How realistic are solar model atmospheres?

T. M. D. Pereira[1,2], M. Asplund[1,3], R. Collet[1,3], I. Thaler[3], R. Trampedach[4], J. Leenaarts[2]

[1] Research School of Astronomy and Astrophysics, Australian National University, Cotter Rd., Weston, ACT 2611, Australia
[2] Institute for Theoretical Astrophysics, University of Oslo, PO Box 1029, Blindern, N–0315, Norway
[3] Max-Planck-Institut für Astrophysik, Postfach 1317, D–85741 Garching b. München, Germany
[4] JILA, University of Colorado and National Institute of Standards and Technology, 440 UCB, Boulder, CO 80309, USA

ABSTRACT

Context. Recently, new solar model atmospheres have been developed to replace classical 1D LTE hydrostatic models and used to for example derive the solar chemical composition.

Aims. We aim to test various models against key observational constraints. In particular, a 3D model used to derive the solar abundances, a 3D MHD model (with an imposed 10 mT vertical magnetic field), 1D NLTE and LTE models from the PHOENIX project, the 1D MARCS model, and the 1D semi-empirical model of Holweger & Müller.

Methods. We confront the models with observational diagnostics of the temperature profile: continuum centre-to-limb variations, absolute continuum fluxes, and the wings of hydrogen lines. We also test the 3D models for the intensity distribution of the granulation and spectral line shapes.

Results. The predictions from the 3D model are in excellent agreement with the continuum centre-to-limb observations, performing even better than the Holweger & Müller model (constructed largely to fulfil such observations). The predictions of the 1D theoretical models are worse, given their steeper temperature gradients. For the continuum fluxes, predictions for most models agree well with the observations. No model fits all hydrogen lines perfectly, but again the 3D model comes ahead. The 3D model also reproduces the observed continuum intensity fluctuations and spectral line shapes very well.

Conclusions. The excellent agreement of the 3D model with the observables reinforces the view that its temperature structure is realistic. It outperforms the MHD simulation in all diagnostics, implying that recent claims for revised abundances based on MHD modelling are premature. Several weaknesses in the 1D hydrostatic models (theoretical and semi-empirical) are exposed. The differences between the PHOENIX LTE and NLTE models are small. We conclude that the 3D hydrodynamical model is superior to any of the tested 1D models, which gives further confidence in the solar abundance analyses based on it.

Key words. Sun: photosphere – line: formation – stars: atmospheres – Sun:abundances – Sun: granulation

1. Introduction

Solar model atmospheres are a cornerstone in stellar astronomy. The Sun is a natural reference when studying other stars, and realistic solar photosphere models are essential to infer solar parameters such as its chemical composition. The wealth of solar data available can be used to rigorously test and constrain photosphere models. This testing provides invaluable insight about the model physics and its degree of realism, paving the way for building realistic models of other stars.

With the significant increases in computational power of recent years, the classical approximations used when modelling stellar atmospheres have started to be challenged. Of these approximations, the most significant are the assumption of a static 1D atmosphere with a mixing length type treatment of convection, and the assumption of local thermodynamical equilibrium (LTE). Although at present no model of a stellar atmosphere is able to relax these two assumptions simultaneously, efforts have been made to tackle each of these approximations individually. On the geometry/convection side, realistic 3D hydrodynamical time-dependent simulations of convection have been developed and used as models of the solar photosphere (*e.g.* Stein & Nordlund 1998; Asplund et al. 2000a; Freytag et al. 2002; Vögler et al. 2004; Carlsson et al. 2004; Caffau et al. 2008b). On the radiative transfer side, Short & Hauschildt (2005) have computed a 1D non-LTE (NLTE) hydrostatic solar model atmosphere with the PHOENIX code (Hauschildt & Baron 1999) following the pioneering work in this regard by Anderson (1989).

The application of 3D solar models to abundance analysis (*e.g.* Asplund et al. 2004; Caffau et al. 2010) has resulted in a revised solar photospheric metallicity of $Z = 0.0134$ (Asplund et al. 2009), substantially smaller than previous canonical values (*e.g.* $Z = 0.0201$ in Anders & Grevesse 1989 and $Z = 0.0169$ in Grevesse & Sauval 1998). The realistic treatment of convection and velocity fields in the 3D models resulted in an excellent agreement between predicted and observed line shapes and bisectors, not possible with the 1D models even with the free parameters of micro- and macro-turbulence (Asplund et al. 2000a). This agreement is a strong indicator of how realistic the model is. Additionally, when compared with observations of the solar granulation at high spatial resolution, the 3D models correctly predict the characteristic size and lifetimes of the granules (*e.g.* Stein & Nordlund 1998; Nordlund et al. 2009). However, the results are still controversial because by using a lower solar metallicity the previous excellent agreement between solar interior models and helioseismology deteriorates significantly (*e.g.* Bahcall et al. 2005; Basu & Antia 2008; Serenelli et al. 2009).

A criticism of the 3D solar models sometimes raised is that while they have been tested against many spectral lines, they lack a thorough testing of temperature structure, such as the continuum centre-to-limb variation and absolute continuum fluxes (Basu & Antia 2008). A correct temperature structure is of the

utmost importance to abundance studies. Using a 1D horizontal and temporal average of the 3D model of Asplund et al. (2000a), Ayres et al. (2006) suggest that the 3D model fails to describe the observed centre-to-limb variation and its temperature gradient is too steep. The first claim is partly dismissed by Koesterke et al. (2008), showing that the 1D average is not a valid approximation of the full 3D model for temperature profiling and that the performance of the 3D model used by Asplund et al. (2000a) and Asplund et al. (2005) in the continuum centre-to-limb variation is comparable to that of theoretical 1D models – a view corroborated by Pereira et al. (2008) and Trujillo Bueno & Shchukina (2009). However, Koesterke et al. (2008) agree with Ayres et al. (2006) in that the temperature gradient of this particular older 3D model is slightly too steep in the continuum forming layers. As mentioned by Asplund et al. (2009), an improved radiative transfer treatment resulted in a new 3D model with a more realistic temperature gradient (less steep than before).

The aim of this work is to systematically test the temperature structure of several models of the solar photosphere, including 3D and 1D NLTE models. We use the new 3D hydrodynamical model employed by Asplund et al. (2009), the 10 mT 3D MHD model of Thaler et al. (in preparation) and recent 1D NLTE and LTE models from the PHOENIX project (Hauschildt & Baron 1999). We also employ the widely used 1D MARCS model (Gustafsson et al. 2008) and the semi-empirical 1D model of Holweger & Müller (1974).

The models used are described in more detail in Sect. 2. To compare their temperature structure with the observations we use three classical tests: the continuum centre-to-limb variation in Sect. 3, the absolute continuum flux distribution in Sect. 4 and the wings of hydrogen lines in Sect. 5. In addition, we also test the 3D model against observations of the continuum intensity distribution and $\Delta I_{\rm rms}$ in Sect. 6. Sect. 7 contains a comparison of the predicted and observed line shapes, including inferred abundances, bisectors and line shifts, for a sample of Fe I and Fe II lines. Our conclusions are given in Sect. 8.

2. Model atmospheres

2.1. 3D Model

We use the same 3D hydrodynamical model solar atmosphere that was adopted by Asplund et al. (2009) for the derivation of photospheric solar abundances. The solar surface convection simulation was performed using the 3D, radiative, hydrodynamical, conservative, STAGGER-CODE (Nordlund & Galsgaard 1995). In the simulation, the equations for the conservation of mass, momentum, and energy are solved together with the radiative transfer equation for a representative volume of solar surface ($6\times6\times3.8$ Mm^3) on a Cartesian mesh with 240^3 numerical resolution. The horizontal grid is equidistant, while the vertical depth scale has a non-constant spacing optimised to better resolve the layers at the photospheric transition where temperature gradients are at their steepest. Open, transmitting, boundaries are assumed at the top and bottom of the simulation domain, and periodic boundary conditions are enforced horizontally. It is important for the lower boundary to be transmitting, to avoid homogenising the otherwise highly asymmetric (between up and down) convective flows.

The simulation domain completely covers the Rosseland optical depth range $-5 \leq \log \tau_{\rm Ross} \leq 7$. The radiative transfer equation is solved using a long-characteristics Feautrier-like scheme down to $\tau_{\rm Ross}\approx300$, and the diffusion approximation is employed in the deeper layers. The main improvement over the simulation

of Asplund et al. (2000a) is the treatment of opacities and line-blocking in particular. In both the old (2005) and new (2009) 3D simulations, continuous and line opacities are included via a statistical method, called *opacity binning* or *multi-group method* (Nordlund 1982): wavelengths are sorted into *opacity bins* according to the strength of the opacity and the corresponding LTE source functions are added together within each bin. In the original binning scheme, Nordlund (1982) used the Rosseland average κ_0 of the opacities in the continuum bin, scaled by a constant factor for each of the other bins, $\kappa_{\rm j} = \kappa_0 10^{j\Delta x}$, in practice assuming $\Delta x = 1$ and $j = 0, \ldots, 11$. The transition to free streaming for optical depths in the continuum bin $\tau_0 \ll 1$ is ensured by an exponential bridging (in τ_0) to an intensity-weighted mean opacity. The multi-group method was further developed by Skartlien (2000), who relaxed the approximation of κ_j just being a scaled κ_0, and instead computed the actual Rosseland average for each bin. This also ensures that the Rosseland mean of the bin-wise opacities converges to the actual Rosseland mean of the monochromatic opacities. The new 3D model has further been improved by sorting opacities into bins not only according to opacity strength but also according to wavelength, and allowing arbitrary bin sizes; a similar binning criterion has been implemented by Caffau et al. (2008a) in the CO^{5}BOLD code. For the present simulation, continuous opacities are taken from Gustafsson et al. (1975, and subsequent updates) and line opacities from the latest MARCS stellar atmosphere package (Gustafsson et al. 2008). The solar chemical composition by Asplund et al. (2005) and the equation-of-state by Mihalas et al. (1988, and subsequent updates) are adopted.

The positions of the bin borders are then optimised with respect to a monochromatic radiative transfer calculation for the simulation's average temperature and density stratification taken on surfaces of constant Rosseland optical depth. The generalisation of the bins and the optimisation reduce the differences between the radiative heating of the monochromatic and the binned solution by a factor of five, to within less than 1%. The effect of radiative heating and cooling on the simulation is therefore faithfully reproduced by this opacity-binning, for the average atmosphere. We have also performed a 3D monochromatic radiative transfer calculation on the full 3D simulation cube for a single snapshot, and found a < 5 K (0.08%) difference in $T_{\rm eff}$ between the monochromatic and the binned solution. Before being subjected to scientific tests, the simulation has been fully relaxed, and has subsequently run for the 45 solar minutes used here (90 snapshots). The relaxation process included extracting energy from radial p-modes, and ensuring that the total flux is statistically constant with depth, that no drifts are present in the thermodynamical quantities at the bottom boundary, and that the vertical grid's resolution is enough for the radiative transfer in the photosphere.

For the calculations presented here the original simulation was interpolated to a coarser $50\times50\times82$ resolution (with a finer vertical depth scale) to save computing time. The effective temperature of the 90 snapshots used is $T_{\rm eff} = 5778 \pm 2$ K, close to the observed value of 5777 ± 3 K (Willson & Hudson 1988).

