

# Comment réussir un projet de *crowdsourcing* d'observations spatialisées en agriculture ?

Par Léo PICHON

Ingénieur agronome, docteur en sciences agronomiques  
et maître de conférences à l'Institut Agro Montpellier

L'agriculture a engagé des transitions (agro-écologique, numérique, etc.) pour répondre aux enjeux complexes et parfois contradictoires auxquels elle est confrontée. La collecte collaborative d'observations spatialisées est un outil pertinent pour construire collectivement les nouvelles connaissances nécessaires à ces transitions. Ce type de projet est pourtant encore peu développé en agriculture. Notre objectif avec cet article est de proposer des éléments de compréhension des spécificités des projets de *crowdsourcing* d'observations spatialisées en agriculture, afin d'identifier les facteurs ou les caractéristiques à favoriser dans ces projets pour qu'ils se développent plus largement. Les structures spatiale et temporelle des phénomènes étudiés, le caractère asynchrone et hétérotopes des jeux de données obtenus et le fait que les participants soient des professionnels semblent être les caractéristiques principales à considérer pour favoriser la réussite de ce type de projet.

## INTRODUCTION

L'agriculture est actuellement confrontée à des enjeux complexes et parfois contradictoires. Pour y répondre, elle a engagé des transitions agro-écologique, climatique, énergétique ou encore numérique qui mobilisent l'ensemble des acteurs de l'agriculture, depuis les agriculteurs et leurs conseillers jusqu'aux chercheurs et aux décideurs politiques (Slimi *et al.*, 2021). Les projets qui mêlent au sein d'un même collectif des chercheurs et des acteurs ayant des points de vue variés semblent particulièrement adaptés à l'émergence de connaissances nouvelles permettant d'accompagner ces transitions (Moneyron *et al.*, 2017). Par exemple, le projet mené par Van Etten *et al.* (2019) est emblématique de cette approche qui mêle un collectif d'agriculteurs et de chercheurs pour générer de la connaissance sur des enjeux complexes. Ces auteurs ont proposé une méthode de recherche qui permet de comparer entre elles plusieurs variétés de semences en impliquant de nombreux agriculteurs qui expérimentent sur leurs fermes et partagent leurs observations. Elle offre des opportunités enthousiasmantes pour générer collectivement de la connaissance et accompagner l'agriculture dans ses transitions. Cette méthodologie, basée sur des informations collectées par la foule, s'appelle le *crowdsourcing*.

D'après Minet *et al.* (2017), il en existe trois types qui dépendent du type de contribution des participants. Tout d'abord, le *crowdsourcing* de connaissance, dans lequel les participants partagent leurs savoirs et leurs expériences, par exemple sous la forme d'une encyclopédie en ligne comme Wikipedia (<https://www.wikipedia.org/>) (Westerman,

2009). Ensuite, le *crowdsourcing* de tâches, dans lequel les participants réalisent, depuis leur ordinateur, des tâches unitaires de toutes natures. Ce type de *crowdsourcing* est beaucoup utilisé pour réaliser des tâches fastidieuses qui ne peuvent être faites que par des humains (Altwaijry *et al.*, 2016). Enfin, le *crowdsourcing* d'observations spatialisées, dans lequel les participants collectent des observations géolocalisées. La démocratisation des récepteurs GNSS, puis des *smartphones* a largement contribué à l'essor de ce type de *crowdsourcing* (Massad et Dalyot, 2018). Cette approche est particulièrement développée dans le domaine du suivi environnemental, par exemple pour caractériser et suivre la répartition spatiale d'une espèce d'intérêt (Sullivan *et al.*, 2009). En agriculture, les projets de *crowdsourcing* d'observations spatialisées présentent un intérêt particulier, car la réalisation d'observations spatialisées et datées est au cœur du métier d'agriculteur et d'agronome (Ebitu *et al.*, 2021), par exemple pour suivre des phénomènes biotiques (suivi de ravageurs...) ou abiotiques (suivi de la sécheresse...) à l'échelle régionale. Malgré le fort potentiel de ces projets (Van de Gevel *et al.*, 2020), leur développement reste encore limité en agriculture (Minet *et al.*, 2017).

