AGRICULTURAL UNIVERSITY OF ATHENS

Dimitrios Kremmydas

Integrated Mathematical Programming and Agent Based Model for Policy Analysis of Greek Agriculture

PhD Dissertation submitted
to the Department of Agricultural Economics & Rural Development

January 2019

Athens, Greece

Supervising committee

Athanasios Kampas, Associate Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

Stelios Rozakis, Associate Professor, School of Environmental Engineering, Technical University of Crete, Greece

Ioannis Athanasiadis, Assistant Professor with the Information Technology group at Wageningen University, Netherlands

Examination committee

Athanasios Kampas, Associate Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

Stelios Rozakis, Associate Professor, School of Environmental Engineering, Technical University of Crete, Greece

Ioannis Athanasiadis, Assistant Professor with the Information Technology group at Wageningen University, Netherlands

Kontstantinos Tsiboukas, Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

George Vlahos, Assistant Professor, Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

Thomas Bournaris, Assistant Professor, Faculty of Agriculture, Forestry and Natural Environment, Aristotle University of Thessaloniki, Greece

Argyris Kanellopoulos, Assistant Professor Department of Social Sciences, Wageningen University, Netherlands

Αφιερωμένο στην γυναίκα μου Σταυρούλα

και στα παιδιά μας Νίκο, Χρήστο και Ουρανία

με τους οποίους, αν και ο ερευνητικός μου βίος είναι πιο δύσκολος, η ζωή μου γεμάτη και ευτυχισμένη

Dedicated to my wife Stavroula

and our children Nikos, Christos and Ourania

with whom my research is slower but my life is happier and full

«ὅτι μὲν οὖν ἐστὶν ἡ ἀρετὴ ἡ ἠθικὴ μεσότης, καὶ πῶς, καὶ ὅτι μεσότης δύο κακιῶν, τῆς μὲν καθ’ ὑπερβολὴν τῆς δὲ κατ’ ἔλλειψιν, καὶ ὅτι τοιαύτη ἐστὶ διὰ τὸ στοχαστικὴ τοῦ μέσου εἶναι τοῦ ἐν τοῖς πάθεσι καὶ ἐν ταῖς πράξεσιν, ἱκανῶς εἴρηται».

Αριστοτέλης – Ηθικά Νικομάχεια 1109a[1]

«Enough has now been said to show that moral virtue is a mean, and in what sense this is so, namely that it is a mean between two vices, one of excess and the other of defect; and that it is such a mean because it aims at hitting the middle point in feelings and in actions.»

Aristotle, Nicomachean Ethics, 1109a[1]

Ευχαριστίες

Σύμφωνα με την λαϊκή σοφία, *η αρχή είναι το ήμισυ του παντός*. Χωρίς τον αρχικό επιβλέποντα της παρούσας διατριβής, Στέλιο Ροζάκη, αυτή δεν θα είχε ξεκινήσει. Αυτοδικαίως λοιπόν του αναλογούν οι μισές και παραπάνω ευχαριστίες. Αν δεν τον είχα συναντήσει, εξαιτίας της κακής μου συνήθειας να ασχολούμαι μόνο με πράγματα που μου αρέσουν, χωρίς να υπολογίζω το κόστος ευκαιρίας, δεν θα είχα ποτέ ξεκινήσει. Το μεράκι που έχει για την ακαδημαϊκή έρευνα, μου άνοιξε τον δρόμο. Επιπλέον χωρίς την καθοδήγηση, τις υποδείξεις και την υποστήριξη του, η διατριβή αυτή δεν θα είχε ολοκληρωθεί.

Θέλω επίσης να ευχαριστήσω τον κύριο Αθανασιο Καμπά, ο οποίος από την στιγμή που ανέλαβε επιβλέπων της διατριβής μου, ανάλωσε αρκετόν από τον χρόνο του σε συζητήσεις για σχετικά θέματα και είχε ουσιαστική συμβολή στην ολοκλήρωση της

Ευχαριστίες θέλω να απευθύνω και στον κύριο Ιωάννη Αθανασιάδη, ο οποίος αν και δεν με γνώριζε προσωπικά, δέχτηκε εξαρχής να είναι μέλος της Τριμελούς Επιτροπής. Τόσο η τακτική επικοινωνία μαζί του όσο και η 4-ήμερη επίσκεψη μου το 2018 στο Πανεπιστήμιο του Wageningen ήταν καθοριστικής σημασίας για το τελικό αποτέλεσμα αυτής της διατριβής.

