

A Report on the Fourth International Workshop on Environmental Applications of Machine Learning (EAML 2004)

27-29 September 2004, Bled, Slovenia

Sašo Džeroski
Jozef Stefan Institute
Jamova 39
SI-1000 Ljubljana, Slovenia
saso.dzeroski@ijs.si

Bernard Ženko
Jozef Stefan Institute
Jamova 39
SI-1000 Ljubljana, Slovenia
bernard.zenko@ijs.si

Marko Debeljak
Jozef Stefan Institute
Jamova 39
SI-1000 Ljubljana, Slovenia
marko.debeljak@ijs.si

The Fourth International Workshop on Environmental Applications of Machine Learning (EAML) was organized in Bled, Slovenia, during the week of September 27 – October 1, 2004, jointly with the Fourth European Conference on Ecological Modelling (ECEM). The aim of these events was to bring together researchers from the areas of ecology, ecological modelling and environmental sciences, on one hand, and the areas of data analysis, data mining and machine learning, on the other. In many respects, they were treated as a single event, namely ECEM/EAML 2004.

As environmental concerns grow and information technology develops, more and more data on the different aspects (physical, chemical, biological, ecological) of the environment are gathered. There is an increasing need to analyse the collected environmental data for different purposes, which include the support for environmental management decisions. Machine learning methods can be used to discover knowledge in environmental data, as well as construct ecological models in an automated fashion, thus providing support for environmental management.

The International Workshop on Environmental Applications of Machine Learning provided a forum for presenting recent advances in applying machine learning and data mining techniques for the analysis of environmental data. EAML 2004 was the fourth workshop on the topic, with previous workshops mainly attracting researchers in ecological modeling interested in the use of machine learning. The ECEM/EAML 2004 conference brought together participants from both communities, including many prominent figures, such as the founders/first editors of the journals of Ecological Modelling and Machine Learning.

The scientific program of ECEM/EAML 2004 consisted of invited lectures as well as oral and poster presentations of accepted contributions. The conference brought together two separate communities which have overlapping research topics. Both events have had separate advisory committees and invited speakers. However, the program chairs and organizing committee as well as the program committee and review process were joint for both events. Also, a single registration fee allowed access to both events.

In addition to the scientific program, each day was concluded with a social event which facilitated the interaction and communication between the two communities. Each of ECEM and EAML lasted for three days (EAML September 27-29, ECEM September 29 to October 1). Wednesday September 29 was a joint/overlap day, with a joint invited speaker, technical sessions of interest to both communities and a joint poster session.

A total of 110 abstracts were submitted until March 20. Each submitted abstract was sent to three members of the program committee for review. The program committee members themselves did the majority of reviews, assisted by some additional reviewers appointed by the committee members. Classification of accepted presentations for oral or poster sessions was based on the reviewers' evaluation of their quality and generality of interest. Revised versions of accepted abstracts were published in two separate proceedings, one for ECEM 2004 (ISBN:961-6303-60-0) and another for EAML 2004 (ISBN: 961-6303-59-7). Both proceedings are available on-line from the website of ECEM/EAML.

Of the 110 submissions, 90 were accepted. Of these, 29 went to EAML and 61 to ECEM. Authors of accepted abstracts (for both oral and poster presentation) were invited to submit full versions of their papers by June 20. About 60 full papers have been already submitted, and they are currently being reviewed by members of the program committee and additional reviewers. The accepted papers will be published in two special issues of the journal Ecological Modelling (one for ECEM 2004, one for EAML 2004).

At ECEM/EAML, opening addresses were given by distinguished guests. These included Zoran Stančič (State secretary, Slovenian Ministry of Education, Science and Sport) and Nada Lavrač (Head of the Department of Knowledge Technologies, Jožef Stefan Institute, Ljubljana), who spoke at the opening of EAML 2004 on September 27. Tarzan Legović (President of European Chapter of ISEM, Rudjer Bosković Institute, Zagreb, Croatia) spoke at the opening of ECEM 2004 on September 29.

The program of EAML 2004 included three invited talks, given by Joseph C. Coughlan (NASA Ames, USA), Cesare Furlanello (ITC-irst, Trento, Italy), and Jacqueline McGlade (European Environment Agency), as well as 22 oral and 7 poster presentations of accepted contributions.

The topics which were covered by the presentations can be clustered in three major (overlapping) groups: analysis of environmental data with various data mining methods and techniques (e.g., computational scientific discovery, decision and regression trees, evolutionary computing, statistical learning, neural networks), application of data mining and machine learning methods to ecological and environmental modelling (e.g., modelling of different types of ecosystems, modelling different aspects of ecosystems and ecological processes, analysis of spatial data with geographic information system (GIS), analysis of different types of environmental data (paleo/historical and recent), and various environmental applications of machine learning for decision support (e.g. water treatment plants, risk assessment, earthquake predictions, climate changes, habitat modelling).

