

# **COST**

Domain Committee "Information and Communication  
Technologies (ICT)"

## **COST Action IC0702**

**Start Date** 01/04/2008

*Combining Soft Computing Techniques and  
Statistical Methods to Improve Data Analysis  
Solutions*

## **MONITORING PROGRESS REPORT**

**Reporting Period:** *from 01/04/2008  
to 28/02/2009*

This Report is presented to the relevant Domain Committee.  
It contains three parts:

- I. Management Report*** prepared by the COST Office/Grant Holder
- II. Scientific Report*** prepared by the Chair of the Management Committee of the Action
- III. Previous versions of the Scientific Report; i.e., part II of past reporting periods***

The report is a "cumulative" report, i.e. it is updated annually and covers the entire period of the Action.

Confidentiality: the documents will be made available to the public via the COST Action web page except for chapter *II.D. Self evaluation*.

Based on the monitoring results, the COST Office will decide on the following year's budget allocation.

**Executive summary (max.250 words):**

The main objective of the COST action IC0702 "Combining Soft Computing Techniques and Statistical Methods to Improve Data Analysis Solutions" is to strengthen the dialogue between the statistics and soft computing research communities in order to cross-pollinate both fields and generate mutual improvement activities. Soft computing, as an engineering science, and statistics, as a branch of mathematics, emphasize different aspects of data analysis. Soft computing focuses on obtaining working solutions quickly, accepting approximations and unconventional approaches. Its strength lies in its flexibility to create models that suit the needs arising in applications (context of discovery, model generation). In addition, it emphasizes the need for intuitive and interpretable models, which are tolerant to imprecision and uncertainty. Statistics is more rigorous and focuses on establishing objective conclusions based on experimental data by analysing the possible situations and their (relative) likelihood (context of justification, model validation). It emphasizes the need for mathematical methods and tools to assess solutions and guarantee performance. Bringing the two fields closer together will enhance the robustness and generalisability of data analysis methods, while preserving the flexibility to solve real-world problems efficiently and intuitively.

**I. Management Report** prepared by the Grant Holder (European Centre for Soft Computing – Spain)



**I.A. COST Action Fact Sheet**

• **COST Action** IC0702 - *Combining Soft Computing Techniques and Statistical Methods to Improve Data Analysis Solutions*

• **Domain** *Information and Communication Technologies (ICT)*

• **Action details:**

**CSO Approval:** 16/11/2007

**End date:** 30/03/2012

**Entry into force:** 10/01/2008

**Extension:** -

• **Objectives**

The main objective of the Action is to strengthen the dialogue between the statistics and soft computing research communities in order to cross-pollinate both fields, generate mutual improvement activities, spark new ideas, and trigger new research lines and projects.

• **Parties:**

Austria 11/04/2008	Netherlands 07/02/2008
Belgium 10/01/2008	Norway 17/03/2008
Cyprus 22/10/2008	Poland 10/01/2008
Estonia 14/05/2008	Portugal 10/04/2008
France 17/01/2008	Spain 10/01/2008
Germany 10/01/2008	Switzerland 21/08/2008
Hungary 22/02/2008	Turkey 03/09/2008
Italy 12/03/2008	United Kingdom 10/01/2008
Lithuania 08/08/2008	

• **Intentions to accept:** None.

• **Other participants:** None

**Chair:** *Dr. Christian BORGELT,*  
*European Centre for Soft Computing,*  
*Calle Gonzalo Gutiérrez Quirós s/n -*  
*33600 Mieres – Spain,*  
[christian.borgelt@softcomputing.es](mailto:christian.borgelt@softcomputing.es)

**DC Rapporteur:** *Professor Leon VAN*  
*DER TORRE, Department of Computer*  
*Science, University of Luxembourg, 6*  
*rue Richard Coudenhove - Luxembourg*  
[leon.vandertorre@uni.lu](mailto:leon.vandertorre@uni.lu)

