A&A 600, A45 (2017)
DOI: 10.1051/0004-6361/201629446

Astronomy & Astrophysics

Uncertainties in the solar photospheric oxygen abundance

M. Cubas Armas[1,2], A. Asensio Ramos[1,2], and H. Socas-Navarro[1,2]

[1] Instituto de Astrofísica de Canarias (IAC), Avda Vía Láctea S/N, 38200 La Laguna, Tenerife, Spain
e-mail: mcubas@iac.es
[2] Departamento de Astrofísica, Universidad de La Laguna, 38205 La Laguna, Tenerife, Spain

Received 1 August 2016 / Accepted 18 January 2017

ABSTRACT

Aims. The purpose of this work is to better understand the confidence limits of the photospheric solar oxygen abundance derived from three-dimensional models using the forbidden [OI] line at 6300 Å, including correlations with other parameters involved.
Methods. We worked with a three-dimensional empirical model and two solar intensity atlases. Bayesian inference was employed as a tool to determine the most probable value for the solar oxygen abundance given the model chosen. We considered a number of error sources, such as uncertainties in the continuum derivation, in the wavelength calibration and in the abundance/strength of Ni.
Results. Our results show correlations between the effects of several parameters employed in the derivation. The Bayesian analysis provides robust confidence limits taking into account all of these factors in a rigorous manner. We obtain that, given the empirical three-dimensional model and the atlas observations employed here, the most probable value for the solar oxygen abundance is $\log(\epsilon_O) = 8.86 \pm 0.04$. However, we note that this uncertainty does not consider possible sources of systematic errors due to the model choice.

Key words. Sun: abundances – Sun: atmosphere – Sun: photosphere – methods: statistical

1. Introduction

The solar chemical composition is still under debate, the case of oxygen (the third most abundant element in the Sun after hydrogen and helium) being particularly important. One of the traditionally accepted abundance sets considers an oxygen abundances of $\log(\epsilon_O) = 8.93 \pm 0.04$ (Grevesse et al. 1984), in the usual astronomical scale referring to hydrogen (where $\log(\epsilon_O) = \log(N_X/N_H) + 12$). This abundance was later downward revised to $\log(\epsilon_O) = 8.83 \pm 0.06$ (Grevesse & Sauval 1998). With this abundance set, there is excellent agreement between solar interior model predictions and the helioseismology. However, using more recent three-dimensional theoretical models, lower solar metallicities have been obtained. For example, Asplund et al. (2004) reached $\log(\epsilon_o) = 8.66 \pm 0.05$, which may spoil the agreement with the helioseismology (see Basu & Antia 2008). We note, however, that a recent study suggests that we might be underestimating interior opacities, which could restore the agreement with helioseismology (Bailey et al. 2015). Given that oxygen is a very important element for stellar interior models, adopting such dramatic revisions would have implications over a broad range of topics in astrophysics.

Further works have not yet converged to a satisfactory resolution of the issue. On one hand, some studies found a relatively low oxygen abundance (e.g., Socas-Navarro & Norton 2007; Pereira et al. 2009; Grevesse et al. 2010) but, on the other hand, high oxygen abundance has also been obtained (e.g., Centeno & Socas-Navarro 2008; Ayres 2008; Socas-Navarro 2015b). Perhaps an intermediate value, such as $\log(\epsilon_O) \approx 8.75$ (e.g., Caffau et al. 2008, 2015), might be able to satisfy all of the constraints if the uncertainties had been underestimated thus far (i.e., the model atmosphere is typically assumed to be perfect except in Socas-Navarro 2011). In conclusion, the oxygen abundance issue remains open and largely controversial.

A key factor in the abundance derivation is the solar atmosphere model used. It seems natural that a three-dimensional model should be preferred over a one-dimensional one. However, when the three-dimensional model is a numerical simulation and the one-dimensional model has been obtained empirically by fitting observations, it is not clear that the former is better than the latter for fitting observations (Ayres 2008).

Deriving element abundances is far from trivial. The results are not determined solely by the solar atmosphere model used. They are also dependent on the solar atlas adopted to fit the model and even some details on the calibration (choice of the continuum level, spectrum rectification, wavelength calibration, etc.). Recently, some studies had noted the differences between the solar atlases (e.g., Caffau et al. 2008, 2009; Doerr et al. 2016) and the need to use more than one in the analysis. Another obvious factor is the value of the line strength employed (parameterized in terms of $\log(gf)$), which has been discussed in some works (e.g., Storey & Zeippen 2000; Johansson et al. 2003).

