

Stellar Evolution Challenge

A. Weiss¹, S. Cassisi², A. Dotter³, Z. Han⁴ and Y. Lebreton⁵

¹Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany
email: aweiss@mpa-garching.mpg.de

²INAF – Osservatorio Astronomico di Collurania, Via M. Maggini, 64100 Teramo, Italy

³Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory,
Hanover, NH 03755, USA

⁴Yunnan Observatory, Chinese Academy of Sciences, Kunming, 650011, China

⁵Observatoire de Paris, GEPI, CNRS UMR 8111, Place J. Janssen, 92 195 Meudon, France

Abstract. We report about the results of the Stellar Evolution Challenge, initiated at the Leiden Lorentz Center meeting “Fine-tuning stellar population models” (June, 2006).

Keywords. stars: general, evolution

1. Introduction

At the Lorentz Center meeting “Fine-tuning stellar population models” in Leiden (June, 2006), a “Stellar Evolution Challenge” for the present meeting was officially announced. Similar to the “Stellar Population Challenge” (Trager, this meeting), the challenge consisted in defining a number of stellar structures (mass, composition, age) for which participants should provide the models obtained by their code. To make the comparisons more meaningful, details of the physical assumptions and ingredients had to be specified and followed as closely as possible.

The idea of this challenge was almost identical to a project of the first author (“Stellar Code Calibration”), which he had started almost a year earlier. The intention of it was to demonstrate that for given physical input stellar evolution codes can be tuned to produce identical models to high accuracy. After such a calibration, model differences between authors could be traced convincingly to differences in the physics employed, such that we could learn indeed about stellar physics when comparing to observed objects. It would also strengthen the confidence of users in stellar models and libraries. The progress of this project, though more than 20 colleagues had been contacted and around 15 had responded positively, has been negligible until the Leiden meeting.

Similar efforts have been under way in the seismology community. The well-known success of the solar model is currently being repeated by the Corot/ESTA (Evolution and Seismic Tools Activity) working group, where a number of tasks with regard to potential targets of the Corot-satellite have to be fulfilled. Task 1 consists of stellar models without diffusion and is similar to the intention of the Stellar Evolution Challenge. (Task 2 consists of the computation of oscillation frequencies, and task 3 of models with diffusion; see the excellent website maintained by Mario Monteiro, www.astro.up.pt/corot for more details.)

In August a “Solar Code Calibration” website became available†, where details about the challenge, input physics, analysis tools and more can be found, and where solutions to the challenge cases can be uploaded. Until a week before this meeting, 5 (five) active

† www.mpa-garching.mpg.de/stars/SCC/StarCode.html

participants had provided models for some or all challenge cases, specified in Sect. 2. The participants are (in alphabetical order; name — code — reference)

- Santi Cassisi — FRANECode used for BASTI library — Pietrinferni et al. (2004)
- Aaron Dotter — Dartmouth Stellar Evolution Program (DSEP) — Bjork & Chaboyer (2006)
- Zhanwen Han — Yunnan Observatory version of Eggleton code — Eggleton (1971)
- Yveline Lebreton — Cesam2k public French code — Morel (1997)
- Achim Weiss — Garching stellar evolution code (GARSTEC) — Weiss & Schlattl (2000)

2. Specifications

2.1. Input physics

For the physics employed in the challenge models it was agreed to use the most simple assumptions and widespread data. In particular they were

- solar metal distribution as in Grevesse & Noels (1993)
- neither mass loss nor diffusion, overshoot, or semiconvection
- Eddington grey atmospheres
- NACRE (Angulo *et al.* 1999) reaction rates
- opacities from Iglesias & Rogers (1996) and Alexander & Ferguson (1994)
- equation of state: either Irwin's FreeEOS (<http://freeeos.sourceforge.net/>), or OPAL tables (Rogers et al. 1996); FreeEOS tables in OPAL-format were provided in 11/06 on the SCC-website; Irwin's FreeEOS is supposed to reproduce the OPAL-tables, but have a wider range of validity.
- Mixing Length Theory description of convection, using a parameter of $\alpha_{\text{MLT}} = 1.6$, and following the formulation in Weiss *et al.* (2004)