2.2. 3D MHD Model

In addition to the 3D simulation, we also test a 3D magneto-hydrodynamical (MHD) simulation from Thaler et al. (in preparation). This simulation was performed using the same code and physical ingredients of the 3D hydrodynamical model of Asplund et al. (2009). It uses 240×240 grid points horizontally and 220 vertically. It has the same horizontal size of 6×6 Mm^2 as
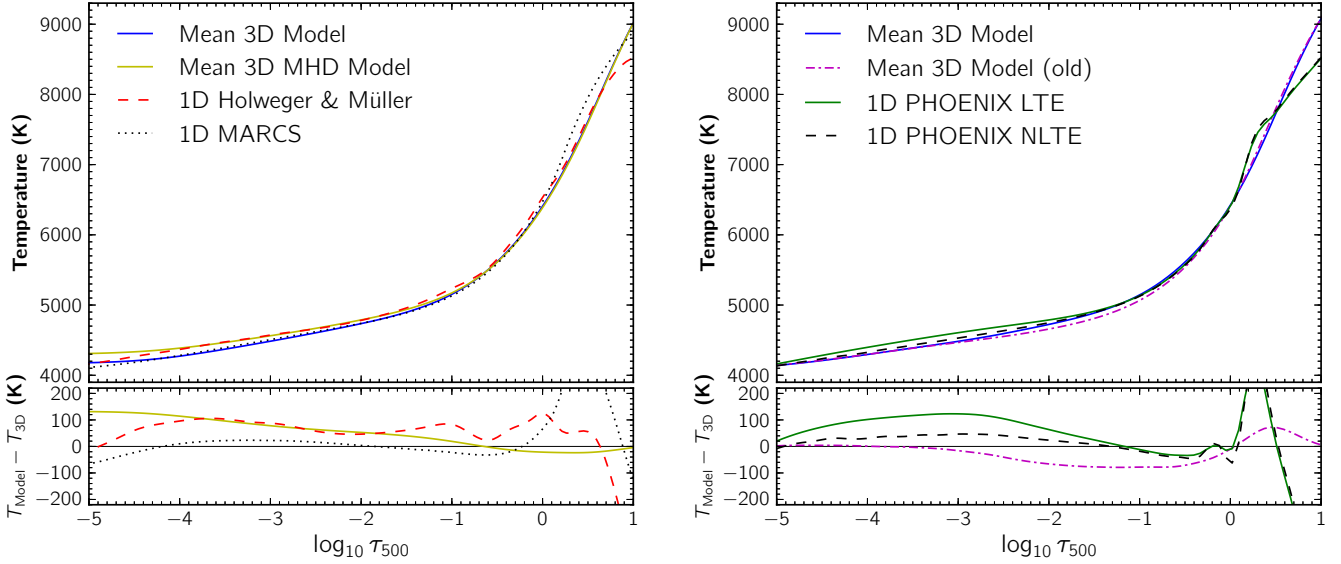
Fig. 1. Temperature structure of the 3D and 1D models, plotted against the optical depth at 500 nm. For the 3D models structure represents the temporal and spatial mean (over τ_{500} iso-surfaces). *Bottom panels:* differences between the 3D model and a given model (legend according to the top panels).

the 3D hydrodynamical model, but a slightly shorter vertical extension of 3.34 Mm. It extends 2.7 Mm into the convection zone and reaches 0.645 Mm up the photosphere. As in the 3D hydrodynamical model, the horizontal grid is equidistant and the vertical depth scale optimised to resolve the photosphere. The boundary conditions are the same as in the Asplund et al. (2009) simulation. On a thermodynamically relaxed hydrodynamical snapshot, a vertical magnetic field of 10 mT was overimposed. The magnetic field is kept vertical at the bottom boundary and tends toward a potential field at the top. This configuration was run for 120 min of solar time. For the results presented here, a sequence of 38 min was used, taken after this simulation had run for 68 min.

The choice of 10 mT for B was made as it is a reasonable value for the quiet sun's mean field strength (e.g. Trujillo Bueno et al. 2006) and because it allows a comparison with the middle MHD model of Fabbian et al. (2010, 2012).

2.3. 1D Models

We use two types of 1D models: theoretical and semi-empirical. The semi-empirical 1D model of Holweger & Müller (1974) was built from a range of observables to reproduce the mean physical quantities of the solar photosphere. Most importantly, it was constructed to follow the observed continuum centre-to-limb variation between 0.5–300 μm and the line depths of ≈900 spectral lines. This is an important detail to note when considering our centre-to-limb variation comparison, although the observations we employ are more recent than the ones available when the Holweger & Müller model was built. Historically, the Holweger & Müller model has been the atmosphere of choice when deriving solar abundances. It assumes hydrostatic equilibrium and does not explicitly include convection.

Of the 1D hydrostatic theoretical model atmospheres we include the LTE, line-blanketed solar MARCS model, which is a reference for the MARCS grid of model atmospheres (Gustafsson et al. 2008). We also include an LTE and an NLTE model from the PHOENIX project (Hauschildt et al. 1999).

These two models have been computed for the solar abundances of Asplund et al. (2005) and the same input physics as for the PHOENIX Gaia grid (Brott & Hauschildt 2005). They differ only in their treatment of atomic level populations with the NLTE model having been computed with a NLTE treatment of H, He, C, N, O, Mg and Fe.

To ensure consistency when using our line formation code, opacities and equation of state, for the 1D models we took the $T(\tau)$ relation from their respective references and integrated P_{gas} in optical depth assuming hydrostatic equilibrium to obtain the pressures and densities that yield the same $T(\tau)$ when using our opacities and equation of state. We note that this is an often overlooked source of error when comparing results for supposedly the same model atmosphere since this pressure-integration is not always carried out. In fact the Holweger & Müller (1974) model is essentially only an updated version of the Holweger (1967) model with a new pressure-integration due to their effects on opacities and thermodynamics.

2.4. Mean stratification

In Fig. 1 we compare the mean temperature structure of the 3D hydrodynamical model with other models. On the left, with the 3D MHD model of Thaler et al. (in preparation), the 1D Holweger & Müller and MARCS models (left) and on the right with the old 3D model of Asplund et al. (2000a), and the PHOENIX LTE and NLTE models. The mean structure of the 3D models was calculated by averaging over τ_{500} iso-surfaces.

The 3D model of Asplund et al. (2000a) is provided here for a quick comparison. This older model was run with a precursor of the stagger-code, and was used to derive the solar abundances of Asplund et al. (2005). The largest difference between the old and new models is the treatment of radiation. The new model has an improved multi-group opacity scheme with 12 bins, whereas the old model had a more approximate scheme and used only 4 bins. From Fig. 1 one can see that the major consequence of the improved scheme is a warmer upper photosphere, but unchanged at the top of the domain, resulting in a shallower temperature
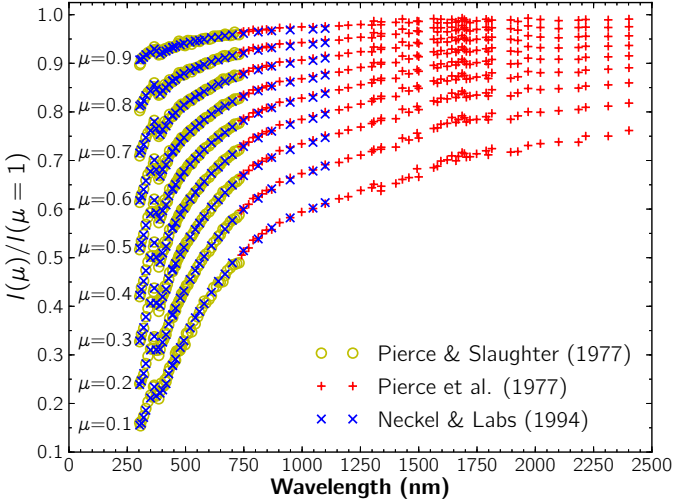
3

Fig. 2. Continuum centre-to-limb observations in the visible and near-infrared, from Pierce & Slaughter (1977), Pierce et al. (1977) and Neckel & Labs (1994). In the wavelength region where the sets overlap we use only Neckel & Labs (1994) for our comparisons.

gradient. This difference in temperature gradient will be most noticeable in the continuum centre-to-limb variations, which we discuss below.

The effect of NLTE in the PHOENIX models seems to be a cooling of the outer layers ($\sim 50\,\mathrm{K}$), with minor differences at other depths. This NLTE cooling goes in the opposite direction of the NLTE effects of other PHOENIX solar NLTE models (Short & Hauschildt 2005, 2009), where the NLTE effects cause a warming in the outer layers as naively expected from reduced line blanketing and surface cooling. This discrepancy seems to be associated with a different choice of atomic species treated in NLTE.

When debating on the advantages and disadvantages of employing a full 3D analysis to stellar photospheres, an important question to ask is: is the advantage brought by the 3D treatment of convection merely a realistic temperature stratification, or are the spatial and temporal inhomogeneities essential to derive accurate abundances? To answer these questions we included another 1D model in the testing, corresponding to the spatial and temporal mean structure of the 3D model (shown in Fig. 1). The horizontal averaging was done on surfaces of equal optical depth rather than geometrical height. This model, hereafter the ⟨3D⟩ model, will enable us to disentangle the effects of the mean temperature structure from the full 3D treatment.

3. Continuum centre-to-limb variations

3.1. Context

The centre-to-limb variations (CLV) of continuum intensities provide a sensitive probe of the solar photosphere. Because the continuum intensity is proportional to the local source function of continuum forming regions, its CLV is a measure of the temperature variation with depth (the closer to the solar limb, the higher up in the atmosphere). The variation of depth can be expressed in terms of $\mu \equiv \cos\theta$, where θ is the heliocentric viewing angle. Normalising $I(\mu)$ by the disk-centre value $I(\mu = 1)$, one has a measure of the temperature gradient of the photosphere around the continuum forming layers. This provides a robust test of models.

3.2. Observations

We make use of the CLV observations of Neckel & Labs (1994) and Pierce et al. (1977). They cover respectively the wavelengths between 303–1099 nm and 740.4–2401.8 nm, as shown in Fig. 2 for $0.1 \leq \mu \leq 0.9$. We compare the models with Neckel & Labs (1994) for $\lambda \leq 1099$ nm and Pierce et al. (1977) for $\lambda > 1099$ nm. Although not used in our comparison, the observations of Pierce & Slaughter (1977), covering the range 303.3–729.7 nm, are also plotted in Fig. 2 and agree very well with Neckel & Labs (1994).

Other CLV observations of longer wavelengths exist, such as Spickler et al. (1996). However, as is visible in Fig. 3 of Spickler et al. (1996), there is a considerable scatter between different observations, especially for $\lambda \gtrsim 4\ \mu$m. Because of these uncertainties we do not include them in this comparison.

3.3. Results and discussion

The model predictions were computed using our 3D LTE line formation code, which was used to obtain the predicted intensity at different inclinations. The intensities were computed for nine different values of $\mu \equiv \cos\theta$ and for each inclination except the vertical they were averaged over four φ-angles in the 3D case. These intensities were interpolated in μ for the inclinations shown. For the 3D model the intensities were computed for each of the 90 snapshots, the final value being the spatial and temporal average.

Results for the visible and near-infrared continuum centre-to-limb variation are shown in Fig. 3. A more compact comparison with Neckel & Labs (1994) is made in Fig. 4, where we plot the (normalised) absolute value of the difference between models and observations, averaged over wavelength for each μ value. Because of uncertainty in the observations (from the difficulty in finding continuum regions), this average is limited to the $400 - 1099$ nm wavelength region.

The results show an excellent agreement between the 3D hydrodynamical model of Asplund et al. (2009) and the observations, particularly when comparing with Neckel & Labs (1994). This agreement is visibly better even than that of the 1D Holweger & Müller model, which is quite remarkable given that it was empirically constructed to fit the centre-to-limb variations. It should be noted, however, that the observational and atomic data have improved much since the construction of the Holweger & Müller model. The application of modern opacities to the Holweger & Müller $T(\tau)$ stratification therefore results in discrepancies not present at its development in 1974. We note also that the spectral resolving power of the solar atlas employed by Holweger & Müller was poorer than the atlases available today, leading to less steep line cores and consequently a too high temperature was inferred from the spectral inversion process; rectifying this shortcoming results in a temperature structure very similar to the mean of the here employed 3D model (N. Grevesse, 2009, private communication). With the infrared observations the agreement with the models is slightly worse. However, observations at these wavelengths seem more uncertain, as can be seen by the increased scatter between data points. It is likely that this additional uncertainty affects the agreement with the models. Indicative of this is the region between 1500–1850 nm, where the observations show some scatter and are consistently higher than the model predictions.

