Principal Component Analysis Approach to Evaluate Instrument Performances in Developing a Cost-Effective Reliable Instrument Network for Atmospheric Measurements

SIMONE LOLLI

Joint Center for Earth Systems Technology, University of Maryland, Baltimore County, Baltimore, and NASA Goddard Space Flight Center, Greenbelt, Maryland

PAOLO DI GIROLAMO

Scuola di Ingegneria, Università della Basilicata, Potenza, Italy

(Manuscript received 29 April 2015, in final form 7 July 2015)

ABSTRACT

Developing a reliable cost-effective instrument network for data measurement is a challenging task for agency decisionmakers. A simple way to fully characterize the performances of an instrument that also considers economical and operational factors—price, maintenance cost, lifetime, etc.—currently does not exist. Through principal component analysis, a method is developed to build a composite index that assigns a single score to each instrument, taking into account all the scientific, economic, and operational aspects. This index will then represent solid help in building and optimizing a cost-effective network, bridging the gap between two very different worlds: the scientific need for precision and economic constraints.

1. Introduction

Earth observation is becoming more and more important for a variety of reasons: among others, to collect important climate-related variable measurements, to check the health of our planet, for pure scientific research and for ground validation of satellite missions (Lolli et al. 2013a,b; Wang et al. 2013; Welton et al. 2002; Holben et al. 1998; Schafer et al. 2002). Since the late 1990s, thanks to the Internet, networks of instruments started to develop (Holben et al. 1998). Some of those networks were spontaneously constituted, as a result of the federation of research institutes owning the same instruments [Micropulse Lidar Network (MPLNET; Welton et al. 2002], or they were designed on purpose (Lund Myhre et al. 2012). However, designing and setting up a cost-effective instrument network to measure a geophysical variable is not an easy task. Often policymakers are facing objective difficulties, as a simple way to fully characterize the performance of an instrument that also considers economical or operational factors—price, maintenance cost, lifetime, etc.—does not exist. Agency decisionmakers then refer to scientists' advice in selecting an instrument instead of another to be deployed in future networks. Nevertheless, the choice can be suboptimal, as there are gaps and different visions between scientists and decisionmakers. The main purpose of this paper is to create a meaningful composite index to fully characterize the performances of an instrument with respect to a measured variable through the principal component analysis (PCA). This robust technique has been widely used in atmospheric physics, that is, to classify and group cirrus clouds (Dionisi et al. 2013) or clustering different air masses (Borchi and Marenco 2002). The composite index will then represent solid help in building and optimizing a cost-effective network, bridging the gap between two very different worlds: the scientific need for precision and economic constraints. The composite index, evaluated on existing networks, may be a useful tool to plan a future strategy, optimizing those stations equipped with lower-performing instruments.

In this paper we consider the created composite index to characterize the performances of instruments dedicated to the measurement of two specific atmospheric parameters,

Corresponding author address: Simone Lolli, NASA Goddard Space Flight Center, Code 612, Greenbelt, MD 20771.
E-mail: simone.lolli@nasa.gov

DOI: 10.1175/JTECH-D-15-0085.1

TABLE 1. (a) Variable indicator table for wind speed measurements. The values of the indicator variables are represented for each instrument (Instr). (b) As in (a), but adjusted for indices' sign discordance and normalized values. Values equal to unity indicate ideal instrument performance on that variable indicator.

a)

Indicators	Instr 1	Instr 2	Instr 3	Instr 4	Instr 5	Instr 6
Data reliability (%)	62	82	68	88	83	95
Dynamic range (m s^{-1})	30	80	50	40	65	120
Data availability (%)	40	70	80	65	82	95
Sensitivity (m s^{-1})	0.7	0.2	0.1	0.09	0.3	0.05
Price ($)	40 000	4000	7000	5000	9000	1500
Cost per hour ($)	15	2.6	10	3	4	1
Lifetime (yr)	0.5	2	1	4	2.2	10
Dimensions (dm^3 kg^{-1})	100	21	32	40	55	5
Daily file size (MB)	30	5	8	7.5	12	2
Operational temp range (°C)	3	30	15	22	10	95

b)

