

## CHARACTERIZATION OF DIFFERENT REGIMES IN NONLINEAR LIQUID CRYSTAL MODELS

I. ORQUÍN-SERRANO<sup>\*,¶</sup>, F. R. VILLATORO<sup>†</sup>, A. FERRANDO<sup>‡</sup>,  
P. FERNÁNDEZ DE CÓRDOBA<sup>\*</sup> and H. MICHINEL<sup>§</sup>

<sup>\*</sup>*Instituto Universitario de Matemática Pura y Aplicada (IUMPA),  
Universitat Politècnica de València, 46022 València, Spain*

<sup>†</sup>*Dept. Lenguajes y Ciencias de la Computación,  
Universidad de Málaga, 29071 Málaga, Spain*

<sup>‡</sup>*Dept. d'Òptica, Universitat de València,  
46100 Burjassot, València, Spain*

<sup>§</sup>*Área de Óptica, Campus de Ourense,  
Universidade de Vigo, 32004 Ourense, Spain*

<sup>¶</sup>*isorser@doctor.upv.es*

Received 2 September 2012

The range of validity of two models for nonlocal nonlinear optics in Nematic Liquid Crystals (NLC) is studied. Particularly the influence of the optical power and the initial position of the beam over its trajectory is studied when launching the beam with an offset in a planar cell. The main difference between both models is the dependence of the orientational angle with the optical field, either linear or nonlinear. The results demonstrate the critical role of the nonlinearity in the propagation of nematicons in NLC planar cells.

*Keywords:* Liquid Crystals; Nematics; Nonlinear optics.

PACS Number(s): 42.65.-k, 42.79.Kr, 42.65.Tg

### 1. Introduction

Nematic Liquid Crystal (NLC) devices are being widely studied in the field of nonlinear optics due to its large nonlinear response.<sup>1,2</sup> It allows to generate nonlinear solutions with no change of shape, the so called nematicons at very low optical powers.<sup>3</sup> Numerous applications are being studied manipulating these nematicons: micro-devices for steering solitons in NLCs using curved surfaces<sup>4</sup> or external light sources,<sup>5</sup> all-optical switching and logic gating with spatial solitons in liquid crystals<sup>6,7</sup> and signal processing by opto-optical interactions between self-localized and free propagating beams in liquid crystals,<sup>8</sup> for example. Besides, the nonlocality exhibited by NLC cells has been shown as an efficient mechanism for

stabilizing optical complex solutions which cannot exist in local nonlinear homogeneous media, as it has been demonstrated numerically in the cases of vortex and multivortex solitons,<sup>9–11</sup> dipole solitons<sup>12</sup> or soliton clusters<sup>13</sup> as well as in experiments like those described in Refs. 14 and 15.

The transverse oscillation of a nematicon induced by the force exerted by the boundary in highly nonlocal nonlinear media when there exists an asymmetry on the launching position of the beam is a well-known phenomenon as it has been pointed in Ref. 16 and experimentally in Ref. 17, as well as for thermal media in Ref. 18. The effect of the boundary force can be seen when the initial position of the light beam is not centered in the cell since it yields an offset in the peak position of the tilt angle which does not coincide with the position of the optical peak. We pretend here to show how the orientational nonlinearity affects this transverse oscillation, since linearized models are usually employed to explain certain properties of the nematicon trajectory such as its period or its amplitude.

Oriental nonlinearity arises, in the more complete nonlinear model we study in this work, by the effective coupling among the three differential equations involved and not from a nonlinear term appearing in the beam propagation method as other commonly used models do.<sup>19</sup> Here we are not restricted to low tilt angle values as it occurs with linearized models.<sup>17</sup>

The contents of the paper are as follows: Sec. 2 shows the two models which are being compared in this work explaining their derivation while in Sec. 3 a numerical comparison studying the transverse oscillation of nematicons is presented. Finally our main conclusions are summarized in Sec. 4.

## 2. Physical Models

Let us consider a NLC device consisting of two glass plates with commercial NLC E7 in between and two indium tin oxide (ITO) transparent electrodes next to the glass plates which permits us to apply external voltages to the liquid crystal. Figure 1 illustrates the geometrical configuration of the NLC device, where the  $z$  co-ordinate is the propagation direction, while  $x$  and  $y$  co-ordinates represent transverse spatial dimensions.<sup>17</sup> Planar alignment imposes director orientation nearly parallel to the glass plates, i.e., tilt angles smaller than two degrees (2). Let us take the optical field

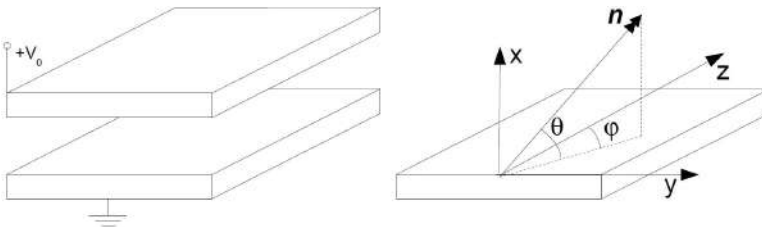


Fig. 1. NLC device. Cell with planar alignment (left). On the right we see the parameters that define the director orientation: the tilt and twist angles.<sup>17</sup>

to be linearly polarized as an extraordinary wave in direction  $x$ . The optical axis (or molecular director) of the NLC is parallel to  $z$  and no twist angle is allowed in the model, since the symmetry in the transverse co-ordinate  $y$  of the electrodes does not allow the appearance of  $y$  components in the electric field, which are responsible for twist deformations. We will focus in lateral propagation of light inside the device, this is, the wavevector nearly parallel to propagation direction,  $z$  co-ordinate. No big deviation angles are allowed since we work with paraxial approximation. Model B (Mod. B) was introduced in the first paper on nematicons.<sup>20</sup> Then, the setup has been being widely studied in the highly nonlocal nonlinear regime.<sup>21–23</sup>

The two models of NLC devices to be studied in this paper are the following. Model A (Mod. A) is the linear model employed in other studies<sup>24,25</sup> which assumes two main hypothesis. First, the orientational equation is linearized, so the model is suitable only for low tilt angle deviations, around a tilt angle of  $\pi/4$ , which maximizes the nonlinear effects. And second, the external electric field distribution will not be introduced in the model so it will not be possible to study the influence of external fields. The resulting governing equations are given by:

$$K \frac{\partial^2 \theta}{\partial x^2} + 2\epsilon |E_x^o|^2 = 0, \quad (1)$$

$$j \frac{\partial E_x^o}{\partial z} = \frac{1}{2} \frac{\partial^2 E_x^o}{\partial x^2} + 2\theta E_x^o, \quad (2)$$

where  $\theta$  is the orientational angle,  $E_x^o$  is the amplitude of the optical field linearly polarized through  $x$  direction,  $K$  is the distortion energy constant,  $j = \sqrt{-1}$  and  $\epsilon$  is the dielectric constant of the medium.

Equation (1) is solved analytically in Ref. 16 for studying the transverse oscillation of a nematicon under the effect of boundary forces when the optical beam is launched with an offset in the transverse co-ordinate. Ehrenfest theorem is employed there in order to study the boundary force over the nematicon as well as its trajectory and period dependence over the optical power and offset arriving to conclusions on the approximate behavior of such phenomenon.

Model B was introduced in the first paper on nematicons<sup>20</sup> and is a nonlinear model mainly that has been also studied, among others, by Beeckman *et al.*<sup>3,22,26</sup> The two hypothesis used in Mod. A are relaxed, so the full nonlinear character of the orientational equation is used and the external electric fields are included. Here, just scalar and paraxial approximations are done on the beam propagation method presented above. The system of equations of this nonlinear models is

$$K \frac{\partial^2 \theta}{\partial x^2} + \frac{1}{2} \epsilon_0 \sin(2\theta) (\Delta \epsilon^s |E_x^s|^2 + \Delta \epsilon^o |E_x^o|^2) = 0, \quad (3)$$

$$\int_0^d E_x^s dx = \int_0^d \frac{D_x^s}{\epsilon_{xx}} dx = -V_0, \quad (4)$$

$$2jk_0 n_0 \frac{\partial E_x^o}{\partial z} = \frac{\partial^2 E_x^o}{\partial x^2} + k_0^2 (\epsilon - n_0^2) E_x^o, \quad (5)$$

where  $E_x^s$  is the static electric field (just  $x$  component),  $D_x^s$  is the corresponding displacement field,  $V_0$  is the potential applied to the electrode,  $k_0 = 2\pi/\lambda_0$  is the wavevector of the light beam,  $n_0$  is the refractive index,  $\epsilon_0$  is the vacuum dielectric constant,  $\Delta\epsilon^s = \epsilon_{\parallel} - \epsilon_{\perp}$  is the birefringence at external field frequencies equal to the difference between the relative (a dimensional) permittivities for parallel and perpendicular directions respect to  $z$ ,  $\Delta\epsilon^o = n_{\parallel}^2 - n_{\perp}^2$  is the optical birefringence equal to the difference between the refraction indices for parallel and perpendicular directions,  $\epsilon(\theta)$  is the nonlinear optical permittivity and  $\epsilon_{xx}(\theta)$  is the nonlinear electrical permittivity. Let us take  $\epsilon = \epsilon_{\perp}^o + \Delta\epsilon^o \sin^2 \theta$ , and  $\epsilon_{xx} = \epsilon_{\perp}^s + \Delta\epsilon^s \sin^2 \theta$  for the nonlinear permittivities.

In this paper the properties of NLC E7 are used:  $K = 12 \times 10^{-12}$  N,  $\epsilon_{\parallel} = 19.6$ ,  $\epsilon_{\perp} = 5.1$ ,  $n_{\parallel} = 1.6954$ , and  $n_{\perp} = 1.5038$ . Hence,  $\Delta\epsilon^s = 19.6 - 5.1 = 15.4$ , and  $\Delta\epsilon^o = n_{\parallel}^2 - n_{\perp}^2 = 1.6954^2 - 1.5038^2 = 0.613$ . We also have used  $n_0 = 1$ ,  $\lambda_0 = 1.064 \mu\text{m}$  and  $\epsilon_0 = 8.8541878176 \times 10^{-12}$  CV $^{-1}$ m $^{-1}$ .

These models are solved numerically in an iterative way so that an auto-consistent solution is determined.<sup>17,26</sup> The procedure is as follows: first, find an auto-consistent solution for static Eqs. (3) and (4). Second, use such solution to calculate the next propagation step in Eq. (5). And third, since this optical field influences again the whole static problem, repeat the first two steps until convergence. Note that Eq. (4) has an analytical solution. However (3) and (5) are solved numerically by a standard finite difference time domain (FDTD) method.<sup>27</sup> Since Eq. (3) is nonlinear an iterative scheme is required, solving a linear system of equations obtained by the FDTD method in every step. Equation (5) is linear so no iteration is required.

### 3. Comparison Between the Models

Our goal is to check to which extent the linear approximation is valid when studying the transverse oscillation mentioned above. In order to compare the results of each model, we take Mod. B to the range of parameters at which it is equivalent to Model A, this is taking  $k_0 = 1$ ,  $\epsilon_{\perp} = 1$ ,  $\Delta\epsilon = 4$ ,  $K = \nu\Delta\epsilon\epsilon_0\theta_{\max}/2$ ,  $\theta \ll 1$ , and  $\theta_{\max} = 1/2$ .

The amplitude of this transverse oscillation depends on the initial position of the beam with its period depending on the optical power and the offset of its initial position.

The dependence on the period of such oscillation on the optical amplitude is shown in Fig. 2 for both models. We can see there are different behaviors between the models. Figure 2(a) shows that the period of the oscillations for Model A evolves monotonically decreasing with the square root of the power density (since the force exerted by the boundary is linear with the power density). For Mod. B, Fig. 2(b) shows a nonmonotonic behavior, with a minimum period close to an amplitude of 0.4 for every value of the offset, as expected since the linear relationship between intensity and index perturbation is no longer valid. Figure 2 also illustrates that



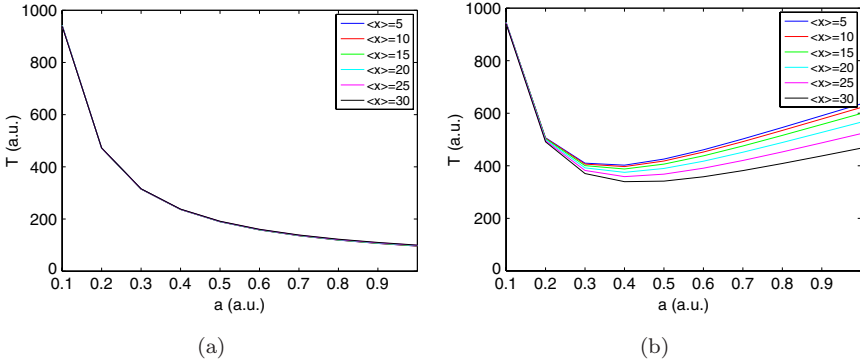


Fig. 2. Period of the transverse oscillation  $T$  versus optical amplitude  $a$  for (a) Model A and (b) Model B.

the initial position does not influence the period of the transverse oscillation for Model A but they do for Mod. B, as stated in Ref. 16 and experimentally in Ref. 17 as well as in the case of thermal media.<sup>18</sup> Model A shows very little dependence of the period over the offset. We can see the nonlinearity of Mod. B induces this minimum, which can be explained in terms of the nonlinear force exerted by the boundary. This force has been calculated numerically through the second derivative of the trajectory of the center of mass. It is known that the force grows when approaching the boundary.<sup>17</sup> Figure 3 shows the dependence on the initial position of the force and its dependence with the optical amplitude. There it can be seen the nonlinearity of the relationship among the force and the optical amplitude, which makes it appear as a minimum responsible of the minimum appearing in the period dependence on the optical amplitude.

The evolution of the amplitude peak of the optical field is shown in Figs. 4 and 5. We see that results are qualitatively (even quantitatively) equal for low powers

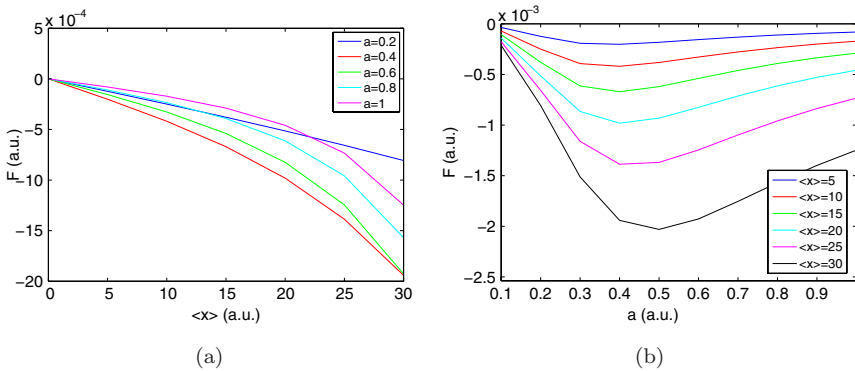


Fig. 3. (a) Force  $F$  exerted by the boundary versus initial position  $\langle x \rangle$  for different optical amplitudes,  $a$ . (b) Force  $F$  versus optical amplitude  $a$  for different offset values  $\langle x \rangle$ .

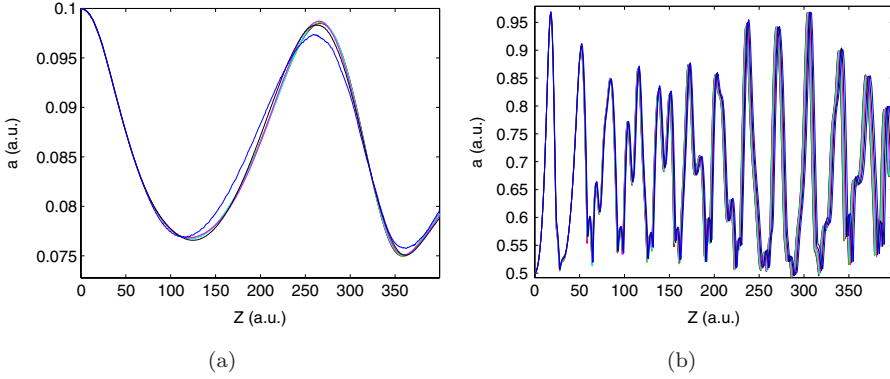


Fig. 4. Evolution of the optical peak amplitude  $a$  for (a)  $a = 0.1$  and (b)  $a = 0.5$  using Model A. Each color represents different values for the offset  $\langle x \rangle$ .

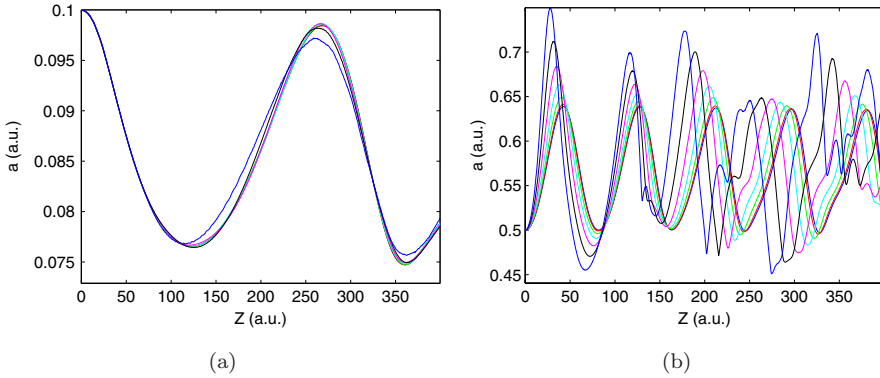


Fig. 5. Evolution of the optical peak amplitude  $a$  for (a)  $a = 0.1$  and (b)  $a = 0.5$  using Mod. B. Each color represents different values for the offset  $\langle x \rangle$ .

where the linear approximation is valid ( $a = 0.1$ ). However differences arise as power increases. We observe three main differences. First, the oscillation frequency of the peak amplitude increases for Mod. A as it does the optical amplitude but does not for Mod. B. This is closely related to the artificial higher values of the director orientation reached in Mod. A. Higher values of the midtilt yields higher values for the effective index. Model A establishes a linear relationship among the potential, or effective index profile, and the midtilt, the value of the tilt at its maximum, as it can be seen in Eq. (2). There it can be seen that the potential term is  $2\theta$ . However Mod. B has a potential term which is proportional to the dielectric permittivity  $\epsilon = \epsilon_{\perp} + \Delta\epsilon \sin^2 \theta$  (Ref. 1) which yields in a bounded saturable value for the midtilt that makes it reach more moderate realistic values than Mod. A. Moreover the greater the values of the effective index the bigger will be the self-focusing effect which explains the high value for the frequency of the optical amplitude oscillation.

Second, Mod. A almost overlap the amplitudes for different offset values, what can be best appreciated in Fig. 4, however, Mod. B shows a different behavior of the peak amplitude when changing the offset (see Fig. 5) since its frequency is much lower and behaviors for different offset values are not so equal as it occurs in Mod. A.

And third, the difference between its maximum value and its minimum one is much higher in Mod. A than in Mod. B as it appears a higher self-focusing effect in Mod. A due to a non-saturable effect of the midtilt. Such effect is an artifact of the linear approximation to the orientational nonlinearity, which is saturable as the midtilt is bounded by  $\pi/2$ .

We must also mention that the presence of transverse oscillations in solitons in nematics was first reported and studied by Peccianti *et al.*<sup>28</sup> and their breathing in Ref. 29. Our results show a good agreement in the weak nonlinear regime. A full comparison in the high nonlinear regime is beyond the scope of the present work and will be addressed in further research.

#### 4. Conclusions

In conclusion, we have investigated the limits of standard liquid crystal models when applied to study a transverse oscillation phenomenon induced by the boundary. Big differences appear when we work in the nonlinear regime where the simplified models are no longer valid. Numerical simulations are in good agreement with related results already published at the time they widen our knowledge of such transverse oscillation.

#### Acknowledgments

F. V. M. acknowledges partial support by Project No. MTM2010-19969 from the Ministerio de Ciencia e Innovación (MICINN) of Spain. A. F. thanks MICINN of Spain by its financial support (TIN2006-12890 and TEC2010-1532). P. F. C. thanks MICINN of Spain by its financial support (TIN2009-12359). The authors would also like to thank Dr. A. S. Desyatnikov for his important suggestions to this work.

#### References

1. I. C. Khoo, *Liquid Crystals: Physical Properties and Nonlinear Optical Phenomena* (John Wiley & Sons, New York, 1992).
2. M. Peccianti and G. Assanto, Nematic liquid crystals: A suitable medium for self-confinement of coherent and incoherent light, *Phys. Rev. E* **65** (2002) 035603.
3. M. Peccianti and G. Assanto, Nematicons, *Phys. Rep.* **516** (2012) 147–208.
4. Y. Izdebskaya, V. Shvedov, A. Desyatnikov, W. Krolikowski and Y. Kivshar, Soliton bending and routing induced by interaction with curved surfaces in nematic liquid crystals, *Opt. Lett.* **35** (2010) 1692–1694.
5. A. Piccardi, G. Assanto, L. Lucchetti and F. Simoni, All-optical steering of soliton waveguides in dye-doped liquid crystals, *Appl. Phys. Lett.* **93** (2008) 171104-1-3.

6. M. Peccianti, C. Conti, G. Assanto, A. De Luca and C. Umeton, All-optical switching and logic gating with spatial solitons in liquid crystals, *Appl. Phys. Lett.* **81** (2002) 3335–3337.
7. S. V. Serak, N. V. Tabiryán, M. Peccianti and G. Assanto, Spatial soliton all-optical logic gates, *IEEE Phot. Tech. Lett.* **18** (2006) 1287–1289.
8. A. Pasquazi, A. Alberucci, M. Peccianti and G. Assanto, Signal processing by opto-optical interactions between self-localized and free propagating beams in liquid crystals, *Appl. Phys. Lett.* **87** (2005) 261104-1–3.
9. A. Minzoni, N. F. Smyth, Z. Xu and Y. Kivshar, Stabilization of vortex soliton beams in nematic liquid crystals, *Phys. Rev. A* **79** (2009) 063808-1–7.
10. A. Minzoni, N. F. Smyth and Z. Xu, Stability of an optical vortex in a circular nematic cell, *Phys. Rev. A* **81** (2010) 033816-1–5.
11. D. Buccoliero, A. Desyatnikov, W. Krolikowski and Y. Kivshar, Spiraling multivortex solitons in nonlocal nonlinear media, *Opt. Lett.* **33** (2008) 198–200.
12. F. Ye, Y. Kartashov and L. Torner, Stabilization of dipole solitons in nonlocal nonlinear media, *Phys. Rev. A* **77** (2008) 043821-1–7.
13. D. Buccoliero, A. Desyatnikov, W. Krolikowski and Y. Kivshar, Laguerre and Hermite soliton clusters in nonlocal nonlinear media, *Phys. Rev. Lett.* **98** (2007) 053901-1–4.
14. A. Alberucci, M. Peccianti, G. Assanto, A. Dyadyusha and M. Kaczmarek, Two-color vector solitons in non local media, *Phys. Rev. Lett.* **97** (2006) 153903-1–4.
15. C. Conti, M. Peccianti and G. Assanto, Complex dynamics and configurational entropy of spatial optical solitons in nonlocal media, *Opt. Lett.* **31** (2006) 2030–2032.
16. A. Alberucci and G. Assanto, Propagation of optical spatial solitons in finite-size media: Interplay between nonlocality and boundary conditions, *J. Opt. Soc. Am. B* **24** (2007) 2314–2320.
17. A. Alberucci, M. Peccianti and G. Assanto, Nonlinear bouncing of nonlocal spatial solitons at the boundaries, *Opt. Lett.* **32** (2007) 2795–2797.
18. B. Alfassi, C. Rotschild, O. Manela, M. Segev and D. N. Christodoulides, Boundary force effects exerted on solitons in highly nonlocal nonlinear media, *Opt. Lett.* **32** (2007) 154–156.
19. A. W. Snyder and D. J. Mitchell, Accessible solitons, *Science* **276** (1997) 1538–1541.
20. M. Peccianti, A. De Rossi, G. Assanto, A. De Luca, C. Umeton and I. C. Khoo, Electrically assisted self-confinement and waveguiding in planar nematic liquid crystal cells, *Appl. Phys. Lett.* **77** (2000) 7–9.
21. G. Assanto, A. Fratallocchi and M. Peccianti, Spatial solitons in nematic liquid crystals: From bulk to discrete, *Opt. Express* **15** (2007) 5249–5259.
22. X. Hutsebaut, C. Cambournac, M. Haelterman, J. Beekman and K. Neyts, Measurement of the self-induced waveguide of a solitonlike optical beam in a nematic liquid crystal, *J. Opt. Soc. Am. B* **22** (2005) 1424–1431.
23. C. Conti, M. Peccianti and G. Assanto, Route to nonlocality and observation of accessible solitons, *Phys. Rev. Lett.* **91** (2003) 073901-1–4.
24. A. Alberucci, G. Assanto, D. Buccoliero, A. Desyatnikov, T. Marchant and N. Smyth, Modulation analysis of boundary induced motion of optical solitary waves in a nematic liquid crystal, *Phys. Rev. A* **79** (2009) 043816-1–8.
25. G. Assanto, N. F. Smyth and W. Xia, Modulation analysis of nonlinear beam refraction at an interface in liquid crystals, *Phys. Rev. A* **84** (2011) 033818-1–7.
26. J. Beekman, K. Neyts, X. Hutsebaut and M. Haelterman, One-dimensional simulation of field-induced director reorientation and lateral light propagation in liquid crystals, *Proc. SPIE* **5249** (2004) 577–585.

27. W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. (Cambridge University Press, 1992).
28. M. Peccianti, A. Fratalocchi and G. Assanto, Transverse dynamics of nematicons, *Opt. Express* **12** (2004) 6524–6529.
29. C. Conti, M. Peccianti and G. Assanto, Observation of optical spatial solitons in a highly nonlocal medium, *Phys. Rev. Lett.* **92** (2004) 113902-1–4.

PAPER • OPEN ACCESS

## Transparent Boundary Condition for Oseen-Frank Model. Application for NLC Cells With Patterned Electrodes

To cite this article: I Orquín-Serrano *et al* 2015 *J. Phys.: Conf. Ser.* **605** 012028

View the [article online](#) for updates and enhancements.

You may also like

- [Meshfree-based physics-informed neural networks for the unsteady Oseen equations](#)  
Keyi Peng, , Jing Yue et al.
- [Linear instability of a perturbed Lamb–Oseen vortex](#)  
Sherwin A Maslowe
- [A slender body theory in Oseen flow obtained by expanding the Oseenlets in the Green's integral representation](#)  
Edmund Chadwick

**PRIME**  
PACIFIC RIM MEETING  
ON ELECTROCHEMICAL  
AND SOLID STATE SCIENCE

HONOLULU, HI  
Oct 6–11, 2024

Abstract submission deadline:  
**April 12, 2024**

**Learn more and submit!**

**Joint Meeting of**  
The Electrochemical Society  
•  
The Electrochemical Society of Japan  
•  
Korea Electrochemical Society

# Transparent Boundary Condition for Oseen-Frank Model. Application for NLC Cells With Patterned Electrodes

I Orquín-Serrano<sup>1</sup>, J Vijande<sup>2</sup>, F R Villatoro<sup>3</sup>, A Ferrando<sup>4</sup>, P Fernández de Córdoba<sup>1</sup>, and H Michinel<sup>5</sup>

<sup>1</sup>Instituto Universitario de Matemática Pura y Aplicada (IUMPA), Universidad Politécnica de Valencia (UPV), Camino de Vera s/n 46022 Valencia, Spain

<sup>2</sup>Departamento de Física Atómica, Molecular y Nuclear, Universidad de Valencia (UV) and IFIC (UV-CSIC), Valencia, Spain.

<sup>3</sup>Departamento de Lenguajes y Ciencias de la Computación, Universidad de Málaga (UMA), Málaga, Spain.

<sup>4</sup>Departament d'Òptica, Universitat de València (UV), 46100 Burjassot, Valencia, Spain.

<sup>5</sup>Optics Lab. University of Vigo, Facultade de Ciencias de Ourense, As Lagoas s/n. Ourense, 32004. Spain.

E-mail: isorser@doctor.upv.es

**Abstract.** In the present work a novel application of Transparent Boundary Conditions (TBC) to nematic liquid crystal cells (NLCC) with planar alignment and a patterned electrode is studied. This device is attracting great interest since it allows soliton steering by optically and externally induced waveguides. We employ the continuum Oseen-Frank theory to find the tilt and twist angle distributions in the cell under the one-constant approximation. The electric field distribution takes into account the whole 2D permittivity tensor for the transverse coordinates. Standard finite difference time domain methods together with an iterative method is applied to find an approximate solution to our coupled problem. A novel class of TBC is used to correctly define the boundary for both the distortion angle problem and the electric field distribution when using patterned electrodes. Thus, we achieve an important decrease of computational needs when solving this kind of problems and we are also capable of exploring weak anchoring conditions for NLCC.

## 1. Introduction

Nematic Liquid Crystal (NLC) Cells for lateral propagation are attracting much attention due, among other reasons, to their capability to form tunable reconfigurable waveguides. They can be formed by an external electric field [1] or by creating a self-consistent waveguide in bulk NLC [2]. These soliton induced waveguides can be used to implement all-optical switching and logic gating [3]. Tunable waveguide arrays can also be obtained in these kind of devices, in which all-optical switching phenomena [4] and optical multiband vector breathers [5] appear.

With respect to electrically induced waveguides, several NLC devices are being investigated. Many of them are based on using a patterned electrode to affect the molecules of the liquid crystal in such a way that the effective refractive index profile obtained yields a graded index



waveguide inside the nematics [6]. Other dispositions, where two electrodes at different voltages are employed, are used for soliton routing [7, 8].

The problem of simulating the optical evolution in a stripped electrode device is a difficult one and involves necessarily a  $2D + 1$  anisotropic beam propagation method that at least takes into account the transverse anisotropy induced by the external electric field. Moreover, vectorial effects appear when the  $y$  component (see Fig. 1) of the external electric field is not negligible and pushes the molecules of the liquid crystal out of the propagation plane, making them twist. This implies changes in the polarization of the optical beam and also losses of the extraordinary solitonic beam. Some approximations are usually employed to avoid this complexity, obtaining good agreement with the experiments. One of these approximations is not considering the twist movement of the molecules. This approximation holds for low voltage values and for those cases where, although voltages are high, light moves far away from the vicinity of the patterned electrode where this twist is noticeable. Hard anchoring conditions for the orientational problem are usually employed. This boundary condition is realistic since the glass plates confining the NLC can be treated in such a way that tilt and twist angle are fixed at the boundary. However we can also work with a non-fixed value for the twist angle next to the glass plates allowing the tilt angle to be fixed. This is also realistic since a suitable treatment of the glass plates that allows this twist movement of the director at the boundary can be found.

In a weak anchoring situation the elastic constant depends on the position, in the proximity of the boundaries [9] and to determine the value of this surface (or anchoring) energy, small deviations close to the Friedericksz transition are usually considered [10]. The anchoring energy for deformations will be neglected as it is much lower than the corresponding energy for tilt deformations (typically, one or even two orders of magnitude)[11]. We will show the numerical boundary treatment in this case, what could be used for further numerical research looking for better characterizing the boundary influence on the elastic constants under weak anchoring conditions as it was done in some studies in the past[12].

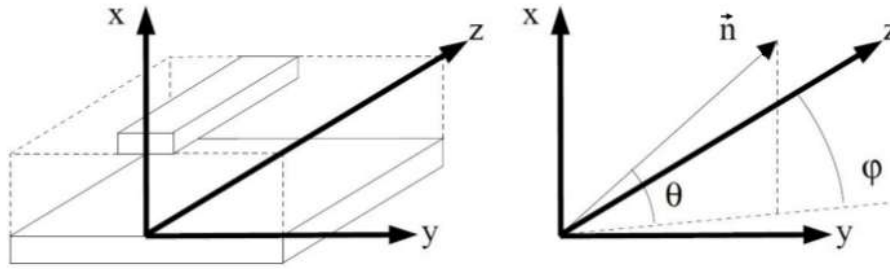
This is the regime we want to show in this paper, where transparent boundary conditions for the twist angle will be necessary in the boundary next to the glass plates. These simulations widens our knowledge of what is actually happening in the area closest to the electrode.

This paper is organized as follows: section 2 presents the physical model, section 3 shows the numerical treatment that allows us to approximate the solution of the problem using the boundary condition explained in section 4. Some results of our numerical model are presented in section 5 for different electrode configurations and finally our conclusions are presented in section 6.

## 2. Physical Model

We want to simulate a NLC device made of two glass plates containing commercial NLC E7 in between indium tin oxide transparent electrodes connected to the glass plates to apply an external voltage to the liquid crystal. It can be seen in Fig. 1 the geometrical properties of such a device;  $z$  coordinate stands for the propagation axis while  $x$  and  $y$  represent transverse spatial dimensions. Tilt angle ( $\theta$ ) appears with any movement of the director along a plane perpendicular to the  $yz$  plane. Twist angle ( $\varphi$ ) is defined in  $zy$  plane and holds for the twist deformations of the director. Planar alignment imposes director orientation nearly parallel to the glass plates, i.e., small tilt angle ( $\theta = 2^\circ$ ).  $V = 0$  is applied to the lower electrode affecting the whole glass plate. Upper one is thinner and doesn't affect the whole upper glass plate, so we will refer to it as the stripped electrode, where  $V \neq 0$  is applied. This physical configuration makes the numerical boundary difficult to work with because of its undefinition where the upper electrode doesn't influence the liquid crystal orientation. This discontinuity on the upper boundary, forced by the discontinuous electrode makes appear an  $y$  component for the electric field, such a component being responsible for the twist angle to appear in the proximities of





**Figure 1.** Sketch of the nematic liquid crystal device with patterned electrode (left) and notation for the planar alignment in the cell (right).

the stripped electrode. So a complete molecular reorientation is needed to explain both angle distributions. The whole orientational problem has already been solved in an iterative way for different NLC configurations where no discontinuous electrode is taken into account [13]. We extend this iterative method to the two dimensional case and we also present the two dimensional electric field distribution in those cases where tilt and twist angle distributions affect the electric field imposed by the electrodes, i.e., taking into account the whole 2D permittivity tensor.

### 2.1. Electric field distribution

We start from the Maxwell equations

$$\left. \begin{array}{l} \nabla D = 0 \\ D = \epsilon E \\ E = -\nabla V \end{array} \right) \rightarrow \nabla (\epsilon \nabla V) = 0, \quad (1)$$

where  $\epsilon_{ij}$  with  $i = x, y, z$  and  $j = x, y, z$  are permittivity tensor terms, given by

$$\epsilon = \begin{pmatrix} \epsilon_{\perp} + \Delta\epsilon \sin^2 \theta & \Delta\epsilon \sin \theta \cos \theta \sin \varphi \\ \Delta\epsilon \sin \theta \cos \theta \sin \varphi & \epsilon_{\perp} + \Delta\epsilon \cos^2 \theta \sin^2 \theta \end{pmatrix}, \quad (2)$$

where  $\epsilon_{\perp}$  is the perpendicular (to the glass plates) component of the permittivity and  $\Delta\epsilon = \epsilon_{\parallel} - \epsilon_{\perp}$  is the birefringence of the medium. Tilt ( $\theta(x, y)$ ) and twist ( $\varphi(x, y)$ ) angles depend only on these transverse coordinates. By doing so we are uncoupling the  $z$  coordinate therefore obtaining a static transverse distribution. We also consider negligible any dependence on  $z$  of the functions, i.e.  $V(x, y, z) = V(x, y)$ ,  $\frac{\partial V}{\partial z} = 0$ ,  $\frac{\partial^2 V}{\partial z^2} = 0$ ,  $\frac{\partial^2 V}{\partial y \partial z} = 0$ ,  $\frac{\partial^2 V}{\partial x \partial z} = 0$ ,  $\frac{\partial \theta}{\partial z} = 0$ ,  $\frac{\partial \varphi}{\partial z} = 0$ .

Substituting the dielectric tensor expressions in Eq. (1) and not considering any dependence on the  $z$  coordinate one obtains for the static two dimensional problem in the transverse plane,

$$a_1 \frac{\partial^2 V}{\partial x^2} + a_2 \frac{\partial^2 V}{\partial y^2} + a_3 \frac{\partial^2 V}{\partial x \partial y} + a_4 \frac{\partial V}{\partial x} + a_5 \frac{\partial V}{\partial y} = 0, \quad (3)$$

with

$$a_1 = \frac{1}{\Delta x^2} (\epsilon_{\perp} + \Delta\epsilon \sin^2 \theta), \quad (4)$$

$$a_2 = \frac{1}{\Delta y^2} (\epsilon_{\perp} + \Delta\epsilon \cos^2 \theta \sin^2 \varphi), \quad (5)$$

$$a_3 = \frac{1}{4\Delta x \Delta y} \Delta\epsilon \sin(2\theta) \sin \varphi, \quad (6)$$

$$a_4 = \frac{1}{2\Delta x} \left( \Delta\epsilon \left( \cos(2\theta) \sin\varphi \frac{\partial\theta}{\partial y} + \cos\theta \sin\theta \left( \cos\varphi \frac{\partial\varphi}{\partial y} + 2\frac{\partial\theta}{\partial x} \right) \right) \right), \quad (7)$$

$$a_5 = \frac{1}{2\Delta y} \left( \Delta\epsilon \left( (\cos^2\theta \sin\varphi - \sin^2\theta \sin\varphi) \frac{\partial\theta}{\partial x} \right. \right. \quad (8)$$

$$\left. \left. - \sin(2\theta) \sin^2\varphi \frac{\partial\theta}{\partial y} + \cos\theta \cos(\varphi) \sin\theta \frac{\partial\varphi}{\partial x} + \cos^2\theta \sin(2\varphi) \frac{\partial\varphi}{\partial y} \right) \right). \quad (9)$$

This is the equation we solve to obtain the electric field distribution. Notice that the dielectric tensor is responsible of the graded index profile which appears in the Beam Propagation Method, this is the way this whole static problem is coupled with the optical problem, which we are not studying here.

Finite Difference method is used to approximate this equation. Centered second order differences are employed for approximating the differential operators. Newton-Raphson method is used to solve the nonlinear equations of the finite difference scheme.

## 2.2. Elastic distortion angles distribution

As it was mentioned before, Oseen-Frank theory is employed to describe the molecular orientation in the NLC cell [14]. The total free energy of the liquid crystal is calculated to obtain its minimum via the Euler Lagrange theory. All three energy terms are considered in this model: the electric energy due to the external voltage, the electric energy due to the optical beam propagating inside the cell, and the elastic distortion energy.

The elongated molecules from the NLC feel any electric field due to its electric polarity, so they react changing its orientation so that it is aligned with the total electric field. Thus, molecules orientation is defined by the director vector  $\bar{n} = (\sin\theta, \cos\theta \sin\varphi, \cos\theta \cos\varphi)$ .

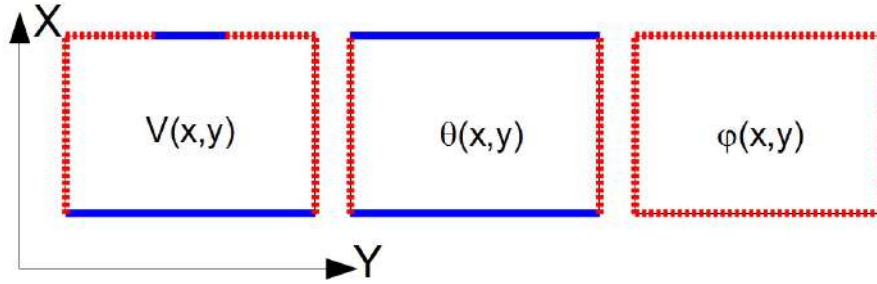
The distortion energy per volume unit considers three different types of distortion, each one with different elastic constants  $K$ . Under *one constant approximation*, all three elastic constants, related to different molecular distortions, are considered to be equal resulting in a more simplified equation. Minimizing the energy density can be done by transforming this problem to the solution of a partial differential equation via Euler-Lagrange theory, yielding:

$$\begin{aligned} & -K \left( \frac{\partial^2\theta}{\partial x^2} + \frac{\partial^2\theta}{\partial y^2} \right) + \sin(2\theta) \left[ -\frac{K}{2} \left( \frac{\partial\varphi}{\partial x}^2 + \frac{\partial\varphi}{\partial y}^2 \right) \right. \\ & + \epsilon_0 \left( -\frac{1}{2} (\Delta\epsilon^s |E_x^s|^2 + \Delta\epsilon^o |E_x^o|^2) + \frac{\Delta\epsilon^s}{4} |E_y^s|^2 (1 - \cos(2\varphi)) \right. \\ & \left. \left. + \frac{1}{2} \Delta\epsilon^o \sin^2\varphi |E_y^o|^2 \right) \right] - \frac{1}{2} \epsilon_0 \cos(2\theta) \sin\varphi \left( \Delta\epsilon^o (E_y^{o*} E_x^o + E_x^{o*} E_y^o) \right. \\ & \left. + 2\Delta\epsilon^s E_x^s E_y^s \right) = 0, \end{aligned} \quad (10)$$

and

$$\begin{aligned} & K \cos\theta \left[ \cos\theta \left( \frac{\partial^2\varphi}{\partial x^2} + \frac{\partial^2\varphi}{\partial y^2} \right) - 2 \sin\theta \left( \frac{\partial\theta}{\partial x} \frac{\partial\varphi}{\partial x} + \frac{\partial\theta}{\partial y} \frac{\partial\varphi}{\partial y} \right) \right] \\ & + \frac{1}{2} \epsilon_0 \cos^2\theta \sin(2\varphi) \left( \Delta\epsilon^o |E_y^o|^2 + \Delta\epsilon^s |E_y^s|^2 \right) \\ & + \frac{1}{4} \epsilon_0 \cos\varphi \sin(2\theta) \left( \Delta\epsilon^o (E_y^{o*} E_x^o + E_x^{o*} E_y^o) \right) \\ & + 2\Delta\epsilon^s E_x^s E_y^s = 0, \end{aligned} \quad (11)$$

that will allow us to solve  $\theta(x, y)$  and  $\varphi(x, y)$  respectively. Typical values we use for the magnitudes involved are  $K = 12 \times 10^{-12}$  N,  $\Delta\epsilon^s = \epsilon_{\parallel}^s - \epsilon_{\perp}^s = 19.6 - 5.1$ ,  $\Delta\epsilon^o = \epsilon_{\parallel}^o - \epsilon_{\perp}^o = n_{\parallel}^2 - n_{\perp}^2 = 1.69^2 - 1.50^2$ .



**Figure 2.** Numerical domain for all three differential equations. Blue solid lines represent hard anchoring condition while red dotted lines hold for transparent boundary condition.

### 3. Numerical Model

All three equations (3, 10 and 11) are approximated by a finite difference method consisting of difference operators which approximates the differential operators to second order, both for first and second derivatives. The corresponding nonlinear boundary value problems require the use of an iterative method such as Newton-Raphson ([15]) in order to find an approximate solution. Following the notation of [16] we write the difference equation which approximates eq. (3)

$$\begin{aligned}
 a_{1_{j,k}} \delta_x^2 V_{j,k} + a_{2_{j,k}} \delta_y^2 V_{j,k} + a_{3_{j,k}} (\delta_{0y} (\delta_{0x} V_{j,k})) \\
 + a_{4_{j,k}} \delta_{0x} V_{j,k} + a_{5_{j,k}} \delta_{0y} V_{j,k} = 0,
 \end{aligned}
 \tag{12}$$

being  $\delta_x^2 V_{j,k} = V_{j-1,k} - 2V_{j,k} + V_{j+1,k}$  and  $\delta_y^2 V_{j,k} = V_{j,k-1} - 2V_{j,k} + V_{j,k+1}$  the second order approximations to the second spatial derivatives and  $\delta_{0x} V_{j,k} = V_{j+1,k} - V_{j-1,k}$  and  $\delta_{0y} V_{j,k} = V_{j,k+1} - V_{j,k-1}$  the second order approximations to the first spatial derivatives. Notice that difference operators  $\delta_{0x}$  and  $\delta_{0y}$  commute. Coefficients  $a_{i_{j,k}}$  represent the discretized version of the respective coefficients of eq. (3). It can be seen that coefficients  $a_{1_{j,k}}$ ,  $a_{2_{j,k}}$  and  $a_{3_{j,k}}$  are point approximations while  $a_{4_{j,k}}$  and  $a_{5_{j,k}}$  require the application of difference operators to the angle derivatives appearing in those coefficients. Thus the boundary condition will appear also in these coefficients.

We can rewrite eq. (12) in a more clear way

$$\begin{aligned}
 A_{NW_{j,k}} V_{j-1,k-1} + A_{N_{j,k}} V_{j,k-1} + A_{NE_{j,k}} V_{j+1,k-1} \\
 + A_{W_{j,k}} V_{j-1,k} + A_{P_{j,k}} V_{j,k} + A_{E_{j,k}} V_{j+1,k} \\
 + A_{SW_{j,k}} V_{j-1,k+1} + A_{S_{j,k}} V_{j,k+1} + A_{SE_{j,k}} V_{j+1,k+1} = 0
 \end{aligned}
 \tag{13}$$

where the coefficients stand for the location of the nodes in the discrete molecule of the difference method. This is a nine point molecule whose coefficients are  $A_{NW_{j,k}} = a_{3_{j,k}}$ ,  $A_{N_{j,k}} = a_{2_{j,k}} - a_{5_{j,k}}$ ,  $A_{NE_{j,k}} = -a_{3_{j,k}}$ ,  $A_{W_{j,k}} = a_{1_{j,k}} - a_{4_{j,k}}$ ,  $A_{P_{j,k}} = -2a_{1_{j,k}} - 2a_{2_{j,k}}$ ,  $A_{E_{j,k}} = a_{1_{j,k}} + a_{4_{j,k}}$ ,  $A_{SW_{j,k}} = -a_{3_{j,k}}$ ,  $A_{S_{j,k}} = a_{2_{j,k}} + a_{5_{j,k}}$ ,  $A_{SE_{j,k}} = a_{3_{j,k}}$  accounting for the north-west, north, north-east, west, central, east, south-west, south and south-east nodes respectively. This is a very convenient notation since the matrix of coefficients of eq. (13) and the matrix representing the jacobian of that difference functional coincides.

Equations (10) and (11) are solved analogously using second order difference operators and Newton-Raphson method. Notice that the numerical scheme of these equations yields a five-point molecule instead of the nine-point molecule appearing in eq. (13) due to the crossed second derivative approximation.

The main drawback of this formalism is that an iterative scheme is needed to reach an autoconsistent solution since equations (3), (10) and (11) are coupled. An iterative scheme can

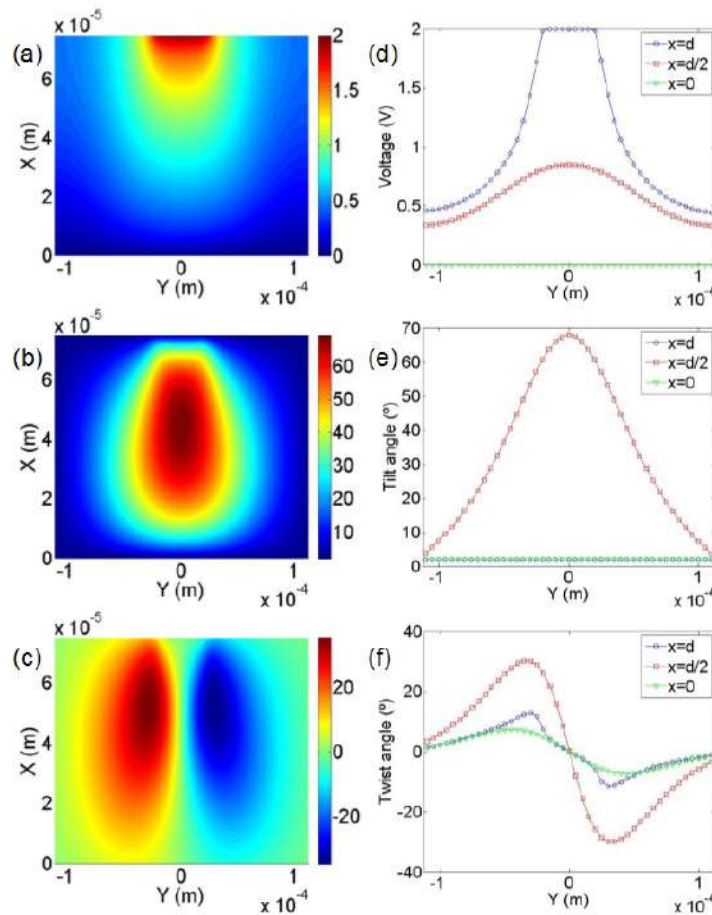
be used to obtain the solution to this coupled problem. First an iterative loop will find the auto consistent solution of the approximated equations of (10) and (11) and then use this solution to find the electric field distribution using (3). Then other iterative loop is needed for finding an autoconsistent solution of both problems the electric and the angular one. Thus, two nested iterative loops scheme have been used to find the whole autoconsistent solution.

#### 4. Transparent Boundary Condition

All three differential equations we solve have different boundary conditions as it can be seen in Fig. 2. First, the calculation for the external electric potential applied over the nematics requires hard anchoring condition wherever we have an electrode fixing the voltage. Where no electrode is located we use transparent boundary condition since it allows no restrictions over the voltage value at the boundary, which is a realistic boundary. Second, with respect to the tilt angle distribution,  $\theta(x, y)$  we use hard anchoring condition for the  $x$  boundary, since that is the area where the nematics is limited by the glass plate, which usually has a rubbing treatment to fix the tilt angle to a low value, typically  $\theta = 2^\circ$ , with no possibility of movement in this boundary. Finally, twist angle  $\varphi(x, y)$  is not fixed in the  $x$  boundary, since the glass plates allows twist movement of the molecules, parallel to the glass plates, after appropriate treatment of the surface. It can be seen in Fig. 2 that TBC are used for all three problems in the  $y$  boundary allowing a finite extent in this transverse coordinate of such device. Gratings can be simulated by employing periodic boundary conditions in  $y$  boundary.

TBC allow us to predict which has to be the value of the function at the boundary using the information we have in the interior domain. From the middle 70's on much work has been done to develop a total non reflecting boundary condition for the Schrödinger equation ([17, 18, 19, 20, 21, 22, 23, 24, 25]). Absorbing boundary conditions are still being studied and its application to dispersive waves phenomena [26] or its behavior with semilinear parabolic equations [27] have been published recently. Different approaches have been proposed, both analytical and numerical, but the work of G. R. Hadley has become one of the most popular results published in this topic because of its easy implementation at the time it offers very good results. Hadley's work presents the derivation of a TBC for the Schrödinger equation, based on physical considerations about the energy flow through the numerical boundaries [28, 29]. We use similar mathematical considerations which allow us to predict the value of the function at the boundary. In this work we extend the use of this TBC to other differential equations. Let us show the calculation of this transparent boundary condition with an example for the external electric potential  $V(x, y)$ . Let  $V_{j,k}$  be the discretized values  $V(j\Delta x, k\Delta y)$  where  $2 \leq j \leq N_X - 1$  and  $2 \leq k \leq N_Y - 1$ ,  $N_X$  and  $N_Y$  being the number of unknowns in the  $x$  and  $y$  direction respectively.

Consider a fixed value for  $j$  in the external electric potential case, say  $j = N_X/2$ . Calculate the set of values  $c_{y_j,k} = \frac{V_{j,k}}{V_{j,k+1}}$ , as the quotient of consecutive values in the  $y$  direction. Index  $k$  ranges from 2 to  $N_Y - 1$ , being  $N_Y$  the number of nodes in the  $y$  direction (with  $N_Y - 2$  unknowns). Notice that boundary values in the  $y$  direction in this set are  $V_{j,1}$  and  $V_{j,N_Y}$ . We could guess the values of the  $y$  boundary if we got a value for  $c_{y_j,1}$  and  $c_{y_j,N_Y-1}$ . Of course this set of values  $c_{y_j,k}$  depend on the value function. In order to obtain these quantities at the boundaries we use a linear extrapolation, saying  $c_{y_j,1} = 2 \cdot c_{y_j,2} - c_{y_j,3}$  and  $c_{y_j,N_Y-1} = 2 \cdot c_{y_j,N_Y-2} - c_{y_j,N_Y-3}$ . Depending on the solution shape at the boundary we must evaluate the convenience of using other higher order extrapolation in order to obtain better accuracy. Thus we can say  $V_{j,1} = V_{j,2} \cdot c_{y_j,1}$  and  $V_{j,N_Y} = V_{j,N_Y-1} / c_{y_j,N_Y-1}$ . We calculate all transparent boundaries in a similar way, using the extrapolated values of the convenient quotient for each case. These transparent boundary values for the external electric potential  $V_{j,k}$  are used in Eq. (3). Notice that not all the values for the quotient vector  $c_{x_j,k}$  (or analogously  $c_{y_j,k}$ ) are employed.



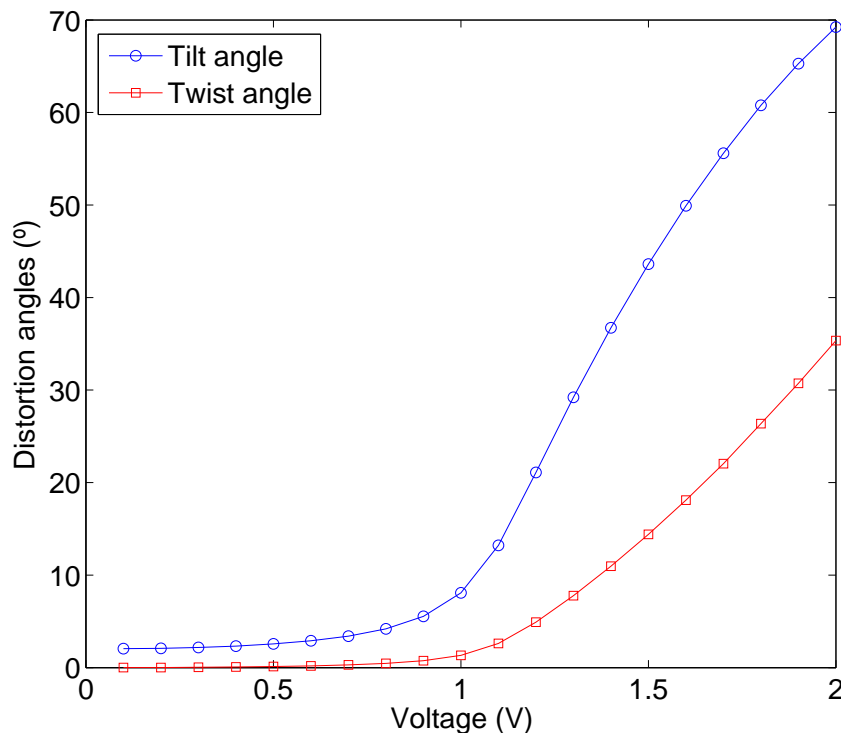
**Figure 3.** (a), (b) and (c) Show potential, tilt and twist angle distribution for  $V = 2$  V respectively. (d), (e) and (f) Show three different sections of (a), (b) and (c) respectively. Blue circles, red squares and green triangles represent cuts of the three dimensional representation on the left around  $x = d$ ,  $x = d/2$  and  $x = 0$  respectively. Notice that some curve might be overlapping the axes or other curve.

## 5. Results

### 5.1. One single electrode

The single electrode device represented in Fig. 1 has been used experimentally in order to guide nematicons in planar devices [30]. Fig. 3 shows the solution for the electric potential, the tilt angle and twist angle for  $V = 2$  V. There it can be seen the effect of the transparent boundary condition. Although optical power affects the molecular orientation (as it can be seen in equations 10 and 11, we perform simulations for different voltages with no optical power, so that it doesn't appear any other electric excitation than the external field. It can be seen in Fig. 4 the effect of the voltage on both the tilt and the twist angle. Both angles show different behavior: tilt angle changes more abruptly than twist angle. Changes begin at voltage  $V = 1$  V since that supposes the minimal amount of energy necessary to start moving the LC molecules, what is known as the Friedericksz threshold. Models which don't take into account twist angle shouldn't be appropriate for describing voltages above one volt since from then on twist angle is no longer negligible. To our knowledge this is the first time that transparent boundary conditions are applied to these equations.

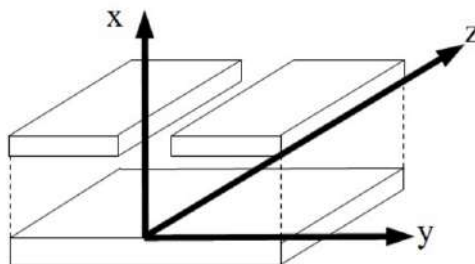
This one electrode device has already been studied to show the lateral propagation of a soliton



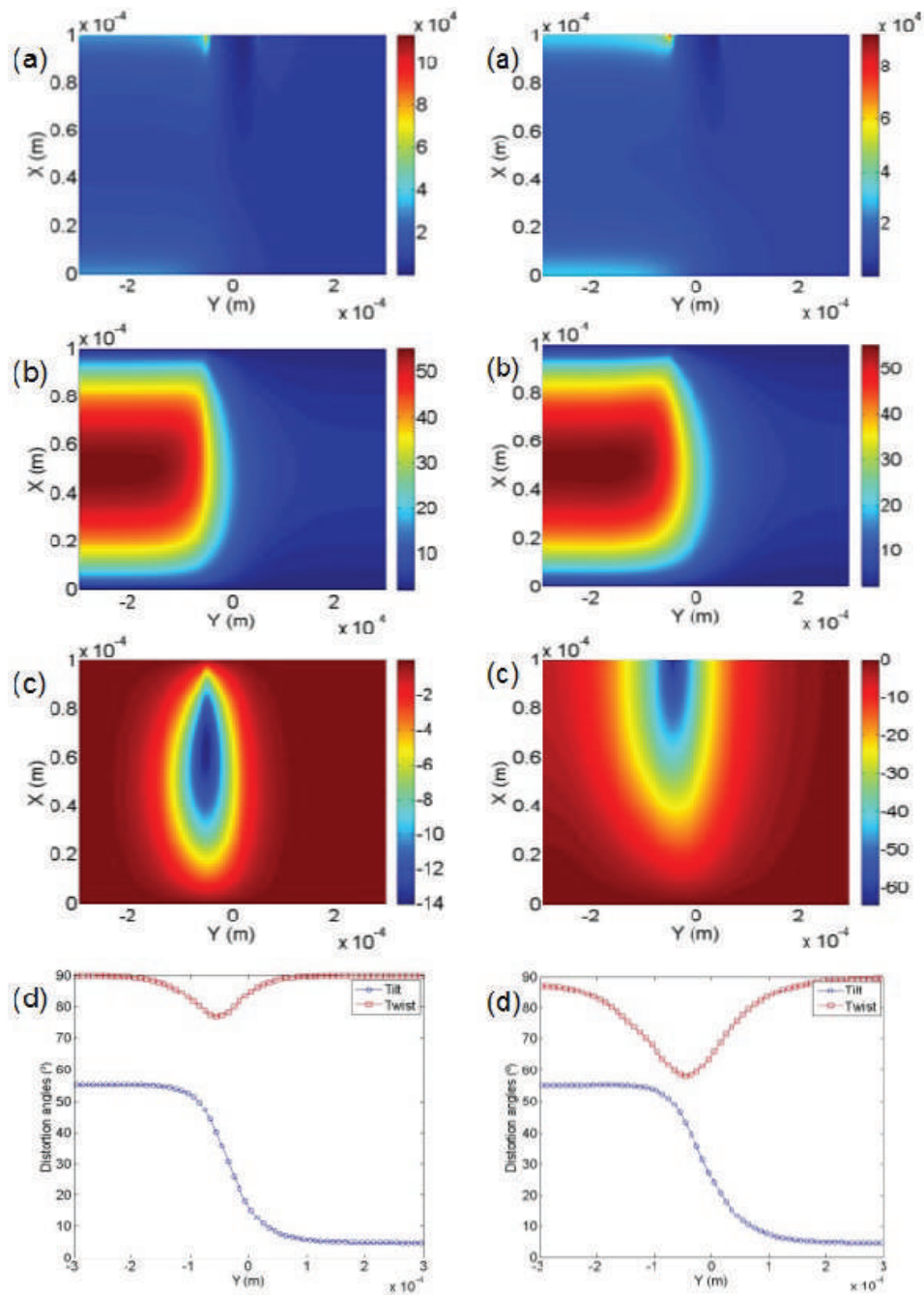
**Figure 4.** Voltage dependence over the peak values for the tilt and twist angles. These are the results of 200 simulations ranging from  $V = 0.01$  V to  $V = 2$  V.

inside the NLCC. However it is said that it is necessary to extend artificially the computational domain in order to solve the electric field distribution inside the cell since nothing is known from the solution at the boundary [30]. The idea is setting a fixed value for the solution in an artificial boundary far enough from the real one so that one can say that it doesn't affect the solution inside the original numerical domain. Our TBC permits us to avoid this artificial extension of the 2D numerical domain, lowering this way the computational needs to solve the whole coupled problem.

5.2. Two different electrodes with gap in between



**Figure 5.** NLC Device with two different electrodes in the upper area.



**Figure 6.** Solution for a two electrode device with Dirichlet boundary conditions at  $x$  boundaries (on the left) and using transparent boundary conditions at  $x$  boundaries for twist angle (on the right). Voltages are  $V_1 = 1.5$  V and  $V_2 = 0.7$  V. (a) Electric field, (b) tilt angle and (c) twist angle. Plot (d) show cross sections of both angles, tilt and twist, at  $x = d/2$ .

NLC Devices for lateral propagation and two electrodes with a gap in between have been studied for demonstrating tunable refraction and reflection of a soliton inside the nematics [31]. Fig. 5 shows which is the device we are referring to. Hard anchoring boundary conditions have been employed for the distortion angles, tilt and twist, in the  $x$  boundary since these values are fixed experimentally using a rubbing technique of the glass plates. Tilt angle is always fixed at the boundary since it is very difficult to allow a tilt movement of the molecules in a plane perpendicular to the glass plate there. However other boundary conditions are also realistic for the twist angle. Particularly, the *hybrid* case where the  $x = 0$  boundary is defined by a hard anchoring condition and the  $x = d$  boundary is transparent, allowing the molecules to twist at the boundary  $x = d$ , is also realistic. We study this case as a clear example of the use of our transparent boundary conditions. Notice that also transparent boundary conditions are employed for all three problems ( $V(x,y)$ ,  $\theta(x,y)$  and  $\varphi(x,y)$ ) in the  $y$  boundaries, allowing thus to study a single device of a finite extent.