In this paper, we present a study of the oxygen abundance in the solar photosphere with the novelty that we applied Bayesian inference to properly disentangle the effects of the relevant parameters involved. In order to do that, we used a three-dimensional empirical model of the solar atmosphere (Socas-Navarro 2011, 2015b) to fit the forbidden [O I] line at 630.03 nm observed in two different solar atlases. In Sect. 2, we describe the solar model used, the syntheses made and the solar atlas observations. A brief summary of the Bayesian inference and a table with our priors are given in Sect. 3. Then, we present the results in Sect. 4 and, finally, some conclusions in Sect. 5.

2. Model atmosphere and observational data

2.1. Solar model used and synthesis

The solar atmosphere model that is used in this work was derived by Socas-Navarro (2011, 2015a) based on observations of the spectro-polarimeter (SP; Lites et al. 2001) of the solar optical telescope onboard the Hinode satellite (Kosugi et al. 2007). The observations have a field of view very close to the disk center and the wavelength range covers 6300.89 to 6303.27 Å. The spectrum at each pixel was inverted using the code NICOLE (Socas-Navarro et al. 2015) to determine a column with the height stratification of temperature, line-of-sight velocity, and magnetic field vector. The reader is referred to the papers cited above for more details on the model.

Our aim is to compare the synthetic line profiles with those observed in different atlases of the solar spectrum. The spatial resolution of these observations is extremely poor (of the order of tens of arcsec). Thus, we mimick this resolution by averaging all profiles synthesized in all pixels of the snapshot. As a consequence, the forward model in this three-dimensional snapshot turns out to be very time consuming. Given that the Bayesian analysis that we explain in the following section requires the evaluation of several tens of thousands of forward models, it is convenient to pre-compute a grid of models and carry out the Bayesian inference using a simulator that just interpolates on the grid of models (O'Hagan 2006). This greatly accelerates the inference by several orders of magnitude.

The precomputed database is built using a Cartesian grid in all parameters of the model. The model parameters can be separated into two different classes: slow parameters, that require the recomputation of the synthesis in the whole snapshot because they affect the radiative transfer, and fast parameters, that can be applied directly to the profiles with simple manipulations. The first two slow parameters are the oxygen and nickel abundances (note that the [O I] feature analyzed here has a Ni I blend). The third slow parameter is an enhancement factor for the velocities in the lower layers of the snapshot. This parameter is introduced because the model of Socas-Navarro (2011) is constructed fitting the strong Fe I lines at 630.1 and 630.2 nm, and it misses some of the dynamics in the lower layers of the atmosphere, precisely where the oxygen line forms. Fortunately, the nearby Sc II line, very similar in strength and formation height to the [O I] line, is an excellent calibration tool for this missing turbulence, which we parameterize as a multiplicative enhancement factor applied on the velocities of the lower layers. For more details, see the discussion in Socas-Navarro (2011). The fast parameters are: two for a linear correction to the continuum, a global velocity shift that takes into account imprecision in the wavelength calibration (this correction is important because a degeneration exists between the ratio of Ni and O abundances; this is explained in more detailed in Sect. 4.2 below) and the uncertainty of the fit.

To end up with a well sampled database, we took ten possible values for each one of the slow parameters, resulting in a grid of 1000 models. The oxygen and nickel abundances were varied between typical values reported in the literature (see e.g., Grevesse et al. 1984; Anders & Grevesse 1989; Grevesse & Sauval 1998; Asplund et al. 2009). Their minimum and maximum values are shown in Table 2 together with those for the enhancement factor.

It is important to stress that we considered the nickel abundance in the analysis because the spectral feature at 630.0 nm is actually the result of a blend between the [O I] line at 630.03 nm and a Ni I line at 630.03 nm. We considered the two major isotopes ^{58}Ni I and ^{60}Ni I, with a $\log(gf)$ value of -2.11 and a fraction of 72% and 28% for each isotope, respectively (Johansson et al. 2003; see Table 1).

For the synthesis with NICOLE, we chose a wavelength range from 630.0 to 630.3 nm, including also the Fe I lines at 630.15 and 630.25 nm. We introduced the Fe I lines because the [O I] line is formed in the far wing of the 630.15 nm Fe I line and we need to include this effect in our synthesis, since it affects the continuum estimate.

We introduce an additional simplification in the forward modeling, given that the model of Socas-Navarro (2011) has 200×200 spatial pixels in the field of view. Synthesizing the wavelength range in each one of the 1000 models in our grid would be very demanding computationally. For this reason, we carried out a Montecarlo selection of a subset of 1000 pixels, as a good compromise between accuracy and computational time, that produces an average profile that is equivalent to the average profile of the whole snapshot. By "equivalent", we mean that the difference between the average intensity profile of the full map and that of the subset is smaller than 0.10% in the oxygen line.

