

Fallback atmosphere models for lidar measurements without corresponding radio-sounding data and modelling of a local atmosphere profile from historic radio-sounding data: Example of the Sao Paulo lidar metropolitan station

Riad Bourayou^{(1,*),} Fabio Juliano Da Silva Lopes^{(2),} Renata Facundes Da Costa^{(2),}
Patricia Rodrigues Ferrini^{(2),} Eliane Gonçalves Larroza^{(2),} Walter Morinobu Nakaema^{(2),}
Gregori Arruda Moreira^{(2),} Felipe Vasconcelos Araujo^{(2),} Bruno Poluha^{(2),}
Eduardo Landulfo⁽²⁾

1. Independent researcher, www.bourayou.com, Weichselplatz 8, 12045 Berlin, Germany.
2. Laser Environmental Applications, Center for Lasers and Applications, Nuclear and Energy Institute, Av. Prof Lineu Prestes 2242, Cidade Universitária 05508-000 São Paulo, Brazil.

(*) Email: riad@bourayou.com

Received / Recibido: 15/02/2014. Revised / Revisado: 05/06/2014. Accepted / Aceptado: 26/06/2014.

DOI: <http://dx.doi.org/10.7149/OPA.47.3.183>

ABSTRACT:

We compare the performance of some available atmospheric models for the atmosphere of Sao Paulo (Brazil) to be used in case of absence of radio-sounding data for the given day. We developed our own model (SPm) from historic radio-sounding data in order to create a local model. By performing inversions of lidar signals (distributed over a year), we could benchmark the performance of the models against radio-sounding data. SPm and ISA-15N show the smallest deviation and represent, therefore, the best fallback models for this southern latitude.

Key words: Lidar, Atmospheric Modelling, Temperature Profile, Aerosol Extinction.

REFERENCES AND LINKS / REFERENCIAS Y ENLACES

- [1]. U.S. Standard Atmosphere, NASA document NASA-TM-X-74335 (1976).
- [2]. E. Thorin, *Design of Algorithms to Extract Atmospheric Aerosol Extinction from Raman Lidar Data*, M.Sc Thesis, University of Linköping (2006)
- [3]. D. J. Diston, *Computational Modelling and Simulation of Aircraft and the Environment*, John Wiley & Sons, Ltd, Chichester, UK(2009). [DOI](#)
- [4]. <http://weather.uwyo.edu/upperair/sounding.html>, Station Campo do Marte SBMT.
- [5]. H. Bai, *A New Resampling Method to Improve Quality of Research with Small Samples*, PhD Thesis, Univ. of Cincinnati (2006).
- [6]. B. E. Wampold, C. J. Drew, *Theory and Application of Statistics*, McGraw-Hill, New York (1990).
- [7]. C. C Tomkins, "An Introduction to Non-parametric Statistics for Health Scientists", *University of Alberta Health Sciences Journal* **3** (1), 20-21 (2006).
- [8]. S. Siegel, N. J. Castellan, *Nonparametric Statistics for the Behavioural Sciences*, McGraw-Hill, New York (1988).

1. Introduction

In the absence of radio-sounding data (RWS), the inversion of the lidar signal can be done using modelled atmosphere profiles. Many available models, for instance US1976 [1], an early

Earlinet Atmosphere [2] and ISA [3], were designed for the northern hemisphere of the globe and therefore their application could be challenging when dealing with LaLiNet (Latin-American Lidar Network) stations. A comparison

with the RWS data for Sao Paulo showed that the measured temperature profile at this latitude significantly departs from the modelled values.

We present a methodology to create local atmospheric models from historical radio-soundings; the straight-forwardness of the principle would make it applicable to any Lidar station of the LaLiNet. In a first approach, only the temperature profile has been modelled, as the pressure profile did correlate well with the standard models. In addition, the use of historic local data will, contrarily to some classical models, take the seasonal variation of the temperature profile into account.

In order to assess the relative performance of the chosen atmospheric temperature models for LIDAR retrievals, real data acquired at the São Paulo metropolitan lidar station will be reconstructed and compared to the case where radio-sounding data have been used.