In the device described in [31] the electrodes are used to create an effective index distribution so that it affects the nematic trajectory in different ways. We have performed the simulations of such device in two different scenarios: total anchoring condition in the  $x$  boundary, and the hybrid case mentioned before. The results are illustrated in Fig. 6, where we show the difference among both configurations. No changes in tilt angle distribution are observed since the calculation remains the same in both cases. However, twist angle distributions differ because of the effect of the transparent boundary condition applied in  $x = d$ . In fact, the amplitude of the twist angle in the latter case is bigger since it is not restricted by the boundary. This also affects the electric field distribution which reaches lower values in that boundary.

The application of a TBC to NLC devices allow us to study new configurations where weak anchoring conditions are relevant. For example in-plane movements of the molecule (those corresponding to the twist deformation) at the boundary are common in twisted nematic liquid crystal cells which may be used as displays, for example.

## 6. Conclusions

We have developed a realistic two dimensional numerical model for the coupled problem of finding the electric field distribution in a planar NLC cell considering the effects of both the tilt and twist director deformations. Oseen-Frank theory for the director distribution has been employed. To our knowledge, transparent boundary conditions are applied for each differential equation for the first time, thus avoiding time-consuming codes where artificial extension of the numerical domain were necessary to simulate devices with patterned electrodes correctly. The use of TBC in the  $y$  direction allow us to simulate one single electrode device properly. In the  $x$  direction TBC are only used for the twist deformation in the single electrode device. It reaches non negligible values close to the upper boundary where the patterned electrode is located. Other devices have been simulated to show the robustness of our code, for example a two electrode device with different voltages has been calculated. A TBC for the twist deformation at all four boundaries has been used to show the differences between the code already solved in the literature and our code, which enlightens a bit more the weak anchoring situation (for the twist angle) in those kind of devices.

## Acknowledgments

A.F. thanks Ministerio de Ciencia e Innovación of Spain by its financial support (TIN2006-12890 and TEC2010-1532). F.R.V. thanks Ministerio de Ciencia e Innovación (MICINN) of Spain by its financial support (TIN2009-12359 and MTM2010-19969). P.F.C. thanks MICINN of Spain by its financial support (TIN2009-12359).



## References

- [1] Peccianti M, De Rossi A, Assanto G, De Luca A, Umetsu C and Khoo I C 2000 *Applied Physics Letters* **77**
- [2] Peccianti M and Assanto G 2001 *Opt. Lett.* **26** 1690–1692
- [3] Peccianti M, Conti C, Assanto G, De Luca A and Umetsu C 2002 *Applied Physics Letters* **81**
- [4] Fratallocchi A, Assanto G, Brzdkiewicz K A and Karpierz M A 2005 *Applied Physics Letters* **86** 051112
- [5] Fratallocchi A, Assanto G, Brzdkiewicz K A and Karpierz M A 2005 *Opt. Lett.* **30** 174–176
- [6] Beeckman J, Neyts K and Haelterman M 2006 *Journal of Optics A: Pure and Applied Optics* **8** 214
- [7] Peccianti M, Dyadyusha A, Kaczmarek M and Assanto G 2008 *Phys. Rev. Lett.* **101**(15) 153902
- [8] Kivshar Y 2006 *Nature Phys.* **2** 729
- [9] Alexe-Ionescu A L, Barberi R, Barbero G and Giocondo M 1994 *Phys. Rev. E* **49**(6) 5378–5388
- [10] Sodr e N, Zola R S, Lenzi E K and Evangelista L R 2006 *Rev. Mex. Rev. Mex. de F sica S* **52** 80–84
- [11] J r me B 2008 *Handbook of Liquid Crystals* (Wiley-VCH Verlag GmbH) pp 535–548
- [12] Allen M P and Frenkel D 1988 *Phys. Rev. A* **37**(5) 1813–1816
- [13] Wang Q, He S, Yu F and Huang N 2001 *Optical Engineering* **40** 2552–2557
- [14] Khoo I C 2006 *Liquid Crystals: Physical Properties and Nonlinear Optical Phenomena* (Cambridge University Press)
- [15] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 2007 *Numerical Recipes: The Art of Scientific Computing* (John Wiley & Sons, Inc.)
- [16] Thomas J W 1995 *Numerical Partial Differential Equations* (Springer, New York)
- [17] Engquist B and Majda A 1977 *Math. Comp.* **31** 629–651
- [18] Grote M and Keller J 1995 *SIAM Journal on Applied Mathematics* **55** 280–297
- [19] Hagstrom T, Warburton T and Givoli D 2010 *Journal of Computational and Applied Mathematics* **234**
- [20] Baskakov V and Popov A 1991 *Wave Motion* **14** 123 – 128
- [21] Lubich C and Sch dle A 2002 *SIAM Journal on Scientific Computing* **24** 161–182
- [22] Kuska J P 1992 *Phys. Rev. B* **46**(8) 5000–5003
- [23] Menza L D 1996 *Applied Mathematics Letters* **9** 55 – 59
- [24] Sch A 2002 *Wave Motion* **35** 181 – 188
- [25] Arnold A, Ehrhardt M and Sofronov I 2003 *Commun. Math. Sci.* **1** 501–556
- [26] Lindquist J M, Neta B and Giraldo F X 2012 *Applied Mathematics and Computation* **218** 6666 – 6676
- [27] Sun Z Z, Wu X, Zhang J and Wang D 2012 *Applied Mathematics and Computation* **218** 5187 – 5201
- [28] Hadley G R 1991 *Opt. Lett.* **16** 624–626
- [29] Hadley G 1992 *Quantum Electronics, IEEE Journal of* **28** 363–370
- [30] Beeckman J, Chaubiska K and Neyts K 2006 *Ferroelectrics* **344** 225–231
- [31] Peccianti M, Dyadyusha A, Kaczmarek M and Assanto G 2006 *Nat Phys* **2** 737–742

Letter

# Integral Involving Bessel Functions Arising in Propagation Phenomena

Yu. L. Ratis, P. Fernández de Córdoba , J. M. Isidro \*  and R. Jaime

Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, 46022 Valencia, Spain; ratis@rambler.ru (Y.L.R.); pfernandez@mat.upv.es (P.F.d.C.); ramon@upr.edu.cu (R.J.)

\* Correspondence: joissan@mat.upv.es

Received: 17 April 2019; Accepted: 8 May 2019; Published: 16 May 2019



**Abstract:** An new integral identity involving the product of two modified Bessel functions  $K_0$  and  $K_1$ , previously unreported, is presented and proved.

**Keywords:** bessel functions; feynman propagators; quantum gravity

## 1. Introduction

Bessel functions are ubiquitous in mathematical physics and engineering. Yet they continue to surprise us with amazing new properties that, to the best of our knowledge, seem not to have been recorded anywhere, at least not in classic literature on the subject [1–3]. In this paper we state and prove one such property. We also speculate on a possible application of this new property to an eventual theory of quantum gravity.

Concerning our use of Bessel functions, we follow the notations and conventions of ref. [2].

## 2. Statement and Proof of Theorem

**Theorem 1.** *The following identity holds:*

$$K_0\left(\sqrt{x^2 + y^2}\right) = \frac{|y|}{\pi} \int_{-\infty}^{\infty} K_0(|x - \zeta|) \frac{K_1\left(\sqrt{y^2 + \zeta^2}\right)}{\sqrt{y^2 + \zeta^2}} d\zeta, \quad (1)$$

where  $x, y \in \mathbb{R}$  with  $y \neq 0$ .

**Proof.** Let us consider the contour integral in the complex  $\zeta$ -plane

$$f_C(x, y) := \frac{|y|}{\pi} \int_C K_0(|x - \zeta|) \frac{K_1\left(\sqrt{y^2 + \zeta^2}\right)}{\sqrt{y^2 + \zeta^2}} d\zeta \quad (2)$$

along a contour  $C$  to be specified presently. For a fixed  $x$ , the change of variables

$$\zeta := z + x \quad (3)$$

transforms (2) into

$$f_C(x, y) = \frac{|y|}{\pi} \int_C K_0(|z|) \frac{K_1\left(\sqrt{y^2 + (z + x)^2}\right)}{\sqrt{y^2 + (z + x)^2}} dz. \quad (4)$$

To begin with, we assume  $y > 0$ . Let the contour  $C$  extend along the straight segment from  $-x + iy$  to  $x + iy$ , while closing on the upper half-plane along a semicircle centred at the point  $(0, y)$  with radius  $R$ . By Cauchy’s theorem

$$f_C(x, y) = 2\pi i \sum \text{Res} \left\{ K_0(|z|) \frac{K_1\left(\sqrt{y^2 + (z+x)^2}\right)}{\sqrt{y^2 + (z+x)^2}} \right\} \tag{5}$$

where the sum extends over all the residues enclosed by  $C$ .

We will need the asymptotics of  $K_0(z)$  and  $K_1(z)$  on the complex plane. As  $z \rightarrow 0$  we have

$$K_0(z) = -\ln z + \dots, \quad K_1(z) = \frac{1}{z} + \dots \tag{6}$$

while for  $z \rightarrow \infty$

$$K_0(z) = e^{-z} \sqrt{\frac{\pi}{2z}} + \dots, \quad K_1(z) = e^{-z} \sqrt{\frac{\pi}{2z}} + \dots, \quad |\arg(z)| < \pi. \tag{7}$$

Away from the negative real axis in the complex plane, the functions  $K_0$  and  $K_1$  are analytic. In Equation (7),  $K_0(|z|)$  and  $K_1(\sqrt{y^2 + (z+x)^2})$  both vanish as  $R \rightarrow \infty$ . Then the sum over residues in Equation (5) reduces to the single residue located at

$$z_0 = -x + iy. \tag{8}$$

In Equation (6), in a neighbourhood of  $z_0$  the function on the right-hand side of (5) behaves as

$$K_0(|z|) \frac{K_1\left(\sqrt{y^2 + (z+x)^2}\right)}{\sqrt{y^2 + (z+x)^2}} \simeq \frac{K_0(|z|)}{y^2 + (z+x)^2} = \frac{K_0(|z|)}{(z+x+iy)(z+x-iy)}. \tag{9}$$

Thus

$$\sum \text{Res} \left\{ K_0(|z|) \frac{K_1\left(\sqrt{y^2 + (z+x)^2}\right)}{\sqrt{y^2 + (z+x)^2}} \right\} = \frac{K_0(|z_0|)}{2iy}. \tag{10}$$

Altogether

$$\lim_{R \rightarrow \infty} f_C(x, y) = \frac{y}{\pi} 2\pi i \frac{K_0(|z_0|)}{2iy} = K_0(|z_0|), \tag{11}$$

which proves the statement made in Equation (1) when  $y > 0$ . An analogous reasoning holds when  $y < 0$ , the contour  $C$  now being closed by a semicircle in the lower half plane and the residue being located at  $z_0 = -x - iy$ , which introduces an extra minus sign. This completes the proof.  $\square$

### 3. Quantum Gravity?

The authors’ own interest in the subject arose out of contact [4–7] with problems in the theory of heat conduction [8] and quantum field theory [9]. Without aspirations to completeness (the subject is too vast to summarise) let us briefly mention one such occurrence of Bessel functions.

In the quantum field theory of a scalar field with mass  $m$  in Euclidean spacetime  $\mathbb{R}^d$  [9] one is interested in the Feynman propagator  $G_d$ . The latter is a Green function for the Klein-Gordon operator on  $\mathbb{R}^d$  satisfying a specific set of boundary conditions. With  $\mathbf{x}$  denoting a vector in  $\mathbb{R}^d$  and  $\|\mathbf{x}\| = \sqrt{\mathbf{x}^2}$  its Euclidean norm, it turns out that

$$G_d(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{\exp(i\mathbf{p}\mathbf{x})}{\mathbf{p}^2 + m^2} d^d \mathbf{p} = \frac{m^{d-2}}{(2\pi)^{d/2}} (m\|\mathbf{x}\|)^{1-d/2} K_{1-d/2}(m\|\mathbf{x}\|). \tag{12}$$

We see that the propagator  $G_d$  depends on the vector argument  $\mathbf{x} \in \mathbb{R}^d$  only through the norm  $\|\mathbf{x}\|$ . The latter is a real number that we may denote by  $x$ . Further setting  $m = 1$  for convenience, we simplify Equation (12) to a function of a single real variable  $x := \|\mathbf{x}\|$ :

$$G_d(x) = \frac{1}{(2\pi)^{d/2}} x^{1-d/2} K_{1-d/2}(x). \tag{13}$$

The Bessel functions  $K_{1-d/2}$  of lowest integer order arise when  $d = 2$  and  $d = 4$ , respectively. When  $d = 2$  we have that the Feynman propagator  $G_2(x)$  is proportional to the function  $K_0(x)$ :

$$G_2(x) = \frac{1}{2\pi} K_0(x). \tag{14}$$

When  $d = 4$ , the identity  $K_{-1}(x) = K_1(x)$  yields

$$G_4(x) = \frac{1}{(2\pi)^2} \frac{K_1(x)}{x}. \tag{15}$$

The new identity (1) is thus reminiscent of the self-reproducing property of integral kernels (such as the propagators  $G_2$  and  $G_4$ ), yet it is definitely different, because it involves propagation in *different* dimensions. In fact we speculate that it might be related to propagation phenomena in an (eventual) theory of quantum gravity. In order to see how this might come about, let us recall the following facts [10,11].

The propagators (14) and (15) have been obtained under the assumption that spacetime remains a continuum all the way down to the shortest classical length, namely zero length. Now there are good theoretical reasons to believe that the continuous nature of spacetime breaks down when one reaches lengths of the order of the Planck scale  $L_P = \sqrt{\hbar G/c^3}$ . Here  $\hbar$  is Planck’s quantum of action,  $G$  is Newton’s gravitational constant, and  $c$  is the speed of light. The numerical value of  $L_P$  is roughly  $10^{-33}$  cm, well beyond what current technology can probe. All this notwithstanding, it has been convincingly argued [10,11] that an effective, or low-energy, manifestation of quantum gravity effects on the propagator functions  $G_d(x)$  can be captured by the simple replacement

$$x = \sqrt{x^2} \longrightarrow \sqrt{x^2 + L_P^2} \tag{16}$$

within the argument of  $G_d$ . In other words, a true propagator that encodes overall quantum-gravity effects will no longer be given by  $G_d(x) = G_d(\sqrt{x^2})$ , but instead by  $G_d(\sqrt{x^2 + L_P^2})$ . Thus even after setting  $x = 0$ , spacetime points will continue to be separated by a nonzero length  $L_P$ . Among other desirable properties, this has the welcome effect that the singularities present in the propagators (14) and (15) as  $x \rightarrow 0$  are smoothed out. For the sake of simplicity let us assume choosing units such that  $L_P = 1$ . Then the *quantum-gravity corrected propagator functions* (14) and (15) become

$$G_2^{(QG)}(x) = \frac{1}{2\pi} K_0(\sqrt{x^2 + 1}), \quad G_4^{(QG)}(x) = \frac{1}{(2\pi)^2} \frac{K_1(\sqrt{x^2 + 1})}{\sqrt{x^2 + 1}}. \tag{17}$$

The above corrected propagators duly reduce to their counterparts (14) and (15) for large  $x$ , while no longer exhibiting a singular behaviour at the origin. This correction is denoted by the superindex QG, standing for *quantum gravity*. As a rule one generally expects quantum effects to mollify (at least some of) the singular behaviour that classical theories may exhibit.

Altogether, substituting Equations (14) and (17) into Equation (1) and setting  $y = 1 = L_P$  in the latter, the identity we have just proved reads

$$G_2^{(QG)}(x) = 4\pi \int_{-\infty}^{\infty} G_2(|\xi - x|) G_4^{(QG)}(\xi) d\xi \tag{18}$$

in terms of scalar propagation functions in Euclidean space. Although reminiscent of the self-reproducing property of Feynman propagators, it is intriguing for several reasons. First, it involves different dimensions (2 and 4). Second, it links propagators in the *presence* of quantum-gravity effects to propagators in the *absence* of quantum-gravity effects. Last but not least, it requires that  $y$  be nonvanishing, so it suggests setting  $y = L_P$ —a key feature of quantum-gravity theories is the existence of a quantum of length.

One possible reading of Equation (18) is the following: starting from the 2-dimensional, quantum-gravity free propagator  $G_2$ , it suffices to compute its convolution with the 4-dimensional, quantum-gravity corrected propagator  $G_4^{(QG)}$ , in order to obtain the quantum-gravity corrected propagator  $G_2^{(QG)}$  in two dimensions. We hope to report on this issue in the future [12].

**Author Contributions:** All authors contributed equally.

**Funding:** This research was supported by grant no. RTI2018-102256-B-I00 (Spain).

**Acknowledgments:** It is a pleasure to thank Felix Finster for technical discussions.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Abramowitz, M.; Stegun, I. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*; Dover: New York, NY, USA, 1974.
2. Gradshteyn, I.; Ryzhik, I. *Table of Integrals, Series and Products*; Elsevier: Burlington, MA, USA, 2007.
3. Watson, G. *A Treatise on the Theory of Bessel Functions*; Cambridge University Press: Cambridge, MA, USA, 1952.
4. Bandos, T.; Montero, A.; Fernández, E.; González-Santander, J.L.; Isidro, J.M.; Pérez, J.; de Córdoba, P.F.; Urchueguía, J. Finite Line-Source Model for Borehole Heat Exchangers: Effect of Vertical Temperature Variations. *Geothermics* **2009**, *38*, 263. [[CrossRef](#)]
5. González-Santander, J.L.; Castañeda-Porras, P.; Ratis, Y.L.; Isidro, J.M.; de Córdoba, P.F. Calculation of Some Integrals Arising in Heat Transfer in Grinding. *Math. Prob. Eng.* **2010**, *2010*, 535801.
6. González-Santander, J.L.; Placeres, J.M.V.; Isidro, J.M. Exact Solution for the Time-Dependent Temperature Field in Dry Grinding: Application to Segmental Wheels. *Math. Prob. Eng.* **2011**, *2011*, 927876. [[CrossRef](#)]
7. Skuratov, D.; Ratis, Y.L.; Selezneva, I.; Pérez, J.; de Córdoba, P.F.; Urchueguía, J. Mathematical Modelling and Analytical Solution for Workpiece Temperature in Grinding. *Appl. Math. Model.* **2007**, *31*, 1039. [[CrossRef](#)]
8. Carslaw, H.; Jaeger, J. *Conduction of Heat in Solids*; Oxford University Press: Oxford, UK, 1959.
9. Bogoliubov, N.; Shirkov, D. *Introduction to the Theory of Quantized Fields*; Wiley: New York, NY, USA, 1980.
10. Padmanabhan, T. Duality and Zero-Point Length of Spacetime. *Phys. Rev. Lett.* **1997**, *78*, 1854. [[CrossRef](#)]
11. Padmanabhan, T. Hypothesis of Path Integral Duality. I. Quantum Gravitational Corrections to the Propagator. *Phys. Rev.* **1998**, *D57*, 6206. [[CrossRef](#)]
12. Finster, F.; Isidro, J.M. Universidad Politécnica de Valencia, Valencia, Spain. Unpublished work, 2019.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Document downloaded from:

<http://hdl.handle.net/10251/37840>

This paper must be cited as:

Reyes, R.; Gamermann ., D.; Montagud Aquino, A.; Fuente Herraiz, D.; Triana, J.; Urchueguía Schölzel, JF.; Fernández De Córdoba Castellá, PJ. (2012). Automation on the Generation of Genome-Scale Metabolic Models. *Journal of Computational Biology*. 19(12):1295-1306. doi:10.1089/cmb.2012.0183.



The final publication is available at

<http://dx.doi.org/10.1089/cmb.2012.0183>

Copyright Mary Ann Liebert

# Automation on the generation of genome scale metabolic models

R. Reyes<sup>a</sup>, D. Gamermann<sup>\*b,c</sup>, A. Montagud<sup>c</sup>, D. Fuente<sup>c</sup>, J. Triana<sup>a</sup>,  
J. F. Urchueguía<sup>c</sup>, and P. Fernández de Córdoba<sup>c</sup>

<sup>a</sup>Universidad Pinar del Río “Hermanos Saíz Montes de Oca”, , Martí  
270, 20100, Pinar del Río, Cuba.

<sup>b</sup>Cátedra Energesis de Tecnología Interdisciplinar, Universidad  
Católica de Valencia San Vicente Mártir, , Guillem de Castro 94,  
E-46003, Valencia, Spain.

<sup>c</sup>Instituto Universitario de Matemática Pura y Aplicada, Universidad  
Politécnica de Valencia,, Camino de Vera 14, 46022 Valencia, Spain.

June 5, 2012

## Abstract

**Background:** Nowadays, the reconstruction of genome scale metabolic models is a non-automatized and interactive process based on decision taking. This lengthy process usually requires a full year of one person’s work in order to satisfactory collect, analyze and validate the list of all metabolic reactions present in a specific organism. In order to write this list, one manually has to go through a huge amount of genomic, metabolomic and physiological information. Currently, there is no optimal algorithm that allows one to automatically go through all this information and generate the models taking into account probabilistic criteria of unicity and completeness that a biologist would consider.

**Results:** This work presents the automation of a methodology for the reconstruction of genome scale metabolic models for any organism. The methodology that follows is the automatized version of the steps implemented manually for the reconstruction of the genome scale metabolic model of a photosynthetic organism, *Synechocystis sp.*

---

\*daniel.gamermann@ucv.es

*PCC6803*. The steps for the reconstruction are implemented in a computational platform (COPABI) that generates the models from the probabilistic algorithms that have been developed.

**Conclusions:** For validation of the developed algorithm robustness, the metabolic models of several organisms generated by the platform have been studied together with published models that have been manually curated. Network properties of the models like connectivity and average shortest mean path of the different models have been compared and analyzed.

*Keywords:* Genome-scale metabolic models, Networks, Connectivity

## 1 Introduction

Since the second half of the twentieth century, the development of molecular biology has allowed fast advances in the understanding of the functions and working principles of cells and unicellular organisms at molecular level. In particular, high throughput experimental techniques of sequencing and analysis of genomic and proteomic information has given birth to rich web based databases of biological information on thousands of organisms, from prokaryotic bacteria to complex organisms like birds and mammals.

One of the new research fields that emerges from this panorama is system biology [1]: the bottom-up approach in order to quantitatively explain the properties of biological systems from the modeling and simulation of the interactions and characteristics of its macromolecular components. From these systematic studies based on mathematical modeling and computational simulations, yet a new discipline appears: synthetic biology, which is focused on the design and construction of “*a la carte*” new biological entities with new biological functions [2]. This new field of biotechnology lies in the limit between biology and engineering and aims at the partial design of modified organisms for different technological applications. For this purpose rational design principles of engineering must be combined with the available biological information and biotechnological techniques. The huge amount of biological information and the complexity of the engineering principles and analysis tools make it evident the need for a good computational platform that aids in the data mining and design development for artificial biological systems.

One of the corner stones of systems biology is the reconstruction of genome scale metabolic models. This means to gather information on all enzymatic reactions that take place in an organism based on genetic infor-



mation available for this organism’s genome. The current state of art of this process requires the effort of a specialist during a long term period (usually one year) in order to collect the available information from many different databases and the literature. Currently there are few software applications specially designed to help in this specific task. Moreover, different available software applications are too specific for a few determined tasks that do not embrace all steps in the whole process, leaving huge gaps that must be filled manually by the researcher.

The genome-scale metabolic reconstruction is the starting point of many different researches and applications, like the determination of the metabolic capacities or the determination of protocols for an optimal growing strategy for some organism and in particular, the search for potential sites for metabolic engineering [4]. The aim of metabolic engineering is the modification and/or introduction of biochemical reactions with technologies like recombining DNA, in order to optimize the production of some metabolites of interest, to redirect metabolic fluxes to new pathways or even to extend the metabolic capacities of an organism for the production of new metabolites. The accomplishment of these objectives for a specific organism depends on a good reconstruction of its metabolism, from which one can study the structure of the metabolic network and the consequences of adding or deleting specific genes [5].

In this context, several projects have been developed for the reconstruction of genome scale metabolic models with different ends, like the production of fuel from cyanobacteria [6] or yeast. Other examples are the genome-scale metabolic reconstruction of the *Burkholderia cenocepacia* J2315 [7] for research of treatment in patients suffering from cystic fibrosis; the *Rhodobacter sphaeroides* [8], capable of producing hydrogen, polyhydroxybutyrate and other biofuels; the *Clostridium beijerinckii* [9] capable of producing butanol.

Genome-scale metabolic models result from the integration of genomic, proteomic and metabolomic information obtained at different experimental levels. The study of these theoretical reconstructions of cell metabolism allows researchers to investigate emergent phenomena in biology, like the feedback control loops that regulate the organisms and other aspects of metabolic and genetic transcription and regulation. The metabolic models constitute an important tool for the comprehension of an organism, its metabolic capacities and prediction of the responses of it to different environmental and genetic changes. Moreover, they facilitate the development of strategies for the engineering of metabolic systems focused in an improvement of the metabolic efficiency.

The process in order to generate the models consists in a first step to

collect all available information on the metabolome of a species, as well as all genes that code for the different enzymes that catalyze each one of the metabolic reactions that take place in cell metabolism. Other aspects that must be taken into account are the coenzymes and cofactors needed for the enzymatic catalysis, the stoichiometry and reversibility of the reactions, information on the biomass composition and metabolic regulation [3]. Among all possible applications of a metabolic model, there is the possibility of evaluating projects for production and optimization of a metabolite of interest [4]. If a model is satisfactorily constructed, it should allow a realistic simulation of the organism's metabolism, submitted to different environmental and genetic perturbations. This simulation would represent, with its natural restrictions, a virtual organism or an "*in silico*" cell in which one can apply different computational algorithms to explore possible flux distributions inside the cell subjected to different environmental conditions and genetic configurations [6]. For the analysis of metabolic models there are several computational tools and algorithms already developed [12]. Those include Flux Balance Analysis (FBA) [11, 12], Minimization of Metabolic Adjustments (MOMA) [13] and Metabolic Flux Analysis (MFA) [14, 15] among others.

Currently, researchers that work in the genome-scale metabolic reconstruction use different computational tools in order to accomplish different tasks. For example, in order to generate the list of biochemical reactions for a specific organism, the software Pathway Tools [16] is a common choice. This software allows the compilation of specific databases for proteins, enzymes and metabolites of a determined organism, and to obtain a rough draft for the network. Nevertheless, it might be said that this software obtains the list of reactions with no regard to associations between metabolic routes and the reactions, or to completeness and unicity criteria for such model. These drawbacks leave a considerable amount of work to be manually done by the biologist that has to consider the reversibility of all reactions, and the analysis of possible repeated reactions inside the model, as well as the inclusion of reactions studied for other organisms that complete specific metabolic pathways, but do not show up in the database due to gaps in the genomic annotation for the organism in hand. Other specific software might be found with different computational algorithms implemented, like the Optgene software that includes an evolutionary programming based method to rapidly identify gene deletion strategies for optimization of a desired phenotypic objective function [40].

The lengthy work of reconstructing a genome-scale metabolic network and analyzing it would be much faster and greatly simplified if one could

find all algorithms and computational tools needed in the same software or platform. The aim of this work is to present a platform developed by our research group in order to automatically generate genome-scale metabolic models. With respect to other software that might be found for this purpose, our platform has the advantage that it automatically takes into account the criteria for unicity of the biochemical reactions, and presents the possibility for the biologist to automatically complete gapped metabolic pathways based in probabilistic criteria and comparison of the same metabolic route in different organisms. Moreover, the platform produces the metabolic generated network in different outputs: either as a SBML file or directly as an OptGene file format that might be directly piped in other analysis software. For the generation of the OptGene file, the biologist is also given the possibility to choose the biomass composition among the metabolites appearing in the model and directly fulfill the restrictions for the flux analysis and balance.

The work is divided as follows: in the next section we explain the algorithm for the generation of genome-scale metabolic models implemented in a web based platform. The algorithm automatically obtains information from the KEGG database for a specific organism and constructs from it the list of reactions in its metabolism. Criteria of unicity and completeness are taken into account in order to cope with different enzymes catalyzing the same reactions or to fill in missing reactions. Section 3 presents an analysis done with the models generated by the platform. The generated OptGene files are used as inputs in different algorithms to study network properties of the reconstructed metabolic models and compare them with manually reconstructed models taken from the literature. In the end we present our conclusions, an overview and future perspectives.

## 2 Algorithm Description

The algorithm for automatically generating metabolic models comprises several steps: the information compilation from free access biological databases, following some interaction of the user with the platform in order to properly select the parameters for the probabilistic criteria and choices for the biomass components and restrictions, and finally application of unicity and completeness criteria and production of the output.

## 2.1 Obtaining the Biological Information

In the last years, applications of biotechnology in different areas of science and technology have considerably increased causing an exponential growth on the available information on different organisms about their genetics, regulation processes and metabolism. Such information, obtained by different techniques with growing efficiency, becomes part of huge databases, many of which are of free access. This information together with the vast published scientific works put in the hands of the researchers a rich ever-growing amount of data and information.

A start point for the genome-scale metabolic reconstruction is to obtain the relevant information about the organism for which the model is going to be generated, namely the list of reactions, genes, metabolites and enzymes present in the studied cell. This information is available from public free access databases like Biocyc [17], Kegg [18], Brenda [19], Uniprot [20], etc. Nevertheless, the lack of quality in some entries of the databases are an inconvenient that one must amend: false positives, false negatives, as well as objects wrongly annotated may pose obstacles in the efforts to compile a meaningful correct list of reactions [21]. As a consequence, the reconstruction must be done under strict control of all and each one of the reactions, the biomass equation must be based in constituent molecules and coherence and integrity of the network must be prerequisites for the generation of a quality and useful model [22].

The first objective of the project is to obtain the relevant rough information. For this aim we have chosen to use Kegg API web service offered by the Kegg database. This service allows one to access the Kegg system via SOAP/WSDL which offers valuable tools in order to access the available information in the mentioned database. These tools are designed for the search of cellular biochemical processes as well as to analyze the universe of genes and completed genomic sequences of thousands of organisms. The users might access Kegg API server by the SOAP technology through HTTP protocol.

The Simple Object Access Protocol (SOAP) defines how two objects in different processes may communicate trough exchange of XML (eXtensible Markup Language) data for a variety of bioinformatic applications. With this protocol, an application running in a machine anywhere in the world can use algorithms, data and resources stored in different servers [23]. Web Services Description Language (WSDL) is based in XML and allows one to have the description of a web service, specifying the abstract interface trough which a client may access the service and the details on how to use

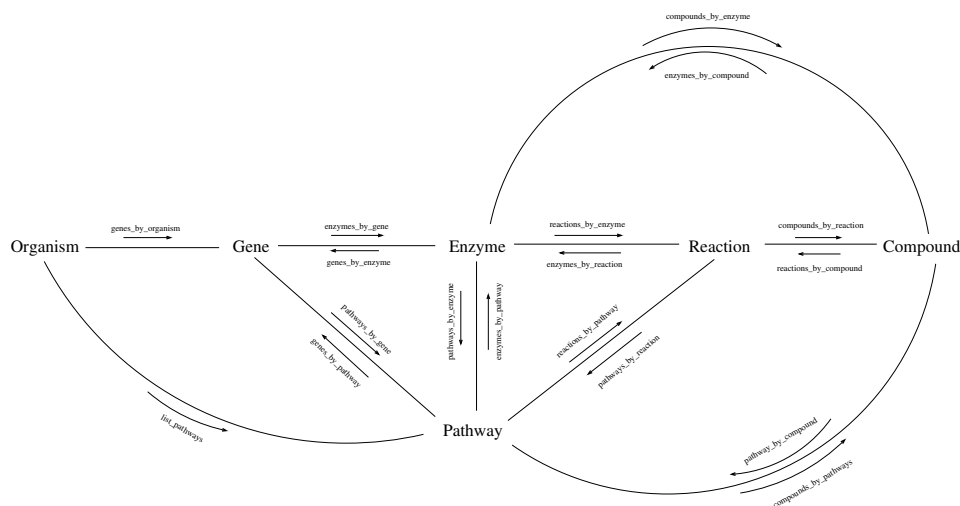


Figure 1: Scheme of KEGG information with the WSDL methods for obtaining the information. The same methods that appear referencing compounds also exist for glycans (which are not shown in the figure).

it [24].

These technologies allowed us to access Kegg API and to construct a service web client using Java [25] as programming language and Netbeans [26] 6.8 as integrated development environment. This way we obtained biological information from the definition of directional data model that relates biological elements for its storage in a database constructed in Postgres [27, 28, 29], taking into account the relationships among them and their importance in the reconstruction of the genome-scale metabolic models. The relevant biological elements identified at this point are the organism, gene, pathway, enzyme, reaction, compound (metabolites and glycans) and the references. In figure 1 we show an schema on how information is organized and related in the KEGG data base, and the methods to obtain the relations between the different elements.

## 2.2 Computational Platform for the Access of Biological Information (COPABI)

All the biological databases available in the internet offer a web application to show their information, as well as methods for exporting this information in different forms. The first aim of our project is to implement an application

that allows biologists to consult the relevant information from the database and to automatically generate metabolic models from it, where probabilist criteria for unicity and completeness are taken into account in order to generate more accurate models with greater quality in an efficient way. The platform resulting from the project is named COPABI from COmputational Platform for the Access of Biological Information.

Every web application requires for its publication a web server which is the responsible for waiting client requests and to answer them. In the case of COPABI the chosen web server application is Apache, an open source code that works under any platform that has become one of the best web servers in terms of efficiency, functionality and speed [30].

In order to select the tools for the implementation of COPABI, it has been taken into account that among the possible web applications there are two major groups: the languages that run from the side of the server, like PHP and the ones that run from the side of the client, like HTML [31], Javascript [32], etc. In the case of COPABI, PHP has been chosen for this is an open source and very popular programming language, especially in web development [33, 34]. As web application framework (WAF) we have used Codeigniter 1.7.3 [35]. The aim of this framework is to aid in the development of code, from a rich sample of libraries, a simple interface and a logical structure for accessing the libraries. Javascript has been used for interactivity of the pages and for validations and the completion of data in the search for biological information.

### 2.3 Formats for the Output

Different analysis tools in system biology make use of different file formats. Therefore, a good choice for the output file format will be determinant in the uses and utility of the generated models. Some examples of file formats used by different databases are: SBML [36] (System Biology Markup Language), FASTA [37] (Fast All), BLAST [38] (Basic Local Alignment Search Tool) and Kegg presents its information in a particular markup language called KGML [39].

The lack of a standard makes more difficult the use of different softwares that usually have specific formats for their inputs and output. Sometimes the user has to go through the lengthy work of converting the information from one format to the other. One of the most versatile formats is certainly SBML, since it is a description language based on XML that can be used to represent models of different biological processes like metabolic network, cellular signaling pathways, genetic regulation networks, among others. A

very useful file format specific for metabolic networks is the one following the input requisites for the OptGene software [40], also called BioOpt and used in BioMet toolbox ([www.sysbio.se/biomet](http://www.sysbio.se/biomet)).

For the COPABI platform these two output file formats have been chosen: the standard SBML level 2 version 1 and the OptGene file format, which can be readily used for flux balance analysis.

## 2.4 Construction of the Models

For the generation of metabolic models for any organism, the applied methodology follows the same steps used in the manual reconstruction of the first model of a photosynthetic organism, the *Synechocystis sp.* PCC6803 [6] also taking into account the probabilist criteria followed in this paper in order to deal with duplicated and missing reactions, which will be discussed later.

About the list of reactions in the reconstruction of the metabolic model, before applying the probabilist criteria one has to properly organize and identify the chemical reactions. Three issues should be noted here:

- First is the compilation of all chemical reactions of a particular pathway present in the organism. There are two kinds of reactions that one has to take into account: most reactions are catalyzed by enzymes and each enzyme has a code called EC (Enzyme Commission), each reaction in the model receives this number as its identifier. On the other hand, there are a few reactions in some pathways that are not catalyzed by enzymes, they are spontaneous reactions, instead. These spontaneous chemical reactions receive as their identifier the name “non-enzymatic” and a number after it indicating the order in which these reactions appeared.
- Next is the related to enzymes that can catalyze different reactions of the same type. This usually happens because different substrates have very similar structure and can couple to the enzyme which performs its catalytic activity. In these cases, next to the EC number of the enzyme, in the identifier comes an extra letter (a,b,c,...).
- In a last step the reversibility (or irreversibility) of each reaction in a pathway is checked from kgml files for every pathway present in an organism.

Example of output from *Synechocystis sp.* PCC6803 in OptGene format (the symbols # indicate comments and we use it for specifying the name of the metabolic pathway):

-REACTIONS  
# Glycolysis / Gluconeogenesis  
1.2.4.1a: Pyruvate + Thiamin diphosphate -> CO2 + 2-(alpha-Hydroxyethyl)thiamine diphosphate  
2.7.1.40a: ATP + Pyruvate -> ADP + Phosphoenolpyruvate  
6.2.1.1a: ATP + CoA + Acetate -> Diphosphate + AMP + Acetyl-CoA  
1.2.1.5a: H2O + NAD+ + Acetaldehyde <-> NADH + Acetate + H+  
1.1.1.2: NADP+ + Ethanol <-> NADPH + H+ + Acetaldehyde  
# Citrate cycle (TCA cycle)  
1.1.1.42a: Oxalosuccinate <-> CO2 + 2-Oxoglutarate  
1.1.1.37: NAD+ + (S)-Malate -> NADH + Oxaloacetate + H+  
2.3.3.1: CoA + Citrate <-> H2O + Acetyl-CoA + Oxaloacetate  
6.2.1.5a: ATP + CoA + Succinate <-> ADP + Orthophosphate + Succinyl-CoA

## 2.5 Probabilistic Criteria

Until this point, the reconstructed model will have exactly the same information stored in the database. The next steps are the implementation of automatic algorithms that will take into account the probabilistic criteria in order to complete missing gaps in some metabolic pathways (completeness) or to exclude duplicated reactions (unicity).

### 2.5.1 Unicity

For the unicity criteria the algorithm identifies reactions that appear more than once and identifies their enzymes. Repeated reactions must be eliminated, the criteria to choose which reaction is eliminated is the following: the enzyme that appears less frequently in the model is not eliminated. As an example we show the reactions:

1.2.1.12a: A + B <-> C + D  
1.2.1.12b: G + E <-> K + L  
1.2.1.12c: P + V -> Y  
1.1.1.1: G + E <-> K + L

In the example enzyme EC1.2.1.12 catalyzes 3 different reactions, one of which is the same as the reaction catalyzed by enzyme EC1.1.1.1. Following the algorithm criteria, reaction 1.2.1.12b is eliminated from the metabolic model.



### 2.5.2 Completeness

The addition of new reactions to a metabolic model is associated to the comparison between the information available about the metabolic reactions in a determined pathway in a specific organism with a general pathway (theoretical one) generated from the compilation of all metabolic reactions present in “all” organisms in nature. In many cases the genomes are not perfectly annotated and some genes are missing from the annotation. This comparison of pathways is a tool to help biologists identify these missing genes.

The reactions associated to gaps identified in a particular pathway are going to be added to the metabolic model if they satisfy the following criteria:

- The reactions have as final product a metabolite belonging to the biomass equation.
- The reactions present in the model in this pathway correspond to, at least, a determined percentage of the whole general pathway.

The COPABI interface allows the user to chose the metabolites that compose the biomass equation, as well as the percentage value used in the second criteria.

The reactions added to the model with these criteria are added to the end of the output after a comment (`# not pres ! ~~IMPORTANT, following X reactions not in sequence!!~~`) and their identifier is the correspondent EC number with the symbol “.” preceding it.

## 3 Results - Validation of the Models

For the validation of the metabolic models generated by the COPABI platform, we have analyzed the models generated for determined organisms with metabolic models manually curated taken from the literature. In a first step we analyze general properties of the generated models (number of metabolites, reactions, ...) and properties of the networks described by the metabolic model.

From the network point of view each metabolite of a model can be thought as a node and each reaction represents links between the metabolites in the lefthand-side with the metabolites on the righthand-side of the reaction equation. These links can be directed if one takes into account the

direction of each reaction and the reversibility of the reactions or undirected if one neglects this information.

In a first step of the analysis, our algorithms make an automatic debug of the metabolic models. This means look for bad reactions, those being reactions without substrates or products (some transport reactions in the sbml files taken from the literature present this issue), reactions where the same metabolite appears as substrate and product or reactions decoupled from the network, meaning that at least one substrate and one product of the reaction appear only in this reaction and nowhere else. All these bad reactions are excluded from the models before any calculation is done, because these bad reactions add up errors and uncertainties to the results.

Each metabolic model now represents a network and, in a first step, we have chosen to work with the directed version of it, meaning that the links connecting two metabolites have direction from the substrate to the product and in reversible reactions the pair of metabolites would have two links of opposite directions connecting them.

For each organism two versions of the metabolic model have been generated with two different values for the parameter appearing in the decision taking process for the completeness criteria, explained in the section before. In one model, the parameter is chosen to be 100% and in the other 10%, two possible extremes. Taking this parameter to be 100% means that only reactions for enzymes that are annotated in the genome of an organism will be written to the model. On the other hand if the parameter is taken to be 10% (a fare low value) means that if one out of ten reactions in a pathway are annotated in the genome, the generated model will have all reactions in this pathway. In the end, for each organism there are three models being studied, two automatically generated by the COPABI platform and one manually constructed taken from published works. The models taken from the literature correspond to the following organisms: the *Synechocystis sp PCC6803* [6], *Synechococcus elongatus PCC7942* [41], *Burkholderia cenocepacia J2315* [7], *Rhodobacter sphaeroides* [8], *Clostridium beijerinckii* [9], *Mycoplasma genitalium* [42], *Lactobacillus plantarum* [43], *Thermotoga maritima* [44] and *Yersinia pestis* [45].

In table 1 we show results for the general comparison of the models. Below we explain each column of the table separately:

- # Met. → The number of different metabolites (or compounds) found in the model.
- # Reac. → The number of reactions present in the model (after excluding bad reactions).

- % Rev  $\rightarrow$  The percentage of the reactions that are reversible.
- % Irr  $\rightarrow$  The percentage of the reactions that are irreversible.
- ASP  $\rightarrow$  The average shortest path. For each pair of metabolites in the model we have used Dijkstra's algorithm to calculate the shortest path connecting them in the network. For all pair of metabolites where the shortest path was found, the average value was calculated (pair of metabolites not connected by any path were left out).
- $\sigma_{\text{ASP}}$   $\rightarrow$  The standard deviation for the ASP calculation.
- $N_R$   $\rightarrow$  The number of pair of metabolites for which a path connecting them was found.
- $N_U$   $\rightarrow$  The number of pair of metabolites for which a path connecting them was not found. One should note that the network is directed, so, metabolites that have no link pointing in their direction can not be reached by any pair and are, therefore, either external metabolites that should be absorbed by the cell from the environment or are badly incorporated to the model.

As one can see from the table, although the networks usually have hundreds of different metabolites, two different metabolites are on average only three steps apart from each other. As a consequence, the whole network should very fast respond to changes in any of the metabolites' concentrations or to environmental perturbations. This closeness of the nodes in the network is known as small world behavior and is a consequence of a property of the network connectivity known as scale free distribution. Metabolic networks are known to follow a free scale distribution for node connectivity, meaning that the number of nodes ( $P$ ) with some number of connections ( $x$ ) follows a power law:  $P(x) \sim x^{-\gamma}$  where  $\gamma$  is usually a number between 2 and 3. From this law one concludes that there are very few nodes with a large number of connections (these are called hubs) and most of the nodes have very few connections.

For studying node connectivity, for each metabolite an algorithm counts in how many reactions it appears as a substrate (or product in reversible reactions). In figure 2 we show results for the three metabolic models of some organisms. It is clear from these plots the tendency of the distribution to follow a power law.

As one can see, all metabolic networks are very similar when their global network properties are studied. Therefore, in order to be able to differentiate

Table 1: General Comparison. The explanation of each column is in the text.

Org.	# Met.	# Reac.	% Rev.	% Irr.	ASP	$\sigma_{ASP}$	$N_R$	$N_U$
syn_lit	803	893	34.49	65.51	3.51	1.15	494446	150363
syn_10	707	718	37.74	62.26	3.2	0.90	355430	144419
syn_100	656	640	36.40	63.60	3.29	0.94	295093	135243
syf_lit	777	847	36.01	63.99	3.55	1.19	475612	128117
syf_10	711	705	37.02	62.98	3.19	0.88	356066	149455
syf_100	655	622	35.05	64.95	3.32	0.95	292390	136635
cbe_lit	732	856	27.22	72.78	3.05	0.82	409910	125914
cbe_10	752	808	40.22	59.78	3.21	0.88	412228	153276
cbe_100	693	733	38.2	61.8	3.33	0.97	335276	144973
tma_lit	583	612	41.67	58.33	3.19	0.96	242290	97599
tma_10	566	614	46.09	53.91	3.06	0.83	250504	69852
tma_100	489	517	44.1	55.9	3.24	0.91	183170	55951
bcj_lit	792	847	27.63	72.37	3.04	0.83	523487	103777
bcj_10	955	1018	37.03	62.97	3.25	0.88	632355	279670
bcj_100	907	948	36.29	63.71	3.32	0.92	564967	257682
mge_lit	342	262	40.08	59.92	3.00	0.99	83279	33685
mge_10	268	254	48.82	51.18	2.89	0.82	54311	17513
mge_100	116	104	55.77	44.23	3.42	1.25	11543	1913
eco_lit	1034	1435	14.91	85.09	3.19	0.90	912770	156386
eco_10	888	1017	39.53	60.47	3.19	0.84	576962	211582
eco_100	846	968	38.95	61.05	3.26	0.87	525007	190709
lpl_lit	513	526	31.37	68.63	2.97	0.82	221786	41383
lpl_10	566	595	41.85	58.15	3.14	0.83	233640	86716
lpl_100	492	512	41.8	58.2	3.23	0.88	173827	68237
rsp_lit	788	863	64.31	35.69	2.74	0.68	593663	27281
rsp_10	869	934	41.65	58.35	3.16	0.82	543333	211828
rsp_100	827	873	40.21	59.79	3.24	0.87	485209	198720
ypk_lit	817	948	29.85	70.15	3.04	0.85	339398	142238
ypk_10	838	945	39.47	60.53	3.18	0.84	520075	182169
ypk_100	779	891	39.62	60.38	3.25	0.89	444404	162437

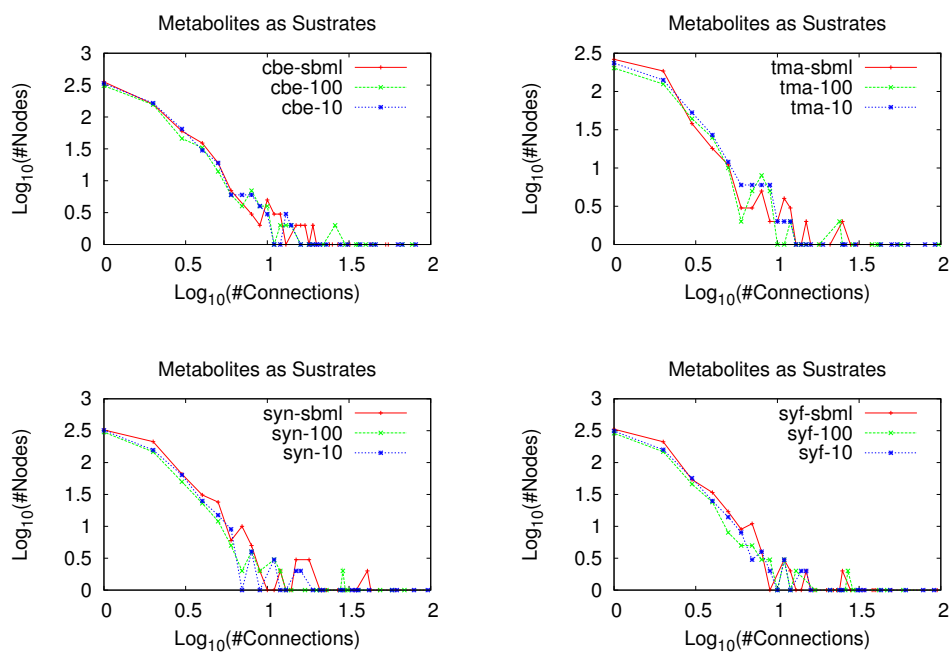


Figure 2: Connectivity distribution for the three metabolic models in different organisms.

the metabolic network of one organism from the network of another different organism, one has to look into the details of the networks, namely specific metabolites and hubs that are particular for each organism.

In order to find these differences we are going to define a similarity parameter in order to measure the degree of similarity of two metabolic networks. Two criteria are taken into account in the definition of this parameter, first the metabolites present in each metabolic network and the degree (number of connections) of each metabolite to all others. Since important metabolites for one organism, might be different from the essential metabolites in another one, we would like to take into account in the comparison also the identity of the metabolites in the different connections and not only the number of connections. This step is rather tricky because the metabolites names used in the metabolic models taken from literature do not follow any standard and the authors of each model chose different abbreviations and names for each compound. For some models, however, the authors have also made available the association of each compound name used in their models with a Kegg identifier. For these models we were able to construct an algorithm that translated the metabolic model to the same standard metabolite names used by kegg and therefore we were able to also compare the metabolite identities.

### 3.1 Network Comparison

Our goal here is to define a distance between two networks. We call it a distance in the sense that the bigger the value of this number, the more different the two networks will be, while the smaller the number is, the closer the networks are.

Given two metabolic networks, each one has a set of metabolite (let's call the set in each network by set  $A$  and set  $B$ ). Among all metabolites in the two networks, there are three different set: metabolites particular to network  $A$ , metabolites particular to network  $B$  and metabolites common to the two networks:

$$A \cup B = \underbrace{(A \cap \bar{B})}_{\text{only in A}} \cup \underbrace{(A \cap B)}_{\text{Common}} \cup \underbrace{(\bar{A} \cap B)}_{\text{Only in B}} \quad (1)$$

Now, let's consider the connections of the metabolites: Each metabolite  $i$  has  $n_i$  connections in total and  $n_{\alpha i}$  connections only to metabolites in the set  $A \cap \bar{B}$ ,  $n_{\beta i}$  connections only to metabolites in the set  $B \cap \bar{A}$  and  $n_{\gamma i}$  connections only to metabolites in the set  $A \cap B$ .

Let's define the number of metabolites in each set and the total number of connections inside each set:

$$N_\alpha = \| A \cap \bar{B} \| \quad (2)$$

$$N_\beta = \| B \cap \bar{A} \| \quad (3)$$

$$N_\gamma = \| A \cap B \| \quad (4)$$

$$N_A = \sum_{i \in A \cap \bar{B}} n_i \quad (5)$$

$$N_B = \sum_{i \in B \cap \bar{A}} n_i \quad (6)$$

$$N_C = \sum_{i \in A \cap B} n_i \quad (7)$$

Here,  $\| C \|$  means the number of elements in the set  $C$ .

Now, for each set, let's sum the proportion of connections of each metabolite to metabolites inside the set, weighted by the inverse of the total number of connections and averaged for all metabolites.

$$p_{Ai} = \frac{n_{\alpha i}}{n_i}$$

$$\alpha = \frac{N_A}{N_\alpha} \sum_{i \in A \cap \bar{B}} \frac{1}{n_i} p_{Ai}$$

and analogously we define  $\beta$  and  $\gamma$  for the metabolites in the other two sets.

The distance between the two networks is defined as:

$$\text{dist} = \frac{\alpha + \beta}{2\gamma}$$

For an identical network,  $\alpha$  and  $\beta$  are zero, so  $\text{dist}=0$ . For two networks which have not a single metabolite in common,  $\gamma=0$  and so  $\text{dist}=\infty$ .

For validating our models we proceeded by doing this rough comparison, calculating the distance between each one of the automatically generated models to each model taken from the literature. In Table 2 we show results for the comparison among 5 models taken from the literature and the automatically generated ones, one can observe that for each organism, the biggest value for the comparison in each column is when comparing the models for the same organism.

Table 2: Distance comparison between automatically generated and manually curated models. In bold face are the largest number in each row, showing that the best comparison between models is when comparing models for the same species.

org	syn_lit	syf_lit	cbe_lit	rsp_lit	ypk_lit	tma_lit
mge_10	1.246	1.254	0.401	0.695	1.003	1.541
lpl_10	0.815	0.755	0.121	0.317	0.476	0.834
syn_10	<b>0.47</b>	0.527	0.183	0.248	0.626	1.065
syf_10	0.54	<b>0.496</b>	0.18	0.255	0.628	0.999
cbe_10	0.697	0.699	<b>0.076</b>	0.212	0.413	0.814
bcj_10	0.708	0.721	0.156	0.183	0.459	1.063
eco_10	0.748	0.799	0.13	0.204	0.387	0.959
tma_10	0.72	0.7	0.103	0.278	0.498	<b>0.636</b>
rsp_10	0.735	0.741	0.157	<b>0.138</b>	0.549	1.103
ypk_10	0.772	0.782	0.12	0.181	<b>0.324</b>	0.882

This is a rough comparison, since the identification of the metabolites' names is not perfect. Moreover, since the automatically generated models have not yet been used in flux balance analysis, there is no distinction between internal and external metabolites, and we used a version of these models where no biomass was defined. This introduces lots of errors and uncertainties in the comparison made, because the models from the literature do have these features defined and they appear, from the point of view of the comparing algorithm, as new and different metabolites for which there will be no counterpart in the automatically generated models. Despite these pitfalls, the features already contained in the models are enough to differentiate the different organisms when compared to models from the literature.

## 4 Discussion

We presented here the metabolic models automatically generated by an algorithm developed by our research group. This algorithm, implemented in the COPABI platform, is able to automatically download genomic, metabolomic and proteomic information from the kegg database and construct, from this information, a genome-scale metabolic model either in OptGene or xml (sbml) file formats. The tools presented here will be soon made available in



the web.

The metabolic models generated have been thoroughly analyzed by standard algorithms in order to calculate average shortest mean path between nodes in the network and connectivity distribution. Commonly observed features in these networks are small world behavior and scale free distribution for the nodes degree.

After a general analysis of the global properties of the networks, we proceeded to compare the automatically generated models with model manually curated we took from published works. For this task a distance parameter between metabolic networks has been defined. The comparison shows that the automatically generated models are consistent with models constructed manually found in the literature.

The reconstruction of genome-scale metabolic models is an important step in different areas of research related to systems biology. Nowadays, this is a lengthy and slow process that might take over a year to be completed for a single organism. The algorithm developed here will certainly speed up the process and help researches to gain access to reconstructed models for any organism in just a few days period.

## **acknowledgements**

The authors would like to thank Alberto Conejero for useful discussions. This work has been funded by MICINN TIN2009-12359 project ArtBioCom from the Spanish Ministerio de Educación y Ciencia.

## References

- [1] Snoep, J. L., F. Bruggeman, B. G. Olivier and H. V. Westerhoff. Towards building the silicon cell: A modular approach. 2006. *BioSystems* 83: 207216.
- [2] Heinemann M. y Panke S. Synthetic biology putting engineering into biology. 2006. *Systems biology* 22: 27902799.
- [3] Frster, J., I. Famili, P. Fu, B. O. Palsson and J. Nielsen. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. 2003. *Genome Res* 13: 244-53.
- [4] Oberhardt, M. A., B. O. Palsson and J. A. Papin. Applications of genome-scale metabolic reconstructions. 2009. *Molecular Systems Biology* 5: 320.
- [5] Edwards, J.S., Ibarra, R.U., and Palsson, B.. 2001. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* 19: 125-130.
- [6] Montagud, A., Navarro, E., F. de Córdoba, P., Urchuegua, J.F. and Raosaheb, K. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. 2010. *Systems biology* 1752-0509/4/156.
- [7] Fang, K., Zhao, H., Sun, Ch., Lam, C. M., Chang, S., Zhang, K., Panda, G., Godinho, M., Martins dos Santos, V. and Wang, J. Exploring the metabolic network of the epidemic pathogen *Burkholderia cenocepacia* J2315 via genome-scale reconstruction. 2011. *Systems biology* 1752-0509/5/83.
- [8] Imam, S. Y. (2011) iRsp1095: A genome scale reconstruction of the *Rhodobacter sphaeroides* metabolic network. *BMC Systems Biology* 5:116
- [9] Milne, C. E.-J.-S. (2011) Metabolic network reconstruction and genome scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052. *BMC Systems Biology*, 5:130
- [10] Patil KR, Akesson M and Nielsen J: Use of genome-scale microbial models for metabolic engineering. 2004. *Current Opinion in Biotechnology*, 15:64-9.

- [11] Varma A, Palsson BO: Metabolic capabilities of Escherichia coli: II. Optimal growth patterns. *J Theor Biol* 1993, 165:503522
- [12] Edwards JS, Ramakrishna R, Schilling CH, Palsson BO: Metabolic flux balance analysis. In *Metabolic engineering*. Edited by Lee SY and Pappoutsakis ET. New York: Marcel Dekker Inc.; 1999: 1357
- [13] Segre D, Vitkup D, Church G: Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A* 2002, 99:15112-7
- [14] Schilling CH, Schuster S, Palsson BO, Heinrich R. Metabolic pathway analysis: basic concepts and scientific applications in the post-genomic era. *Biotechnology Progress* 1999;15: 296303.
- [15] Varma A, Palson BO. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product excretion in wild-type Escherichia coli W3110. *Applied Environmental Microbiology* 1994;60:372431
- [16] Karp PD, Paley S, Romero P: The Pathway Tools software. *Bioinformatics* 2002, 18 Suppl 1:S225-32
- [17] Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahren D, Tsoka S, Darzentas N, Kunin V, Lopez-Bigas N: Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Research* 2005, 19:6083-89.
- [18] Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y: KEGG for linking, genomes to life and the environment. *Nucleic Acids Res* 2008, 36:480-484
- [19] Chang A, Scheer M, Grote A, Schomburg I, Schomburg D: BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucleic Acids Res* 2009, 37:Database issue 2009
- [20] The UniProt Consortium: The universal protein resource (UniProt). *Nucleic Acids Res* 2008, 36:190-195
- [21] Weise S, Grosse I, Klukas C, Koschtzki D, Scholz U, Schreiber F, Junker BH: Meta-All: a system for managing metabolic pathway information. *BMC Bioinformatics* 2006, 7:465








- [22] Feist AM, Herrgard MJ, Thiele I, Reed JL, Palsson BO: Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 2009, 7:129143.
- [23] James Snell, Doug Tidwell & Pavel Kulchenko. 2001. *Programming Web Services with SOAP*. O'Reilly Media 2001. Chapter 3
- [24] Erik Christensen, F.C., Greg Meredith, Sanjiva Weerawarana. *Web Services Description Language (WSDL) 1.1*. 15 March 2001 [cited 2009 September 25]; Available from: <http://www.w3.org/TR/wsdl>.
- [25] *Big Java 4th Edition for Java 7 and 8, International Student Version* by Cay S. Horstmann. Publisher: Wiley Edition ISBN 978-0-470-55309-1 February 2010
- [26] Myatt, Adam (February 21, 2008). *Pro Netbeans IDE 6 Rich Client Platform Edition (First ed.)*. Apress. pp. 491. ISBN 1-59059-895-4.
- [27] *PostgreSQL 9.0 High Performance* by Gregory Smith. Publisher: Packt Publishing ISBN 184951030X (ISBN13: 9781849510301) October 2010
- [28] Bourne, P. 2005. Will a biological database be different from a biological journal? *PLoS Comput Biol*, Vol. 1, Nr. 3 (August 2005), p. 179-181.
- [29] Altman RB, "Building successful biological databases", *Brief. Bioinformatics* 5 (1): 4-5, March 2004. PMID 15153301.
- [30] *Developing Web Services with Apache Axis2 (3rd edition)* by Kent Kallok Tong. Publisher: TipTec Development (July 21, 2008) ISBN-10: 9993792918 ISBN-13: 978-9993792918
- [31] *HTML, XHTML, and CSS, Sixth Edition: Visual QuickStart Guide, 6th Edition* by Elizabeth Castro. Publisher: Peachpit Press (16 August 2006). ISBN-10: 0-321-43084-0 ISBN-13: 978-0-321-43084-7
- [32] *JavaScript Patterns* by Stoyan Stefanov  
Publisher: O'Reilly Media; 1 edition (September 28, 2010) ISBN-10: 0596806752 ISBN-13: 9780596806750
- [33] *PHP 5 Power Programming* by Andi Gutmans, Stig Bakken, Derick Rethans.  
Publisher: Prentice Hall (Oct 27, 2004) ISBN-10: 0-13-147149-X ISBN-13: 978-0-13-147149-8

- [34] Advanced Php Programming: Developing Large Scale Web Applications With Php 5 by George Schlossnagle.  
Publisher: Sams (August 2007)  
ISBN-10: 0672329239 ISBN13: 9780672329234
- [35] CodeIgniter for Rapid PHP Application Development: Improve your PHP coding productivity with the free compact open-source MVC CodeIgniter framework! by David Upton.  
Publisher: Packt Publishing Ltd. (July 2007).  
ISBN: 978-1-847191-74-8
- [36] Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novre N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J and Wang J. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. 2003. PubMed, PMID: 12611808
- [37] Biological Sequence Format: Fasta Format, Stockholm Format, Fastq Format by Books Llc  
Publisher: General Books LLC, 2010. ISBN-10: 1158522533 ISBN-13: 9781158522538
- [38] Christiam Camacho, Thomas Madden, George Coulouris, Ning Ma, Tao Tao, and Richa Agarwala. BLAST Help. NCBI Help Manual Bethesda (MD): National Center for Biotechnology Information (US); 2008.
- [39] Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucl. Acids Res. 28, 27-30.
- [40] Patil et al. (2005) Evolutionary programming as a platform for in silico metabolic engineering. BMC Bioinformatics, 6:308
- [41] J. Triana, A. Montagud, D. Gamermann, J.F. Urchueguía and P. Fernandez de Córdoba. Genome-scale reconstruction and *in silico* analysis of the *Synechococcus elongatus PCC7942* metabolic network. In preparation.