2.2. Challenge models

2.2.1. Corot/ESTA cases

Since the Corot/ESTA group is already providing models for their Task 1 cases, it was decided to adopt three of them. These were

(a) $M = 0.9 M_{\odot}$; $Y_0 = 0.28$, $Z_0 = 0.02$; evolved on the main-sequence to $X_c = 0.35$ (Corot case 1.1)

(b) $M = 1.2 M_{\odot}$; $Y_0 = 0.26$, $Z_0 = 0.01$; evolved to the subgiant phase, $M_{\text{cHe}} = 0.1 M_{\odot}$; the He-core is defined as the central region with $X < 0.01$ (Corot case 1.3)

(c) $M = 3.0 M_{\odot}$; $Y_0 = 0.28$, $Z_0 = 0.01$ evolved into the early MS, when $X_c = 0.69$ (from $X_0 = 0.71$; Corot case 1.6)

For these models the complete structure is requested; the GONG-format (specifications available from the SCC- and Corot-websites) is used for data exchange.

2.2.2. Zero-age main-sequence models

Models along the ZAMS for two compositions are to be provided. The evolution should start with homogeneous pre-main sequence models, and the ZAMS was defined as being reached once for the first time the nuclear energy provides more than 99% of the radiated luminosity. Global (L , T_{eff} , R) and central (T_c , ρ_c) properties are requested. The cases are

- ZAMS 1: $Y = 0.280$, $Z = 0.020$
- ZAMS 2: $Y = 0.250$, $Z = 0.001$
- mass values: 0.8, 1.0, 1.3, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 7.0, 10., 15., 20., 30., 40., 50., 60 M_{\odot} ; or any subset of these.

2.2.3. Evolution

Three cases were selected to see how the evolution of two low-mass and one massive star agree. The first two cases are to be followed from the ZAMS (definition as above) into He-ignition (up to the point where $L_{\text{He}} = 10^6 L_{\odot}$), the last one to the end of He-burning. Again, surface and central properties (including abundances of carbon and oxygen) are compared. The cases are

- $M = 1.0 M_{\odot}$; $Y_0 = 0.28$, $Z_0 = 0.02$
- $M = 20.0 M_{\odot}$; $Y_0 = 0.28$, $Z_0 = 0.02$
- $M = 0.8 M_{\odot}$; $Y_0 = 0.25$, $Z_0 = 0.0001$

3. Results

In the following we will present some of the comparisons. Since the plots are done with an automated machinery, not all of them are optimally designed. If models from a participant are missing, this is because they were not provided. Due to the very late delivery of some models, there was no time for a second iteration to improve the agreement between different sources. In that sense this is what a user of stellar models can expect in terms of agreement of different codes once the model details have been specified. Only in cases of clear technical mistakes (such as using the wrong mass or mixing up the meaning of data), models have been replaced.

3.1. ZAMS models

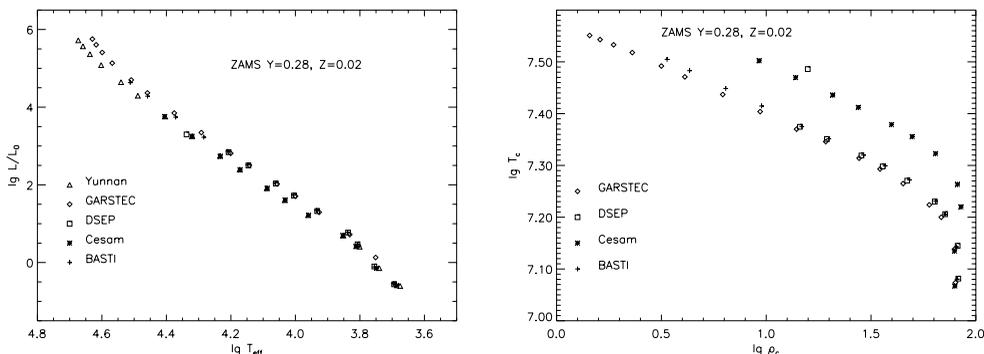


Figure 1. Left: HRD for ZAMS models with $Y = 0.28$, $Z = 0.02$ (for mass values see Sect. 2.2.2) Right: Central values of T_c vs. ρ_c for the same models as in the left figure