The inclusion of a 10 mT magnetic field in the 3D model somewhat degrades the agreement with continuum centre-to-limb variation as seen in Fig. 3 and 4. This difference can be
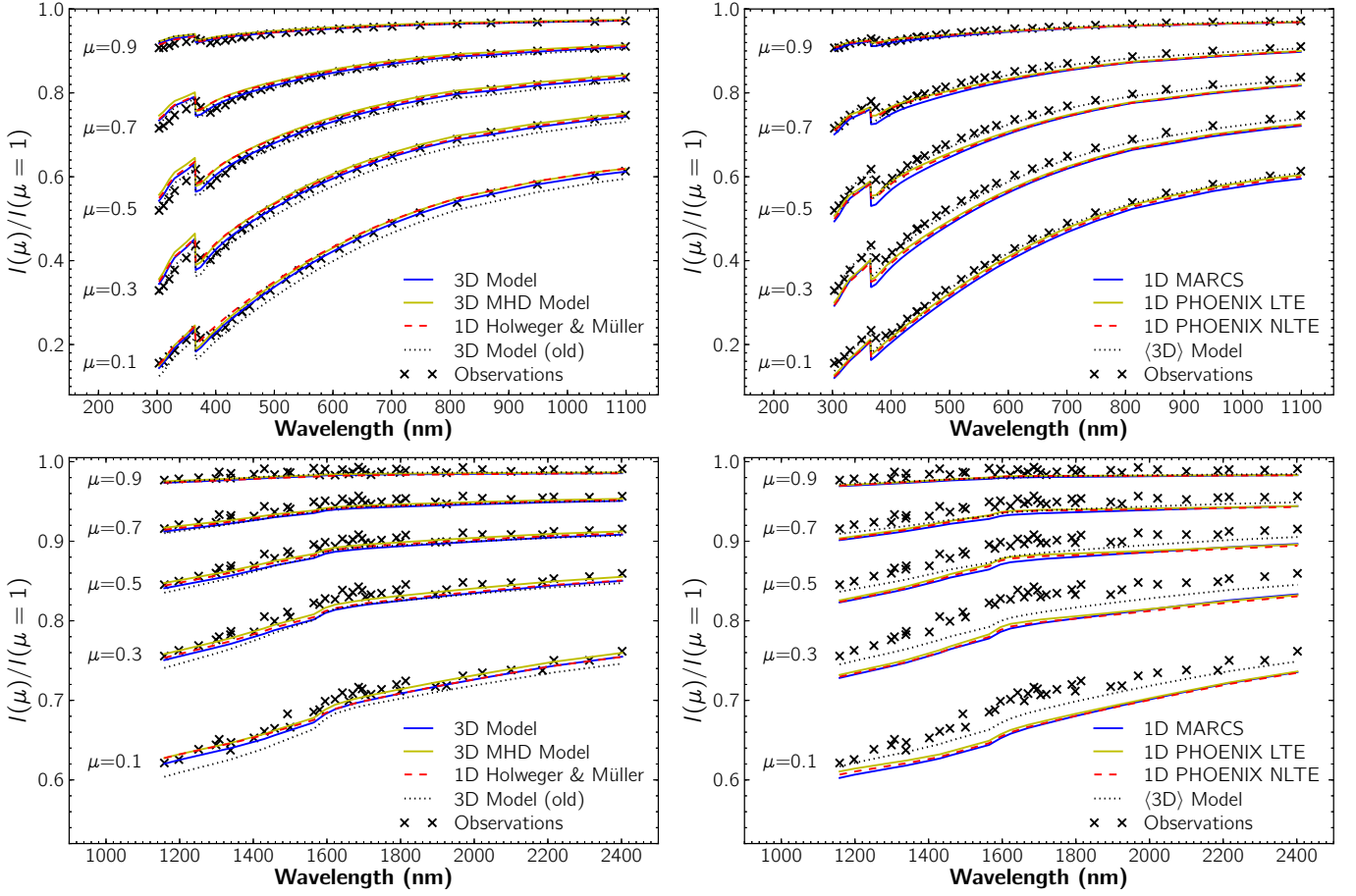
Fig. 3. Centre-to-limb variations in the continuum intensity. *Top panels:* comparison with the visible/infrared observations of Neckel & Labs (1994). *Bottom panels:* comparison with the near-infrared observations of Pierce et al. (1977), for wavelengths between 1158.35–2401.8 nm.

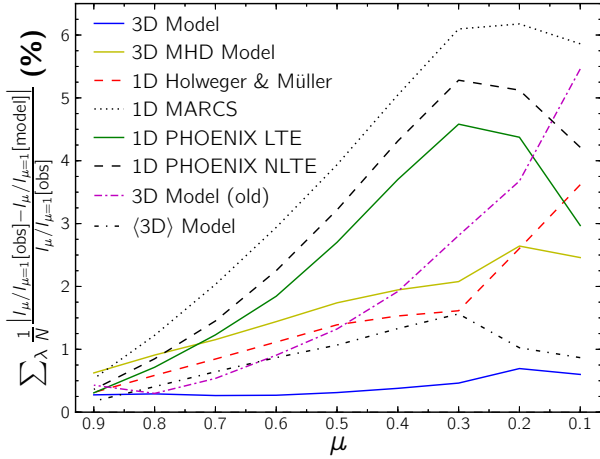**Fig. 4.** Normalised differences between observations and models in the centre-to-limb variation, averaged over wavelength as a function of μ (see text). Comparison with Neckel & Labs (1994) for $400 < \lambda < 1099$ nm.

traced to the shallower temperature gradient (Fig. 1) in the MHD model. The 3D MHD model performs slightly worse than the Holweger & Müller model.

The agreement with the theoretical 1D models is not as good. It is interesting to note in Fig. 4 that LTE models of

MARCS and PHOENIX have the same trend with μ, although the PHOENIX model performs slightly better. The results for the PHOENIX NLTE model depart only slightly from the LTE model results. The NLTE cooling of the outer layers seen in Fig. 1 causes a slightly steeper temperature gradient, which leads to a worse agreement with the observed centre-to-limb variations. The overall structure and dependence with μ remains essentially the same for both PHOENIX models as well as for the MARCS model, as seen in Figs. 3 and 4, due to the similarity in $T(\tau)$ for $-2 < \log \tau < 0$, the layers largely tested with continuum CLV.

Compared to other models, the differences between the 3D and $\langle 3D \rangle$ models are small, meaning that the mean temperature gradient is the main driver of the continuum CLV behaviour. Nevertheless, the 3D model predictions agree even closer with the observations, confirming the results of Koesterke et al. (2008), although we find a smaller "3D–$\langle 3D \rangle$" difference. Looking at Fig. 3, the $\langle 3D \rangle$ model lies slightly below the 3D model, in other words the effect of the atmospheric inhomogeneities increases $I(\mu)/I(\mu = 1)$. One would therefore expect that if spatial and temporal inhomogeneities were added to the Holweger & Müller model, its predictions would lie further above the observations. This indicates that the temperature gradient of the Holweger & Müller model is too shallow compared to the Sun.

We compare also with the old 3D model. While this is a realistic model that reproduces the observed line shifts and shapes

(Asplund et al. 2000a), its steeper temperature gradient has a noticeable effect on the continuum CLV. Its predictions fare worse when compared to the observations (but still better than the 1D models). We do not use this old model in the other observational tests.

In summary, the confrontation with continuum centre-to-limb observations reveal that the 3D hydrodynamical model is very realistic. The ⟨3D⟩ hydrodynamical model, the 3D MHD and the Holweger & Müller model all perform slightly worse with too little limb-darkening, while the MARCS and PHOENIX 1D theoretical models all predict much too strong centre-to-limb variation due to a too steep temperature gradient.

4. Absolute flux distribution

4.1. Context

An independent test of the temperature structure of solar models is provided by the observed absolute continuum fluxes. They provide an absolute temperature scale for the photospheric layers.

4.2. Observations

Several observations of the absolute solar flux/irradiance are available, either space-based (*e.g.* Colina et al. 1996; Thuillier et al. 2004) or Earth-based (*e.g.* Kurucz et al. 1984; Brault & Neckel 1987; Kurucz 2005, all obtained with the high-resolution NSO/Kitt Peak Fourier transform spectrometer). The space-based observations provide a spectrum clean of terrestrial absorption features, but at the cost of a lower spectral resolution.

To extract the observed continuum fluxes we used an absolute flux atlas divided by a normalised flux atlas (for the points deemed to be near the continuum). We used the Kurucz (2005) irradiances and normalised flux atlases, available for the interval of 300–1000 nm. They are a recent reduction of the data used to produce the atlas of Kurucz et al. (1984), and have been carefully adjusted to remove the telluric absorption features. The choice of using the Kurucz (2005) atlases instead of space-based observations was made because of its consistency between absolute and normalised fluxes.

If another irradiance atlas was used, the slight differences in line strengths or wavelength mismatches between the irradiance and normalised flux atlases would cause some scatter in the continuum fluxes, which would have to be smoothed out (see Ayres et al. 2006, where such an approach is followed). To avoid these uncertainties we use the Kurucz (2005) atlases that, being produced from the same data, have a consistent continuum determination and wavelength calibration between the irradiance and normalised flux atlases.

The observed continuum flux was obtained as follows. First the irradiance is converted to flux at the solar surface using the multiplicative factor of $[(1 \text{ AU})/R_\odot]^2 = 46202$. Then the wavelengths of the (near) continuum points in the normalised flux are identified. These are defined as $F_\lambda > 0.99 \, F_\lambda^{\mathrm{m}}$, where F_λ^{m} is a local maximum in 5 nm windows. Finally the absolute flux is divided by the normalised flux for the continuum high wavelengths. This ratio is linearly interpolated to a coarser resolution of 1 nm. A few spurious points were manually removed. The Kurucz (2005) irradiance has been normalised to the total solar irradiance of Thuillier et al. (2004). Following the discussion in Ayres et al. (2006) we have readjusted the continuum fluxes by −0.4%, to account for the more accurate total solar irradiance
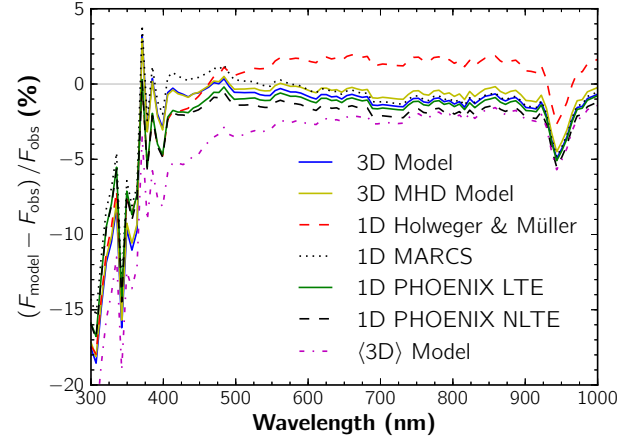

Fig. 6. Continuum flux differences between the models and the observations. For $\lambda \lesssim 450$ nm the observations are not very reliable because of difficulties in continuum placement. The feature at $\lambda \approx 950$ nm is likely to be caused by uncorrected telluric absorption in the observations.

of Kopp et al. (2005). In Fig. 5 we show the original observed fluxes along with our derived continuum fluxes.

4.3. Results and discussion

The predicted continuum fluxes were computed with our LTE line formation code. The disk-integrated fluxes are computed using eight μ-angles and four φ-angles, a total of 32 angles; in addition the disk-centre intensity was calculated. The μ-angles and weights are taken from a Gaussian quadrature and are thus not identical to those used for the centre-to-limb study presented in Sect. 3. For the 3D models the fluxes are computed and spatially and temporally averaged for the 90 simulation snapshots. For 1D models we used the same nine μ-angles, including disk centre.

The resulting fluxes for all models are shown in Fig. 5. A differential comparison is also shown in Fig. 6. There is an excess of flux for $\lambda \lesssim 370$ nm in the observed continuum fluxes when compared with the original flux (with lines) or any of the models. This difference, also evident in Fig. 6, seems to be caused by a too high continuum placement in the normalised Kurucz (2005) flux atlas (*e.g.*, higher than in Kurucz et al. 1984; Brault & Neckel 1987). Being a region very crowded with lines, the discrepancy between observations and models likely arises from the difficulty in finding the continuum level (which is systematically overestimated), and not from a failure of the models.

Overall, the 1D MARCS model is the best at reproducing the observed flux, although the differences between different models are small. The Holweger & Müller predictions are consistent with a correct T_{eff}, but its flux distribution has a different shape. Between 350–450 nm it has less flux, but beyond $\lambda \approx 500$ nm it shows an excess flux when compared with the observations. The 3D model consistently predicts slightly less flux than the observations, but nevertheless has a good agreement and reproduces the flux distribution well. The PHOENIX LTE model behaves similarly to the 3D model, but with less flux at the peak of the distribution. The differences between the NLTE and LTE PHOENIX models are very small, not surprising given that very little flux comes from the regions where the NLTE cooling is more efficient.

In the comparison between ⟨3D⟩ and 3D, one finds considerably less flux coming from the ⟨3D⟩. This is because the τ_{500}
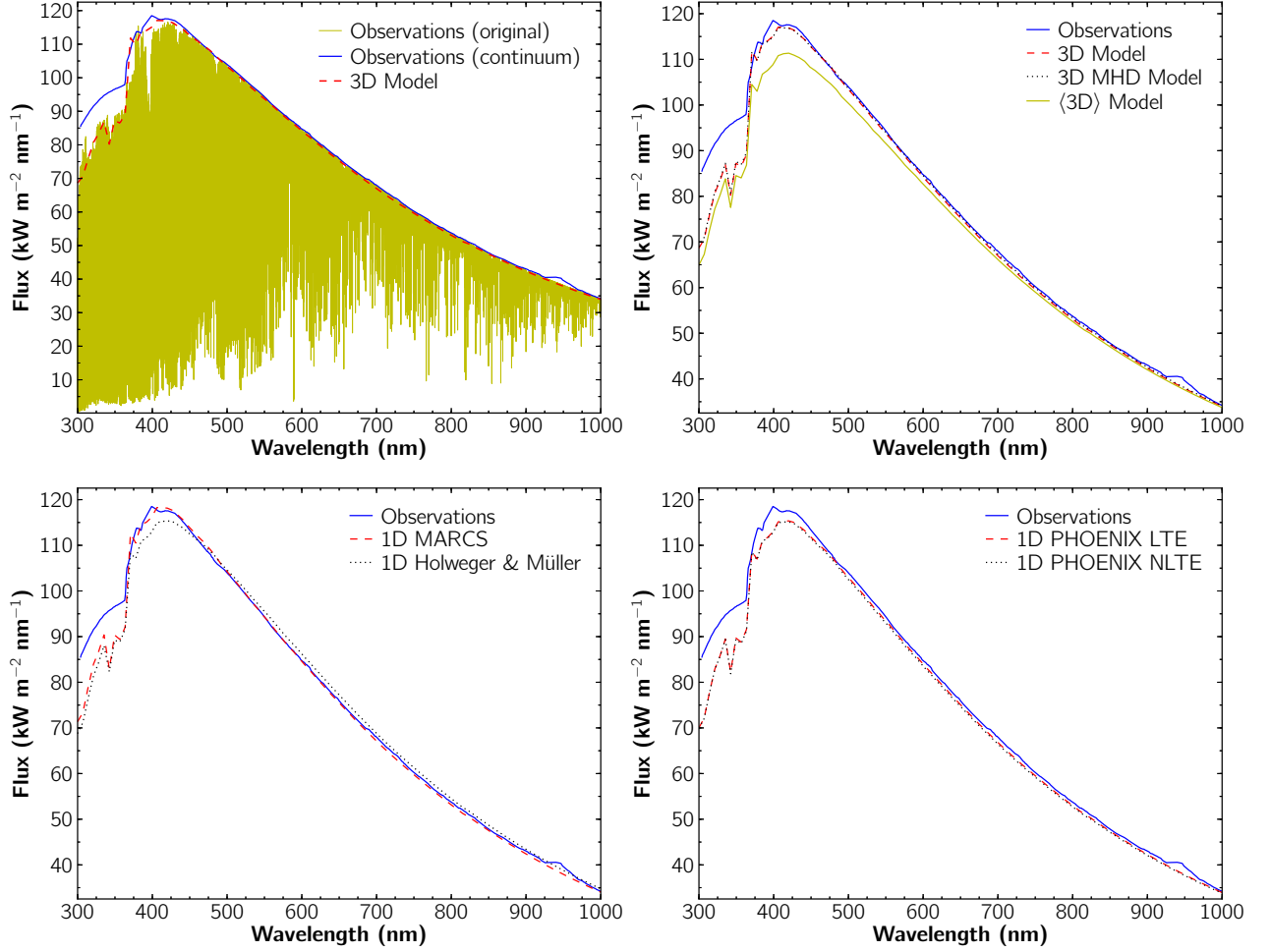
Fig. 5. Absolute fluxes for the models and observations. *Top left:* Comparison of the original fluxes of Kurucz (2005), our derived continuum fluxes (see text), and the predicted continuum fluxes from the 3D model. *Other panels:* observed continuum fluxes and predicted continuum fluxes for several models.