- [42] Suthers PF, Dasika MS, Kumar VS, Denisov G, Glass JI, et al. (2009) A Genome-Scale Metabolic Reconstruction of *Mycoplasma genitalium*, iPS189. PLoS Comput Biol 5(2): e1000285. doi:10.1371/journal.pcbi.1000285
- [43] B. Teusink, A. Wiersma, D. Molenaar, C. Francke, W.M. de Vos, R.J. Siezen and E.J. Smid. Analysis of Growth of *Lactobacillus plantarum* WCFS1 on a Complex Medium Using a Genome-scale Metabolic Model. THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 52, pp. 4004140048 (2006)
- [44] Ying Zhang, Ines Thiele, Dana Weekes, Zhanwen Li, Lukasz Jaroszewski, Krzysztof Ginalski, Ashley M. Deacon, John Wooley, Scott A. Lesley, Ian A. Wilson, Bernhard Palsson, Andrei Osterman, and Adam Godzik. Three-dimensional Structural View of the Central Metabolic Network of *Thermotoga maritima*. Science. 2009 September 18; 325(5947): 15441549. doi:10.1126/science.1174671
- [45] Ali Navid and Eivind Almaas. Genome-scale reconstruction of the metabolic network in *Yersinia pestis*, strain 91001. Mol. BioSyst., 2009, 5, 368375

Article

# Analysis of the Influence Subjective Human Parameters in the Calculation of Thermal Comfort and Energy Consumption of Buildings

Roberto Robledo-Fava <sup>1,\*</sup>, Mónica C. Hernández-Luna <sup>1</sup>, Pedro Fernández-de-Córdoba <sup>1</sup>,  
Humberto Michinel <sup>2</sup>, Sonia Zaragoza <sup>3</sup> and A Castillo-Guzman <sup>4</sup>  
and Romeo Selvas-Aguilar <sup>4</sup>

<sup>1</sup> UMPA, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022 Valencia, Spain; moher@doctor.upv.es (M.C.H.-L.); pfernandez@mat.upv.es (P.F.-d.-C.)

<sup>2</sup> Applied Physics Department, Universidade de Vigo, As Lagoas s/n, 32004 Ourense, Spain; hmichinel@uvigo.es

<sup>3</sup> Escuela Politécnica Superior, Universidade da Coruña, Mendizábal s/n, 15403 Ferrol, Spain; szaragoza@udc.es

<sup>4</sup> Universidad Autónoma de Nuevo León, Facultad de Ciencias Físico-Matemáticas, Av. Universidad S/N, Ciudad Universitaria, San Nicolas de los Garza, Nuevo León 66455, Mexico; arturo.castillogz@uanl.edu.mx (A.C.-G.); romeo.selvasag@uanl.edu.mx (R.S.-A.)

\* Correspondence: rorobfa@doctor.upv.es; Tel.: +52-818-250-7386

Received: 11 March 2019; Accepted: 16 April 2019; Published: 23 April 2019

**Abstract:** In the present work, we analyze the influence of the designer's choice of values for the human metabolic index (met) and insulation by clothing (clo) that can be selected within the ISO 7730 for the calculation of the energy demand of buildings. To this aim, we first numerically modeled, using TRNSYS, two buildings in different countries and climatologies. Then, we consistently validated our simulations by predicting indoor temperatures and comparing them with measured data. After that, the energy demand of both buildings was obtained. Subsequently, the variability of the set-point temperature concerning the choice of clo and met, within limits prescribed in ISO 7730, was analyzed using a Monte Carlo method. This variability of the interior comfort conditions has been finally used in the numerical model previously validated, to calculate the changes in the energy demand of the two buildings. Therefore, this work demonstrated that the diversity of possibilities offered by ISO 7730 for the choice of clo and met results, depending on the values chosen by the designer, in significant differences in indoor comfort conditions, leading to non-negligible changes in the calculations of energy consumption, especially in the case of big buildings.

**Keywords:** Monte Carlo method; ISO 7730; TRNSYS

## 1. Introduction

The building envelope is the interface between the outdoor environment and the interior of buildings and its primary function is to act as a physical barrier to offer a comfortable place to develop different activities, in exchange for a certain demand for energy [1–6]. 40% of the primary energy consumption in the world and 17% of carbon dioxide emissions are caused by the building sector [7–10]. Thus, power consumption in buildings is one of the leading causes, among other adverse effects, of ozone layer depletion, global warming, and climate change [11,12]. Besides, residential buildings are often mentioned as one of the most profitable areas to decrease global power demand [13,14].

A rapid search on scientific databases shows that comfort in buildings is addressed by different areas, such as Engineering ( $\approx 39\%$ ) energy ( $\approx 14\%$ ), environmental ( $\approx 10\%$ ), computer ( $\approx 7.5\%$ ) and social sciences ( $\approx 7\%$ ). Most of the analysis emphasizes the fact that adequate levels of indoor comfort

(thermal, acoustic, lighting, and air quality) are essential to guarantee the health and welfare of the occupants and to prevent adverse problems such as disease, among others [15].

Some recent examples of the application of comfort analysis concepts in singular environments include a synagogue design [16], strategies for low-cost housing [17], the evaluation of a modern low-energy office [18] or materials for hip-protective pads [19]. Thermal comfort is thus one of the most important elements in the relationship with the estimation of indoor environment quality, and it has been defined in parallel by ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) Standard 55 and ISO (International Standard Organization) Standard 7730 as “*the condition of the mind in which satisfaction is expressed with the thermal environment*” [20,21].

Thus, comfort ranges will depend on the physical and human parameters of its occupants and therefore, should be established according to the specificity of each region or country [22]. Well established international standards from countries and regions such as USA [23] or Europe [24], restrict comfort intervals based on numerical indices, such as the predicted mean vote (PMV) and the predicted percentage dissatisfied (PPD) [21,24,25], both defined in the ISO 7730 [26]. In the present work, we precisely focus on ISO 7730, which is the mandatory standard in Europe. In this norm, the designer has some freedom, within a given range, to choose the adequate clo and met parameters, based on his/her observation of the occupants of the building under analysis [27]. Thus, as we will explain in detail below, we have analyzed with a Monte Carlo code, the possible variations in energy demand calculations, due to aleatory designer’s choices *within the range provided by ISO 7730*. In this way, we quantify the impact of the choice of met and clo on energy demand, according to the norm. We must stress that other norms such as EN 15251-2007 or ISO 17772-1-2017 have been used in studies related with occupants modeling [28]; however, they are not mandatory in many countries, such as Spain or Mexico, that we analyzed in the present work, where ISO 7730 is the standard.

*The above parameters are used to predict the thermal sensation as a function of the four classical thermal environmental parameters (air temperature, mean radiant temperature, air velocity, and humidity) and two subjective human parameters: activity and clothing; meaning activity the intensity of the physical tasks a person is developing (expressed through the metabolic rate index) and clothing standing for the total thermal resistance from the skin to the outer surface of the clothed body [26,29].* This comfort estimate is evidently more significant for buildings located in extreme weather regions, either cold, such as northern Europe, Canada or Asia [30,31] or warm such as Africa, South-East Asia or Latin America [32]. On the other hand, the use of HVAC (Heating, Ventilation and Air Conditioning) systems is strongly related with thermal comfort and energy consumption and it is well known that in countries with extreme weather conditions, HVAC systems may represent more than one half of the total power demand of a single dwelling [33,34].

Building Information Modeling (BIM)-based simulations are commonly used as an effective tool to analyze simultaneously energy consumption and thermal comfort in buildings [35]. Some building simulation programs such as Energy Plus, ESP-r or TRNSYS are software used to calculate comfort levels inside of buildings [36,37]. TRNSYS (TRaNsient SYStems simulation) is one of the most widely established programs, providing a flexible and graphical simulation tool, based on FORTRAN code, originally developed in 1975 by the University of Wisconsin [38].

Over the last few years, multiple energy modeling studies have been carried out, taking into account the level of insulation of clothing and the level of activity [39–41]. In these studies, the relationship between clo and met with energy consumption has been analyzed under different perspectives [42–44]. There have also been many papers relating the parameters of thermal comfort with the most widely used international standard as is the case of the ASHRAE 55 [45–47] or ISO 7730 [48,49].

A factor of considerable uncertainty is the modeling behavior of the occupant [50]. Adaptive thermal comfort, which relies on changing clothing levels to gain a wider comfortable temperature range has been also widely researched [51]. Clothing levels cannot be sensed electronically and therefore surveys and observations have been commonly used [41] and in recent years methodological advances have



been developed that make dynamic simulations of building-occupant systems more tractable, such as visualization systems [52] agent-based models [53] or adaptive building simulations [54] among others. However, a significant problem when simulating the performance of a building is still the precise determination of the complete requirements of the models of the behavior of the occupants [55,56].

While most of the available simulators are focused on the performance of electrical and thermal networks, TRNSYS can similarly well be used to model other dynamic systems such as traffic flow, or even biological processes [57,58]. Studies using TRNSYS are reported in the literature as comparing cooling strategies for an office in different European climates [59], thermal comfort in an indoor swimming pool [60], or operating performance in cooling mode of a nearly zero energy building in China [61].

In this paper, we use TRNSYS for analyzing two different buildings: one in Spain and another one in Mexico. The numerical simulations have been carried out by multi-zone analysis, including technical/physical properties of the buildings, focusing on the cold (winter) and warm (summer) conditions in these countries, as there is more influence on thermal comfort. The novelty of our approach mostly lies in the combination of such BIM-based computer simulations with a sensitivity analysis (SA), to determine which input parameters have the maximum impact on the output [29].

The SA can be done by different popular methods such as Morris, regression, variance-based, or Monte Carlo, among others [62,63]. In our case, the SA based on Monte Carlo method [49] offers the possibility of prioritizing and fixating inputs when considering multiple outputs, in comparison with other methods [64,65]. Several studies can be found in the literature with SA performed on the effects of technical and physical parameters on the energy consumption of buildings and simulation programs [49,66,67]. However, to our knowledge, less attention has been paid to the effect of subjective human parameters [51] and thus, this aspect will be our main focus for the present work. Therefore, as a final step, we will show the significant impact of the designer's subjective choices on the calculation of the energy demand of buildings, depending on the election of the amounts of activity or clothing, based on the environmental ergonomic standard ISO 7730. We consider that our SA of the above-mentioned parameters using the Monte Carlo method will help the designers to define the most appropriate parameters to be considered, depending on the aim and type of the numerical modeling performed.

## 2. Case Study

In the present work, we analyze two non-residential buildings localized in different continents (America and Europe), with particular characteristics in their structure, use, and surrounding climate. The first building is a recent one in Cambre, in north-western Spain (latitude 43°18'22.55" N, longitude 8°17'36.13" W) located at a height of 99 m above sea level. The second is a building situated in San Nicolás de los Garza, in north-eastern Mexico (latitude 27°32'02" N, longitude 99°58'33"), at a height of 239 m above sea level.

### 2.1. Climate Conditions

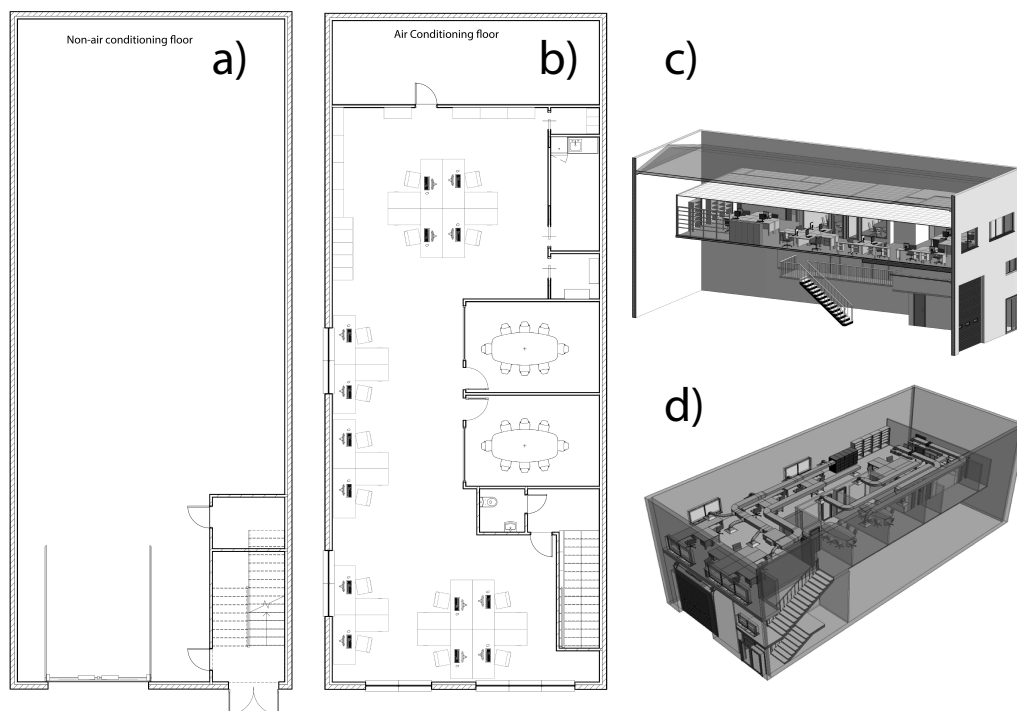
The Map of Köppen-Geiger Map [68] define different regions classifications in function of the central values of temperature and humidity, the Spanish location is classified in the Cfb region (C—temperate, f—without dry season, b—warm summer), where Oceanic Climate has a soft summer with a minimum rainfall and does not exceed of 22 °C on average in the warmest month, and at least four months averaging above 10 °C. No significant precipitation variation between seasons.

With respect to Mexico, according to the same classification, San Nicolás de los Garza can be considered to be BSh (BS—semi-arid, h—hot). A semi-arid (or steppe) region receives precipitation below potential evapotranspiration, but not extremely. Semi-arid climates tend to support short or scrubby vegetation, with large zones usually dominated by either grasses or shrubs. Also, the average annual temperature is higher than 18 °C. From the databases available in TRNSYS, the web platform

Equus [69], and Meteonorm [70], we obtain the values of different climatic variables such as air temperature, relative humidity, direct irradiance (solar) and diffuse horizontal irradiance.

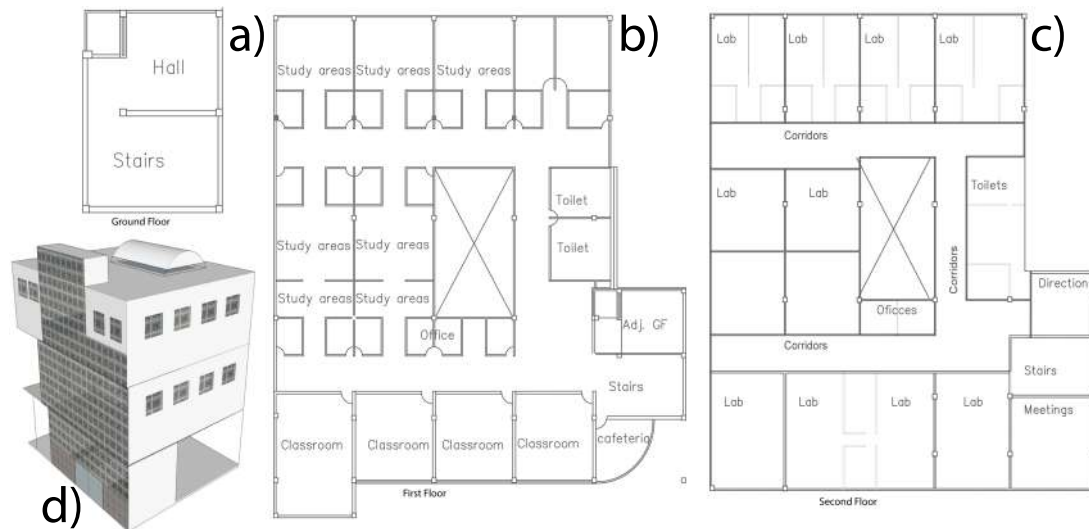
## 2.2. Building Structure And Use

The base of the first building is of rectangular shape with an area of 218.14 m<sup>2</sup> (see Figure 1) and its main wall faces southwest (216° N). The building is distributed on two floors (defined as ground and first), being the lower level used for storage and the upper zone for offices, meeting rooms, and kitchen, among others uses. This building is the headquarters of a private company specialized in air-conditioning systems. The working hours, considered from Monday to Friday throughout the year, are from 08:30 a.m. to 06:00 p.m.



**Figure 1.** Architectural blueprints of the first case study building (Spain): (a) ground floor; (b) first floor. Numerical model geometry: (c) Northeast view and (d) Southwest view.

For the second case, the plant of the building has an “L” shape, covering an area of 1214 m<sup>2</sup> in each of its three heights that we define as: ground floor (GF), first floor (FF) and second floor (SF). The GF corresponds to a parking lot without vertical enclosures and a reception area, while the other two floors are classrooms, meeting rooms, laboratories, offices, etc. as is shown in Figure 2. The building hosts a research center on physics and mathematics (CICFIM), which is located behind of the School of Physics and Mathematics of the Universidad Autónoma de Nuevo León (Mexico). Its principal access is on the facade facing east, next to the campus stadium, to the west. This building starts the activities at 7:00 a.m., ending at 9:00 p.m. from Monday to Friday. Saturday’s schedule is from 10:00 a.m. to 2:00 p.m.



**Figure 2.** Architectural blueprints of the second case study building (Mexico): (a) ground floor and (b) first floor (c) second floor and (d) Numerical model geometry: Northeast view.

The building in Spain is made with three different types of envelopes: the vertical facing walls, the floor, and the roof. In Table 1 we indicate that their respective thicknesses ( $t$ ) are 46.5 cm, 30.4 cm and 185 cm. The building in Mexico Table 2 has four types of envelopes: internal (9 cm) and external (35 cm) walls, as well as 10 cm-thick divisions. In this case, the same materials were used in the floor and the roof, with a total thickness of 40 cm.

In both Tables 1 and 2, the physical characteristics of each layer of material are specified, such as thermal transmittance ( $U$  – value), thermal conductivity ( $\kappa$ ), heat capacity ( $C_p$ ), density ( $\rho$ ) and resistance ( $R$ ) in specific cases.

**Table 1.** Thermal and structural properties of adopted constructive solutions for case 1 (Spain).


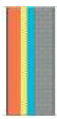


Wall Types	Structure	Layer	$t$ (cm)	$\kappa$ (kJ/mK)	$C_p$ (kJ/kgK)	$\rho$ (g/cm <sup>3</sup> )	$R$ (hm <sup>2</sup> K/kj)
External Walls		Cement roughcast	1.5	5.040	1.10	2.00	--
		Concrete block	20	1.764	1.10	1.20	--
		Air chamber	5.0	--	--	--	0.05
		Plasterboard	20	0.900	1.00	0.90	--
$U$ – value = 0.637 W/m <sup>2</sup> K							
Floor		Ceramic Brick	25	4.104	0.90	1.25	--
		Compressed concrete	5.0	1.00	1.10	1.50	--
		Air chamber	0.4	0.612	1.40	1.20	--
$U$ – value = 2.402 W/m <sup>2</sup> K							
Roof		Galvanized Metal	2.5	180.1	1.50	7.85	--
		Air chamber	180	--	--	--	0.05
		Perlite ceiling	2.5	0.187	1.50	0.12	--
$U$ – value = 0.934 W/m <sup>2</sup> K							

For both buildings, the properties of the windows are described in Table A1. In the first building (Spain) there are two types of zones, depending on whether there is air conditioning or not (see Table A2) in Appendix A. The second edifice (Mexico) is distributed in three partitions, because of the different uses and temperatures inside. Among its main characteristics are frame percentage, thermal transmittance ( $U$ ) and solar gain ( $G$  – Value).

In Spain, the HVAC system consists on an inverter heat pump ( $2 \times 1$ ) and a split air–air unit, with a total refrigeration capacity of 17.2 kW, a heat power of 19.5 kW and an installed power of 6.5 kW. The diffusion of air is carried out by ducts and impulse/return grids. For the building in Mexico which

a conventional technology, based on air-air heat pumps with 26 exterior units and indoor cassette units and a total cooling capacity of 349.4 kW. The heat output is 447.7 kW and the installed power is 166.4 kW.

**Table 2.** Thermal and structural properties of adopted constructive solutions for case 2 (Mexico).

Wall Types	Structure	Layer	t (cm)	$\kappa$ (kJ/mK)	$C_p$ (kJ/kgK)	$\rho$ (g/cm <sup>3</sup> )	R (hm <sup>2</sup> K/kJ)
Internal Walls		Plasterboard	1.5	0.900	1.00	0.90	--
		Mineral wool	6.0	0.130	1.03	0.05	--
		Plasterboard	1.5	0.900	1.00	0.90	--
$U - \text{value} = 0.512 \text{ W/m}^2\text{K}$							
External Walls		Plasterboard	1.0				
		Perforated brick	9.0	1.080	1.00	0.80	--
		Extruded polystyrene	5.0	2.736	1.00	1.60	--
		polystyrene	5.0	0.122	1.45	0.03	--
		Air Chamber	14	--	--	--	0.05
		Hollow brick	1.0	1.764	0.92	1.20	--
		Cement mortar	1.0	5.040	1.05	2.00	--
$U - \text{value} = 0.441 \text{ W/m}^2\text{K}$							
Floor & Roof		Concrete	3.0	4.140	1.00	1.80	--
		Cement mortar	1.0	5.040	1.05	2.00	--
		Extruded polystyrene	6.0	0.122	1.45	0.03	--
		Reinfor. concrete	15	8.280	1.00	2.30	--
		Pebble	15	2.916	0.92	1.70	--
$U - \text{value} = 0.450 \text{ W/m}^2\text{K}$							
Divisions		Aluminum	3.0	575	0.90	2.80	--
		Mineral wool	4.0	0.11	1.00	0.05	--
		Aluminum	3.0	575	0.90	2.80	--
$U - \text{value} = 3.296 \text{ W/m}^2\text{K}$							

### 3. Methodology

The main objective of this research is to show the consequences on the results of BIM-based simulations of subjective decisions made by the designers when using the environmental ergonomic standard ISO 7730 to calculate thermal comfort ranges and its impact on the energy demand of buildings, as well as its economic repercussion. *The thermal sensation experienced by a human being is mainly due to the overall thermal balance of the body, which depends on the physical activity (met) and clothing (clo) of the subject, as well as by the environmental parameters: air temperature, radiant temperature, air speed, and air humidity [71].* We must stress that the definition of the parameters of human clothing and activity is done subjectively by the designer and therefore we refer to these quantities as Subjective Human Parameters (SHPs). The use of biometric data from wearable devices has been recently proposed as an interesting option for metabolic rate estimation in thermal comfort analysis [72].

The PMV and PPD indices, according to Fanger's method [71] are commonly used to express warm and cool discomfort for the body as a whole, although thermal dissatisfaction may also be caused by unwanted heating or cooling of one particular part of the body (local discomfort). Therefore, when the previous factors are known, the thermal sensation for the occupants can be predicted, using a seven-point scale ranging from cold (−3) to hot (3), being 0 the value for “neutral sensation” [73]. The PMV may be computed from tables or may be measured by an instrument [74]. The methodology proposed in this paper consists of obtaining first the thermal performances of the buildings and then performing the SA for the SHPs used in the thermal comfort calculation, with the purpose of identifying which quantity has more influence on the BIM-based simulation outputs. Figure 3 shows a scheme of the procedure applied, consisting of the following steps, for both cases studied:

1. Real-time tracking of the main variables using Equis web platform.
2. TRNSYS modeling, getting the indoor temperature.
3. Verification of the thermal simulation with measured temperature from Equis database.
4. Mathematical estimation of PMV and PPD indices, according to ISO 7730.

5. SA (Monte Carlo) to determine which parameter (activity and clothing) values yield significant variation on PMV and PPD outputs.
6. New set-point temperatures are established, depending on the changes in the thermal categories in the previous step.
7. TRNSYS recalculation of the energy demand for each set-point temperature from the Monte Carlo method.
8. Evaluation of the variations in the energy demand for the different options.

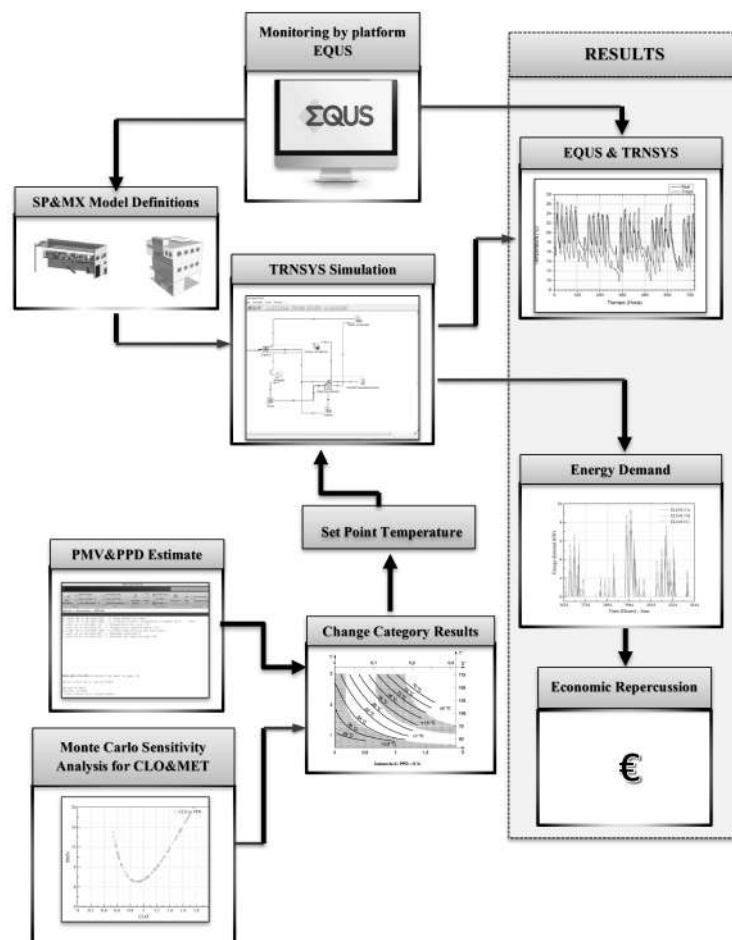


Figure 3. Scheme of the methodology followed.

### 3.1. Monitoring by Web Platform

A web platform for energy management of buildings named “Equis”, was developed with the purpose of control, maintenance and monitoring of energy consumptions and environmental variables as temperature, humidity and CO<sub>2</sub> levels. This tool also includes the possibility of remotely controlling any device by SCADA (Supervisory Control and Data Acquisition) system, ensuring the quality 6 of the statistical results. Equis collects in its database the information sent by the different sensors, allowing us to know the variation of the physical variables remotely, as access to these data can be obtained with any computer or other devices such as smartphones, which are connected to the Internet, making it very simple the treatment of the data, which can be done in the same device accessing to the values, in real time [69].

### 3.2. Trnsys Modeling of The Buildings

Through dynamic simulation tools, such as TRNSYS, we can define and evaluate the comfort and energy demand of the two buildings accurately; in both cases, we work with multi-zone building

models. Notice that TRNSYS is integrated into two parts: TRNBuild, which reads and processes an input file, solving and plotting the system; and TRNSYS Studio, which is a library of components that can be used to model different systems. An advantage of using TRNSYS is the possibility to convert complex problems as multi-zone building models into smaller components, to allow the designer to establish the connections between the building and numerous other subsystems/components in the simulation environment [75].

To verify the energy demands of the structures, after analyzing the sensitivity with the Monte Carlo method of the set-point temperatures of the air-conditioning systems with the met and clo parameters, it has opted for its simulation in the program TRNSYS. To simulate the building, both in Spain and in Mexico, all the necessary inputs, dimensions, location, construction characteristics, etc. have been introduced into the program. This simulation will be verified later comparing the results that the TRNSYS throws in certain indoor temperature conditions, with the actual measurements of both the environmental conditions and the interior temperatures. Once the model is validated, we can calculate the variation of the energy demand in both buildings as indicated by the Monte Carlo SA to the parameters studied in ISO 7730.

Entering the buildings models in TRNSYS was done in two steps: first, in TRNBuild we implemented the different zones and all its elements as location, orientations, walls, layers of materials, gains, ventilation, infiltration, schedules, and comfort in TRNBuild. The second step corresponds to include in TRNSYS models as weather, inputs such as the building previously described and outputs that will show through graphics the thermal performance and energy demands of the buildings. Thus, first, we obtained the database of the buildings, which includes information such as orientations of the facades, materials for the walls and windows, the energy demand of the installed equipment, the working hours, the occupation, etc. The process of the definition of the buildings in TRNBuild includes the following information:

- Location and orientations. This information was described in Section 2.2
- Structure of the building: the materials detailed in Tables 1, 2 and A1 were assigned from the software database.
- The zones described in Table A2, were created as WAC and AC for the first case, GF-FF, SF, and SL for the second case in the TRNBuild Manager.
- In Table A3 we show the different types of schedules created for each building, we set a value of 1 in the periods where there is activity in the building, in any other case, including the weekend, the value is 0.
- In Table A4 we summarize the infiltrations for summer and winter in both cases. This information was calculated based on ASHRAE Standard 55.
- For calculating the ventilation we use standard EN 13779, ASHRAE 62 R. It was estimated a value for each zone which is shown in Table A2.
- The internal gains of the buildings are shown in Tables 3 and 4.
- The values of clothing factor, metabolic rate, relative air velocity were introduced in TRNSYS based on the ISO Standard 7730.

For estimation of comfort, it is important to note that thermal sensations are different between persons sharing the same environment. Perceptions of human beings about thermal comfort are very different because even people who stay in similar places have different opinions, because there exist many elements that impact the perception of human beings.

Thermal comfort is the main goal and has a strong influence in the thermal behavior of a building, for this reason, the values of clothing factor, metabolic rate, relative air velocity were introduced in TRNSYS, based on the ISO Standard 7730. In Table 5, we show the numbers for both cases in summer and winter. We decided to set the same values of clothing and activity, because the characteristics of the occupants are very similar. It should be noted that these values are subjective decisions on the part of the designer. The analysis of the impact of these decisions is one of the main goals of this work.



**Table 3.** Input data of internal gains in the AC zone on Spain case.

Gains	Description	Schedule	Total Energy Rate (W)
Occupancy	15 Adults seated doing office work	Weekly 1	2250
Computers	16 PC with monitor	Weekly 1	2240
Artificial Lighting	37 fluorescent lamps in 218.12 m <sup>2</sup>	Weekly 1	3600
Other Gains	HVAC	Weekly 1	3332
	Coffee machine (10 cups)	Weekly 2	1500
	Copy Machine (office type)	Weekly 3	1060
	Microwave	Weekly 2	600
	Refrigerator	All time	322
	Plotter	Weekly 1	250
	TV	Weekly 1	90

**Table 4.** Input data of internal gains in the different zones on Mexico case.

Gains	Zones		Description	Schedule	Total Energy Rate (W)	
	GF-FF	SF			GF-FF	SF
Occupancy	63	51	Adults seated doing office work	Weekly A	9450	7650
Computers	54	50	PC with monitor	Weekly A	7650	7000
Artificial Lighting	37	9	Fluorescent lamps (64 W)	Weekly A	2304	576
	60	46	Fluorescent lamps (80 W)	Weekly A	4800	3680
	64	62	Fluorescent lamps (40 W)	Weekly A	2560	2480
Other Gains			HVAC	Weekly B	21,320	
	5	3	Coffee machine (10 cups)	Weekly A	1500	
	2	1	Water Cooler	Weekly B	1060	
	1	--	Microwave	Weekly A	600	
	2	1	Refrigerator	All time	322	
	10	12	Plotter	Weekly A	250	
	2	--	TV	Weekly A	90	

**Table 5.** Input data of comfort in the two buildings, according to ISO 7730.

Parameters	Description	Value
Clothing factor	Summer 1: panties, t-shirt, shorts, thin socks, sandals	0.30 clo/0.050 m <sup>2</sup> K/W
	Summer 2: underpants, t-shirt, light pants, thin socks, shoes	0.50 clo/0.080 m <sup>2</sup> K/W
	Summer 3: underwear, shirt, pants, socks, shoes	0.70 clo/0.110 m <sup>2</sup> K/W
	Winter: underwear, shirt, pants, thermal jacket, socks, shoes	1.20 clo/0.185 m <sup>2</sup> K/W
Metabolic Rate	Rest, seated	1.0 met/58 W/m <sup>2</sup>
	Seated, light work (office, home, school, laboratory)	1.2 met/70 W/m <sup>2</sup>
External Work	In general, the external work is around 0	0 met
Relative air velocity	The air velocity relative to the person	0.3 m/s

#### 4. Simulation of Building Models

In this section, it is presented the dynamic thermal simulation model in TRNSYS Studio of the building previously described in TRNBuild, for this purpose we use multi-zone building model "TYPE 56" that can be connected to a large number of other components, including weather data, HVAC systems, occupancy schedules, controllers, output functions, etc. Other models used were Type 62, Type 109-TMY2, and Type 65.

Under ISO Standard 7730, the evaluation of the levels of comfort was calculated using a mathematical MATLAB software package for the interactive numerical study of dynamical systems.

The change of PMV and PPD in base of its parameters such as metabolic rate, clothing insulation, and air temperature, is easier of visualizing in MATLAB than in TRNSYS because you can modify its parameters independently and realize studies more specific such SA.

### Monte Carlo Sensitivity Analysis

The principal aim of the SA using a Monte Carlo code is to obtain the variation of the PPD and PMV concerning the possible choices of the clo and met that are indicated in the annexes of the 7730 for thermal comfort calculations. In this way, by changing the values of clo and met according to what is reported in the tables of ISO 7730, we try to find out how the environmental category varies and, consequently, the set-point temperature. The SA is extensively suitable for associating input parameters with the overall building performance (for example, the temperature or demand for heating or cooling) [64]. The individual contribution from specific input can be determined by SA to prepare for future optimization of energy, climate, and economic performance [76,77]. SA methods can typically be classified to the local SA where input parameters are varied one at a time and the global SA where all inputs are changed simultaneously.

Sampling-based methods for uncertainty and SA have become very popular. The sampling-based Monte Carlo method is a procedure that performs a model repeated times with random samples generated from input distributions. It provides approximate solutions to both uncertainty and sensitivity analyses by making statistical modeling investigations [78]. The method can manage complex black box models irrespective of the linearity and continuity, and generate a probability distribution for each output depending on input distribution types [62].

The analysis under consideration can be represented by a function  $y(x)$  where:

$$\mathbf{y}(x) = [y_1(x), y_2(x), \dots, y_{n_Y}(x)] \quad (1)$$

and

$$\mathbf{x} = [x_1, x_2, \dots, x_{n_X}] \quad (2)$$

designates the outputs and inputs of the analysis, respectively. Successively, sensitivity in  $x$  results in a corresponding sensitivity in  $y(x)$ . Furthermore, the distributions

$$\mathbf{D} = [D_1, D_2, \dots, D_{n_X}] \quad (3)$$

must be well-defined to characterize the uncertainty associated with the elements of  $\mathbf{x}$ , where  $D_i$  is the distribution associated with  $x_i$  for  $i = 1, 2 \dots n_X$ .

The most common way for applying SA in building performance analysis consists on the following steps: determine input variations, create building energy models, run energy models, collect simulation results, run SA and presentation of SA results [64]. Determining the probability distributions of input parameters is the first step in any SA, which depends on the election of sensitivity method. In the present study, Monte Carlo SA is used, considering  $N$  input parameters defining the system:

$$O_i(P_1, P_2, \dots, P_N) \quad i = 1, 2, \dots, M. \quad (4)$$

Suppose that the parameter  $P_j$  can change within a range given by  $\Delta P_j$ . To analyze the influence of the parameter  $P_j$  in the objective  $O_i$ , we proceed first generating a uniform random number distribution  $r \in [0, 1]$ , thus:

$$P_j^1 = (P_j - \Delta P_j) + 2r\Delta P_j, \quad P_j^1 \in [P_j - \Delta P_j, P_j + \Delta P_j], \quad (5)$$

gives us the first random value  $P_j^1$ . With the value of  $N$  randomly generated parameters  $P_j^1 (j = 1, 2, \dots, N)$ , we obtain a value of  $O_i^1$ , which is the first evaluation of the objective  $O_i$ . We repeat this procedure up to the desired number of random evaluations. We can randomly change each one



of the parameters or we can apply random variations to one of the parameters keeping constant the value of the rest.

For this work, the objectives  $O_i$  are the PMV and PPD indices, which are a function of different parameters mentioned in the previous sections. To analyze the influence on the thermal performance of the buildings, the  $P_j$  parameters are CLO and MET. Depending on the admissible ranges for PMV and PPD, three kinds of comfort zones or categories of thermal requirements are defined by UNE-EN ISO 7730 as: category I (or class A) (PPD < 6%, i.e.,  $-0.2 < PMV < 0.2$ ), category II (or class B) (PPD < 10%, i.e.,  $-0.5 < PMV < 0.5$ ) and category III (or class C) (PPD < 15%, i.e.,  $-0.7 < PMV < 0.7$ ). The ranges of recommended air temperatures for different types of buildings depending on the previous categories are shown in Table 6 [26]. Based on the results of the SA, we can establish the set-point temperature for the cooling and heating seasons.

**Table 6.** The range of recommended air temperatures for offices and classrooms, according to ISO 7730.

Type of Building	Activity (W/m <sup>2</sup> )	Category	Temperature (°C)	
			Summer	Winter
Classrooms	70	A	24.5 ± 1.0	22.0 ± 1.0
Main Hall		B	24.5 ± 1.5	22.0 ± 2.0
Offices		C	24.5 ± 2.5	22.0 ± 3.0
Conferences room				

We use the SA based on the Monte Carlo method, randomly generating a uniform distribution for clo from 0 to 2 and met from 0.8 to 4, since in this range the PMV index lies between  $-2$  and  $+2$ , taking into account that this is the range suggested by the standard. A uniform distribution corresponds to the case of a random variable that can only have values between two extremes  $a$  and  $b$ , so that all the intervals of the same length (within  $(a, b)$ ) have the equal probability.

The above definition shows that the density function must be the same value for all points within the interval  $(a, b)$  (and zero outside the range), i.e.,

$$f_X(\chi) = \left\{ \begin{array}{l} \frac{1}{b-a} \quad \text{si } \chi \in (a,b) \\ 0 \quad \text{si } \chi \notin (a,b) \end{array} \right\} \quad (6)$$

The results of the SA on the comfort index (PMV and PPD) for the clothing and metabolic rate are shown in Figures 4 and 5, respectively. The thermal category changes according to PMV and PPD values, which in turn depend on the subjective decisions of clo and met. Figure 4 shows that for category A, the clo values vary from 0.46 to 0.66, for B of 0.34 to 0.46 clo and 0.66 to 0.87 clo. Category C is between 0.27 to 0.34 clo and 0.87 to 1.04 clo. For the case of met and PPD, in Figure 5 we show that category A lies between met values from 0.83 to 1.04. For the case of category B, it remains from 0.8 to 0.83 met and 1.04 to 1.16 clo. Finally, category C corresponds to the range from 1.16 to 1.31 met.

An attractive alternative in this type of SA, it is to use a normal distribution, which is one of the theoretical distributions best studied in statistical texts and more in practice, also called Gaussian distribution. Its importance is mainly due to the frequency with the variables other than natural and everyday phenomena, approximately, this distribution.

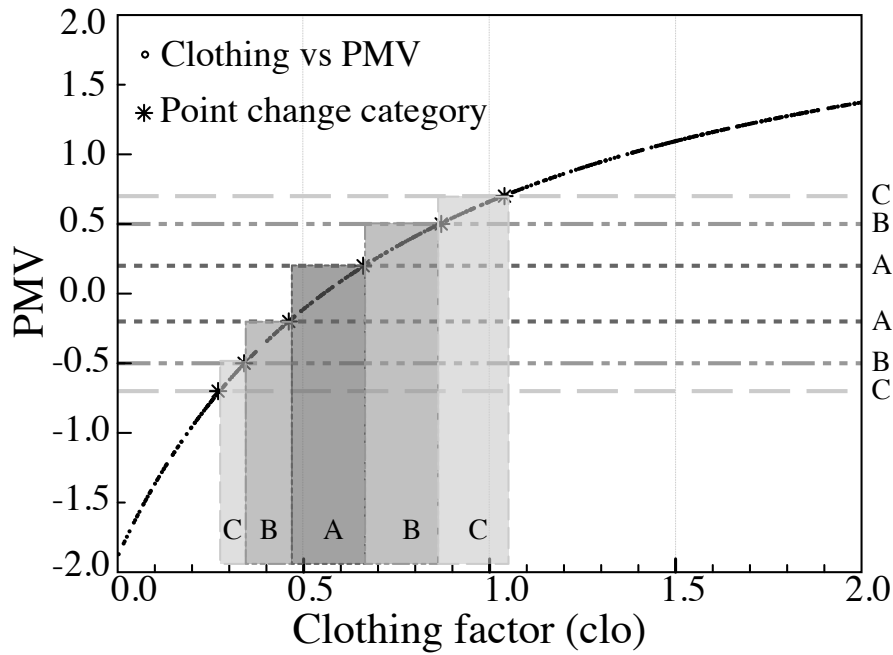


Figure 4. Change of level of thermal comfort in base of sensitivity analysis of clothing factor on PMV.

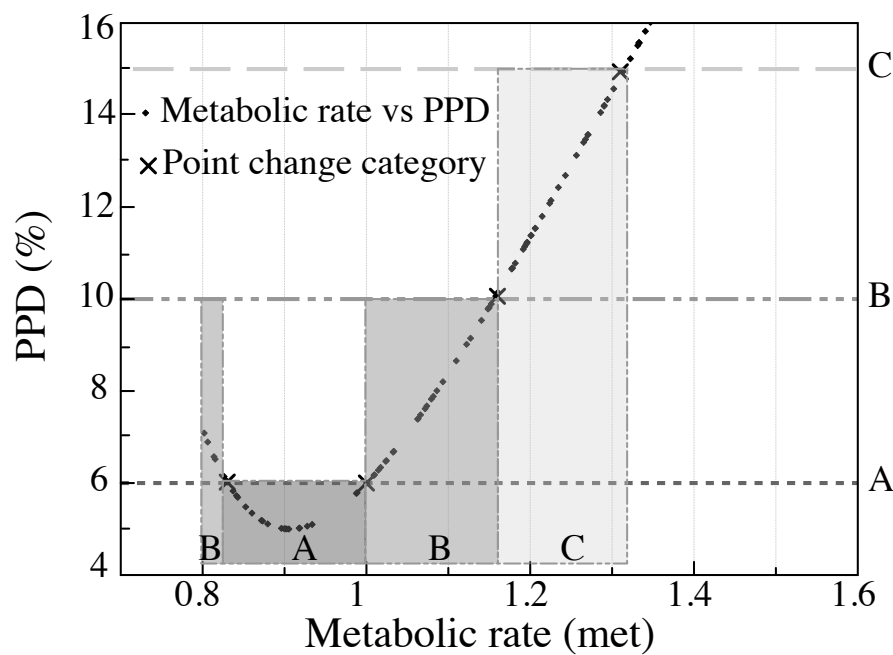


Figure 5. Change of level of thermal comfort in base of sensitivity analysis of metabolic rate on PPD.

The normal distribution is determined by two parameters, their mean and their standard deviation, usually denoted by  $\mu$  y  $\sigma$  [79]. With this notation, the density of the distribution Gaussian is given by the equation:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ \frac{-1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right\}; \quad -\infty < x < \infty \quad (7)$$

For the case of a normal distribution, we have randomly generated values for clo with  $\mu = 1.0$ , setting the metabolic factor to 1.2 met and, varying the standard deviation  $\sigma_{clo} = 0.025, 0.05, 0.075, 0.1$  and 0.125 . For these values, we obtain the distribution of the PMV and PPD values and, by adjusting to a normal distribution, we obtain the values of  $\mu$  and  $\sigma$  for the PMV and PPD. Similarly, for the case

of the met (metabolic factor) a  $\mu = 1.2$  has been established, setting the clothing factor to 1.0 clo and varying the standard deviation  $\sigma_{met} = 0.025, 0.05, 0.075, 0.1$  and  $0.125$ .

That is to say, we see easily as variations in the values of clo and met, they are transferred through the model to variations of PMV and PPD, and therefore, we can see the sensitivity of the model to these variations of the clo and met parameters.

That is, we have obtained the relation of the variations of the clothing (clo) and activity (met) with the widths ( $\sigma$ ) of the PMV and PPD, with the objective of observe which of the two parameters that the designer enters in a subjective manner, produces a more significant impact on the output of the model.

In the Figure 6a, the relationship between the widths of the distributions of clo and met and the value of  $\sigma_{PMV}$  is presented. For example, for a width of 5 % with respect to the central value of clo  $\mu = 1$  and ( $\sigma_{clo} = 0.10$ ), a value of  $\sigma_{PMV}$  of 0.150 is obtained and for a width of 5 % with respect to the central value  $\mu = 1.2$  of met ( $\sigma_{met} = 0.10$ ), a value of  $\sigma_{PMV}$  of 0.184 is obtained. Therefore, the impact of the metabolic factor (met) is higher than that of the clothing factor (clo) for the PMV case.

In the same way, for each percentage of variation in the values of clo and met we have a value of  $\sigma_{PPD}$ , which is shown in the Figure 6b, for example: for a width of 12.5 % the value of  $\sigma_{PPD}$  is 1.990 for a variation in clo and 2.594 if the parameter met varies. Also, the values predict that the impact of the metabolic factor (met) is greater than the clothing factor (clo) in the PPD.

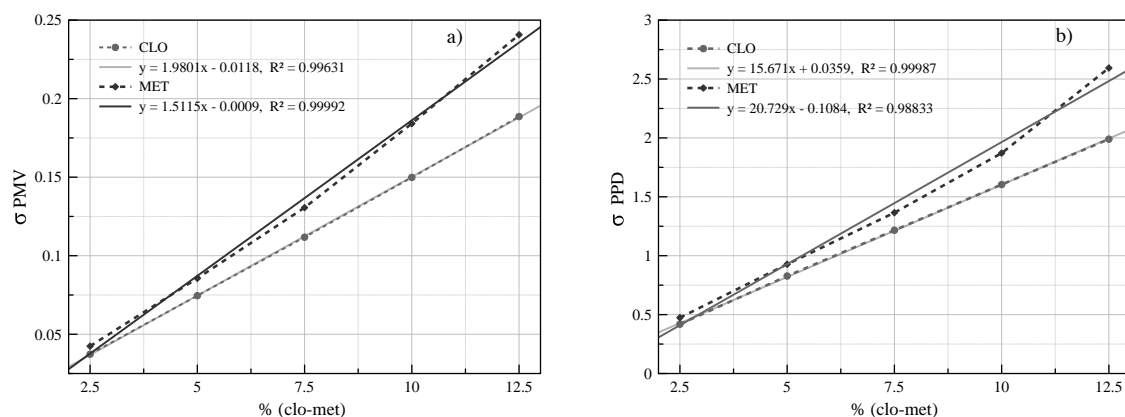


Figure 6. Association between variations of clothing and activity with the widths of (a)  $\sigma_{PMV}$  (b)  $\sigma_{PPD}$ .

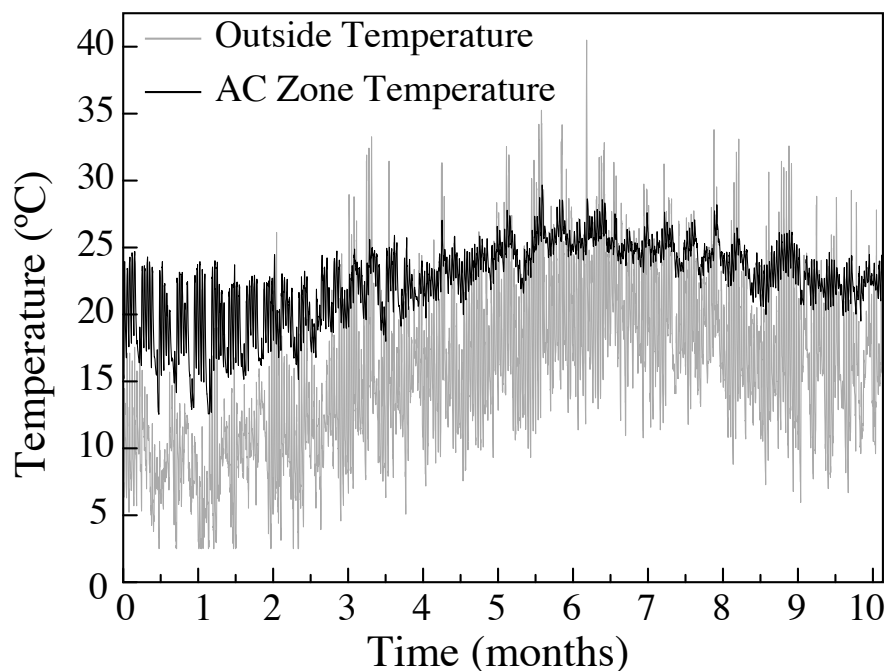
## 5. Results

In this section, we discuss the results of the simulations of both buildings. First, we will show the verification of the monitoring using the data from Equis and the TRNSYS simulation of each building. Then, after a second simulation, we will get the energy demand using the calculated results for the category change, depending on the estimation of the PMV and PPD values as well as the sensitivity analysis of the CLO and MET variables. Finally, we analyze the economic impact of energy demand savings.

As noted before, this work joins BIM-based energy simulations and SA of two non-residential buildings in two seasons (summer and winter) under the international standard ISO 7730. This norm has been studied in previous papers from different perspectives: the effect of personal factors in comfort assessment [80], energy simulation [27], thermal comfort in outdoor urban spaces [81] or the validity of ISO-PMV for predicting comfort votes in everyday thermal environments [82], among others. However, to our knowledge, the effect on the energy demand of the diversity of possibilities offered by ISO 7730 for the choice of clo and met results, depending on the values chosen by the designer, was not previously calculated.

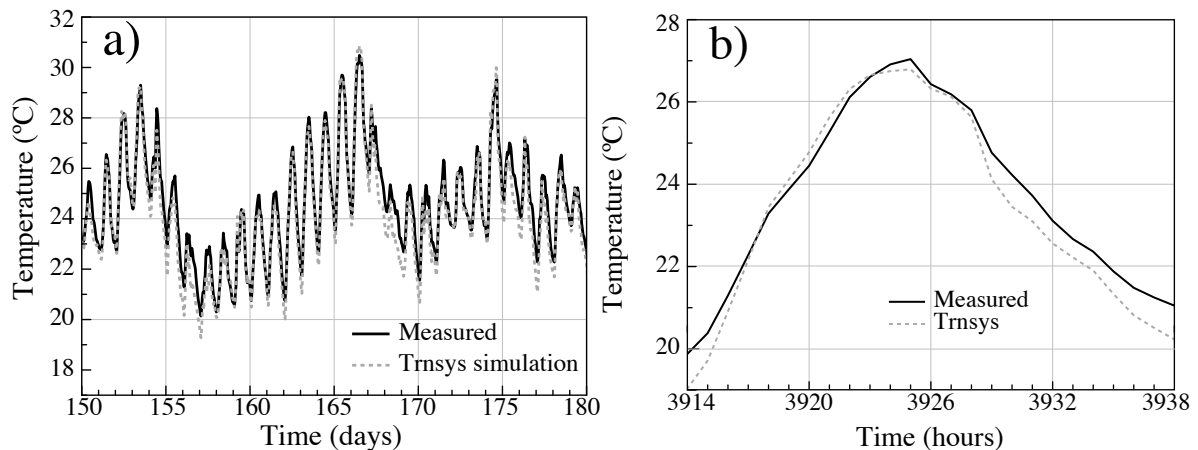
### 5.1. Real-Time Monitoring and Validation of The Models

Real-time tracking made online with Equis demonstrates the reliability of the measurements that validate the computational simulations. In Figure 7 we show the temperatures measured from 1 January to 31 October 2015 of the air-conditioning zone of the building in Spain, located  $43^{\circ}18'$  north and  $8^{\circ}17'$  west. In the first month (January, winter season) the lowest temperature of the period is  $12.53^{\circ}\text{C}$  registered on Saturday 17. The highest temperature ( $30.48^{\circ}\text{C}$ ), was achieved on Monday, 15 June (summer). The mean temperature was  $22.43^{\circ}\text{C}$ . For the study of the outside temperature, we observe that the cold and hot temperatures are between  $2^{\circ}\text{C}$  and  $41^{\circ}\text{C}$ , approximately. The mean temperature outside the building was  $15.49^{\circ}\text{C}$ .



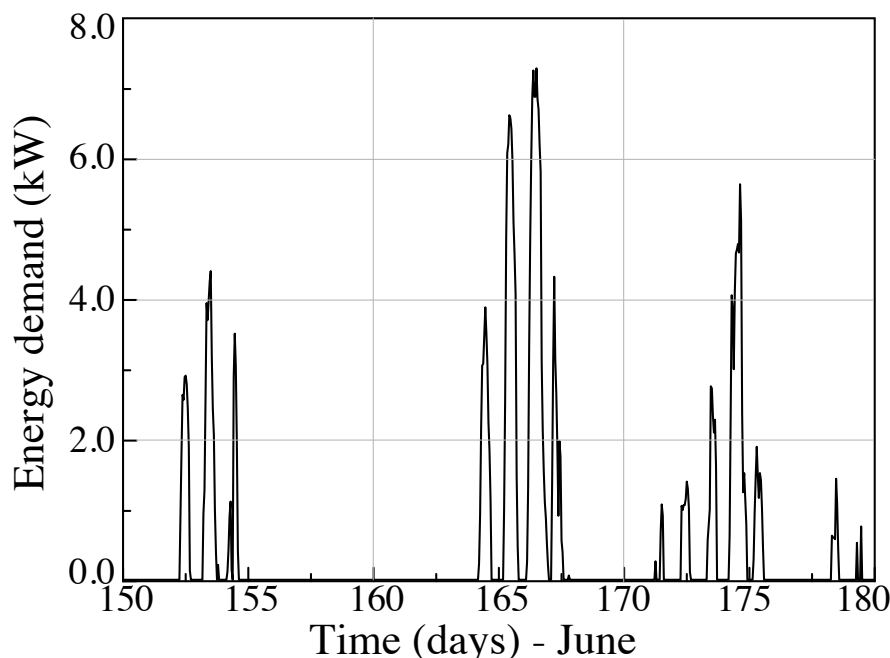
**Figure 7.** Internal and external temperatures measured by the web platform Equis. In the case of the outside temperature, we observe that the cold and hot temperatures are between  $2^{\circ}\text{C}$  and  $41^{\circ}\text{C}$ , approximately. The mean temperature was  $15.49^{\circ}\text{C}$ .

The comparison between the real data and the simulations is required to have reliable simulations of the buildings. Currently, a simulation of a building's energy consumption is usually considered "calibrated" if the criteria set out by ASHRAE Guideline 14 is met [83]. Accordingly, the model can be considered "calibrated," according to current international criteria to accept BES models, provided that considerable accordance is reached between the measured and simulated data [84]. We must stress that there is no unique model that meets the previous criteria and therefore, several alternative models of the same building can be considered to be "calibrated". Moreover, it should be noted that our calibration is very accurate for the simulated environment (e.g., temperature profiles). As an example, the parallelism among experimental data and the modeling for internal air temperature for the month of June from the hour 3624 to 4344, is shown in Figure 8a. For the period of June, the maximum difference between the simulated and measured temperatures is  $1.47^{\circ}\text{C}$ , being the mean deviation  $0.49^{\circ}\text{C}$ . These values have no influence with respect to indoor comfort, showing that the results of our model are basically correct.



**Figure 8.** Comparison between the real data obtained by Equis and theoretical simulations results of the first case of study for the month of June (a) and a single day of June (b).

In Figure 8b we show the temperature of a single day of June (24 h from hour 3914 to 3938 of the simulation). We can observe in more detail the difference between the calculated and the measured temperature for the period, where the minimum variation is of  $0.04\text{ }^{\circ}\text{C}$  (obtained in working hours) and the mean difference is of  $0.35\text{ }^{\circ}\text{C}$ . The maximum variation is  $0.72\text{ }^{\circ}\text{C}$ . These results were obtained within the situation described in the previous sections. In particular, during the occupancy periods, the metabolic rate (met) of the occupiers was selected as to be “seated light work”, which corresponds to  $70\text{ W/m}^2$  or 1.2 met. The clothing factor (clo) was established to 0.7 for the summer period and 1.2 for winter. Notice that a clothing with factor 1.2 clo corresponds to: underwear, shirt, pants, thermal jacket, socks, shoes. Summer clothing of 0.7 clo corresponds to underwear, shirt, pants, socks, shoes [26]. Table 5 shows the input that was used for the simulation for the base case scenario. In Figure 9 we can observe the energy demand in the initial configuration, where the set point was established in  $23.8\text{ }^{\circ}\text{C}$ .



**Figure 9.** Energy demand of the building in initial configuration of thermal comfort parameters.

## 5.2. Evaluation of the Variations in the Energy Demand

The main goal of the application of a SA in this work is to visualize in a simpler and more detailed way than in TRNSYS, the impact of subjective decisions made by the designer on the values

of insulation of clothes and activity for thermal comfort, which are defined in the TRNSYS simulation program. Based on ISO 7730, the insulation of clothes (clo) can be estimated directly from the combination of clothes or by adding the partial insulation values for each clothes. For example, typical value of clothes insulation for summer (cooling) and winter (heating) are 0.5 and 1 clo, respectively. Therefore, these quantities can be determinant for the different types of clothes that the occupants use and thus, the designers have a greater margin of error when determining by subjective decisions the values of clothing parameters.

Once we have made the variation of clo and met, we obtain variations in the PMV and PPD indices, these values of the indices allow us to define the thermal category based on the definition of the ISO 7730 standard. This category establishes the range of recommended temperatures for these input parameters of comfort. In this permissible temperature range, we can select a value of set-point temperature, to introduce it to our validated BIM model and thus obtain the new energy demand.

For instance: summer light clothing insulation of 0.5 clo includes items such as underpants, short-sleeved shirt, light pants, thin socks, or shoes. In our case, we choose for both cases of study a mean value of 0.7 clo that includes underwear, shirt, pants, socks, shoes, which are also considered light clothing for the same season in this type of buildings. Analogously, the estimation of the value of activity or metabolic rate (met), which represents a heat production depending on the activity level, also depends on the subjective decision made by the designer. According to ISO 7730 for an activity corresponding to this type of buildings such as office, school, and laboratory, the activity parameter can vary from 1.0 to 1.6 met. For this work we assumed a value of 1.2 met for the energy simulations on TRNSYS.

Thus, it is of fundamental importance to determine the impact of the mentioned subjective choices of clothing and metabolic factor, on the values of the PMV and PPD indexes, which show the level of thermal comfort and local thermal comfort criteria, reporting the environmental requirements that are considered acceptable for global thermal comfort, including the local discomfort.

We must notice that a variation in the clo values of summer clothing, from 0.3 to 0.8, causes a change in category from A to C. In Figure 10, the impact on the calculation of energy demand is clearly shown that has the change of environmental category.

According to the SA results and the results of Table 6, we established in TRNSYS the new values for clo and set-point temperature. The first configuration was set for a value of clo of 0.7 with a set-point temperature of 23.5 °C, which belongs to category A. For category B, we used a temperature of 24 °C and clo of 0.5. This configuration was initially defined for the validation of the TRNSYS simulation previously shown. Finally, a set-point temperature of 25 °C for category C and clo of 0.3 was considered. As an example, for the month of June, from Figure 10 for clo values of 0.3, 0.5 and 0.7, we obtain a total energy demand of 347.19 kW, 565.05 kW, and 885.2 kW, respectively.

In Table 7 it can be observed that changes in the energy demand due to variations in the designer's criteria in climates such as Mexico, can reach up to 22% depending on the choice of clo and up to 23% in energy demand, depending on the choice of the activity rate made by the designer. As it can be seen after the analysis of the Monte Carlo method, the BIM calculation is very sensitive to both the appropriate choice of the met and the clo choice. In the case of the clo we have studied variations of daily clothing in summer with clo 0.3 (panties, t-shirt, shorts, thin socks, sandals), clo 0.5 (underpants, t-shirt, light pants, thin socks, shoes), clo 0.7 (underwear, shirt, pants, socks, shoes).

All previous values of clo are assignable to the attire of the occupants of both buildings studied both in the case of Spain and in Mexico, in summer. However, after analyzing it by the Monte Carlo method and calculating the environmental category by the TRNSYS program, we see that in both buildings for a room 0.3 the comfort in the buildings studied is in the category A ranges; for a clo of 0.5 in category B and if the clo is 0.3, in category C (Figure 4). This variation in the category, due to the uncertainty in the choice of the clo on the part of the designer, can have, consequently, alterations in the energy demand.

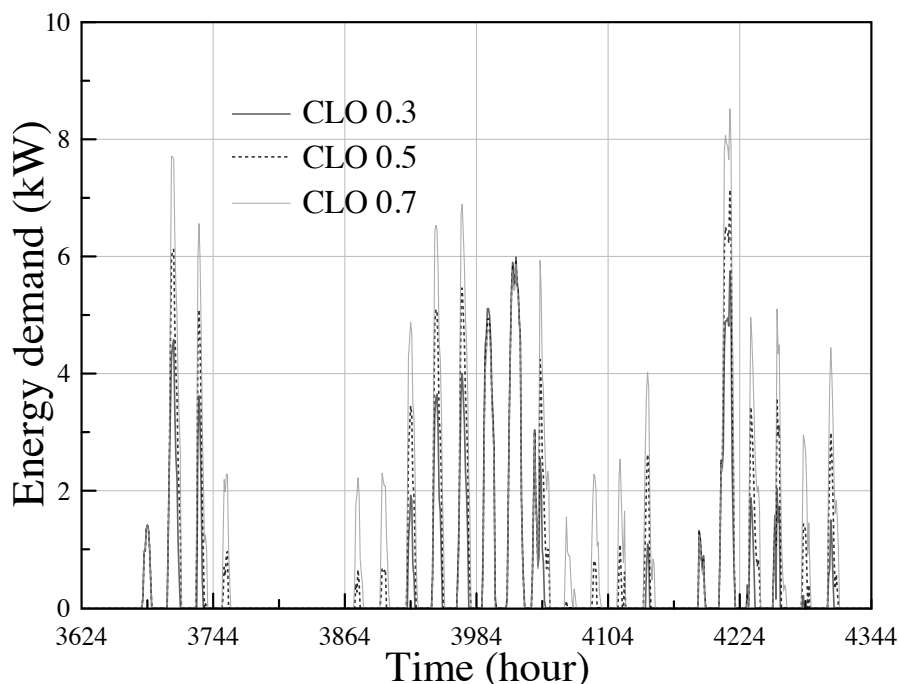


Figure 10. Energy demand for several values of clothing factor.