Ευχαριστίες επίσης απευθύνω και στον κύριο Κωνσταντίνο Τσιμπούκα, διευθυντή του εργαστηρίου Διοίκησης Γεωργικών Επιχειρήσεων και Εκμεταλλεύσεων, στο οποίο οργανικά ανήκω, του οποίου οι συμβουλές και η καθοδήγηση ήταν πολύ σημαντικές για την συμπερίληψη του κλάδου της αιγοπροβατοτροφίας στο μοντέλο της διατριβής. Η επαφή μου μαζί του, μου έδωσε την ευκαιρία να πλησιάσω εγγύτερα το αντικείμενο της έρευνας μου, και να ξεβολευτώ κάπως από την πολυθρόνα της στείρας ακαδημαϊκής προσέγγισης της γεωργικής δραστηριότητας (armchair economics).

Acknowledgments

I own gratitude to Dr Stelios Rozakis without whom this Dissertation would have never begun. He was all way along the PhD path, supporting and advising me.

I also want to thank Dr Thanasis Kampas, Dr Ioannis Athanasiadis and Professor Kostas Tsiboukas for helping me accomplish various aspects of this dissertation.

Table of Contents

Εκτεταμένη Περίληψη

Η μαθηματική/οικονομική προτυποποίηση για την αξιολόγηση αγροτικής πολιτικής θα μπορούσε να διακριθεί σε δύο μεγάλες κατηγορίες: Πρότυπα υποδείγματα (μοντέλα) τα οποία είναι εστιασμένα στο μακρο-επίπεδο (π.χ. μοντέλα μερικής ή γενικής ισορροπίας) και πρότυπα υποδείγματα τα οποία προσομοιώνουν την συμπεριφορά της γεωργικής εκμετάλλευσης στο μικροοικονομικό επίπεδο και απεικονίζουν τον τομέα με βάση τη συμπεριφορά των επιμέρους ανεξάρτητων μονάδων λήψης απόφασης (εκ των κάτω - bottom-up προσέγγιση). Η δεύτερη προσέγγιση είναι κατάλληλη για να αποτελέσει την βάση ανάπτυξης ενός υποδείγματος δρώντων υποκειμένων[1] (agent based model), μια μεθοδολογική καινοτομία που εμπλουτίζει και διευρύνει τις δυνατότητες του υποδείγματος το οποίο θα χρησιμοποιηθεί για ανάλυση πολιτικής.

Σε ευρωπαϊκό επίπεδο, περιφερειακά και τομεακά μοντέλα που υιοθετούν τη δεύτερη προσέγγιση και στηρίζονται σε υποδείγματα εκμεταλλεύσεων (Υ.ΕΚΜ) έχουν κατά καιρούς χρησιμοποιηθεί για να υπολογιστούν οι επιπτώσεις των διάφορων σεναρίων της εξέλιξης της Κοινής Αγροτικής Πολιτικής (ΚΑΠ), για παράδειγμα Ackril et al., 2001, Guinde et al., 2005, Wilson et al., 2003, Britz et al., 2012, Van Ittersum et al., 2008, Galko et. al., 2011, Reidsma et al., 2018. Για την Ελλάδα οι αναλύσεις εστιάζουν κυρίως σε καλλιέργειες των οποίων τα καθεστώτα άλλαξαν δραστικά όπως στον καπνό και το βαμβάκι. Ενδεικτικά αναφέρουμε μοντέλα με κλασσικό γραμμικό προγραμματισμό (Mattas et al., 2006), πολυκριτηριακές

[1] Είναι σύνηθες στην Ελληνική βιβλιογραφία ο όρος "Agent based" να μεταφράζεται ως «Πράκτορες Λογισμικού». Για παράδειγμα ο όρος «πράκτορες» στην μηχανή αναζήτησης ελληνικών ψηφιακών βιβλιοθηκών openarchives.gr επιστρέφει 135 σχετικά με "agent based modeling" αποτελέσματα. Επίσης η αναζήτηση στην " Βάση Τηλεπικοινωνιακών Όρων" της Ελληνικής Εταιρείας Ορολογίας (ΕΛΕΤΟ) επιστρέφει στην λέξη "πράκτορας" την λέξη "agent".