2. Methods

2.1. Lidar data: acquisition and reconstruction

Elastic lidar data are routinely acquired in the Metropolitan Station of the IPEN (city of Sao Paulo, Brazil), and we extracted a subset of measurements for the year 2007. For each month, a clear-sky day was picked where radio-sounding data were also available. The acquired lidar signals were averaged in order to deliver a single measured profile

Only the 532 nm channel of the dataset was used, after the gluing of the PC and analogue profiles. The original vertical resolution of 15 m was dithered to 60 m (for higher altitudes) in order to compensate for SNR loss. The baseline drift at the far end of the lidar signal (a condition where the mean of the signal at very high altitudes is excessively offset) was also corrected for.

The aerosol extinction was reconstructed using the Klett algorithm, independently for each atmospheric model and for the RWS data.

2.2. Generation of the SPm local model

For the generation of our atmosphere model for the city of Sao Paulo (Brazil), called SPm, mean

temperature profiles had to be calculated over the years where RWS data were accessible (2006-2011) from the nearest RWS-station of Campo de Marte, representing a volume of roughly 730 profiles per year. The monthly data were downloaded from the archive of the University of Wyoming [4] and saved in hypertext-markup language format, before being parsed using a custom Matlab (The Mathworks, Natick, MA) routine that extracted the profiles and checked for outliers. The temperature profiles were subsequently interpolated to the standardized vertical resolution of 15 m between the altitudes 722 m to 16 km. For this study, the maximal altitude was limited to 16 km due to the scarcity of the available RWS dataset above this height.

For every altitude, the mean value of the daily temperature was calculated, and this synthetic monthly average of the temperature profile was modelled over three altitude ranges by different polynomial orders. Figure 1 shows the chosen polynomial fit types, in a case where all RWS measurements reached above 25 km.

Thus, for each month between January 2006 and December 2011, we have calculated a mean temperature profile, and through fitting with polynomials the latter is described by a set of 11 coefficients.

The next step is to find the mean profile for each of the 12 months of the year, that is to say,

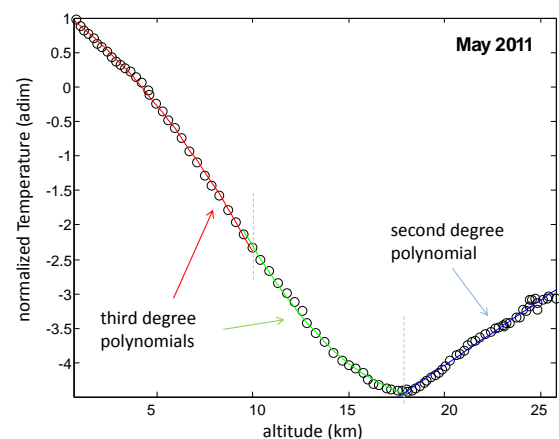


Fig. 1. The monthly-averaged temperature profile of the atmosphere has been modelled piece-wise using polynomials. To illustrate the resulting curve, the mean temperature profile for the month of May of 2011 (symbols) has been overlaid on the fitted polynomials.

generate the mean profile for January based on the six occurrences from 2006 to 2011. For this, we would take the mean of the 11 coefficients, but each coefficient has only six occurrences. The statistical parametric methods would not allow for a trustworthy analysis: the distribution function of the 6 values is unknown, and outliers are not to be excluded. It is known that in this frame a resampling method is appropriate for the statistic analysis. We settled on the conventional bootstrap method [5], as it is proven and well documented [6-8]. Our “surrogate population” contains 6 values (for a given coefficient, for a given month, over the 6 years), from which 64 bootstrap samples were created, in order to estimate of the shape of the distribution of the mean. This procedure, done for each of the 12 months of the year, also provides the 95% confidence interval for each of the 11 coefficients. Figure 2 shows the coefficients that were inferred for the (monthly) polynomial fits.