iso-surface averaging per construction does not preserve the effective temperature. In the 3D model a reasonable amount of flux will be emitted from the snapshots with a higher T_{eff}, as $F \propto T_{\mathrm{eff}}^4$. The $\langle 3D \rangle$ model is not averaged in T^4 and will have a lower T_{eff}. This is perhaps one of the strongest arguments against constructing modified 1D models that recover the mean structure of realistic convection: T_{eff} will not be preserved. One can argue that the $\langle 3D \rangle$ model could be averaged in T^4. The issue of how to average 3D models will be addressed in a forthcoming paper; tests performed so far indicate that for the purposes of line formation there is little difference between T- and T^4-averaged models.

The flux differences in Fig. 6 are quantified in Table 1, where we show the root mean square differences between the models and observations, summed over the region of 375–975 nm, and normalised to the results from the 3D model. Again one can see that the MARCS model gives the best flux predictions, followed by the 3D MHD and 3D HD model. The PHOENIX and Holweger & Müller models perform similarly while interestingly the $\langle 3D \rangle$ model shows the largest differences.

Table 1. Root mean square differences between the models and the observations, between 375–975 nm.

Model	$\Delta F^2 / \Delta F_{3D}^2$
3D Model	1.00
3D MHD Model	0.78
1D Holweger & Müller	3.31
1D MARCS	0.67
1D PHOENIX LTE	2.96
1D PHOENIX NLTE	3.98
$\langle 3D \rangle$ Model	15.00

5. Hydrogen lines

5.1. Context

The first two lines in the hydrogen Balmer series lines, Hα and Hβ, have traditionally been used for temperature determination in late-type stellar atmospheres. Their wings are typically formed in the region around $-2 \lesssim \log \tau_{500} \lesssim 0.5$, deeper than most of the other spectral lines in late-type stars (Fuhrmann et al. 1993), while their cores are formed in the chromosphere. The

7

fact that for late-type stars these lines are more sensitive to temperature than they are to gravity, metallicity or indeed the hydrogen abundance has established them as a popular tool for temperature determinations (Fuhrmann et al. 1993, 1994; Barklem et al. 2002; Behara et al. 2009). The shapes of hydrogen lines, because of their large depths of formation, are however dependent on the convection efficiency (e.g. Ludwig et al. 2009).

Also relevant to this work are the lines of the Paschen series. With E_{low} = 12.088 eV, these lines are formed even deeper than the Balmer series. However, owing to their longer wavelengths and lower level populations (and consequently weaker wings), they have not been used as extensively as Hα and Hβ. We include in our comparisons the Paγ and Paβ lines, which are just inside the wavelength range of the high-resolution FTS atlases of the Sun.

Departures from LTE can be important for hydrogen lines (Przybilla & Butler 2004; Barklem 2007). While the far wings are formed in deep, hot regions where LTE conditions largely prevail, the LTE assumption is no longer valid for the cores of the strong lines. However, Barklem (2007, hereafter PB07) shows that in the case of Balmer lines, the NLTE effects can even extend into the line wings, depending on the adopted rates for collisions with H atoms. In their LTE and NLTE study of solar H lines, Przybilla & Butler (2004) find a weakening of the NLTE effects for the Paschen lines. To obtain realistic line profiles, we perform NLTE line synthesis as detailed in Sect. 5.3.

5.2. Observations

For the solar observations of hydrogen lines we use the high-resolution FTS normalised flux atlas of Kurucz (2005), and the FTS disk-centre atlas of Brault & Neckel (1987). For our comparison we study the Balmer lines Hα and Hβ along with the Paschen lines Paγ and Paβ.

In addition to the disk-centre atlas, Brault & Neckel (1987) have also made available a flux atlas. For the lines considered here it is essentially identical to the Kurucz (2005) atlas, which is not surprising since it was produced from the same raw FTS data. However, we would like to point out that the flux atlas of Brault & Neckel (1987) has a slightly different normalisation, with a lower continuum level. This difference amounts to $\approx 0.4-0.6\%$ of the normalised flux. For a consistent comparison of the flux and disk-centre intensity profiles, we have increased the continuum level of the Brault & Neckel (1987) disk-centre atlas by 0.5%, so that it matches that of the Kurucz flux atlas. This choice of continuum will not affect the differential results between models, only their relative standing to the observations. Our choice of continuum falls on the Kurucz atlas because it is a more recent (and hopefully more accurate) reduction of the same data. However, the 0.5% continuum error, which corresponds to a T_{eff} difference of ≈ 40 K in the Balmer lines, is within the uncertainties of this analysis (including the errors from the observations, model, broadening, NLTE effects, etc.). Barklem et al. (2002, hereafter PB02) discuss in detail the several uncertainties associated with the calculation of hydrogen lines, to which one should also add the uncertainties from the choice of hydrogen collisions (PB07).

The Paβ line, at 1281.8 nm, is not available in the Brault & Neckel (1987) disk-centre atlas nor in the Kurucz (2005) atlas, which only extend to 1250.8 nm and 1098 nm, respectively. For this line we use only the Kurucz et al. (1984) flux atlas, which is based on the same raw data as Kurucz (2005), but covers a somewhat larger wavelength range. However, the normalisation of the Kurucz et al. (1984) atlas in the region around Paβ is prob-
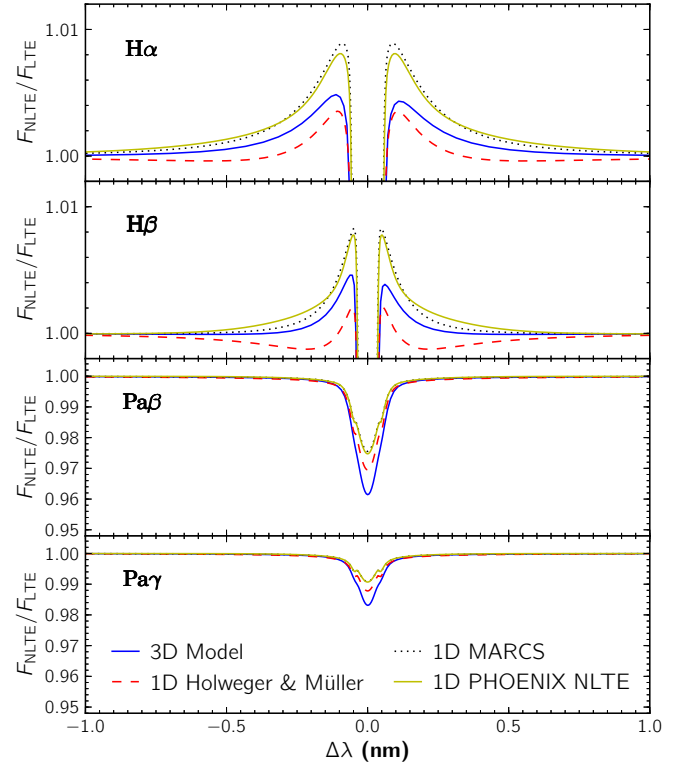

Fig. 7. Flux ratios between NLTE and LTE line profiles, for the hydrogen lines considered and for the different models. Wavelength difference $\Delta\lambda$ measured from the line core. Results for the PHOENIX NLTE model atmosphere (not shown) are indistinguishable from the corresponding LTE model in this figure. For Hα and Hβ the ratio for the line core is not shown, to emphasise the NLTE effects in the wings, formed in the photosphere.

lematic. To compensate for this we re-normalised this part of the atlas, finding a suitable continuum level from a polynomial fit to nearby continuum high points.

5.3. Synthetic profiles

The computation of the H line profiles has been done allowing for departures from LTE. We use a 20-level hydrogen model atom (19 H I levels plus continuum) based on the atom of PB07, which was adapted from Carlsson & Rutten (1992). The collisional cross-sections by PB07 are employed. These include inelastic collisions with electrons and hydrogen atoms (using the cross sections of Soon 1992, assuming $E = E_{cm}$), mutual neutralisation and Penning ionisation. To speed up the calculations, especially in the 3D case, we have neglected the bound-bound transitions starting with a lower level $n >= 6$, thus including only 80 bound-bound transitions (out of a total of 171 in the original model atom). Tests with 1D models show that the effects of removing these lines are negligible.

Line profiles were computed using our LTE code and the MPI-version of the MULTI3D code (Leenaarts & Carlsson 2009). To save computational time, full 3D NLTE line formation was performed only on eight snapshots of the simulation. Using MULTI3D to compute the LTE and NLTE line profiles we obtain the wavelength-dependent NLTE/LTE ratio, averaged for these eight snapshots. The final line profiles are then obtained by using our LTE code to compute the 3D LTE line profiles for all the 90 snapshots in the simulation and then multiplying them

by the average NLTE/LTE ratio for each wavelength. The very small variation between the NLTE/LTE ratios for the eight snapshots indicates that this procedure is a very good approximation. For the 1D models the computational requirements are unimportant, and they are not time-dependent, but for consistency we use the same procedure of the NLTE/LTE ratio for each model.

We present the hydrogen line profiles for flux (disk-averaged intensity) and disk-centre intensity. After convergence of the level populations, the flux profiles have been computed using a total of 32 inclined rays (8 μ-angles, 4 φ-angles); a vertical ray was used for the disk-centre intensity profiles. A rotational velocity of $v_{\rm rot} = 1.8$ km s^{-1} was used for the disk-integration. For the 1D models a microturbulence of 1.0 km s^{-1} and a macroturbulence of 2.5 km s^{-1} (consistent with values derived from a sample of Fe I lines) were used, although these choices have little effect on the H line wings. The hydrogen opacity is calculated using the HLINOP routine (Barklem & Piskunov 2003; PB07). This ensures a proper treatment of self-broadening (following Barklem et al. 2000) and Stark broadening (following Stehlé & Hutcheon 1999).

5.4. Results

We present the H line profile results in Figs. 8 and 9 for flux and disk-centre intensity respectively. To quantify the differences between the observations and synthetic profiles a χ^2 approach was carried out in the following way. First, regions close to the line-wing continuum were identified in the observations. These regions are indicated in grey in Fig. 8 and were used for both disk-centre and flux profiles. The differences between observations and synthetic profiles were calculated in these regions. For a tangible quantification of these differences, we have estimated the change in $T_{\rm eff}$ that each model would need to best match the observations of each line. This was achieved using several MARCS models with $5500 \lesssim T_{\rm eff} \lesssim 6000$ (all with $\log g = 4.44$ and [Fe/H]=0.0), whose hydrogen line profiles were calculated and used to derive an intensity ratio $I(\lambda, T_{\rm eff})/I(\lambda, T_{\rm eff} = 5777 \,{\rm K})$. This intensity ratio was then multiplied by each synthetic line profile, to obtain the approximate line profile of each model for an arbitrary $T_{\rm eff}$. This approach is only an approximation, as the variation of the hydrogen lines with $T_{\rm eff}$ will vary from model to model. Nevertheless, it is good enough for this purpose. Using an optimisation procedure we calculated the $T_{\rm eff}$ that, for each set of observations, minimises the reduced χ^2, defined as $1/N \cdot \sum (I_{\rm obs} - I_{\rm model})^2 /\sigma^2$, where N is the number of wavelength points minus the degrees of freedom (one, in this case) and σ^2 the measurement error, which we assume to be constant. These results are shown in Table 2. Also shown is the reduced χ^2 for the $T_{\rm eff}$ adjusted line profiles, summed over all the observations, and normalised by the value for the 3D model. This gives a measure of the goodness of the fits.

