

Focus

Un système OLAP pour l'analyse de données de lutte intégrée : application à la culture de l'olivier

La lutte intégrée pour la protection des cultures génère de très nombreuses données qui aujourd'hui encore ne sont pas stockées dans les systèmes d'information de gestion agricole. Or ces données sont très utiles pour aider les agriculteurs à améliorer les performances environnementales et économiques de leur exploitation. L'auteur de ce focus explore l'utilisation des systèmes d'entrepôt de données et de traitement analytique en ligne (OLAP) pour stocker et analyser ce type de données.

L

Le secteur agroalimentaire est confronté à des défis mondiaux. Le premier défi concerne l'alimentation d'une population mondiale qui, selon les prévisions de l'Organisation des Nations unies, atteindra 9,3 milliards de personnes en 2050. Le deuxième défi est de satisfaire la demande

croissante des consommateurs pour des produits de haute qualité, obtenus par des chaînes agroalimentaires plus durables (Grunert, 2005). Pour y répondre, les agriculteurs doivent augmenter la qualité et la quantité de leurs productions, en réduisant les impacts sur l'environnement grâce à de nouvelles stratégies et outils de gestion. Dans ce contexte, les systèmes d'information de gestion agricole (SIGF) jouent un rôle important. Un SIGF est défini comme un « système d'information pour la collecte, le traitement, le stockage et la diffusion des données pour la gestion efficace des opérations et des fonctionnalités d'une exploitation agricole ». Selon Fountas *et al.* (2005), les SIGFs offrent des fonctionnalités de gestion des équipements, d'inventaire, de traçabilité et vente des productions et de gestion des ressources humaines. En particulier, l'intégration des outils de lutte intégrée dans les SIGFs apparaît comme un instrument nécessaire pour aider les agriculteurs à faire face aux défis de l'agriculture durable. La lutte intégrée (LI) nécessite l'utilisation simultanée de différentes techniques de protection des cultures pour lutter contre les insectes, les agents pathogènes et les mauvaises herbes, par le biais d'une approche écologique et économique (Prokopy, 2003).

Les données de LI ne sont pas stockées dans les SIGFs existants. Par conséquent, il est difficile de rechercher les meilleures pratiques durables en fonction des données de LI des campagnes passées. Or, ces données sont très utiles pour aider les agriculteurs à améliorer les performances environnementales et économiques de leur

exploitation. Pour combler ce manque d'informations, nous explorons dans ce travail l'utilisation des systèmes d'entrepôt de données (ED) et de traitement analytique en ligne (OLAP) pour l'analyse de données de LI.

Dans ce travail, nous nous sommes intéressés à un jeu de données provenant de quelques agriculteurs qui collaborent avec l'université de Foggia (Italie) dans le cadre d'un projet de recherche en cours. Ces dernières années, les systèmes OLAP ont été appliqués avec succès dans plusieurs domaines tels que le marketing, la santé... et aussi dans l'agriculture (Bimonte, 2016). Ce sont des outils web qui permettent l'analyse de données à distance, sur la base de tableaux croisés dynamiques. Sur ce dernier point, ils se rapprochent donc des outils Excel®. Cependant à la différence de Excel®, les outils OLAP sont connectés à des entrepôts de données permettant l'analyse de très gros volumes de données (des millions, voire des milliards de données). Enfin, contrairement aux bases de données classiques, les entrepôts de données structurent les données avec une modélisation particulière qui rend leur traitement très performant.

Modèle OLAP pour l'analyse de données de lutte intégrée

Dans cette section, nous décrivons la LI contre les ravageurs de l'olivier que nous utilisons dans ce travail, ainsi que le modèle OLAP que nous proposons pour l'analyse de données de LI.

Lutte antiparasitaire intégrée pour l'olivier

La directive-cadre 2009/128/CE (*European Parliament*, 2009) sur l'utilisation durable des pesticides oblige tous les utilisateurs professionnels de pesticides à appliquer l'approche de LI dans les pays membres européens, depuis le 1^{er} janvier 2014.



❶ L'olivier est une culture relativement rustique, mais qui peut être fortement attaquée par la mouche *Bactrocera oleae*, son principal ravageur.

© Y. Laaroussi (Irstea)

Dans ce travail, nous utilisons la LI comme un indicateur de durabilité qui assure aux exploitations des meilleures performances économiques et environnementales. L'idée principale est de vérifier que les agriculteurs appliquent les spécifications de la LI.

La LI spécifique de chaque culture est composée de deux parties :

- une partie générale : ensemble de recommandations concernant tous les aspects de la gestion des cultures, telles que les restrictions de fertilisation, la gestion des sols, la fréquence d'irrigation, etc. ;
- les règles de défense : toutes les opérations autorisées (agronomiques et chimiques) pour protéger la culture des organismes nuisibles.