Indicators	Instr 1	Instr 2	Instr 3	Instr 4	Instr 5	Instr 6
Data reliability (%)	0.62	0.82	0.68	0.88	0.83	0.98
Dynamic range (m s^{-1})	0.25	0.66	0.416	0.33	0.54	0.8
Data availability (%)	0.4	0.7	0.8	0.65	0.82	0.95
Sensitivity (m s^{-1})	0.01	0.1	0.25	0.07	0.125	1
Price ($)	0.036	0.36	0.21	0.28	0.16	0.95
Cost per hour ($)	0.04	0.26	0.066	0.22	0.16	0.73
Lifetime (yr)	0.05	0.2	0.1	0.40	0.22	1
Dimensions (dm^3 kg^{-1})	0.02	0.095	0.062	0.05	0.036	0.94
Daily file size (MB)	0.11	0.66	0.416	0.44	0.28	0.73
Operational temp range (°C)	0.03	0.3	0.15	0.22	0.1	0.95

TABLE 2. Eigenvalues and rotated eigenvalues of the PCA performed on the indicator variables in Table 1.

	Initial eigenvalues			Eigenvalues of rotated extracted components		
Component	Total	Variance (%)	Cumulative (%)	Total	Variance (%)	Cumulative (%)
1	8.08	80.8	84.1	6.14	61.4	61.4
2	1.23	10.2	91	2.96	29.6	91
3	0.44	4.4	95.4			
4	0.33	3.3	98.7			
5	0.13	1.3	100			

wind speed and planetary boundary layer height, by considering a case based on real instruments available on the market.

2. Methodology

The composite index should be robust, reliable, universal, and relevant for policymakers (Greyling 2013). Each single instrument, with respect to a well-defined measurement parameter of interest (wind speed, temperature, etc.), can be represented as a point in hyperspace, where the dimensions are the different variable indicators—cost, lifetime, measurement accuracy, dimensions, power consumption, etc. Assigning a unique score to an instrument implies linking together all these different variables, reducing complexity and then dimensions. The index is built assuming a linear dependence among the different variable indicators. The main challenge is to determine an appropriate weighting method to apply to each single variable. Equal weighting is leading to an ineffective composite index (Greyling 2013). Among all the statistical methods, for our purposes, the PCA is best choice, as it avoids duplication of information using the orthogonal components and permits building a synthetic index, selecting only the meaningful variables (Greyling 2013; Somarriba and Pena 2009).

a. PCA description

PCA is a well-known multivariate statistical technique employed to simplify the complexity of a given initial dataset. Its main purpose is to reduce a certain number of original variables (representing the main characteristics of the analyzed phenomenon) into latent variables. The new variables are a linear combination of the old ones and are ordered by decreasing variance. The first new variable, projected on the first axis of the new Cartesian system, shows the higher variance of its component scores. The second variable, orthogonal to the first one, will be projected onto the second axis, and so on. Complexity reduction is reached by analyzing only the principal components (with respect to the variance).

Mathematically, PCA is defined in Eq. (1) as

$$\mathrm{PCA} = A \times X, \tag{1}$$

where $\mathbf{PCA} = \{\mathrm{PC}_i\}$ is the column vector representing the principal components, $\mathbf{A} = \{a_{i,j}\}$ is a matrix where $a_{i,j}$ represents the weight for the ith component and the jth indicator, and $\mathbf{X} = \{X_j\}$ is the set of variable indicators. The principal component weights given by covariance matrix are the eigenvectors.

b. Performing PCA

To perform PCA, the variable indicators should be sufficiently correlated. The Kaiser–Meyer–Olkin (KMO; Tabachnick and Fidell 2007) test is a measure of the sampling adequacy of the set of variable indicators, that is, the degree of correlation of the variables. The factor analysis is not appropriate for KMO test values less than 0.5, while values close to unity indicate that a factor analysis is highly effective.