To help connect the differences between models and observations with the temperature structure of the models we calculated the temperature gradient, defined as:

$$\nabla T \equiv \frac{{\rm d} \log_{10} T}{{\rm d} \log_{10} \tau},\qquad(1)$$

with the optical depth τ evaluated at 500 nm. For the disk-centre intensity profiles of Hα, Hβ, and Paγ, the temperature gradient was calculated for three regions, corresponding to the formation regions of three wavelength points in the wings of the line profiles. The three wavelength points λ_1, λ_2, and λ_3 are defined as the velocity shifts from the line cores of respectively -609,
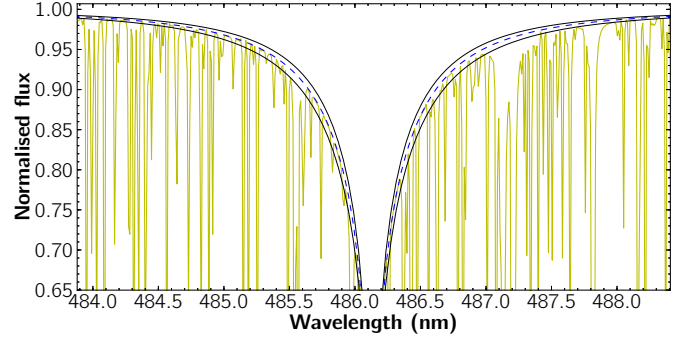

Fig. 10. Effect of multiple blends on the wings of Hβ. Two synthetic Hβ flux line profiles of a 1D MARCS model are shown: using only hydrogen (*dashed blue line*), and including 220 other atomic lines (*solid yellow line*). Results for MARCS models with $T_{\rm eff} + 100$ K and -100 K are also shown (*lower and upper thin black lines, respectively*).
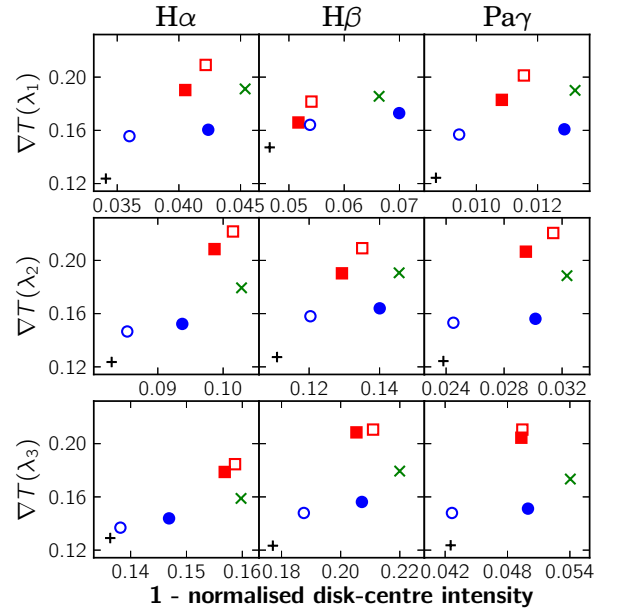

Fig. 11. Temperature gradient at depths corresponding to the formation layers of three wavelengths. Models shown are the 3D (*blue filled circles*), ⟨3D⟩ (*blue open circles*), Holweger & Müller (*black plus signs*), MARCS (*green crosses*), PHOENIX LTE (*red filled squares*), and PHOENIX NLTE (*red open squares*).

-316, and -196 km s^{-1}. For each wavelength the contribution function is calculated, and ∇T taken as the average of a small depth range around the peak of the contribution function. For the 3D model this is done on a column by column basis, and the final value for ∇T taken as the mean of all the 1D columns in all snapshots, weighted by the continuum intensity of each column. The resulting values of ∇T, plotted against line depression of the disk-centre profiles, are shown in Fig. 11. (The wavelength points used are also shown as vertical lines in Fig. 9.)

5.5. Discussion

5.5.1. NLTE effects

The NLTE effects on the hydrogen lines are quantified in Fig. 7, where the NLTE/LTE flux ratios are shown. The main conse-
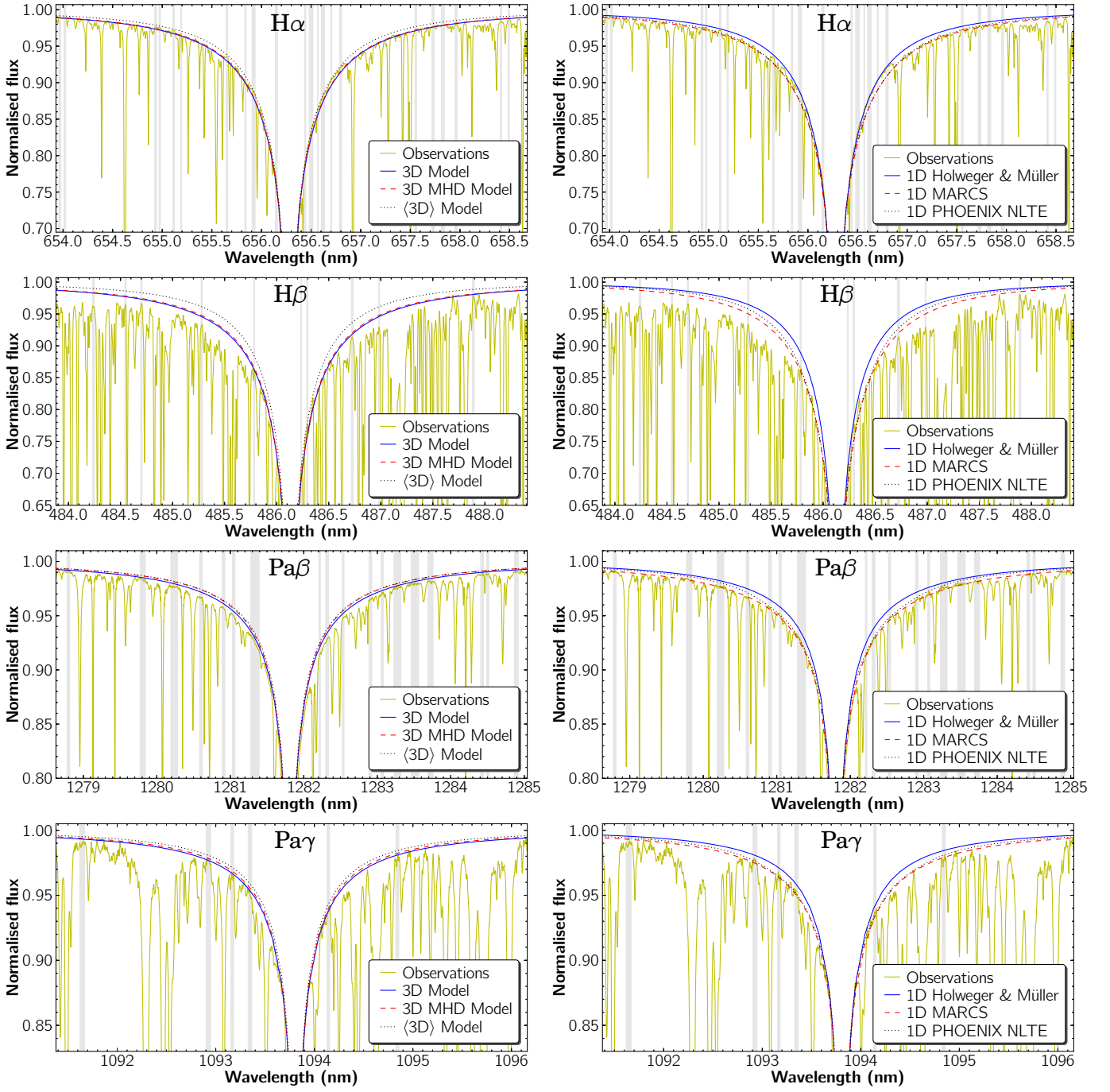
Fig. 8. Normalised flux profiles for the H lines Hβ, Hα, Paγ, and Paβ. Compared with the solar observations of Kurucz (2005). The exception is Paβ, where the Kurucz et al. (1984) atlas is used because the Kurucz (2005) atlas does not cover these wavelengths. For this line the continuum was re-normalised (see text). Synthetic profiles were computed in NLTE. The regions used in the χ^2 analysis are indicated in grey.

quence of the NLTE effects in the H lines is a deeper core, which is of less interest here since it is little sensitive to the photospheric stratification but rather determined by chromospheric conditions. However, as noted by PB07 and shown in Fig. 7, there are non-negligible effects on the wings of the Balmer lines, making the wings weaker compared to the LTE case, at least with our particular choice of H collisions for the NLTE calculations (see discussion in Barklem 2007)). For the Paschen lines the NLTE effects are much smaller, and they cause only a deeper core.

Compared with the results of PB07, we find a weaker NLTE effect in the wings of the Balmer lines. Using a similar MARCS model, the same recipe for the collisional rates, and a similar code for NLTE radiative transfer (MULTI), PB07 finds a maximum NLTE excess flux on the wings of Hα of about 2.5%, whereas we find around 1%. The origin of this difference is our inclusion of line-blanketing for photo-ionisation transitions. This additional source of opacity, in particular in the UV, leads to a decrease in the photo-ionisation rates, bringing the level populations closer to LTE. Line-blanketing was not included in the
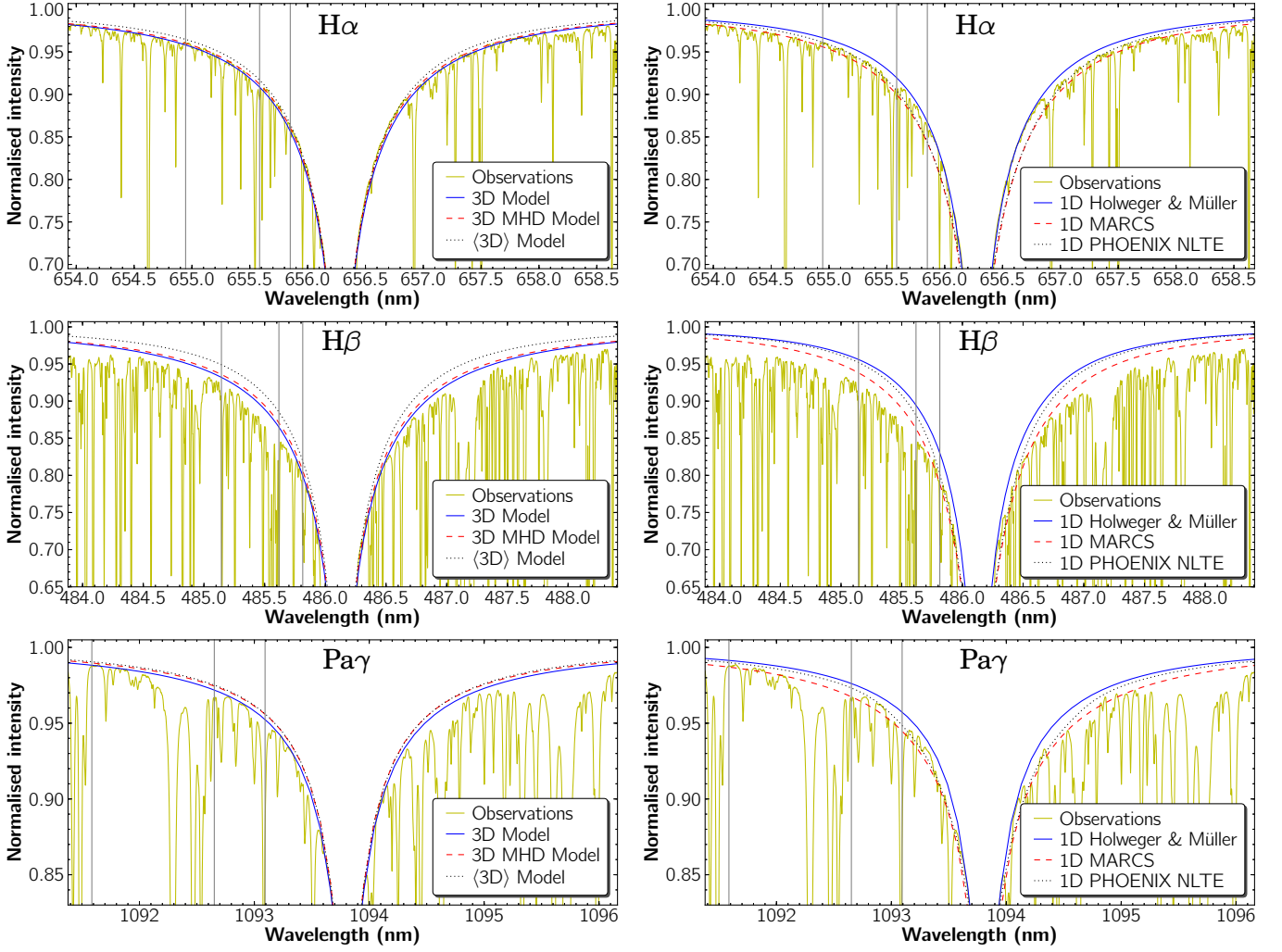
Fig. 9. Normalised disk-centre intensity profiles for the H lines Hβ, Hα, and Paγ. Compared with the solar observations of Brault & Neckel (1987). Synthetic profiles were computed in NLTE. The three vertical lines correspond to the wavelengths λ_1, λ_2, and λ_3 used to make Fig. 11.

