Table of Contents

Volume 3, Issue 3, May/ June 2009.

Pages

Smart Sim Selector: A Software for Simulation Software Selection

Ashu Gupta   guptashu1@rediffmail.com
Senior Lecturer
Department of Computer Applications
Apeejay Institute of Management
Rama Mandi-Hoshiarpur Road
Jalandhar-144007 (India)
Mob: 98158-91562

Dr. Rajesh Verma   dr.rajeshverma1@rediffmail.com
Assistant Professor
Lovely School of Business
Lovely Professional University
Phagwara (Punjab) India

Dr. Kawaljeet Singh   singhkawaljeet@rediffmail.com
Director, University Computer Centre
Punjabi University
Patiala, Punjab (India)

Abstract

In a period of continuous change in global business environment, organizations, large and small, are finding it increasingly difficult to deal with, and adjust to the demands for such change. Simulation is a powerful tool for allowing designers imagine new systems and enabling them to both quantify and observe behavior. Currently the market offers a variety of simulation software packages. Some are less expensive than others. Some are generic and can be used in a wide variety of application areas while others are more specific. Some have powerful features for modeling while others provide only basic features. Modeling approaches and strategies are different for different packages. Companies are seeking advice about the desirable features of software for manufacturing simulation, depending on the purpose of its use. Because of this, the importance of an adequate approach to simulation software selection is apparent.

Smart Sim Selector is a software developed for the purpose of providing support for users when selecting simulation software. Smart Sim Selector consists of a database which is linked to an interface developed using Visual Basic 6.0. The system queries a database and finds a simulation package suitable to the user, based on requirements which have been specified. This paper provides an insight into the development of Smart Sim Selector, in addition to the reasoning behind the system.

Keywords: Simulation, Simulation software, Selection, Rating

1. INTRODUCTION

Growing competition in many industries has resulted in a greater emphasis on developing and using automated manufacturing systems to improve productivity and to reduce costs. Due to the complexity and dynamic behavior of such systems, simulation modeling is becoming one of the most popular methods of facilitating their design and assessing operating strategies. Increasing popularity of simulation techniques has resulted in an increase in the number of simulation packages available in the market. Currently the market offers a variety of discrete-event simulation software packages. For large international companies with their own simulation team it is often hard to select new discrete event simulation software. In simulation software selection problems, packages are evaluated either on their own merits or in comparison with other packages. The wrong selection of an inappropriate package can result in significant financial losses as well as the disruption of simulation projects. However, appropriate assistance in simulation software selection might reduce the scale of such problems. Many studies have taken place concerning the selection of simulation software.

2. RESEARCH IN SIMULATION SOFTWARE SELECTION

The starting point for the research was to review previous studies on the selection of simulation software tools. There are many studies on simulation software selection, some of which are as described below:

Banks and Gibson (1997) suggested some considerations to be made while purchasing the simulation software like accuracy and detail, powerful capabilities, fastest speed, demo-solution of problem, opinions of companies with similar applications, attending the user group meetings.

Hlupic and Paul (1999) presented criteria for the evaluation of simulation packages in the manufacturing domain together with their levels of importance for the particular purpose of use. They suggested general guidelines for software selection. They pointed that not to expect a particular package to satisfy all criteria. However, it is indicated which criteria are more important than others, according to the purpose of software use. These guidelines can be used both by users that are looking for a suitable simulator to buy, and by developers of such simulators who wish to improve existing versions of simulators, or who wish to try to develop a new, better manufacturing simulator.

Nikoukaran *et al.* (1999) created a framework of criteria to be considered when evaluating discrete-event simulation software. This framework is structured, and pays attention to a rich set of criteria on which simulation packages can be compared. It is, however, difficult to base a decision for a large multinational company on these criteria, as it is only a comparison, without weighing and without a method to determine the relative weights, and the weight differences between parts of the simulation team.

Tewoldeberhan *et al.* (2002) proposed a two-phase evaluation and selection methodology for simulation software selection. Phase one quickly reduces the long-list to a short-list of packages. Phase two matches the requirements of the company with the features of the simulation package in detail. Different methods are used for a detailed evaluation of each package. Simulation software vendors participate in both phases.

Seila *et al.* (2003) presented a framework for choosing simulation software for discrete event simulation. By evaluating about 20 software tools, the proposed framework first tries to identify the project objective, since a common understanding of the objective will help frame discussions with internal company resources a well as vendors and service providers. It is also prudent to define long-term expectations. Other important questions deal with model dissemination across the organization for others to use, model builders and model users, type of process (assembly lines, counter operations, material handling) the models will be focused, range of systems represented by the models etc.

Popovic *et al.* (2005) developed criteria that can help experts in a flexible selection of business process management tools. They classified the simulation tools selection criteria in seven categories: model development, simulation, animation, integration with other tools,

analysis of results, optimization, and testing and efficiency. The importance of individual criteria (its weight) is influenced by the goal of simulation project and its members (i.e., simulation model developers and model users).

Many other studies have been carried out involving the selection of simulation software such as **Bovone *et a1.* (1989)** and **Holder (1990)**. All of these studies are geared towards a more theoretical approach of the problems involving the selection of simulation software. Smart Sim Selector attempts to put these issues into practice and aid the user in various difficulties that may arise when selecting a suitable simulation package.

This paper is structured as follows: Subsequent to a description of the criteria used for software selection, and justifications for choosing criteria, a description of the database and the way in which the database handles a query are provided. Later sections provide an insight into the software itself. Issues related to an overview of the software and the way in which a query is processed by Smart Sim Selector is addressed. The methodology which Smart Sim Selector uses to recommend most suitable packages is also discussed. Finally, conclusions outline the main findings of this research; highlight the benefits that can be derived from Smart Sim Selector as well as its future developments.

3. SELECTION CRITERIA

The type of criteria required to evaluate simulation software was an important issue that was addressed throughout the development of Smart Sim Selector. Research within this topic was undertaken and several important studies were focused upon **(Tocher, 1965), (Cellier, 1983), (Banks *et al.* 1991),** and **(Law and Kelton, 1991).** However, majority of these studies were theoretical and unsuitable as not many of them provided a critical evaluation of the software products under consideration. In addition, it was difficult to compare the evaluation of different studies because they did not use common criteria.

The criteria used for the Smart Sim Selector was taken from **Verma *et al.* (2009)** who proposed more than 210 evaluation criteria divided into 14 main groups: General Information, Coding Aspects, Compatibility, User Support, Financial & Technical Features, General Features, Modeling Assistance, Visual Aspects, Efficiency, Testability, Experimentation, Statistical facilities, Input/Output Capabilities, Analysis Capabilities. The main reason for incorporating data from such a wide range of categories is to ensure that Smart Sim Selector will be able to produce accurate results.

4. DATABASE DESIGN

The database is considered to be the "engine-room" of Smart Sim Selector. It holds the information related to evaluation details of each package. In addition, the queries concerning the suitability of packages are also generated within the database. The initial requirements of the database were largely concerned with not only the ability to store and access data, but also to ensure that the database could be easily modified and maintained. This is an issue of great importance because in the near future the number of packages which are evaluated will be increased. The data has been collected from Automobile manufacturers in North India which includes 18 automobile companies.

The database is designed in a conventional manner, with information being stored about each package. A particular data table shown in Figure 1 contains information regarding the simulation software package such as type of package, purpose of simulation, and price. Another table designed, holds information regarding the name of the package and a unique index which holds the pack-id number. A query is generated after the user has finalized the requirements and priority rating.

Software_ID	1	2	3	4	5	6	7	8	9
1	4	5	5	4	3	3	3	4	4
10	5	5	5	5	5	5	5	5	5
11	5	4	3	3	3	3	4	4	5
12	6	5	5	5	6	6	6	6	6
13	5	5	5	6	6	6	6	6	6
14	4	3	3	3	2	2	4	4	4
2	5	4	4	5	5	3	4	3	3
3	4	4	4	4	4	4	5	6	6
6	6	4	4	4	3	2	3	4	4
7	4	4	5	4	3	3	3	2	2
8	5	4	3	3	4	5	3	3	4

Figure 1: One of the Tables within the Smart Sim Selector Database

5. OVERVIEW OF SMART SIM SELECTOR

Smart Sim Selector was designed using Visual Basic 6.0, because it allowed for the creation of effective and user-friendly interface. Visual Basic 6.0 has a built in Access Engine which is easily capable of handling all database queries. This was ideal for Smart Sim Selector which required a relational database to store information regarding various packages. At the moment, Smart Sim Selector stores information about 11 packages, based on 210 different criteria. There is no bias towards a specific supplier, which is one reason why Smart Sim Selector may be seen as an effective tool in the simulation software selection process. Figure 2 shows the starting screen of Smart Sim Selector.

Figure 2: Staring Screen of Smart Sim Selector

Ashu Gupta, Rajesh Verma & Kawaljeet Singh
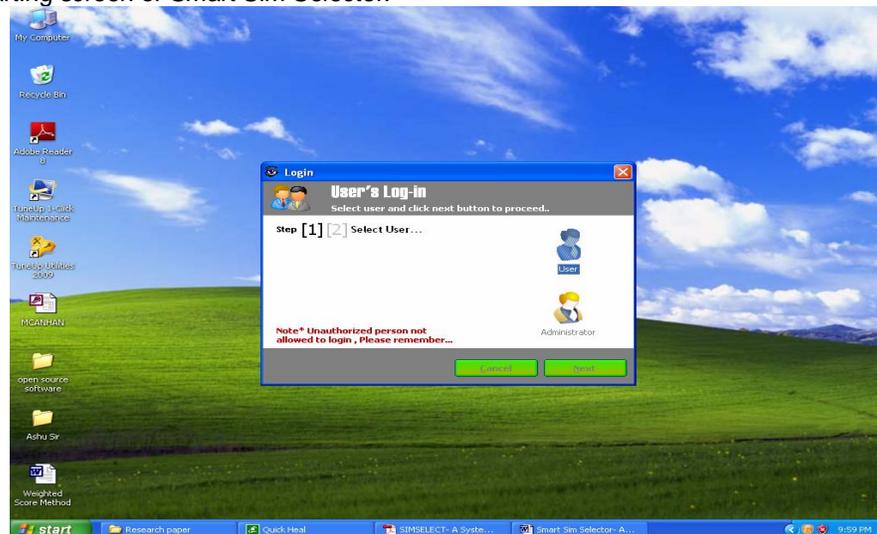
5.1 Generating a Query
In order to generate a query using Smart Sim Selector, the user is required to input data regarding the requirements of the desired system. The main menu of the system is shown in Figure 3. The options provided by the menu are Administrator Work, Suggest Software, Comparison of Softwares, and Exit.
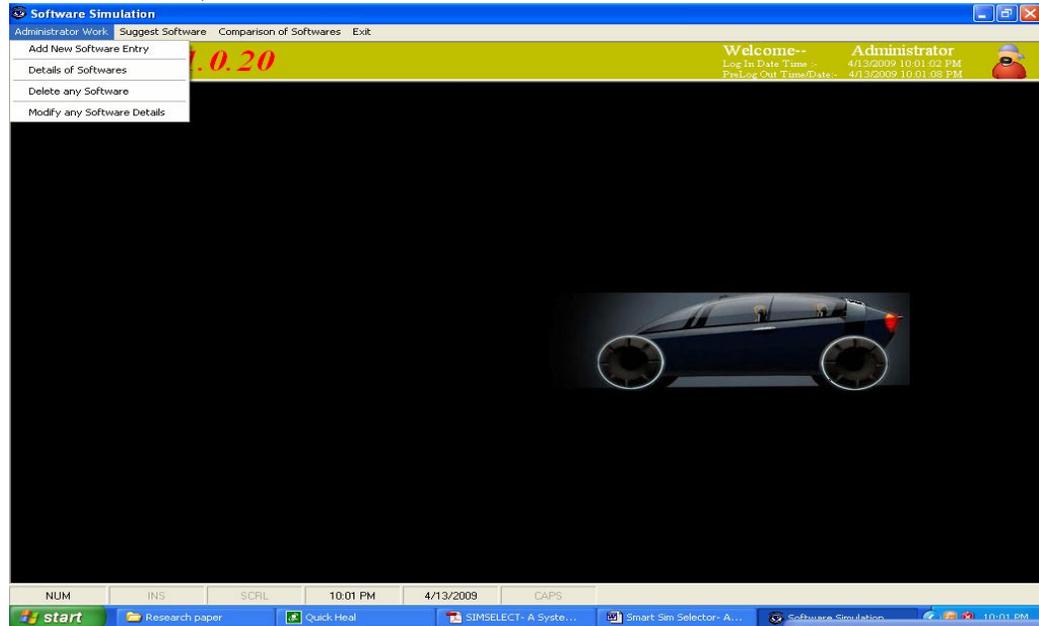
Figure 3: Smart Sim Selector Menu System

Administrator can add, delete and modify any software in the database. Also the option 'Details of the Softwares' will display the details of all the softwares contained in the database to the user and administrator. Once the administrator selects to add new software, the screen as shown in Figure 4 will be displayed.
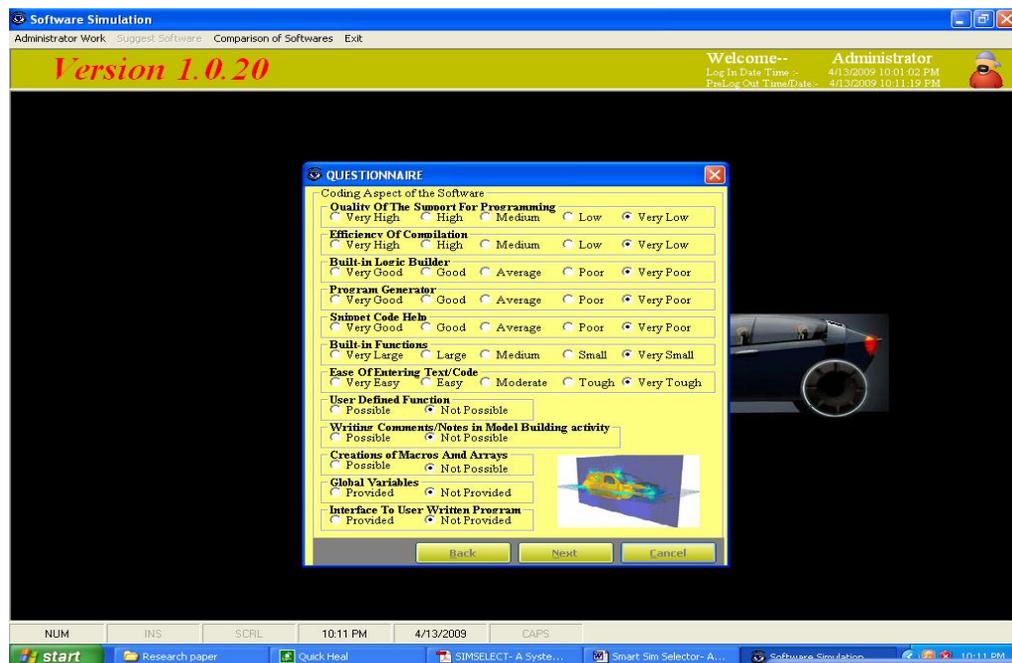
Figure 4: Screen for adding new software in Smart Sim Selector database

5.2 COMPARISON OF SOFTWARES

In order to compare the simulation softwares, a rating of these has been established. This rating is based on an analysis of the simulation softwares. As such it should be considered as a relative measure of quality of these softwares from the perspective of groups of criteria rather than as an absolute value.

There are total 14 groups of features i.e. general information**,** coding aspects, software compatibility, user support, general features, modeling assistance, visual aspects, efficiency, testability, financial and technical features, experimentation facilities, statistical facilities, input and output capabilities, analysis capabilities. The value (out of 10) of these groups of features is calculated for all the simulation softwares in the database. Figure 5 shows the comparison screen.

$$\text{Evaluated Value} = \frac{\text{Calculated Value} \times 10}{\text{Maximum Value}}$$

Where Maximum Value= Sum of highest possible values that can be selected in a particular group of features

and Calculated Value= Sum of actual values selected in a particular group of features

Features / Software Name	MOLDEX 3D	NX-IDEAS	NASTRAN	ProModel
Coding Aspects	7.234 (High)	7.66 (High)	6.809 (High)	9.149 (Very High)
Software Compatibility	3.571 (Poor)	2.857 (Poor)	3.571 (Poor)	5 (Medium)
User Support	7.5 (High)	7.647 (High)	6.176 (High)	6.176 (High)
General Features	7.794 (High)	7.5 (High)	6.176 (High)	6.765 (High)
Modeling Assistance	8.409 (Very High)	8.409 (Very High)	5.227 (Medium)	9.545 (Very High)
Visual Aspect	8.095 (Very High)	8.69 (Very High)	5.833 (Medium)	8.81 (Very High)
Efficiency	8.226 (Very High)	7.742 (High)	7.581 (High)	7.742 (High)
Testability	7.857 (High)	8.571 (Very High)	6.429 (High)	6.714 (High)
Experimentation Facilities	8.261 (Very High)	8.696 (Very High)	7.391 (High)	5.652 (Medium)
Statistical Facilities	7.6 (High)	7.6 (High)	5.2 (Medium)	6.8 (High)
Input/Output Capabilities	7.463 (High)	7.91 (High)	6.716 (High)	6.716 (High)
Analysis Capabilities	10 (Very High)	10 (Very High)	6.667 (High)	10 (Very High)
Financial & Technicle Features	6.383 (High)	7.234 (High)	5.426 (Medium)	7.447 (High)

Figure 5: Screen showing Comparison of the Softwares

6. SUGGESTING THE BEST SUITABLE SIMULATION SOFTWARE

Smart Sim Selector has three different techniques for selecting the best suitable software for the user as shown in Figure 6:

6.1 Analytic Hierarchy Process

6.2 Weighted Score Method

6.3 TOPSIS (Technique of Order Preference by Similarity to Ideal Solution)
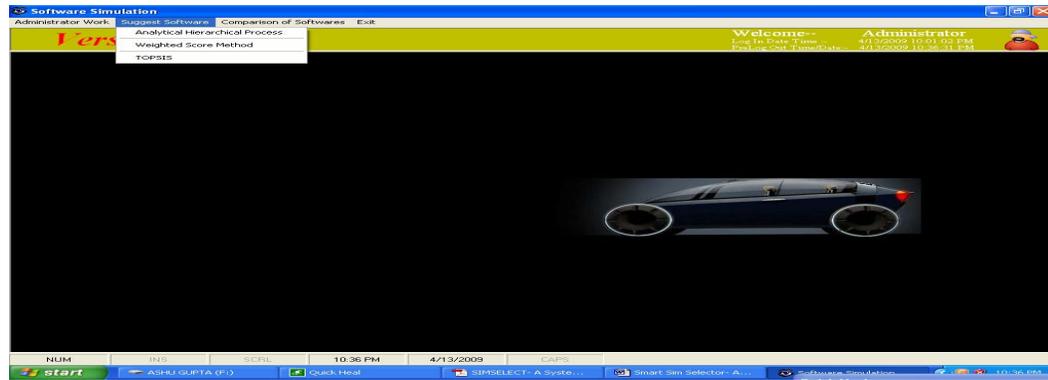
Ashu Gupta, Rajesh Verma & Kawaljeet Singh

Figure 6: Different Techniques for Software Selection

6.1 The Analytic Hierarchy Process and Simulation Software Selection

The AHP separates the evaluation decision into hierarchy levels and attempts to reduce the inconsistencies in human judgment. It was originally used for socio-economic and political situations but of late it has proved useful for judgmental decision making in other areas, such as the selection of equipment for ice breakers (**Hannan et al., 1983**), the selection of materials handling equipment (**Frazelle, 1985**) and perhaps more relevant, the selection of manufacturing software (**Williams, 1987**) and scheduling software (**Williams and Trauth, 1991**). Further applications, along with a good exposure of AHP, are given by **Partovi et al. (1990)** and **Zahedi (1986)**.

In using the AHP technique all the criteria are compared in a pairwise way, using Saaty's intensities of importance (**Saaty and Roger, 1976**) shown in Table 1, in order to establish which criteria are more important than others. The values are then placed in a matrix and the normalized principal eigenvector is found to provide the weighting factors which provide a measure of relative importance for the decision maker. The next step is to make pairwise comparisons of all alternative with respect to each criterion. Final rankings of the alternatives are made by multiplying the critical weights of the alternatives by the critical weights of the criteria. The alternative with the highest score is then deemed to be the preferred choice. Figure 7 shows the screen for AHP.

Table 1: Saaty's Intensities of Importance

| Intensity of Importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Weak importance of one over another	The judgement is to favor one activity over another, but it is not conclusive
5	Essential or strong importance	The judgement is to strongly favor one activity over another
7	Demonstrated importance	Conclusive judgement as to the importance of one activity over another
9	Absolute importance	The judgement in favor of one activity over another is of the highest possible order of affirmation
Reciprocals of above non-zero numbers	If activity i has one of the above non-zero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with i	

Ashu Gupta, Rajesh Verma & Kawaljeet Singh
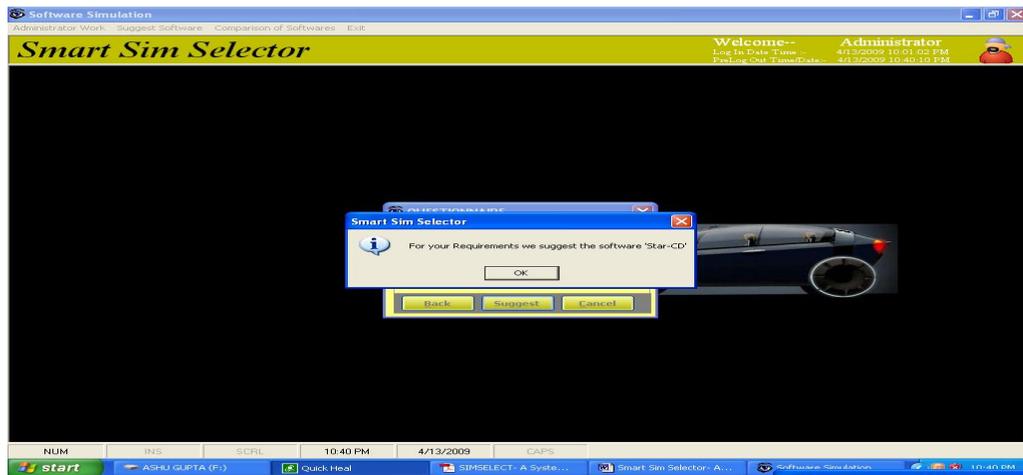
Figure 7: Best Software Suggested

6.2 Simulation Software Selection using Weighted Score Method

The steps used in Weighted Score method are:

1. **Determine the criteria for the problem:** We have 14 different groups of features to satisfy this step
2. **Determine the weight for each criterion:** This value will be taken by the user when he will select the criteria and their importance. Figure 8 shows this screen.
3. **Obtain the score of alternative i using each criteria j for all i and j as follows:**
   Let S_{ij} be the score of software i using group of criteria j
   W_j be the weight for group of criteria j
   S_i score of alternative i is given as:

   $$S_i = \sum_j W_j S_{ij}$$

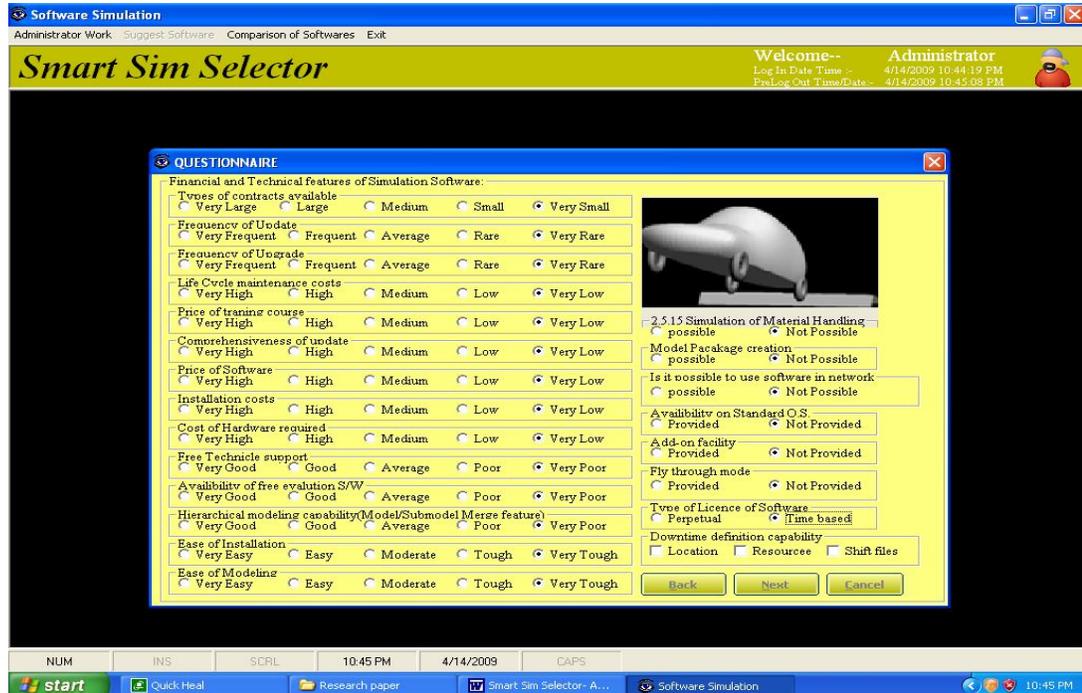   **The simulation software with the best score is selected.**

Ashu Gupta, Rajesh Verma & Kawaljeet Singh

Figure 8: Weightage to each criteria given by the user

6.4 TOPSIS (Technique of Order Preference by Similarity to Ideal Solution)

In this method two artificial alternatives are hypothesized:
1. Ideal alternative: the one which has the best level for all attributes considered.
2. Negative ideal alternative: the one which has the worst attribute values.
3. TOPSIS selects the alternative that is the closest to the ideal solution and farthest from negative ideal alternative.

TOPSIS assumes that we have m alternatives and n criteria and we have the score of each alternative with respect to each criterion.
1. Let xij score of alternative i with respect to criterion j
    We have a matrix X = (xij)   m×n matrix.
2. Let J be the set of benefit criteria (more is better)
3. Let J' be the set of negative criteria (less is better)

Step 1: Construct normalized decision matrix: This step transforms various attribute dimensions into non-dimensional attributes, which allows comparisons across criteria.

Normalize scores or data as follows:

$r_{ij} = x_{ij}/ (\Sigma x^2_{ij})$ for i = 1, …, m; j = 1, …, n

Step 2: Construct the weighted normalized decision matrix:
➢ Assume we have a set of weights for each criteria w_j for j = 1,…n.
➢ Multiply each column of the normalized decision matrix by its associated weight.
➢ An element of the new matrix is: $v_{ij} = w_j r_{ij}$

Step 3: Determine the ideal and negative ideal solutions:
➢ **Ideal solution:**
    $A^* = \{ v_1^* , …, v_n^*\}$, where
    $v_j^* =\{ max (v_{ij}) \text{ if } j \in J ; min (v_{ij}) \text{ if } j \in J' \}$
➢ **Negative ideal solution:**
    $A' = \{ v_1' , …, v_n' \}$, where $v' = \{ min (v_{ij}) \text{ if } j \in J ; max (v_{ij}) \text{ if } j \in J' \}$

Step 4: Calculate the separation measures for each alternative.

> ➢ **The separation from the ideal alternative is:**
> $$S_i^* = [\Sigma (v_j^* - v_{ij})^2]^{\frac{1}{2}} \qquad i = 1, ..., m$$
> ➢ **Similarly, the separation from the negative ideal alternative is:**
> $$S_i' = [\Sigma (v_j' - v_{ij})^2]^{\frac{1}{2}} \qquad i = 1, ..., m$$

Step 5: Calculate the relative closeness (C_i^*) to the ideal solution
$$C_i^* = S_i' / (S_i^* + S_i') , \qquad 0 < C_i^* < 1$$
Select the option with Ci* closest to 1

7. SUMMARY AND CONCLUSIONS

Smart Sim Selector is a tool which can assist the user in the simulation software selection process. Although the system may seem limited in terms of the number of packages currently evaluated and stored within the database, it still represents an explicit attempt to combat some of the problems involved in simulation software selection.

Smart Sim Selector offers an alternative view of selecting simulation software as it is unbiased and unrelated to any software vendor or supplier. Each simulation package has been evaluated and stored in the database against criteria which covers a variety of issues. The criteria do not favor a single package, and the database will be increased to cover more criteria and simulation packages in the near future.

In addition to further expansion of the database, Smart Sim Selector will be distributed to various educational and commercial institutions involved in simulation, in order to get more feedback on possible improvements. Smart Sim Selector also requires enhancements to be implemented which will allow for the system to actually distinguish between the various levels of criteria that a simulation package can offer. One way to do this would involve creating another data table which takes into account such concepts. For example, after specifying that a statistical distribution fitting mechanism would be required, the user would be able to distinguish how many statistical distributions would be sufficient. The system would then take this into account rather than just offer a package which offered a statistical distribution fitting mechanism but not knowing how many distributions were provided.

Similarly an enhancement is required to allow Smart Sim Selector to take into account user comments at various stages so that the user actually inputs information rather than to select from a number of choices offered by the system. The user should be able to input comments regarding the type of animation that is required, the level of coding.
Despite the limitations, Smart Sim Selector represents a step forward to more explicit assistance in the simulation software selection process.

References
Banks, J., Aviles, E., McLaughlin, J. R., & Yuan, R. C. (1991). The Simulator: New Member of the Simulation Family. *Interfaces, 21*(2): 76-86.
Banks, J., & Gibson, R. R. (1997). Selecting Simulation Software. *IIE Solutions, 29*(5): 30-32.
Bovone, M., Ferrari, D. V. and Manuelli, R. (1989). How to Choose an Useful Simulation Software. In D. M. Smith, J. Stephenson, & R. N. Zobel (Eds.), *Proceedings of the 1989*

European Simulation Multiconference (pp. 39-43). SCS, San Diego: Society for Computer simulation International.

Cellier, F. E. (1983). Simulation Software: Today and Tomorrow. In J. Burger, & Y. Jarny (Eds.), *Proceedings of the IMACS International Symposium (pp. 3-19),* Amsterdam: Elsevier Science Publishers.

Hlupic, V., & Paul, R. J. (1999). Guidelines for Selection of Manufacturing Simulation Software. *IIE Transactions, 31*(1): 21-29.

Holder, K. (1990). Selecting Simulation Software. *OR Insight, 3*(4): 19-24.

Law, A. M., & Kelton, W. D. (1991). Simulation Modeling and Analysis. Singapore: McGraw-Hill.

Nikoukaran, J., Hlupic, V., & Paul, R. J. (1999). A Hierarchical Framework for Evaluating Simulation Software. *Simulation Practice and Theory, 7*(3): 219-231.

Popovic, A., Jaklic, J., & Vuksic, V. B. (2005). Business Process Change and Simulation Modeling. *System Integration Journal, 13*(2): 29-37.

Seila, A. F., Ceric, V., & Tadikamalla, P. (2003). Applied Simulation Modeling. Australia: Thomson Learning.

Tewoldeberhan, T. W., Verbraeck, A., Valentin E., & Bardonnet, G. (2002). An Evaluation and Selection Methodology for Discrete-Event Simulation Software. In E. Ycesan, J. L. Snowdon, J. M. Charnes, & J. Wayne (Eds.), *Proceedings of the 2002 Winter Simulation Conference* (pp. 67-75). Boston: Kluwer Academic Publisher.

Tocher, K. D. (1965). Review of Simulation Languages. *Operational Research Quarterly, 16*(2): 189-217.

Verma, R., Gupta, A., & Singh, K. (2009). A Critical Evaluation and Comparison of Four Manufacturing Simulation Softwares. *International Journal of Science, Engineering and Technology, 5*(1): 104-120.

Implementing a Functional ISO 9001 Quality Management System in Small and Medium-Sized Enterprises

Andres Sousa-Poza   asousapo@odu.edu
Department of Engineering Management
and Systems Engineering
Old Dominion University
Norfolk, 23529, USA

Mert Altinkilinc   malti001@odu.edu
Department of Engineering Management
and Systems Engineering
Old Dominion University
Norfolk, 23529, USA

Cory Searcy   cory.searcy@ryerson.ca
Department of Mechanical and Industrial
Engineering
Ryerson University
Toronto, M5B 2K3, Canada

Abstract

This conceptual paper provides guidance for the implementation of a functional ISO 9001 quality management system (QMS) in small and medium-sized enterprises (SMEs). To help a SME understand its starting point, four initial states for QMS implementation are defined. Five paths for moving the QMS from the initial state to the desired state are described. To support the transition from the initial to the desired state, some key considerations in implementing a QMS in SMEs are discussed. The paper is based on site visits and implementation assistance the authors have provided to several SMEs. It is anticipated the paper will help managers in SMEs understand the process of implementing ISO 9001 and help them avoid the development of a paper-driven QMS that provides limited value.

Key words: Small and medium sized enterprises, quality management, ISO 9001

1. INTRODUCTION

Small and medium-sized enterprises (SMEs) play a critical role in economies around the world. To remain competitive, SMEs must be capable of delivering high quality products and services on-time at a reasonable cost. In response to these competitive pressures and customer demand, many SMEs have developed ISO 9001 quality management systems (QMS).

ISO 9001 is an international standard that specifies the basic requirements for a QMS. The two primary objectives of the standard are to help an organization demonstrate its ability to meet customer and regulatory requirements and to enhance customer satisfaction. To that end, the standard contains key requirements clauses focusing on (1) the QMS in general, (2) management responsibility, (3) resource management, (4) product realization, and (5) measurement, analysis, and improvement. Originally released in 1987, the standard was updated in 1994, 2000, and

2008. As of December 31, 2007, at least 951,486 certificates had been issued worldwide certifying verified compliance with ISO 9001 [1]. ISO 9001 is supported by a broader family of standards. These include ISO 9000 (QMS fundamentals and vocabulary), ISO 9004 (QMS guidelines for performance improvements), and ISO 19011 (guidelines for quality and/or environmental management systems auditing).

There is general agreement in the literature that SMEs possess unique characteristics that will influence the implementation of a QMS. In one widely-cited representative study, Ghobadian and Gallear [2] compared the common characteristics of SMEs with large organizations on the basis of organizational structure, procedures, behavior, processes, people, and contact. For example, relative to large organizations SMEs tend to have fewer layers of management, informal rules and procedures, a fluid culture, a simple planning and control system, modest human and financial resources, and a limited customer base [2]. While there is no universal definition of a SME, a widely used classification is that they consist of fewer than 500 employees.

Recognizing that there are differences between SMEs and large organizations, several papers have been published focusing on the application of ISO 9001 and broader total quality management (TQM) principles to SMEs. The literature reveals that there are contradictory results on whether or not implementing ISO 9001 helps improve the performance of SMEs. For example, in a study of Australian SMEs, Rahman [3] found that there is no significant difference between SMEs with and without ISO 9001 certification with respect to TQM implementation and organizational performance. However, in a study of Iranian SMEs, Bayati and Taghavi [4] concluded that acquiring ISO 9001 certification improved the performance of the companies studied. Similarly, Koc [5] found that SMEs gain significant benefits by implementing ISO 9000 standards. For further information on ISO 9001 and TQM implementation in SMEs, the interested reader is referred to Lewis *et al.* [6], Poksinska *et al.* [7], and Pinho [8].

There are a number of challenges that SMEs face in implementing an effective ISO 9001 QMS. Building on the discussion in Ghobadian and Gallear [2], some of the representative challenges SMEs may face include a lack of financial and human resources, inadequate technical knowledge of quality management, a lack of knowledge of formalized systems, and a lack of experience in internal auditing. Although the heterogeneous nature of SMEs makes it difficult to generalize, managers and employees in SMEs are typically well focused on core activities such as production, sales, and customer service. However, they are generally not knowledgeable about QMS requirements or on how to improve processes. As Temtime [9] explains, "…the majority of SMEs understand the importance of both TQM and planning activities for their survival and growth. However, their approach to both is generally unsystematic, non-formal and short-term oriented". An unsystematic approach is likely to exacerbate the problems faced by SMEs when attempting to introduce an ISO 9001 QMS.

The development of a systematic approach to ISO 9001 implementation should help managers in SMEs increase the possibility of a successful implementation. In particular, guidance is required on recognizing the initial state, recognizing the desired state, and defining a path to navigate the transition. This paper presents a conceptual model with those needs in mind. The paper is organized into five sections. Immediately following this introduction, four initial states of a QMS in SMEs are introduced. This provides the basis for the third section focusing on how the QMS can be moved from the initial state to the desired state. In the fourth section some of the key considerations in implementing ISO 9001 in SMEs are highlighted. The paper finishes with a brief summary and conclusions.

2. ESTABLISHING THE INITIAL STATE OF THE QMS
The implementation of an ISO 9001 conformant system must recognize that it is but a step in a long-term development of a continually improving QMS. This is depicted in Figure 1. Unfortunately, it is often the case that ISO 9001 is taken as a means to an end, where the implementation of a QMS is not the primary objective, rather certification is. As a result, SMEs may end up with stacks of documentation waiting to be processed that adds no value, but cost.

According to the requirements of ISO 9001, an organization must develop only six documented procedures: (1) control of documents, (2) control of quality records, (3) internal audits, (4) control of non-conformities, (5) corrective action, and (6) preventative action. A quality manual and several records are also required. The development of other procedures, work instructions, and other documents is largely at the discretion of the organization. From the very beginning of the process, it is therefore essential that SMEs establish a balanced view between a short-term focus (marketing/sales) and a long-term focus (achieving company-wide quality awareness through TQM). ISO documentation should be considered as an enabler along that way and SMEs must guard against the creation of unnecessary documentation.
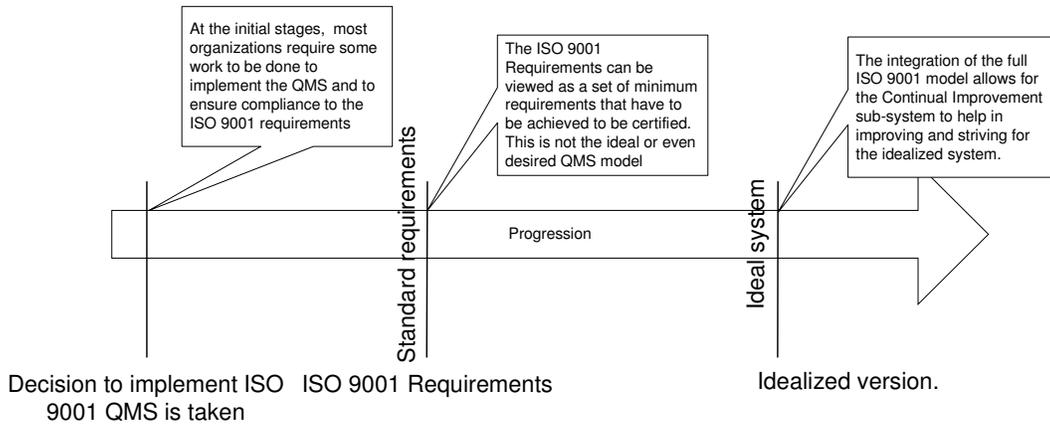
At the initial stages, most organizations require some work to be done to implement the QMS and to ensure compliance to the ISO 9001 requirements

The ISO 9001 Requirements can be viewed as a set of minimum requirements that have to be achieved to be certified. This is not the ideal or even desired QMS model

The integration of the full ISO 9001 model allows for the Continual Improvement sub-system to help in improving and striving for the idealized system.

Standard requirements

Ideal system

Progression

Decision to implement ISO   ISO 9001 Requirements
9001 QMS is taken

Idealized version.

FIGURE 1: The QMS Implementation Continuum

However, even when such a view is adopted, many SMEs struggle to move from their initial state to a fully functional ISO 9001 QMS. Over the last several years, we have been involved in ISO 9001 implementation projects in seven different SMEs. The SMEs have ranged in size from approximately 20 employees to 500 employees. The SMEs have been drawn from a variety of sectors in Virginia, including manufacturing, distribution, and services. Based on our experience, we developed a schematic of initial states of an organization in terms of the *existence* and *functionality* of the QMS as shown in Figure 2. Throughout this paper, existence is equated with the documentation required by the standard while functionality is equated with an effectively operated QMS that leads to increased customer satisfaction and continuous improvement of business results. A successful QMS must be fully functional and appropriately documented. With that in mind, there are four main states in which SMEs can be located in the beginning of the implementation process:
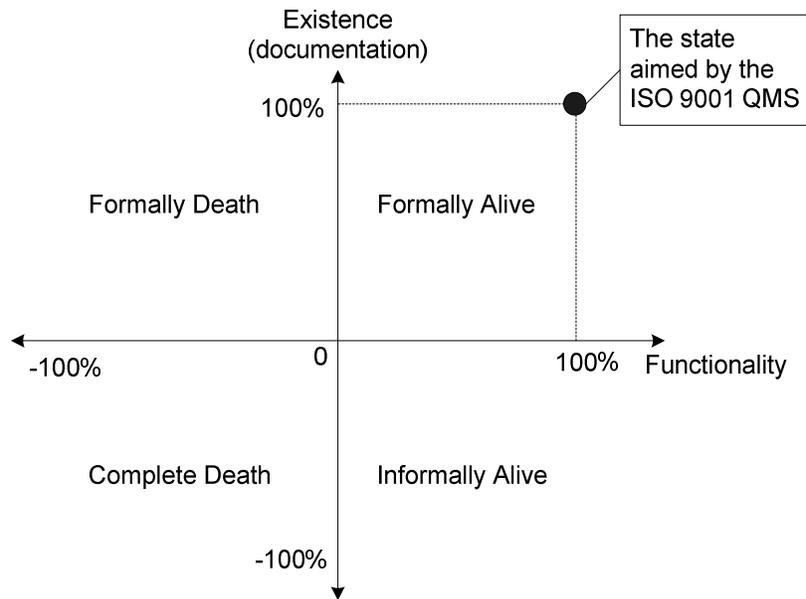
FIGURE 2: Schematic of Initial States in terms of Existence and Functionality of ISO 9001 QMS

1. *Complete Death: No documentation, no functioning.*
This is the state in which there is no indication of the existence and functionality of the QMS. No documentation exists and no processes are in place to help ensure the quality of the product. Relatively few companies will find themselves in this situation.

2. *Informally Alive: No documentation, some level of functioning.*
Many SMEs exhibit an organic structure characterized by an absence of standardization and the prevalence of loose and informal working relationships. SMEs operating in this state are more likely to rely on people rather than a system. In such situations, key personnel may resist documentation for two key reasons "(1) documentation is considered a waste of time and (2) documentation of processes and procedures makes the individual less dependable" [2]. SMEs in this state perform some or all of the processes required by ISO 9001 and the QMS may function fairly well. However, they are not willing and ready to document those processes unless there is a cultural change lead by top management.

3. *Formally Death: Some level of documentation, no functioning.*
SMEs categorized in this state have documented processes and procedures at some degree, however, the documents are generally not followed and do not necessarily reflect the actual manner in which the organization undertakes its operations and management. This situation highlights the fact that the mere existence of documentation does not necessarily lead to a functional QMS. Moreover, such a situation may help perpetuate the view that ISO 9001 is a way for SMEs to market their products and services but that implementation of the standard requires stacks of documents that offer no value.

4. *Formally Alive: Some level of documentation, some level of functioning.*
Each SME considered in this state, achieves a unique combination of the existence and functionality of processes and procedures that may or may not be required by ISO 9001. As illustrated in Figure 2, this situation is closest to the desired state of full functionality (100%) of the ISO 9001 QMS and full documentation (100%) of this functionality.

3. MOVING THE QMS FROM THE INITIAL STATE TO THE DESIRED STATE

The unique characteristics of each SME will affect the implementation process and the constituent elements of each step in different ways. After determining its initial state, the organization must formulate the unique implementation approach that will take it to the ideal state depicted in Figure 3. Building on that premise, Figure 4 shows the different ways that a SME can move from its initial state to the desired state of a 100% "Formally Alive" ISO 9001 QMS. Each of the transitions in Figure 4 is discussed below.
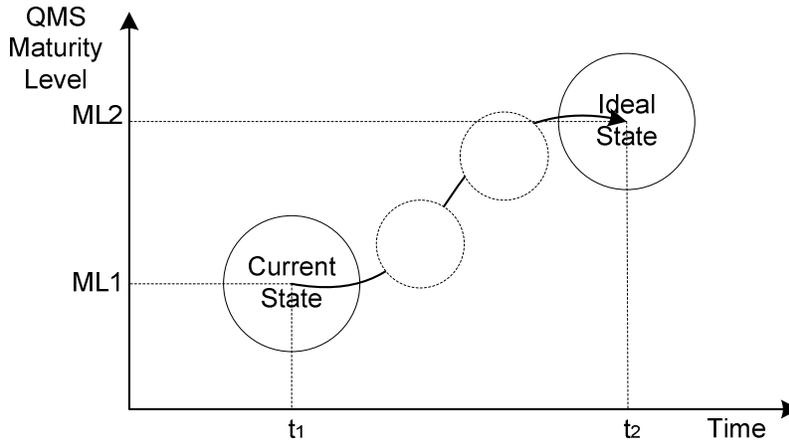
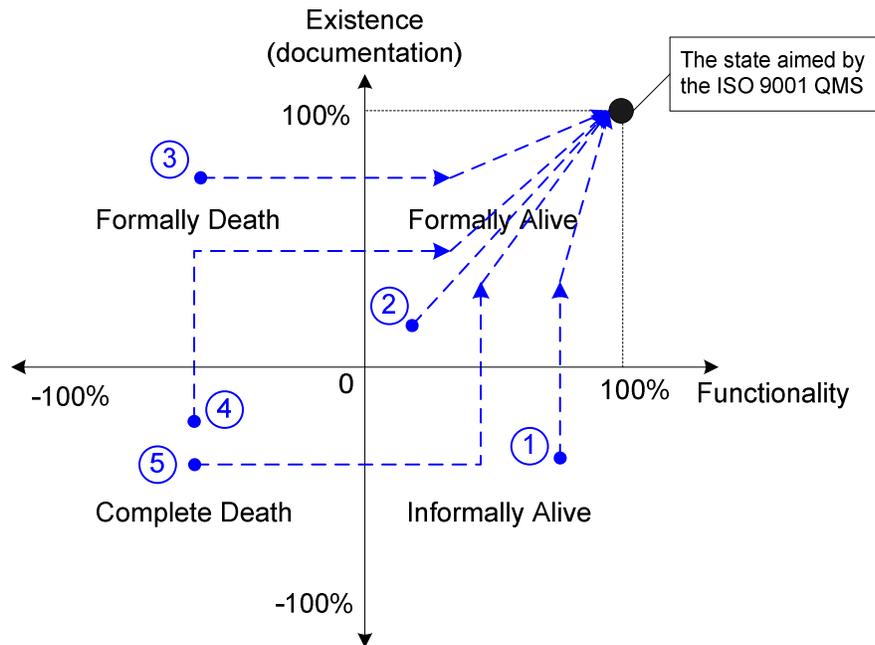**Figure 3:** Transition from Initial State to Ideal State

Figure 4: Directions to Formulate an Implementation Approach

1. Informally Alive ➔ 100% Formally Alive:
One advantage of being a SME in today's global economy is flexibility and agility in the face of changing conditions. The need to react quickly may help create a culture in some SMEs that is against establishing clear processes and procedures. Processes and procedures in this case are initially informal and in the minds of individuals. The first thing that needs to be done in this

situation is to help top management understand the importance of having a fully functional and documented QMS and the organizational changes that are required to achieve it. Once that is accomplished, training must be provided to increase the organization-wide awareness of the QMS and to facilitate the development of the required documentation. The selection of a competent management representative to communicate and continually reinforce the vision to the whole organization is particularly critical.

2. Formally Alive ➜ 100% Formally Alive:
For SMEs that achieve some degree of functionality and documentation in the management of their quality system, the transition to a fully functional and documented QMS will be less dramatic than in the other cases. The SMEs in this state are likely to have a strong base to build on, including an organizational culture influenced by a committed top management that is supportive of the needed education and training. Considerations to begin the transition to a 100% Formally Alive QMS, such as documenting the current processes that are informally done or initiating the realization of processes that have not been previously performed, will depend on the maturity level of the organization.

3. Formally Death ➜ 100% Formally Alive:
In this state, SMEs already have the documentation required for ISO 9001 certification. However, having the required documentation in place only ensures the "existence" of the QMS. This "Formal Death" of the QMS was relatively common in SMEs certified according to the 1994 version of ISO 9001. To implement a ISO 9001:2008 QMS that will help the organization increase customer satisfaction and continuous improvement, the SME must overcome several critical challenges. These include a short-term perspective of top management, the pressure to gain ISO 9001 certification as soon as possible, over reliance on external consultants, lack of spending on internal training, lack of QMS planning, and a product rather than a process focus.

4. and 5. Complete Death ➜ 100% Formally Alive:
A typical characteristic of SMEs in this state is "underorganization", a term explained by Yauch and Steudel [10]. In this context, underorganization describes a company that lacks structured systems, policies, and procedures. Organizational members accept this condition as a normal pattern and do not necessarily want a more structured system. "Problems associated with underorganized systems include the dissipation of human energy due to a lack of direction, unavailability of relevant information, and inefficient or undependable task accomplishment" [10]. Therefore, one of the biggest obstacles to the introduction of new processes and procedures in these particular SMEs is management realization and the ability of owner managers to modify their behavior and management style. A key starting point in the resolution of these problems is being explicit and disciplined about goals and objectives, roles and responsibilities, and promoting commitment to the QMS at all levels. The time required for these cultural changes to happen is usually longer than the time "desired" for the organization to be ISO 9001 certified. Figure 4 illustrates that SMEs in this state are therefore at risk of ending up with one of two states in the pursuit of a fully functional and documented QMS (Formally Alive): (1) Formally Death, which is represented along Path 4 or (2) Informally Alive, which is represented along Path 5.

4. KEY CONSIDERATIONS IN IMPLEMENTING ISO 9001 IN SMEs
There are many issues that must be addressed in moving the QMS from the initial state to the desired state. For example, all organizations implementing ISO 9001 will need to consider the unique culture within the organization, its size, and the resources available. Beyond those widely discussed points, three issues that merit particular attention are (1) consideration of the QMS as a parallel function, (2) training, and (3) auditing. Key points associated with these issues are discussed below.

Consideration of QMS as Parallel Function
In the case of all of the transitions depicted in Figure 4, real benefits from the QMS are more likely to be experienced if the QMS is implemented directly into the core structure of the organization. SMEs must be cautious against establishing a QMS that is run separately in

parallel to its other systems. In SMEs, the parallel subsystem most commonly exhibits itself as a separate Quality Assurance, or in some cases, ISO 9001 department. Possible reasons for this may include the existence of rigid departmental boundaries in some SMEs or overemphasis on core activities. As Yauch and Steudel [10] note, SMEs tend to focus their attention on "…necessary routine activities (such as sales, production, shipping, etc.) rather than activities aimed at improving processes or systems." If a SME insists on establishing a separate quality department, its level of effectiveness can be increased by embedding the QMS in widely-used organizational systems where practical. The integration is largely a function of how well the QMS manages to share information with other subsystems and its ability to align with the policies, norms, goals, and values in place throughout the organization.

Training
In SMEs, training and staff development is more likely to be *ad hoc* and small scale because of modest human and financial resources and the absence of a specific training budget. To prevent the problems arising from lack of education and training, two things must be done:

1. Education of Top Management: The centralization of decision-making processes within many SMEs means that the management can either be the main stumbling block to change or the main catalyst for change. Therefore, any approach to ISO 9001 implementation must involve considerable education for the top management of the organization to create awareness and understanding of the implementation process as a change initiative. Implementing a fully functional and documented QMS requires motivation by top management to appreciate, achieve, and implement the necessary measures to meet the standards' criteria.
2. Education and Training of Employees: SMEs are often under pressure to quickly gain ISO 9001 registration. Meeting the requirements of the standard in a short period of time can prove a formidable obstacle for a small company. Since most SMEs do not possess the needed expertise internally, they may be inclined to hire external experts to provide the necessary technical expertise and manpower. However, having a functioning and documented QMS requires more than that. It requires ensuring that all employees in the organization clearly know what is expected of them and how they can contribute to the attainment of their organizations' goals. This will likely require the preparation and implementation of a training plan tailored specifically to the unique characteristics and maturity level of the SME.

Auditing
As emphasized throughout the paper, a QMS is not going to produce the expected results unless it is fully functional. While auditing must therefore verify the existence of the necessary documentation, it must also focus on the functionality of the QMS. The measurement of the functionality and the qualitative and financial impacts of a QMS have been the subject of several studies, including Kaynak [11]. Among the categories used to measure functionality and performance improvement, two are particularly noteworthy for our purposes: management commitment and employee involvement. A QMS cannot be functional in the absence of those two characteristics. Therefore, as a minimum, internal and external auditors should continually verify top management's commitment to increased company-wide quality awareness and improvement in addition to employee involvement in the design, implementation, operation, and improvement of quality related processes and procedures.

5. SUMMARY AND CONCLUSIONS
A wide variety of methodologies and techniques have been used to implement ISO 9001. However, successful implementation of ISO 9001 in SMEs can be elusive. One of the key reasons for this is that many SMEs overlook the complexity of the implementation processes and the organizational changes that are needed to ensure the QMS is fully functional.

A fully functional QMS leads to increased customer satisfaction and continuous improvement of business results. Although the existence of documentation is a key requirement of a functional

ISO 9001 QMS, it is not in itself sufficient. To develop and implement a fully functional ISO 9001 QMS, it is essential that a SME correctly identifies the initial state of its QMS and the path it will follow to achieve the desired state.

This paper uniquely contributed to this challenge by identifying four main initial states and five paths from the initial to the desired state. The explicit identification of the initial states and paths will help managers within SMEs to understand the process of implementing ISO 9001 and the fundamental issues that they must address. This should help SMEs to prevent implementation failures within their organizations. In particular, the paper cautions that ISO 9001 cannot be a documentation-driven system. Although this may indicate the QMS exists, it does not guarantee its functionality.

REFERENCES

1. International Organization for Standardization (ISO). "*The ISO Survey 2007*", International Organization for Standardization, Geneva, 2008.

2. A. Ghobadian, and D. Gallear, "*TQM and organization size*", International Journal of Operations and Production Management, 17(2):121-163, 1997.

3. S. Rahman, "*A comparative study of TQM practice and organizational performance of SMEs with and without ISO 9000 certification*", International Journal of Quality and Reliability Management, 18(1):35-49, 2001.

4. A. Bayati and A. Taghavi, "*The impacts of acquiring ISO 9000 certification on the performance of SMEs in Tehran*", The TQM Magazine, 19(2):140-149, 2007.

5. T. Koc, "*The impact of ISO 9000 quality management systems on manufacturing*", Journal of Materials Processing Technology, 186:207-213, 2007.

6. W.G. Lewis, K.F. Pun, and T.M. Lalla, "*Exploring soft versus hard factors for TQM implementation in small and medium-sized enterprises*", International Journal of Productivity and Performance Management, 55(7):539-554, 2006.

7. B. Poksinska, J.A.E. Eklund, and J.J. Dahlgaard, "*ISO 9001: 2000 in small organizations: lost opportunities, benefits and influencing factors*", International Journal of Quality and Reliability Management, 23(5):490-512, 2006.

8. J.C Pinho. "*TQM and performance in small medium enterprises: the mediating effect of customer orientation and innovation*", International Journal of Quality and Reliability Management, 25(3):256-275, 2008.

9. Z.T. Temtime. "*The moderating impacts of business planning and firm size on total quality management practices*", The TQM Magazine, 15(1):52-60, 2003.

10. C.A. Yauch, and H.J. Steudel, "*Cellular manufacturing for small businesses: key cultural factors that impact the conversion process*", Journal of Operations Management, 20(5):593-617, 2002.

11. H. Kaynak, "*The relationship between total quality management practices and their effects on firm performance*", Journal of Operations Management, 21(4):405-435, 2003.

Sousa-Poza, Mert Altinkilinc & Cory Searcy

Study of the thermal behavior of a synchronous motor with permanent magnets

Souhil Seghir-Oualil
seghir.ouali@univ-valenciennes.fr
LME, UVHC, Le Mont Houy 59 313
Valenciennes Cedex 9, France

Souad Harmand
souad.harmand@univ-valenciennes.fr
LME, UVHC, Le Mont Houy 59 313
Valenciennes Cedex 9, France

Daniel Laloy
daniel.laloy@jeumontelectric.com
Jeumont Electric, 27 rue de l'Industrie,
59 460 Jeumont, France

Abstract

The present study deals with the thermal behavior of an electric motor used in naval propulsion. A permanent three-dimensional model based on a nodal approach was developed to predict the thermal behavior of the machine in permanent state. The various heat transfer modes playing a role (mainly the conduction, the convection and the mass transfer) are taken into account. The three dimensional developed model provides information (in terms of heat fluxes or temperatures) for the whole machine. A parametric study using this model is carried out aiming at two objectives. The first one consists in an evaluation of the influence of several factors on the temperature distribution. The second one aims at obtaining the inception of the thermal optimization and evaluating the more efficient cooling techniques or solutions.

Key words: Convective heat transfer, electrical synchronous motor, cooling enhancement, infrared thermography, thermal modeling

Nomenclature

a	thermal diffusivity = $\lambda/\rho c_p$	$m^2 s^{-1}$
c_p	specific heat	$J\ kg^{-1}K^{-1}$
D	rotor diameter	m
e	air-gap thickness	m
e_{eq}	equivalent thickness = $2S_{mo}/P_{mo}$	m
G	thermal conductance	WK^{-1}
g	gravity	$9.81\ m\ s^{-2}$
H	fin width	m
h	heat transfer coefficient	$W\ K^{-1}m^{-2}$
L	length of the rotor	m
P	perimeter	m
Q_m	mass flow rate	$kg\ s^{-1}$
R	radius	m
S	surface	m^2
T	temperature	°C
Z	axial coordinate	m
U, V	velocity	$m\ s^{-1}$
Dimensionless numbers		
Nu	Nusselt number	
Pr	Prandtl number = $\mu c_p/\lambda$	
Ra	Rayleigh number = $g\beta\ (T_p-T_\infty)\ D_i^3/a\nu$	
Re_a	axial Reynolds number = $2eU_a/\nu$	
Re_r	Reynolds number = $R_c U_\infty/\nu$	
Ta	Taylor number = $2e^{0.5}R_{ro}^{1.5}\omega/\nu$	

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

Greek symbols

α	angle	rad
β	thermal coefficient of volumetric expansion	K^{-1}
ϕ	heat flux	W
λ	thermal conductivity	$W\,K^{-1}m^{-1}$
η	geometrical report/ratio of form = R_{st}/R_{ro}	
μ	dynamic viscosity	$kg\,m^{-1}s^{-1}$
ν	kinematic viscosity	m^2s^{-1}
ρ	density	$kg\,m^3$
ω	rotational speed	$rad\,s^{-1}$

Subscripts and superscripts

a	axial
c	frame
cond	conduction
conv	convection
flui	fluidic
i	interior
m	averaged
mo	wet
op	optimal
p	wall
r	rotational
ro	rotor
st	stator
∞	ambient or far from the wall

1. INTRODUCTION

Since the electric motors beginning and the will to increase their performances, such machines are confronted to the heating problems. Some parts of these machines can reach significant and disallowed temperatures entailing their deterioration. These parts are in particular the winding insulators, the bearings or the magnets. These latter are particularly sensitive to high temperatures since they can loose a part of their magnetization when they are exposed to a too high temperature. In order to preserve the optimal conditions ensuring the good operation of the motor, it is imperative to well understand and control the thermal behavior of these motors. This good understanding needs the ideal design of coolers to ensure the operating temperatures not exceed the acceptable limits of used materials.

The synchronous motors operation is based on the existence of a turning magnetic field which can be obtained by the circulation of a controlled current into stator winding. Only a part of the electric power is converted into mechanical power, the other part entails a warm up of the machine and constitutes the losses to be evacuated. To predict the lifetime of an electric motor and to improve its efficiency, it is necessary to carefully investigate its thermal design in order to determine the best cooling solutions.

In this work, a three-dimensional numerical code based on a nodal approach was developed in order to better apprehend the thermal behavior of the machine. The model provides information about the thermal behavior of the motor letting obtain the whole machine temperature.