Table 7. Cumulated energy demand between June and September, with varying values of clo and met, for different set temperatures based on the categories of thermal comfort.

Category (Case)	Set temp	Cumulated Energy Demand (kWh/m <sup>2</sup> ) for clo & met (June to September)					
		0.3 clo	0.5 clo	0.7 clo	0.8 met	1.0 met	1.2 met
A (Spain)	23 °C	46.01	46.32	46.51	48.02	48.20	48.85
B (Spain)	24 °C	44.67	44.69	44.91	46.61	46.72	46.84
C (Spain)	25 °C	43.86	43.93	43.99	45.84	45.89	45.95
A (Mexico)	23 °C	119.53	120.76	122.03	150.35	152.85	155.36
B (Mexico)	24 °C	108.42	109.42	110.44	126.43	128.71	131.02
C (Mexico)	25 °C	99.67	100.44	101.23	106.09	106.85	108.93

As it can be appreciated from the values of Table 7 the change is more than 20%, for the building in Mexico. For the building in Spain, the variation in the calculation of the energy demand, despite having a higher sensitivity, is much lower (only 6%). This discrepancy is due to the different volumes of the two buildings. The higher the volume, the greater the energy demand and therefore the variations of the environmental categories that determine the thermal comfort inside have a much greater impact.

With respect to the met, according to ISO 7730, the cases of a technical office (main activity of the building of Spain) and a University (main activity of the building of Mexico) correspond to met 1 (sitting rest), and met 1.2 (office activity, school, home, laboratory). As it can be appreciated from the values of Table 7, the higher variations in the calculation of energy demand in buildings have been again in the case of Mexico (more than 20%). In the case of the building in Spain the oscillation is only of 6.6%.

## 6. Conclusions

In the present work, we have analyzed, within the mandatory standard ISO 7730, the influence of the designer's choice of values for human metabolic index and insulation by clothing that can be selected for the estimation of the energy demand of structures. We have demonstrated that the diversity of possibilities offered by ISO 7730 for the choice of clo and met results in significant



differences in indoor comfort conditions, leading to non-negligible changes in the calculations of energy consumption.

It has been clearly outlined that the higher the cooling demands, the greater the sensitivity, generating greater differences according to the designer's choice. Since ISO 7730 is an international norm, given the great variety of existing climates, it is vital to take into account this peculiarity of the system at the time of its application and in the choice of the designer, who must be familiar with the procedure. The international norm ISO 7730 is a reference in most of the countries of the world and nevertheless requires of an experienced designer for an appropriate application. Therefore, we have shown in the present work the considerable impact of the designer's subjective choices on the calculation of the energy demand of buildings, as well as the extra costs cumulated, depending on the election of the amounts of activity or clothing, based on the environmental ergonomic standard ISO 7730. We consider that our SA of the parameters mentioned above using the Monte Carlo method will help to define the most appropriate metrics to be admitted, depending on the aim and type of the modeling performed.

We have calculated that the effect of to the different possibilities of election within the ISO 7730 can reach a quarter of the energy demand, showing that this standard is susceptible to the designer's choice of the clo and met. As the ISO 7730 is the only standard which applies to the building regulations in many parts of the world, we conclude that the norm could be revisited from the perspective of energy efficiency, maybe including recommendations from other standards such as EN 15251-2007 or ISO 17772-1-2017, designed explicitly for providing environmental input parameters for design and assessment of energy performance of buildings.

**Author Contributions:** Conceptualization, R.R.F., M.C.H.-L. and S.Z.; Methodology, M.C.H.-L. and H.M.; Software, R.R.F. and P.F.-d.-C.; Validation, S.Z., P.F.-d.-C. and H.M.; Formal Analysis, M.C.H.-L. and R.R.F.; Investigation, R.R.F., S.Z., A.C.-G. and M.C.H.-L.; Resources, P.F.-d.-C., A.C.-G. and R.S.-A. ; Data Curation, R.R.F., R.S.-A. and M.C.H.-L.; Writing-Original Draft Preparation, S.Z., R.R.F., M.C.H.-L. and H.M.; Writing-Review & Editing, A.C.-G., R.S.-A., H.M., S.Z., M.C.H.-L., R.R.F. and P.F.-d.-C.; Visualization, A.C.-G., M.C.H.-L. and R.S.-A.; Supervision, P.F.-d.-C., H.M. and S.Z.; Project Administration, S.Z. and P.F.-d.-C.; Funding Acquisition, P.F.-d.-C., A.C.-G. and R.S.-A.

**Funding:** This work was partially funded by grants OHMERA MAT2017-86453-R, FIS2017-83762-P and ENE2015-71333-R from MINECO (Spain). R. Robledo and M. Hernández were supported by CONACYT grants 298503 and 296471, respectively. We also thanks to supporting given by the project number INFRA-187906 from the Mexican National Council of Science and Technology-CONACYT.

**Acknowledgments:** The authors acknowledge Fridama Instalaciones, S.L. and Facultad de Ciencias Físico-Matemáticas, U.A.N.L., for the access to the data used in the present work.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

In this appendix section, some tables used to complete the information of the manuscript are shown.



**Table A1.** Thermal and structural properties in windows.

	Window Types	Glass	Thickness (mm)	% Frame	U ( $\text{Wm}^{-2}\text{K}^{-1}$ )	G-Value
Case 1 (Spain)	Type A	double low emissivity glass	4/6/4	15	3.44	0.76
	Type 1	Single solar control glass	6	5	5.73	0.482
Case 2 (Mexico)	Type 2	Single clear glass	6	15	5.73	0.837
	Type 3	Double clear glass	6/12/6	10	3.21	0.722
	Type 4	Single clear glass	6	35	5.73	0.837
	Type 5	Single clear glass	2	5	5.87	0.888
	Type 6	Single solar control glass	6	25	5.73	0.482

**Table A2.** Distribution of the cases of study.

	Zones	Partitions	Characteristics
Case 1 (Spain)	Without air-conditioning (WAC)	Storage and air chamber of the roof.	Area (253.36 m <sup>2</sup> ) Volume (1910 m <sup>3</sup> ) Capacitance (2292.71 kJ/°K)
	Air-conditioned (AC)	2 Offices, 2 meeting rooms, small storage kitchen, rack, access and cleaning room.	Area (218.13 m <sup>2</sup> ) Volume (610.78 m <sup>3</sup> ) Capacitance (732.94 kJ/°K)
Case 2 (Mexico)	Zone GF-FF	GF: the reception of the building. FF: labs, offices, toilets, study areas, dining room.	Area (1142.68 m <sup>2</sup> ) Volume (3366.99 m <sup>3</sup> ) Capacitance (4040.39 kJ/°K)
	Zone SF	Labs, offices, toilets computer center, cleaning room, meeting room.	Area (1142.68 m <sup>2</sup> ) Volume (3225.0 m <sup>3</sup> ) Capacitance (3870.0 kJ/°K)
	Zone SL	Unused open space (Skylight)	Area (84.0 m <sup>2</sup> ) Volume (462.0 m <sup>3</sup> ) Capacitance (554.4 kJ/°K)

**Table A3.** Types of schedules in the buildings.

	Schedule Type	Hours	Use Factor	
Case 1 (Spain)	Daily	Daily 1	08:00 to 18:30	1
		Daily 2	13:00 to 16:00	1
		Daily 3	08:00 to 13:00 16:00 to 18:30	1 1
	Weekly	Weekly 1	Monday to Friday	Daily 1
		Weekly 2	Monday to Friday	Daily 2
		Weekly 3	Monday to Friday	Daily 3
Case 2 (Mexico)	Daily	Daily A	09:00 to 21:00	1
		Daily B	07:00 to 21:00	1
		Daily C	10:00 to 14:00	1
	Weekly	Weekly A	Monday to Friday Saturday	Daily A Daily C
		Weekly B	Monday to Sunday	Daily B

**Table A4.** Infiltrations and ventilation for each zone in summer and winter.

Case of Study	Zones	Infiltrations		Ventilation
		Air Change per Hour		Air Change per Hour
		Summer	Winter	
Case 1: Spain	WAC	0.143	0.031	5.00
	AC	0.066	0.067	1.65
Case 2: Mexico	GF-FF	3.39	1.80	8.00
	SF	3.41	1.81	8.00
	SL	25.27	13.41	10.0

## References

- Hemsath, T.L.; Bandhosseini, K.A. Sensitivity analysis evaluating basic building geometry's effect on energy use. *Renew. Energy* **2015**, *76*, 526–538. [CrossRef]
- Griego, D.; Krarti, M.; Hernandez-Guerrero, A. Energy efficiency optimization of new and existing office buildings in Guanajuato, Mexico. *Sustain. Cities Soc.* **2015**, *17*, 132–140. [CrossRef]
- Lin, H.W.; Hong, T. On variations of space-heating energy use in office buildings. *Appl. Energy* **2013**, *111*, 515–528. [CrossRef]
- Pikas, E.; Thalfeldt, M.; Kurnitski, J. Cost optimal and nearly zero energy building solutions for office buildings. *Energy Build.* **2014**, *74*, 30–42. [CrossRef]
- Terés-Zubiaga, J.; Campos-Celador, A.; González-Pino, I.; Escudero-Revilla, C. Energy and economic assessment of the envelope retrofitting in residential buildings in Northern Spain. *Energy Build.* **2015**, *86*, 194–202. [CrossRef]
- Lee, J.; Kim, J.; Song, D.; Kim, J.; Jang, C. Impact of external insulation and internal thermal density upon energy consumption of buildings in a temperate climate with four distinct seasons. *Renew. Sustain. Energy Rev.* **2017**, *75*, 1081–1088. [CrossRef]
- IEA. *World Energy Outlook 2017*; Organisation for Economic Co-Operation and Development, OECD: Paris, France, 2017.
- Anderson, J.E.; Wulfhorst, G.; Lang, W. Energy analysis of the built environment—A review and outlook. *Renew. Sustain. Energy Rev.* **2015**, *44*, 149–158. [CrossRef]
- Abdelaziz, E.; Saidur, R.; Mekhilef, S. A review on energy saving strategies in industrial sector. *Renew. Sustain. Energy Rev.* **2011**, *15*, 150–168. [CrossRef]
- Nejat, P.; Jomehzadeh, F.; Taheri, M.M.; Gohari, M.; Majid, M.Z.A. A global review of energy consumption, CO<sub>2</sub> emissions and policy in the residential sector (with an overview of the top ten CO<sub>2</sub> emitting countries). *Renew. Sustain. Energy Rev.* **2015**, *43*, 843–862. [CrossRef]
- Balaras, C.A.; Drousa, K.; Dascalaki, E.; Kontoyiannidis, S. Heating energy consumption and resulting environmental impact of European apartment buildings. *Energy Build.* **2005**, *37*, 429–442. [CrossRef]
- Pérez-Lombard, L.; Ortiz, J.; Pout, C. A review on buildings energy consumption information. *Energy Build.* **2008**, *40*, 394–398. [CrossRef]
- Galvin, R. Thermal upgrades of existing homes in Germany: The building code, subsidies, and economic efficiency. *Energy Build.* **2010**, *42*, 834–844. [CrossRef]
- Creyts, J.C. Reducing US Greenhouse Gas Emissions: How Much at What Cost?: US Greenhouse Gas Abatement Mapping Initiative. *McKinsey & Co.* 2007. Available online: <https://www.mckinsey.com/business-functions/sustainability/our-insights/reducing-us-greenhouse-gas-emissions> (accessed on 1 March 2018).
- Chappells, H.; Shove, E. Debating the future of comfort: Environmental sustainability, energy consumption and the indoor environment. *Build. Res. Inf.* **2005**, *33*, 32–40. [CrossRef]
- Geva, A.; Saaroni, H.; Morris, J. Measurements and simulations of thermal comfort: A synagogue in Tel Aviv, Israel. *J. Build. Perform. Simul.* **2014**, *7*, 233–250. [CrossRef]
- Nguyen, A.T.; Reiter, S. Passive designs and strategies for low-cost housing using simulation-based optimization and different thermal comfort criteria. *J. Build. Perform. Simul.* **2014**, *7*, 68–81. [CrossRef]

18. Rey Martínez, F.J.; Chicote, M.A.; Peñalver, A.V.; González, A.T.; Gómez, E.V. Indoor air quality and thermal comfort evaluation in a Spanish modern low-energy office with thermally activated building systems. *Sci. Technol. Built Environ.* **2015**, *21*, 1091–1099. [[CrossRef](#)]
19. Wardiningsih, W.; Troynikov, O. Force attenuation capacity and thermophysiological wear comfort of vertically lapped nonwoven fabric. *J. Text. Inst.* **2017**, *109*, 1–9. [[CrossRef](#)]
20. Ashrae, A.S. *Standard 62-1989, Ventilation for Acceptable Indoor Air Quality*, Atlanta, GA; American Society of Heating, Refrigerating, and Air Conditioning Engineers, Inc.: New York, NY, USA, 1989.
21. MacArthur, J.; Arens, E.; Gonzalez, R.; Berglund, L.; Spain, S.; Madsen, T.; Oleson, B.; Reid, K. HVAC is for people. *Ashrae Trans.* **1986**, *92*, 5–64.
22. Oropeza-Perez, I.; Petzold-Rodriguez, A.H.; Bonilla-Lopez, C. Adaptive thermal comfort in the main Mexican climate conditions with and without passive cooling. *Energy Build.* **2017**, *145*, 251–258. [[CrossRef](#)]
23. *Standard 55 Thermal Environmental Conditions For Human Occupancy*; ASHRAE: New York, NY, USA, 2010.
24. CEN. *15251-Criteria for the Indoor Environment, Including Thermal, Indoor Air Quality (Ventilation), Light And Noise, 2006*; CEN: Brussels, Belgium, 2006.
25. Matzarakis, A.; Mayer, H.; Iziomon, M.G. Applications of a universal thermal index: Physiological equivalent temperature. *Int. J. Biometeorol.* **1999**, *43*, 76–84. [[CrossRef](#)]
26. *Standard 7730. Ergonomics of the Thermal Environment—Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the Pmv and Ppd Indices and Local Thermal Comfort Criteria*; International Organization for Standardization: Geneva, Switzerland, 2005.
27. Peeters, L.; De Dear, R.; Hensen, J.; D’haeseleer, W. Thermal comfort in residential buildings: Comfort values and scales for building energy simulation. *Appl. Energy* **2009**, *86*, 772–780. [[CrossRef](#)]
28. Ahmed, K.; Akhondzada, A.; Kurnitski, J.; Olesen, B. Occupancy schedules for energy simulation in new prEN16798-1 and ISO/FDIS 17772-1 standards. *Sustain. Cities Soc.* **2017**, *35*, 134–144. [[CrossRef](#)]
29. Antoniadou, P.; Papadopoulos, A.M. Occupants’ thermal comfort: State of the art and the prospects of personalized assessment in office buildings. *Energy Build.* **2017**, *153*, 136–149. [[CrossRef](#)]
30. Oropeza-Perez, I.; Østergaard, P.A. Potential of natural ventilation in temperate countries—a case study of Denmark. *Appl. Energy* **2014**, *114*, 520–530. [[CrossRef](#)]
31. Zhang, L.; Zhang, L.; Wang, Y. Shape optimization of free-form buildings based on solar radiation gain and space efficiency using a multi-objective genetic algorithm in the severe cold zones of China. *Sol. Energy* **2016**, *132*, 38–50. [[CrossRef](#)]
32. Lei, J.; Yang, J.; Yang, E.H. Energy performance of building envelopes integrated with phase change materials for cooling load reduction in tropical Singapore. *Appl. Energy* **2016**, *162*, 207–217. [[CrossRef](#)]
33. Chen, C.W.; Lee, C.W.; Lin, Y.W. Air Conditioning—Optimizing Performance by Reducing Energy Consumption. *Energy Environ.* **2014**, *25*, 1019–1024. [[CrossRef](#)]
34. Sivak, M. Potential energy demand for cooling in the 50 largest metropolitan areas of the world: Implications for developing countries. *Energy Policy* **2009**, *37*, 1382–1384. [[CrossRef](#)]
35. Attia, S.; Hensen, J.L.; Beltrán, L.; De Herde, A. Selection criteria for building performance simulation tools: contrasting architects’ and engineers’ needs. *J. Build. Perform. Simul.* **2012**, *5*, 155–169. [[CrossRef](#)]
36. Crawley, D.B.; Lawrie, L.K.; Winkelmann, F.C.; Buhl, W.F.; Huang, Y.J.; Pedersen, C.O.; Strand, R.K.; Liesen, R.J.; Fisher, D.E.; Witte, M.J.; et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy Build.* **2001**, *33–34*, 319–331. [[CrossRef](#)]
37. Attia, S.; De Herde, A. Early design simulation tools for net zero energy buildings: A comparison of ten tools. In Proceedings of the Conference 12th International Building Performance Simulation Association, Sydney, Australia, 14–16 November 2011.
38. Klein, S.A. *TRNSYS-A Transient System Simulation Program*; University of Wisconsin-Madison, Engineering Experiment Station Report; University of Wisconsin-Madison: Madison, WI, USA, 1988; pp. 38–12.
39. Newsham, G.R. Clothing as a thermal comfort moderator and the effect on energy consumption. *Energy Build.* **1997**, *26*, 283–291. [[CrossRef](#)]
40. Hensen, J.L.; Lamberts, R. *Introduction to Building Performance Simulation*; Building Performance Simulation for Design and Operation: London, UK, 2011; pp. 365–401.
41. Schiavon, S.; Lee, K.H. Dynamic predictive clothing insulation models based on outdoor air and indoor operative temperatures. *Build. Environ.* **2013**, *59*, 250–260. [[CrossRef](#)]

42. Lee, Y.S.; Malkawi, A.M. Simulating multiple occupant behaviors in buildings: An agent-based modeling approach. *Energy Build.* **2014**, *69*, 407–416. [[CrossRef](#)]
43. Kang, D.H.; Mo, P.H.; Choi, D.H.; Song, S.Y.; Yeo, M.S.; Kim, K.W. Effect of MRT variation on the energy consumption in a PMV-controlled office. *Build. Environ.* **2010**, *45*, 1914–1922. [[CrossRef](#)]
44. Luo, M.; Cao, B.; Zhou, X.; Li, M.; Zhang, J.; Ouyang, Q.; Zhu, Y. Can personal control influence human thermal comfort? A field study in residential buildings in China in winter. *Energy Build.* **2014**, *72*, 411–418. [[CrossRef](#)]
45. De Dear, R.; Brager, G.S. *Developing an Adaptive Model of Thermal Comfort And Preference*; UC Berkeley: Berkeley, CA, USA, 1998.
46. De Dear, R.J. A global database of thermal comfort field experiments. *ASHRAE Trans.* **1998**, *104*, 1141.
47. Manu, S.; Shukla, Y.; Rawal, R.; Thomas, L.E.; de Dear, R. Field studies of thermal comfort across multiple climate zones for the subcontinent: India Model for Adaptive Comfort (IMAC). *Build. Environ.* **2016**, *98*, 55–70. [[CrossRef](#)]
48. Hwang, R.L.; Shu, S.Y. Building envelope regulations on thermal comfort in glass facade buildings and energy-saving potential for PMV-based comfort control. *Build. Environ.* **2011**, *46*, 824–834. [[CrossRef](#)]
49. Ioannou, A.; Itard, L. Energy performance and comfort in residential buildings: Sensitivity for building parameters and occupancy. *Energy Build.* **2015**, *92*, 216–233. [[CrossRef](#)]
50. Hong, T.; Taylor-Lange, S.C.; D'Oca, S.; Yan, D.; Corgnati, S.P. Advances in research and applications of energy-related occupant behavior in buildings. *Energy Build.* **2016**, *116*, 694–702. [[CrossRef](#)]
51. Yan, D.; O'Brien, W.; Hong, T.; Feng, X.; Gunay, H.B.; Tahmasebi, F.; Mahdavi, A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **2015**, *107*, 264–278. [[CrossRef](#)]
52. Chen, Y.; Liang, X.; Hong, T.; Luo, X. Simulation and visualization of energy-related occupant behavior in office buildings. *Build. Simul.* **2017**, *10*, 785–798. [[CrossRef](#)]
53. Putra, H.C.; Andrews, C.J.; Senick, J.A. An agent-based model of building occupant behavior during load shedding. *Build. Simul.* **2017**, *10*, 845–859. [[CrossRef](#)]
54. Thomas, A.; Menassa, C.C.; Kamat, V.R. Lightweight and adaptive building simulation (LABS) framework for integrated building energy and thermal comfort analysis. *Build. Simul.* **2017**, *10*, 1023–1044. [[CrossRef](#)]
55. Lindner, A.J.; Park, S.; Mitterhofer, M. Determination of requirements on occupant behavior models for the use in building performance simulations. *Build. Simul.* **2017**, *10*, 861–874. [[CrossRef](#)]
56. Laurent, J.G.C.; Samuelson, H.W.; Chen, Y. The impact of window opening and other occupant behavior on simulated energy performance in residence halls. *Build. Simul.* **2017**, *10*, 963–976. [[CrossRef](#)]
57. Kuznik, F.; Virgone, J.; Johannes, K. Development and validation of a new TRNSYS type for the simulation of external building walls containing PCM. *Energy Build.* **2010**, *42*, 1004–1009. [[CrossRef](#)]
58. Klein, S.; Beckman, W.; Mitchell, J.; Duffie, J.; Duffie, N.; Freeman, T.; Mitchell, J.; Braun, J.; Evans, B.; Kummer, J.; et al. *TRNSYS 16—A TRaNsient System Simulation Program, User Manual*; Solar Energy Laboratory, University of Wisconsin-Madison: Madison, WI, USA, 2004.
59. Salvalai, G.; Pfafferott, J.; Sesana, M.M. Assessing energy and thermal comfort of different low-energy cooling concepts for non-residential buildings. *Energy Convers. Manag.* **2013**, *76*, 332–341. [[CrossRef](#)]
60. Lebon, M.; Fellouah, H.; Galanis, N.; Limane, A.; Guerfala, N. Numerical analysis and field measurements of the airflow patterns and thermal comfort in an indoor swimming pool: A case study. *Energy Effic.* **2017**, *10*, 527–548. [[CrossRef](#)]
61. Zhang, S.; Jiang, Y.; Xu, W.; Li, H.; Yu, Z. Operating performance in cooling mode of a ground source heat pump of a nearly-zero energy building in the cold region of China. *Renew. Energy* **2016**, *87*, 1045–1052. [[CrossRef](#)]
62. Rodríguez, G.C.; Andrés, A.C.; Muñoz, F.D.; López, J.M.C.; Zhang, Y. Uncertainties and sensitivity analysis in building energy simulation using macroparameters. *Energy Build.* **2013**, *67*, 79–87. [[CrossRef](#)]
63. Basinska, M.; Koczyk, H.; Szczechowiak, E. Sensitivity analysis in determining the optimum energy for residential buildings in Polish conditions. *Energy Build.* **2015**, *107*, 307–318. [[CrossRef](#)]
64. Tian, W. A review of sensitivity analysis methods in building energy analysis. *Renew. Sustain. Energy Rev.* **2013**, *20*, 411–419. [[CrossRef](#)]
65. Lomas, K.J.; Eppel, H. Sensitivity analysis techniques for building thermal simulation programs. *Energy Build.* **1992**, *19*, 21–44. [[CrossRef](#)]

66. Breesch, H.; Janssens, A. Performance evaluation of passive cooling in office buildings based on uncertainty and sensitivity analysis. *Sol. Energy* **2010**, *84*, 1453–1467. [[CrossRef](#)]
67. Ruiz Flores, R.; Bertagnolio, S.; Lemort, V. Global sensitivity analysis applied to total energy use in buildings. In Proceedings of the 2nd International High Performance Buildings Conference, Purdue, West Lafayette, IN, USA, 16–19 July 2012.
68. Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci. Discuss.* **2007**, *4*, 439–473. [[CrossRef](#)]
69. Balvís, E.; Sampedro, Ó.; Zaragoza, S.; Paredes, A.; Michinel, H. A simple model for automatic analysis and diagnosis of environmental thermal comfort in energy efficient buildings. *Appl. Energy* **2016**, *177*, 60–70. [[CrossRef](#)]
70. Remund, J. *Meteonorm: Irradiation data for every place on Earth*. Bern2014: Switzerland, 2014. Available online: <https://meteonorm.com> (accessed on 20 September 2018).
71. Fanger, P.O. Thermal environment—Human requirements. *Environmentalist* **1986**, *6*, 275–278. [[CrossRef](#)]
72. Hasan, M.H.; Alsaleem, F.; Rifaie, M. Sensitivity study for the PMV thermal comfort model and the use of wearable devices biometric data for metabolic rate estimation. *Build. Environ.* **2016**, *110*, 173–183. [[CrossRef](#)]
73. Fanger, P.O. Thermal comfort. Analysis and applications in environmental engineering. In *Thermal Comfort. Analysis and Applications in Environmental Engineering*; Danish Technical Press: Copenhagen, Denmark, 1970.
74. Madsen, T. *Description of Thermal Manikin for Measuring Thermal Insulation Values of Clothing*; Thermal Insulation Report; Technical University of Denmark: Lyngby, Denmark, 1976.
75. DIN. 13779: *Ventilation For Non-Residential Buildings-Performance Requirements for Ventilation and Room-Conditioning Systems*; DIN: Berlin, Germany, 2007.
76. Ascione, F.; Bianco, N.; De Stasio, C.; Mauro, G.M.; Vanoli, G.P. Multi-stage and multi-objective optimization for energy retrofitting a developed hospital reference building: A new approach to assess cost-optimality. *Appl. Energy* **2016**, *174*, 37–68. [[CrossRef](#)]
77. Echenagucia, T.M.; Capozzoli, A.; Cascone, Y.; Sassone, M. The early design stage of a building envelope: Multi-objective search through heating, cooling and lighting energy performance analysis. *Appl. Energy* **2015**, *154*, 577–591. [[CrossRef](#)]
78. Lam, J.C.; Hui, S.C. Sensitivity analysis of energy performance of office buildings. *Build. Environ.* **1996**, *31*, 27–39. [[CrossRef](#)]
79. Pértegas Díaz, S.; Pita Fernández, S. La distribución normal. *Cad Aten Primaria* **2001**, *8*, 268–274.
80. Havenith, G.; Holmér, I.; Parsons, K. Personal factors in thermal comfort assessment: Clothing properties and metabolic heat production. *Energy Build.* **2002**, *34*, 581–591. [[CrossRef](#)]
81. Nikolopoulou, M.; Baker, N.; Steemers, K. Thermal comfort in outdoor urban spaces: Understanding the human parameter. *Sol. Energy* **2001**, *70*, 227–235. [[CrossRef](#)]
82. Humphreys, M.A.; Nicol, J.F. The validity of ISO-PMV for predicting comfort votes in every-day thermal environments. *Energy Build.* **2002**, *34*, 667–684. [[CrossRef](#)]
83. ASHRAE. Guideline 14-2002. *Meas. Energy Demand Sav.* **2002**, *22*, 32–43.
84. Coakley, D.; Raftery, P.; Keane, M. A review of methods to match building energy simulation models to measured data. *Renew. Sustain. Energy Rev.* **2014**, *37*, 123–141. [[CrossRef](#)]



## Research Article

# Calculation of Some Integrals Arising in Heat Transfer in Geothermics

**J. L. G. Santander,<sup>1</sup> P. Castañeda Porras,<sup>2</sup> J. M. Isidro,<sup>3</sup>  
and P. Fernández de Córdoba<sup>3</sup>**

<sup>1</sup> *Departamento de Ciencias Experimentales y Matemáticas, Universidad Católica de Valencia, 46002 Valencia, Spain*

<sup>2</sup> *Departamento de Matemáticas, Universidad de Pinar del Río, 20200 Pinar del Río, Cuba*

<sup>3</sup> *Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, 46022 Valencia, Spain*

Correspondence should be addressed to J. L. G. Santander, martinez.gonzalez@ucv.es

Received 28 September 2010; Accepted 29 November 2010

Academic Editor: Miguel Cerrolaza

Copyright © 2010 J. L. G. Santander et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We calculate some integrals involved in the temperature field evaluation of the ground, heated by a borehole heat exchanger. This calculation allows a faster computation of that component of the temperature field which involves the periodic oscillation of the ambient temperature or the ambient heat flux.

## 1. Introduction

Ground-coupled heat pumps (GCHPs) are an attractive choice of system for heating and cooling buildings [1]. By comparison with standard technologies, these pumps offer competitive levels of comfort, reduced noise levels, lower greenhouse gas emissions, and reasonable environmental safety. Furthermore, their electrical consumption and maintenance requirements are lower than those required by conventional systems and, consequently, have lower annual operating cost [2].

A GCHP exchanges heat with the ground through a buried U tube loop. Since this exchange strongly depends on the thermal properties of the ground, it is very important to have a knowledge of these properties when designing GCHP air-conditioning systems. Methods to estimate them include literature searches, conducting laboratory experiments on soil/rock samples, and/or performing field tests. The *in situ* tests are based on studying the thermal response of the borehole heat exchanger to a constant heat injection or extraction. The outputs of the thermal response test are the inlet and outlet temperature of the heat-carrier



fluid as a function of time. From these experimental data, and with an appropriate model describing the heat transfer between the fluid and the ground, the thermal conductivity of the surroundings is inferred.

The results presented in this study are based on a model of the underground heat transfer due to a borehole heat exchanger assuming that it works as a line heat source of finite depth inside the ground. We assume as well a constant power of length unit for the heat source considered and a periodic oscillation of the ambient temperature or heat flux on the ground surface.

This paper is organized as follows. Section 2 describes the equations that model the heat transfer inside the ground due to a borehole heat exchanger. Section 3 solves the integrals given in literature for periodic oscillations of the ambient temperature or heat flux, on the ground surface. Section 4 shows that our results agree with the literature in some particular cases, such as the solution on the ground surface and the solution of the quasistationary regime.

## 2. The Equations

Let us consider the heat equation in cylindrical coordinates, in which  $T(r, z, t)$  is the temperature field,

$$\frac{\partial T}{\partial t} - k\nabla^2 T = \frac{Q_z}{C} \delta(r) [H(z) - H(z - L)]. \quad (2.1)$$

On the one hand, the thermal parameters of (2.1) are the ground thermal diffusivity  $k(m^2s^{-1})$ , the heat flow per length unit  $Q_z(Wm^{-1})$ , and the volumetric heat capacity of ground  $C(Jm^{-3}K^{-1})$ . On the other hand, the geometry of the heat source, depicted in Figure 1, has been taken into account introducing the Heaviside function  $H$  for its finite length  $L$  and the Dirac delta distribution  $\delta$  for its infinitely thin width. Equation (2.1) is subject to the initial condition

$$T(r, z, 0) = T_0 + k_{\text{geo}}z, \quad z \geq 0, \quad (2.2)$$

where  $T_0$  is the undisturbed ground temperature and  $k_{\text{geo}}$  ( $Km^{-1}$ ) is the geothermal gradient. We may consider that (2.1) is subject to a Dirichlet boundary condition, which represents a time-dependent temperature on the surface  $z = 0$ ,

$$T(r, 0, t) = \psi(t), \quad t \geq 0, \quad (2.3)$$

or to a Neumann boundary condition, which represents a time-dependent heat flux on the surface  $z = 0$ ,

$$-\lambda \partial_z T(r, 0, t) = \varphi(t), \quad t \geq 0, \quad (2.4)$$

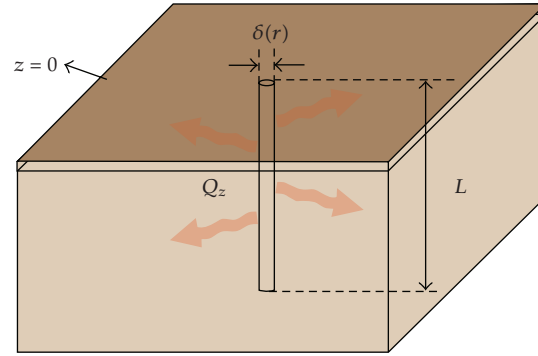


Figure 1: Borehole heat exchanger model.

where  $\lambda(Wm^{-1}K^{-1})$  is the ground thermal conductivity. We split the Cauchy problem posed in (2.1), (2.2), and (2.3) or (2.4) in three terms [3], so that

$$T(r, z, t) = v_d(r, z, t) + v_0(z, t) + v_s(z, t), \quad (2.5)$$

and where the functions  $v_0$ ,  $v_s$  and  $v_d$  satisfy the Cauchy problems presented below.

### 2.1. The Problem for $v_d$

The partial differential equation that  $v_d(r, z, t)$  satisfies is

$$\frac{\partial v_d}{\partial t} - k\nabla^2 v_d = \frac{Q_z}{C} \delta(r) [H(z) - H(z - L)], \quad (2.6)$$

subject to a homogeneous initial condition,

$$v_d(r, z, 0) = 0, \quad z \geq 0, \quad (2.7)$$

and to a homogeneous Dirichlet boundary condition,

$$v_d(r, 0, t) = 0, \quad t \geq 0, \quad (2.8)$$

or to a Neumann boundary condition, homogenous as well,

$$\partial_z v_d(r, 0, t) = 0, \quad t \geq 0. \quad (2.9)$$



The solution to (2.6), (2.7), and (2.8) is [4]

$$v_d(r, z, t) = \frac{Q_z}{8\pi\lambda} \int_{r^2/4kt}^{\infty} du \frac{e^{-u}}{u} \times \left\{ 2 \operatorname{erf}\left(\frac{z}{r}\sqrt{u}\right) + \operatorname{erf}\left(\frac{L-z}{r}\sqrt{u}\right) - \operatorname{erf}\left(\frac{L+z}{r}\sqrt{u}\right) \right\}, \quad (2.10)$$

and the solution to (2.6), (2.7), and (2.9) is

$$v_d(r, z, t) = \frac{Q_z}{8\pi\lambda} \int_{r^2/4kt}^{\infty} \frac{e^{-u}}{u} \left\{ \operatorname{erf}\left(\frac{z+L}{r}\sqrt{u}\right) - \operatorname{erf}\left(\frac{z-L}{r}\sqrt{u}\right) \right\} du, \quad (2.11)$$

that we will present in a future work.

## 2.2. The Problem for $v_0$

The partial differential equation that  $v_0(z, t)$  satisfies is

$$\frac{\partial v_0}{\partial t} = k \frac{\partial^2 v_0}{\partial z^2}, \quad (2.12)$$

subject to the initial condition

$$v_0(z, 0) = T_0 + k_{\text{geo}}z, \quad z \geq 0, \quad (2.13)$$

and to a Dirichlet boundary condition,

$$v_0(0, t) = 0, \quad t \geq 0, \quad (2.14)$$

or to a Neumann homogeneous boundary condition

$$\partial_z v_0(0, t) = 0, \quad t \geq 0. \quad (2.15)$$

The solution to (2.12), (2.13), and (2.14) is [5, Equation 2.4.13],

$$v_0(z, t) = T_0 \operatorname{erf}\left(\frac{z}{2\sqrt{kt}}\right) + k_{\text{geo}}z, \quad (2.16)$$

and the solution to (2.12), (2.13), and (2.15) is [5, Equation 2.9.7]

$$v_0(z, t) = T_0 + k_{\text{geo}} \left\{ z \operatorname{erf}\left(\frac{z}{2\sqrt{kt}}\right) + \frac{2\sqrt{kt}}{\sqrt{\pi}} \exp\left(-\frac{z^2}{4kt}\right) \right\}, \quad (2.17)$$

where the error function is defined as

$$\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du. \quad (2.18)$$

### 2.3. The Problem for $v_s$

The partial differential equation that  $v_s(z, t)$  satisfies is

$$\frac{\partial v_s}{\partial t} = k \frac{\partial^2 v_s}{\partial z^2}, \quad (2.19)$$

subject to an homogeneous initial condition

$$v_s(z, 0) = 0, \quad z \geq 0, \quad (2.20)$$

and to the time-dependent Dirichlet boundary condition

$$v_s(0, t) = \varphi(t), \quad t \geq 0, \quad (2.21)$$

or to the time-dependent Neumann boundary condition

$$-\lambda \partial_z v_s(0, t) = \varphi(t), \quad t \geq 0. \quad (2.22)$$

The solution to (2.19), (2.20) and (2.21) is [5, Equation 2.5.1.]

$$v_s(z, t) = \frac{2}{\sqrt{\pi}} \int_{z/2\sqrt{kt}}^{\infty} \varphi\left(t - \frac{z^2}{4ku^2}\right) e^{-u^2} du, \quad (2.23)$$

and the solution to (2.19), (2.20) and (2.22) is [5, Equation 2.9.9]

$$v_s(z, t) = \frac{\sqrt{k}}{\lambda \sqrt{\pi}} \int_0^t \frac{\varphi(t-u)}{\sqrt{u}} e^{-z^2/4ku} du. \quad (2.24)$$

## 3. Harmonic Analysis of the Boundary Condition

We may consider that on the ground surface we have a periodic temperature or heat flux oscillation, due to daily or annual cycles. Therefore,  $\varphi(t)$  is a periodic function of a certain period  $\tau$ , continuous and bounded, that we may expand in Fourier series

$$\varphi(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\omega t) + b_n \sin(n\omega t), \quad (3.1)$$

where the frequency is  $\omega = 2\pi/\tau$ , and the coefficients  $a_n$  and  $b_n$  are given by

$$\begin{aligned} a_0 &= \frac{2}{\tau} \int_{-\tau/2}^{\tau/2} \psi(u) du, \\ a_n &= \frac{2}{\tau} \int_{-\tau/2}^{\tau/2} \psi(u) \cos(n\omega u) du, \\ b_n &= \frac{2}{\tau} \int_{-\tau/2}^{\tau/2} \psi(u) \sin(n\omega u) du, \end{aligned} \quad (3.2)$$

where  $n \in \mathbb{N}$ .

### 3.1. Dirichlet Boundary Condition

Substituting (3.1) in (2.23), we obtain

$$\begin{aligned} v_s(z, t) &= \frac{a_0}{2} \operatorname{erfc}\left(\frac{z}{2\sqrt{kt}}\right) \\ &+ \frac{2}{\sqrt{\pi}} \sum_{n=1}^{\infty} a_n \int_{z/2\sqrt{kt}}^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du \\ &+ b_n \int_{z/2\sqrt{kt}}^{\infty} \sin\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du, \end{aligned} \quad (3.3)$$

where  $\operatorname{erfc}(z) = 1 - \operatorname{erf}(z)$  is the complementary error function. The integrals given in (3.3) for  $t \geq 0$  may be rewritten as

$$\begin{aligned} \int_{z/2\sqrt{kt}}^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du &= \operatorname{Re}\left[e^{in\omega t} I_{D,n}(z, t)\right], \\ \int_{z/2\sqrt{kt}}^{\infty} \sin\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du &= \operatorname{Im}\left[e^{in\omega t} I_{D,n}(z, t)\right], \end{aligned} \quad (3.4)$$

where we have defined

$$I_{D,n}(z, t) := \int_{z/2\sqrt{kt}}^{\infty} \exp\left(-\frac{in\omega z^2}{4ku^2} - u^2\right) du. \quad (3.5)$$

In order to calculate (3.5), let us apply the integral [6, Equation 7.4.33]

$$\int \exp\left(-ax^2 - \frac{b}{x^2}\right) dx = \frac{\sqrt{\pi}}{4\sqrt{a}} \left[ e^{2\sqrt{ab}} \operatorname{erf}\left(\sqrt{a}x + \frac{\sqrt{b}}{x}\right) + e^{-2\sqrt{ab}} \operatorname{erf}\left(\sqrt{a}x - \frac{\sqrt{b}}{x}\right) \right] + \text{const}, \quad (3.6)$$

taking  $a = 1$ ,  $b = in\omega z^2/4k$ , and  $z \geq 0$ , so that

$$I_{D,n}(z, t) = \frac{\sqrt{\pi}}{4} \left[ \exp\left(\frac{z\sqrt{in\omega}}{\sqrt{k}}\right) \operatorname{erfc}\left(\frac{z}{2\sqrt{kt}} + \sqrt{in\omega t}\right) + \exp\left(\frac{-z\sqrt{in\omega}}{\sqrt{k}}\right) \operatorname{erfc}\left(\frac{z}{2\sqrt{kt}} - \sqrt{in\omega t}\right) \right]. \quad (3.7)$$

Finally, substituting (3.7) in (3.4), we obtain for  $z, t \geq 0$ ,

$$\int_{z/2\sqrt{kt}}^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{4} \operatorname{Re}\left\{e^{in\omega t} [A_n^+(z, t) + A_n^-(z, t)]\right\}, \quad (3.8)$$

$$\int_{z/2\sqrt{kt}}^{\infty} \sin\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{4} \operatorname{Im}\left\{e^{in\omega t} [A_n^+(z, t) + A_n^-(z, t)]\right\},$$

where we have defined

$$A_n^{\pm}(z, t) := \exp\left(\pm \frac{z\sqrt{in\omega}}{\sqrt{k}}\right) \operatorname{erfc}\left(\frac{z}{2\sqrt{kt}} \pm \sqrt{in\omega t}\right). \quad (3.9)$$

As far as we know, the integrals given in (3.8) are not reported in the literature.

### 3.1.1. Neumann Boundary Condition

Substituting (3.1) in (2.24), we obtain

$$v_s(z, t) = \frac{a_0\sqrt{k}}{2\lambda\sqrt{\pi}} \int_0^t \frac{e^{-z^2/4ku}}{\sqrt{u}} du \quad (3.10)$$

$$+ \frac{\sqrt{k}}{\lambda\sqrt{\pi}} \sum_{n=1}^{\infty} a_n \int_0^t \frac{\cos[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du \quad (3.11)$$

$$+ b_n \int_0^t \frac{\sin[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du. \quad (3.12)$$

Integrating by parts and performing the substitution  $\zeta = z/2\sqrt{ku}$ , considering  $z, t \geq 0$ , we may calculate the integral given in (3.10)

$$\int_0^t \frac{e^{-z^2/4ku}}{\sqrt{u}} du = 2\sqrt{t} \exp\left(\frac{-z^2}{4kt}\right) - \frac{\sqrt{\pi}z}{\sqrt{k}} \operatorname{erfc}\left(\frac{z}{2\sqrt{kt}}\right). \quad (3.13)$$

In order to calculate the integrals given in (3.11) and (3.12), let us perform the substitution  $u = v^2$ , considering  $t \geq 0$ , so that

$$\begin{aligned} \int_0^t \frac{\cos[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du &= 2 \int_0^{\sqrt{t}} \cos[n\omega(t-v^2)] e^{-z^2/4kv^2} dv \\ &= 2 \operatorname{Re} \left[ e^{in\omega t} I_{N,n}(z,t) \right], \end{aligned} \quad (3.14)$$

and similarly

$$\int_0^t \frac{\sin[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du = 2 \operatorname{Im} \left[ e^{in\omega t} I_{N,n}(z,t) \right], \quad (3.15)$$

where we have defined

$$I_{N,n}(z,t) := \int_0^{\sqrt{t}} \exp \left( -in\omega v^2 - \frac{z^2}{4kv^2} \right) dv. \quad (3.16)$$

Taking  $a = in\omega$ ,  $b = z^2/4k$ , and  $z \geq 0$  in (3.6), we may rewrite (3.16) as

$$\begin{aligned} I_{N,n}(z,t) &= \frac{\sqrt{\pi}}{4\sqrt{in\omega}} \left[ \exp \left( \frac{-z\sqrt{in\omega}}{\sqrt{k}} \right) \operatorname{erfc} \left( \frac{z}{2\sqrt{kt}} - \sqrt{in\omega t} \right) \right. \\ &\quad \left. - \exp \left( \frac{z\sqrt{in\omega}}{\sqrt{k}} \right) \operatorname{erfc} \left( \frac{z}{2\sqrt{kt}} + \sqrt{in\omega t} \right) \right]. \end{aligned} \quad (3.17)$$

Finally, substituting (3.17) in (3.14) and (3.15), for  $z, t \geq 0$ , we obtain

$$\begin{aligned} \int_0^t \frac{\cos[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du &= \frac{\sqrt{\pi}}{2} \operatorname{Re} \left\{ \frac{e^{in\omega t}}{\sqrt{in\omega}} [A_n^-(z,t) - A_n^+(z,t)] \right\}, \\ \int_0^t \frac{\sin[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du &= \frac{\sqrt{\pi}}{2} \operatorname{Im} \left\{ \frac{e^{in\omega t}}{\sqrt{in\omega}} [A_n^-(z,t) - A_n^+(z,t)] \right\}. \end{aligned} \quad (3.18)$$

As far as we know, the integrals given in (3.18) are not reported in the literature.

#### 4. Particular Cases

As a consistency check, let us verify that in certain particular cases the new integrals given in (3.8) and (3.18) are reduced to integrals reported in the literature. These particular cases have to do with the quasistationary regime and the solution on the ground surface.

### 4.1. Quasistationary Regime

Since there is a periodic oscillation on the ground surface, the surroundings of the borehole heat exchanger never reach a stationary regime. However, according to [5, Section 2.6], we may define a quasistationary regime in which the ground temperature field is stabilized periodically. In this quasistationary regime, (3.8) becomes

$$\int_0^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{4} \operatorname{Re}\left\{e^{in\omega t} A_{D,n}(z)\right\}, \quad (4.1)$$

$$\int_0^{\infty} \sin\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{4} \operatorname{Im}\left\{e^{in\omega t} A_{D,n}(z)\right\}, \quad (4.2)$$

where we have defined

$$A_{D,n}(z) := \lim_{t \rightarrow \infty} [A_n^+(z, t) + A_n^-(z, t)]. \quad (4.3)$$

Similarly, (3.18) becomes

$$\int_0^{\infty} \frac{\cos[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du = \frac{\sqrt{\pi}}{2} \operatorname{Re}\left\{\frac{e^{in\omega t}}{\sqrt{in\omega}} A_{N,n}(z)\right\}, \quad (4.4)$$

$$\int_0^{\infty} \frac{\sin[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du = \frac{\sqrt{\pi}}{2} \operatorname{Im}\left\{\frac{e^{in\omega t}}{\sqrt{in\omega}} A_{N,n}(z)\right\},$$

where we have defined

$$A_{N,n}(z) := \lim_{t \rightarrow \infty} [A_n^-(z, t) - A_n^+(z, t)]. \quad (4.5)$$

#### 4.1.1. Dirichlet Case

Taking into account that  $\operatorname{erfc}(\infty) = 0$  and  $\operatorname{erfc}(-\infty) = 2$ , we may calculate the limit given in (4.3), so that (4.1) becomes

$$\begin{aligned} \int_0^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du &= \frac{\sqrt{\pi}}{2} \operatorname{Re}\left\{e^{in\omega t} \exp\left(\frac{-z\sqrt{in\omega}}{\sqrt{k}}\right)\right\} \\ &= \frac{\sqrt{\pi}}{2} e^{-z\sqrt{n\omega/2k}} \operatorname{Re}\left\{e^{i(n\omega t - z\sqrt{n\omega/2k})}\right\}, \end{aligned} \quad (4.6)$$

where we have substitute  $\sqrt{i} = (1+i)/\sqrt{2}$ . Thus, for  $z, t \geq 0$ ,

$$\int_0^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{2} e^{-z\sqrt{n\omega/2k}} \cos\left(n\omega t - \frac{z\sqrt{n\omega}}{\sqrt{2k}}\right). \quad (4.7)$$

Similarly, we may obtain

$$\int_0^{\infty} \sin\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du = \frac{\sqrt{\pi}}{2} e^{-z\sqrt{n\omega/2k}} \sin\left(n\omega t - \frac{z\sqrt{n\omega}}{\sqrt{2k}}\right). \quad (4.8)$$

Notice that (4.7) and (4.8) are sinusoidal in  $t$ , with an amplitude exponentially decreasing in depth  $z$ . In the literature, we may find [7, Equation 3.928.1-2].

$$\begin{aligned} \int_0^{\infty} e^{-p^2 x^2 - q^2/x^2} \sin\left(a^2 x^2 + \frac{b^2}{x^2}\right) dx &= \frac{\sqrt{\pi}}{2r} e^{-2rs \cos(A+B)} \\ &\quad \times \sin(A + 2rs \sin(A+B)), \\ \int_0^{\infty} e^{-p^2 x^2 - q^2/x^2} \cos\left(a^2 x^2 + \frac{b^2}{x^2}\right) dx &= \frac{\sqrt{\pi}}{2r} e^{-2rs \cos(A+B)} \\ &\quad \times \cos(A + 2rs \sin(A+B)), \end{aligned} \quad (4.9)$$

where  $a^2 + p^2 > 0$  and

$$\begin{aligned} r &= \sqrt[4]{a^4 + p^4}, & s &= \sqrt[4]{b^4 + q^4}, \\ A &= \frac{1}{2} \tan^{-1}\left(\frac{a^2}{p^2}\right), & B &= \frac{1}{2} \tan^{-1}\left(\frac{b^2}{q^2}\right), \end{aligned} \quad (4.10)$$

thus, taking  $p = 1$ ,  $q = 0$ , and  $a = 0$ ,

$$\begin{aligned} \int_0^{\infty} e^{-x^2} \sin\left(\frac{b^2}{x^2}\right) dx &= \frac{\sqrt{\pi}}{2} e^{-\sqrt{2}b} \sin(\sqrt{2}b), \\ \int_0^{\infty} e^{-x^2} \cos\left(\frac{b^2}{x^2}\right) dx &= \frac{\sqrt{\pi}}{2} e^{-\sqrt{2}b} \cos(\sqrt{2}b). \end{aligned} \quad (4.11)$$

Therefore, rewriting the left side of (4.7) as

$$\begin{aligned} \int_0^{\infty} \cos\left(n\omega t - \frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du &= \cos(n\omega t) \int_0^{\infty} \cos\left(\frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du \\ &\quad + \sin(n\omega t) \int_0^{\infty} \sin\left(\frac{n\omega z^2}{4ku^2}\right) e^{-u^2} du, \end{aligned} \quad (4.12)$$

and applying (4.11), taking  $b = \sqrt{n\omega}z/2\sqrt{k}$ , ( $z \geq 0$ ), we eventually get the same result as (4.7). Similarly, we may check the result given in (4.8).

#### 4.1.2. Neumann Case

In a similar way as in the previous subsection, we may calculate (4.4) for  $z, t \geq 0$ , arriving at

$$\int_0^{\infty} \frac{\cos[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du = \frac{\sqrt{\pi} e^{-z\sqrt{n\omega/2k}}}{\sqrt{n\omega}} \cos\left(n\omega t - \frac{z\sqrt{n\omega}}{\sqrt{2k}} - \frac{\pi}{4}\right), \quad (4.13)$$

$$\int_0^{\infty} \frac{\sin[n\omega(t-u)]}{\sqrt{u}} e^{-z^2/4ku} du = \frac{\sqrt{\pi} e^{-z\sqrt{n\omega/2k}}}{\sqrt{n\omega}} \sin\left(n\omega t - \frac{z\sqrt{n\omega}}{\sqrt{2k}} - \frac{\pi}{4}\right). \quad (4.14)$$

Once again, (4.13) and (4.7) are sinusoidal in  $t$ , with an amplitude exponentially decreasing in depth  $z$ . In the literature, we may find [7, Equation 3.957.1-2]

$$\begin{aligned} \int_0^{\infty} x^{\mu-1} e^{-\beta^2/4x} \cos ax \, dx &= \left(\frac{\beta}{2\sqrt{a}}\right)^{\mu} \left[ e^{-i\pi\mu/4} K_{\mu}(\beta e^{i\pi/4}\sqrt{a}) + e^{i\pi\mu/4} K_{\mu}(\beta e^{-i\pi/4}\sqrt{a}) \right], \\ \int_0^{\infty} x^{\mu-1} e^{-\beta^2/4x} \sin ax \, dx &= i \left(\frac{\beta}{2\sqrt{a}}\right)^{\mu} \left[ e^{-i\pi\mu/4} K_{\mu}(\beta e^{i\pi/4}\sqrt{a}) - e^{i\pi\mu/4} K_{\mu}(\beta e^{-i\pi/4}\sqrt{a}) \right], \end{aligned} \quad (4.15)$$

where  $\text{Re } \beta > 0$ ,  $\text{Re } \mu < 1$  and  $a > 0$ . Therefore, taking in (4.15),  $\mu = 1/2$ ,  $a = n\omega$ ,  $\beta = z/\sqrt{k}$ , ( $z \geq 0$ ), and knowing that the Macdonald function of order  $1/2$  is [8, Equation 5.5.5],

$$K_{1/2}(z) = \sqrt{\frac{\pi}{2z}} e^{-z}, \quad (4.16)$$

we eventually get the same results as (4.13) and (4.14).

### 4.2. Solution on the Ground Surface

Notice that in the special case  $z = 0$ , (3.8) becomes trivial,

$$\begin{aligned} \int_0^{\infty} \cos(n\omega t) e^{-u^2} du &= \frac{\sqrt{\pi}}{2} \cos(n\omega t), \\ \int_0^{\infty} \sin(n\omega t) e^{-u^2} du &= \frac{\sqrt{\pi}}{2} \sin(n\omega t), \end{aligned} \quad (4.17)$$

while (3.18) are reduced to

$$\begin{aligned} \int_0^t \frac{\cos[n\omega(t-u)]}{\sqrt{u}} du &= \sqrt{\pi} \text{Re} \left\{ \frac{e^{in\omega t}}{\sqrt{in\omega}} \text{erf}(\sqrt{in\omega t}) \right\}, \\ \int_0^t \frac{\sin[n\omega(t-u)]}{\sqrt{u}} du &= \sqrt{\pi} \text{Im} \left\{ \frac{e^{in\omega t}}{\sqrt{in\omega}} \text{erf}(\sqrt{in\omega t}) \right\}. \end{aligned} \quad (4.18)$$



Since [8, Equation 2.4.1]

$$\frac{\operatorname{erf}(\sqrt{i}x)}{\sqrt{i}} = \sqrt{2} \left[ C\left(\sqrt{\frac{2}{\pi}}x\right) - iS\left(\sqrt{\frac{2}{\pi}}x\right) \right], \quad x \in \mathbb{R}, \quad (4.19)$$

where

$$\begin{aligned} C(z) &:= \int_0^z \cos \frac{\pi t^2}{2} dt, \\ S(z) &:= \int_0^z \sin \frac{\pi t^2}{2} dt \end{aligned} \quad (4.20)$$

are the *Fresnel integrals*; then (4.18) may be expressed as

$$\begin{aligned} \int_0^t \frac{\cos[n\omega(t-u)]}{\sqrt{u}} du &= \sqrt{\frac{2\pi}{n\omega}} \left[ \cos(n\omega t) C\left(\sqrt{\frac{2n\omega t}{\pi}}\right) + \sin(n\omega t) S\left(\sqrt{\frac{2n\omega t}{\pi}}\right) \right], \\ \int_0^t \frac{\sin[n\omega(t-u)]}{\sqrt{u}} du &= \sqrt{\frac{2\pi}{n\omega}} \left[ \sin(n\omega t) C\left(\sqrt{\frac{2n\omega t}{\pi}}\right) - \cos(n\omega t) S\left(\sqrt{\frac{2n\omega t}{\pi}}\right) \right]. \end{aligned} \quad (4.21)$$

The results given in (4.21), agrees with [7, Equation 2.653.1-2]

$$\begin{aligned} \int \frac{\sin x}{\sqrt{x}} dx &= \sqrt{2\pi} S(\sqrt{x}), \\ \int \frac{\cos x}{\sqrt{x}} dx &= \sqrt{2\pi} C(\sqrt{x}). \end{aligned} \quad (4.22)$$

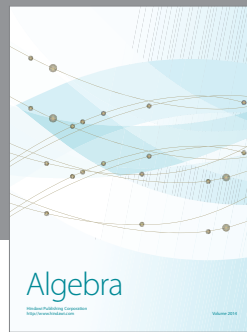
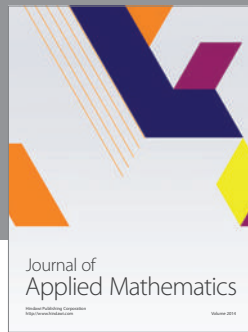
## Acknowledgments

The authors wish to thank the financial support received from Generalitat Valenciana under Grant no. GVA 3012/2009 and from Universidad Politécnic de Valencia under Grant no. PAID-06-09. This work has been partially supported by the Structural Funds of the European Regional Development Fund (ERDF).

## References

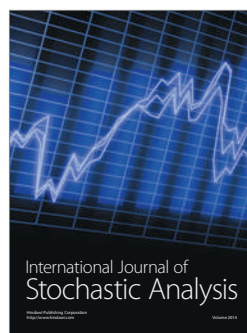
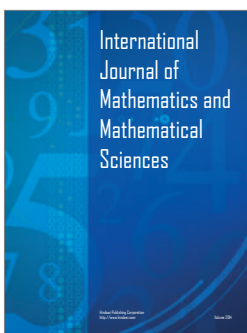
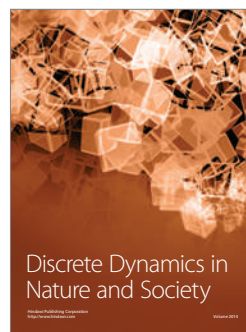
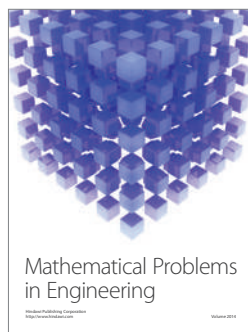
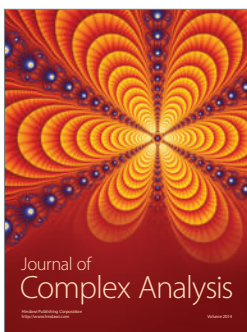
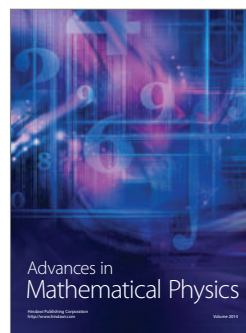
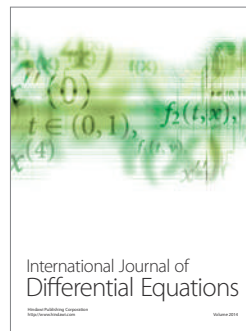
- [1] J. F. Urchueguía, M. Zacarés, J. M. Corberán, A. Montero, J. Martos, and H. Witte, "Comparison between the energy performance of a ground coupled water to water heat pump system and an air to water heat pump system for heating and cooling in typical conditions of the European Mediterranean coast," *Energy Conversion and Management*, vol. 49, no. 10, pp. 2917–2923, 2008.
- [2] J. W. Lund, "Ground source (geothermal) heat pumps," in *Course on Heating with Geothermal Energy: Conventional and New Schemes. World Geothermal Congress*, P. J. Lineau, Ed., pp. 1–21, Kazuno, Japan, 2000, Short Courses.

- [3] T. V. Bandos, A. Montero, E. Fernández et al., "Finite line-source model for borehole heat exchangers: effect of vertical temperature variations," *Geothermics*, vol. 38, no. 2, pp. 263–270, 2009.
- [4] P. Eskilson, *Thermal analysis of heat extraction boreholes*, Ph.D. thesis, Department of Mathematical Physics, University of Lund, Lund, Sweden, 1987.
- [5] H. S. Carslaw and J. C. Jaeger, *Conduction of Heat in Solids*, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, NY, USA, 2nd edition, 1988.
- [6] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications Paperback, New York, NY, USA, 1972.
- [7] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Elsevier/Academic Press, Amsterdam, Netherlands, 7th edition, 2007.
- [8] N. N. Lebedev, *Special Functions and Their Applications*, Dover, New York, NY, USA, 1972.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>



Article

## Generation and Evaluation of a Genome-Scale Metabolic Network Model of *Synechococcus elongatus* PCC7942

Julián Triana<sup>1</sup>, Arnau Montagud<sup>2,†</sup>, Maria Siurana<sup>2</sup>, David Fuente<sup>2</sup>, Arantxa Urchueguía<sup>3</sup>, Daniel Gamermann<sup>4</sup>, Javier Torres<sup>5</sup>, Jose Tena<sup>5</sup>, Pedro Fernández de Córdoba<sup>2</sup> and Javier F. Urchueguía<sup>2,\*</sup>

<sup>1</sup> Universidad de Pinar del Río “Hermanos Saíz Montes de Oca”, Pinar del Río 20100, Cuba; E-Mail: j triana@upr.edu.cu

<sup>2</sup> Instituto Universitario de Matemática Pura y Aplicada, Universitat Politècnica de València, València 46022, Spain; E-Mails: armontag@mat.upv.es (A.M.); masiupa@upv.es (M.S.); dafueher@upv.es (D.F.); pfernandez@mat.upv.es (P.F.C.)

<sup>3</sup> Biozentrum, Universität Basel, 4056 Basel, Switzerland; E-Mail: arantxa.urchueguia@gmail.com

<sup>4</sup> Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre 91501-970, Brazil; E-Mail: danielg@if.ufrgs.br

<sup>5</sup> Instituto Universitario de Medio Ambiente y Ciencia Marina, Universidad Católica de Valencia “San Vicente Mártir”, València 46001, Spain; E-Mails: javier.torres@ucv.es (J.To.); josetena@ucv.es (J.Te.).

<sup>†</sup> Current address: INSERM U900, Institut Curie, 26 rue d’Ulm, Paris 75005, France.

\* Author to whom correspondence should be addressed; E-Mail: jfurchueguia@fis.upv.es; Tel.: +34-963-879-895.

Received: 26 February 2014; in revised form: 5 August 2014 / Accepted: 12 August 2014 /

Published: 20 August 2014

---

**Abstract:** The reconstruction of genome-scale metabolic models and their applications represent a great advantage of systems biology. Through their use as metabolic flux simulation models, production of industrially-interesting metabolites can be predicted. Due to the growing number of studies of metabolic models driven by the increasing genomic sequencing projects, it is important to conceptualize steps of reconstruction and analysis. We have focused our work in the cyanobacterium *Synechococcus elongatus* PCC7942, for which several analyses and insights are unveiled. A comprehensive approach has been used, which can be of interest to lead the process of manual curation and genome-scale

metabolic analysis. The final model, *iSyf715* includes 851 reactions and 838 metabolites. A biomass equation, which encompasses elementary building blocks to allow cell growth, is also included. The applicability of the model is finally demonstrated by simulating autotrophic growth conditions of *Synechococcus elongatus* PCC7942.

