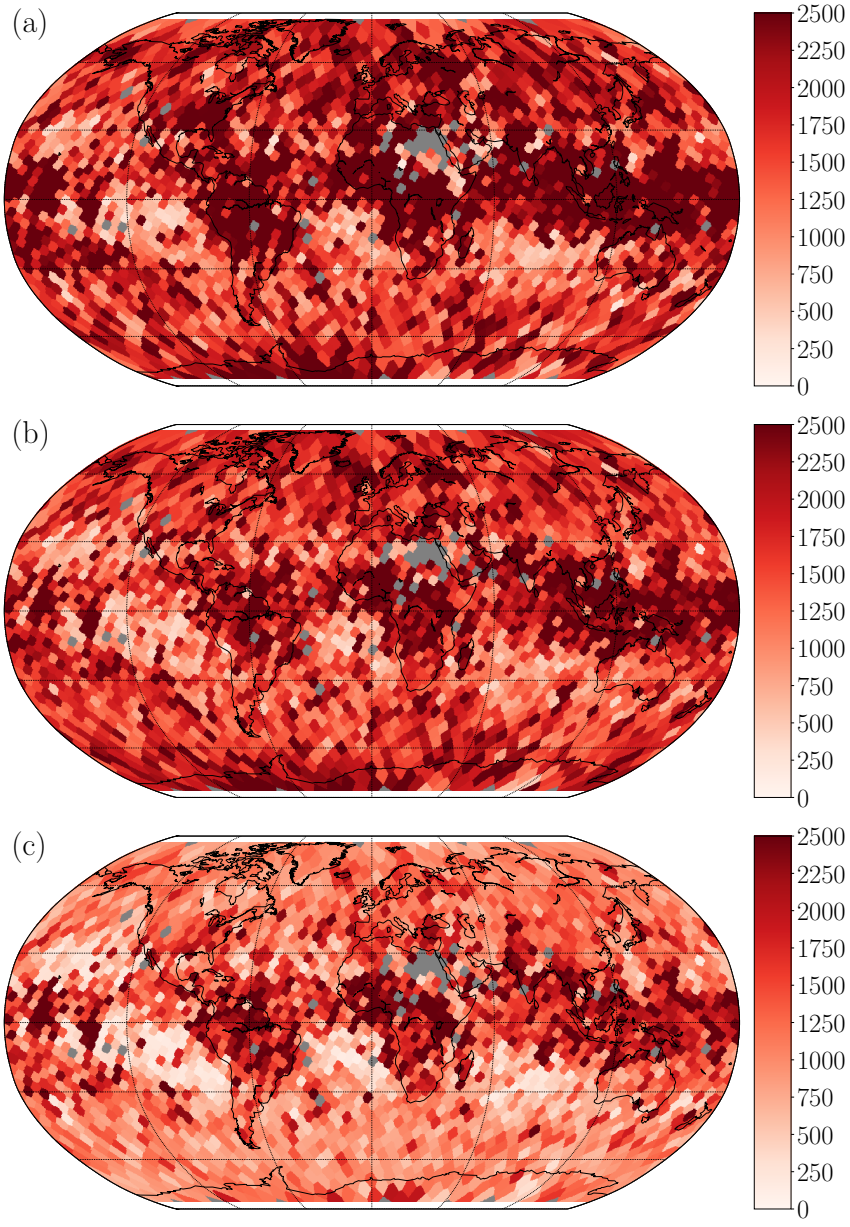

Figure 5. Mean absolute error in meters compared to CALIOP height. From the top a) PPS-v2014, b) MODIS-C6, and c) NN-AVHRR. Results are calculated for bins evenly spread out 250km apart. Bins with less than 10 cloudy pixels are excluded (plotted in dark grey). The final validation and testing under development data (see Table 1) are included to get enough pixels.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

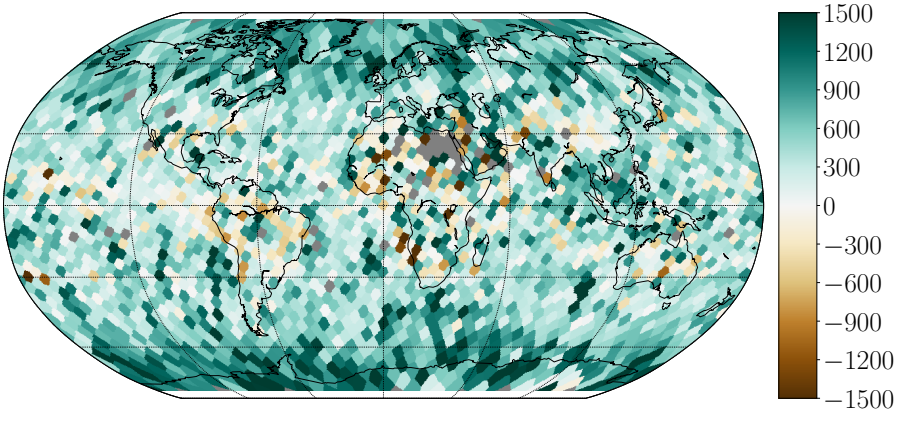

Figure 6. Mean absolute error difference in meters between MODIS-C6 and NN-AVHRR compared to CALIOP. Results are calculated for bins evenly spread out 250km apart. Bins with less than 10 cloudy pixels are excluded (plotted in dark grey). Dark green means NN-AVHRR is 1.5km better than MODIS-C6, dark brown means MODIS-C6 is 1.5km better than NN-AVHRR. The final validation and testing under development data (see Table 1) are included to get enough pixels.

Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2017-443
Manuscript under review for journal Atmos. Meas. Tech.
Discussion started: 30 January 2018

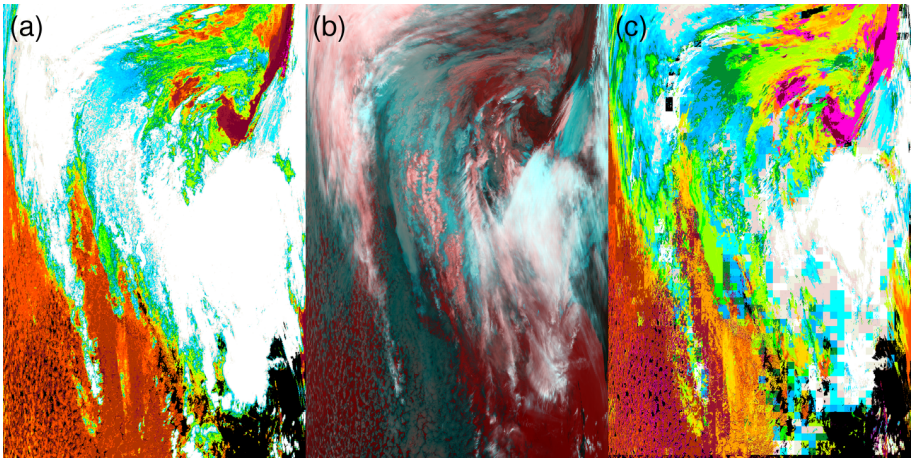

Figure 7. Comparing the cloud top height from the NN-AVHRR (left) to PPS-v2014 (right) with a RGB in the middle using channels at 3.7µm, 11µm, 12µm. Cloud retrievals below 2km are red, brown or purple. Cloud retrievals between 2 and 5km are orange and yellow. Cloud retrievals between 5 and 8km are green and blue. White pixels are cloud retrievals above 8km. Notice that the NN-AVHRR is smoother, contain less nodata and that the small high ice clouds in the lower part of the figure are better captured. This is from MODIS on Aqua 14th of January 2010, 00:05UTC.