In summary, we constructed a grid of 1000 models where each model has 1000 pixels. For each one of them, we synthesized the spectral lines of interest using NICOLE, taking into account the atomic information compiled in Table 1. For the sake of reproducibility, the table also includes the references for the $\log(gf)$ values. The $\log(gf)$ value for the Fe I at 630.25 nm was taken from Socas-Navarro (2011). The radiative, Stark and van der Waals damping parameters are also provided for each line except the Fe lines. For those, we give instead the α and σ damping parameters using the method of Anstee & O'Mara (1995) obtained with the code of Barklem et al. (1998).

2.2. Observational data

In this study, we also intend to assess the influence of the choice of solar observation on the abundance determination. We used two different intensity atlases of the solar photosphere that have a good signal-to-noise ratio and have been widely used in previous works (e.g., Caffau et al. 2008, 2009; Ayres 2008; Maiorca et al. 2009).

The first one is the disk center intensity atlas of the solar spectrum from 3000 Å to 10 000 Å by Delbouille et al. (1973). The spectrum was obtained at the International Scientific Station of the Jungfraujoch and may be downloaded from the BASS2000 web server[1]. The second atlas that we considered was the Kitt Peak intensity Fourier Transform Spectrometer (FTS) atlas by Neckel & Labs (1984). This atlas was produced with the FTS instrument at the McMath telescope and spans the wavelength range from 3290 to 12 510 Å.

3. Bayesian analysis

We employ the formalism of Bayesian inference (e.g., Gregory 2005) to quantify the uncertainty in our conclusions derived from the use of a prescribed model, a bunch of free parameters, and a set of priors (information known a priori). The aim of the Bayesian analysis is to compute the posterior probability distribution associated with all model parameters by taking into account the information provided by the observations and all our a priori information. It relies on two fundamental tools. The first one is the Bayes theorem that describes very simply

[1] http://bass2000.obspm.fr/solar_spect.php

Table 1. Adopted atomic parameters.

Ion	Wavelength [nm]	Ex. Pot. [eV]	$\log(gf)$	Conf. (lower)	Conf. (upper)	$\gamma_{\rm rad}$	$\gamma_{\rm Stark}$	$\gamma_{\rm Waals}$	σ	α
[O I]	630.0304	0.000	–9.717[a]	^{3}P$_2$	^{1}D$_2$	0.0	0.05	1.00	...	...
Ni I	630.0335	4.266	–2.253[b]	^{3}D$_1$	^{3}P$_0$	2.63	0.054	1.82	...	...
Ni I	630.0355	4.266	–2.663[b]	^{3}D$_1$	^{3}P$_0$	2.63	0.054	1.82	...	...
Sc II	630.0678	1.507	–1.898[c]	^{3}P$_2$	^{3}D$_2$	2.30	0.05	1.30	...	...
Fe I	630.15012	3.654	–0.718[c]	^{5}P$_2$	^{5}D$_2$	...	...	...	834.4	0.243
Fe I	630.24940	3.686	–1.13	^{5}P$_1$	^{5}D$_0$	...	...	...	850.2	0.239

Notes. $\gamma_{\rm rad}$, $\gamma_{\rm Stark}$ and $\gamma_{\rm Waals}$ are the radiative, the Stark and the van der Waals damping parameters (units 10^8 rad s^{-1}). The $\log(gf)$ are taken from: [(a)] Storey & Zeippen (2000), [(b)] Johansson et al. (2003), [(c)] VALD database.

how the prior information is updated to the posterior information with the acquisition of new observations. If we denote the model parameters of interest with the vector $\boldsymbol{\theta}$, and the observations with D, the Bayes theorem states that

$$p(\boldsymbol{\theta}|D, I) = \frac{p(\boldsymbol{\theta}|I)p(D|\boldsymbol{\theta}, I)}{p(D|I)}. \quad (1)$$

In the Bayes theorem, $p(\boldsymbol{\theta}|I)$ is the prior probability distribution of θ, that encodes all the a-priori information we know about the parameters (i.e., if a quantity is positive or if a certain region of the space of parameters is more probable than others, etc.). $p(D|\boldsymbol{\theta}, I)$ is the likelihood, which encodes the information about the model parameters that can be extracted from the observations. $p(\boldsymbol{\theta}|D, I)$ is the posterior distribution, that describes everything that we know about the model parameters. $p(D|I)$ is the model evidence, which is an unimportant constant in our case, because it does not depend on $\boldsymbol{\theta}$. Finally, I refers to any important context information that is necessary for the inference. In our case, for instance, the specific radiative transfer model that we use. The second fundamental tool is the marginalization that is used to obtain the posterior distribution for any parameter taking into account the uncertainties and correlations with other parameters:

$$p(\theta_1|D, I) = \int \mathrm{d}\theta_2 \ldots \mathrm{d}\theta_n p(\boldsymbol{\theta}|D, I). \quad (2)$$

In this work, we examined seven free parameters that we consider necessary for explaining all the expected variability of the line profiles: the oxygen abundance (which is the main parameter of interest), the abundance of nickel (necessary to reproduce the line shape because the oxygen line of interest is blended with a Ni I line), a global wavelength shift to account for possible inaccuracies in the wavelength scale, an enhancement factor for the dynamics that is described above, two free parameters that are used to correct the continuum level and its slope, and finally a parameter σ that describes the (unknown) uncertainty of the observations and also quantifies the quality of our modeling.

