



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

D2.3. Módulo de fusión de *Big Data*



Deliverable Number	D2.3
Lead Beneficiary	AUTH
Authors	AUTH, IDENER
Work package	WP2
Delivery Date	31/08/2022 (M36)
Dissemination Level	Público

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), IDENER Team
Deliverable reviewers	IDENER Team

Version History

Version	Descripción	Organización	Fecha
0.1	Propuesta	AUTH	10 Mayo 2022
0.2	Aprobación	IDENER	15 Mayo 2022
0.3	Inclusión de contenido (Primer borrador)	AUTH	19 Julio 2022
0.4	Revisión y comentarios	IDENER	07 Agosto 2022
0.5	Implementación de las correcciones	IDE	20 Agosto 2022
1.0	Exportación final y formato (Versión Final)	AUTH	31 Agosto d2022

Resumen ejecutivo

AGRICORE es un proyecto de investigación financiado por la Comisión Europea dentro de la convocatoria RUR-04-2018 parte del programa H2020, que propone una forma innovadora de aplicar la modelización basada en agentes para mejorar la capacidad de los responsables políticos para evaluar el impacto de las medidas relacionadas con la agricultura dentro y fuera del marco de la Política Agrícola Común (PAC).

Este producto presenta el módulo de fusión de datos AGRICORE, que permite la integración e integración y mezcla de conjuntos de datos individuales (obtenidos previamente por el módulo de extracción de datos) para constituir conjuntos de datos enriquecidos que se utilizan para el funcionamiento de los diferentes módulos de AGRICORE.

La principal operación de fusión de datos necesaria para la implementación de un caso de uso de AGRICORE es la necesaria para producir los agentes sintéticos que representan las explotaciones agrícolas y ganaderas en estudio. En concreto, se necesita un artefacto matemático para generar los valores que luego se asignan a los atributos que componen cada uno de los agentes.

La herramienta matemática elegida para realizar esta función es la Red Bayesiana (BN). Este entregable presenta el algoritmo o los algoritmos de construcción de redes bayesianas que se han desarrollado y o mejorado para ser utilizado en el proyecto AGRICORE.

Para probar estos algoritmos, se han generado cuatro muestras sintéticas de explotaciones agrícolas en tres regiones NUTS2 y una subregión NUTS3 de Grecia. Este entregable presenta estos casos de ejemplo incluyendo las agregaciones de variables específicas, la estructura de la BN resultante para cada caso, y la evaluación del ajuste de la muestra sintética generada con respecto a la muestra de referencia real.

Conclusiones

Este producto presenta el Módulo de Fusión de Datos (DFM) de la suite AGRICORE. El DFM es el responsable de acceder al almacén de datos para acceder a los conjuntos de datos individuales previamente extraídos y transformados por el MDE, así como sus metadatos (caracterización estadística y relaciones prohibidas) obtenidos también a través del MDE. Una vez cargados los datos necesarios, el DFM ejecuta una serie de procedimientos para generar conjuntos de datos enriquecidos mediante la integración de los conjuntos de datos individuales.

Estos conjuntos de datos enriquecidos se utilizan para varios procesos en AGRICORE. El más importante es la construcción de agentes anónimos que forman la población sintética que posteriormente se simula por el motor ABM. Esta construcción requiere generar, para cada agente, valores pseudoaleatorios y asignarlos a cada uno de sus atributos. Dado que las variables asociadas a determinados atributos muestran correlación con las variables de otros atributos, la asignación de un valor para un atributo condiciona el rango de valores asignables a otros atributos. Por lo tanto, es necesario disponer de un objeto matemático para determinar el orden de generación de los atributos, así como las densidades de probabilidad conjuntas de estos atributos.

El artefacto matemático elegido para AGRICORE es la Red Bayesiana. Este producto describe el algoritmo Max-Min Hill Climbing (MMHC) y las variantes incorporadas a él para adaptarlo a las necesidades particulares del proyecto.

Para probar el rendimiento de las BNs construidas con el MMHC, se han implementado 4 casos de ejemplo a nivel regional (NUTS2) y subregional (NUTS3) pertenecientes al caso de uso griego del proyecto AGRICORE. En concreto, a partir de las submuestras regionalizadas de la RICA griega se construyeron submuestras sintéticas equivalentes y se analizó su bondad de ajuste.

Los resultados muestran que el ajuste es muy preciso en los cuatro casos, lo que hace que el procedimiento propuesto sea muy de la población sintética (GPS). Los siguientes pasos son la integración y empaquetamiento de los scripts de construcción de BN para su ejecución desde el SPG, y las pruebas de generación de poblaciones sintéticas completas para los 3 casos de uso contemplados en el proyecto.