2. DESCRIPTION OF THE STUDIED ENGINE

Figures 1 and 2 present respectively an axial and a radial cross-section of the studied engine. This kind of motor is used for the naval propulsion. It is mainly composed by two parts: the rotor (rotating part) and the stator (static part) separated by the very thin air-gap (some millimeters). The stator is made by a succession of steel sheets insulated electrically. It is at the origin of important electromagnetic dissipations. These sheets are axially grooved and the slots are equally distributed

following the angle. In these latter, insulated coppers conductors are inserted to compose the statoric winding. These conductors are rolled up around the magnetic teeth. The exceeding insulated copper part is called the end-winding. The rotor consists in an empty rotating cylinder placed inside the stator and maintained with two side-plates. Permanent magnets equally distributed with the angle are stuck at its periphery. The rotor and stator are placed together into a frame made of steel. The latter aims at protect and maintain the various parts of the machine.
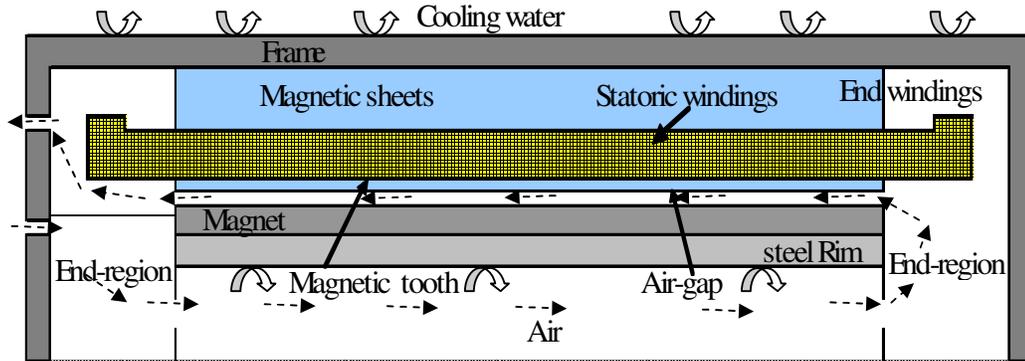
Figure 1: Axial cross-section of the motor
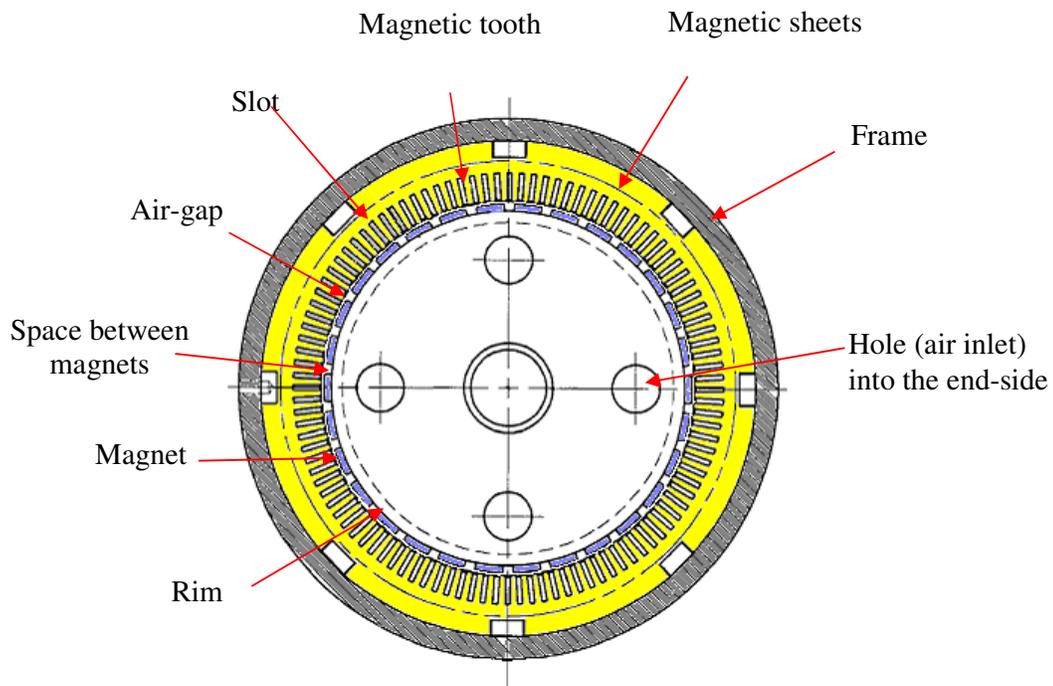
Figure 2: Radial view of the motor

This high power motor is of a great compactness. The most part of the absorbed electric power is converted into mechanical power and drives the motor into rotation. The residual part (losses) of this energy must be evacuated since it entails the heating up of the machine. These losses constituting the heating sources are mainly known as the Joule effect in the windings of the stator, and the iron losses located mainly in stator sheets and very partially into the magnets of the rotor. The cooling of this electrical machine is ensured both by a very low airflow rate into the empty rotor and by the cooling water which flows at its external periphery.

3. NUMERICAL MODEL

Thermal designing needs a good knowledge of the phenomena playing a role. The predominant phenomena occurring in the studied electric motor are: the conductive heat exchanges between the various solids parts of the machine, the convective heat exchanges between the walls and the fluids, and also the heat exchanges due to the airflow. Due to the low temperature levels, radiation can be neglected. A nodal approach was chosen to model the thermal behavior of this engine. This method, rather simple to implement in steady state operation, lets quickly calculate the temperature field into the whole machine using a reasonable node networks and a reduced storage capacity.

To establish a nodal thermal model of this synchronous motor, a nodal network of the machine was performed and divides the machine into numerous solid and fluid isotherm blocks which is represented by a node. The fluid area connections are either represented by convective thermal resistances, either by fluidic thermal resistances, either both. Conduction only occurs between solid area nodes. Each connection between neighbor solid nodes is represented by a conductive thermal resistance.

3.1 Nodal analysis

The nodal method was initially developed for the electrical network analysis, and was quickly applied successfully to thermal systems [1-3]. This approach consists in dividing the system in a certain number of elementary volumes assumed as isotherm and represented by a node. Heat transfers between these nodes are modeled using thermal conductance, heat or sink sources and/or imposed temperature conditions. Then this method consists on the one hand to represent the various elements of the system (rotor, stator, airflow…) by isothermal nodes in order to build an equivalent network and on the other hand to identify the various heat transfer modes occurring in this network. After having discretized the domain, the equations of heat transfers associated to each node can be written. These equations represent the energy conservation within the element of volume related to a given node. Thus, this energy balance expressed the equality between the internal energy variation (within the considered volume connected to this node) and the energy exchanged with other nodes using the various heat transfer modes. In steady state, for a node (i) connected with the nodes (j), the heat balance can then be written as:

$$\sum_{j\neq i} G_{ij}^{cond}(T_j - T_i) + \sum_{j\neq i} G_{ij}^{conv}(T_j - T_i) + \sum_{j\neq i} G_{ij}^{flui}(T_j - T_i) + \phi_i = 0 \tag{1}$$

In this equation ϕ_i referred to an internal heat source, and G_{ij}^{cond}, G_{ij}^{conv}, G_{ij}^{flui} referred respectively to the conductive, convective and fluidic (due to the airflow) conductances between the node i and the nodes j. T_i and T_j referred to the temperature of the node i or j. The radiative exchanges are neglected due to the nature of the walls and the temperature levels into the motor. Applying this equation to the whole machine and using a matrix form, (1) can then be written:

$$\left[G_{ij} \right] \times \left[T_i \right] = \left[\phi_i \right] \tag{2}$$

In this last equation, [G_{ij}] is the matrix of the thermal conductances, [T_i] is the temperature vector, and [ϕ_i] is the heat source /sink vector. The temperatures field ([T_i]) into the whole motor can then easily be obtained since:

$$\left[T_i \right] = \left[G_{ij} \right]^{-1} \times \left[\phi_i \right] \tag{3}$$

The specific computer code was developed for modeling of the thermal behavior. The resolution is carried out in two steps. The first step consists in the (nodal) network building and the calculation of both the thermal conductances and the heat sources. The second step consists in the inversion of the matrix to obtain the temperatures.

3.2 Model description

An electric machine constitutes a complex thermal system, but presents several symmetries and periodicities which were used in the model to decrease the node number. Table 1 gathers the various geometric elements as well as various heat flux and conductances used in the model.

heat transfert type	conductance and heat flux	Scheme
Conduction *(cylindrical element)*	$$G_{ij} = \frac{\alpha_{ij}\ell_{ij}}{\dfrac{\left|\ln(R_{ij}/R_{ii})\right|}{\lambda_{ij}} + \dfrac{\left|\ln(R_{jj}/R_{ij})\right|}{\lambda_{ji}}},$$ $$\phi = G_{ij}(T_i - T_j)$$	
Conduction *(plane element)*	$$G_{ij} = \frac{S_{ij}}{\dfrac{e_{ij}}{\lambda_{ij}} + \dfrac{e_{ji}}{\lambda_{ji}}}, \quad \phi = G_{ij}(T_i - T_j)$$	
Convection	$G_{ij} = h_{ij}S_{ij}, \quad \phi = h_{ij}S_{ij}(T_i - T_j)$	
Fluid flow	$G_{ij} = Q_m c_p, \quad \phi = Q_m c_p (T_i - T_j)$	

Table 1: Expressions used for G_{ij}
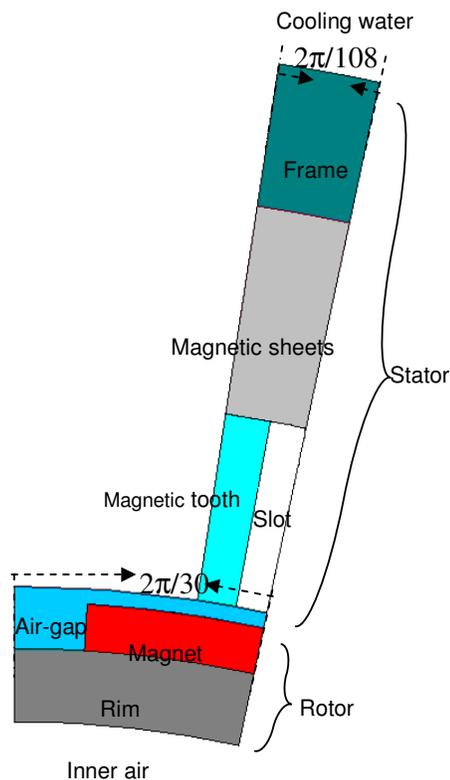
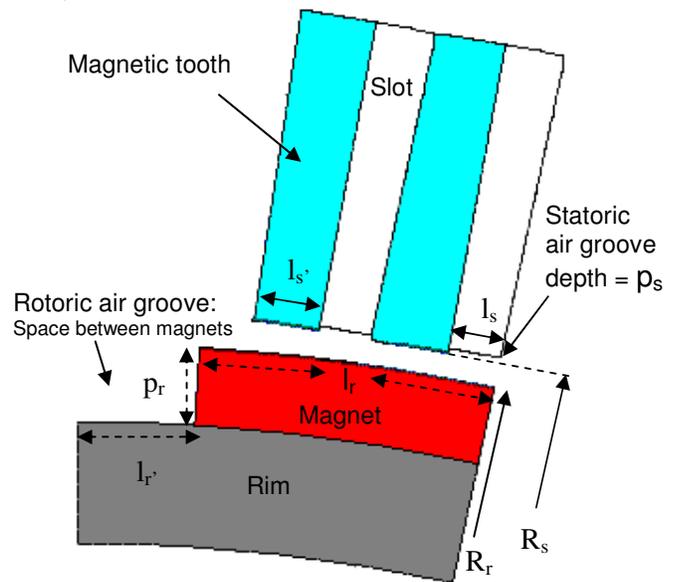**Figure 3 : 2D Model for the stator and the rotor**

Figure 4: Radial cross-section of the air-gap (close up)

Earlier works on the rotating machines carried out by Bertin [4] and Vasilescu [5] showed it was possible to only study an elementary sector instead of the whole machine. Using geometrical considerations, only 1/108th of the stator and 1/30th of the rotor were modeled. These two parts are then coupled using the mathematical coupling suggested by Bertin [4]. As shown on the figures 3 and

4, each elementary pattern of the stator is composed with a slot, a magnetic tooth, a portion of the magnetic sheets and the frame. For the rotor, the studied pattern consists in a magnet, an air space between magnets and a portion of the rim. With these considerations, the studied structure presented on the figure 3 can be divided in 7 independent node blocks.

3.3 Boundary conditions

On the radial plane presented on the figure 3, convective boundary conditions are applied to the internal and external surfaces (inner air and water). For side surfaces, Neumann's type boundary conditions ($\partial T/\partial n=0$) are applied to take into account the symmetries. The temperature of the cooling water (T_{water}) is 40 °C whereas the air temperature at the rotor inlet (T_{air}) is 45 °C. An adiabatic condition is applied at the two end-sides (in the axial direction) of the machine.

3.4 Thermal conductivities, losses and convective heat transfer coefficients

Parts	Material	Thermal conductivity (W/mK)	
Windings	Copper	380	
taping	enamel, glass fiber, epoxy resin	0.15	
turn insulation	glass fiber, mica, epoxy resin	0.15	
Ribbon	Dacron®, epoxy resin	0.15	
spacer	glass fiber, epoxy resin	0.15	
Impregnation	epoxy resin	0.15	
Stick		0.15	
Magnets	NdFeB or frit SmCo	9	
Frame, Rim	Steel	50	
Magnetic sheets	Steel, electric insulation on both faces	25 (in the radial cross section)	2 (in the axial cross section)

Table 2: Materials used and their physical properties

The materials constituting the different parts of the electrical motor, and their thermal conductivities are gathered on table 2. In the studied motor, the windings need a specific attention since they are not isotropic materials and are subjected to very severe thermal constraints. These windings are made of copper, enamel, epoxy resin, Dacron® and glass fiber. Their thermal resistances are determined by considering they are a 3D thermal network (serial and parallel resistance association in the radial, ortho-radial and axial direction) of their different composing materials. In the radial and ortho-radial direction, the insulating materials govern the values of equivalent thermal conductivities whereas this value is governed by the high conductivity of the copper in the axial direction. The values obtained for the thermal conductivities in the three directions are gathered on table 3.

Direction	Radial	Ortho-radial	Axial
Thermal conductivity (W/mK)	0.59	0.71	237

Table 3: final windings (windings+insulation) thermal conductivity

During operation, a significant number of losses are generated in the various components. The table 4 gathers their localizations as well as their experimental values. The determination of the convective thermal conductances needs the knowledge of the convective heat transfer coefficients. These coefficients are often determined using correlations available in literature. Most of the studies [4,5] about the modeling of the electric motors suggest to take into account the convection in a very simplified way, by imposing a constant heat transfer on all the wall in contact with the fluid.

Location of the losses	losses (W)

Iron losses into the statoric windings (air-gap side)	17550
Iron losses into the statoric windings (frame side)	15150
Iron losses into the end-windings	6100
Iron losses into the teeth	21000
Iron losses into the frame	13300
Losses into the magnets	3000
Total losses	**76100**

Table 4: Location and values of the losses in the motor

Keeping in mind that the engine frame is underwater and that airflow flows inside an empty rotor as well as in its air-gap, 4 convective heat transfer coefficients have to be identified: into the rotor, into the air-gap, at the periphery of the frame and onto the two internal face of the end-regions.

Air in the air-gap:
The studied air-gap consists in a cylinder rotating inside a static one with a ribbed air-gap where air flows axially. The rotor has 30 axial slots equally spaced on its external surface (depth: 18.1 mm, width: 28 mm) whereas the stator has 108 axial slots equally distributed on its inner circumference (depth: 0.8 mm, width: 12.3) (see figure 4). In such a configuration, the flow and the convective heat transfer are both controlled by the Taylor number and the axial Reynolds number respectively characterizing the rotation and the airflow effects. Several authors proposed correlations giving the Nusselt number in an air-gap with an axial airflow [6-10]. These correlations obtained experimentally were often carried out on quite specific engines. So, with these correlations, it is difficult to accurately identify the heat transfer coefficients in the configuration investigated here. Initially, the correlation of Tachibana and Fukui [6] is used since it is the less restrictive (in term of hydrodynamic parameters and of geometry). It entails to consider a smooth air-gap even if the studied configuration corresponds to a grooved rotor. The case of grooved air-gaps was investigated experimentally by some authors [8-10]. The influence of the existence or not of the slots is only discussed in a qualitative way and some appreciations can be found. Unfortunately, no general relations adaptable to this study are obtained. Thus, the heat transfer coefficient is estimated using the correlation (4) corresponding to the smooth air-gap. In this correlation, an equivalent air-gap thickness is used to take into account the slots (rotor and stator).

$$\mathrm{Nu}_a = 0.015\left(1+2.3\frac{2e_{eq}}{L}\right)\eta^{0.45}\,\mathrm{Re}_a^{0.8}\,\mathrm{Pr}^{0.33}$$

$$\mathrm{Nu}_r = 0.046\times\mathrm{Ta}^{0.66}\times\mathrm{Pr}^{0.33}$$

$$\mathrm{Nu} = \mathrm{Nu}_r + \mathrm{Nu}_a \tag{4}$$

$$\text{where: } e_{eq} = 4\frac{S_m}{P_m} = 4\frac{\pi(R_s^2 - R_r^2) + 108p_s l_s - 30p_r l_r}{108(l_s + l_{s'} + 2p_s) + 30(l_r + l_{r'} + 2p_r)}$$

Using the geometrical and dynamic characteristics of the engine the value of the average heat transfer coefficient is obtained, and

$$h = \frac{\lambda \mathrm{Nu}}{2e_{eq}} = 22\,\mathrm{W/m^2 K} \, .$$

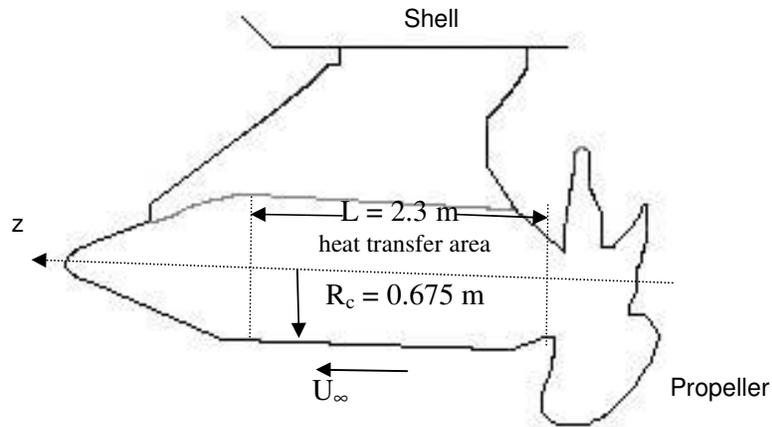*Convective heat transfer between the water (external cooling agent) and the frame:*

Figure 5: POD version Motor

The frame represents the external envelope of the machine. It is in direct contact with the cooling surroundings water. In the case of an on-board motor, the water circulation is carried out using an external pump. It flows through a heat exchanger before flowing through the frame of the engine. In the case of a POD configuration, the engine is directly underwater and the frame is then in direct contact with sea water at the maximum relative speed of 20 knots, *i.e.* 10.3 m/s. The local convective heat transfer can then be evaluated using the Stasiulevichius correlation [11].

$$Nu_z = 0.0106 Re_z^{0.86} \left[1 + 0.37 \left(\frac{z}{R_c} \right)^{0.8} Re_r^{-0.2} \right]^{0.14} \left(\frac{T_c}{T_\infty} \right)^{-0.25} \tag{5}$$

With $\quad Re_z = \dfrac{U_\infty z}{\nu} > 2 \times 10^6$

The total averaged heat transfer for the whole zone is thus obtained by integration of this local Nusselt number over the length where the heat transfer occurred (see fig. 5): $Nu_m = \dfrac{1}{L} \int_0^L Nu_z dz$. The table 5 gathers the values of the heat transfer coefficients for both the configurations (on-board or POD).

Configuration	water speed (m/s)	$Re_l = \dfrac{1}{L} \int_0^L \dfrac{zU_\infty}{\nu} dz$	$h_m = \dfrac{\lambda Nu_m}{L}$
on board motor	2	10^7	2757 W/m²K
POD version	10.3	$3.5\ 10^7$	11575 W/m²K

Table 5: Heat transfer coefficients between the frame and the cooling water

Convective heat transfer inside the rotor:
The airflow inside the rotor is due to the coupling of rotation from the rotor and the axial airflow caused by the blower. Few works about the identification of the convective heat exchanges in such a configuration can be found. Among them, some authors [12, 13] showed the rotation of the cylinder (rotor) causes a destabilizing effect on the laminar flow which becomes turbulent. Thus, the rate of heat transfer between the wall and the airflow becomes higher. Other authors [14-16] conclude the rotation of the cylinder reduces this heat transfer rate if the axial air inflow at the inlet is ever turbulent. It was however observed that these various studies cover a relatively limited range since they are valid only for high values of the geometrical ratio L/D_i. The few results available in the literature do not allow to accurately identify the trends or to obtain suitable correlations connecting the Nusselt number to both the axial and rotational Reynolds numbers. A model was developed to understand the heat transfer mechanisms in such a configuration. This study proposes a correlation at the high rotational

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

speed and the low axial airflows, where the transfer only depends on the rotational Reynolds number [17].

$$Nu = 2.85 \times 10^{-4} \, Re^{1.19} \tag{6}$$

$$Re = \frac{\omega D_i^2}{2\nu} > 2.77 \times 10^5$$

4. END-REGIONS STUDY
To accurately model the end-regions using a nodal approach, it is of paramount interest to know their aerodynamics. Aiming at that, CFD simulations of the internal flows of these end-regions were realized. These simulations were carried out using the commercial CFD software Fluent. The results obtained lets then identify parameters used in the nodal modeling.

4.1 Geometrical description of the end-regions and grid
The stator of the studied engine has 108 slots. However due to the periodicity, the symmetries of the problem, and the nature of the flows, an angular sector of 23.3 ° was retained for this numerical study and allows a reduction in the cells number of the mesh and thus of the calculation cost. Figures 6 and 7 present respectively the simplified geometries for the two end-regions: the left end-region where the air inflows from the air-gap, and the right end-region where the air inflows from the rotor before outflow through the air-gap. On the real machine, the air flows into the rotor through 4 holes equally distributed on the rotor end-side (revolving wall). Then the air flows through the air-gap and is evacuated outside through 7 holes drilled in the frame. Due to the existence of these four holes in the rotor end-side, the exact modeling of the real machine requires to model at least a quarter of the machine. This option increases the number of cells in an excessive way. So, an equivalent annular surface is used instead of the real geometry (4 holes). The same approach was adopted for the 7 holes letting the air flows outside. The mass conservation lets determined in both cases the equivalent surfaces. Thus, in the simplified model for the right end-region, the air (at the temperature T_∞) flows in by an annular space located the revolving part (rotor end-side) and flows out by an annular space (air-gap). The internal flow is supposed to be incompressible, turbulent, and three-dimensional. The physical properties of the fluid are assumed as constant.
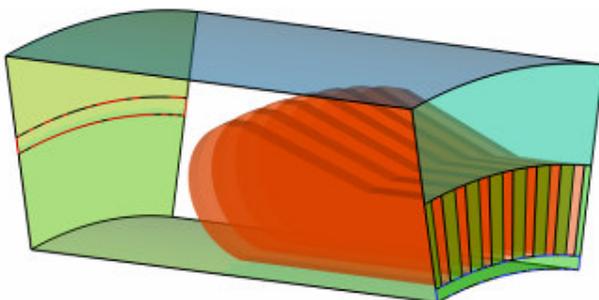

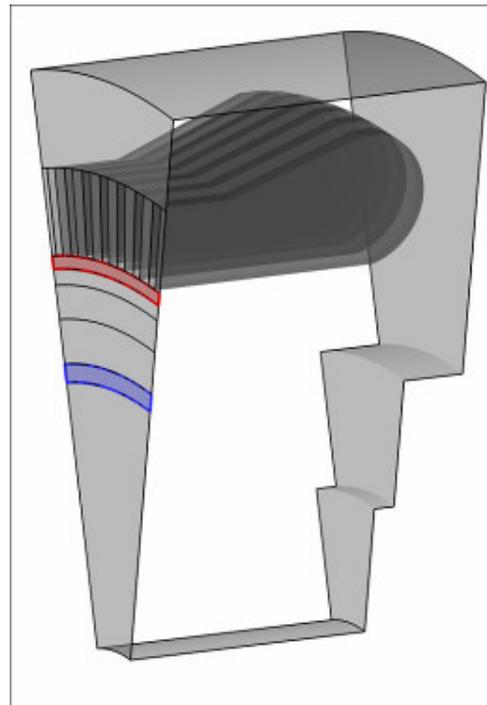Figure 6: Geometry of the left end-region   Figure 7: Geometry of the right end-region

The geometry and the grid were built using Gambit. This grid (quadrilateral and tetrahedral elements) has approximately 530 000 cells in the right end-region and 900 000 in the left one. These significant cell numbers are due to the alternation between big size and small size volumes (space between end-windings, air-gap,…). Figures 8 and 9 present the grid of these two end-regions.
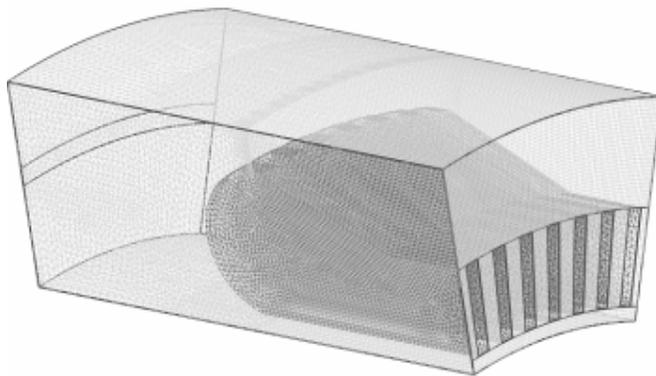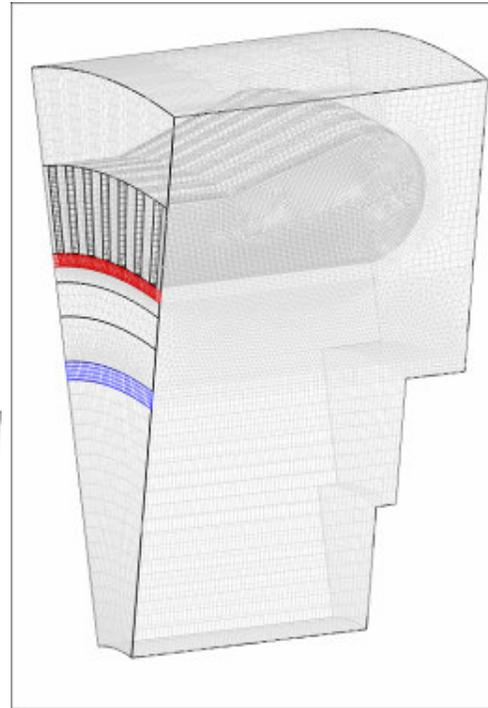

Figure 8: Mesh of the left end-region

Figure 9: Mesh of the right end-region

The thermal and aerodynamic boundary conditions of the problem are gathered into tables 6 and 7 for both the end-regions. These temperatures are estimated using a nodal modeling. Simulations were carried out for only one operation point characterized by an inlet mass flow rate of 0.18 kg/s and a rotational speed of 250 rpm. This imposes the inlet air velocity in the end-regions and the rotational speed of the revolving walls. The direction of the inlet velocity vector is imposed using the components of the velocity vector (high tangential component).

Part	Thermal	Motion
Rotor end-side	45°C	Rotation, 250 rpm
Rim	129°C	Rotation, 250 rpm
Magnets	130°C	Rotation, 250 rpm
Magnetic tooth	97°C	Static
End windings	174°C	Static
Magnetic sheets	66°C	Static
Frame	40°C	Static
Bearings	Adiabatic	Static
shaft	Adiabatic	Rotation, 250 rpm
Air inlet	45°C (=T)	Q_m =0.18 kg/s imposed direction

Table 6: Boundary conditions of the right end region

Part	Thermal	Motion
Magnetic tooth	93°C	Static
End windings	146°C	Static
Magnetic sheets	64°C	Static
Frame	40°C	Static
Bearings	Adiabatic	Static
Air inlet	90°C (=T)	Q_m =0.18 kg/s imposed direction

Table 7: Boundary conditions of the left end region

4.2 Numerical model

Fluent is based on the finite volumes method to express the mass, the momentum, as well as the energy conservation (Navier-Stokes equations). The turbulent phenomena are taken into account using the well-known k- RNG model. This statistical model uses two transport equations (closure equations) for the turbulent kinetic energy (k) and for the dissipation rate ( ). This model, in its RNG formulation (based on the theory of the renormalization group), was chosen because it has an option well adapted to the types of studied flows ("swirl dominated flows") and was suitable for many configurations. For this numerical model, the wall laws were also used ($y^+ \approx 30$). The discretization scheme chosen for the pressure is the "PRESTO!" model. This model is recommended [18] in the case of rotating flows. The first order "Upwind" scheme was used for the equations of moment, energy and transport (turbulent kinetic energy and rate dissipation). For the pressure-velocity coupling, the SIMPLE model was retained for its wide range of use. The convergence criteria are based on the values of the residues. The solution is supposed to be converged when all the residues are lower than 10^{-7}. This convergence was also checked by verifying the energy and mass balances.
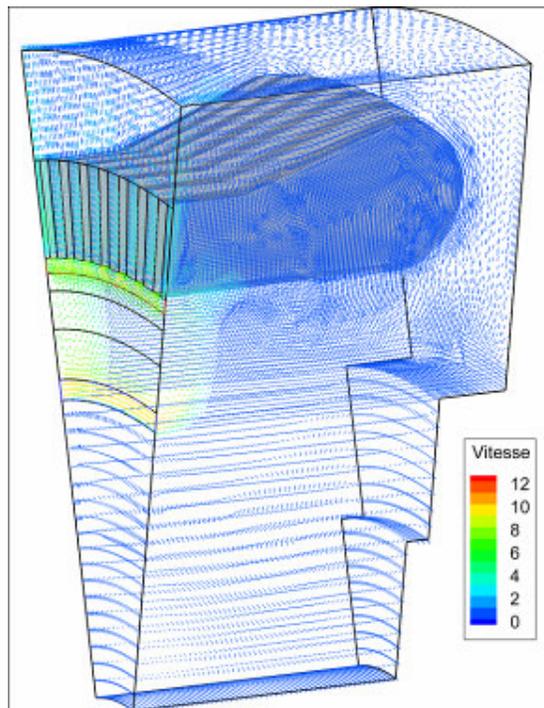
4.3 CFD results

Figure 10: Velocity vector field in the right end-region
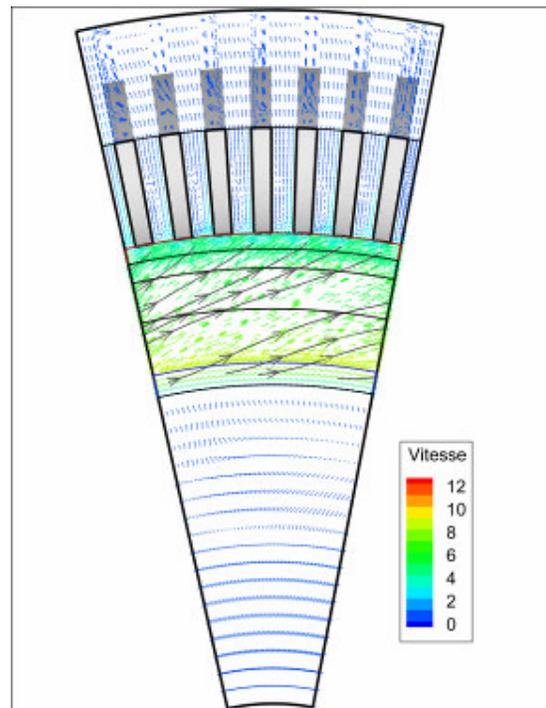

Figure 11: Velocity field in a radial cross section of the right end-region

Figure 10 shows the velocity vectors field in the right end-region letting obtain internal aerodynamics in particular around the end-windings. This figure shows the main part of the inlet airflow directly flows into the air-gap. Maximum velocity magnitudes are located on the area between the rotor inlet and the outlet through the air-gap (speed ranging from 6 to 11 m/s). Velocities in the other areas are definitely lower (between 0 and 1.5 m/s). The importance of the tangential component clearly appears on figure

11 which represents the velocity field (in a radial plan). A swirling effect of the flow can be observed there. It is caused by both the rotation and the driving effect of the various layers of air (viscous effects).
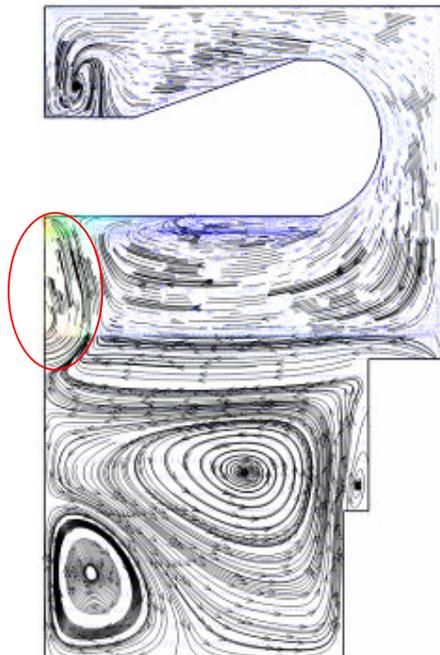
Air-gap

(air outlet)

End plate

(Air inlet)

Figure 12: Velocity field in an axial cross section of the right end-region (through an end-winding)
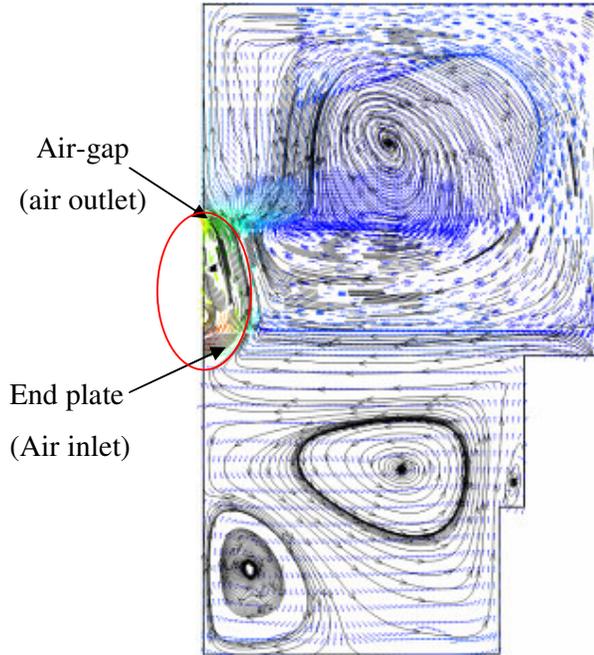
Figure 13: Velocity field in an axial cross section of the right end-region (between two end-windings)

Figures 12 and 13 present respectively the velocity field in an axial plane crossing an end-winding and in an axial plane between two end-windings. These figures confirm what appeared on figure 10, namely that the majority of the inlet airflow flows immediately into the air-gap and a negligible part flows around the end-windings.

Contra-rotating recirculation areas where the velocity magnitude is low can also be identified. One is located below the end-windings, where the velocities remain very low. It is mainly caused by the particles flowing slowly downward from the upper part. These last are slowed down due to the presence of the main airflow flowing out of this end-region through the air-gap. These figures also point out the presence of two contra-rotating recirculation areas in the lower part of this end-region. These two low intensity recirculation areas occur in the whole lower part of the right end-region.
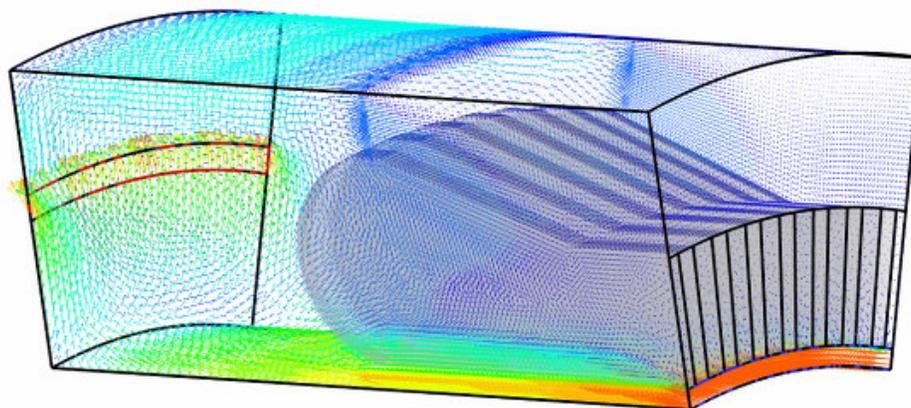
Vitesse: 0.5 1 1.5 2 2.5 3 3.5 4 4.5 5 5.5

Figure 14: 3D velocity field in the left end-region

Figure 14 shows the velocity vector field in the left end-region. The most part of the inlet airflow flows directly outside the machine without crossing through the end-windings. Due to the important tangential velocity component, an important swirling effect of the air can be noted. This figure also shows that the maximum velocity magnitude zones are under the end-windings and at the outlet, whereas the minimal velocity magnitude areas are located between the end-windings and in the upper part of the left end-region (above the end-windings).
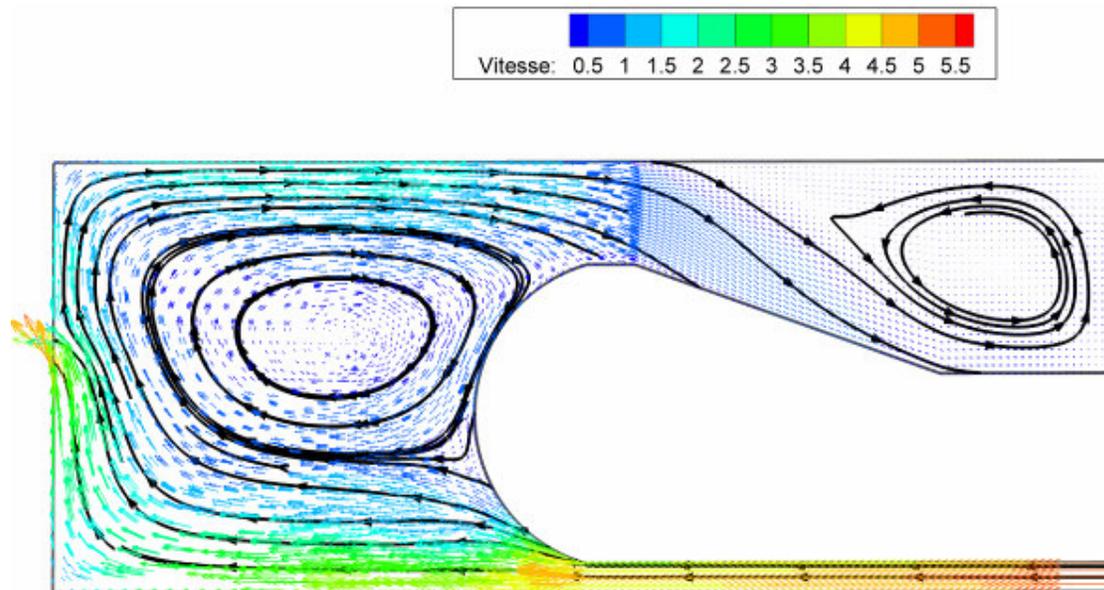

Figure 15: Velocity field in an axial cross section of the left end-region (through an end-winding)
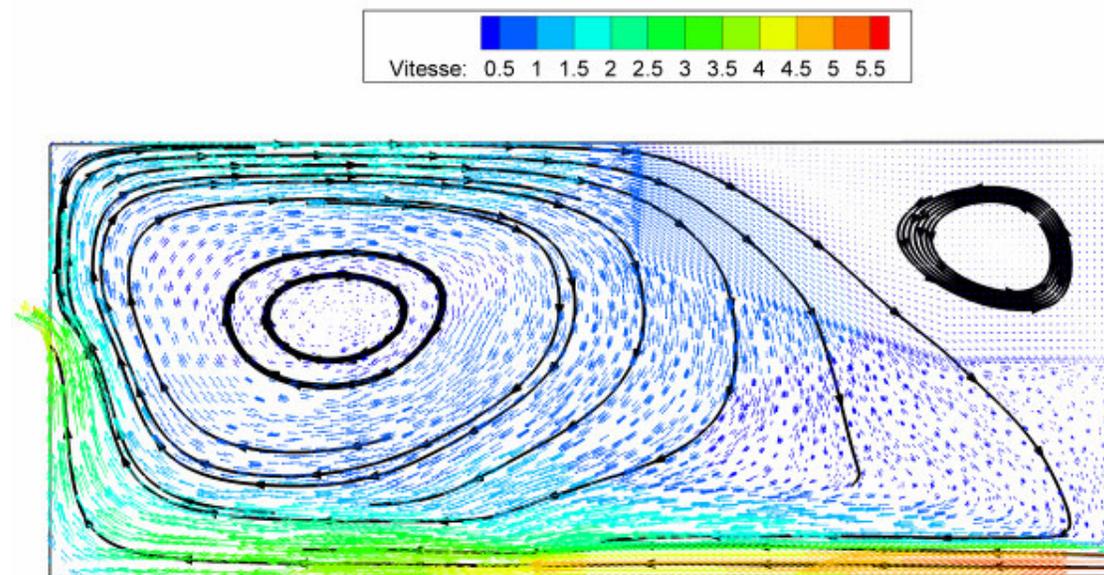

Figure 16: Velocity field in an axial cross section of the left end-region (between two end-windings)

Figures 15 and 16 present the velocity vectors field and the path lines in axial plane of the left end-region. The previous observations can be found again, namely that the majority of the airflow penetrating into the left end-region flows directly outside without crossing the end-windings. In particular, a very small part of the main flow tends to flow upwards to the frame and reached the top of the left end-region. This part flows along the top of the frame before flowing downwards along the vertical wall of the end-region. Thus, a recirculation area can be observed. This last, of low magnitude,

occurs in the whole upper part of the end-region. Finally, the flow is composed of three zones, the main flow which flows directly outside, and two contra- rotating recirculation areas.

The CFD results show the quasi-totality of the inlet airflow flows directly outside without crossing through the end-windings what is harmful for cooling. Convective heat exchanges between the walls of this area and the inner air can thus be legitimately assumed as controlled by the free convection.
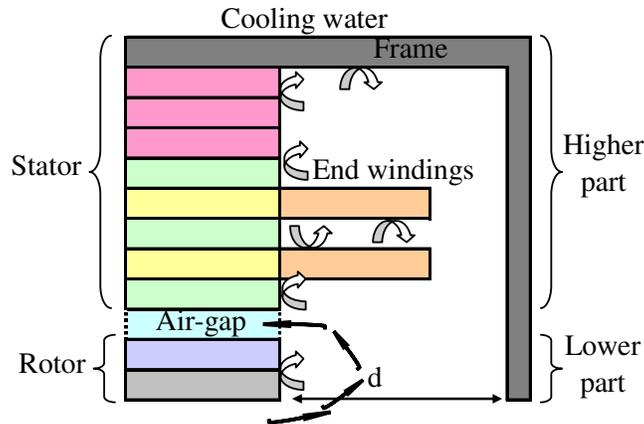

Figure 17: Convective heat transfer in the right end-region
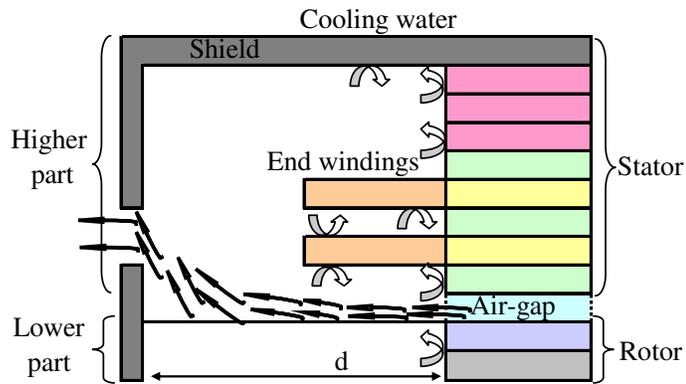

Figure 18: Convective heat transfer in the left end-region

To simplify the nodal model, the end-windings are modeled by two parallel parts, as shown on the figures 17 and 18. In the upper part of the end-regions, three kind of convective heat transfer coefficient can be identified. They are gathered in table 8. The characteristic length used to identify the convective heat transfer coefficient $h = \dfrac{\lambda Nu}{\ell}$ is the element length.

Free convection	vertical flat plate	horizontal plate heated from below	horizontal flat plate heated from above
Authors	Mc Adams [19]	Mc Adams [19]	Fishenden et Saunders [20]
Correlation	$Nu = 0.59Ra^{0.25}$	$Nu = 0.27Ra^{0.25}$	$Nu = 0.54Ra^{0.25}$
Validity	$10^4 < Ra < 10^9$	$3 \times 10^5 < Ra < 3 \times 10^{10}$	$10^5 < Ra < 10^7$

Table 8: Correlations used

The rotor end-side is considered in this work as a rotating disc placed at a distance d of a static wall. According to Kreith's results [21], for the values d/R greater than 0.2, the interaction between the rotating and the static part is very low and the Nusselt number is the same as the one obtained at the surface of a single rotating disc in still air. In our case, the ratio d/R is close to 0.5 and the Reynolds number is about 10^5 for a rotating speed of 250 rpm. According to Popiel and Boguslawski [22], this condition corresponds to a transition regime from laminar to turbulent and the convective heat transfer coefficient is obtained using the following correlation

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

$$\text{Nu} = 0.0188\,\text{Re}^{0.8}, \qquad \text{where}\ \ \text{Re} = \frac{\omega R}{\nu}$$

Thus, the convective heat transfer coefficient obtained in this case is about 24 W/m²K.

5. VALIDATION OF THE MODEL AND RESULTS
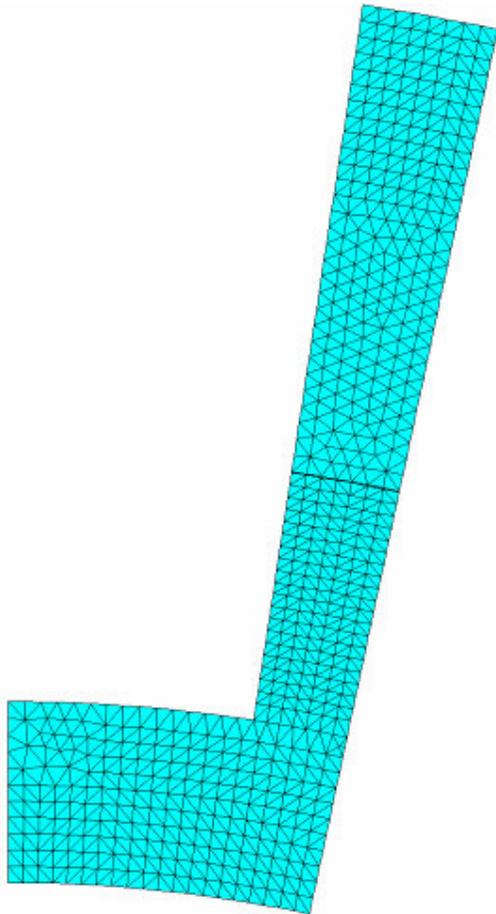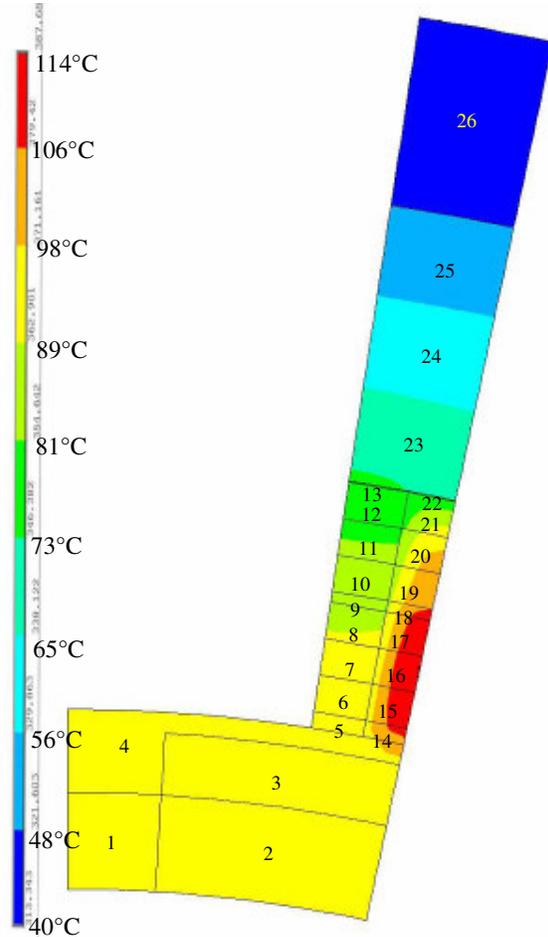### 5.1 Nodal network

Figure 19: finite element mesh   **Figure 20: Thermal cartography (Ansys) and nodal model mesh**

The validation is carried out on a 2D model made of basic elements taking into account the geometry of each part (number of poles, slots, teeth…) as shown on the figures 3-4. In the nodal model, the (thermal) network must take care of three criteria: heterogeneity of materials, heat (or sink) sources and isotherm areas. The first two criteria are easy to respect on the contrary to the last one. Ansys (the finite elements computer code) was used to identify isotherm areas (third criterion). The convergence of Ansys was obtained with a grid containing 1925 nodes and shown on the figure 19. The nodal network was build by gathering in the same area all the consecutive elements where the temperature is the same at about 2°C. This nodal network plots on the figure 20 has 1 node for the magnets, the air-gap and the frame, 2 nodes for the rim, 9 nodes for the slot and the tooth and 3 nodes for the magnetic sheet. In such conditions, only the slot and the tooth need to be refined to respect the temperature inhomogeneity inside these parts. The complete 2D thermal network has 26 nodes. The figure 21 compares the results obtained for the temperature using both models (nodal or Ansys). The plots have the same trends and no significant differences can be noticed pointing out the suitability of the developed model.
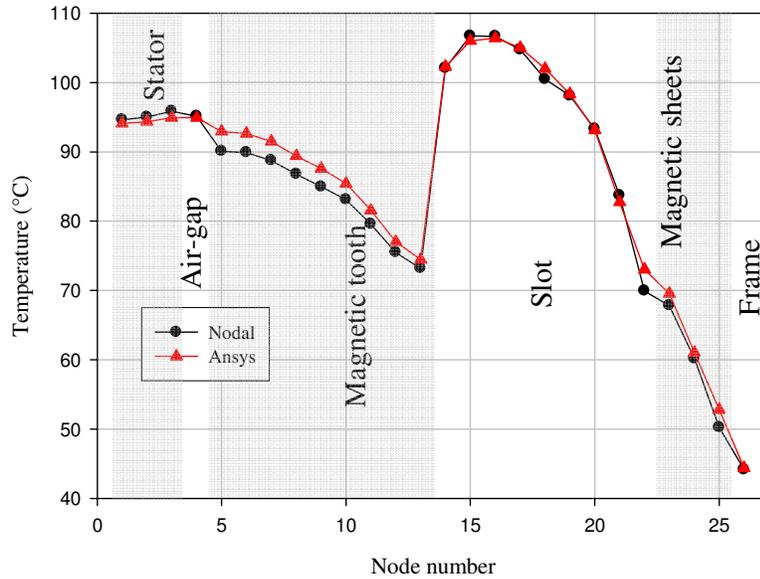
Figure 21: Comparison of the temperatures obtained using both methods (nodal and finite elements)

5.2 Results and discussion

The results presented thereafter correspond to a 3D model of the complete synchronous electrical engine. This 3D model takes into account the end-windings. In the longitudinal direction, 8 calculus sections -(r,) plane- over the entire length of the machine (2.3 m) were used. A greater number of sections was tested and does not change significantly the temperature field into the whole machine. In order to well model the end-regions of the engine, the end-windings were axially divided into three sections. Thus, a 378 thermal nodes network (14 sections x 27 nodes per section) was used. It can be noticed that by considering the whole machine (symmetries…), a 36456 calculation nodes network has to be considered, which need more CPU time but remains reasonable to model the complete machine.

Cooling water   40°C

40	40	40	45	44	44	44	44	44	44	45	40	40	40
46	46	49	56	54	53	53	53	53	54	56	48	46	46
83	75	69	66	62	61	61	61	61	63	66	71	79	84
127	113	91	75	71	69	69	69	70	71	76	98	117	126
161	153	126	79	73	72	71	71	72	74	79	129	144	154
186	173	149	100	89	86	85	86	86	89	100	140	165	177
188	177	154	112	100	96	95	96	97	100	113	146	169	180
190	180	158	118	105	101	100	101	102	106	119	150	172	182
197	189	164	121	108	104	103	103	105	109	122	161	179	189
205	194	170	126	113	109	108	108	109	114	128	163	186	198
205	194	171	129	115	111	110	110	112	116	130	164	187	198
204	193	171	128	115	112	111	111	113	117	131	164	187	198
177	171	153	122	112	110	110	111	112	116	126	157	175	185
111	110	110	110	111	112	114	116	119	121	123	125	152	162
100	103	103	103	111	118	125	130	134	136	137	127	142	150
87	101	102	102	110	117	124	129	133	136	137	128	139	145
51	59	73	87	98	108	115	121	125	127	129	128	138	144

Adiabatic (left side) — Adiabatic (right side)

Air inlet 45°C →

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

Frame   Stator windings   Rim

Magnetic sheets   End-windings   Air-gap   Magnetic tooth
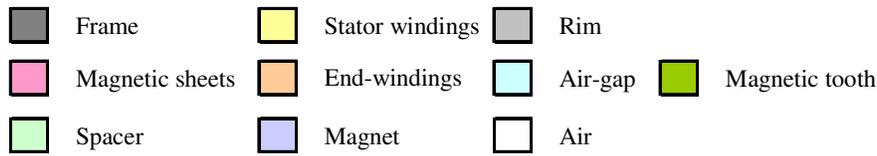
Spacer   Magnet   Air

Figure 22: Temperature field (°C)

The temperature field of the machine resulting from modeling is shown on figure 22.
Using this thermal cartography, figure 23 plots the temperature variations of the magnet, of the air-gap and of the copper. The maximum values of temperature which must not be exceeded in order to guarantee the correct engine operation are for the magnets T-magnet<90 °C and for the cooper T-copper<130 °C. With these considerations, the hottest parts of the machine are located close to the end-windings and the magnets. The airflow rate imposed into the machine is very low and does not allow a correct cooling of the end-windings. It is to note that the air temperature into the rotor increases along the axis since a part of the energy released by the rotor is transferred to the inner airflow (into the empty rotor). The same trend for the temperature of the magnets can be observed. It can be explained by the warm up of the inner air temperature into the rotor. Indeed, due to this inner air warm up, the cooling efficiency decreases and then the magnets are less cooled down at the outlet of the rotor. The temperatures observed on the frame are about 44 °C. This is due to the proximity of water at 40°C (most unfavorable case) and the high convective heat transfer coefficient between the water and the frame. In contrast with that, the rotor temperatures are far from the inflow temperature 45°C. This is due to the losses in the magnets, which heat up the air in the rotor and the low value of the airflow rate. Concerning the winding, the copper temperature remains almost constant except in the end-windings.
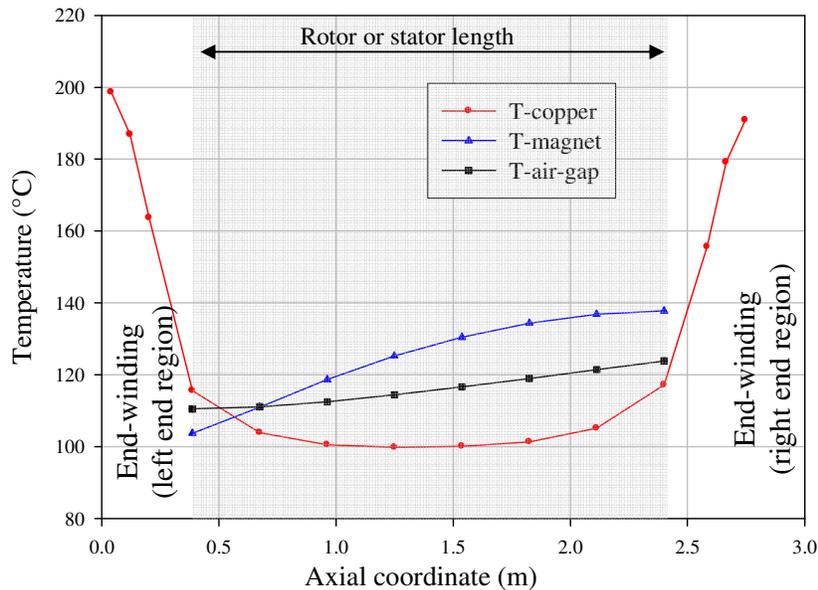

Figure 23: Temperature evolution versus the axial coordinate

Globally, the temperatures found are harmful for the correct operation of the engine, resulting in reconsidering its cooling principle. Aiming at that, a sensibility study of the temperatures of the hottest zones is necessary to determine the parameters which influence the cooling of the machine.

5.3 Study of the sensibility
An analysis of the sensibility is carried out in order to identify the parameters which influence the cooling of the machine. This analysis allows studying the temperature variation of the hot zones namely the magnets and the end-windings versus various parameters such as the airflow rate and the heat transfer coefficients in the rotor, in the air-gap and in the frame. For these parametric studies, only one quantity varies whereas the others are assumed as constants and considered at their reference value. These different quantities vary in a range where they have a physical meaning.
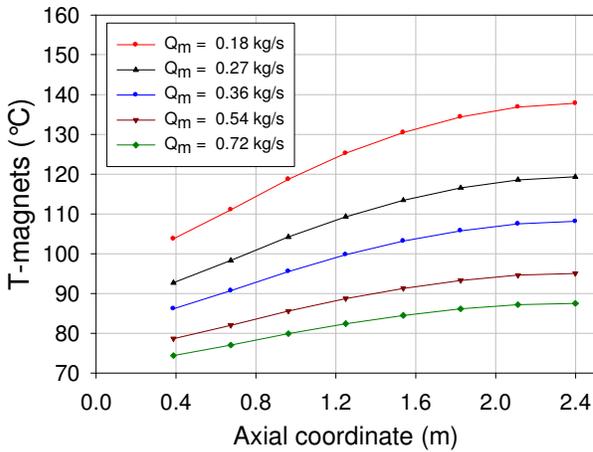
Figure 24: Airflow influence on the axial evolution of the magnet temperature.
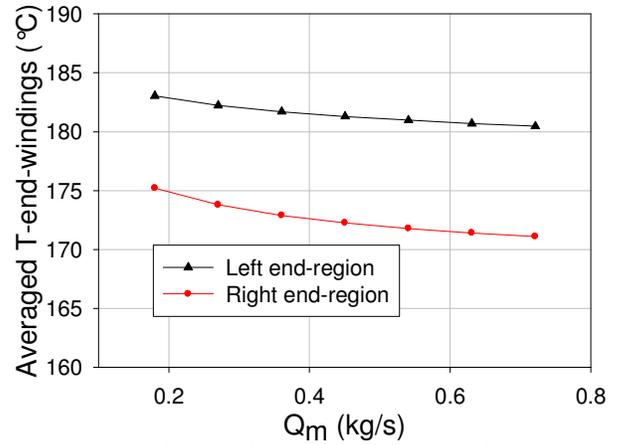

Figure 25 : Airflow influence on the end-winding temperature

The airflow rate plays an important role in the estimation of the convective heat transfer between a solid node and a fluid one, but also in the heat balance of the two consecutive air nodes. Figure 24 illustrates the effects of the air flow rate on the magnets temperature. All the plots have similar trends. The temperature increases with the axial coordinate. The areas of high sensibility are in the regions far from the rotor inlet. The magnet temperature decreases when the airflow rate augments. For instance, an augmentation of the airflow rate from 0.18 to 0.72 kg/s entails a magnet temperature reduction ranging between 30 (inlet) and 50 °C (outlet). Globally, the magnets temperature is very sensitive to the airflow rate variation. To ensure a correct cooling of the magnets, an airflow rate equals to 0.72 kg/s would be more appropriate. Figure 25 presents the evolution of the end-windings temperature for both side of the machine as a function of the airflow rate. These temperatures are weakly sensitive to the airflow variation. It can be explain using the air velocity vectors field which shows the end-windings are not crossed by the airflow. Then the transfer in the upper regions of the end-regions is only controlled by free convection. For instance, an augmentation of the airflow rate from 0.18 to 0.72 kg/s entails a temperature drop equals to about 4 °C.
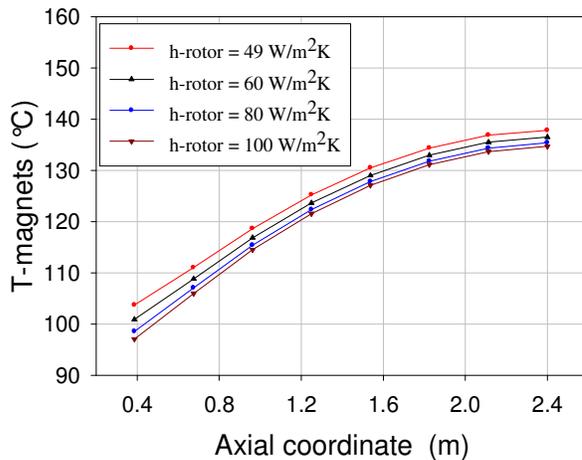

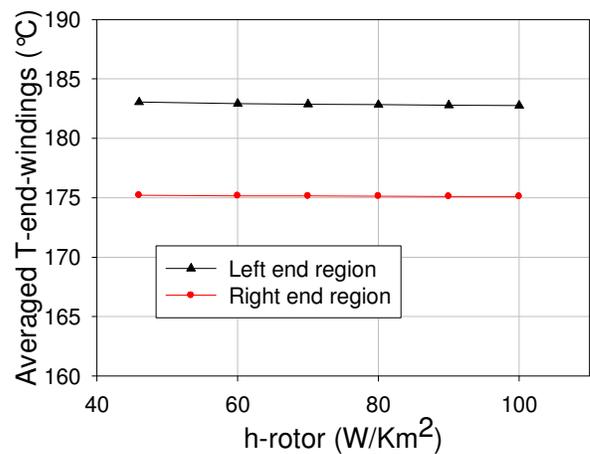**Figure 26 : Evolution of the magnet temperature with h-rotor**

Figure 27 : Evolution of the end-winding temperature with h-rotor

Figure 26 presents the axial evolution of the magnet temperature versus the convective heat transfer coefficient into the rotor: h-rotor. Whatever the values of this coefficient, the magnet temperature increases when the air flows inside the rotor. These values are lower when h-rotor increases. This decrease remains all the same limited due to the significant temperature of the air in the rotor. In such a situation, the increase of h-rotor is not high enough to lower the magnet temperature in a suitable way. The same conclusions can be drawn for the end-windings temperatures which are weakly sensitive to the h-rotor variation, as shown on the figure 27. The important increase in the heat

transfer coefficient (from 45 to 100 W/m²K) only reduces to about 1 °C the end-windings averaged temperature.
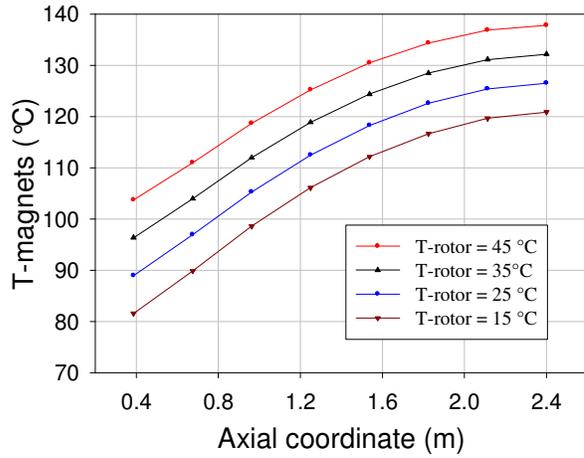

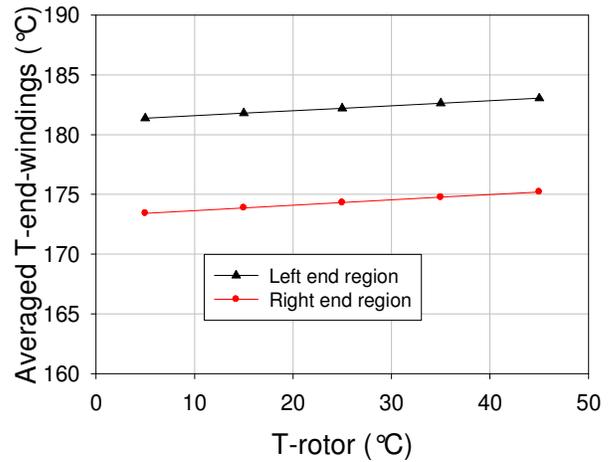Figure 28: Evolution of the magnet temperature with T-rotor

Figure 29: Evolution of the end-winding temperature with T-rotor

Figure 28 represents the axial evolution of the magnet temperature with the inlet temperature of air: T-air_inlet. For the values considered, similar trends can be noted, but the magnet temperature is quite different. The lower the air inlet temperature is, the lower the magnet temperature. Even if a certain improvement can be noted, the reduction in the inlet temperature of air cannot ensure a suitable cooling of the magnets since they have not to exceed 90 °C. The evolution of the end-windings temperature with T-air_inlet is plotted on the figure 29. The inlet temperature of air has a very weak influence on the end-winding temperature since this parameter induces a temperature drop which does not exceed 2 °C.
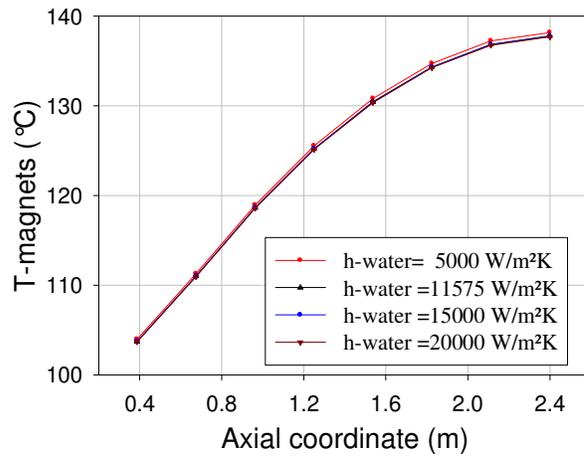

Figure 30: Evolution of the magnet temperature with h-water
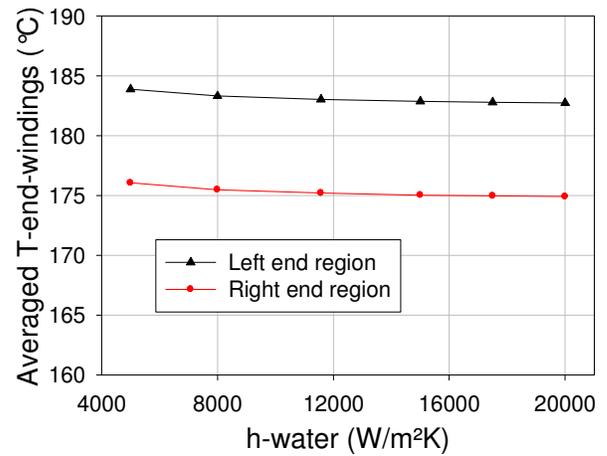
Figure 31: Evolution of the end-winding temperature with h-water

The variation of the heat transfer coefficient is caused by the speed variation of ship in the case of the POD version and the speed of cooling water in the case of an on-board motor. The influence of h-water on the magnet temperature is plotted on the figure 30. It clearly appears that the magnet temperature is not dependant to this coefficient variation since the plots of the temperature evolution along the axis are almost identical. On the figure 31, it is also found that the end-winding temperature is not dependent of h-water. Indeed, the temperature remains close to 183 °C for the end-windings in the left end-region and of 175 °C for the other one. For both cases, the temperature reduction does not exceed 2 °C. This previous analysis lets identify the parameters which influence the thermal behavior of the machine. From this study, it can be concluded that only the airflow rate plays an

important role on the cooling of the magnets whereas the end-windings are not dependant to this airflow.