The metal-rich ZAMS models are shown in Fig. 1 (left: HRD; right: central conditions). The agreement in the HRD (left panel) is acceptable; typical differences are 0.1 dex (25%) in luminosity and 0.03 dex (7%) in T_{eff} . In terms of central conditions (right panel) the Cesam2k models clearly separate from the others (hotter and denser), although the overall agreement in central temperature is of the order of 0.01 dex (2%). Y. Lebreton (private communication) looked deeper into this problem and found that the ZAMS-criterion is causing the selection of a more evolved model in the Cesam-code. At the HRD location, where, for example, the GARSTEC code defines a ZAMS model, the Cesam-model has only about 95% of the total luminosity produced by nuclear processes, although the central hydrogen content is very similar. If we compare this model to the GARSTEC and BASTI ZAMS models ($M = 1.3 M_{\odot}$), the Cesam-model agrees to the level of 0.001 dex with the GARSTEC one and falls in between this and the BASTI

model. The $5 M_{\odot}$ DSEP-model appears to suffer from a similar problem. Clearly, we need a more accurate and stable definition of the ZAMS. We also conclude that the differences seen in Fig. 1 will then be reduced even further.

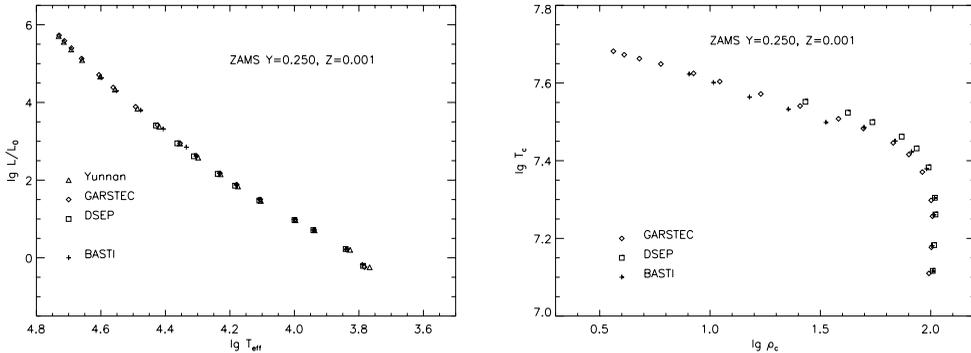


Figure 2. As Fig. 1, but for the ZAMS models with $Y = 0.25$, $Z = 0.001$

In case of the metal-poor ZAMS ($Z = 0.001$) the agreement in all quantities is of the order of 0.01 dex (2%). However, no Cesam models were provided, and in both cases no central density values for the Yunnan models were available.

3.2. Corot cases

For the Corot cases comparison with the Corot/ESTA models is possible. After some iterations, all models for case 1.1 ($M = 0.9 M_{\odot}$ MS star) provided there agree (with the exception of one code) in terms of luminosities within 0.5%, and within 1.5% in effective temperature. The model ages vary by at most 5% (see Lebreton & Monteiro, ESTA meeting, Nice, Sept. 2005).

Table 1 shows the comparison in our case. During the meeting first models from the Geneva group (G. Meynet, private communication) were received and partially included in this table. In addition, the influence of using the nuclear reaction rates by Adelberger *et al.* (1998) instead of the standard NACRE rates is shown for the GARSTEC code (case GARSTEC/A). Overall, if we ignore the Yunnan models, ages agree within 1%, luminosities within 1.6%, T_{eff} within 2.0% and central values within 1.1%. The influence of the Adelberger rates is of similar order. The main deviation of the Yunnan model is in age (+10%) and ρ_c (3%).

