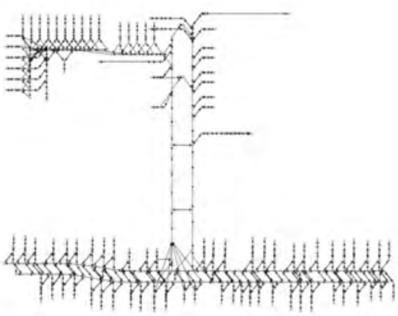
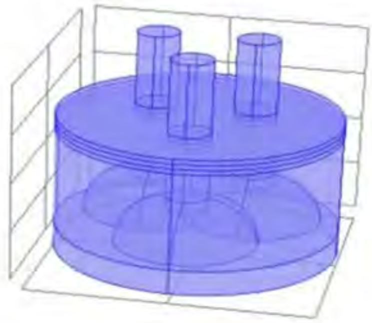
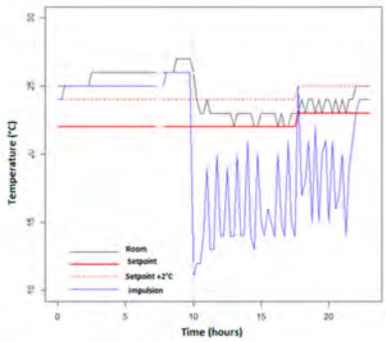
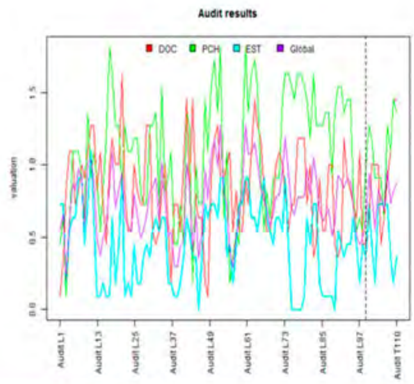
122 European Study Group with Industry









Ricardo Cao Abad
Carlos Parés Madroñal
Javier Roca Pardiñas
Julio Rubio García
Mª Pilar Salgado Rodríguez
Editors

Facultad de Matemáticas | Santiago de Compostela
September 19 - 23, 2016

[math-in].net
Red Española Matemática-Industria

ITMATI

PROCEEDINGS OF THE

122 EUROPEAN STUDY GROUP WITH INDUSTRY (122 ESGI)

Santiago de Compostela, September 19 - 23, 2016

Editors

Ricardo Cao Abad
University of A Coruña
ricardo.cao@udc.es

Carlos Parés Madroñal
University of Malaga
pares@uma.es

Javier Roca Pardiñas
University of Vigo
roca@uvigo.es

Julio Rubio García
University of La Rioja
julio.rubio@unirioja.es

Mª Pilar Salgado Rodríguez
University of Santiago de Compostela
mpilar.salgado@usc.es

PROCEEDINGS OF THE

122 EUROPEAN STUDY GROUP WITH INDUSTRY

This volume contains the proceedings of the 122 European Study Group with Industry (122 ESGI) held in Santiago de Compostela, Spain, from 19th to 23th September 2016.

The 122 ESGI was jointly organized by the Technological Institute for Industrial Mathematics (ITMATI) and the Spanish Network for Mathematics and Industry (math-in).

Index

Introduction

Initiated in Oxford in 1968, Study Groups with Industry provide a forum for industrial scientists to work alongside academic mathematicians on problems of direct industrial relevance. They are an internationally recognized method of technology and knowledge transfer between academic mathematicians and industry, usually lasting one week.

The success of the ESGI lies in its unique format which has been copied around the world, and which allows Mathematicians to work on reduced groups to study problems presented by industry. These problems arise from any economic sector thanks to the versatility of Mathematics.

The objective is to present the capabilities of Mathematics and its applicability in a large part of the challenges and needs that industry presents. It aims to bring small, medium and large companies a technology with great potential, used by highly qualified researchers and which does not require large investments to use.

Therefore, collaboration between industry experts and researchers is key to address technological innovation issues by using successful mathematical techniques. ESGI contributes to the promotion of mathematics and helps companies to use Mathematics to improve their processes.

The goals which want to be reached at the ESGI are:

- found solutions and insights into existing industrial problems;
- established lasting and productive working links between research applied mathematicians and industry;
- propose new lines of research based on business challenges;
- reinforce the importance of mathematics in industry and mathematical profiles companies;
- stimulated greater awareness in the wider community of the power of mathematics in providing solution paths to real-world problems.

Finally, it should be pointed out that 65 researchers, students, professors and company technicians contributed to a successful 122 ESGI.

Santiago de Compostela on 23^{th} September, 2016

Members of the Scientific Committee:

- Ricardo Cao Abad. Professor of Statistics and Operations Research at the University of A Coruña. Affiliated researcher of ITMATI.
- Carlos Parés Madroñal. Professor of Applied Mathematics at the University of Malaga. Vice-President of the Spanish Network for Mathematics and Industry(math-in).
- Javier Roca Pardiñas. Associate Professor of Statistics and Operations Research at the University of Vigo. Affiliated researcher of ITMATI.
- Julio Rubio García. Professor of Computer Science and Artificial Intelligence at the University of La Rioja. Member of math-in.
- Mª Pilar Salgado Rodríguez. Associate Professor of Applied Mathematics at the University of Santiago de Compostela. Affiliated researcher of ITMATI.

Fleet planning of AGVs

Academic Coordinator Julio González Díaz
University University of Santiago de Compostela

Business Coordinators Jesús Enrique Sierra García
Company Academic and Scientific Relations Manager

Specialist Juan José Salazar González.
Company Professor of Statistics and Operations Research University of La Laguna

Team Xián Álvarez López(University of Santiago de Compostela), Brais González Rodríguez(University of Santiago de Compostela), Ángel Manuel González Rueda(University of Santiago de Compostela), Patricio Reyes Valenzuela(Technological Institute for Industrial Mathematics), Alberto Rodríguez Vázquez(University of Santiago de Compostela), Jorge Rodríguez Veiga(Technological Institute for Industrial Mathematics), Juan Vidal Puga(University of Vigo).

Fleet planning of AGVs

Julio González Díaz[1]

Xián lvarez López, Brais González Rodrez, Ángel Manuel González Rueda,

Patricio Reyes Valenzuela, Alberto Rodríguez Vázquez, Jorge Rodríguez Veiga, Juan José Salazar González, Jesús Enrique Sierra García, Juan Vidal Puga

Abstract

This document contains a description of the main challenges one would face to design an intelligent system that allows to dynamically control the automatic guided vehicles (AGVs) of some facility such as manufacturing plants, warehouses, ...

The system should be able to dynamically adjust to changes such as new orders and unexpected events. It should also anticipate congestion problems and avoid collisions and deadlocks. The first goal to achieve by this system would be to outperform the one currently in use by the company that proposed this challenge. Performance would be measured in both the quality of the solution and the versatility to accommodate different types of facilities.

To illustrate the complexity of the challenge, some of the ingredients of the associated mathematical programming problem are discussed for the static version of the optimization problem at hand. Then, a heuristic algorithm is proposed for the dynamic version, along with a series of potential enhancements.

1 Introduction

The optimization of the operations that take place in facilities such as logistic warehouses and distribution centers is a topic of great relevance for many companies. This is specially so because of the increasing capabilities and complexity of these facilities over the recent years, at a pace that is nothing but accelerating. Because of the complexity and relevance of these problems, they are attracting attention from a wide variety of fields such as mathematics, engineering and computer science. Refer to [3] and [1] for two recent surveys on the topic.

Because of the above complexity, the efficient operation of such facilities is often a multi-layered problem, with many subproblems. Most of these subproblems are themselves complicated combinatorial problems and, hence, the simultaneous and coordinated optimization of all of them is normally out of reach. Therefore, one should identify the critical layers for a given company

[1] julio.gonzalez@usc.es

or for a given facility and attack the corresponding subproblems. It is worth emphasizing that these problems are already challenging in a static setting and that, in most real facilities, optimal decisions must be adjusted dynamically to accommodate new orders or to react to unexpected events.

One goal of the present document is to present a brief overview of the aforementioned subproblems and a more detailed discussion of those that were identified as more relevant for the partner company of the study group in which this work took place.

The focus in this challenge is on the routing and scheduling of vehicles in facilities were these vehicles are automatically guided (AGVs), which requires taking into consideration potential issues such as collisions and deadlocks. Refer to [7] and [4] for two surveys that deal specifically with design and control of systems with AGVs.

Once the main challenges are identified, a heuristic algorithm will be discussed. The need for an algorithm that can be run dynamically on real time makes it unrealistic to hope for the existence of some sort of exact algorithm that is fast enough.

2 Challenge description

As we have already mentioned, the purpose of this challenge is the design of an intelligent system that allows to dynamically control the automatic guided vehicles (AGVs) of some facility such as manufacturing plants, warehouses,... The system should run on real time, reacting dynamically to new orders, cancelations, and different types of unexpected events. Because of the use of AGVs, special attention should be put on the prevention on collisions and deadlocks, two issues that are not so critical when vehicles are not automatically guided.

2.1 Different layers of the problem

Next we present a brief description of the main subproblems that arise when trying to do an efficient design and operation of a general facility. We refer the reader to the surveys [3] and [1] to get a deeper discussion on them.

Warehouse layout. The design of the layout of the warehouse is a matter of great relevance, since the efficiency of future operations crucially depends on it. In this document we take it as given, although a complete decision system should eventually be able to account for this aspect of the problem as well.

Slotting/storing strategies. Depending on the type of facility, the strategy chosen to store the different goods can have a big impact on the final costs. For instance, one can decide to always put the goods in the nearest empty slots or to reserve the most accessible places to the goods that are expected to moved earlier. For this document these strategies are also taken as given.

Guide-path layout. In the case of facilities with AGVs, one also has to define the guide-path layout, to specify the “virtual” lanes to be followed by the vehicles. Again, this layout is taken as given for the present document.

Figure 1 represents a typical layout, obtained from the encoding of one real facility handled by the company that set forth this challenge.

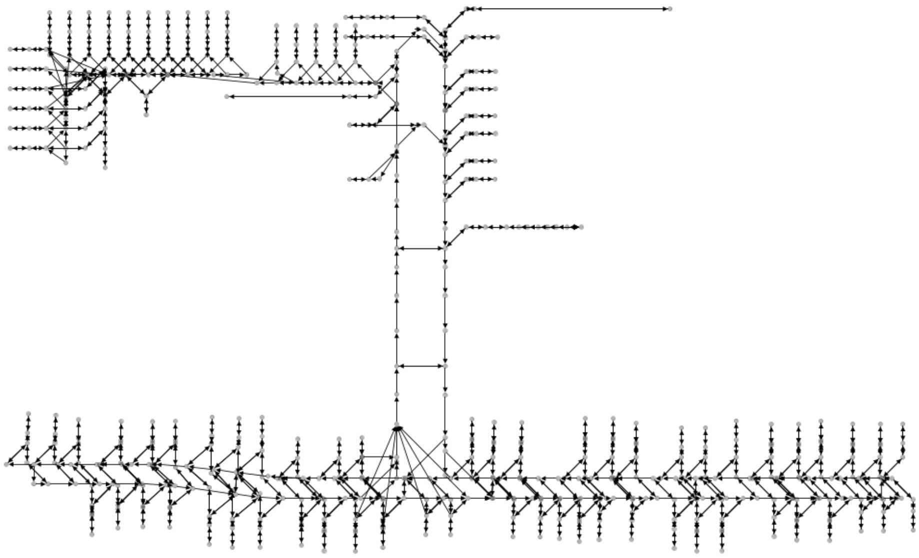


Figure 1: Layout of one of the warehouses of the company.

Size and composition of the fleet. Another important aspect is the composition of the fleet. In the case of different types of vehicles, one has to decide the number of vehicles of each type. Again, for the present document the composition of the fleet is taken as given. Yet, one immediate application of an algorithm that allows to efficiently control a given fleet would be to compare the performances of different potential fleet compositions, in order to help deciding which one to choose.

Control of AGVs. This is the part of the problem in which we focus in this document. The optimal control of the vehicles is a highly complex optimization problem, in which one has to take five different decisions: assignment of AGVs to orders, spacial routing of the AGVs, temporal scheduling of the chosen routes, charging decisions, and idle vehicle positioning. Ideally, one should address all these decisions simultaneously in order to get a globally optimal solution, but the complexity is such that one may have to settle with a sequential resolution of them.

Assignment of vehicles to orders. At any given time, there will be a set of orders in the system that need to be processed by the AGVs, and the AGVs may themselves be idle or in their way to process some of the orders.

At each such time, one has to decide how to assign the different orders (which might be prioritized) to the available AGVs. Which assignment is better depends on the ensuing optimal routing of the chosen vehicles, which is handled by the next subproblems.

Vehicle routing (space). The first part of the routing would consist on deciding the route of the different vehicles, but abstracting from any time considerations.

Vehicle scheduling (time). Then, one should decide on the time in which in each part of the route is traversed, ensuring that collisions are avoided.

Battery management. Orders typically consists on picking a given good and moving it to a new location, but AGVs may be assigned other tasks such as moving to a charging station. The decision of when to charge each AGV is also an important one, and it has associated a routing and scheduling problem as well.

Positioning of idle vehicles. Finally, one should also decide where to position vehicles that have no assigned task. For instance, one could just leave the idle vehicles where they completed their last task, unless that can block the way for another AGVs or move them to charging stations. Moreover, one can have different "parking locations" scattered through the layout, and then decide which is the most appropriate one for each idle vehicle, in which case a new routing and scheduling problem would arise.

3 Static optimization: Outline of the mathematical programming model

In this section we present a possible modeling of the static optimization problem for the control of AGVs. We do not pretend to be fully rigorous or exhaustive here. We just present some hints at the modeling variables and some of the constraints to illustrate the difficulties of such a modeling.

3.1 Basic elements of the model

First, we need a description of the guide-path layout, which is done by the *spatial graph* $G = (N, S)$, where $N = \{1, \dots, |N|\}$ contains the *nodes*: structural nodes, order pick-up stations, order-delivering stations, charging stations and parking locations for the AGVs. The set S contains the *edges*, which we assume are bidirectional. The graph contains many edges, since the length of each edge is fixed to the distance an AGV can cover in a unit of time. The reason for this modeling approach is because edges will be used to control collisions just by ensuring that we will never have two AGVs in the same edge. For instance, an example like the one in Figure 1 has 552 nodes and 1100 edges.

In order to define the scheduling of the AGVs when covering their assigned routes, we also need to track time. Let $T = \{1, \dots, |T|\}$ denote the time periods.

In this static version of the problem we assume that we know in advance all the orders we will face in the $|T|$ periods under consideration. Let $P = \{(o_p, d_p, t_p, w_p) \in N \times N \times T \times [0, 1]\}$ denote the set of orders where

- o_p denotes the pick-up node for the order,
- d_p denotes the delivery node for the order,
- t_o denotes the time period in which the order arrives to the system, and

- w_p denotes the priority.

We assume that the set of vehicles is given by $K = \{1, \ldots, |K|\}$. In particular, we are assuming that the fleet of vehicles is homogeneous.