6. COOLING TECHNIQUES OR SOLUTIONS

This part aims to analyze the efficiency of several solutions to enhance the cooling. To decrease the magnet temperature, it would be suitable to improve the heat transfer between the rotor and the air. Aiming at that, three solutions are studied. The first one implies to increase the airflow rate. This solution requires the installation of more powerful and/or more efficient fans. This solution is not acceptable for economic considerations as well as noise, weight and cumbersome issues. The second solution consists in increasing the convective conductance (h S) using a slight airflow rate augmentation and the use of fins into the rotor. The third and last solution consists in introduce inside the rotor a pipe in which flows cooling water. This pipe plays the role of a heat exchanger between the internal air of the rotor and the cooling agent (water). Concerning the end-windings, none of the tested solutions improve the cooling. Another solution consists in install a thermal bridge between the end-windings and the frame, letting dissipate a part of the end-winding energy to the external water (through the thermal bridge and then through the frame).
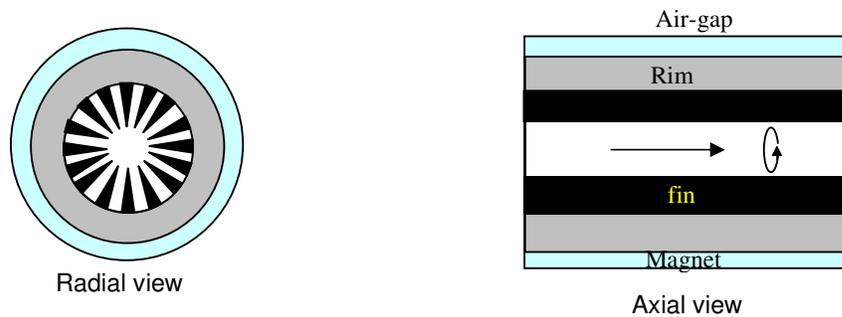
6.1 Rotor with longitudinal fins

Figure 32: Rotor with triangular longitudinal fins

The heat exchange from the rotor to the inner air can be enhanced by increasing the rim surface using fins [23]. This solution is presented on the figure 32. In the proposed configuration, the rotor has longitudinal fins. These fins are parallels to the direction of the airflow rate. The following heat balance can then be written:

$$\phi = hS_r(T_p - T_\infty) + \eta hS_m(T_p - T_\infty) = (G_r + G_a)(T_p - T_\infty)$$

In this last equation, S_r is the contact surface between the rotor and the air and S_m referred to the wetted surface of the fin. The part of the rotor with fin, presented on the figure 33, can be thermally modeled as two conductances (series association) G_a and G_r. The first one refers to the fin contribution and the other one to the rotor without fin contribution (figure 34).
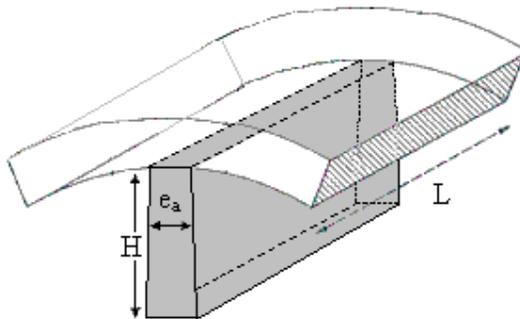
Figure 33: Piece of the rotor with a fin
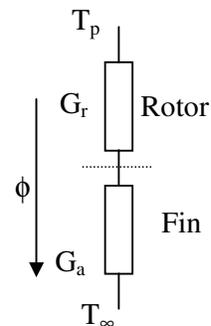
Figure 34: Equivalent scheme

The role of the fins is to dissipate the (thermal) energy of the rotor as well as ensuring a quite good compromise between its thermal and aeraulic performance, its mass and its cumbersome. G_a is related to both the geometrical characteristics and the thermal fin conductivity. It depends on the three following quantities: the thermal conductivity of the fin ( ), its thickness (e_a) and its width H. The

heat transfer coefficient h_i is assumed as constant. For considerations connected with the fin installation, 30 fins were used (one fin per magnet). Taking into account rotor dimensions and this fin number, the sizes of the fins are limited by $e_a < \dfrac{2\pi R_i}{30} = 0.0869$ m and $H < R_i = 0.415$ m. To determine the optimal fin profile, in term of magnet cooling, two kinds of fins were analyzed: rectangular crossed section fins and triangular ones. The thermal equivalent scheme, the conductance and the efficiency of each considered fin are gathered in the table 9.

fin profile	Rectangular fin	Triangular fin
Scheme	e_a — H	e_a α — H
Conductance	$G_{rc} = \eta_{rc} h L (e_a + 2H)$	$G_{tr} = \eta_{tr} h \dfrac{2HL}{\cos(\alpha)}$
Efficiency [24]	$\eta_{rc} = \dfrac{e_a}{\delta(e_a + 2L)} \dfrac{(1+\delta)\exp(\chi H) - (1-\delta)\exp(-\chi H)}{(1+\delta)\exp(\chi H) + (1-\delta)\exp(-\chi H)}$	$\eta_{tr} = \dfrac{1}{\tau\sqrt{H}} \dfrac{I_1(2\tau\sqrt{H})}{I_0(2\tau\sqrt{H})}$
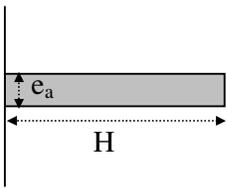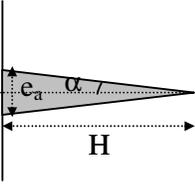
with $\qquad \delta = \sqrt{\dfrac{he_a}{2\lambda}}, \qquad \chi = \sqrt{\dfrac{2h}{\lambda e_a}} \qquad$ et $\qquad \tau = \sqrt{\dfrac{2Hh}{e_a \lambda \cos(\alpha)}}$

Table 9: fin profile

Figure 35 plots the variation of the averaged magnet temperature versus the fins mass for the two considered fin profiles. The plots have similar decreasing trends and let conclude that for an imposed fin mass, the triangular fins are more efficient in terms of cooling than the rectangular ones.
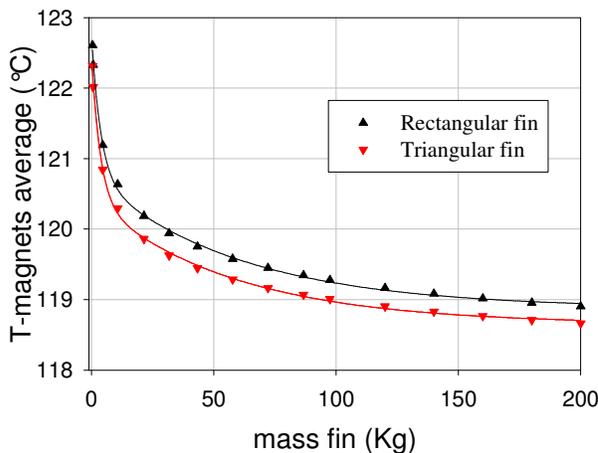
Figure 35: Evolution of the magnet temperature with the mass of the fin
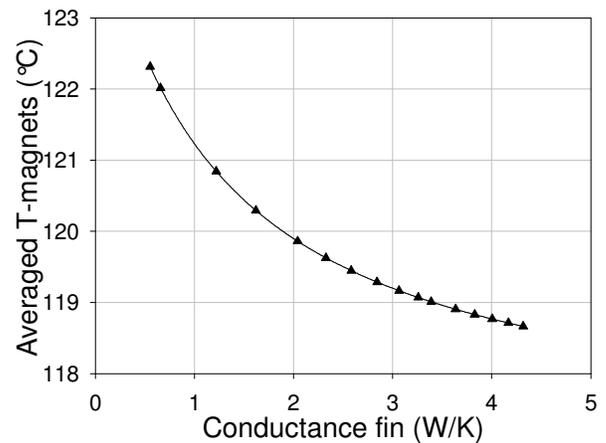
Figure 36: Evolution of the magnet temperature with the conductance of the fin

Figure 36 plots the variation of the averaged magnet temperature versus the conductance G_a. This figure shows that the fin presence entails a reducing in the magnet temperature. The greater the fin conductance is, the lower the magnet temperature.

Aiming at reduce the technical problems involved by the manufacture and the installation of the fins (economic and cumbersome considerations), minimal size of the fins ensuring the suitable magnet cooling is seek. To identify the optimal mass value which entails a maximal heat exchange, the reduced sensibility coefficient S_m was investigated. This coefficient is defined as:

$$S_m = m \left| \frac{\partial \mathrm{Taimant}}{\partial m} \right|$$
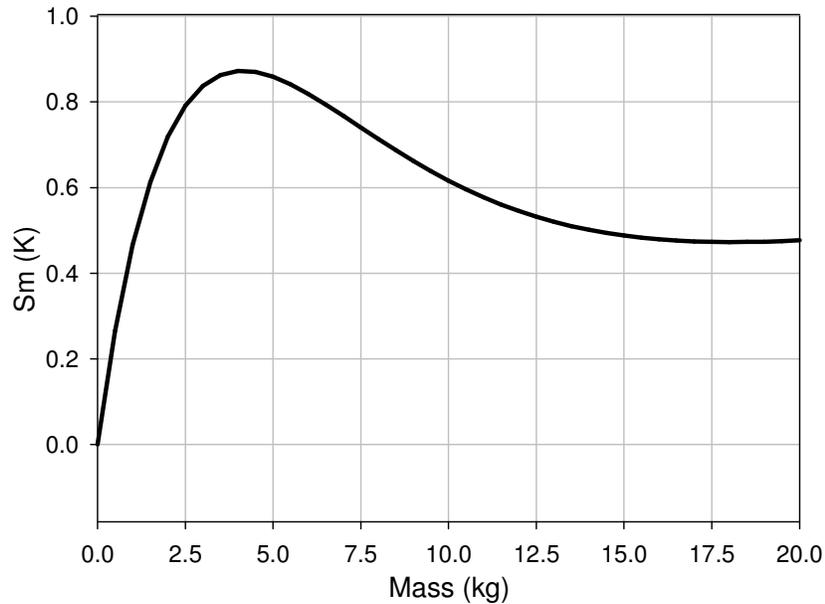
Figure 37: Reduced sensibility of T-magnets to the mass of the fin

Figure 37 plots the evolution of S_m versus the mass of the fin. This reduced sensibility is maximal for a fin mass m equals to about 4 kg. The correlations of Vriendt [24] let obtain the optimal pair (e_a, H) for a given mass. These 2 correlations are written below:

$$\frac{H}{e_a} = 0.6547 \sqrt{\frac{2\lambda}{he_a}} \tag{8}$$

$$m = 0.5 \times \rho \times H \times e_a \times L = 4.1\,\mathrm{kg} \tag{9}$$

Using these two correlations, the optimal pair (e_a, H) is obtained. The retained configuration is a rotor with 30 steel triangular fins, 3 mm thick and 75 mm width.

The axial temperature evolution of the magnets for the different airflow rates and fin configurations (with or without) is plotted on the figure 38. The reference case namely smooth rotor with an axial airflow rate equals to Q_m=0.18 kg/s, the case where the airflow rate inside a smooth rotor is 0.72 kg/s and a case where the airflow rate inside a rotor with fins is 0.59 kg/s are studied. Even if the fins do not entail a real enhancement, they improve the magnet cooling for reasonable airflow rate: Q_m =0.59 kg/s.
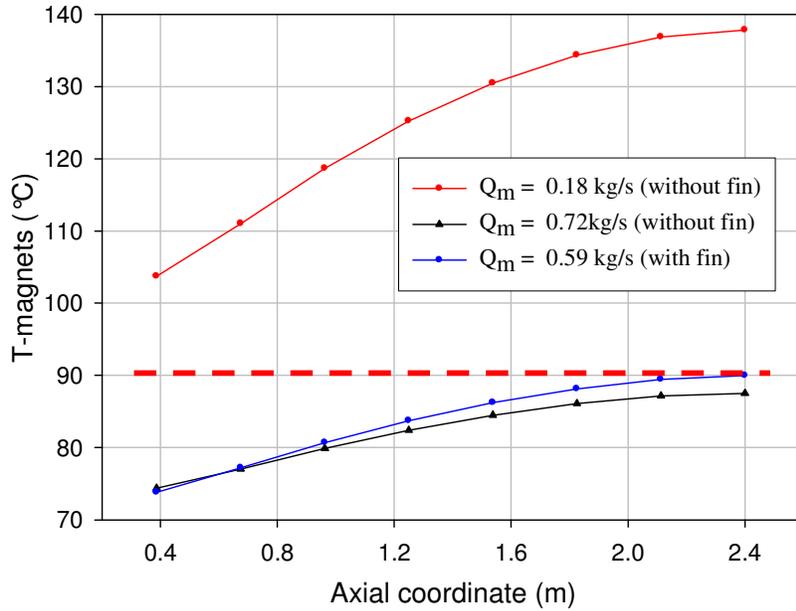
Figure 38: Axial evolution of the magnet temperature
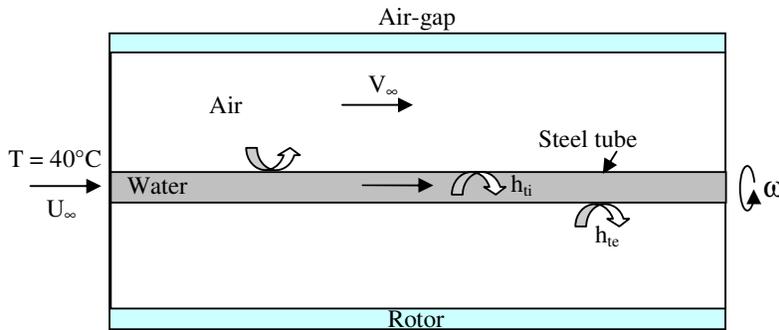
Figure 39: Rotor with a static water cooling tube

6.2 heat exchanger
The magnet temperature lowering can also be obtained by enhancing the inner air temperature flowing into the rotor. An auxiliary empty pipe placed into the rotor and acting as a heat exchanger (with for instance water as cooling agent) can ensure this function. The axial airflow into the rotor is then cooled down. The chosen pipe is a 2.3 m length pipe corresponding to the iron length of the machine and is made of steel. It is 5 mm thick, and its inner diameter is 100 mm (fig. 39). This thickness lets obtain both the suitable thermal efficiency and the correct mechanical resistance. The figure 40 presents the equivalent thermal network suitable for such a configuration. The nodal resolution needs the identification of the convective heat transfer coefficients h_{ti} and h_{te}. h_{ti} corresponds to the convective heat transfer coefficient between the wall and the cooling agent flowing into the pipe (water) and h_{te} is the one between the wall and the inner air into the rotor.

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy
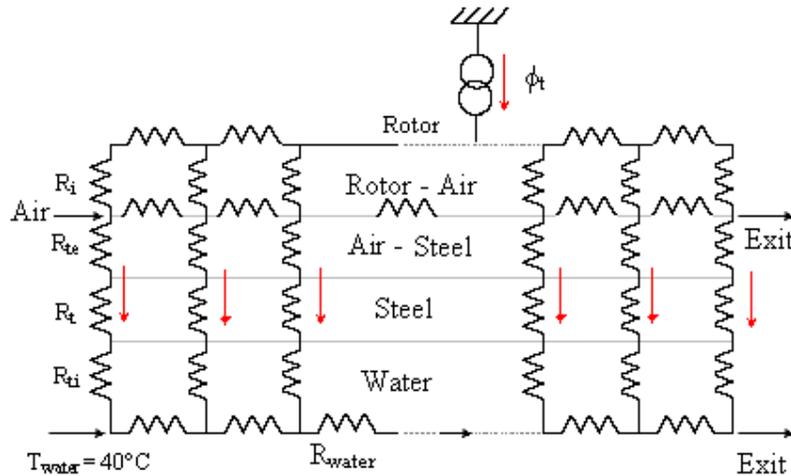
Figure 40: Equivalent scheme.

The convective heat transfer coefficient between the water and the internal surface of the pipe is calculated using the following correlation of Mc-Adams [19] adapted to the forced convection in turbulent pipe flow. This correlation takes into account the geometry via the ratio L/D_t.

$$\mathrm{Nu} = 0.023\,\mathrm{Pr}^{0.33}\left(1+\left(\frac{D_t}{L}\right)^{0.7}\right)\mathrm{Re}^{0.8} \quad \text{and} \quad \mathrm{Re} = \frac{D_t U_\infty}{\nu_{water}}$$

This correlation remains valid for a Prandtl number ranging between 0.7 and 100, an axial Reynolds number ranging from 10^4 and 1.2×10^5. Using this relation in the studied case, the heat transfer coefficient is:

$$h_{ti} = \frac{\lambda_{water}\mathrm{Nu}}{D_t} = 23511\,\mathrm{W/m^2K}$$

The flow outside the pipe results from the coupling of the air driven into rotation by the rim of the rotor (viscous effects) and the axial airflow into the rotor. It is difficult to identify the convective heat transfer coefficient, since no correlations corresponding to this specific case have been found. However, two cases of heat transfer are relatively similar to this problem. These two cases are presented thereafter and concern the case of the free convection around a horizontal tube and the case of the rotating pipe in still air.

Free convection around a horizontal tube
As the rotor diameter is considerable (Di = 830 mm) and the airflow velocity very low, V = 0.3 m/s, the convective transfer can be assumed in a first approach as controlled by laws of free convection. It thermally represents the worse case. The Nusselt number is then estimated using the following Mc Adams correlation [19] based on the Rayleigh number:

$$\mathrm{Nu} = 0.53\mathrm{Ra}^{0.25} \quad \text{for } 10^4 < \mathrm{Ra} < 10^9$$

To evaluate Nu and Ra the reference geometrical quantity is the pipe diameter. Then $\mathrm{Ra} = 9.4\times10^5$ and $h_{te} = 4.3\,\mathrm{W/m^2K}$. Other thermo-physic quantities are taken at the film temperature, *i.e.* $(T_p+T)/2$. T is the averaged air temperature into the rotor and is obtained using an iterating process and the previous model.

Rotating pipe into still air
In this second case, the air driven into rotation by the rotor rim controls the convective heat transfer. Many correlations exist in the case of rotating pipe in still air. If it is assumed that the heat transfers are not too different in both cases, value for the convective heat transfer coefficient can be obtained. For instance, the Becker correlation [25] presented thereafter lets obtain a value for this quantity. This correlation can be written as:

$$Nu = 0.133\,\mathrm{Re}^{0.66}\,\mathrm{Pr}^{0.33} \quad \text{with} \quad 800 < \mathrm{Re} = \frac{\omega D^2}{2\nu} < 10^5$$

Using the physical properties of the air taken at the film temperature $(T_P+T)/2$, the Reynolds number is equal to $\mathrm{Re}_r = \dfrac{\omega D^2}{2\nu} = 7.4 \times 10^3$, and then $h_{te} = 11.5 \ \mathrm{W/m^2K}$.
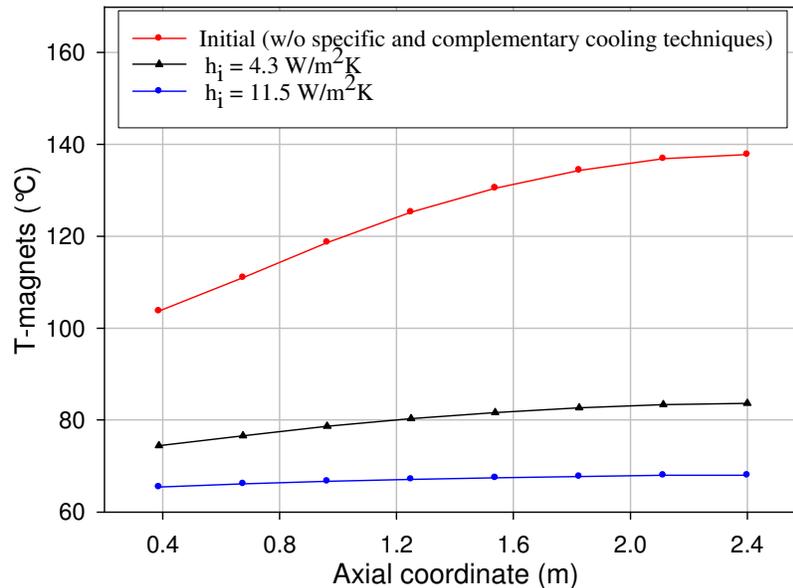
Figure 41: Axial evolution of the magnet temperature.

The results plotted on the figure 41 correspond to the axial evolution of the magnet temperature for both the configurations (with and without pipe). For the case with the pipe, the magnet temperature is evaluated using h_{ti} calculated with the both assumptions made previously. The presence of the pipe enhances considerably the heat transfer resulting in a magnet temperature decreasing. The losses in the rotor are transferred to the internal airflow and then to water flowing into the tube. It involves a considerable enhancement of the heat transfer between the rotor and the air. The results obtained with the model show that the use of this kind of technique lowers the magnet temperature of 44 °C if h_{it} = 4.3 W/m² K (first assumption) and 57 °C if h_{it} = 11.5 W/m²K (second assumption). This technique is very interesting in a POD configuration since it does not require the use of an additional heat exchanger to cool down the water flowing into the pipe (direct use of the sea water).
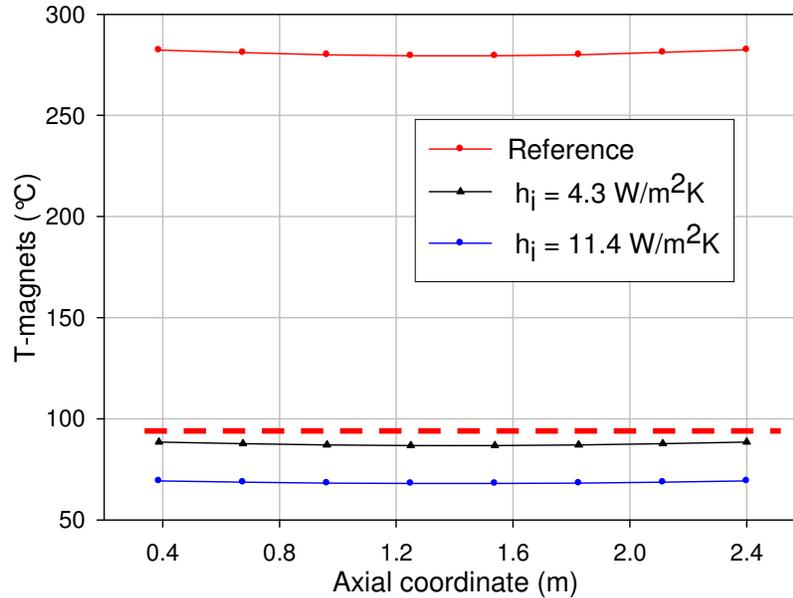
Figure 42: Axial evolution of the magnet temperature without mass flow rate

Among the machine architectures interesting Jeumont S.A-Areva, the case of the self-ventilated machines can be quoted. They are perfectly waterproof and do not require any external forced ventilation. This kind of architecture is often used in the high rotational speed machines. However, it can be interesting to analyze the influence of no airflow rate on the thermal behavior of the studied machine even if it is driven at lower rotation speed (=250 rpm). Figure 42 shows the axial evolution of the magnet temperature in a configuration without inner airflow. The plot shows the cooling is highly improved when compared to the reference configuration. With this technique and despite the airflow withdrawal, the temperatures are lowered and remains within the thermal operating conditions. This auxiliary tube is thus an efficient solution to cool down this part.

6.3 Thermal bridge
The thermal behavior of the end-windings is quite weakly influenced by the inner aeraulics. To ensure a suitable cooling, the frame can be both connected to these end-windings and to the water using a copper thermal bridge as shown on the figure 43. The correct thermal connection between the end-windings and the thermal bridge is ensured using resin layers which is an electrical insulation but not a thermal one (= 5 W/mK). The thickness of the resin layers ranges between 0.3 and 10 mm. Using this solution, the quasi-motionless air playing an important thermal insulation role (= 0.025 W/mK) is replace by a material of greater thermal conductivity. The heat transfer from the end-windings towards the frame (through the thermal bridge) and then the cooling water is then enhanced. The figure 43 also permits to compare the end-windings temperature in both cases, *i.e.* with and without the thermal bridge. It is clearly shown the bridge enhances the cooling. Indeed, with the thermal bridge, the heat transfer between the end-windings and the frame is improved since, for the end-windings, a reduction of 34 °C in the average temperature is observed for the case of a crown type thermal bridge and a diminution of 95 °C for the case of a cylindrical type thermal bridge. The ring type bridge case is more efficient in the cooling of the upper part of the end-windings since a decrease of the average temperature equals to 49 °C can be noticed. However, the lower part of the end-windings is not significantly cooling and the temperature diminution in this part is lower than 18 °C. Even if the heat transfer is enhanced, hot spots mainly located in the lower part exist. This remark shows the thermal bridge must still be improved in order to dissipate more heat. Choice of the cylindrical type bridges ended with a copper ring and covering with resin the whole end-windings is better. Indeed, in this case, the maximal end-windings temperature is T_{max} = 96 °C. No hot spots were observed. Due to this thermal bridge, the heat transfer from the end-windings towards the frame is intensified. The thermal energy of the end-windings is then evacuated outside and a significant temperature lowering is to note.

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

| Without thermal bridge | With ring type thermal bridge | With cylindrical type thermal bridge |
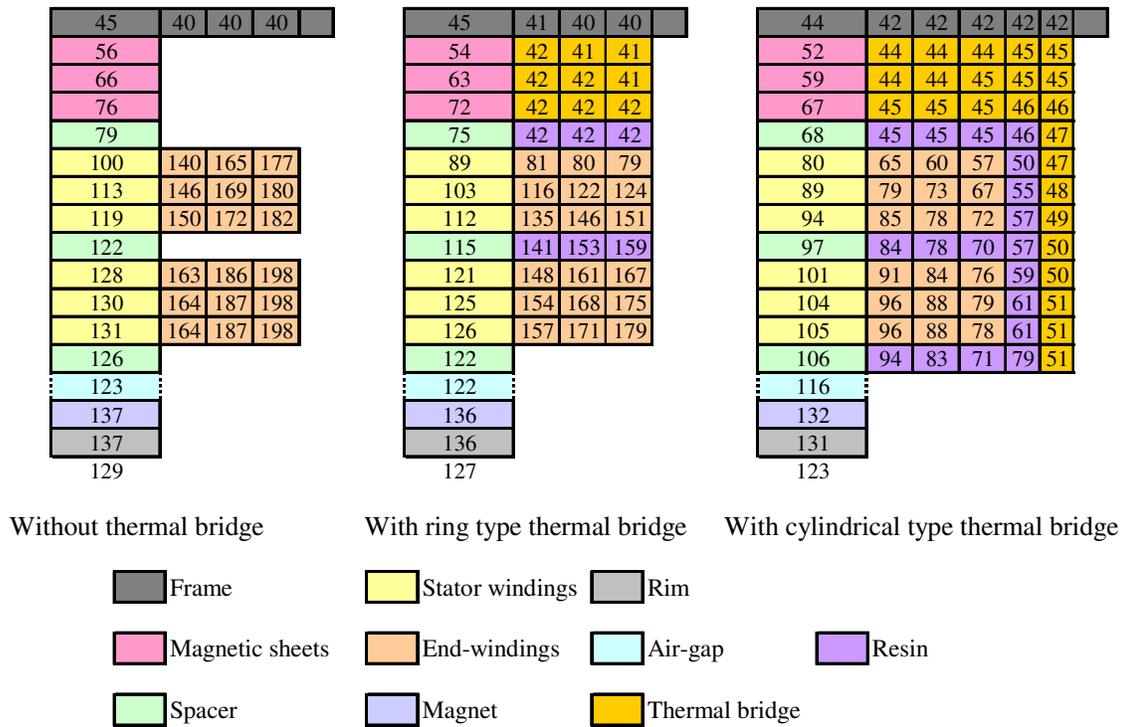
Figure 43: Thermal behavior of the end-windings (Temperatures in °C)

7. CONCLUSION

This work deals with a thermal study of a permanent magnets synchronous motor. A nodal approach was developed in order to simulate the thermal behavior of this machine. This model permits to better understand the thermal behavior, to limit the temperatures of the different components, and to predict the thermal response of the engine in extreme operating conditions. It also permits to design and develop new cooling systems. The results show that the hottest parts are located in the magnets and end-windings where the temperatures exceed the operational limits. A particular attention was focused on these parts in order to analyze their thermal behavior with the sensible parameters variation. These sensibility studies lets propose technical cooling solutions able to ensure optimized temperature values in the whole machine. Several cooling solutions were investigated. In order to decrease in a significant way the magnet temperature, the interior of the rotor can be equipped with longitudinal fins. The use of an auxiliary pipe into the rotor operating as a heat exchanger is interesting to maintain the magnet temperature far from the critical temperature. The utilization of thermal bridges for the end-winding cooling lets improve the heat transfers and results in the decrease of the end-winding temperature. Implementation of these cooling solutions limits the temperatures into the whole machine to values lower than expected.

Acknowledgement

References

[1]   W.W. Soroka: Analog methods in computation and simulation. Ed. McGraw Hill, p: 390 (1954).

[2]   J.B. Saulnier : La modélisation thermique et applications aux transferts couplés et au contrôle actif. Thèse de l'université de Poitiers (1980).

[3]   J.B. Saulnier, A. Alexandre : La modélisation thermique par la méthode nodale : Ses principes, ses succès et ses limites. Rev. Gén. Therm. Fr, 280, 363-372 (1985).

[4]   Y. Bertin : Analyse des transferts de chaleur dans un moteur électrique asynchrone. Développement d'un environnement CAO et modélisations thermoaréauliques. Thèse de l'université de Poitiers (1987).

Souhil Seghir-OualiI, Souad Harmand & Daniel Laloy

[5]   C. Vasilescu : Modélisation du transfert de chaleur au sein des machines électriques tournantes- dimensionnement et optimisation de leur système de refroidissement. Thèse de l'université de Paris VI (2002).

[6]   F. Tachibana, S. Fukui : Convective heat transfer of the rotational an axial flow between two concentric cylinders. Bulletin of JSME 7 (26) 385-391 (1964).

[7]   P.R.N. Childs, A.B. Turner : Heat transfer on the surface of a cylinder rotating in an annulus at high axial and rotational Reynolds numbers. Congres Brighton (1994).

[8]   M. Bouafia, Y. Bertin, J. B. Saulnier : Analyse expérimentale des transferts de chaleur en espace annulaire étroit et rainuré avec cylindre intérieur tournant. Int. J. Heat Mass Transfer, N°10, p : 1279-1291 (1998).

[9]   I. Peres, Y. Bertin : Caractérisation des échanges de chaleur dans un espace annulaire lisse et rainuré de type entrefer moteur électrique. Société française des thermiciens (1994).

[10]   M. Bouafia, A. Ziouchi, Y. Bertin, J. B. Saulnier : Étude expérimentale et numérique des transferts de chaleur en espace annulaire sans débit axial et avec cylindre intérieur tournant. Int. J. Therm. Sci. n°38, p: 547-559 (1999).

[11]   J. Gosse : Guide technique de thermique. Dunod Paris, p :168-169 (1981).

[12]   G. Reich: Strömung und wärmeübertragung in einem axial rotierenden Rohr, Doctoral Thesis, TH Darmstadt (1988).

[13]   G. Reich, B. Weigand, H. Beer: Fluid flow and heat transfer in an axially rotating pipe-II. Effect of rotation on laminar pipe flow, Int. J. Heat Mass Tran. 32 (3) 563-574 (1989).

[14]   J.N. Cannon, W.M. Kays: Heat transfer to a fluid flowing inside a pipe rotating about its longitudinal axis, J. Heat Transfer. 91 135-139 (1969).

[15]   G. Reich, H. Beer: Fluid flow and heat transfer in an axially rotating pipe-I. Effect of rotation on turbulent pipe flow, Int. J. Heat Mass Tran. 32 (3) 551-562 (1989).

[16]   B. Weigand, H. Beer: Fluid flow and heat transfer in an axially rotating pipe subjected to external convection, Int. J. Heat Mass Tran. 35 (7) 1803-1809 (1992).

[17]   S. Seghir-Ouali, D. Saury, S. Harmand, O. Phillipart, D. Laloy: Convective heat transfer inside a rotating cylinder with axial air flow. Int. J. Therm. Sci. (in press).

[18]   Fluent v.6.0, User's guide (2001).

[19]   W.H Mc Adams: Transmission de la chaleur. Dunod Paris (1961).

[20]   M. Fishenden, O.A Saunders: An introduction to heat transfer. Oxford (1950).

[21]   F. Kreith: Convection heat transfer in rotating systems. Vol. 21, Ed. Irvine T.F et Hartnett J.P (1968).

[22]   C.Z.O Popiel, L. Boguslawski : Local heat transfer coefficients on the rotating disk in still air. Int. J. Heat Mass Transfer, Vol 18, p : 167-170 (1975).

[23]   A. Bontemps : Échangeur de chaleur, Description des échangeurs. Techniques de l'ingénieur, B2 341 (1995).

[24]   André B. De Vriendt : La transmission de la chaleur volume I, tome 1. Gaëtan Morin éditeur (1984).

[25]   Kurt M. Becker: Measurement of convective heat transfer from a horizontal cylinder rotating in a tank of water. Int. J. Heat Mass Transfer, Vol 6, p: 1053-1062 (1963).

Finite Element Investigation of Hybrid and Conventional Knee Implants

Habiba Bougherara   habiba.bougherara@ryerson.ca
Department of Mechanical and
Industrial Engineering
Ryerson University
Toronto, M5B 2K1, Canada

Ziauddin Mahboob   zmahboob@ryerson.ca
Department of Aerospace Engineering
Ryerson University
Toronto, M5B 2K1, Canada

Milan Miric   mmiric@ryerson.ca
Department of Mechanical and
Industrial Engineering
Ryerson University
Toronto, M5B 2K1, Canada

Mohamad Youssef   m3yousse@ryerson.ca
Department of Mechanical and
Industrial Engineering
Ryerson University
Toronto, M5B 2K1, Canada

Abstract

Total Knee arthroplasty (TKA) procedures relieve arthritic pain and restore joint function by replacing the contact surfaces of the knee joint. These procedures are often performed following arthritic degeneration of the joint causing the patient pain. Cobalt-chrome, stainless steel (316L grade) and titanium alloys are widely used in the majority of distal femoral implants in TKA procedures. The use of such stiff materials causes stress shielding (i.e. a lack of mechanical stresses being experienced by the bone surrounding the implant) leading to gradual bone loss and implant failure. The aim of this paper is to develop a new hybrid knee implant which combines a polymer-composite (CF/PA-12) with an existing commercial implant system (P.F.C.® Sigma™) made from stainless steel. This hybrid implant is expected to alleviate stress shielding and bone loss by transferring much more load to the femur compared to conventional metallic implants. Results of the FEA simulations showed that the CF/PA-12 lined femoral component generated almost 63% less in peak stress compared to the regular stainless steel component, indicating more load transfer to the bone and consequently alleviating bone resorption.

1. INTRODUCTION

Implant designs for TKA procedures replace the arthritic surfaces of the knee joint, resurfacing the femoral trochlear groove and both lateral and medial condyles. Polyethylene spacers are attached replacing the tibial articular surface through a plate of porous-finish metal, acting as the contact point for the femoral condyles. In procedures where total condylar prostheses are used the patella is also resurfaced, however this does not affect the performance of the femoral implant [1, 2].

Cobalt-chrome, stainless steel, and titanium alloys are used in the majority of distal femoral implants in TKA procedures [3-5]. In addition trabecular metal, a tantalum based biomaterial (80% porosity) with a crystalline microstructure similar to that of trabecular bone (cancellous bone) is used for contact surfaces requiring direct bone apposition [6]. With regards to the tibial plates specifically, porous metal (trabecular metal) allows for bone ingrowth and implant stability [6]. In addition this porosity encourages soft tissue growth and supports vascularisation of adjacent tissue.

Polyethylene spacers attached to the tibial plate are allowed to rotate on the tibial plate reducing abnormal wear and fatigue compared to fixed articular surface implants. Regardless of these advancements, wear of polyethylene is the limiting factor in long term performance of the implant and the cause of the majority of the 37,544 revision surgeries performed in 2005 [3]. The development of polyethylene has progressed since the 1970s to the ultra high molecular weight polyethylene (UHMWPE) currently used [7].

The use of high strength, high stiffness titanium and stainless steel alloys in distal femoral implants causes stress absorption in the implant shielding the femur from physiologic stress and loading. As such, abnormal stress/loading patterns develop along the shaft of the femur leading to bone degeneration. In addition, the titanium and stainless steel alloys used offer poor osseoinduction and biocompatibility [8].

In this study, the femoral component of the implant is the focus in improving the design. The existing designs for polyethylene spacers, tibial plates and patella implants are not modified. Thus, the intent of this study was to develop a femoral component that would promote natural stress distribution within the femur and exhibit osseoinductive properties.

Previous studies of orthopaedic implants by one of the current authors [9] used a carbon fibre based polymer matrix composite to resurface hip joints. The implant designed used a composite to line the contact surfaces of a stainless steel implant reducing stress shielding and promoting physiologic loading. Carbon fibre based composites have been shown to provide significantly better osseoinduction than titanium. The CF/PA-12 (Carbon fibre reinforced polyamide 12) composite developed by one of the current authors in previous studies demonstrated excellent fatigue life under loading exceeding by several orders of magnitude that of the knee under natural gait [9]. In designing a biomimetic distal femoral implant for TKA a similar approach was adopted using a carbon fibre based composite to line the bone apposition surfaces of the implant.

a   b
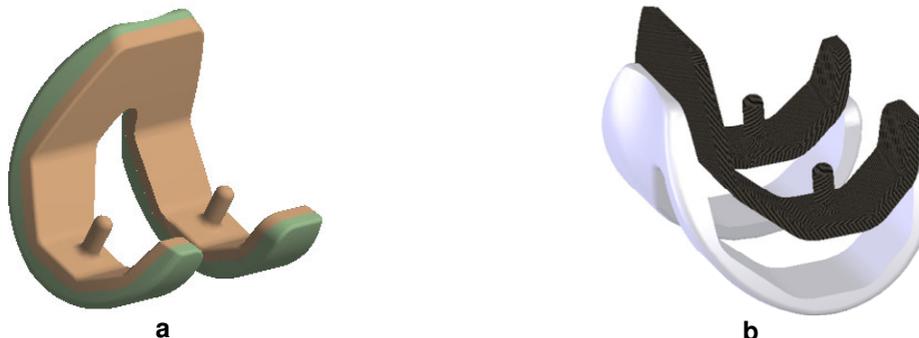
FIGURE 1: CAD depiction of hybrid CF/PA-12 lining on femoral component; a) Assembled view, b) Exploded view

Analysis of the performance of the implant was conducted using finite element analysis and compared to a 316L grade stainless steel implant of the same geometry. The validity of the results of this design concept was assessed based on the magnitude of stress in the tibial component's UHMWPE layer.

2. GENERATING THE MODEL

CAD model of femur and tibia

Computed tomography (CT) scans of a composite fourth generation femur and tibia supplied by Sawbones Worldwide (Items #3406 and #3402 respectively, Pacific Research Laboratories, Vashon, WA, USA) were performed at intervals of 0.5 mm along the length of the bones [10, 11]. FIGURE 2 and FIGURE 3 display the original femur and tibia models, respectively. The dimensions are given in TABLE 1. Using MIMICS® Medical Imaging Software (The Materialise Group, Leuven, Belgium) the cross sectional geometries of both the femur and tibia were exported into SolidWorks 2007 (Dassault Systèmes SolidWorks Corp, Concord, MA, USA) as independent files. With the aid of 'SPLINES' in SolidWorks 2007, the 'LOFT' function was used to generate the CAD geometry of both bones. Care was taken to maintain both cancellous and cortical bone geometries within both the femur and tibia.
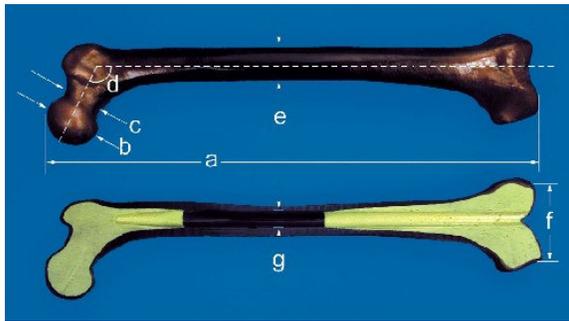
FIGURE 2: Fourth generation left femur [10]　　　**FIGURE 3:** Fourth generation left tibia [11]
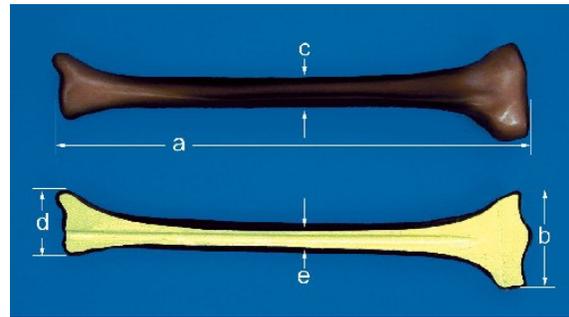
Dimensions	Femur [mm]	Tibia [mm]
a	485	405
b	52	84
c	37	28
d	120	58
e	32	10
f	93	n/a
g	16	n/a

TABLE 1: Dimensions of the fourth generation femur and tibia models [10, 11]

CAD model of implant for TKA

The CAD model of the press-fit condylar P.F.C.® Sigma™ Knee System (DePuy Orthopaedics Inc, Johnson & Johnson, Warsaw, IN, USA) used was generated in SolidWorks 2007 based on the specific dimensions of the 71 M/L x 65 A/P sized femoral implant. Likewise, the geometries of the UHMWPE layer and tibial plate were based on the same DePuy implant system.

The modified implant concept developed by the authors, which is the innovation in this study, consists of a layer of CF/PA-12 that is around half the thickness of the original femoral implant.

This hybrid model has a final assembled geometry that is exactly the same as the original metal implant.

3. FINITE ELEMENT ANALYSIS

Assembly model for FEA

Assembly of the implant and its placement into the bone was conducted in the 'ASSEMBLY' window of SolidWorks 2007. The configuration established was focused on direct axial loading of the implant and thus only concerned the femoral and tibial bone mass within 20 cm of the tibial articular surface. This was done to ensure that uncompromised physiologic loading was simulated in the areas of focus. Thus the assembly only contained the implant, the distal end of the femur, and the proximal end of the tibia. The assembled geometry was exported into 'DesignModeler' of ANSYS® Workbench 10.0 (Canonsburg, PA, USA) software suite. The exported assembly is shown in **FIGURE 4**.
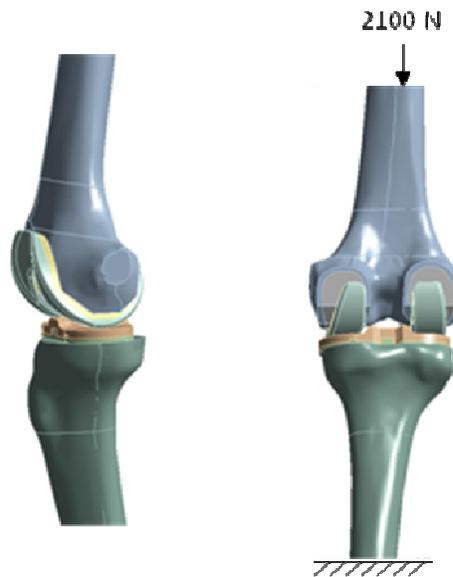
FIGURE 4: Geometry of the implant-bone system and boundary conditions

Mesh generation in ANSYS

From the 'DesignModeler' window in ANSYS® Workbench 10.0 the model was exported into the Simulation window where the mesh was generated based on 10 node quadratic tetrahedral elements sized at 5 mm. The generated mesh contained a total of 59037 nodes as a result of 38821 total elements. The meshed assembly of the bone-implant is shown in FIGURE 5.

FIGURE 5: Meshed assembly of implant system configuration

Material properties

Properties for the 4th generation composite femur and tibia were taken from the manufacturer [12], given in TABLE 2, and verified with previous studies [13-18]. The macro-structure and material properties of CF/PA-12 were based on earlier studies [9, 19, 20].

	Simulated cortical bone	Simulated cancellous bone	
		Solid	**Cellular**
Density [g/cc]	1.64	0.27	0.32
Compressive strength [MPa]	157	6.0	5.4
Compressive modulus [GPa]	16.7	0.155	0.137
Tensile strength [MPa]	106	n/a	n/a
Tensile modulus [GPa]	16.0	n/a	n/a

TABLE 2: Material properties of simulated cortical bone and cacellous bone [12, 16]

Carbon fibre (CF) weight fraction [%]	68
Polyamide 12 (PA-12) weight fraction [%]	32
CF volume fraction	0.55
PA-12 volume fraction	0.45
Density of CF [g/cc]	1.78
Density of PA-12 [g/cc]	1.03
Theoretical density of CF/PA-12 composite [g/cc]	1.443
Modulus of elasticity [GPa]	E_x = 3.0; E_y = 10.7; E_z = 10.7
Shear modulus [GPa]	G_{yz} = 2.0; G_{zx} = 2.5; G_{xy} = 2.5
Poisson ratio	$v_{yz} = v_{zx} = v_{xy}$ = 0.3

TABLE 3: Material properties of the CF/PA-12 composite [9, 19, 20]

Simulation and solution

Axial loading and restraints were applied to the meshed assemblies as described in the section "Assembly model for FEA". The assembly was restrained by restricting motion along all three axes at the distal-most cut end of the tibia. This assembly was loaded axially with 2100 N at the proximal-most end of the femur (FIGURE 5), representing approximately 3 times a nominal body weight of 70 kg. Many researchers have estimated that the maximum compressive load on the knee joint during natural gait fall within 2 to 4 times the body weight [3, 21-27], so 2100 N is reasonably typical test load. This compares well with the 2200 N used by Chu [24], two 1000 N compressive forces used by Miyoshi et al. [23], 2200 N axial force used by Godest et al. [25],

2300 N used by Halloran et al. [26], and 2000 N used by Villa et al. [27]. Bonded contact was used for all connected parts.

4. RESULTS

Validation of the FE model

Comparison between the hybrid and the conventional implants

Stress distribution contours of the 316L stainless steel and the hybrid implant are shown in FIGURE 6. Stress distribution in the composite CF/PA-12 layer alone is shown in FIGURE 7.
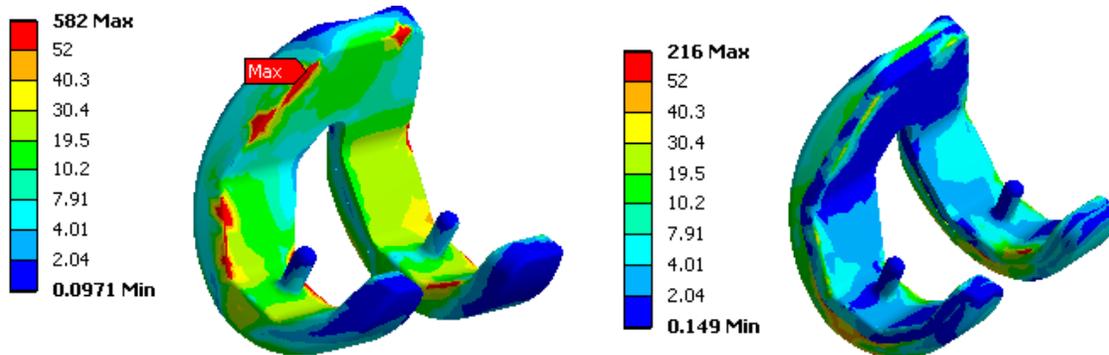
FIGURE 6: Stress distribution contours (MPa) in the 316L implant (left) and the hybrid implant (right)

The 316L implant generated minimum and maximum stresses of 0.0971 MPa and 582 MPa respectively. The hybrid implant generated a higher minimum stress of 0.149 MPa, and a much lower maximum stress of 216 MPa.
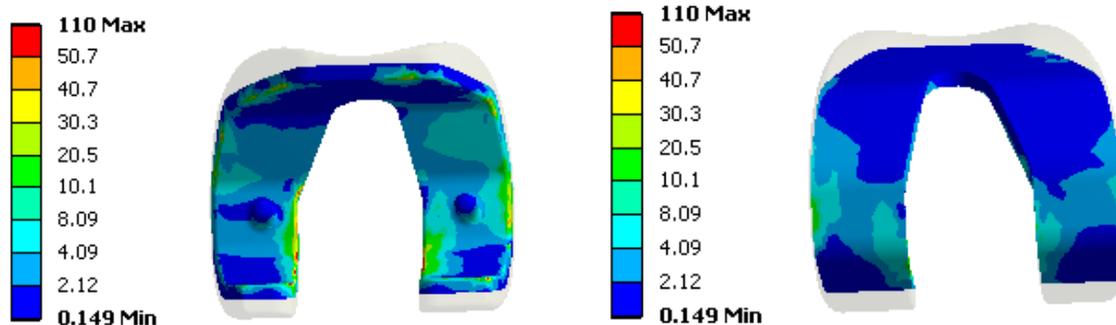
FIGURE 7: Stress distribution contours (MPa) in the CF/PA-12 layer; posterior view (left) and anterior view (right)

The stress distribution in the CF/PA-12 portion of the femoral component indicates a minimum stress of 0.149 MPa, and a maximum stress of 110 MPa. These values, when compared to the stresses in the assembled hybrid implant, show that the maximum stress generated in the hybrid femoral component (216 MPa) is not in the CF/PA-12 layer.
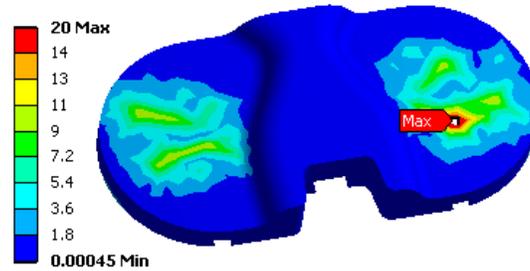
FIGURE 8: Stress distribution contours (MPa) in the UHMWPE tibial plate

The stresses generated in the UHMWPE layer peaked at 20 MPa, with most of the layer largely unaffected by the loading.

5. DISCUSSION

The stress distribution within the UHMWPE layer of the tibial component was used to validate the performance of the modelled implant assembly. Studies by Miyoshi et al. [23] have demonstrated peak stresses of 23.3 MPa in the UHMWPE tibial component, under a combined loading of 2000 N on the tibial plate. The axial load of 2100 N used in this study at the proximal end of the implant assembly produced maximum stresses in the UHMWPE layer of up to 20 MPa, as shown in FIGURE 8. This agrees well with the results of Miyoshi et al.

The FEA indicated considerably lower peak stresses in the CF/PA-12 lined femoral component when compared with the 316L stainless steel model. As shown in FIGURE 6, a peak stress of 582 MPa was generated in the 316L implant, compared to a much lower peak stress of 216 MPa generated in the hybrid implant. This decrease in the stress is due to the flexibility of the composite layer (i.e., at a maximum elastic modulus of 10.7 GPa, CF/PA-12 is 19 times more flexible than stainless steel at 193 GPa [28]).

This 63% (nearly two-thirds) reduction in peak stress tends to indicate that stress shielding, which is a transfer of load from the bone to the implant, can be expected to be much less in the hybrid implant. This reduction of overall peak stresses within the implant confirms that more load is being transferred to the bone when attached to a hybrid CF/PA-12 lined implant. This is an extremely significant improvement in exposing the femoral bone to natural physiologic loading.

6. CONCLUSION & FUTURE WORK

The results, as discussed in the previous section, indicate that a hybrid implant will reduce stress shielding and subsequent bone resorption. In turn, this will accommodate better osseoinduction and lead to longer implant life. This study is specifically geared towards a finite element analysis, and the stress transfer in the model is validated by comparing the stresses generated in the tibial UHMWPE layer to published results of implant systems under similar loading conditions [23].

Further validation of the CAD and FEA model used in this study can be performed to support the inferences derived. Some of the current authors are conducting an experimental stress analysis study on the actual commercial implant to ensure that the FEA results of the metal implant match the experimental results. The experimental study is expected to use composite femur and tibia specimens (supplied by Sawbones Worldwide [10-12]) attached to the implant system. Once it has been verified that the numerical metal implant model generates stresses comparable to the experimental one, it can be safely concluded that any subsequent modifications to the model (i.e. applying a composite layer) will produce reliable results. As shown by a previous study [9], CF-PA/12 has displayed promising results in improving the stress shielding effects in hip arthroplasty, and this study seeks to show that the same composite will prove to be similarly advantageous in knee arthroplasty.

Bougherara, H; Mahboob, Z; Miric & M; Youssef

7. REFERENCES

[1] R. J. K. Khan, P. Khoo, D. P. Fick, R. R. Gupta, W. Jacobs and D. J. Wood. *"Patella Resurfacing in Total Knee Arthroplasty"*. Cochrane Database of Systematic Reviews: Protocols, John Wiley & Sons Ltd, (1):2009.

[2] A. V. Maheshwari, P. G. Tsailas, A. S. Ranawat and C. S. Ranawat. *"How to Address the Patella in Revision Total Knee Arthroplasty"*. Knee, 16(2):92-97, 2009.

[3] B. C. Carr, and T. Goswami. *"Knee Implants – Review of Models and Biomechanics"*. Materials & Design, 30(2):398-413, 2009.

[4] P. T. Nielsen, E. B. Hansen and K. Rechnagel. *"Cementless Total Knee Arthroplasty in Unselected Cases of Osteoarthritis and Rheumatoid Arthritis : A 3-Year Follow-Up Study of 103 Cases"*. The Journal of Arthroplasty, 7(2):137-143, 1992.

[5] R. Skripitz, and P. Aspenberg. *"Tensile Bond between Bone and Titanium. A Reappraisal of Osseointegration"*. Acta Orthopaedica Scandinavica, 69(3):315-319, 1998.

[6] S. D. Stulberg. *"Bone Loss in Revision Total Knee Arthroplasty: Graft Options and Adjuncts"*. The Journal of Arthroplasty, 18(3, Part 2):48-50, 2003.

[7] S.M. Kurtz. *"The UHMWPE Handbook: Ultra-high Molecular Weight Polyethylene in Total Joint Replacement"*. Elsevier Academic Press, San Diego; London, pp. 379 (2004).

[8] G. H. Van Lenthe, M. C. De Waal Malefijt and R. Huiskes. *"Stress Shielding After Total Knee Replacement may Cause Bone Resorption in the Distal Femur"*. Journal of Bone and Joint Surgery - Series B, 79(1):117-122, 1997.

[9] H. Bougherara, M. Bureau, M. Campbell, A. Vadean and L. Yahia. *"Design of a Biomimetic Polymer-Composite Hip Prosthesis"*. Journal of Biomedical Materials Research Part A, 82A(1):27-40, 2007.

[10] Sawbones Worldwide. 2009. *"Large Left Fourth Generation Composite Femur (Item #3406)"*, Sawbones.com. Pacific Research Laboratories, Inc. Accessed January 12, 2009, <http://www.sawbones.com/products/product.aspx?1937>.

[11] Sawbones Worldwide. 2009. *"Large Left Fourth Generation Composite Tibia (Item #3402)"*, Sawbones.com. Pacific Research Laboratories, Inc. Accessed January 12, 2009, <http://www.sawbones.com/products/product.aspx?1934>.

[12] Sawbones Worldwide. 2009. *"Composite Bones: Typical Properties"*, Sawbones.com. Pacific Research Laboratories, Inc. Accessed January 16, 2009, <http://www.sawbones.com/products/bio/composite.aspx>.

Bougherara, H; Mahboob, Z; Miric & M; Youssef

[13] A. C. M. Chong, E. A. Friis, G. P. Ballard, P. J. Czuwala and F. W. Cooke. *"Fatigue Performance of Composite Analogue Femur Constructs Under High Activity Loading".* Annals of Biomedical Engineering, 35(7):1196-1205, 2007.

[14] A. C. M. Chong, F. Miller, M. Buxton and E. A. Friis. *"Fracture Toughness and Fatigue Crack Propagation Rate of Short Fiber Reinforced Epoxy Composites for Analogue Cortical Bone".* Journal of Biomechanical Engineering, 129(4):487-493, 2007.