Table 1. Comparison of challenge case “Corot 1.1”, an $M = 0.9 M_{\odot}$ MS star

Model	age (Myr)	L/L_{\odot}	T_{eff}	T_c (10^7 K)	ρ_c
GARSTEC	6886	0.6273	5460	1.447	150.56
GARSTEC/A	6899	0.6267	5452	1.444	149.89
BASTI	6849	0.6338	5466	1.448	151.98
Cesam	6827	0.6276	5440	1.446	150.30
Yunnan	7559	0.6377	5354	1.463	156.16
Geneva	6843	0.6273	5423		

Since the absolute values of the interior structure variables agree very well, we show in Fig. 3 *relative* differences with respect to the Cesam2k model (which serves as our reference case to the Corot model comparison). The left panel shows hydrogen profile differences as function of absolute radius, the right one luminosity differences vs. absolute

pressure. Most models agree to better than a few per mille with respect to their hydrogen profile, i.e. their nuclear evolution, and within a few per cent in the comparison of their thermal as function of mechanical structure (right panel), and as such our models reach the same level of agreement as the Corot models do. The deviant Yunnan model needs further clarification, but given the large age difference we suspect that the reason lies in an insufficient temporal resolution (see below for more on this point).

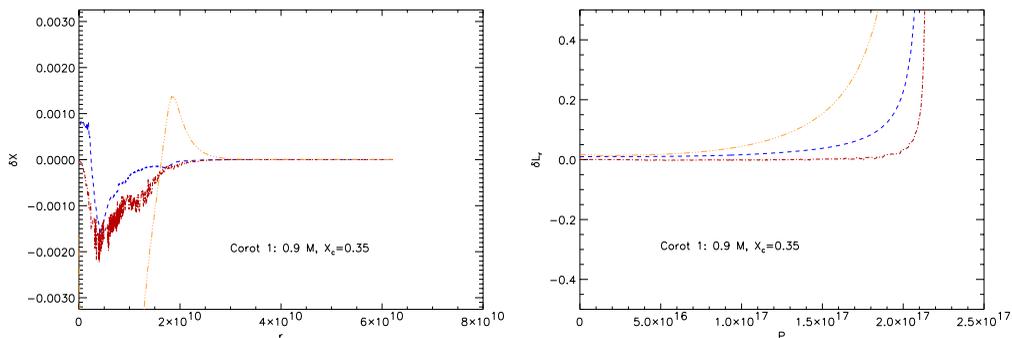


Figure 3. Comparison of interior structure of evolved models with $M = 0.9 M_{\odot}$, and $X_c = 0.35$ (Corot 1.1). All relative differences are with respect to the Cesam2k model. The lines correspond to BASTI (blue, dashed), GARSTEC (red, dash-dot) and Yunnan (orange, dot-dot-dot-dash) solutions. Left: relative hydrogen abundance differences as function of absolute radius; right: relative luminosity differences as function of pressure.

In case of the $1.2 M_{\odot}$ model, which was evolved into the subgiant phase (helium core mass of $0.1 M_{\odot}$), the agreement is not as good. Ages agree within 3% without the GARSTEC model (7% including it), L and T_{eff} within 2% resp. 0.6%, the latter value being surprisingly small, if one considers that the models have already quite developed convective envelopes. The interior comparison (Fig. 4) shows a slightly discrepant H-profile in the Yunnan model, which could be traced back to a different definition of “He core” (Han, private communication), and is therefore an artifact of not following exactly the challenge specifications. The ^3He -profiles agree surprisingly well, given the fact that the abundance of this isotope is very sensitive to temperature history and extremely low within the He core.

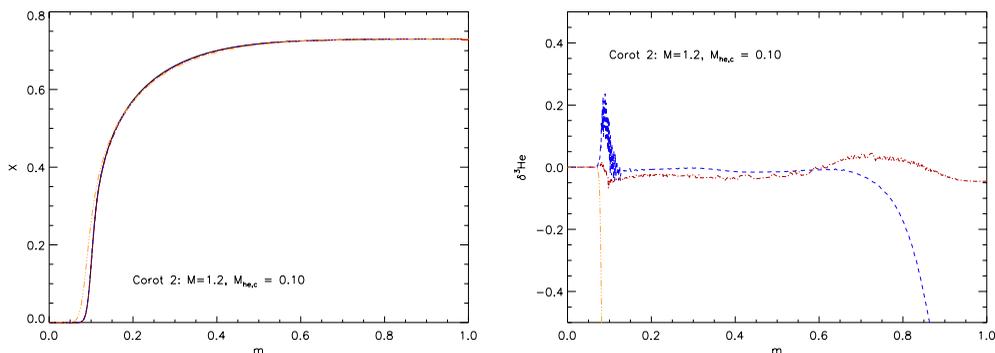


Figure 4. Comparison of interior structure of evolved models with $M = 1.2 M_{\odot}$, and $M_{\text{cHe}} = 0.1 M_{\odot}$ (Corot 1.3). Left: absolute hydrogen abundance, and right: relative ^3He differences as function of relative mass. Line types and colors have the same meaning as in Fig. 3