The first EAML invited talk, "An overview of ecological modeling and machine learning research within the U.S. National Aeronautics and Space Administration" given by Joseph C. Coughlan, NASA Ames, USA, summarized NASA research activities to understand global habitability, and to quantify the processes and fluxes between the Earth's vegetation and the biosphere. He presented some of the research on data mining in support for environmental sciences, that NASA has funded within its Intelligent Systems project. He gave a specific example in ecological forecasting: predicting land surface properties given weather forecasts and nowcasts, using the Terrestrial Observation and Prediction System.

Besides the invited talk, the first day featured presentation on a variety of topics. This included the analysis of data on forests, water and waste water, and forested watersheds. Two papers on paleoecology, the ecology of organisms living in the past, were also presented. The final session of the day was on the topic of modeling climate and air pollution and concluded with a talk on the analysis of radon content in thermal water gasses for the prediction of earthquakes.

The second EAML invited talk "GIS-based predictive models for ecology" presented by Cesare Furlanello, ITC-irst, Trento, Italy, discussed how machine learning methods can be integrated within a Geographical Information System (GIS) for the development of new approaches in ecology research. He presented a risk-mapping system applied to two cases. He first described a risk mapping system for tick-borne diseases in the Trentino region of the Italian Alps, and second a predictive risk model for deer-vehicle collisions developed for the Wildlife Management and Road Transportation Services of the same region. The talk also presented different methods for variable importance analysis, classification with combined models and the resulting roe deer-vehicle accident risk maps.

On the same day, a number of contributions were presented dealing with habitat models, constructed from GIS data by using machine learning methods. On the aquatic side, this included habitat models for freshwater fish and decapods. On the terrestrial side, habitat models were constructed for soil insects, foliage-dwelling spiders, and red deer.

The scientific program of the joint ECEM/EAML day started by the invited talk of Jacqueline McGlade, executive director of the European Environment Agency (EEA), titled "Spatial assessments of Europe's environment".

The first part of the talk presented the requirements for properly examining issues such as the impact of climate change, loss of biodiversity, environmental threats to human health, the long-term effects of infrastructure development on Europe's landscapes. Namely, the EEA needs to analyse changes in Europe's environment across a wide range of scales and media (water, air, soil, etc.). The second part of the talk presented land accounts for Europe that are being implemented by the EEA. The purpose of land accounts is to observe, qualify and quantify the cover of land resulting from ecosystem and land use.

The joint day included several technical sessions on topics of interest to both communities. These included a session on qualitative modelling, as well as a session on automated process-based modelling of ecosystems, using both measured data and domain knowledge. The latter attracted a lot of attention from prominent figures in the ecological modelling community and the topic was judged as a very promising direction for further research. The day was concluded by the joint poster session.

The remaining two days were devoted to the ECEM conference. Broder Breckling, University of Bremen, Germany, gave an invited talk on Thursday, September 30: the talk was titled Individual-based models as tools for ecological theory and application - Understanding the emergence of organisational properties in ecological systems. Bai-Lian Li, University of California Riverside, USA, gave an invited talk on October 1: the talk (given in a teleconferencing session) was titled Modeling ecological complexity: Challenges and opportunities. Besides the invited talks on individual-based modelling and ecosystem complexity, many contributions to ECEM on these topics were presented.

Both communities had separate community meetings (EAML on September 28, and ECEM on September 30) where the future of the fields, and primarily future workshops/conferences were discussed. Participants were asked for suggestions about the organisation of such events, review process, format of publication, and publicity of both communities. The discussion about future ECEM and EAML meetings/conferences was started, to be continued in appropriate forums after the event.

In summary, ECEM/EAML 2004 succeeded in providing a forum for presenting the most recent and high quality work on the topics of ecological modelling and environmental applications of machine learning. It enabled stimulating discussions between members of each of the two communities. More importantly, it facilitated the exchange of knowledge and ideas between the two communities. Hopefully we will see much more of this in the future.

The ECEM/EAML 2004 conference was organised by the Jožef Stefan Institute, Ljubljana. It was attended by 98 participants from 20 countries. The conference was financially supported by Slovenian Ministry for Education, Science and Sport, ISEM - The International Society for Ecological Modeling, KD-net - The Knowledge Discovery Network, and PASCAL - Pattern Analysis, Statistical Modelling and Computational Learning: The network of excellence.

The slides of the oral presentations and the proceedings of both events are available at the conference web page (<http://www-ai.ijs.si/SasoDzeroski/ECEMEAML04/>).