Ωστόσο, αν και για την Επιστήμη της Πληροφορικής ο όρος αποδίδει το σημαινόμενο σωστά, για τις κοινωνικές επιστήμες ο όρος καθίσταται αδόκιμος.

Πιο συγκεκριμένα, από το λεξικό merriam-webster (http://www.merriam-webster.com/dictionary/agent) καταγράφονται τα εξής σημαινόμενα της λέξης ***Agent****: «****1:*** *one that acts or exerts power* ***2a****: something that produces or is capable of producing an effect : an active or efficient cause [...]* ***4:****one who is authorized to act for or in the place of another: as a (a): a representative, emissary, or official of a government <crown agent> <federal agent> (b): one engaged in undercover activities (as espionage) : spy <secret agent> ».*

Όπως διαβάζουμε στο λεξικό Μπαμπινιώτη, *«****Πράκτορας****: ο συμβαλλόμενος στη σύμβαση πρακτορείας ο οποίος αναλαμβάνει έναντι αμοιβής την επιμέλεια και διεκπεραίωση του αντισυμβαλλόμενου και στο πλαίσιο της οποίας διενεργεί κάθε απαραίτητη νομική ή υλική πράξη στο όνομα και για λογαριασμό του πρακτορευόμενου».* Αντιστοιχεί δηλαδή η λέξη "πράκτορας" στην έννοια #4 της αγγλικής λέξης "agent".

Έτσι, στο επιστημονικό πεδίο της πληροφορικής, οι "πράκτορες" (λογισμικού) αναλαμβάνουν να κάνουν αυτόνομα μία εργασία σαν "αντιπρόσωποι" ενός ανθρώπου. Είναι λοιπόν δόκιμος ο όρος "πράκτορας" καθώς εννοιολογικά αντιστοιχεί στην έννοια #4 του λεξικού merriam-webster.

Ωστόσο στο πεδίο της οικονομικής επιστήμης ο όρος αυτός υπονοεί το σημαινόμενο της ερμηνείας #1 και #2a του αγγλικού λεξικού, " *one that acts or exerts power /: something that produces or is capable of producing an effec* ". Όμως όπως είδαμε η ελληνική λέξη «Πράκτορας», ανεξαρτήτως νοηματικού πλαισίου, δεν έχει την έννοια των #1 και #2a αλλά του #4.

Έτσι, στις κοινωνικές επιστήμες, η απόδοση του αγγλικού όρου "agent" είναι πιο δόκιμο να γίνει ως «υποκείμενο» ή «δράστης» ή «μονάδα δράσης» ή «αυτουργός».

μεθόδους (Manos et al. 2009), θετικά υποδείγματα που ενσωματώνουν φθίνουσα ζήτηση (Rozakis et al., 2008) ή αύξουσα συνάρτηση κόστους με Θετικό Μαθηματικό Προγραμματισμό (Petsakos and Rozakis, 2009) στην αντικειμενική συνάρτηση. Οι πολυκριτηριακές μέθοδοι με αυτόματο υπολογισμό της συνάρτησης χρησιμότητας του παραγωγού (Amador et al. 1999) και ο Θετικός Μαθηματικός Προγραμματισμός έχουν κυριαρχήσει στη βιβλιογραφία για την ανάλυση των επιπτώσεων της ΚΑΠ καθώς αυξάνουν την εγκυρότητα των Υ.ΕΚΜ σε σχέση με τις κλασσικές μεθόδους.