[15] J. T. Dunlap, A. C. M. Chong, G. L. Lucas and F. W. Cooke. *"Structural Properties of a Novel Design of Composite Analogue Humeri Models".* Annals of Biomedical Engineering, 36(11):1922-1926, 2008.

[16] A. D. Heiner. *"Structural Properties of Fourth-Generation Composite Femurs and Tibias".* Journal of Biomechanics, 41(15):3282-3284, 2008.

[17] R. Zdero, M. Olsen, H. Bougherara and E. H. Schemitsch. *"Cancellous Bone Screw Purchase: A Comparison of Synthetic Femurs, Human Femurs, and Finite Element Analysis".* Proceedings of the I MECH E Part H Journal of Engineering in Medicine, 2221175-1183(9), 2008.

[18] M. Papini, R. Zdero, E. H. Schemitsch and P. Zalzal. *"The Biomechanics of Human Femurs in Axial and Torsional Loading: Comparison of Finite Element Analysis, Human Cadaveric Femurs, and Synthetic Femurs".* Journal of Biomechanical Engineering, 129(1):12-19, 2007.

[19] M. Campbell, M. Bureau and L. Yahia. *"Performance of CF/PA12 Composite Femoral Stems".* Journal of Materials Science: Materials in Medicine, 19(2):683-693, 2008.

[20] M. Campbell, J. Denault, L. Yahia and M. N. Bureau. *"CF/PA12 Composite Femoral Stems: Manufacturing and Properties".* Composites Part A: Applied Science and Manufacturing, 39(5):796-804, 2008.

[21] J. B. Morrison. *"The Mechanics of the Knee Joint in Relation to Normal Walking".* Journal of Biomechanics, 3(1):51-61, 1970.

[22] J.N. Insall, and W.N. Scott. *"Surgery of the knee".* Churchill Livingstone, New York, (2001).

[23] S. Miyoshi, T. Takahashi, M. Ohtani, H. Yamamoto and K. Kameyama. *"Analysis of the Shape of the Tibial Tray in Total Knee Arthroplasty using a Three Dimension Finite Element Model".* Clinical Biomechanics, 17(7):521-525, 2002.

[24] T. Chu. *"An Investigation on Contact Stresses of New Jersey Low Contact Stress (NJLCS) Knee using Finite Element Method".* Journal of Systems Integration, 9(2):187-199, 1999.

[25] A. C. Godest, M. Beaugonin, E. Haug, M. Taylor and P. J. Gregson. *"Simulation of a Knee Joint Replacement during a Gait Cycle using Explicit Finite Element Analysis".* Journal of Biomechanics, 35(2):267-275, 2002.

[26] J. P. Halloran, A. J. Petrella and P. J. Rullkoetter. *"Explicit Finite Element Modeling of Total Knee Replacement Mechanics".* Journal of Biomechanics, 38(2):323-331, 2005.

[27] T. Villa, F. Migliavacca, D. Gastaldi, M. Colombo and R. Pietrabissa. *"Contact Stresses and Fatigue Life in a Knee Prosthesis: Comparison between in Vitro Measurements and Computational Simulations".* Journal of Biomechanics, 37(1):45-53, 2004.

[28] Lenntech. 2009. *"Stainless Steel 316L",* Lenntech.com. Lenntech Water Treatment & Air Purification Holding B.V. Accessed January 7, 2009, <http://www.lenntech.com/Stainless-steel-316L.htm>.

The Effect of High Zeta Potentials on the Flow Hydrodynamics in Parallel-Plate Micro-Channels

A. Elazhary   umelazha@cc.umanitoba.ca
Department of Mechanical and
Manufacturing Engineering
University of Manitoba
Winnipeg, Manitoba, Canada R3T 5V6

H.M. Soliman   hsolima@cc.umanitoba.ca
Department of Mechanical and
Manufacturing Engineering
University of Manitoba
Winnipeg, Manitoba, Canada R3T 5V6

Abstract

This paper investigates the effect of the EDL at the solid-liquid interface on the liquid flow through a micro-channel formed by two parallel plates. The complete Poisson-Boltzmann equation (without the frequently used linear approximation) was solved analytically in order to determine the EDL field near the solid-liquid interface. The momentum equation was solved analytically taking into consideration the electrical body force resulting from the EDL field. Effects of the channel size and the strength of the zeta-potential on the electrostatic potential, the streaming potential, the velocity profile, the volume flow rate, the apparent viscosity, and the friction factor are presented and discussed. Results of the present analysis, which are based on the complete Poisson-Boltzmann equation, are compared with a simplified analysis that used a linear approximation of the Poisson-Boltzmann equation.

Keywords: Micro-fluidics, Fully developed, Laminar flow, Parallel-plate micro-channels, High zeta-potentials, Analytical solutions.

1. INTRODUCTION

Micro-scale fluid devices are increasingly becoming an attractive alternative to the conventional flow systems because of their compactness and large surface-to-volume ratio. These micro-scale devices are candidates for applications in heat transfer augmentation, micro-electronics and micro-electro-mechanical systems (MEMS), miniaturized chemical reactors and combustors, aerospace, and biomedical systems. Therefore, it is important to enhance our understanding of the relevant phenomena associated with fluid flow in micro-channels.

Due to the small sizes of these micro-channels, some surface phenomena (such as electrostatic forces and surface roughness) become significantly important. The present work is concerned with the effect of the electrostatic force associated with the electric double layer (EDL) on the fluid flow in micro-channels. Almost all solid surfaces have electrostatic charges, positive or negative, with different intensities. The fact that similar charges repel and different charges attract is the reason for the formation of the EDL. Consider a situation of negatively-charged surfaces

bounding a micro-channel carrying a pressure-driven liquid. The negative charges of the solid surface attract the positive charges in the liquid, while repelling the negative charges. As a result, an electric double layer (EDL) is created. The fluid flow under the influence of the pressure gradient pushes the charges in the diffuse layer towards the end of the channel giving rise to an electrical current called the streaming current. Consequently, the potential difference between the two ends of the channel generates an electrical current in the opposite flow direction known as the induction current. At steady-state conditions, a potential difference, namely the streaming potential, is generated between the two ends of the channel. The induction current carries charges and molecules in the opposite direction of the flow creating extra impedance to the flow motion which is called the electro-viscous effect. The maximum strength of the electrostatic charges occurs at the surface and is called the zeta-potential, ξ. The strength of the electrostatic potential, ψ, declines exponentially as we move away from the surface.

The existence of the EDL phenomenon has been known for over a century. Debye and Hückel [1] in 1923 linearized the exponential Boltzmann ion energy distribution and solved for the distribution of the electric potential in a solution at low wall potential. Their analysis, now known as the Debye-Hückel approximation, is valid only for situations where the zeta-potential is sufficiently low $\left(\xi \leq 25 \mathrm{mV} \right)$. Later, Burgreen and Nakache [2] determined the ψ-distribution for electrokinetic flow in parallel-plate micro-channels at high ξ. The ψ-distribution was determined in terms of elliptic integral functions of the first kind. Mala et al. [3] used the Debye-Hückel approximation and solved analytically for the velocity distribution, volume flow rate, and friction factor during fully-developed laminar flow in a parallel-plate micro-channel. Li [4] extended the analysis to the geometry of two-dimensional rectangular channels. He determined the ψ-distribution, velocity distribution, and volume flow rate from an analytical linear solution (that used the Debye-Hückel approximation) and a numerical nonlinear solution (that used the complete Poisson-Boltzmann equation). Li noted large deviations between the two solutions, particularly in the region close to the channel walls. Chen et al. [5] investigated the fluid flow characteristics for developing, pressure-driven, liquid flow in parallel-plate micro-channels. The mathematical model (Poisson-Boltzmann, Nernst-Planck, continuity and Navier-Stokes equations) was solved numerically by means of a finite-volume method. For the micro-tubes geometry, Rice and Whitehead [6] solved for the fully-developed, laminar velocity distribution using the Debye-Hückel approximation and noted that, near the wall, the negative electrostatic force caused by the migration of ions can exceed the positive force due to the pressure gradient resulting in a region of back flow near the wall. Levine et al. [7] extended the work in [6] to conditions of high zeta-potentials. A review of the electro-kinetic effects on the flow hydrodynamics in micro-channels can be found in the recent book by Li [8].

The objective of the present investigation was to generate an analytical solution for the velocity distribution during steady, laminar, fully-developed flow of liquids in parallel-plate micro-channels using the full Poisson-Boltzmann equation. Therefore, the present analysis is expected to be valid for practical situations of high zeta-potentials. The results from the present analysis will be compared with the simplified solution (that uses the Debye-Hückel approximation) developed by Mala et al. [3] in order to assess of the effects of this approximation on the fluid-flow characteristics.

2. MATHEMATICAL FORMULATION

The geometry under consideration is shown schematically in Fig. 1, whereby a micro-channel is formed between two parallel plates separated by a distance $2a$. An incompressible Newtonian aqueous 1-1 electrolyte of uniform dielectric constant ε flows in the micro-channel under the influences of an imposed pressure gradient dp/dz and a uniform zeta-potential x at both walls. A solution for the fluid-flow characteristics was obtained based on an analysis that is applicable to high x. The present analysis was conducted under the following assumptions:

1. The fluid is incompressible and Newtonian with constant thermo-physical properties.

2. The flow is considered to be steady, laminar, and fully developed.

3. The channel width (W) is much larger than channel height ($2a$), therefore, the flow is considered to be one-dimensional.

4. Non-slip conditions apply at both walls.

Electrostatic Potential Field

The electrostatic potential field (ψ) in the fluid region is governed by the Poisson–Boltzmann equation, which can be written as [3, 5]:

$$\frac{d^2\psi}{dx^2} = \frac{2n_o z_v e}{\varepsilon\varepsilon_o}\sinh\left(\frac{z_v e\psi}{k_b T}\right). \tag{1}$$

The parameters n_o, z_v, e, ε_o, k_b, and T are the bulk concentration of ions, valence of ions, electron charge, permittivity of vacuum, Boltzmann constant, and absolute temperature, respectively.

Introducing the following dimensionless parameters:

$$\bar{\psi} = \frac{z_v e\psi}{k_b T}, \quad \bar{x} = \frac{x}{a}, \quad \text{and} \quad k = \left[\frac{2n_o z_v^2 e^2}{\varepsilon\varepsilon_o k_b T}\right]^{\frac{1}{2}}, \tag{2}$$
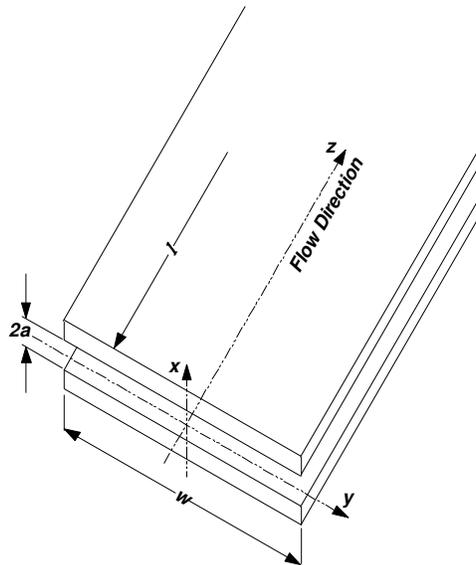
FIGURE 1: Geometry and Coordinate System

A. Elazhary & H.M. Soliman

the dimensionless form of the Poisson–Boltzmann equation can be expressed as

$$\frac{d^2\bar{\psi}}{d\bar{x}^2} = (ka)^2 \sinh(\bar{\psi}).$$

(3)

The parameter k is the Debye–Hückel parameter and $(1/k)$ is normally referred to as the characteristic thickness of the EDL. Equation (3) is subject to the following boundary conditions:

$$\frac{d\bar{\psi}}{d\bar{x}} = 0 \quad \text{at} \quad \bar{x} = 0, \quad \text{and} \quad \bar{\psi} = \bar{\xi} \quad \text{at} \quad \bar{x} = 1,$$

(4)

where, $\bar{\xi} = \dfrac{z_v e \xi}{k_b T}$.

Integrating Eq. (3) and applying the boundary condition at the center of the channel, we get

$$\frac{d\bar{\psi}}{d\bar{x}} = \sqrt{2}(ka)\big[\cosh(\bar{\psi}) - \cosh(\bar{\psi}_o)\big]^{1/2},$$

(5)

where $\bar{\psi}_o$ is the dimensionless electrostatic potential at the center of the channel. Assuming that the half-thickness of the channel is greater than the EDL thickness (i.e., $ka > 1$), we may set the parameter $\bar{\psi}_o$ in Eq. (5) to zero in order to facilitate the integration. The lowest value of ka used in the present analysis is 5.5. Integrating Eq. (5), apply the boundary condition at the channel wall, and rearranging, we get

$$\bar{\psi} = 4 \tanh^{-1}\left[\tanh\left(\frac{\bar{\xi}}{4}\right) e^{-ka(1-\bar{x})}\right].$$

(6)

The predicted $\bar{y}$ - $\bar{x}$ field from Eq. (6) as a function of (ka) and $\bar{x}$ will be compared with a numerical solution of Eq. (3) in order to confirm its validity.

Velocity Field

For steady, laminar, one-dimensional flow between parallel plates, the momentum equation (including the effect of EDL) has the following form:

$$\mu \frac{d^2 v_z}{dx^2} - \frac{dp}{dz} + E_z \rho = 0 ,$$

(7)

where ρ is the charge density and it is defined as

$$\rho = -2(n_o z_v e) \sinh\left(\frac{z_v e \psi}{k_b T}\right),$$

(8)

and E_z is the electric field strength. The product $E_z r$ in Eq. (7) represents an electric body force and its effect appears to be the opposite of that of the pressure gradient. Utilizing the following dimensionless parameters:

$$G_1 = \frac{n_o k_b T}{L\left(-\dfrac{dp}{dz}\right)}, \quad \bar{v}_z = \frac{v_z}{v_o}, \quad \text{and} \quad \bar{E}_s = \frac{L E_z}{\xi}, \tag{9}$$

where v_o is a reference velocity given by $v_o = \dfrac{1}{\mu}\left(-\dfrac{dp}{dz}\right)a^2$, and substituting from Eq. (3), the momentum equation reduces to

$$\frac{d^2 \bar{v}_z}{d\bar{x}^2} - \frac{2 G_1 \bar{\xi} \bar{E}_s}{(ka)^2} \frac{d^2 \bar{\psi}}{d\bar{x}^2} + 1 = 0. \tag{10}$$

Equation (10) is subject to the following boundary conditions:

$$\frac{d\bar{\psi}}{d\bar{x}} = \frac{d\bar{v}_z}{d\bar{x}} = 0 \quad \text{at} \quad \bar{x} = 0, \quad \text{and} \quad \bar{\psi} = \bar{\xi} \text{ and } \bar{v}_z = 0 \quad \text{at} \quad \bar{x} = 1 \tag{11}$$

Integrating Eq. (10) twice and imposing boundary conditions (11), the following velocity field was obtained:

$$\bar{v}_z = \frac{1}{2}\left(1 - \bar{x}^2\right) - \frac{2 G_1 \bar{\xi}^2 \bar{E}_s}{(ka)^2}\left(1 - \frac{\bar{\psi}}{\bar{\xi}}\right), \tag{12}$$

where the electrostatic potential $\bar{y}$ is given by Eq. (6). The first term on the right-hand-side of Eq. (12) corresponds to the velocity component induced by the pressure gradient and the second term corresponds to the retardation due to the EDL. In order to calculate the velocity from Eq. (12), the streaming potential $\bar{E}_s$ must be determined.

Streaming Potential

The streaming current that is generated due to the transport of charges by the liquid flow can be calculated from the following integral:

$$I_s = \int_{A_c} v_z \rho \, dA_c. \tag{13}$$

Introducing the dimensionless parameters

$$\bar{\rho} = \frac{\rho}{n_o z_v e} = -2\sinh(\bar{\psi}) \quad \text{and} \quad \bar{I}_s = \frac{I_s}{2 v_o n_o z_v e a}, \tag{14}$$

the streaming current equation can be written in the following dimensionless form:

$$\bar{I}_s = -2\int_0^1 \bar{v}_z \sinh(\bar{\psi}) \, d\bar{x}. \tag{15}$$

A. Elazhary & H.M. Soliman

Substituting from Eqs. (6) and (12) and performing the integration, the dimensionless steaming current can be written as

$$\bar{I}_s = I_1 - \left(\frac{4G_1\bar{\xi}\bar{E}_s}{(ka)^2}\right)I_2 - \left(1 - \frac{4G_1\bar{\xi}^2\bar{E}_s}{(ka)^2}\right)I_3 , \tag{16}$$

Where

$$I_1 = \frac{4}{(ka)^3}\left\{-ka\left[\ln\left(\frac{1+\eta e^{ka}}{1-\eta e^{ka}}\right) + \frac{\eta e^{ka}}{(\eta e^{ka})^2 - 1}\right] + \left[\text{Li}_2\left(\eta e^{ka}\right) - \text{Li}_2\left(-\eta e^{ka}\right)\right] - \left[\text{Li}_2\left(\eta\right) - \text{Li}_2\left(-\eta\right)\right]\right\} \tag{17a}$$

$$I_2 = \frac{8}{ka}\left[\frac{1 - 2\eta e^{ka}\tanh^{-1}\left(\eta e^{ka}\right)}{(\eta e^{ka})^2 - 1} - \frac{1 - 2\eta\tanh^{-1}\left(\eta\right)}{\eta^2 - 1}\right] , \tag{17b}$$

$$I_3 = -\frac{4}{ka}\left[\frac{\eta e^{ka}}{(\eta e^{ka})^2 - 1} - \frac{\eta}{\eta^2 - 1}\right] , \tag{17c}$$

$$h = \tanh(\bar{x}/4)e^{-ka} , \tag{17d}$$

and Li$_2$ (β) is the Poly-Logarithmic function of second order defined by

$$\text{Li}_2(b) = -\int_0^b \frac{\ln(1-t)}{t}dt . \tag{17e}$$

An equilibrium state occurs when the streaming current is equal to the conduction current, i.e.,

$$I_c + I_s = 0 , \tag{18}$$

which can be written in the following dimensionless form:

$$\bar{I}_c + \left(\frac{G_2(ka)^2}{\bar{\xi}}\right)\bar{I}_s = 0 , \tag{19}$$

where $G_2 = \dfrac{L\varepsilon\varepsilon_o(-dp/dz)}{2\mu\lambda_o}$ and the dimensionless conduction current is $\bar{I}_c = \bar{E}_s\bar{A}_c/L$. Substituting Eq. (16) into Eq. (19) and rearranging, we get

$$\bar{E}_s = \frac{\dfrac{G_2(ka)^2}{\bar{\xi}}(I_3 - I_1)}{1 - 4G_1G_2\left(I_2 - \bar{\xi}I_3\right)} . \tag{20}$$

A. Elazhary & H.M. Soliman

Apparent Viscosity

The volume flow rate can be calculated from the relation

$$Q = 2W \int_0^a v_z \, dx \, . \tag{21}$$

In dimensionless form, Eq. (21) can be written as

$$\bar{Q} = \frac{Q}{2aW v_o} = \bar{v}_{zm} \, , \tag{22}$$

where the mean axial velocity, $\bar{v}_{zm}$, is given by

$$\bar{v}_{zm} = \int_0^1 \bar{v}_z \, d\bar{x} = \frac{1}{3} + \frac{4G_1 \overline{\xi E_s}}{(ka)^3} \left[\frac{(ka)\bar{\xi}}{2} + \left\{ \text{Li}_2 \left(\eta e^{ka} \right) - \text{Li}_2 \left(-\eta e^{ka} \right) \right\} - \left\{ \text{Li}_2 \left(\eta \right) - \text{Li}_2 \left(-\eta \right) \right\} \right] . \tag{23}$$

For poiseuille flow between two parallel plates without the EDL effect, the volumetric flow rate is expressed as

$$Q = \frac{2(-dp/dz) a^3 W}{3 \mu_a} \, , \tag{24}$$

where μ_a stands for the apparent viscosity. Equation (24) can be written as

$$\bar{Q} = \frac{\mu}{3 \mu_a} \, . \tag{25}$$

Comparing Eqs. (22) and (25), we get after rearranging

$$\frac{\mu_a}{\mu} = \frac{1}{3 \bar{v}_{zm}} = \frac{1}{3 \bar{Q}} \, . \tag{26}$$

Friction Factor

The friction factor is normally defined as

$$f = \frac{2a(-dp/dx)}{\rho v_{zm}^2} \, , \tag{27}$$

and Reynolds number is given by

$$\text{Re} = 4a\rho_f v_{zm}/\mu \, . \tag{28}$$

Combining Eqs. (27) and (28), and using the definition of G_1 given by Eq. (9), we get

$$f \, \text{Re} = \frac{8}{\bar{v}_{zm}} = \frac{8}{\bar{Q}} \tag{29}$$

where $\bar{v}_{zm}$ is given by Eq. (23). Because μ/μ_a and f Re are directly proportional to $\bar{Q}$, as given by Eqs. (26) and (29), respectively, only the results of $\bar{Q}$ will be presented later.

3. RESULTS AND DISCUSSION

All the results presented in this paper correspond to an infinitely diluted aqueous 1:1 electrolyte solution ($n_o = 6.022 \times 10^{20}$ m^{-3}) at $T = 298$ K ($\varepsilon = 80$, $\lambda_o = 1.264 \times 10^{-7}$ Ω^{-1}m^{-1}, and $\mu = 0.9 \times 10^{-3}$ N s m^{-2}). A fixed pressure gradient $(-dp/dz) = 3$ atm/cm was used, while different values of a and ξ were utilized in generating the results. In the following figures, the results corresponding to the full model are based on the present analysis, while the results corresponding to the simplified model are based on the analysis of Mala et al. [3].

Figure 2 shows a comparison between three solutions of $\bar{\psi}$-distribution corresponding to $\xi = 240$ mv $\left(\bar{\bar{\xi}} = 9.35\right)$ and $a = 1.7$ μm ($ka = 5.54$). The three solutions correspond to the full model given by Eq. (6), the simplified model of Mala et al. [3], and an accurate numerical solution of Eq. (3) together with boundary conditions (4). The close agreement between the full model and the numerical solution indicates that Eq. (6) produces accurate values of $\bar{\psi}$ throughout the channel at high zeta-potentials, thus confirming that the assumption of $\bar{\psi}_o = 0$ in Eq. (5) does not have a strong influence on the $\bar{\psi}$-distribution. On the other hand, the simplified model corresponding to the Debye–Hückel approximation deviates significantly from the other two sets of results.
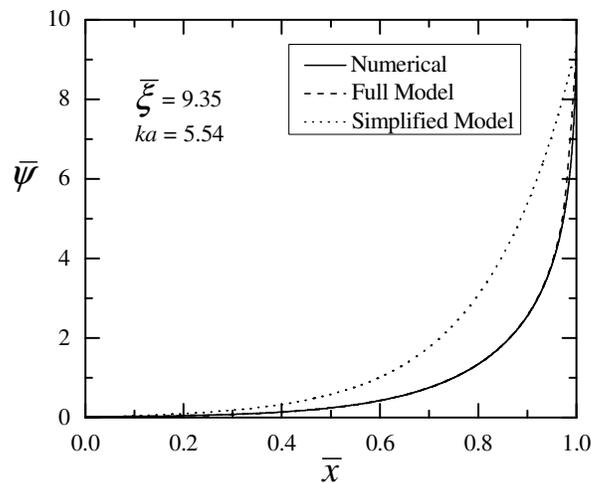

FIGURE 2: Validation of Eq. (6) for $\bar{\psi}$.

A. Elazhary & H.M. Soliman

The influence of $\bar{\bar{\xi}}$ on the $\bar{\psi}$-distribution at ka = 5.5 is shown in Fig. 3. Both the simplified and the full model predict that $\bar{\psi}$ increases with an increase in $\bar{\bar{\xi}}$; however, the simplified model overestimates the value of $\bar{\psi}$ in all cases. Significant deviations can be seen between the two models at higher values of $\bar{\bar{\xi}}$. It is clear from Fig. 3 that $\bar{\bar{\xi}}$ is the key parameter in determining the magnitude of the deviation between the two models and that even for a narrow channel with ka = 5.5, the deviation can be small if $\bar{\bar{\xi}}$ is small (e.g., $\bar{\bar{\xi}} = 2$). The present result showing the simplified model over- predicting ψ at high ξ is consistent with previous results [4].
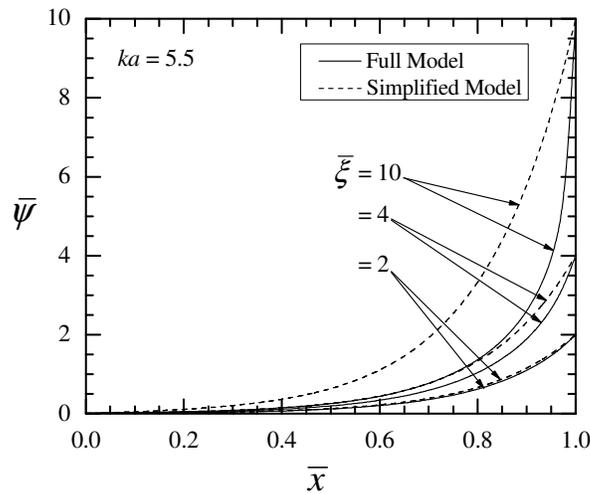
FIGURE 3: Effect of $\bar{\bar{\xi}}$ on $\bar{\psi}$ at ka = 5.5.

Variation of the streaming potential, $\bar{E}_s$, with $\bar{\bar{\xi}}$ and (ka) is shown in Fig. 4. Both the simplified and the full models predict that $\bar{E}_s$ decreases as $\bar{\bar{\xi}}$ increases or (ka) decreases. An increase in ξ results in more ions being attracted in the double layer and less ions carried downstream by the flowing fluid and thus a lower E_s. Also, a decrease in (ka) decreases the volume between the plates resulting in less ions carried downstream and consequently a lower E_s. Figures 4 indicates that the simplified solution overestimates the value of $\bar{E}_s$ in all cases, consistent with the overestimation of $\bar{\psi}$ shown earlier. The deviation between the two solutions increases as $\bar{\bar{\xi}}$ increases.

The velocity distributions predicted by both the simplified and the full models for the range $0 \leq \bar{\bar{\xi}} \leq 10$ are shown in Fig. 5 for ka = 5.5. The profile marked $\bar{\bar{\xi}}$ = 0 represents the velocity component induced by the pressure gradient without EDL retardation. Figure 5 shows that, as $\bar{\bar{\xi}}$

increases, the full model predicts significant retardation of the flow velocity throughout the cross-section and a region of reversed flow near the wall up to $\bar{\bar{\xi}}$ = 2. The deviation between the full model and the simplified model is insignificant at $\bar{\bar{\xi}}$ = 0.5 and small at $\bar{\bar{\xi}}$ = 2. Surprisingly, the magnitude of velocity retardation in the core region of the flow decreases as $\bar{\bar{\xi}}$ increases from 2 to 10, while increased velocity retardation continues near the wall. As well, the near-wall zone (where increasing retardation continues) shrinks in size as $\bar{\bar{\xi}}$ increases. This result is consistent with the near exponential decrease of $\bar{E}_s$ with $\bar{\bar{\xi}}$, shown earlier in Fig. 4. It is important to note that the simplified model predicted an opposite trend in terms of the effect of $\bar{\bar{\xi}}$ on the velocity retardation over the range 2 ≤ $\bar{\bar{\xi}}$ ≤ 10. These results are consistent with those of Burgreen and Nakache [2] who reported that the retarding-flow component is less than 10% of the pressure-induced flow component for ka > 20, while for ka = 1.6, the retarding-flow component can be as much as 68% of pressure-induced flow component at $\bar{\bar{\xi}}$ = 4. The results in [4] also indicate that for ka > 1, the magnitude of the retarding component decreased as $\bar{\bar{\xi}}$ increased from 4 to 10.
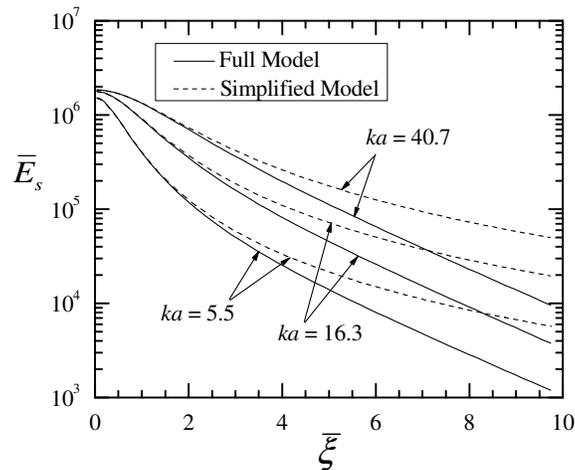
FIGURE 4: Variation of $\bar{E}_s$ with $\bar{\bar{\xi}}$ and (ka).

Predictions of the dimensionless volume flow rate, $\bar{Q}$, are shown in Fig. 6. The full model predicts that $\bar{Q}$ decreases with $\bar{\bar{\xi}}$ due to velocity retardation down to a minimum value at about $\bar{\bar{\xi}}$ = 2.14 for ka = 5.5 and about $\bar{\bar{\xi}}$ = 3.31 for ka = 40.7, and then increases with $\bar{\bar{\xi}}$. The reduction in volume flow rate due to EDL is more significant for low values of ka; the maximum reduction is about 63% for ka = 5.5 and only about 9.6% for ka = 40.7. The trends in these results are consistent with the velocity results presented earlier. Values of μ_a/μ can be deduced from Fig. 6 using Eq. (26). As well, values of f Re can be deduced from Fig. 6 using Eq. (29).
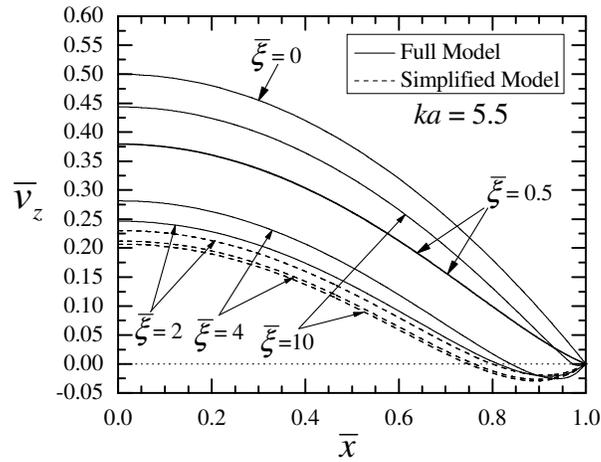
FIGURE 5: Velocity Profiles for $ka = 5.5$ and various $\bar{\xi}$.
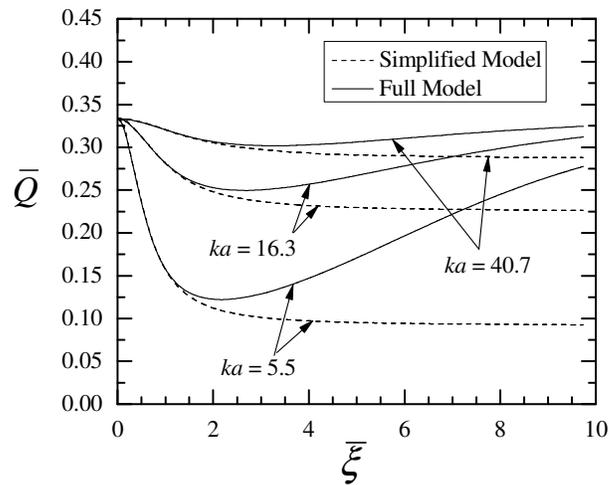
FIGURE 6: Variation of $\bar{Q}$ with $\bar{\xi}$ and (ka).

4. CONCLUSIONS

An analytical solution was developed for steady, laminar, fully-developed flow of liquids in a micro-channel formed by two parallel plates under the influence of an electric double layer (EDL). The solution was based on the full Poisson-Boltzmann equation and the results include the distribution of the electrostatic potential, ψ, the streaming potential, $\bar{E}_s$, the velocity distribution,

the volumetric flow rate, the apparent viscosity, and the friction factor. Comparisons were made between the predictions from the present (full) model and those from a simplified model based on the linearized Debye-Hückel approximation. The following conclusions can be drawn from the present results:

1. Both the simplified and the full models predict that $\bar{\psi}$ increases with an increase in $\bar{\bar{\xi}}$ or a decrease in (ka); however, the simplified model overestimates the value of $\bar{\psi}$ in all cases. The deviations between the two models increase as $\bar{\bar{\xi}}$ increases.

2. Both the simplified and the full models predict that $\bar{E}_s$ decreases as $\bar{\bar{\xi}}$ increases or (ka) decreases. The simplified solution overestimates the value of $\bar{E}_s$ in all cases, consistent with the overestimation of $\bar{\psi}$, and the deviation between the two solutions increases as $\bar{\bar{\xi}}$ increases.

3. The full model predicts that the magnitude of velocity retardation increases as $\bar{\bar{\xi}}$ increases up to $\bar{\bar{\xi}}$ = 2.14 (for ka = 5.5) and $\bar{\bar{\xi}}$ = 3.31 (for ka = 40.7), beyond which the velocity retardation decreases with a further increase in $\bar{\bar{\xi}}$. On the other hand, the simplified model predicts that the velocity retardation increases continuously as $\bar{\bar{\xi}}$ increases.

4. The full model predicts that the volumetric flow rate decreases as $\bar{\bar{\xi}}$ increases down to a minimum at $\bar{\bar{\xi}}$ = 2.14 (for ka = 5.5) and $\bar{\bar{\xi}}$ = 3.31 (for ka = 40.7), beyond which the volumetric flow rate increases with a further increase in $\bar{\bar{\xi}}$. On the other hand, the simplified model predicts that the volumetric flow rate decreases continuously as $\bar{\bar{\xi}}$ increases. The trends of the behaviours of the apparent viscosity and the friction factor follow the trend of the volumetric flow rate.

ACKNOWLEDGEMENT

5. REFERENCES

1. P. Debye, E. Hückel. "Zur theorie der elektrolyte. I. Gefrierpunktserniedrigung und verwandte erscheinungen" Physik. Z., 24: 185-208, 1923

2. D. Burgreen, F.R. Nakache. "Electrokinetic flow in ultrafine capillary slits". J. Physical Chemistry, 68: 1084-1091, 1964

3. G.M. Mala, D. Li, C. Werner, H.-J. Jacobasch, Y.B. Ning. "Flow characteristics of water through a microchannel between two parallel plates with electrokinetic effects". Int. J. Heat Fluid Flow, 18: 489-496, 1997

4. D. Li. "Electro-viscous effects on pressure-driven liquid flow in microchannels". Colloids and Surfaces, A: Physicochemical and Engineering Aspects, 195: 35-57, 2001

A. Elazhary & H.M. Soliman

5. X.Y. Chen, K.C. Toh, J.C. Chai, C. Yang. "Developing pressure-driven liquid flow in microchannels under the electrokinetic effect". Int. J. Eng. Sci., 42: 609-622, 2004

6. C.L. Rice, R. Whitehead. "Electrokinetic flow in a narrow cylindrical capillary". J. Physical Chemistry, 69: 4017-4023, 1965

7. S. Levine, J.R. Marriott, G. Neale, N. Epstein. "Theory of electrokinetic flow in fine cylindrical capillaries at high zeta-potentials". J. Colloid and Interface Science, 52: 136-149, 1975

8. D. Li. "Electrokinetics in Microfluidics". Elsevier Academic Press, 2004.

MFBLP Method Forecast for Regional Load Demand System

Zuhairi Baharudin   zuhairb@petronas.com.my

Electrical and Electronics Department
University Technology PETRONAS
31750 Tronoh, Perak, Malaysia

Nidal Kamel   nidalkamel@petronas.com.my

Electrical and Electronics Department
University Technology PETRONAS
31750 Tronoh, Perak, Malaysia

Abstract

Load forecast plays an important role in planning and operation of a power system. The accuracy of this forecasted value is necessary for economically efficient operation and also for effective control. This paper describes a method of modified forward backward linear predictor (MFBLP) for solving the regional load demand of New South Wales (NSW), Australia. The method is designed and simulated based on the actual load data of New South Wales, Australia. The accuracy of discussed method is obtained and comparison with previous methods is also reported.

Keywords: Short term load forecasting (STLF), autoregressive (AR), Modified, Linear Predictor, Autoregressive moving average (ARMA), Burg.

---

1. INTRODUCTION

Short term prediction of future load demand is important for the economic and secure operation of power systems. Fundamental operation functions such as unit commitment, hydro-thermal coordination, interchange evaluation, scheduled maintenance and security assessment require a reliable short term load forecast (STLF).

Throughout the paper the term "short" is used to imply prediction times of the order of hours. The time boundaries are from the next hour, or possibly a half-hour, up to 168 h. The basic quantity of interest in STLF is the hourly integrated total system load. Owing to the importance of the STLF, research in this area in the past two decades has resulted in the development of numerous forecasting methods [1, 2, 2-10]. One of the STLF methods that received significant attention in literature for more than 20 years and a large number of estimation methods is the autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA). Both models are also known as Box-Jenkins have more degrees of freedom than the autoregressive, so greater latitude in its ability to generate diverse time-series shapes is therefore, expected of its estimators. Unfortunately, this is not always the case, because of the nonlinear nature required of algorithms that must simultaneously estimate the moving average and autoregressive parameters of the models. This phenomenon finally produces low accuracy in forecast of the model algorithm based from the assumption below:

Though the all pole models have less degree of freedom than ARMA and ARIMA, they exhibit major advantages. First of all, all-pole models have been found to provide a sufficiently accurate representation for many different types of signals in many different applications. Another reason for the popularity is the special structure which leads to fast and efficient algorithms for finding the all-pole parameters. In this

paper, the modified forward backward linear predictor (MFBLP) is introduced as autoregressive (AR) based solution to STLF problem. The method is discussed and the performance is tested and compared with Box-Jenkins ARIMA [5] (assuming that the performance of ARIMA is better than ARMA), artificial neural network (ANN) [11] and Burg's algorithm [12]. A load demand data record from New South Wales (NSW), Australia is used for the model designed purpose and validation process [13].

2. METHODOLOGY

The methods highlight in the paper is the non-weather type model. This means that, it takes only the historical load data as an input to the model. In general, the time dependent non-weather model uses past and latest load behavior to extrapolate the sample data prior to the forecast stage. The hourly load data are observed to identify the data behavior. Then, the data need to examine whether the data consist of seasonality or trend. The non stationary behavior should be remedied before the estimating process takes place.

In the previous literature [5, 14] had suggested that the method of difference the data can be one of the steps to achieve stationary white-noise in the data series. The next step is to examine the residuals of the estimation process. The residuals should be white-noise (stationary) [15] then, it possible to run the key steps in the methodology, which is the forecast. The summary of the steps taken in the methodology is depicted in the Figure 1.

2.1. Data observation
The historical load data used in this study has been gathered from the National Electricity Market Management Company Limited (NEMMCO) [13]. The data series is the actual hourly data of the system load demand for the region of New South Wales (NSW), Australia. The regional data series for NSW system demand covering the period from 1st January 2005 to 31st December 2007. Figure 2 and 3 shows the time series plot (sequence plot) for the estimation sample and forecast validation data respectively.
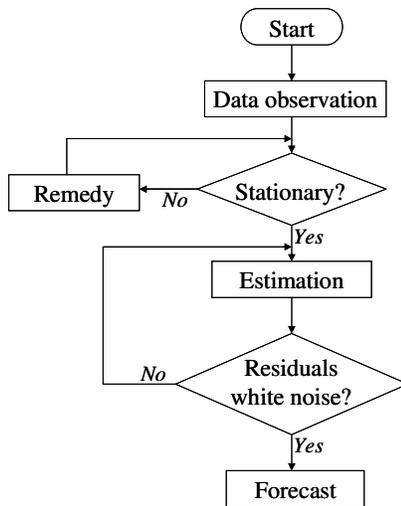

FIGURE 1: The general procedure of the STLF.
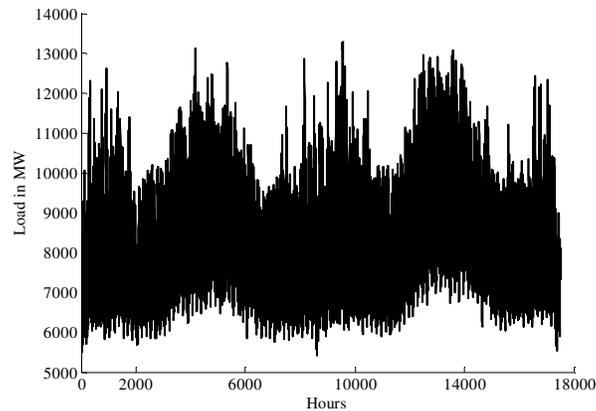
FIGURE 2: NSW hourly data from 1.1.2005 to 31.12.2006.
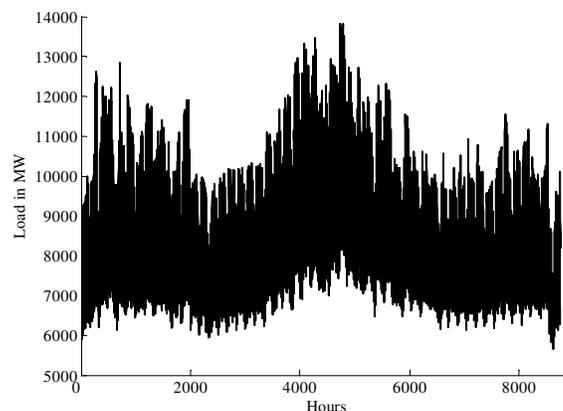
FIGURE 3: NSW hourly data from 1.1.2007 to 31.12.2007

Figure 2 and 3 clearly depict that the data series contains seasonality. Moreover, in Figure 2, the data show slightly positive trends. The seasonality is formed by the cycle of the load demand behavior, which is contributed from daily, weekly and yearly load demand patterns. It is clearly seen that, the data series is having non-zero mean and non-constant variance [12]. This suggests that a further remedy needs to apply to the data in order to achieve the zero mean and constant variance.

2.2. Data pre-processing

The important issue in time series forecast is to clarify the behavior of zero mean and constant variance of the given data series, which is the main assumption in this paper. According to Figure 1, the estimation sample data in Figure 2 need the necessary remedy to remove the seasonality and trend in the data. One of the effective solutions for this is differentiating the data as suggested in [5, 14]. Consequently, suggest the triple difference as a method to remove the non-stationary behavior in the data.

By assuming the actual load demand data series as L_t, hence, the considered difference is as follows:

$$x_t = L_t - L_{t-1} \qquad (1)$$
$$y_t = x_t - x_{t-24} \qquad (2)$$
$$z_t = y_t - y_{t-168} \qquad (3)$$

The first difference when $k = 1$ eliminate the small trend that develops in Figure 2. The difference of $k = 24$ and later by $k = 168$ eliminate the daily and weekly seasonality respectively. Tentatively the model can be described as follows:

$$\nabla_1 \nabla_{24} \nabla_{168} L_t = e_t \qquad (4)$$

$$(1 - B)(1 - B^{24})(1 - B^{168})L_t = e_t \qquad (5)$$

Where B denotes the normal back-shift operator, i.e. $BL_t = L_{t-1}$ or $B^k L_t = L_{t-k}$ and the term e_t is the white noise process of the residuals. For simplicity the expression in (5) can be reduced as:

$$z_t = e_t \qquad (6)$$

The sequence plot and the autocorrelation plot (ACF) of z_t are depicted in Figure 4 and 5 respectively. It is clearly seen from the Figure 4, the new data series show stationary and most importantly the data are now zero mean and constant variance. The data autocorrelation is considered negligible, and it is well in the bound of 95% confidence level. The confidence level is given by [15],

$$Confidence\ level_{95\%} = \frac{\pm 1.96}{\sqrt{n}} \qquad (7)$$

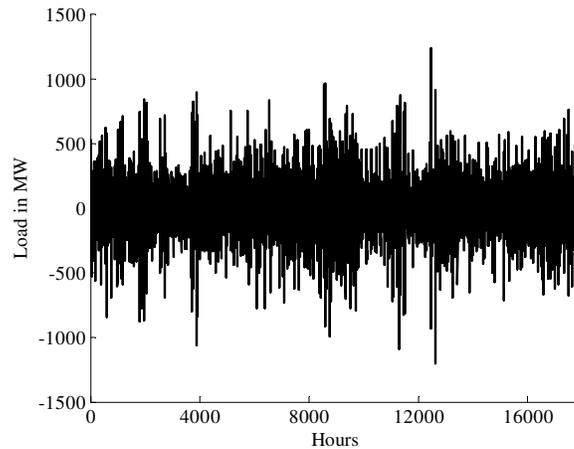Where n is the number of samples.
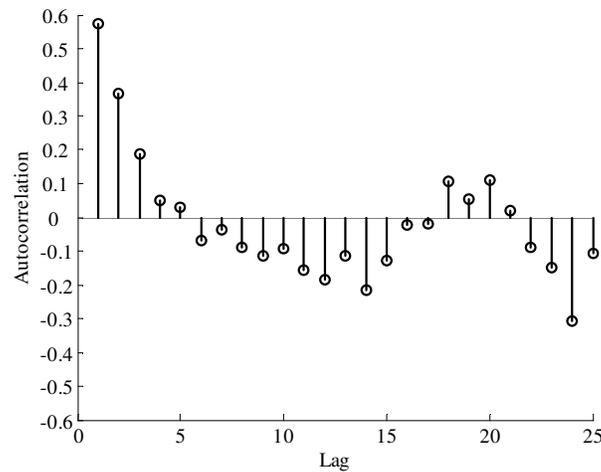

FIGURE 4: Sequence plot of z_t.

FIGURE 5: ACF of z_t for the first 25 lags.

2.3. The MFBLP algorithm

The *M*-point data sequence of $z(1)$, $z(2)$, … $z(M)$ is to be used to estimate the *p*th AR filter coefficients. Since, with AR algorithms the order of the model is proportional to the length of data record and in order to avoid using large orders with long data records, let's consider the segmentation of the *M*-points data sequence into Q segments of N samples each.

Assume one segment of data out of the available Q segments. Because forward and backward linear predictions have similar statistical information, it seems reasonable to combine the linear prediction error statistics of both directions in order to generate more error points [16]. The net result should be an improved estimate of the autoregressive parameters. The (N-p) forward and the (N-p) backward linear prediction samples of the non-windowed case may be written in matrix form as follows:

$$\mathbf{D_q f} = \mathbf{w} \qquad (8)$$

Where the 2×(N-p) forward-backward linear prediction data matrix is defined as,

$$\mathbf{D_q} = \begin{bmatrix} x(N-1) & x(N-2) & \cdots & x(0) \\ x(N) & x(N-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(N-2) & x(N-3) & \cdots & x(N-p) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(N-p) & x^*(N-p+1) & \cdots & x^*(N-1) \end{bmatrix} \qquad (9)$$

Let **w** denotes the desired response at the predictor output, defined as,

$$\mathbf{w_q} = \begin{bmatrix} x(p) & x(p+1) & \cdots & x(N-1) & x^*(1) & x^*(2) & \cdots & x^*(N-p-1) \end{bmatrix}^T \qquad (10)$$

Meanwhile the following vector notation for the forward linear prediction coefficients is given as,

$$\mathbf{f} = \begin{bmatrix} a_p^f(1) & a_p^f(2) & \ldots & a_p^f(p) \end{bmatrix}^T \qquad (11)$$

Since the forward and backward prediction coefficients for a stationary random process are simply complex conjugate of one another, the vector of the desired response at the predictor output, once again can be rewritten in matrix form as in Eq. (8). By forming the data matrix $\mathbf{D_q}$ in corresponds to each data segment, and arranging the resultant matrices in the following form,

$$\mathbf{D} = \begin{bmatrix} \mathbf{D_1} \\ \mathbf{D_2} \\ \vdots \\ \mathbf{D_Q} \end{bmatrix} \qquad (12)$$

The corresponding predicted vector to matrix **D** is defined as,

$$\mathbf{w} = \begin{bmatrix} \mathbf{w_1} \\ \mathbf{w_2} \\ \vdots \\ \mathbf{w_Q} \end{bmatrix} \qquad (13)$$

From Eq.(12) and Eq.(13), the modified equations of the predictors, can be simplify as,

$$\mathbf{D f} = \mathbf{w} \qquad (14)$$

Where **D** is the data series, **f** is the predicted coefficients and **w** is the predicted response (signal). A well established method of least squares [17, 18] is used to obtain a solution to Eq. 14 for the predictor coefficient vector **f**. The solution will guarantee the minimum sum [16] of the squared values of the predicted errors (residuals).

2.4. Estimation

By referring to the Eq. (14), the linear predicted errors are given by,

$$\mathbf{e} = \mathbf{w} - \mathbf{Df} \qquad (15)$$

and the sum of squared errors is given by,

$$\zeta = \mathbf{e}^{\mathbf{H}}\mathbf{e} \qquad (16)$$

By substituting Eq. (14) into Eq. (16), the equation may explicitly express the dependence of the sum of squared errors on the predictor coefficients, as follows

$$\zeta = \mathbf{w}^{\mathbf{H}}\mathbf{w} - \mathbf{w}^{\mathbf{H}}\mathbf{Df} - \mathbf{f}^{\mathbf{H}}\mathbf{D}^{\mathbf{H}}\mathbf{w} + \mathbf{f}^{\mathbf{H}}\mathbf{D}^{\mathbf{H}}\mathbf{Df} \qquad (17)$$

Now, by differentiating Eq. (17) with respect to the **f**, hence the following gradient vector is determined,

$$\frac{\partial \zeta}{\partial \mathbf{f}} = -2\,\mathbf{D}^{\mathbf{H}}\mathbf{w} + 2\,\mathbf{D}^{\mathbf{H}}\mathbf{Df} \qquad (18)$$

$$= -2\,\mathbf{D}^{\mathbf{H}}\mathbf{e}$$

As it is clear the sum of squared errors reaches its minimum value when the gradient vector is zero. Then from Eq. (18), the equation immediately deduces to,

$$\mathbf{D}^{\mathbf{H}}\mathbf{Df} = \mathbf{D}^{\mathbf{H}}\mathbf{w} \qquad (19)$$

Thus, the predictor coefficients that give the least squared errors are obtained as a solution to Eq. (19). However, this solution is unique only when the matrix **D** is full rank. When this condition is satisfied the matrix $\mathbf{D}^{\mathbf{H}}\mathbf{D}$ is nonsingular and the solution is unique, given as,

$$\mathbf{f} = \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathbf{H}}\mathbf{w} \qquad (20)$$

Thus the predictor coefficients that give the least squared errors are obtained as a solution to Eq. (20). However, this solution is unique only when the nullity of the matrix **D** is zero [19]. The nullity of a matrix denoted as null (.) is defined as the dimension of the matrix null space. In other words the least-squares solution is unique when the matrix **D** is of full rank. When this condition is satisfied, the *p*-by-*p* matrix $\mathbf{D}^{\mathbf{H}}\mathbf{D}$ is nonsingular and the solution is unique, given as,

$$\mathbf{f} = \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathbf{H}}\mathbf{w} \qquad (21)$$

$$= \mathbf{D}^{\#}\mathbf{w}$$

Where **D** is called the pseudo-inverse of the matrix **D**, given as,

$$\mathbf{D}^{\#} = \left(\mathbf{D}^{\mathbf{H}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathbf{H}} \qquad (22)$$

2.5. Forecast

Once the process estimates the suggested model, the predicted coefficients, **f**, are determined. With the coefficients, the forecast load for the region of NSW will be determined. One year actual data from 1.1.2007 to 31.12.2007 are used as a validation data or model forecast performance indicator. This validation data series is the actual hourly load data and can be described as,

$$v(t) = \left[v(1)\ v(2) \cdots v(N)\right] \qquad (23)$$

The forecast horizons are set as follows: 1-hour, 24-hour, 48-hour and 168-hour step ahead. This means that the coefficients, **f**, are applied in determining the *t*-th step ahead by using the autoregressive (AR) method. The AR a model for *p*-th order of coefficients, **f**, for the 1-hour ahead is given by,

$$f(1) = \left[a(1)z(t) + a(2)z(t-1) \cdots a(p)z(t-p+1)\right] \qquad (24)$$

The process is reversed to determine the actual forecast value for, $F(1)$, and is given by,

$$F(1) = (1-B)(1-B^{24})(1-B^{168})f(1) \qquad (25)$$

Thus the forecast error is stated as,

$$e_t = v(1) - F(1) \qquad (26)$$

Therefore, for the *t*-th step ahead forecast, the AR (*p*) model can be described as,

$$F(t) = (1-B)(1-B^{24})(1-B^{168})f(t) \qquad (27)$$

Where

$$f(t) = a_p(B)z_t \qquad (28)$$

The term B is the backshift operator, defined by $Bz_t = z_{t-1}$, and hence $B^k z_t = z_{t-k}$.

It is noticed that the Eq. 25 are assumed as a 'starter' forecast value, where this formulation is used to determine the next value of forecast whenever $t > 1$. The accuracy of the total forecast is depending on this starter value and how accurate the process estimates in producing the AR coefficients, $a(p)$. As mentioned earlier, by using the MFBLP algorithm and estimation by the least squares methods, the errors of the forecast are considered minimum.

3. THE EXPERIMENT

The simulations considered in this paper consist of computing the forecast of 1, 24, 36 and 168 hours ahead. The system load demand data that are applied in this work are gathered from NEMMCO. For the estimation process, the data are range from 1.1.2005 to 31.12.2006 and for the forecast validation the data are taken from 1.1.2007 to 31.12.2007. Both mentioned data are hourly sequence load demand data. This data series is considered as a raw data because no attempts are done to smooth the data significant from the abrupt change in weather, extreme weather condition or even the effect of public holiday which are very significant for the particular affected week of the load demand. The main objective for this is to test and confirm that the non-weather AR-model of Eq. 14 and Eq. 28 could deliver an acceptable forecast result. For a comparison purpose, several previous methods are in comparison to benchmark the proposed method described in the paper. All models in this work are simulated in a computer with a MATLAB version 7.2.0.232 and Pentium 4 CPU 3.20GHz with 0.5GB of RAM.

3.1. Errors indicator

The error analysis in estimation and forecast processes are represented by the mean absolute percentage errors (MAPE), and it is given by,

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|x_i - \hat{x}_i|}{x_i}\times 100\% \quad (29)$$

3.2. The segmentation parameter

The main significant finding for the MFBLP is the formulation of data segmentation. The data are segmented into Q segments of N samples each. As described previously, the total point's data after it has been remedied to remove the seasonality and trends consist of 17520 points of data. Consequently, provide several combinations of segmentation parameters (SP) of Q and N. Four SP combinations are investigated and the results for the average MAPE during an estimation process for each SP combination are depicted in the Table I.

Table I: Average MAPE for four combinations of SP.

Segmentation parameters		Estimation
Q	N	Average MAPE, %
2	8760	0.97
12	1460	0.82
24	730	0.75
48	365	0.91

From Table 1 it is clearly found that the best SP combinations are Q = 24 and N = 730. The assumption for the SP combination finding is that $Q*N$ must equal to total number of data points, which in this case 17520.

3.3 Estimation the number of order coefficients

The MFBLP in Eq. 9 and the simplified Eq. 14 are used to determine the estimation of the data. By implementing the least squares, it can provide the minimum sum squares of the errors. As mentioned before, the works in this paper highlight the non-weather AR (AR) methods in defining the model order number and its coefficients. In [20] suggested that for the optimum number of model order, it should not exceed 10% of the total estimation sample data. This would result from the initial draw of the parameters for the model should be less than 1752. However, in the actual experiment, the model order number is much less than the suggested one [20]. Figure 6 depicts the relation between the estimation errors and the model order number.
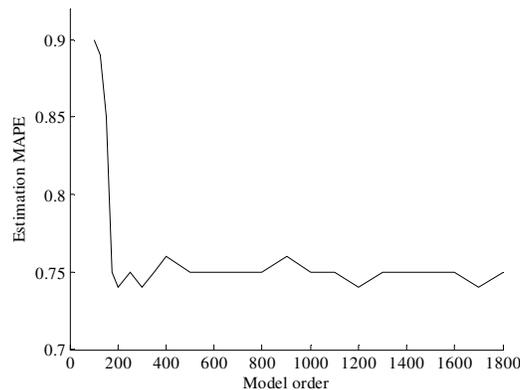
FIGURE 6: The relation of estimation MAPE and the model order number.

It is clearly seen that, the MAPE converges after a certain number of order. The lowest MAPE is 0.74% at order 200. There is a slight ripple of MAPE averaging at 0.75%. By using the AICC [21] the order with the minimum value is selected, and it is found that the best number of the order model is 175 as indicated in the Figure 7. The residuals of the estimated model are calculated and examined with the selected model order. Figure 8 depicts the residuals for the estimated samples and Table II describes the model performance with the specified parameters. The white noise of the residuals is supported by checking the ACF value which is well inside the 95% confidence interval bound. This is a clear indication that the residuals are stable, controllable and most importantly stationary with zero mean and constant variances.
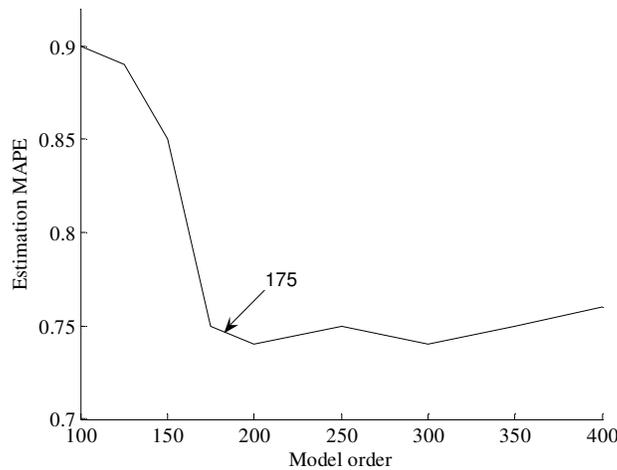
FIGURE 7: The estimation MAPE and the model order number up to 400.
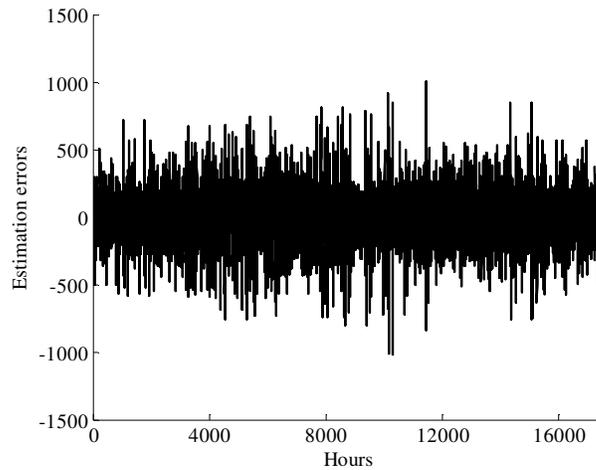
FIGURE 8: The residuals of the estimated models.

TABLE II: The number of order parameter selected and the performance of estimated models.

No. of samples	No. of order parameter	Residuals
17520	175	0.75%

3.4. Model forecast

The forecast model is described in previous section and can be represented by the Eq. 14, Eq. 16 and Eq. 19. The models emphasize the methods of non-weather AR with the application of high number of order coefficients. The model order coefficients are defined as,

$$a(p) = \{a(1), a(2), ..., a(p-1), a(p)\} \qquad (30)$$

From Table II it is found that the model order coefficients p, is equal to 175. Therefore, the forecast values are calculated by using the back shift operator as described in Eq. 27.

The forecast horizons simulated in this experiment are 1, 24, 36 and 168-hours ahead. For every forecast horizon, the value of forecast is determined and being validated with the actual (out of sample) data covering the period from 1.1.2007 to 31.12.2007. For the comparison purpose, the proposed method is compared with the multiplicative ARIMA [5], artificial neural network (ANN) [11], Burg's algorithm [12] and also with the naïve-random walk ARIMA (0,1,0) as a basic benchmarking.

4. APPLICATION RESULTS

Figure 9 depicts the model performance for the forecast duration of the year 2007. Meanwhile in Figure 10 describes the seasonal MAPE of 168-hours (one week) ahead forecast.
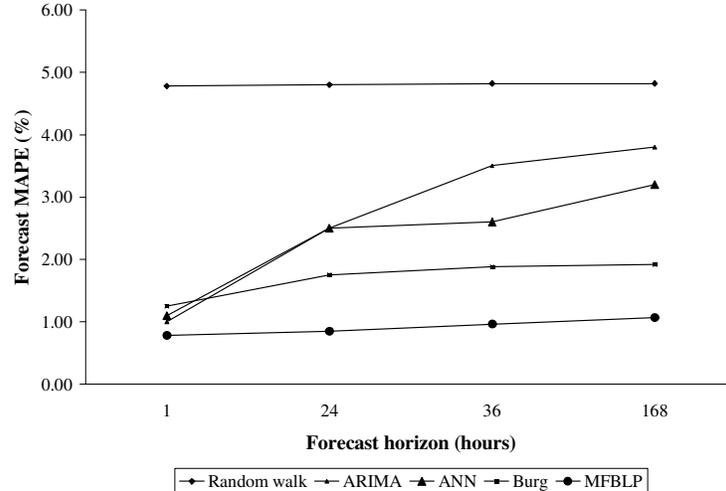
Figure 9: The MAPE plotted against forecast horizons for the NSW region of one year forecast period.
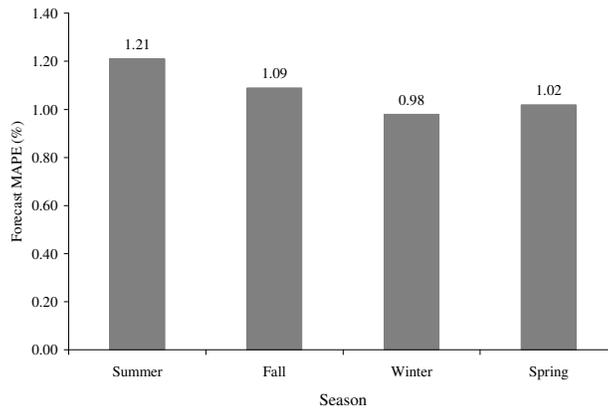
Figure 10: The MAPE plotted against seasons for the NSW one year forecast period.