In any Bayesian inference, it is crucial to define the priors. In our case, we define the priors as flat (meaning there is no preference) for almost all parameters in the ranges shown in Table 2. The ranges have been chosen taking into consideration the results obtained in previous studies (see e.g., Grevesse et al. 1984; Anders & Grevesse 1989; Grevesse & Sauval 1998; Asplund et al. 2009) but trying not to discard parts of the space of parameters that might be compatible with the observations. The only non-flat prior is that of σ, for which we adopted a Jeffreys prior, $p(\sigma) \propto \sigma^{-1}$, because it is a scale parameter that can potentially have values spanning several orders of magnitude (e.g., Gregory 2005).

Table 2. Priors selected for each parameter of our analysis.

Parameter	Range	Type
$\log(\epsilon_{\rm o})$	(8.55, 9.20)	Flat
$\log(\epsilon_{\rm Ni})$	(5.84, 6.36)	Flat
v [km s^{-1}]	(–1, 1)	Flat
Enhancement factor	(0.5, 2.3)	Flat
σ	(0.0001, 0.02)	Jeffreys
Cont. slope [$I/I_{\rm c}$ Å^{-1}]	(–0.1, 0.1)	Flat
Cont. y intercept [$I/I_{\rm c}$]	(0.9, 1.1)	Flat

For computing the likelihood, we assume the following generative model:

$$O(\lambda_i) = I(\boldsymbol{\theta}, \lambda_i) + n(\lambda_i), \quad (3)$$

which states that the observations for the ith wavelength point, $O(\lambda_i)$, can be modeled with $I(\boldsymbol{\theta}, \lambda_i)$ plus some uncertainty that has zero mean and variance σ^2. Following the standard approach (e.g., Gregory 2005), the likelihood has the form of an uncorrelated multivariate normal distribution, which reduces to:

$$p(D|\boldsymbol{\theta}, I) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{-[O(\lambda_i) - I(\boldsymbol{\theta}, \lambda_i)]^2}{2\sigma^2}\right], \quad (4)$$

where n is the number of observed wavelengths.

We sample the posterior distribution using the nested sampling technique by Skilling (2004)[2].

4. Results

4.1. Line fit

Figure 1 shows our fits of the [O I] and Sc II lines in our model (black lines) with respect to the atlases (red). The left panel uses the Neckel atlas and the right one uses the Delbouille atlas. Instead of a single fit, we provide a sampling of models that are compatible with the observations. They are obtained by synthesizing models with parameters extracted from the posterior distribution. Figure 1 clearly shows that our model fits both atlases very well, even in very weak lines such as those shown here. However, the models present some discrepancies with the observations in the continuum region in between the two lines, where the models set the continuum to a higher level. Given that we see this effect in both atlases, it might be produced by some unknown absorption in this region. The lines of interest are very weak, thus, this region could affect our abundance determination indirectly by setting the continuum level to an incorrect value.

[2] We use the implementation of the Python package `nestle`, which can be obtained from `http://kbarbary.github.io/nestle/`

