Publication Year	2015
Acceptance in OA@INAF	2020-05-12T07:53:31Z
Title	EChO spectra and stellar activity - I. Correcting the infrared signal using simultaneous optical spectroscopy
Authors	MICELA, Giuseppina
DOI	10.1007/s10686-014-9430-1
Handle	http://hdl.handle.net/20.500.12386/24712
Journal	EXPERIMENTAL ASTRONOMY
Number	40

Experimental Astronomy

EChO spectra and stellar activity - I. Correcting the infrared signal using simultaneous optical spectroscopy

--Manuscript Draft--

Manuscript Number:	EXPA-D-14-00015R2
Full Title:	EChO spectra and stellar activity - I. Correcting the infrared signal using simultaneous optical spectroscopy
Article Type:	SI: TI EchO
Keywords:	Exoplanets; Atmosphere; Activity; EChO
Corresponding Author:	Giuseppina Micela, Ph.D. INAF Palermo, ITALY
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	INAF
Corresponding Author's Secondary Institution:	
First Author:	Giuseppina Micela, Ph.D.
First Author Secondary Information:	
Order of Authors:	Giuseppina Micela, Ph.D.
Order of Authors Secondary Information:	
Abstract:	Stellar activity is the major astrophysical limiting factor for the study of planetary atmospheres. Its variability and spectral characteristics may affect the extraction of the planetary signal even for moderately active stars. A technique based on spectral change in the visible band was developed to estimate the effects in the infrared due to star activity. This method has been purposely developed for the EChO mission which had the crucial characteristics of monitoring simultaneously a broadband from visible to infrared. Thanks to this capability the optical spectrum, whose variations are mainly due to stellar activity, has been used as in an instantaneous calibrator to correct the infrared spectrum. The technique is based on principal component analysis which significantly reduces the dimensionality of the spectra. The method was tested on a set of simulations with realistic photon noise. It can be generalized to any chromatic variability effects provided that optical and infrared variations are correlated.
Response to Reviewers:	I thanks the referee for his/her comments on the new version of the paper. I expanded the final section to better discuss the possible developments of the work, as suggested by the referee All the changes suggested by the referee have been implemented

Noname manuscript No.
(will be inserted by the editor)

EChO spectra and stellar activity – I. Correcting the infrared signal using simultaneous optical spectroscopy

Giuseppina Micela

Received: date / Accepted: date

Abstract Stellar activity is the major astrophysical limiting factor for the study of planetary atmospheres. Its variability and spectral characteristics may affect the extraction of the planetary signal even for moderately active stars. A technique based on spectral change in the visible band was developed to estimate the effects in the infrared due to star activity. This method has been purposely developed for the EChO mission which had the crucial characteristics of monitoring simultaneously a broadband from visible to infrared. Thanks to this capability the optical spectrum, whose variations are mainly due to stellar activity, has been used as in an instantaneous calibrator to correct the infrared spectrum. The technique is based on principal component analysis which significantly reduces the dimensionality of the spectra. The method was tested on a set of simulations with realistic photon noise. It can be generalized to any chromatic variability effects provided that optical and infrared variations are correlated.

Keywords Exoplanets · Atmosphere · Activity · EChO

1 Introduction

The EChO (Exoplanet Characterization Observatory) mission has been proposed to the ESA Cosmic Vision program as a medium-size mission (Tinetti et al. 2012). The scientific objective of EChO is spectral observations of a portfolio of exo-planetary atmospheres in order to determine their physical conditions. The observational methodology is based on differential spectroscopy

---

G. Micela
INAF - Osservatorio Astronomico di Palermo
Piazza del Parlamento 1, 90134 Palermo, Italy
Tel.: +39-091233111
Fax: +39-091233444
E-mail: giusi@astropa.inaf.it

in and out of transit, which leaves a signal of the order of 10^{-4} of the stellar one. Such a tiny signal requires a ”perfect” knowledge of the stellar spectrum, which has to be known (at least) at the same level of the planetary signal.