We refined the model in order to be able to generate the temperature profile for any arbitrary day in the year. At this step of our analysis, the “level of detail” is in the order of the month: we have modelled each month as a block. In the further steps we want to ensure that the SPM model generates a distinct temperature profiles for each day. We thus make the hypothesis of smooth variations of the real temperature profile from one day to another. This smooth variation would translate into smooth changes in the values of the coefficients, so we use a smooth interpolant to extend the 12 values (one value per month) to 365 (one value per day). The interpolant for each of the 11 coefficients, a Fourier series of the third order that generates a cyclic model with a period of a year and assert the continuity from the month of December to January, was least-square fitted over the 12 monthly values of the coefficients.

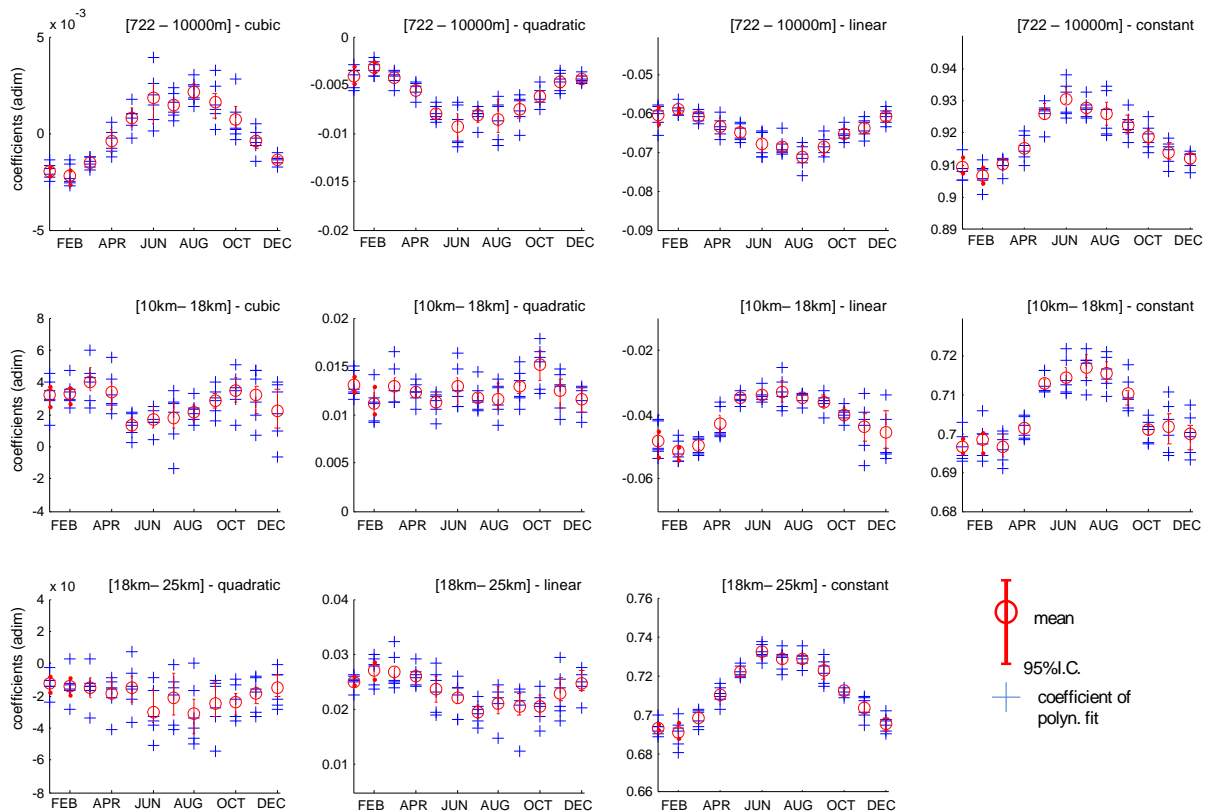


Fig. 2. Each month (in abscissa) is modelled through a set of 11 coefficients (hence 11 sub-figures). The blue crosses show the 6 values of a given coefficient for a given month over the observation period of 6 years. Using a resampling method, their mean (red circle) and 95% interval of confidence values (red bars) could be calculated, so that the monthly values for all polynomial coefficients could be derived.