Table 2. Estimated effective temperature differences between the models and observations.

	ΔT_{eff} [K]							$\langle \Delta T_{\text{eff}} \rangle^*$	$\chi^2/\chi^2_{\text{3D}}{}^\dagger$
Model	Hα		Hβ		Paβ	Paγ		[K]	
	Flux	$I(\mu = 1)$	Flux	$I(\mu = 1)$	Flux	Flux	$I(\mu = 1)$		
3D Model	−50	−53	10	11	66	41	25	**7**	1.0
3D MHD Model	−37	−21	29	49	105	69	66	**37**	1.1
⟨3D⟩ Model	18	−1	153	136	120	114	69	**87**	2.1
1D Holweger & Müller	−10	7	174	196	126	106	82	**97**	2.5
1D MARCS	−98	−156	−9	−30	−17	−7	−74	**−56**	2.3
1D PHOENIX LTE	−75	−115	50	81	53	28	−6	**2**	3.7
1D PHOENIX NLTE	−77	−128	37	59	36	14	−27	**−12**	3.7

* $\langle \Delta T_{\text{eff}} \rangle$ denotes the simple mean of the seven ΔT_{eff} values.
† The χ^2 value is the sum of the individual reduced χ^2, normalised by the 3D model results.

MULTI version 2.2 used by PB07, but is in MULTI version 2.3 and in MULTI3D. If we switch off line-blanketing in MULTI3D, we obtain essentially the same results as PB07.

5.5.2. Temperature gradient

The results for the temperature gradient ∇T in Fig. 11 show a correlation between the line strength and ∇T. Typically, the higher ∇T, the stronger the normalised line profiles as one would naively expect. In most cases the Holweger & Müller, ⟨3D⟩, and PHOENIX models seem to fall on the same linear rela-

11

tion, while the 3D and MARCS models, showing a similar relation, have a lower ∇T for similar line strengths. When compared with the Holweger & Müller, the higher ∇T of the PHOENIX and MARCS models is consistent with their predictions for the wings being much stronger than those of the Holweger & Müller model. However, the ∇T alone is not enough to explain the differences between MARCS and PHOENIX models, with the latter having usually a larger ∇T but a smaller T_{eff} correction.

Between the PHOENIX models, the differences in the ΔT_{eff} corrections are consistent with what is seen in ∇T: the NLTE model has a higher ∇T, and consequently also larger ΔT_{eff} corrections (positive or negative). Between the Balmer lines, the temperature gradient of the PHOENIX models is lower for Hβ. This is a consequence of the abrupt change in temperature gradient that these models show at $\log_{10} \tau_{500} \approx 0.3$ (see Fig. 1). Hβ, being formed deeper, is formed mostly on the flatter side of this knee, whereas Hα is mostly formed on the steeper side of the knee. This helps explain why these models predict Hα to be too strong, while predicting Hβ to be too weak. The MARCS model, that does not show this 'knee', predicts both Balmer lines to be stronger than the observations.

Perhaps the most interesting departure from the ∇T relation with line strength is the case of the 3D and $\langle 3D \rangle$ models. These models have a very similar ∇T, but the predictions of the $\langle 3D \rangle$ model are of much weaker lines than the 3D model. For all the tests in the hydrogen lines, the $\langle 3D \rangle$ is closer to the Holweger & Müller model. This indicates that the spatial and temporal variations of the 3D model are important to describe the shapes and strengths of the hydrogen lines due to the very large non-linearity in the line formation of these high excitation lines. Ludwig et al. (2009) reached similar conclusions in their analysis of H lines using 3D CO^{5}BOLD models and stressed that the "3D effect" in terms of T_{eff} depends sensitively on the particular adopted mixing length parameters in the 1D model atmospheres.

5.5.3. Comparison with observations

Overall, no single model seems to reproduce the observations perfectly. All of the models tested require different T_{eff} corrections for different lines. This is particularly true for the Balmer lines, whose T_{eff} corrections often have opposite signs. Most models predict a too strong Hα line, while at the same time Hβ is not strong enough compared to observations. Given the higher number of blending features in Hβ one wonders if this effect is not caused by a continuum depression because of all the blends, not considered in the hydrogen-only synthetic profiles. To make sure the single-line approximation is valid, we performed spectral synthesis in this region for the 1D MARCS model, including 220 additional atomic lines in the calculations. Data for these lines were extracted from the VALD database (Piskunov et al. 1995; Kupka et al. 2000), selected as the strongest lines in this region. The results, shown in Fig. 10, indicate that the blends have a negligible effect in lowering the local continuum of the wings of Hβ, validating our single-line approximation.

The MARCS model predicts line profiles that are too strong in all cases, requiring a negative T_{eff} correction. The Holweger & Müller model suffers from the opposite effect: its predictions are considerably weaker than the observations, except for Hα, when they seem to be just about right. Both PHOENIX models seem to behave similary to the MARCS in Hα, but then have mostly positive T_{eff} corrections for the Paschen lines, a likely consequence of the variations in their temperature gradients in the deepest atmospheric layers, as discussed before. The 3D model predicts

the wings of Hα to be stronger and the Paschen line wings to be weaker than the observations, but its prediction for Hβ agrees very well with the observations.

The Balmer line results of the MARCS model can be compared with the LTE results of PB02. Although PB02 used an earlier MARCS model (Asplund et al. 1997), slightly different regions for χ^2 with the 1984 Kurucz flux atlas, and possibly slightly different input physics, it is nevertheless relevant to compare our results with theirs. Calculating the LTE MARCS Balmer profiles for the same flux atlas used by PB02, we derive a T_{eff} from each Balmer line. For Hβ our determined T_{eff} is in good agreement with PB02, but for Hα our T_{eff} is 75 K lower. While difficult to pinpoint an exact cause, this difference is likely to come from differences in the input physics used to calculate the line profiles.

Except for the $\langle 3D \rangle$ model, the synthetic profiles of the theoretical models agree better for flux than disk-centre intensity. The biggest difference in fitted T_{eff} for intensity/flux is found for Hα in particular for the MARCS model. This is a hint of shortcomings in the description of the solar temperature profile in the deeper layers of these models, as the disk-centre intensity profiles are formed deeper than the flux profiles. The Holweger & Müller model shows a different trend: its predictions generally agree better for the intensity profiles, and the difference between the corrections for flux and intensity is smaller. It is reassuring to find that the 3D model gives the smallest variation between flux and intensity corrections, again an indication of a realistic temperature profile.

For all the line profiles considered, the 3D and the PHOENIX LTE and NLTE models give the smallest variation for $\langle \Delta T_{eff} \rangle$, the mean of the T_{eff} corrections. While a measure of the internal consistency of the models, $\langle \Delta T_{eff} \rangle$ is a simple approximation that does not take into account the shape of the line profiles. With the highest χ^2 values of the models tested, the predictions of the PHOENIX models do not reproduce the shape of the observed line profiles very well, while the 3D model performs the best also in this regard. The Holweger & Müller is the model with the highest $\langle \Delta T_{eff} \rangle$. The difference $\langle 3D \rangle - 3D$ is significant for the hydrogen lines: a $\langle \Delta T_{eff} \rangle$ difference of about 80 K, and a worse χ^2 for the $\langle 3D \rangle$ model. Overall, the 3D model stands out from the other models. While not perfectly describing the observations, in particular Hα and Paβ, it gives the smallest variation in T_{eff} between flux and intensity profiles, one of the smallest variations of T_{eff} among the different lines, and its predicted line shapes agree very well with the observations as evidenced by the lowest χ^2 among the models.

6. Continuum intensity distribution

6.1. Context

Another relevant diagnostic of our simulations is the continuum intensity distribution and contrast, ΔI_{rms}, of the granulation. A comparison of these with observations will reveal how well the simulations capture the differences in radiative transfer between the up- and the downflows.

6.2. Observations

Because solar observations are made with instruments with a finite resolution and subject to other effects such as straylight, it is difficult to ascertain what the solar intensity distribution and granulation contrast really is. All instrumental effects need to be carefully considered in order to make meaningful comparisons
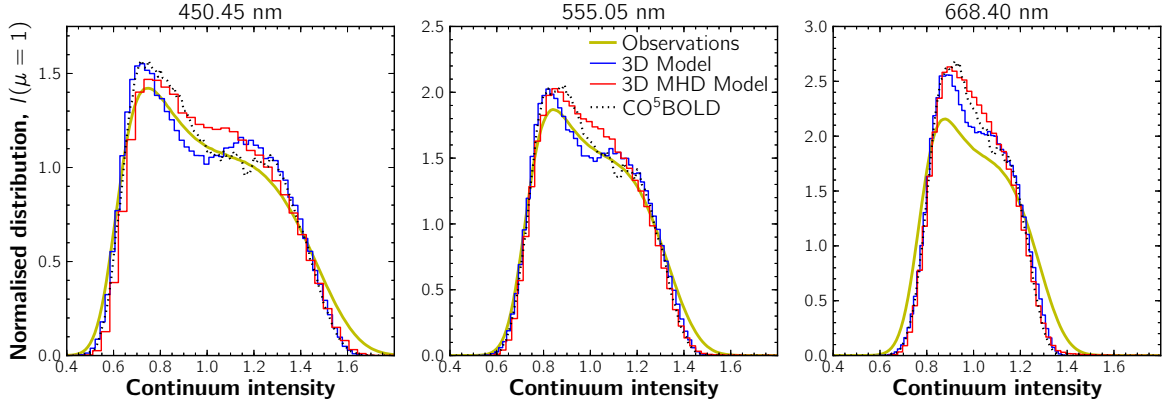
Fig. 12. Continuum intensity distributions for the solar disk-centre for three wavelengths, as a function of the normalised continuum intensity for the 3D model (*thin blue solid line*), the 3D MHD model (*thin red solid line*), compared with observations (*thick yellow line*). Also shown is the prediction from the CO⁵BOLD 3D solar model (*dotted line*) taken from Fig. 5 of WR09. The intensity distributions for our 3D models were averaged over all the snapshots.

with 3D simulations, but accurately compensating for all optical and straylight effects is a difficult task.

Several studies have found the observed $\Delta I_{\rm rms}$ to be lower than that of the 3D simulations even after consideration of atmospheric and instrumental seeing effects (*e.g.* Uitenbroek et al. 2007; Kiselman 2008). For the Hinode Solar Optical Telescope (SOT, Tsuneta et al. 2008), Danilovic et al. (2008) have characterised in detail the instrumental effects for the Spectro-Polarimeter (SP) instrument, and Wedemeyer-Böhm & Rouppe van der Voort (2009, hereafter WR09) have done the same for its Broad Filter Imager (BFI). Both of these studies find that when the instrumental degradation is carefully modelled, the continuum intensity distribution and $\Delta I_{\rm rms}$ agree very well with the predictions from 3D models. WR09 in particular test several types of 3D models and find an overall good agreement with the space-based observations of Hinode/SOT.

For our comparison we use the Hinode/BFI observations employed by WR09. These observations were taken in three BFI channels in three wide-band filtergrams with central wavelengths of 450.45 nm, 555.05 nm, and 668.40 nm. The FWHM of the filter transmission profiles are respectively 0.22 nm, 0.27 nm, and 0.31 nm. The observations were obtained in the period between November 2006 and February 2008. For our comparison, we use the values from WR09 for the observations deconvolved with the instrumental profile (in an attempt to cancel out the image degradation), and compare them with the raw predictions from the simulations (no image degradation applied).

6.3. Results and discussion

We compare our 3D models with the observations and the CO⁵BOLD 3D model results from WR09. CO⁵BOLD is independent of the stagger-code we employed, using different numerical methods and atomic physics. Results for the disk-centre intensity distributions at three wavelengths are shown in Fig. 12. The $\Delta I_{\rm rms}$ values for the same wavelengths are given in Table 3. Here $\Delta I_{\rm rms}$ was calculated in the same way as WR09, following their equation (1), and averaged in time for all the snapshots considered.