**Science Officer:** *Mr. Gian Mario*  
*MAGGIO, [gmmaggio@cost.esf.org](mailto:gmmaggio@cost.esf.org)*

**Administrative Officer:** *Ms. Sophie*  
*BEAUBRON, [sbeaubron@cost.esf.org](mailto:sbeaubron@cost.esf.org)*

- **Action Web site:** <http://www.cost-ic0702.org/> **Grant Holder Representative**  
*Dr. Raul del Coso, [raul.delcoso@softcomputing.es](mailto:raul.delcoso@softcomputing.es)*

- **Working Groups**

- **Working Group A – Model selection and validation**

- Michel Verleysen (leader), Ingunn Berget, Christian Borgelt, Christakis Charalambous, Ana Colubi, Thierry Denoeux, Cristian Gatu, Gil González Rodriguez, Mario Guarrancino, Martin Høy, Raimundas Jasinevicius, Uzay Kaymak, Frank Klawonn, Erricos Kontoghiorghes, Rudolf Kruse, Sven Laur, Irini Moustaki, Tormod Næs, Wolfgang Näther, Gabriella Pasi, Bernd Reusch, Matthias Steinbrecher, Konstantin Tretjakov, Wolfgang Trutschnig, Bülent Tütmez, Jaak Vilo

- **Working Group B – Bio-inspired metaheuristics**

- João Miguel da Costa Sousa (leader), Bernard de Baets, Jürgen Branke, Christakis Charalambous, Cristian Gatu, Raimundas Jasinevicius, Janusz Kacprzyk, Uzay Kaymak, Frank Klawonn, Erricos Kontoghiorghes, Bernd Reusch, Özlem Türksen, Bülent Tütmez

- **Working Group C – Statistics with imperfect and incomplete data**

- Maria Angeles Gil (leader), Stefan van Aelst, Bernard de Baets, Ingunn Berget, Christian Borgelt, Simone Borra, Ricardo Cao, Ana Colubi, Etienne Come, Renato Coppi, Fabio Crestani, Thierry Denoeux, Didier Dubois, Janos Fodor, Manfred Gilli, Gil Gonzalez Rodriguez, Mario Guarracino, Martin Høy, Janusz Kacprzyk, Uzay Kaymak, Rudolf Kruse, Sven Laur, Trevor Martin, Irini Moustaki, Tormod Næs, Wolfgang Näther, Gabriella Pasi, João Miguel da Costa Sousa, Matthias Steinbrecher, Konstantin Tretjakov, Wolfgang Trutschnig, Rosanna Verde, Michel Verleysen, Reinhard Viertl, Jaak Vilo, Peter Winker