Beside the instrumental and poissonian fluctuations, the variations due to stellar activity are the main sources of astrophysical noise and may affect the measurements of planetary atmospheres, specially during the primary transit. The hypothesis that justifies the subtraction of the spectra taken in and out of the transit is that the planet occults a portion of the stellar surface that has the same flux of the average stellar surface, thus reducing the transit to a purely geometrical effect. Unfortunately activity is a coloured phenomenon, therefore this hypothesis does not hold true for active stars. Several phenomena may affect the transit observations, including granulation, flares, pulsations, spots, and faculae. Granulation cell size is of the order of 10^8cm (between approximately 5×10^8 and 0.3×10^8cm from mid-F to dM stars, Beeck et al. 2013) much smaller than the size of the transiting planets targets of EChO. Therefore granulation is not a problem for planetary observations since its effect is averaged on the area occulted by the planet. This is true also for atmospheres, whose observation is integrated in a circular annulus having the internal radius larger than the typical scale of granulation. Pulsations are not stochastic phenomena, therefore they may be modeled using the out of transit observations to correct those made during the transit. Flares are stochastic events that may last from few to several minutes, therefore they may take place on the same timescale as transits. Their occurrence may be monitored by specific indicators (i.e. Hα emission), though flares are difficult to correct.

Faculae have a typical temperature close to that of the unperturbed surface, therefore their effects on planetary radius determination is minimal. Spots, on the contrary may have a size comparable with that of the projected area of the planets and have a temperature contrast of several hundreds of degrees, which produces an important effect on the planetary measure. Their timescale evolution is on the stellar rotation time and may affect the comparison between the measures taken during and out of the transit, as well as from different transits.

In fact the presence of cool spots modifies the transit depth in spectral observations taken during the primary transit inducing an error in the planetary radius determination. In the case of not occulted spot the net effect is an overestimated planetary radius since the flux of the eclipsed stellar surface is higher than the average flux measured out of transit (lowered by the presence of a cool stellar spot). If a spot is occulted by the planet a typical bump will occur during the transit since the flux of the occulted spot is lower than the rest of the surface. In the latter case the resulting radius will be underestimated.

Furthermore, since activity is a coloured phenomenon, it affects the extracted atmospheric spectrum or may even mimic the existence of a planetary atmosphere, if not properly accounted for. Analysis of the variation timescale will help in disentangling the origin of the observed variability, and a careful spectral analysis will allow us to correct for activity-induced spectrum distortion.

A method is herewith presented to correct the stellar spectrum for activity effects that takes advantage of the large simultaneous wavelength EChO coverage and moderate spectral resolution. The method uses the shape of the spectrum in the visible band to predict and correct the shape of the spectrum in the infrared.

In order to calibrate this method, a set of spectra is needed to cover as much as possible the range of possible activity levels; the ideal set will be given by the out of transits spectra collected by EChO itself. This will be the best data set to adapt the method to a given star, with the reasonable assumption that stellar spectra during the transits do not behave in a different way than out of transit. The calibration of the method will be easily refined and updated by accumulating data during the EChO lifetime.

Today, in the absence of EChO observed spectra or an equivalent set of data, the method has been developed and tested on a set of simplified models of spotted stars, with the idea that it may be generalized to any kind of model or data. The only assumption behind it, is that the spectral variations in the NIR band are correlated to those in the visible, therefore the proposed approach can be applied to any kind of variations with this property.

In a companion paper (Scandariato and Micela 2014) a different analysis has been presented, specifically developed for the case of dM stars, based on a large set of observed medium resolution spectra.

2 The approach

Activity is prominent specially in the blue spectral window and has smaller effects (even 4-5 times smaller) in the infrared band (i.e. Ballerini et al. 2012) where the planetary spectrum is concentrated. Advantage of the broad band coverage of EChO can be taken to correct the infrared spectrum using the visible part of the spectrum. In practice the latter will be used as an instantaneous calibrator to correct the spectrum in the IR band in order to recover the planetary signal uncontaminated by stellar activity. This approach can be justified by the fact that the spectral variation observed in the optical band may be attributed almost completely to the star activity whereas the planet contribution is much smaller. In fact, while the effect of a stellar spot having properties similar to those observed in the active Sun, is 0.7 -1% of the observed flux in V band, depending on the spectral type (decreasing down to 0.3 - 0.6% in K band, Ballerini et al. 2012), the effect of a planetary atmosphere during a primary transit is not higher than 10^{-4} of the stellar flux (a hot Jupiter with an extended atmosphere).