**Keywords:** genome-scale metabolic network reconstruction; systems biology; metabolic pathways; flux balance analysis; biological databases

---

## 1. Introduction

*Synechococcus elongatus* PCC7942 is considered a model organism since the early 1970s, when successful transformations of exogenous DNA were performed for the first time in a cyanobacterium [1]. In particular, it has been used as a paradigm for the study of circadian rhythms in prokaryotes as it has been demonstrated that prokaryotes are capable of measuring time [2]. *S. elongatus* has a rod-shaped appearance, is oligotrophic having the ability to survive in freshwater environments with low nutrients and is considered an obligate autotroph [3,4]. The genus *Synechococcus* is among the most important photosynthetic bacteria in the marine environment as it accounts, after different estimates, for about 25% of the primary production in marine habitats [5]. Like all cyanobacteria, *Synechococcus elongatus* uses CO<sub>2</sub> as carbon source and light as energy source, which explains the interest in exploring its potential as a photo-biological cell factory for the production of valuable compounds for various applications. Potential applications are broad in this sense and research has been focused in the production of diverse metabolites of industrial interest, such as different types of biofuels [6] like hydrogen [7], among others. Its role as a model organism and the unique properties of this photosynthetic prokaryote illustrate why *Synechococcus elongatus* PCC7942 constitutes an interesting target for metabolic engineering and the benefit of developing, for the first time, a genome-scale metabolic model of this bacterium.

The development of genome sequencing and genetic mapping together with omics-science paved the way towards the quantitative study of biological systems. Thus, systems biology has emerged as a promising predictive science on a large and quantitatively deep scale [8], aiming at engineering of metabolic pathways and their capabilities [9]. Biological systems dynamics are inherently nonlinear and show functional synergies that may lead to emerging properties [10]. The construction of metabolic networks is, not only a compilation of chemical reactions, but also a gathering of exchange ratios, metabolic fluxes, and other type of biological constraints that make possible the *in silico* analyses of the organism's behaviour. These analyses have been used by researchers to design metabolic engineering strategies in a variety of problems [11–14].

The genome-scale metabolic network reconstruction is based on genetic information available on the organism of interest. The *Synechococcus elongatus* PCC7942 genome was sequenced, annotated and published in 1980 [15–17]. In order to build a meaningful model, experimental data are required together with established knowledge such as physiological and biochemical information accessible from the literature, journal articles, experiments and databases,. In certain cases, lack of clarity and quality in published data, such as mistakes in entries, false negatives and false positives undermines

the quality of the reconstructed models making their simulations worthless [18]. The relationships between complex metabolic processes usually falls to properly determine the processing of substrates into products and their stoichiometry, if a transformation is spontaneous or catalyzed by enzymes or if cofactors are involved. Moreover, the subcellular localization of the reactions and some thermodynamic aspects such as irreversibility must be known [19].

The usefulness of genome-scale metabolic models has been demonstrated through several computational analyses. Constraint-based approaches, such as flux balance analysis (FBA), are among the most common ones used to simulate phenotypic behavior under imposed physiological and/or genetic conditions [20–22]. FBA aims to obtain, through the optimization of a cellular objective (usually growth), the space of allowable flux distributions of a biological system under steady-state conditions. The optimization problem is subject to a set of constraints associated with lower and upper bounds in every reaction, which are defined by thermodynamic and experimental data. Finally, the resulting flux distribution can be contrasted with *in vivo* information and, thus, the metabolic model can be used for further analyses [23,24].

Presently, the process of reconstruction is long and arduous mainly due to its manual construction and proper quality-control check [25]. Some efforts have been directed to automate the metabolic reconstruction, or at least some parts of it, in order to cut down the time needed for such an endeavor. However, these efforts have been hampered due to problems in database information consistency and genome annotations [26]. Thereby, resulting algorithms are still unable to generate quality metabolic networks models as a basis for predictive analysis [25]. Several protocols have been published to define in detail each one of the steps of a proper reconstruction, as well as the software packages and databases that can assist in this labor [19,25–27]. For instance, Thiele and Palsson described in a very instructive way the process of debugging and validation [24].

Here, we present the manually curated metabolic reconstruction for *Synechococcus elongatus* PCC7942. The current model, *iSyf715* features a detailed biomass equation including all the building blocks that are needed for a flux distribution simulation. Moreover, FBA analysis is performed to assess the accuracy of the model and to explore possible biotechnological production strategies.

With today's energy shortage concerns, having seemingly infinite energy source of light represents an interesting avenue for research and development. Therefore, *iSyf715* is a step towards the development of photo-biological production platforms for the synthesis of several compounds of industrial interest.

## 2. Results and Discussion

In the reconstruction of a genome-scale metabolic model the reliability of the model depends critically on the quality of the data used. Nowadays, several databases (Supplementary File S1) can be used to obtain reliable metabolic reactions and annotated genome sequences for the organism of interest. The reconstruction of the model of *Synechococcus elongatus* PCC7942, termed *iSyf715* in the present work, started with an exhaustive data search and gathering of genome sequences and annotation files from the NCBI Entrez Gene database [28]. The rationale behind the name is that the “*i*” of the name refers to an *in silico* model, followed by the organism database identifier (e.g., KEGG ID) and then the number of genes whose information is included in the model.

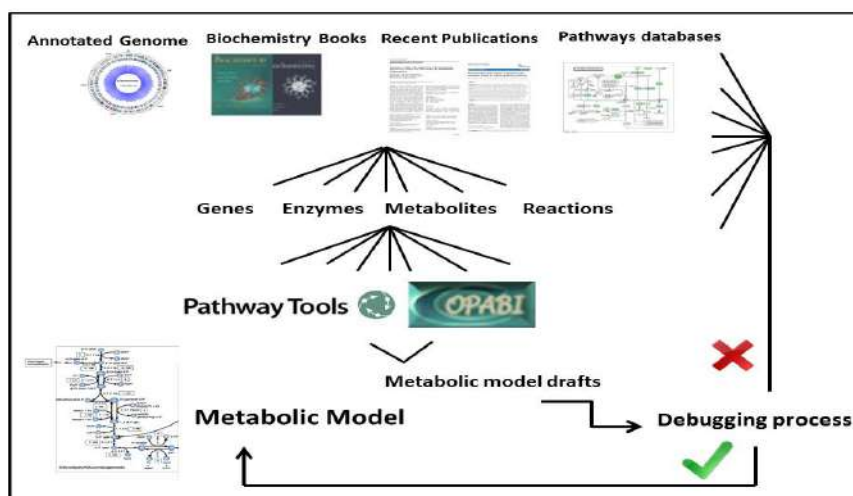


The files, of which descriptions can be found in Table 1, were used as input for the software Pathway Tools [29] in order to build a database with all the genes, proteins and metabolites present in the cyanobacteria. The list of reactions and cognate genes was then retrieved with this software. Alternatively, we used the COPABI computational platform to build a similar database and to automatically generate the metabolic model following probabilistic criteria of uniqueness and completeness [30]. The algorithm allows identifying and filling gaps in a given pathway, choosing a completeness value comparing the available information of our specific metabolic reactions (e.g., reactions related with buildings blocks) to a general pathway (meaning, ideally, “all” metabolic reactions in nature, or the meta-metabolism). This completeness value is a probability that the missing reaction can occur in our metabolic network. Furthermore, this platform gives the possibility to exclude duplicated reactions, allowing the inclusion and correction of many isoenzymes, through the uniqueness value. This strategy was conceived in order to double check the results of both software. Figure 1 summarizes the whole process.

**Table 1.** Summary of genome features of *Synechococcus elongatus* PCC7942.

	Chromosome	Plasmid pANL	Plasmid pANS
<b>Length of DNA (base pairs)</b>	2,695,903	46,366	7835
<b>G+C (%)</b>	~55.47	52.9	~59
<b>RNA genes</b>	54	-	-
<b>rRNA genes</b>	6	-	-
<b>tRNA genes</b>	45	-	-
<b>Other RNA genes</b>	3	-	-
<b>Protein genes</b>	2856	50	8
<b>With predicted function</b>	1682	17	-
<b>Without predicted function</b>	1174	33	-
<b>Total genes</b>	2906	50	8

**Figure 1.** Genome-scale metabolic reconstruction process.



The first draft included 672 genes coding for 540 enzymes that participated in 898 reactions. At this point we thoroughly checked the model identifying reactions that had no corresponding enzyme-coding

gene assigned, but that needed to be included in the model, as well as removing reactions related to genetic replication, gene expression and cell division that were not meant to be included in the model [19]. Characteristics of this first draft and the final model version can be studied in Table 2. We verified EC numbers and stoichiometry of the reactions with several databases, such as KEGG pathway [31] and MetaCyc [32], as well as a complete literature examination from different biochemistry books [33–36]. If no conclusive outcome came from these sources, certain published metabolic models, such as *Synechocystis* sp. PCC6803 [11,12], cyanobacteria of the same phylogenetic Phylum as *Synechococcus elongatus* PCC7942 served as a reference to solve these issues.

**Table 2.** Distribution of elements in the database retrieved from Pathways Tools and the final model, *iSyf715*.

<b>General Overview</b>	<b><i>iSyf715</i></b>
Genes	715
Metabolic reactions	851
Metabolites	838
Enzymes	530
Multimeric enzymes and enzymatic complexes	79
<b>Reactions overview</b>	
Reversible reactions	326
Irreversible reactions	525
<b>Reactions with assigned genes</b>	<b>735</b>
Enzymatic conversion	710
Protein-mediated transport (active and passive-mediated transports)	25
<b>Reactions with no cognate genes</b>	<b>116</b>
Non-enzymatic conversion (spontaneous)	13
Passive transport reactions (simple diffusion)	16
EC reactions not annotated	76
Unassigned reactions	11

In order to render chemical conversions coherent, all elements were balanced. Protons have been balanced, by accounting the total number of each chemical element on both sides of the associated-reactions in model. Electrons were balanced in the chemical reactions, even though we are unable to know the reducing state of metabolites in many reactions. We considered the principle of conservation of reducing power and have corrected the REDOX reactions based on an approximate balance of *electron-donor* and *electron-acceptor*. As some of the reactions included in these databases are usually reported in a non-specific form (e.g., an electron acceptor or an alcohol), corresponding organism-specific metabolites had to be identified [25]. In addition, reactions catalyzed by multimeric enzymes or enzymatic complexes were described as a single reaction [19]. The BRENDA database [37] helped to identify 325 reactions that were found to be reversible in the model. If no conclusive evidence was reported, reactions were set to be reversible.

Through these analyses, we assessed the possibility of including missing cofactors (e.g., water molecule or hydrogen ion, among others) in some reactions, like the ones catalyzed by hydrolases, oxidoreductases or transferases. If state of the art was unable to specify a single cofactor requirement, like NADH or NADPH, two reactions were included in the metabolic network.



No lumped reactions were left in the model to enable the tracing of the reactions' fluxes. Several reactions were found to be necessary for the synthesis of monomers, precursors or building blocks, but had no corresponding enzyme-coding gene assigned. We included these reactions to allow the formation of biomass, which was an objective function of the simulations. Whenever such a reaction was included in the model, it was mentioned explicitly.

In the biomass reaction, proteins, nucleic acids, lipids, carbohydrates, and other essential organic compounds, are drained together in a virtual reaction that evolves a mole of biomass. The ratios of each one of these precursors, ideally determined experimentally, are added as stoichiometries in the reaction.

Additionally, some transport systems across the membrane such as: phosphate, water, sulphate, nitrate, ammonia, as well as carbon monoxide and hydrogen peroxide transport, were included in the model and properly bounded. Some of the reversible reactions involving NADH and NADPH were constrained to be irreversible so that spurious transhydrogenation was controlled.

Another essential point in the debugging process was the removing of internal loops that are thermodynamically infeasible, for instance futile cycles, like substrate cycles described in [33] and Type III-extreme pathway [38]. Blocking these reactions is crucial since several constraint-based approaches, such as Flux Balance Analysis [20], do not account for regulation, thus, futile cycles cannot be shut down otherwise and simulations could retrieve unnatural flux behaviors.

The resulting network of this reconstruction process encompasses all known metabolites that take place in *Synechococcus elongatus* PCC7942 and consists of 851 metabolic reactions and 838 metabolites (see Table 2). The bulk of reactions are catalyzed by 530 enzymes encoded by 715 genes. The presence of protein complexes and multimeric enzymes, explains the differences between the number of enzymes and genes. Additionally, a set of reactions with no cognate genes is present in *iSyf715*: 13 non-enzymatic (spontaneous) conversions, 16 simple diffusion reactions, and 11 unassigned reactions (the majority according to the KEGG report). During the reconstruction process, 54 external metabolites and 40 exchange reactions were included. In short, a total of 76 reactions not annotated in the genome were included in the model on the basis of biochemical evidence or physiological considerations. Examples of these are the genes that encode for malate synthase (EC 2.3.3.9) and isocitratelase (EC 4.1.3.1), whose enzymatic activities have been measured [39], but do not have a cognate ORF associated to them, and whose presence is necessary to complete the glyoxylate shunt.

The final model includes central metabolic pathways, such as the glycolysis/gluconeogenesis pathway, the Calvin-Benson cycle, the pentose phosphate pathway, incomplete reactions within the tricarboxylic acid cycle (TCA), as well as the complete set of anabolic pathways involved in the biosynthesis of chlorophyll, glycogen, amino acids, lipids, nucleotides, vitamins, cofactors, *etc.* Pathways for glyoxylate synthesis (via ribulose-1, 5-bisphosphate carboxylase/oxygenase and the shunt across TCA cycle), and amino sugars metabolism are also included.

Photosynthetic electron transfer associated with the thylakoid membrane is represented as a set of 10 separate reactions, including light captured by photosystem II (PSII) and photosystem I (PSI), electron transfer between the two photosystems, and cyclic electron transfer which involves PSI and ferredoxin.

Working model files can be obtained in Supplementary information. Additionally, *iSyf715* model on SBML format was deposited in BioModels Database [40] and assigned the identifier MODEL1407310000.

### 2.1. Network Topology and Connectivity Analysis

The network topology analysis of *iSyf715* can help in the understanding of how metabolites and their interactions determine their metabolic function in the cell. As many studies have shown, most of these networks are scale-free and thereby the nodes connection can be estimated by a power-law degree distribution [41–44]. Most connected metabolites distribution in the *iSyf715* network and a comparison with their presence in other microbial genome-scale metabolic networks can be seen in Table 3.

**Table 3.** Most connected metabolites in *S. elongatus* PCC7942 and other genome-scale metabolic models.

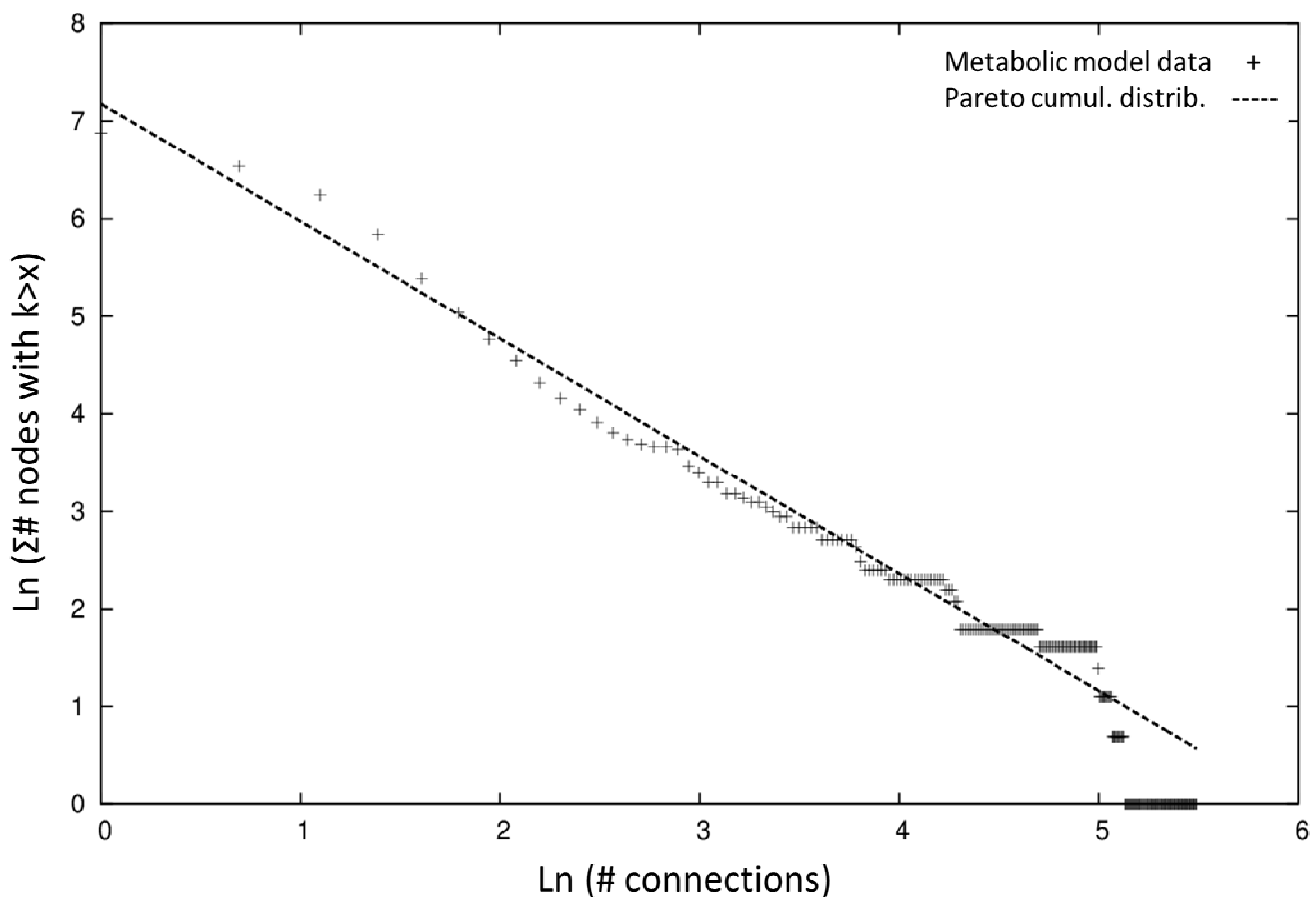
Metabolic Hubs	Neighbors in <i>iSyf715</i>	Neighbors in <i>iSyn811</i> [12]	Neighbors in <i>E. coli</i> [45]	Neighbors in Yeast [19]
H <sub>2</sub> O	243	219	697	-
Phosphate	169	112	81	113
ADP	159	111	253	131
ATP	148	136	338	166
H <sup>+</sup>	149	153	923	188
Diphosphate	110	84	28	-
CO <sub>2</sub>	69	72	53	66
AMP	74	21	86	48
NADPH	74	68	66	57
NADP <sup>+</sup>	72	68	39	61
L-glutamate	52	44	52	56
NAD <sup>+</sup>	46	52	79	58
NADH	45	48	75	52
oxygen O <sub>2</sub>	45	40	40	31
S-adenosyl-L-methionine	37	28	18	19
Ammonia	44	28	22	-
coenzyme A	29	23	71	39
Pyruvate	32	20	61	20
L-glutamine	30	21	18	23
Glutathione	32	26	17	10
S-adenosyl-L-homocysteine	25	24	12	14

A few metabolites are involved in many reactions and are widely used by diverse metabolic machineries; on the other hand, many metabolites have very few connections. These highly connected metabolites are often referred as metabolic hubs; these are the metabolites around which metabolism is organized in a bow tie manner [46]. Unsurprisingly, water is the well-connected compound. Its role both as a substrate or product in reactions, such as reduction-oxidation, hydrolysis and condensation, is well known. As we can see in Table 3, other most well connected metabolites include carrier molecules like: ADP, ATP, NADP<sup>+</sup>, NAD<sup>+</sup>, phosphate, and oxygen, which are related to energy

transport, energy storage and redox pathways; a few amino acids, peptides and their precursors (L-glutamate, L-glutamine and glutathione); and key components in the porphyrin and chlorophyll metabolism (S-adenosyl-L-methionine/S-adenosyl-L-homocysteine). Additional well connected metabolites are ammonia, coenzyme A and pyruvate, which constitute either the substrates or products of many central metabolic pathways, like glycolysis, tricarboxylic acid (TCA) cycle, glyoxylate shunt, and amino acids metabolism.

For the connectivity distribution analysis of *iSyf715* we used a systematic mathematical approach: the Pareto's law in terms of the cumulative distribution function ( $P(K > k) \sim k^{-\gamma+1}$ ) to get a proper fit [47,48]. We used the cumulative distribution rather than picturing a log-log scale plot of the distribution of connections counts among number of nodes, so that the distribution tail smoothed out in the cumulative distribution and no data were "obscured" as in the logarithmic binning procedure [49]. Our analysis of the data from Table 3 leads us to think the *iSyf715* network is reasonably characterized by a power-law distribution ( $\gamma = 2.203$ ) with high non-uniformity, as we can see at the cumulative distribution towards the right of Figure 2. The biological significance of this hierarchical connectivity is said to be related to an evolutionary process, where the hubs were the first compounds that were present in the earliest cells predecessor's metabolism [50]. Metabolic hubs represent effective targets for metabolic engineering and should be considered in the design of strategies for the production of other metabolites.

**Figure 2.** Connectivity distribution of the *iSyf715* metabolic model using the cumulative distribution towards the right via Pareto cumulative distribution.



## 2.2. Simulation of the Model

We performed a constraint-based flux simulation by using the Flux Balance Analysis (FBA) [20] algorithm. This approach has been used to simulate exponential growth in bacteria and assumes that the cells grow optimally using a given amount of carbon and energy source. Another assumption is the steady state of intracellular reactions in the exponential phase. As the number of equations (reactions) is much larger than the number of variables (metabolites), we have an underdetermined system of equations, which is solved by applying a steady state and optimizing for an objective function (cell growth or biomass evolution). This optimization was done with linear programming, which requires the establishment of boundary parameters for the uptake. The solution for this mathematical problem (see Methods) is a flux vector, where each reaction has a flux value. Model validation usually focuses on testing whether the growth capabilities or any particular objective flux corresponds to a given set of experimental flux data.

*Synechococcus elongatus* PCC7942 is an obligate photoautotroph organism, thus we defined a set of constraints for this growth conditions (detailed in Supplementary file S1, Table S2). A two-step optimization procedure was applied as in previous works [11]. The first step was the maximization of biomass growth while the light intake was unconstrained. Next, the maximum growth value was incorporated as a constraint to minimize the light uptake rate (the second step). This was designed with the aim to estimate physiologically meaningful photon uptake values that tallied experimental growth measurements. Biomass synthesis, a theoretical abstraction for cellular growth, is considered as a drain of some metabolite intermediates, into a general biomass component [11]. We looked upon information about weight fractions of macromolecules and monomers to reflect the composition of biomass. Frequently, data related to the relative amounts of these metabolites are not available in literature, or the published information is shown in a particular physiological condition not usable for our goals. As long as possible, the weight fractions were updated to reflect the specifics of *Synechococcus elongatus* PCC7942. The quantities measured in other phylogenetically related biological systems could be a close approximation to the metabolic reality of the concerned organism. In the present work we have used the biomass composition reported in Table 4. For details, please find a more detailed explanation of biomass equation in Methods and Supplementary File S1.

For the estimation of the theoretical maximum illumination we calculated the surface area per weight of biomass. We considered a spheroid geometry of the cell with a length of  $3.57 \pm 0.12 \mu\text{m}$ , a width of  $1.47 \pm 0.09 \mu\text{m}$  and a dry weight of  $3.87 \pm 0.03 \text{ ng}$  [51] and we assumed an irradiance value of  $0.156 \text{ mE m}^{-2} \text{ s}^{-1}$  in a 12:12 hours photoperiod as the value with the highest growth rate experimentally determined [51]. Taking into account these data the theoretical maximum illumination that would reach the cell membrane in our model was estimated as  $1.96 \text{ mE gDW}^{-1} \text{ h}^{-1}$ . The first optimization was carried out by constraining the  $\text{CO}_2$  and  $\text{HCO}_3^-$  uptake rates at  $1.99 \text{ mmol gDW}^{-1} \text{ h}^{-1}$  [52] and maximizing growth. Then we imposed that specific growth rate as a condition and we minimized the photon uptake rate. That is, finding the minimal quantity of photons needed for the metabolism to work with the experimentally determined carbon uptake. In this way, we estimated physiologically meaningful photon uptake values, which do tally with experimental growth measurements [11,53]. The specific growth rate for *iSyf715* resulted to be  $0.05987 \text{ h}^{-1}$  with a photon uptake rate of  $0.1 \text{ mE gDW}^{-1} \text{ h}^{-1}$  on each photosystem.

Due to the biphasic nature of cyanobacterium growth, we look for reported data for exponential growth phase with which to compare our simulation. The maximum specific growth rate of *Synechococcus elongatus* PCC7942 has been reported from  $0.0519 \text{ h}^{-1}$  to  $0.0551 \text{ h}^{-1}$  [52,54], despite scarce information about studies on optimization of the specific growth rate of this prokaryote. The slight difference of our simulated growth with the experimental data could be the result of several factors, including regulation, stress and feedback inhibition, which cannot be captured in constraint-based stoichiometric models. Moreover, the growth of many laboratory strains are not consistent with the computed optimal by FBA because they are not necessarily evolved for growth maximization [55,56]. These results are, therefore, an overall acceptable validation of the genome-scale metabolic model.

**Table 4.** Biomass composition of the *S. elongatus* PCC7942 metabolic model.

Metabolites	mmol/gDW	Metabolites	mmol/gDW
Proteins	0.000459	<b>Ribonucleotide</b>	
<b>Carbohydrates</b>		s	
Glycogen	0.53439	AMP	0.140389293
<b>Antenna</b>		UMP	0.140389293
<b>chromophores</b>		GMP	0.123745851
Zeaxanthin	0.00079	CMP	0.123745851
Beta-carotene	0.000875	<b>Lipids</b>	
Trans-lycopene	0.00820225	14C-lipid	0.028
Chlorophyll a	0.0057	16C-lipid	0.0042
Phycocyanobiline	0.0285	18C-lipid	0.00448
<b>Deoxyribonucleotides</b>		(9Z)16C-lipid	0.0066
dATP	0.0201156	(9Z)18C-lipid	0.00625
dTTP	0.0201156		
dGTP	0.02538445		
dCTP	0.02538445		

Furthermore, we studied split-flux reactions, nodes in the network where flux can go in two or more directions and that are very good checkpoints to see if complex metabolic behavior has been properly captured by the metabolic model (Figure 3 depicts the relevant reactions vector). The flux landscape of the model shows a flux distribution directed towards the CO<sub>2</sub> fixation at the Calvin-Benson cycle, as expected of a photosynthetic cell that uses light and carbon substrates to build complex carbohydrate molecules. The solution space reveals that the autotrophic growth flows in the gluconeogenic direction (see Figure 3). Unsurprisingly, RuBisCO (reaction “4.1.1.39b”) and carbonic anhydrase (reaction “4.2.1.1b”) have high fluxes as they are entry points of carbon dioxide and carbonic acid, the carbon sources of the system. The demand of ATP and NADPH is supplemented by the photosynthetic processes. Depending on the needs of each of them, photosynthesis will be linear or include a cyclic part if more ATP is required. *iSyf715* shows significant flows over the ferredoxin-NADP<sup>+</sup> reductase (reaction “\_1.18.1.2”) and across the ATPase (reaction “\_3.6.3.14”), the governing complex in the photophosphorylation process. Thus, *iSyf715* is flexible enough as to incorporate ATP/NADPH ratios, following the hypothesis that the existence of both electron transports must be essential for efficient photosynthesis [57].



“5.3.1.9a,b”), fructose 1,6-bisphosphate phosphatase (reaction “3.1.3.11”) and phosphoribulokinase (reaction “2.7.1.19”), as well as those that produce D-ribulose-1,5-bisphosphate and catalyzed by ribose-5-phosphate isomerase (reaction “5.3.1.6”) and ribulose-phosphate 3-epimerase (reaction “5.1.3.1”), are active, as should be in photoautotrophic conditions. Finally, we would like to focus on the high flux value of the reversible reaction “\_1.3.99.1” catalyzed by succinate dehydrogenase. In this case the direction implies the succinate oxidation to fumarate reducing ubiquinone. The synthesis of fumarate is an essential reaction since it represents an intermediate node in many metabolic pathways that yield building blocks for biomass formation; such as uridine monophosphate (UMP) and aspartate. It is known that this cyanobacterium does not have abundance of complex morphological characteristics [62], therefore, the reduced ubiquinone could be oxidized by other processes, such as photosynthesis electronic transfer, contributing to the formation of NADPH and ATP without triggering their classical synthesis reactions’ fluxes.

The genome-scale metabolic model of *Synechococcus elongatus* PCC7942, *iSyf715*, is available in Supplementary files S2, S3 and S4 and its results can be traced in Supplementary file S5.

### 3. Methods

#### 3.1. Genome-Scale Metabolic Network Reconstruction

Pathway Tools [29] and COPABI [30] software were used to construct a specific database of genes, proteins, enzymes and metabolites. All the genome and annotation files of *Synechococcus elongatus* PCC7942 were downloaded from the NCBI Entrez Genome [28] repository as of 24 February, 2011. The network was retrieved by Pathway Tools and was checked using COPABI filling gaps algorithms and different databases, like KEGG pathway database [31], MetaCyc [32], and BRENDA Enzyme database [37]. See Supplementary File S1 for a complete list.

*iSyf715* model on SBML format was deposited in BioModels Database [40] and assigned the identifier MODEL1407310000.

#### 3.2. Linear Programming for Flux Balance Analysis

All biochemical reactions in *iSyf715* were formulated as a stoichiometric model, such as:

$$S \cdot v$$

Here,  $S$  is a stoichiometric matrix that contains the stoichiometric coefficients from all internal (balanced) metabolites, whereas  $v$  is a flux vector that corresponds to the columns of  $S$ . Using an approach known as flux balance analysis or FBA [20], the former equation can be solved using linear programming and a given set of experimentally-driven constraints. In FBA the number of reactions is often larger than the number of metabolites, making the system underdetermined. Generally, in order to solve the problem and obtain a feasible solution for the intracellular fluxes, a steady state and an optimization criterion on metabolic balances has to be imposed. Steady state can be imposed so that:

$$S \cdot v = 0$$

In addition, the optimization criterion can be set by maximizing one biochemical reaction, such as the biomass evolving equation. Additionally, boundary conditions have to be used to be able to simulate the steady state.

For instance,

$$\begin{aligned} \text{Max } (v_j) \quad & \text{subject to } S \cdot v_j = 0 \quad \forall j \in N \\ & v_{j, \text{irr}} \in \mathbb{R}^+ \\ & v_{j, \text{rev}} \in \mathbb{R} \\ & v_{j, \text{const}} \in \mathbb{R}, v_{\min} < v_{j, \text{const}} < v_{\max} \\ & v_{j, \text{uptake}} \in \mathbb{R}, v_{\min} < v_{j, \text{uptake}} < v_{\max} \end{aligned}$$

Where  $v_j$  is the rate of the  $j^{\text{th}}$  reaction. The elements of the flux vector  $v$  were constrained for the definition of reversible and irreversible reactions,  $v_{j, \text{rev}}$  and  $v_{j, \text{irr}}$ , respectively. Additionally, two set of equations were established,  $v_{j, \text{const}}$ , constraining metabolic reactions, and  $v_{j, \text{uptake}}$ , uptake reactions, which were bound through experimentally determined values from the literature. Biomass synthesis was considered as a drain of precursors or building blocks into hypothetical biomass component. Flux through biomass synthesis reaction, being the biomass formation rate, is directly related to growth of the modeled organism. Simulations were performed using the OptGene software [63], later termed BioOpt [64,65]

More information about the modeling mechanics can be found in Montagud *et al.*, 2010 [11]. Readers will find in Supplementary file S2 the model file where all reactions and simulated constraints are depicted and ready to be used in OptGene software [63].

### 3.3. Biomass Composition

Biomass synthesis was considered as a drain of precursors or building blocks into a hypothetical biomass component. Precursors considered were amino acid, carbohydrates, chromophores, nucleic acids, and lipids. All data were retrieved from published information as described in Supplementary File S1. Flux through the biomass synthesis reaction, being the biomass formation rate, is directly related to growth of the modeled organism [20]. Table 4 shows the biomass composition that was considered in the current metabolic model and a detailed description can be found in Supplementary File S1.

## 4. Conclusions

We have successfully reconstructed the first genome-scale metabolic network for *Synechococcus elongatus* PCC7942, called *iSyf715*. The curated model represents an up-to-date database that encompasses all knowledge available in public databases, scientific publications and textbooks on the metabolism of this cyanobacterium. The model has been compiled in OptGene and SBML to enable its use with different software.

Our model includes 851 metabolic reactions and 838 metabolites, as well as the information from 715 genes. Moreover, we identified 76 enzymatic reactions needed for the correct function of the



metabolism, but with no annotated cognate gene. These genes are interesting targets for experimental studies as we have seen that their presence is required in order to build up the basic cellular components.

From the topological perspective the characteristics of the model are very similar to other published organisms' providing support for an evolutionary study of the structure and organizational properties of metabolic networks, in the line of recent works [66]. The connectivity analysis of the network model using the Pareto cumulative distribution shows scale-free behavior with a high non-uniformity and a hierarchical connectivity of the metabolites, which is typical of biological networks and points towards functional properties discussed in other works [46].

Flux balance analysis of the model was applied in order to simulate the autotrophic growth rate of the cyanobacteria. The *iSyf715* was able to simulate growth, which approached experimental values as well as characteristic reactions directions and split fluxes specific of this cyanobacterium. Moreover, typical flux characteristics of photosynthetic cells, such as active Calvin-Benson cycle or production of glycogen through the ADP-D-glucose synthesis were observed in our simulations.

In conclusion, the genome-scale metabolic network of *Synechococcus elongatus* PCC7942 (*iSyf715*) will be a valuable tool for industrial applications and fundamental research. Its use will allow the study and experimentation of this cyanobacterium as a possible light-driven production chassis for metabolites of industrial interest, in the line of *Synechocystis* sp. PCC6803 [11,12], *Cyanothece* sp. ATCC 51142 [67] or *Synechococcus* sp. PCC 7002 [68].

## Acknowledgments

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 308518 (CyanoFactory), from the Spanish Ministerio de Educación Cultura y Deporte grant FPU12/05873 through the program FPU and from the Universitat Politècnica de València grant Contratos Predoctorales FPI 2013.

## Supplementary Information

Supplementary materials can be accessed at: <http://www.mdpi.com/2218-1989/4/3/680/s1>.

- Supplementary File S1: Table S1, Table S2 and detailed description of biomass reaction.
- Supplementary File S2: *iSyf715* model with constraints ready to be simulated in OptGene/BioOpt.
- Supplementary File S3: *iSyf715* model with gene/reaction association in Excel.
- Supplementary File S4: *iSyf715* model in SBML.
- Supplementary File S5: *iSyf715* metabolic flux values.

## Author Contributions

Julián Triana, Daniel Gamermann, Arnau Montagud, Javier Torres, Jose Tena, Pedro Fernández de Córdoba and Javier F. Urchueguía designed the research. Julián Triana and Arnau Montagud performed the genome-scale metabolic network reconstruction. Julián Triana and Daniel Gamermann studied the network topology analyses. Julián Triana, Arantxa Urchueguía, Maria Siurana and David Fuente performed the model simulations. Javier Torres, Jose Tena, Pedro Fernández de Córdoba and Javier F. Urchueguía

conceived of the study. Julián Triana, Arnau Montagud, Arantxa Urchueguía and Javier F. Urchueguía drafted the manuscript. All authors contributed to, read and approved the final manuscript.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Shestakov, S.V.; Khyen, N.T. Evidence for genetic transformation in blue-green alga *Anacystis nidulans* R2. *Mol. Gen. Genet.* **1970**, *107*, 372–375.
2. Andersson, C.R.; Tsinoremas, N.F.; Shelton, J.; Lebedeva, N.V.; Yarrow, J.; Min, H.; Golden, S.S. Application of bioluminescence to the study of circadian rhythms in cyanobacteria. *Method. Enzymol.* **2000**, *305*, 527–542.
3. Waterbury, J.B.; Watson, S.W.; Valois, F.W.; Franks, D.G. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. *Can. Bull. Fish Aquat. Sci.* **1986**, *214*, 71–120.
4. Rippka, R.; Deruelles, J.; Waterbury, J.B.; Herdman, M.; Stanier, R.Y. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J. Gen. Microbiol.* **1979**, *111*, 1–61.
5. Scanlan, D.J.; Nyree, J.W. Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiol. Ecol.* **2000**, *40*, 1–12.
6. Ducat, D.C.; Way, J.C.; Silver, P.A. Engineering cyanobacteria to generate high-value products. *Trends Biotechnol.* **2011**, *29*, 95–103.
7. Montagud, A.; Gamermann, D.; Fernández de Córdoba, P.; Urchueguía, J.F. *Synechocystis* sp. PCC6803 metabolic models for the enhanced production of biofuels. *Crit. Rev. Biotechnol.* **2013**, *8551*, 1–15.
8. Snoep, J.L.; Bruggeman, F.; Olivier, B.G.; Westerhoff, H.V. Towards building the silicon cell: A modular approach. *BioSystems* **2006**, *83*, 207–216.
9. Papin, J.A.; Price, N.D.; Wiback, S.J.; Fell, D.A.; Palsson, B.Ø. Metabolic pathways in the post-genome era. *Trends Biochem. Sci.* **2003**, *28*, 250–258.
10. Szallasi, Z.; Periwal, V.; Stelling, J. In *Modeling in Cellular System Biology: From Concepts to Nuts and Bolts*; The MIT Press: Cambridge, MA, USA, 2006.
11. Montagud, A.; Navarro, E.; Fernández de Córdoba, P.; Urchueguía, J.F.; Patil, K.R. Reconstruction and analysis of genome-scale metabolic model of a photosynthetic bacterium. *BMC Syst. Biol.* **2010**, *4*, 156–172.
12. Montagud, A.; Zelezniak, A.; Navarro, E.; de Córdoba, P.F.; Urchueguía, J.F.; Patil, K.R. Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium *Synechocystis* sp. PCC6803. *Biotechnol. J.* **2011**, *6*, 330–342.
13. Park, J.M.; Kim, T.Y.; Lee, S.Y. Genome-scale reconstruction and *in silico* analysis of the *Ralstonia eutropha* H16 for polyhydroxyalkanoate synthesis, lithoautotrophic growth, and 2-methyl citric acid production. *BMC Syst. Biol.* **2011**, *5*, 101–112.

14. Milne, C.B.; Eddy, J.A.; Raju, R.; Ardekani, S.; Kim, P.-J.; Senger, R.S.; Jin, Y.-S.; Blaschek, H.P.; Price, N.D. Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052. *BMC Syst. Biol.* **2011**, *5*, 130–145.
15. Van den Hondel, C.A.; Verbeek, S.; Van der Ende, A.; Weisbeek, P.J.; Borrias, W.E.; van Arkel, G.A. Introduction of transposon Tn901 into a plasmid of *Anacystis nidulans*: Preparation for cloning in cyanobacteria. *Proc. Natl. Acad. Sci.* **1980**, *77*, 1570–1574.
16. Van der Plas, J. Identification of replication and stability functions in the complete nucleotide sequence of plasmid pUH24 from the cyanobacterium *Synechococcus* sp. PCC7942. *Mol. Microbiol.* **1992**, *6*, 653–664.
17. Chen, Y.C.; Holtman, K.C.; Magnuson, R.D.; Youderian, P.A.; Golden, S.S. The complete sequence and functional analysis of pANL, the large plasmid of the unicellular freshwater cyanobacterium *Synechococcus elongatus* PCC7942. *Plasmid* **2008**, *59*, 176–192.
18. Weise, S.; Grosse, I.; Klukas, C.; Koschützki, D.; Scholz, U.; Schreiber, F.; Junker, B.H. Meta-All: A system for managing metabolic pathway information. *BMC Bioinform.* **2006**, *7*, 465.
19. Förster, J.; Famili, I.; Fu, P.; Palsson, B.Ø.; Nielsen, J. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **2003**, *13*, 244–253.
20. Stephanopoulos, G.; Aristidou, A.A.; Nielsen, J.H. *Metabolic engineering: principles and methodologies*; Academic Press: San Diego, CA, USA, 1998.
21. Price, N.D.; Papin, J.A.; Schilling, C.H.; Palsson, B.Ø. Genome-scale microbial *in silico* models: The constraints-based approach. *Trends Biotechnol.* **2003**, *21*, 162–169.
22. Durot, M.; Bourguignon, P.Y.; Schachter, V. Genome-scale models of bacterial metabolism: Reconstruction and applications. *FEMS Microbiol. Rev.* **2009**, *33*, 164–190.
23. Price, N.D.; Reed, J.L.; Palsson, B.Ø. Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2004**, *2*, 886–897.
24. Orth, J.D.; Thiele, I.; Palsson, B.Ø. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248.
25. Thiele, I.; Palsson, B.Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **2010**, *5*, 93–121.
26. Feist, A.M.; Herrgard, M.J.; Thiele, I.; Reed, J.L.; Palsson, B.Ø. Reconstruction of biochemical networks in microorganisms. *Nat. Rev. Microbiol.* **2009**, *7*, 129–143.
27. Notebaart, R.A.; van Enckevort, F.H.; Francke, C.; Siezen, R.J.; Teusink, B. Accelerating the reconstruction of genome-scale metabolic networks. *BMC Bioinform.* **2006**, *7*, 296.
28. NCBI. Entrez Genome for *Synechococcus elongatus* PCC7942. Available online: [http://www.ncbi.nlm.nih.gov/genome/430?project\\_id=58045](http://www.ncbi.nlm.nih.gov/genome/430?project_id=58045) (accessed on 24 February 2011).
29. Karp, P.D.; Paley, S.; Romero, P. The Pathway Tools software. *Bioinformatics* **2002**, *18*(Suppl 1), S225–S232.
30. Reyes, R.; Gamermann, D.; Montagud, A.; Fuente, D.; Triana, J.; Urchueguía, J.F.; de Córdoba, P.F. Automation on the generation of genome-scale metabolic models. *J. Comput. Biol.* **2012**, *7*, 1295–1306.
31. Kanehisa, M.; Goto, S.; Furumichi, M.; Tanabe, M.; Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucl. Acids Res.* **2010**, *38*, 355–360.

32. Caspi, R.; Foerster, H.; Fulcher, C.A.; Karp, P.D. MetaCyc: A multiorganism database of metabolic pathways and enzymes. *Nucl. Acids Res.* **2006**, *34*, 511–516.
33. Voet, D.; Voet, J.G. *Biochemistry*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
34. Nelson D.L.; Cox M.M. *Lehninger Principles of Biochemistry*; Freeman, W.H., Ed.; W. H. Freeman: New York, NY, USA, 2008.
35. Berg, J.M.; Tymoczko, J.L.; Stryer, L. *Biochemistry*; W.H. Freeman & Palgrave MacMillan: Basingstoke, Hampshire, UK, 2011.
36. Mathews, C.K.; van Holde, K.E.; Appling, D.R.; Anthony-Cahill, S.J. *Biochemistry*; Prentice Hall: New York, NY, USA, 2012.
37. Chang, A.; Scheer, M.; Grote, A.; Schomburg, I.; Schomburg, D. BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucl. Acids Res.* **2008**, *36*, 190–195.
38. Schilling, C.H.; Letscher, D.; Palsson, B.Ø. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. *J. Theor. Biol.* **2000**, *203*, 229–248.
39. Pearce, J.; Carr, N.G. The metabolism of acetate by the blue-green algae, *Anabaena variabilis* and *Anacystis nidulans*. *J. Gen. Microbiol.* **1967**, *49*, 301–313.
40. Li, C.; Donizelli, M.; Rodriguez, N.; Dharuri, H.; Endler, L.; Chelliah, V.; Li, L.; He, E.; Henry, A.; Stefan, M.I.; *et al.* BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst. Biol.* **2010**, *4*, 92.
41. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512.
42. Albert, R.; Jeong, H.; Barabási, A.L. Error and attack tolerance of complex networks. *Nature* **2000**, *406*, 378–382.
43. Barabási, A.L.; Bonabeau, E. Scale-free networks. *Sci. Am.* **2003**, *288*, 50–59.
44. Barabási, A.L.; Oltvai, Z.N. Network biology understanding the cell's functional organization. *Nat. Rev. Genet.* **2004**, *5*, 101–113.
45. Feist, A.M.; Henry, C.S.; Reed, J.L.; Krummenacker, M.; Joyce, A.R.; Karp, P.D.; Broadbelt, L.J.; Hatzimanikatis, V.; Palsson, B.Ø. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* **2007**, *3*, 121.
46. Csete, M.; Doyle, J. Bow ties, metabolism and disease. *Trends Biotechnol.* **2004**, *22*, 446–450.
47. Hardy, M. “Pareto’s Law”. *Math. Intell.* **2010**, *32*, 38–43.
48. Newman, M.E.J. “Power laws, Pareto distributions and Zipf’s law”. *Contemp. Phys.* **2005**, *46*, 323–351.
49. Adamic, L.A.; Huberman, B.A. “Zipf’s law and the Internet”, *Glottometrics* **2002**, *3*, 143–150.
50. Wagner, A.; Fell, D.A. The small world inside large metabolic networks. *Proc. R. Soc. B.* **2001**, *268*, 1803–1810.
51. Rosales-Loaiza, N.; Ortega, J.; Mora, R. Influencia de la salinidad sobre crecimiento y composición bioquímica de la cianobacteria *Synechococcus* sp. *Cienc. Mar.* **2005**, *31*, 349–355.
52. Kajiwara, S.; Yamada, H.; Ohkuni, N.; Ohtaguchi, K. Design of the bioreactor for carbon dioxide fixation by *Synechococcus* PCC7492. *Energy Convers. Manag.* **1997**, *38*, 529–532.

53. Shastri, A.A.; Morgan, J.A. Flux balance analysis of photoautotrophic metabolism. *Biotechnol. Prog.* **2005**, *21*, 1617–1626.
54. Kuan, D. Growth optimization of *Synechococcus elongatus* PCC7942 in lab flask and 2D photobioreactor. Master Thesis, The University of British Columbia, Vancouver, BC, Canada, 2013. Available online: [https://circle.ubc.ca/bitstream/handle/2429/45010/ubc\\_2013\\_fall\\_kuan\\_david.pdf?sequence=1](https://circle.ubc.ca/bitstream/handle/2429/45010/ubc_2013_fall_kuan_david.pdf?sequence=1) (accessed on 14/08/14).
55. Lewis, N.E.; Hixson, K.K.; Conrad, T.M.; Lerman, J.A.; Charusanti, P.; Polpitiya, A.D.; Adkins, J.N.; Schramm, G.; Purvine, S.O.; Lopez-Ferrer, D.; *et al.* Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **2010**, *6*, 390.
56. Imam, S.; Yilmaz, S.; Sohmen, U.; Gorzalski, A.S.; Reed, J.L.; Noguera, D.R.; Donohue, T.J. iRsp1095: A genome-scale reconstruction of the *Rhodobacter sphaeroides* metabolic network. *BMC Syst. Biol.* **2011**, *5*, 116.
57. Munekaga, Y.; Hashimoto, M.; Miyaka, C.; Tomizawa, K.-I.; Endo, T.; Tasaka, M.; Shikanai, T. Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature* **2004**, *429*, 579–582.
58. Chen, Y.; Daviet, L.; Schalk, M.; Siewers, V.; Nielsen, J. Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Met. Eng.* **2013**, *15*, 48–54.
59. Shi, S.; Chen, Y.; Siewers, V.; Nielsen, J. Improving production of malonyl coenzyme A-derived metabolites by abolishing Snf1-dependent regulation of Acc1. *mBio* **2014**, *5*, e01130–e01140.
60. Krivoruchko, A.; Serrano-Amatriain, C.; Chen, Y.; Siewer, V.; Nielsen, J. Improving biobutanol production in engineered *Saccharomyces cerevisiae* by manipulation of acetyl-CoA metabolism. *J. Ind. Microbiol. Biotechnol.* **2013**, *40*, 1051–1056.
61. Buchanan, B.B.; Gruissem, W.; Jones, R.L. *Biochemistry and Molecular Biology of Plants*; American Society of Plants Physiologists: Poole, UK, 2001.
62. Robertson, B.R.; Tezuka, N.; Watanabe, M.M. Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 861–871.
63. Patil, K.R.; Rocha, I.; Förster, J.; Nielsen, J. Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinform.* **2005**, *6*, 308.
64. Cvijovic, M.; Olivares-Hernández, R.; Agren, R.; Dahr, N.; Vongsangnak, W.; Nookaew, I.; Patil, K.R.; Nielsen, J. BioMet Toolbox: genome-wide analysis of metabolism. *Nucleic Acids Res.* **2010**, *38 Suppl*, W144–W149.
65. BioOpt software. Available online: <http://biomet-toolbox.org/index.php?page=downtools-bioOpt> (accessed on 14 August 2014).
66. Gamermann, D.; Montagud, A.; Conejero, J.A.; Urcheguía, J.F.; de Córdoba, P.F. New approach for phylogenetic tree recovery based on genome-scale metabolic networks. *J. Comput. Biol.* **2014**, *21*, 508–519.
67. Vu, T.T.; Stolyar, S.M.; Pinchuk, G.E.; Hill, E.A.; Kucek, L.A.; Brown, R.N.; Lipton, M.S.; Osterman, A.; Fredrickson, J.K.; Konopka, A.E.; *et al.* Genome-scale modeling of light-driven reductant partitioning and carbon fluxes in diazotrophic unicellular cyanobacterium *Cyanothece* sp. ATCC 51142. *PLoS Comput. Biol.* **2012**, *8*, e1002460.

68. Hamilton, J.J.; Reed, J.L. Identification of functional differences in metabolic networks using comparative genomics and constraint-based models. *PLoS One* **2012**, *7*, e34670.

© 2014 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

## THE ROLE OF AGE AND EMOTIONAL VALENCE IN WORD RECOGNITION: AN EX-GAUSSIAN ANALYSIS

Amparo MORENO-CID<sup>2</sup>, Carmen MORET-TATAY<sup>1</sup>, Tatiana Quarti IRIGARAY<sup>3</sup>,  
Irani I. L. ARGIMON<sup>3</sup>, Michael MURPHY<sup>4</sup>, Marcin SZCZEBINSKI<sup>4</sup>,  
David MARTÍNEZ-RUBIO<sup>1</sup>, María José BENEYTO-ARROJO<sup>1</sup>,  
Esperanza NAVARRO-PARDO<sup>2</sup>, Pedro FERNÁNDEZ<sup>5</sup>

<sup>1</sup> Departamento de Neuropsicobiología, Metodología y Psicología Social  
Facultad de Psicología, Magisterio y Ciencias de la Educación  
Universidad Católica de Valencia San Vicente Mártir  
Sede de San Juan Bautista, 175, 46008, Valencia, Spain  
E-mail: carmenmoret@gmail.com

<sup>2</sup> Universidad de València, Spain

<sup>3</sup> Pontificia Universidad Católica de Río Grande del Sul, Brazil

<sup>4</sup> University College Cork, Ireland

<sup>5</sup> Universidad Politécnica de València, Spain

*Abstract:* The aim of this work is to evaluate the roles of age and emotional valence in word recognition in terms of ex-Gaussian distribution components. In order to do that, a word recognition task was carried out with two age groups, in which emotional valence was manipulated. Older participants did not present a clear trend for reaction times. The younger participants showed significant statistical differences in negative words for target and distracting conditions. Addressing the ex-Gaussian  $\tau$  parameter, often related to attentional demands in the literature, age-related differences in emotional valence seem not to have an effect for negative words. Focusing on emotional valence for each group, the younger participants only showed an effect on negative distracting words. The older participants showed an effect regarding negative and positive target words, and negative distracting words. This suggests that the attentional demand is higher for emotional words, in particular, for the older participants.

*Key words:* word recognition, emotional valence, ex-Gaussian components

In recent years, the scientific community has increased its interest in the influence that emotion can exercise on cognitive processes

such as memory and attention (Bradley, Greenwald, Petry, & Lang, 1992; Schmidt, 2012; Schmidt & Saari, 2007). Moreover, the literature shows a robust effect of emotional valence on memory. As indicated by Ferré (2003), the emotional effect on memory has been repeatedly proven using several types of stimuli.

---

### *Acknowledgments*

We would like to thank the anonymous reviewers and Reza Shah for their invaluable help and comments.

Focusing on the effect of emotional content on word recognition, the evidence shows more interference in emotionally valent (positive or negative) than neutral conditions for university students (Anderson, 2005; Dresler, M'eriau, Heekeren, & van der Meer, 2009; Pratto & John, 1991). On the other hand, emotional words are better remembered than neutral ones (Altarriba & Bauer, 2004; Brierley, Medford, Shaw, & David, 2007; Buchanan, Etzel, Adolphs, & Tranel, 2006; Ferré, 2003; Ferré, García, Fraga, Sánchez-Casas, & Molero, 2010; Herbert, Junghofer, & Kissler, 2008; Herbert & Kissler, 2010; Kensinger, 2008; Kensinger & Corkin, 2003; Kissler, Herbert, Peyk, & Junghofer, 2007; Schmidt & Saari, 2007). However, with aging, it seems that negative memories vanish more easily than positive memories (Abrisqueta-Gomez, Bueno, Oliveira, & Bertolucci, 2002; Berntsen & Rubin, 2002; Emery Hess, 2008); a clear difference as compared to younger participants in the literature. Mather and Cartensen (2003) explained this by claiming that older participants try to improve their well-being in terms of memory. Another explanation offered by Gross et al. (1997), proposes that people learn to regulate their negative emotions better as they get older.

Regarding the role of attention in the field of memory, the effect of emotion on memory may be mediated by attentional demands. Reisberg and Heuer (1995) argued that attention might be higher (as opposed to neutral) during the encoding process for emotional stimuli. Research reviewed above is based on reaction time data. They are typically analyzed using ANOVA (ANalysis Of VAriance). An alternative strategy (in terms of statistical analysis) to deal with this variable is by its characterization: one option

being an ex-Gaussian distribution fit. Luce (1986) has argued that the ex-Gaussian function provides a good fit to multiple empirical response time distributions. The ex-Gaussian function is the combination of two functions; a Gaussian (normal) and an exponential distribution. Several authors (Lemus-Zúñiga, Navarro-Pardo, Moret-Tatay & Pocinho, 2015; Moret-Tatay, Irigaray, Oliveira, & Argimon, 2015; Moret-Tatay et al., 2014; Moret-Tatay et al., in press; Navarro-Pardo, Navarro-Prados, Gamermann, & Moret-Tatay, 2013) have discussed the benefits of this approach in terms of statistical analysis. The ex-Gaussian distribution is characterized by three parameters,  $\mu$ ,  $\sigma$  and  $\tau$ . The  $\tau$  parameter is particularly interesting, as some have argued that it represents specific related cognitive processes such as attention. The study of Leth-Steensen, King Elbaz and Douglas (2000) is particularly relevant in this context. The research authors found differences in  $\tau$  parameter between children with ADHD and controls. They proposed several interpretations for this parameter, from a defective effort control mechanism role to developmental delay or processing problems. However, Matzke and Wagenmakers (2009) argued that the relationship of the ex-Gaussian distribution parameters to specific cognitive processes was not straightforward after a study of simulation. Further research along these lines is necessary. Nevertheless, there is a clear advantage of the ex-Gaussian distribution fit in that it allows the inclusion of all data without employing trimming or cut-off techniques which could leave interesting data out from a study.

The aim of the current study is to examine the roles of emotional valence and age in word recognition, in terms of ex-Gaussian components. To achieve this, a word recog-



dition task was carried out with older and younger participants. Stimuli were derived using a B-Pal, or Busca Palabras (Davis & Perea, 2005): a tool for deriving psycholinguistic properties of words. It is a Spanish language version of an earlier English language instrument (Davis, 2005). For our purposes, it provides information about emotional valence of words, among other aspects. We were interested in emotional valence, in terms of positive, negative and neutral. Reaction times were analyzed through the traditional analysis of variance and an alternative strategy: an ex-Gaussian distribution fit. It is expected that the ex-Gaussian modeling of the distribution of reaction time data will offer a superior description and will allow for more straightforward interpretation of the results.

### Method

#### Participants

A sample of 33 university students took part in experiment 1 (28 women and 5 men, with an average age of 21.15 years and  $SD = 1.60$ ); while a sample of 33 senior university students took part in experiment 2 (29 women and 4 men, with an average age of 65.39 years and  $SD = 5.20$ ). All the participants had normal or corrected to normal vision, were native Spanish speakers and did not report cognitive or neurological disorders.

#### Materials

The stimuli employed were a selection of words from the Busca Palabras database (Davis & Perea, 2005). A total of 90 words were divided into three sets of 30 stimuli, based on their scores on emotional valence (positive, negative or neutral); see Table 1. Employing the same classification as Moret-Tatay et al. (2014), stimuli rated 4 or lower were considered negative, rated 4 to 6 to be neutral, and rated above 6 to be positive. From the 90 words selected, 45 words were designated as target (and appeared in the first and second block) and the other 45 as distractors (and appeared only in the second block).

#### Procedure

Participants were tested in a quiet room, in groups of three or four. The presentation of stimuli and recording of response times were controlled by a Windows operating system through the DMDX software (Forster & Forster, 2003). The experiment consisted of two phases. In the first phase, 45 target stimuli were randomly presented (divided into 15 stimuli for each of the three valence categories) with short exposures of 2 seconds each. In the second phase (15 minutes after the participants were distracted by performing Stroop tasks), 45 target stimuli plus the 45 distracting stimuli were randomly pre-

Table 1 *Average valence for the selected words in the different sets (standard deviation in parenthesis)*

		Neutral	Negative	Positive
Valence	Target	4.77 (0.20)	2.86 (0.62)	6.62 (0.53)
	Distracting	4.89 (0.40)	2.84 (0.78)	7.20 (0.62)

sented. Each word was presented until the participant gave a response or 2000 ms passed. The participants were instructed to press a button (labelled "Yes") if the stimulus was a target stimulus, and press another button (labelled "No") if the stimulus was a distractor stimulus. The participants were also instructed to respond as quickly as possible while maintaining a reasonable level of accuracy. The session lasted approximately 30 minutes.

#### Statistical Design and Data Analysis

Valence (neutral, negative and positive) was manipulated as within group variables. Two different analysis procedures were carried out: a classical ANOVA and a fitting of the data to an ex-Gaussian distribution function. The ex-Gaussian distribution fit was

carried out by data sets, which were distributed in intervals in order to create a histogram. Differences between parameters from the ex-Gaussian distribution fit were analyzed regarding their uncertainties (errors) as confidence interval lengths for each parameter.

#### Results

The statistical analysis was performed using SPSS 20. Table 2 presents the average reaction times (ms), error rates and standard deviation for each group of words, and for each age group.

In the ANOVA for latency analyses, RTs less than 250 ms and over 1800 ms were excluded (less than 2% of the data set). The 1800 ms cut-off point was adopted for consistency with earlier studies in the field (Moret-Tatay & Perea, 2011; Perea, Moret-

Table 2 Response time averages (ms), error rate and standard deviation (SD) for different experimental conditions

Images	Neutral	Negative	Positive
Younger			
Target	847.13	822.99	841.07
SD	112.22	92.53	137.60
Errors	24%	26%	26%
Distracting			
Target	937.29	897.83	932.68
SD	116.81	129.84	152.16
Errors	26%	16%	14%
Older			
Target	923.65	923.24	888.78
SD	176.24	177.02	144.02
Errors	23%	32%	21%
Distracting			
Target	1029.72	1006.72	1022.23
SD	197.47	209.50	199.33
Errors	28%	26%	33%

Tatay, & Panadero, 2011; Perea, Panadero, Moret-Tatay, & Gómez, 2012). A mixed 2 X 2 X 3 ANOVA was conducted, with valence a repeated measure and both age group and target/distractor as independent subject factors. The ANOVA was also carried out on both older and younger participants' correct RTs.

The ANOVA of the global data set showed that differences in emotional valence were close to but did not reach statistical significance,  $F_{(2,128)} = 2.63$ ;  $MSE = 6018.46$ ;  $\eta^2 = 0.04$ ;  $p = 0.08$ . In the case of distracting and target conditions, target stimuli were processed faster ( $M = 874.48$  ms) than distractors ( $M = 971.08$  ms), and these differences reached statistical significance:  $F_{(1,64)} = 41.30$ ;  $MSE = 22364.27$ ;  $\eta^2 = 0.39$ ;  $p < 0.05$ . It is concerning that the analyses of variance revealed significant differences in variability of response times between younger and older participants. Homogeneity of variance across populations is one of the assumptions of analysis of variance. Therefore, a violation of this assumption calls into question any conclusions drawn from this analysis. This is why ANOVA for both groups was conducted separately.

For younger university students, the ANOVA on the latency data showed that differences in valence were statistically significant,  $F_{(2,64)} = 4.43$ ;  $MSE = 4316.57$ ;  $\eta^2 = 0.12$ ;  $p < 0.05$ . The neutral stimuli ( $M = 892.21$  ms) were processed slower than the positive ones ( $M = 886.84$  ms) which, in turn were processed slower than the negatives ones ( $M = 860.41$  ms). In the case of distracting and target conditions, differences also reached statistical significance:  $F_{(1,32)} = 23.83$ ;  $MSE = 15195.46$ ;  $\eta^2 = 0.42$ ;  $p < 0.05$ . Target stimuli were processed faster ( $M = 837.06$  ms) than distractors ( $M = 922.60$  ms).

Any interaction or error differences did not reach statistical significance ( $F < 1$ ).

The ANOVA of the latency of the older participants' data, for valence, did not reach the significance level ( $F < 1$ ). In the case of distracting and target conditions, these differences reached statistical significance:  $F_{(1,32)} = 19.42$ ;  $MSE = 29533.88$ ;  $\eta^2 = 0.38$ ;  $p < 0.05$ . Target stimuli were processed faster ( $M = 911.89$  ms) than distractors ( $M = 1019.56$  ms). Any interaction or error differences did not reach statistical significance ( $F < 1$ ).

Secondly, we proceeded to characterize the reaction times using an ex-Gaussian distribution fit. The distribution is completely characterized by the three parameters  $\mu$ ,  $\sigma$  and  $\tau$ . The parameters  $\mu$  and  $\sigma$  are the average and standard deviation of the Gaussian component, while  $\tau$  is the parameter that characterizes the exponential component. One should bear in mind that  $\mu$  and  $\sigma$  are not the average and standard deviation of the ex-Gaussian distribution, which should be calculated via the three parameters that describe the distribution:  $M = \mu + \tau$ .