It is clearly shown that from Figure 9 and 10, the methods of non-weather AR capable to describe its performance outstandingly. It is worth to investigate the model performance by weekly results. In order to obtain the results, only one selected season is chosen. The summer season is selected because, during this duration the load demand variations are high. Therefore, the forecast process is somehow difficult. It is assumed that other seasons show slightly simpler in forecasting and their MAPE results are considered lower than the summer season. Figure 11 describes the model accuracy for the random selected one week forecast during the summer season. The forecast horizons apply for 1, 24, 36 and 168-hours ahead. Meanwhile, the MAPE of the daily accuracy for the 168-hours ahead forecast horizon is depicted in Table III. The minimum and maximum for the 24-hours forecast MAPE during the week is highlighted in Table IV and Table V respectively.
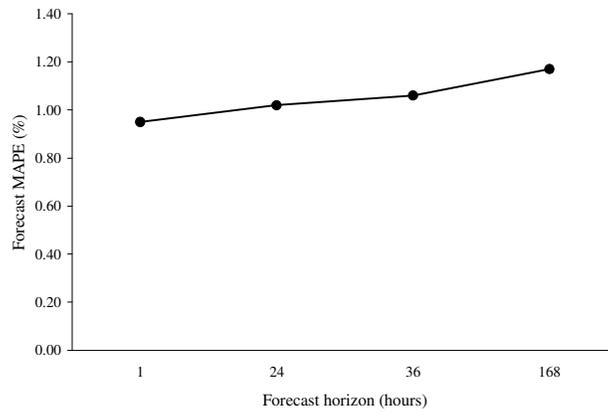
Figure 11: The forecast MAPE plotted against forecast horizons for the one week forecast duration in summer.

TABLE III: Daily MAPE for one week in the summer.

Forecast day	MAPE (%)
Day 1	1.25
Day 2	0.58
Day 3	0.78
Day 4	1.33
Day 5	1.38
Day 6	1.44
Day 7	1.34

TABLE IV: The hourly MAPE minimum of the week.

Forecast hour	MAPE (%)	Forecast hour	MAPE (%)
1	0.30	13	0.47
2	0.68	14	0.81
3	0.70	15	0.17
4	0.73	16	0.39
5	1.11	17	0.58
6	0.86	18	0.76
7	0.02	19	0.04
8	0.28	20	0.23
9	0.66	21	0.05
10	0.85	22	0.59
11	0.80	23	1.33
12	0.68	24	0.75

TABLE V: The hourly MAPE maximum of the week.

Forecast hour	MAPE (%)	Forecast hour	MAPE (%)
1	0.09	13	0.36
2	0.86	14	0.53
3	0.46	15	1.26
4	0.47	16	0.10
5	0.23	17	0.85
6	0.09	18	0.74
7	1.87	19	1.85
8	3.60	20	0.49
9	4.16	21	2.49
10	2.49	22	1.70
11	4.30	23	1.60
12	1.79	24	2.07

5. CONCLUSIONS

In this paper the non-weather autoregressive (AR) methods are proposed and discussed. The methods highlight the application of MFBLP and the estimation of model order coefficients by using the least squares approach. The experiment uses the hourly electricity load demand data gathered form NEMMCO. The estimation process uses two years hourly load data and one year hourly data for the forecast validation purpose. The forecast horizon for 1, 24, 36 and 168-hours ahead are examined to explain the workable of model forecast. It is worth mentioning that, there is no attempt to smooth the data prior to the estimation which is significant from the weather change and public holiday. This proves that the methods manage to find the accurate model coefficients to estimate and forecast the given data series.

The MAPE versus the forecast horizons for previous suggested approaches are also determined. The results show that the proposed MFBLP method is superior in others. For all forecast horizons MFBLP method performs remarkable results than others. Perhaps it could be a strong indicative of represent the performance of all suggested method. MFBLP shows a very promising result and perhaps, it can be used to solve any of the time series forecasting problems. In future work, it would be interesting to consider the public holidays and the obvious outliers prior to the estimation-forecast processes. This probably would enhance the forecast MAPE. Consequently, it would benefit the utility in reducing the operational cost and most importantly could avoid the energy being waste. Hence the daily usage power could be optimized significantly from less fuel being burnt and helping to reduce the pollution in our environment.

6. References

[1] Qia Ding, "Long-Term Load Forecast Using Decision Tree Method," *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES,* pp. 1541-1543, 2006.

[2] D. Asber, S. Lefebvre, J. Asber, M. Saad and C. Desbiens, "Non-parametric short-term load forecasting," *International Journal of Electrical Power & Energy Systems,* vol. 29, pp. 630-635, 10. 2007.

[3] D. W. Bunn and E. D. Farmer, *Comparative Models for Electrical Load Forecasting.* United States: John Wiley and Sons Inc.,New York, NY, 1985 Jan 01, pp. Pages: 186.

[4] D. W. Bunn, "Forecasting loads and prices in competitive power markets," *Proceedings of the IEEE,* vol. 88, pp. 163-169, 2000.

[5] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time Series Analysis: Forecasting & Control (3rd Edition).* New York: Prentice Hall; 3rd edition (February 28, 1994), 1994, pp. 592 pages.

[6] Z. Baharudin and N. S. Kamel, "AR modified covariance in one week ahead short term load forecast," *IASTED Proceeding Power and Energy Systems,* vol. 606-070, 2008.

[7] M. El-Telbany and F. El-Karmi, "Short-term forecasting of Jordanian electricity demand using particle swarm optimization," *Electric Power Systems Research,* vol. 78, pp. 425-433, 3. 2008.

[8] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas and P. Liatsis, "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models," *Energy,* vol. 33, pp. 1353-1360, 9. 2008.

[9] N. Amjady, "Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method," *Power Systems, IEEE Transactions on,* vol. 22, pp. 333-341, 2007.

[10] D. Asber, S. Lefebvre, M. Saad and C. Desbiens, "Modeling of Distribution Loads for Short and Medium-Term Load Forecasting," *Power Engineering Society General Meeting, 2007. IEEE,* pp. 1-5, 2007.

[11] G. A. Darbellay and M. Slama, "Forecasting the short-term demand for electricity: Do neural networks stand a better chance?" *International Journal of Forecasting,* vol. 16, pp. 71-83, 0. 2000.

[12] N. Kamel and Z. Baharudin, "Short term load forecast using Burg autoregressive technique," *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on,* pp. 912-916, 2007.

[13] NEMMCO. January 5, 2008). National electricity market management company; australia wholesale market and power system operator. *2008(January),*

[14] G. E. P. Box and G. C. Reinsel, *Time Series Analysis Forecasting and Control.* San Francisco: Holden Day, 1976,

[15] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting.* ,2nd ed.Springer-Verlag New York, Inc., 2002,

[16] N. S. Kamel, "High resolution sonar signal processing for detection of arrivals estimation," June 1993.

[17] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Review of linear least squares theory," in *Time Series Analysis Forecasting and Control* ,Third Edition ed.Anonymous Prentice Hall, 1994, pp. 286-304.

[18] M. H. Hayes, "The least squares method," in *Statistical Digital Signal Processing and Modeling* ,First ed.Anonymous USA: John Wiley & Sons, Inc., 1996, pp. 131-133.

[19] S. Haykin, *Adaptive Filter Theory.* New Jersey: Prentice Hall Inc, Englewood Cliff, 1986,

[20] S. de Waele and P. M. T. Broersen, "Order selection for vector autoregressive models," *Signal Processing, IEEE Transactions on,* vol. 51, pp. 427-433, 2003.

[21] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on,* vol. 19, pp. 716-723, 1974.

Abdallah SALEM, Zohra KARDOUS & Naceur BENHADJ BRAIEK

On the State Observer Based Stabilization of T-S Systems with Maximum Convergence Rate

Abdallah SALEM   abdallah.salem@isegf.rnu.tn
Laboratory Process Study and Control
Tunisia Polytechnic School,
BP. 743, 2078, La Marsa, Tunisia

Zohra KARDOUS   zohra.kardous@enit.rnu.tn
Laboratory Process Study and Control
Tunisia Polytechnic School,
BP. 743, 2078, La Marsa, Tunisia

Naceur BENHADJ BRAIEK   naceur.benhadj@ept.rnu.tn
Laboratory Process Study and Control
Tunisia Polytechnic School,
BP. 743, 2078, La Marsa, Tunisia

Abstract

This paper presents improved relaxed stabilization conditions and design procedures of state observer based controllers for continuous nonlinear systems in T-S model representation. First, the T-S model approach for nonlinear systems and some stabilization results are recalled. New stabilization conditions are obtained by relaxing those derived in previous works in this field. The asymptotic and exponential stabilization are considered with the maximization of the convergence rate. Design procedures for stabilizing T-S observer based controller using the concept of PDC (Parallel Distributed Compensation) and improved relaxed stabilization conditions are proposed.

Keywords: Continuous T-S systems, Observer based controller, PDC.

1. INTRODUCTION

The design of state feedback control, as well as the design of state observer, for nonlinear systems, has been actively considered during the last decades in many works using the Takagi-Sugeno (T-S) models [1], [2], [3], [4], [5].

The T-S model approach consists to construct nonlinear or complex dynamic systems that cannot be exactly modelled, by means of interpolating the behaviour of several LTI (Linear Time Invariant) submodels. Each submodel contributes to the global model in a particular subset of the operating space [2], [6], [7], [8].

Note that this modelling approach can be applied for a large class of physical and industrial processes as electrical machines and robot manipulators [9], [10], [11].

Abdallah SALEM, Zohra KARDOUS & Naceur BENHADJ BRAIEK

Recently, T-S observer based controller has attracted increasing attention, because it can provide a suitable solution to the control of plants that are complex and ill-defined and have immeasurable state variables [12], [13], [14], [15],

The T-S observer based controller has been considered to develop some systematic design algorithms to guarantee the stability and specific performances for the T-S model based systems [16], **[**17], [18].

The synthesis of the observer based controller can be considered as a convex problem and solved by Linear Matrix Inequalities (LMI) optimization techniques [19]. In spite of the advantages of LMI, the existence of a solution that satisfies the sufficient conditions is not guaranteed, especially, when the number of submodels increases or if many constraints are added such as control performance, the problem may become infeasible [20].

In attempt to avoid this situation, in some works relaxed stabilization conditions are derived to minimize the conservatism on LMIs [2], [3], [21]. However, the maximization of the convergence rate hasn't been considered.

This paper extends these works by proposing new relaxed conditions stabilization and design procedures for the observer based controller, using the concept of Parallel Distributed Compensation (PDC), with maximization of the convergence rate for the T-S model systems. An optimization tool is then used instead of LMIs.

This paper is organized as follows. Section 3 presents the structure of T-S models and recalls previous stability results. In Section 4, the observer design for T-S model is presented. In Section 5, we derive improved stabilization conditions and new design procedures of T-S observer based controller. To illustrate the proposed approaches a numerical example is considered in Section 6.

2. NOTATIONS

In this paper, we denote the minimum and maximum eigenvalues of a matrix X respectively by $\lambda_{\min}(X)$ and $\lambda_{\max}(X)$, the symmetric positive definite matrix X by $X > 0$ (the symmetric positive semidefinite matrix X by $X \geq 0$) and the transpose of X by X^T .
The following notations are also considered:

$$\sum_{i,j}^{n} x_i x_j = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j \ , \ \sum_{1\leq i<j}^{n} x_i x_j = \sum_{i=1}^{n}\sum_{i<j}^{n} x_i x_j \ \ \text{and} \ \ \sum_{i,j,k,l}^{n} x_i x_j x_k x_l = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} x_i x_j x_k x_l \ .$$

3. T-S MODEL AND STABILITY RESULTS

3.1 T-S model representation
A T-S model is based on the interpolation of several LTI local models as follows [2], [22]:

$$\begin{cases} \dot{x}(t) = \sum_{i=1}^{n} \mu_i\big(z(t)\big)\big(A_i x(t) + B_i u(t)\big) \\ y(t) = \sum_{i=1}^{n} \mu_i\big(z(t)\big)C_i x(t) \end{cases} \qquad (1)$$

where n is the number of submodels, $x(t)\in \mathbf{R}^p$ is the state vector, $y(t)\in \mathrm{R}^l$ is the output vector, $u(t)\in \mathrm{R}^m$ is the input vector, $z(t)\in \mathbf{R}^q$ is the decision variables vector and $\mu_i\big(z(t)\big)$ is the activation function.
$A_i \in \mathbf{R}^{p\times p}$, $B_i \in \mathbf{R}^{p\times m}$ and $C_i \in \mathbf{R}^{l\times p}$ are respectively the state matrix, the input matrix and the output matrix.

Different classes of models can be considered with respect to the choice of the decision variables and the type of the activation function.

In this paper, all the decision variables of the T-S model (1) are assumed measurable.

Each linear consequent equation represented by $\left(A_i x(t) + B_i u(t)\right)$ is called "subsystem" or "submodel".

The normalized activation function $\mu_i\left(z(t)\right)$ corresponding to the ith submodel is such that [6], [23], [24]:

$$\begin{cases} \sum_{i=1}^{n} \mu_i\left(z(t)\right) = 1 \\ \mu_i\left(z(t)\right) \geq 0 \quad \forall\, i \in \{1,...,n\} \end{cases} \tag{2}$$

3.2 Basic stabilization conditions

Let us consider the system (1) in its autonomous form, then we have:

$$\dot{x}(t) = \sum_{i=1}^{n} \mu_i\left(z(t)\right) A_i x(t) \tag{3}$$

Stabilization conditions of system (3) are derived using Lyapunov approach. So, the equilibrium of the T-S control system described by (3) is globally asymptotically stable if there exist a common positive definite matrix P such that [25]:

$$A_i^T P + P A_i < 0 \quad \text{for } 1 \leq i \leq n \tag{4}$$

4. OBSERVER DESIGN FOR T-S MODEL

In order to estimate the non measurable state variables of the T-S model (1), a T-S observer can be designed using PDC technique [7]. In this case, the global observer is obtained by interpolation of the local linear observers, associated to the different submodels.

For the T-S observer design, it is supposed that the decision variables are measurable and the T-S model of system (1) is locally detectable *i.e.* all the pairs $\left(A_i, C_i\right); i = 1,...,n$ are detectable.

The T-S observer is written as follows [2], [3]:

$$\begin{cases} \dot{\hat{x}}(t) = \sum_{i=1}^{n} \mu_i\left(z(t)\right)\left[A_i \hat{x}(t) + B_i u(t) + L_i\left(y(t) - \hat{y}(t)\right)\right] \\ \hat{y}(t) = \sum_{i=1}^{n} \mu_i\left(z(t)\right) C_i \hat{x}(t) \end{cases} \tag{5}$$

where $\hat{x}(t)$ is the estimated state vector and the activation function is the same than that used in the T-S model verifying (2).

One considers the following state estimation error defined as :

$$\varepsilon(t) = x(t) - \hat{x}(t) \tag{6}$$

From (1) and (6), the state estimation error dynamic is described by the following equation:

$$\dot{\varepsilon}(t) = \sum_{i,j} \mu_i\left(z(t)\right)\mu_j\left(z(t)\right)\left(A_i - L_i C_j\right)\varepsilon(t) = \sum_{i,j} \mu_i\left(z(t)\right)\mu_j\left(z(t)\right) R_{ij}\varepsilon(t) \tag{7}$$

where :

$$R_{ij} = A_i - L_i C_j \tag{8}$$

The T-S observer is required to satisfy $\hat{x}(t) \to x(t)$ when $t \to \infty$, this condition is guaranteed when the error $\varepsilon(t)$ converges to zero.

5. OBSERVER BASED CONTROLLER DESIGN FOR T-S MODEL

When the estimated state $\hat{x}(t)$ is available, we can consider the global control law with PDC technique as follows:

$$u(t) = -\sum_{i=1}^{n} \mu_i(z(t)) K_i \hat{x}(t) \qquad (9)$$

From (5), (6) and (9), one obtains:

$$\begin{cases} \dot{\hat{x}}(t) = \sum_{i,j} \mu_i(z(t)) \mu_j(z(t))(A_i - B_i K_j)\hat{x}(t) + \sum_{i,j} \mu_i(z(t)) \mu_j(z(t)) L_i C_j \varepsilon(t) \\ \hat{y}(t) = \sum_{i=1}^{n} \mu_i(z(t)) C_i \hat{x}(t) \end{cases} \qquad (10)$$

The augmented system is given by:

$$\dot{X}(t) = \sum_{i,j} \mu_i(z(t)) \mu_j z(t) H_{ij} X(t) = \sum_{i=1}^{n} \mu_i^2(z(t)) H_{ii} X(t) + 2\sum_{i<j} \mu_i(z(t)) \mu_j(z(t)) \left(\frac{H_{ij} + H_{ji}}{2} \right) X(t) \quad (11)$$

where:

$$X(t) = \begin{bmatrix} \hat{x}(t) \\ \varepsilon(t) \end{bmatrix}, \quad H_{ij} = \begin{bmatrix} A_i - B_i K_j & L_i C_j \\ 0 & A_i - L_i C_j \end{bmatrix} = \begin{bmatrix} G_{ij} & L_i C_j \\ 0 & R_{ij} \end{bmatrix} \qquad (12)$$

whith:

$$G_{ij} = A_i - B_i K_j \qquad (13)$$

The equation (11) makes appear the dominant submodels characterized by the matrices H_{ii} and

the coupled submodels characterized by the matrices $\left(\dfrac{H_{ij} + H_{ji}}{2} \right)$.

5.1 Asymptotic Stability

The T-S system described by (11) is globally asymptotically stable if there exist a common positive definite matrix P such that [2]:

$$\begin{cases} H_{ii}^T P + P H_{ii} < 0, & 1 \le i \le n \\ \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P\left(\dfrac{H_{ij} + H_{ji}}{2} \right) \le 0, & 1 \le i < j \le n \\ \forall (i,j) / \mu_i(z(t)) \mu_j(z(t)) \ne 0, \forall t \end{cases} \qquad (14)$$

One notes that the conditions (14) are conservative, because they require the stability of all the submodels (dominants and coupled). This result shows that the stabilization analysis of the T-S observer based controller system is reduced to a problem of finding a common matrix P. If n is large, it might be difficult to find a common P satisfying the conditions (14).

To reduce the conservatism, in the reference [2] relaxed conditions which require only the stability of the dominant submodels have been proposed. These conditions are recalled in the following theorem:

Theorem 1 [2]: Assume that the number of rules that fire for all t is less than or equal to s where $2 \le s \le n$. The equilibrium of the T-S system described by (11) is asymptotically stable in the large if there exist a common positive definite matrix P and a common positive semi definite matrix Q such that:

$$\begin{cases} H_{ii}^T P + P H_{ii} + (s-1)Q < 0, & 1 \le i \le n \\ \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P\left(\dfrac{H_{ij} + H_{ji}}{2} \right) - Q \le 0, 1 \le i < j \le n \\ \forall (i,j) / \mu_i(z(t)) \mu_j(z(t)) \ne 0, \forall t \text{ and } s > 1 \end{cases} \qquad (15)$$

5.2 Exponential stability

It is important to consider not only stabilization, but also other control performances such as speed of response, which is related to the decay rate, also called degree of stabilization and defined to be the largest $a > 0$ such that:

$$\lim_{t\to\infty} e^{at} \left\| X(t) \right\| = 0 \tag{16}$$

holds for all nonzero trajectories $X(t)$ of the system (11).

The condition (16) is equivalent to have:

$$\dot{V}(X(t)) \le -2aV(X(t)) \tag{17}$$

where :

$$V(X(t)) = X^T(t) P X(t) \tag{18}$$

is a quadratic Lyapunov function with $P > 0$.

The condition (17) has to be verified for all trajectories and leads to the inequality:

$$\left\| X(t) \right\| \le e^{-at} K(P) \left\| X(0) \right\| \tag{19}$$

where:

$$K(P) = \left(\frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} \right)^{1/2} \tag{20}$$

and $a > 0$ is the minimum decay rate.

The inequality (19) guarantees the global exponential stability of (11).

In [3], conditions of global exponential stability of system (11) have been derived and the minimum decay rate of the system has been characterized. These results are recalled in the following theorem:

Theorem 2 [3]: Suppose that there exist a common positive definite matrix P and a common positive semi definite matrix Q such that:

$$\begin{cases} H_{ii}^T P + P H_{ii} + \left(s - \dfrac{1}{2} \right) Q < 0, & 1 \le i \le n \\[2mm] \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P \left(\dfrac{H_{ij} + H_{ji}}{2} \right) - \dfrac{Q}{2} \le 0, & 1 \le i < j \le n \\[2mm] \forall (i,j) / \mu_i(z(t)) \mu_j(z(t)) \ne 0, \forall t \text{ and } s > 1 \end{cases} \tag{21}$$

Then the closed loop T-S model described by (11) is globally exponentially stable.

The minimum decay rate in this case is:

$$a_c = \frac{\lambda_{\min}(Q)}{4\lambda_{\max}(P)} \tag{22}$$

Note that the conditions (15) of theorem 1 and those (21) of theorem (2) can be unified in the following form:

$$\begin{cases} H_{ii}^T P + P H_{ii} + (s - \beta) Q < 0, & 1 \le i \le n \\[2mm] \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P \left(\dfrac{H_{ij} + H_{ji}}{2} \right) - \beta Q \le 0, & 1 \le i < j \le n \\[2mm] \forall (i,j) / \mu_i(z(t)) \mu_j(z(t)) \ne 0, \forall t \text{ and } s > 1 \end{cases} \tag{23}$$

with $\beta = 1$ or 0.5.

Remark 1: When β takes 1, one obtains the asymptotic stability conditions (15) and when β is replaced by 0.5, one obtains the conditions (21).

5.3 Main results: generalized and improved relaxed stability conditions

Let us note that two questions arise about the unified conditions (23):
- When no solution exists for $\beta = 0.5$, could the conditions (23) be relaxed to obtain an exponential stability solution?

- How to maximize the minimum decay rate when the exponential stability is guaranteed?

In this work, we have been interested by these two points and we have proved that the conditions (23) can be extended for any β such that $0 < \beta < 1$. Then the following theorem can be stated:

Theorem 3: Assume that the number of submodels simultaneously activated is s such that $2 \leq s \leq n$. The system described by (11) is globally exponentially stable, if there exist a common positive definite matrix P, a common positive semi definite matrix Q and a scalar $0 < \beta < 1$ such that:

$$\begin{cases} H_{ii}^T P + P H_{ii} + (s - \beta) Q < 0, \qquad 1 \leq i \leq n \\ \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P \left(\dfrac{H_{ij} + H_{ji}}{2} \right) - \beta Q \leq 0, \ 1 \leq i < j \leq n \\ \forall (i,j) \, / \, \mu_i(z(t)) \mu_j(z(t)) \neq 0, \forall t \ \text{and} \ s > 1 \end{cases} \qquad (24)$$

Then, the minimum decay rate is:

$$v_c = (1 - \beta) \frac{\lambda_{\min}(Q)}{2 \lambda_{\max}(P)} \qquad (25)$$

Proof. To prove the theorem 3, we use the following lemma 1:

Lemma 1 [2], [22]: Assume that the number of submodels simultaneously activated is s such that $2 \leq s \leq n$, then:

$$(s-1) \sum_{i=1}^{n} \mu_i^2 - 2 \sum_{1 \leq i < j}^{n} \mu_i \mu_j \geq 0$$

$$\text{and} \ s \sum_{i=1}^{n} \mu_i^2 \geq 1 \ \text{where} \ \sum_{i=1}^{n} \mu_i = 1, \ \mu_i \geq 0 \qquad (26) \qquad \square$$

Multiplying the first term of (24) by μ_i^2, the second by $2\mu_i\mu_j$ and adding up all terms for $i = 1$ to n, we get:

$$\sum_{i=1}^{n} \mu_i^2 \left[H_{ii}^T P + P H_{ii} + (s - \beta) Q \right] + \sum_{1 \leq i < j}^{n} \mu_i \mu_j \left[(H_{ij} + H_{ji})^T P + P (H_{ij} + H_{ji}) - 2\beta Q \right] < 0$$

Since $\sum_{i,j}^{n} \mu_i \mu_j H_{ij} = \sum_{i=1}^{n} \mu_i^2 H_{ii} + \sum_{1 \leq i < j}^{n} \mu_i \mu_j (H_{ij} + H_{ji})$, then the previous inequality is equivalent to :

$$\sum_{i,j}^{n} \mu_i \mu_j \left(H_{ij}^T P + P H_{ij} \right) + \left[(s - \beta) \sum_{i=1}^{n} \mu_i^2 - 2\beta \sum_{1 \leq i < j}^{n} \mu_i \mu_j \right] Q < 0$$

Assume that $H = \sum_{i,j}^{n} \mu_i \mu_j H_{ij}$, one obtains:

$$H^T P + P H + \left[s \sum_{i=1}^{n} \mu_i^2 - \beta \left(\sum_{i=1}^{n} \mu_i^2 + 2 \sum_{1 \leq i < j}^{n} \mu_i \mu_j \right) \right] Q < 0$$

or

$$H^T P + P H + \left(s \sum_{i=1}^{n} \mu_i^2 - \beta \right) Q < 0$$

From (26), we have $s \sum_{i=1}^{n} \mu_i^2 \geq 1$, and then it comes out:

$$H^T P + P H + (1 - \beta) Q \leq H^T P + P H + \left(s \sum_{i=1}^{n} \mu_i^2 - \beta \right) Q < 0$$

Consider the quadratic Lyapunov function $V(X) = X^T P X$ which is positive since $P > 0$. Then its derivative is given by:

$$\dot{V}(X) = \dot{X}^T P X + X^T P \dot{X} = X^T (H^T P + P H) X$$

Thus, it comes out : $\dot{V}(X) + (1-\beta) X^T Q X < 0$.

In the other hand, we have:

$$0 < \lambda_{\min}(P)\|X\|^2 \le X^T P X = V(X) \le \lambda_{\max}(P)\|X\|^2 \text{ and } 0 \le \lambda_{\min}(Q)\|X\|^2 \le X^T Q X \le \lambda_{\max}(Q)\|X\|^2$$

These two double inequalities yield the following one:

$$\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} V(X) \le X^T Q X \le \frac{\lambda_{\max}(Q)}{\lambda_{\min}(P)} V(X)$$

Since $0 < \beta < 1$, we get:

$$\dot{V}(X) + (1-\beta)\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} V(X) \le \dot{V}(X) + (1-\beta) X^T Q X < 0$$

Then, if there exist P and Q and a scalar β such that $0 < \beta < 1$ and the conditions (24) of theorem 3 are verified, the T-S system described by (11) is globally exponentially stable with minimum decay rate given by:

$$v_c = \left[(1-\beta)\frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)}\right] \text{ where } \left[(1-\beta)\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}\right] > 0 \qquad \square$$

Remark 2: The generalized conditions (24) of theorem 3 are:
-   less conservative than those of theorem 2 (corresponding to $\beta = 0.5$) if $0.5 < \beta < 1$;
-   more conservative than those of theorem 2 if $\beta < 0.5$.

and the minimum decay rate can reach important values greater than that obtained for $\beta = 0.5$.

5.4 Proposed procedures for observer based exponential stabilization of T-S system

The observer based stabilization of T-S systems can be leaded using two procedures : the first one is based on the separation principle to synthesise the observer and the controller gains and the second one aims to the maximisation of the decay rate.

5.4.1 Separation principle based procedure

Using the separation principle [3], [18], the conditions of theorem 3 (or those of theorem 2 when $\beta = 0.5$) are developed to determine the state feedback and the observer gains. The resulted procedure is summarized as the separated inequalities presented in the following theorem 4:

Theorem 4: If there exist positive symmetric definite matrices P_1, P_2, Q_1 and Q_2 such that:

$$\begin{cases} P_1 > 0, P_2 > 0, Q_1 > 0, Q_2 > 0 \\ G_{ii}^T P_1 + P_1 G_{ii} + (s-\beta) Q_1 < 0, \ 1 \le i \le n \\ \left(\dfrac{G_{ij} + G_{ji}}{2}\right)^T P_1 + P_1\left(\dfrac{G_{ij} + G_{ji}}{2}\right) - \beta Q_1 \le 0, 1 \le i < j \le n \\ R_{ii}^T P_2 + P_2 R_{ii} + (s-\beta) Q_2 < 0, \ 1 \le i \le n \\ \left(\dfrac{R_{ij} + R_{ji}}{2}\right)^T P_2 + P_2\left(\dfrac{R_{ij} + R_{ji}}{2}\right) - \beta Q_2 \le 0, 1 \le i < j \le n \\ \forall (i,j) / \mu_i(z(t)) \mu_j(z(t)) \ne 0, \forall t \text{ and } s > 1 \end{cases} \qquad (27)$$

then one can always find a quadratic Lyapunov function which prove the global exponential stability of the augmented system (11).

Proof. Consider the Lyapunov function $V(X) = X^T P X$ with the following structure of P and Q :

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & \sigma P_2 \end{bmatrix}, Q = \begin{bmatrix} Q_1 & 0 \\ 0 & \sigma Q_2 \end{bmatrix} \tag{28}$$

with $P_1 > 0$, $Q_1 > 0$, $P_2 > 0$, $Q_2 > 0$ and $\sigma \in {}^{+*}$.

To prove (27), one can proceed with the same proof in [3] and obtains that $\sigma \geq \max(\sigma_1, \sigma_2)$ where:

$$\sigma_1 = \frac{\lambda_{\min}\left[P_1 L_i C_i \left[R_{ii}^T P_2 + P_2 R_{ii} + (s-\beta)Q_2 \right]^{-1} \left(L_i C_i \right)^T P_1 \right]}{\lambda_{\max}\left(G_{ii}^T P_1 + P_1 G_{ii} + (s-\beta)Q_1 \right)}$$

$$\sigma_2 = \frac{\lambda_{\min}\left[\begin{matrix} P_1 \left(L_i C_j + L_j C_i \right) \begin{bmatrix} \left(R_{ij} + R_{ji} \right)^T P_2 + \\ P_2 \left(R_{ij} + R_{ji} \right) - 2\beta Q_2 \end{bmatrix}^{-1} \\ \left(L_i C_j + L_j C_i \right)^T P_1 \end{matrix} \right]}{\lambda_{\max}\left[\left(G_{ij} + G_{ji} \right)^T P_1 + P_1 \left(G_{ij} + G_{ji} \right) - 2\beta Q_1 \right]} \tag{29}$$

In order to simplify the resolution of the bilinear inequalities (27) we consider the following variables change:

$$X_1 = P_1^{-1}, \ M_i = K_i X_1, \ Y_1 = P_1^{-1} Q_1 P_1^{-1} \text{ and } N_i = P_2 L_i \tag{30}$$

Then, one obtains the following Generalized Eigenvalues Problem (GEVP) in X_1, Y_1, Q_2, P_2, M_i, N_i and β :

$$\begin{cases} X_1 > 0, \ P_2 > 0, \ Y_1 > 0, \ Q_2 > 0 \\ X_1 A_i^T - M_i^T B_i^T + A_i X_1 - B_i M_i + (s-\beta)Y_1 < 0, \ 1 \leq i \leq n \\ X_1 \left(A_i + A_j \right)^T + \left(A_i + A_j \right) X_1 - M_j^T B_i^T - M_i^T B_j^T - B_i M_j - B_j M_i - 2\beta Y_1 \leq 0, 1 \leq i < j \leq n \\ A_i^T P_2 + P_2 A_i - C_i^T N_i^T - N_i C_i + (s-\beta)Q_2 < 0, \ \ 1 \leq i \leq n \\ \left(A_i + A_j \right)^T P_2 + P_2 \left(A_i + A_j \right) - C_j^T N_i^T - C_i^T N_j^T - N_i C_j - N_j C_i - 2\beta Q_2 \leq 0, \ 1 \leq i < j \leq n \end{cases} \tag{31}$$

Note that for a given scalar β, the constraints (31) are Linear Matrix Inequalities (LMI) in X_1, Y_1, Q_2, P_2, M_i and N_i.

5.4.2 Maximization of the decay rate

From the generalized exponential stability conditions of theorem 3, one can look for a control law (9) maximizing the minimum decay rate v_c. This problem can be solved with respect to the scalar β such that $0.5 < \beta < 1$, the common positive definite matrix P, the common positive semi definite matrix Q, the state feedback gains $K_i, i = 1, ..., n$ and the observer gains $L_i, i = 1, ..., n$ as follows:

Abdallah SALEM, Zohra KARDOUS & Naceur BENHADJ BRAIEK

$$\begin{cases} \text{Maximize } v_c = \left[(1-\beta) \dfrac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)} \right] \text{ subject to:} \\[2mm] 0.5 < \beta < 1, \ P > 0, \ Q \geq 0 \\[2mm] H_{ij} = \begin{bmatrix} A_i - B_i K_j & L_i C_j \\ 0 & A_i - L_i C_j \end{bmatrix}, \ 1 \leq i \leq n \text{ and } 1 \leq j \leq n \\[4mm] H_{ii}^T P + P H_{ii} + (s-\beta)Q < 0, \qquad 1 \leq i \leq n \\[2mm] \left(\dfrac{H_{ij} + H_{ji}}{2} \right)^T P + P\left(\dfrac{H_{ij} + H_{ji}}{2} \right) - \beta Q \leq 0, 1 \leq i < j \leq n \\[4mm] \forall (i,j) \, / \, \mu_i(z(t))\mu_j(z(t)) \neq 0, \forall t \text{ and } s > 1 \end{cases} \qquad (32)$$

The maximization problem (32) can be solved using the optimization tools of MATLAB as the *fmincon*.

6. NUMERICAL EXAMPLE

We consider the T-S system composed by two subsystems studied in [21] and characterized by:

$$\begin{cases} \dot{x}(t) = \displaystyle\sum_{i=1}^{2} \mu_i(z(t))(A_i x(t) + B_i u(t)) \\ y(t) = \displaystyle\sum_{i=1}^{2} \mu_i(z(t))C_i x(t) \end{cases} \qquad (33)$$

where:

$$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}, \ A_1 = \begin{bmatrix} 2 & -10 \\ 1 & 0 \end{bmatrix}, \ A_2 = \begin{bmatrix} 15 & -10 \\ a & 0 \end{bmatrix},$$

$$B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ B_2 = \begin{bmatrix} b \\ 0 \end{bmatrix} \text{ and } C_1 = C_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

$$\mu_1 = \begin{cases} 1 - \dfrac{|x_1(t)|}{3}, \forall x_1(t) \in [-3 \ \ 3] \\ 0 \qquad \text{otherwise} \end{cases}, \qquad \mu_2 = 1 - \mu_1$$

a and b are the system parameters.

We assume that the number of the submodels simultaneously activated is $s = 2$.

With this example of multimodel system, we will show that the derived result in this paper can be used for two different goals:

- the enlargement of the system parameters variation area in which the stability of the multimodel system is guaranteed;
- the improvement of the decay rate of the system.

6.1 Enlargement of the system parameters variation area with the guaranteed stabilization

We study the multimodel stabilization of the system (33) with respect to the (a,b) parameters variation. The applied feedback control law is given by:

$$u(t) = -(\mu_1 K_1 + \mu_2 K_2)\hat{x}(t) \qquad (34)$$

where K_1 and K_2 are the local feedback gains determined such that the poles of the local controlled subsystems are placed to the values -1 and -2.

The local observers gains L_1 and L_2 are determined such that the poles of the local estimation error dynamic are placed to the values -2 and -4.

Thus we have $K_1 = \begin{bmatrix} 5 & -8 \end{bmatrix}$, $L_1 = \begin{bmatrix} 8 \\ 0.2 \end{bmatrix}$ and the gains K_2 and L_2 depend on the values of a and b parameters.

To determine the (a,b) area in which the stability of the controlled T-S system (33) is guaranteed, we vary the parameters a and b ($a > 0$ and $b > 0$) and then we verify the stability conditions of theorem 2 (for $\beta = 0.5$) and those of theorem 3 (for $0.5 < \beta < 1$).

The figures 1 and 2 show the feasible (a,b)-area corresponding to the conditions given respectively by theorem 2 and theorem 3. The mark (*) indicates the stability conditions feasibility.
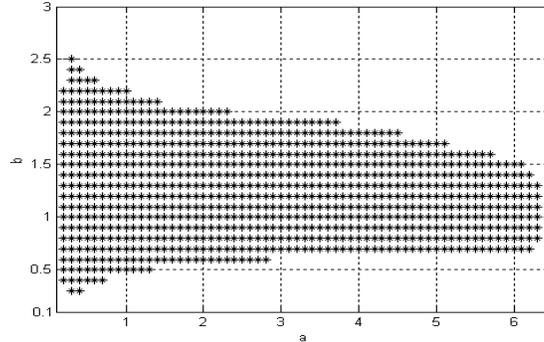

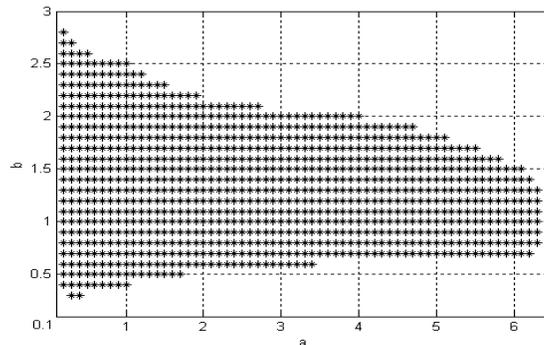**FIGURE 1:** Feasibility (a,b)-area for the stability conditions of theorem 2 ($\beta = 0.5$).

FIGURE 2: Feasibility (a,b)-area for the stability conditions of theorem 3 ($\beta = 0.9$).

From the figures 1 and 2, it can be noted that the theorem 3 ($0.5 < \beta < 1$) leads to relaxed conditions compared to those of theorem 2. Indeed the feasibility (a,b)-area of the figure 2 corresponding to the application of theorem 3 is clearly larger than that of the figure 1 obtained by the application of theorem 2.

6.2 Maximization of the decay rate

We consider now the system (33) with $a = b = 2$, and we search to maximize the decay rate

$$v_c = \left[(1-\beta) \frac{\lambda_{\min}(Q)}{2\lambda_{\max}(P)} \right]$$ subject to the optimization problem (32).

The maximization is leaded with respect to P and Q matrices, the feedback gains K_1 and K_2 and the observers gains L_1 and L_2 for different values of β. The obtained values are the following :

- For $\beta = 0.5$

$$K_1 = \begin{bmatrix} 55.53 & 31.19 \end{bmatrix}, \; K_2 = \begin{bmatrix} 46.27 & 18.492 \end{bmatrix}, \; L_1 = \begin{bmatrix} 18.69 \\ -0.722 \end{bmatrix}, \; L_1 = \begin{bmatrix} 45.088 \\ -3.542 \end{bmatrix}$$

$$v_{c\max} = 0.046$$

$$P = \begin{bmatrix} 50.621 & 39.337 & -19.409 & -0.338 \\ 39.337 & 482.69 & -19.278 & 70.198 \\ -19.409 & -19.278 & 36.821 & 20.259 \\ -0.338 & 70.198 & 20.259 & 125.31 \end{bmatrix}, \; Q = \begin{bmatrix} 186.62 & -99.644 & -12.877 & 10.071 \\ -99.644 & 196.66 & 14.082 & -10.328 \\ -12.877 & 14.082 & 95.853 & -0.0137 \\ 10.071 & -10.328 & -0.0137 & 94.084 \end{bmatrix}$$

- For $\beta = 0.55$

$$K_1 = \begin{bmatrix} 38.732 & 27.381 \end{bmatrix}, \; K_2 = \begin{bmatrix} 79.457 & 15.051 \end{bmatrix} \;, \; L_1 = \begin{bmatrix} 13.062 \\ -0.463 \end{bmatrix}, \; L_2 = \begin{bmatrix} 40.503 \\ -0.645 \end{bmatrix}$$

$$v_{c\max} = 0.122$$

$$P = \begin{bmatrix} 15.972 & 11.274 & -3.075 & 2.150 \\ 11.274 & 582.36 & -13.46 & 24.851 \\ -3.075 & -13.46 & 44.326 & 31.224 \\ 2.150 & 24.851 & 31.224 & 417.78 \end{bmatrix}, \; Q = \begin{bmatrix} 387.58 & 1.366 & 4.069 & 54.824 \\ 1.366 & 318.68 & 2.304 & 1.968 \\ 4.0694 & 2.304 & 367.87 & 24.644 \\ 54.824 & 1.968 & 24.644 & 371.51 \end{bmatrix}$$

It comes out from these results that the decay rate has been improved $(v_{c\max} = 0.122)$ with the parameter $\beta = 0.55$ in comparison with that obtained for $\beta = 0.5$ $(v_{c\max} = 0.046)$. This conclusion shows the importance of the new proposed stability conditions corresponding to $0.5 < \beta < 1$ and given by theorem 3.

The figure 3 shows the simulation results (x_1 and x_2 state variables behaviour) of the system (33) controlled with the multimodel law (34) in both cases of local gains: the gains values corresponding to $\beta = 0.5$ and the gains values corresponding to $\beta = 0.55$.
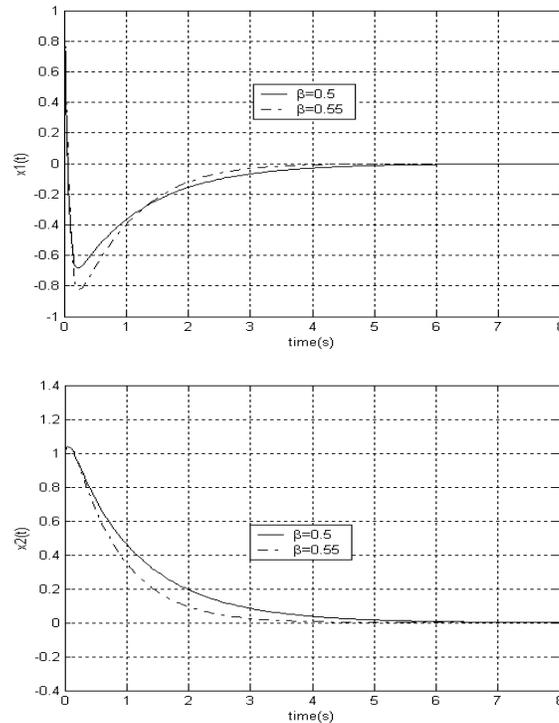

FIGURE 3: State variables behaviour for $\beta = 0.5$ and $\beta = 0.55$.

It appears on the simulation curves that the dynamic of the system is faster for the control law corresponding to $\beta = 0.55$, which confirm the conclusion obtained with the comparison study of the decay rates.

7. CONCLUSION

In this paper, improved approaches are suggested for the quadratic stabilization of observer based controlled T-S systems.

These approaches which aim to relax some results reached in previous works can be applied to the stabilization of nonlinear systems represented by T-S models, using the concept of parallel distributed compensation.

Two design procedures of the improved stabilization synthesis have been proposed. The first one is based on the separation principle between the controller and observer gains determination and it is formulated as a Generalized Eigenvalues Problem. The second one aims to the maximization of the decay rate of the exponential stability of the controlled system and it leads to an optimization problem.

A comparison study of the results derived in this work with previous ones has shown the importance of the proposed approaches for the enlargement of the availability domain of the stabilization conditions and the performance improvement of the stabilized nonlinear systems.

8. REFERENCES

1. N. Benhadj Braiek, F. Rotella. "State observer design for a class of nonlinear system". Journal of Systems Analysis, Modelling and Simulation (SAMS), 17:265-277, 1995

2. K. Tanaka, T. Ikeda, H. O. Wang. "Fuzzy regulators and fuzzy observers: relaxed stability conditions and LMI-based designs". IEEE Trans. on Fuzzy Systems, 6:250-265,1998

3. M. Chadli, D. Maquin and J. Ragot. "Observer-based controller for the Takagi-Sugeno models". In Proceeding of the IEEE International Conference of SMC. Tunisia, 2002

4. J. Zhang and M. Fei. "Analysis and design of robust fuzzy controllers and robust fuzzy observers of nonlinear systems". In Proceeding of the 6th World Compress on Intelligent Control and Automation. China, 2006

5. R. Murray-Smith and T. A. Johansen. "Multiple model approaches to modelling and control", Edition Taylor and Francis, (1997)

6. T. Takagi, M. Sugeno. "Fuzzy identification systems and its applications to modeling and control". IEEE Trans. Syst. Man Cybern., 15:116-132,1985

7. H.O. Wang, K. Tanaka, M. F. Griffin. "An approach to fuzzy control of nonlinear systems: Stability and design issues". IEEE Trans. Fuzzy Systems, 4:14-23, 1996

8. C.-S. Ting. "Stability analysis and design of Takagi-Sugeno systems". Information Sciences, 176:2817-2845, 2006

9. B.B. Kook, W. Chul Ham. "Adaptative Control of Robot Manipulator Using Fuzzy Compensator'", IEEE Transactions on Fuzzy Systems, 8:718-737, 2000

10. A. Salem, A. S. Tlili, N. Benhadj Braiek. "On the Polytopic and Multimodel State Observers of Induction Motors'", Journal of Automation and Systems Engineering (JASE), 2(4): 235-247, 2008

11. A. Salem, Z. Kardous, A.S. Tlili and N. Benhadj Braiek. "Observateur Multimodèle avec Stabilité Garantie de Systèmes à Paramètres Variants: Application à une Machine Asynchrone". Cinquième Conférence Internationale d'Électrotechnique et d'Automatique JTEA 2008. Hammamet, Tunisie, 02-04 Mai , 2008

12. Ch-Sh Ting. "An adaptive fuzzy observer-based approach for chaotic synchronization". International Journal of Approximate Reasoning, 39(1):97-114, 2005

13. A. Sala, C. Arino. "Asymptotically necessary and sufficient conditions for stability and performance in fuzzy control: Applications of Polya's theorem". Fuzzy Sets and Systems, 158:2671-2686, 2007

14. S.-W. Kau, C.-M. Yang, H.-W. Chen, W.- J. Xie, H.-J. Lee and C.-H. Fang. "Stabilization and Observer-Based H∞ Control for T-S Fuzzy Systems: An Improved LMI Approach". In Proceeding of the European Control Conference. Greece, 2007

15. J. Zhang, Y. Xia, P. Shi and J. Qiu. "H∞ output feedback control design for discrete-time fuzzy systems". Journal of Optimal Control Application and Methods, Published Online: 23 Oct 2008

16. X. Liu, Q. Zhang. "New approaches to H∞ controller designs based on fuzzy observers for fuzzy systems via LMI". Automatica, 39:1571-1582, 2003

17. C. Lin, Q.-G. Wang, T. H. Lee. "Improvement on observer-based H∞ control for T-S fuzzy systems". Automatica, 41:1651-1656, 2005

18. X. Ma, Z. Sun, Y. He. "Analysis and design of fuzzy controller and fuzzy observer". IEEE Trans. Fuzzy Syst., 6:41-51, 1998

19. J. Kim, D. Park. "LMI-based design of stabilizing fuzzy controllers for nonlinear systems described by Takagi-Sugeno model". Fuzzy Sets Syst., 122:73-82, 2003

20. L. Luoh. "New Stability analysis of fuzzy systems with robust approach". Math. Comput. Simul., 59:335-340, 2002

21. E. Kim, H. Lee. "New approaches to relaxed quadratic stability condition on fuzzy control systems". IEEE Transactions on Fuzzy Systems, 8:523-533, 2000

22. M. Chadli, D. Maquin and J. Ragot. "Stabilisation of Takagi-Sugeno models with maximum convergence rate". In Proceeding of the IEEE International Conference on Fuzzy Systems. Hungary, 2004

23. Z. Kardous, N. Benhadj Braiek and A. Elkamel. "On the residue formulation in multimodel control". In Proceeding of the IEEE International Conference on SMC. Tunisia, 2002

24. Z. Kardous, N. Benhadj Braiek, A. Elkamel and P. Borne, "On the quadratic stabilization in discrete multimodel control". Conference on Control Applications CCA'2003. Istanbul, 2003

25. K. Tanaka, M. Sugeno. "Stability analysis and design of fuzzy control systems". Fuzzy Sets Systems, 45:136-156, 1992

A Comparative Evaluation and Design of Input Current Waveshaping Methods for Rectifier

Hussein A Kazem   h.kazem@soharuni.edu.om
Faculty of Engineering/Electrical &
Computer Department
Sohar University
Sohar, PO Box 44, Post code 311,
Sultanate of Oman

Abstract

This paper presents a comparative evaluation of passive current waveshaping methods for single-phase rectifier. The simulation results show that the new method can further lower the input current components and THD_i of a single-phase diode rectifier as compared with the three passive current waveshaping methods by P. D. Ziogas [5], Ji Yanchao [4] and H. A. Kazem [6]. The relevant input current and voltage waveforms, the input current total harmonic distortion and the input power factor value are derived from the computer simulation or theoretical analysis. A design example is provided and simulation results have been verified on 500 mW experimental models.

Keywords: Current Waveshaping, Rectifier, Power Factor Corrections, Harmonics.

1. INTRODUCTION

The major polluters of power system harmonics are the rectifier circuits. The economic advantages that these rectifiers presently enjoy in the marketplace can change overnight by the imposition of stricter harmonic standards. With this threat in mind, many researchers have started to probe into the possibilities of passive and active waveshaping methods. Single-phase rectifiers (Fig. 1) have the problems of poor power quality in terms of injected current harmonics, resultant voltage distortion and poor power factor at input ac mains and slowly varying rippled dc output at load end, low efficiency, and large size of ac and dc filters, [1]-[2].
A growing number of current waveshaping methods applied to single-phase rectifier are now available including active, passive and hybrid methods, [3]. Among the proposed passive waveshaping methods, the Ziogas method in [4] is superior to the others in reducing the input current harmonic components and improving the input power factor. On the bases of the Ziogas method, Ji Yanchao [5] proposed an improved method, which can further improve the input current waveform and therefore has a better input power factor. On the bases of the two methods Hussein [6] proposed anther improvement to reduce THD_i. This paper presents a comparative evaluation of the three previews method.

2. CONVENTIONAL SINGLE-PHASE RECTIFIER

The nature of rectifiers either it is conventional (Fig. 1) or switch mode types, all of them contribute to high THD_i and low efficiency to the power system. However, this method has the disadvantage of generating pulsed as line currents drawn from the ac distribution network. Under the condition that the input peak voltage is 1 pu (12 volt) and fundamental frequency is 1 pu (50 Hz), the relevant input voltage and current waveforms of the conventional rectifier and its input

current Fourier analysis results are obtained as shown in Fig. 2, which makes it clear that the input current contains a very large quantity of harmonics.
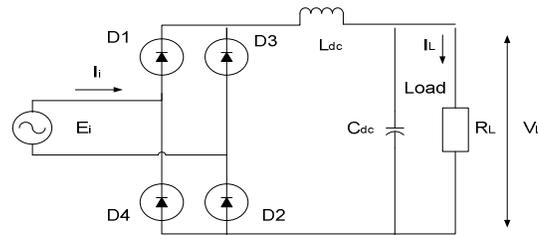

FIGURE 1: Conventional single-phase diode rectifier topology.

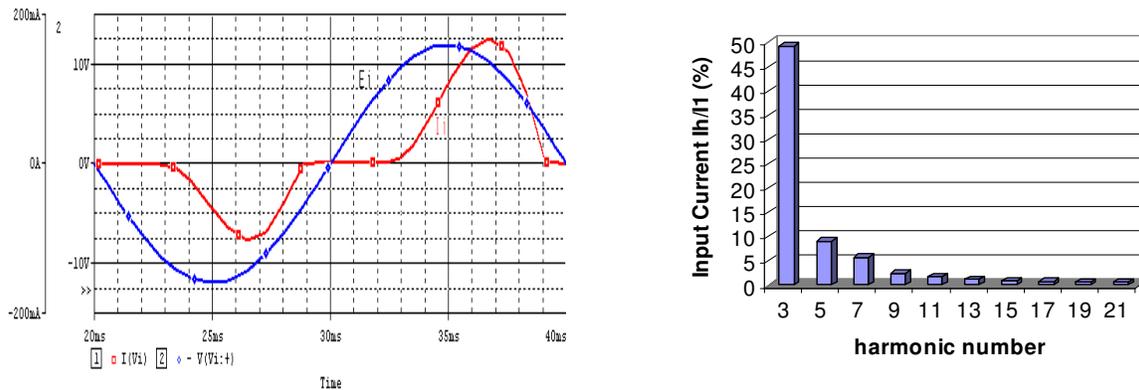(a)   (b)

FIGURE 2: a. simulated waveform of input current and voltage with conventional rectifier topology, b. Harmonic spectrum of input current with conventional rectifier.

Due to the presence of the considerable distortion power, the power factor of the conventional topology is very low. It is found that the power factor to deliver 1.0 pu power P_r is only about 0.698. This conventional method has many disadvantages, including:

1) High input current harmonic component and THD_i is 55.16% also 3rd harmonic is 49.3%;

2) Low input power factor, the maximum value of which to deliver 1.0 pu P_r is only about 0.698;

3) Low conversion efficiency.

3. PASSIVE WAVESHAPING FOR SINGLE-PHASE RECTIFIER

In 1990 passive waveshaping method proposed by P.D. Ziogas (Fig. 3), which uses an input L_r-C_r parallel resonant tank to remove the third harmonic component from the input current. The input power factor increases because the third harmonic component is the main reason of the low input power factor. The relevant input current and voltage waveforms of the novel diode rectifier and Fourier analysis result of the input current waveform are illustrated in Fig. 4. It is shows that the parallel resonant tank can effectively eliminate the third harmonic component and reduce the rectifier input THD_i. The advantages of the novel method over the conventional method include:

1) Lower the input current THD$_i$, which is about 30.26% also 3rd harmonic is 11.51%.

2) Higher input power factor, the maximum value of which to deliver 1.0 pu P$_r$ is only about 0.931.

3) Increase efficiency of the rectifier.

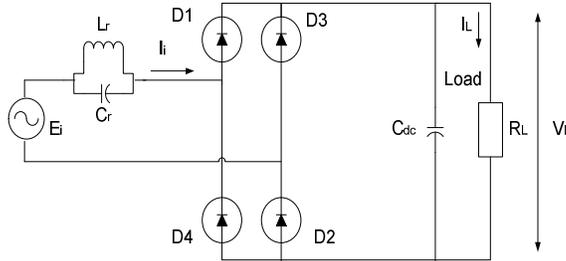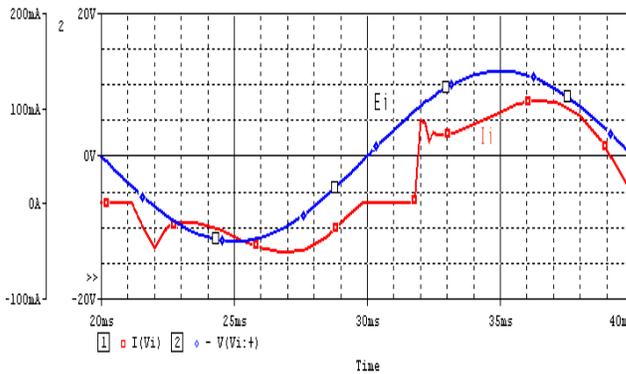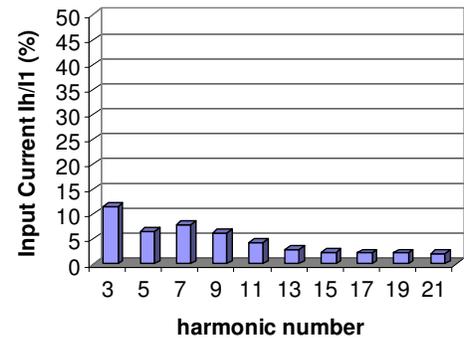**FIGURE 3:** Ziogas single-phase diode rectifier topology.

(a)   (b)

FIGURE 4: a. simulated waveform of input current and voltage with Ziogas rectifier topology, b. Harmonic spectrum of input current with Ziogas rectifier.

4. IMPROVED METHOD-1 FOR SINGLE-PHASE RECTIFIER

To further lower the input current THD$_i$ of the novel diode rectifier, Ji Yanchoa [4] proposed improved method (Fig. 5) by place a capacitor C$_b$ in parallel between the parallel resonant tank and the rectifier bridge. When C$_r$ has a value of 7.93µF or 0.39 pu and L$_r$ is 141mH or 0.31 pu, the value of C$_b$ is selected such that the input power factor at the rated output power reaches its peak value. The input current and voltage waveforms and variation of the input power factor with the value of C$_b$ at the rated output power is shown in Figs. 6 & 7 respectively. It is clear from Fig. 7 that for the rated output power, the value of C$_b$ should be selected to be 2.5µF or 0.11 pu. Under this condition, the relevant input power factor approaches its maximum value of 0.967.

The advantages of the improved method-1 over the novel and conventional methods include:

1) Lower the input current THD$_i$, which is about 27.80% also 3rd harmonic is 9.59%.

2) Higher input power factor, the maximum value of which to deliver 1.0 pu P$_r$ is only about 0.935.

3) Increase efficiency of the rectifier.

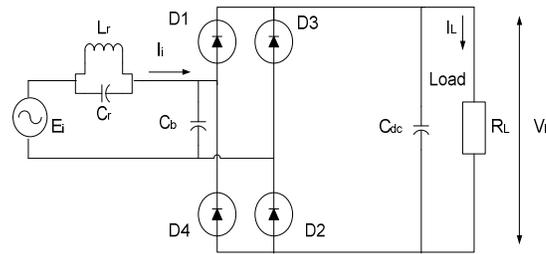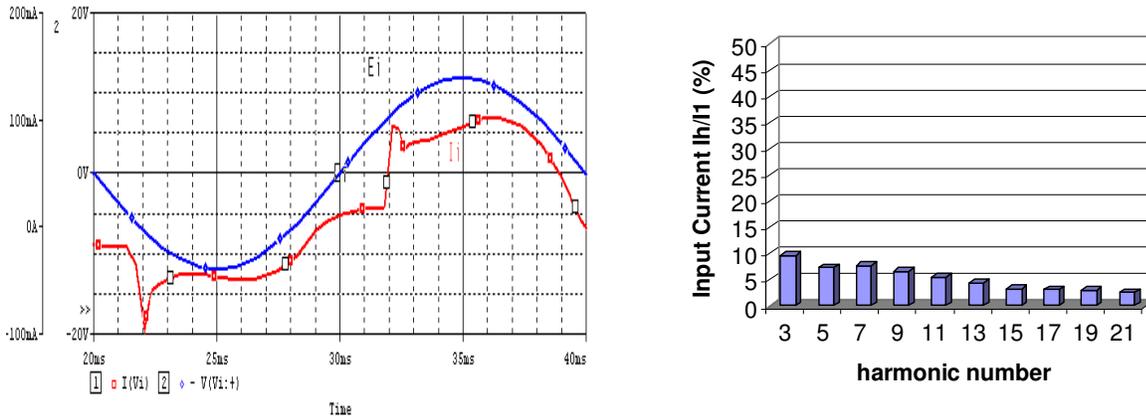**FIGURE 5:** Improved-1 single-phase diode rectifier topology.

(a)   (b)

FIGURE 6: a. simulated waveform of input current and voltage with improved-1 rectifier topology, b. Harmonic spectrum of input current with improved-1 rectifier.
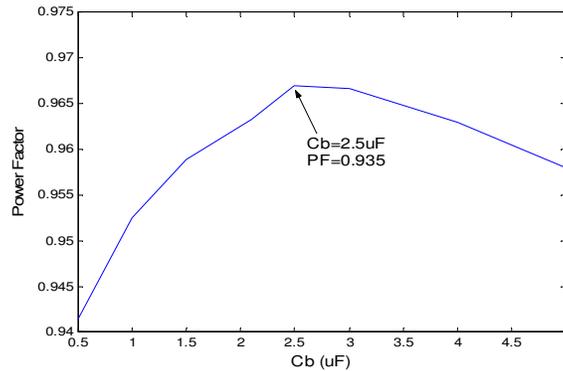

FIGURE 7: Variation of PF with C_b at rated load

5. IMPROVED METHOD-2 FOR SINGLE-PHASE RECTIFIER

To further lower the input current THD_i of the novel and improved method-1 diode rectifier, an improved method-2 is proposed by place inductance L_o in series with the output of the rectifier (Fig. 8). When C_r has a value of 7.93μF or 0.39 pu, L_r is 141mH or 0.31 pu, and C_b 2.5μF or 0.11 pu, L_o is selected such that the input power factor at the rated output power reaches its peak value.

The input current & voltage waveforms and harmonic spectrum and variation of the input power factor with the value of L_o at the rated output power are shown in Figs. 9 & 10 respectively. It is

clear from Fig. 10 that for the rated output power; the value of L_o should be selected to be 0.35 mH or 0.76×10^{-3} pu. Under this condition, the relevant input power factor approaches its maximum value of 0.969.

The advantages of the improved method-2 over improved method-1, novel and conventional methods include:

1) Lower the input current THD_i, which is about 25.03% also 3rd harmonic is 8.40%.

2) Higher input power factor, the maximum value of which to deliver 1.0 pu P_r is only about 0.969.

3) Increase efficiency of the rectifier.