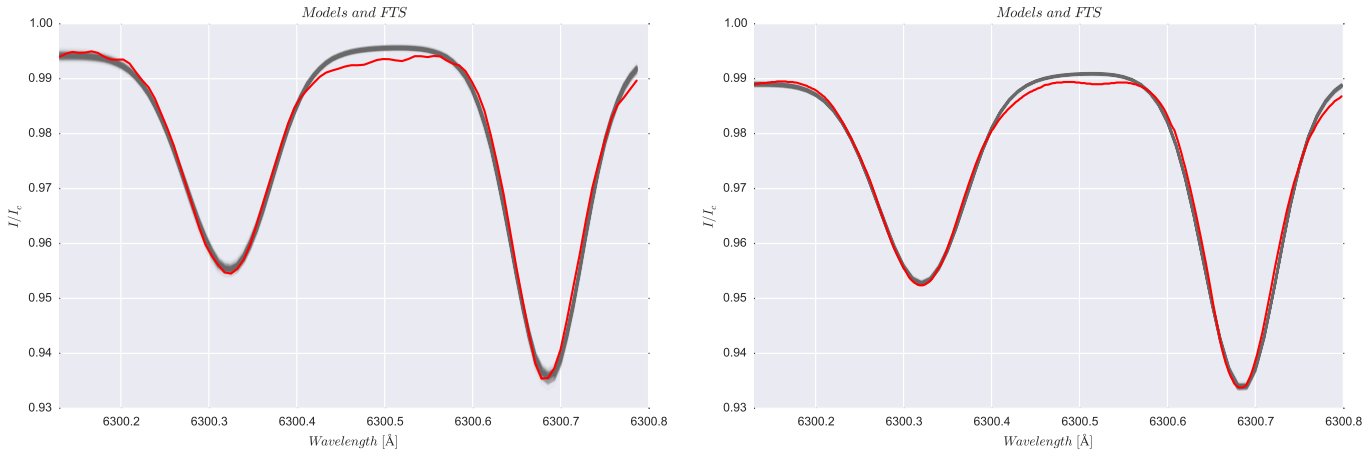

Fig. 1. Some representative fits of the [O I] line (*left*) and Sc II line (*right*) are shown in black. The red profile corresponds to the Neckel solar atlas in the left panel and to the Delbouille atlas in the right panel.

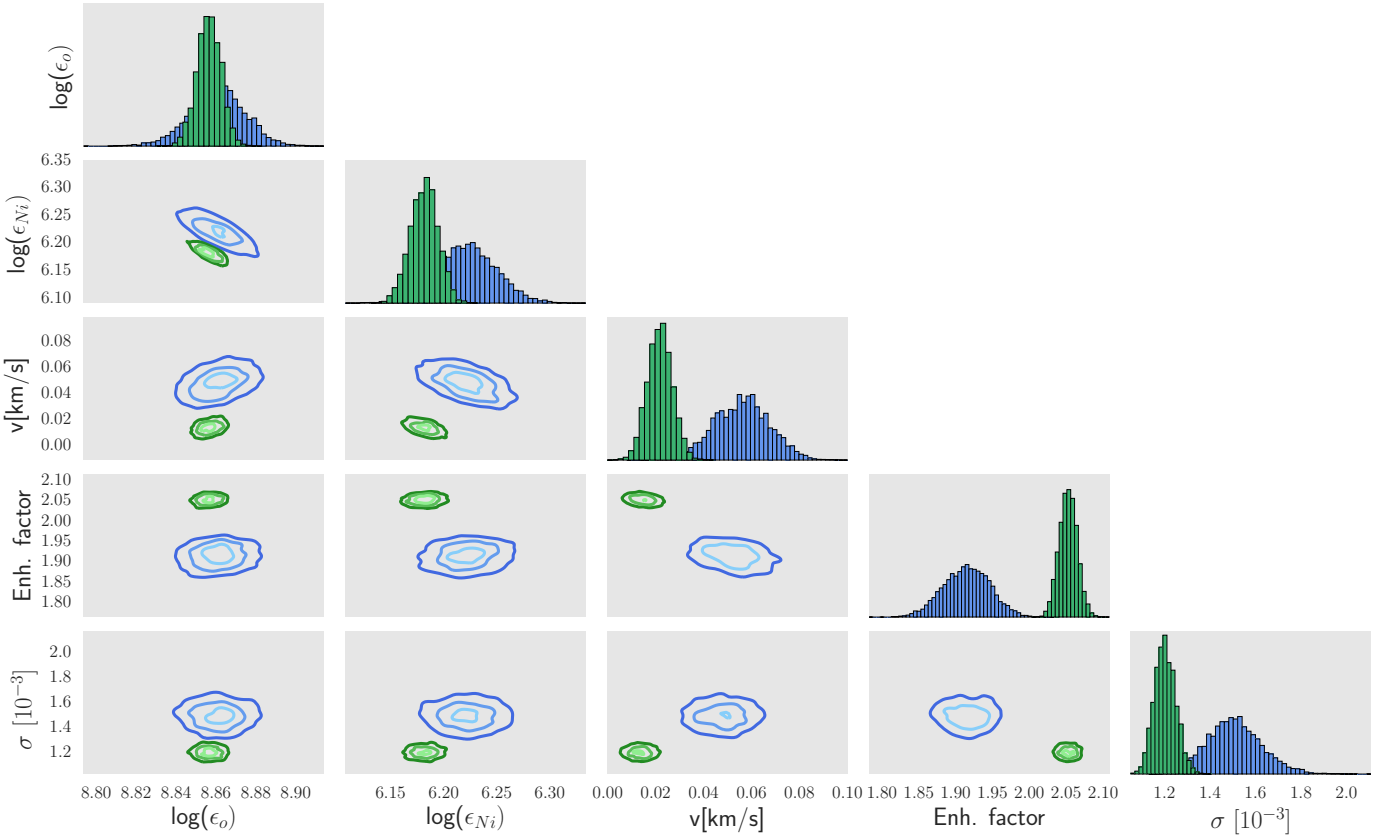

Fig. 2. Histograms and correlations of the posterior distributions of the main five parameters (in blue for Neckel solar atlas and green for Delbouille atlas). *From left to right, and top to bottom*, we show: oxygen abundance, nickel abundance, velocity (in km s^{-1}), enhancement factor and σ.