L'objectif de cet article est de proposer des éléments de compréhension des spécificités des projets de *crowdsourcing* d'observations spatialisées en agriculture qui peuvent expliquer leur développement encore relativement faible au regard de l'intérêt qu'ils présentent pour l'agriculture. Pour cela, cet article énonce d'abord les facteurs génériques identifiés dans la littérature comme favorisant le succès des projets de *crowdsourcing*. Il décrit ensuite les spécificités des projets de *crowdsourcing* d'observations spatialisées en agriculture, afin d'identifier les facteurs ou les caractéristiques à favoriser dans ces projets pour qu'ils soient des succès et qu'ils se développent plus largement.

## LES FACTEURS DE SUCCÈS DES PROJETS DE *CROWDSOURCING*

Un projet de *crowdsourcing* est considéré comme un succès si les contributions des participants permettent de répondre à la question posée. Ce succès dépend principalement d'un nombre important de contributions, et de contributions de bonne qualité (Rechenberger *et al.*, 2015).

### La quantité de contributions

Les participants doivent être motivés pour garantir un nombre de contributions collectées important. La motivation des participants est certainement le facteur de participation qui a été le plus étudié dans la littérature. Batson *et al.* (2002) ont proposé un *framework* qui identifie quatre types de motivation chez les contributeurs.

#### *Motivation égoïste*

Les participants s'impliquent pour répondre à leur propre intérêt. Pour eux, la contribution peut alors donner lieu à une satisfaction directe (en s'amusant, etc.) ou indirecte (ex. : en obtenant de l'argent, de l'information, de la reconnaissance, etc.) (Kaufmann *et al.*, 2011). On parle respectivement de motivations égoïstes intrinsèque et extrinsèque.

#### *Motivation altruiste*

Les participants s'impliquent pour répondre à l'intérêt d'autres personnes pour lesquelles ils ressentent de l'empathie.

#### *Motivation collectiviste*

Les participants s'impliquent pour répondre à l'intérêt d'un groupe ou d'un collectif. Dans ce cas, le projet collectif est au centre, et c'est la volonté de créer du bien commun qui guide l'implication des participants.

## Motivation pour un principe

Les participants s'impliquent au nom de principes moraux et d'intérêts supérieurs comme la liberté ou la justice. Le projet collectif n'est qu'une conséquence non intentionnelle ou un simple moyen pour servir cet intérêt.

Dans chaque projet de *crowdsourcing* et pour chaque participant, ces quatre types de motivation sont présents, mais leurs proportions respectives dépendent du type de projet et de la manière dont les objectifs sont présentés (Land-Zandstra *et al.*, 2016). De plus, ces proportions peuvent évoluer au cours du temps. L'égoïsme est souvent le type de motivation principal lors d'une première contribution, puis l'altruisme et le collectivisme peuvent jouer un rôle plus important dans le choix des participants de continuer à contribuer (Rotman *et al.*, 2012).

## La qualité des contributions

- Bordogna *et al.* (2014) décrivent deux stratégies pour assurer des contributions de qualité dans les projets de *crowdsourcing*. La première stratégie, aussi appelée par certains auteurs "quality assurance" (Jonietz *et al.*, 2017), consiste à mettre en place des mécanismes qui se déroulent avant la collecte (stratégie *ex ante*) et qui garantissent que les données collectées seront de bonne qualité. Par exemple, ces mécanismes peuvent être la mise en place de formations pour les participants, la définition d'un vocabulaire commun ou de protocoles de collecte stricts (Criscuolo *et al.*, 2016). Kelling *et al.* (2019) insistent sur l'importance de collecter des informations, non seulement sur le phénomène étudié mais aussi sur le processus de collecte lui-même. D'après eux, la collecte de métadonnées associées aux contributions des participants (qui contribue, à quel endroit, à quelle date, selon quel protocole, etc.) n'améliore pas directement la qualité des données collectées, mais elle facilite leur compréhension et leur analyse *a posteriori*.
- La seconde stratégie, aussi appelée par certains auteurs "quality assesment" (Jonietz *et al.*, 2017), consiste à évaluer, une fois qu'elles ont été collectées (stratégie *ex post*), la qualité des données afin de pouvoir éventuellement éliminer celles qui sont de mauvaise qualité. Cette qualité peut être évaluée soit directement de façon experte (Criscuolo *et al.*, 2016), soit par la communauté en évaluant la réputation des participants (Jabeur *et al.*, 2018), soit, enfin, de façon automatique en étudiant les caractéristiques du jeu de données obtenu (Touya *et al.*, 2017) ou en comparant les contributions à des données de référence (Massad et Dalyot, 2018). Dans la littérature, ces trois approches sont fréquemment combinées (Fonte *et al.*, 2017).