In order to avoid collisions of AGVs, we need to track the position of each AGV at each moment of time. In order to do so, we define the space time graph, given by $H = (V, A)$, where $V = (N \times T)$ and edges are ordered pairs of elements of V such as $\big((i,3),(j,4)\big)$ to represent a movement from node i to node j in period 3. We slightly abuse notation and write $(n_1, n_2, t) \in H$ to represent the arc corresponding with the movement, at time t, from n_1 to n_2.

3.2 Outline of the mathematical programming problem

Now we describe some of the elements of a tentative mathematical programming problem.

Variables

The variables of the model would be the following

$x_{kpvwt} = \{$ 1 if vehicle k moves from v to w at time t with order p assigned to it0otherwise.

$y_{kp} = \{$ 1 if order p has been assigned to vehicle k0otherwise.

Note that, given the above variables, the resulting problem will be an integer programming problem with a very large number of variables. In particular, if the time discretization is at the level of seconds, modeling a 24 hour labor day with 20 different vehicles and 200 different orders with the graph in Figure 1 would result in around $20 * 200 * 502 * 502 * 20 * 24 * 3600 \approx 8 * 10^{13}$ binary variables, which is too much for any state of the art solver nowadays.

Objective function

The main goal is to minimize the total travel time of the vehicles

$$\min \sum_p \sum_k \sum_v \sum_w \sum_t time_{kvw} \cdot x_{kpvwt},$$

where $time_{kvw}$ denotes the time need for vehicle k to move from v to w. This would allow to account for a heterogeneous fleet. Moreover, one might even have to consider that the speed of a vehicle may depend on whether it is loaded or not, which would call for additional parameters.

Most likely the best modelling approach for the priorities of the different orders would be as a soft constraint, so there could be additional terms in the objective function.

Constraints

We should now be able to state all the constraints for the model. This contraints would be, for instance,

- Assignment of vehicles to orders, $\forall k \in K, \forall p \in P, \forall t \in T$,

$$\sum_{(n_1,n_2,t)\in H} x_{kpn_1n_2t} \leq y_{kp}.$$

- A vehicle can only handle one order at a time, $\forall k \in K, \forall t \in T$,

$$\sum_{p\in P} \sum_{(n_1,n_2,t)\in H} x_{kpn_1n_2t} \leq 1.$$

- Each order can only be assigned to one vehicle, $\forall p \in P$,

$$\sum_{k\in K} y_{kp} = 1.$$

The above contraints are quite simple ones, but there are much more constraints some of which are significantly more involved such as i) no two vehicles can be in the same edge at the same time (avoid collisions), ii) the variables x must define feasible routes for the vehicles in the space-time graph H if a vehicle picks an order, the vehicle has to deliver the order, iii) the order cannot be delivered before the pick-up, iv) battery management constraints within some minimum and maximum battery levels. For the sake of illustration we present a tentative description of a couple of these constraints:

- The delivery node must be visited after visiting the pick-up node, $\forall k \in K, \forall p = (o_p, d_p, t_p, w_p) \in P, \forall t_2 \geq t_p$,

$$-\sum_{(o_p,n_1,t_1)\in H:t_p\leq t_1\leq t_2} x_{kpo_pn_1t_1} + \sum_{(n_2,d_p,t_3)\in H:t_2\leq t_3} x_{kpn_2d_pt_3} \geq 0.$$

- Avoid collisions, $\forall k_1, k_2 \in K, \forall n_1, n_2 \in N, \forall t, \forall (n_3, n_4, t') \in I_{n_1,n_2,t}$,

$$x_{k_1pn_1n_2,t} + x_{k_2pn_3n_4t'} \leq 1,$$

  where $I_{n_1,n_2,t}$ contains the edges in the space-time graph in which there cannot be AGVs when at time t an AGV is moving from n_1 to n_2. In a nutshell, $I_{n_1,n_2,t}$ is an *exclusion zone* associated to the movement on an AGV from n_1 to n_2 at time t. The formal definition of $I_{n_1,n_2,t}$ would depend on the specifics of the given facility, layout, and AGVs.

Even though the above modeling approach for the mathematical programming problem is far from being the most efficient one, it should be clear that the associated optimization problem is indeed a very hard nut to crack. Moreover, if on top of this we need that the resulting algorithm is run on real time every time a new order enters into the system, every time an AGV completes an order, and every time an unexpected event occurs, it seems that we have to settle for some heuristic algorithm. In the next section we describe one such algorithm.

4 Dynamic optimization: Heuristic algorithm

The need for a heuristic algorithm in this context is not something new. The existing literature on the topic has long ago acknowledged this fact. The core idea of the algorithm we present in this section builds upon the approach in [6].

4.1 The base algorithm

We now outline the idea of what could be a base algorithm for the real time optimization of control of AGVs.

- **Vehicle assignment.** Greedy heuristic to compute which available AGV is closest to the pick-up point of the current order (shortest path in the *spatial graph*).
- **Routing and scheduling.** The *core* of the algorithm:
  - Computation, in each step, of the shortest path of the current AGV to undergo its current task (in the *space-time graph*):
    - **Collisions**, **deadlocks** and **congestion** are controlled.
    - **Battery level** should be accounted for.
    - **Safeguard implementation** must be carefully analyzed (this could be a departure from existing literature), related to the $I_{n_1,n_2,t}$ sets mentioned above.
  - The system should continuously track the status of the AGVs, to check for delays and other contingencies. If needed it should decide how to **reroute/reassign**.
- **Battery management.** Vehicles under a certain threshold should be sent to the nearest empty charging location.
- **Idle vehicles.** If a vehicle has no assigned order or task, then it should go to the nearest empty parking location.

We are not claiming that this algorithm we have just outlined would lead high quality solutions, since there are many crucial junctures at which it behaves very miopically. In particular, vehicle assignment is done myopically and completely decoupled from the ensuing routing/scheduling. On the other end, this algorithm should be very fast, so it can be run on real time.

In the next section we present a few potential enhancements, whose priority should be decided based on the performance of the base algorithm.

4.2 Potential enhancements of the algorithm

Vehicle assignment

The computation of the shortest paths should be done directly in the *space-time graph*. This is more costly from the computational point of view, but current traffic conditions in the system would be accounted for.

A second enhancement to reduce the myopia of the base algorithm would be to compute multiple assignments at once, instead of doing it sequentially. There are different levels of sophistication that could be incorporated here, also related to the above enhancement dealing with the space-time graph:

- One could compute a matrix of the *shortest path distances* in the space graph between available AGVs and pick-up points of the current order queue and perform an optimal *assignment*.
- The same as above, but in the space-time graph.

- What seems extremely hard would be to be able to compute the above shortest paths in the space-time graph *simultaneously*, *i.e.*, solving the multiple shortest path problem in the space-time graph.[2] Then, traffic would be taken into account and we would have full integration of *assignment-routing-scheduling*.

Vehicle routing

A first approach would be to search for some *refinement* of the time-windows computed in the space-time graph that allows to correct the *miopic* nature of the base algorithm (schedules are computed one at a time). One option here would be to work with the classic *2-opt* and *3-opt* moves, in the spirit of the seminal paper for the traveling salesman problem (see [5]).

If the assignment stage results in more than one assignment one could try to behave in a less miopic way:

- Try all the *different orderings* of AGVs when computing schedules and take the best one.
- Try a subset of the above orderings obtained randomly or through some heuristic.
- Identify bottlenecks and *reconfigure* some of the edges to mitigate them.

5 Conclusions

The design of an intelligent system that allows to dynamically control the automatic guided vehicles is a very challenging and complex optimization problem (highly combinatorial).

The base algorithm outlined in Section 4.1 should be relatively easy to implement. Interenstingly, the company behind this challenge already has a "simulator" that would help with the validation of the algorithm and, moreover, provide an excellent benchmark for comparison with the algorithm currently in use. The extensions should then be prioritized to improve the performance of the algorithm in those parts where it has proven to be weaker.

The company also has data on real operations in a very accessible format, which would also facilitate validation and integration of the different versions of the algorithm.

References

[1] H. J. Carlo, Iris F. A. Vis, and K. J. Roodbergen. Transport operations in container terminals: Literature overview, trends, research directions and classification scheme. *European Journal of Operational Research*, 2014.

[2] Tali Eilam-Tzoreff. The disjoint shortest paths problem. *Discrete Applied Mathematics*, 85:113–138, 1998.

[2] This has been called *the disjoint shortest paths problem*, which is known to be an NP-complete problem, but it might be possible to solve it relatively efficiently if the maximum number of simultaneous disjoint paths to compute is fixed. A good starting point for the design of such an algorithm could be [2].

[3] J. Karasek. An overview of warehouse optimization. *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems*, 2:111–117, 2013.

[4] Tuan Le-Anh and M. B. M. De Koster. A review of design and control of automated guided vehicle systems. *European Journal Of Operational Research*, 171:1–23, 2006.

[5] S. Lin and B. W. Kernighan. An effective heuristic algorithm for the traveling-salesman problem. *Operations Research*, 21(2):498–516, 1973.

[6] Rolf H. Möhring, Ekkehard Köhler, Ewgenij Gawrilow, and Björn Stenzel. *Operations Research Proceedings 2004*, chapter Conflict-free Real-time AGV Routing, pages 18–24. Springer, Berlin, Heidelberg, 2005.

[7] Iris F. A. Vis. Survey of research in the design and control of automated guided vehicle systems. *European Journal of Operational Research*, 170:677–709, 2006.

Information processing from inspections of quality management

Academic Coordinators Rosa M. Crujeiras
University Associate Professor at the Department of Statistics and Operations Research in University of Santiago de Compostela and Affiliated researcher of ITMATI.

Business Coordinator Hugo Caridad Cordeiro
Company CEO iAuditoria.

Information processing from inspections of quality management

Rosa M. Crujeiras, Adela Martínez Calvo

Abstract

iAuditoria is a tool for conducting audits and/or inspections from a checklist previously established by the user, all in a cloud environment. Taking the IFS Food as a reference standard, a discrete time Markov Chain modelling approach is analysed for predicting the outcome of an audit at item level. A particularized model formulation is assessed through simulations, presenting some results which could be included as outputs in iAuditoria, providing the users with tools for a better understanding of their systems which could help in preventing failures.

1 Introduction

Nowadays, organizations and institutions provided themselves with quality management systems which help in preventing failures in manufactured products and problems when delivering services to customers. These systems comprise also quality control processes, which are usually put into practice by means of quality audits. A quality audit is a systematic examination process on the quality system which tries to assess the performance of an organization according to certain standards. One of the organizations that establishes such standards for quality audits is the International Feature Standards (IFS[1]), which has developed eight different food and non–food standards, covering the processes along the supply chain. Specifically, IFS Food is a standard for the auditing of companies that process food or companies that pack loose food products. It is only applied where the product is processed or handled, or if there is a danger of product contamination during the primary packaging. The IFS Food standard has a reference checklist, where a series of items are classified according to the following fields: documental registries, correct hygiene practices and structural issues. Some examples of items will be shown in next sections. The aforementioned checklist can be used as a basis for proceeding with a quality audit, for an internal or an external examiner. The audit process requires then checking all the items, giving a value to each of them and computing a final score for the audit itself. These processes are usually tedious and the use of ICTs can help in systematizing and simplifying the auditor work.

[1]International Feature Standards: https://www.ifs-certification.com/

iAuditoria ICT[2] is a tool for conducting audits and/or inspections from a checklist previously established by the user. Each user can designed a checklist, with items organized in different blocks as for the IFS Food standard, and different values can be assigned individually and by groups. For instance, each item can be valued on an A-B-C-D scale and weights can be also put to each grade (e.g. A=5, B=4, C=1, D=0). Different weights can be also assigned for grouping fields, and with all these information, the user receives a final result of the audit. This result can be favourable or non-favourable, and in each case, different types of deviations from the standard can be identified. The requirements to meet each final grade can be also set by the user (for instance, 100% of the items with A for favourable with no deviations). Summarizing: the auditor uses iAuditoria to perform the audit and to apply the (selected, self-configured) audit model. For the audit model, the user sets the fields weight, the items score and the requirements to meet the final result, which also includes a global score based on the previous settings. The goal of this work is, once a series of audits has been performed, try to predict the outcome of the next audit at item level. This will allow to follow the evolution of the system in such a way that the user can identify critical items and act to prevent failures of mistakes.

This report is organized as follows. The (mathematical) problem formulation is introduced in Section 2, and particularized for iAuditoria. Section 3 shows some simulation experiments, describing the simulation scenario and presenting some results which could be included as outputs in iAuditoria. Some final remarks on model validation, limitations and an extended model with a multivariate perspective are briefly discussed in Section 4.

2 Problem formulation

A statistical predictive approach based on Markov Chains will be considered to model the evolution of item values, which can be used in a second step for modelling the evolution of the system. Consider a system that evolves randomly over time and suppose that the state of this system is registered at times $n = 0, 1, 2\ldots$. Denote by X_n the state of the system at time n, taking values on a space E (state space). The discrete time process $\{X_n,\ n \geq 0\}$ is said to satisfy the Markov condition, and then such a process is a discrete time Markov Chain (DTMC) if, for all $n \geq 0$:

$$\mathcal{P}(X_{n+1} = x_{n+1} | X_1 = x_1, \ldots, X_n = x_n) = \mathcal{P}(X_{n+1} = x_{n+1} | X_n = x_n), \quad (1)$$

where $\mathcal{P}$ stands for probability. The reader is referred to [3] for a complete review on DTMCs. Intuitively, condition (1) (that is, the Markov property), indicates that the *present* of the system (the state at instant n) only depends on the *immediate past* (state at instant $(n-1)$), regardless previous evolution. Furthermore, the DTMC is said to be homogeneous if the probabilities in (1) do not depend on the moment n. Hence:

$$\mathcal{P}(X_{n+1} = j | X_n = i) = p_{ij}, \quad i, j \in E. \quad (2)$$

The probabilities in (2) are called transition probabilities, and they provide the chances of passing from state i to state j in a single step. Those probabilities can

[2]iAuditoria: https://www.iauditoria.com

be organized in matrix form, obtaining P, a matrix of dimension $m \times m$, being m the number of different states in E and entries $P(i,j) = p_{ij}$, for $i,j = 1,\ldots,m$.

Further considerations on DTMC modelling can be found in [3]. For the problem under study, it is worth mentioning that from the observation of the system along time, it is possible to compute the transition probabilities (conditional probabilities, given a specific state), with their corresponding standard errors and confidence intervals for the true underlying values. Transition probabilities estimation was initially studied by [1], who introduced the maximum likelihood estimator as well as some tools for hypothesis testing. More recently, [5] analyses different estimators for DTMCs.

Modelling the results of different audits for a fixed item by means of DTMC would allow to complete the audit report produced by the iAuditoria tool: as mentioned above, it is possible to estimate transition probabilities and to *attach* to the derived estimates an uncertainty quantification measure, by means of a standard error or by constructing confidence intervals. In addition, exploratory tools applied on the estimated quantities (transition probabilities and standard errors) would allow to identify *critical* items. Finally, given the configuration of weights (by fields and by item outcome), an estimation of a final score can be also computed. In the next section, the specific formulation of the model for the iAuditoria setting will be presented.

