



**AGENT-BASED
SUPPORT TOOL FOR
THE DEVELOPMENT
OF AGRICULTURE POLICIES**

D2.3 Big Data Fusion Modul



Deliverable Number	D2.3
Lead Beneficiary	AUTH
Authors	AUTH, IDE
Work package	WP2
Delivery Date	31/08/2022
Dissemination Level	Public

www.agricore-project.eu



The Agricore project has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No. 816078





Document Information

Project title	Agent-based support tool for the development of agriculture policies
Project acronym	AGRICORE
Project call	H2020-RUR-04-2018-2019
Grant number	816078
Project duration	1.09.2019-31.8.2023 (48 months)
Deliverable Authors	Michail Tsagris (AUTH), Vangelis Tzouvelekas (AUTH), Dimitrios Natos (AUTH), Konstadinos Mattas (AUTH)
Deliverable Reviewers	IDENER Team

Version History

Version	Description	Organisation	Date
1.0	Deliverable complete	AUTH	30/08/2022
1.0 GER	German summary translation	AXIA	31/08/2022

Executive Summary

AGRICORE ist ein Forschungsprojekt, das von der Europäischen Kommission im Rahmen der Ausschreibung RUR-04-2018 finanziert wird, Teil des H2020-Programms, das einen innovativen Weg zur Anwendung agentenbasierter Modellierung vorschlägt, um die Fähigkeit der politischen Entscheidungsträger zu verbessern um die Auswirkungen von agrarbezogenen Maßnahmen innerhalb und außerhalb des Rahmens der Gemeinsamen Agrarpolitik (GAP) zu evaluieren. In diesem Beitrag wird das AGRICORE-Datenfusionsmodul vorgestellt, das die Integration und Verschmelzung einzelner Datensätze (die zuvor mit dem Datenextraktionsmodul gewonnen wurden) zu angereicherten Datensätzen erlaubt, die für den Betrieb der verschiedenen AGRICORE-Module verwendet werden.

Die wichtigste Datenfusionsoperation, die für die Umsetzung eines AGRICORE-Anwendungsfalls erforderlich ist, ist die für die Erstellung der synthetischen Agenten, der zu untersuchenden Landwirtschafts- und Viehzuchtbetriebe. Insbesondere wird ein mathematisches Artefakt benötigt, um die Werte zu erzeugen, die dann Attributen zugewiesen werden, aus denen sich die einzelnen Agenten zusammensetzen.

Das mathematische Werkzeug, das für diese Funktion gewählt wurde, ist das Bayessches Netz (BN). Dieser Beitrag stellt den/die Algorithmus(e) zur Konstruktion eines Bayessches Netz vor, der/die entwickelt und/oder verbessert wurde(n), um im Rahmen des AGRICORE-Projekts verwendet zu werden.

Um diese(n) Algorithmus(e) zu testen, wurden vier synthetische Stichproben von landwirtschaftlichen Betrieben in drei NUTS2-Regionen und einer NUTS3-Unterregion Griechenlands erstellt. In diesem Bericht werden diese Beispielfälle vorgestellt einschließlich der Aggregationen spezifischer Variablen, die Struktur des resultierenden BN für jeden Fall und die Bewertung der Übereinstimmung der generierten synthetischen Stichprobe mit der realen Basisstichprobe.

Conclusions

In diesem Beitrag wurde das Datenfusionsmodul (DFM) des AGRICORE-Suites' vorgestellt. Das DFM greift auf die einzelnen Datensätze zu die zuvor vom DEM extrahiert und transformiert wurden, sowie auf deren Metadaten (statistische Charakterisierung und verbotene Beziehungen), die ebenfalls über das DEM gewonnen wurden. Sobald die erforderlichen Daten geladen sind, führt das DFM eine Reihe von Prozeduren durch, um durch die Integration der einzelnen Datensätze angereicherte Datensätze zu erzeugen. Diese angereicherten Datensätze werden für verschiedene Prozesse in AGRICORE verwendet. Der wichtigste davon ist die Konstruktion von anonymisierten Agenten, die die synthetische Population bilden, die anschließend von der ABM-Engine simuliert wird. Diese Konstruktion erfordert für jeden Agenten die Erzeugung von Pseudo-Zufalls-Werten und diese dann jedem seiner Attribute zuzuordnen. Da die Variablen, die mit bestimmten Attributen eine Korrelation mit den Variablen anderer Attribute aufweisen, bedingt die Zuweisung eines Wertes für ein Attribut den Bereich der Werte, die anderen Attributen zugewiesen werden können. Daher ist es notwendig, ein mathematisches Objekt zu haben, um die Reihenfolge zu bestimmen in der die Attribute erzeugt werden, sowie die gemeinsamen Wahrscheinlichkeitsdichten dieser Attribute.

Das für AGRICORE gewählte mathematische Artefakt ist das Bayessches Netz. Dieser Beitrag beschreibt den Max-Min Hill Climbing (MMHC) Algorithmus und die Varianten, die in ihn eingebaut wurden, um ihn an die besonderen Bedürfnisse des Projekts anzupassen.

Um die Leistung der mit dem MMHC erstellten BNs zu testen, wurden 4 Beispielfälle auf regionaler (NUTS2) und subregionaler (NUTS3) Ebene implementiert, die zum griechischen Anwendungsfall des AGRICORE-Projekts gehören. Im Einzelnen wurden auf der Grundlage der regionalisierten Teilstichproben des griechischen FADN, äquivalente synthetische Teilstichproben konstruiert und ihre Anpassungsgüte analysiert.

Die Ergebnisse zeigen, dass die Anpassung in allen vier Fällen sehr genau ist, was das vorgeschlagene Verfahren sehr vielversprechend für die Anwendung im Synthetischen Populationsgenerator (SPG) macht. Die nächsten Schritte sind die Integration und Paketierung der BN-Konstruktionsskripte für die Ausführung durch den SPG und das Testen der Generierung vollständiger synthetischer Populationen für die 3 im Projekt vorgesehenen Anwendungsfälle.