Having a sufficient number of broad band spectra of a star observed at different levels of activity, it is possible to calibrate the technique specifically for that star. EChO will provide us with such spectra. In the meantime, this approach can be developed and tested on a grid of purposely built models. The following sections present the models and the method to correct spectra which will be obtained by ECHO. The method will be tested through simulations of a set of spectra of *real* stars with a given SNR. These tests will be used to verify the capability of the method to recover the infrared spectrum, based on the visible band.

2.1 Active star models

This section presents how a set of stellar spectra at several levels of activity is modeled. Although reasonably realistic, the model claims no completeness (see Ribas et al. 2014, for more complete active star models), since its aim is the development and calibration of the method. While the general behavior will reflect that of the real spectra, there is no claim to reproduce the spot details. The models of active stars are based on the technique used by Ballerini et al. (2012) to estimate multi-band photometric effects of stellar activity, as extended to spectroscopy.

The modeling of the stellar activity is based on the assumption that activity is dominated by spots at T=$T_{\rm s}$ covering a fraction f (filling factor) of the stellar surface

The stellar flux in the presence of spots can be expressed as in Eq. (1), where other effects are neglected, including the limb darkening (see Ballerini et al. 2012 for a discussion on these effects) and plages.

$$F_\lambda(\mathrm{obs}) = (1-f)\times F_\lambda(T_*) + f\times F_\lambda(T_{\rm s}) \tag{1}$$

T_* is the effective stellar temperature for a given spectral type star in the absence of spots, $T_{\rm s}$ is the spot temperature, smaller than T_* and f is the filling factor, i.e. the fraction of the projected stellar disk covered by the spot(s). Fluxes $F_\lambda(T)$ are taken from models of Allard et al. (2011).

In the following the flux refers to the (0.55 -1.0 μm) band. The blue cutoff is imposed by the instrument requirement while the red cutoff is a compromise between having a band sufficiently broad and maintaining a small as possible overlap with the band ”of interest” for planetary atmosphere. The EChO goal to extend the blue band to 0.4μm would be a further advantage since the stellar activity is specially evident in the bluest part of the spectrum.

For the present application for a given stellar photospheric temperature a grid of spotted stars has been created as in Eq. (1). The filling factor (f) is defined in the interval f = [0.001 − 0.101] and has a step of $\mathit{\Delta} f = 0.001$. $T_{\rm s}$ spans the $[T_* - 1500 - T_*]$ interval with a step of 100 K. Models from Allard et al. (2013), BT- Settl (http://phoenix.ens-lyon.fr/Grids/BT-Settl/) have been used . Spectra corresponding to temperatures not present in the Phoenix library are generated by interpolating spectra with the closest temperatures.

Spectra are degraded to a resolution R=300 to match the EChO resolution, by filtering the original high resolution spectra with a Gaussian with $\sigma = \lambda/300$.

2.2 Principal Components

The proposed method is based on the multivariate statistical technique known as Principal Component Analysis (PCA, Jollife 1986). Principal Components are linear combinations of the original variables rotated in a space of ortoghonal variables. Since the new variables are chosen with the goal of maximizing the variance of first components, it is possible, establishing a threshold to the explained variance, to reduce, even significantly, the dimensionality of the space. The technique reduces the noise, normally represented by the less significant components, and takes into account for possible dependencies among the original variables. The technique of PCA has been already used in an astrophysical context to classify spectra (e.g. Collura et al. 1995; Connolly et al. 1995, Folkes et al. 1999, Hojnacki et al. 2008).

In the present application the "observations" are a set of spectra representing all the possible spectra of the star at different levels of activity. The original variables are the set of fluxes observed in each spectral bin (resolution element), therefore the principal components are a linear combination of these fluxes. The effectiveness of the method is strongly dependent on the choice of the set of "observations" used to derive the principal components. The best choice should be a set of truly observed spectra of the target star at several activity levels, ideally taken during an entire activity cycle. This implies the need for a high precision spectroscopic, high S/N observing campaign on the EChO targets before launch. Alternatively, the spectra obtained by EChO out of the transits may give a set of data well suited for calibrating the method. In order to test and calibrate the proposed method, the spectral grid is obtained as explained above. Even if the present application is based on a model, the method could be applied in a completely non-parametric way if calibration relies on observed spectra. It must be stressed that the goal of the present work is not retrieving the spot parameters but developing a method able to estimate the spectral change in the infrared band, from the change observed in the visible.