We tried to minimize this effect by introducing the Sc II line in the analysis, which then sets a much better continuum level estimation. We have plans to refine this approach in the future by using a modeling that is able to absorb these deficiencies (similar to the nonparametric approach based on Gaussian Processes of Czekala et al. 2015).

4.2. Marginal posterior distributions

The joint and marginal posterior distributions for the five main parameters (oxygen and nickel abundances, velocity, enhancement factor and σ) are shown in Fig. 2. The blue distributions correspond to the Neckel atlas while the green distributions correspond to the Delbouille atlas. We did not show the distribution for the two parameters of the continuum because they are nuisance parameters that are only needed to have a good fit of the line, but they do not provide relevant information.

For both atlases, the joint and marginal posterior distributions show a Gaussian-like shape, where those corresponding to the Neckel atlas always present a larger width. This width indicates that the uncertainties in the abundances inferred from the Neckel atlas are larger than those obtained from the Delbouille atlas. Moreover, the marginal distribution for σ also shows that this parameter is smaller for the case of the Delbouille atlas.

Table 3. Parameters median values and deviations for both atlases.

Atlas	$\log(\epsilon_o)$	$\log(\epsilon_{Ni})$	v [km s^{-1}]	Enh. factor	σ
Neckel	8.861 ± 0.014	6.22 ± 0.03	0.05 ± 0.01	1.92 ± 0.03	0.0015 ± 0.0001
Delbouille	8.856 ± 0.006	6.18 ± 0.01	0.014 ± 0.005	2.05 ± 0.01	0.00120 ± 0.00005

Table 4. Comparison of the results for the different experiments.

			$\log(\epsilon_o)$	$\log(\epsilon_{Ni})$	v [km s^{-1}]	Enh. factor	σ
Both lines	Neckel	prior Ni	8.864 ± 0.013	6.21 ± 0.03	0.05 ± 0.01	1.92 ± 0.03	0.0015 ± 0.0001
		no prior	8.861 ± 0.014	6.22 ± 0.03	0.05 ± 0.01	1.92 ± 0.03	0.0015 ± 0.0001
	Delbouille	prior Ni	8.858 ± 0.006	6.18 ± 0.01	0.014 ± 0.005	2.05 ± 0.01	0.00120 ± 0.00005
		no prior	8.856 ± 0.006	6.18 ± 0.01	0.014 ± 0.005	2.05 ± 0.01	0.00120 ± 0.00005
Only [O I]	Neckel	prior Ni	$8.910^{+0.015}_{-0.019}$	6.02 ± 0.05	$0.17^{+0.03}_{-0.04}$	1.70 ± 0.04	0.00052 ± 0.00005
		no prior	$8.940^{+0.007}_{-0.012}$	$5.90^{+0.06}_{-0.04}$	$0.23^{+0.01}_{-0.02}$	1.74 ± 0.04	0.00050 ± 0.00005
	Delbouille	prior Ni	$8.937^{+0.004}_{-0.007}$	$5.88^{+0.04}_{-0.03}$	$0.171^{+0.007}_{-0.013}$	1.99 ± 0.02	0.00036 ± 0.00005
		no prior	8.940 ± 0.002	$5.85^{+0.02}_{-0.01}$	$0.178^{+0.003}_{-0.004}$	2.00 ± 0.01	0.00036 ± 0.00002

Thus, our model seems to provide more constrained model parameters and better fits for the Delbouille atlas.

It is encouraging that the marginal posterior for the oxygen abundance is consistent in both atlases, with that inferred from the Delbouille atlas being contained inside the uncertainty associated with the Neckel atlas. Further, it is interesting that the maximum marginal a posteriori (the peak of the marginal posterior) is located at approximately the same value for the two cases.

This is not the same with the remaining parameters. The marginal posteriors corresponding to the Delbouille atlas tend to be shifted towards smaller values, except for the enhancement factor, where the opposite happens. In this last parameter, the marginal distributions even indicate that the inferred enhancement factors are not compatible (at least to three standard deviations).

If we look at the panels with the joint posteriors (showing the correlation between pairs of parameters), we see a clear correlation between the abundances of oxygen and nickel. This correlation is a direct consequence of the fact that the [O I] line is a blend; we can obtain an equally good fit if we take a model with a higher nickel abundance and a lower oxygen abundance, and vice versa. Furthermore, these two parameters (oxygen and nickel abundances) are correlated with the velocity shift. This correlation makes sense for small velocities, since a wavelength shift displaces the line and one can still obtain a good fit by appropriately modifying the abundances. For example, a blue shift can be compensated with a lower nickel abundance and a higher oxygen abundance.