The Bartlett's test of sphericity is a test performed on the correlation matrix to verify how close it is to the identity matrix: the closer the correlation matrix is to the identity matrix, the more the variable indicators are uncorrelated. Values of the test that are less than 0.05 indicate that a factor analysis can be effectively applied to the dataset. Once the correlation among the indicators is verified, a certain number of components (fewer than the initial variable indicators), which represent the data variability, are identified. If the variable indicators have different units, then the correlation matrix is a better choice with respect to the covariance matrix to extract the components. Using the normalized correlation matrix (centered matrix, where for each row the means are subtracted off and divided by standard deviation) guarantees that all the data have equal weighting, independent of their measurement units. There are several techniques to extract the appropriate number of components (Greyling 2013). In this paper we consider only those components that contribute (cumulatively) to a variance greater than 90%. The scree plot is an important visual tool to help in this choice. The scree plot depicts each eigenvalue versus the component number. The right number of components is

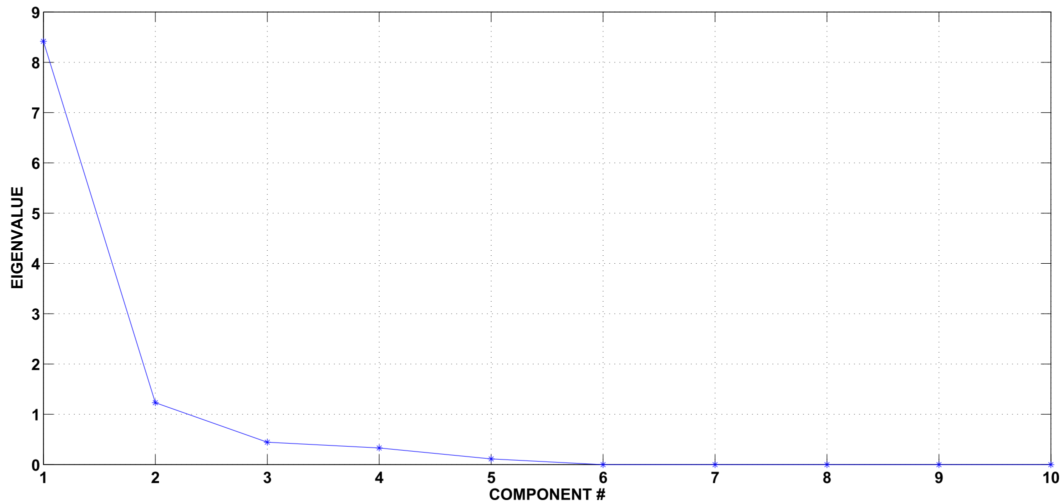

FIG. 1. Scree plot.

chosen at a discontinuity point (elbow on the plot) where the remaining eigenvalues are about all the same size and negligible. Once the components are extracted, the so-called factor loadings for each of the indicators on the components are calculated. The factor loadings represent the correlations between a variable indicator and the latent components. Interpretation of the results can be difficult, and then it is standard practice to rotate the components. Rotation helps to find a clear pattern of the factor loadings, minimizing the number of individual variables having a high loading on the same component (Nardo et al. 2005). An ideal situation is the one in which each variable indicator loads only onto one component. Nevertheless, the complexity of the variables is high, causing to load on more than one component.

c. *Composite index building*

Consider the situation represented by a government agency intending to set up a network of instruments to measure the wind speed at principal airports. We assume that six manufacturers available on the market produce devices capable of measuring this parameter. The variable indicators (10 in this example; Table 1) should be carefully chosen in relation to the wind speed measurement, as the result is dependent on this choice. If policymakers believe that instrument size and portability are not an issue, then the dimensions variable can be removed from Table 1. The result will be influenced by this choice (see also section 3). The KMO test, performed on indicators in Table 1, has a value of 0.78, while the Bartlett test returns a value smaller than 0.05, rejecting the null hypothesis. Both tests indicate that it is pertinent to perform PCA. For instrument 6 the indicator variables are set arbitrarily close to those of an ideal instrument, while for instrument 1 we set values representing overall bad performance. Table 1 highlights that some indicators are contradictory: an ideal instrument should cost 0 and have infinite lifetime. To build a robust index, all the variables are set in a way that higher values are close to the values of an ideal instrument. Table 2 shows the eigenvalues of the PCA performed on the indicator variables of Table 1. Kaiser's criterion (Kaiser 1958) or the eigenvalue rule and the scree plot guide in the selection for determining the significant number of components.

The eigenvalue, or the variance, for each principal component indicates the percentage variation of the principal component in the total dataset. According to Kaiser's criterion, we extract only those components with eigenvalues greater than 1.