In order to determine the principal components of spectra the wavelength range [0.55-1.0]μm, was used, normalized to its integral flux to detect shape variations taking in no account absolute flux. In this implementation principal components are derived by centering the variables (the flux in each spectral bin) to their means. Normalization with respect to standard deviation, as commonly applied in principal component analysis, is not done, since it is intended to maintain the information on flux range in the different parts of the spectrum.

In all the explored cases two components are sufficient to explain at least 90% of the variance. The first component is linked to the slope of the spectrum whereas the higher order components are related to features of the spectrum. Examples of the two first components for three representative cases are reported in Figure 1.

The two components are directly connected with the model parameters, temperature and filling factor of the spot. These relations are shown for the

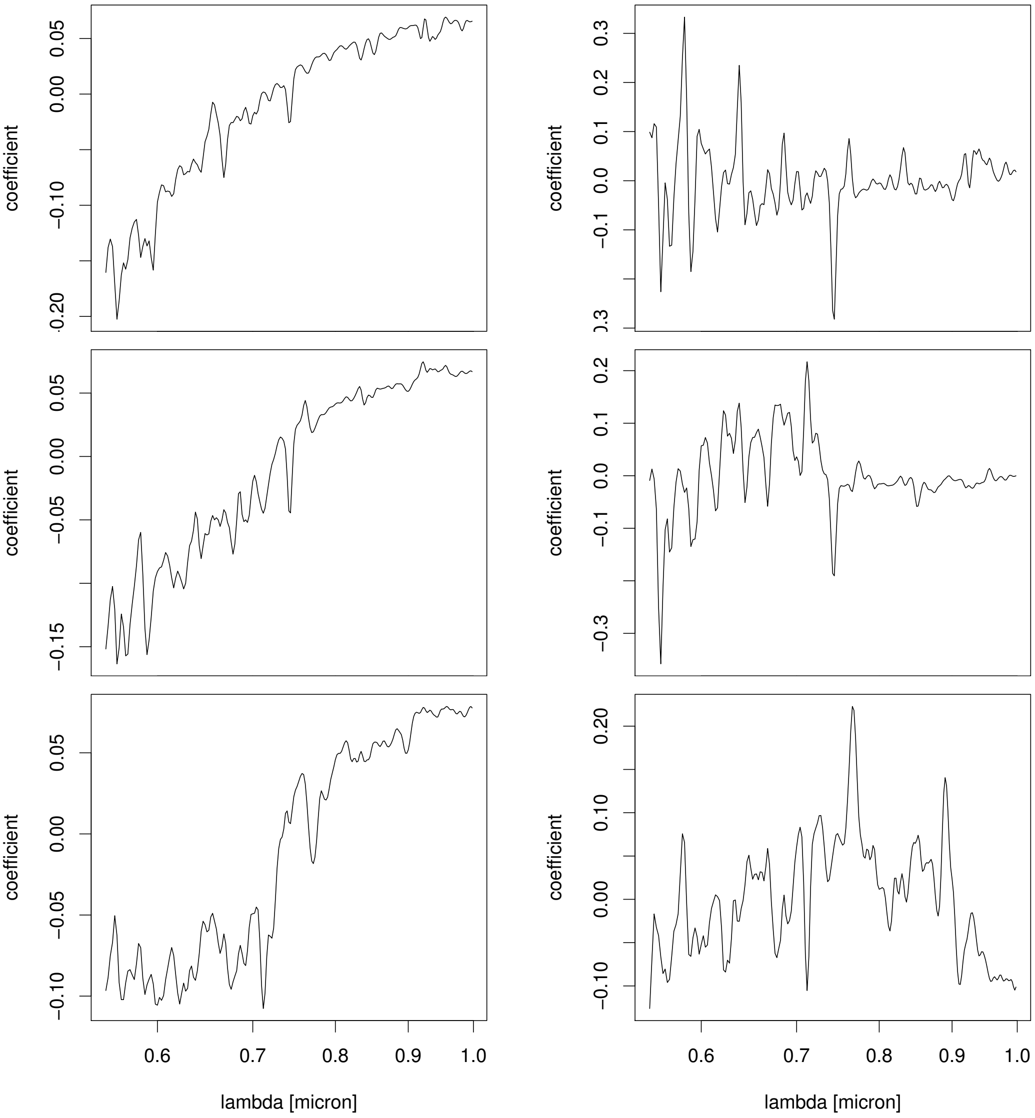

Fig. 1 First two components (left-right) for T_*=6000 K (top); T_*=5200 K (middle); T_*=4200 K (bottom)

case with $T_* = 5200$ K in Figure 2. The two other studied cases ($T_* = 6000$K and $T_* = 4200$K) present qualitatively similar relations.

The first component alone is not sufficient to determine univocally the spectral shape, here parameterized by both the spot temperature and the filling factor, and presents a degeneracy in the two parameters. By combining the first

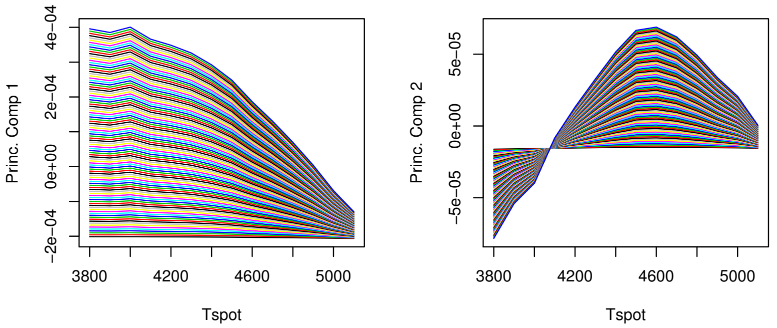

Fig. 2 Dependence of first (left) and second (right) component on spot temperature for a star with T_{*}=5200 K. Different curve lines correspond to different filling factors, the highest in the upper part of the plot.

Table 1 Summary of parameters of the grid used to derive the Principal Components

	G0	mid-G	K5
T_{*}	6000	5250	4200