Dans ce travail, nous nous intéressons plus particulièrement aux règles de défense incluses dans les spécifications techniques de la région des Pouilles (Italie) pour la culture de l'olivier (photo ❶).

Les données de la LI pour la protection de la culture de l'olivier dans la région des Pouilles concernent les entrées suivantes :

- ravageur : le type de ravageur de la culture de l'olivier. Par exemple, *Bactrocera oleae*, qui est le ravageur le plus dangereux pour l'olivier dans le monde (Navrozidis *et al.*, 2000) ;
- critères d'intervention : le seuil d'intervention représente la limite avant laquelle l'organisme nuisible ne cause pas de dommages économiques aux cultures ;
- substance active et auxiliaire : la liste des substances actives autorisées et des auxiliaires pouvant être utilisées pour lutter contre l'organisme nuisible ;
- nombre maximum de traitements : le nombre maximum de traitements admis pour une molécule spécifique. Par exemple, les produits à base de *Beauveria bassiana*¹ ne sont pas limités, car il s'agit d'un organisme naturel

sans impact sur l'environnement. En revanche, le diméthoate et l'imidaclopride sont limités à deux et un traitement par an respectivement, en raison de la toxicité des molécules ;

- limites d'utilisation et remarques : la liste des autres restrictions éventuelles pour une molécule spécifique, telles que la restriction temporelle. Par exemple, les produits à base d'imidaclopride ne sont autorisés qu'après la phase phénologique de floraison.

Modèle OLAP

Le modèle OLAP organise les données en fonction de dimensions et de faits. Les dimensions représentent les axes d'analyse et sont organisées en hiérarchies. Une hiérarchie est composée de niveaux qui définissent les granularités d'analyse. Les faits sont les sujets d'analyse et sont décrits par des mesures. En général, les mesures sont des valeurs numériques qui sont agrégées à l'aide de fonctions d'agrégation numérique sur les niveaux de dimension. Les données stockées sont ensuite explorées à l'aide des opérateurs OLAP, qui permettent de naviguer dans les hiérarchies (Roll-Up et Drill-Down), et de sélectionner un sous-ensemble de données de l'ED (Slice and Dice).

Le modèle multidimensionnel présente cinq dimensions :

- temps : dimension temporelle avec la hiérarchie, jour, mois, année ;
- campagne : qualifie la campagne, par exemple 2009-2010 ;
- exploitation : décrit l'organisation des exploitations. Cette hiérarchie permet d'analyser tous les traitements effectués à différentes échelles spatiales. Par exemple, une même ferme peut traiter deux parcelles de la même culture de différentes manières au cours d'une même campagne, afin d'évaluer la meilleure stratégie de gestion pour les années à venir ;
- ravageurs : représente les nuisibles (*Bactrocera oleae*, *Spilotea oleagina*, etc.) qui sont organisés en types, par exemple mauvaises herbes, champignons, insectes, bactéries, etc. ;

1. Champignon du sol utilisé en lutte biologique.

• traitements : concerne les différents types d'interventions autorisées dans la spécification technique de la LI. Les mesures et leurs agrégations sont :

• nombre de traitements : détaille le nombre de traitements effectués (par exemple, l'application de diméthoate au cours de la campagne 2009-2010) ;

• respect de la LI : mesure booléenne indiquant quand l'action effectuée respecte les règles de la LI. Par exemple, l'utilisation du glyphosate à 30,8 % n'est pas autorisée. Par conséquent, la valeur de mesure pour ce traitement est « NON » ;

• type d'erreur : décrit si le traitement est conforme aux règles de la LI. C'est une valeur d'énumération :

– permis : le traitement est autorisé (par exemple, l'utilisation de cuivre pour lutter contre le ravageur *S. oleagina*) ;
 – substance active (SA) non permise : le traitement n'est pas autorisé (par exemple, l'utilisation de glyphosate à 30,8%) ;

– nombre d'interventions dépassé : le nombre de traitements dépasse la quantité maximale de traitement autorisée, par exemple la troisième intervention basée sur le diméthoate visant à contrôler *B. oleae* n'est pas autorisée ;
 – SA non autorisée dans le temps : le traitement effectué n'est pas autorisé pendant cette période. Par exemple, les produits à base d'imidaclopride sont interdits avant la phase phénologique de floraison ;

– erreurs : plusieurs erreurs se produisent. Il est donc nécessaire de procéder à une analyse approfondie pour obtenir le type d'erreur spécifique pour un traitement unique.