### ***I.B. Management Committee member list***

<i>Name</i>	<i>Country</i>	<i>E-mail</i>
Prof. Reinhard Viertl	Austria	<a href="mailto:r.viertl@tuwien.ac.at">r.viertl@tuwien.ac.at</a>
Prof. Peter Filzmoser	Austria	<a href="mailto:P.Filzmoser@tuwien.ac.at">P.Filzmoser@tuwien.ac.at</a>
Prof. Bernard de Baets	Belgium	<a href="mailto:bernard.debaets@ugent.be">bernard.debaets@ugent.be</a>
Prof. Michel Verleysen	Belgium	<a href="mailto:michel.verleysen@uclouvain.be">michel.verleysen@uclouvain.be</a>
Prof. Erricos Kontoghiorghes	Cyprus	<a href="mailto:erricos@dcs.bbk.ac.uk">erricos@dcs.bbk.ac.uk</a>
Prof. Christakis Charalambous	Cyprus	<a href="mailto:bachris@ucy.ac.cy">bachris@ucy.ac.cy</a>
Prof. Jaak Vilo	Estonia	<a href="mailto:aak.vilo@ut.ee">aak.vilo@ut.ee</a>
Prof. Thierry Denoeux	France	<a href="mailto:tdenoeux@hds.utc.fr">tdenoeux@hds.utc.fr</a>
Prof. Didier Dubois	France	<a href="mailto:dubois@irit.fr">dubois@irit.fr</a>
Prof. Frank Klawonn	Germany	<a href="mailto:f.klawonn@fh-wolfenbuettel.de">f.klawonn@fh-wolfenbuettel.de</a>
Prof. Rudolf Kruse	Germany	<a href="mailto:kruse@iws.cs.uni-magdeburg.de">kruse@iws.cs.uni-magdeburg.de</a>
Prof. János Fodor	Hungary	<a href="mailto:fodor@bmf.hu">fodor@bmf.hu</a>
Dr. Mario Guarracino	Italy	<a href="mailto:mario.guarracino@na.icar.cnr.it">mario.guarracino@na.icar.cnr.it</a>
Prof. Renato Coppi	Italy	<a href="mailto:renato.coppi@uniroma1.it">renato.coppi@uniroma1.it</a>
Prof. Raimundas Jasinevicius	Lithuania	<a href="mailto:raimund@ifko.ktu.lt">raimund@ifko.ktu.lt</a>
Dr. Olga Kurasova	Lithuania	<a href="mailto:kurasova@ktl.mii.lt">kurasova@ktl.mii.lt</a>
Prof. Uzay Kaymak	Netherlands	<a href="mailto:kaymak@few.eur.nl">kaymak@few.eur.nl</a>
Prof. Tormod Næs	Norway	<a href="mailto:tormod.naes@matforsk.no">tormod.naes@matforsk.no</a>
Prof. Janusz Kacprzyk	Poland	<a href="mailto:kacprzyk@ibspan.waw.pl">kacprzyk@ibspan.waw.pl</a>
Prof. Olgierd Hryniewicz	Poland	<a href="mailto:Olgierd.Hryniewicz@ibspan.waw.pl">Olgierd.Hryniewicz@ibspan.waw.pl</a>
Prof. João M. da Costa Sousa	Portugal	<a href="mailto:jmsousa@ist.utl.pt">jmsousa@ist.utl.pt</a>
Dr. Christian Borgelt	Spain	<a href="mailto:christian.borgelt@softcomputing.es">christian.borgelt@softcomputing.es</a>
Prof. Maria Angeles Gil	Spain	<a href="mailto:magil@uniovi.es">magil@uniovi.es</a>
Dr. Raul del Coso	Spain	<a href="mailto:raul.delcoso@softcomputing.es">raul.delcoso@softcomputing.es</a>
Prof. Fabio Crestani	Switzerland	<a href="mailto:fabio.crestani@lu.unisi.ch">fabio.crestani@lu.unisi.ch</a>
Prof. Manfred Gilli	Switzerland	<a href="mailto:Manfred.Gilli@unige.ch">Manfred.Gilli@unige.ch</a>
Prof. Bülent Tütmez	Turkey	<a href="mailto:btutmez@inonu.edu.tr">btutmez@inonu.edu.tr</a>
Prof. Trevor Martin	United Kingdom	<a href="mailto:trevor.martin@bristol.ac.uk">trevor.martin@bristol.ac.uk</a>