T_{s}	4500-5900	3700-5100	2700-4100
f	0.001 - 0.101	0.001 - 0.101	0.001 - 0.101

two components it is possible to break this degeneracy and recover the original spectrum. Projecting an observed spectrum on the principal component space should enable the spot parameters to be derived and therefore the spectrum of the original grid that better matches the visible part of the observed spectrum to be identified.The infrared part of this spectrum will be the best estimate of the ”true”infrared spectrum taking into account stellar activity. The plane defined by the two components for the case with T_{*}=5200 K is shown in Figure 3.

The following analysis has been performed assuming the three stellar temperatures reported in Table 1.

3 Simulations

In order to test the method capability to infer the infrared spectrum in the presence of spots a set of star-spot combinations has been simulated. In agreement with the EChO requirement on signal to noise in the visible band SNR=200 per resolution element at 0.9 μm was assumed(considering only photon noise, and scaling the SNR at the other wavelengths assuming a flat instrument response). For each combination of T_{s} and f, 1000 simulations were generated and the resulting spectra projected in the principal component space. In order to recover the spot parameters and to save computational

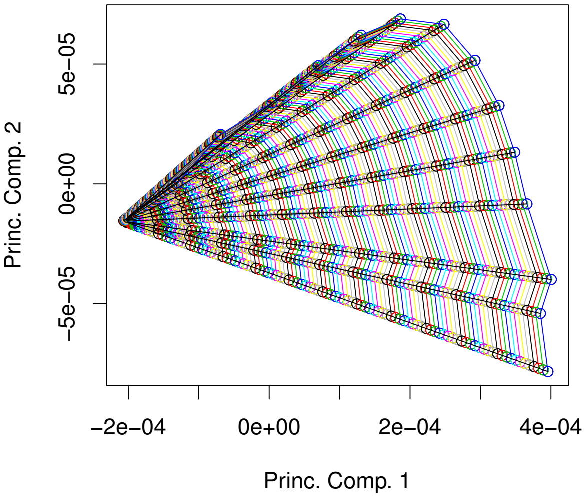

Fig. 3 Relation between the first two components for T_*=5200 K. Different lines correspond to different filling factors (the highest in the right of the plot).

time a fit was not performed, but, for simplicity, the closest point of the grid was chosen. The resulting sensitivity is therefore limited by the resolution of the grid itself: Δf=0.001 and ΔT=100 K, respectively. A more sophisticated fitting procedure may be performed to obtain more finer tuned results. As expected the best results are obtained for large filling factors and low $T_{\rm s}$.