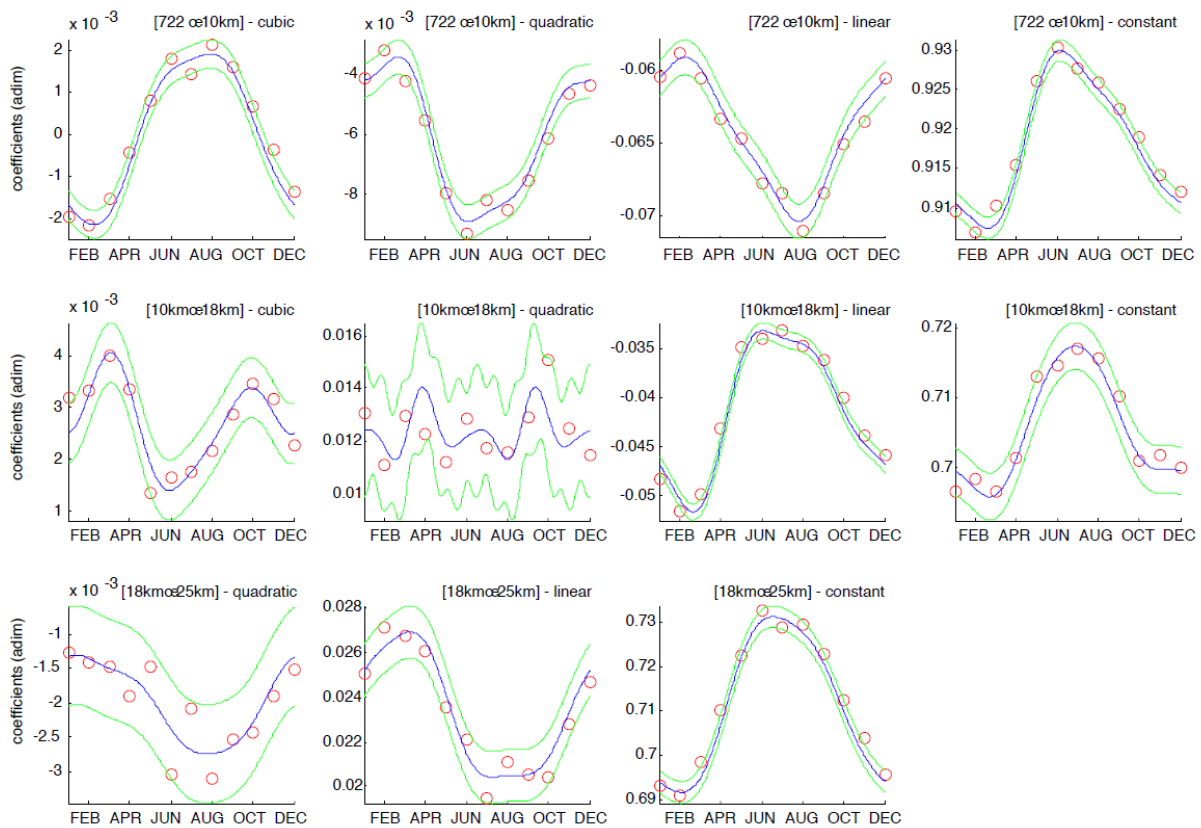


Fig. 3. The annual variation of the 11 coefficients (hence 11 sub-figures) was smoothed using a Fourier series interpolant (blue curve, the green dotted lines show the standard deviation) that respects the continuity between the end of December and the beginning of January.

The description of the interpolant was stored to allow for the generation of a synthetic temperature profile for any chosen date/hour in the year. On Fig. 3 we can observe the seasonal variation of the temperature profile, which is reflected clearly in the cyclic change of the coefficients of the SPM model along the year.

3. Benchmarking the models

We generated a profile of the pressure and of the temperature using the models SPM, US-1976, Earlinet and ISA-15N, while a unique pressure profile was used. A preliminary comparison of the ISA models with the radio-soundings showed that the ISA-15N was the more appropriate than its ISA-siblings (for instance: 30N and 45N), so only this model was retained. Figure 4 shows how the chosen synthetic temperature models mimic an arbitrary sample radio-sounding measurement on an arbitrarily chose day of the beginning of the year 2007.