We find that our 3D model reproduces the observations well. Its $\Delta I_{\rm rms}$ is slightly higher than for the CO⁵BOLD model, but otherwise results from the two models are very close, which is encouraging. The observed contrast is slightly higher than

predicted from our 3D model, probably because (as noted by WR09) the observations seem to span a wider range of intensities and in particular have a more pronounced 'tail' at high intensities. The double peaked structure of the intensity distribution represents the brightness from the inter-granular lanes (highest peak) and the granules (lower peak). Compared to the observations, the 3D model has a pronounced minimum in the distribution between the two peaks, which can be attributed to the lack of magnetic features (such as bright points, ribbons, and other structures). The 3D MHD model, on the other hand, has a distribution that is closer to the observed, with no local minimum between the two peaks. This is probably because the bright points and other structures tend to blur the sharp transition from granule to inter-granular lane, and populate the distribution with intensities below the peak attained in the hottest granules. The reduced contrast of the MHD model is consistent with its different stratification and shallower temperature gradient (warmer upper layers), when compared to the 3D model with no magnetic fields. Nevertheless, the 3D MHD model still does not show the high-intensity tail seen in the observations.

Table 3. Disk-centre $\Delta I_{\rm rms}$ for the deconvolved observations, our 3D model and the CO⁵BOLD 3D model.

$\Delta I_{\rm rms}$ (%)	λ (nm)		
	450.45	555.05	668.40
Observations[a]	26.7 ± 1.3	19.4 ± 1.4	16.6 ± 0.7
3D model	25.4 ± 0.8	18.6 ± 0.6	14.3 ± 0.5
3D MHD model	24.1 ± 0.8	17.7 ± 0.6	13.7 ± 0.5
CO⁵BOLD[a]	25.0 ± 0.1	18.1 ± 0.1	13.8 ± 0.1

[a] From WR09.

7. Fe abundances and line shapes

7.1. Context

The detailed shapes of photospheric absorption lines carry crucial information about the atmospheric conditions. The strengths of weak spectral lines mainly reflect the temperature structure

13

in the line-forming regions, at least in the framework of LTE, while stronger lines become increasingly sensitive to the velocity field due to desaturation effects. All observed spectral lines show asymmetries due to the presence of (anti-)correlations between temperature and velocities in the up- and downflows. The warmer upflows lead to high continuum intensities and blueshifted, strong line profiles due to the steep temperature gradients while lines from downflows are red-shifted, weak and have low continuum intensities; the line strengths are normally heavily biased towards the upflows, also because of their typically larger area coverage. The spatially averaged profiles thus become skewed, resulting in lines with typically blue-shifted cores and C-shaped bisectors for stars like the Sun (e.g. Dravins 1982; Asplund et al. 2000a, and Fig. 15). Naturally, 1D hydrostatic model atmospheres are unable to explain such line asymmetries and furthermore require the introduction of two additional free parameters: micro-turbulence to obtain realistic broadening of partly saturated lines and macro-turbulence to get reasonable line widths even if the detailed shape of observed lines can never be fully recovered. With a self-consistent convective velocity field, 3D models like those employed here do not need to invoke micro- or macro-turbulence to obtain an excellent agreement with observed line shapes. Realistic 3D simulations achieve this from first principles, without recourse to adjustable parameters (Asplund et al. 2000a).

Fabbian et al. (2010, 2012) have recently investigated the impact of magnetic fields on Fe spectral line formation in the quiet Sun, in particular in terms of the inferred solar photospheric Fe abundance. Using a series of 3D MHD simulations of varying magnetic field strengths computed with the Stagger-code, they investigated the 3D LTE line formation of 28 Fe I lines and found noticeable differences: lines with small or negligible Zeeman-broadening are still affected by the different mean temperature structures in 3D models with magnetic fields. For an average vertical field strength of 10 mT, they found that the derived Fe I-based abundance is $\approx$ 0.05 dex higher than without magnetic fields, the exact effect depending on the particular line in question. If this holds true, it would mean that the solar chemical composition presented by Asplund et al. (2009) would have to be revisited given that it was determined using the same non-magnetic 3D model as we are studying herein.

7.2. Observations

We make use of the solar FTS disk-centre intensity atlas of Brault & Neckel (1987) to measure the observed line shifts and line bisectors for a sample of Fe I and Fe II lines. The FTS atlas is on an absolute radial velocity scale as the relative motion between Sun and Earth is known precisely and has been corrected for the solar gravitational redshift of $633\,\mathrm{m\,s^{-1}}$ (for light intercepted on Earth, Lindegren et al. 1999). The line list stems from Asplund et al. (2009) and is augmented with additional lines from Asplund et al. (2000a) The necessary laboratory wavelengths to place the measured line shifts and bisectors on a velocity scale comes from Nave et al. (1994) for Fe I and from Johansson (1998, private communication, see also Nave 2012) for Fe II.

7.3. Results and discussion

We performed 3D LTE radiative transfer calculations of Fe I and Fe II lines using the 3D hydrodynamical solar model atmosphere also employed by Asplund et al. (2009) as well as a 3D MHD
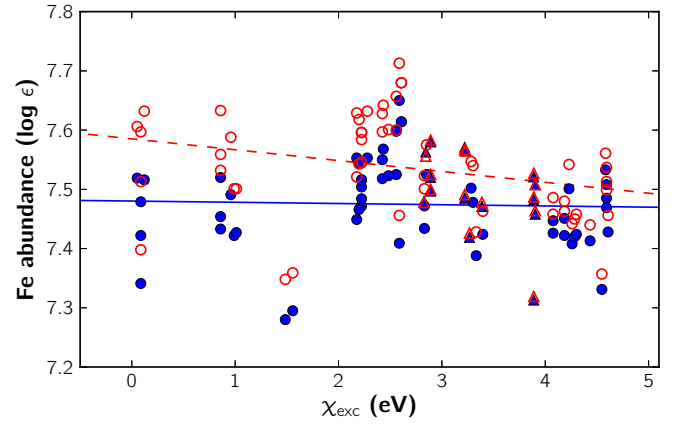

Fig. 13. Iron abundances derived from Fe I lines (circles) and Fe II lines (triangles), as a function of excitation potential (χ_{exc}) for the 3D model (blue, filled symbols) and the 3D MHD model (red, open symbols). The lines show linear fits to the abundance relation for Fe I lines, for the 3D model (blue solid) and 3D MHD model (red dashed).

simulation with an average vertical magnetic field of 10 mT from Thaler et al. (in preparation); we have also ensured that the corresponding 0 T simulation of Thaler et al. produces Fe lines indistinguishable from those of the Asplund et al. (2009) model. The theoretical disk-centre intensity profiles have been spatially and temporally averaged; the time sequence corresponds to 45 min of solar time with snapshots every minute. No micro- or macro-turbulent broadening entered the 3D line formation calculations but the resulting averaged line profiles were convolved with a Gaussian corresponding to the finite resolving power of the Brault & Neckel (1987) solar atlas. Polarisation and Zeeman-splitting have not been considered in the radiative transfer calculations.

The solar Fe abundance has been derived from each of our Fe lines using both the 3D hydrodynamical model and the 3D MHD model. For simplicity, we have adopted the equivalent widths given in Scott et al. (2013, in preparation), and, when not available there, in Asplund et al. (2000b). The magnetic fields can impact the line strengths in two ways: directly via Zeeman-splitting and indirectly via the different atmospheric stratifications, especially the temperature structure. We are not in a position to exactly quantify the former effect but the latter will dominate for all of our lines due to their small Landé factors and relatively short wavelengths Fabbian et al. (2012). Had we considered Zeeman splitting the derived Fe abundances would be < 0.02 dex *smaller* for the few lines with non-negligible Landé factors in our line sample (Fabbian et al. 2012).

Fig. 13 presents a comparison of the Fe abundances from the two 3D models. Because the MHD model is slightly warmer in the higher atmospheric layers (Fig. 1), one would expect the inferred Fe I abundance to be higher than without consideration of magnetic fields in the convection simulation. Also, the Fe II abundances should show very small differences given their greater formation depths. We find both of these to be true: the MHD mean Fe I abundance is 0.06 dex higher than without magnetic field, in line with the findings of Fabbian et al. (2012), while there is little difference for Fe II. However, as is clear from Fig. 13 there is distinct trend with excitation potential for the 3D LTE Fe I abundances with the MHD model, which is not present with the model of Asplund et al. (2009). It should be noted that this trend cannot be removed by departures from LTE because
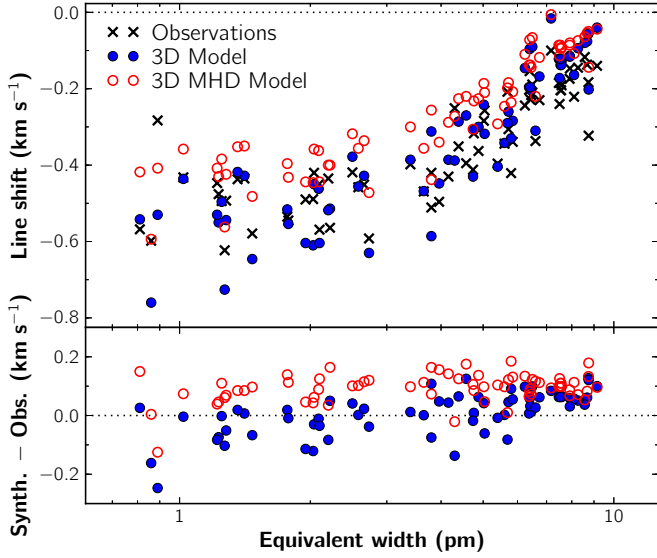
Fig. 14. *Upper panel:* Observed (crosses) line shifts for a sample of Fe I and Fe II lines for disk-centre intensity as a function of line strength together with predictions from the 3D model (blue, filled circles) and the 3D MHD (100 Gauss) model (red, open circles). *Lower panel:* Differences between predicted and observed line shifts.

those are positive and more so for low excitation lines, as demonstrated using the temporally and spatially averaged ⟨3D⟩ model (Bergemann et al. 2012; Lind et al. 2012).

For both the observed and predicted line profiles, the line centres were determined using a cubic spline around the wavelength points with minimum intensity. The solar gravitational redshift (of light intercepted at Earth) of $633\,\mathrm{m\,s^{-1}}$ was subtracted to obtain the observed central wavelengths on an absolute wavelength scale. Fig. 14 shows the observed and predicted Fe I and Fe II line shifts relative to the adopted laboratory wavelengths of the lines. In line with previous findings, weaker lines have a more pronounced convective blue-shift due to the larger depths of formation where convection and the anti-correlations between temperature and velocity are the largest; the cores of stronger lines become progressively less blue-shifted such that Fe lines with an equivalent width of $\sim 10\,\mathrm{pm}$ have nearly vanishing line shifts (Asplund et al. 2000a). The agreement between predicted and observed line shifts is very satisfactory for the 3D hydrodynamical model as demonstrated in the lower panel of Fig. 14: $30 \pm 60\,\mathrm{m\,s^{-1}}$ for our Fe I lines and $-50 \pm 70\,\mathrm{m\,s^{-1}}$ for Fe II. As also found by Asplund et al. (2000a) the stronger Fe lines tend to have slightly underestimated convective blue-shifts; Fe I lines with equivalent widths < 6 pm have a mean difference of only $9\,\mathrm{m\,s^{-1}}$. Given the slightly deviating behaviour of two of the weakest lines (Fe I 669.9 nm and Fe II 562.7 nm) one could suspect that they are more affected by blends or erroneous laboratory wavelengths than the average line. The predictions from the 3D MHD model have the correct qualitative behaviour but have systematically too little convective blue-shifts; the mean difference for Fe I is $100 \pm 50\,\mathrm{m\,s^{-1}}$.

A comparison between observed and predicted line bisectors tell a similar story as the line shifts. Solar disk-centre intensity line profiles show a characteristic C-shaped bisector (weaker lines tend to show only the upper part) with a typical velocity span of 300-600 $\mathrm{m\,s^{-1}}$ with the exact shape depending on the line formation height and temperature/velocity sensitivity (Asplund
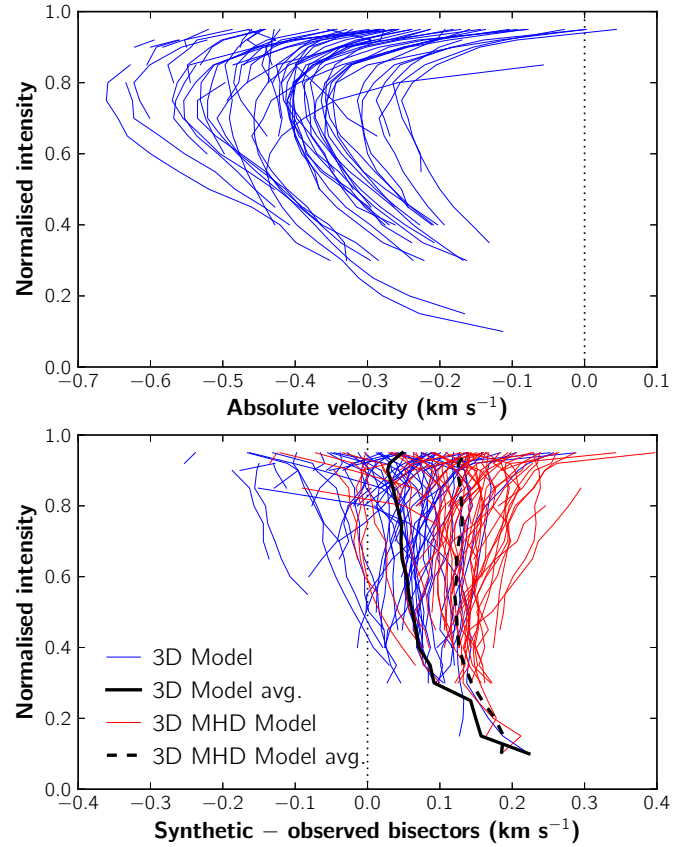

Fig. 15. *Upper panel:* Observed disk-centre intensity bisectors for a sample of Fe I and Fe II lines. *Lower panel:* Differences between predicted and observed bisectors. The thick lines represent the average difference over all bisectors.

et al. 2000a). Fig. 15 shows the *differences* between the predicted and observed bisectors for our sample of Fe lines; ideally these differences should manifest themselves as vertical lines at zero velocity offset. The agreement is very satisfactory for the 3D hydrodynamical model while the bisectors based on the 3D MHD are not sufficiently blue-shifted, in line with the line centre comparison.