Since it was not intended to model the star in detail whereas it was desired to correct for the activity or any other chromatic variations, retrieving precisely the filling factor and spot temperature was not necessary. Therefore it was assessed which spectrum of the grid, as represented by the first two principal components, was more similar to the true spectrum referred to as the ”best fitted” one. The fit procedure uses the visible spectrum alone and, assuming that the variations in the visible band are correlated with those at longer wavelengths, the NIR part of the best fitted spectrum is the best prediction for the infrared spectrum. This approach can be used on any grid independently of the origin of the spectra on which is based, models or observations. The only assumption behind the method is that the visible and infrared variations are correlated.

The capability of the method was tested on a few cases by comparing the original spectral ”distortion”, defined as:

$$\frac{F_\lambda^{*,u}}{F_\lambda^{*,sp}} \tag{2}$$

and the residual distortion after applying the above procedure, computed as:

$$\frac{F_\lambda^{*,best}}{F_\lambda^{*,sp}} \tag{3}$$

where $F_\lambda^{*,u}$ is the spectrum of the star neglecting the presence of the spot, $F_\lambda^{*,sp}$ is the "true" stellar spectrum considering the spot (i.e. the simulated spectrum), and $F_\lambda^{*,best}$ is the spectrum estimated as a result of this method (the "best fit" above).

Fig. 4 shows the result for three cases (stars with T_*=6000, 5200, 4200; ΔT=700K; f=0.01). The solid black line in the panel marks the uncorrected spectral distortion, whereas the red lines mark the residual distortion obtained by correcting the spectrum with the NIR best fitted spectrum derived from the above procedure. The lines are the 25% and 75% percentiles obtained by 1000 simulations with S/N=200 per resolution element in the visible band.

The figure shows that the spectral distortion may be reduced significantly, with the simple assumption that variability in the visible and NIR are correlated, independently from the detailed origin of the variations.

4 Summary and conclusions

A method to correct the infrared spectrum for the effects due to spots, taking advantage of simultaneous observation of the visible spectrum was developed. The method projects the observed visible spectrum on a space defined by the principal components computed from a grid of spectra. While in the ideal case the grid may be built from a series of observed spectra of the same star observed at different levels of activity at several time scales, the approach herewith discussed created the grid from a set of simplified models of spotted stars.

The first two components are sufficient to describe most of the variance. These two components are related to the spot parameters by identifying the combination of spot temperature and filling factor. In order to test the method, a set of simulated spectra was generated, assuming SNR = 200 per resolution element in the visible band and it was verified that in all explored cases the NIR spectrum was reproduced pretty well , assuming as corrected infrared spectrum, that of the best fitted visible spectrum.

The main assumption behind this technique is that the activity-induced variations in the NIR band are correlated with those observed in the visible. This is a reasonable assumption consistent with what is known on stellar activity independently of the details (spot, plages, limb darkening, large spots vs. several small spots, granulation, etc) that should make the method robust. A test on the validity of this assumption may be based on a campaign of mutiwavelength simultaneous observations, provided the needed instrumental precision in all bands. In principle a set of spectral windows, particularly sensitive to the activity can be used, replacing the role of the visible spectra. The