In order to benchmark the models, the aerosol extinction was computed using atmospheric values from the models and then compared to the values obtained using the experimental RWS data as a reference. A profile-to-profile ratio of the extinction coefficient was drawn by dividing the values obtained using a given model by the values obtained using the time-nearest radio-sounding. A rather homogeneous value was found throughout all altitudes. This result could be explained by the fact that the models do not differ significantly for low altitudes, but rather above 10km, where the measured signal was weak as we chose to analyse clear sky days. It is thus to assume that, for the study of cirrus clouds, the models would still deliver converging extinction values, as the models US1976 and Earlinet only depart significantly from the RWS data above an altitude around 12 km.

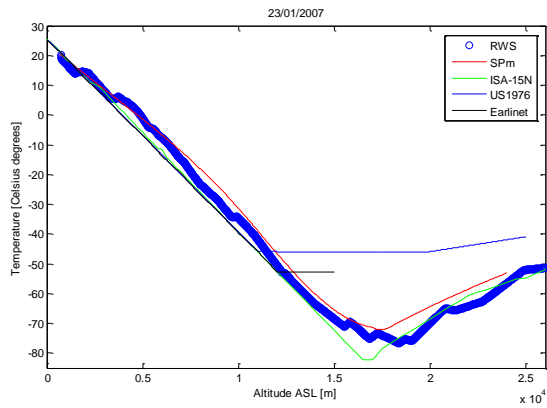


Fig. 4. Radio-sounding data of the day January 23rd of 2007 compared to the benchmarked models (solid lines).

The quality in the results delivered by the models is further explored using statistical quantities comparing them to the reference dataset obtained using RWS; we use the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{1}{2} \sum_{i=1}^N (\alpha_i^{RWS} - \alpha_i^{model})^2}, \quad (1)$$

the mean relative error (MRE) expressed in percent:

$$MRE = 100 \sum_{i=1}^N \frac{|\alpha_i^{model} - \alpha_i^{RWS}|}{\alpha_i^{RWS}}, \quad (2)$$

and the mean absolute error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |\alpha_i^{model} - \alpha_i^{RWS}|, \quad (3)$$

where N is the total number of altitudes considered in the lidar retrieval vector and α is the extinction at the altitude i (subscript) for the given model (superscript). The results for the profile chosen for each month are compiled in the figures

The RMSE on the retrieved extinction profiles (Fig. 5) show that the model ISA-15N gives closest results to these retrieved using the RWS data, closely followed by the SPm model (with a yearly mean RMSE of respectively 4.43×10^{-8} and 5.39×10^{-8}). The associated root mean squared errors are approximately one order of magnitude smaller compared to other models. There is an unexplained poorer performance for all models for the months of January and December.

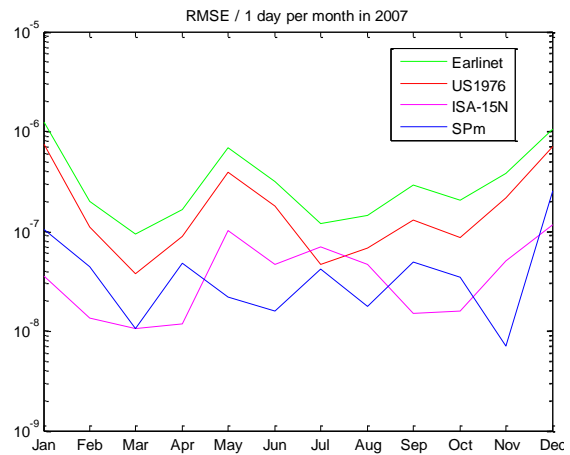


Fig. 5. Root mean squared error of the aerosol extinction retrieved using the models compared to when using radio-soundings. The SPm and ISA-15N show the best performance.