FIGURE 8: Improved-2 single-phase diode rectifier topology.
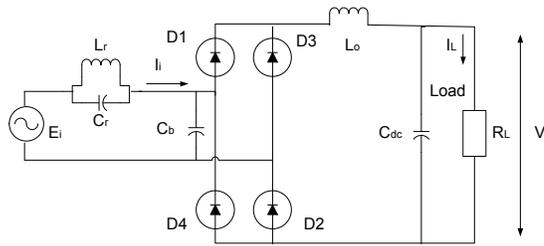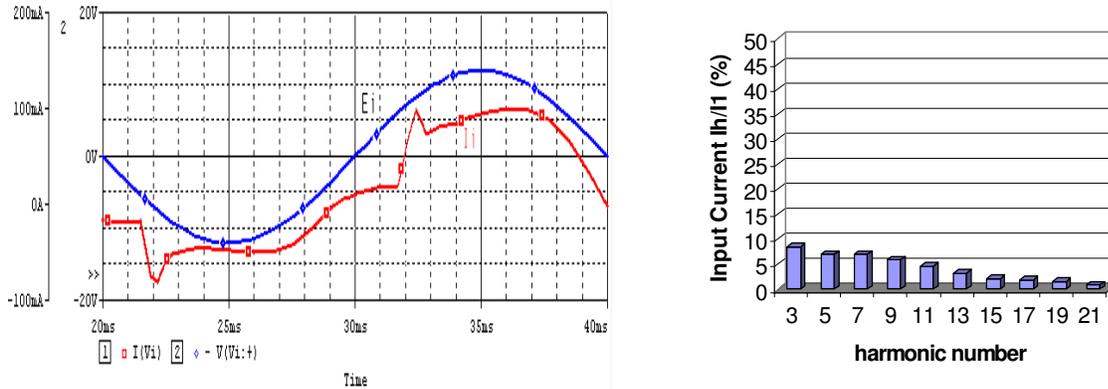

(a)   (b)

FIGURE 9: a. simulated waveform of input current and voltage with improved-2 rectifier topology, b. Harmonic spectrum of input current with improved-2 rectifier.
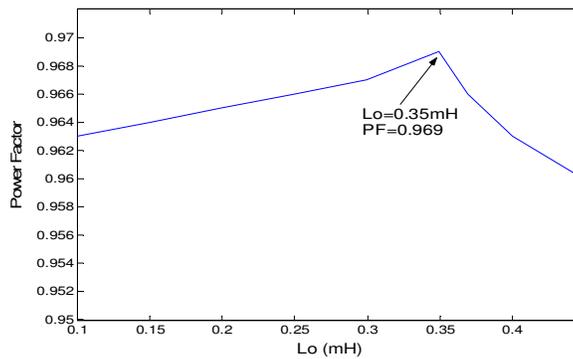

FIGURE 10: Variation of PF with L_o at rated load.

Table 1 illustrates a comparison between the four cases. It is clear seen that improved method-2 have better power factor and less THD_i.

	conventional	novel	Improved-1	Improved-1
PF	0.698	0.931	0.935	0.969
THDi	55.16%	30.26%	27.80%	25.03%
3rd	49.3%	11.51%	9.59%	8.40%

TABLE 1: PF and THDi for the four cases.

6. DESIGN EXAMPLE AND EXPERIMENTAL RESULTS

A. Design example

To illustrate the validly of the simulation analysis in the previous sections, the following design example is presented. The rectifier has the following specifications:
V_s = 8.5 rms ≡ 1.0 pu;
P_r = 500 mW ≡ 1.0 pu;
Output voltage V_L ripple = 5%.
From these values
1 pu angular frequency = $2\pi f$ = 314 rad/sec;
1 pu current = 0.5/8.5 = 59 mA
1 pu impedance = 8.5/0.059 = 143.8Ω
1 pu inductance = 143.8/314 =457.8mH
1 pu capacitance = 1/(143.8×314) = 22.13μF

The value of DC Filter Capacitor C_o
The value of C_o for 5% harmonic on the capacitor voltage at the optimum operating from [3] is given by

$$C_O = \frac{100 \times I_{o,2}}{2\omega V_{L,o} \times 5\%}$$ (1)

Where

$V_{L,o}$: The dc average value of the output voltage.

$I_{o,2}$: The rms value of the 2nd harmonic output current.

The value of C_o (assuming 5%) can be calculated by using (1). Its value is 102.8μF or 4.61 pu.

The value of ac Compensation Capacitor C_b
From section-IV: C_b = 0.11×22.13=2.5μF.

The value of dc Filter Inductor L_o
From the previews section-V: L_o = 0.00076×457.8= 0.35 mH.

B. Experimental Results

To verify the predicted results obtained in the previews sections, a 500 mW experimental diode rectifier was implemented with the following circuit parameters: C_r= 8μF, L_r=150mH, C_b=2.2μF, L_o=0.3mH, C_o=100μF, R_L=150Ω. The experimental waveforms of the input voltage and current are shown in Figs. 11, which are obtained under the condition that the output power is 1.0 pu, the input rms voltage is 8.5V and its frequency is 50Hz. Evaluation of Figs. 11 and 9a shows that the simulation results are in close agreement with the experimental results.
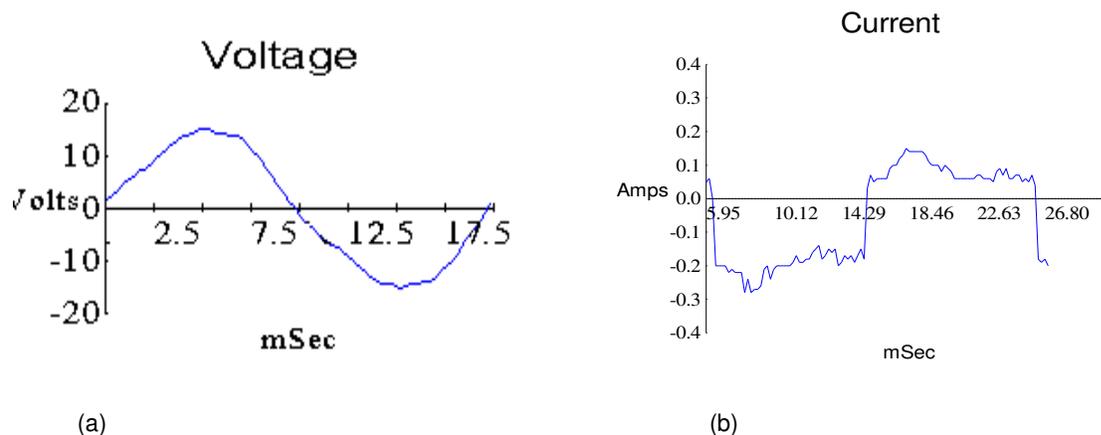
(a)   (b)

FIGURE 11: an Experimental input voltage waveform of the improved method-2 rectifier, b.
Experimental input current waveform of the improved method-2 rectifier

7. CONSLUSION & FUTURE WORK

A novel passive input current waveshaping method for single-phase rectifiers has been proposed and the relevant waveforms of the input current and voltage obtained from computer simulation and the spectrum of the input current obtained from Fourier analysis have been shown in this paper. Also, an improved method-1 to reduce THD_i and increase power factor was developed by installing a parallel capacitor C_b between the parallel resonant tank and the rectifier bridge. For further reduction in the input current THD_i and increase power factor achieved by install a series inductor with the output of the rectifier. The validity of the simulation results and the feasibility of the improved method have been verified on a 500 mW laboratory prototype unit.

8. REFERENCES

[1] Omar Stihi and Boon-Tech Ooi, "*A Single-Phase Controlled-Current PWM Rectifier*", IEEE Transactions on Power Electronocs, Vol. 3, No. 4, 1988, pp. 453-459.

[2] Hussein A. Kazem, Abdulhakeem A. Alblushi, Ali. S. Aljabri & Khmais H. Alsaidi, "*Simple and Advanced Models for Calculating Single-Phase Diode Rectifier Line-Side Harmonics*", Transactions on Engineering, Computing and Technology, Vol.9, November 2005, pp. 179-183, ISBN 975-98458-8-1.

[3] Hussein A. Kazem, "*Input Current Waveshaping Methods Applied to Single-Phase Rectifier*", International Conference on Electrical Machines and Systems 2007, October 8-11, 2007, Seoul, Korea.

[4] Ji Yanchao, Liang Xiaobing, Liu Zhuo, Jin Jisheng & Liu Xinhua, "*An improved Passive Input Current Waveshaping Method for Single-Phase Rectifier*", Industrial Electronics, Control and Instrumentation, IEEE IECON, Vol. 2, 1996, pp 695-699.

[5] Atluri Rama Prasad, Phoivos D. Ziogas & Stefanous Manias, "*A Novel Passive Waveshaping Method for Single-Phase Diode Rectifiers*", IEEE Transactions on Industrial Electronics, Vol. 37, No. 6, December 1990, pp 521-530.

[6] Hussein A. Kazem, "*An Improved Method of Passive Input Current Waveshaping for Single-Phase Rectifier*", 5th International Conference CPE 2007, May, 2007, Poland, pp. 1-4.

Grid Free Lagrangian Blobs Vortex Method with Brinkman Layer Domain Embedding Approach for Heterogeneous Unsteady Thermo Fluid Dynamics Problems

Carmine GOLIA
carmine.golia@unina2.it

Faculty of Engineering, Department of
Aerospace and Mechanical Engineering
Second University of Naples
Via Roma 29, Aversa, 81031 Italy

Bernardo BUONOMO
bernardo.buonomo@unina2.it

Faculty of Engineering, Department of
Aerospace and Mechanical Engineering
Second University of Naples
Via Roma 29, Aversa, 81031 Italy

Antonio VIVIANI
antonio.viviani@unina2.it

Faculty of Engineering, Department of
Aerospace and Mechanical Engineering
Second University of Naples
Via Roma 29, Aversa, 81031 Italy

Abstract

Modeling unsteady thermal–viscous flows inside/around complicated geometries containing multiphase sub-systems (fluid–porous–solid) and multi-physics phenomena (diffusion; forced/free/mixed convection; time variations of velocity, temperatures and heat fluxes sources; still and moving bodies) is an ambitious challenge in many applications of interest in science and engineering. Scope of this exploratory work is to investigate if the combination of a grid free Lagrangian Blob method with a Brinkman layer domain embedding approach can be useful for the preliminary analysis of heterogeneous unsteady thermal buoyant problems, where easiness, readiness, short computational times, good qualitative and sufficient quantitative accuracy are the most important aspects. In this work we couple a grid free unsteady Lagrangian Thermal-Vortex Blob method with a double penalization method that considers solid bodies contoured by a fictitious buffer thin boundary layer, described by a porous Brinkman model. The model problem in this study is the interaction of a thermal buoyant plume with a solid body, still or in motion. Both solid body and Brinkman boundary layer are described by volume penalization applied by an unsteady mask method. After a description of this novel approach, preliminary analyses for validations are presented for various thermal buoyant steady/unsteady problems relative to thermal and thermal-vortical patches, fixed, free and in presence of a still or moving body. Comments on pros, contras and further work, conclude the paper.

Keywords: Buoyant Plume/Body Interaction, Lagrangian Methods, Thermal/Vortex Blobs, Grid Free, Volume Penalization, Brinkman Domain Embedding

1. INTRODUCTION

The Brinkman domain embedding approach introduces a penalty term [treated as a continuous] directly in the set of global equations over the whole flow domain, this term takes different order of magnitudes according to the particular sub domain. This approach makes the viscous body analysis almost independent on the grid domain, coordinate system, since it allows the analysis of heterogeneous problems with quite complicated geometries by using simple mask functions, and it can be used almost independently on the solution method used (FEM, FDM, Spectral method, etc.).

The advantages of this formulation are:

- Single (Cartesian or other) mesh approach to the global field, avoiding body-fitted unstructured mesh,
- Physical description of the sub domain (shape, position, size, dynamic and diffusive parameters) by simple mask functions [that can be still or can move according to the dynamics of the body].
- The warranted continuity of the diffusive fluxes across the interfaces implicitly embedded in the global equation. This automatically takes account of conjugated heat transfer problems and of coupling problems among heterogeneous media.
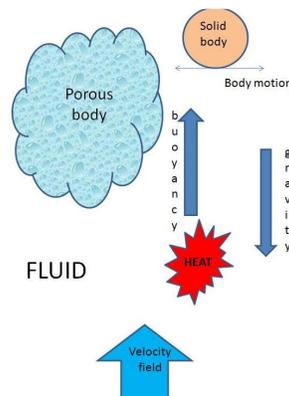

FIGURE 1: Embedding Domains.

This approach, proposed by Arquis & Caltagirone [1] and validated in Ref.s [2], [3], [4], [5], was widely used as penalization techniques for the simulation of flow around solid bodies in the contest of grid dependant computation methods, i.e. adaptive wavelet method by Schneider *et al*. ([6], [7], [8], [9], [10]); hybrid wavelet method by Vasilyev and Kevlahan [11], Kevlahan *et al.* [12]; spectral methods by Kevlahan and Ghidaglia [13].

The method has also been successfully used to simulate transitional and turbulent flows by an array of cylinders using Cartesian grids by Bruneau and Mortazavi [14] and extended to compressible flows by Chiavassa and Donat [15].

2. PRELIMINARIES

The physical idea of the penalization technique is to model the obstacle as a porous medium with porosity tending to zero. This corresponds to a Brinkman-type model with a permeability that varies according to the space domain. We can consider the whole field, treated as a global Brinkman porous medium, with properties locally varying according to the particular body treated as a specific porous media. The Navier-Stokes system is modified by adding a supplementary

term with the idea of forcing the velocity to satisfy the no-slip condition on the boundary of the obstacle. The new system is then solved in an obstacle–free computational domain.

Momentum (Navier-Stokes):

$$\frac{\partial \underline{V}}{\partial t} + \underline{V} \bullet \nabla \underline{V} = -\nabla p + \frac{1}{Re} \nabla \bullet \left[\mu \left(\nabla \underline{V} + \nabla \underline{V}^t \right) \right] - \left[\frac{\mu \left(\underline{V} - \underline{V}_{body} \right)}{Re\,Da\,K} \right] + \frac{1}{\rho_{rif}} (\Delta \rho)_T \underline{g} + \underline{f} \tag{1}$$

The value of the local specific permeability, K, models the specific embedded medium, i.e.:

$$K(t,\underline{r}) = \begin{cases} K_f & \to +\infty & \text{if } \underline{r} \in \text{fluid region} \\ K_p & & \text{if } \underline{r} \in \text{porous region} \\ K_s & \to 0^+ & \text{if } \underline{r} \in \text{solid region} \end{cases} \tag{2}$$

- In case of fluid, the "Darcy drag" is negligible with respect to the other terms and (1) reduces to the classical Navier Stokes equation.
- In case of solid, we consider a medium with porosity almost unitary and permeability nearly zero.
- In case of porous region, the "Darcy drag" term takes the role of a penalty term that imposes low velocity, so that the convective terms become negligible and (1) reduces to the classical Brinkman equation used in porous media theory:

$$\nabla p - \frac{1}{Re} \nabla \bullet \left[\mu \left(\nabla \underline{V} + \nabla \underline{V}^t \right) \right] + \left[\frac{\mu (\underline{V})}{Re\,Da\,K_p} \right] = \frac{1}{\rho_o} (\Delta \rho)_T \underline{g} + \underline{f} \tag{3}$$

Note that the prolongation of the flow parameters [pressure, velocity and temperature inside the solid media] makes the continuity conditions on the velocity and temperature field and diffusive fluxes satisfied on and across the boundaries.

Problems may arise from the fact that in the regions where penalization constraints are imposed, the formulation is first order [as in all penalizations] and the equations become very stiff to integrate in time.

2.1 Brinkman double penalization method

Angot *et al.* [3] first proposed a "double penalization method" where a solid body is bounded by a porous Brinkman layer. This approach was used by Bruneau and Mortazavi [16, 17] to compute the viscous flow around a ground vehicle surrounded by a thin layer of porous material. Carbou [18], [19] explained the good performances of this method characterizing the boundary layer that appears over the obstacle, and in [20] by using a BKW method performed an asymptotic expansion of the solution when a little parameter, measuring the thickness of the thin layer and the inverse of the penalization coefficient, tending to zero, proved that the method is equivalent to the computations using greater thickness model by a standard Brinkman Model

2.2 Lagrangian Blob Method

The blob concept [21] is based on an attempt to make discrete the free space Dirac representation of a generic function " f " at the location (x) and time (t):

$$f(\underline{x},t) = \int f(\underline{x}',t) \, \delta(\underline{x} - \underline{x}') \, d\underline{x}' \tag{4}$$

This representation can be made discrete, in an Eulerian formulation, as follows:

$$f(\underline{x}_p,t) = \int_{\mathscr{V}} f(\underline{x}_q,t) \, W\left(\frac{\underline{x}_p - \underline{x}_q}{h}, h \right) d\mathscr{V} \tag{5}$$

where the kernel function W($\underline{r}$, h) satisfies given properties in order that, in the limit h→0, the two representations must coincide (h can be regarded as the grid size).

The Lagrangian Blob formulation is aimed to represent the value of the given field function " f " for the p-th particle that at time (t) is located in the position (x_p). This requires further modifications of eq. (5) by considering, for a finite number of particle within a cluster around the pth-particle, the following definition:

$$f\left(\underline{x}_p, t\right) = \sum_{\substack{q \in \text{Cluster} \\ \text{around} p}} f\left(\underline{x}_q, t\right) \; W\left(\frac{x_p - x_q}{h}, h\right) \Delta Vol_q \tag{6}$$

where ΔVol_q is the finite elementary volume associated with the q-th particle.
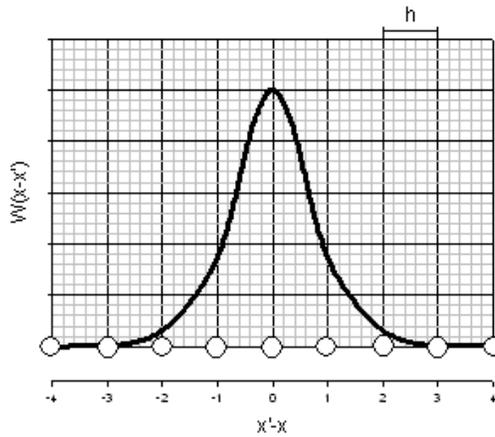

FIGURE 2: Kernel Function for the Blob Representation.

It is easy to observe that the approximations deriving from the Lagrangian Blob formulation depends on three facts:

1. The choice of the kernel function W(r, h) that must be rapidly decay with r (either Gaussian like or a polynomials compact over x/r usually equal to 1-2),
2. The computation of the integral in eq. (5) (approximated by a simple summation),
3. The finite number of particles considered in the cluster (depending on the cluster's radius).

It must be pointed out that:

• in order to allow a suitable overlap of the volumes associated to adjoining blobs a further "mollification" is performed as:

$$f_p\left(\underline{r}_p, t\right) = \sum_{\substack{q \in \text{Cluster} \\ \text{around } p}} f_q\left(\underline{r}_q, t\right) W\left(\frac{r_p - r_q}{\sigma}, h\right) \Delta Vol_q \tag{7}$$

where " $\sigma = h \, \gamma$ " is defined as the blob's radius and "γ " is an overlap parameter (usually of the order of 1 – 1.5).

• In order to be dimensionally consistent the kernel function W(r, h) must have the dimension of the inverse of a Volume (according to the dimension "d" of the space considered). Since it is usual to consider as kernels non-dimensional functions rapidly decaying, the "mollification" takes the form:

$$f_p\left(\underline{r_p},t\right) = \sum_{\substack{q\in \text{Cluster} \\ \text{around } p}} \frac{1}{\sigma^d} W\left(\frac{\underline{r_p} - \underline{r_q}}{\sigma},h\right)\left\{f_q\left(\underline{r_q},t\right)\Delta Vol_q\right\} = \sum_{\substack{q\in \text{Cluster} \\ \text{around } p}} \frac{1}{\sigma^d} W\left(\frac{\underline{r_p} - \underline{r_q}}{\sigma},h\right) F_q\left(\underline{r_q},t\right) \tag{8}$$

where $F_q(r_q,t) = f_q \Delta Vol_p$ is considered as the "field intensity" of the variable "f_q" over the finite volume ΔV_p.

Note that the blob representation is defined as a space average (i.e. it is an implicit sub-scale model), that is: the Lagrangian blob method for the Helmholtz formulation of the complete Navier-Stokes equations, under the limit of validity of the Boussinesq approximation, represents the space averaged equivalent equations:

$$\begin{cases} \dfrac{\partial \overline{\omega}}{\partial t} + \underline{\nabla}\bullet\left(\overline{\underline{V}\,\overline{\omega}}\right) = \underline{\nabla}\bullet\left(\overline{\omega\,\underline{V}}\right) + \nu\,\nabla^2\overline{\omega} - \beta\left[\underline{g}\bullet\underline{\nabla}\overline{T}\right] \\[2mm] \dfrac{\partial \overline{T}}{\partial t} + \underline{\nabla}\bullet\left(\overline{\underline{V}\,\overline{T}}\right) = \alpha\,\nabla^2\overline{T} \end{cases} \tag{9}$$

So that, if the "h" parameter and the time integration step are small enough with respect to turbulent scale, the Lagrangian Blob method does not need any turbulent stress term, and it can be considered to represent a DNS method.

The Blob particles representation is linked to important Lagrangian conservation properties. For a general advection problem Lu=g , the blob representation satisfy the "self-adjoint-ness" condition:

$$\int_0^T \left\langle Lu\left(\bullet,t\right),W\left(\bullet,t\right)\right\rangle dt = \left\langle u\left(\bullet,0\right),L*W\left(\bullet,0\right)\right\rangle + \int_0^T \left\langle g(t),L*W\left(\bullet,t\right)\right\rangle dt \tag{10}$$

where $\left\langle Lu,W\right\rangle$ is the (space averaged) blob representation of the operator Lu, with L* its adjoint.

Obviously the Lagrangian Blob formulation of a differential problem needs the representation of all the differential operators in the problems, namely divergence, gradient and Laplacian. This may be a problem that is solved in various manners, as it will be reported later.

In conclusion Blob methods represent exact weak solution for any admissible test function (local averaged equation), i.e blob particle method achieve some (implicit) sub-grid scale model. Blobs are "Dirac particles" that directly translate with flow and transport extensive properties. They move according to velocity field, and exchange momentum and energy with neighborhood blob particles according to diffusive processes.

Blob methods then differ from classical grid techniques since they do not involve projection of the equation in a finite dimensional space. They are suitable for any advection and diffusion problem such as: Navier Stokes equations ; Multiphase media; Gas dynamics; Vlasov-Maxwell equation.

3. THE MODEL PROBLEM

The objective is to consider an obstacle–free computational domain and in the frame work of a low order unsteady scheme we are aimed to use a vorticity– velocity– temperature particle blob approach with a Lagrangian formulation that is typically second order.

The present approach considers the analysis of the Helmholtz equations that are obtained as the curl of the momentum term of the Navier Stokes. In this formulation the pressure disappear and the velocity field is implicitly solenoidal. The problem of solenoidality that deals with the vorticity field does not arise in 2D problems that are the scope of this explorative analysis.

This approach avoid the imposition of the divergence free velocity field and of the asymptotic pressure closure condition needed for fixing mass flow in buoyant problems that is very difficult to impose and represent one of the weak points of many industrial codes.

The formulation, Golia *et al.* [22], [23], [24], [25], furthermore makes use of a splitting technique of Chorin [26], [27], that separates explicitly convective and diffuses steps, recasting the problem in a hyperbolic one for the particle and in parabolic problems for the diffusive and source phenomena that occur along each particle path.

We do not consider flows around solid body by Brinkman double penalization, Fig. (3). In this regard we shall consider 3 zones (fluid, Brinkman layer, solid body) that are defined by a time varying mask function. Then for the generic p-th particle located at r_p at time t, we must solve the following initial value problem (in case of constant diffusive properties):

Convective steps (trajectory of the p-th particle):

$$\frac{d\,r_p}{dt} = U(\underline{r}_p, t) \tag{11}$$

Diffusive Step (variations of the field properties of the p-th particle along its trajectory):

$$\frac{\partial \underline{\omega}_p}{\partial t} = \begin{cases} \dfrac{1}{Re}\nabla^2\underline{\omega}_p - \beta\left[\underline{g}\wedge\underline{\nabla}T_p\right] & \text{fluid} \\[3mm] \dfrac{1}{Re}\nabla^2\underline{\omega}_p - \left[\dfrac{\underline{\nabla}\wedge\left(\underline{V}_p - \underline{V}_{body}\right)}{\eta}\right] & \text{Brink. Layer} \end{cases} \tag{12}$$

$$\underline{\omega}_p = \underline{\nabla}\wedge\underline{V}_{body} \qquad \text{Solid Body}$$

$$\frac{\partial T_p}{\partial t} = \frac{1}{Re\,Pr}\nabla^2 T_p \tag{13}$$

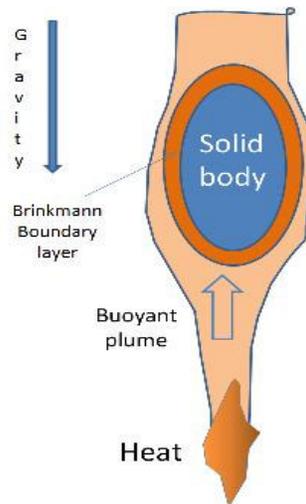**FIGURE 3:** The Brinkman Layer.

Note that:
- in the buoyant term 'β' denotes the thermal coefficient of expansion
- η <<1 is a penalization parameter
- we have not considered the dissipative term in the energy equation since it is very small and it is usually discarded in case of buoyancy.
- we do not consider buoyancy in the Brinkman Layer

- the error introduced, see [3], is expected to be of the order $\eta^{-1/4}$

The Blob representation of the eqns. (11), (12), (13) - detailed elsewhere [22], [23], [24], [25] - considers as field unknowns the p-th particle local "Vortex intensity" $\Gamma_p = \omega_p \, \Delta V_p$ (i.e. local velocity circulation) and the p-th particle local "Thermal intensity" $\Theta_p = T_p \, \Delta V_p$.

The velocity field needed in (12) is computed using an Helmholtz decomposition, i.e. as sum of a potential velocity field and a vortical velocity field computed with a generalized Biot-Savart law. This is a weak point of the Blob Vorticity Methods since the velocity of a given blob, induced by all the N vortex blobs present in the field, represents a classical N-Body problem that require $O(N^2)$ calculations. The deriving computational burden is obviously not acceptable in problems where N increase very rapidly in time. To overcome this fact, we use an "in house" Adaptive Fast Multipole Method (FMM) [28] that results to be a O(N) algorithm. The FMM was devised to be capable of self organize in order to reach optimal computation conditions while N is varying and able warrant an imposed error level on the calculation of the velocity field induced by the vortex blobs. The blob representations of the Laplacian diffusive terms and of the Gradient terms present in (12) and (13) are performed according to the Particle Strength Exchange (PSE) method proposed by Degond & Mas-Gallic [29]:

$$\nabla^2 \underline{\Gamma}_p = \sum_{q=1}^{N} \left(\underline{\Gamma}_p - \underline{\Gamma}_q \right) \mathrm{Lapl}\left(\underline{r}_p - \underline{r}_q, h \right) \tag{14}$$

$$\underline{\nabla} \, \Theta_p = \sum_{q=1}^{N} \left(\underline{r}_p - \underline{r}_q \right) \left(\Theta_p + \Theta_q \right) \mathrm{Grd}\left(\underline{r}_p - \underline{r}_q, h \right) \tag{15}$$

The kernels Lapl(•,•) and Grd(•,•) used here are the high order ones according to Eldredge et alia [30]. The PSE discretization of the differential operators are quite accurate for blobs regularly distributed in the field and away from boundaries/discontinuities. Usually, after a number of integration steps, to avoid inaccuracy due to the typical Lagrangian distortion of the particle field, a re-gridding process is performed to project the field on a regular mesh. New particles are allocated on the new mesh points and their values are interpolated from the cluster of the neighborhood ones, on the old disordered lattice, with a suitable kernel function:

$$F_p^{new} = \sum_{q=1}^{\substack{\text{Neighborhood} \\ \text{Cluster}}} F_q^{old} \, W_{regrid}\left[\left(\underline{r}_p^{new} - \underline{r}_q^{old} \right), h \right] \tag{16}$$

The kernel $W_{regrid}(\bullet,\bullet)$ used here is compact over r/h=2.5 and is third order accurate [31].

After the discretization of each term of (11), (12), (13), it results in an initial value problem (I.V.P.) problem that requires the time integration of N-equations. This is performed, in general with a 2nd order algorithm.

It must be said that for particles located within the Brinkman Layer the problem is extremely stiff. The time integration can produce oscillations or inaccuracies if the time step "dt" is not enough smaller then the value of the penalty parameter "η". This is a burden since in order to have sufficient accuracy the penalty parameter "η" must be at least of order of 10^{-4}.

In this work we use different strategies for the time integration for particles located in particular volume phase. For fluid particle we use a 3rd order Adams-Bashforth method that computes the "n+1 time step" values using the old values at the "n time step" and at the "n-1 time step".

For particles located in the Brinkman Layer we use a first order implicit method. Future work will test various methods to improve accuracy in order to use larger values of the ratios of time steps over penalty parameters.

4. THE VALIDATION

Validations are needed to ascertain the preliminary accuracy of the original Thermal-Blob method before the use of the Double Brinkman Penalization. In the following we shall present various tests that consecutively validate each particular segments of the method and of the relative code.

4.1 Inviscid elliptical vortex blob patches

Scope of this test is to validate the accuracy of the routines for computation of the velocity field induced by the vortex blobs and for the time integration of the blob's particle paths.

We do consider a classical inviscid Elliptical Vortex Blob Patches (Lambs' vortex) that [32] will rotate steadily and as rigid body, according to the value of the (constant) patch vorticity and to the values of the semi axes a, b, with the angular velocity:

$$\Omega_{rot} = \omega_{patch} \frac{ab}{(a+b)^2} \tag{17}$$

The time frames depicted in Fig. 4 show the almost perfect rigid body rotation (no regrinds), thus validating the computation of the induced velocity and of the time integration routines.
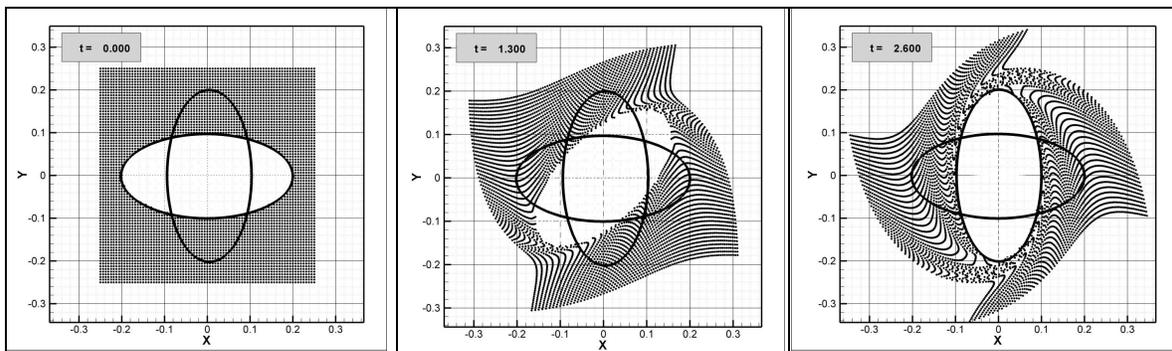

FIGURE 4: Inviscid Elliptic Patch at Various Times.
Nx=Ny=81; dx=dy=0.00625; dt=0.00125, $\omega_{patch} = \pi$, Vel. Acc.= 10^{-4}

4.2 Free buoyant thermal blob patches

Scope of this test is the validation of the routines that compute Laplacian (diffusion processes) and gradients, the correctness of the buoyant term and the effectiveness of the regrid processes.

We consider an initially Thermal Free Patch in a gravity field undergoing buoyant and diffusion processes. The patch rising will split in two anti rotating cores due to the vorticity created by the buoyancy term in eqn.(12). Geometrical symmetry is expected and global vorticity must be conserved to zero.

Fig. 5 represents the real computational field and the evolution of the blobs particles, colored according to their temperature. It is interesting to note as the Lagrangian formulation allows the code to be completely adaptively and grid free. The blobs naturally go where needed and the code automatically creates new particles according to the vorticity generations.

The run corresponds to the following parameters: Gr=0.65107; Pr=0.748; h=0.01; dt=0.0125. At t=0, the initial, number of blobs is 2511 and grows spontaneously up to 8919 , at t=5.
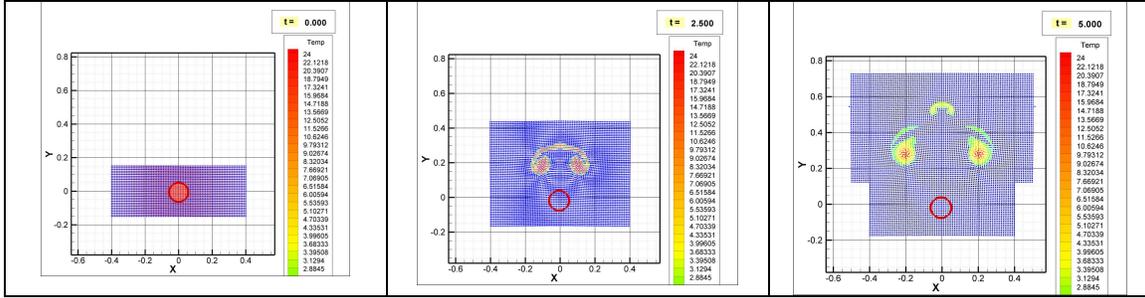
FIGURE 5: Free Plume Convection of Thermal Free Patch Initially Circular.

4.3 Free buoyant thermal-vortex blob patches

Scope of this test is the validation of the coupling of the thermal buoyant process with the dynamics effects of a non zero initial vortex patch. Here we consider an initially circular Free Patch of blob particles, having both Thermal and Vortex content, placed in a gravity field, that undergoes buoyant and diffusion processes.

Due to the vorticity created by the buoyancy term in (12), the rising patch will split in two anti rotating cores. But in this case geometrical symmetry is not maintained due to the fact that global vorticity must be conserved to the non zero initial value.
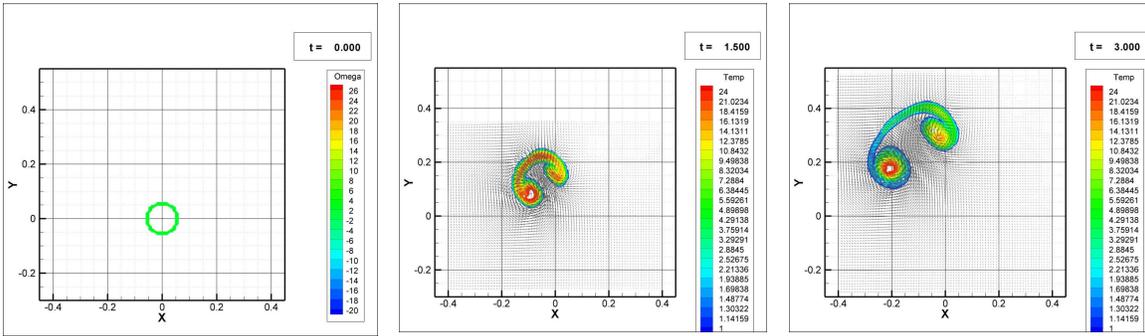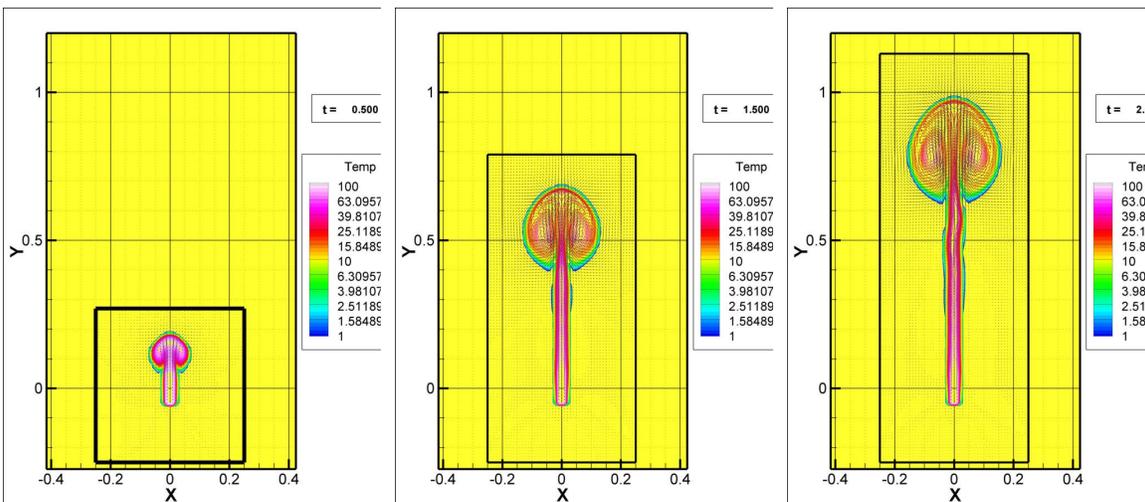

FIGURE 6 : Free Plume Convection of Vortex-Thermal Free Patch Initially Circular.
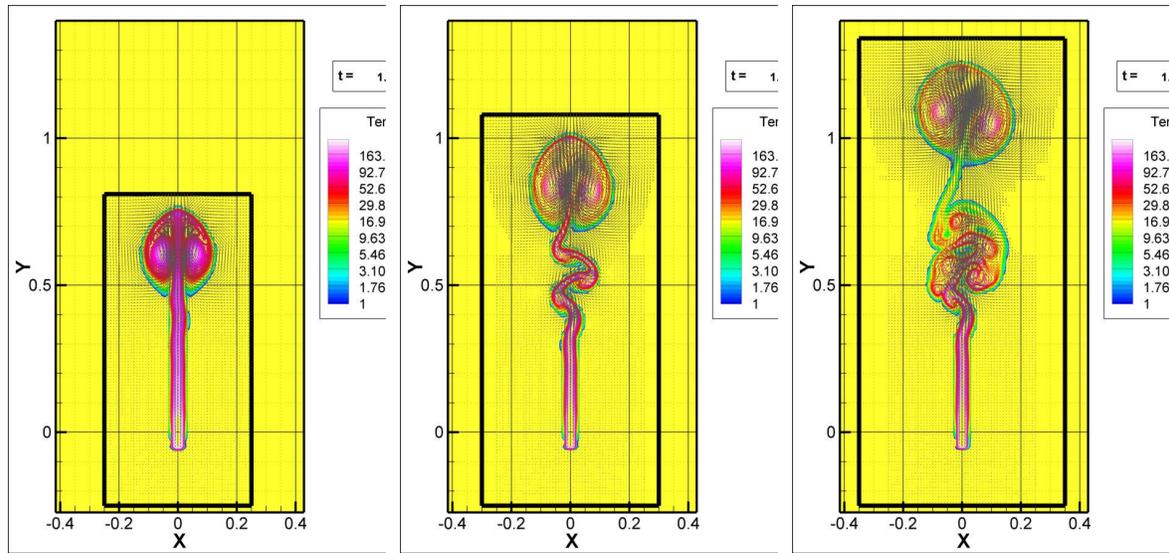
FIGURE 7: Free Buoyant Rising Plume from a Thermal Blob Patch at Different Grashof Numbers.

As shown in [33], integral balance laws impose according to the sign of the initial vorticity, that one of the two cores rotates more (i.e. will have larger vorticity) than the other.

Fig. 6 reports the evolution of the iso-vorticity contours. The computational parameters are: Pr=0.748; h=0.1, dt = 0.00125, V_{ref}=0.125 10-3; L_{ref} = 0.1188, Rey=1, regrid process is performed after 10 time steps.

It is interesting to note that, as predicted, the initial circular blob path rising is split in two counter-rotating ones, the left one having a large vorticity content of the other, then having higher angular rotation velocity. Since, the global of vorticity content must be constant as the initial one, the rotation difference of the two cores will vary in time. Therefore the iso-vorticity contours undergo higher relative changes with respect to the thermal ones.

4.4 Fixed buoyant thermal blob patch

Scope of the test is to verify the free buoyant rising plume that stems from a Thermal Blob Patch which is maintained fixed and at a constant temperature. The same Patch will be hereafter considered to generate a thermal plume that shall interact with solid bodies (still or moving).

Fig.7 reports the thermal contours of the rising plume at various times, for different Grashof numbers: the top row refers to Gr=0.25 10^8, the bottom row refers to Gr=0.625 10^8. Temperature level refers to temperature rise with respect to the asymptotic one. Both case are relative to an elliptical fixed thermal patch having a=0.0198 and b=3a. The run corresponds to the following parameters: Gr = 0.65 10^7 ; $Rey_{\Omega D}$ = 0.243 10^4 ; Pr = 0.748 ; h = 0.01 ; Δt = 0.0125. The initial (at t=0) number of blobs is 41311, and grows spontaneously up to 67220, at t=3,.

Even if the value of the reference blob distance "h" is not small enough to reach turbulent scale, it is evident, from the figure, the capacity of the method and of the code to catch the instability of the rising plume that is verified at a certain rising distance from the standing thermal patch.

5. PLUME-BODY INTERACTIONS

In this paragraph we consider the interaction of the thermal buoyant plume (described in the previous paragraph) with a body using the Brinkman Layer as per Fig. 3. In the following, we shall present the interactions of the rising plume with a still blunt body, with a still slender body and with a slender moving body (linear motion and rotation).

5.1 Plume interaction with a still blunt body

The data of the thermal plume are as the ones of the top row of Fig. 7.

The body is an ellipse b/a=3 staying still above a thermal patch that originate a buoyant plume. The body is at zero temperature rise and the Brinkman layer has a thickness twice the blob nominal distance "h".

Fig. 8 shows three frames of a video clip referring to the touching of the plume mushroom with the Brinkman Layer of the body and two other significant later states, namely the breaking of the mushroom and the lateral migration of the two counter-rotating vortices generated by the interaction.
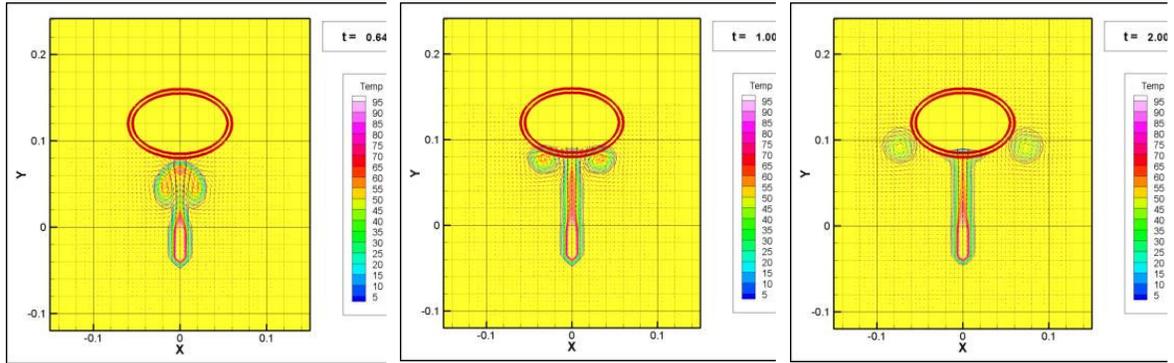

FIGURE 8: Plume Interaction with a Still Blunt Body Simulated with Brinkman Layer.

5.2 Plume interaction with a still slender body

The same comments as the ones of paragraph 5.1 apply, but here the elliptical body, as depicted in Fig. 9, is rotated by 90° normally to the gravity vector to represent a slender obstacle for the rising plume. In this case a quite lighter interaction is expected and the migration of the two counter-rotating vortices generated by the interaction is quite faster. The technique used to highlight the temperature contours emphasizes the automatic dynamic development of the computation field that, at the initial stages, does not contain the body.
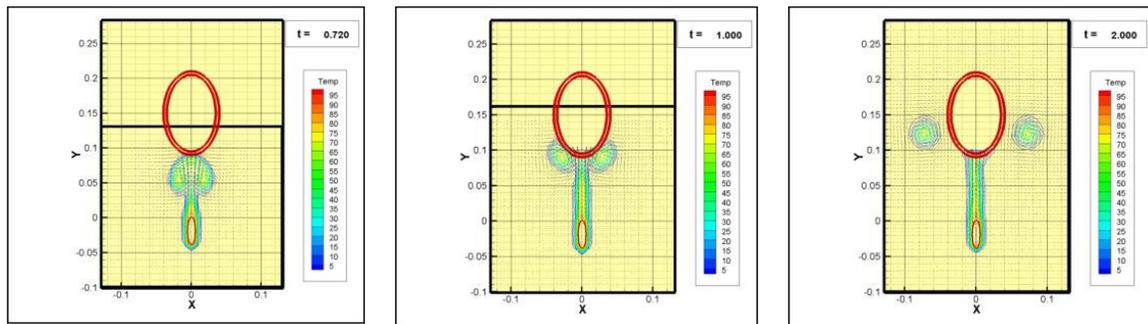
FIGURE 9: Plume Interaction with a Still Slender Body Simulated with Brinkman Layer.

5.3 Plume interaction with a transverse moving body

The data of the thermal plume are as the ones of the top row of Fig. 7. The body is an ellipse with b/a=3, oscillating horizontally (normally to the gravity vector) with an amplitude of 0.2 and a period of 0.82. The body is at zero temperature rise and the Brinkman layer has a thickness twice the blob nominal distance "h".

The images in Fig. 10 report the blobs position and the velocity vector. In the top row the blobs are colored according to the local vorticity, in the bottom row the blobs are colored according the local temperature rise.

As a check, the invariance of the zero total vorticity over a full period of oscillation is verified. The oscillatory motion of the solid body is synchronized such that at time t=1 the plume top mushroom hits the bottom of the body.
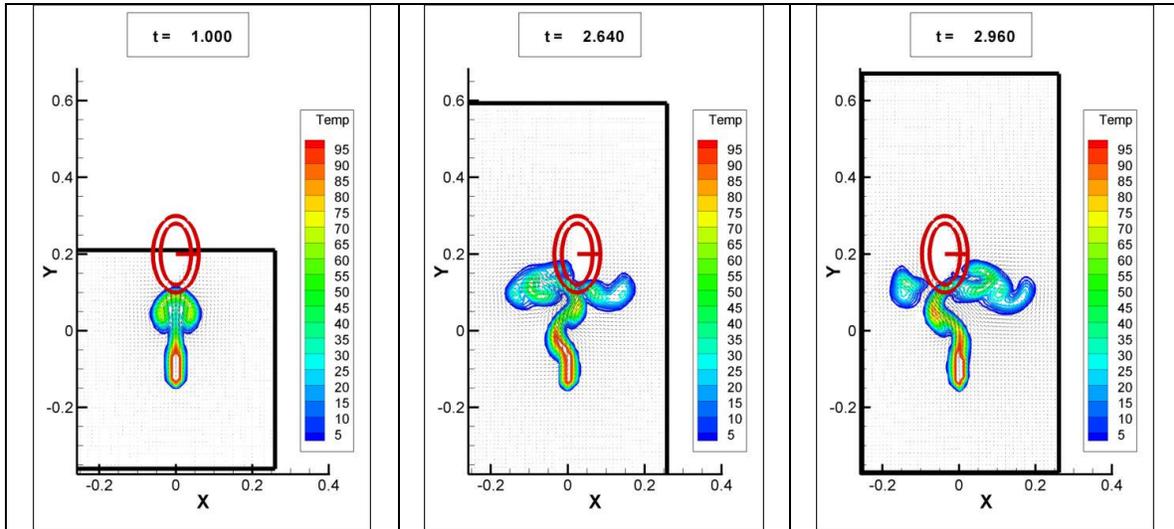

FIGURE 10: Thermal Plume Interacting with a Slender Body Oscillating Laterally (contours).
Figure 10 reports some shots that reveal the strike and the relative hysteresis phenomena of the flow pattern. The t = 2.64 shot refers to the ellipse moving west ward, the t = 2.96 shot refers to the ellipse moving eastwards. Obviously the running of the video clip can better reveal the dynamics of the interaction and the realistic behaviors of such a complicate unsteady interaction.

5.4 Plume interaction with a rotating body

In this case we consider a thermal plume generated by a thermal fixed particle cluster posed below a circular body that rotates with a given angular speed Ω_b anti-clockwise.

Fig.11 refers to the case Ω_b= 10 [rad/s] with hot blob cluster aligned below the rotating body.

Fig.12 refers to the case Ω_b= 20 [rad/s] with hot blob cluster aligned below the rotating body.

Fig.13 refers to the case Ω_b= 30 [rad/s] with hot blob cluster misaligned below the rotating body.

The six shots of each Figure refer to different times and display blobs colored according the specific temperature value. We can notice:

FIGURE 11: Thermal Plume Interacting with a Slender Body Rotating Anti-Clockwise at 10 rad/s. Plume Generated by a Thermal Fixed Cluster Aligned with the Rotating Body. Blobs Colored with the Specific Thermal Level.
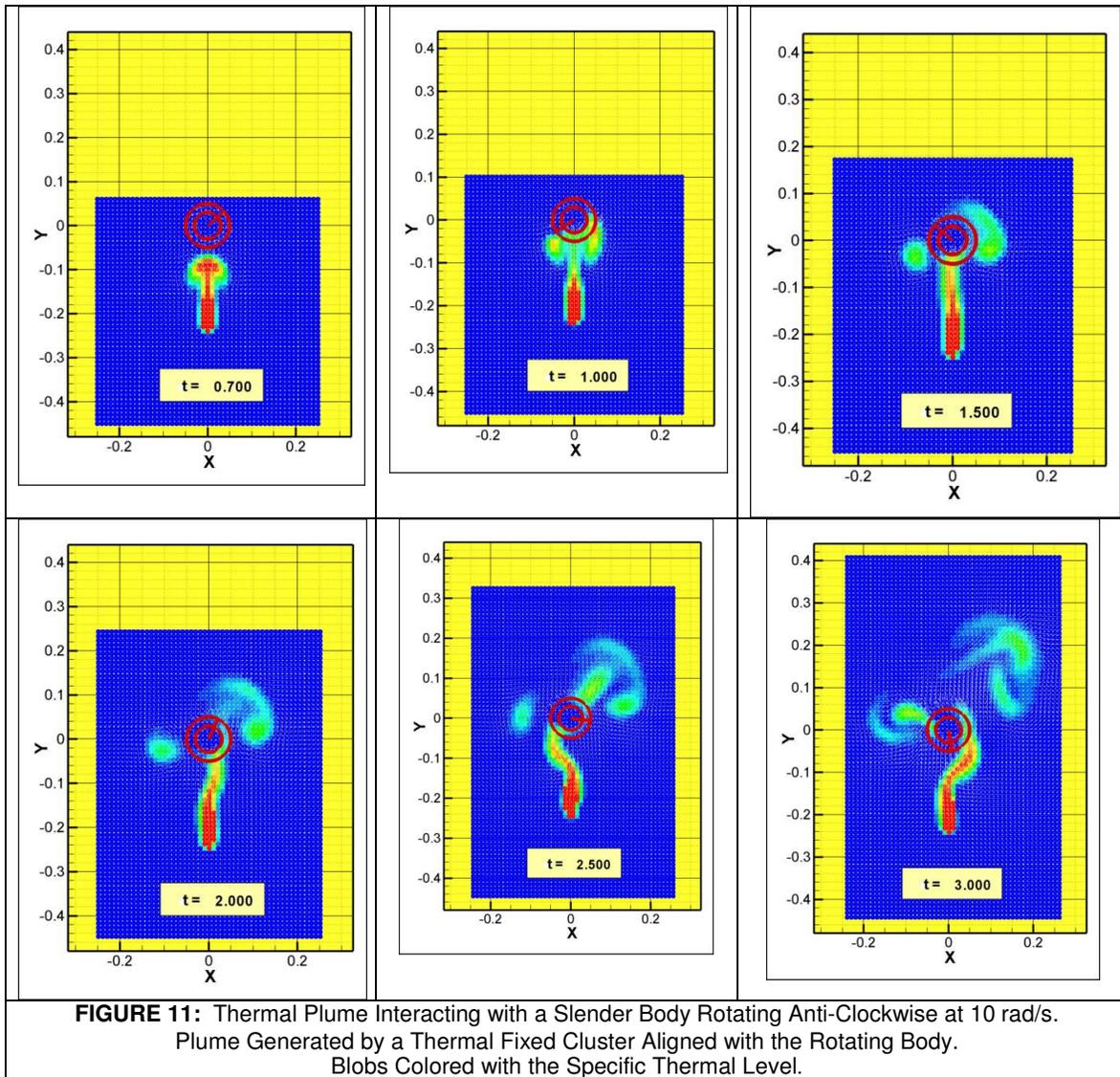
- the capacity of the approach to generate automatically blobs, as needed, that moves in the region where thermal and vorticity effects are relevant.
- the different dynamic behavior of the plume:
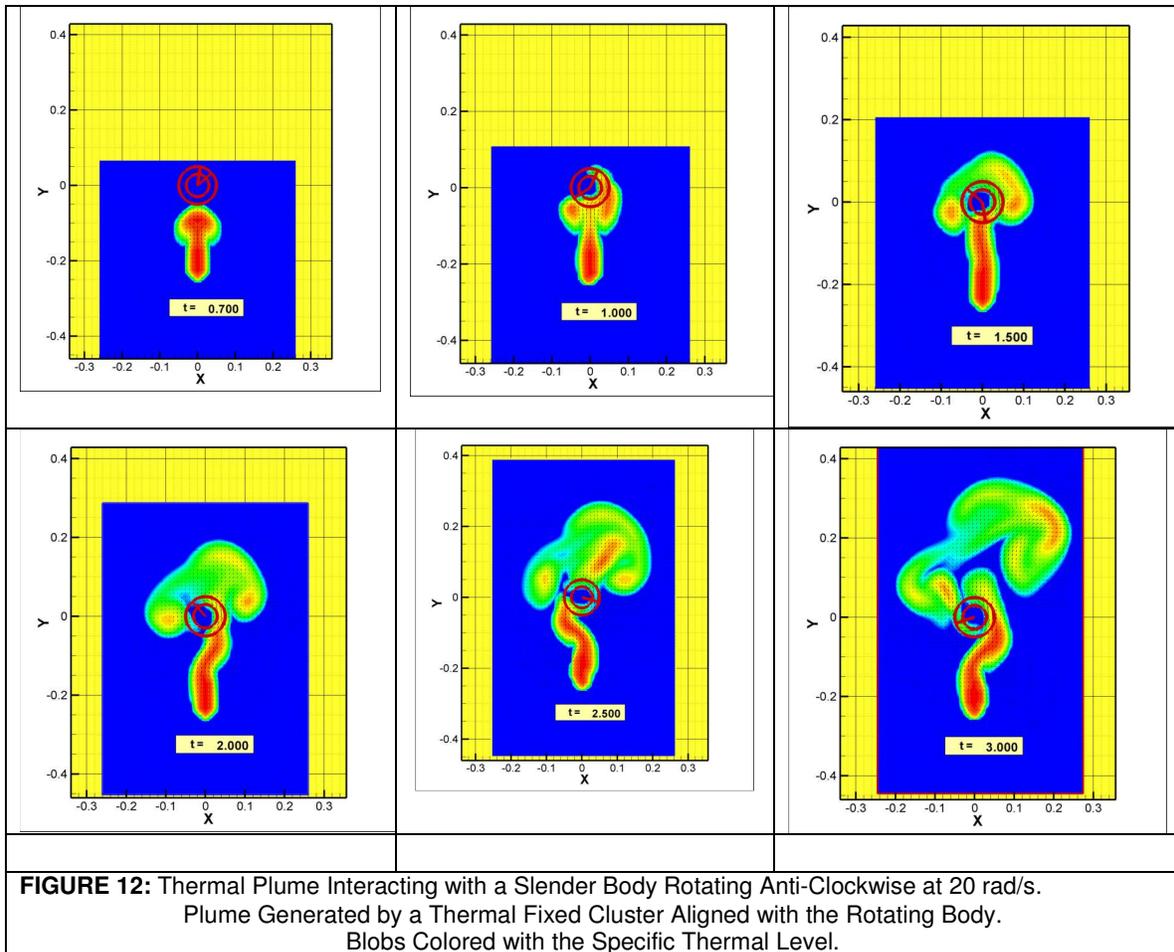 - In the initial stage the steam of the plume is linked with the mushroom cap that contains anti-rotating vortical structures: the sting furnish energy to the cup enough to overcame the loss of thermal and kinetic energy due to heat conduction and to viscous effects: it derives a certain stability for the sting;
  - In the latter stage a pinch-off phenomena is verified: the sting separates for the cup that is no longer feden;
  - It derive that the cup tends to decrease the rotation speed and to fragmentize whereas the sting loses it stability and oscillates laterally (as a candle) due, from one side, to the rotation of the top body and from the other side, to the need to obey to conservation of its global vorticity;
  - This is evident in the last shots of all figures, but in particular in Fig. 13 where the sting is first moved away by the anti-clockwise rotation of the body and after is attracted.

FIGURE 12: Thermal Plume Interacting with a Slender Body Rotating Anti-Clockwise at 20 rad/s.
Plume Generated by a Thermal Fixed Cluster Aligned with the Rotating Body.
Blobs Colored with the Specific Thermal Level.

6. CONCLUSIONS & FUTURE WORK

The use of the Brinkman penalization method to simulate the layer over a body considered as a porous media appears compatible with the grid free Lagrangian Vortex and Thermal Method developed by the authors. This method seems quite sound for simulating the viscous interaction of plumes and bodies since is based on fluid dynamics considerations that confine diffusive phenomena in segregated regions near the body.
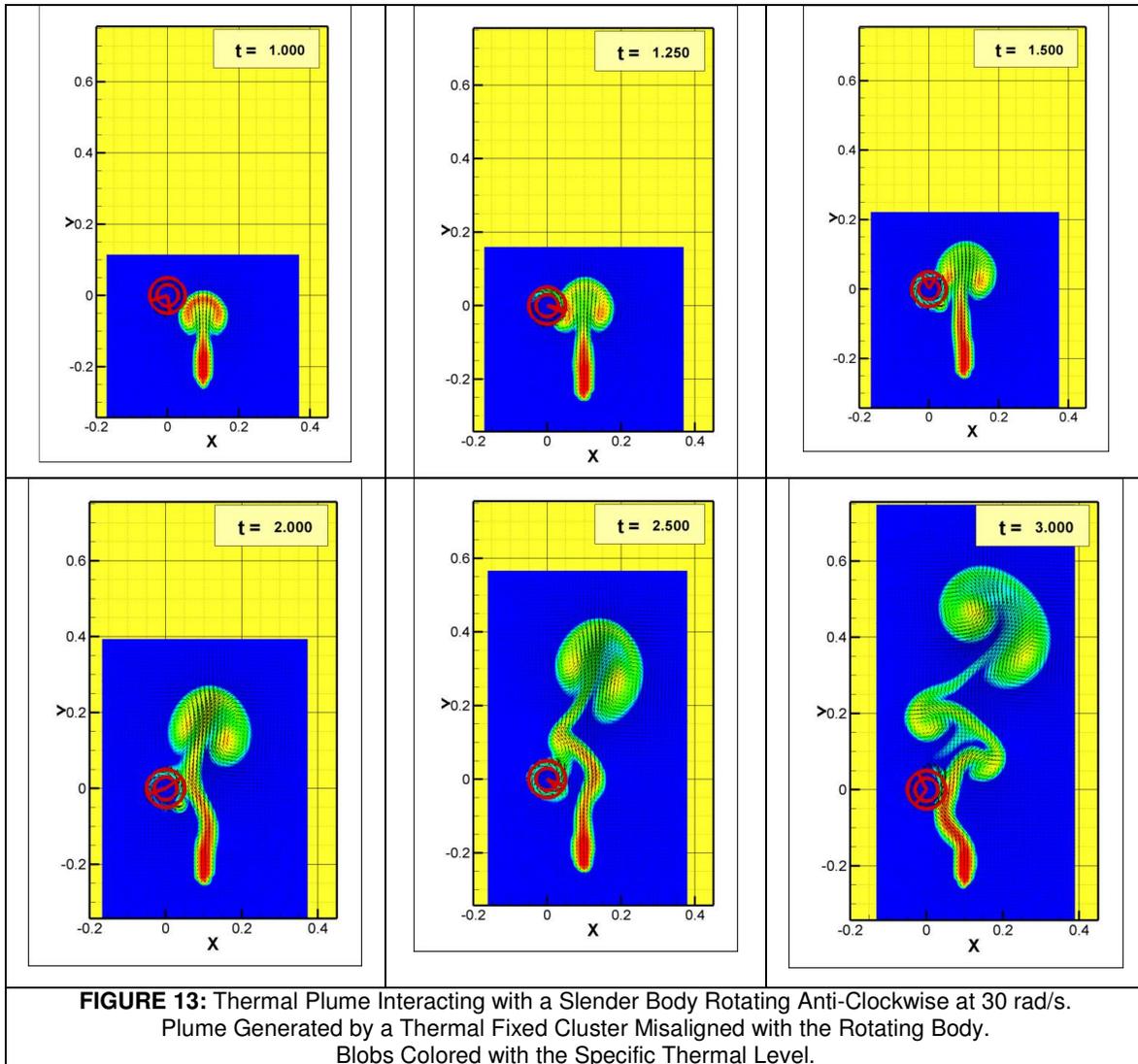
The first experience is positive but further works are required mainly in relation to the time integration of the diffusion processes for particles within the Brinkman Layer. The use of implicit scheme was experienced to be necessary, but further analysis is recommended in order to optimize peculiar schemes to improve the ratio of the time step interval over the penalization parameter. It must be underlined that dealing with unsteady motion of the body was quite simple and the relative effort limited mainly to the development of an efficient masking procedure, and to the of the time integration step that must be small enough with respect to the oscillation period.