To fit the distribution is to find the optimal values for the parameters  $\mu$ ,  $\tau$  and  $\sigma$  that best describe the experimental data. For this purpose, we used the fitting function of the open-source software Gnuplot. With this software, the fit of any mathematical function to any data set can be obtained straightforwardly by a single function call, but given the amount of data in the present work, the need to prepare the data (distribute it in intervals) and the fact that many different data sets had to be fitted, a python script was employed. This script automatically reads a set of data (reaction times), groups this data in intervals, creating a histogram and interacts with the Gnuplot software in order to fit

an ex-Gaussian function to the data points. Both, distributions fits and graphics, were carried out by the command-line program GNU plot 4.2 (via the python script published in Navarro-Pardo et al., 2013). The Gnuplot software employs the Levenberg-Marquardt algorithm, also known as damped least-square method (Marquardt, 1963), in order to find the optimal parameters that minimize the square of the difference between a given data set ( $x_i$  and  $y_i$ ) and a target function which depends on some parameters. The algorithm is an interactive procedure that readjusts the set of parameters in each interaction. First of all one should define a goodness of fit func-

tion, in order to reflect the quality of the fit. The goodness of fit can be evaluated through the residual variance (the most widely used method in behavioral sciences,  $\chi^2/\text{degrees of freedom}$ ). Smaller values are preferable and show a better fit. Table 3 shows the different parameters obtained by the fitting procedure and Figures 1 and 2 show the graphical representation of the histograms together with their fit, for each condition.

If we focus on the uncertainties (errors – defined as a confidence interval length for each parameter) from Table 3, we can compare the parameters for the different condi-

Table 3  $\mu$ ,  $\sigma$  and  $\tau$  parameters with their uncertainty in parenthesis (error), dof (degrees of freedom) and the goodness of fit by  $\chi^2/\text{dof}$  for younger and older university students

Task/Parameters	$\mu$	$\sigma$	$\tau$	dof	$\chi^2/\text{dof}$
<i>Younger</i>					
Target					
Neutral	622.86 ± 12.34	96.17 ± 9.84	211.09 ± 17.46	26	0.63
Negative	615.48 ± 14.81	86.55 ± 12.02	194.42 ± 20.74	21	1.01
Positive	593.41 ± 9.05	77.80 ± 7.72	232.38 ± 14.89	23	0.45
Distracting					
Neutral	663.06 ± 23.53	109.42 ± 17.56	289.18 ± 41.94	27	1.37
Negative	681.14 ± 13.57	98.07 ± 11.17	210.38 ± 19.39	28	0.89
Positive	691.57 ± 18.63	98.79 ± 15.96	222.12 ± 27.36	25	1.59
<i>Older</i>					
Target					
Neutral	650.90 ± 18.12	126.63 ± 13.85	265.63 ± 27.79	25	0.75
Negative	688.87 ± 39.34	175.28 ± 28.15	184.63 ± 47.39	27	1.44
Positive	607.34 ± 15.66	91.34 ± 13.79	282.97 ± 27.92	27	1.04
Distracting					
Neutral	722.78 ± 25.91	131.42 ± 23.26	345.89 ± 45.22	27	1.16
Negative	798.02 ± 39.27	147.29 ± 26.78	144.35 ± 45.89	27	1.90
Positive	818.08 ± 51.88	189.21 ± 34.32	164.01 ± 62.59	26	1.60

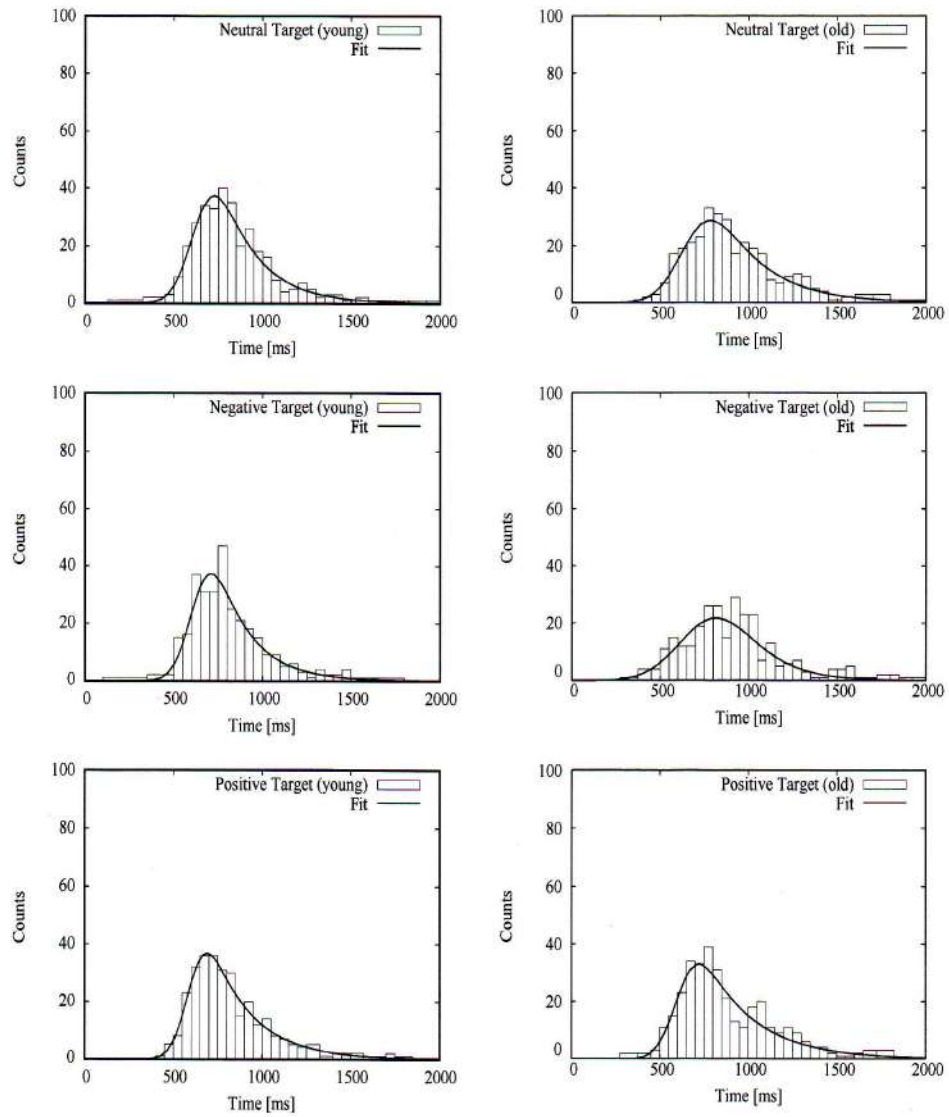


Figure 1. Emotional word recognition graphics for younger and older university students in the target condition

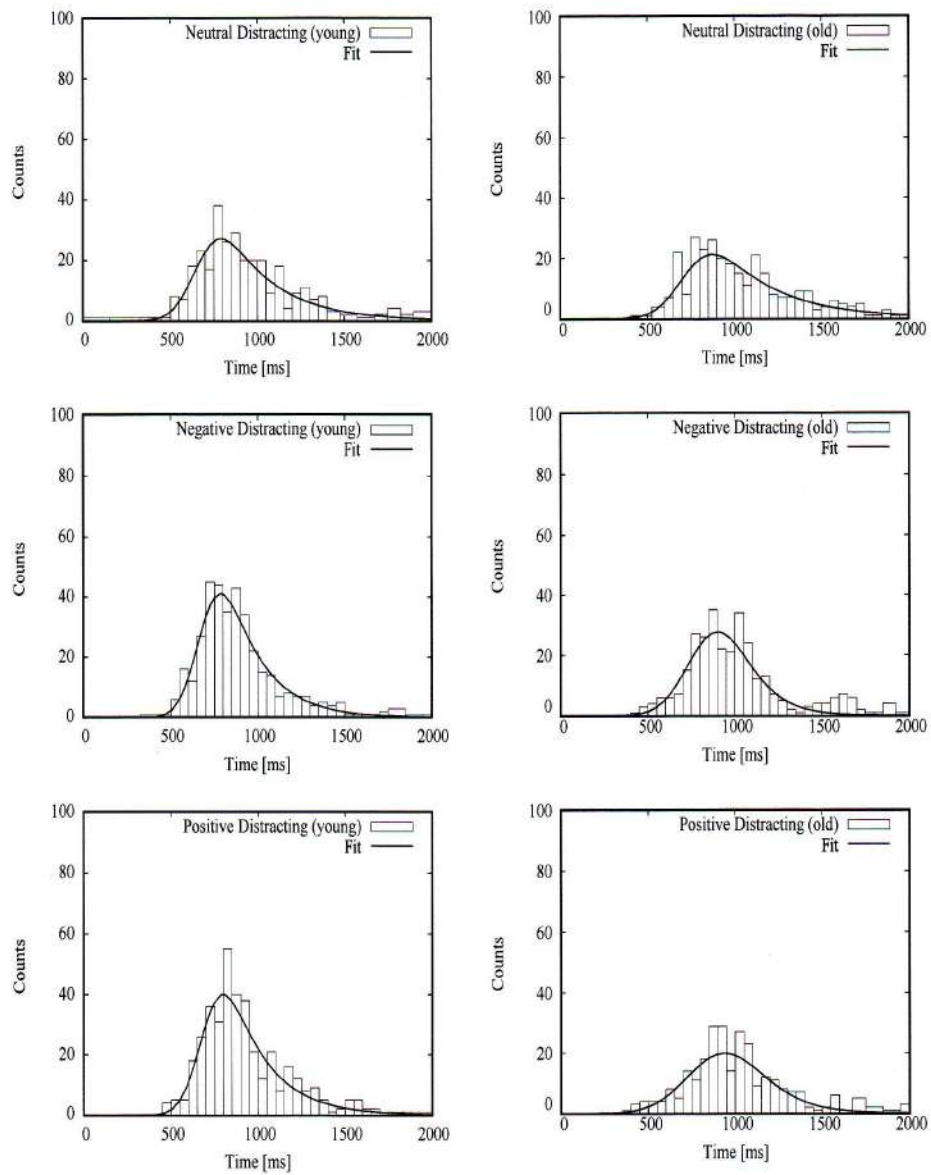


Figure 2. Emotional word recognition graphics for younger and older university students in the distracting condition

tions. If we compare the  $\tau$  parameter differences in conditions by age, we find differences in positive targets between younger and older participants. Here, the differences between older and younger participants in this condition (50.59) were higher than the uncertainties sum (42.81). Furthermore, it is also possible to find differences in the neutral target condition between younger and older participants for  $\tau$  parameter. In this case, the differences between older and younger participants in this condition (54.54) were also higher than the uncertainties sum (45.25).

Moving on to the differences between levels of emotional valence, the  $\tau$  parameter seems not to have special properties for the younger group, except between the negative and neutral distracting conditions (78.8), which was higher than the uncertainties sum (61.33). Older participants showed a difference between the negative and neutral target conditions (81), which also was higher than the uncertainties sum (75.18). Nevertheless, the role of valence was clear for the distracting words in this group: the difference between negative and neutral condition was 201.54, and its uncertainties sum was 91.11, while the difference between positive and neutral conditions was 181.81 and its uncertainties sum was 107.81.

### Discussion

The aim of the current study is to examine the roles of emotional valence and age in word recognition, in terms of ex-Gaussian components. To achieve that, a recognition task, employing a short retention interval, was carried out. This recognition task was analyzed not only by the classical analysis of variance, but also through the character-

ization of the reaction times via an ex-Gaussian distribution fit, which allows the analysis of the conditions in terms of parameters. The logic of this analysis is to reexamine the existing hypotheses cited in the literature in terms of ex-Gaussian parameters, and in particular, the  $\tau$  parameter often related to attention.

Younger participants showed shorter RTs to negative words for target and distracting condition. Older participant RTs were shorter for neutral and negative target condition and positive distracting condition. However, the results only reached the statistical significance level for younger participant RTs. Thus, the emotional effect was clear for younger students but it was not for the older participants. During the traditional analysis some data was trimmed to minimize the effect of outliers. This is a popular technique in Cognitive Psychology. However, it is not clear where it is better to set the limit between what is noise and what is signal. An alternative could be an ex-Gaussian fit, where all data set is employed. In particular, the differences in this second analysis seemed clearer for the emotional words than in case of the traditional one.

In terms of parameters, age-related differences in emotional valence seem not to have special properties for negative words. Focusing on emotional valence for each group, the younger participants only showed an effect for the negative condition in the distracting words. This pattern was clear for the older participants. They showed an effect for negative target words, plus positive and negative distracting words.

Finally, if we interpret the  $\tau$  parameter as an indicator of attentional demands, these results seem to reinforce the idea that emotional stimuli have a remarkable effect on at-

tention, and thereby, on other processes, such as memory.

### Conclusions

Several studies claim that emotional words might capture more attention than neutral ones (Algom, Chajut, & Lev, 2004; Huang, Baddeley, & Younger, 2008; Phelps et al., 1997). The RTs for distracting negative words were similar to the neutral ones in the case of the older participants. These results are consistent with Mather and Carstensen (2003), reinforcing the idea of age differences in the processing of negative and positive information (Carstensen et al., 2000). Another possible explanation (Gross et al., 1997) could be related to how people learn to regulate their negative emotions, which might be more effective for older than younger adults. Comparing word recognition to image recognition, some research (León et al., 2010; Moret-Tatay et al., 2014) found slower latencies towards negative stimuli valence employing the IAPS battery.

However, addressing the question of the ex-Gaussian components, the negative words seem not to have special properties for age. This suggests that there are no age differences for this valence. Focusing on emotional valence globally (negative and positive words), the pattern is clearer for the older than the younger group. It suggests that the role of valence has a higher effect for the older adults.

This work also presents potential limitations. There are remarkably high error rates, which indicates that the task might be difficult for participants. For future research it would be interesting to develop easier series of experiments that might also examine the role of arousal in terms of parameters,

and possible interactions with valence. Future studies may include a series of experiments that would examine the role of  $\tau$  parameter in the context of negative images, which would try to explain if such stimuli attract attention more readily or just hold attention for a longer period of time.

Received January 7, 2014

### References

- Abrisqueta-Gomez, J., Bueno O. F. A., Oliveira, M. G., Bertolucci P. H. (2002). Recognition memory for emotional pictures in Alzheimer's patients. *Acta Neurologica Scandinavica*, 105(1), 51-54.
- Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, 133, 323-338.
- Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *The American Journal of Psychology*, 117, 389-410.
- Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology: General*, 134, 258-281.
- Berntsen, D., & Rubin, D. C. (2002). Emotionally charged autobiographical memories across the life span: The recall of happy, sad, traumatic and involuntary memories. *Psychology and Aging*, 17(4), 636-652.
- Brierley, B., Medford, N., Shaw, P., & David, A. S. (2007). Emotional memory for words: Separating content and context. *Cognition and Emotion*, 21, 495-521.
- Bradley, M. M., Greenwald, M. K., Petry, M. C., & Lang, P. J. (1992). Remembering pictures: Pleasure and arousal in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(2), 379-390.
- Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of autonomic arousal and semantic relatedness on memory for emotional words. *International Journal of Psychophysiology*, 61, 26-33.
- Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. (2000). Emotion experience in



- the daily lives of older and younger adults. *Journal of Personality and Social Psychology*, 79, 644-655.
- Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37, 65-70.
- Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods*, 37(4), 665-671.
- Dresler, T., Mériaux, K., Heekeren, H. R., & van der Meer, E. (2009). Emotional Stroop task: Effects of word arousal and subject anxiety on emotional interference. *Psychological Research*, 73, 364-371.
- Emery, L., Hess, T. M. (2008). Viewing instructions impact emotional memory differently in older and young adults. *Psychology and Aging*, 23, 2-12. doi: 10.1037/0882-974.23.1.2.
- Ferré, P. (2003). Effects of level of processing on memory for affectively valenced words. *Cognition and Emotion*, 17, 859-880.
- Ferré P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition and Emotion*, 24, 760-785.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, and Computers*, 35, 116-124.
- Gross, J. J., Carstensen, L. L., Pasupathi, M., Tsai, J., Skorpen, C. G., & Hsu, A. Y. C. (1997). Emotion and aging: Experience, expression, and control. *Psychology of Aging*, 12, 590-599.
- Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, 45, 487-498.
- Herbert, C., & Kissler, J. (2010). Motivational priming and processing interrupt: Startle reflex modulation during shallow and deep processing of emotional words. *International Journal of Psychophysiology*, 76, 64-71.
- Huang, Y. M., Baddeley, A., & Younger, A. W. (2008). Attentional capture by emotional stimuli is modulated by semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 328-339.
- Kensinger, E. A. (2008). Age differences in memory for arousing and nonarousing emotional words. *Journal of Gerontology*, 63B, 13-18.
- Kensinger, E. A., Brierley, B., Gowdon, J. H., & Corkin, S. (2002). Effects of normal aging and Alzheimer's disease on emotional memory. *Emotion*, 2(2), 118-134.
- Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory and Cognition*, 31, 1169-1180.
- Kensinger, E. A., & Schacter, D. L. (2006). Processing emotional pictures and words: Effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, 6, 110-126.
- Kissler, J., Herbert, C., Peyk, P., & Junghofer, M. (2007). Early cortical responses to emotional words during reading. *Psychological Science*, 18, 475-480.
- Lemus-Zúñiga, L. G., Navarro-Pardo, E., Moret-Tatay, C., & Pocinho, R. (2015). Serious games for elderly continuous monitoring. In *Data Mining in Clinical Medicine* (pp. 259-267). Springer New York.
- León, F. G., Martínez, J. M. A., Hernández, L. M., Cruz, J. S., Meilán, J. J. G., Ramos, J. C., & Sáez, E. P. (2010). Emoción y memoria de reconocimiento: La discriminación de la información negativa como un proceso adaptativo. *Psicothema*, 22(4), 765-771.
- Leth-Steensen, C., King Elbaz, Z., & Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica* 104, 167-190.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2), 431-441. doi:10.1137/0111030
- Mather, M., & Carstensen, L. L. (2003). Aging and attentional biases for emotional faces. *Psychological Science*, 14(5), 409-415.
- Matzke, D., & Wagenmakers, E. J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review*, 16(5), 798-817.
- Moret-Tatay, C., Irigaray, T. Q., Oliveira, C., & Argimon, I. I. L. (2015). Reaction Times as a dependent variable of memory: Future lines of research through an ex-Gaussian fit. *Horizons in Neuroscience Research*, 17.

- Moret-Tatay, C., Moreno-Cid, A., Argimon, I. I. D. L., Quarti Irigaray, T., Szczerbinski, M., Murphy, M., Vázquez-Martínez, A., Vázquez-Molina, J., Sáiz-Mauleon, B., Navarro-Pardo, E., & Fernández de Córdoba Castellá, P. (2014). The effects of age and emotional valence on recognition memory: An ex-Gaussian components analysis. *Scandinavian Journal of Psychology, 55*(5), 420-426.
- Moret-Tatay, C., Navarro-Pardo, E., Fernández-Muñoz, J.J., Vázquez-Martínez, A., Vázquez-Molina, J., & Civera-Mollá, C. (in press). The role of feedback on learning: Evidence through ex-Gaussian components. *Procedia - Social and Behavioral Sciences*.
- Moret-Tatay, C., & Perea, M. (2011). Is the go/no-go lexical decision task preferable to the yes/no task with developing readers? *Journal of Experimental Child Psychology, 110*(1), 125-132.
- Navarro-Pardo, E., Navarro-Prados, A. B., Gamermann, D., & Moret-Tatay, C. (2013). Differences between younger and older university students on lexical decision task: Evidence through an ex-Gaussian approach. *Journal of General Psychology, 140*(4), 251-268.
- Perea, M., Moret-Tatay, C., & Panadero, V. (2011). Suppression of mirror generalization for reversible letters: Evidence from priming. *Journal of Memory and Language, 65*(3), 237-246.
- Perea, M., Panadero, V., Moret-Tatay, C., & Gómez, P. (2012). The effects of inter-letter spacing in visual-word recognition: Evidence with younger normal readers and developmental dyslexics. *Learning and Instruction, 22*(6), 420-430.
- Phelps, E. A., LaBar, K. S., & Spencer, D. D. (1997). Memory for emotional words following unilateral temporal lobectomy. *Brain and Cognition, 35*, 85-109.
- Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention grabbing power of negative social information. *Journal of Personality and Social Psychology, 61*, 380-391.
- Reisberg, D., & Heuer, F. (1995). Emotion's multiple effects on memory. In J. L. McGaugh, N. M. Weinberger, & G. Lynch (Eds.), *Brain and memory: Modulation and mediation of neuroplasticity* (pp.84-92) New York: Oxford University Press.
- Schmidt, S. R. (2012). Memory for emotional words in sentences: The importance of emotional contrast. *Cognition & Emotion, 26*(6), 1015-1035. doi:10.1080/02699931.2011.63198
- Schmidt, S. R., Saari, B. (2007). The emotional memory effect: Differential processing or item distinctiveness? *Memory & Cognition, 35*(8), 1905-1916.

#### VPLYV VEKU A EMOČNEJ VALENCIE NA ZNOVUPOZNÁVANIE SLOV: EX-GAUSSOVA ANALÝZA

C. Moret-Tatay, A. Moreno-Cid, T. Q. Irigaray, I. I. L. Argimon, M. Murphy, M. Szczerbinski, D. Martínez-Rubio, M. J. Beneyto-Arrojo, E. Navarro-Pardo, P. Fernández

**Súhrn:** Cieľom práce je posúdiť vplyv veku a emočnej valencie na znovupoznávanie slov v rámci ex-Gaussových distribučných komponentov. Dvomi vekovými skupinám sme administrovali test znovupoznávania slov, v ktorom sme manipulovali emočnou valenciou. U mladších respondentov sa prejavili štatisticky významné rozdiely pri negatívnych slovách v experimentálnej podmienke a v podmienke s distrakciou. U starších respondentov sme v odpovedových časoch nezistili jasnú tendenciu. Vzhľadom na ex-Gaussovský parameter  $\tau$ , ktorý sa v literatúre často spája s nárokmi na pozornosť, vekovo podmienené rozdiely v emočnej valencii nemali žiaden vplyv na negatívne slová. Ak sa zameriame na emočnú valenciu v oboch skupinách, u mladších respondentov sa prejavil vplyv v negatívnych distrakčných slovách. U starších respondentov sa prejavil vplyv v negatívnych a pozitívnych cieľových slovách a negatívnych distrakčných slovách. Z toho vyplýva, že nároky na pozornosť sú vyššie pre emočne nabitá slová, najmä u starších respondentov.

# TRABAJO DE FIN DE GRADO

Grado en Biotecnología

Curso 2018-2019

## Caracterización de los factores de transcripción que regulan la expresión de *AHA2* en *Arabidopsis thaliana*



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Autora: Mireia Usó Agost  
Tutor: Eduardo Bueso Ródenas  
Cotutor: Gaetano Bissoli

Tipo Licencia: Licencia Creative Commons “Reconocimiento no Comercial –Sin Obra Derivada”



València, Julio 2019

## RESUMEN

QQS es un gen huérfano que tiene un papel relevante en el metabolismo del almidón incrementando el ratio nitrógeno/carbono, un rasgo agronómicamente muy interesante que se puede conseguir sobreexpresando QQS en arroz, soja y maíz (Li et al., 2009, Li et al., 2015). Además, QQS está relacionado con la protón ATPasa de la membrana plasmática (en *Arabidopsis*, AHA), responsable del crecimiento celular y la toma de nutrientes. Micromatrices de pérdida de función de las dos principales isoformas de AHA, *aha1-6* y *aha2-4*, muestran niveles bajos de mRNA de QQS (Haruta et al., 2012). Una posterior publicación asigna a QQS la tarea de facilitar la entrada al interior del núcleo el factor de transcripción NF-YC4. Este pertenece a la familia NF-YCs, los cuales se comportan como inhibidores transcripcionales de la bomba de protones. Posteriormente, analizando las secuencias de zonas promotoras de los AHAs se pone de manifiesto que AHA2 tiene 5 cajas CCAAT típicas de los NF-Ys, mientras que AHA1 tiene solo un par de ellas. El análisis de expresión mediante qRT-PCR de líneas de ganancia de función de QQS y NF-YC4 (NF-YC4 OE) muestran inhibición específica de la expresión del gen AHA2. En las mismas condiciones, la pérdida de función de QQS y el triple mutante *nf-yc3,4,9* tienen aumentadas su expresión. El factor de activación clave de ese gen se llama PTM, que procede de la pared del cloroplasto y reconoce una región de 63 pares de bases. El análisis de expresión del mutante *ptm2* (At5g22760), un gen de la familia del ya descrito PTM, muestra una expresión reducida de AHA2, mientras que en líneas constitutivas con una versión truncada muestran sobreexpresión de la misma. Todo esto, nos ha llevado a comprobar si PTM2 reconoce la caja del promotor AHA2 mediante un ensayo de One Hybrid en levadura.

**Palabras clave:** H<sup>+</sup>-ATPasa (AHA2), QQS, PTM2, NF-YC, Yeast One-Hybrid.

## ABSTRACT

QQS is a gene that has a relevant role in the starch metabolism by increasing the nitrogen / carbon ratio, an agronomically interesting feature that can be reached by overexpressing QQS in rice, soybeans and corn (Li et al., 2009, Li et al., 2015) . In addition, QQS is related to the proton ATPase of the plasma membrane (in *Arabidopsis*, AHA), responsible for cell growth and nutrient intake. Microarrays of loss of function of the two main isoforms of AHA, *aha1-6* and *aha2-4*, show low levels of QQS mRNA (Haruta et al., 2012). A subsequent publication assigns to QQS the task of facilitating the entry into the core of the transcription factor NF-YC4. This is a member of the NF-YCs family, which behave as transcriptional inhibitors of the proton pump. Subsequently, analyzing the sequences of the AHAs promoters, it is revealed that AHA2 has 5 CCAAT boxes typical of the NF-Ys, while AHA1 has only a couple of them. qRT-PCR expression analysis of QQS and NF-YC4 function gain lines (NF-YC4 OE) indicates specific inhibition of AHA2 gene expression. Under the same conditions, the loss of the QQS function and the triple mutant *nf-yc3,4,9* have increased expression. The activation factor of that gene is called PTM, which is derived from the reduction of chloroplast and a region of 63 base pairs is recognized. The analysis of the expression of the mutant *ptm2* (At5g22760), a gene of the family of the already described PTM, shows a reduced expression of AHA2, while in the constitutive lines with a truncated version show the overexpression of AHA2. Because of that, we are checking if PTM2 recognizes the AHA2 promoter box by means of a One Hybrid test in yeast.

**Keywords:** H<sup>+</sup>-ATPase (AHA2), QQS, PTM2, NF-YC, Yeast One-Hybrid.

## **AGRADECIMIENTOS**

A Gaetano, por enseñarme tantísimas cosas en el laboratorio, por contagiarme la pasión por la biología molecular y por nunca dejar de creer en mí.

A Edu, por ayudarme hasta el último día tanto con este trabajo como con mis decisiones profesionales.

A Joan y a todos los demás con los que he compartido laboratorio por la ayuda, los ánimos y el buen ambiente de trabajo.

A mi familia y a todos mis amigos, en especial a Nerea y a Sergi, por su apoyo más incondicional y sincero.

# ÍNDICE

<b>1. INTRODUCCIÓN</b> .....	1
<b>1.1. AHA2 (Autoinhibited H<sup>+</sup>-ATPase)</b> .....	1
1.1.1. Funciones fisiológicas.....	2
<b>1.2. Factores de regulación de AHA a nivel postraduccional</b> .....	3
<b>1.3. Factores de regulación de AHA a nivel transcripcional</b> .....	4
1.3.1. NF-YC.....	4
1.3.2. QQS.....	5
1.3.3. PTM2.....	6
<b>1.4. One Hybrid</b> .....	7
<b>1.5. Antecedentes</b> .....	8
<b>2. OBJETIVOS</b> .....	9
<b>3. MATERIALES Y MÉTODOS</b> .....	10
<b>3.1. Material biológico</b> .....	10
3.1.1. <i>Arabidopsis thaliana</i> .....	10
3.1.2. <i>Escherichia coli</i> .....	10
3.1.3. <i>Saccharomyces cerevisiae</i> .....	10
<b>3.2. Vectores de clonación</b> .....	11
<b>3.3. Manipulación y crecimiento de <i>E. Coli</i></b> .....	14
3.3.1. Preparación de <i>E. coli</i> competentes y transformación.....	15
<b>3.4. Manipulación y crecimiento de <i>S. cerevisiae</i></b> .....	15
3.4.1. Preparación de <i>S. cerevisiae</i> competentes y transformación.....	15
<b>3.5. Medios de cultivo</b> .....	15
3.5.1. Murashige y Skoog (MS) y solución nutritiva.....	15
3.5.2. Luria-Bertani (LB).....	16
3.5.3. SOC (Super Optimal Broth).....	16
3.5.4. YPDA (Yeast extract-Peptone Dextrosa Adenina).....	16
3.5.5. Medio sintético SD.....	16
<b>3.6. Primers</b> .....	17
<b>3.7. Reacción en cadena de la Polimerasa (PCR)</b> .....	18
<b>3.8. Extracción RNA, conversión a cDNA y qRT-PCR</b> .....	18
<b>3.9. Electroforesis</b> .....	19
<b>3.10. Recombinación</b> .....	19
<b>3.11. Minipreparaciones de plásmidos</b> .....	19
<b>3.12. Digestiones</b> .....	19
<b>3.13. Ligaciones</b> .....	20
<b>3.14. Secuenciación</b> .....	20
<b>3.15. One Hybrid</b> .....	20

3.15.1. Construcción PTM2 y transformación en levadura THY.AP4.....	20
3.15.2. Construcción promotor AHA2 y transformación Y187.....	21
3.15.3. Ensayo Yeast One-Hybrid (Y1H).....	21
<b>4. RESULTADOS.....</b>	<b>22</b>
<b>4.1. Regulación de AHA2 por NF-YC y QQS.....</b>	<b>22</b>
4.1.1. Cuantificación de la expresión de AHA1, AHA2 y AHA3.....	22
4.1.2. Cuantificación de la expresión de AHA1, AHA2 en condiciones de ayuno y tras tratamiento con sacarosa.....	23
<b>4.2. Regulación de AHA2 por PTM2.....</b>	<b>24</b>
4.2.1. Expresión de AHA2 en mutantes de PHD.....	24
4.2.2. Clonaje de promotor de AHA2 en pHIPPIE y pHIS2.....	25
4.2.3. Clonaje de PTM2 en pGADT7-GW.....	26
4.2.4. Yeast One-Hybrid (Y1H).....	29
<b>5. DISCUSIÓN.....</b>	<b>31</b>
<b>5.1. Regulación negativa de AHA2.....</b>	<b>31</b>
<b>5.2. Regulación positiva de AHA2.....</b>	<b>31</b>
<b>6. CONCLUSIONES.....</b>	<b>33</b>
<b>7. BIBLIOGRAFÍA.....</b>	<b>34</b>



## Índice de Figuras

<b>Figura 1</b> Esquema de la H <sup>+</sup> -ATPasa.....	1
<b>Figura 2</b> Formación del complejo de los NF-YA, B y C según Myers & Holt, 2018.....	5
<b>Figura 3</b> Modelo de inducción de QQS según Li et al., 2015.....	6
<b>Figura 4</b> Componentes del ensayo Y1H.....	7
<b>Figura 5</b> Esquema del plásmido pDONR/Zeo.....	11
<b>Figura 6</b> Esquema del plásmido pGEM®-T Easy.....	12
<b>Figura 7</b> Esquema del plásmido pHIS2.....	13
<b>Figura 8</b> Esquema del mecanismo de integración de pHIPPIE en el genoma de levadura.....	13
<b>Figura 9</b> Esquema del plásmido pGADT7-GW con los sitios de restricción.....	14
<b>Figura 10</b> qRT-PCR de AHA1,2 y 3.....	22
<b>Figura 11</b> qRT-PCR de AHA1 y 2 tras tratamiento con sacarosa.....	23
<b>Figura 12</b> qRT-PCR de AHA2 con líneas mutantes de PTM.....	24
<b>Figura 13</b> Gel de electroforesis tras digerir y tratar los vectores con fosfatasa alcalina.....	25
<b>Figura 14A.</b> Gel de electroforesis de la PCR para comprobar el sentido del inserto en pHIPPIE y PHIS2.....	26
<b>Figura 14B.</b> Gel de electroforesis de comprobación de la integración del vector pHIPPIE en el genoma de levadura.....	26
<b>Figura 15A.</b> Reacción BP.....	27
<b>Figura 15B.</b> Electroforesis en gel de agarosa de la digestión parcial del plásmido pDONR/Zeo con el gen PTM.....	27
<b>Figura 16A.</b> Reacción Gateway LR entre el vector pDONR/Zeo y el pGADT7-GW.....	28
<b>Figura 16B.</b> Gel de agarosa de la digestión con HindIII y EcoRV de pGADT7 de <i>E.Coli</i> transformadas.....	28
<b>Figura 17</b> Ensayo final del Y1H.....	30

## Índice de abreviaturas

H<sup>+</sup>-ATPasa, AHA2: protón ATPasa  
SOS1: Salt Overly Sensitive 1  
QQS : Qua Quine Starch  
NF: Nuclear Factor  
PHD: Plant Homeodomain  
FLC: FLOWERING LOCUS C  
CO: CONSTANS  
Y1H: Yeast One-Hybrid  
AD: Activation Domain  
BD: Binding Domain  
MCS: MultiCloning Site  
PCR: Polymerase Chain reaction  
qRT-PCR: Quantitative Real Time PCR

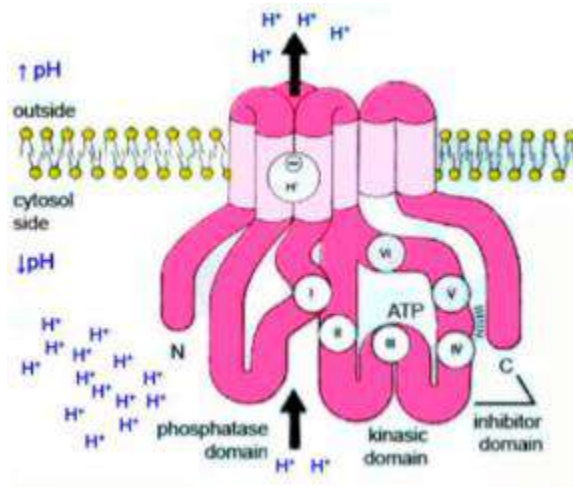
# 1. INTRODUCCIÓN

## 1.1.AHA2 (Autoinhibited H<sup>+</sup>-ATPase)

La protón ATPasa (H<sup>+</sup>-ATPasa) es una proteína transmembrana que funciona como bomba de exportación de protones (Duby & Boutry, 2008). Esta ATPasa crea un gradiente electroquímico de protones que conduce a la toma de nutrientes y regula el crecimiento mediante la modulación del pH intracelular (neutro/básico) y extracelular (ácido). (**Figura 1**) (Morsomme & Boutry, 2000).

La H<sup>+</sup>-ATPasa pertenece a la extensa familia de las ATPasas P. En *Arabidopsis* hay 11 isoformas de H<sup>+</sup>-ATPasas, llamadas desde AHA1 a AHA11 (Yuan et al., 2017), las cuales suelen expresarse en tipos celulares específicos. Un ejemplo es AHA3, con gran importancia en el desarrollo del polen (Robertson et al., 2004). No obstante AHA1 y AHA2 son las H<sup>+</sup>-ATPasas más expresadas en toda la planta. Estas dos participan en la señalización celular modificando el pH citosólico cuando hay estrés abiótico y biótico.

Estas dos isoformas son las más importantes ya que el doble mutante de estas no es viable. Por otra parte, mutantes simples de cada una de las isoformas (*aha1-6* y *aha 2-4*) inducen la expresión del mismo gen, QQS (Qua Quine Starch) (Haruta & Sussman, 2012).



**Figura 1.** Estructura de la H<sup>+</sup>-ATPasa. La H<sup>+</sup>-ATPasa en un estado poco activo donde expulsa al apoplasto un limitado número de protones manteniendo un pH citosólico ligeramente alcalino (7,25).

Se conoce que esta proteína de membrana está codificada por aproximadamente 10 genes, aunque difiere según las especies (Morsomme & Boutry, 2000).

AHA2 está constituida por 948 residuos de aminoácidos, los cuales se pliegan de manera que conforman 10 segmentos en forma de alfa hélice hidrofóbicas (**Figura 1**). La actividad catalítica se encuentra en la parte citoplasmática (dominios A, N y P) donde además hay dos dominios terminales hidrofóbicos que participan en la regulación. El dominio C-terminal, de unos 100 residuos aproximadamente, inhibe la bomba de protones interfiriendo con estos dominios catalíticos. (Palmgren et al., 1991). Dependiendo de los dominios terminales, la H<sup>+</sup>-ATPasa estará en dos estados: autoinhibido, donde la hidrólisis de ATP y el transporte de H<sup>+</sup> están acoplados débilmente, e inducido, donde esta hidrólisis de ATP y transporte de H<sup>+</sup> están fuertemente acoplados (Falhof et al., 2016).

Además de en el transporte activo de protones y el consecuente transporte de otros compuestos, esta proteína participa en muchas otras funciones esenciales para las plantas.

### **1.1.1. Funciones fisiológicas** (Morsomme & Boutry, 2000)

#### **-Transporte secundario**

La H-ATPasa es la activadora principal del transporte secundario (transportadores y canales) ya que genera el gradiente electroquímico necesario para poder absorber los compuestos que la célula necesita a través de la membrana plasmática. La mayoría de estos sistemas son de simporte, como en el caso de la toma de potasio por canales tipo Shaker, aunque también encontramos antiportadores como el de H<sup>+</sup>/Na<sup>+</sup> tipo SOS1 (Falhof et al., 2016).

#### **-Salinidad**

La sal es tóxica para las plantas, y por ello, estas han desarrollado varios mecanismos para evitar su acumulación utilizando los transportadores secundarios.

Las plantas no tienen bombas de Na<sup>+</sup> o Na<sup>+</sup>/K<sup>+</sup> ATPasa como tienen los animales, aunque las células vegetales sean más sensibles a la salinidad del medio que los animales. Sin embargo, las plantas sí son capaces de expulsar al medio el Na<sup>+</sup> mediante antiportadores Na<sup>+</sup>/H<sup>+</sup> (SOS1 en *Arabidopsis*), que dependen directamente del gradiente que genera la H<sup>+</sup>ATPasa a través de la membrana plasmática.

La compartimentalización y la extrusión (expulsión) de los iones tóxicos como las sales, en la vacuola se produce como respuesta a la acumulación de estos mismos iones en el citosol.

### **-Apertura de estomas**

Los estomas son unos pequeños orificios que se encuentran en las hojas de las plantas y son los responsables del intercambio de gases ( $O_2$ ,  $CO_2$  y  $H_2O$ ). Están rodeados por dos células oclusivas, que controlan la apertura de los orificios mediante los cambios en su turgencia regulados a su vez por la hiperpolarización de la membrana plasmática y la acidificación del medio extracelular por parte de la  $H^+$ -ATPasa (Ilan, 1994). La función de la  $H^+$ -ATPasa es proporcionar el potencial electroquímico necesario para el movimiento de los iones a través de las membranas y controlar la apertura y cierre de los canales de iones (como el de  $Na^+$ ), lo cual explica porqué la  $H^+$ -ATPasa está altamente expresada en estas células que rodean los estomas.

### **-Regulación pH intracelular y elongación celular**

La teoría del crecimiento ácido dicta que los protones expulsados por la  $H^+$ -ATPasa bajan el pH del apoplasto. Esto provoca que se debiliten los puentes de hidrógeno de las microfibrillas de la pared celular vegetal y que se desplacen por la presión intracelular aumentando el volumen celular, por lo que la activación de la  $H^+$ -ATPasa promueve la elongación de las células y consecuentemente el crecimiento de la planta (Dubey & Boutry, 2008). Además de esto, la acidificación del apoplasto induce también la toma de  $K^+$  por la hiperpolarización de la membrana, lo cual produce cambios osmóticos permitiendo que haya un flujo de agua a través de las acuaporinas, favoreciendo la elongación celular (Morsomme & Boutry, 2000).

## **1.2. FACTORES DE REGULACIÓN DE AHA A NIVEL POSTRADUCCIONAL**

Las modificaciones post-traduccionales como las fosforilaciones tienen un papel muy importante en muchos tipos de funciones celulares. Se conoce que alrededor de un tercio de todas las proteínas de eucariotas se fosforilan en algún momento para regular su actividad, estabilidad u organización espacial. Las fosforilaciones son un paso clave en la transducción de señales por cascada y son mediadas por kinasas (Rudashevskaya et al., 2011).

En general, hay varios sitios de activación e inhibición por fosforilación, que dependen principalmente del extremo C-terminal y de la posición donde se dan estas fosforilaciones (Falhof et al., 2016). No obstante, no todos los sitios de fosforilación están determinados.

Según la literatura (Haruta et al., 2018), hay varias fosforilaciones en la H<sup>+</sup>-ATPasa, no obstante, la más estudiada ha sido la del penúltimo residuo, en la Treonina 947 (Thr-947). Esta fosforilación estabiliza la H<sup>+</sup>-ATPasa creando un sitio de unión para proteínas 14-3-3, que se unen fuertemente a otras proteínas para estabilizar estructuras proteicas (Rosenquist et al., 2000).

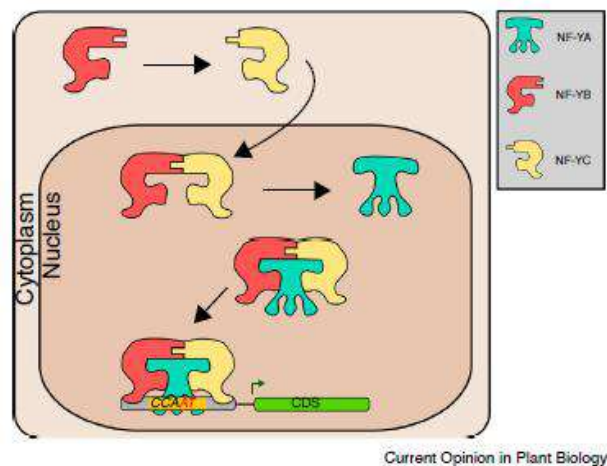
A pesar de que existen muchos estudios de regulación postraducciona de AHA2, a nivel de regulación transcripcional no existe apenas literatura.

### **1.3. FACTORES DE REGULACIÓN DE AHA A NIVEL TRANSCRIPCIONAL**

#### **1.3.1. NF-YC**

Los “Nuclear Factor-Y” (NF-Y) están involucrados en muchos aspectos de la fisiología de las plantas como en el tiempo de floración o en la respuesta al estrés (Myers & Holt, 2018). Estos NF-Y son subunidades de 3 familias proteicas distintas altamente conservadas en eucariotas (Kumimoto et al., 2010): NF-YA, NF-YB y NF-YC. Se conoce que estos tres factores forman un complejo heterotrimérico que reconoce la secuencia consenso CCAAT de la zona promotora de los genes que regula y se une a ella. No obstante, hay más proteínas o complejos proteicos que se unen a estas cajas CCAAT, es decir, que hay otros factores que contribuyen a la unión de los NF-Y. El factor NF-YA es el que reconoce la secuencia y se une a ella y los NF-YB y NF-YC tienen unos dominios específicos de unión entre ellos y con la tercera subunidad de NF-YA (**Figura 2**).

Según Kumimoto et al., 2010, NF-YC3, NF-YC4 y NF-YC9 interactúan con CONSTANS, un factor de transcripción que regula el tiempo de floración. CONSTANS interactúa con estos NF-YC para activar al FLOWERING LOCUS T (FT), provocando la reducción del tiempo de floración. Estos tres NF-YC también están relacionados con la señalización de fitohormonas como el ácido abscísico y las giberelinas (Myers & Holt, 2018).



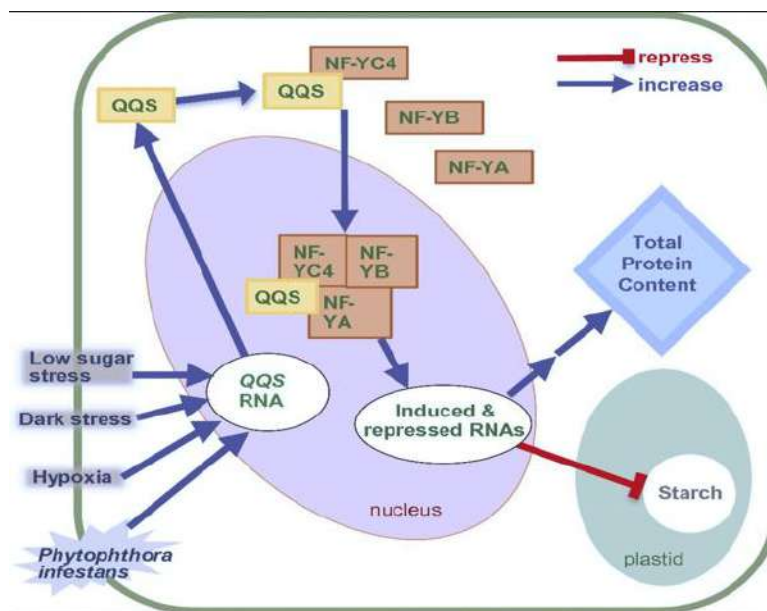
**Figura 2.** Formación del complejo de los NF-YA, B y C según Myers & Holt, 2018. La unión de los NF-YB y NF-YC ocurre en el citoplasma y seguidamente este complejo entra en el núcleo para unirse a su vez con NF-YCA. Así, estos tres factores reconocen las secuencias CCAAT de los genes a regular.

### 1.3.2. QQS

El gen Qua-Quine Starch (QQS) es un gen huérfano, es decir, que solo se encuentra en una especie, en este caso *Arabidopsis thaliana*, cuya función se relaciona con la inhibición de la síntesis de almidón. La sobreexpresión de este gen en *Arabidopsis* aumenta el contenido total proteico de la célula y disminuye el de almidón.

QQS incrementa el ratio nitrógeno/carbono, un rasgo agronómicamente muy interesante que se puede conseguir sobreexpresándolo en arroz, soja y maíz (Li et al., 2009, Li et al., 2015).

En Li et al., 2015 se demostró, aunque no se conoce el mecanismo molecular todavía, que QQS, que se encuentra en el citosol, interacciona con el factor NF-YC4 formando un complejo que se desplaza al núcleo y modula la expresión de diferentes genes, dado que en ausencia de QQS, NF-YC4 no realiza su tarea específica (**Figura 3**).



**Figura 3.** Modelo de inducción de QQS según Li et al., 2015. El QQS citosólico forma un complejo con NF-YC4, se desplaza al núcleo y modula la expresión de diferentes genes.

### 1.3.3. PTM2

PTM es un factor de transcripción del tipo PHD (“plant homeodomain”) con dominios transmembrana del cloroplasto que reconoce una región de 63 pares de bases y se activa por la rotura proteolítica transmitiendo múltiples señales retrógradas hacia el núcleo. Los “PHD zinc fingers” son estructuras conservadas en proteínas que modifican la cromatina, así como también median interacciones moleculares en la transcripción de genes (Sanchez & Zhou, 2011).

Se sabe que PTM libera su fragmento N-terminal para asociarse con la cromatina y suprimir la transcripción del FLOWERING LOCUS C (FLC). Esto en definitiva hace que la planta florezca antes, ya que el FLC es un represor de la floración. Además el promotor de FLC tiene una caja muy parecida a la que tiene AHA2 (Feng et al., 2016).

El gen At5g22760 también es un gen de la familia de los PHD pero que no está descrito. En este trabajo se denomina PTM2 por la similitud con el gen PTM.



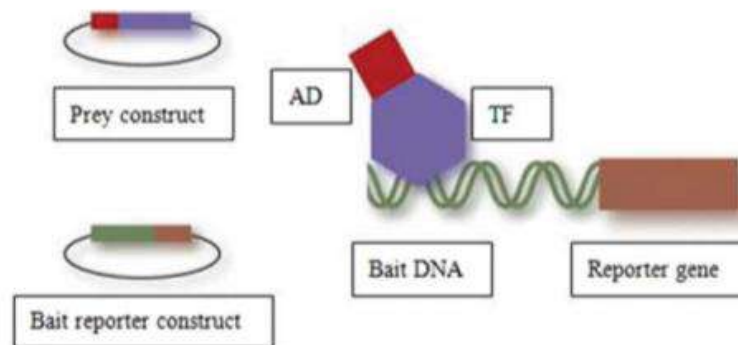
## 1.4. YEAST ONE-HYBRID (Y1H)

La “yeast one-hybrid” (Y1H) es una técnica derivada de la “yeast two-hybrid” que sirve para analizar interacciones DNA-proteína gracias a la expresión de genes reporteros. En este sistema, se detectan interacciones en sistemas eucariotas, lo que permite el estudio de factores de transcripción en animales y en plantas. Estos factores de transcripción se conforman normalmente de dos dominios: uno que se une al DNA (DNA-binding domain o BD) y otro que activa la transcripción (activation domain o AD) (Sun et al., 2011).

El sistema Y1H está formado por dos componentes (**Figura 4**, Sun et al., 2011): la “presa” (prey) y el “cebo” (bait). La llamada “presa” es la proteína híbrida que se forma con nuestra proteína de interés conjugada con el factor GAL4, y el “cebo” será el DNA al que se unirá la “presa”, que en caso de que esta interacción ocurra, expresará el gen reportero.

Ambos componentes están codificados en dos vectores diferentes que se introducirán en una misma levadura. Entonces si nuestro factor de transcripción codificado en un vector, se une al DNA codificado por el otro vector, el AD inducirá la expresión de la proteína reportera. Cabe señalar que esta técnica solo nos informará de si ocurre la interacción de la proteína con el DNA, no de si este factor es activador o inhibidor, ya que el AD de la levadura activa igualmente al gen reportero.

Las cepas utilizadas de levaduras para este ensayo pueden tener varias mutaciones en genes auxotróficos. Los más comunes son el triptófano (*trp1*), la leucina (*leu2*), la histidina (*his3*) y el uracilo (*ura3*), así como el gen *lacZ*, que también se usa para seleccionar las colonias con las construcciones deseadas mediante los fenotipos según los medios de selección.



**Figura 4.** Componentes del ensayo Y1H. Arriba se ve el plásmido presa que da como producto el factor de transcripción de interés (TF) unido al dominio de activación (AD). Abajo, el vector que da lugar al DNA de estudio y el gen reportero que se expresará en el caso de que esta interacción TF-DNA ocurra.

## **1.5. ANTECEDENTES**

Buscando indicios de la regulación de la H<sup>+</sup>ATPasa por factores de transcripción, se probaron los miembros de la familia PHD, entre los cuales solo el At5g22760 parecía regular la expresión. FLC es regulado por PTM, por lo que es posible que haya una interacción con la caja de este gen, y se sabe que el promotor de FLC tiene una caja muy parecida a la que tiene AHA2 (Feng et al., 2016).

A este gen se le denominó PTM2, por su similitud con el ya descrito PTM (Sun et al., 2011).

## 2. OBJETIVOS

**Objetivo general:** estudiar factores de transcripción que puedan regular la expresión de AHA2.

**Objetivos específicos:**

- Estudiar la expresión de AHA2 en mutantes de pérdida de función *ptm2* para conocer si es un activador o inhibidor de AHA2.
- Comprobar si existe interacción directa entre este factor PTM2 y el promotor AHA2.

## 3. MATERIAL Y MÉTODOS

### 3.1. MATERIAL BIOLÓGICO

#### 3.1.1. *Arabidopsis Thaliana*

Para estos ensayos, se utilizó el ecotipo Columbia 0 (Col-0) de la planta *Arabidopsis thaliana*. El triple mutante de *Arabidopsis nf-yc3,4,9* ha sido concedido por el laboratorio de Ben F. Holt III (Kunimoto et al., 2010). La pérdida de función (de microRNA artificial) amiQQS viene del laboratorio de Wurtele (Li et al., 2009). La sobreexpresión de NF-YC4 viene de la colección de factores de transcripción TRANSPLANTA (Coego et al., 2014) y pedida con los códigos N2102545 y TPT\_5.63470.1A al NASC (Nottingham, UK). Siempre se pidieron del NASC la línea de pérdida de función *ptm* (At5g35210), en concreto pedimos la línea N666425 correspondiente a SALK\_073799C (homocigota). La línea de pérdida de función *ptm2* (At5g22760) identificada como N843295 y SAIL\_1170\_D01 (homocigota).

Para el cultivo in vitro de las semillas, se utilizaron placas del medio sólido MS (Murashige y Skoog, 1962), se sellaron con cinta micropore y crecieron a 23°C en condiciones de día largo.

Tanto el factor de transcripción PTM2 como el fragmento del promotor de *AHA2* clonados para el ensayo Y1H provienen del transcriptoma y del genoma de *Arabidopsis thaliana* respectivamente.

#### 3.1.2. *Escherichia Coli*

Se utilizó la cepa DH5α de *Escherichia Coli* para manejar los genes clonados en distintos vectores plasmídicos.

#### 3.1.3. *Saccharomyces cerevisiae*

Para el ensayo de Y1H se han utilizados dos cepas haploides de *Saccharomyces cerevisiae* cada una con una construcción y luego gracias a los determinantes mating MATα y MATa se ha conseguido juntarlas para obtener la copresencia de la construcción cebo y presa en levadura diploide Y187 y THY.AP4.

La Y187 es un MATα que presenta el siguiente genotipo: *ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4Δ, met-, gal80Δ, MEL1, URA3::GAL1UAS- GAL1TATA-lacZ*. Debido a estas mutaciones, la cepa incluye las siguientes auxotrofías: triptófano, leucina, histidina, adenina y metionina. Esta cepa es la que se utilizó para ser transformada con el plásmido

pHIS2 con el promotor de *AHA2* y *pHIS2* plásmido vacío como control. Se utilizó la misma cepa para transformar e integrar el plásmido pHIPPIE en el genoma de levadura tanto la construcción con el promotor de *AHA2* y la construcción sin promotor. Y la cepa *THY.AP4* lleva *determinante sexual MATa* y ha sido transformada con *pGALT7-PTM2* y lleva *ura3, leu2, lexA::lacZ::trp1, lexA::HIS3, lexA::ADE2* (Obrdlik et al., 2004).

### 3.2. VECTORES DE CLONACIÓN

Se utilizaron diversos vectores para clonación y subclonación del promotor de *AHA2* y del factor de transcripción *PTM2*.

**-pDONR/Zeo:** El vector pDONR/Zeo (ThermoFisherScientific™)(Figura 5) es un vector con la tecnología Gateway diseñado para generar clones con el gen de interés flanqueado con unos extremos attL, que se recombinarán con otro clon o producto de PCR flanqueado con extremos attB. Después de esta reacción BP, se podrá recombinar con otros vectores con tecnología Gateway con una reacción de recombinación LR. El vector contiene además unos terminadores T1 y T2 para proteger el gen clonado, unos primers M13 en sentido y en antisentido para la secuenciación del inserto, los sitios de recombinación anteriormente mencionados, el gen *ccdB* para la selección negativa, el origen de replicación pUC para *E.coli*, y resistencias a cloramfenicol y kanamicina o zeocina.

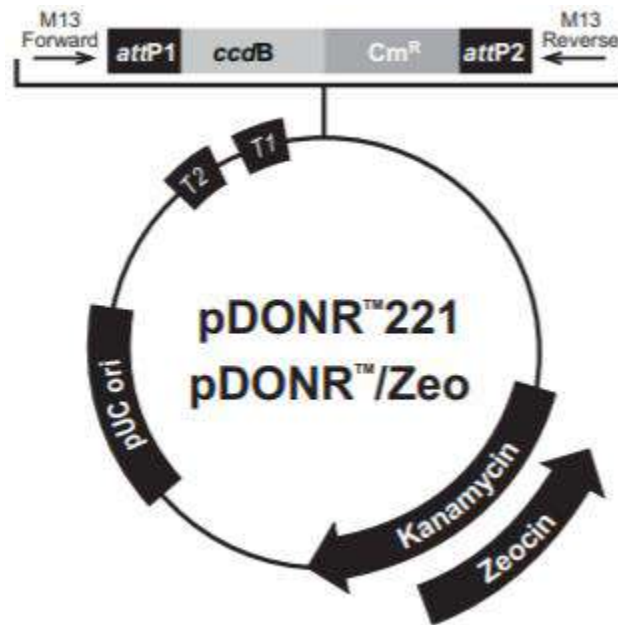
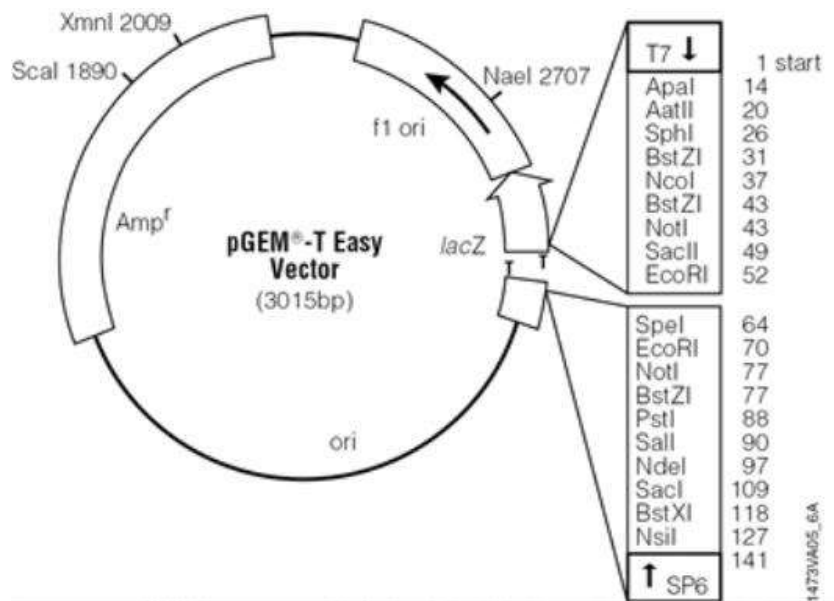


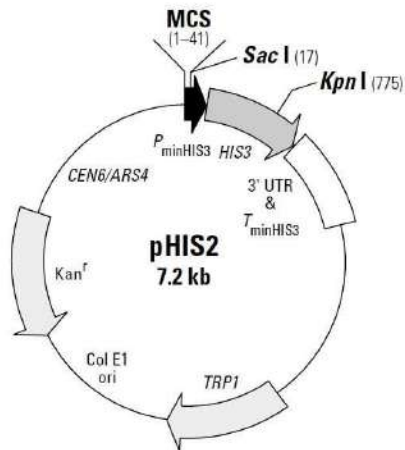
Figura 5. Esquema del plásmido pDONR/Zeo.

**-pGEM®-T Easy:** El vector pre-linealizado pGEM®-T Easy (Promega) se ha usado para subclonar el promotor de AHA2 entre -186 y -253 par de bases antes del codón de inicio ATG. Flanqueando el sitio de inserción se encuentran múltiples sitios de restricción (**Figura 6**) y además contiene extremos 3'-T para que sea compatible con los con los productos de PCR que añaden a los extremos una adenina como ocurre por ejemplo con la Taq polimerasa. Otra característica relevante de este plásmido es que las colonias que han incorporado un inserto en la reacción de ligación interrumpirán la región codificante para el péptido  $\alpha$  (LacZ) para la  $\beta$ -galactosidasa. Las colonias recombinantes serán blancas, mientras que las que no hayan captado el inserto serán azules, por la expresión de este péptido. Este vector también contiene un origen de replicación para *E.coli*, un gen de resistencia a ampicilina y un par de promotores para hacer posible la secuenciación del inserto (T7 y SP6).



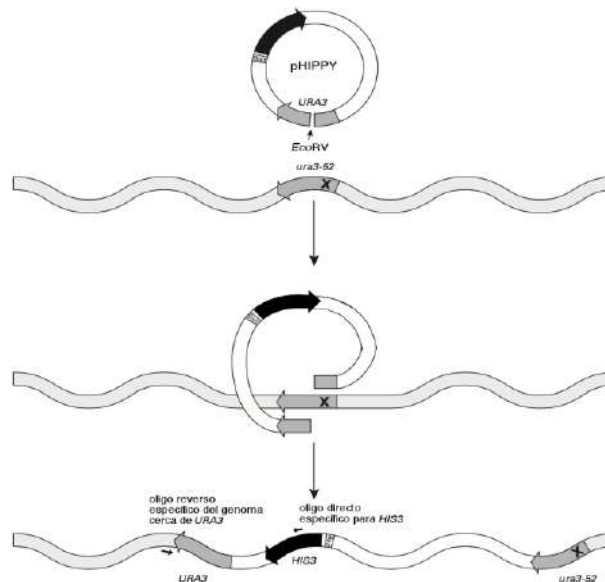
**Figura 6** . Esquema del Plásmido pGEM®-T Easy. Al lado derecho, el listado sitios de restricción que contiene el multiple cloning site (MCS).

**- pHIS2:** Se trata de un vector diseñado para ensayos de Y1H (**Figura 7**) que permite clonar fácilmente la zona promotora del gen del interés, en nuestro caso AHA2. Además, contiene un gen reportero HIS3 sin su promotor nativo, pero con un promotor mínimo y un sitio de múltiple conación para permitir el inserto de las secuencias de interés. Este plásmido lleva dos orígenes de replicación, uno para *E.coli* (Col E1) y uno para *S. cerevisiae* (ARS4/CEN6), además de un gen de resistencia a kanamicina y el gen de la auxotrofia para Triptófano (TRP1).



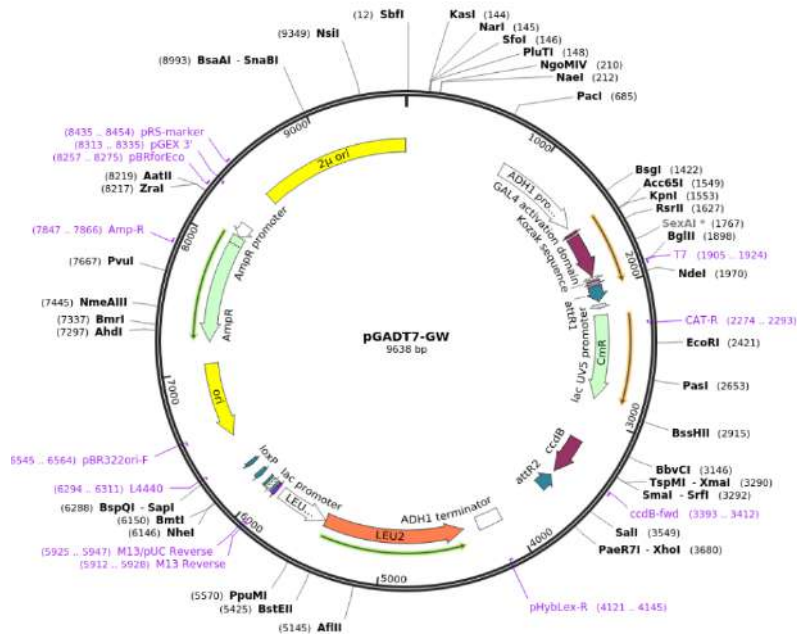
**Figura 7.** Esquema del vector pHIS2. Se muestran las partes detalladas y algunos sitios de cortes de restricción (Sac I, KpnI).