Analyses OLAP

Grâce au modèle OLAP (et à son implémentation dans un entrepôt de données), les décideurs/agriculteurs peuvent répondre à différents types de questions :

• spécifications techniques : conformité aux spécifications techniques de la LI pour les cultures d'olives fournies par la région des Pouilles. Dans ce cas, le modèle permet d'enquêter sur le respect de la spécification technique. Ces requêtes peuvent être regroupées en deux classes :

– les requêtes « inter-exploitations » comparent les traitements entre différentes exploitations sur des données historiques. On peut citer ces exemples de requêtes : « Visualiser tous les traitements chimiques des fermes » (figure 1), « Visualiser les traitements de *Bactrocera olea* des exploitations respectant les spécifications techniques », « Visualiser tous les traitements chimiques non autorisés des exploitations ne respectant pas les spécifications techniques », etc. Pour des raisons de confidentialité, les informations les plus détaillées à l'échelle de l'exploitation ne sont disponibles que pour son responsable ;

– les requêtes « intra-exploitation » comparent les traitements entre différentes parcelles d'une même ferme, comme par exemple « Visualiser tous les traitements chimiques effectués sur les parcelles de la ferme de M. Dupont » (figure 2), « Visualiser les traitements qui ne respectent pas les spécifications techniques des parcelles de l'exploitation », etc. Il est donc possible d'analyser le respect des spécifications techniques à une échelle inférieure de celle de l'exploitation et d'identifier exactement où un mauvais traitement a été effectué ;

1 Exemple de requête « inter-exploitations » et « spécifications techniques » (Zaza et al., 2017).

Company	Treatments	Time	Error type	Respect of IPM	Number of Treatments
+Bimonte	Copper	+2016	Allowed	OK	2
	Deltamethrin	+2016	Errors	NO	2
	Dimethoate	+2016	Exceeded nb inter.	NO	5
	glyphosate 30,8%	+2016	A. S. not allowed	NO	1
	Pyraclostrobin	+2016	A. S. not allowed in time	NO	1
+2	Copper	+2016	Allowed	OK	5
	Dimethoate	+2016	Allowed	OK	3

2 Exemple de requête « intra-exploitation » et « spécifications techniques » (Zaza et al., 2017).

Company	Treatments	Time	Error type	Respect of IPM	Number of Treatments
+Plot1	Copper	-2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
	Deltamethrin	-2016	A. S. not allowed	NO	1
		+10/2016	A. S. not allowed	NO	1
	Dimethoate	-2016	Allowed	OK	2
		+06/2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
	glyphosate 30,8%	-2016	A. S. not allowed	NO	1
	Pyraclostrobin	-2016	A. S. not allowed in time	NO	1
	+Plot2	Copper	-2016	Allowed	OK
+09/2016			Allowed	OK	1
Deltamethrin		-2016	A. S. not allowed	NO	1
		+10/2016	A. S. not allowed	NO	1
Dimethoate		-2016	Exceeded nb inter.	NO	3
		+06/2016	Allowed	OK	1
		+09/2016	Allowed	OK	1
		+10/2016	Exceeded nb inter.	NO	1

3 Exemple de requête « intra-exploitation » et « gestion » (Zaza et al., 2017).

Company	Treatments	Pests	Number of Treatments
+Zaza	-agronomic	<i>P. syringae</i> pv. <i>savastanoi</i>	1
	Pruning	<i>P. syringae</i> pv. <i>savastanoi</i>	1
	-chemical	<i>Spilotea oleagina</i>	5
		<i>Bactrocera olea</i>	5
	Copper	<i>Spilotea oleagina</i>	5
	Dimethoate	<i>Bactrocera olea</i>	3
	Imidacloprid	<i>Bactrocera olea</i>	2

4 Exemple de requête à l'échelle de la parcelle (Zaza et al., 2017).

Company	Treatments	Number of Treatments			
		Pests			
		+ bacteria	+ fungi	+ insects	+ weed
+Plot11	-agronomic	1			
	Pruning	1			
+Plot12	+chemical		3	2	
	+agronomic				
+Plot12	+agronomic		2		
	+chemical			3	

• gestion : analyse détaillée des stratégies de LI. Dans ce deuxième groupe de requêtes, des analyses détaillées sont fournies à l'échelle « intra-exploitation ». En particulier, les différentes stratégies de protection des cultures sont étudiées selon différents points de vue, tels que le nombre de traitements par parcelle, le type d'intervention pour chaque ravageur, etc. Les agriculteurs et/ou les décideurs peuvent répondre à des questions de ce type : « Visualiser le nombre et le type de traitements effectués à l'échelle de l'exploitation pour chaque organisme nuisible » (figure 3), « Visualiser le nombre de traitements agronomiques et chimiques réalisés pour lutter contre les champignons nuisibles », « Visualiser tout le traitement de tous les parasites à l'échelle de la parcelle », etc. Ce type de requêtes permet à l'utilisateur d'effectuer différentes analyses, par exemple en comparant différentes stratégies de protection des cultures réalisées dans différentes parcelles, ou en déterminant quel organisme nuisible a causé le plus de dégâts afin de mieux contrôler la prochaine campagne.