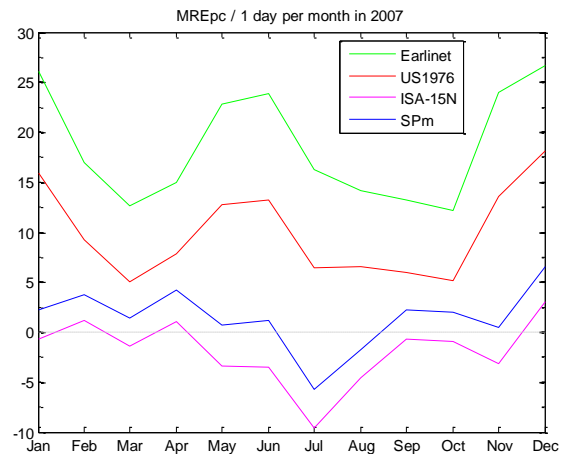


Fig. 6. Mean relative error of the aerosol extinction retrieved using the models compared to when using radio-soundings. The SPm and ISA-15N show the best performance.

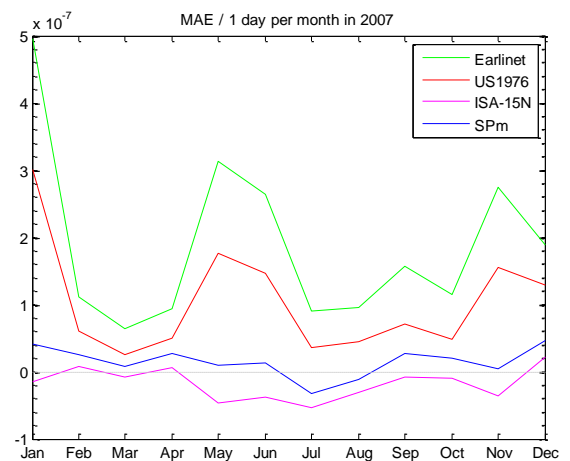


Fig. 7. Mean absolute error of the aerosol extinction retrieved using the models compared to when using radio-soundings. The SPm and ISA-15N show the best performance.

The results of the MRE, shown on Fig. 6, acknowledge the good performance of the models ISA-15N and SPm. We note that the ISA-15N model tends to under-estimate the extinction, while SPm shows the opposite trend (with a yearly mean MRE of respectively -1.89% and 1.44%). This trend mirrors in the MRE (Fig. 7, with a yearly mean MAE of respectively -1.73×10^{-8} and 1.48×10^{-8}), showing the superior performance of the two models.

4. Discussion

The temperature profiles from the models SPm and ISA-15N showed a significantly better match to the radio-soundings than both other models for the latitude of São Paulo, and a similar finding was observed for the days used in this study. They were the only models reaching a MRE below 5%. The US-1976 and Earlinet models show, in average, poor performance for the southern latitude of São Paulo. The SPm model had the smallest MAE and MRE, but it is noticeable that the ISA-15N model shows a consistent performance, sometimes even superior to the one offered by the SPm model, although the latter derives from local radio-sounding data. A possible reason is the very low number of years used to create the SPm model; we were forced to use non-parametric statistics and this might create a bias. The advantage of the methodology presented here is that it is possible to create a rather consistent local model for the atmospheric temperature while starting with a rather small historical RWS dataset. SPm can be refined each year by adding the latest RWS measurements until reaching the maturity of the standard atmosphere models.

Comparing the RWS data to the SPm profile show that the polynomial used for the lowest altitudes (from 722 m to 10 km) is too smooth and sometimes departs significantly for the RWS profile. Although this is a possible way of improvement, the ISA-15N is piece-wise linear and did show a very respectable performance. Besides, the RWS data show that the shoulder around 16km varies in altitude, and the SPm should better account for this fact.

5. Conclusions

The models SPm and ISA-15N showed a significantly better performance for LIDAR retrieval above São Paulo than both other models throughout the year. Although the ISA-15N model does not consider seasonal variations, its performance would make it suitable for modelling the atmosphere of São Paulo, particularly for the months December to May, where it surpassed the SPm model. Nevertheless, the methodology used to create the SPm model could be used to create a fallback atmospheric profile for Lidar stations where the standard atmospheres could not provide a relevant profile for the temperature. Further work would involve the creation of such synthetic atmospheres for the stations of the LaLiNet, the consolidation of the model by including the historic data measured after the year 2011, and the investigation of the factors that would reduce the efficiency of the model.

Acknowledgements

RB acknowledges funding by the FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), project 2012/00809-1.