In our example, the first two components explain more than 90% of the variation of the dataset, which is an acceptable value of the explained variance to be used in further analysis. Figure 1 shows the scree plot, and Fig. 2 shows the first and the second principal component factor loadings after varimax rotation (Kaiser 1958). The higher factor loadings in the first component (Fig. 2, top panel) are for those indicators more related to the economic and operational aspects of the instrument, such as price, lifetime, and cost per hour. The first component can then be labeled as the "economic and operational index." Differently, the second component shows higher loadings on those indicators more related to the measurement itself, such as accuracy, data availability, daily data size, and the measurement range. For this reason, this component is

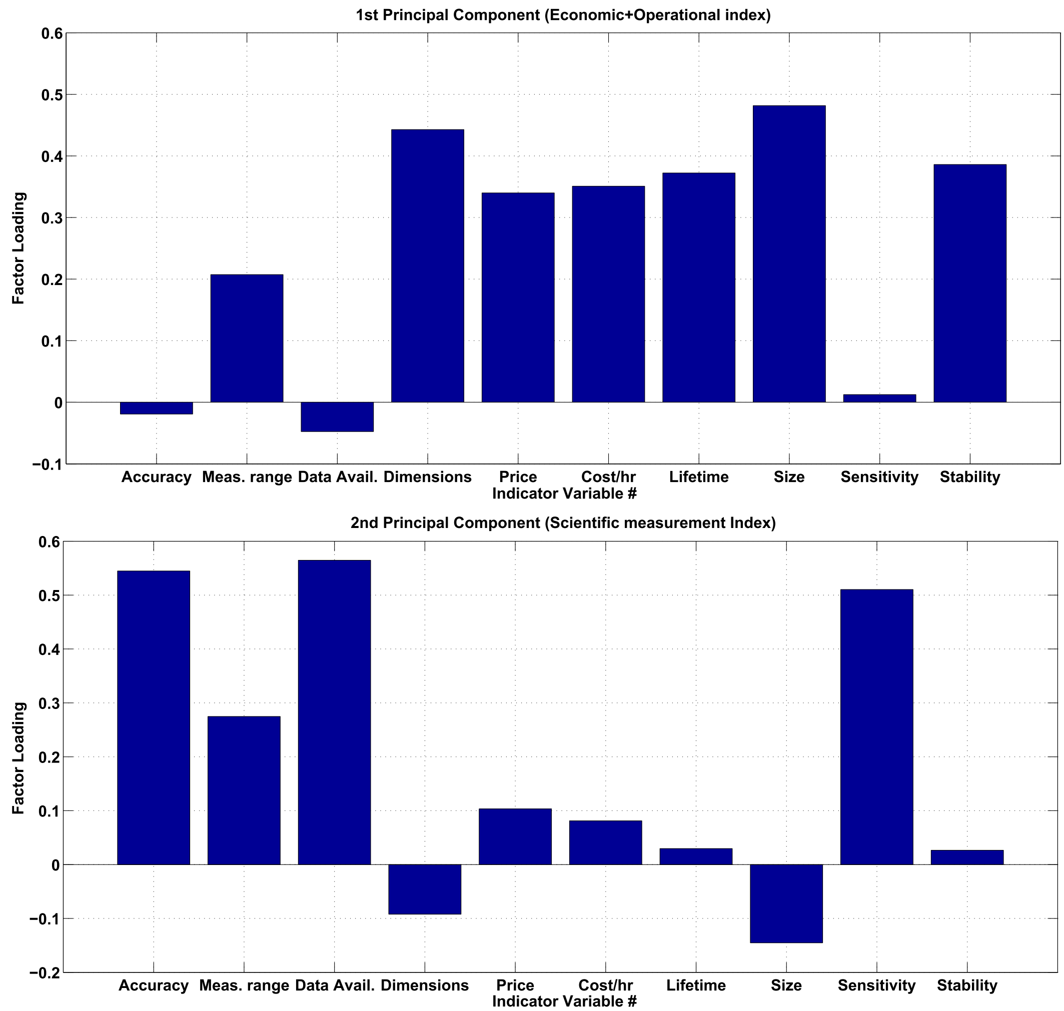

FIG. 2. (top) First and (bottom) second PC factor loadings.

labeled "scientific measurement index" (Fig. 2, bottom panel).