## LES SPÉCIFICITÉS DES PROJETS DE CROWDSOURCING D'OBSERVATIONS SPATIALISÉES EN AGRICULTURE

Les projets de *crowdsourcing* d'observations spatialisées en agriculture possèdent des spécificités par rapport aux autres projets de *crowdsourcing* qui sont dues à la fois au type de contributions (observations spatialisées) et au contexte agricole dans lequel ces contributions sont collectées.

### Les phénomènes étudiés sont structurés dans le temps et dans l'espace

Les phénomènes étudiés répondent à des problématiques agricoles comme la présence de maladies, de ravageurs (Awuor *et al.*, 2019) ou de stress abiotiques (Dell'Acqua *et al.*, 2018). Ces phénomènes sont souvent dépendants de facteurs du milieu comme le sol (e.g. pédologie, topométrie) ou le climat (e.g. température, pluviométrie). Or, ces facteurs

du milieu sont généralement organisés dans l'espace (Colaço *et al.*, 2019). Ils peuvent également évoluer avec le temps et présenter une structure temporelle. Cela signifie que, pour la variable étudiée, deux observations collectées à des dates et sur des sites proches ont une probabilité plus grande de présenter des valeurs semblables que deux observations éloignées dans le temps et dans l'espace. De plus, les phénomènes étudiés sont généralement saisonniers et dynamiques. Naturellement, les participants qui cherchent à observer un phénomène ont donc tendance à réaliser leurs observations pendant la période durant laquelle celui-ci est observable. Il en résulte que le nombre d'observations au cours du temps varie très fortement et de manière cyclique dans ce type de jeux de données (Prudic *et al.*, 2017).

## Les jeux de données sont asynchrones et hétérotopes

Les projets de *crowdsourcing* d'observations spatialisées fonctionnent généralement avec une stratégie de collecte opportuniste (Kelling *et al.*, 2019). La conséquence majeure est que les observations sont collectées à des sites et à des dates qui sont déterminés par les participants uniquement en fonction de leurs contraintes ou de leurs motivations individuelles. Ces sites et ces dates ne sont pas connus à l'avance et ne sont pas nécessairement optimaux pour l'observation du phénomène d'intérêt. Les jeux de données qui en résultent sont donc :

- asynchrones, c'est-à-dire que les observations faites sur deux sites différents sont généralement faites à deux dates différentes ;
- hétérotopes, c'est-à-dire que les observations faites à deux dates différentes sont généralement situées sur deux sites différents.

Ces caractéristiques sont pratiquement inexistantes dans les approches expérimentales classiques en agriculture. L'expérimentation repose habituellement sur des dispositifs randomisés avec répétition, où les mesures sont systématiquement effectuées à la même date et répétées sur les mêmes placettes, ce qui simplifie grandement l'analyse des données. Le caractère asynchrone et hétérotope de ces jeux de données est donc fondamental, car il va conditionner les méthodes de traitement des données qui permettent d'inférer des connaissances. Dans ce contexte, la collecte de métadonnées sur les observations est un enjeu majeur, car elle permet d'enregistrer et de connaître *a posteriori* les sites et les dates de collecte afin de gérer ce caractère asynchrone et hétérotope.