The problem for iAuditoria

Consider the collection of items of an audit checklist, which are classified in the three fields presented in the Introduction[3]: DOC (documental registries), CHP (correct hygiene practices) and STR (structural issues). The items are denoted by $i = 1,\ldots,k_1,k_1+1,\ldots,k_2,k_2+1,\ldots,k_3$, each block corresponding to each one of the previous fields. For item k, define the DTMC X_n^k = state of item k at n–th audit, $n \geq 0$ (X_n^0 denotes the initial state of the item). The state space for every item is $E_k = E = \{A,B,C,D\}$ (the value of the audit for each item). The Markov condition and time homogeneity can be expressed as follows for each item k and considering $e_0, e_1, e, e' \in E$, for all $n \geq 0$:

$$\mathcal{P}(X_{n+1}^k = e | X_0^k = e_0, X_1^k = e_1, \ldots, X_n^k = e') = \mathcal{P}(X_{n+1}^k = e | X_n^k = e').$$

The transition matrix (the same for all items) is given by:

$$P = P^k = \begin{pmatrix} p_{AA}^k & p_{AB}^k & p_{AC}^k & p_{AD}^k \\ p_{BA}^k & p_{BB}^k & p_{BC}^k & p_{BD}^k \\ p_{CA}^k & p_{CB}^k & p_{CC}^k & p_{CD}^k \\ p_{DA}^k & p_{DB}^k & p_{DC}^k & p_{DD}^k \end{pmatrix}$$

Once the model is established, the elements of the DTMC must be estimated based on audit observations, and the model itself must be also validated. Specifically, for each item k, once a sequence of values (from audits) has been observed, P^k can be estimated as follows. Consider $e, e' \in E$, two different states. The maximum likelihood estimator proposed by [1] is computed as

$$\widehat{p}_{ee'}^k = \frac{n_{ee'}}{\sum_{u \in E} n_{eu}},$$

[3]In Figures, CHP and STR are replaced by PCH and EST, corresponding to the Spanish spelling of the fields denomination.

being $n_{ee'}$ the observed frequency of transitions from e to e'. The corresponding standard error for this estimator (see [5]) can be computed as

$$\mathrm{SE}(\widehat{p}^k_{ee'}) = \frac{\widehat{p}^k_{ee'}}{\sqrt{n_{ee'}}}.$$

The asymptotic normality of $\widehat{p}_{ee'}$ allows for the construction of confidence intervals, and then it is possible to attach some uncertainty to the obtained estimates. Hence, confidence intervals (LL, UL) can be computed based on normal quantiles as:

$$LL = \widehat{p}^k_{ee'} - q_{\alpha/2}\mathrm{SE}(\widehat{p}^k_{ee'}), \quad UL = \widehat{p}^k_{ee'} + q_{\alpha/2}\mathrm{SE}(\widehat{p}^k_{ee'})$$

where $q_{\alpha/2}$ is the $(1-\alpha/2)$ quantile of a standard normal distribution. For $\alpha = 0.05$, $q_{\alpha/2} = 1.96$, giving a confidence interval with 95% probability coverage. The standard error of the estimators can be also computed by bootstrap methods, as proposed by [7] and denoted by $\mathrm{SE}^{boot}(\widehat{p}^k_{ee'})$.

3 Some simulation experiments

Since it has not been possible to obtain real data from audits in practice, the approach followed to illustrate the potentiality of the method is the design and implementation of a simulation study. In the next sections, the simulation scenario will be described in detail, trying to resemble as much as possible a real situation. In subsequent sections, some simulation results will be provided with the aim of presenting which type of information would be susceptible of being included as outputs in iAuditoria.

Simulation scenario

Taking the IFS standard as a reference, three fields have been considered in an audit procedure, which will be used for item classification. Specifically, items are classified by their impact/relation with documental registries (DOC), correct hygiene practices (CHP) and structural issues (STR). For simulations, just 9 items are considered (3, 4 and 2, in each field). The proposal is presented in the following list, and they are an extract of the IFS standard.

DOC: Documental registries

Item 1: There are registries about internal audits.

Item 2: Specifications about raw materials and products will be at employees disposal.

Item 3: There are registries related to PCH incidences.

CHP: Correct hygiene practices

Item 4: There is a procedure for controlling the correct hand hygiene.

Item 5: Cuts and grazes must be covered with coloured bandages with a metal strip.

Item 6: Jewellery, piercings and watches are not allowed.

Item 7: Toilets cannot be accessed directly from working areas where products are handled.

STR: Structural issues

Item 8: Hand washing installations have potable water, liquid soap and hand-drier.

Item 9: All working areas will be suitably illuminated.

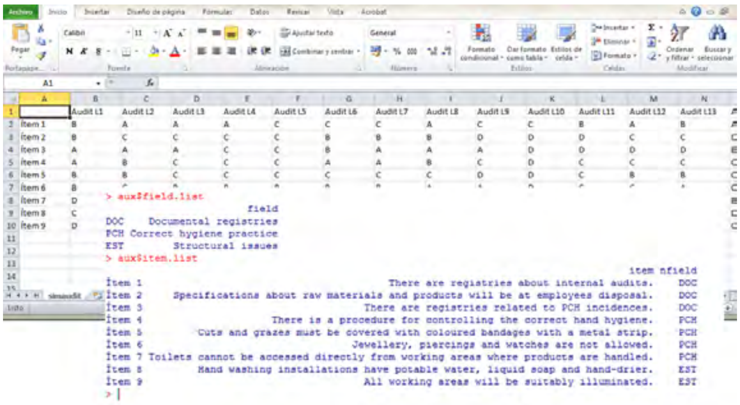


Figure 1: Example of dataset for the simulated scenario. Excel file (output from iAuditoria) and R dataframe elements.

In Figure 1, a partial view of the data in an Excel file (output from iAuditoria) and the R dataframe elements are shown. It is easy to see in the data snapshot that items are valued on a scale A, B, C, D, which are the states of the DTMC for each item. Items are organized by rows and audits are recorded by columns. The same transition probability matrix has been considered for all the items (see Table 1). In an initial observation of the system, all states are equally possible. With these premises, 110 audits have been simulated. A transition plot for item 9 is shown in Figure 2. Note that the transition matrix is stochastic, meaning that the sum by rows is equal to 1. The interpretation of the matrix elements is quite simple, just bearing in mind expression (2): for instance, from the values in the first row, it is clear that if the current state of an item is A, the probability that it remains at A in the next audit is 0.5. The probability of passing from A to D in a single audit is 0.05. Recall that this is just a transition matrix for data simulation. Another way of summarizing the transition matrix is by using a graph, as the one shown in Figure 2.

All simulations have been carried out in R [4] using the recently released package `markovchain` (see [6] for an introduction). This package contains functions for handling and modelling DTMC data. The statistical analysis included in the package allows for estimation of transition matrices (point estimates for the transition probabilities and corresponding standard errors, as well as confidence intervals), prediction and validation tests.

	A	B	C	D
A	0.5	0.3	0.15	0.05
B	0.2	0.5	0.25	0.05
C	0.1	0.25	0.5	0.15
D	0.05	0.15	0.3	0.5

Table 1: Transition matrix for audit simulations with $E = \{A, B, C, D\}$. Note that the rows of the matrix sum up to 1 (stochastic matrix).

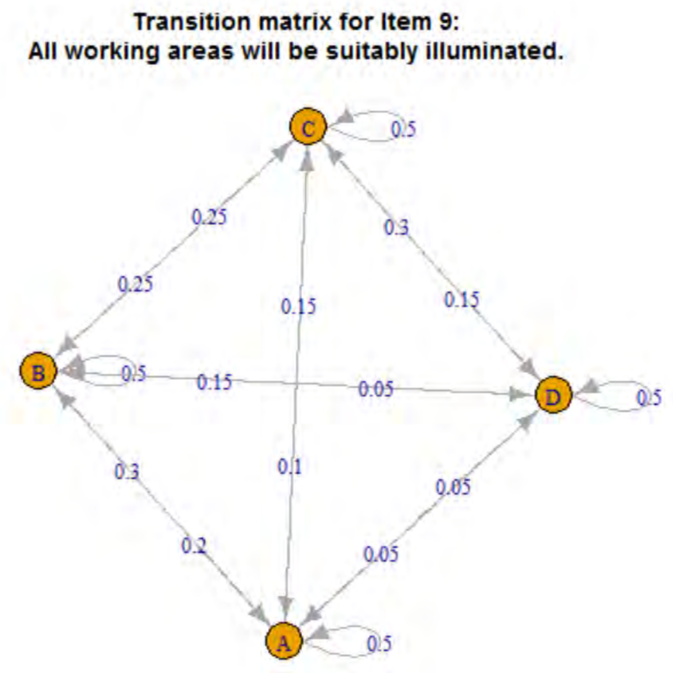


Figure 2: Representation of transition matrices for Item 9: All working areas will be suitably illuminated. States are marked in orange. Arrows indicate the possible transitions (also for remaining states). Transition probabilities displayed close to corresponding arrows.

Simulation results

For illustration purposes, consider that each item is valued A in audit n and the goal is to analyse evolution of the item, specifically, its behaviour in the next audit $(n+1)$. The possible values for audit $(n+1)$ are the ones contained in E (that is, A, B, C or D), but it would be useful to know which ones have a higher probability. Specifically, it is possible to estimate the probability of passing from A to B. Results are shown in Table 2, where maximum likelihood

	MLE	Bootstrap
Item 1	0.353 (0.144)	0.323 (0.022)
Item 2	0.364 (0.128)	0.336 (0.038)
Item 3	0.571 (0.165)	0.525 (0.029)
Item 4	0.071 (0.071)	0.096 (0.026)
Item 5	0.615 (0.218)	0.705 (0.049)
Item 6	0.228 (0.081)	0.257 (0.017)
Item 7	0.280 (0.106)	0.341 (0.049)
Item 8	0.292 (0.110)	0.324 (0.026)
Item 9	0.250 (0.144)	0.273 (0.061)

Table 2: Estimates of P_{AB} probabilities with maximum likelihood and bootstrap approaches for all items. Standard errors in brackets.

	LL-MLE	UL-MLE	LL-Boot	UL-Boot
Item 1	0.116	0.589	0.287	0.359
Item 5	0.257	0.973	0.624	0.786
Item 9	0.012	0.487	0.172	0.373

Table 3: Some examples of 95% confidence intervals for a transition probability from A to B. Results for maximum likelihood (MLE) and bootstrap (Boot) approaches. LL: lower limit. UL: upper limit.

and bootstrap estimators (see [5] for further details on these ones), with their corresponding standard errors, are provided.

This previous exercise could be done with all the transition probabilities in the matrix, and confidence intervals can be as well constructed. Examples for some items are provided in Table 3.

Apart from estimating the probability of passing from one state to another (from A and B, for example) it is also interesting to predict which is the most likely state given the current situation. For instance, if the result for the item is A on the present audit, for the simulated data, the most probable state for the next audit is shown in Table 4. Note that the results with maximum likelihood and bootstrap methods coincide for all items, except for item 9.

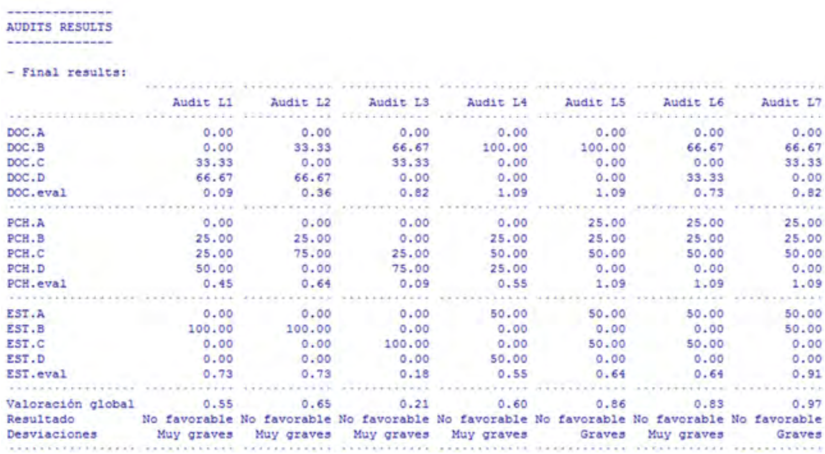
Reporting results

Different summaries of the results can be obtained with the methods presented above, but these ones are provided on an item scale. As mentioned in the Introduction, iAuditoria allows the user not only to fix the scores of the A-B-C-D scale for item values, but also the weights of the different fields and the requirements that must be met for obtaining the final result of the audit. This information can be also easily integrated in the R code, reporting the results shown in Figure 3. In this case, just for the first seven audits, final points for all the items and for the fields, as well as a global value, are provided.

This information could be directly facilitated by the tool as a final summary, but a deeper insight on the results could be also provided to the user. For instance, the evolution of the different field values could be shown as in Figure 4.

	MLE	Bootstrap
Item 1	C	C
Item 2	B	B
Item 3	C	C
Item 4	C	C
Item 5	B	B
Item 6	C	C
Item 7	A	A
Item 8	B	B
Item 9	D	C

Table 4: Predictions for the next audit, given that each item has A in the current one. Maximum likelihood and bootstrap results.


AUDITS RESULTS

- Final results:

	Audit L1	Audit L2	Audit L3	Audit L4	Audit L5	Audit L6	Audit L7
DOC.A	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DOC.B	0.00	33.33	66.67	100.00	100.00	66.67	66.67
DOC.C	33.33	0.00	33.33	0.00	0.00	0.00	33.33
DOC.D	66.67	66.67	0.00	0.00	0.00	33.33	0.00
DOC.eval	0.09	0.36	0.82	1.09	1.09	0.73	0.82
PCH.A	0.00	0.00	0.00	0.00	25.00	25.00	25.00
PCH.B	25.00	25.00	0.00	25.00	25.00	25.00	25.00
PCH.C	25.00	75.00	25.00	50.00	50.00	50.00	50.00
PCH.D	50.00	0.00	75.00	25.00	0.00	0.00	0.00
PCH.eval	0.45	0.64	0.09	0.55	1.09	1.09	1.09
EST.A	0.00	0.00	0.00	50.00	50.00	50.00	50.00
EST.B	100.00	100.00	0.00	0.00	0.00	0.00	50.00
EST.C	0.00	0.00	100.00	0.00	50.00	50.00	0.00
EST.D	0.00	0.00	0.00	50.00	0.00	0.00	0.00
EST.eval	0.73	0.73	0.18	0.55	0.64	0.64	0.91
Valoración global	0.55	0.65	0.21	0.60	0.86	0.83	0.97
Resultado	No favorable	No favorable	No favorable	No favorable	No favorable	No favorable	No favorable
Desviaciones	Muy graves	Muy graves	Muy graves	Muy graves	Graves	Muy graves	Graves

Figure 3: An example of audit results. Fields for item classification identified by horizontal blocks. Audits summarized by columns. Bottom rows: final value and qualitative result.

This type of plots could help the user to identify which field (documental registries, hygiene practices or structural issues) requires further attention in order to prevent severe failures.

The estimates obtained for the transition probabilities can be represented in a matrix as shown in Figure 5. A colour legend is used to identify the different states of the item in the next audit, presenting their estimated transition probabilities accompanied by a coloured circle with size proportional to their values. This information, with a similar graphical display, could be introduced in the iAuditoria report.