- **pHIPPIE:** Este plásmido es una variante del plásmido anterior pHIS2 (Clonetech). Es plásmido tiene el gen de auxotrofía para uracilo (URA3) en lugar del de triptófano para poder integrarse mediante recombinación homóloga (de la mutación *ura3-52*) al genoma de la levadura transformada y sirve para hacer el ensayo Y1H (**Figura 8**). Para conseguir este cambio, cortamos el vector pHIS2 con *EcoRI* y *EcoRV* y el plásmido pCM259 con *EcoRI* y *PmeI* y los ligamos. En el fragmento de pHIS2, nos queda el reportero HIS3 y en el fragmento de pCM259 nos queda URA3 y la resistencia a ampicilina/carbenicilina.



**Figura 8.** Esquema del mecanismo de integración del plásmido pHIPPIE. Esto ocurre mediante las regiones de URA3 del plásmido y *ura3-52* del genoma de levadura.

- **pGADT7-GW**: Es un plásmido con la tecnología Gateway® (**Figura 9**) que sirve como vector “presa” en ensayos Yeast Two/One Hybrid. Esto es porque contiene el dominio GAL4 de activación que se pega a la proteína recombinante resultante. Tiene un promotor de levadura que genera un alto número de copias (2 $\mu$ ) y un gen de auxotrofia para la leucina (LEU2). Posee dos sitios de recombinación attR flanqueando el gen ccdB, que provoca la muerte de la célula, para la reacción LR que sustituirá el inserto de interés por este gen ccdB. También permite su uso en bacterias gracias al origen de replicación en *E.coli* y dos genes de resistencia a antibióticos, ampicilina y cloramfenicol.



**Figura 9.** Esquema del vector pGADT7-GW con los sitios de restricción.

### 3.3. Manipulación y crecimiento de *E.Coli*

Cultivamos bacterias *E.coli* a 37°C durante 24 horas. Para los medios sólidos utilizamos medio LB con agar y el antibiótico necesario para la posterior selección de colonias con los plásmidos de interés. El medio líquido (sin agar) lo utilizamos posteriormente para crecer las colonias seleccionadas en el medio sólido, incubándolo en agitación a 200 rpm a 37°C para conseguir una buena aireación.



### **3.3.1. Preparación de *E.coli* competentes y transformación**

Partimos de glicerizados de células competentes de *E.coli* (previamente preparada con método (Inoue et al., 1990) para transformarlas con nuestros plásmidos. Estos tubos glicerizados se descongelaron, ya que se encontraban a  $-80^{\circ}\text{C}$ , y se incubaron con los vectores de interés en hielo durante 30 minutos.

Para que nuestro DNA plasmídico entrara en las células se procedió a realizar un choque térmico, que consiste en incubar durante 50 segundos las bacterias a  $42^{\circ}\text{C}$ . Tras esto, añadimos 300  $\mu\text{L}$  en medio LB y mantuvimos en agitación 1 hora a  $37^{\circ}\text{C}$ , con la finalidad de recuperar las bacterias del estrés térmico anterior.

Para acabar, sembramos las bacterias en medio sólido de LB sobre una placa con el antibiótico correspondiente para la posterior selección de las células transformadas.

### **3.4. Manipulación y crecimiento de *Saccharomyces cerevisiae***

Las levaduras se mantienen 48 horas a  $28^{\circ}\text{C}$  en medio rico YPDA líquido en agitación a 200 rpm para su correcto crecimiento hasta 1 OD. Después se hizo un lavado con agua y uno con AcLiTE ( $\text{Li}^+$  acetato 0,1M, 10mM Tris-HCl pH7,6, EDTA 1mM) y añadimos el carrier (o ssDNA) para conseguir que estas células sean competentes.

#### **3.4.1. Preparación de *S. cerevisiae* y transformación de *Saccharomyces cerevisiae***

La transformación de levaduras también se llevó a cabo por choque térmico. En este caso, ha de ser un choque más agresivo ya que las levaduras tiene más resistencia a la lisis. Realizamos el choque en presencia de PEGLiTE (Polietilenglicol 4000 al 40% p/v,  $\text{Li}^+$  acetato 0,1M, EDTA 1mM, 10mM Tris-HCl pH7,6) durante 30 minutos a  $28^{\circ}\text{C}$  y luego 20 minutos a  $42^{\circ}\text{C}$ . Luego se extienden en placas con medios selectivos SD con los aminoácidos que se necesite para seleccionar el/los plásmidos de interés.

### **3.5. Medios de cultivo**

#### **3.5.1. Murashige y Skoog (MS)**

Se preparó medio con sales Murashige y Skoog con la cantidad pertinente tanto para medios líquidos como sólidos. Para el medio sólido, sin sacarosa, 10mM de ácido 2-[Nmorpholino]ethanesulfonic (MES) ajustado a pH 6 con Tris básico y con un 1% de agar. (Murashige y Skoog, 1962).

Para las muestras de qRT-PCR las plantas se crecieron en medio nutritivo (Naranjo et al., 2003) sólido y para experimentos de ayuno medio líquido.

Para compensar la presión osmótica que causa la sacarosa al 1%, a las plantas sin tratamiento (después del ayuno) se le ha añadido a la solución nutritiva el 3% de PEG6000.

### **3.5.2. Luria-Bertani (LB)**

El medio Luria-Bertani (LB) es el medio que utilizamos para el crecimiento de *E.coli*. La composición de este medio es: triptona al 1%, extracto de levadura al 0,5% y NaCl al 1% ajustando el pH con NaOH a 7,2. En el caso de realizar el medio en placa, se le añade 20 g/L para que sea sólido, se autoclava y se añaden los antibióticos pertinentes para la selección bajo campana para mantener la esterilidad. Los antibióticos que usamos fueron la kanamicina a 50 µg/mL para las construcciones de pHIS2, la carbenicilina 100µg/mL.

### **3.5.3. SOC (Super optimal broth)**

Este medio lo utilizamos para el crecimiento de las bacterias tras el estrés del choque térmico. La composición es: triptona al 2%, extracto de levadura al 0,5%, 10 mM de NaCl y 2,5 mM de KCl. Ajustamos a pH 7 con NaOH, autoclavamos y cuando se enfrió un poco, añadimos 10mM de MgCl<sub>2</sub> y 20 mM de glucosa, ambas estériles.

### **3.5.4.YPDA (Yeast extract-Peptone Dextrosa Adenina)**

YPDA es el medio más completo para las levaduras y su composición es: 20 g/L de peptona y 10 g/L de extracto de levadura. A esto se le añade HCl para ajustar el pH a 5,8 y se enrasa con agua hasta 1 L para después esterilizar en el autoclave. Tras la esterilización y enfriamiento a unos 50°C, se le añade 50 ml glucosa al 20%.

### **3.5.5. Medio sintético SD**

Otro medio para levadura es el medio SD, este no es el mejor medio para las condiciones de crecimiento, pero es útil a la hora de seleccionar las levaduras que hayan incorporado el plásmido de interés. Su composición es: Yeast Nitrogen Buffer (YNB) al 0,7%, glucosa al 2%, 0,05M de ácido succínico a pH 5,5 y un 2% de agar, además de los aminoácidos que añadiremos para que la levadura crezca, pero en los que omitiremos los aminoácidos para la selección por auxotrofia, que dependerán en cada caso del plásmido con el que transformemos.

### 3.6. Primers

Para amplificar la secuencia de nuestro promotor de AHA2 a partir del DNA genómico de Arabidopsis, necesitamos diseñar unos primers, que fueron los siguientes:

AHA2 forward: GATTAAGAACTTAAAGAGGCAG

AHA2 reverse: GCATAGCTGAGACGTTGTCTT

Y para amplificar el cDNA (es cDNA porque en levadura , donde se hará Y1H, no podemos poner intrones, tiene que estar ya procesado para que de la proteína) de la secuencia de PTM2, necesitamos diseñar otros cebadores:

At5g22760\_0027s : ATGGAAGGAAAGGTGGCTAACCTAGA

At5g22760\_1540STOPas: TCATACCAGGACAACCTTTTGAATGT

Los cebadores para las qRT-PCR:

PP2AA3\_AT1G13320\_F (gen referencia) :5'-ACCTGCGGTAATAACTGCATCTA-3'

PP2AA3\_AT1G13320\_R (gen referencia) :5'-CCGAACATCAACATCTGGGTC-3'

AHA1Forward: 5'-GAACTAATGGGTGCTGTGTATCT-3'

AHA1Reverse: 5'-CCAGGACGTTTCGACAAATGA-3'

AHA2Forward: 5'-CCAGGCCATCATGACTGTTATT-3'

AHA2Reverse: 5'-GTAAGTACACCGCACCCATTAG-3'

AHA3Forward: 5'-CGTTACATACTAGCCGGAACAG-3'

Los cebadores para comprobar la integración del plásmido pHIPPIE:

231HIS3F (específico vector integración): 5'-ATCGTTGAGTGCATTGGTGAC-3'

closeURA3\_0660as (específico genoma levadura): 5'-TCTAATTTGTGAGTTTAGTATACAT-3'

### 3.7. Reacción en cadena de la Polimerasa (PCR)

Para las reacciones de PCR del promotor de AHA2 se utilizaron los siguientes reactivos: tampón de PCR 10x (BioTools), 0,2 mM dNTPs, 1  $\mu$ M primer directo, 1  $\mu$ M primer reverso, DNA (0,1-20 ng/ $\mu$ L), DNA polimerasa Taq (1 U/ $\mu$ L) (BioTools) y agua desionizada hasta un volumen de 20  $\mu$ L/reacción.

Las reacciones tuvieron lugar en un termociclador bajo las siguientes condiciones:

- 1 ciclo de desnaturalización a 95 °C de 5 min.
- 30 ciclos de amplificación: 30 segundos de desnaturalización a 95 °C, 30 segundos de hibridación de primers utilizando una temperatura de anillamiento de los oligonucleótidos entre 57 y 62 °C, dependiendo del par de primers utilizados en cada caso y extensión a 72 °C durante 1 min por cada kilobase de DNA que se fuera a amplificar.
- 1 ciclo de extensión final a 72 °C durante 5 min.

Sin embargo, para amplificar el cDNA de ptm1 utilizamos la DNA Polimerasa Phusion (New England BioLabs), en lugar de la Taq Polimerasa anterior. Esto es porque la Phusion tiene 50 veces más fidelidad de copia (tasa de error  $4,4 \cdot 10^{-7}$ ) que la DNA polimerasa Taq y posee la actividad 3'-5' exonucleasa, actividad correctora por la que se evita arrastrar errores durante la reacción.

Los reactivos fueron: tampón HF 5x (New England BioLabs), 0,2 mM dNTPs, 0,5  $\mu$ M primer directo, 0,5  $\mu$ M primer reverso, DNA genómico (~2 ng/ $\mu$ L), 3% DMSO, DNA polimerasa Phusion (1 U/ $\mu$ L), agua MiliQ estéril hasta alcanzar un volumen final de 50  $\mu$ L/reacción. Y las condiciones en termociclador:

- 1 ciclo de desnaturalización a 98 °C de 3 min.
- 30 ciclos de amplificación: 10 segundos de desnaturalización a 98 °C, 15 segundos de hibridación de primers a 57° C y 15 segundos de extensión a 72 °C.
- 1 ciclo de extensión final a 72 °C durante 3 min.

### 3.8. Extracción RNA, conversión a cDNA y qRT-PCR

Para realizar las PCR cuantitativas a tiempo real (qRT-PCR) se tuvo que extraer RNA de las plantas según el protocolo de Oñate-Sánchez & Vicente-Carbajosa, 2008. Tras esto, normalizamos la concentración del RNA en todas las muestras con la ayuda del NanoDrop. Después seguimos el protocolo ("Maxima First Strand cDNA Synthesis Kit for RT-qPCR") de ThermoFisherScientific para sintetizar el cDNA a partir del RNA. La reacción de síntesis se mantuvo a 65°C durante 45 minutos para un volumen de 9,6 $\mu$ L y un 0.4 $\mu$ L de molde, que se realizó en la máquina "Applied Biosystems 7500 Fast", con el sistema Cmb (5X Pyro Taq EvaGreen qPCR).

### **3.9. Electroforesis**

Utilizamos la electroforesis como comprobación tanto después de amplificar DNA por reacciones de PCR tanto como después de reacciones de ligación y digestión. Los geles de agarosa al 0,7% contenían también 1,5%, tampón TBE (45 mM Tris base, 45 mM ácido bórico y 1 mM EDTA) y bromuro de etidio (500 µg/L). Para cargar las muestras en los pocillos del gel añadimos 1/5 volúmenes de tampón de carga LB x6 (40% sacarosa, 0,25% azul de bromofenol y 0,1 M EDTA). Utilizamos también como referencia para todas las electroforesis el marcador GeneRuler1kb para ver los tamaños aproximados de las bandas de interés.

Una vez cargado el gel, se deja correr a un máximo de 10V/cm con el mismo tampón TBE y se visualizan y fotografían en un transiluminador con luz UV.

En el caso de querer extraer la banda, como fue en el caso de la clonación de AHA2 en pHIPPIE y pHIS2, utilizamos el kit GENE CLEAN® Turbo (MP Biomedicals) para purificarla y poder manipularla.

### **3.10. Recombinación**

La reacción de Gateway® según el protocolo de la casa comercial Invitrogen se utilizaró para insertar el producto PCR de PTM2 en el plásmido pDONR/Zeo y la reacción LR para poner el inserto PTM2 en el plásmido final pGADT7-GW.

### **3.11. Minipreparaciones de plásmidos**

Purificamos los plásmidos incorporados y seleccionados en E.coli mediante el kit NucleoSpin® Plasmid (Macherey-Nagel). Para el caso de la purificación de los plásmidos para levadura se utilizó el mismo protocolo, solo que se hizo un paso previo de lisis celular con bolas de cristal de 0,5 mm de diámetro y agitación con vórtex durante 10 minutos, ya que las levaduras son mucho más resistentes a la rotura que las bacterias.

### **3.12. Digestiones**

Hemos realizado varias digestiones con enzimas de restricción diferentes con el fin de comprobar que las minipreparaciones de los plásmidos eran correctas tras correr un gel y ver el tamaño de las bandas obtenidas, o con el fin de subclonar los fragmentos de DNA. En este caso, pudimos cortar el inserto de interés de un plásmido, correrlo en un gel para comprobarlo y extraerla para seguir trabajando con ella. Los reactivos son: tampón 10X (con BSA), el enzima de restricción pertinente (2U/µL) de la casa Thermo Fisher Scientific, el DNA plasmídico (1 µg/µL), y agua desionizada hasta un volumen final de 10? O 20 µL. Después se deja incubando durante una hora a 37°C.

### **3.13. Ligaciones**

Tanto en la clonación de AHA2 como en la clonación del factor de transcripción PTM2 se realizaron ligaciones. Estas con la finalidad de ligar un fragmento dentro de un plásmido. Para ello se utiliza el enzima T4 DNA Ligasa con el tampón pertinente y se dejó una hora a temperatura ambiente para que llevara a cabo la reacción.

### **3.14. Secuenciación**

La secuenciación fue un recurso vital al que recurrimos para comprobar si nuestros insertos estaban en los plásmidos deseados. Mandamos las muestras al propio servicio de secuenciación del IBMCP, a una concentración de 200 ng/μL en un volumen de 10 μL. También mandamos los primers correspondientes a 5 μM, o en el caso de utilizar alguno de los primers típicos (T7, M13) que tiene el servicio, lo indicamos en el informe.

Tras obtener los resultados de las secuencias, los analizamos con herramientas bioinformáticas para ver que todo era correcto (no mutaciones por PCR o por el clonaje) y en qué sentido se insertaba nuestro fragmento de interés. Una de estas herramientas fue el BLAST (Basic Local Alignment Search Tool) con las bases de datos TAIR (<http://www.arabidopsis.org>) y PLAZA ([https://bioinformatics.psb.ugent.be/plaza/versions/plaza\\_v4\\_dicots/](https://bioinformatics.psb.ugent.be/plaza/versions/plaza_v4_dicots/)).

### **3.15. One Hybrid**

#### **3.15.1. Construcción PTM2 y transformación en levadura THY.AP4**

La primera transformación que realizamos fue la de nuestro factor de transcripción PTM2 en el vector p-GADT7-GW. Primero subclonamos el factor de transcripción en el vector pDONR/Zeo mediante tecnología Gateway con una reacción BP y después de comprobar que la inserción fue correcta, transferimos este factor PTM2 al vector final pGADT7-GW mediante una reacción LR también por la tecnología Gateway.

Tras esto, transformamos las levaduras por choque térmico con nuestra construcción pGADT7-GW. Para seleccionar las levaduras que habían sido transformadas con nuestro vector, plaqueamos sobre placas con medio SD sin leucina, que es el gen de auxotrofia específico del vector pGADT7-GW.

### **3.15.2. Construcción promotor AHA2 y transformación en levadura Y187**

En este caso, la construcción fue la de los vectores pHIPPIE y pHIS2 con el promotor de AHA2. Para ello, ligamos nuestro inserto AHA2 en el plásmido pGEMT-easy con el enzima ligasa. Después linealizamos los vectores pHIPPIE y pHIS2, los tratamos con fosfatasa alcalina según el protocolo de la casa Roche a 37°C durante 30 minutos para evitar la autoligación de los vectores. Comprobamos mediante la digestión del plásmido pGEMT-Easy que nuestro inserto era correcto y recuperamos la banda del gel para poder seguir trabajando con ella. Esta banda, perteneciente al promotor de AHA2 se ligó en ambos vectores con una ligasa. Para ver qué vectores habían aceptado el inserto y en qué sentido, realizamos una PCR con dos combinaciones de primers. Tras esto, pasamos a la transformación de las levaduras con nuestras construcciones. En el caso de pHIS2 este entra en la levadura, mientras que en el pHIPPIE, el plásmido se integra en el genoma mediante la recombinación homóloga con la región de URA3. El plásmido pHIPPIE posee el gen URA3, que en la levadura está mutado, lo cual no permite a la levadura crecer sin uracilo. Para que se integrara el vector, se cortó con EcoRV, que corta dentro de la auxotrofia para el uracilo. Así, linealizamos el vector, y si este se integra en el genoma, consigue reconstruir el gen de URA3, por lo que la levadura será capaz de crecer sin uracilo en el medio mínimo SD.

Nuestra cepa de levadura Y187 es auxótrofa para el triptófano, histidina, lisina, leucina, adenina y uracilo, por lo que al medio SD le añadimos todos estos aminoácidos a excepción del uracilo.

### **3.15.3. Ensayo Yeast One-Hybrid (Y1H)**

Las colonias diploides crecidas después de un par de días en medio mínimo selectivo líquido (-UL o -WL) habían llegado a saturación. Por mayor exactitud medimos la OD (densidad óptica) a 660 nm y llevamos todas las muestras con agua a una OD final de 0,07 con para que todas estuvieran con la misma concentración inicial de células. Para este ensayo gastamos una placa multipocillos para poner de forma ordenada nuestras muestras con 4 repeticiones independientes en distintos pocillos. En cada pocillo había 200uL de células e hicimos goteos gastando placas petri de SD (-UL o -WL) para comprobar si hay crecimiento regular (control) y placas petri (-ULH o -WLH) para medir la posible interacción de nuestro factor de transcripción con el promotor de interés o autoactivación si creciera con el vector vacío.

## 4. RESULTADOS

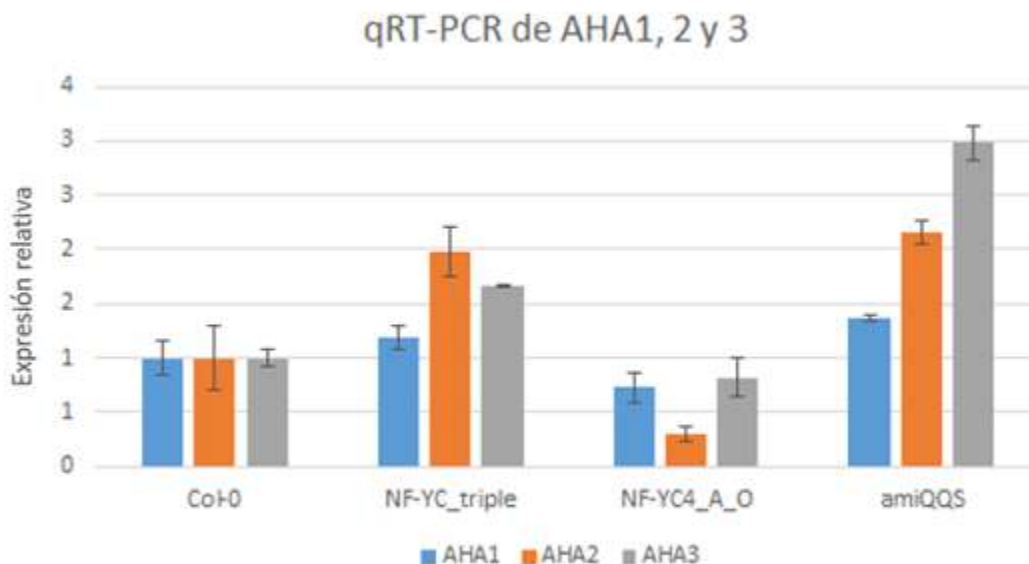
### 4.1. REGULACIÓN DE AHA2 POR NF-YC Y QQS

#### 4.1.1. Cuantificación de la expresión de AHA1, AHA2 y AHA3

Mediante una qRT-PCR se quiso comprobar si la expresión de las ATPasa está modificada en las líneas de *Arabidopsis* mutantes para los genes de QQS y NF-YC.

Se escogieron las 3 isoformas más relevantes de AHA y las siguientes líneas: Col-0, línea control, NF-YC\_triple, línea mutante de pérdida de función para los tres factores NF-YC, NF-YC4\_A\_O, línea mutante con sobreexpresión de NF-YC4, y por último, amiQQS. Esta última se trata de una línea mutante generada mediante la tecnología de los microRNAs (miRNA) artificiales, que silencian el gen diana utilizando pequeños fragmentos de RNA.

Se observa en la **Figura 10**, que en el mutante de los tres factores NY-FC se induce la expresión, especialmente de AHA2 y en cambio en la sobreexpresión de NF-YC4 disminuye su expresión. La línea mutante de QQS, en la que se inhibe la expresión de QQS, también aumenta la expresión de las tres isoformas de AHA. La línea Col-0 fue utilizada como referencia.



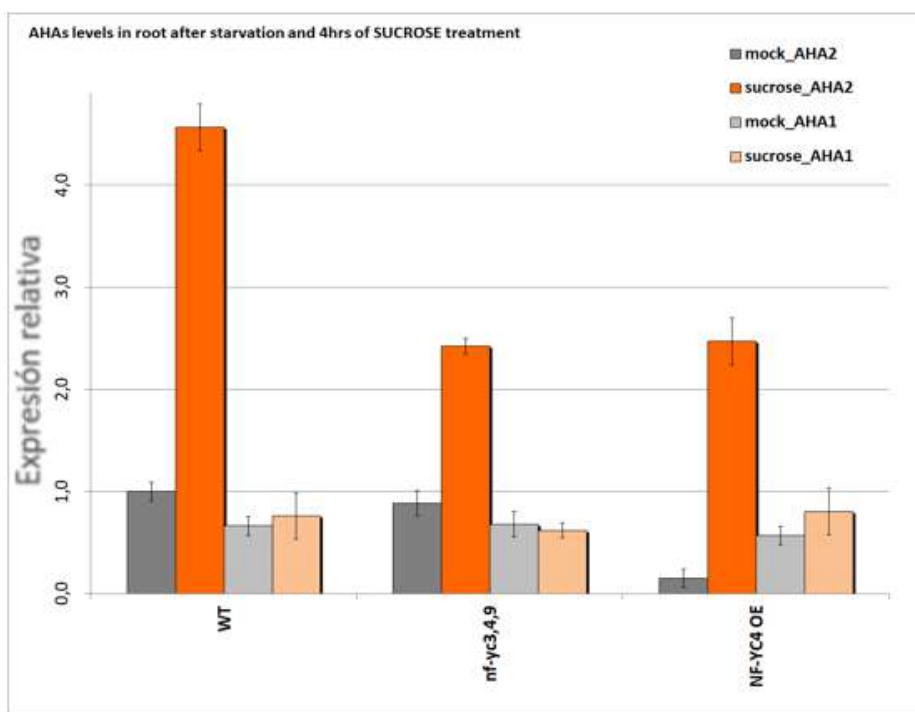
**Figura 10.** qRT-PCR de AHA1,2 y 3. Cuantificación de la expresión de las 3 isoformas más relevantes de AHA (AHA1, AHA2, AHA3) en diferentes líneas de *Arabidopsis* (Col-0, NF-YC\_triple, NF-YC4\_A\_O, amiQQS) mediante una qRT-PCR.



#### 4.1.2. Cuantificación de la expresión de AHA1, AHA2 en condiciones de ayuno y tras tratamiento con sacarosa

La sacarosa es detectada por proteínas de membrana que desencadenan cascadas de fosforilación. Como último paso se pueden fosforilar factores de transcripción, los cuales regularían la expresión de distintos genes, entre los cuales se podría encontrar las protón ATPasas (Niittyä et al., 2007).

Con el fin de comprobar esta hipótesis, realizamos una qRT-PCR de diferentes líneas de Arabidopsis para las isoformas AHA1 y AHA2. Utilizamos una planta wild-type (WT), una línea mutante para los tres factores de transcripción NF-YC (*nf-yc3,4,9*), y otra con la sobreexpresión de el factor NF-YC4 (NF-YC4\_OE). Todas las plantas fueron sometidas a un periodo de ayuno y después tratadas con sacarosa. Los resultados fueron los que se observan en la **Figura 11**. Los valores para AHA1 no varían tanto con tratamiento como sin tratamiento, pero los valores de AHA2 se disparan en la planta control Col-0 (WT).



**Figura 11.** qRT-PCR de AHA1 y 2 tras tratamiento con sacarosa. Resultados de la qRT-PCR donde se ve la expresión relativa de AHA1 y AHA2 tras el tratamiento con sacarosa en las diferentes líneas mutantes.

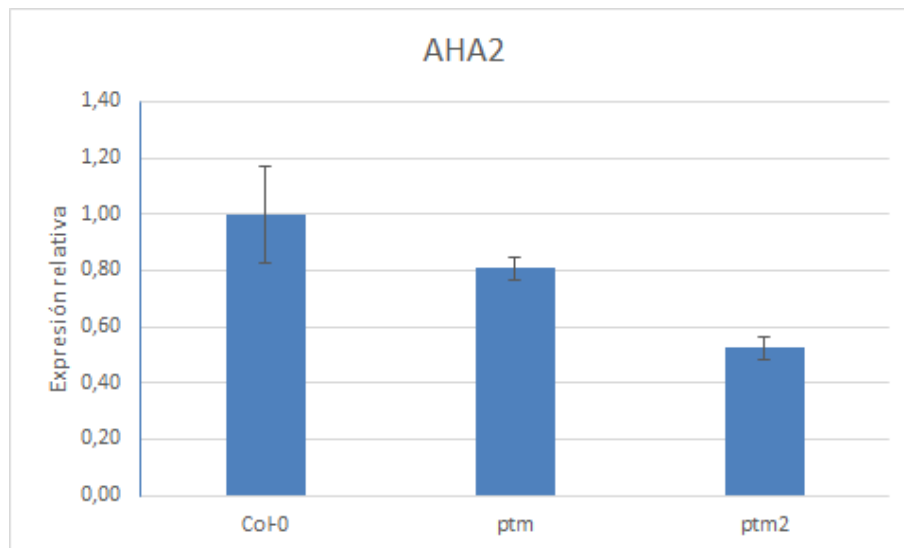
## 4.2. REGULACIÓN DE AHA2 POR PTM2

### 4.2.1. Expresión de AHA2 en mutantes de PHD

En primer lugar, realizamos una qRT-PCR de AHA2 en dos mutantes de los genes de tipo PHD: el ya descrito PTM y el At5g22760, que nosotros denominamos PTM2 por ser de la misma familia que el anterior.

Respecto a Col-0, que son las plantas de Arabidopsis que utilizamos como control, pudimos observar que la expresión de AHA2 en los mutantes *ptm*. La expresión relativa de AHA2 se tomó como referencia en las plantas control y le asignamos un valor de 1. La expresión, en el caso de las líneas mutantes *ptm*, tal como se observa en el gráfico (**Figura 12**), fue inferior a los valores del control. Esto nos indicó, que el gen PTM está involucrado en la expresión de AHA2, activándola, ya que en los mutantes donde *ptm* y *ptm2* no se expresan, la expresión de AHA2 también disminuye.

Este fue el ensayo clave por el que decidimos estudiar al gen At5g22760 o PTM2 mediante un ensayo Y1H para comprobar si este factor de transcripción, que no está caracterizado, interacciona con el promotor de AHA2.

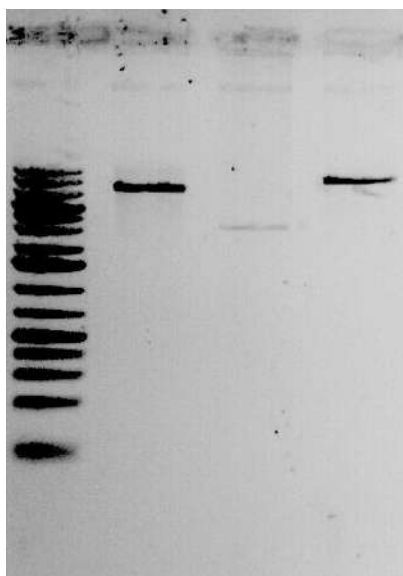


**Figura 12.** qRT-PCR de AHA2 con líneas mutantes de PTM. En la primera columna está el ecotipo Col-0 que sirve como referencia. *ptm* y *ptm2* son los mutantes en los que estos genes están inhibidos. Podemos observar la expresión relativa de AHA2 en Col-0 asignada con valor 1, mientras que en los mutantes disminuye la expresión, siendo menor en el mutante *ptm2*.

#### 4.2.2. Clonaje de promotor de AHA2 en pHIPPIE y pHIS2

Mediante una PCR con los cebadores se amplificó el fragmento del promotor de AHA2 y lo subclonamos en el vector pGEMT-Easy mediante una ligación. La inserción se comprobó mediante secuenciación.

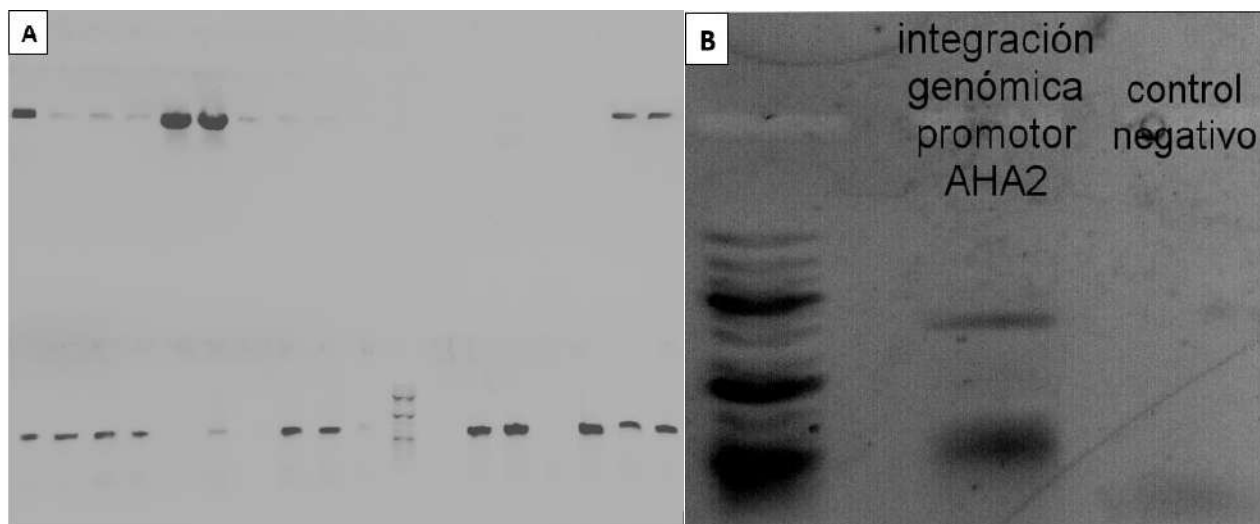
A continuación, procedimos a clonar este inserto en los vectores finales pHIS2 y pHIPPIE para el ensayo de Y1H. Para ello, cortamos los vectores con el enzima EcoRI, debido a que solo presentan un sitio de corte de este enzima. Para prevenir la autoligación del mismo vector, tratamos según el protocolo con fosfatasa alcalina y comprobamos en un gel de electroforesis (Figura 12). Una vez vimos que nuestros vectores estaban bien cortados, extrajimos la banda que correspondía a nuestro fragmento de interés (**Figura 13, columna 3**) y lo insertamos mediante una ligación con la ligasa.



**Figura 13.** Gel de electroforesis tras digerir y tratar los vectores con fosfatasa alcalina. La primera columna es la del marcador, la segunda es la digestión parcial de pHIS2, la tercera es la digestión de pGEMT-Easy y la cuarta la digestión de pHIPPIE.

Una vez hicimos la ligación, hicimos una PCR para comprobar que el inserto estaba insertado y en qué sentido se encontraba. Para ello utilizamos dos combinaciones de primers: ambas con el primer reverso del vector en HIS3. Una llevaba el primer *forward* del inserto y la otra combinación, el primer inverso del inserto. Corrimos un gel de electroforesis (figura 14A) con las muestras (10 de pHIPPIE y 6 de pHIS2). Asumimos que las muestras correctas fueron la pHIPPIE 5 en sentido forward, y el pHIS2 2 y 4 en sentido reverso según la imagen del gel.

Una vez conseguimos nuestros plásmidos finales con el inserto de interés, transformamos las levaduras de la cepa Y187 para el ensayo final de Y1H. En el caso de el plásmido pHIS2, este entra en la levadura y en el caso del pHIPPIE, el plásmido recombina mediante las regiones de URA y se integra en el genoma de la levadura. Para comprobar que pHIPPIE se había integrado en el genoma de la levadura correctamente, hicimos una PCR y corrimos un gel. (**Figura 14B**).



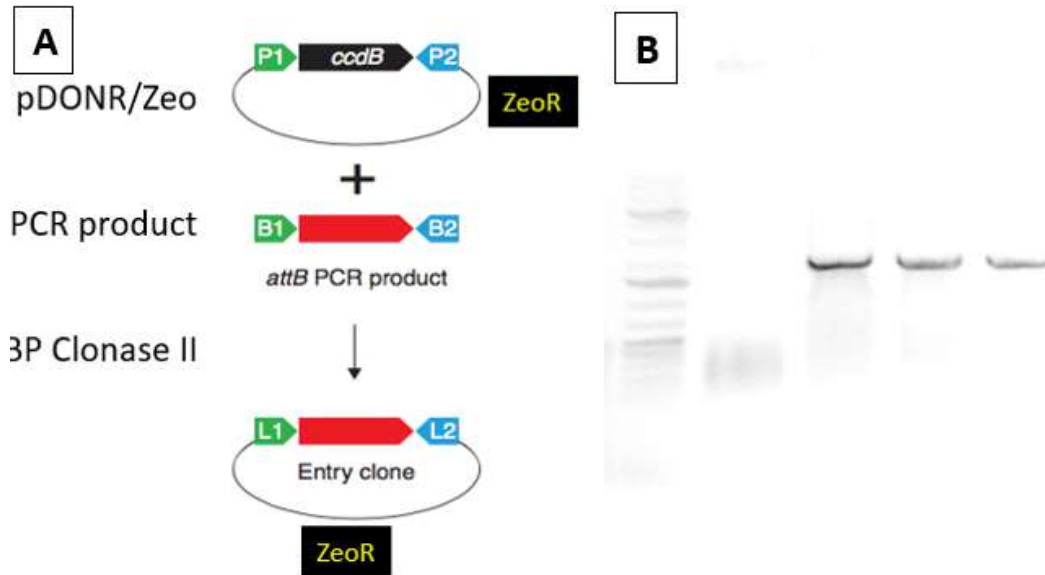
**Figura 14. A.** Gel de electroforesis de la PCR para comprobar el sentido del inserto en pHIPPIE y pHIS2. Hay dos líneas de pocillos, la de arriba con la combinación de cebadores de sentido *forward* para el inserto y la de abajo con la combinación para el sentido reverso. Las muestras válidas son las que presentan una banda en solo una de las dos líneas de combinaciones de cebadores. pHIPPIE5 solo muestra banda arriba, con lo cual se encuentra en sentido *forward* y pHIS2 2 y 4 solo se encuentran en la línea de abajo, por lo que el inserto lo tendrán en sentido reverso. **B.** Gel de electroforesis de comprobación de la integración del vector pHIPPIE en el genoma de levadura. La banda superior en la segunda columna es el fragmento correspondiente a AHA2 (aprox 3500 bp) y la inferior es un producto secundario, algo que puede suceder en PCR tan complejas.

#### 4.2.3. Clonaje de PTM2 en pGADT7-GW

El primer paso fue subclonar el gen de PTM2 en el vector pDONR/Zeo mediante una reacción BP (**Figura 15.A**).

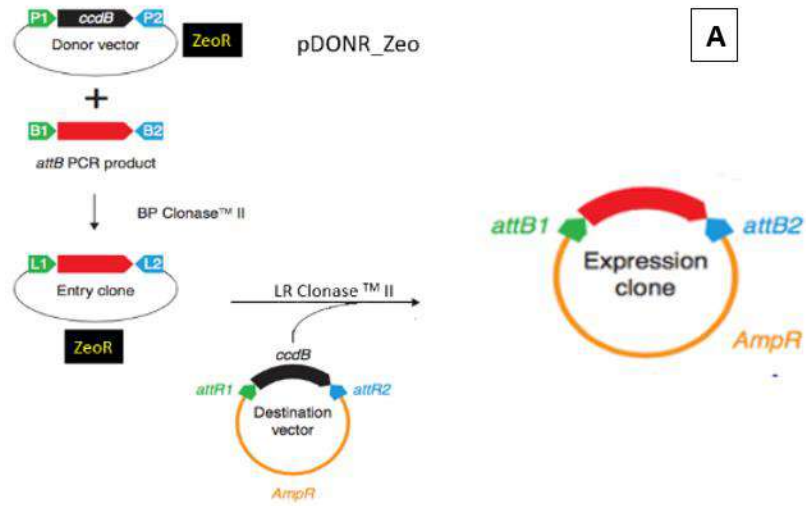
Para comprobar que nuestro gen de interés se insertó, se realizó una digestión parcial con el enzima MunI, que tiene un sitio de restricción dentro del inserto, por lo que si el plásmido ha integrado el inserto, se cortará.

Una vez hecha la digestión parcial, se hizo una electroforesis en gel de agarosa en la que se debería esperar para el plásmido con el inserto una banda de 3300 bp. Esto es porque restando al total del vector (4291 bp) las pares de bases que existen entre los sitios de recombinación attp1 y attp2 (2500) y sumando lo que ocupa nuestro gen PTM2 (1540), nos da 3300 bp. En nuestro gel pusimos 4 muestras, 3 de las cuales dieron el fragmento esperado (**Figura 15.B**).

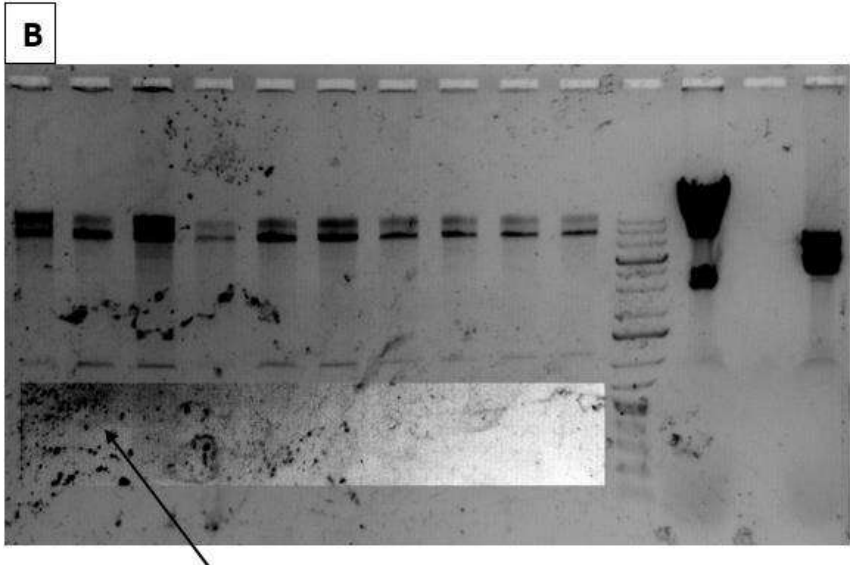


**Figura 15. A.** Reacción BP. Con la enzima “BP clonase II” entre el vector pDONR/Zeo y nuestro gen de interés. **B.** Electroforesis en gel de agarosa de la digestión parcial del plásmido pDONR/Zeo con el gen PTM. En la primera columna se encuentra el marcador molecular y en las 4 siguientes, las diferentes muestras de la digestión. Solo las 3 últimas dan la banda esperada a 3300 bp aproximadamente.

Tras tener asegurado que nuestro gen PTM2 se insertó en el vector de subclonación, pasamos este mismo gen al vector final pGADT7 mediante una reacción Gateway LR. (**Figura 16.A**) Para comprobar que pGADT7 tiene el inserto PTM2 transformamos *E.coli*, y realizamos una digestión parcial con EcoRV y HindIII. Los sitios de restricción de HindIII son dos fuera del inserto, y EcoRV solo tiene uno dentro de este. En el gel de agarosa se observa una banda de un tamaño alrededor de 400 bp (flecha de la **Figura 16.B**) que correspondería al fragmento que se encuentra entre los sitios de corte HindIII y EcoRV, indicando que el vector contiene el inserto. Luego comprobamos que efectivamente sí se encontraba en el vector secuenciando con el promotor T7 del pGADT7.



A



B

**Imagen.16 A.** Reacción Gateway LR entre el vector pDONR/Zeo y el pGADT7-GW. **B.** Gel de agarosa de la digestión con HindIII y EcoRV de pGADT7 de *E.Coli* transformadas. La flecha indica la banda que corresponde al inserto PTM2.

#### 4.2.4. Yeast One-Hybrid (Y1H)

El ensayo Y1H está formado por dos componentes: la “presa” (prey) y el “cebo” (bait). La llamada “presa” es la proteína híbrida que se forma con nuestra proteína de interés (en nuestro caso PTM2) conjugada con el factor GAL4, y el “cebo” será el DNA al que se unirá la “presa”, que en caso de que esta interacción ocurra, expresará el gen reportero (en nuestro caso HIS3).

En la **Figura 17** se observa nuestro resultado final del Y1H. Las combinaciones que realizamos fueron:

- pHIPPIE-AHA2 con pGADT7-PTM2
- pHIPPIE sin promotor (vacío) con pGADT7-PTM2
- pHIS2-AHA2 con pGADT7-PTM2
- pHIS2 sin promotor (vacío) con pGADT7-PTM2
- pGADT7-GW con CK2 $\alpha$ 1
- pGBTK con CK2 $\beta$ 1

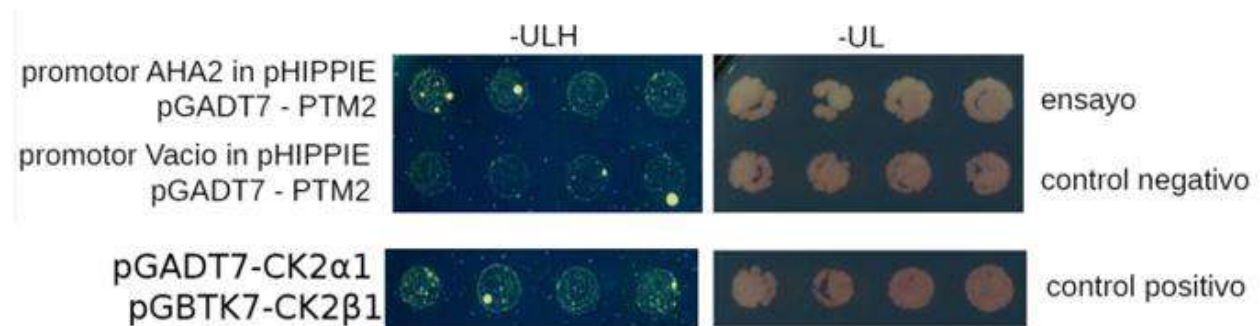
El vector “presa” pGADT7-GW está presente en todas las combinaciones por lo que ningún medio tendrá leucina. pGADT7-GW lleva el gen para la auxotrofia de la leucina, por lo que si el plásmido se transcribe correctamente la levadura será capaz de crecer sin la leucina. En el caso de estar presente el vector pHIPPIE (“cebo”), el medio no llevará ni uracilo ni histidina, ya que este plásmido contiene los genes para la auxotrofia para estos dos aminoácidos. En el caso de estar presente el plásmido pHIS2 (“cebo”), la histidina y triptófano no estarán presentes en el medio por la misma razón.

En el caso de la combinación de pHIPPIE con pGADT7-GW se replicaron en medio SD sin leucina ni uracilo, donde se deberían ver todas las colonias creciendo ya que ambos compuestos moleculares del ensayo están presentes. Y también en medio sin leucina, ni uracilo, ni histidina, donde deberían crecer solo las colonias en las que interaccione el factor de transcripción con el promotor, debido a que se transcribirá el gen reportero de la histidina, lo que permitirá crecer a la levadura. En el caso de la combinación de pHIS2 con pGADT7-GW las colonias crecieron en un medio sin triptófano y sin leucina, en el cual deberían crecer todas las colonias ya que los dos plásmidos que confieren la capacidad de sintetizar estos aminoácidos por ellos solos están presentes. Y en el medio en el que además se le quita también la histidina, deberían crecer colonias en el caso de que la interacción del promotor AHA2 y PTM2 exista.

La interacción entre el promotor AHA2 y el factor PTM2 con GAL4 permitiría la transcripción y traducción de histidina, lo cual permitiría que las levaduras crecieran en un medio en ausencia de este aminoácido. Pero puede darse el caso de algún resultado positivo porque se auto-active alguno de los componentes, que descartaremos por los controles negativos.

Por otra parte, contamos con dos controles positivos, CK2 $\alpha$ 1 y CK2 $\beta$ 1, dos kinasas que interactúan entre ellas según (Riera et al., 2011) comprobado por un ensayo de doble híbrido. Estos dos controles son muy importantes ya que en el caso de que todo nos hubiera salido negativo, es decir, que no hubiera crecido nada, no sabríamos si es porque la interacción que estamos testando no ocurre o si es por algún fallo en la metodología, como por ejemplo una mala selección de los aminoácidos para los medios selectivos.

En nuestro ensayo Y1H, se puede observar según la **Figura 17**, los diferentes medios con combinaciones anteriormente descritas. Si nuestro factor de transcripción interacciona con el promotor AHA2, las levaduras deberían crecer en los medios correspondientes sin histidina, no así, las mismas combinaciones pero con un plásmido vacío. Así pues, nosotros consideramos que el resultado es positivo cuando en el círculo donde hemos depositado las células se ve el crecimiento de una manera global y no de colonias aisladas. En cambio, cuando el vector pHIPPIE está vacío, parece que hay crecimiento también pero es mucho más débil que en el caso del pHIPPIE con el inserto. Es por esto, que consideramos que nuestra combinación de pHIPPIE con pGADT7-GW es positiva, por lo que aceptamos que hay una interacción de DNA-proteína.



**Figura 17.** Ensayo final del Y1H. El cuadro -UL (medio SD sin uracilo y sin leucina) es para ver qué levaduras han adquirido el plásmido, dado que salen colonias en los dos controles, todas las levaduras han adquirido los plásmidos correspondientes. El cuadro -ULH (sin uracilo, leucina e histidina) es para ver si hay interacción. Si crecen las levaduras es porque son capaces de crecer sin histidina, que significa que hay interacción, ya que es por ella por la que se transcribe el gen de la histidina y les permite crecer en un medio sin dicho aminoácido. Se han omitido las imágenes de pHIS2 ya que no salían resultados conclusivos. CK2 $\alpha$ 1 y CK2 $\beta$ 1, son el control positivo del experimento para saber que las condiciones son correctas y no ha habido ningún error metodológico. Hay interacción segura, por lo que podemos comparar esta interacción con la de nuestro promotor y nuestro factor de transcripción.



## 5. DISCUSIÓN

### 5.1. REGULACIÓN NEGATIVA DE AHA2

La H<sup>+</sup>-ATPasa es una proteína muy interesante para el estudio debido a sus múltiples y muy importantes funciones a nivel fisiológico. Esto ha hecho que no falten estudios en cuanto a su regulación; no obstante, la mayoría de estos estudios se limitan a la regulación post-traducciona, dejando a la regulación transcripcional en un segundo plano. Es por ello nuestro interés en centrarnos en la regulación de esta proteína a nivel transcriptómico. En este trabajo se utilizaron distintas líneas mutantes en las cuales se determinó la expresión de las tres isoformas de la protón ATPasa AHA1, AHA2 y AHA3.

La sobreexpresión de los factores NF-YC, así como su inhibición y la de QQS, confirman que estos se tratan de reguladores negativos de la H<sup>+</sup>-ATPasa, lo cual pudimos comprobar con los estudios de la expresión relativa de esta proteína de membrana. Estos resultados concuerdan con la investigación de Haruta et. al 2012. Además, también pudimos observar como los azúcares también participan en esta regulación de la H<sup>+</sup>-ATPasa. Con otro ensayo de medición de la expresión relativa de AHA2, pudimos comprobar que tras tratar las plantas en ayuno con un poco de sacarosa, esta disminuía su expresión. La sacarosa por tanto, es un componente importante de regulación. La activación de AHA2 tras el tratamiento con sacarosa nos indica que tras este estrés de ayuno con el posterior tratamiento con azúcar, la protón ATPasa es capaz de inducir su actividad. Este mecanismo sería especialmente relevante para el crecimiento de la planta.

### 5.2. REGULACIÓN POSITIVA DE AHA2

Según nuestros ensayos de qRT-PCR pudimos observar como los mutantes en PTM y PTM2 tienen una expresión relativa menor de AHA2. Especialmente relevante sería en el caso de *ptm2*. Este hecho fue el punto de partida que nos impulsó a querer comprobar si existía una interacción entre este factor de transcripción y el promotor de AHA2.

El estudio de la posible interacción entre el factor de transcripción PTM2 y el promotor de AHA2 se llevó a cabo mediante un ensayo de Yeast One-Hybrid, debido a que este sistema estaba ya puesto a punto en nuestro laboratorio y a las ventajas que ofrece (Sun et al., 2011) respecto a otros sistemas de detección de este tipo de interacciones. Este sistema también nos permite detectar interacciones DNA-proteína que no son activadoras, lo que permite detectar proteínas represoras, a diferencia que en los ensayos de luciferasa. Otra ventaja es que este ensayo Y1H es compatible con muchas bibliotecas ya existentes, ya que la mayoría de ensayos Y1H pueden usar “presas” híbridas y bibliotecas construídas para Y2H.

Además Y1H puede detectar interacciones específicas de diferentes isoformas mediante el diseño de plásmidos que codifique para las diferentes isoformas de interés.

Otro ensayo que podría haber sido útil es el ChIP-seq, ya que se basa en la unión de anticuerpos específicos a la proteína de interés, que por el uso de estos ya resulta muy caro para nuestro presupuesto. Otra opción habría sido realizar plantas transgénicas con GFP y detectar esta proteína GFP para ver cuándo se expresa, lo cual es un proyecto factible y que posiblemente se realice en el laboratorio.

Tras realizar el ensayo Y1H pudimos observar que la interacción de PTM2 con el promotor AHA2 parece producirse, aunque como hemos comentado previamente, las condiciones de crecimiento de las levaduras no fueron las mejores. Esta interacción, podría mejorarse repitiendo con otras condiciones de crecimiento.

## 6. CONCLUSIONES

1. Se ha observado mediante el estudio de la expresión relativa de AHA2 en diferentes líneas mutantes de los factores NF-YC estudiados, como del gen QQS, que estos participan en la regulación transcripcional inhibiendo a la isoforma 2 de la H<sup>+</sup>-ATPasa en *Arabidopsis thaliana*.
2. En cambio, PTM2, es un factor de transcripción de la familia de los PHD, que participa también en la regulación de AHA2 en *Arabidopsis thaliana*, pero activándola.
3. El promotor de AHA2 en *Arabidopsis thaliana* parece sí interactuar con el factor de transcripción PTM2, aunque sería necesario repetir el ensayo Y1H en mejores condiciones para las células e incluso complementarlo con otros estudios que pudieran confirmar con más certeza que esta interacción existe en la naturaleza.

## 7. BIBLIOGRAFÍA

**Duby, G., & Boutry, M. (2008).** The plant plasma membrane proton pump ATPase: a highly regulated P-type ATPase with multiple physiological roles. *Pflügers Archiv - European Journal Of Physiology*, 457(3), 645-655.

**Coego, A., Brizuela, E., Castillejo, P., Ruíz, S., Koncz, C., del Pozo, J.C., Piñeiro, M., Jarillo, J.A., Paz-Ares, J., León, J., and The TRANSPLANTA Consortium (2014).** The TRANSPLANTA collection of Arabidopsis lines: a resource for functional analysis of transcription factors based on their conditional overexpression. *The Plant Journal* 77: 944–953.

**Falhof, J., Pedersen, J., Fuglsang, A., & Palmgren, M. (2016).** Plasma Membrane H<sup>+</sup>-ATPase Regulation in the Center of Plant Physiology. *Molecular Plant*, 9(3), 323-337.

**Feng, P., Guo, H., Chi, W., Chai, X., Sun, X., & Xu, X. et al. (2016).** Chloroplast retrograde signal regulates flowering. *Proceedings Of The National Academy Of Sciences*, 113(38), 10708-10713.

**Haruta, M., & Sussman, M. (2012).** The Effect of a Genetically Reduced Plasma Membrane Protonmotive Force on Vegetative Growth of Arabidopsis. *Plant Physiology*, 158(3), 1158-1171.

**Haruta, M., Gaddameedi, V., Burch, H., Fernandez, D., & Sussman, M. R. (2018).** Comparison of the effects of a kinase-dead mutation of FERONIA on ovule fertilization and root growth of Arabidopsis. *FEBS Letters*, 592, 2395–2402.

**Ilan, N. (1994).** External pH effects on the depolarization-activated K channels in guard cell protoplasts of *Vicia faba*. *The Journal Of General Physiology*, 103(5), 807-831.

**Inoue, H., Nojima, H., and Okayama, H. (1990).** High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96: 23–28.

**Kumimoto, R., Zhang, Y., Siefers, N., & Holt, B. (2010).** NF-YC3, NF-YC4 and NF-YC9 are required for CONSTANS-mediated, photoperiod-dependent flowering in *Arabidopsis thaliana*. *The Plant Journal*, 63(3), 379-391.

**Li, L., Foster, C.M., Gan, Q., Nettleton, D., James, M.G., Myers, A.M., and Wurtele, E.S. (2009).** Identification of the novel protein QQS as a component of the starch metabolic network in *Arabidopsis* leaves. *The Plant Journal* 58: 485–498

**Li, L., Zheng, W., Zhu, Y., Ye, H., Tang, B., & Arendsee, Z. et al. (2015).** QQSorphan gene regulates carbon and nitrogen partitioning across species via NF-YC interactions. *Proceedings Of The National Academy Of Sciences*, 112(47), 14734-14739.

**Morsomme, P., & Boutry, M. (2000).** The plant plasma membrane H<sup>+</sup>-ATPase: structure, function and regulation. *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, 1465(1-2), 1-16.

**Murashige, T. and Skoog, F. (1962).** A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. *Physiologia Plantarum* 15: 473–497.

**Myers, Z., & Holt, B. (2018).** NUCLEAR FACTOR-Y: still complex after all these years?. *Current Opinion In Plant Biology*, 45, 96-102.

**Naranjo, M. A., Romero, C., Bellés, J. M., Montesinos, C., Vicente, O., & Serrano, R.(2003).** Lithium treatment induces a hypersensitive-like response in tobacco. *Planta*, 217(3), 417–424.

**Niittylä, T., Fuglsang, A., Palmgren, M., Frommer, W., & Schulze, W. (2007).** Temporal Analysis of Sucrose-induced Phosphorylation Changes in Plasma Membrane Proteins of *Arabidopsis*. *Molecular & Cellular Proteomics*, 6(10), 1711-1726.

**Obdlik, P., El-Bakkoury, M., Hamacher, T., Cappellaro, C., Vilarino, C., & Fleischer, C. et al. (2004).** K<sup>+</sup> channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. *Proceedings of the National Academy of Sciences* 101: 12242–12247.

**Oñate-Sánchez, L., & Vicente-Carbajosa, J. (2008).** DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Research Notes*, 1(1), 93.

**Palmgren MG, Sommarin M, Serrano R, Larsson C. (1991).** Identification of an autoinhibitory domain in the C-terminal region of the plant plasma membrane H(+)-ATPase. *J Biol Chem*. 266(30):20470-5.

**Riera, M., Irar, S., Vélez-Bermúdez, I., Carretero-Paulet, L., Lumbreras, V., & Pagès, M. (2011).** Role of Plant-Specific N-Terminal Domain of Maize CK2 $\beta$ 1 Subunit in CK2 $\beta$  Functions and Holoenzyme Regulation. *Plos ONE*, 6(7), e21909.

**Robertson, W., Clark, K., Young, J., & Sussman, M. (2004).** An *Arabidopsis thaliana* Plasma Membrane Proton Pump Is Essential for Pollen Development. *Genetics*, 168(3), 1677-1687.

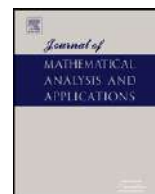
**Rosenquist, M., Sehnke, P., Ferl, R., Sommarin, M., & Larsson, C. (2000).** Evolution of the 14-3-3 Protein Family: Does the Large Number of Isoforms in Multicellular Organisms Reflect Functional Specificity?. *Journal Of Molecular Evolution*, 51(5), 446-458.

**Rudashevskaya, E., Ye, J., Jensen, O., Fuglsang, A., & Palmgren, M. (2011).** Phosphosite Mapping of P-type Plasma Membrane H<sup>+</sup>-ATPase in Homologous and Heterologous Environments. *Journal Of Biological Chemistry*, 287(7), 4904-4913.

**Sanchez, R., & Zhou, M. (2011).** The PHD finger: a versatile epigenome reader. *Trends In Biochemical Sciences*.

**Sun, X., Feng, P., Xu, X., Guo, H., Ma, J., & Chi, W. et al. (2011).** A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. *Nature Communications*, 2(1).

**Yuan, W., Zhang, D., Song, T., Xu, F., Lin, S., & Xu, W. et al. (2017).** *Arabidopsis* plasma membrane H<sup>+</sup>-ATPase genes AHA2 and AHA7 have distinct and overlapping roles in the modulation of root tip H<sup>+</sup> efflux in response to low-phosphorus stress. *Journal Of Experimental Botany*, 68(7), 1731-1741.



## Perturbation analysis of the heat transfer in porous media with small thermal conductivity

F.R. Villatoro<sup>a,\*</sup>, J. Pérez<sup>b,c</sup>, J.L.G. Santander<sup>d</sup>, M.A. Borovsky<sup>e</sup>, Yu.L. Ratis<sup>f</sup>, E.A. Izzheurov<sup>e</sup>, P. Fernández de Córdoba<sup>b,c</sup>

<sup>a</sup> Dpto. de Lenguajes y Ciencias de la Computación, Universidad de Málaga, Málaga, Spain

<sup>b</sup> Dpto. de Matemática Aplicada, Universidad Politécnica de Valencia, Valencia, Spain

<sup>c</sup> Instituto Universitario de Matemática Pura y Aplicada, Universidad Politécnica de Valencia, Valencia, Spain

<sup>d</sup> Universidad Católica de Valencia, C/ Guillem de Castro 106, 46003 Valencia, Spain

<sup>e</sup> Aviation Engines, Samara State Aerospace University named after Academician S.P. Korolev, Samara, Russia

<sup>f</sup> Department of Informatics, Samara State Aerospace University named after Academician S.P. Korolev, Samara, Russia

### ARTICLE INFO

#### Article history:

Received 24 June 2009

Available online 25 August 2010

Submitted by P. Broadbridge

#### Keywords:

Two-phase heat transfer

Perturbation methods

Laplace transform

Low thermal conductivity

### ABSTRACT

An approximate analytical solution for the one-dimensional problem of heat transfer between an inert gas and a porous semi-infinite medium is presented. Perturbation methods based on Laplace transforms have been applied using the solid thermal conductivity as small parameter. The leading order approximation is the solution of Nusselt (or Schumann) problem. Such solution is corrected by means of an outer approximation. The boundary condition at the origin has been taking into account using an inner approximation for a boundary layer. The gas temperature presents a discontinuous front (due to the incompatibility between initial and boundary conditions) which propagates at constant velocity. The solid temperature at the front has been smoothed out using an internal layer asymptotic approximation. The good accuracy of the resulting asymptotic expansion shows its usefulness in several engineering problems such as heat transfer in porous media, in exhausted chemical reactions, mass transfer in packed beds, or in the analysis of capillary electrochromatography techniques.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Nusselt [1,2] obtained the analytical solution for the problem of heat transfer in plug flow heat exchangers when the axial heat conduction is neglected. The same solution, but written in different notation, was rediscovered by Anzelius [3], Schumann [4], and Furnas [5], among others, for the heating (cooling) of one-dimensional porous media (packed bed) by the passage of a hot (cool) fluid. In fact, the corresponding mathematical equations model a large number of physical processes, for example, mass transfer in packed beds, exhausted chemical reactions, capillary electrochromatography, to name only a few [6]. Surprisingly, such a solution has been rediscovered several times during 20th century, using alternative mathematical formulations which are completely equivalent [7,8], stressing its importance and usefulness in practical applications.