8. Conclusions

Realistic solar atmospheres are of paramount importance for our understanding of not just the Sun but also of observations of other stars. The Sun provides an ideal test bench to test the physical ingredients of the models, which if successful can then be applied to other stars with some confidence. A critical requirement for a realistic model is that its thermodynamic quantities such as temperature, density and pressure match those of the real Sun. In this work we have undertaken a systematic study of the temperature structure of several solar models, using several key observational tests: continuum centre-to-limb variation, absolute continuum fluxes, wings of hydrogen lines, and also the intensity fluctuations over the granulation and detailed line shapes and asymmetries.

In all diagnostics we find that the 3D model reproduces the observations very well. This is especially true for the centre-to-limb variations, where its remarkable agreement surpasses even that of the semi-empirical Holweger & Müller model, which was built to fit the centre-to-limb variations. The 3D model also performs very favourably against the absolute continuum fluxes ob-

15

servations. For the hydrogen lines, the 3D model predicts the wings of the Hα line to be slightly stronger than the observations, but on the other hand provides a very good agreement for the other lines, and the best overall agreement of all the models tested. In terms of the continuum intensity fluctuations over the solar granulation, it is reassuring to find that the 3D model reproduces the observed intensity distribution and ΔI_{rms} well. The 3D model also predicts line shifts and asymmetries that agree very well with observations, which further supports its high degree of realism given the great sensitivity of the exact line shapes on the atmospheric conditions and line formation process.

In light of the work of Fabbian et al. (2010, 2012), we also calculated the predictions of a simulation with an average vertical magnetic field of 10 mT (the 3D MHD model). Regarding the Fe line asymmetries, shifts, and abundances, the 3D MHD model agrees slightly less well with observations, suggesting that either the effects of magnetic fields have been overestimated or that it is missing some ingredient that counteracts the consequences of the magnetic fields for the Fe line formation. Together with the evidence from the other diagnostics, it implies that at this stage there is no justification to prefer the solar abundances derived from the current generation of 3D MHD solar models over the 3D-based analysis of Asplund et al. (2009); our results suggest that the 3D MHD Fe abundance corrections advocated by Fabbian et al. (2010, 2012) are over-estimated.

The 1D theoretical models agree well with the observed absolute continuum fluxes, especially the MARCS model. However, both the MARCS and the PHOENIX models predictions for the centre-to-limb variations are consistently below the observations, both in the visible and in the infrared, which we attribute to a too steep temperature gradient. Such 1D hydrostatic models obviously cannot predict any line asymmetries or intensity contrasts. We find that the small difference in the temperature structure between the PHOENIX LTE and NLTE models does not translate into any significant difference in our comparison. Their results are very similar. If anything, the NLTE model performs slightly worse against the observational tests. This is likely to result from its somewhat steeper temperature gradient, due to NLTE cooling of the outer layers.

The agreement between the predictions from the 3D model and the observations demonstrates its very high degree of realism in its temperature stratification. Together with its realistic velocity fields and treatment of convection as exemplified by the line asymmetries, it places the 3D modelling in an excellent position to perform chemical abundance studies. It is noteworthy that it greatly outperforms any of the investigated 1D models, both theoretical flavours such as MARCS and PHOENIX and the semi-empirical Holweger & Müller model. There is thus no justification to continue to rely on the inferred solar abundances from 1D-based analyses when a significantly improved alternative is available.

References

Anders, E. & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
Anderson, L. S. 1989, ApJ, 339, 558
Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, ed. T. G. Barnes, III & F. N. Bash, 25–+
Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, A&A, 417, 751
Asplund, M., Grevesse, N., Sauval, J., & Scott, P. 2009, ARA&A, 47, 481
Asplund, M., Gustafsson, B., Kiselman, D., & Eriksson, K. 1997, A&A, 318, 521
Asplund, M., Nordlund, Å., Trampedach, R., Allende Prieto, C., & Stein, R. F. 2000a, A&A, 359, 729
Asplund, M., Nordlund, Å., Trampedach, R., & Stein, R. F. 2000b, A&A, 359, 743
Ayres, T. R., Plymate, C., & Keller, C. U. 2006, ApJS, 165, 618
Bahcall, J. N., Basu, S., Pinsonneault, M., & Serenelli, A. M. 2005, ApJ, 618, 1049
Barklem, P. S. 2007, A&A, 466, 327
Barklem, P. S. & Piskunov, N. 2003, in IAU Symposium, ed. N. Piskunov, W. W. Weiss, & D. F. Gray, 28P–+
Barklem, P. S., Piskunov, N., & O'Mara, B. J. 2000, A&A, 363, 1091
Barklem, P. S., Stempels, H. C., Allende Prieto, C., et al. 2002, A&A, 385, 951
Basu, S. & Antia, H. M. 2008, Phys. Rep., 457, 217
Behara, N. T., Ludwig, H.-G., Steffen, M., & Bonifacio, P. 2009, in American Institute of Physics Conference Series, Vol. 1094, American Institute of Physics Conference Series, ed. E. Stempels, 784–787
Bergemann, M., Lind, K., Collet, R., Magic, Z., & Asplund, M. 2012, MNRAS, 427, 27
Brault, J. W. & Neckel, H. 1987, ftp://ftp.hs.uni-hamburg.de/pub/outgoing/FTS-Atlas
Brott, I. & Hauschildt, P. H. 2005, in ESA Special Publication, Vol. 576, The Three-Dimensional Universe with Gaia, ed. C. Turon, K. S. O'Flaherty, & M. A. C. Perryman, 565–+
Caffau, E., Ludwig, H.-G., Bonifacio, P., et al. 2010, A&A, 514, A92
Caffau, E., Ludwig, H.-G., Steffen, M., et al. 2008a, A&A, 488, 1031
Caffau, E., Sbordone, L., Ludwig, H.-G., et al. 2008b, A&A, 483, 591
Carlsson, M. & Rutten, R. J. 1992, A&A, 259, L53
Carlsson, M., Stein, R. F., Nordlund, Å., & Scharmer, G. B. 2004, ApJ, 610, L137
Colina, L., Bohlin, R. C., & Castelli, F. 1996, AJ, 112, 307
Danilovic, S., Gandorfer, A., Lagg, A., et al. 2008, A&A, 484, L17
Dravins, D. 1982, ARA&A, 20, 61
Fabbian, D., Khomenko, E., Moreno-Insertis, F., & Nordlund, Å. 2010, ApJ, 724, 1536
Fabbian, D., Moreno-Insertis, F., Khomenko, E., & Nordlund, Å. 2012, A&A, 548, A35
Freytag, B., Steffen, M., & Dorch, B. 2002, Astronomische Nachrichten, 323, 213
Fuhrmann, K., Axer, M., & Gehren, T. 1993, A&A, 271, 451
Fuhrmann, K., Axer, M., & Gehren, T. 1994, A&A, 285, 585
Grevesse, N. & Sauval, A. J. 1998, Space Science Reviews, 85, 161
Gustafsson, B., Bell, R. A., Eriksson, K., & Nordlund, Å. 1975, A&A, 42, 407
Gustafsson, B., Edvardsson, B., Eriksson, K., et al. 2008, A&A, 486, 951
Hauschildt, P. H., Allard, F., & Baron, E. 1999, ApJ, 512, 377
Hauschildt, P. H. & Baron, E. 1999, Journal of Computational and Applied Mathematics, 109, 41
Holweger, H. 1967, Zeitschrift fur Astrophysik, 65, 365
Holweger, H. & Müller, E. A. 1974, Sol. Phys., 39, 19
Kiselman, D. 2008, Physica Scripta Volume T, 133, 014016
Koesterke, L., Allende Prieto, C., & Lambert, D. L. 2008, ApJ, 680, 764
Kopp, G., Lawrence, G., & Rottman, G. 2005, Sol. Phys., 230, 129
Kupka, F. G., Ryabchikova, T. A., Piskunov, N. E., Stempels, H. C., & Weiss, W. W. 2000, Baltic Astronomy, 9, 590
Kurucz, R. L. 2005, Memorie della Societa Astronomica Italiana Supplement, 8, 189
Kurucz, R. L., Furenlid, I., Brault, J., & Testerman, L. 1984, Solar flux atlas from 296 to 1300 nm (National Solar Observatory Atlas, Sunspot, New Mexico: National Solar Observatory, 1984)
Leenaarts, J. & Carlsson, M. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 415, Astronomical Society of the Pacific Conference Series, ed. B. Lites, M. Cheung, T. Magara, J. Mariska, & K. Reeves, 87–+
Lind, K., Bergemann, M., & Asplund, M. 2012, MNRAS, 427, 50
Lindegren, L., Dravins, D., & Madsen, S. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 185, IAU Colloq. 170: Precise Stellar Radial Velocities, ed. J. B. Hearnshaw & C. D. Scarfe, 73
Ludwig, H.-G., Behara, N. T., Steffen, M., & Bonifacio, P. 2009, A&A, 502, L1
Mihalas, D., Dappen, W., & Hummer, D. G. 1988, ApJ, 331, 815
Nave, G. 2012, MNRAS, 420, 1570
Nave, G., Johansson, S., Learner, R. C. M., Thorne, A. P., & Brault, J. W. 1994, ApJS, 94, 221
Neckel, H. & Labs, D. 1994, Sol. Phys., 153, 91

Nordlund, Å. 1982, A&A, 107, 1

Nordlund, Å. & Galsgaard, K. 1995, A 3D MHD Code for Parallel Computers, Tech. rep., Astronomical Observatory, Copenhagen University

Nordlund, Å., Stein, R. F., & Asplund, M. 2009, Living Reviews in Solar Physics, 6, 2

Pereira, T. M. D., Asplund, M., & Trampedach, R. 2008, in Precision Spectroscopy in Astrophysics, ed. N. C. Santos, L. Pasquini, A. C. M. Correia, & M. Romaniello, 313–314

Pierce, A. K. & Slaughter, C. D. 1977, Sol. Phys., 51, 25

Pierce, A. K., Slaughter, C. D., & Weinberger, D. 1977, Sol. Phys., 52, 179

Piskunov, N. E., Kupka, F., Ryabchikova, T. A., Weiss, W. W., & Jeffery, C. S. 1995, A&AS, 112, 525

Przybilla, N. & Butler, K. 2004, ApJ, 610, L61

Serenelli, A. M., Basu, S., Ferguson, J. W., & Asplund, M. 2009, ApJ, 705, L123

Short, C. I. & Hauschildt, P. H. 2005, ApJ, 618, 926

Short, C. I. & Hauschildt, P. H. 2009, ApJ, 691, 1634

Skartlien, R. 2000, ApJ, 536, 465

Soon, W. H. 1992, ApJ, 394, 717

Spickler, P. T., Benner, D. C., & Russel, III, J. M. 1996, Sol. Phys., 165, 23

Stehlé, C. & Hutcheon, R. 1999, A&AS, 140, 93

Stein, R. F. & Nordlund, Å. 1998, ApJ, 499, 914

Thuillier, G., Floyd, L., Woods, T. N., et al. 2004, Advances in Space Research, 34, 256

Trujillo Bueno, J., Asensio Ramos, A., & Shchukina, N. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 358, Astronomical Society of the Pacific Conference Series, ed. R. Casini & B. W. Lites, 269

Trujillo Bueno, J. & Shchukina, N. 2009, ApJ, 694, 1364

Tsuneta, S., Ichimoto, K., Katsukawa, Y., et al. 2008, Sol. Phys., 249, 167

Uitenbroek, H., Tritschler, A., & Rimmele, T. 2007, ApJ, 668, 586

Vögler, A., Bruls, J. H. M. J., & Schüssler, M. 2004, A&A, 421, 741

Wedemeyer-Böhm, S. & Rouppe van der Voort, L. 2009, A&A, 503, 225

Willson, R. C. & Hudson, H. S. 1988, Nature, 332, 810