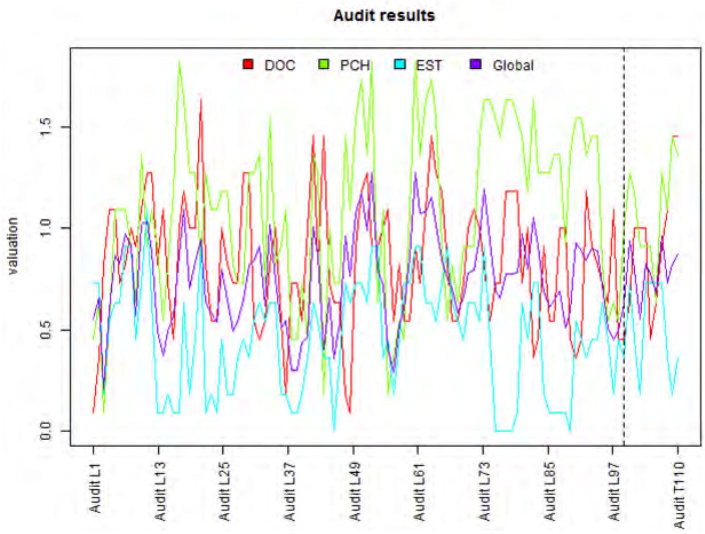


Figure 4: Audit results for different fields and global result along the 110 simulated audits. Red: documental registries (DOC). Green: correct hygiene practices (CHP). Light blue: structural issues (STR). Dark blue: global result.

4 Model validation, limitations and alternative approach

Up to now, the proposed methodology has been applied on simulated data. For a proper application in practice, it is necessary to check that the assumptions imposed on the model are met. Specifically, the Markov property (1) must be tested, along with the order of the chain (one–step, two–steps,...). Homogeneity in time needs also to be tested. The simulated DTMCs satisfy these three requirements. In practice, `markovchain` includes functions for the previous purposes: `verifyMarkovProperty`, `assessOrder` and `assessStationarity`. However, if validation fails, a different strategy (conditioned by the unfulfilled requirements) should be designed.

A feature that seems quite obvious in this setting where there is a list of different DTMCs, associated to a collection of items classified by fields, is that the evolution of these chains may be related. That is, items belonging to the same field may present a similar evolution or can be related, or this could also happen with items that, although belonging to different fields, may refer to the same aspect of the system under study. An idea is to consider the whole set of DTMCs as a multivariate Markov Chain, as proposed and analysed by [2] and [8].

For presenting the multivariate DTMCs model, the (marginal) state space of each chain is the same as before $E = \{A, B, C, D\}$. The multivariate DTMC is defined as: $\mathbf{X}_n$ = states of items $1, \ldots, K$ at n–th audit, $n \geq 0$. Observations of n audits can be organized in a matrix of size $(n \times K)$, being K the number of items. The Markov condition (1) in this new context is formulated as follows: for K sequences (items) with m possible states ($m = 4$ in our example), the

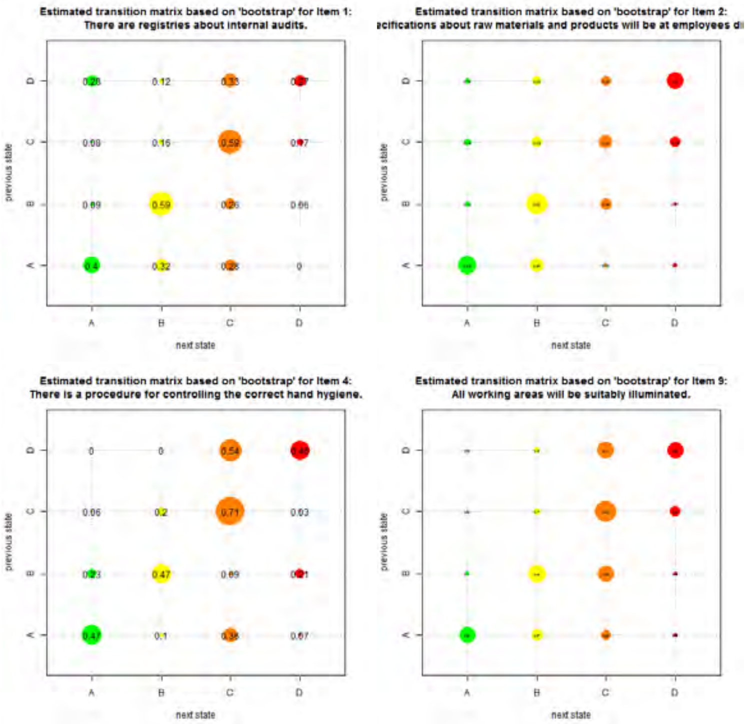


Figure 5: Transition matrices for items 1 and 2 (top row) and 4 and 9 (bottom row). Colours identify the different states: green=A, yellow=B, orange=C and red=D. Size of circles is proportional to the assigned probability.

state probability distribution of the j–th sequence at time $(n+1)$ depends on the state probability distributions of the K sequences (including itself) at time n.

Then, for modelling one–step ahead transition probabilities:

$$p^{(j)} = \sum_{k=1}^{K} \lambda_{kj} P^{(kj)} p_0^{(k)}, \quad \lambda_{kj} \geq 0, \sum_{k=1}^{K} \lambda_{kj} = 1 \tag{3}$$

where $p^{(j)'} = (\mathcal{P}(X_1^j = A), ..., \mathcal{P}(X_1^j = D)) = (p_A^j, p_B^j, p_C^j, p_D^j)$, $p_0^{(k)}$ is the initial probability of the k–th sequence and $P^{(kj)}$ is the one–step transition probability matrix from the states at time n in sequence k to the states at time $(n+1)$ in sequence j. By vectorising the probability distribution vectors with components $p^{(j)'} = (p_A^j, p_B^j, p_C^j, p_D^j)$, and constructing a block–matrix with the λ and P components, (3) can be written in matrix form:

$$\begin{pmatrix} p^{(1)} \\ p^{(2)} \\ \vdots \\ p^{(K)} \end{pmatrix} = \begin{pmatrix} \lambda_{11}P^{(11)} & \lambda_{12}P^{(12)} & \dots & \lambda_{1k}P^{(1K)} \\ \lambda_{21}P^{(21)} & \lambda_{22}P^{(22)} & \dots & \lambda_{2K}P^{(2K)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{K1}P^{(K1)} & \lambda_{K2}P^{(K2)} & \dots & \lambda_{KK}P^{(KK)} \end{pmatrix} \begin{pmatrix} p_0^{(1)} \\ p_0^{(2)} \\ \vdots \\ p_0^{(K)} \end{pmatrix}.$$

Under the imposed conditions on the λ_{kj}, an estimator for these parameters is proposed by [2]. For the previous dataset, the estimated transition probability matrix is shown in (5)

	A	B	C	D
A	0.083	0.292	0.458	0.167
B	0.091	0.424	0.218	0.250
C	0.125	0.218	0.406	0.250
D	0.150	0.250	0.250	0.350

Table 5: Transition matrix for audit simulations by multivariate DTMC modelling. State space: $E = \{A, B, C, D\}$.

No comparisons between the results obtained by the two methods have been performed, and certainly this last one requires further study. Nevertheless, the challenge once the modelling strategy has been analysed is to apply the proposal to real data observed from different audits.

References

[1] T.W. Anderson and L.A. Goodman. Statistical inference about markov chains. *The Annals of Mathematical Statistics*, pages 89–110, 1957.

[2] W-K. Ching, M.K. Ng, and E.S. Fung. Higher-order multivariate markov chains and their applications. *Linear Algebra and its Applications*, 428(2-3):492–507, 2008.

[3] V. G Kulkarni. *Modeling and analysis of stochastic systems.* CRC Press, 2016.

[4] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2016.

[5] M. Skuriat-Olechnowska. Statistical inference and hypothesis testing for markov chains with interval censoring. *Master Theses, TUDelft*, 2005.

[6] G.A. Spedicato, T.S. Kang, S.B. Yalamanchi, M. Thoralf, and D. Yadav. Package: markovchain. easy handling discrete time markov chains. *Version 0.6.5.1 (2016-09-10).*

[7] I. Teodorescu. Maximum likelihood estimation for markov chains. *arXiv preprint arXiv:0905.4131*, 2009.

[8] C. Wang, T.Z. Huang, and W-K. Ching. A new multivariate markov chain model for adding a new categorical data sequence. *Mathematical Problems in Engineering*, 2014, 2014.

Efficient Simulation Method for Current and Power distribution in Electrical Furnaces

---

Academic Coordinator Mª Pilar Salgado Rodríguez
University University of Santiago de Compostela and Technological Institute for Industrial Mathematics.

Business Coordinator Svenn Anton Halvorsen
Company Teknova AS

Team Eva Barazón (University of Santiago de Compostela), Juan Bedoya (University of Santiago de Compostela), Cristina Caravaca (University of Santiago de Compostela), Miguel Carballo (University of Santiago de Compostela), Dolores Gómez (University of Santiago de Compostela and Technological Institute for Industrial Mathematicals),Egil Herland (Teknova AS), María Hidalgo (University of Santiago de Compostela), Aida Martínez (University of Santiago de Compostela), Begoña Nicolás (University of Santiago de Compostela), Álvaro Pais (University of Santiago de Compostela) and Miguel Yermo (University of Santiago de Compostela)

Efficient simulation method for current and power distribution in electrical furnaces

Mads Fromreide[1], *Dolores Gómez*[2], *Svenn Anton Halvorsen*[3], *Egil Herland*[4], *Pilar Salgado*[5]

Abstract

The aim of this study group has been to analyze two different low-frequency electromagnetic models used for computing the power in three-phase electrical furnaces. A simplified AC model, which combines three DC solutions and takes into account the phase shift between the electrodes, is compared with a full AC model which also includes induction effects.The results obtained show that the simplified AC model yields accurate results for the power and current distribution in some regions of the furnace. The analysis developed is an important step to limit the need for using more complicated models in order to obtain accurate results.

1 Introduction

Teknova AS is technology-based industrial research institute established in the southern part of Norway. Teknova has a close cooperation with the Norwegian industrial environment and its core competences include the fields of Smart Instrumentation, Modelling and Simulation and Clean Technologies.

In 2015, Teknova received a research grant from the Research Council of Norway for founding the project "Electrical Conditions and their Process Interactions in High Temperature Metallurgical Reactors (ElMet)". Teknova AS is the project owner of ElMet which also includes the following partners: Norwegian University of Science and Technology (NTNU), Technological Institute for Industrial Mathematics (ITMATI), the University of Oxford, ALCOA Norway AS, Elkem AS and Eramet Norway AS. The interdisciplinary research group combines the experience in mathematical modeling, the understanding of metallurgical process and the industrial experience to analyze and improve industrial challenges which arise in ferroalloy, silicon and aluminium production.

The problem presented to the ESGI 122 lies in the framework of ElMet project and it is related to the electromagnetic modeling of three-phase alternating current (AC) furnaces which are devoted to the production of ferromanganese (FeMn) and other metallurgical processes. The problem and the objectives are described in more detail in the next section.

[1]Teknova AS; mads.fromreide@teknova.no
[2]University of Santiago de Compostela and ITMATI; mdolores.gomez@usc.es
[3]Teknova AS; sah@teknova.no
[4]Teknova AS; Egil.Herland@teknova.no
[5]University of Santiago de Compostela and ITMATI; mpilar.salgado@usc.es

2 Challenge description

Three-phase AC furnaces are used in the production of different feroalloys such as ferromanganese and ferrosillicon. During the ESGI 122, Teknova has focused on furnaces used in the production of FeMn. Figure 1-left shows a sketch of this kind of furnaces. They consist of a cylindrical shell with three cylindrical electrodes buried in the charge material. The raw materials are loaded at the top of the furnace and an alternating current is applied through the top of the three electrodes with 120° electrical phase shift between the electrodes. The electrodes, consisting of good conductors, carry the electric current to the center of the furnace where the main chemical reactions take place. Indeed, the energy developed in the furnace is used to melt and reduce oxides to metal. The interior of the furnace may be divided into different regions which are described in Figure 1-right; notice that we can distinguish a first part with charge components (mixture) at different temperatures, the coke bed where ore, slag and fluxes are molten and finally, the metal. The term coke bed is usually employed to denote the enriched area between the electrode tip and the metal bath, where the oxides are liquid.

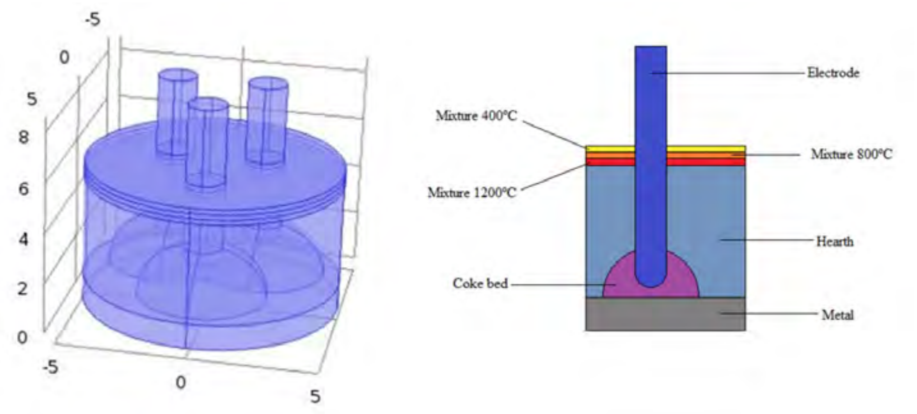


Figure 1: Sketch of three-phase AC furnace (left). Main conducting zones found in furnaces used in FeMn (right).

Previous efforts found in literature to model the multiphysics behavior of electrical furnaces covers aspects of heat transfer, electromagnetic and fluid dynamics. However, an identified gap is the accurate understanding of the effects of three-phase alternating current, including how the associated power distribution governs the chemical reactions and temperature shape profile. An efficient method for computing three-phase current paths and power distribution is required to close this knowledge gap.

Previous electromagnetic simulations within metallurgy have often applied direct current (DC) to study large and three-phase smelting furnaces. Such DC simulations are valid for "only one instant of time" and it is not straightforward to interpret the results for real furnaces; see, for instance [5]. Recently, it was shown that an approximate harmonic solution for three-phase alternating currents (AC) can be described as a superposition of three DC solutions, taking the phase shift between the three electrodes into account. The ESGI group has focused on comparing this simplified model with a genuine AC three-phase model including electromagnetic induction. The simplified AC approximation seems to be valid for small furnaces since induction effects typically are limited in such systems. The main objective during the ESGI 122 has been to clarify how well it works for larger furnaces, when induction effects are significant, typically in the electrodes and molten metal regions of the furnace.

The group splitted into three teams which focused on the following tasks:

- To understand and exploit the simplified low frequency AC model developed by Teknova.
- To develop a genuine (full) AC model and compare the results obtained with the simplified and the full models for different operating conditions.
- To explore the possibility of reducing the computational cost of the full AC model by using symmetry, rotating or periodic conditions.

  In this last task it was not possible to arrive to any satisfactory solution. This is why no section concerning to this point has been included in this report.

All of the models have been implemented in the Comsol Multiphysics® 5.2 software.

