**INTERNATIONAL KARS CONGRESS ON SCIENTIFIC RESEARCH**
Agriculture, Biology, Zoology, Veterinary, Natural Sciences, Animal Farming, Food, Iksad Institute, February 17-19, 2023, Kars Turkiye

**JOKER#2 or derivation of lacking data in biological systems**

László Pitlik (<https://orcid.org/0000-0001-5819-0319>), Dániel Váradi (<https://orcid.org/0000-0001-9610-8566>),

e-Mails: pitlik@my-x.hu, danielvaradi140@gmail.com

Kodolányi János University and MY-X research team Hungary

*Keywords: modelling, testing, validating, consistence, artificial intelligence, knowledge representation, ceteris paribus*

**Abstract**

History of the project: The SPEL (Sektorales Produktions- und Einkommensmodell der Landwirtschaft = Sectoral Production and Income Model for Agriculture, see: <https://op.europa.eu/portal2012-portlet/html/downloadHandler.jsp?identifier=84b7fd4b-7fd9-4446-98b0-500f176f891f&format=pdfa1b&language=en&productionSystem=cellar&part=>) is a sophisticated framework for a consistence-oriented data asset management (since 1980). The Hungarian activities on this research field is an important part of the education about integrated information systems. The most relevant characteristics of the SPEL system is the strong consistence, where raw data are finetuned if biological rules can not be reflected in a direct way.

Goals/Tasks: The presentation and the full-text publication have to demonstrate, how robust biological patterns can be modelled based on a non-causal, but multi-layered consistence-oriented approach (JOKER#2). The robustness of these estimations is not only a question of numerical approximation, the experiments (as tasks) should also be capable of covering the hidden system-logic.

Solutions: JOKER#2 is the name of a new software-concept ensuring automated estimations in a context-free way (c.f. <https://tr.discoveranatolia.org/_files/ugd/614b1f_5d2ae9f2566d4177a40f63debab3f684.pdf#page=29>). JOKER#1 was constructed by Dobó independent but parallel to the starting years of the SPEL system. The context-free characteristics of JOKER#2 (here and now in form of an MS Excel-Solver-application) can be finetuned through arbitrary constraints ensuring consistence-oriented impacts in the non-causal optimization processes.

There are further own approaches (different forms of similarity analyses – like production functions, anti-discriminative models, explorative models – COCO: <https://miau.my-x.hu/myx-free/>) for detecting data anomalies and for estimation lacking data positions. Similarity analyses can be seen as a kind of causal modelling with additive or multiplicative structures. Similarity analyses are capable of handling always one single phenomenon. JOKER#2 can be used for arbitrary phenomena in a parallel way. Similarity analyses are using staircase functions as knowledge representation. JOKER#2 has not a single visible knowledge representation characteristic (see: non-causal modelling). Similarity analyses can produce unlimited estimation for the even-observed attribute. The recent version of JOKER#2 can deliver estimations max. only 100 positions. In case of JOKER#2 the lacking positions do not have to follow a specific pattern (only one single attribute/column and one single object/row should have real data). JOKER#2 is not a kind of simple proportion-driven calculation scheme (see demo materials with the keyword of ‘kazah’).

Experiments: For this study, a lot of existing patterns were involved into optimization tasks in order to prove, what kind of biological rules can be approximated in a robust form and what kind of relationships can not be interpreted not even based on the potential constraints? The examined rules are less or more complex (physical and/or monetary) balances (in specific cases in form of chained annual relationships between data).

Conclusions: The similarity analyses (in the training phases) are always capable of detecting (additive) balances based on a limited amount of cases. Similarity analyses can deliver two production functions and two anti-discriminative models in case of one single lacking data. JOKER#2 produced one single approximation. The best model in case of one single lack is COCO STD as production function for one raw attribute as Y-variable. The second-best approximation is JOKER#2, because the analysed random pattern had a very wide-ranged robustness (quasi limitless) and JOKER#2 was capable of approximating the real targeted value compared to this wide instability-range closer than the COCO Y0 (as third competitor). Parallel, it was necessary to explore the impact of the data volume concerning the competitiveness of the alternative models: more examples (data: 20 to 100 objects) did not always produce better approximations for the test cases. Explorative models, anti-discriminative models, production functions were error-free again. Y0-model based on more data is better. STD-model based on more data is not better. Explorative models delivered irrational estimations – especially based on raw data because of lack of balance-oriented directions. JOKER#2 with more data ensured better estimations than with less data. Regression models are also error-free – not only in case of pre-defined customized signs (directions) for the X-variables (c.f. real causality). Initializing with doubled attribute sets also lead to error-free interpretations in the learning phases – but without real causalities. Multiplicative knowledge representations deliver low error-levels, but not error-free approximations.

Future: More balances and more complex balances should be examined in future in order to compare causality vs. flexibility.

Demo materials: <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=kazah>, <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=spel>, <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=mszr>, <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=aszm>, <https://miau.my-x.hu/miau/296/spel_joker2.xlsx>, https://miau.my-x.hu/miau/296/spel\_joker2\_negative.xlsx