### ***I.C. Overview activities and expenditure***

**1Apr 2008-30Jun 2008**

**Total Action Budget: 21.965,00€**

**Remaining Action Commitment:**

#### **Meetings**

Meeting Type	Number				Total
WG	1				<b>12.339,43</b>
DC	1				<b>919,95</b>

#### **STSM**

Beneficiary	Number				Total
					<b>0</b>

#### **Workshops**

Beneficiary	Number				Total
E.Kontoghiorghes	1				<b>3.000,00</b>

#### **General Support Grants**

Beneficiary	Number				Total
					<b>0</b>

#### **Schools**

Title	Number				Total
					<b>0</b>

#### **Dissemination**

Title	Number				Total
					<b>0</b>

#### **Others**

Type	Total
Secretarial support	<b>2.438,91</b>

**Action Total : 18.698,29**

**Further details can be found in the financial reports**



## 1Jul 2008-28Feb 2009

Total Action Budget: 114.782,84€

Remaining Action Commitment: 85.500,00€

### Meetings

Meeting Type	Number				Total
WG & MC	1				<b>19.477,56</b>

### STSM

Beneficiary	Number				Total
					<b>0</b>

### Workshops

Beneficiary	Number				Total
					<b>0</b>

### General Support Grants

Beneficiary	Number				Total
					<b>0</b>

### Schools

Title	Number				Total
					<b>0</b>

### Dissemination

Title	Number				Total
					<b>0</b>

### Others

	Total
Bank charges	<b>51,00</b>
Secretarial Support	<b>2.929,28</b>

Action Total : **22.457,84**

Further details can be found in the financial reports



**1Mar 2009-30Jun 2009**

**Remaining Action Commitment: 85.500,00€**

**Meetings**

Meeting Type	Number				Total
WG&MC	1				<b>27.500,00</b>

**STSM**

Beneficiary	Number				Total
Several	9				<b>18.000,00</b>

**Workshops**

Beneficiary	Number				Total
					<b>0</b>

**General Support Grants**

Beneficiary	Number				Total
					<b>0</b>

**Schools**

Title	Number				Total
Spring school	1				<b>34.783,00</b>

**Dissemination**

Title	Number				Total
					<b>0</b>

**Others**

	Total
Secretarial Support	<b>12.042,00</b>

**Action Total : 92.325,00**



## **II. Scientific Report**

More additional material, in particular the slides of several scientific presentations that were given at work group meetings and in the COST Sessions at ERCIM 2008, but also agendas, the Action poster, and other presentations, can be found on the Action web site: <http://www.cost-ic0702.org/>

### **II.A. Innovative networking**

The Action started its networking activities with the Kickoff Meeting, held in Brussels (Belgium), March 31, 2008, where the organizational structure and the work groups were set up. This initial meeting already brought together several people who had not worked together before or did not even have direct contact before this Action was established.

In order to give the Action a quick start, three special sessions, one for each of the three working groups, were organized as part of the 1<sup>st</sup> Workshop of the ERCIM Work Group on Computing and Statistics, which was co-located with the 2<sup>nd</sup> Int. Workshop on Computational and Financial Econometrics (June 19-21, 2008, Neuchâtel, Switzerland). These sessions took place on June 19 and June 20, 2008. A total of 14 presentations were given by members of the Action's working groups. The purpose of these sessions was twofold: in the first place, they provided an opportunity for an initial exchange about the research areas of the work group members. Secondly, the co-location with two workshops that are strongly oriented towards statistics made it possible to attract attention from statisticians to soft computing techniques and combinations of statistical and soft computing approaches. Both purposes were successfully achieved, in particular, because the sessions attracted several other participants of the co-located workshops. A large part of the presentations of these sessions can be found at <http://www.cost-ic0702.org/ercim08.html>.

The exchange that was initiated in Neuchâtel was continued and intensified in two work group meetings in Warsaw (Poland), October 20-21, 2008, and Palma de Mallorca (Spain), March 16-17, 2009. Some of the presentations (slides) can be found on the Action web page: <http://www.cost-ic0702.org/meetings.html>. A difference to the Neuchâtel sessions was that more time was allocated for questions and discussions, which was well received by the participants and used extensively for animated discussions about the presentations.

The main research results were and will be achieved in the short-term scientific missions, of which nine were or will be completed until the end of the COST year 2008/2009 (see II.C for details). The subject areas of these STSMs covered applications of statistical and soft computing techniques in three main areas: (1) information retrieval (novelty detection, summarization, topic detection and tracking, opinion detection), (2) statistics with fuzzy data (hypothesis tests with imperfect data, linear regression with imprecise response, robust fuzzy regression), (3) knowledge discovery in complex and structured domains (mining stream data, change mining, spatial data analysis, feature selection). All of these projects address challenging problems and can be expected, on completion, to provide significant contributions to scientific progress.