In this preliminary analysis we did not consider the thermal interaction between the plume and the thermal characteristics of the body (considered at zero temperature rise). This is a quite large field of analysis that comprise, among other, all the various thermal conditions on the body, and the insertion of heat flux within the body itself and others.

It is obvious that further work should be reserved also to forced and mixed thermal convection and to real porous bodies. The working agenda is certainly quite full.

We do want to conclude this paper by stating that the first experience is definitely positive since the mix of the adaptive Lagrangian Blobs method with Brinkman penalization generates a

computational environment quite handy and efficient for the analysis of quite complicated problems of real interest in engineering science.

FIGURE 13: Thermal Plume Interacting with a Slender Body Rotating Anti-Clockwise at 30 rad/s. Plume Generated by a Thermal Fixed Cluster Misaligned with the Rotating Body. Blobs Colored with the Specific Thermal Level.

7. REFERENCES

1. Arquis,E. , Caltagirone,J-P., Vincent,S. "*Sur le conditions hydrodynamique au voisinage d'une interface millieux fluid-millieux poreux: application à la convection naturelle*". C.R. Acad. Sci. Paris, Series IIb, 299: 1-4, 1984

2. Angot, Ph.. "*A Mathematical and numerical modelling for a fictious domain method with penalized immersed boundary conditions*". HDR thesis, Universitè Mediterranée, Aix-Marseille,1998

3. Angot, Ph., Bruneau C.H., Fabrie, P.. "*A penalization method to take into account obstacles in incompressible viscous flows*". Numerische Mathematik, 81(4): 497-520, 1999

4. Caltagirone,J-P., Vincent,S : Sur une methode de penalisation tensorielle pour la resolution des equations de Navier-Stokes, C.R. Acad. Sci. Paris, Series IIb, 329: 607-613, 2001

5. Khadra, K., Angot,Ph., Parneix, S., Caltagirone,J-P.. "*Fictitious domain approach for the numerical modeling of Navier-Stokes equations*". Int. J. Numer. Methods in Fluids, 34 (8): 651-684, 2000

6. Schneider,K. ,Farge, M. *."Coherent Vortex Simulation (CVS) of a 2D bluff body flow using an adaptive wavelet method with penalization".* Advances in Turbulence, 9: 471-474, 2002
7. Schneider, K. ,Farge, M. *"Adaptive Wavelet Simulation of a flow around an impulsively started cylinder using penalization".* Applied and Computational Harmonic Analysis, 12: 374-380, 2002
8. Schneider,K. ,Farge, M. *"Numerical simulation of dipole-wall interaction using an adaptive wavelet discretizatuion with volume penalization".* ENUMATH (eds. Bermudiz et alia), Springer, 822-830 (2005)
9. Schneider,K. ,Farge, M. *"Numerical simulation of transient flow behaviour in tube bundles using volume penalisation method".* Journal of Fluids and Structures, 20 (4): 555-566, 2005
10. Schneider,K. *"Numerical simulation of the transient flow behaviour in chemical reactors using penalization method".* Computers and Fluids, 34(10): 1223-1238, 2005
11. Vasilyev,O.V , Kevlahan, N.K.R. *"Hybrid wavelet collocation-Brinkman penalization method for complex geometry flow".* Int. J. Numer. Methods in Fluids, 40(3-4): 531-538, 2002
12. Kevlahan, N.K.R., Vasilyev,O.V., Goldstein, D., Jay, A. *"A adaptive wavelet collocation method for fluid-structure interaction at high Reynolds number".* SIAM J. Sci. Comp., 26 (6): 1894-1915, 2005
13. Kevlahan, N.K.R., Ghidaglia JM. *"Computation of turbulent flow past array of cylinders using spectral method with Brinkman penalization".* Eur. J. Mech. B - Fluids, 20:333-350, 2001
14. Bruneau,C.H. , Mortazavi I. *"Control of Vortex Shedding Around a Pipe section Using a Porous Shealth".* International Journal of Offshore and Polar Engineering, 16: 90-96, 2006
15. Chiavassa G. , Donat R. *"Point Value Multiscale Algorithms for 2D Compressible Flow".* SIAM J. Sci. Comp., 23 (3) : 805-823, 2001
16. Bruneau,C.H. , Mortazavi I. *"Controle passif d'ecoulement autour d'obstacles par les milieux poreaux".* C.R. Acad. Sci. Paris, Series IIb, 329 :517-521, 2001
17. Bruneau,C.H. , Mortazavi I. *"Passive control of the flow around a square cylinder using porous media".* Int. J. Numer. Methods in Fluids, 45(4): 415-433, 2004
18. Carbou,G. , Fabrie P *"Boundary Layer for a penalization method for viscous compressible flow".* Adv. Differential Equations, 8 :1453-1461, 2003
19. Carbou,G. *"Penalization method for viscous incompressible flow around a porous thin layer".* Nonlinear Anal. Appl., 5: 815-855, 2004
20. Carbou,G. "Brinkman Model and Double Penalization Method for the Flow around a Porous Thin Layer". J. Math. Fluid. Mech., 10: 126-158, 2008
21. Cottet,G.H. and Koumoutsakos,P.D. *"Vortex Methods: Theory and Practice".* Cambridge University Press, 2000
22. Golia,C., Buonomo,B., Manca,O., Viviani, A. *"A Vortex-Thermal Blobs Method For Unsteady Buoyancy Driven Flows".* ASME-IMECE, Anaheim (CA), 2004, Paper IMECE-2004-59180
23. Golia,C., and Buonomo,B. *"An Effective Blob Approach to Unsteady Thermal Buoyant Flow".* CMEM, Malta, 2005
24. Golia,C., and Buonomo,B. *"Numerical Simulation of Unsteady Natural Convection by Blobs Methods".* 60th ATI Congress, Roma, 2005
25. Golia,C. and Buonuomo,B. *"On the accuracy of integral representation of differential operators in meshless Lagrangian blob method".* CMEM, Praga, 2007.
26. Chorin, A.J. *"Numerical Study of Slightly Viscous Flow".* Journal of Fluid Mechanics 57 (4) : 785-796, 1973
27. Chorin, A.J. *"Vortex sheet approximation".* J. Comp. Phys., 27 (3) : 428-442, 1978
28. Greengard L., Rokhlin, V. *"A fast algorithm for particle simulation".* J. Comp. Phys., 73 (2): 325-348, 1987
29. Degond,P. and Mas-Gallic,S. *"The weighted Particle method for Convection-Diffusion Equations, part.1: the case of an isotropic viscosity".* Math.s of Computation, 58 (188) : 485-507, 1989
30. Eldredge,J.D., Leonard,A. and Colonius,T. *"A General Deterministic Treatment of Derivatives in Particle Methods".* J. Comp. Phys., 180 (2): 686-709, 2002
31. Meijering, E., Unser, M. *"A Note on Cubic Convolution Interpolation".* IEEE Trans. Image Process, 12 (4): 477-479, 2003
32. Batchelor G.K. *"An Introduction to Fluid Dynamics".* Cambridge Univ. Press, 1967

33. Riccardi G., Golia C. *"Integral Balance Laws for 2D buoyancy driven flows".* VII SIMAI Congress, Venice, 2004.

Heat Transfer Analysis for a Winged Reentry
Flight Test Bed

Antonio Viviani   antonio.viviani@unina2.it
Dipartimento di Ingegneria Aerospaziale e Meccanica (DIAM)
Seconda Università di Napoli (SUN)
via Roma 29, I-81031 Aversa, Italy

Giuseppe Pezzella   g.pezzella@cira.it
Centro Italiano Ricerche Aerospaziali (CIRA)
via Maiorise, I-81043 Capua, Italy

Abstract

In this paper we deal with the aero-heating analysis of a reentry flight demonstrator helpful to the research activities for the design and development of a possible winged Reusable Launch Vehicle. In fact, to reduce risks in the development of next generation reusable launch vehicles, as first step it is suitable to gain deep design knowledge by means of extensive numerical computations, in particular for the aero-thermal environment the vehicle has to withstand during reentry. The demonstrator under study is a reentry space glider, to be used both as Crew Rescue Vehicle and Crew Transfer Vehicle for the International Space Station. It is designed to have large atmospheric manoeuvring capability, to test the whole path from the orbit down to subsonic speeds and then to the landing on a conventional runway. Several analysis tools are integrated in the framework of the vehicle aerothermal design. Between the others, we used computational analyses to simulate aerothermodynamic flowfield around the spacecraft and heat flux distributions over the vehicle surfaces for the assessment of the vehicle Thermal Protection System design. Heat flux distributions, provided for equilibrium conditions of radiation at wall and thermal shield emissivity equal to 0.85, highlight that the vehicle thermal shield has to withstand with about 1500 [kW/m^2] and 400 [kW/m^2] at nose and wing leading edge, respectively. Therefore, the fast developing new generation of thermal protection materials, such as Ultra High Temperature Ceramics, are available candidate to built the thermal shield in the most solicited vehicle parts. On the other hand, away from spacecraft leading edges, due to the low angle of attack profile followed by the vehicle during descent, the heat flux is close to values attainable with conventional heat shield. Also, the paper shows that the flying test bed is able to validate hypersonic aerothermodynamic design database and passenger experiments, including thermal shield and hot structures, giving confidence that a full-scale development can successfully proceed.

Keywords: Atmospheric Reentry, Nonequilibirum Hypersonic Flow, Aerothermochemistry, Aeroheating, Thermal Protection System, Ultra High Temperature Ceramics.

Nomenclature

D	Aerodynamic Drag, N
h	Specific enthalpy, J/kg
$\vec{j}$	Mass diffusion flux vector, kg/m^{2}s
k	Gas thermal conductivity, W/mK
Kn	Knudsen number
L	Aerodynamic Lift, N
M	Molar mass, kg/kgmole

$\hat{n}$	Unit normal
$\dot{q}$	Heat flux, W/m^2
Q	Integrated heat load, J/m^2
$\vec{r}$	Position vector, m
R	Radius, m
$\Re$	Universal gas constant, J/kgmole K
Re	Reynolds number
S	Surface, m^2
t	Time, s
T	Temperature, K
v	Flight-path speed, m/s

Greek Symbols

α	Angle of attack, deg
γ	Recombination coefficient
ρ	Density, kg/m^3
σ	Stefan-Boltzmann constant, W/m^2K

Subscripts

fm	Free molecular
FP	Flat Plate
LE	Leading edge
N	Nose
r	Reference value
s	Species
w	Wall
∞	Freestream conditions
0	Total (reservoir) conditions

1. INTRODUCTION

The need for a safer access to space dictates the review of operational capabilities and hence of design approach for manned reentry vehicles of next generation [1–3]. Up to now several hypersonic shapes have been investigated for use in recoverable space systems. Initial efforts focused on low Lift-to-Drag ratio (L/D), as Apollo spacecraft. Current systems such as the Space Shuttle, still fly at relatively low L/D (<1.5) but are reusable. Research has shown that reentry vehicle designs with high L/D could be designed to take advantage of aerodynamic lift during reentry. Higher L/D is desirable because it increases the area from which a re-entering vehicle can be recovered (e.g. reentry window) [4]. In fact, the times and locations from which the Space Shuttle can return are limited by its down-range and cross-range capability after reentry. A vehicle with high aerodynamic efficiency would have enhanced down-range and cross-range capability (more than twice that of the Space Shuttle). This leads to increase significantly the operational flexibility of reentry space launcher, for which the ability to reach demanding orbits and to return quickly for reuse will be critical to their operations and economics.

This paper presents results of the thermal protection analysis and design of a vehicle designed to serve as flying test bed in the framework of a reusable Two-Stage-To-Orbit (TSTO) space launch system development.

Considerable technological progress and strong multidisciplinary design optimization (MDO) approach, validated by in flight operations, are needed. Up to now, however, no practical experience exists in development and operation of a Reusable Launch Vehicle (RLV) in Europe, even if many candidate concepts for a future RLV have been proposed.

Although considerable progress has been achieved in hypersonics flow computations, and large wind tunnels exist, e.g. the "Scirocco" PWT (Plasma Wind Tunnel) at CIRA, this is by far not sufficient for the design of an operational space vehicle. For example, none of the upper atmosphere and hypersonic regime conditions have been explored in Europe for a winged vehicle. Further, space launch systems are still regarded as both costly and unsafe. To avoid substantial risks in developing a next generation RLV, it seems advisable to gain first a practical RLV design knowledge by in-flight demonstrations of scaled low cost prototype vehicle. To this end, we have integrated several analysis tools in the conceptual design process of a sharp edge Reentry Test Bed (RTB) able to perform a return to Earth.

RTB is a reentry space glider, with large atmospheric manoeuvring capability, expected to test the complete path from the orbit down to subsonic speeds and then landing on a conventional runway. Thus, it

will be able to validate hypersonic aerothermodynamic design database and passenger experiments, including Thermal Protection System (TPS) and hot structures.

The flight test bed, named VIPER-G1, is a prototype winged vehicle, embodying the critical technologies and the features of an operational system, giving confidence that a full-scale development can successfully proceed, with particular attention both on Crew Rescue Vehicle (CRV) and Crew Transfer Vehicle (CTV) development for the International Space Station (ISS).

The vehicle detailed design, however, is beyond the scope of this work and the mission and system requirements will be defined only at the level able to prove the concept feasibility. The overall goal is to perform a return back to Earth in a gentle winged descent (lifting return), involving both angle of attack (AoA) and bank angle modulation, keeping constant the temperature of the nose stagnation in the radiation equilibrium conditions (i.e. radiative cooling), restricting the g peak experienced to less than a tenth above normal ground-level values and, finally, with a wider than usual "reentry window" that would permit landing at any one of the many choices of airfields [5].

With a high cross range the vehicle can more easily meet the requirements for the return of an injured or sick crew member to an approved trauma center within few hours. Increasing a hypersonic vehicle L/D ratio, effectively increases the footprint in which the vehicle can safely maneuver and land. With an increased footprint, the vehicle can account for unforeseen reentry anomalies or for weather constraints at the prime landing sites and change course to an available site.

The key technologies, we have selected, include:
- low wing loading;
- high lift;
- sharp leading edges;
- advanced thermal protection system;
The assessment of the vehicle aeroheating environment is reported in the next paragraphs.

2. VEHICLE CONCEPT OVERVIEW

The vehicle concept consists of the following main elements: main body with a sharp nose, low aspect ratio wing, vertical stabilizer, wing and body flaps.

Several subsystems are necessary for the RTB flight: the Reaction Control System (RCS) to control the attitude of the demonstrator when the aerodynamic control surfaces are ineffective; the landing parachute, the avionics, electrics, and hydraulics systems; an advanced TPS covering, in particular, the nose and the leading edges of wing and tail.

In the present study, particular attention is focused on the vehicle TPS concept. From the Space Shuttle experience, the TPS has been identified as one of the area where large improvements can be made regarding operations and maintenance costs. Its mass fraction, equal to about 15% of the vehicle empty weight, is large enough to be targeted for potential mass saving. This mass saving, together with operational reentry trajectories, characterized by low inertial loads (i.e. vehicle with lightweight primary structures), lead to vehicles with low wing loading. Finally, with low wing loading and advanced aerodynamic configuration (e.g. high L/D) high performance vehicles can be attained.

The sharp leading edges can stand at hypersonic conditions if bulk self standing shapes Ultra High Temperature Ceramics (UHTC) are provided [6-7]. UHTC shapes spread out the region of high surface temperature over which radiative cooling is effective [8-10].

For the remaining vehicle parts, instead, Metal Matrix Composite (MMC), such as for example PM1000, are expected to be used. Practically, the TPS becomes very integrated with the vehicle "cold" structure.

This compares with, and differs from, the usual method of thermal protection that makes use of a leading edge, having significant radius of curvature (i.e. blunt bodies), made of refractory material such as, e.g., Reinforced Carbon-Carbon (RCC) as in the Space Shuttle case.
However, it is known that conduction-assisted radiative-cooling is an adequate thermal protection in reentry only if the heating rates are relatively low. This implies that the vehicle wing loading (W/S_r) must be kept small.

 The above considerations, however, are by far for an exhaustive design approach for reentry vehicles.

The adopted concept is a conventional wing-body configuration with a sharp-leading-edges double-delta planform as basic shape. The configuration has a distinct wing body design with a blended wing body interface and a flat bottom surface to increase the overall hypersonic performance of the RTB. The three-view drawing of the vehicle are presented in Fig. 1

The forebody, instead, is characterized by a simple cone-sphere geometry with a sharp nose; while the overall vehicle length and the position of the wing are dictated by center of gravity (c.g.) position with respect to the aerodynamic center of pressure. In order to obtain satisfactory vehicle stability and controllability for major parts of the flight envelope the c.g. has to be close to vehicle centre of lift. So, the wing is placed near

the rear of the body and is also swept to assure best performance with respect to the supersonic drag and aerodynamic heating.
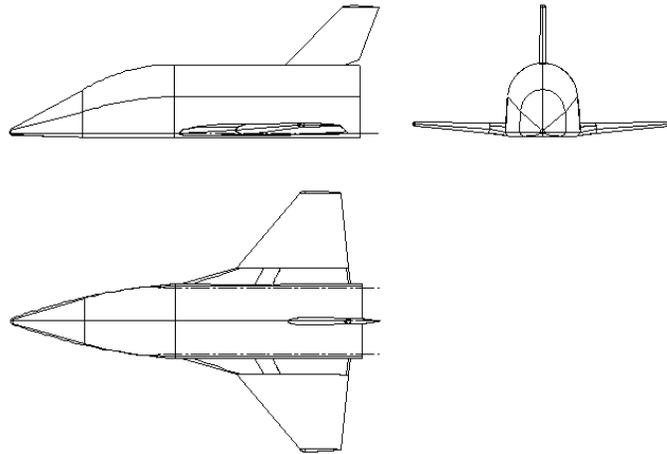

FIGURE 1 : Three View Drawing.

Furthermore, a dihedral angle is provided in order to enhance flight stability. The airfoil shape is maintained from root to wing tip and provides leading edge with small radius of curvature to reduce wave drag and to take advantage of the boundary layer thickening.

The control surface on the wings are elevons which must serve as ailerons, elevators, and flaps (due to heat buildup at the joints, split surfaces must be avoided). Therefore, the control surfaces deflect differentially and must be large enough to provide stability at landing speed, without sacrificing too much lift **Error! Reference source not found.**.

The vertical tail provides the vehicle sideslip stability. Note that along the low risk reentry the tail is effective unlike the classical reentry where, at high AoA, it is shielded from the flow providing no control [2].

In order to improve later on the vehicle aerodynamic and stability along the atmospheric descent, a body flap, extending from the base of the vehicle, is provided (at hypersonic speeds a surface behind the vehicle c.g. balances the nose up pitching moment provided by the vehicle forebody).

This additional control surface is also able to protect the engine nozzle, against aerodynamic heating when the spacecraft is flying at an AoA.

One of the most important inputs in the vehicle sizing process is the size requirement of the payload accommodation bay of the Expendable Launch Vehicle (ELV). In fact this requirement determines the overall dimensions of the demonstrator.
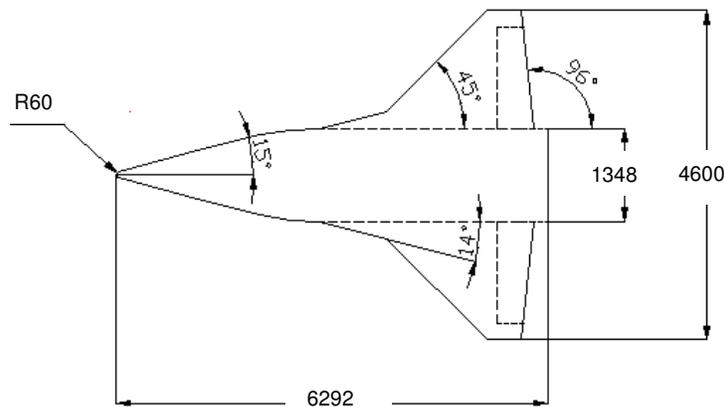

FIGURE 2 : RTB Dimensions, in [mm].

To take advantage of small low cost ELV, the following RTB dimensions (see Fig. 2) are used:

- total length (tail included) [m]:   6.400;
- fuselage length (L) [m]:   6.292;
- maximum fuselage width [m]:   1.348;
- wingspan [m]:   4.600;
- planform area [m^2]:   11.840;
- wing area (S_r) [m^2]:   9.344;
- nose radius (R_N) [m]:   0.060;
- wing leading edge radius (R_{WN}) [m]:   0.040.

The fineness ratio of the fuselage is approximately 5 while the wing aspect ratio is 2.26. A sweep angle of 45° provides a delta wing that does not extend far behind the fuselage and, in addition, minimizes the aeroheating on the sharp leading edge [12]. In order to increase stability, five degrees of dihedral angle is adopted. The vertical tail sweep angle is 45 degrees. The 3D Computer Aided Design (CAD) model of the vehicle is shown in Fig. 3.
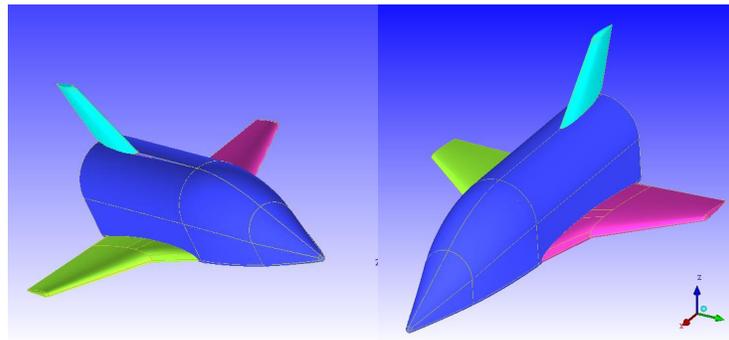

FIGURE 3 : The RTB CAD Image.

The following percentage data of the hypersonic flight demonstrator are provided (see Fig. 4):
- Structure: 25.0 %,
- Subsystems: 53.0 %,
- TPS: 16.0 %,
- Propellant: 6.0 %
keeping in mind that this vehicle represents a trade-off between a compact structure required to reduce reentry heating and a sleeker structure that can glide with an acceptable glide ratio and sink rate.
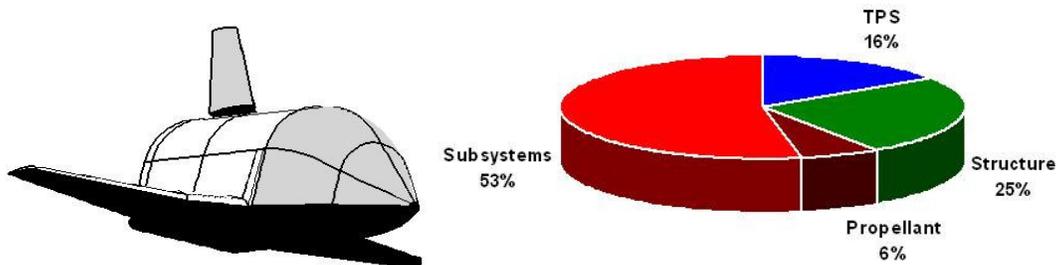

FIGURE 4 : Vehicle TPS Layout and Mass Budget.

3. REENTRY FLIGHT SCENARIO

CFD simulations of the flowfield past the RTB require several freestream conditions as for example Mach number, density, temperature, gas composition, and vehicle trim conditions (e.g. AoA and/or Angle of Sideslip - AoS). Those complete set of CFD input data are provided by the reentry flight scenario (e.g. at every point along the descent trajectory).

To this scope, the vehicle reentry flight scenario has been evaluated. A 6 degree-of-freedom (dof) model is used for describing the dynamics of the RTB and the vehicle performances have been determined by the complete trajectory simulations, from reentry to landing, as done by the code ENTRY developed in [13].

ENTRY performs the descent analysis once the complete data set of the vehicle aerodynamic coefficients, as function of the AoA, AoS, Reynolds number, and Mach number is provided.
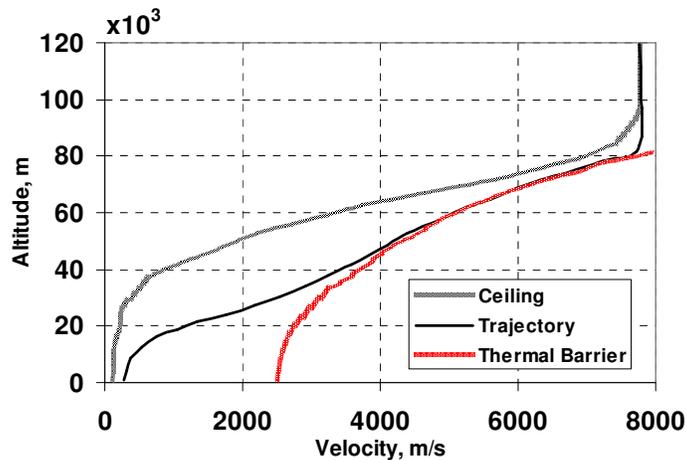

FIGURE 5 : Vehicle Reentry Scenario.

Fig. 5 shows the reentry design trajectory used to assess the vehicle aerothermodynamic environment for TPS design scopes; the reentry corridor is also shown. Time histories of reentry altitude and AoA are summarized in Fig.6.
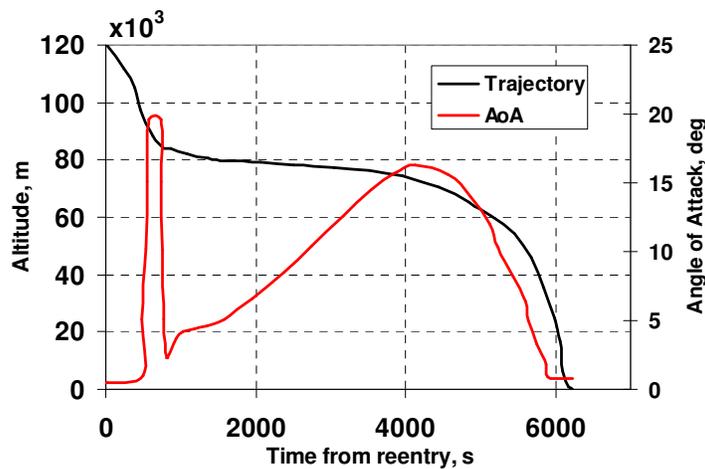

FIGURE 6 : Time History of Altitude and AoA.

As shown in Fig. 5 and Fig.6 , the reentry scenario envisages an enduring atmospheric gliding flight, during which the vehicle flies for a long period very close to the thermal barrier (e.g. the floor of reentry corridor). This extended-duration trajectory present a formidable challenge to TPS design due to the potential for large heat soaks into the airframe.

The vehicle, however, takes advantage of radiative equilibrium cooling i.e., the total aeroheating to the wall (both conductive and diffusive heating) is assumed to be equal to that re-radiated from the surface.

During the period of maximum aeroheating, as shown in Figs. 7 and 8, the vehicle Guide Navigation and Control (GNC) strategy is based on the AoA modulation. Therefore, the flight conditions at the AoA peak determines the freestream conditions to consider in the numerical computations of the vehicle aeroheating, as will be well explained hereinafter.

The aerothermodynamic environment of the airframe fuselage, nose, and leading edges have been calculated using again the ENTRY engineering code, as recognized in the next paragraph.

4. AEROHEATING ASSESSMENT

During reentry the RTB suddenly heats due to the dissipation, in the boundary layer, of its high energy (potential and kinetic) by friction with the atmosphere.

Knowing the freestream density, flight speed, nose and wing leading edge radii, and enthalpy variation with temperature, the stagnation point heat flux can be computed, thus providing a preliminary assessment of the vehicle aeroheating environment.

The simplest method for estimating hypersonic aerodynamic heating refers to the following general relationship [14], [15]:

$$\dot{q} = C \, \rho_\infty^a \, v_\infty^b \tag{1}$$

where $\dot{q}$ is in [W/m^2], if the free-stream density ρ is given in [Kg/m^3], and the flight velocity v is given in [m/s].

With that formula one is able to describe the heat transfer at vehicle leading edges (i.e. nose, wing, and tail) and to flat bottom vehicle surfaces (i.e. fuselage and wing) if appropriate values for **C**, **a** and **b** are provided. If:

$$\begin{cases} C = 1.83\text{x}10^{-4} \dfrac{1}{\sqrt{R_N}} \left(1 - \dfrac{h_w}{h_0} \right) \\ a = \dfrac{1}{2}, \quad b = 3 \end{cases} \tag{2}$$

the stagnation point heat transfer at fuselage nose ($\dot{q}_N$) is described according to Ref. [15]; R_N is the nose radius in meters, $h_w = c_{pw} T_{rw}$ and $h_o = \dfrac{v_\infty^2}{2} + c_p T_\infty \cong \dfrac{v_\infty^2}{2}$ are the wall and total enthalpies, respectively [15].

In order to perform the vehicle windward surface heating analysis ($\dot{q}_{FP}$), if we consider the laminar flat plate case and the condition that $M_\infty \sin\phi > 1$, it follows that:

$$\begin{cases} C = 2.53\text{x}10^{-5} \dfrac{\sin\phi}{\sqrt{x}} \sqrt{\cos\phi} \left(1 - \dfrac{h_w}{h_o} \right) \\ a = \dfrac{1}{2}, \quad b = 3.2 \end{cases} \tag{3}$$

where ϕ is the local body angle with respect to the free-stream (i.e. the local surface inclination plus the AoA), x is the distance, in meters, from the stagnation point measured along the body surface [14].

This last correlation formula highlights how it is preferable to have a vehicle windside flat surface because ϕ, in this case, is only the AoA.

To estimate the heat transfer at wing and tail ($\dot{q}_{LE}$) (i.e. vehicle swept leading edge) one can take advantage of both the above correlation formulas. In fact :

$$\dot{q}_{LE} = \sqrt{\left(\frac{1}{\sqrt{2}} \dot{q}_N \cos\Delta \right)^2 + \left(\dot{q}_{FP} \sin\Delta \right)^2} \cos\alpha \tag{4}$$

where Δ is the leading edge sweep angle [14].

Naturally, the validity of this analysis is reasonable as long as flight conditions are such that the boundary layer theory is valid. Note that if we let the characteristic length be a running distance x from the nose or leading edge of the vehicle, then the Knudsen number becomes infinite when x=0 (Kn=λ/x, where λ is the mean free molecular path). To address the effects of rarefaction on the nose heating calculation, the free molecules (fm) heating rate ($\dot{q}_{fm}$) is estimated by Eq. (1) with:

$$\begin{cases} C = \dfrac{St_o}{2}\left(1 - \dfrac{h_w}{h_o}\right) \\ a = 1, \quad b = 3 \end{cases} \qquad (5)$$

The Stanton number at the stagnation point (St_o) is a function of the Cheng and Chang rarefaction parameter (K_r^2), that is close to the Reynolds number behind the bow shock wave [16]. In Ref [16] it is shown that K_r^2, in the range of interest, leads to $\dot{q}_{fm} > \dot{q}$. So the continuum formulas, Eq. (1) and Eq. (4), are conservative.
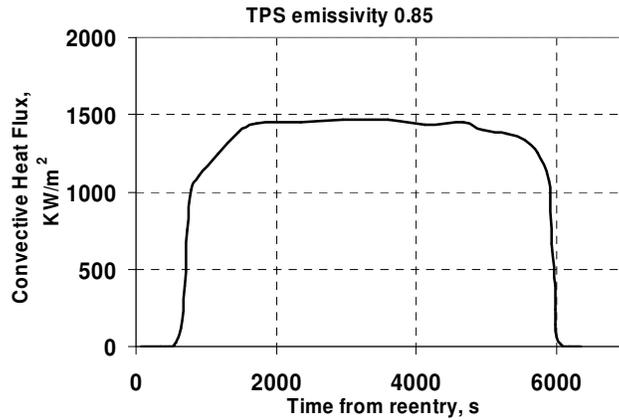

FIGURE 7 : Heat Flux Profile During Reentry.

The vehicle aerothermal environment as predicted by Eq. (1) is summarized in Fig. 7, where the convective heat flux, with the wall in radiative equilibrium condition, is reported for the vehicle nose.

The time history of the corresponding radiative equilibrium temperature can be recognized in Fig.8, supposing that the TPS emissivity is equal to 0.85.
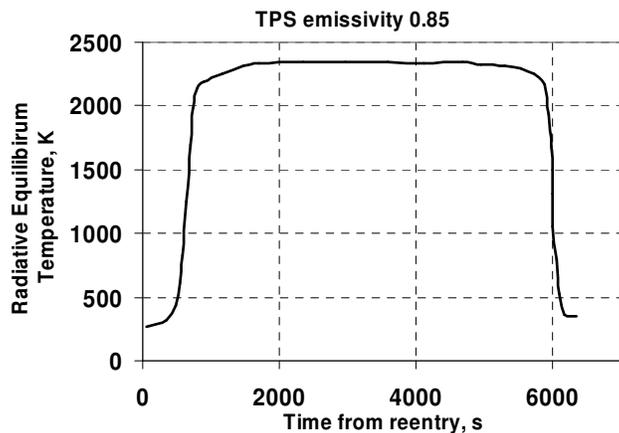

FIGURE 8 : Radiative Equilibrium Temperature.

This overall analysis appears to be conservative if one takes into account the effect of material thermal conductivity of the TPS [17].

However, the heat flux profile is only one criterion in the design of the TPS. For example, also the pressure is important since it determines the TPS structural strength, to maintain the vehicle aerodynamic shape [18]. The other important criterion to consider is the area under the heat flux curve (i.e. the integrated heating rate over the time allotted by the descent trajectory). This area is also referred to as integrated heat load (IHL):

$$Q(\vec{r}) = \int_{t_i}^{t_f} \dot{q}(\vec{r}, t) \, dt \qquad (6)$$

In fact, while the peak heat flux guides the selection of Thermal Protection Material (TPM) able to withstand the heat flux peak, the IHL determines the thickness of the TPM (i.e. the TPS thickness determines the thermal budget that the structure has to manage to protect the "cold" vehicle structure). Therefore, in the framework of TPS vehicle design, one should select on the vehicle reentry trajectory, a number of points for computational analysis (control points), able to replicate the area under the heat pulse. The leftmost point (early on the trajectory) on the heat pulse is chosen from continuum considerations, (i.e., the Knudsen number based on the vehicle characteristic dimension does not exceed 0.001 and the convective heat flux has a "reasonable value"). On the contrary the rightmost point on the heat pulse is chosen from either a desired dynamic pressure, or Mach number limit (e.g., $M_\infty \leq 4$), while the other points are distributed between these two. These engineering evaluations, however, are not sufficient to guarantee a reliable estimation of the vehicle aeroheating. More accurate computation is needed by using CFD analyses, as done in the next paragraph.

4.1 Aerothermal Analysis by CFD

As far as the heat flux distribution over the vehicle surface is concerned, a trajectory check point (freestream conditions at the flight point) has been considered, in the framework of a trajectory-based design approach. This point, shown in Fig.9, represents the most challenging flight conditions from the aeroheating point of view, since the vehicle is flying along the thermal barrier at the higher AoA (see Fig.5 and Fig.6). The freestream conditions are Mach$_\infty$=22.4, H$_\infty$=73.1 [Km], and AoA=16.5 [deg], as given in Table 1.

TABLE 1 : Freestream Conditions for CFD Analysis.

Altitude [km]	Pressure [Pa]	Temperature [K]	Density [Kg/m³]	Mach [-]	AoA [deg]
73.1	3.28	212.3	5.38×10^{-5}	22.4	16.5

4.1.1 Heat flux modeling

The aeroheating analysis of a vehicle reentering from space demands different physical approach to implement in the numerical simulations. In fact, as recognized in Fig. 5 and Fig. 9, the vehicle experiences different flight regimes ranging from free molecular flow at very high altitudes to complete continuum deep in the atmosphere. Most of the heating occurs, however, in the continuum regime, where the CFD approach, based on the integration of the Navier-Stokes equations (with appropriate physical models for shock layer processes), is adequate [19].

Generally speaking the air is modeled as a mixture of several species (up to 11 species), each one assumed to be thermally perfect, and the main requirements for the models are: *thermodynamic* and *transport* (mass, momentum, and energy) properties of the constituent species; accurate representation of reactions (e.g. *reaction mechanism*), and their associated rates (e.g. *chemical kinetics*) in the shock layer; the models for *thermal nonequilibrium*, if necessary, as in the case of substantial radiative heating.

A rough estimation of how many chemical species to consider in the numerical computation may be found analyzing the post-shock species distributions provided by a simple one-dimensional equilibrium computation for a normal shock corresponding to the freestream conditions at the trajectory points.

Moreover, there are three important modeling issues that need further attention within aeroheating analysis: *gas-surface interaction*, *transition*, and *turbulence* [20].

At the vehicle surface, apart from the usual "no-slip" and zero normal pressure gradient boundary conditions at the wall, mass and energy balance equations are necessary to represent the interaction of the gas and surface (e.g. heterogeneous reactions).

The mass balance equations are obtained considering that the flux due to mass diffusion is balanced by the production of molecular species through recombination of atoms:

$$\vec{J}_{s_w} \cdot \hat{n}_w = \rho_{s_w} \gamma_{s_w}(T_w) \sqrt{\frac{\Re T_w}{2\pi M_s}} \qquad (7)$$

When $\gamma_s=1$, the wall permits complete recombination of atoms at the surface (e.g. Fully Catalytic Wall – FCW). If $\gamma_s=0$, the surface does not permit recombination and it is said to be noncatalytic (NCW). The heat released due to recombination at wall is maximum for FCW, and zero for NCW. However, for a real TPM, the recombination coefficient is between the two extremes, i.e., $0 \leq \gamma_s \leq 1$ and is a characteristic of the material.

Further, recombination coefficients are functions of temperature, i.e., $\gamma_s =\gamma_s(T)$. As preliminary TPS design criteria, the conservative assumption of a FCW is preferred since one can expect maximal heat release from recombination. Note that in the case of air, the surface is assumed to be noncatalytic to NO and permits only recombination of N and O. Secondly, radiative equilibrium is assumed to exist at the wall.

Therefore, the energy balance equation at the surface, neglecting in-depth conduction through the TPS, reads:

$$-k_w \vec{\nabla} T \cdot \hat{n}_w + \sum_{s=1}^{n_s} h_s(T_w) \vec{J}_{s_w} \cdot \hat{n}_w = \sigma \varepsilon_w T_w^4 \qquad (8)$$

where the emissivity, in general, depends on the temperature and the type of TPM, i.e., $\varepsilon_w=\varepsilon_w(T_w, TPM)$.

The other important issues that affect the convective heating at the wall are transition and turbulence. Since the onset of transition cannot be predicted *a priori*, the results from laminar computations are post-processed for boundary-layer momentum thickness and edge Mach number. The ratio of the momentum thickness Reynolds number to the edge Mach number is used as a guide to determine the onset of transition empirically through correlation of computed laminar boundary-layer parameters (notably the momentum thickness) with experimental data. Note that this is only one of many criteria, and assumes the body is smooth. Irregularities in the surface — either roughness or steps/gaps — could cause transition to occur earlier.
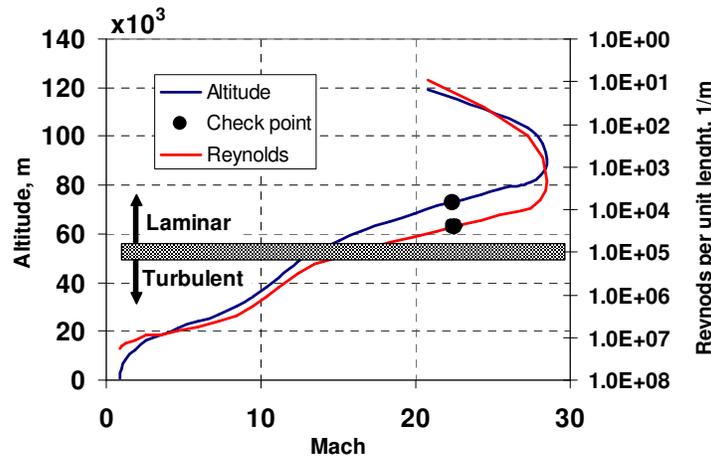

FIGURE 9 : Reentry trajectory in Altitude – Mach and unit Reynolds-Mach planes.

Assuming that onset of transition can be determined using the momentum thickness Reynolds number criterion, the length of the transition region must be predicted. For the lack of a good transition model, the assumption of a fully turbulent flow is usually made, and an algebraic turbulence model is used, e.g., the Baldwin Lomax model corrected for compressibility — reasonably good for attached flows but not so for separated leeside flows. Such an assumption can be excessively conservative in the design of the TPS — simply due to predicted high levels of heating in forward part of the configuration.

The shaded region in Fig. 9 refers to unit Reynolds number equal to 10^5 above which the flow is assumed fully turbulent. Thus, a selected flight condition lying under this boundary (see Fig. 9) requires a turbulent flow computation. It is worth nothing that the unit Reynolds number, during the critical aeroheating phase of atmospheric descent, is less than the threshold value, i.e. it is above the dashed region of Fig. 9. Therefore, for the computations reported in this work the issues of transition and turbulence are neglected.

4.1.2 Grid generation

In the framework of the flow field computational analysis a relevant issue is to build an accurate volume mesh. The multiblock mesh on the vehicle surface is shown in Fig.10.
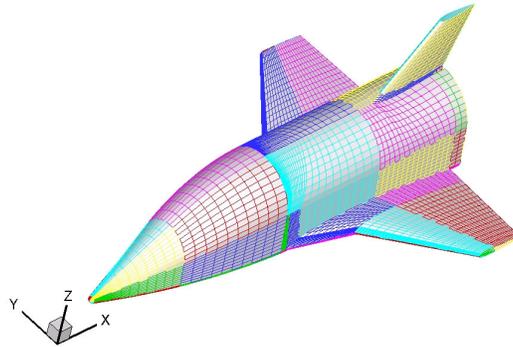

FIGURE 10 : Multiblock Mesh on the Vehicle Surface.

Since no sideslip flow has been taken into account, the grid has been built only over either the port or starboard half of the vehicle. The distribution of surface grid points is dictated by the level of resolution required in various areas, e.g., the bow shock-wing shock interaction region for a winged vehicle requires fine resolution. The distribution of grid points in the wall-normal direction are driven by the freestream Reynolds number in order to require adequate spacing to resolve the thin shear layer bound to the wall. The grid is then tailored for the freestream conditions at the selected trajectory point (see Table 1).

4.1.3. CFD results
CFD analyses of the vehicle along the descent flight path are needed to confirm (and/or to correct) the results of the engineering approach followed above; and to assess that, during reentry, the radiative equilibrium temperature distribution of the skin of the remaining vehicle structure (fuselage and wing), made of conventional material, be below the tolerable limits (i.e. about 1500 [K] for the PM1000).

The flow conditions considered for the numerical simulations are summarized in Table 1. Computation is performed at (conservative) steady state conditions and solves the flowfield using a local radiative equilibrium wall-boundary condition. Moreover, simulations are performed at laminar flow conditions with the flow modeled as perfect gas with constant specific heat, thermal conductivity described by kinetic theory and viscosity derived from Sutherland law; these assumptions are acceptable since the air can be assumed as "frozen" gas mixture because of the streamlined vehicle configuration. Numerical computations show that the flowfield is dominated, in fact, by the presence of a weak bow shock wave, as shown in Fig.11.
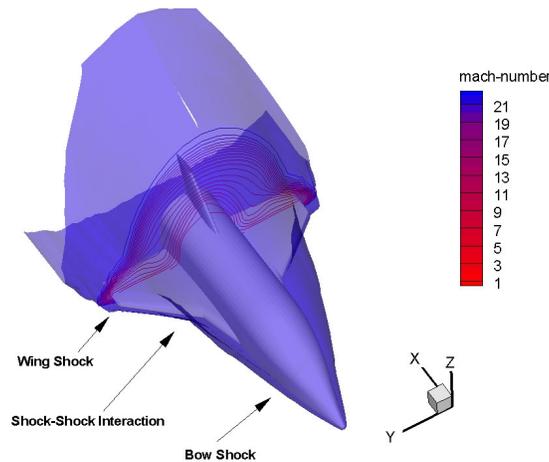

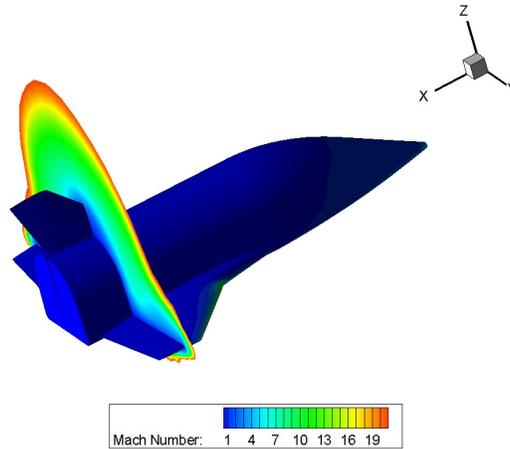**FIGURE 11 :** Shock Shape with Mach Number Contours.

FIGURE 12 : Mach Number Contours in the Vehicle Cross-Section at x=5.5 [m].

Looking at Fig.11 and Fig.12 one can appreciate the shock shape that occurs ahead of vehicle at this flight conditions, with the Mach number contour field on the vehicle cross section at x=5.5 m (see Fig.12). As one can see, the shock surface envelopes the vehicle, and since the vehicle features a small radius of curvature, the stand-off distance is quite small. Therefore, the bow shock may impinge on wing leading edge and vertical fin thus increasing locally the heat flux (overheating) that the vehicle thermal shield has to withstand.
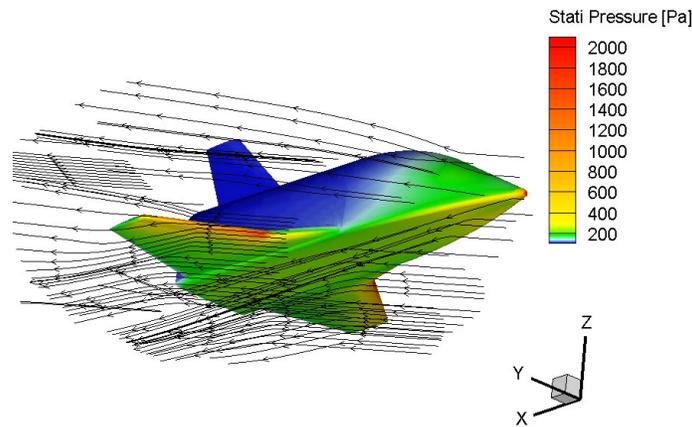

FIGURE 13 : Flow-field Streamlines with Static Pressure Contours on the Vehicle.
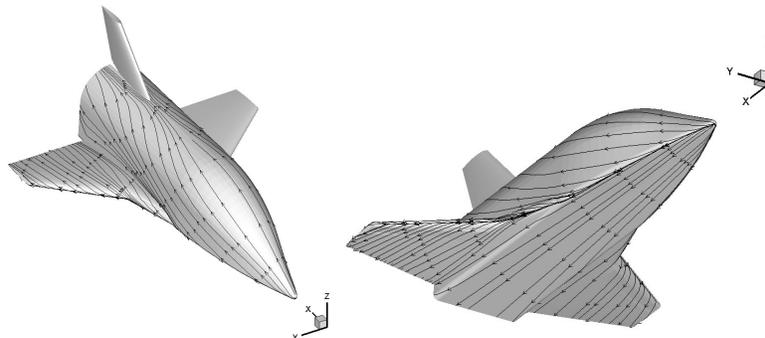

FIGURE 14 : Vehicle Skin Friction Patterns.

The flow-field streamlines around the vehicle can be recognized in Fig.13, where the pressure distribution on the vehicle surface is also reported. In Fig.14 it is shown the surface streamlines distribution on the whole

vehicle configuration. As one can see on the fuselage forebody and strake, 3D effects are present as highlighted by the curvature of the lines at the wall, whereas the wing part between the tip and the strake experiences a quasi 2D flowfield.

The three-dimensional distributions of pressure and heating on the vehicle were evaluated to establish TPS design guidelines. For example, both Fig.15 and Fig.16 report the static pressure distributions on the vehicle surface. In particular, Fig.16 shows the detail of pressure distribution on wing leading edge, thus highlighting the effects (e.g. pressure spike) of the bow shock impingement (known as Shock-Shock Interaction) on Wing Leading Edge (WLE), (see also Fig.11).
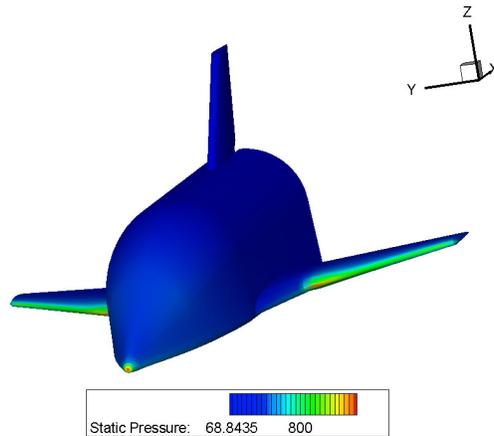

FIGURE 15 : Contours of Static Pressure [Pa].
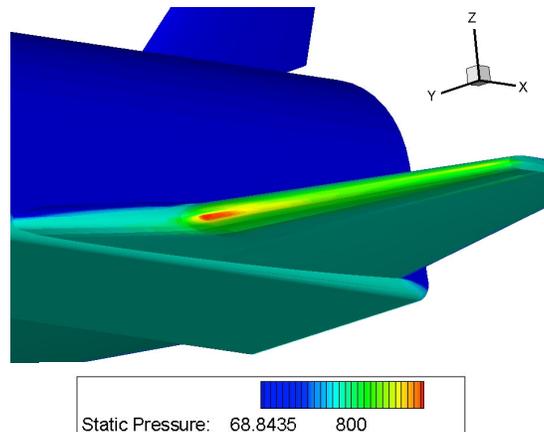
FIGURE 16 : Contours of Static Pressure [Pa]; Detail on the Wing Leading Edge.

Figure 17, Fig. 18 and Fig. 19 show the surface distribution of the radiative equilibrium temperature. The TPS emissivity is supposed equal to 0.85. These distributions highlight that, on the spacecraft, the aeroheating is comparable with that of windward. Therefore, it would be reasonable to design the leeward TPS distribution similar to that of windside.

As a result, careful attention has to be made in defining the vehicle TPS layout. In fact, for the conventional reentry vehicle such as the Space Shuttle, the peak heating conditions occur when the spacecraft is flying at 40 deg of AoA, thus exposing to the oncoming flow only the windward side that is the most heated vehicle part. It is worth to note that, as shown in Fig.17, the CFD confirms the results provided by Eq.(1) for the vehicle stagnation region.

Then, candidate TPS materials may be selected to accommodate the maximum radiation equilibrium temperatures shown in Fig. 17, Fig. 18 and Fig. 19, taking into account that: the actual surface temperatures

of the vehicle will be lower than these radiation equilibrium temperatures because of heat-sink effects of the structure; and transient thermal analyses is mandatory in order to determine TPS thickness of the vehicle thermal shield.
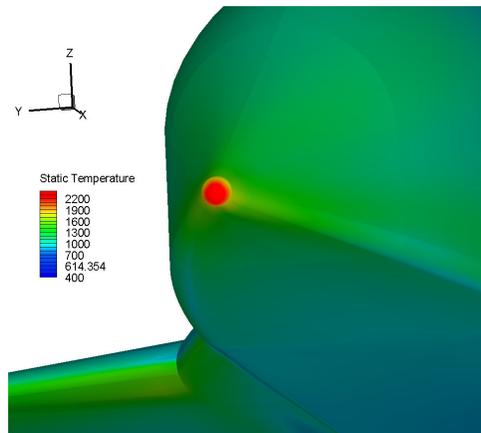
FIGURE 17 : Contours of Radiative Equilibrium Temperature [K]; Detail of the Nose Region.
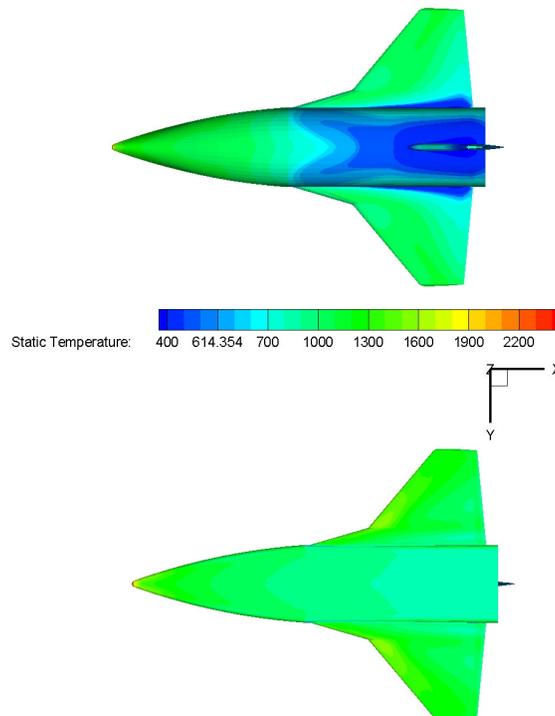
FIGURE 18 : Contours of Radiative Equilibrium Temperature [K] on the
Leeside (up) and Windside of the Vehicle (down).

Moreover, it can be concluded that an analysis of SSI with overloads (pressure and heat flux) at impingement are mandatory for a reliable vehicle TPS layout as vehicle design proceeds, since difference on heat flux distribution are expected moving in span wise direction toward the tip (see Fig. 19). Indeed, the part of the wing that is behind the vehicle bow shock experiences different flowfield conditions with respect to the freestream ones: the most heated zone of the wing is the one inside the bow shock.

Note that the position of the high heat load region localized on the WLE by SSI changes along the atmospheric descent, depending on Altitude, Mach and AoA reentry profiles (i.e. it varies from point to point along the trajectory). Moving along the leading edge, in the span wise direction, there is a quite constant temperature value due to the constant leading edge radius.

Static Temperature [K]:   400   589.474 778.947 968.421 1157.89 1347.37   1590

FIGURE 19 : Contours of Radiative Equilibrium Temperature [K] with
Skin Friction Lines; Detail on the Wing Leading Edge.

A developmental TPM concepts, known as Ultra High Temperature Ceramics (UHTC) were identified that seems to offer significant advantages in thermal shielding capabili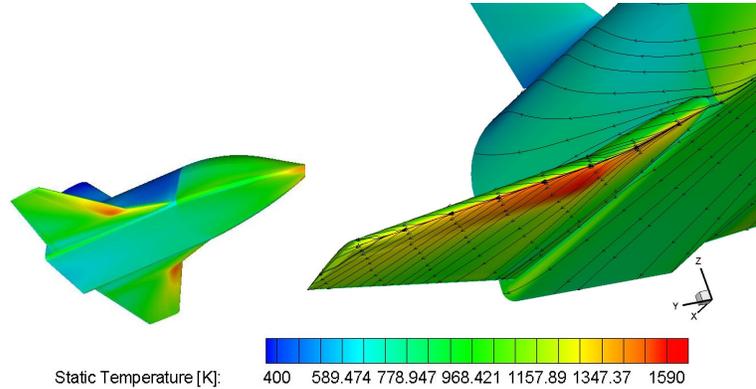ty when the temperature involved are in exceed of 1700 [K] [20]. UHTC, in fact, are able to withstand to temperature up to 3000 [K] as the HfB_2 (Hafnium diboride). Therefore they are very promising for the operability of the future hypersonic vehicles and represent an enabling technology for next generation RLVs. These TPM concepts are being developed by the NASA Ames Research Center as part of the "next generation reusable TPS" effort [6], [20]. An advanced metallic TPS concept with structure fabricated from Inconel alloy honeycombs was also considered for the vehicle surfaces away from leading edges.

Finally, in Fig. 20 and Fig. 21 it is shown the bow shock shape in a xy plane passing through the vehicle WLE. One can clearly appreciated where the SSI takes place on the WLE.

Note how the bow shock is close to the vehicle forebody due to its slender configuration. This confirms that the real gas effects are expected to be negligible compared with those take place for bluff body configuration such as that of Space Shuttle.
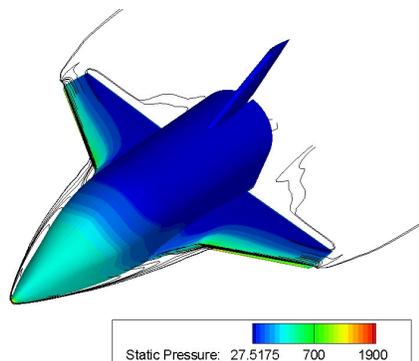

Static Pressure: 27.5175   700   1900

Figure 20 : Contours of Static Pressure [Pa] with Shock Shape Trace
in a xy Plane Passing through the Vehicle WLE.

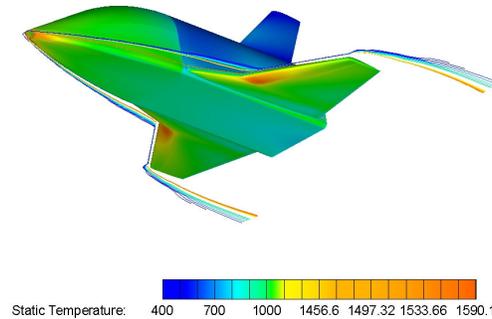Static Temperature: 400 700 1000 1456.6 1497.32 1533.66 1590.1

FIGURE 21 : Contours of Radiative Equilibrium Temperature [K] with Shock Shape Trace
in a xy Plane Passing through the Vehicle WLE.

5. CONCLUSIONS

The aero-thermal environment of a TSTO flying test bed has been assessed both from the engineering based and CFD based approaches. Computations with the air modelled as perfect gas highlight that the vehicle aeroheating is more severe than that of existing reentry vehicles, since the RTB features a streamlined configuration and flies at low angle of attack in order to improve spacecraft aerodynamics.

Also, new technologies for RLV, for which strong necessity of in-flight testing exists, were identified. Although there are many mature TPS materials, only advanced thermal shield using TPM such as UHTC, being developed, will be able to withstand the radiative equilibrium temperatures reached during reentry. A major task of this work has been the study of vehicle concepts and related enabling technologies required for the goal of a highly safe return to Earth.

Finally, we observe that the RTB provides a completely novel approach to the reentry problem. It has some very attractive features, although some are yet unproven, that are mostly derived from its low wing loading. There appears a strong case for continuing the current project study: the development of a European reusable launch vehicle, for which hypersonic flight demonstration is an indispensable next step.

6. REFERENCES

[1]   Dumbacher D., *"NASA's Second Generation Reusable Launch Vehicle Program Introduction, Status and Future Plans".* 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Indianapolis, Indiana, July 7-10, 2002. AIAA-2002-3613.

[2]   Whitmore S., Dunbar B., *"Orbital Space Plane, Past, Present, and Future".* AIAA International Air and Space Symposium: The Next 100 Years, Dayton, Ohio, July 14-17, 2003. AIAA-2003-2718

[3]   Rasky D., *"Access from Space: A New Perspective on NASA's Space Transportation Technology Requirements and Opportunities".* Space 2004 Conference, San Diego, California, Sep. 28-30, 2004. AIAA-2004-6103.

[4]   Reuther, J., Kinney, D., Smith, S., Kontinos, D., Gage, P., Saunders, D., *"A Reusable Space Vehicle Design Study Exploring Sharp Leading Edges".* AIAA Thermophysics Conference, 35th, Anaheim, CA, June 11-14, 2001. AIAA-2001-2884.

[5]   Saunders, D., Allen, G. Jr., Gage, P., Reuther, J., *"Crew Transfer Vehicle Trajectory Optimization".* 35th AIAA Thermophysics Conference, Anaheim, CA, June 11-14, 2001. AIAA-2001-2885.

[6]   Kolodziej, P., Bowles, J. V., Roberts, C., *"Optimizing Hypersonic Sharp Body Concepts from a Thermal Protection System Perspective".* 8th AIAA International Space Planes and Hypersonic Systems and Technologies Conference, Norfolk, VA, Apr. 27-30, 1998, paper AIAA-1998-1610.

[7]   Chen, Y. K., Milos, F. S., Bull, J. D., Squire, T. H.,   *"Integrated analysis tool for ultra-high temperature ceramic slender-body reentry vehicles".* 37th Aerospace Sciences Meeting, Reno, NV, Jan. 11-14, 1999. AIAA-1999-350.

[8]   Kontinos D. A., Gee K., Prabbu D. K., *"Temperature constraints at the sharp leading edge of a Crew Transfer Vehicle".* 35th AIAA Thermophysics Conference, Anaheim, CA, June 11-14, 2001.

[9]   Lee C. H., Park S. O., *"Aerothermal performance constraints for small radius nosetip at high altitude".* AIAA/NAL-NASDA-ISAS 10th International Space Planes and Hypersonic Systems and Technologies Conference, Kyoto, Japan, Apr. 24-27, 2001. AIAA-2001-1823.

[10] Jeroen Buursink, Kees Sudmeijer, *"Experimental Studies of an Enhanced Radiation Cooling System"*. Space 2004 Conference and Exhibit, San Diego, California, Sep. 28-30, 2004. AIAA-2004-5827.

[11] Smith S., Reuther J., Kinney D., Saunders D., *"Low speed aerodynamics and landing characteristics of Sharp-class Crew Transfer Vehicle concepts"*. 35th AIAA Thermophysics Conference, Anaheim, CA, June 11-14, 2001.paper AIAA-2001-2888.

[12] Gomg L., Ko W.L., Quinn R.D., "*Thermal Response of Space Shuttle Wing During Reentry Heating"*. NASA TM 85907, June 1984.

[13] Viviani A., Pezzella G., Cinquegrana D., *"Aerothermodynamic Analysis of an Apollo-like Reentry Vehicle"*. AIAA-2006-8082.