The composite index is built by grouping together the indicators into intermediate composite indices (Nicoletti et al. 2000), depending on the total number of components (2 in our example). Once the two intermediate indices are constructed, they are combined together. The best choice to build the intermediate indices is the geometric aggregation, as the single indicators have different units. Moreover, poor performance in some indicators cannot be compensated by high values in other indicators, as it happens with additive aggregation (Nardo et al. 2005). The composite index I is built as

$$I = \sum_{i=1}^{N=2} w_i \left(\prod_{q=1}^{10} x_q^{w_{q,i}} \right), \qquad (2)$$

where x_q represents the normalized values of the indicators for the considered instrument (columns in Table 1b), while each squared factor loading $w_{q,i}$ (squared values of factor loadings values in Fig. 2) represents the portion of the total unit variance (100%) of the indicator on the ith component. This means that the instrument "size" in the first component accounts for 23% in the intermediate composite index labeled economic and operational index, while "data availability" accounts for 33% in the scientific measurement index (values in Fig. 2). Each intermediate composite index contributes to the main composite index with a weight of w_i, equal to 0.67[6.14(6.14 + 2.96)] and 0.33[2.96(6.14 + 2.96)], respectively, calculated from the eigenvalues of Table 2.

The score of each single instrument is then shown as a percentage (Fig. 3), with 100% being the maximum score. To attain 100%, the instrument has to score the ideal value for each indicator in Table 2. As expected, instrument 6 shows the best performance, while instrument 1 has the lowest score. Each instrument can