Finally, the early MS star with $M = 3 M_{\odot}$ is compared. Here, although HRD position and central temperature and pressure values agree at the per cent level or better, age differences of 15% must be noticed. Ages vary from 13.94 Myr (BASTI) to 16.24 Myr (GARSTEC). I think this is connected with insufficient temporal accuracy, but also with the presence of a convective core (see below). Figure 5 shows again two comparisons: the hydrogen profile (left) which displays different chemical profiles at the edge of the convective core, stressing this conjecture, and relative luminosity differences (right), which are limited to below 5%, even for the center, where L_r/L becomes rather small, and relative differences are thus emphasized.

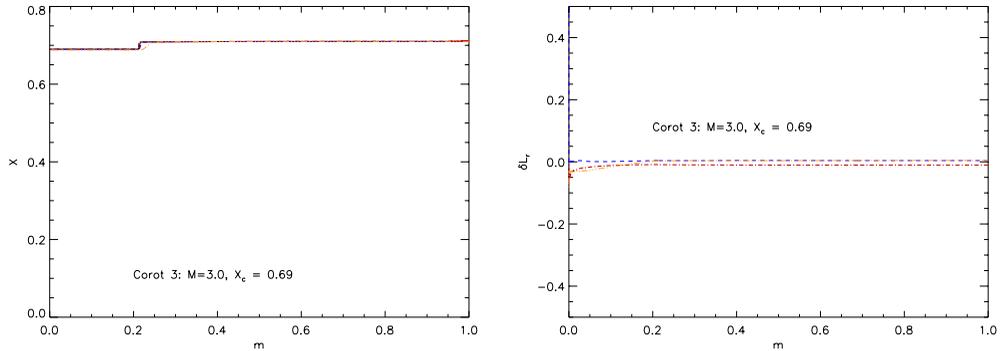


Figure 5. Comparison of interior structure of the early main-sequence models with $M = 3.0 M_{\odot}$, and $X_c = 0.69$ (Corot 1.6). Left: absolute hydrogen abundance, and right: relative luminosity differences as function of relative mass. Linetypes and colors have the same meaning as in Fig. 3

3.3. Evolution cases

We finally turn to the evolution cases, beginning with the $1.0 M_{\odot}$ star of solar-like composition ($Z = 0.02$). Fig. 6 shows the HRD (left) and the $L(t)$ -plot (right). Effective temperatures agree very well within only 15 K, reflecting the use of identical input physics and a consistent MLT treatment. The exception is the Yunnan model, which is definitely cooler than any other model. We suspect that this is due to the MLT-implementation in the Eggleton-code. Concerning stellar lifetimes, they vary by as much as 4% ($5 \cdot 10^8$ yrs) between the BASTI and the GARSTEC code. This can be traced back to insufficient temporal resolution in the latter calculation. To reduce the lifetime by another 4-5%, timesteps on the MS had to be increased by another factor of 10 (on average). This will have to be checked also for the other codes participating.

The case of the second low-mass star ($M = 0.8 M_{\odot}$, $Z = 0.0001$) is not shown here, as it resembles very much the former one. Lifetime differences are close to 10% ($1.2 \cdot 10^9$ yrs), with the same systematics as in Fig. 6. At the TO, temperatures differ more, up to 140 K (2%). For these stars, internal properties and RGB features (bump, tip) have not yet been compared.