[14] Tauber E., *"A review of High-Speed, Convective, Heat-Transfer Computation Methods"*. NASA TP-2914, Jul 1989.

[15] Anderson J. D., *"Hypersonic and High Temperature Gas Dynamics"*. McGraw-Hill (series in Aeronautical and Aerospace Engineering), 1989.

[16] Cheng H.K., Chang A.L. *"Stagnation Region in Rarefied High Mach Number Flow*". AIAA J.,1(1),1963.

[17] Loomis M., Palmer G., *"Pre-flight CFD analysis of arc jet and flight environments for the SHARP-B2 flight experiment"*. 39th Aerospace Sciences Meeting, Reno, NV, Jan. 8-11, 2001. AIAA-2001-982.

[18] Kinney, D. J., Bowles, J. V., Yang, L. H., Roberts, C. D., *"Conceptual design of a SHARP–CTV"*. AIAA Thermophysics Conference, 35th, Anaheim, CA, June 11-14, 2001. AIAA-2001-2887.

[19] Zuniga F., Cliff S., Kinney D., Hawke V., Tang C., Smith S., *"Vehicle Design of a Sharp CTV Concept Using a Virtual Flight Rapid Integration Test Environment"*. AIAA Atmospheric Flight Mechanics Conference and Exhibit, Monterey, California, Aug. 5-8, 2002.AIAA-2002-4881.

[20] Johnson S., Gasch M., Leiser D., Stewart D., Stackpool M., Thornton J., Espinoza C*., "Development of New TPS at NASA Ames RC"*. 15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference, Dayton, Ohio, Apr. 28-1, 2008. AIAA-2008-2560.

Sea Water Air Conditioning [SWAC]: A Cost Effective Alternative

A.F. Elsafty   aelsafty@aast.edu

*Associate professor-College of Engineering
and Technology/Mechanical Engineering Department
Arab Academy for Science, Technology
and Maritime Transport- AASTMT
Alexandria, POBox: 1029, EGYPT –
http://www.aast.edu*

L.A. Saeid   loai_Saeid@natgas.com.eg

*GIS, Design Section head,
National Gas Company,
Alexandria,
EGYPT - http://www.natgas.com.eg*

Abstract

The energy demand for air conditioning is quite extensive due to the hot and humid summer climate in Egypt. The rapid increase in non industrial electricity consumption is due to the rural electrification and the presence of many buildings air conditioned in summer using electricity.

Deep cold ocean and seawater is a valuable natural resource that can be used for energy production, cooling, desalination, aquaculture and agriculture. The most economically viable use of this deep water is to air-condition buildings through a Sea Water Air Conditioning (SWAC) system.

This study reports the results of a technical and economical assessment of the potential for using (SWAC) other than conventional vapor compression systems to air condition hotels at a new tourists resort called "Sahl-Hasheesh",18km south of Hurghada, Egypt.

This study analyzed and sized the major components of the Sea Water Air Conditioning (SWAC) system, determined the operational performance, and estimated the probable costs. The economic analysis was based on two different methods, the simple pay back and the net present value (NPV) method.

The results showed that the SWAC system is the preferred option for its short payback period as well as the minimum net present value when being applied at Sahl-Hasheesh area. Large energy savings approaching 80% compared to conventional. This is in addition to the low greenhouse gas emissions.

Keywords: *Seawater, HVAC, Sahl-Hasheesh, Economical Study.*

Nomenclatures

Symbol	Description	S.I Units	Symbol	Description	S.I Units
CW	Chilled water	-	i	Interest rate	-
Dep C	Depreciation Coefficient	-	IC	Initial cost	-
Dep K	Depreciation at the year K	-	L	Pipe length	m
HDPE	High Density Polyethylene	-	m^o	Mass flow rate	kg/s
NPV	Net Present Value	-	NOx	Mono nitrogen oxides	-

PLC	Programmable logic controller	-	SOx	Sulphur oxides	-
QA/C	Air conditioning load	kW	SW	Seawater	-
RC	Running Cost	-	VAS	Vapor Absorption System	-
SDR	Standard Dimension Ratio	-	VCS	Vapor Compression System	-
TOR	Ton Of Refrigeration				

1. INTRODUCTION

All current air conditioning systems depend mainly on electricity or heat source to operate their various components, the two main known types of air conditioning are the vapor compression system (electric operated) and the absorption system (Heat operated).

A growing number of scientists and engineers have become concerned about global climate change. This phenomenon shows a strong correlation to human use of fossil fuels. Exponential growth in the build-up of combustion products trapped within Earth's atmosphere is implicated as the primary cause of the "Greenhouse Effect". [1]

The amount of greenhouse gas emissions is expressed in tons of "carbon dioxide equivalents". CO2 has the largest warming potential and as such is used as an index for other greenhouse gases. Egypt ranked the 27th over the world in producing CO_2 emissions year 2006 [2].

Furthermore, one of the main fossil fuels, petroleum, is a finite resource and has been a focus of international conflict. This may be considered as another main reason for searching for new and renewable energy sources or lowering the power consumption.

In order to make a reasonable assessment of the technical and economic feasibility of deep sea water air conditioning, three options were investigated in the economical study.

The first is the use of a conventional air conditioning system. This option provides a baseline for the other options being investigated. The second option is the use of deep seawater only and the third option involves the use of a hybrid system using both a sea water air conditioning and a conventional chiller in series where part of the AC demand will be held by the SWAC and the rest of the demand by the chiller.

Additionally, this study introduces a new air conditioning technique in the Middle East region known as Sea Water Air Conditioning (SWAC) system utilizing a renewable energy source to reduce the electricity consumption. The reduction of the greenhouse gases was estimated.

2. CASE STUDY

SITE (LOCATION)

SAHL HASHEESH [SH] (located 18km south of Hurghada - Upper Egypt) is to be a Resort Community project of a scale and scope that is unprecedented in the region. The project promises to become an Integrated Resort destination of world-class standards.

The typical meteorological year database for Sahl-Hasheesh was used to estimate the gross cooling load for the hotels. HVAC Load Explorer program [3] was used to calculate the air conditioning demands of the on duty hotels and the results were checked manually.

SH is planned to contain 24 hotels at the seaside. Survey has been taken for the site's hotels. Air conditioning load was calculated for "Pyramisa hotel" (on duty hotel) using the (HVAC Load Explorer program) [3] and based on the area of each hotel and the numbers of rooms, the air conditioning load for the other hotels was estimated according to ASHRAE [4,5].

It turns that the sum of all the loads was estimated to be 26,500 TOR. Figure 1 shows the AC load map for Sahl-Hasheesh.

FIGURE 1: Sahl-Hasheesh Master Plan & Load Map

SWAC SYSTEM

A Sea Water Air Conditioning district cooling system consists of a cold seawater supply line, a heat exchanger (at the shoreline), and a closed cycle fresh water distribution system, all with appropriate pumps as shown in Figure 2.
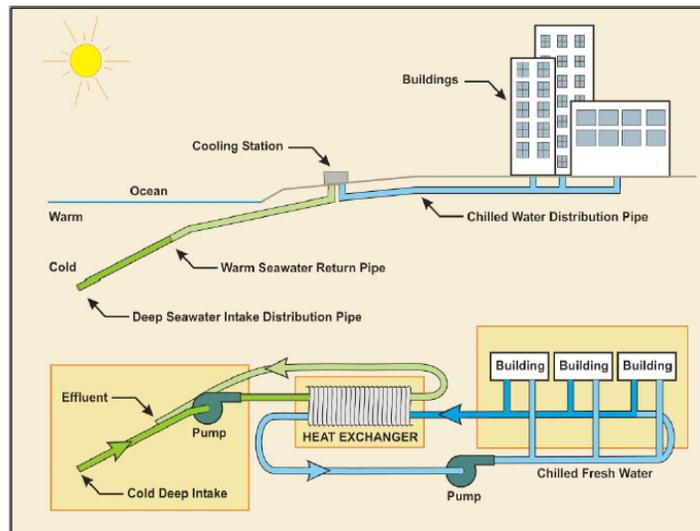

FIGURE 2: SWAC Schematic diagram

Cold seawater is drawn from 600 meters deep at a temperature of 7°C. It follows a long pipeline that lies along the seabed about 500 meters from shore. Pump station delivers the water into a cold water distribution pipe buried under the beach then after getting warmed from the heat exchange.

It is well known that the discharge of thermal heat into seawater imposes an environmental and biological impact on the marine life. This impact is called "Cold shock". Hence, the seawater is then pumped back to the sea through an effluent pipe at 200 meters depth to avoid biological effect. On the other side the chilled distribution closed loop exchange heat with the air to be conditioned.

For air conditioning hotels in SH, 9-10°C water is needed to circulate inside the buildings based on Pyramisa hotel request [6], taking into consideration the low relative humidity in SH resulting in low latent cooling load.

To obtain this low temperature for the fresh water, a lower temperature of 7°C shall be drawn out of the sea circulating through the heat exchanger.

For this purpose an approximate temperature depth profile at SH shown in Figure 3 was obtained for the red sea from the navy forces [7].

Local Bathymetry
From Figure 3, a depth of about 600 m is determined to achieve the desired water temperature (7°C). The SWAC system technical evaluation method begins by outlining coastal regions where the 600 m bathymetric contour (7°C) lies within the minimum distance from shore. For this purpose, bathymetry map for Sahl-Hasheesh area was obtained from the (Egyptian naval forces). By the aid of this map and after the site visit, a pipeline schematic was designed starting from the point at depth 600 m up to the shore at the pump station.
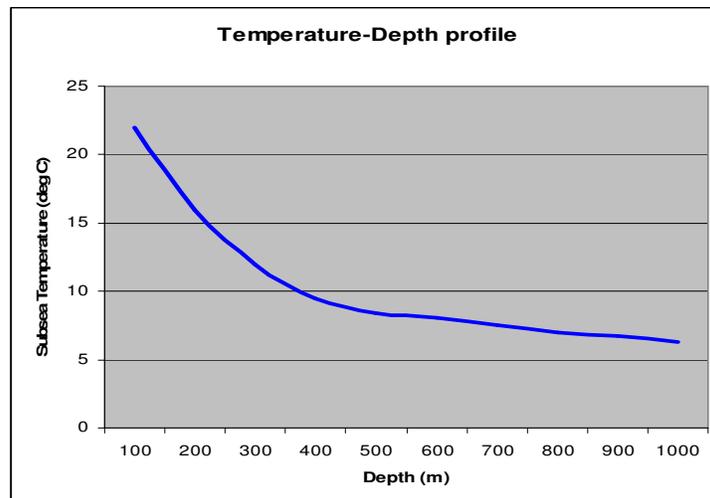
FIGURE 3: Sahl-Hasheesh Temperature-Depth profile

Seawater Pipe
A primary concern regarding pipeline placement is the impact upon SH marine preserve. Two possible pipeline paths have been investigated in this study. Both routes A and B are shown in Figure 4.

Route A is the most direct path between the distribution system onshore and the offshore source, requires the shorter tunneling, not far away from the marine preserve area, and has a landing at the north end of the distribution system. Route B provides cold water to the center of the SH

Elsafty, A. F. & Saeid, L.

distribution system, thus splitting the seawater flow onshore and allowing smaller pipelines in addition not supplying all the hotels from a single line. Route B provides an alternate pipe landing and an alternative pump station location based on the site visit.

HDPE 1000mm (40"), SDR 11 (Standard Dimension Ratio refers to the outside diameter divided by the wall thickness of the pipe with thickness above 76mm (3") is used to withstand the external pressure on the pipe.
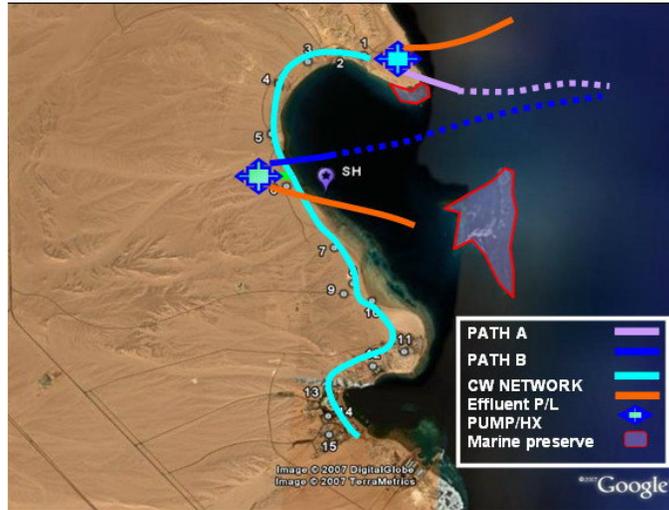
FIGURE 4: Seawater Pipe Route scenarios

Distribution Network
The pump station shall be located halfway of the distribution network in order to decrease the pipe diameters and to secure the availability of AC at least for half of the hotels.

The two lines shown in Figure 5 represent the distribution network feeding the customers where the (North pipeline) represents a pipeline feeding 12,250 TOR and the (South pipeline) represents a pipeline feeding 14,250 TOR of the total expected air conditioning demands in SH.

FIGURE 5: Distribution network pass

Elsafty, A. F. & Saeid, L.

Each pipe of the above two pipes is divided into two sections with determined lengths and known flow rates feeding the hotels in order to design the network.

Based on the required flow rates and the recommended velocity in pipes, a set range of diameters was chosen from 400mm (16") up to 760mm (30") to be the interest in this study.

Head losses were calculated for each section after determining the flow type and Reynolds's number and the friction factor , hence the corresponding pumping power was estimated and the optimum network diameter was established to be 760 mm. [8].

3. RESULTS & DISCUSSIONS

Technical Analysis
A piping schematic is constructed connecting the district loads and the cold-water supply pipe via a heat exchanger.

Computer program is used to optimize the piping network for best ratios of capital expenditure versus displaced electricity, which is the driving variable.

The head losses in all distribution, deep seawater supply, and return pipelines are determined based on optimal pipe sizes that are staggered throughout the system. The required pressures are set for each user and total system pressure is computed to provide all customers the minimum desired pressure differential. The cold water intake pipe is sized based on flow and allowable suction pressure and pump station elevation limitations. The wall thickness of the intake pipe is set to prevent collapse and, if necessary, stiffeners are added to the pipelines. Finally, the pumps are sized based on total fresh water and seawater flows and the total heads for these systems.

Most of the deep seawater intake pipelines designed for by (Makai, 1994) use polyethylene as the pipeline material [9]. Polyethylene has significant advantages for these pipelines in that it is inert and will neither corrode nor contaminate the water. Polyethylene lengths are heat fused together to form a long, continuous pipeline with joints that are as strong as the pipeline itself. Polyethylene has excellent strength and flexibility and is buoyant in water. For the distribution network material, a comparison between the polyethylene, Steel and PVC pipes was done.

For the (SWAC) system, the best choice of heat exchanger is a modular titanium plate heat exchanger with gasket joints in a counter-flow configuration.

SWAC System Summary
Table 1 summarizes the contents of the SWAC system in Sahl-Hasheesh

Distance from shore to the point of the desired depth (m)	520
SW pipe diameter (mm/inches)	1000/40
SW pipe length (m)	10,000
SW velocity in pipe (m/sec)	5
SW pumping power (MW)	4.12
Distribution network (length in meter / diameter in mm (inches))	8900/760(30)
CW pumping power (MW)	5.32

TABLE 1: SWAC system summary.

Economical Analysis
The costs associated with the proposed SWAC system are primarily related to the initial capital expenditure. This, in turn, is related to the distance to the cold water, the temperature of that water, the extent and location of the onshore distribution loop, and the sizes of all pipelines.

Operating costs are related to amount of pumping power required. This is related to the amount of water to be pumped and the size and length of the pipelines.

Initial Costs

Item	Cost ($)
Seawater Pipe	29,200,347
Effluent Pipe	10,220,121
Distribution Network	4,458,900
SW Pump and Sump	27,176,074
Heat Exchanger	13,894,500
CW Pump	1,250,323
20% Contingency	17,240,053
Total	**103,440,318**

TABLE 2: SWAC system initial costs in US$

The costs of the conventional systems are estimated according to "An Introduction to Absorption Cooling" by Harwell, 1999.

The initial cost of the distribution network ($ 4,458,900) shall be added to the vapor compression system as added to the SWAC system.

INITIAL COST			
	Single-Effect VAS	**Double-Effect VAS**	**VCS**
Vapor Absorption System VAS &Vapor Compression System VCS			
Machine Capacity (TOR)	26500		
Cost $ / TOR	684	936	504
Total Cost ($)	18,126,000	24,804,000	13,356,000
Life Time (yr)	16	16	8
Heat Rejection Equipment			
Cost $ / TOR	227	205	151
Total cost ($)	6,015,500	5,432,500	4,001,500
Life Time (yr)	16	16	16

TABLE 3: Conventional Systems Initial Costs in US$

Running Costs

Tables 4 and 5 show the SWAC and conventional running costs respectively. For the vapor absorption system, knowing that it is a heat operated system and a large amount of fuel (Solar energy, Natural gas) has to be located continuously at SH which is hard to obtain because there is no source for such a fuel there in addition applying a solar system is beyond the scope of this study and would be inapplicable due to the very high initial costs required.

Economical comparison was based on two methods: Simple Pay Back and the Net Present Value.

Item	Cost ($)
Seawater Pump	975,785
Chilled water pump	1,258,622

Maintenance & Labor cost	400,000
Total	2,634,407

TABLE 4: SWAC System Running Costs in US$

RUNNING COSTS			
Price of kWh ($/kWh)	**0.045**		
VAS& VCS machines	**Single-Effect VAS**	**Double-Effect VAS**	**VCS**
kWh/TOR/yr	108	108	6224
Total Use (kWh/yr.)	2,862,000	2,862,000	164,954,550
Annual cost $	**128,790**	**128,790**	**7,422,955**
Cooling Water Pump			
Cooling Water Pump motor Efficiency	0.68		
kWh/TOR/yr	739	506	493
Total Use (kWh/yr.)	19,599,400	13,409,000	13,064,500
Annual Cost $	**881,973**	**603,405**	**587,903**
Cooling Tower Fans			
Fans Efficiency	0.6		
Fan partial use factor	0.4		
kWh/TOR/yr	588	480	392
Total Use (kWh/yr.)	15,582,000	12,720,000	10,388,000
Annual Cost $	**701,190**	**572,400**	**467,460**
Natural gas consumption			
NG m3/hr	8,480	8,480	0
m3 / year	44,570,880	44,570,880	0
Gas unit cost $/ m3	0.045454545	0.045454545	0
Annual Gas cost $	**2,025,949**	**2,025,949**	**0**
Total Annual Operating Cost $	**3,737,902**	**3,330,544**	**8,478,317**

TABLE 5: Conventional Systems Running Costs in US$

   SIMPLE PAY BACK
An energy investment simple payback period is the amount of time it will take to recover the initial investment in energy savings, dividing initial installed cost by the annual energy cost savings and is calculated according to the following equation: [10]

Pay Back Period = Difference in initial cost / Saving in Running Cost **(1)**

Difference in initial cost = 81,623,918 $
Saving in Running Cost (1st year) = 5,843,910 $

Pay Back Period =   11 years

NET PRESENT VALUE TECHNIQUE [NPV] [11]
The NPV method determines the worth of a project over time, in today's dollars. Unlike the payback method, NPV also accounts for the savings that occur after the payback period. To calculate the NPV, The following factors were taken into consideration. [10,11]

It is also important to note that if the cold seawater pipeline were to fail, than cooling for the hotel would not be available until the pipeline could be restored. Of course no one would pay top dollar to stay in a hot and humid hotel room, so the loss of the pipeline can result in significant losses in revenue. Emergency portable vapor compression units could be rented while the pipelines are repaired; however, these costs were not included in this study.

$$\textbf{Dep C} = \frac{(1+i)^{n} \times i}{(1+i)^{n} - 1} \; \textbf{' Depreciation coefficient'} \tag{2}$$

Where i is the interest rate and is taken as 0.1

$$\text{DEPRECIATION=INITIAL COST* Dep C} \tag{3}$$

$$\text{Dep K=DEPRECIATION+RUNNING COST} \tag{4}$$
But ;
$$\textbf{Net Present Value=Dep K*}((1+i)^{-k}) \tag{5}$$

Where: n is the number of years [life period] and is taken as 15 years, k is the current year
Total Net Present Value is the accumulated NPV throughout the life period of the project or i.e., it is the sum of NPV of all 30 years.

The results show that although the SWAC system requires high initial costs nearly about seven times that for the CCS, its running cost is 15% of that for the CCS.

In Addition to the less maintenance needed for running this system either than the conventional systems and the thermal energy storage that can be utilized from the effluent cold seawater.

	SWAC	VCS
IC	103,440,318	21,365,500
RC (1st Year)	2,634,407	8,478,317
NPV (30 Years)	143,078,436	193,002,429

TABLE 6: NPV for SWAC and VCS systems

Environmental Analysis
Greenhouse gas emissions from power production will be reduced by the following quantities shown in Table 7. [1]

Pollutant Name	kg/kWh	kWh Savings	Reduction in Tons
CO2	0.714		99,070
CO	0.000365		51
CH4	0.00168	138,753,559	233
NOX	0.00125		173
N2O	0.0000169		2
SOx	0.00379		526
Solid Waste	0.0863		11,974

TABLE 7: Greenhouse Gases Emissions Reductions

Hybrid System

This study was based on the temperature-depth profile shown in Figure 3. There may be unfavorable temperature variations at the intake site; if the SWAC system fails to provide this 7°C seawater to the heat exchanger, the temperature of the chilled fresh water exiting the heat exchanger will rise up and further chilling through an auxiliary chiller should be needed. This may be solved by implementing a Hybrid system in which part of the air conditioning load should be provided by the SWAC system and the rest part by an auxiliary chiller which will be automatically operated by a programmable logic controller (PLC) taking a signal from a temperature sensor.

Figure 6 shows the result NPV for different scenarios of the Hybrid system where the percentages are changed among the system.
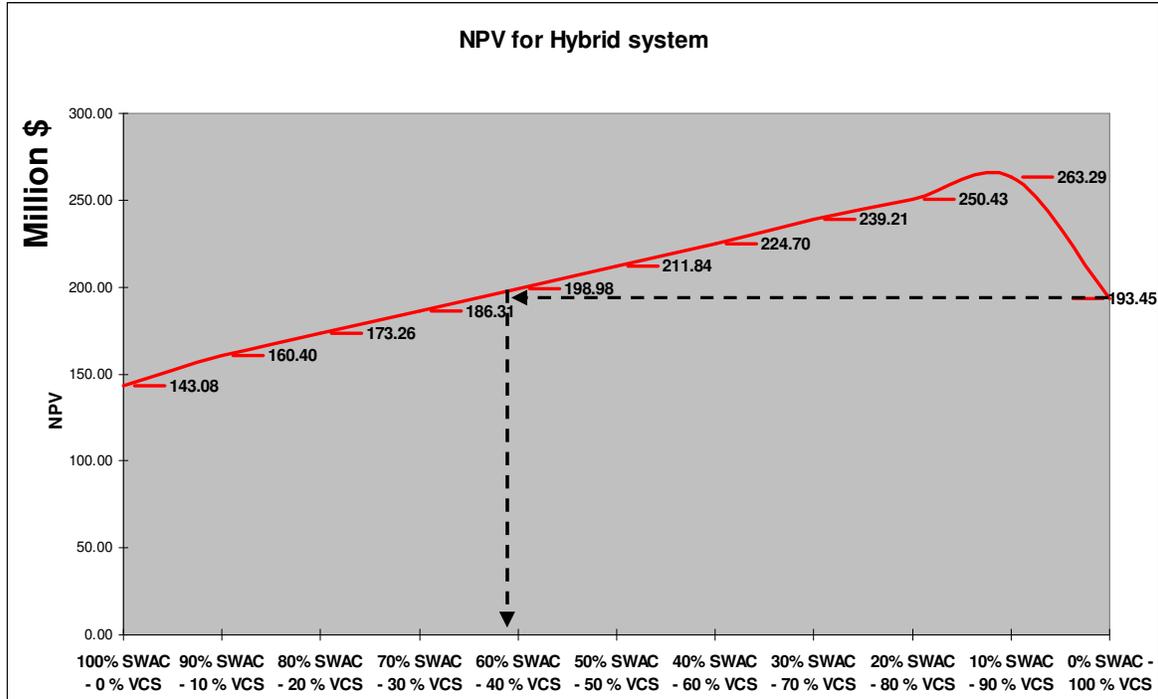

FIGURE 6: Hybrid system for Sahl-Hasheesh

At the end of the economic analysis the impact of the electricity unit rate and the life time of the project were studied.

Economic Sensitivity: Impact of electricity rate and project life time

Electricity rate is a major factor in determining the profitability of a SWAC system. For this study 0.25 Egyptian pounds (0.045$) was assumed as a cost for the kWh based on year 2007 costs. Any variation in the electricity unit cost will result in changes in the results.

Figure 7 illustrates the impact of the electricity rates on the net present values for the SWAC and VCS systems. Thirty years was chosen as a project life time for the air conditioning system at SH and which the author assumed that the life time shall exceeds this value since SH is a very huge project and shall always require air conditioning for the hotels.
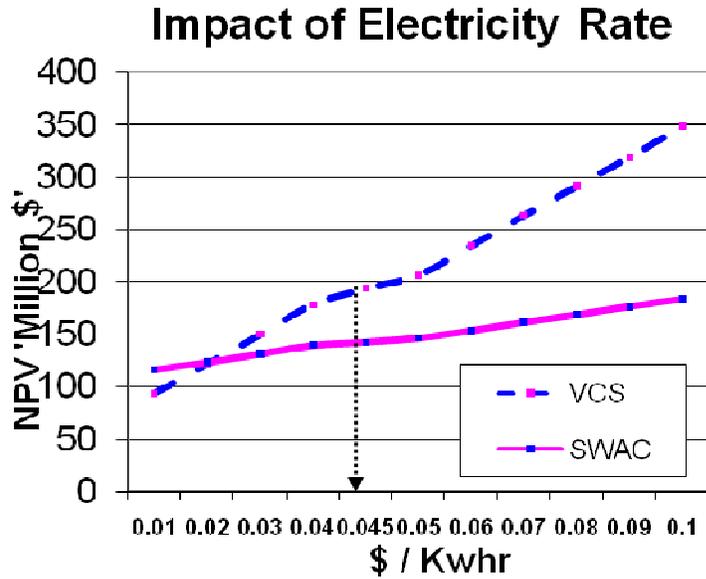
Impact of Electricity Rate

FIGURE 7: Impact of electricity rate on NPV

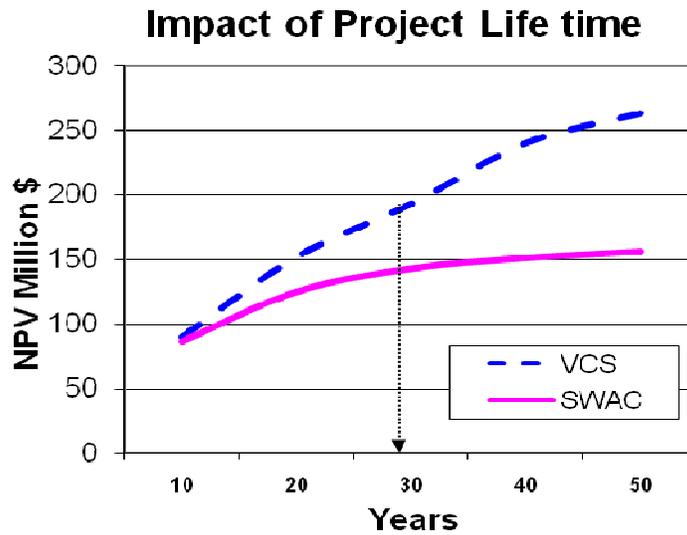Increasing the project life time shall result in more economical SWAC system than VCS as illustrated in Figure 8.

Impact of Project Life time

FIGURE 8: Impact of life time on NPV

4. CONSLUSIONS

Sea water air conditioning is an established technology being applied in an innovative way. The cold sea water air conditioning has merit over conventional vapor compression air conditioning systems. This merit is for hotels located in regions of

the world where access to cold seawater is at a minimum and there is year-round high humidity.

This study dealt with the design and economical investigation for the Sea Water Air Conditioning (SWAC) System for a very high cooling load area [Sahl-Hashish, Egypt]. It is concluded that:

- SWAC is technically feasible in Sahl-Hasheesh. On the other hand, cold sea water air conditioning is not considered economically feasible for tropical cooling loads less than 5000 TOR.
- The major challenge is crossing the Marine Preserve.
- If water below 8°C is required, the flattening bathymetry suggests that auxiliary chillers would be more cost-effective.
- Bathymetry and site specific temperatures need to be collected; differences from the values assumed could introduce unexpected challenges.
- Greater Independence from Energy Price Escalation - In a world of rapidly increasing energy prices, SWAC costs (which are capital dominated) are relatively flat compared to that of energy intensive conventional AC systems. Users will have a known and relatively flat future AC cost.
- Short economic payback period.
- Reduction of electricity use. The SWAC system reduces the annual electric energy usage with 75% compared to on-site chillers. At a peak demand of 26,500 tons of refrigeration (TOR) in SH and 60% as a utilization factor the SWAC system will reduce the electric energy usage by 138,745 MWh per year.
- Reduction of air pollution due to greenhouse gases emissions reduction.

5. REFERENCES

1. Marland, G., T.A. Boden, and R. J. Andres, 2001, "Global, Regional, and National CO2 Emissions." In Trends: A Compendium of Data on Global Change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

2. EIA, April 2008; 'Energy information administration, official energy statistics from the US government. (http://tonto.eia.doe.gov/country/country_energy_data.cfm?fips=EG)

3. McQuiston, F.C., J.D. Parker, and J.D. Spitler.2005. "Heating, Ventilating, and Air Conditioning Analysis and Design", "HVAC Load Explorer", Software Guide, Sixth Edition. New York: John Wiley and Sons.

4. ASHRAE, 1997. ASHRAE handbook-Fundamentals, Atlanta: American Society of Heating Refrigeration and Air conditioning Engineers, Inc.

5. ASHRAE, 1999. ASHRAE handbook-HVAC Applications, Atlanta: American Society of Heating Refrigeration and Air conditioning Engineers, Inc.

6. Elkader Hany, 2007, Pyramisa hotel maintenance engineer, Sahl-Hasheesh, Personal communications.

7. Abdelmoneem Y., 2006; Hydrographic - Navy forces, Personal Communication

Elsafty, A. F. & Saeid, L.

8. Haaland, S.E. March 1983, "Simple and explicit formulas for the friction factor in turbulent pipe flow," Fluids Eng., as referenced in White, Frank. Fluid Mechanics, 3rd ed. McGraw Hill. New York, 1994, p317

9. MOE, 1994. Makai Ocean Engineering, Inc.. "Sea Water Air Conditioning for Hawaii Phase 1: West Beach, Oahu", The State of Hawaii, Department of Business, Economic Development and Tourism, Energy Division, Honolulu

10. DeGarmo E. P., Sullivan W. G. and Bontadelli J. A., 1993; "Engineering Economy,"9th ed., MacMillan Publishing Company, New York

11. Harwell, Didcot and Oxford shire, 1999, "Good practice guide – An introduction to absorption cooling".

Development of a Bench-Top Air-to-Water Heat Pump Experimental Apparatus

H. I. Abu-Mulaweh   mulaweh@engr.ipfw.edu
Mechanical Engineering Department
Purdue University at Fort Wayne
Fort Wayne, IN 46805, USA

Abstract

A bench-top air-to-water heat pump experimental apparatus was designed, developed, and constructed for instructional and demonstrative purposes. This air-to-water heat pump experimental apparatus is capable of demonstrating thermodynamics and heat transfer concepts and principles. This heat pump experimental setup was designed around the vapor compression refrigeration cycle. This experimental apparatus has an intuitive user interface, reliable, safe for student use, and portable. The interface is capable of allowing data acquisition by a computer. A PC-based control system which consists of LabVIEW and data acquisition unit is employed to monitor and control this experimental laboratory apparatus. This paper provides details about the development of this unit and the integration of the electrical/electronic component and the control system.

Keywords: Heat pump, laboratory apparatus. water heater

1. INTRODUCTION

One of the important applications of the subjects of thermodynamics and heat transfer is heat pump systems. Exposing thermal engineering students to heat pump will enhance their understanding of the principles and concepts of thermodynamics and heat transfer.

There are many are many heating, ventilating, and air-conditioning (HVAC) systems, but very few of them are appropriate for undergraduate education [1, 2]. Some of these systems have computer data logging equipment. However, the computer data logging applications are exclusively standalone and not compatible with PC based data processing. In addition, these systems tend to be rather large and expensive. Recently, Abu-Mulaweh [3] has designed, developed, and constructed a bench-top air-conditioning experimental apparatus for instructional and demonstrative purposes.

Acquiring new instructional laboratory apparatus is a challenge due to typical budgetary limitations. In addition, the apparatus designed by companies specializing in education equipment may not exactly reflect the educational objective intended by the faculty. These obstacles had forced us to seek and search different venues to acquire "high tech" experimental laboratory apparatus for demonstrating heating and refrigeration processes. It was decided to develop and build a cost effective system that can be employed to demonstrate and monitor refrigeration cycle, as well as some fundamental concepts in heat transfer, thermodynamics, and heat exchangers.

Hot water heaters come in various sizes and either gas fired or electric. Using a heat pump water heater to supply hot water for residential and commercial usage, is a much more efficient and energy conservative method. Heat pump water heaters can be either water source or air source. The heat pump is an electrically powered mechanical device that transfers heat from a lower-temperature source to a higher-temperature body, such as an air conditioner. The feasibility and the effectiveness of heat pump water heaters have been examined in the past [4-7].

The coefficient of performance (COP) of heat pump water heaters is typically in the order of three. This implies that the energy consumption can theoretically be reduced by two-thirds when resistance heaters are replaced by heat pumps. The installed electrical capacity is also reduced by almost two-thirds due the COP. To the building owner, this would mean a reduction in both the direct cost of units of energy consumed and the monthly peak demand charges.

The replacement of resistance heaters with heat pumps water heaters will also result in a reduction in the peak electrical demand imposed on the national electricity supply grid. To the utility, this could mean a reduction in the marginal cost of supplying each new unit of power since the need to build new power stations may be deferred.

2. SYSTEM SPECIFICATIONS

The design process that the students follow in the capstone senior design project is the one outlined by Bejan et al. [8] and Jaluria [9]. The first essential and basic feature of this process is the formulation of the problem statement. The formulation of the design problem statement involves determining the requirements of the system, the given parameters, the design variables, any limitations or constraints, and any additional considerations arising from safety, financial, environmental, or other concerns.

In order for the bench-top air-to-water heat pump to function as a useful piece of lab equipment, the following requirements and specifications need to be met. These include requirements that will make the heat pump useful for demonstrating thermal science and fluid dynamics principles as well as ensure the heat pump will operate safely.

- Construction – The air-to-water heat pump is to be designed to operate based on a vapor-compression cycle.

- Instrumentation – The instrumentation requirements have two distinct sets of necessary specifications.

  1. The heat pump must be fully instrumented with autonomous gages so that it may demonstrate its principles without needing to be hooked up to an outside computer.
  2. Although it may operate without an external computer, the heat pump must also be outfitted with a data acquisition (DAQ) bus that can be connected to an external DAQ board or software system. This bus must be able to supply to the external DAQ system the measurements that will be shown on the onboard instrumentation. In addition, the measurements must be logged by the DAQ.

- Safety – The safety considerations deal primarily with the fact that the design requires both large amounts of electrical equipment and liquids to be in close proximity. For this reason the following are required of the electrical design scheme:

  1. Residual current circuit breaker.
  2. Combined double pole main switch and overload cut-out.
  3. All components connected to common earth conductor.

3. EXPERIMENTAL APARATUS

A bench-top air-to-water heat pump, shown in Figure 1, was designed, developed, and constructed for instructional and demonstrative purposes. This heat pump was designed around the vapor compression refrigeration cycle. The bench-top air-to-water heat pump has an intuitive user interface, reliable, safe for student use, and portable. The interface is capable of allowing data acquisition by an existing laboratory computer. The unit is capable of warming the water 10ºC in an open or closed configuration, and cool it back down if desired. The system fully controls and monitors the fluid properties at key points in the refrigerant and water loops.

FIGURE 1: Experimental apparatus

3.1 Mechanical Concepts Selection

To select properly sized components for the heat pump, parametric studies were completed using the EES model and plots of the important operating variables were generated. Because the entire heat pump apparatus needs to operate off of the power provided by a single 115V, 15A wall circuit, the amount of steady-state power required by the compressor was limited to 800 W (2732.14 BTU/hr). This power constraint is imposed in order to leave sufficient reserve power for start-up energy requirements as well as sufficient power to run the rest of the powered devices that will be used on the heat pump. For the heat pump to meet its water heating performance requirements the water flowing through the condenser must experience at least a 10 °C (18 °F) temperature rise between the inlet and exit.

3.2 Refrigeration Cycle Components

Condensing Unit: A Tecumseh model AEA4440YXAXW water cooled condensing unit was chosen as the basis for the heat pump. At the nominal operating conditions anticipated, this

condensing unit will deliver 0.53 GPM (3.34E-5 m³/s) of water raised 13.5°C (24.3°F) when paired with a 4600 BTU/hr (1346.93 W) evaporator.

Evaporator & Fan: Knowing the requirements of the evaporator in terms of heat, a model 012-0850 evaporator was chosen from www.Rparts.com. This evaporator is rated for 7500 BTU/hr (2196.08 W) at an evaporating temperature difference of 15°C (27°F) and airflow of 200 CFM (0.09 m³/s). The evaporating temperature difference is defined as the difference in temperature between the air entering the evaporator and the saturation temperature of the refrigerant within the evaporator. The evaporator is oversized to make sure that the compressor is never starved for refrigerant and oil, which is emulsified in the refrigerant. Excess evaporator heat can be reduced and adjusted for by decreasing the air flow across the coils or by blocking of a portion of the heat exchange surface with dampers. The selected fan for the evaporator is the Falcon model from Lytron. It is capable of providing up to 300 CFM (0.14 m³/s) at 0.36 amps of running current.

Expansion Valve: The selection of the expansion valve, located right before the evaporator, is based on the capacity of the system. The target capacity of the evaporator is approximately 4600 BTU/hr (1346.93 W) which corresponds to 0.383 tons of refrigeration. Based on this value, a Danfoss expansion valve model 095-0200 (orifice part No. 095-0003) was selected. This model is externally equalized and the orifice is rated for up to 0.5 tons (6000 BTU/hr) of refrigeration.

Filter Dryer: To remove any moisture or contaminants that could potentially damage the system, a filter dryer from Sporlan was chosen (# 020-0052, Mod.C-082). This model will also neutralize any acid that might initially form from the interaction between the refrigerant and coils.

Sight Glass: A Sporlan sight glass (#077-0100, Mod. SA12) was chosen to provide a visual to warn of and diagnose excessively high moisture levels.

Tubing: Copper tubing of 3/8 inch (0.95 cm) inner diameter was selected for the refrigerant.

3.3   Water Cycle Components
Water Pump: The pumping force for the circulation of the water loop will come from a Delphi – Laing DC Pump, (Mod. DDC-1TPMP). This pump can supply water at a maximum pressure of 22 psi (151.68 kPa), and a maximum flow of 1.75 GPM (1.10E-4 m³/s) at 13.2 VDC. It operates up to a temperature of 60ºC (140ºF).

Water Reservoir and Tubing: An insulated water cooler was used as a water reservoir. Bulkhead fittings were mounted on the cooler for pipes connections. The water piping was constructed from copper tubing.

Radiator & Fans: To cool the water back down, an appropriate radiator was selected to fit the requirements of this cycle. The radiator needs to move 6500 BTU/hr (1903.27 W) to cool the water down close to room temperature from 113ºF (45ºC), with a flow rate of 1 GPM (6.31E-5 m³/s). The component is made by Lytron (Mod. M14-240). It has capacity of 373 BTU/hr ºF at the specified flow rate, with an air flow of 550 CFM (0.26 m³/s). The radiator is equipped with two fans with a capability of 450 CFM (0.21 m³/s) each, at a running current of 0.5 amps per fan.

4. INTEGRATION OF THE ELECTRICAL/ELECTRONIC COMPONENT
All of the low voltage electrical components were tested for functionality before assembly into their respective circuits. This was done by applying voltages to each chip and checking output for correctness. After these individual components were tested, each of the circuits were temporarily built on a breadboard and tested to verify that they worked as expected. This was proven by applying inputs of voltages or frequencies from lab equipment and measuring the output voltages with a handheld meter. The microprocessors were also programmed and tested.

In the second phase of the building process, the final assembly of the low voltage electrical system was performed. All of the circuits were laid out, built, and soldered together on a perforated proto-board. The circuits were then retested with lab equipment to verify that building and soldering was completed correctly. Once the circuits were built and soldered, more testing was done to determine the accuracy and percentage error associated with these circuits.

The fourth phase comprised the high voltage wiring of the mechanical system. All high voltage components were wired with a common ground, the ground fault interrupter (GFI) was placed on the high voltage wire coming directly from the wall outlet to protect against large surges of current, and the double pole single throw (DPST) circuit breaker was placed directly following the GFI. In addition to these original specifications, two lights were added to show the status of the high voltage power. The first light demonstrates that the system is connected to the high voltage power and the second light confirms the system is on and operational.

The final integration of the electrical/electronic system with the mechanical system consisted of running wires from the pressure circuits to the pressure transducers, connecting the thermocouple wires to the temperature circuits, and applying low voltage power to the water pump. The final step in the electrical/electronic building process was testing the DAQ functionality. Wires were run from the control box on the heat pump unit connecting to the DAQ board via a printer cable.

After the entire system was built and operational it was then connected to the SCB-68 break out box and then to the DAQ card, via the on-board DB-25 connector and modified printer cable. It was observed that the sensor signals dropped due to the voltage divider circuit created when the DAQ was applied to the system. To remedy this problem, fifteen LM741 operational amplifiers were employed as a non inverting amplifier and voltage followers. The non-inverting amplifier was used to amplify (magnitude of ten) the water flow signal, as it was observed that the DAQ was having difficulty accurately collecting the small signal. This signal isolation system was added to the cable that connects the control box to the DAQ.

5. INTERFACE AND CONTROL SYSTEM

The control requirements of this system consist of monitoring two of the pressures (the condenser pressure and the evaporator pressure) and switching the compressor and the evaporator fan on and off, depending on the levels of the pressure transducers. In order to control this system there are a few integral components; 1.) A PIC16F687, the microprocessor that will read run the control program. 2.) Two Solid State Relays (P/N 4062RL), used for switching on and off the motors. 3.) Two pressure transducer outputs (P/N 2CP5-50-1). The microcontroller runs the program that is flowcharted in Figure 2. This program utilizes the onboard 10-bit analog to digital converters to convert the pressure transducers' voltages into a usable digital value.   The microprocessor compares the values that are produced from the pressure transducers to some preset values and then output control signals to the solid state relays to control the motors.

The microcontroller is programmed using the flow chart shown in Figure 2. This program starts both of the relays at zero output logic and initially checks the condenser pressure. If the condenser pressure is above the set point, both of the motors will remain at zero. The relays will remain at zero until the condenser pressure is below another preset value. Once this pressure drops to the set point the microprocessor will check the evaporator pressure, if this pressure is below a certain value the compressor will remain at zero and the evaporator pressure will continue to be polled until it rises to the specified suitable temperature. Once this temperature is achieved the compressor will be turned on and the evaporator pressure will be checked. If this pressure becomes too large the evaporator fans will remain at zero and the system will poll the evaporator pressure until it falls below the set point.   Once the set point is achieved the

evaporator fan will be turned on and the loop will continue to iterate at the check condenser pressure stage until the system is turned off.
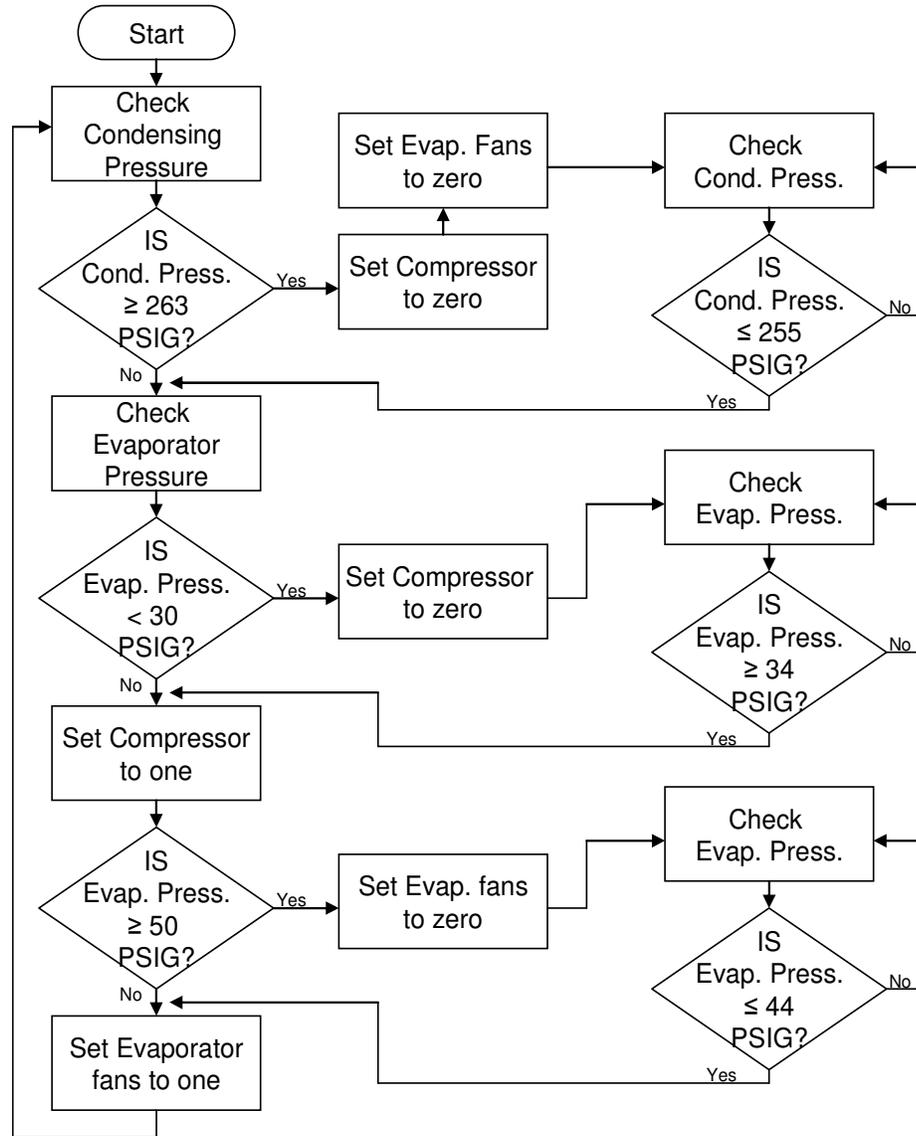

FIGURE 2: Flow Chart for Controller Program

6. PERFORMANCE OF THE APPARATUS

The heat pump experimental apparatus was designed to operate as an open water loop (external water supply) and as a closed water loop (onboard supply). Figure 3 the sensors locations for both the refrigerant and the water loops. The measurements are obtained from either the onboard instrument display or from a DQA.
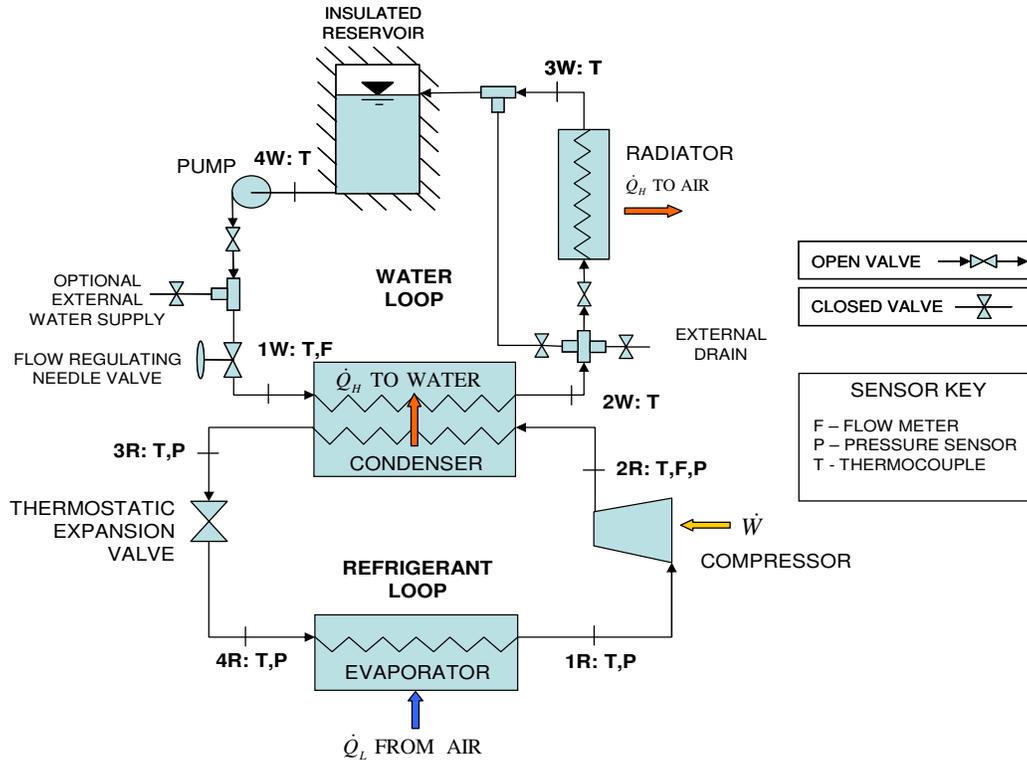
FIGURE 3: Sensor locations

6.1 Measurements

In each run the following measurements can be obtained:
1. Refrigerant Temperature – The temperatures at each of the four points in the vapor compression cycle will be measured using K-type thermocouples.
2. Refrigerant Pressure – The pressure at each of the four points in the vapor compression cycle will be measured using pressure transducers.
3. Refrigerant Flow Rate – The volume flow rate of the refrigerant will be measured using a paddle-wheel flow sensor.
4. Water Temperature – The temperature of the water will be measured at the inlet and outlet of the condenser and in the reservoir using K-type thermocouples.
5. Water Flow Rate – The volume flow rate of the water will be measured using a turbine-type flow sensor.
6. Ambient Temperature – Ambient temperature will be measured using a K-type thermocouple as well.

6.2 Calculations

From the measured data the following calculations can be performed:

1. Mass flow rate of water

$$\dot{m}_w = \dot{V} \cdot \rho.$$

2. Heat transferred in the condenser

$$\dot{Q}_H = \dot{m}_w (h_{2w} - h_{1w})$$

3. Using the heat transfer to the water and points R2 and R3 in the refrigerant loop (see Fig.3), the mass flow of the refrigerant can be determined.

$$\dot{m}_R = \frac{\dot{Q}_H}{h_{2R} - h_{3R}}$$

4. Heat transferred in the evaporator

$$\dot{Q}_L = \dot{m}_R (h_{1R} - h_{4R})$$

5. Calculate the work done by the compressor.

$$\dot{W}_{compressor} = \dot{m}_R (h_{2R} - h_{1R})$$

6. Isentropic efficiency of the compressor.

$$\eta_{comp} = \frac{h_{2sR} - h_{1R}}{h_{2R} - h_{1R}}$$

7. Determine the coefficient of performance.

$$CoP = \frac{\dot{Q}_H}{\dot{W}_{comp}}$$

6.3 Sample Results
The heat pump can be set up in four different configurations:
   1. open loop
   2. closed loop
   3. closed system
   4. system cooling

Open Loop: The open loop configuration makes use of water from a tap. Hoses are provided with the unit. The water comes in to the unit, travels through the heat exchanger and exits the heat pump. Water is expelled into a drain.

Closed Loop: The closed loop configuration does not require an outside water source. In this configuration water travels from the reservoir through the water loop and back into the reservoir. Heat added in the condenser is dissipated by the radiator.

Closed System: This configuration bypasses the radiator and allows the heat gained in from the condenser to be stored in the reservoir. The closed system does not require an outside water source.

System Cooling: Once the reservoir has been heated using the Closed System, it can be cooled down. This can be accomplished by using one of the hoses provided and turning off the compressor while letting the pump and radiator run.

Steady-state measurements, closed loop configuration are included in Table 1. To insure repeatability, three different runs were carried out. In each run measurements of the water temperatures, water flow rate, refrigerant temperatures and pressures were obtained (refer to Fig. 3 for sensor locations).

Figure 4 and 5 show pressure-enthalpy and temperature-entropy plots of the system, respectively. These plots indicate the vapor-compression cycle is operating correctly, with slight amounts of superheat and sub-cooling. Dotted lines in the P-h plot demonstrate the deviations from ideal refrigeration cycles. The irreversible process of the expansion through the valve and heat transfer to and from components account for this phenomena. Although in ideal situations, the expansion between R3 and R4 (see Figures 3) is a constant enthalpy process, the increase in enthalpy between these two states only amounts to approximately 12%. In the T-s diagram (Figure 4) these irreversibilities are depicted by dotted lines. In a standard cycle, the compression between R1 and R2 (see Figures 3) is an isentropic process. In this case, there is a 3% increase in entropy from R1 to R2 due to the irreversible nature of this process.

Location	Run 1	Run 2	Run 3
Ambient Temp. (°C)	23.6	23.8	24.1
W1 (°C)	28.2	28.2	28.9
W2 (°C)	37.9	39.0	38.2
W3 (°C)	27.5	27.8	28.3
W4 (°C)	27.8	27.8	28.5
Water Flow (gpm)	0.60	0.53	0.62
R1 (°C)	10.9	11.1	11.4
R2 (°C)	82.5	81.3	81.2
R3 (°C)	38.3	39.3	38.8
R4 (°C)	5.7	6.1	6.3
Refrigerant Flow (gpm)	0.101	0.101	0.101
P1 (psia)	59.1	59.0	59.9
P2 (psia)	170.8	176.5	174.1
P3 (psia)	171.0	175.3	172.8
P4 (psia)	57.6	58.2	57.8

TABLE 1: Steady-state measurements, closed loop configuration.
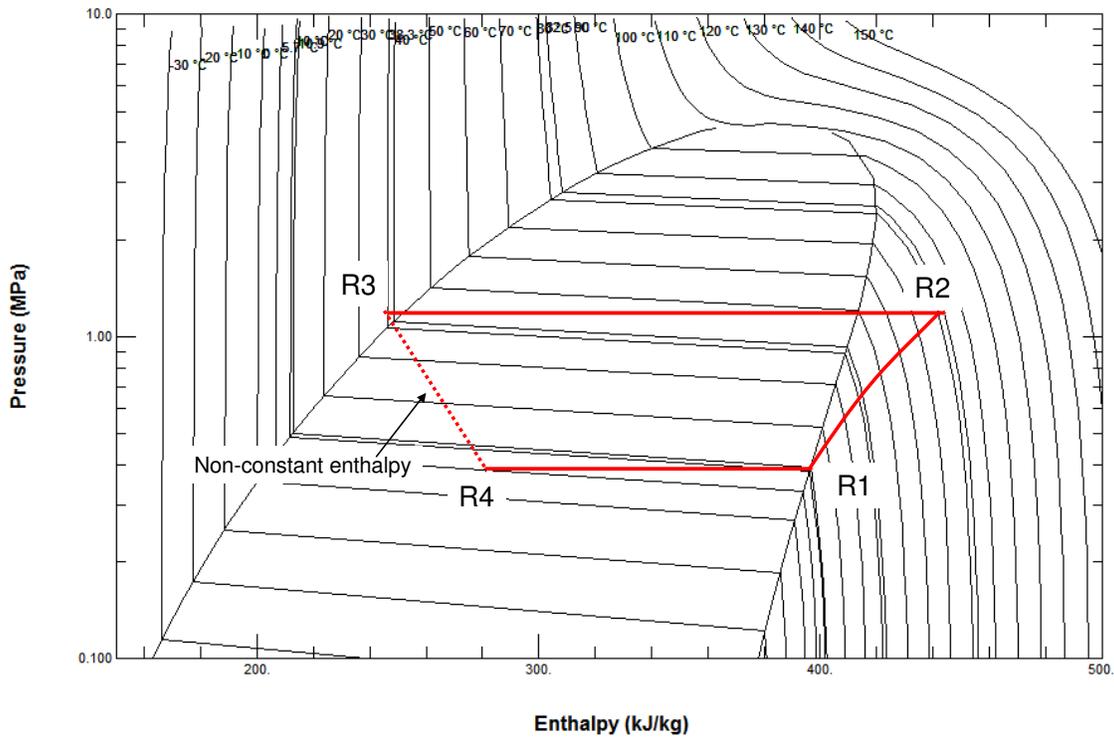

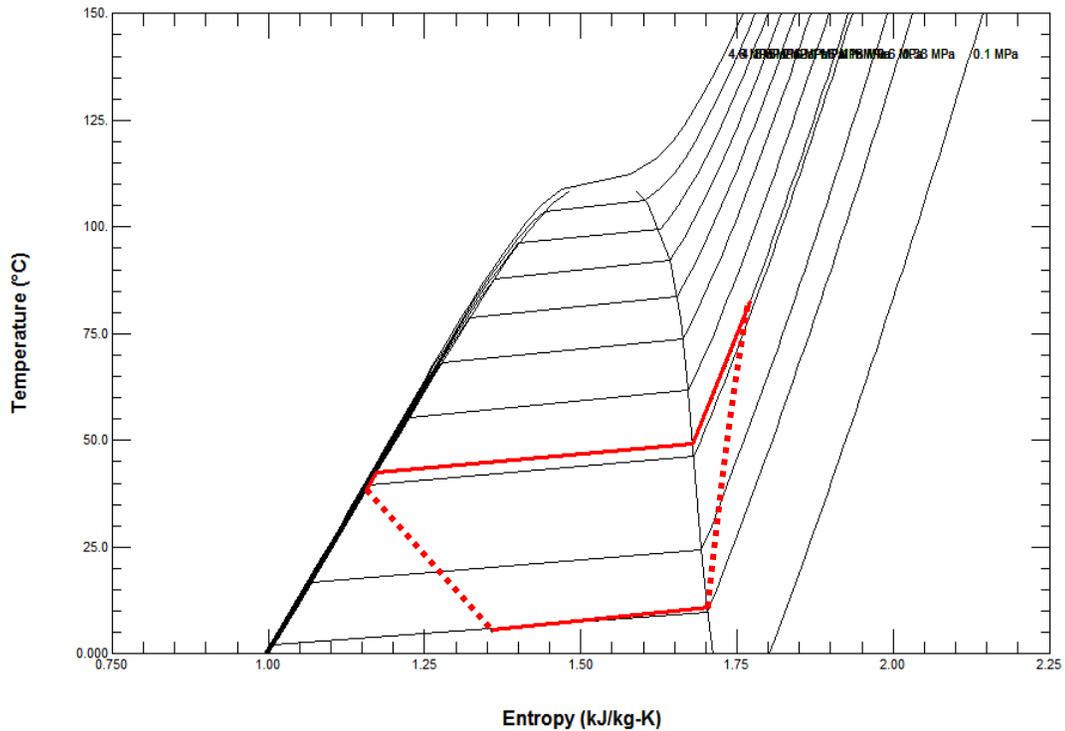**FIGURE 4:** Operating points of the refrigeration system shown on a pressure-enthalpy chart

FIGURE 5: Operating points of the refrigeration system shown on a temperature-entropy chart

7. CONSLUSION

A portable air-to-water heat pump experimental apparatus was developed and built to demonstrate actual refrigeration cycle, as well as some fundamental concepts in heat transfer, thermodynamics, and heat exchangers. The system is fully instrumented and numerous aspects of the refrigeration cycle can be easily monitored. This experimental apparatus has an intuitive user interface, reliable, safe for student use. The interface is capable of allowing data acquisition by a computer. The unit is also instrumented with autonomous gages to be able to demonstrate heat transfer concepts/principles and thermodynamics processes without needing to be hooked up to a computer.

8. ACKNOWLEDGEMENT

9. REFERENCES

1. Hamden Engineering Corporation, Mechanical Engineering Program, East Longmeadow, MA, USA, 2001.
2. Cuissons Educational Technology, Thermodynamics, TRA/2, distributed by Feedback Inc., Hillsborough, NC, USA,1997.

3. H.I. Abu-Mulaweh, "*Development and Performance Validation of Portable Air-Conditioning System Experimental Apparatus*", In Proceedings of The Third International Conference on Thermal Engineering: Theory and Applications, Amman, Jordan, May 21-23, 2007.
4. B.A. Pritchard, W. A. Beckman, and J.W. Mitchell, "*Heat Pump Water Heaters for Restaurant Applications*". International Journal of Ambient Energy, 12: 59-68, 1991.
5. P. Neksa, H. Rekstad, G.R. Zakeri, and P.A. Schiefloe, "*CO_2-Heat Pump Water Heater: Characteristics, System Design and Experimental Results*". International Journal of Refrigeration, 21: 172-179, 1998.
6. B.J. Huang and J.P. Chyng, "*Integral-Type Solar-Assisted Heat Pump Water Heater*". Renewable Energy, 16: 731-734, 1999.
7. C.W. Wood, "*Unsteady Temperature Distributions in Vertical Storage Tanks Heated with Heat pumps*". Heat Transfer Engineering, 19: 43-52, 1998.
8. A. Bejan, G. Tsatsaronis, and M. Moran, "*Thermal Design & Optimization*", *John Wiley & Sons, Inc.*, New York, 1996.
9. Y. Jaluria, "*Design and Optimization of Thermal Systems*", *McGraw-Hill*, New York, 1998.