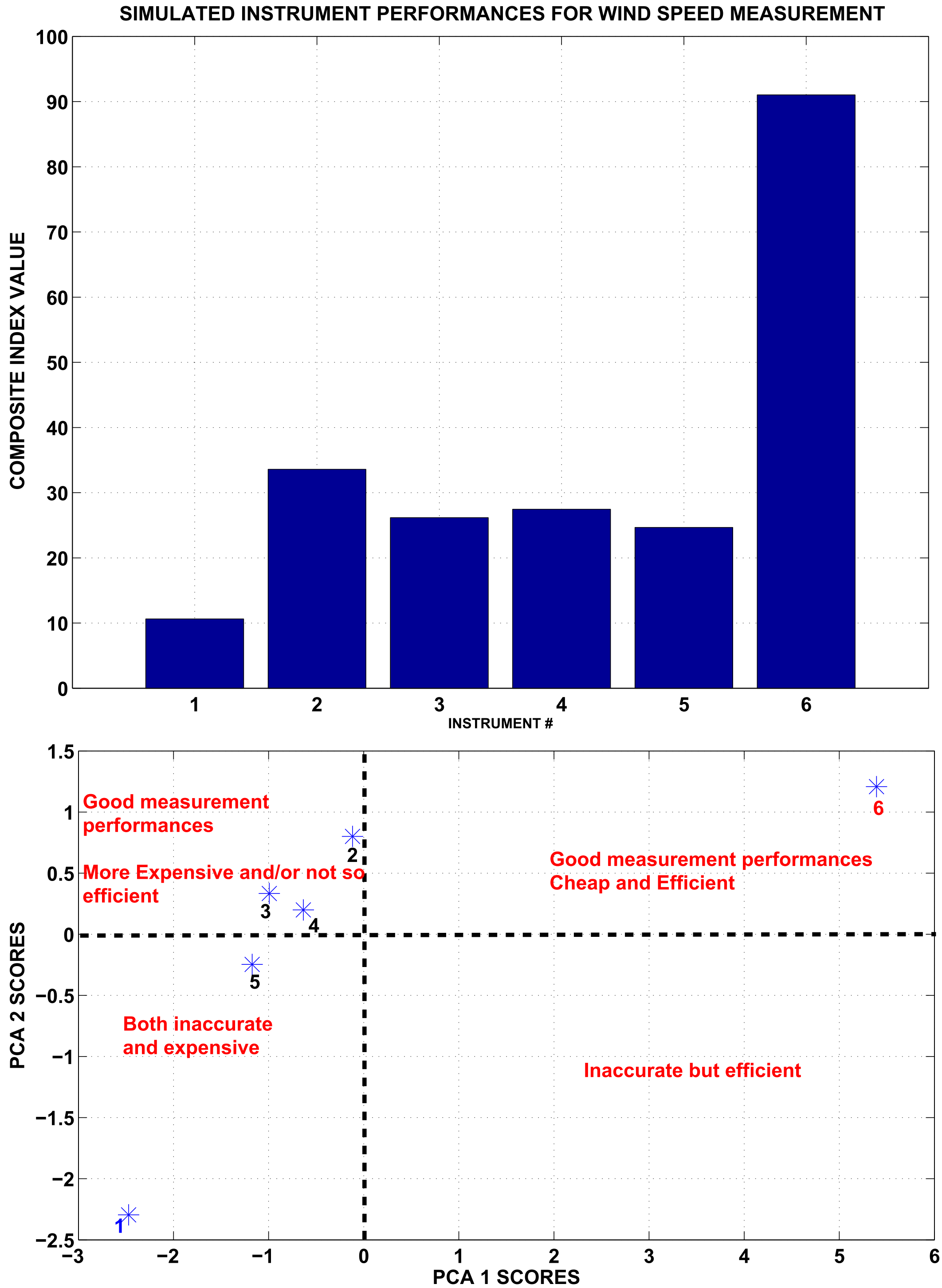

FIG. 3. (top) Scores for each of the different instruments. (bottom) Representation of instruments in the PCA bidimensional space. The number below each asterisk corresponds to the instrument indicator values in Table 2.

now be represented as a point in a reduced bidimensional space of the two extracted principal components. Recalling that the first component is related to the economical and operational aspects of the instrument, higher factor loadings are related to cheaper and more efficient instruments. The second component is related to the accuracy and other scientific aspects of the measurements. Higher values are then related to more

TABLE 3. Indices for instruments involved in ICOS campaign to measure PBL height.

Indicators	ALS300	CHM15K	CL31
Data reliability (%; Haeffelin et al. 2012)	67	62	58
Data availability (%)	95	86	78
Operational temp range (°C)	65	100	95
Price ($)	200 000	45 000	30 000
Cost × year ($)	1800	260	320
Lifetime (yr)	0.5	5	5
Dimensions ($dm^3\,kg^{-1}$)	2143	26 250	3820
Daily file size (MB)	400	50	80
Measurement resolution (m)	1.5	5	5

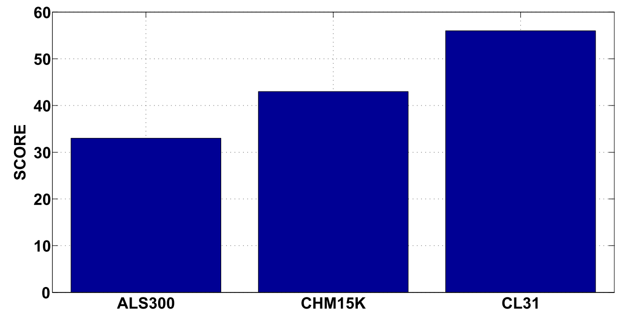

FIG. 4. Scores for instruments involved in the ICOS campaign in Trainou (2011) and Mace Head (2009).

accurate and precise measurements. Instrument 6 in the first quadrant has both positive indices: this means that the instrument is efficient, not costly with high measurement accuracy. The instruments in the second quadrant (instruments 2, 4, 5) still measure the wind speed accurately, but they are more expensive and/or not so efficient (i.e., thermal stability requirements are more strict, or the dimensions are bigger). Selecting an instrument from the third quadrant (instruments 1 and 3) is a bad choice, as both those instruments are not accurate and are expensive and inefficient. In the fourth quadrant the instruments have less accuracy, but are still not expensive and very efficient. An instrument lying in the third quadrant (poor choice) for wind measurement may be an excellent choice for the measurement of a different atmospheric parameter. This simulation of instrument performances with respect to wind speed measurement is also proof of the composite index's robustness. Increasing or decreasing the instrument performances, the composite index changes accordingly, reaching the score of 100 for an ideal instrument.