3 Mathematical modeling

3.1 Time-harmonic Maxwell equations and electric field regimes

The behavior of electric and magnetic fields are described by means of Maxwell equations, which are written in differential form as follows:

$$\frac{\partial \boldsymbol{\mathcal{D}}}{\partial t} - \mathbf{curl}\, \boldsymbol{\mathcal{H}} = -\boldsymbol{\mathcal{J}}, \tag{1}$$

$$\frac{\partial \boldsymbol{\mathcal{B}}}{\partial t} + \mathbf{curl}\, \boldsymbol{\mathcal{E}} = \mathbf{0}, \tag{2}$$

$$\operatorname{div} \boldsymbol{\mathcal{B}} = 0, \tag{3}$$

$$\operatorname{div} \boldsymbol{\mathcal{D}} = \rho, \tag{4}$$

where $\boldsymbol{\mathcal{B}}$ is the magnetic induction, $\boldsymbol{\mathcal{H}}$ the magnetic field, $\boldsymbol{\mathcal{J}}$ the current density, $\boldsymbol{\mathcal{E}}$ the electric field, $\boldsymbol{\mathcal{D}}$ the electric displacement and ρ the charge density. All of the vector fields depend on the time t and the spatial variable $\mathbf{x} \in \mathbb{R}^3$. These equations are completed with the constitutive laws; here, we will assume linear and isotropic materials and then,

$$\boldsymbol{\mathcal{B}} = \mu \boldsymbol{\mathcal{H}}, \tag{5}$$

$$\boldsymbol{\mathcal{D}} = \varepsilon \boldsymbol{\mathcal{E}}, \tag{6}$$

where μ and ε denote, respectively, the magnetic permeability and the electric permittivity of the material. In conducting media, the electric behavior is characterized by means of the Ohm's law

$$\boldsymbol{\mathcal{J}} = \sigma \boldsymbol{\mathcal{E}}, \tag{7}$$

where σ is the electrical conductivity of the material which usually strongly depends on the temperature. In non conducting materials, σ is null and $\boldsymbol{\mathcal{J}} = \mathbf{0}$.

The modeling of many industrial problems does not require to solve the complete system of Maxwell equations and, depending on the operation conditions, some terms may be neglected. The models used in this work are obtained from different approximations of the full Maxwell system and this is why we will start summarizing the main results obtained from the non-dimensionless analysis described in [7].

First of all, since the current supplied to the electrodes is alternating, it varies sinusoidally in time and we can use a time-harmonic approach due to all the materials have a linear behavior. Namely, we can consider that all fields vary harmonically with time on the form

$$\boldsymbol{\mathcal{F}}(\mathbf{x}, t) = \mathrm{Re}\,[\mathrm{e}^{i\omega t}\,\mathbf{F}(\mathbf{x})], \tag{8}$$

where i is the imaginary unit, $\mathbf{F}(\mathbf{x})$ is the complex amplitude of field $\boldsymbol{\mathcal{F}}$ and ω is the angular frequency, $\omega = 2\pi f$, f being the frequency of the alternating current. Thus, we can work in terms of the complex amplitudes for all the electromagnetic fields.

On the other hand, let us introduce the following quantities:

- δ, the skin depth, defined by

  $$\delta = \sqrt{\frac{1}{\pi\mu\sigma f}}. \tag{9}$$

  We notice that when working with alternating current, the electric current tends to become distributed within a conductor in such a way that the current density is higher near the surface of the conductor. This fact is called skin effect, and in cylindrical electrodes with strong skin effects, the skin depth is approximately equal to the depth where the current density has fallen to 1/e from its value at the surface.

- λ, the wavelength, given by

  $$\lambda = \frac{2\pi c}{\omega} \tag{10}$$

  with c being the speed of light.

- L, a characteristic length scale.

The non-dimensionless analysis stated in [7] establishes different regimes for the electric field depending on the scale parameter L:

i) $\lambda \ll L$: electromagnetic waves;

ii) $\lambda \gg L$ and δ not too large, alternating current (AC), where we can neglect the term $\partial\boldsymbol{\mathcal{D}}/\partial t$ in equation (1); thus, this equation in the time-harmonic case is then reduced to

$$\mathbf{curl}\,\boldsymbol{\mathcal{H}} = \boldsymbol{\mathcal{J}},$$

or $\mathbf{curl}\,\mathbf{H} = \mathbf{J}$ in terms of complex amplitudes. The set of equations obtained from the Maxwell equations by using this simplification is the so-called *eddy currents model* which will be described in detail in Section 3.3. Moreover, in this case, we distinguish the following situations:

– High frequency: $\delta \ll L$;
– Low frequency: $\delta \sim L$ (same order of magnitude).

iii) $\lambda \gg L$ and $\delta \gg L$, direct current approximation (DC), which leads to an approximation where the electric field is conservative at each instant of time, and then $\mathbf{curl}\,\boldsymbol{\mathcal{E}} = \mathbf{0}$.

For electrical smelting furnaces, we have that $\lambda \gg L$ and the electromagnetic waves can be neglected. However, a DC approximation strongly depends on the value of the parameter L/δ which is different in each material. Actually, the computation of the electromagnetic fields in the furnace usually requires the approximation of a genuine alternating current model in a 3D domain. Since these 3D computations are complex and lengthy, Halvorsen et al. [6] have recently proposed a simplified AC model based on solving several DC problems and use tools related with the time-harmonic model. The authors have advanced in this paper that the approximation can be acceptable in some parts of the furnace but not in the electrodes and the metal. Here, we will analyze the validity of this approximation in coke beds and hearth.

To attain this goal, we summarize the AC simplified model in the next section. Then, we introduce the full AC model and compare the results obtained with both of them.

3.2 The AC low-frequency approximation

In [6], the authors have developed a very fast model based on a DC approximation to compute the electromagnetic field in a three-phase AC furnace. The idea is to solve three-basis problems where the potential is equal to 1 volt on the top of one electrode and null on the two others. Then, taking into account the linear behavior, the three solutions can be combined to take into account the phase-shift between electrodes, but ignoring the electromagnetic induction. Next, we summarize the main points of this approximation.

Let us consider a three-dimensional bounded domain Ω_{C} composed by the electrodes and the conducting materials around. As we explained above, the DC approximation leads to consider that the electric field $\boldsymbol{\mathcal{E}}$ is conservative at each time t and we can write

$$\boldsymbol{\mathcal{E}}(\mathbf{x},t) = -\,\mathbf{grad}\,\mathcal{U}(\mathbf{x},t) \text{ in } \Omega_{\mathrm{C}}.$$

From the Ohm's law and the current conservation, div $\boldsymbol{\mathcal{J}} = 0$, we obtain the classical equation for a DC model:

$$-\operatorname{div}(\sigma\,\mathbf{grad}\,\mathcal{U}) = 0 \text{ in } \Omega_{\mathrm{C}}.$$

This model is completed by applying the electric potential (voltage) at the top of each electrode and zero normal current at all other boundaries. Moreover, in the time-harmonic case, we could write

$$\mathcal{U}(\mathbf{x},t) = \operatorname{Re}\left[\mathrm{e}^{i\omega t}\, U(\mathbf{x})\right].$$

Since the problem is linear, the proposal described in [6] consists in defining three DC basis problems in real variable and build the complex solution U as a linear combination of the solutions $\tilde{U}_i$ of the three basis problems. Namely, for $i = 1, 2, 3$, we define

$$-\operatorname{div}(\sigma \,\mathbf{grad}\, \tilde{U}_i) = 0 \text{ in } \Omega_C, \tag{11}$$

$$\tilde{U}_i = \delta_j^i \text{ on } \Gamma_C^j, \quad j = 1, 2, 3, \tag{12}$$

$$-\sigma \,\mathbf{grad}\, \tilde{U}_i \cdot \mathbf{n} = 0 \text{ on } \partial\Omega_C \setminus (\Gamma_C^1 \cup \Gamma_C^2 \cup \Gamma_C^3), \tag{13}$$

where Γ_C^j represents the top of the electrode j and δ_j^i is the Kronecker delta function, equal to 1 if $i = j$ and null if $i \neq j$. Figure 2 shows a view of the electrodes and coke beds and the voltage imposed in one of the basis problems.

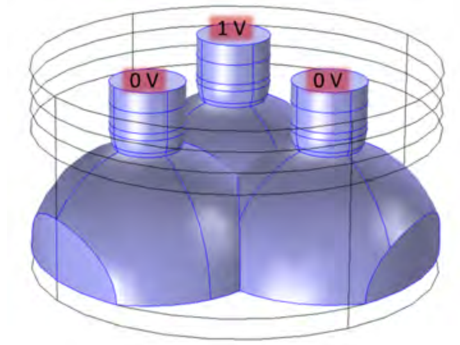


Figure 2: Electrodes and coke beds in the DC problem (11)-(13). Potential imposed in one of the basis problem.

Then, if we know the amplitudes of the voltages in each electrode, u_{1a}, u_{2a} and u_{3a}, we can build the complex potential U by taking into account the phase-shift between the electrodes:

$$U(\mathbf{x}) = u_{1a}\tilde{U}_1(\mathbf{x}) + u_{2a}\mathrm{e}^{-\frac{2\pi}{3}i}\tilde{U}_2(\mathbf{x}) + u_{3a}\mathrm{e}^{\frac{2\pi}{3}i}\tilde{U}_3(\mathbf{x})$$

and consequently, the electric field at each time,

$$\boldsymbol{\mathcal{E}}(\mathbf{x}, t) = -\,\mathbf{grad}\,\mathcal{U}(\mathbf{x}, t) = -\,\mathbf{grad}\,\mathrm{Re}\,[\mathrm{e}^{i\omega t}\, U(\mathbf{x})].$$

The time averaged power density in a cycle and at each point $\mathbf{x}$, will be computed in this simplified model by using the expression:

$$dP_{dc}(\mathbf{x}) = \frac{|\mathbf{J}|^2}{2\sigma},$$

where $\mathbf{J}(\mathbf{x}) = \sigma\mathbf{E}(\mathbf{x}) = -\sigma\,\mathbf{grad}\,U(\mathbf{x})$.

3.3 The time-harmonic eddy currents model

As we advanced above, the eddy currents model is obtained from Maxwell equations by neglecting the term including the electric displacement in Ampere's law. Moreover, taking into account that the electric field is not required in non conducting materials, the time-harmonic eddy current model leads to solve the following equations defined in the whole space:

$$\mathbf{curl}\,\mathbf{H} = \mathbf{J}, \tag{14}$$
$$i\omega\mathbf{B} + \mathbf{curl}\,\mathbf{E} = \mathbf{0}, \tag{15}$$
$$\operatorname{div}\mathbf{B} = 0, \tag{16}$$

with the constitutive laws $\mathbf{B} = \mu\mathbf{H}$ and $\mathbf{J} = \sigma\mathbf{E}$; see, for instance, Chapter 1 of [2]. To solve these equations by using a finite element method, we restrict them to a simply connected 3D bounded domain Ω consisting of two parts, Ω_{C} and Ω_{D}, occupied by conductors and dielectric, respectively; see Figure 3. The conducting domain will be the same as that used for the previous DC problem while the dielectric will be the air around Ω_{C} which is added in order to impose suitable boundary conditions on Γ. The domain Ω is assumed to have a connected boundary Γ. We denote by $\Gamma_{\mathbf{C}}$ and Γ_{D} the open surfaces such that $\bar{\Gamma}_{\mathrm{C}} := \partial\Omega_{\mathrm{C}} \cap \Gamma$ is the outer boundary of the conducting domain and $\bar{\Gamma}_{\mathrm{D}} := \partial\Omega_{\mathrm{D}} \cap \Gamma$ that of the dielectric domain. We also denote by $\mathbf{n}$ a unit normal vector to a given surface.

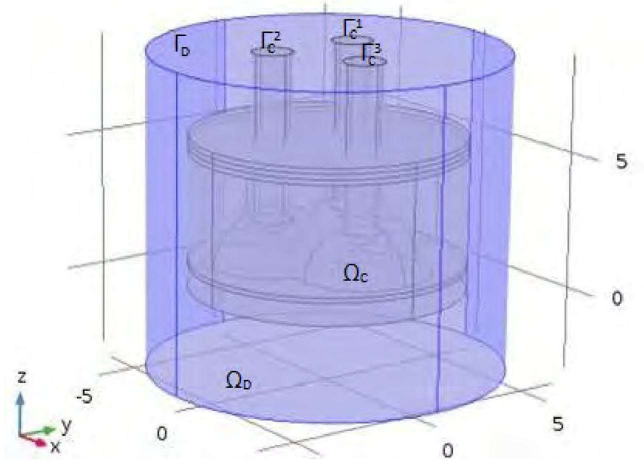


Figure 3: Sketch of the domain for the full AC model.

The model must be completed with suitable boundary conditions on Γ, and we consider the following ones:

$$\mathbf{E} \times \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_{\mathbf{C}}, \tag{17}$$
$$\mu\mathbf{H} \cdot \mathbf{n} = 0 \quad \text{on } \Gamma. \tag{18}$$

Condition (17) means that the electric current enters the domain Ω_{C} perpendicular to the boundary whereas condition (18) implies that the magnetic field is tangential to the boundary.

On the other hand, we assume that the outer boundary of the conductor $\partial\Omega_{\mathrm{C}} \cap \Gamma$ has three disjoint connected components $\Gamma_{\mathbf{C}}^{n}$, which corresponds to the top of the electrodes where the current or the voltage drops will be prescribed.

In particular, we will fix a ground in the first electrode, where we will impose a null potential and on the two other electrodes we will prescribe a current, i.e.,

$$\int_{\Gamma_{\mathbf{C}}^n} \mathbf{J}\cdot\mathbf{n}\,\mathrm{dS} = \mathrm{I}_n, \quad n = 2,3.$$

Notice that in this model, with 120° phase shift between the electrodes, $\mathrm{I}_3 = \mathrm{I}_2 e^{\frac{2\pi}{3}i}$.

To understand the idea of fixing the potential on the top of one of the electrodes, we notice that from (18), we have that $\mathbf{curl}\,\mathbf{E}\cdot\mathbf{n} = 0$ on Γ. Then there exists a smooth function V such that $\mathbf{E}\times\mathbf{n} = -\nabla \mathrm{V}\times\mathbf{n}$ on Γ. Moreover, the boundary condition (17) implies that V must be constant on each connected component of $\Gamma_{\mathbf{C}}$. In particular, we will fix the value of this constant potential to zero on $\Gamma_{\mathbf{C}}^1$; notice that the constant values on $\Gamma_{\mathbf{C}}^2$ and $\Gamma_{\mathbf{C}}^3$ will be obtained as a results.