Toutes ces requêtes peuvent être déclenchées à différentes échelles spatiales, temporels et thématiques à l'aide des hiérarchies de dimensions et des opérateurs OLAP. Par exemple, il est possible de formuler les questions suivantes : « Visualiser le nombre et le type de traitement effectué pour lutter contre les nuisibles par parcelle et exploitation » (figure 4), en appliquant un opérateur Drill-Down via le tableau croisé dynamique de l'outil OLAP.

Conclusion

Motivés par l'absence d'outil d'aide à la décision pour l'analyse historique des données de lutte intégrée, nous présentons un modèle OLAP, ainsi que son implémentation dans un entrepôt de données. Le système appliqué à la culture de l'olivier dans la région des Pouilles permet aux agriculteurs et aux décideurs de vérifier le respect des règles de défense des spécifications techniques de la LI.

Comme illustré dans l'article, le système OLAP a permis d'obtenir des informations spatio-temporelles détaillées relatives aux exploitations. Deux groupes de requêtes ont été développés et testés sur des données issues de quelques exploitations dans le cadre d'un projet de recherche. Le premier concerne les requêtes de conformité aux spécifications techniques, en enquêtant sur le respect des règles de défense à l'échelle inter et intra-exploitation. Le second comprend les requêtes de gestion permettant l'analyse des stratégies de protection des cultures selon différents paramètres, tels que le nombre de traitements pour la parcelle, le type d'intervention pour chaque ravageur, etc.

Dans une prochaine étape, nous voulons appliquer notre approche à un jeu de données provenant d'autres exploitations pour le valider dans un contexte applicatif réel, et par la même, tester les motivations des agriculteurs à mettre à disposition leurs données, pour certaines confidentielles, pour ce type d'analyses.

L'objectif à plus long terme est de mettre au point un modèle avancé OLAP qui à partir d'indices permettra de comparer les données de LI de différentes fermes ayant des conditions climatiques et agronomiques similaires. ■

L'auteur

Sandro BIMONTE

Université Clermont Auvergne, Irstea,
UR TSCF, Centre de Clermont-Ferrand,
9 avenue Blaise Pascal CS 20085,
F-63178 Aubière, France.

sandro.bimonte@irstea.fr

EN SAVOIR PLUS...

- **BIMONTE, S.**, 2016, Current approaches, challenges, and perspectives on spatial OLAP for agri-environmental analysis, *IJAEIS*, n° 7, p. 32-49.
- **EUROPEAN PARLIAMENT**, 2009, Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides., *Off. J. Eur. Union*, n° 309, p. 71-86.
- **FOUNTAS, S., CARLI, G., SØRENSEN, C.G., TSIROPOULOS, Z., CAVALARIS, C., VATSANIDOU, A., LIAKOS, B., CANAVARI, M., WIEBENSOHN, J., TISSERYE, B.**, 2015, Farm management information systems: Current situation and future perspectives, *Comput. Electron. Agric.*, n° 115, p. 40-50.
- **GRUNERT, K.G.**, 2005, Food quality and safety: consumer perception and demand, *Eur. Rev. Agric. Econ.*, n° 32, p. 369-391.
- **KIMBALL, R., ROSS, M.**, 2013, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, Wiley, New York.
- **NAVROZIDIS, E.I., VASARA, E., KARAMANLIDOU, G., SALPIGGIDIS, G.K., KOLIAIS, S.I.**, 2000, Biological control of *Bactrocera oleae* (Diptera: Tephritidae) using a Greek *Bacillus thuringiensis* isolate, *J. Econ. Entomol.*, n° 93, p. 1657-1661.
- **PROKOPY, R.J.**, 2003, Two decades of bottom-up, ecologically based pest management in a small commercial apple orchard in Massachusetts, *Agr. Ecosyst. Environ.*, n° 94, p. 299-309.
- **ZAZA, C., BIMONTE, S., GALLO, C., FACILONGO, N., LA SALA, P., CONTÒ, F.**, 2017, OLAP Analysis of Integrated Pest Management's Defense Rules: Application to Olive Crop in Apulia Region, *ICCSA (2)*, p. 31-44.