3. A real case: The PBL height measurement

Several projects have been started to coordinate the implementation of a network infrastructure to measure an important atmospheric state parameter: the planetary boundary layer (PBL) height at a synoptic scale (Milroy et al. 2012, 2011; Boquet et al. 2009). Tests and measurement campaigns were carried out in the frame of the Integrated Carbon Observation System (ICOS) program (Haeffelin et al. 2012) to assess which instrument is the most suitable to be deployed for a future operative network. The lidar/ceilometer PBL measurements were retrieved from two 1-month campaigns: at Trainou, France, in October 2011 and at Mace Head, Ireland, in June 2009 (Haeffelin et al. 2012). During these campaigns, three instruments were tested: the Aerosol Lidar System 300 (ALS300) lidar (Lolli et al. 2011), and the CHM 15K and CL31 (Munkel at al. 2007) ceilometers. During these campaigns, in situ radiosoundings, launched up to seven times per day, were taken as an independent reference for PBL height measurement. The method is applied to determine which instrument shows a higher composite index score related to the PBL measurements. The indicators are reported in Table 3 and the scores are plotted in Fig. 4. The instrument accuracy is evaluated through the number of consistent retrievals (the difference between the radisounding and the instrument is less than 300 m, expressed in percent at daytime only; Haeffelin et al. 2012). The CL31 has the best score (56). In the Haeffelin et al. (2012) study, only the scientific aspects are considered: the 9% drop in measurement quality with respect to ALS300 (58% of CL31 vs 67%) does not support choosing an instrument that is much more expensive with high maintenance costs and poorer lifetime. CHM 15K is a valid alternative.

Its lower score is mainly due to its huge dimensions and weight. The sensitivity to the initial variables plays a fundamental role and can drastically change the results. Running the analysis again without the dimension indicator put CHM 15K in first place. For this reason, the policymakers should exhaustively take into account all the different aspects carefully to select the most appropriate original variable indicators.

4. Conclusions

A method to evaluate different instrument performances with respect to a selected atmospheric parameter to be measured has been demonstrated. While the method in this paper is applied to evaluate the performances of instruments measuring a specific atmospheric parameter (wind speed/PBL height), the approach can be easily extended to evaluate the performances of instruments in general (with respect to a measurement parameter).

This method assigns a unique composite index, ranging from 0 to 100 (ideal instrument), to each instrument.

The performance index is a powerful tool that links together economical, operational, and measurement aspects related to the instrument versus the parameter to be measured.

The PCA method is applied to reduce the dimensions, finding correlations among the indicator variables. Any principal component links together linearly the different indicator variables, maximizing the variance of the dataset, in a descending order. From each extracted component then an intermediate composite index is constructed. The aggregated instrument performance index is then computed, adding the weighted scores (that depend on component variance) of each selected principal component.

This method permits the evaluation of an instrument with respect to all different aspects—that is, economical, scientific, and operational—helping policymakers to develop or evaluate a cost-effective reliable instrumental network.

As an example, the proposed method is applied to instruments dedicated to the measurement of two specific atmospheric parameters: the wind speed and the PBL height—in this latter case, considering a case based on real instruments available on the market. The initial variable indicators play a fundamental role in the result and should be carefully selected depending on project needs.

REFERENCES

Boquet, M., J. P. Cariou, L. Sauvage, S. Lolli, R. Parmentier, and S. Loaec, 2009: Atmospheric Boundary Layer and Clouds wind speed profile measurements with the new compact long range wind Lidar WindCube(TM) WLS70. *Geophysical Research Abstracts,* Vol. 11, Abstract EGU2009-10541. [Available online at http://meetingorganizer.copernicus.org/EGU2009/EGU2009-10541.pdf.]

Borchi, F., and A. Marenco, 2002: Discrimination of air masses near the extratropical tropopause by multivariate analyses from MOZAIC data. *Atmos. Environ.,* **36,** 1123–1135, doi:10.1016/S1352-2310(01)00576-3.

Dionisi, D., P. Keckhut, G. L. Liberti, F. Cardillo, and F. Congeduti, 2013: Midlatitude cirrus classification at Rome Tor Vergata through a multichannel Raman–Mie–Rayleigh lidar. *Atmos. Chem. Phys.,* **13,** 11 853–11 868, doi:10.5194/acp-13-11853-2013.

Greyling, T. A., 2013: A composite index of quality of life for the Gauteng city-region: A principal component analysis approach. GCRO Reports, Occasional Paper 7, 47 pp. [Available online at http://www.gcro.ac.za/sites/default/files/reports/occasional_paper_7.pdf.]