Thus, the problem defined in Ω, can be written as follows:

$$i\omega\mathbf{B} + \mathbf{curl}\,\mathbf{E} = \mathbf{0}, \tag{19}$$

$$\mathbf{curl}\,\mathbf{H} = \mathbf{J}, \tag{20}$$

$$\operatorname{div}\mathbf{B} = 0, \tag{21}$$

$$\mathbf{B} = \mu\mathbf{H}, \tag{22}$$

$$\mathbf{J} = \sigma\mathbf{E}, \tag{23}$$

$$\mathbf{E}\times\mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_{\mathbf{C}}, \tag{24}$$

$$\mu\mathbf{H}\cdot\mathbf{n} = 0 \quad \text{on } \Gamma, \tag{25}$$

$$\mathrm{V} = 0 \;\; \text{on } \Gamma_{\mathbf{C}}^1, \tag{26}$$

$$\int_{\Gamma_{\mathbf{C}}^n} \mathbf{J}\cdot\mathbf{n}\,\mathrm{dS} = \mathrm{I}_n, \quad n = 2,3 \tag{27}$$

$$\mathrm{V} = \text{constant on } \Gamma_{\mathbf{C}}^n \quad n = 2,3. \tag{28}$$

This AC model can be approximated by using different formulations (see for instance Chapter 8 from [2] or Chapter 11 from [3]). In this work we have solved it in COMSOL by using the AC/DC module which considers a formulation based on magnetic vector potential/scalar electric potential (the well known $\mathbf{A}/\mathrm{V}$ – formulation). We recall here that the magnetic vector potential is defined from the equation (21) which implies that there exists a vector field $\mathbf{A}$, such that

$$\mathbf{B} = \mathbf{curl}\,\mathbf{A}.$$

On the other hand, the electric scalar potential V in the conducting domain is defined from the definition of $\mathbf{A}$ and the Faraday's law $\mathbf{curl}(\mathbf{E} + i\omega\mathbf{A}) = \mathbf{0}$ which implies that $\mathbf{E} + i\omega\mathbf{A} = -\,\mathbf{grad}\,\mathrm{V}$. The gauge condition $\operatorname{div}\mathbf{A} = 0$ joint with suitable boundary conditions are added to ensure the uniqueness of the magnetic vector potential. Namely, a detailed explanation of the $\mathbf{A}/\mathrm{V}$ – formulation with the boundary conditions detailed above to impose currents and voltages drops can be found in [1] or [3]; in the geometric framework described above, the problem is summarized as follows:

Given complex numbers $\mathrm{I}_n,\ n = 2, 3$, *find a vector field* $\mathbf{A}$ *defined in* Ω, *and a scalar field* V *defined in* Ω_{C} *and constant in* $\Gamma^2_{\mathbf{C}}, \Gamma^3_{\mathbf{C}}$, *such that*

$$\sigma(i\omega\mathbf{A} + \mathbf{grad}\,\mathrm{V}) + \mathbf{curl}(\frac{1}{\mu}\,\mathbf{curl}\,\mathbf{A}) = \mathbf{0} \text{ in } \Omega, \tag{29}$$

$$\operatorname{div}\mathbf{A} = 0 \text{ in } \Omega, \tag{30}$$

$$\mathbf{A} \times \mathbf{n} = 0 \text{ on } \Gamma, \tag{31}$$

$$\sigma(i\omega\mathbf{A} + \mathbf{grad}\,\mathrm{V}) \cdot \mathbf{n} = 0 \text{ on } \partial\Omega_{\mathrm{C}} \setminus \Gamma_{\mathbf{C}}, \tag{32}$$

$$\mathrm{V} = 0 \text{ on } \Gamma^1_{\mathbf{C}}, \tag{33}$$

$$\int_{\Gamma^n_{\mathbf{C}}} \sigma(i\omega\mathbf{A} + \mathbf{grad}\,\mathrm{V}) \cdot \mathbf{n}\,\mathrm{dS} = -\mathrm{I}_n,\ n = 2, 3. \tag{34}$$

The time averaged power density in a cycle and at each point $\mathbf{x}$, will be computed in this full AC model by using the expression:

$$dP_{ac}(\mathbf{x}) = \frac{|\mathbf{J}|^2}{2\sigma},$$

where $\mathbf{J}(\mathbf{x}) = \sigma\mathbf{E}(\mathbf{x}) = -\sigma(i\omega\mathbf{A}(\mathbf{x}) + \mathbf{grad}\,\mathrm{V}(\mathbf{x}))$.

4 Numerical solution

In this section we present some results obtained by comparing the AC simplified model with the full AC model by using Comsol Multiphysics® 5.2 software. The geometry of the furnace and the files corresponding to the simplified AC model have been provided by Teknova.

4.1 Physical data and geometrical parameters

Table 1 lists the values assigned to the electrical conductivity for the different materials considered in the furnace (see Figure 1-right). The relative magnetic permeability has been set to one for all the materials. Notice that the electrical conductivity in electrodes and metal is very high and, therefore, the skin depth is very small in these zones; consequently, we cannot expect a good approximation by the simplified AC model in this part. Thus, we have focused on comparing the power dissipated in coke beds and hearth for two kind of geometries: one of them with the coke bed of each electrode without contact and the other one with joined coke beds (see Figure 4). Table 2 shows the main geometric parameters used in both simulations.

Concerning the source current, the reference value to compare the simplified and the full AC models has been the value of the RMS current entering each electrode which is equal to 339.25kA in all cases. We emphasize that this reference value is obtained from the DC model for a given potential and we cannot use directly the potential due to the meaning of potentials is different for the simplified and the full model.

Figure 4: View of the coke beds from the top: separated (left) and joined (right.

Material	Conductivity (S/m)	Skin depth (m)
Mixture 400 °C	0.075	259.90
Mixture 800 °C	0.15	183.77
Mixture 1200 °C	15	18.38
Hearth	15	18.38
Coke beds	500	3.18
Electrodes	150000	0.18
Metal bath	150000	0.18

Table 1: Values of electrical conductivity and skin depth for the different furnace zones.

Parameter	Dimension (m)
Height of the Mixture layer at 400 °C	0.2
Height of the Mixture layer at 800 °C	0.2
Height of the Mixture layer at 1200 °C	0.2
Furnace diameter	10
Distance between electrodes	4
Distance from electrode tip to metal bath	1
Height of hearth layer	4
Height of metal layer	1
Electrode diameter (separated coke beds)	1
Coke bed: x axis (separated coke beds)	1.5
Coke beds: y axis (separated coke beds)	1.5
Height of coke beds (separated coke beds)	1.5
Electrode diameter (joined coke beds)	1.5
Coke bed: x axis (joined coke beds)	2.5
Coke beds: y axis (joined coke beds)	2.5
Height of coke beds (joined coke beds)	2
Height of air layer (over and under the furnace)	3

Table 2: Geometrical parameters in the furnace.

4.2 Numerical results: furnace with separated coke beds

As we advanced in the previous section, the key parameter to analyze the quality of the approximation of the AC simplified model is L/δ, since this factor determines the strength of the induction effects. From a modelling point of view, the increase in induction effects obtained by scaling the size of the furnace by a factor of α is similar to the increase in induction effects obtained by scaling the frequency with a factor of α^2. In the simulations, we have exploited this by changing the frequency in order to investigate different strength of the induction effects, instead of changing the size of the furnace, but the results of both approaches should be similar. In the case of separated coke beds the values for L are 1.5 m and 4 m in coke beds and hearth. Figure 5-left shows the value of L/δ in coke beds and hearth in terms of the frequency.

The power dissipated in coke beds and hearth in megawatts has been computed for frequencies of $50, 250$ and 500Hz; the values are detailed in Table 3. In particular, the dissipated power in each subdomain $\mathcal{O}$ is obtained as the integral of the power density: $\int_{\mathcal{O}} dP_{dc}$ and $\int_{\mathcal{O}} dP_{ac}$ for the simplified and full model, respectively.

An increase of the frequency should be accompanied with a refinement of the mesh near the surface to capture the skin effect. However, due to the high computational effort in 3D and the limited time during the ESGI we have worked with the same mesh in all cases. For this reason, we emphasize that the results show qualitative conclusions but more effort would be needed in order to analyze the quantitative values.

	AC simplified	AC full f = 50Hz	AC full f = 250Hz	AC full f = 500Hz
Coke beds	34.11	34.10	35.99	40.17
Hearth	1.93	3.38	34.71	113.38

Table 3: Power (MW) in coke beds and hearth for the different values of frequency. Separated coke beds.

Figure 5: L/δ vs. the frequency (left). Relative percentual error in terms of power vs. frequency (right). Separated coke beds.

Figure 5 shows, on the left, the behavior of L/δ vs. the frequency and the errors in the power on the right. Notice that the most important errors can be found in hearth. In particular, we can appreciate that even for values of $L/\delta < 1$ in one of the materials (for instance, the coke beds) the power approximation can be poor in the hearth. Indeed, we have seen that the changes in the current distribution are significant due to the skin effect; see, for instance, Figures 6 and 7, which show the current density in hearth obtained with the simplified model and with the full AC model, respectively. In the case of 500 Hz the current concentrates more on the top part of the hearth and reach values much greater in this zone.

4.3 Numerical results: furnace with joined coke beds

Here, we consider the case when the coke beds of the three electrodes are in contact. The values for L are 2 m and 4 m in coke beds and hearth, respectively. Figure 8 shows the value of L/δ in coke beds and hearth in terms of the frequency. Also, in Figs. 9-12, the current densities for the hearth and the coke bed can be compared for the two different models. In this case, the errors are greater in the hearth but in the coke bed the errors are more important than in the previous case for higher frequencies. In particular, in Figure 12 one can appreciate that for higher frequencies the skin effect is very important in the coke beds where the current density concentrates at the top part near the joined zone between them.

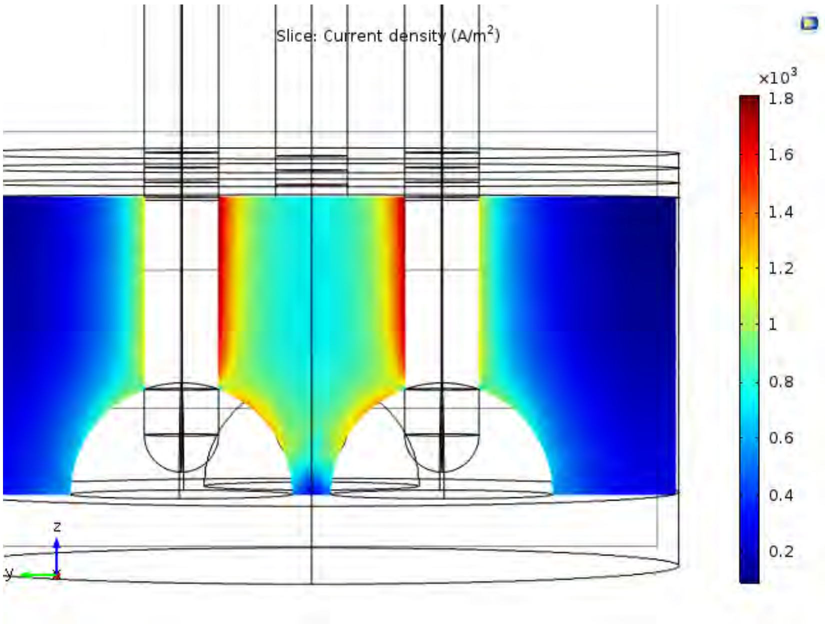


Figure 6: Simplified AC model. Current density in hearth. Maximum: 1.8e3 A/m^2. Separated coke beds.

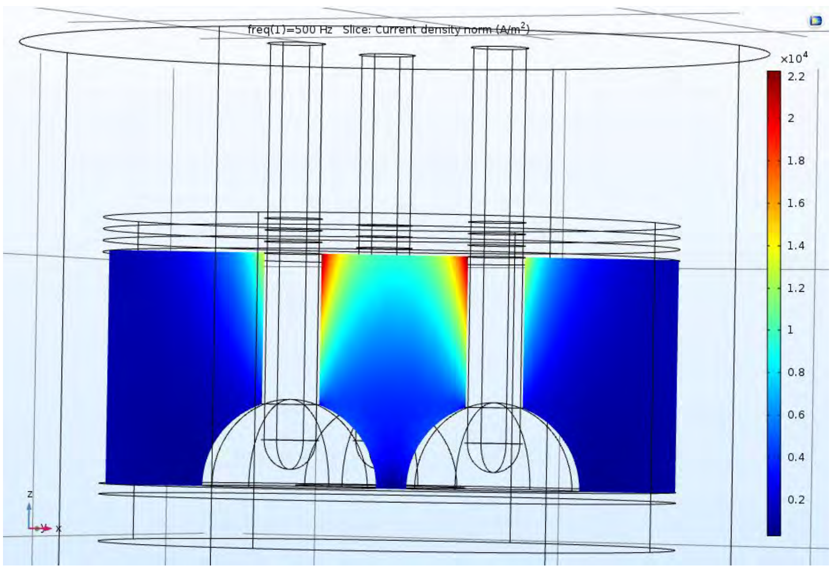


Figure 7: AC full model (f = 500 Hz). Current density in hearth. Maximum: 2.2e4 A/m^2. Separated coke beds.

	AC simplified	AC full f = 50Hz	AC full f = 250Hz	AC full f = 500Hz
Coke beds	44.57	45.24	55.71	67.80
Hearth	2.08	3.25	22.14	65.00

Table 4: Power (MW) in coke beds and hearth for the different values of frequency. Joined coke beds.

Figure 8: L/δ vs. the frequency (left). Relative percentual error in terms of power vs. frequency (right). Joined coke beds.

5 Conclusions

The main objective during the meeting has been to establish limits where a simplified AC model can be used to determine the power distribution in an electric furnace depending on the size of the system and the frequency of the current. The group has used Comsol Multiphysics as the main tool and the results have led to the following conclusions:

- The AC simplified model is cheaper than the full AC model in terms of computational effort (approx. 2.5 minutes versus 7 minutes for a coarse mesh).
- The validity of the AC simplified model strongly depends on the value of L/δ both in the coke beds and hearth. This effect is stronger if the coke beds are joined.
- Even for values of $L/\delta < 1$ in one of the materials (for instance, the coke beds) the power approximation can be poor in the hearth.
- For the most relevant situation at f=50 Hz, the error associated with using the simplified model is small in the coke bed.

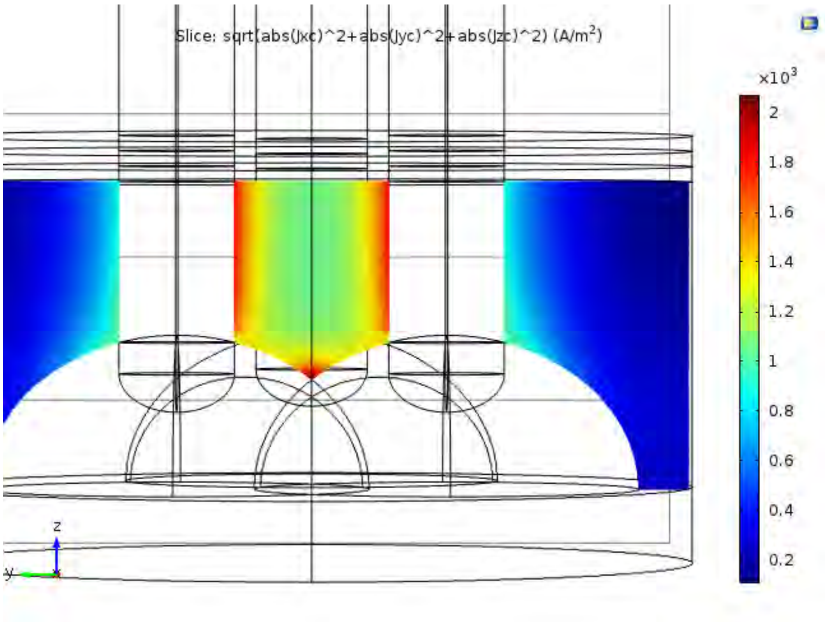


Figure 9: Simplified AC model. Current density in hearth. Maximum: 2e3 A/m^2. Joined coke beds.