Finally, we show the $20 M_{\odot}$ star ($Z = 0.02$) in Fig. 7. Only three groups provided models, but large differences, in particular in the post-MS phase, are obvious. Lifetimes vary again by up to 8% ($5 \cdot 10^5$ yrs) and luminosities by 12% or 0.05 dex. The central hydrogen abundance evolution reflects the different core evolutions; in particular the BASTI model experiences an episode of additional fuel mixed into the core. It is plausible that all differences can be traced back to the treatment of the convective boundary and

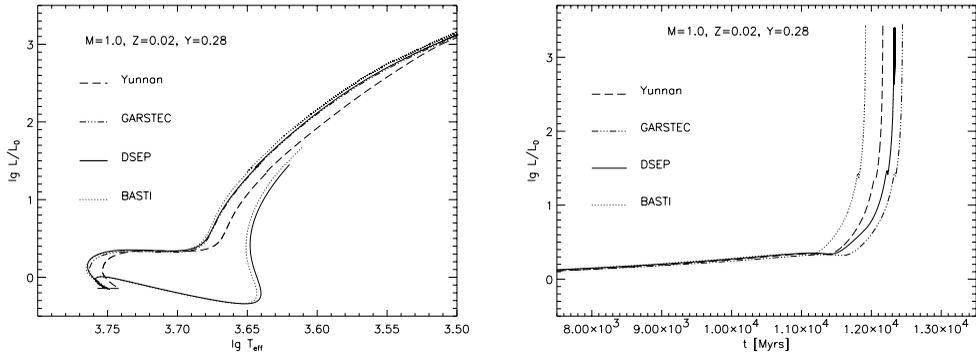


Figure 6. Evolution of a $1.0 M_{\odot}$ model with $Y = 0.28$, $Z = 0.02$ from the ZAMS to the tip of the RGB. In some cases the pre-MS evolution is shown as well. Left: HRD; right: $L(t)/L_{\odot}$

the evolution of the convective core. This will require further and more detailed analysis, and will be more difficult once convective overshooting is taken into account, too.

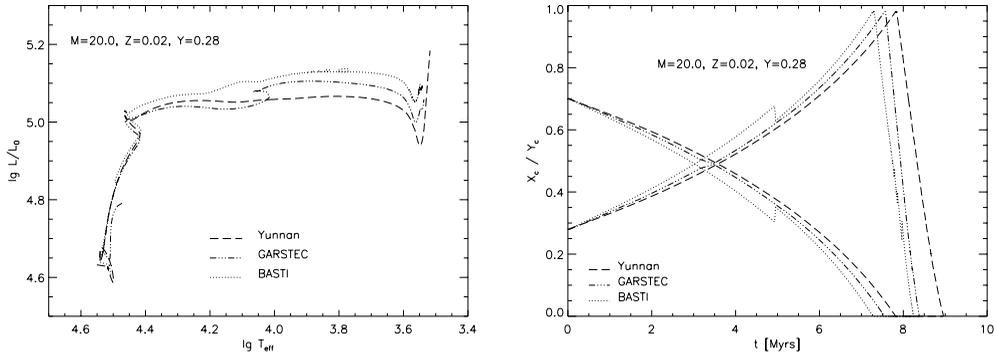


Figure 7. Evolution of a $20.0 M_{\odot}$ model with $Y = 0.28$, $Z = 0.02$ from the ZAMS to the end of core helium burning. Left: HRD; right: central hydrogen and helium abundance as function of time.

4. Discussion and Conclusions

What have we learned from these first exercises? The Corot cases have been reproduced surprisingly well even by the first attempts in most cases. On the other hand, the ZAMS models differ too much, which in part or completely may be due to the insufficiently unique definition of what a ZAMS model is. The temporal resolution of the calculations is not yet sufficient, although most codes reach an accuracy of a few per cent, which in general can be considered to be sufficient. For this challenge, however, we would like to achieve an agreement at the 1% level, and this seems to require much better time resolution on the main sequence, with up to 2000 models or more during central hydrogen burning.