A considerable success of the Action is the organization of a spring school (also supported by COST Actions IC0602 and IC0801) on "Reasoning and Decision Making under Uncertainty and Imprecision", which will be held on May 18-20, 2009, at the European Centre for Soft Computing, Mieres, Spain (see <http://www.cost-ic0702.org/springschool/> for more information). Despite the short time that was available for the organization, it was possible to put together a fairly high class program, which not only furthers the networking activities within COST Action IC0702, but also initiates and strengthens links to COST Actions IC0602 and IC0801, which cover related areas. At the time of this writing we already have more than 30 registrations and thus can be confident that the spring school will be a significant success.

To date no national or European proposal or projects could be created. The main reason for this is that it proved even more difficult than expected to find common grounds for the diverse set of researchers gathered in this action. Even though difficulties were foreseen, the necessary exchange about the individual research areas

proved to be very time consuming and a serious challenge. Nevertheless, a subgroup of working group C has been formed that more intensely pursues the goal of putting together a European Union research proposal.

### ***II.B. Inter-disciplinary networking***

The Action is inter-disciplinary at its core, since its goal is to bring together the disciplines of statistics and soft computing. Basically all of what was said under II.A therefore also applies to this section. Particularly noteworthy are the links that were established together with COST Actions IC0602 and IC0801 through the organization of a spring school that is supported by all three actions (more information can be found at <http://www.cost-ic0702.org/springschool/>). Possibilities to extend and strengthen these connections will be explored during the COST Annual Progress Conference in Santander (Spain), June 6-8, 2009, where the Chairs of these COST Actions plan to take the opportunity to meet and discuss future collaborations.

### ***II.C. New networking***

Up to now, the Action has brought together a total of 48 researchers from different backgrounds and has triggered a lively exchange between them. They come from 18 countries and include 7 women (15%). (This number does not include the four female early stage researchers executing short-term scientific missions, see below).

The Action started with 10 countries signing the MoU and has now 17 countries. Furthermore, we expect to have a new country (Bulgaria) joining the Action in the following months. Furthermore, we started with 14 initial WG members and we have now 48 members of WGs actively participating in the meetings we have organised.

In the COST year 2008/2009 a total of nine STSMs have been approved. Eight of these STSMs were/will be carried out by early stage researchers, four of which are women.

- **STSM-COST-IC0702-0801** (period: 13.04.-24.04.2009)  
Visitor: Mr Az Azrinudin Alidin (early stage researcher), University of Lugano, Switzerland  
Host: Prof. David Losada, University of Santiago de Compostela, Spain
- **STSM-COST-IC0702-0802** (period: 08.06.-13.06.2009)  
Visitor: Prof. Fabio Crestani, University of Lugano, Switzerland  
Host: Prof. Fazli Can, Bilkent University, Turkey
- **STSM-COST-IC0702-0803** (period: 04.05.-15.05.2009)  
Visitor: Ms Shima Gerani (early stage researcher), University of Lugano, Switzerland  
Host: Prof. Maatern De Rijke, University of Amsterdam, Netherlands
- **STSM-COST-IC0702-0804** (period: 23.03.-23.04.2009)  
Visitor: Ms Ana Belen Ramos-Guajardo (early stage researcher), European Centre for Soft Computing, Mieres, Spain  
Host: Prof. Thierry Denoeux, Université Technologie de Compiègne, France
- **STSM-COST-IC0702-0805** (period: 07.01.-07.02.2009)  
Vistor: Mr Antonio Balzanella (early stage researcher), University of Naples "Federico II", Italy  
Host: Dr. Yves Lechevallier, INRIA-Rocquencourt, France
- **STSM-COST-IC0702-0806** (period: 01.03.-31.03.2009)  
Visitor: Ms Maria Brigida Ferraro (early stage researcher), University of Rome "La Sapienza", Italy  
Host: Dr. Gil Gonzalez-Rodriguez, European Centre for Soft Computing, Mieres, Spain
- **STSM-COST-IC0702-0807** (period: 23.05.-25.06.2009)  
Visitor: Dr. Bülent Tütmez (early stage researcher), Inonu University, Turkey  
Host: Prof. Edzer Pebesma, University of Münster, Germany
- **STSM-COST-IC0702-0808** (period: 10.05.-10.06.2009)