## Les participants sont des professionnels

Les participants aux projets de *crowdsourcing* d'observations spatialisées contribuent à ces projets dans le cadre de leur activité professionnelle (Minet *et al.*, 2017). Ils sont donc majoritairement des acteurs du secteur agricole (agriculteurs, conseillers, etc.), et non pas des amateurs enthousiastes ou le grand public comme c'est le cas dans de nombreux projets de *crowdsourcing* (ex. : Prudic *et al.*, 2017). Cette caractéristique des participants influe sur leurs motivations, qui sont plutôt égoïstes et extrinsèques, c'est-à-dire qu'en contribuant, ils cherchent à répondre à un besoin en lien avec leur activité professionnelle – *e.g.* connaître les performances d'une variété (Van Etten *et al.*, 2019). Cette caractéristique des participants influe également sur leur connaissance des phénomènes étudiés. En tant que professionnels du secteur, ils contribuent à l'observation de phénomènes qu'ils connaissent bien. Il peut donc être fait l'hypothèse que leur niveau d'expertise est globalement plus élevé que dans d'autres projets de *crowdsourcing*, bien qu'à notre connaissance aucun article n'ait étudié cette hypothèse en détail. Enfin, les services d'extension sont un des acteurs clés dans les projets de *crowdsourcing* d'observations spatialisées. Comme pour l'adoption de n'importe quelle innovation en agriculture (ex. : de Souza Filho *et al.*, 2021), ils jouent un rôle important dans la contribution des agriculteurs aux projets de *crowdsourcing*, car ils bénéficient de leur confiance, ils sont généralement

bien implantés sur le territoire, et ils permettent d'assurer un lien entre les participants et l'organisation (Ryan *et al.*, 2018).

## CONCLUSION

Les projets de *crowdsourcing*, en particulier ceux qui permettent la collecte collaborative d'observations spatialisées, présentent un intérêt fort pour la construction de nouvelles connaissances et pour l'accompagnement de l'agriculture dans sa transition. Ces projets possèdent des spécificités qui influencent leur réussite et limitent aujourd'hui leur diffusion. Pour développer ce type de projets dans les années à venir, il est nécessaire d'explorer des approches et des méthodes favorisant leur réussite. Concernant le nombre de contributions, la mise au point d'approches actionnant les différents leviers de motivation des participants professionnels agricoles semble être un enjeu fort. Par ailleurs, les phénomènes étudiés ont des caractéristiques qui donnent lieu d'envisager des approches originales pour évaluer la qualité *ex post* des données collectées. Leurs structures spatiale et temporelle ainsi que leur saisonnalité permettent d'estimer des valeurs attendues pour les phénomènes étudiés. Cela offre la possibilité d'identifier des observations qui s'éloignent de ces valeurs attendues, soit parce que le phénomène étudié présente effectivement une anomalie locale, soit parce que l'observation considérée est une erreur ou une saisie malveillante. L'exploration de ce type d'approches devrait permettre le développement des projets de *crowdsourcing* d'observations spatialisées en agriculture.

## BIBLIOGRAPHIE

ALTWAIJRY H., TRULLS E., HAYS J., FUA P. & BELONGIE S. (2016), "Learning to match aerial images with deep attentive architectures", Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, pp. 3539-3547, <https://doi.org/10.1109/CVPR.2016.385>

AWUOR F., OTANGA S., KIMELI V., RAMBIM D. & ABUYA T. (2019), "E-Pest surveillance: Large scale crop pest surveillance and control", 2019 IST-Africa Week Conference, IST-Africa 2019 Nairobi, Kenya, pp. 1-8, <https://doi.org/10.23919/ISTAFRICA.2019.8764824>

BATSON C. D., AHMAD N. & TSANG J. A. (2002), "Four motives for community involvement", *Journal of Social Issues*, 58(3), pp. 429-445, <https://doi.org/10.1111/1540-4560.00269>

BORDOGNA G., CARRARA P., CRISCUOLO L., PEPE M. & RAMPINI A. (2014), "On predicting and improving the quality of volunteer geographic information projects", *International Journal of Digital Earth*, 9(2), pp. 134-155, <https://doi.org/10.1080/17538947.2014.976774>