Concerning the differences between codes there is no systematic effect except for the Yunnan version of the Eggleton code. However, this comes as no surprise: the Eggleton code (Eggleton 1971) is a highly efficient code with a different numerical scheme, which relies in large parts on simple analytical approximations to the underlying physics. It

therefore does not implement the physics specification of this project with respect to the EOS and nuclear reaction rates, and neither uses our MLT formulation. It allows to follow the evolution of a low mass star from the ZAMS to the tip of the RGB with only a few hundred models, and resolves the models on a movable mesh of 200 grid points (for comparison, the GARSTEC code uses of order 1000 grid points at least). While this makes it highly suitable for quick overview calculations, it may be the reason for the differences found here.

Effective temperatures agree rather well in most cases, which is the consequence of consistent opacities and convection descriptions used in most codes. However, at the TO of stars with shallow convective envelopes, the differences are still too large. This may also be related to the treatment of convective borders, which is presently the most urgent problem to be solved.

During this challenge it became obvious that everyone of the participants learned more about his code and models. Some detected problems in very specific evolutionary phases, some realized insufficient resolution, some had to redefine quantities like the mass of the helium core. It also became obvious that the production of reliable sets of stellar tracks, needed for population synthesis, requires a calibration of the code first. All participants are willing to continue with this project, and others (G. Meynet, Geneva, and D. VandenBerg, Victoria) will join. The message to the population synthesis community is, we think, the following: For given physics specifications and after an easy-to-do calibration test, modern codes are able to agree within the per cent level (both in global and in structural properties). However, the calibration is needed and should be done. We will continue and extend the calibration project, since the “Challenge” is still there.

Acknowledgements

A.W. acknowledges additional discussions with D. VandenBerg. Most of the work presented here was done during a Visiting Professorship at Dartmouth College; A.W. thanks the college, the Harris German-Dartmouth program, and B. Chaboyer and his students in particular for their hospitality. S.C. and Z.H. thank M. Salaris and A. Pietrinferni, respectively X. Chen for their help in preparing the challenge.

References

- Adelberger, E., Austin, S., Bahcall, J., Balantekin, A., Bogaert, G., & Buchmann, L., 1998, *Rev. Mod. Phys.* 70, 1265
- Alexander, D. R. & Ferguson, J. W., 1994, *ApJ* 437, 879
- Angulo, C., Arnould, M., Rayet, M., Descouvemont, P., Baye, D., Leclercq-Willain, C., *et al.*, 1999, *Nucl.Phys. A* 656, 3
- Bjork, S. R. & Chaboyer, B., 2006, *ApJ* 641, 1102
- Eggleton, P. P., 1971, *MNRAS* 151, 351
- Grevesse, N. & Noels, A., 1993, *Phys. Scripta* T47, 133
- Iglesias, C. A. & Rogers, F. J., 1996, *ApJ* 464, 943
- Morel, P., 1997, *A&A Supplement* 124, 597
- Pietrinferni, A., Cassisi, S., Salaris, M., & Castelli, F., 2004, *ApJ* 612, 168
- Rogers, F. J., Swenson, F. J., & Iglesias, C. A., 1996, *ApJ* 456, 902
- Weiss, A., Hillebrandt, W., Thomas, H.-C., & Ritter, H., 2004, *Cox and Giuli's Principles of Stellar Structure*, by A. Weiss, W. Hillebrandt, H.-C. Thomas, H. Ritter. Cambridge, UK: Cambridge Scientific Publishers, 2004.
- Weiss, A. & Schlattl, H., 2000, *A&A Supplement* 144, 487

Discussion

GUSTAFSSON: Did you expect the great differences in MS lifetime?. Did all groups use essentially the same methods in computing the time steps, and in applying higher order corrections when stepping?.

WEISS: No, I did not expect such large differences, as the physics involved is rather well known and simple. Concerning time steps, everybody claims that these are fine enough to not influence lifetimes. Higher order schemes are not used, to the best of my knowledge, although we would have that option in the GARSTEC code.

PELETIER: Are you not worried that none of the models that you discussed are 3D?.

WEISS: Not at all as these are low mass stars. 3D effects are only important for high mass stars if the effects of convection, rotation and mass loss are included; or for convective envelopes and atmospheres with respect to stellar spectra. See the talks by Meynet and Gustafsson.



The speaker (left), Brian Marsteller, Ilani Loubser, Robert Blum and Hyun-Chul Lee during a dinner.