Haeffelin, M., and Coauthors, 2012: Evaluation of mixing-height retrievals from automatic profiling lidars and ceilometers in view of future integrated networks in Europe. *Bound.-Layer Meteor.,* **143,** 49–75, doi:10.1007/s10546-011-9643-z.

Holben, B. N., and Coauthors, 1998: AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sens. Environ.,* **66,** 1–16, doi:10.1016/S0034-4257(98)00031-5.

Kaiser, H. F., 1958: The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* **23,** 187–200, doi:10.1007/BF02289233.

Lolli, S., L. Sauvage, S. Loaec, and M. Lardier, 2011: EZ Lidar™: A new compact autonomous eye-safe scanning aerosol Lidar for extinction measurements and PBL height detection. Validation of the performances against other instruments and intercomparison campaigns. *Opt. Pura Apl.,* **44,** 33–41.

——, A. Delaval, C. Loth, A. Garnier, and P. H. Flamant, 2013a: 0.355-micrometer direct detection wind lidar under testing during a field campaign in consideration of ESA's ADM-Aeolus mission. *Atmos. Meas. Tech.,* **6,** 3349–3358, doi:10.5194/amt-6-3349-2013.

——, E. J. Welton, and J. R. Campbell, 2013b: Evaluating light rain drop size estimates from multiwavelength micropulse lidar network profiling. *J. Atmos. Oceanic Technol.,* **30,** 2798–2807, doi:10.1175/JTECH-D-13-00062.1.

Lund Myhre, C., and Coauthors, 2012: Recommendations for a composite surface-based aerosol network. WMO GAW Rep. 207, 58 pp. [Available online at http://www.wmo.int/pages/prog/arep/gaw/documents/Final_GAW_207_WEB.pdf.]

Milroy, C., and Coauthors, 2011: On the ability of pseudo-operational ground-based light detection and ranging (LIDAR) sensors to determine boundary-layer structure: Intercomparison and comparison with in-situ radiosounding. *Atmos. Meas. Tech. Discuss.,* **4,** 563–597, doi:10.5194/amtd-4-563-2011.

——, and Coauthors, 2012: An assessment of pseudo-operational ground-based light detection and ranging sensors to determine the boundary-layer structure in the coastal atmosphere. *Adv. Meteor.,* **2012,** 929080, doi:10.1155/2012/929080.

Munkel, C., N. Eresmaa, J. Rasanen, and A. Karppinen, 2007: Retrieval of mixing height and dust concentration with lidar ceilometer. *Bound.-Layer Meteor.,* **124,** 117–128, doi:10.1007/s10546-006-9103-3.

Nardo, M., M. Saisana, A. Saltelli, S. Tarantola, A. Hoffman, and E. Giovannini, 2005: Handbook on constructing composite indicators: Methodology and user guide. OECD Statistics Working Papers 2005/03, 109 pp., doi:10.1787/533411815016.

Nicoletti, G., S. Scarpetta, and O. Boylaud, 2000: Summary indicators of product market regulation with an extension to employment protection legislation. OECD Economics Department Working Papers 226, 87 pp., doi:10.1787/215182844604.

Schafer, J. S., B. N. Holben, T. F. Eck, M. A. Yamasoe, and P. Artaxo, 2002: Atmospheric effects on insolation in the Brazilian Amazon: Observed modification of solar radiation by clouds and smoke and derived single scattering albedo of fire aerosols. *J. Geophys. Res.,* **107,** 8074, doi:10.1029/2001JD000428.

Somarriba, N., and B. Pena, 2009: Synthetic indicators of quality of life in Europe. *Soc. Indic. Res.,* **94,** 115–133, doi:10.1007/s11205-008-9356-y.

Tabachnick, B. G., and L. S. Fidell, 2007: *Using Multivariate Statistics.* 5th ed. Pearson Education, 980 pp.

Wang, S. H., and Coauthors, 2013: Origin, transport, and vertical distribution of atmospheric pollutants over the northern South China Sea during the 7-SEAS/Dongsha Experiment. *Atmos. Environ.,* **78,** 124–133, doi:10.1016/j.atmosenv.2012.11.013.

Welton, E. J., and Coauthors, 2002: Measurements of aerosol vertical profiles and optical properties during INDOEX 1999 using micropulse lidars. *J. Geophys. Res.,* **107,** 8019, doi:10.1029/2000JD000038.