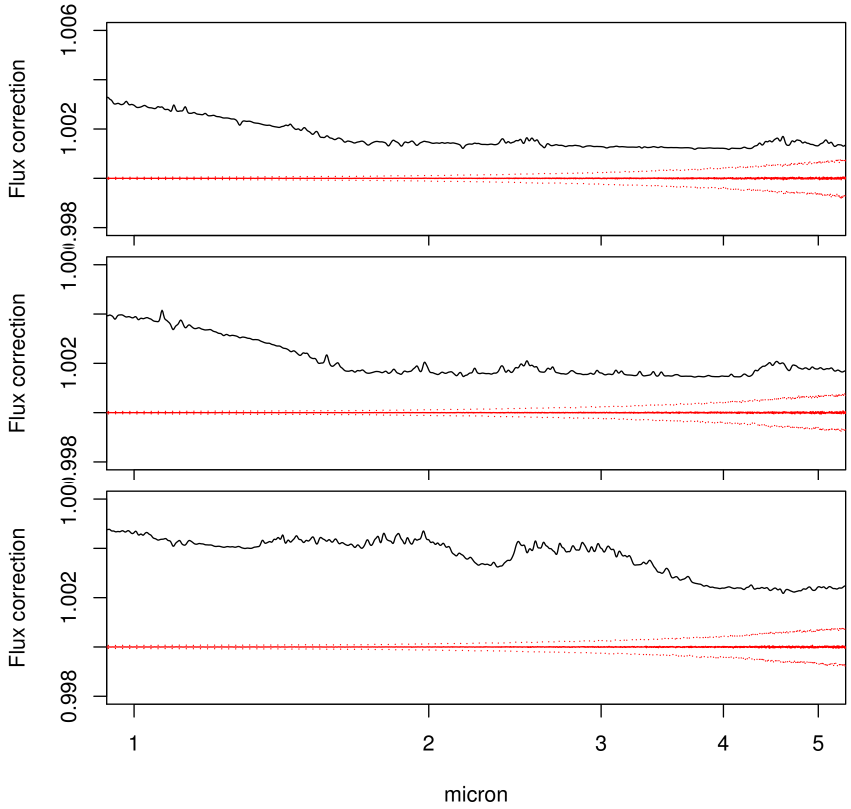

Fig. 4 Distortion of the near infrared spectrum for three cases (stars with T_*=6000, 5200, 4200; ΔT=700K; f=0.01). The solid black line in the panel marks the spectral distortion without any correction. The red lines correspond to the 25, 50, 75th percentiles of 1000 simulations

most obvious ones include the chromospheric features such as Hα or CaII . However in this case it is needed to note that the effects to be corrected are generated at several layers of the atmosphere, not only chromosphere.

In this paper a grid of modeled spectra was used for simplicity. In order to avoid all the uncertainties present in models, a grid may be purposely built for each star, from truly observed spectra obtained at different levels of activity. EChO itself will provide such spectra during its life from the out of transit observations. It is also possible to organize focused ground observational campaigns on EChO targets before launch. This will provide an initial grid that will be updated as new EChO observations are obtained. Obviously, the finer the grid is, the better will be the accuracy achieved. It should be stressed that the relevant properties are not the absolute flux of the spectra, but their shape variations and how the variations in visible band are correlated with those in the infrared.

Acknowledgements:

Partial support by the ASI/INAF contract I/022/12/0 is acknowledged. We wish to thank Giovanna Tinetti and Marcell Tessenyi for their useful discussions, Donatella Randazzo for her help in revising the English language, and the referee for his/her comments which have greatly improved this paper.

References

1. Allard, F., Homeier, D., Freytag, B.: Model Atmospheres From Very Low Mass Stars to Brown Dwarfs. 16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, 448, 91 (2011)
2. Ballerini, P., Micela, G., Lanza, A. F., Pagano, I., A&A, 539, A140 (2012)
3. Beeck B., Cameron R. H., Reiners, A.; Schüssler, M., , A&A. 558, 49 (2013)
4. Collura A., Micela G., Sciortino S., Harnden F. R., Jr., Rosner R., ApJ, 446, 108 (1995)
5. Connolly A. J., et al., AJ, 110, 1071 (1995)
6. Folkes S., et al., MNRAS, 308, 459 (1999)
7. Hojnacki S. M., Micela G., Lalonde S. M., Feigelson E. D., Kastner J. H., StMet, 5, 350 (2008)
8. Jollife I.T., Principal Components (New York: Wiley) (1986)
9. Ribas, I., EspAstr. THIS VOLUME (2014)
10. Scandariato G., Micela G., ExpAstr. THIS VOLUME (2014)
11. TInetti G. et al. , ExpAstr, 34, 311 (2012)