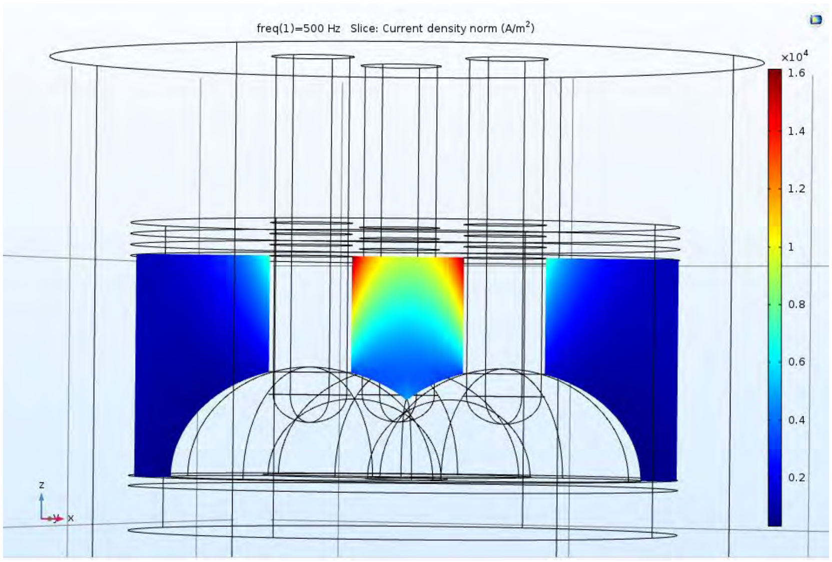


Figure 10: AC full model (f = 500 Hz). Current density in hearth. Maximum: 1.6e4 A/m^2. Joined coke beds.

- The full AC model is more expensive at higher frequencies to capture the skin effect.

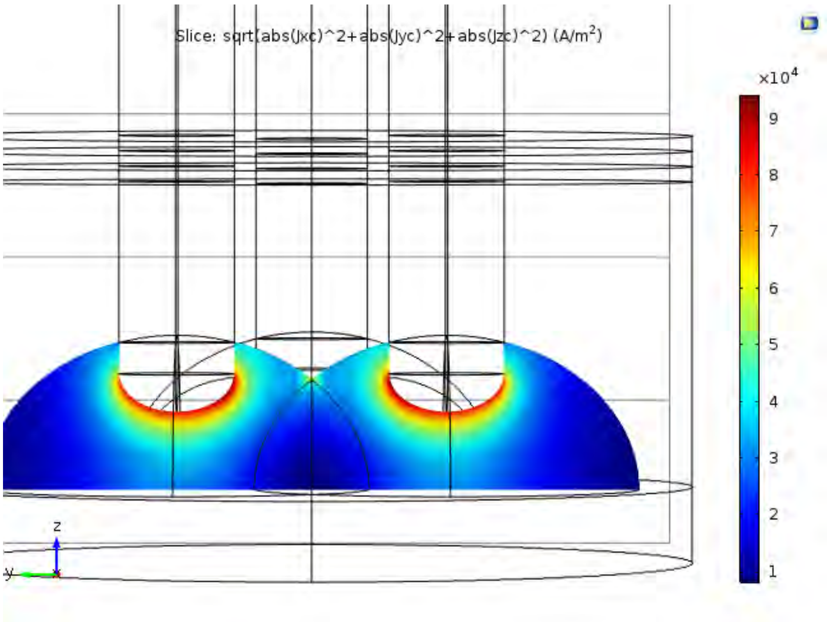


Figure 11: Simplified AC model (f = 500 Hz). Current density in coke beds. Maximum: 9.e4 A/m^2. Joined coke beds.

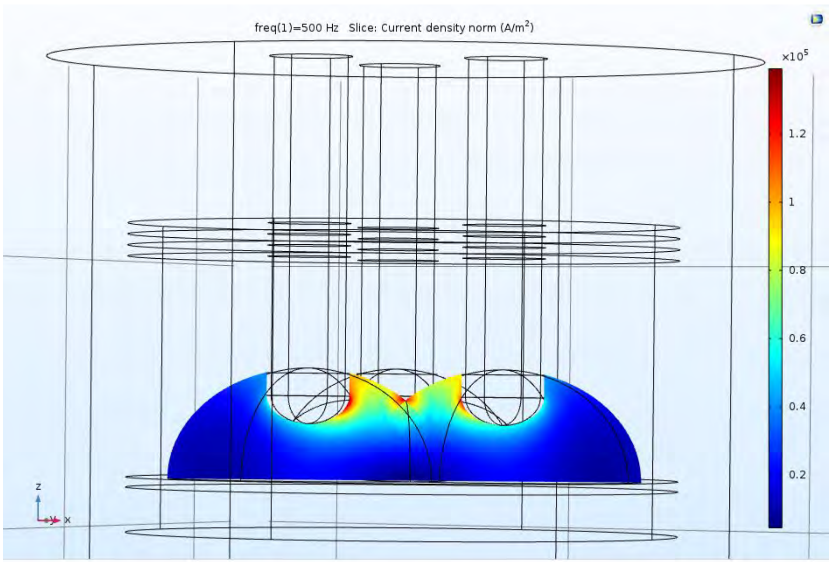


Figure 12: AC full model (f = 500 Hz). Current density in hearth. Maximum: 1.3e5 A/m^2. Joined coke beds.

References

[1] A. Alonso Rodríguez. *Formulation via vector potentials of eddy-current problems with voltage or current excitation.* In: Communications in Applied and Industrial Mathematics **2**: 1, 2011. Available at: http://caim.simai.eu/index.php/caim/article/view/369

[2] A. Alonso Rodríguez and A. Valli. *Eddy current approximation of Maxwell equations, ser. MS&A. Modeling, Simulation and Applications*, Springer-Verlag, Milan, Italia, vol. 4, 2010.

[3] A. Bermúdez, D. Gómez, P. Salgado. *Mathematical models and numerical simulation in electromagnetism*, Springer-Verlag, Milan, Italia, vol. 74, 2014.

[4] A. Bermúdez, R. Rodríguez, P. Salgado. FEM for 3D eddy current problems in bounded domains subject to realistic boundary conditions. An application to metallurgical electrodes, *Computer Methods in Applied Mechanics and Engineering*, **12**, (2005) no. 1, pp. 67–114.

[5] M. Dhainaut. *Simulation of the electric field in a submerged arc furnace*, Proceedings of the Tenth International Ferroalloy Congress, pp. 605–613, 2004.

[6] S.A. Halvorsen, H.A.H. Olsen, M. Fromreide. *An efficient simulation method for current and power distribution in 3-phase electrical smelting furnaces*, Proceedings of the 17th IFAC Symposium on Control, Optimization and Automation in Mining, Mineral and Metal Processing, Vienna, Austria 2016.

[7] R. Schlanbusch, S.A., Halvorsen, S. Shinkevich, D. Gómez. Electrical scale-up of metallurgical processes. COMSOL Conference, Cambridge, UK, 2014. http://www.comsol.com/

Creation of an "Oracle" to support facilities management

Academic Coordinator Javier Roca Pardiñas
University University of Vigo and Technological Institute of Industrial Mathematics

Business Coordinator Pedro Perez Gabriel
Company Managing Director of Ecomanagement Technology S.L.

Creation of an “Oracle” to support facilities management

D.H. Heinisch Rodicio P. Pérez Gabriel, N. Vilela Barreira, J. Roca Pardiñas

Abstract

The aim of this project was to apply mathematical algorithms to help predicting issues in air conditioning systems which are deployed in shops all around the world, suggesting the best possible solution to an employee via a central monitoring platform which is already existing. The system should also adapt to changing conditions (external temperature) and take into account what can be considered the optimal comfort in each individual shop. Possible solutions were analyzed and the results as well as future work lines are gathered in this paper.

1 Introduction

Ecomanagement Technology S.L., or short, EcoMT is a company which offers solutions for remote control and monitoring of multisite facilities with some particular characteristics: relative simplicity, high amount of similar facilities, low financial resources and available maintenance and operation staff. As a result, energy would be used in a more efficient manner, thus resulting in lower expenses for the company, enviroment friendlier facilities and the ability to keep up with the advancing legal constraints. They developed a telemanagement and monitoring platform (OTEA) that aims at comprehensive management of different facilities. Further improvements raise this platform to the field of expert systems, which final aim is to take inteligent decisions for the facilities it manages without human intervention and at a low implementation cost. As for today, EcoMT's platform has been analyzing and managing clima and lighting in over 2,000 establishments with more than 600,000 variables and three trillion data for over five years. The members of EcoMT have designed a road map to capture their aims for the OTEA platform as shown in Figure 1. The first two steps (Smart and Guru OTEA) implement efficient measurements and some predictive models. The last evolution, Sovereign OTEA, would fit the above defined final aim of expert systems to perfection. However, this project focuses around Oracle OTEA, a first step towards human independent monitoring and decision making. The road map for the design of this platform is summarized in Figure 2.

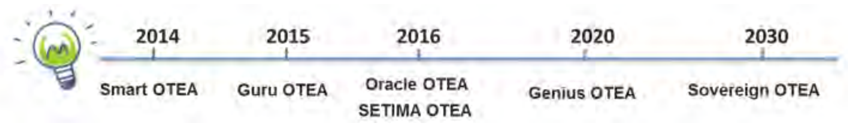


Figure 1: "Road map" to Sovereign OTEA

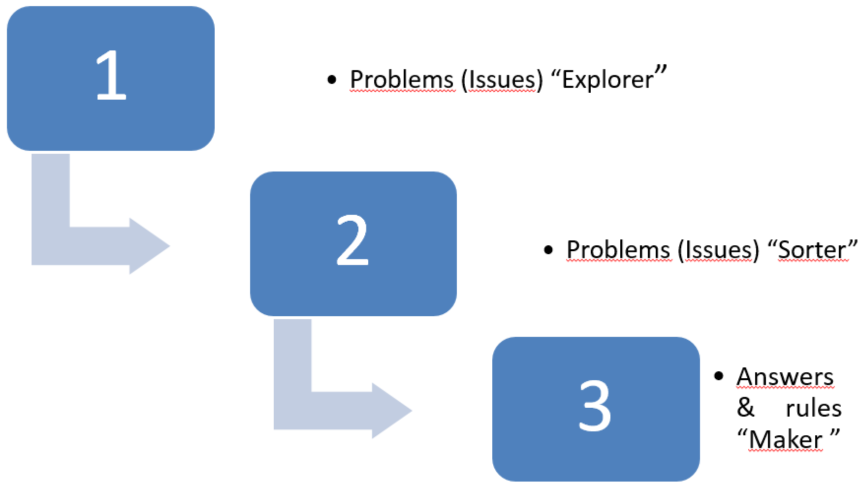


Figure 2: "Road map" to implement OTEA Oracle

This project in particular focuses on the air conditioning system in shops. At first, we define comfort conditions, and based on these, which conditions are to be detected as issues. Next, we compute an algorithm to find these problems. Moreover, we analyze the available variables to see which ones show high correlations with the ones determining comfort, and based on the selected variables, we propose an issue sorter. To conclude we present the results and future work lines.

Note that the time available was very limited and thus most of the work will need revision and polishing, while the development of many ideas will remain for future work.

2 Air conditioning system

To get a deeper understanding of the previously mentioned aims, it is helpful to first understand how the air conditioning system is deployed, and how it is monitored. Let's take the air-water system as an example. It is illustrated in Figure 3 and consists of three main elements:

- Thermal generation equipment: a chiller or heat pump which respectively cools or heats water in a closed circuit.

- Water distribution: a network of isolated pipes that distribute water, cooled or heated by the thermal generation equipment, to the fan coil units.
- Fan coil units (or air conditioners): they use the thermal energy in the water to adapt the temperature of the air which is blown into the room.

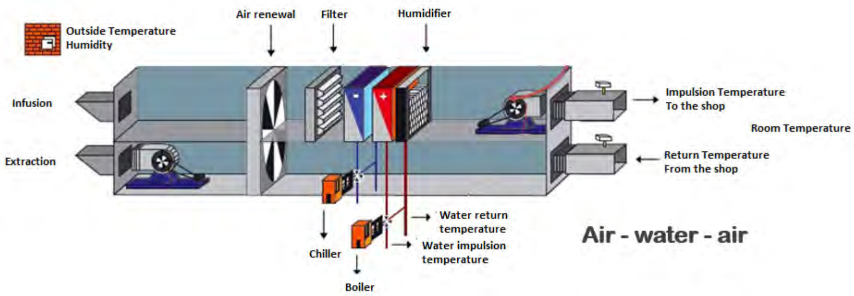


Figure 3: HVAC air conditioning system

Making use of this image, in the next section we will explain the different variables that we find ourselves with in this problem.

3 Variables

One of the first things to be considered is that each shop has a different number of air conditioning machines, which don’t have to be of the same kind as shown above, a control center and several measuring instruments. The latter will provide information on different temperatures and power consumptions. The data obtained through some of the measuring instruments may be shared between several machines in the same area of the shop, but we are not assured to have every temperature data for every machine. In Table 3 we summarize the different power consumption and temperature variables, highlighting the ones we consider to be most relevant at first. Also note that for each air conditioner we may know all or only some of the following.

There are other variables which are not obvious to see at first but we will use further along this project, e.g., a boolean variable that states if there has been an issue the previous day(s),which day of the week it is or if it is a holiday. Last but not least, other variables that weren’t regarded in this project, could surely add important information if handled correctly,such as the occupation and isolation of the shop.

Now that we defined the different variables, in the next section we procede to discuss what comfort conditions are, as well as issues.

Variable	Name	Description
Room temperature	T_{amb}	Temperature at the measuring instrument inside the shop which is closest to the air conditioner
External temperature	T_{ext}	Temperature obtained through professional instruments in a neighborhood of the shop
Impulsion temperature	T_{imp}	Temperature with which the air conditioner blows air into the shop
Return temperature	T_{ret}	Temperature with which the air conditioner blows air out of the shop
Setpoint temperature	T_{con}	Temperature set via the control platform which should ensure best working or shopping conditions
General consumption	P_{gen}	Total power consumption of the shop (includes lighting and others)
Climate consumption	P_{clim}	Power consumption only due to air conditioning

Table 1: Temperature and power consumption variables

4 Defining "comfort" and "issue"

The task of defining these two terms may look easy, but when you take into consideration that they may vary from one country to another, and even from one shop to the other just accross the street, it becomes clear that a universal definition is at least challenging to obtain.

Starting with "comfort conditions", these refer to the room temperature, but there is no way to assure that every person regards the same as comfortable. For the sake of keeping the project simple to be able to advance with the other problems, we focused on summer periods, when most issues take place. Also only considering shops in Spain for the time being, we define "comfort conditions" as flat 24 ℃, but this could be altered from the monitoring platform or even the shop supervisor. This temperature is the one the air conditioning system will aim to reach (T_{con}).

Now, what is an "issue"? We proposed many definitions which all could have been valied, but at the end we decided that we want to minimize the number of undiscovered incidents, leaving the detection of false positives and further sorting to another algorithm. For this reason, we define "issue" as an positive deviation from the setpoint temperature that isn't isolated, i.e., lasts for a prolonged period of time. In the next section we explain the approach we took for detecting issues, noting that as stated beforehand, it surely needs some polishing which like most of this project, we regard as future work.