All marginal posterior distributions have a Gaussian-like shape. Therefore, we summarize them in Table 3 by providing the median and the uncertainty defined by the percentiles 16 and 84 (equivalent to the standard 1σ uncertainty in the Gaussian case). It is clear in the table that the most probable value for the oxygen abundance is the same (down to the second decimal) in both atlases. The value obtained from this study is, hence, $\log(\epsilon_O) = 8.86 \pm 0.01$. This value is within the category of high solar oxygen abundances: lower than the 8.93 ± 0.02 value in Grevesse et al. (1984), but compatible with the 8.83 ± 0.06 revision in Grevesse & Sauval (1998).

4.3. More experiments

In order to quantify how much the impact of the continuum level and the correlation between the oxygen and nickel abundances are, we carried out two more studies. They are also intended to verify the robustness of our results.

First, we carried out the same analysis as before but reducing the wavelength range to take into account the [O I] line only. Although we do not display the fits or the marginal posteriors for simplicity, the summary of the results is displayed in Table 4 (see corresponding rows labeled as "Only [O I]" and "no prior"). For completeness, in this table, we repeat the results of the previous section for an easier comparison (labeled as "Both line" and "no prior"). Thus, modifying the range of wavelengths, the most probable oxygen abundance changes from $\log(\epsilon_O) = 8.86$, a value compatible with Grevesse & Sauval (1998), to $\log(\epsilon_O) = 8.94$, a value compatible with Grevesse et al. (1984). Because we are neglecting the wing of the Sc II line and the inferred abundance is different, we conclude that a good estimate of the continuum is crucial for the estimation of abundances. However, this large oxygen abundance seems improbable in light of the very low nickel abundance inferred, which is incompatible with all previous reports in the literature (that usually give $\log(\epsilon_{Ni}) > 6$).

The second test was motivated by the low nickel abundance of the previous results. In this case, we repeated the analysis, this time including a different prior for the nickel abundance. This prior was set to a Gaussian distribution with mean $\mu = 6.17$ and standard deviation of $\sigma = 0.07$, as reported in Scott et al. (2009). We intentionally increased the value of σ to decrease the informativeness of the prior and let the data drive the results. The results of this new analysis are also shown in Table 4 (with labels "prior Ni"). The results show that the inference implementing only the [O I] line is not reliable because, when including the Gaussian prior for the nickel abundance, the modification in the oxygen and nickel abundances is greater than the previously quoted uncertainties. This fundamentally means that the result is strongly dependent on the prior. However, this is not the case when using the [O I] and Sc II lines together, where we obtain results that are essentially insensitive to the prior.

4.4. Abundance and log(gf) factor

We are aware that an important source of uncertainty in the determination of abundances are the atomic parameters of the spectral lines, specifically the value of $\log(gf)$. We have tried to use the most recent determinations, but it is true that a slightly different value of $\log(gf)$ inside its uncertainty would produce different oxygen and nickel abundances. Therefore, one can consider our

determination of the abundance to be indeed an inference over the product $gf\epsilon_O$, which remains valid for weak lines.

The gf value used in this work (Johansson et al. 2003) is accepted as the most accurate, but its uncertainty can go up to 10%. This uncertainty translates into $\log(gf) = -9.717 \pm 0.043$. If this is taken into account, it would induce an uncertainty in the estimated oxygen abundance of $\log(\epsilon_O) = 8.86 \pm 0.04$.

5. Conclusions

There are many parameters involved in the determination of abundances. To end up with an accurate result, it is crucial to put emphasis on appropriately considering all of them. This is the motivation of this work and it represents a first step towards a reliable determination of the solar abundances. Several conclusions can be extracted from our work:

- We have used a very flexible generative model that contains seven parameters. Some of these parameters are nuisance parameters that are of no diagnostic interest but which are necessary for a good modeling of the spectral line. Including these nuisance parameters and marginalizing over them is crucial for a reliable determination of abundances.
- Including nearby spectral lines turns out to be very important, because they modify the continuum level. We have found that a good characterization of the continuum level in these weak lines is crucial for the inference of abundances.
- We have found some differences when applying exactly the same modeling to different atlases. Therefore, we find it may be convenient to infer solar abundances using the largest panoply of observations. However, we find a reliable oxygen abundance with only 0.003 dex difference between the studies with the two solar atlases.
- A reliable independent determination of the spectral line atomic parameters is very important. Otherwise, this uncertainty propagates accordingly onto the inferred abundances.
- Given the interdependencies among all model parameters and the uncertainty in the observations, one should pursue a fully Bayesian approach taking into account all effects simultaneously. Ideally, such an approach should also be considered for aggregating all sources of uncertainty and checking, through a meta-analysis, for the compatibility of all the inferred abundances that exist in the literature.