COLAÇO A. F., MOLIN J. P., ROSELL-POLO J. R. & ESCOLÀ A. (2019), "Spatial variability in commercial orange groves. Part 2: Relating canopy geometry to soil attributes and historical yield", *Precision Agriculture*, 20(4), pp. 805-822, <https://doi.org/10.1007/s11119-018-9615-0>

CRISCUOLO L., CARRARA P., BORDOGNA G., PEPE M., ZUCCA F., SEPPI R. *et al.* (2016), "Handling quality in crowdsourced geographic information", in CAPINERI C. *et al.* (éd.), *European Handbook of Crowdsourced Geographic Information*, London, Ubiquity Press, pp. 57-74, <https://doi.org/10.5334/bax.e>

DE SOUZA FILHO H. M., VINHOLIS M. M. B., CARRER M. J. & BERNARDO R. (2021), "Determinants of adoption of integrated systems by cattle farmers in the State of Sao Paulo, Brazil", *Agroforestry Systems*, 95(1), pp. 103-117, <https://doi.org/10.1007/s10457-020-00565-8>

- DELL'ACQUA F., IANNELLI G. C., TORRES M. A. & MARTINA M. L. V. (2018), "A novel strategy for very-large-scale cash-crop mapping in the context of weather-related risk assessment, combining global satellite multispectral datasets, environmental constraints, and in situ acquisition of geospatial data", *Sensors*, 18(2), <https://doi.org/10.3390/s18020591>
- EBITU L., AVERY H., MOURAD K. A. & ENYETU J. (2021), "Citizen science for sustainable agriculture – A systematic literature review", *Land Use Policy*, 103, 105326, <https://doi.org/10.1016/j.landusepol.2021.105326>
- FONTE C. C., ANTONIOU V., BASTIN L., ESTIMA J., ARSANJANI J. J., LASO BAYAS J.-C. *et al.* (2017), "Assessing VGI Data Quality", in FOODY G. (éd.), *Mapping and the Citizen Sensor*, London, Ubiquity Press, pp. 137-163, <https://doi.org/https://doi.org/10.5334/bbf.g>
- JABEUR N., KARAM R. & MELCHIORI M. (2018), "A comprehensive reputation assessment framework for volunteered geographic information in crowdsensing applications", *Personal and Ubiquitous Computing*, 23, pp. 669-685, <https://doi.org/https://doi.org/10.1007/s00779-018-1122-9>
- JONIETZ D., ANTONIO V., SEE L. & ZIPF A. (2017), "Highlighting current trends in volunteered geographic information", *ISPRS International Journal of Geo-Information*, 6(7), pp. 1-8, <https://doi.org/10.3390/ijgi6070202>
- KAUFMANN N., SCHULZE T. & VEIT D. (2011), "More than fun and money. Worker motivation in crowdsourcing – A study on mechanical turk", *Proceedings of the Seventeenth Americas Conference on Information Systems*, pp. 1-11, <https://doi.org/10.1145/1979742.1979593>
- KELLING S., JOHNSTON A., BONN A., FINK D., RUIZ-GUTIERREZ V., BONNEY R. *et al.* (2019), "Using semistructured surveys to improve citizen science data for monitoring biodiversity", *BioScience*, 69(3), pp. 170-179, <https://doi.org/10.1093/biosci/biz010>
- LAND-ZANDSTRA A. M., DEVILEE J. L. A., SNIK F., BUURMEIJER F. & VAN DEN BROEK J. M. (2016), "Citizen science on a smartphone: Participants' motivations and learning. Public understanding of science", 25(1), pp. 45-60, <https://doi.org/10.1177/0963662515602406>
- MASSAD I. & DALYOT S. (2018), "Towards the crowdsourcing of massive smartphone assisted-GPS sensor ground observations for the production of digital terrain models", *Sensors*, 18(3), 898, <https://doi.org/10.3390/s18030898>
- MINET J., CURNEL Y., GOBIN A., GOFFART J.-P. P., MÉLARD F., TYCHON B. *et al.* (2017), "Crowdsourcing for agricultural applications: A review of uses and opportunities for a farmsourcing approach", *Computers and Electronics in Agriculture*, 142(A), pp. 126-138, <https://doi.org/10.1016/j.compag.2017.08.026>
- MONEYRON A., LALLEMAND J. F., SCHMITT C., PERRIN M., SOUSTRE-GACOUGNOLLE I. & MASSON J. E. (2017), "Linking the knowledge and reasoning of dissenting actors fosters a bottom-up design of agroecological viticulture", *Agronomy for Sustainable Development*, 37(5), <https://doi.org/10.1007/s13593-017-0449-3>
- PRUDIC K. L., MCFARLAND K. P., OLIVER J. C., HUTCHINSON R. A., LONG E. C., KERR J. T. & LARRIVÉE M. (2017), "eButterfly: Leveraging massive online citizen science for butterfly conservation", *Insects*, 8(2), 53, <https://doi.org/10.3390/insects8020053>
- RECHENBERGER T., JUNG V. M. E., SCHMIDT N. & ROSENKRANZ C. (2015), "Utilizing the crowd – A literature review on factors influencing crowdsourcing initiative success", PACIS 2015 *Proceedings*, p. 250, <http://aisel.aisnet.org/pacis2015%0Ahttp://aisel.aisnet.org/pacis2015/250>