5 Issue detection

The algorithm we propose to detect issues is relatively simple. Using the definition we gave above, we proceed to integrate the positive difference of room temperature and setpoint temperature, considering that a problem has arised if the result exceeds a certain threshold. Short:

$$T_{dif} = \max\left(T_{amb} - T_{con}, 0\right)$$

$$\text{if} \int_{t_i}^{t_{i+1}} T_{dif} \geq \delta_i \quad \text{then we have an issue}$$

We left the index i so that we could define several alarms throughout the day, each having its own threshold. Adjusting the values for a particular shop, we used shorter time windows and a higher threshold at the beggining of the workday (10 AM). This is a crutial moment since the air conditioning system wasn't running beforehand (theoretically) and the integral should yield a greater result. Further, we used a four hour sliding window throughout the rest of the day, together with a final alarm for the entire day (which could be argued to be unnecessary at the end). The following table shows these alarms together with the selected thresholds.

period	**tolerance/hour (℃)**
10h-11h	2.00
10h-12h	2.00
10h-13h	1.50
[t,t+4]h	1.50
10h-22h	1.25

Table 2: Alarms

To obtain the threshold which the integral shouldn't surpass you simply multiply the tolerance per hour by the length of the period.

Once an alarm goes off, the next step is for the issue sorter to analyze what can be the cause. However, let's first ilustrate what behaviour in room temperature could lead to this, using two examples in which we already applyed the algorithm.

5.1 Issue detection results

In the following figures we present the behaviour of a shop's temperatures throughout the day. The first one refers to a day when air conditioning functioned as intented, and therefore there weren't any alarms. The second one shows a problematic day, seing in the bottom which alarms would have gone off.

On the one hand, in Figure 4 we see that when the workday starts, the temperature with which the air conditioning system blows air into the shop falls to around 14 ℃, causing room temperature to be at around 1 ℃ over setpoint, which we consider to be acceptable due to how the system works. You can also notice how general consumption went up at around 6:30 in the morning, even

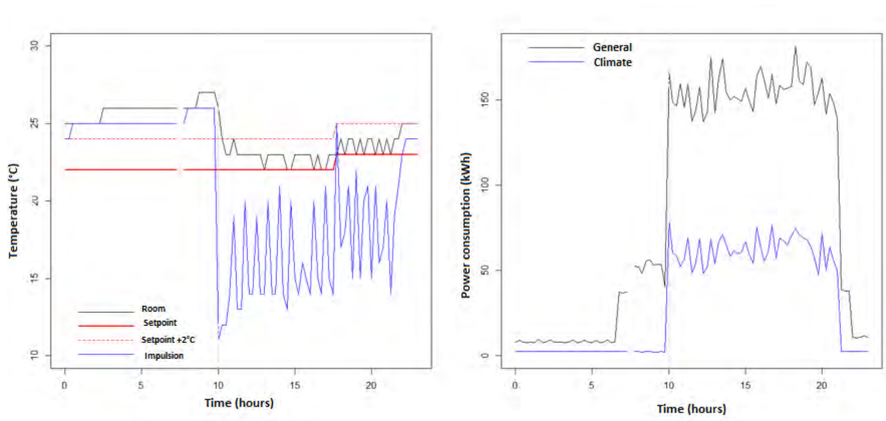


Figure 4: A normal day. Room temperature (Black) is compared to setpoint (red) and setpoint + 2 (red dotted). Also impulsion temperature is represented to see the system is working correctly.

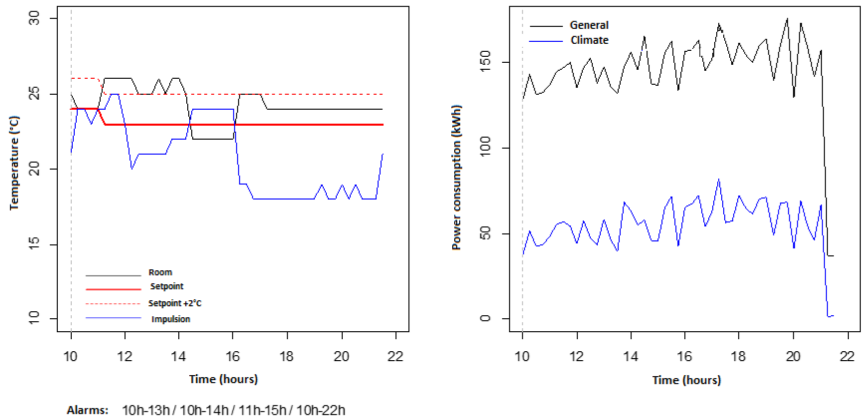


Figure 5: A day with issues. Room temperature (Black) is compared to setpoint (red) and setpoint + 2 (red dotted). Also impulsion temperature is represented to see the system is working correctly.

though climate consumption didn't, the reason for this being that lighting was turned on (probably to do some cleaning or loading/unloading). The energy of the light bulbs caused room temperature to rise before shop opening, which in turn accounts for more severe starting conditions towards air conditioning. Knowing this, we decided to introduce another possible variable for our issue sorter: room temperature at shop opening.

On the other hand, Figure 5 shows how issues in the air conditioning system were correctly detected as issues. In particular, in the first hours of the work day the system didn't blow air into the shop with an adequate temperature and thus room temperature was too high to be considered as acceptable. After these few hours the system started working as intendet, and past 3 PM the alarms stop going off as was to be expected (let aside the alarm for the entire day).

Having seen this, we decided to conduct a few different studies, one of them focusing on how the weekday affects the likelihood of the alarms to go off. In Figure 6 it is clear that on mondays and thursdays there were far less issues (value of the integral surpassing the threshold marked as the red line). This is due to the fact that these days, cleaning staff turned the air conditioning system on several hours before the shop opened, accounting for better initial conditions. From this figure you can also infer that on days the shop is closed (sundays) the alarm would go off most of the times but shouldn't be regardad. However, this will be analyzed by the problem sorter we introduce in the next section. We also see that there would be many positives in the first alarms of the day, but as the day goes on only atypical values (probably real issues) would be detected by our algorithm.

This fact motivated us to try and study the effect of room temperature at opening, which is shown in Figure 7.

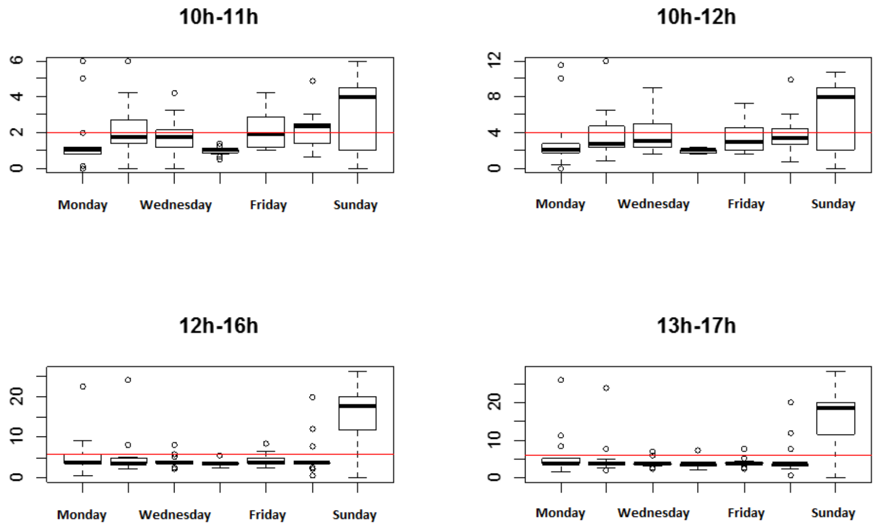


Figure 6: Result of the integral depending on the weekday

There is an evident growing trend in the result of the integral when compared to the room temperature at opening, specially noticable when above 28 °C. We assume that as this temperature rises even more, the air conditioning system wouldn't be able to reach the setpoint temperature even by a larger margin, and this cases should also be carefully analyzed by the issue sorter which we now present.

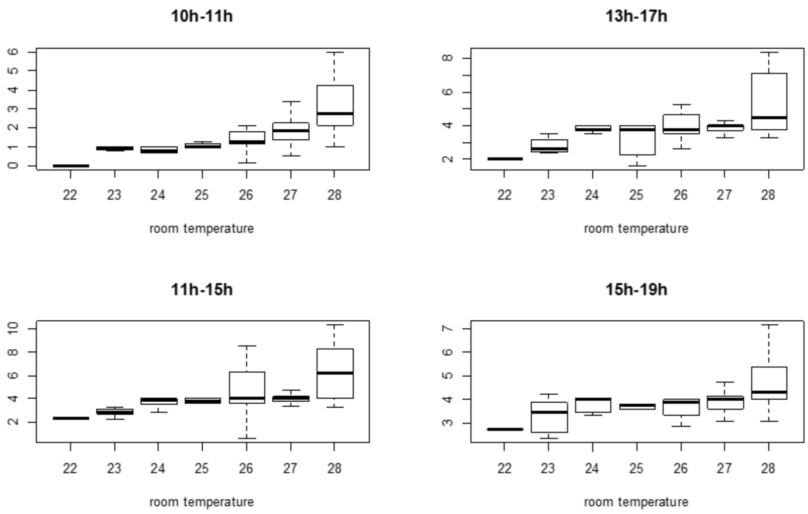


Figure 7: Result of the integral depending on room temperature at opening

6 Issue sorter

Throughout the previous sections we pointed out which variables we considered to be important for further analysis. Before we go ahead and present our proposed issue solver, let us highlight a few things:

Intuition told us that the external temperature should have a big impact in interior temperature conditions and thus in possible problems. However, the relation is not immediate due to buldings having high temperature inertia. Therefore we selected the room temperature at opening as an indicator of general temperature conditions in the recent past and compare it to a reference temperature which we consider adequate.

We would also like to point out that the variable we consider to be most important in identifying if the air conditioning system is working as intented is its impulsion temperature. Here we took into account that room temperature might seem too high but the air conditioning system is working perfectly, the problem being that the conditions are too severe to act against.

This being said, we present a flow diagram of how the issue sorter should behave once an alarm has gone off.

We have not implemented the corresponding algorithm, so further polishing and additions are left for future work once again.

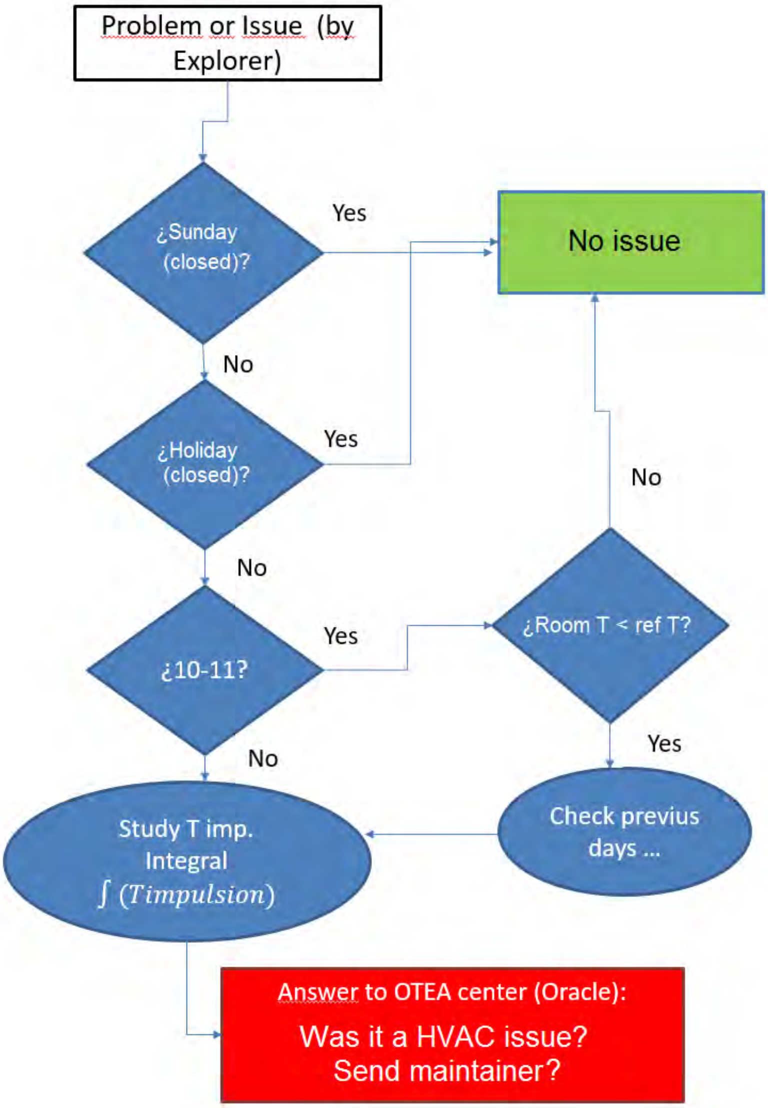


Figure 8: Flow chart for the issue sorter

7 Conclusions and future work

- First and most importantly, we presented an issue detection algorithm based on integrating temperature deviations, which seems to work as intendet for the selected shop. In order to make this algorithm valid for every shop more variables have to be introduced to this algorithm (opening hours,particular comfort conditions) and the thresholds should be obtained by a method which takes the particular conditions of the shop into consideration (variable means to find which temperature deviation can be considered as normal). We also believe that the selected time intervals are not optimal and could be adapted depending on the computational power the system has available.

- The proposed issue sorter is still subject to change, and each of the blocks in the presented diagram still has to be developed, having many thing left to discuss, e.g., how can we know which day is a holiday for each particular shop? It is also not clear if an integral method based on impulsion temperature alone would be optimal to determine if the HVAC is funtioning correctly, but we consider this to be an acceptable starting point.

- The last step in the OTEA Oracle, the answers and rules "Maker", wasn't even considered for the time being, since the first two steps have to be working perfectly for this last system to be productive. Once the previous algorithms are succesfully developed, we believe that even more variables should be taken into consideration, to optimize the time spend by maintainers having to fix issues personally.

- Finally, we would like to state that other mathematical approaches were taken into consideration, one being functional analysis, but due to the lack of time this was disregarded for this study. However, it could be further analyzed if the project would go on. Also, stochastic methods should be considered to try and make use of nondeterministic variables such as shop occupation, or to predict missing variables.

Bibliography

- Breiman, L; Friedman, J.H., Olshen, R.A.; Stone, C.J. (1984). Classification and regression trees. Monterey, CA: Wadsworth and Brooks/Cole Advanced Books and Software.

- Hastie, T., Tibshirani, R., and Friedman, J.H. (2003). The Elements of Statistical Learning. Springer.

- R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

- Russell, S.J.; Norvig, P (2004). Artificial Intelligence: A Modern Approach. Pearson.

- Siegel, R. Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, Or Die. Wiley.

Acknowledgements

The Scientific Committee wishes to thank to the company speakers, the academic coordinators and the researchers of each working team for their invaluable contributions to the scientific success of the 122 European Study Group with Industry.

This 122 ESGI was framed as activity in ITMATI. It has been supported by the Galician Agency for Innovation and the Ministry of Economy and Finance in the framework of the Spanish Strategy for Innovation at Galicia.

Finally, we would also like to express our gratitude to Adriana Castro Novo, technology transfer and innovation manager at ITMATI, Rubén Gayoso Taboada, manager of ITMATI, Andrés Prieto, Associate Professor in the University of A Coruña and Affiliated Researcher to ITMATI, Peregrina Quintela Estévez, director of ITMATI and president of math-in and Fe Sampayo Fernández, technology translator of math-in, whose meticulous work and dedication have contributed to the success of this 122 European Study Group with Industry.

INDEX

Universidade Vigo

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

UNIVERSIDADE DA CORUÑA