Despite the many factors taken into account, we found a good agreement in our results for the oxygen abundance. Thus, we can conclude that, based on the three-dimensional model used, the most probable value for the oxygen solar abundance is $\log(\epsilon_0) = 8.86 \pm 0.04$. This value is classified as a high solar oxygen abundance and is compatible with the results of Grevesse & Sauval (1998).

Finally, a caveat is in order. We note that all results given in this paper rely on the empirical model of Socas-Navarro (2011, 2015a) and the radiative transfer of NICOLE being correct. If these were found to be incorrect or inaccurate, the conclusions could change. The advantage of the Bayesian framework lies in its transparency (every probability distribution in Eq. (1) is conditioned on the a priori information I).

Acknowledgements. Financial support by the Spanish Ministry of Economy and Competitiveness through projects AYA2014-60476-P and Consolider-Ingenio 2010 CSD2009-00038 are gratefully acknowledged. A.A.R. acknowledges financial support through the Ramón y Cajal fellowship. M.C.A. acknowledges Fundación la Caixa for the financial support received in the form of a Ph.D. contract. This research has made use of NASA's Astrophysics Data System Bibliographic Services. This work has made use of the VALD database, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna. We acknowledge the community effort devoted to the development of the following open-source Python packages that were used in this work: `numpy`, `matplotlib`, `nestle`, and `seaborn`.

References

Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Anstee, S. D., & O'Mara, B. J. 1995, MNRAS, 276, 859
Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, A&A, 417, 751
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Ayres, T. R. 2008, ApJ, 686, 731
Bailey, J. E., Nagayama, T., Loisel, G. P., et al. 2015, Nature, 517, 56
Barklem, P. S., Anstee, S. D., & O'Mara, B. J. 1998, PASA, 15, 336
Basu, S., & Antia, H. M. 2008, Phys. Rep., 457, 217
Caffau, E., Ludwig, H.-G., Steffen, M., et al. 2008, A&A, 488, 1031
Caffau, E., Maiorca, E., Bonifacio, P., et al. 2009, A&A, 498, 877
Caffau, E., Ludwig, H.-G., Steffen, M., et al. 2015, A&A, 579, A88
Centeno, R., & Socas-Navarro, H. 2008, ApJ, 682, L61
Czekala, I., Andrews, S. M., Mandel, K. S., Hogg, D. W., & Green, G. M. 2015, ApJ, 812, 128
Delbouille, L., Roland, G., & Neven, L. 1973, Atlas photometrique du spectre solaire de [lambda] 3000 a [lambda] 10 000, Universite de Liège, Institut d'Astrophysique
Doerr, H.-P., Vitas, N., & Fabbian, D. 2016, A&A 590, A118
Gregory, P. C. 2005, Bayesian Logical Data Analysis for the Physical Sciences (Cambridge: Cambridge University Press)
Grevesse, N., & Sauval, A. J. 1998, Space Sci. Rev., 85, 161
Grevesse, N., Sauval, A. J., & van Dishoeck, E. F. 1984, A&A, 141, 10
Grevesse, N., Asplund, M., Sauval, A. J., & Scott, P. 2010, Ap&SS, 328, 179
Johansson, S., Litzén, U., Lundberg, H., & Zhang, Z. 2003, ApJ, 584, L107
Kosugi, T., Matsuzaki, K., Sakao, T., et al. 2007, Sol. Phys., 243, 3
Lites, B. W., Elmore, D. F., Streander, K. V., et al. 2001, in UV/EUV and Visible Space Instrumentation for Astronomy and Solar Physics, eds. O. H. Siegmund, S. Fineschi, & M. A. Gummin, Proc. SPIE, 4498, 73
Maiorca, E., Caffau, E., Bonifacio, P., et al. 2009, PASA, 26, 345
Neckel, H., & Labs, D. 1984, Sol. Phys., 90, 205
O'Hagan, A. 2006, Reliability Engineering & System Safety, 91, 1290
Pereira, T. M. D., Asplund, M., & Kiselman, D. 2009, A&A, 508, 1403
Scott, P., Asplund, M., Grevesse, N., & Sauval, A. J. 2009, ApJ, 691, L119
Skilling, J. 2004, in AIP Conf. Ser. 735, eds. R. Fischer, R. Preuss, & U. V. Toussaint, 395
Socas-Navarro, H. 2011, A&A, 529, A37
Socas-Navarro, H. 2015a, NICOLE: NLTE Stokes Synthesis/Inversion Code, Astrophysics Source Code Library
Socas-Navarro, H. 2015b, A&A, 577, A25
Socas-Navarro, H., & Norton, A. A. 2007, ApJ, 660, L153
Socas-Navarro, H., de la Cruz Rodríguez, J., Asensio Ramos, A., Trujillo Bueno, J., & Ruiz Cobo, B. 2015, A&A, 577, A7
Storey, P. J., & Zeippen, C. J. 2000, MNRAS, 312, 813