The inclusion of a finite axial heat conduction in the Nusselt problem complicates its theoretical analysis. In fact, physical information in Nusselt problem is carried downstream from the inlet by the flow, but information can also travel upstream even for small thermal conductivity. Exact analytical approaches cannot be applied in such a case, hence either numerical or perturbation methods are required. The first ones have been widely applied [9–11], but the second ones are scarce in the

\* Corresponding author. Fax: +34 952132816.

E-mail address: [villa@lcc.uma.es](mailto:villa@lcc.uma.es) (F.R. Villatoro).

scientific literature. Up to the authors knowledge, the only application of perturbation methods in this context is the work of Kuznetsov [12–14] where a perturbation method based on Fourier series is applied to the heating of a two-dimensional rectangular packed beds using a two equation model for the fluid temperature and the difference between temperatures of the fluid and the solid phases. In these works, the inverse of the product of the fluid-to-particle heat transfer coefficient between the solid and fluid phases by the specific surface area is used as small parameter. Under this assumption, the temperature difference between both phases is found to be small compared to the difference between the inlet temperature of the fluid and the initial temperature of solid.

The main contribution in this paper is the development of an asymptotic approximate solution valid for small thermal conductivity under the assumption of a semi-infinite medium. In our approximation the temperature difference between both phases is not assumed to be small. The contents of the paper are as follows. Next section presents the mathematical formulation of the problem and its physical assumptions. Nusselt solution, the leading order solution for both the fluid and the solid phases, is recalled in Section 3. The outer solution is determined in Section 4. Such a solution does not take into account the boundary condition at the origin, so an inner solution is required, which is determined in Section 5. The solution has a front to be dealt with an internal layer which is studied in Section 6. In Section 7 a numerical method based on the method of characteristics is developed in order to validate the accuracy of the asymptotic solution. Finally, Section 8 is devoted to the general conclusions and further lines of research, and Appendix A is included with certain technical details.

## 2. Mathematical model and its assumptions

Let us consider the heat transfer between a fluid and a porous solid under the following assumptions: the fluid is an incompressible Newtonian fluid, with negligible viscous dissipation, negligible heat conduction among the fluid particles, and the fluid motion is only in the axial direction of the solid, from the inlet to the outlet, in quasi-steady conditions; the solid has a constant porosity and negligible radial temperature gradient, with only an axial temperature gradient. This two-phase problem is modelled through the system of coupled partial differential equations [15],

$$p\rho_f c_f \left( \frac{\partial T_f}{\partial t} + v_f \frac{\partial T_f}{\partial x} \right) = -h(T_f - T_s), \quad (1)$$

$$(1-p)\rho_s c_s \frac{\partial T_s}{\partial t} = h(T_f - T_s) + (1-p)\lambda_s \frac{\partial^2 T_s}{\partial x^2}, \quad (2)$$

where  $x$  is the axial position along the medium,  $t$  is time, the subscripts  $f$  and  $s$  stand for the fluid and the solid matrix, respectively,  $T$  is the temperature,  $\rho$  is the density,  $c_s$  is the specific heat of the solid,  $c_f$  is the specific heat at constant pressure of the fluid,  $v_f$  is the fluid flow velocity,  $h$  is the heat transfer coefficient between the fluid and the solid,  $\lambda_s$  is the thermal conductivity, and  $p$  is the connected void fraction of the solid, i.e.,  $(1-p)$  is its “effective” porosity.

Let us assume, as initial condition, that both phases start in thermal equilibrium, i.e., with the same temperature equal to the ambient one ( $T_{s_0}$ ),

$$T_f(0, x) = T_{s_0}, \quad x > 0, \quad (3)$$

$$T_s(0, x) = T_{s_0}, \quad x > 0. \quad (4)$$

Let us also assume that the fluid is injected in the solid at  $x = 0$ , with constant flow velocity and temperature,

$$T_f(t, 0) = T_{f_0}, \quad t > 0, \quad (5)$$

and apply a Robin boundary condition at the inlet boundary of the solid matrix given by

$$\lambda_s \frac{\partial T_s}{\partial x}(t, 0) = h_b(T_s(t, 0) - T_f(t, 0)), \quad (6)$$

where  $h_b$  is the boundary heat transfer coefficient between the fluid and the solid. Finally, let us also assume that the length of the solid medium ( $L$ ) is large enough such that it can be considered as semi-infinite ( $L \rightarrow \infty$ ). In such a case, the outlet boundary condition is given by

$$\lim_{x \rightarrow \infty} T_s(t, x) = \lim_{x \rightarrow \infty} T_f(t, x) = 0. \quad (7)$$

Eqs. (1)–(7) may be nondimensionalized by introducing the following variables

$$\tau = \frac{ht}{(1-p)\rho_s c_s}, \quad y = \frac{nhx}{p\rho_f c_f v_f},$$

where

$$n = \frac{p\rho_f c_f}{(1-p)\rho_s c_s},$$



and defining  $\chi(\tau, y)$  and  $\varepsilon(\tau, y)$  through the relationships

$$T_s = T_{s_0} + \chi(\tau, y)(T_{f_0} - T_{s_0}),$$

$$T_f = T_{s_0} + \varepsilon(\tau, y)(T_{f_0} - T_{s_0}),$$

resulting in the dimensionless form

$$n\left(\frac{\partial \varepsilon}{\partial \tau} + \frac{\partial \varepsilon}{\partial y}\right) = \chi - \varepsilon, \tag{8}$$

$$\frac{\partial \chi}{\partial \tau} - \beta^2 \frac{\partial^2 \chi}{\partial y^2} = \varepsilon - \chi, \tag{9}$$

with initial conditions

$$\varepsilon(0, y) = 0, \quad \chi(0, y) = 0, \quad y > 0, \tag{10}$$

and boundary conditions

$$\varepsilon(\tau, 0) = 1, \tag{11}$$

$$\frac{\partial \chi}{\partial y}(\tau, 0) = \gamma(\chi(\tau, 0) - \varepsilon(\tau, 0)), \tag{12}$$

$$\lim_{y \rightarrow \infty} \chi(\tau, y) = \lim_{y \rightarrow \infty} \varepsilon(\tau, y) = 0, \tag{13}$$

where, in Eqs. (8)–(13),

$$\beta^2 = (1 - p)\lambda_s h \left(\frac{n}{p\rho_f c_f v_f}\right)^2 \quad \text{and} \quad \gamma = \frac{h_b p \rho_f c_f v_f}{h n \lambda_s}.$$

Note that  $h_b \geq 0$  implies that  $\gamma \geq 0$ .

In order to solve the quarterplane problem given by Eqs. (8)–(13), the Laplace transform method can be used. Let  $\hat{\varepsilon}(s, y)$  and  $\hat{\chi}(s, y)$  be the Laplace transform on  $\tau$  of  $\varepsilon(\tau, y)$  and  $\chi(\tau, y)$ , respectively. The Laplace transform of Eqs. (8)–(13) yields

$$n\left(s\hat{\varepsilon}(s, y) + \frac{\partial \hat{\varepsilon}(s, y)}{\partial y}\right) = \hat{\chi}(s, y) - \hat{\varepsilon}(s, y), \tag{14}$$

$$s\hat{\chi} - \beta^2 \frac{\partial^2 \hat{\chi}}{\partial y^2} = \hat{\varepsilon}(s, y) - \hat{\chi}(s, y), \tag{15}$$

with boundary conditions

$$\hat{\varepsilon}(s, 0) = \frac{1}{s}, \tag{16}$$

$$\frac{\partial \hat{\chi}(s, 0)}{\partial y} = \gamma\left(\hat{\chi}(s, 0) - \frac{1}{s}\right). \tag{17}$$

The analytical solution of the problem in Laplace space given by Eqs. (14)–(17) can be obtained in implicit form by means of using Mathematica's DSolve command yielding

$$\hat{\varepsilon}(s, y) = \sum_{i=1}^3 e^{\frac{R_i y}{\beta^2 n}} \left( \frac{\beta^2 n(\gamma(ns + 1)\beta^2 + n + \gamma R_i)}{\beta^2(n(sn + n - 2R_i s) - 2R_i) - 3R_i^2} - \frac{(\gamma n\beta^2 + R_i)s((ns + 1)\beta^2 + R_i)\hat{C}}{\beta^2(n(sn + n - 2R_i s) - 2R_i) - 3R_i^2} \right),$$

$$\hat{\chi}(s, y) = \sum_{i=1}^3 e^{\frac{R_i y}{\beta^2 n}} \left( \frac{\beta^2 n(\gamma(ns + 1)\beta^2 + n + \gamma R_i)}{\beta^2(n(sn + n - 2R_i s) - 2R_i) - 3R_i^2} - \frac{(\gamma n\beta^2 + R_i)s((ns + 1)\beta^2 + R_i)\hat{C}}{\beta^2(n(sn + n - 2R_i s) - 2R_i) - 3R_i^2} \right),$$

where  $R_i$  are the roots of the cubic polynomial

$$n^2 s(sn + n + 1)\beta^4 + R(n(sn + n - Rs) - R)\beta^2 - R^3 = 0,$$

and  $\hat{C}$  is a constant of integration to be determined by the application of Eq. (13). Up to the author's knowledge, the inverse Laplace transform of this solution cannot be obtained analytically and even, if possible, the resulting expression will be very cumbersome to be useful in practice. The authors have tried, without success, the use of a numerical implementation of the Laplace transform inversion (all the algorithms implemented in Ref. [16]). Instead, the numerical integration is straightforward in the time domain [17], as shown in the next section.

The main goal of this paper is the development of an approximate asymptotic solution for both the gas and the solid temperatures when  $\beta^2 \ll 1$  based on the, widely known, exact solution for  $\beta = 0$  of Eqs. (8)–(12) and the use of singular perturbation methods [18,19]. Numerical results, to be presented in Section 7, indicate that the solid temperature for  $\beta^2 > 0$  presents a boundary layer of size  $O(\beta)$  at  $y = 0$  due to the Robin boundary condition. Moreover, at  $y = \tau$ , due to the infinite speed of propagation of the heat in the solid, the sharp front propagating at constant speed in Nusselt solution is smoothed out by an interior layer. The asymptotic method has to deal with both boundary and interior layer approximations.

### 3. Leading order solution for $\beta = 0$

Eqs. (8)–(13) for  $\beta = 0$  are usually referred to as either Nusselt or Schumann problem in the quarterplane, whose exact solution is widely known. Let  $\chi_0$  and  $\varepsilon_0$  be such solution, and  $\hat{\varepsilon}_0(s, y)$  and  $\hat{\chi}_0(s, y)$ , respectively, their Laplace transforms in  $\tau$ . The solution of Eqs. (14)–(16) for  $\beta = 0$  yields

$$\hat{\varepsilon}_0(s, y) = \frac{e^{-ys}}{s} \exp\left(-\frac{sy}{n(s+1)}\right), \quad (18)$$

$$\hat{\chi}_0(s, y) = \frac{e^{-ys}}{s(s+1)} \exp\left(-\frac{sy}{n(s+1)}\right). \quad (19)$$

The evaluation of the Bromwich integral for the inverse Laplace transform of Eqs. (18) and (19) yields [6, Eqs. (3)–(5)]

$$\varepsilon_0(\tau, y) = \theta(\tau - y)e^{-y/n} \left\{ e^{-\tau+y} I_0\left(2\sqrt{\frac{y(\tau-y)}{n}}\right) + \int_0^{\tau-y} e^{-u} I_0\left(2\sqrt{\frac{yu}{n}}\right) du \right\}, \quad (20)$$

$$\chi_0(\tau, y) = \theta(\tau - y) \exp\left(-\frac{y}{n}\right) \int_0^{\tau-y} \exp(-u) I_0\left(2\sqrt{\frac{yu}{n}}\right) du, \quad (21)$$

and

$$\varepsilon_0(\tau, y) - \chi_0(\tau, y) = \theta(\tau - y) \exp\left(y - \frac{y}{n} - \tau\right) I_0\left(2\sqrt{\frac{(\tau-y)y}{n}}\right). \quad (22)$$

Integration by parts in Eq. (20), recalling that  $I_0' = I_1$ , and  $I(0) = 1$ , results in

$$\varepsilon_0(\tau, y) = \theta(\tau - y) \exp\left(-\frac{y}{n}\right) \left\{ 1 + \int_0^{\tau-y} \exp(-u) I_1\left(2\sqrt{\frac{yu}{n}}\right) \sqrt{\frac{y}{un}} du \right\}. \quad (23)$$

Figs. 2 and 3, left plots, show some examples of the numerical evaluation of Eqs. (21) and (23). Further plots and comments can be found elsewhere [8,20].

### 4. Outer approximation

Let us apply a singular perturbation method to Eqs. (14)–(17) in Laplace space and then consider the Laplace inversion of the resulting solution. The results of the numerical simulations shown in Section 7 suggests the appearance of a boundary layer at  $y = 0$  with a width of  $O(\beta)$ , hence both outer and inner asymptotic expansions are required. Let us take

$$\hat{\varepsilon}(s, y) = \hat{\varepsilon}^{(o)}(s, y) + \hat{\varepsilon}^{(i)}(s, Y), \quad \hat{\chi}(s, y) = \hat{\chi}^{(o)}(s, y) + \hat{\chi}^{(i)}(s, Y), \quad (24)$$

where the stretching variable  $Y = y/\beta$  is suggested by the numerical results, and the superindexes  $(o)$  and  $(i)$  correspond to, respectively, outer and inner expansions.

Let us formally expand the outer solutions as function of  $\beta^2$  as

$$\hat{\varepsilon}^{(o)}(s, y) = \hat{\varepsilon}_0(s, y) + \beta^2 \hat{\varepsilon}_1(s, y) + \beta^4 \hat{\varepsilon}_2(s, y) + O(\beta^6), \quad (25)$$

$$\hat{\chi}^{(o)}(s, y) = \hat{\chi}_0(s, y) + \beta^2 \hat{\chi}_1(s, y) + \beta^4 \hat{\chi}_2(s, y) + O(\beta^6). \quad (26)$$

Substituting the outer solutions into both governing equations and boundary conditions, and comparing the coefficients of powers of  $\beta^2$  yields, at leading order,

$$n \left( s \hat{\varepsilon}_0(s, y) + \frac{\partial \hat{\varepsilon}_0(s, y)}{\partial y} \right) = \hat{\chi}_0(s, y) - \hat{\varepsilon}_0(s, y), \quad (27)$$

$$s \hat{\chi}_0 = \hat{\varepsilon}_0(s, y) - \hat{\chi}_0(s, y), \quad (28)$$

$$\hat{\varepsilon}_0(s, 0) = \frac{1}{s}, \quad (29)$$

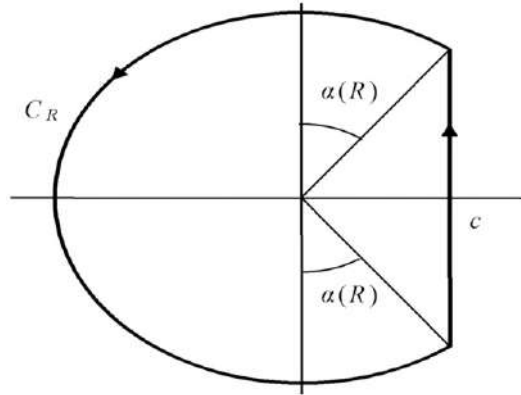


Fig. 1. Simple closed contour C.

and at  $O(\beta^{2j})$ ,

$$n \left( s\hat{\varepsilon}_j(s, y) + \frac{\partial \hat{\varepsilon}_j(s, y)}{\partial y} \right) = \hat{\chi}_j(s, y) - \hat{\varepsilon}_j(s, y), \tag{30}$$

$$s\hat{\chi}_j = \hat{\varepsilon}_j(s, y) - \hat{\chi}_j(s, y) + \frac{\partial^2 \hat{\chi}_{j-1}}{\partial y^2}, \tag{31}$$

$$\hat{\varepsilon}_j(s, 0) = 0, \tag{32}$$

for  $j = 1, 2, \dots$ , where the Robin boundary condition (17) cannot be satisfied, resulting in a boundary layer appearing at  $y = 0$  to be dealt with the inner expansion of the solution.

The solution of Eqs. (27)–(29) yields Eqs. (18) and (19), i.e., Nusselt’s solution. For  $j = 1$ , the solution of Eqs. (30)–(32) yields

$$\begin{aligned} \hat{\varepsilon}_1(s, y) &= \frac{s^2(sn + n + 1)^2}{n^3(s + 1)^4} y \hat{\varepsilon}_0(s, y) \\ &= \frac{sy}{n(s + 1)^2} \left( 1 + \frac{2}{n(s + 1)} + \frac{1}{n^2(s + 1)^2} \right) \exp\left(-\frac{sy}{n} \left( n + \frac{1}{s + 1} \right)\right), \end{aligned} \tag{33}$$

$$\hat{\chi}_1(s, y) = \frac{s^2(sn + n + 1)^2(sn + n + y)}{n^3(s + 1)^5} \hat{\varepsilon}_0(s, y). \tag{34}$$

High-order terms in the expansions (25) and (26) can also be determined straightforwardly.

The Laplace inversion of Eqs. (33) and (34) can proceed as follows. By applying the inversion formula of the Laplace transform to Eq. (33), then

$$\begin{aligned} \varepsilon_1(\tau, y) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \exp(s\tau) \hat{\varepsilon}_1(s, y) ds \\ &= \theta(\tau - y) \frac{1}{2\pi i} \int_C \frac{sy}{n(s + 1)^2} \exp\left(s(\tau - y) - \frac{sy}{n(s + 1)}\right) \left\{ 1 + \frac{2}{n(s + 1)} + \frac{1}{n^2(s + 1)^2} \right\} ds, \end{aligned} \tag{35}$$

and  $C$  is the simple closed contour shown in Fig. 1, with  $c > 0$ . The integration in the  $C_R$  arc of  $C$  is null if and only if  $\tau - y > 0$  as shown in Appendix A.

Recalling that the generating function of the modified Bessel functions  $I_k$  is given by [23, p. 376]

$$\exp\left(\frac{z}{2} \left( \sigma + \frac{1}{\sigma} \right)\right) = \sum_{k=-\infty}^{\infty} \sigma^k I_k(z),$$

then the exponential function in the integrand of Eq. (35) may be written as

$$\begin{aligned} \exp\left((s+1-1)(\tau-y) - \frac{(s+1-1)y}{n(s+1)}\right) &= \exp\left(y-\tau - \frac{y}{n}\right) \exp\left((s+1)(\tau-y) + \frac{y}{n(s+1)}\right) \\ &= \exp\left(y-\tau - \frac{y}{n}\right) \sum_{k=-\infty}^{\infty} \sigma^k I_k(z), \end{aligned} \quad (36)$$

where

$$z = 2\sqrt{\frac{(\tau-y)y}{n}}, \quad \sigma = \frac{s+1}{g}, \quad g = \sqrt{\frac{y}{n(\tau-y)}}. \quad (37)$$

The substitution of Eq. (36) in Eq. (35) results in

$$\begin{aligned} \varepsilon_1(\tau, y) &= \theta(\tau-y) \frac{e^{y-\tau-y/n}}{2\pi i} \int_C \frac{1}{n(g\sigma)^2} \left( \sum_{k=-\infty}^{\infty} \sigma^k I_k(z) \right) \left\{ (g\sigma-1)y \left( 1 + \frac{2}{ng\sigma} + \frac{1}{(ng\sigma)^2} \right) \right\} g d\sigma \\ &= \frac{y}{n} \theta(\tau-y) \frac{e^{y-\tau-y/n}}{2\pi i} \int_C \left( \sum_{k=-\infty}^{\infty} \sigma^k I_k(z) \right) \left\{ \frac{1}{\sigma} + \frac{(2-n)}{gn} \frac{1}{\sigma^2} - \frac{(2n-1)}{(gn)^2} \frac{1}{\sigma^3} - \frac{1}{g^3 n^2} \frac{1}{\sigma^4} \right\} d\sigma, \end{aligned} \quad (38)$$

which can be expanded to

$$\begin{aligned} \varepsilon_1(\tau, y) &= \frac{y}{n} \theta(\tau-y) e^{y-\tau-y/n} \left\{ \frac{1}{2\pi i} \int_C \left( \sum_{k=-\infty}^{\infty} \sigma^{k-1} I_k(z) \right) d\sigma + \frac{(2-n)}{gn} \frac{1}{2\pi i} \int_C \left( \sum_{k=-\infty}^{\infty} \sigma^{k-2} I_k(z) \right) d\sigma \right. \\ &\quad \left. - \frac{(2n-1)}{(gn)^2} \frac{1}{2\pi i} \int_C \left( \sum_{k=-\infty}^{\infty} \sigma^{k-3} I_k(z) \right) d\sigma - \frac{1}{g^3 n^2} \frac{1}{2\pi i} \int_C \left( \sum_{k=-\infty}^{\infty} \sigma^{k-4} I_k(z) \right) d\sigma \right\}. \end{aligned} \quad (39)$$

The use of Cauchy's residue theorem [24] yields

$$\varepsilon_1(\tau, y) = \frac{y}{n} \theta(\tau-y) e^{y-\tau-y/n} \left\{ I_0(z) + \frac{2-n}{gn} I_1(z) - \frac{2n-1}{g^2 n^2} I_2(z) - \frac{1}{g^3 n^2} I_3(z) \right\}. \quad (40)$$

The application of the same procedure to Eq. (34), omitting further details, results in

$$\begin{aligned} \chi_1(\tau, y) &= \theta(\tau-y) e^{y-\tau-y/n} \left\{ I_0(z) + \frac{2-n+y}{gn} I_1(z) - \frac{y}{g^4 n^3} I_4(z) + \frac{1+2y-n(2+y)}{g^2 n^2} I_2(z) \right. \\ &\quad \left. + \frac{y-n(1+2y)}{g^3 n^3} I_3(z) \right\}. \end{aligned} \quad (41)$$

## 5. Inner approximation at $y=0$

Let us determine the inner approximation  $\hat{\varepsilon}^{(i)}(s, Y)$ , where  $Y = y/\beta$ , and  $\partial/\partial y = (1/\beta)\partial/\partial Y$ . This solution deals with the Robin boundary condition (17), which has not been taken into account by the outer approximation determined in the last section. The equations for the inner solution are obtained by substitution of Eq. (24) into Eqs. (13)–(17), yielding

$$n \left( s\beta \hat{\varepsilon}^{(i)}(s, Y) + \frac{\partial \hat{\varepsilon}^{(i)}(s, Y)}{\partial Y} \right) = \beta \left( \hat{\chi}^{(i)}(s, Y) - \hat{\varepsilon}^{(i)}(s, Y) \right), \quad (42)$$

$$s \hat{\chi}^{(i)} - \frac{\partial^2 \hat{\chi}^{(i)}}{\partial Y^2} = \hat{\varepsilon}^{(i)}(s, Y) - \hat{\chi}^{(i)}(s, Y), \quad (43)$$

$$\hat{\varepsilon}^{(i)}(s, 0) = 0, \quad (44)$$

$$\frac{\partial \hat{\chi}^{(i)}(s, 0)}{\partial Y} = \beta \gamma \left( \hat{\chi}^{(i)}(s, 0) - \frac{1}{s} \right) + \beta \left( \gamma \hat{\chi}^{(0)}(s, 0) - \frac{\partial \hat{\chi}^{(0)}(s, 0)}{\partial y} \right), \quad (45)$$

$$\lim_{Y \rightarrow \infty} \hat{\varepsilon}^{(i)}(s, Y) = \lim_{Y \rightarrow \infty} \hat{\chi}^{(i)}(s, Y) = 0. \quad (46)$$

Formally expanding the inner solution as function of  $\beta$  yields

$$\hat{\varepsilon}^{(i)}(s, Y) = \beta \tilde{\varepsilon}_1(s, Y) + \beta^2 \tilde{\varepsilon}_2(s, Y) + \beta^3 \tilde{\varepsilon}_3(s, Y) + O(\beta^4), \quad (47)$$

$$\hat{\chi}^{(i)}(s, Y) = \beta \tilde{\chi}_1(s, Y) + \beta^2 \tilde{\chi}_2(s, Y) + \beta^3 \tilde{\chi}_3(s, Y) + O(\beta^4). \quad (48)$$

Substituting these expression into Eqs. (42)–(45) and comparing the coefficients of powers of  $\beta$  yields, at  $O(\beta)$ ,

$$\frac{\partial \tilde{\varepsilon}_1(s, Y)}{\partial Y} = 0, \tag{49}$$

$$s\tilde{\chi}_1(s, Y) - \frac{\partial^2 \tilde{\chi}_1(s, Y)}{\partial Y^2} = \tilde{\varepsilon}_1(s, Y) - \tilde{\chi}_1(s, Y), \tag{50}$$

$$\tilde{\varepsilon}_1(s, 0) = 0, \tag{51}$$

$$\frac{\partial \tilde{\chi}_1(s, 0)}{\partial Y} = -\frac{\gamma}{s} - \gamma \hat{\chi}_0(s, 0) - \frac{\partial \hat{\chi}_0(s, 0)}{\partial y} = \frac{1 - \gamma}{s + 1} + \frac{1}{n(s + 1)^2}, \tag{52}$$

$$\lim_{Y \rightarrow \infty} \tilde{\varepsilon}_1(s, Y) = \lim_{Y \rightarrow \infty} \tilde{\chi}_1(s, Y) = 0, \tag{53}$$

and, at  $O(\beta^j)$ , for  $j = 1, 2, 3, \dots$ ,

$$n \frac{\partial \tilde{\varepsilon}_j(s, Y)}{\partial Y} = -n s \tilde{\varepsilon}_{j-1}(s, Y) + \tilde{\chi}_{j-1}(s, Y) - \tilde{\varepsilon}_{j-1}(s, Y), \tag{54}$$

$$s\tilde{\chi}_j(s, Y) - \frac{\partial^2 \tilde{\chi}_j(s, Y)}{\partial Y^2} = \tilde{\varepsilon}_j(s, Y) - \tilde{\chi}_j(s, Y), \tag{55}$$

$$\tilde{\varepsilon}_j(s, 0) = 0, \tag{56}$$

$$\lim_{Y \rightarrow \infty} \tilde{\varepsilon}_j(s, Y) = \lim_{Y \rightarrow \infty} \tilde{\chi}_j(s, Y) = 0, \tag{57}$$

with Eq. (45) yielding for even  $j$ ,

$$\frac{\partial \tilde{\chi}_j(s, 0)}{\partial Y} = \gamma \tilde{\chi}_{j-1}(s, 0), \tag{58}$$

and for odd  $j = 2k + 1$ ,

$$\frac{\partial \tilde{\chi}_j(s, 0)}{\partial Y} = \gamma \tilde{\chi}_{j-1}(s, 0) + \gamma \hat{\chi}_k(s, 0) - \frac{\partial \hat{\chi}_k(s, 0)}{\partial y}. \tag{59}$$

The solution of Eqs. (49)–(53) yields

$$\tilde{\varepsilon}_1(s, Y) = 0, \tag{60}$$

$$\tilde{\chi}_1(s, Y) = \frac{(1 - (\gamma - 1)n(s + 1)) \sinh(Y\sqrt{s + 1})}{n(s + 1)^{5/2}}, \tag{61}$$

and the solution of Eqs. (54)–(57), for  $j = 2$ , gives

$$\tilde{\varepsilon}_2(s, Y) = 0, \tag{62}$$

$$\tilde{\chi}_2(s, Y) = \frac{((\gamma - 1)n(s + 1) - 1)Y \sinh(Y\sqrt{s + 1})}{2n^2(s + 1)^{7/2}}, \tag{63}$$

where the integration constants were selected in order to satisfy Eq. (57). The expressions for the solution with  $j > 2$  of Eqs. (54)–(57) can be easily obtained, although they are cumbersome and long.

The inverse Laplace transform of Eqs. (60)–(63), after using elementary Laplace transform properties and tables, yields

$$\tilde{\varepsilon}_1^{(i)}(t, Y) = 0, \tag{64}$$

$$\tilde{\chi}_1^{(i)}(t, Y) = e^{-t} \operatorname{erfc}\left(\frac{Y}{2\sqrt{t}}\right) \frac{Y(Y^2 - 6(\gamma - 1)n + 6t)}{6n} - e^{-\frac{Y^2}{4t} - t} \sqrt{t} \frac{Y^2 - 6(\gamma - 1)n + 4t}{3n\sqrt{\pi}}, \tag{65}$$

$$\tilde{\varepsilon}_2^{(i)}(t, Y) = 0, \tag{66}$$

and

$$\begin{aligned} \tilde{\chi}_2^{(i)}(t, Y) = & \frac{e^{-\frac{Y^2}{4t} - t} \sqrt{t} Y}{120n^2 \sqrt{\pi}} ((Y^2 + 2t)(Y^2 + 16t) - 20(\gamma - 1)n(Y^2 + 4t)) \\ & - \frac{e^{-t} Y^2}{240n^2} \operatorname{erfc}\left(\frac{Y}{2\sqrt{t}}\right) \left( Y^4 + 20(-\gamma n + n + t)Y^2 + 60t(t - 2(\gamma - 1)n) \right). \end{aligned} \tag{67}$$

Similarly, high-order terms for  $\tilde{\varepsilon}_j^{(i)}(t, Y)$  and  $\tilde{\chi}_j^{(i)}(t, Y)$  can be obtained.

**6. Internal layer approximation at  $y = \tau$**

First, let us introduce some notations. Given a bounded function  $f(z)$ , let us define

$$[f](z) = f(z^+) - f(z^-) = \lim_{\theta \uparrow z} f(\theta) - \lim_{\theta \downarrow z} f(\theta) = f^+(z) - f^-(z),$$

where  $f^+(z) = f(z)$  for  $z > 0$ , and  $f^-(z) = f(z)$  for  $z < 0$ . These definitions extends straightforwardly to functions of two variables.

The solid temperature  $\xi(\tau, y)$ , being the solution of a parabolic equation, must be smooth. However, the outer solution presents a nonsmooth front at  $y = \tau$ ; it is continuous but their derivatives have a jump discontinuity

$$\left[ \frac{\partial^k \chi^{(0)}}{\partial y^k} \right](\tau, \tau) = \frac{\partial^k \chi^{(0)}}{\partial y^k}(\tau, \tau^+) - \frac{\partial^k \chi^{(0)}}{\partial y^k}(\tau, \tau^-) \neq 0, \quad k \geq 1. \tag{68}$$

To deal with such a difficulty, an internal layer with a width of  $O(\beta)$  must be introduced and the method developed in Ref. [21] can be applied. Let us introduce the change of variables  $\xi = (y - \tau)/\beta$  and take

$$\begin{aligned} \varepsilon(\tau, y) &= \varepsilon^{(0)}(\tau, y) + \varepsilon^{(i)}(\tau, Y) + \varepsilon^{(l)}(\tau, \xi), \\ \chi(\tau, y) &= \chi^{(0)}(\tau, y) + \chi^{(i)}(\tau, Y) + \chi^{(l)}(\tau, \xi), \end{aligned}$$

where the internal layer solutions are given by

$$\varepsilon^{(l)}(\tau, \xi) = \bar{\varepsilon}_0(\tau, \xi) + \beta \bar{\varepsilon}_1(\tau, \xi) + \beta^2 \bar{\varepsilon}_2(\tau, \xi) + O(\beta^3), \tag{69}$$

$$\chi^{(l)}(\tau, \xi) = \bar{\chi}_0(\tau, \xi) + \beta \bar{\chi}_1(\tau, \xi) + \beta^2 \bar{\chi}_2(\tau, \xi) + O(\beta^3). \tag{70}$$

Theorem 2 in Ref. [21] is applicable in our case resulting in, for even  $m$ ,

$$\bar{\varepsilon}_m^+(\tau, \xi) = -\bar{\varepsilon}_m^-(\tau, -\xi), \quad \bar{\chi}_m^+(\tau, \xi) = -\bar{\chi}_m^-(\tau, -\xi), \tag{71}$$

both for  $\xi > 0$ , and, for odd  $m$ ,

$$\bar{\varepsilon}_m^+(\tau, \xi) = \bar{\varepsilon}_m^-(\tau, -\xi), \quad \bar{\chi}_m^+(\tau, \xi) = \bar{\chi}_m^-(\tau, -\xi), \tag{72}$$

also both for  $\xi > 0$ . Therefore, we will determine only the internal layer solutions for  $\xi > 0$ , i.e.  $\bar{\varepsilon}_k^+$  and  $\bar{\chi}_k^+$ , and then use either Eq. (71) or (72) for  $\xi < 0$ . This is the simplest procedure since  $\varepsilon^{(0)}(\tau, y) = \chi^{(0)}(\tau, y) = 0$  for  $y > \tau$ . Here on, in order to simplify the notation, the superindex  $+$  will be omitted.

Substituting the interior layer expansions (69) and (70) into Eqs. (8)–(13) yields for  $\bar{\varepsilon}_0(\tau, \xi)$  and  $\bar{\chi}_0(\tau, \xi)$ , the equations given by

$$n \frac{\partial \bar{\varepsilon}_0}{\partial \xi} = 0, \quad \frac{\partial \bar{\chi}_0}{\partial \tau} - \frac{\partial^2 \bar{\chi}_0}{\partial \xi^2} + \bar{\chi}_0 = \bar{\varepsilon}_0, \tag{73}$$

with initial conditions

$$\bar{\varepsilon}_0(0, \xi) = 0, \quad \bar{\chi}_0(0, \xi) = 0, \quad \xi > 0, \tag{74}$$

and boundary conditions

$$\bar{\chi}_0(\tau, 0) = [\chi_0](\tau, \tau), \quad \lim_{\xi \rightarrow \infty} \bar{\chi}_0(\tau, \xi) = 0, \tag{75}$$

and similarly, for  $\bar{\varepsilon}_m(\tau, \xi)$  and  $\bar{\chi}_m(\tau, \xi)$ , with  $m = 1, 2, \dots$ , the following equations result

$$n \frac{\partial \bar{\varepsilon}_m}{\partial \xi} = \bar{\chi}_{m-1} - \bar{\varepsilon}_{m-1} - n \frac{\partial \bar{\varepsilon}_{m-1}}{\partial \tau}, \tag{76}$$

$$\frac{\partial \bar{\chi}_m}{\partial \tau} - \frac{\partial^2 \bar{\chi}_m}{\partial \xi^2} + \bar{\chi}_m = \bar{\varepsilon}_m, \tag{77}$$

with initial conditions

$$\bar{\varepsilon}_m(0, \xi) = 0, \quad \bar{\chi}_m(0, \xi) = 0, \quad \xi > 0, \tag{78}$$

boundary conditions

$$\lim_{\xi \rightarrow \infty} \bar{\chi}_m(\tau, \xi) = 0 = \lim_{\xi \rightarrow \infty} \bar{\varepsilon}_m(\tau, \xi), \tag{79}$$

and, for even  $m = 2k$ ,

$$\bar{\chi}_{2k}(\tau, 0) = -\frac{1}{2}[\chi_k](\tau, \tau), \tag{80}$$

and for odd  $m = 2k - 1$ ,

$$\frac{\partial \bar{\chi}_{2k+1}}{\partial \xi}(\tau, 0) = -\frac{1}{2}\left[\frac{\partial \chi_k}{\partial \xi}\right](\tau, \tau). \tag{81}$$

Note that the jumps in  $\chi_k$  and its first spatial derivative are smoothed out by using, respectively, Dirichlet (Eq. (80)) and Neumann (Eq. (81)) boundary conditions for the internal layer approximations of, respectively, even and odd order, as suggested in Ref. [21].

The solution of Eq. (76) is given by

$$\bar{\varepsilon}_m(\tau, \xi) = \frac{1}{n}\bar{\varepsilon}_m(\tau, 0) + \frac{1}{n}\int_0^\xi \left(\bar{\chi}_{m-1}(\tau, \eta) - \bar{\varepsilon}_{m-1}(\tau, \eta) - n\frac{\partial \bar{\varepsilon}_{m-1}}{\partial \tau}(\tau, \eta)\right) d\eta, \tag{82}$$

where

$$\bar{\varepsilon}_m(\tau, 0) = -\int_0^\infty \left(\bar{\chi}_{m-1}(\tau, \eta) - \bar{\varepsilon}_{m-1}(\tau, \eta) - n\frac{\partial \bar{\varepsilon}_{m-1}}{\partial \tau}(\tau, \eta)\right) d\eta,$$

such that the boundary condition (79) is satisfied. Note that the initial condition (78) is also satisfied if both  $\bar{\varepsilon}_{m-1}$  and  $\bar{\chi}_{m-1}$  do it.

The solution  $\bar{\chi}_m(\tau, y)$  of Eqs. (77)–(79), for even  $m = 2k$ , when Eq. (80) is satisfied, is given by [22, Section 1.1.3-4]

$$\bar{\chi}_m(\tau, \xi) = -\frac{1}{2}\int_0^\tau [\chi_k](t, t)\frac{\xi}{\tau-t}G(\xi, 0, \tau-t) dt + \int_0^\tau \int_0^\infty \bar{\varepsilon}_m(t, \eta)(G(\xi, \eta, t) - G(\xi, -\eta, t)) d\eta dt, \tag{83}$$

and for odd  $m = 2k + 1$ , when Eq. (81) applies, is given by [22, Section 1.1.3-5]

$$\bar{\chi}_m(\tau, \xi) = \int_0^\tau \left[\frac{\partial \chi_k}{\partial \xi}\right](t, t)G(\xi, 0, \tau-t) dt + \int_0^\tau \int_0^\infty \bar{\varepsilon}_m(t, \eta)(G(\xi, \eta, t) + G(\xi, -\eta, t)) d\eta dt, \tag{84}$$

where, in both Eqs. (83) and (84),

$$G(\xi, \eta, t) = \frac{e^{-t}}{2\sqrt{\pi t}} \exp\left(-\frac{(\xi - \eta)^2}{4t}\right).$$

The solutions of Eqs. (73)–(75) are  $\bar{\varepsilon}_0(\tau, \xi) \equiv \bar{\chi}_0(\tau, \xi) \equiv 0$ , since  $[\chi_0](\tau, \tau) = 0$ , cf. Eq. (21). Hence,  $\bar{\varepsilon}_1(\tau, \xi) \equiv 0$ . Since Eq. (21) yields

$$\left[\frac{\partial \chi_0}{\partial \xi}\right](\tau, \tau) = -e^{-\tau/n},$$

then

$$\bar{\chi}_1(\tau, \xi) = \frac{e^{-\tau/n}}{2\sqrt{\pi}} \int_0^\tau \frac{1}{\sqrt{z}} \exp\left(\frac{z}{n} - z - \frac{\xi^2}{4z}\right) dz. \tag{85}$$

Following the same procedure results in

$$\bar{\varepsilon}_2(\tau, \xi) = \bar{\varepsilon}_2(\tau, 0) - \frac{e^{-\tau/n}}{2n} \int_0^\tau \exp\left(-z + \frac{z}{n}\right) \operatorname{erf}\left(\frac{\xi}{2\sqrt{z}}\right) dz, \tag{86}$$

where

$$\bar{\varepsilon}_2(\tau, 0) = \begin{cases} \frac{\tau}{2}e^{-\tau}, & n = 1, \\ \frac{1}{2(n-1)}(e^{-\tau/n} - e^{-\tau}), & n \neq 1. \end{cases}$$

The solution  $\bar{\chi}_2(\tau, \xi)$  can be determined by taking into account that, by using Eq. (41),

$$[\chi_1](\tau, \tau) = e^{-\tau/n},$$

resulting in

$$\begin{aligned} \bar{\chi}_2(\tau, \xi) &= \frac{\xi e^{-\tau/n}}{4\sqrt{\pi}} \int_0^\tau z^{-3/2} \exp\left(-z + \frac{z}{n} - \frac{\xi^2}{4z}\right) dz \\ &+ \int_0^\tau \int_0^\infty \frac{\bar{\varepsilon}_2(\tau - z, \eta)}{2\sqrt{\pi z}} (e^{\eta\xi/z} - 1) e^{-z} \exp\left(-\frac{(\eta + \xi)^2}{4z}\right) d\eta dz. \end{aligned} \quad (87)$$

Similarly, high-order terms for  $\varepsilon^{(l)}(\tau, \xi)$  and  $\chi^{(l)}(\tau, \xi)$  can be obtained, although the resulting expressions become more cumbersome as the order grows.

## 7. Presentation of results

In order to validate the asymptotic method developed in this paper, let us numerically solve Eqs. (8)–(13) as follows. The domain  $(\tau, y) \in [0, T] \times [0, L]$  is meshed by using  $\tau_m = m\Delta\tau$ ,  $m = 0, 1, \dots, M$ , and  $y_j = j\Delta y$ ,  $j = 0, 1, \dots, J$ , where  $\Delta\tau = T/M$  and  $\Delta y = L/J$ . Eq. (8) is solved by means of the method of characteristics, i.e., by numerically solving the two ordinary differential equations

$$\frac{dy}{d\tau} = 1, \quad \frac{d\varepsilon}{d\tau} = \frac{\chi - \varepsilon}{n}.$$

Taking  $\Delta\tau = \Delta y$ , the exact solution is recovered by using an explicit Euler method in time given by

$$\varepsilon_j^{m+1} = \varepsilon_{j-1}^m + \frac{\Delta\tau}{n} (\chi_{j-1}^m - \varepsilon_{j-1}^m), \quad m \geq 0, \quad j \geq 1, \quad (88)$$

with initial and boundary conditions as

$$\varepsilon_0^m = 1, \quad m \geq 0, \quad \varepsilon_j^0 = 0, \quad j \geq 1.$$

Eq. (9) is solved by the Crank–Nicolson scheme by using a homogeneous Dirichlet boundary condition at the left side of the spatial interval, i.e.,

$$\frac{\chi_j^{m+1} - \chi_j^m}{\Delta\tau} - \beta^2 \frac{\chi_{j+1}^{m+1/2} - 2\chi_j^{m+1/2} + \chi_{j-1}^{m+1/2}}{\Delta y^2} = \varepsilon_j^{m+1/2} - \chi_j^{m+1/2}, \quad (89)$$

for  $0 \leq j < J$  and  $m \geq 0$ , where, as usual,

$$\chi_j^{m+1/2} = \frac{\chi_j^{m+1} + \chi_j^m}{2}, \quad \varepsilon_j^{m+1/2} = \frac{\varepsilon_j^{m+1} + \varepsilon_j^m}{2},$$

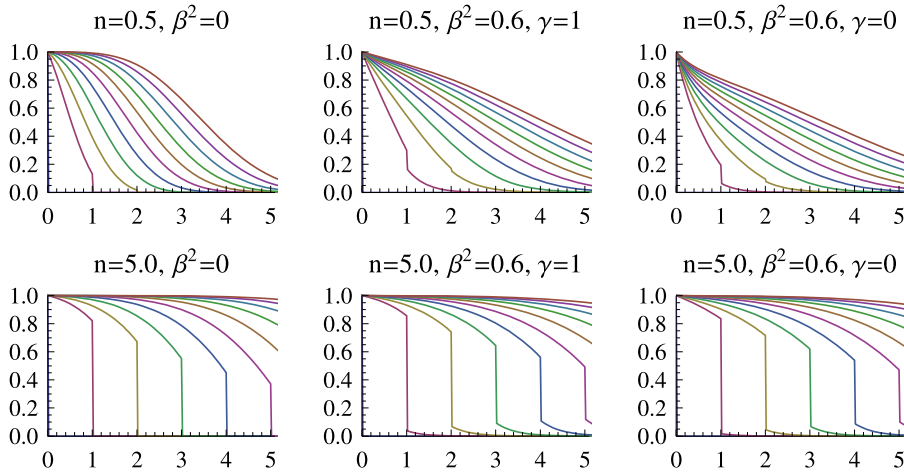
and the initial and boundary conditions are discretized as follows

$$\begin{aligned} \chi_j^0 &= 0, \quad j \geq 0, \quad \chi_j^m = 0, \quad m \geq 0, \\ \frac{\chi_1^{m+1/2} - \chi_{-1}^{m+1/2}}{\Delta y} &= \gamma (\chi_0^{m+1/2} - \varepsilon_0^{m+1/2}). \end{aligned}$$

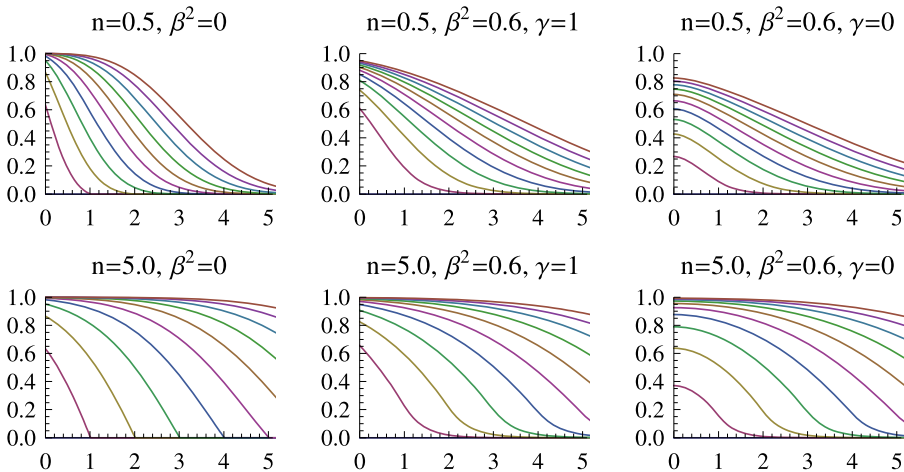
Note that a fictitious node ( $j = -1$ ) has been used in order to the Robin boundary condition be discretized to second-order in  $\Delta y$ .

Let us consider the heating of the solid by the gas, i.e.,  $T_{c_0} < T_{r_0}$ , since a Dirichlet boundary condition equal to unity is used at  $y = 0$ . The cooling process for  $T_{c_0} > T_{r_0}$ , may also be easily analyzed, but omitted here for the sake of brevity. Figs. 2 and 3 show the temporal evolution of the nondimensional temperatures  $\varepsilon(\tau, y)$  and  $\chi(\tau, y)$ , respectively, as functions of  $y$  at times  $\tau = 1, 2, 3, \dots, 10$ , corresponding to curves from left to right in each plot, calculated using the numerical method with  $\Delta\tau = \Delta y = 0.05$  (very similar results have been obtained for  $\Delta\tau = \Delta y = 0.1$ ). These figures compare the solution for  $\beta = 0$  and  $\beta = 0.6$ ; such a large value for the thermal conductivity has been used in order to highlight the main features of the solutions, not so clear for smaller values of  $\beta$ . The gas temperature  $\varepsilon(\tau, y)$  presents a sharp front propagating at constant speed with an amplitude decreasing in time whose initial size is larger as  $n$  grows (compare top and bottom plots in Fig. 2). For  $\beta^2 = 0$ , both gas and solid temperatures are exactly null ahead of the front (for  $y > \tau$ ). However, for  $\beta^2 > 0$ , as Fig. 3 clearly shows, the solid temperature  $\chi(\tau, y)$  has a smooth profile with non-null values for  $y > \tau$ , hence the gas





**Fig. 2.** Plots of the numerical solution  $\epsilon(\tau, y)$  for  $\beta^2 = 0$  (left plots) and  $\beta^2 = 0.6$  (middle and right ones), for  $\gamma = 1$  (middle plots) and  $\gamma = 0$  (right ones), and for  $n = 0.5$  (top plots) and  $n = 5.0$  (bottom ones). The abscissa in every plot is  $y$  and the curves from left to right correspond to times  $\tau = 1, 2, 3, \dots, 10$ .



**Fig. 3.** Plots of the numerical solution  $\chi(\tau, y)$  for  $\beta^2 = 0$  (left plots) and  $\beta^2 = 0.6$  (middle and right ones), for  $\gamma = 1$  (middle plots) and  $\gamma = 0$  (right ones), and for  $n = 0.5$  (top plots) and  $n = 5.0$  (bottom ones). The abscissa in every plot is  $y$  and the curves from left to right correspond to times  $\tau = 1, 2, 3, \dots, 10$ .

temperature  $\epsilon(\tau, y)$  has also a non-null value ahead from the front, as shown in Fig. 2. Note that the features downstream of the solution decreases in size as the conductivity of the solid  $\beta^2$  is diminished down to zero, not illustrated in the plots.

The results shown in Figs. 2 and 3 show the boundary layer at  $y = 0$ , clearly visible in the solid temperature, and the internal layer at  $y = \tau$ , visible in solid and gas temperatures, both layers have been taken into account by the nondimensional asymptotic approximations for the solid and the gas temperatures developed in this paper. Including terms up to second-order,  $O(\beta^2)$ , the resulting expression for the solid temperature is given by Eqs. (21), (41), (65), (67), (85), and (87), i.e.

$$\begin{aligned} \chi(\tau, y) = & \chi_0(\tau, y) + \beta \tilde{\chi}_1^{(i)}(t, y/\beta) + \beta\theta(y - \tau)\bar{\chi}_1(\tau, (y - \tau)/\beta) \\ & + \beta\theta(\tau - y)\bar{\chi}_1(\tau, (\tau - y)/\beta) + \beta^2 \chi_1(\tau, y) + \beta^2 \tilde{\chi}_2^{(i)}(t, Y) \\ & + \beta^2\theta(y - \tau)\bar{\chi}_2(\tau, (y - \tau)/\beta) - \beta^2\theta(\tau - y)\bar{\chi}_2(\tau, (\tau - y)/\beta) \\ & + O(\beta^3), \end{aligned} \tag{90}$$

and for the gas temperature is given by Eqs. (23), (40), and (86), i.e.

$$\epsilon(\tau, y) = \epsilon_0(\tau, y) + \beta^2 \epsilon_1(\tau, y) + \beta^2\theta(y - \tau)\bar{\epsilon}_2(\tau, (y - \tau)/\beta) - \beta^2\theta(\tau - y)\bar{\epsilon}_2(\tau, (\tau - y)/\beta) + O(\beta^3). \tag{91}$$

Let us note that the numerical evaluation of Eqs. (21), (23), (85), (86), and (87) requires the use of any numerical quadrature rule; here, adaptive Gauss–Kronrod quadrature [25, p. 154] has been used.

**Table 1**

Maximum absolute errors between the asymptotic approximation and the numerical solution for the solid ( $\chi(\tau_m, y_j)$ ) and gas ( $\varepsilon(\tau_m, y_j)$ ) temperatures, cf. Eqs. (90) and (91). The error are evaluated at times  $\tau = 1, 2, 3, \dots, 10$ , for all the spatial nodes  $y_j, j = 1, 2, \dots, 4000$ , where  $\Delta\tau = \Delta y = 0.005$ .

$\max  \chi(\tau_m, y_j) - \chi_j^m , \tau_m = 1, 2, \dots, 10$									
$n$	$\gamma$	$\beta$							
		0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60
0.5	0.0	0.00090	0.0053	0.012	0.049	0.11	0.21	0.34	0.50
	0.5	0.00090	0.0053	0.010	0.038	0.090	0.17	0.28	0.42
	1.0	0.00089	0.0053	0.010	0.029	0.072	0.14	0.25	0.38
5.0	0.0	0.0021	0.016	0.032	0.061	0.14	0.24	0.37	0.54
	0.5	0.0021	0.016	0.032	0.060	0.13	0.24	0.37	0.53
	1.0	0.0021	0.016	0.032	0.059	0.13	0.23	0.36	0.52
$\max  \varepsilon(\tau_m, y_j) - \varepsilon_j^m , \tau_m = 1, 2, \dots, 10$									
$n$	$\gamma$	$\beta$							
		0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60
0.5	0.0	0.00073	0.0038	0.014	0.053	0.11	0.20	0.31	0.45
	0.5	0.00073	0.0030	0.011	0.041	0.088	0.16	0.25	0.37
	1.0	0.00072	0.0023	0.0085	0.031	0.070	0.13	0.21	0.32
5.0	0.0	0.00076	0.0033	0.013	0.053	0.12	0.21	0.33	0.47
	0.5	0.00076	0.0032	0.013	0.052	0.12	0.21	0.32	0.47
	1.0	0.00076	0.0032	0.013	0.052	0.12	0.21	0.32	0.46

Table 1 presents the maximum absolute error between the asymptotic solution given by Eqs. (90) and (91) and the results of the numerical method given by Eqs. (88) and (89). As shown in the table, for small values of  $\beta$  the error of the asymptotic solution is of the order of  $\beta^2$  as expected for a uniform asymptotic expansion. In fact the absolute error is smaller than 14% for  $\beta$  as large as 0.30 and smaller than 3.2% for  $\beta \leq 0.10$ . For  $\beta \leq 0.05$ , the error shown in the table is mainly due to the Crank–Nicolson scheme used for the solid temperature, being independent of other parameters of the problem. Table 1 also shows the error for extremely large values of  $\beta$  in order to illustrate that the accuracy of the asymptotic solution degrades in such a case. Plots of the asymptotic solution, omitted here for the sake of brevity, show that the error is distributed uniformly in the whole spatial interval where the solution has been calculated. In fact, the error decreases as time elapses, since the boundary and internal layers have no influence on the solution for  $\tau \gg 1$ , being the leading order outer solution an excellent approximation. In applications where only the solution for large values of  $\tau$  is required, the outer asymptotic solution can be used with success.

## 8. Conclusions

The approximate analytical solution of the problem of heat transfer between an inert gas and a low thermal conductivity porous solid has been obtained by means of asymptotic methods under the assumptions that the gas has negligible conductivity, the fluid motion is only in the axial direction of the solid, from the inlet to the outlet, and the medium is semi-infinite. Both a boundary layer at the origin and an internal layer at the gas front position has been taken into account. The new solution for the gas presents a sharp front whose amplitude decreases with time with a non-null value downstream of the front. The solid temperature shows continuous derivatives up to the asymptotic order of approximation. In practical applications, such as packed beds, where  $\beta^2$  is small and  $n$  is larger than unity, the most important solution is that of the solid matrix, for which the leading order approximation incurs in the largest error. The new analytical solution presented in this paper has good accuracy in such a case, hence it has the potential to be incorporated in simulators currently based on Nusselt/Schumann solution. Moreover, the structure of the sharp front shown in the analytical solution presented in this paper for the gas is difficult to be acquainted by numerical methods (unless Lagrangian methods are used), so our solution can be useful in order to validate numerical schemes for heat transfer problems between fluids and porous media incorporating non-negligible thermal conductivity for the solid.

The extension of our analytical solution to multiphase problems with more than two phases is an interesting topic for further research. The asymptotic method used in this paper can also be applied in order to approximate the solution for a finite length solid medium, although the method Fourier series must be used instead of the Laplace transform. In practical applications, a square wave or periodic sine wave is considered in the inlet. The extension of the present results in such a case is currently under study.

## Acknowledgments

The authors thanks the remarks and suggestion of the reviewers which have greatly improved the contents of the paper. This work has been funded by Project ENE2008-00599 from the Spanish Ministry of Science and Innovation.

**Appendix A**

Eq. (35) can be written as

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \widehat{\Delta\varepsilon_0}(s, y) e^{s\tau} ds = \frac{1}{2\pi i} \int_C \widehat{\Delta\varepsilon_0}(s, y) e^{s\tau} ds - \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{C_R} \widehat{\Delta\varepsilon_0}(s, y) e^{s\tau} ds,$$

where  $C$  is the simple closed contour shown in Fig. 1. Let us show that

$$\lim_{R \rightarrow \infty} \int_{C_R} \widehat{\Delta\varepsilon_0}(s, y) e^{s\tau} ds = 0 \iff \tau - y > 0. \tag{92}$$

Every term in Eq. (92) has the general form

$$\rho_a = \lim_{R \rightarrow \infty} \int_{C_R} \exp\left(s(\tau - y) - \frac{sy}{n(s+1)}\right) \frac{s}{(s+1)^a} ds, \tag{93}$$

where  $a = 2, 3, 4$ . Introducing absolute values,

$$|\rho_a| \leq \lim_{R \rightarrow \infty} \int_{C_R} \left| \exp\left(s(\tau - y) - \frac{sy}{n(s+1)}\right) \right| \frac{|s|}{|s+1|^a} ds. \tag{94}$$

In polar coordinates  $s = R \exp(i\theta)$ ,  $|s| = R$ ,

$$|s+1| = \sqrt{R^2 + 2R \cos(\theta) + 1}, \tag{95}$$

$$ds = iR \exp(i\theta) d\theta, \quad |ds| = R d\theta, \tag{96}$$

and the exponential within the integral (94), can be expressed as  $\exp(\omega)$  with

$$\omega = R \exp(i\theta)(\tau - y) - \frac{Ry \exp(i\theta)}{n(R \exp(i\theta) + 1)}. \tag{97}$$

Since  $|\exp(\omega)| = \exp(\text{Re}\{\omega\})$ , where

$$\text{Re}\{\omega\} = R \cos(\theta)(\tau - y) - \frac{Ry(R + \cos(\theta))}{n(R^2 + 2R \cos(\theta) + 1)}, \tag{98}$$

then Eq. (94) yields

$$|\rho_a| \leq \lim_{R \rightarrow \infty} \int_{\pi/2-\alpha(R)}^{3\pi/2+\alpha(R)} \frac{R^2}{(R^2 + 2R \cos(\theta) + 1)^{a/2}} \exp(\text{Re}\{\omega\}) d\theta. \tag{99}$$

Fig. 1 shows that

$$\lim_{R \rightarrow \infty} \alpha(R) = 0,$$

and, since  $a \geq 2$  and  $\cos(\theta) < 0$  for  $\theta \in (\pi/2, 3\pi/2)$ , then

$$\lim_{R \rightarrow \infty} \frac{R^2}{(R^2 + 2R \cos(\theta) + 1)^{a/2}} \exp(\text{Re}\{\omega\}) = 0, \tag{100}$$

if and only if  $\tau - y > 0$ . Therefore, one gets

$$\lim_{R \rightarrow \infty} \int_{C_R} \exp\left(s(\tau - y) - \frac{sy}{n(s+1)}\right) \frac{s}{(s+1)^a} ds = 0 \iff \tau - y > 0.$$

## References

- [1] W. Nusselt, Der Wärmeübergang im Kreuzstrom, *Z. Ver. Deut. Ing.* 55 (1911) 2021–2024.
- [2] W. Nusselt, Eine neue Formel für den Wärmedurchgang im Kreuzstrom, *Tech. Mech. U. Thermodyn.* 1 (1930) 417–422.
- [3] A. Anzelius, Über Erwärmung vermittels durchströmender Medien, *ZAMM Z. Angew. Math. Mech.* 6 (1926) 291–294.
- [4] T.E.W. Schumann, Heat transfer: A liquid flowing through a porous prism, *J. Franklin Inst.* 208 (1929) 405–416.
- [5] C.C. Furnas, Heat transfer from a gas stream to a bed of broken solids, *Trans. Am. Inst. Chem. Eng.* 24 (1930) 142–193.
- [6] A. Klinkenberg, Heat transfer in cross-flow heat exchangers and packed beds, *Ind. Eng. Chem.* 46 (1954) 2285–2289.
- [7] B.S. Baclic, P.J. Hegggs, On the search for new solutions of the single-pass cross-flow heat-exchanger problem, *Int. J. Heat Mass Transf.* 28 (1985) 1965–1976.
- [8] J. Łach, W. Pieczka, On the general-solution to a certain class of heat and or mass-transfer problems, *Int. J. Heat Mass Transf.* 28 (1985) 1976–1981.
- [9] P.-X. Jiang, Z.-P. Ren, Numerical investigation of forced convection heat transfer in porous media using a thermal non-equilibrium model, *Int. J. Heat Fluid Flow* 22 (2001) 102–110.
- [10] F. Kuwahara, M. Shirota, A. Nakayama, A numerical study of interfacial convective heat transfer coefficient in two-energy equation model for convection in porous media, *Int. J. Heat Mass Transf.* 44 (2001) 1153–1159.
- [11] O. Laguerre, S. Ben Amara, G. Alvarez, D. Flick, Transient heat transfer by free convection in a packed bed of spheres: Comparison between two modelling approaches and experimental results, *Appl. Thermal Eng.* 28 (2008) 14–24.
- [12] A.V. Kuznetsov, An investigation of a wave of temperature difference between solid and fluid phases in a porous packed bed, *Int. J. Heat Mass Transf.* 37 (1994) 3030–3033.
- [13] A.V. Kuznetsov, Comparison of the waves of temperature difference between the solid and fluid phases in a porous slab and in a semi-infinite porous body, *Int. Commun. Heat Mass Transf.* 22 (1995) 499–506.
- [14] A.V. Kuznetsov, A perturbation solution for heating a rectangular sensible heat storage packed bed with a constant temperature at the walls, *Int. J. Heat Mass Transf.* 40 (1997) 1001–1006.
- [15] D.A. Nield, A. Bejan, *Convection in Porous Media*, Springer-Verlag, New York, 1999.
- [16] J. Abate, P.P. Valkó, Multi-precision Laplace transform inversion, *Int. J. Numer. Meth. Engng.* 60 (2004) 979–993.
- [17] W.F. Ames, *Numerical Methods for Partial Differential Equations*, third ed., Academic Press, New York, 1992.
- [18] J. Kevorkian, J.D. Cole, *Multiple Scale and Singular Perturbation Methods*, Springer-Verlag, New York, 1996.
- [19] A.N. Nayfeh, *Perturbation Methods*, John Wiley & Sons, New York, 1973.
- [20] J. Pérez, F.R. Villatoro, J.L.G. Santander, M.A. Borovsky, Yu.L. Ratis, E.A. Izzheurov, P. Fernández de Córdoba, Heat transfer between a gas and an ultralow thermal conductivity porous structure, *Appl. Math. Comput.* 204 (2008) 687–693.
- [21] S.-D. Shih, On a class of singularly perturbed parabolic equations, *ZAMM Z. Angew. Math. Mech.* 81 (2001) 337–345.
- [22] A.D. Polyanin, *Handbook of Linear Partial Differential Equations for Engineers and Scientists*, Chapman & Hall/CRC, Boca Raton, FL, 2002.
- [23] N.N. Lebedev, *Special Functions and Their Applications*, Dover Publications, New York, 1972.
- [24] M.Ya. Antimirov, A.A. Kolyshkin, R. Vaillancourt, *Complex Variables*, Academic Press, 1998.
- [25] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, second ed., Cambridge University Press, Cambridge, UK, 1992.