Visitor: Ms Katharina Tschumitschew (early stage researcher),  
University of Applied Sciences Braunschweig/Wolfenbüttel, Germany  
Host: Dr. Detlef Nauck, British Telecom Intelligent Systems Research Center,  
Ipswich, United Kingdom

- **STSM-COST-IC0702-0901** (period: 24.05.-15.06.2009)

Visitor: Dr. Giorgio Luciano (early stage researcher), University of Genova, Italy  
Host: Prof. Tormod Næs, Oslo University, Norway

From those STSMs that were already completed at the time of this writing and for which scientific reports have been delivered, drafts of research papers are already available. These papers will be completed and submitted to conference and journals in the course of the year 2009.

In addition to the STSMs, early stage researchers are targeted by the spring school on “Reasoning and Decision Making under Uncertainty and Imprecision”, which the Action organized in cooperation with COST Actions IC0602 and IC0801 on May 18-20, 2009, at the European Centre for Soft Computing, Mieres, Spain (more information can be found at <http://www.cost-ic0702.org/springschool/>). In addition, a summer course “Soft Computing and Statistics” has been organized in conjunction with the 2009 Fuzzy Systems Association World Congress (IFSA 2009) and the 2009 Conference of the European Society for Fuzzy Logic and Technology, to be held on July 17-20, 2009, at the Technical University of Lisbon, Portugal (more information can be found at <http://www.cost-ic0702.org/summercourse/>).

#### ***II.D. Self evaluation***

Even though it was expected that due to the highly inter-disciplinary nature of the COST Action it would not be easy to persuade people from different backgrounds to work together, this proved to be an even more serious challenge. Even within each domain (statistics and soft computing) and work group, it took several meetings to convey a sufficiently clear idea of the individual research areas the participants are working in. As a consequence, our expectations to start working on project proposals (EU or other) within the first year proved to be unrealistic. However, some progress in this direction has been made and a subgroup of work group C has been formed, which is now actively pursuing the goal to prepare an FP7 project proposal. Nevertheless it will likely still take some time before it is matured enough to be submitted.

Despite these problems, the scientific exchange between work group members that was triggered and sustained at the meetings in Neuchâtel (Switzerland), June 19-20, 2008 (in conjunction with the ERCIM conference), Warsaw (Poland), October 20-21, 2008, and Palma de Mallorca (Spain), March 16-17, 2009, was very fruitful, and resulted lively and animated discussions, which everyone was eager to continue in future meetings.

Motivating enough people to consider going on a short-term scientific mission also proved to be less easy than we had hoped. It took some time before first applications were submitted and could be evaluated. Nevertheless, by repeated advertising among the work group members of the action, we finally managed to fall only one short of the envisioned number of 10 STSMs in the first year (9 STSMs were carried out or are scheduled and will be completed in the COST year 2008/2009). We hope that in the future, once such a start has been made, it will become easier to encourage enough people to go on a short-term scientific mission.

From those short-term scientific missions that have already been completed at the time of writing of this report and for which the scientific reports are therefore already available, we see that a fruitful exchange has been stimulated. These STSMs led to drafts of scientific papers which, in the course of this year, will be completed and submitted to conferences or journals.

#### ***III. Previous scientific report(s)***

Not applicable, as this is the first monitoring progress report.