ROTMAN D., PREECE J., HAMMOCK J., PROCITA K., HANSEN D., PARR C. *et al.* (2012), "Dynamic changes in motivation in collaborative citizen-science projects", *Proceedings of the ACM 2012 Conference on Computer-Supported Cooperative Work*, New York, NY, USA, Association for Computing Machinery, pp. 217-226, <https://doi.org/10.1145/2145204.2145238>

RYAN S. F., ADAMSON N. L., AKTIPIS A., ANDERSEN L. K., AUSTIN R., BARNES L. *et al.* (2018), "The role of citizen science in addressing grand challenges in food and agriculture research", *Proceedings of the Royal Society B*, vol. 285, <https://doi.org/http://dx.doi.org/10.1098/rspb.2018.1977>

SLIMI C., PROST M., CERF M. & PROST L. (2021), "Exchanges among farmers' collectives in support of sustainable agriculture: From review to reconceptualization", *Journal of Rural Studies*, 83, pp. 268-278, <https://doi.org/https://doi.org/10.1016/j.jrurstud.2021.01.019>

SULLIVAN B. L., WOOD C. L., ILIFF M. J., BONNEY R. E., FINK D. & KELLING S. (2009), "eBird: A citizen-based bird observation network in the biological sciences", *Biological Conservation*, 142(10), pp. 2282-2292, <https://doi.org/10.1016/j.biocon.2009.05.006>

TOUYA G., ANTONIOU V., OLTEANU-RAIMOND A.-M. M. & VAN DAMME M.-D. D. (2017), "Assessing crowdsourced POI quality: Combining methods based on reference data, history, and spatial relations", *ISPRS International Journal of Geo-Information*, 6(3), 80, <https://doi.org/10.3390/ijgi6030080>

VAN DE GEVEL J., VAN ETTEN J. & DETERDING S. (2020), "Citizen science breathes new life into participatory agricultural research. A review", *Agronomy for Sustainable Development*, 40(5), <https://doi.org/10.1007/s13593-020-00636-1>

VAN ETTEN J., BEZA E., CALDERER L., VAN DUIJVENDIJK K., FADDA C., FANTAHUN B. *et al.* (2019), "First experiences with a novel farmer citizen science approach: Crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (Tricot)", *Experimental Agriculture*, 55(S1), pp. 275-296, <https://doi.org/10.1017/S0014479716000739>

WESTERMAN W. (2009), "Epistemology, the sociology of knowledge, and the wikipedia userbox controversy", *Folklore and the Internet*, pp. 123-158, <https://doi.org/https://doi.org/10.2307/j.ctt4cgrx5.9>