Με δεδομένο, πρώτον ότι η ΚΑΠ είναι μία πολύπλοκη πολιτική με διαφορετικά αποτελέσματα για ανταγωνιστικές ομάδες, και δεύτερον, τις ευρείες διαβουλεύσεις που λαμβάνουν χώρα σε εθνικό και ευρωπαϊκό επίπεδό σε σχέση με αλλαγές που, από την κυοφορία μέχρι και την έγκριση τους από τα αρμόδια όργανα της ΕΕ, αποτελούν αντικείμενο διαπραγμάτευσης, η τεχνοκρατική ανάλυση των εναλλακτικών σεναρίων πολιτικής καθίσταται πολύτιμη για όλους τους εμπλεκόμενους φορείς, προσφέροντας όχι μόνο αποτελεσματικότερο διάλογο αλλά και τεκμηριωμένη άποψη στα ευρωπαϊκά όργανα και δεξαμενές σκέψης (Pezaros, 2000). Παρόλα αυτά στην χώρα μας, οι παρεμβάσεις των θεσμικών φορέων στην διαμόρφωση της εξέλιξης της Κοινής Αγροτικής Πολιτικής (ΚΑΠ) δεν φαίνεται να στηρίζονται στα παραπάνω επιστημονικά εργαλεία (Klonaris and Vlahos, 2012).

Η οικονομική προτυποποίηση της ελληνικής γεωργικής παραγωγής έχει επιχειρηθεί επανειλημμένα στο παρελθόν με σκοπό προβλέψεις και αναλύσεις επιπτώσεων πολιτικής (ex ante ή ex post αντίστοιχα): Mantziaris et al. (2017), Giannakis et al. (2014), Manos et al. (2013), Sintori (2012), Petsakos (2012), Efstratoglou et al. (2011), Rozakis (2010), Rezitis and Stavropoulos (2010), Petsakos et al. (2009), Manos et al. (2009), Rozakis et al. (2008), Katranidis (2002). Το πλήθος αυτών των απλά ενδεικτικών αναφορών, αποδεικνύει ότι δεν υπάρχει έλλειμμα ούτε στην ακαδημαϊκή έρευνα ούτε στο επιστημονικό προσωπικό της χώρας σε σχέση με την διερεύνηση των συνεπειών των αλλαγών της αγροτικής πολιτικής στην Ελληνική γεωργία. Ωστόσο οι προσπάθειες αυτές δεν φαίνεται να μπορούν να διεισδύσουν παρά σε περιορισμένο βαθμό στην διαδικασία αναθεώρησης της ΚΑΠ στο θεσμικό και πολιτικό επίπεδο. Πιθανοί λόγοι αυτής της αδυναμίας είναι, αφενός η μικρή κλαδική και γεωγραφική κάλυψη των εν λόγω υποδειγμάτων και αφετέρου η αδυναμία χειρισμού τους από μη ειδικούς.

Η ύπαρξη ενός, κατά το μέτρο του εφικτού, ολοκληρωμένου υποδείγματος της ελληνικής γεωργίας, εύχρηστου και διαδραστικού, θα μπορούσε να γεφυρώσει το χάσμα μεταξύ ακαδημαϊκής έρευνας και των θεσμικών φορέων σε σχέση με τον σχεδιασμό και την εφαρμογή της αγροτικής πολιτικής. Ένα τέτοιο υπόδειγμα θα πρέπει να κινείται σε δύο άξονες: Αφενός να συμπεριλαμβάνει ένα σημαντικό κομμάτι της γεωργικής δραστηριότητας, αξιοποιώντας την υπάρχουσα εμπειρία και τα επιμέρους τομεακά υποδείγματα (οριζόντιος άξονας). Αφετέρου, σε ακολουθία με τις τρέχουσες επιστημονικές τάσεις, να εξετάζει τα διαφορετικά επίπεδα (φυσικό, βιολογικό, οικονομικό, κοινωνικό) στα οποία επιδρά η αγροτική πολιτική (κάθετος άξονας).

Η βασική λοιπόν συμβολή της παρούσας διδακτορικής διατριβής είναι η ανάπτυξη ενός ολοκληρωμένου υποδείγματος της ελληνικής γεωργίας το οποίο μπορεί να χρησιμοποιηθεί για την εκ των προτέρων (ex ante) αξιολόγηση σεναρίων αγροτικής πολιτικής. Αυτό κατέστη εφικτό δίνοντας πρωτότυπες λύσεις σε ζητήματα σύνδεσης των πρωτογενών δεδομένων του Δικτύου Γεωργικής Λογιστικής (ΔΙΓΕΛΠ, FADN) με ένα επίσης πρωτότυπο μοντέλο μαθηματικού προγραμματισμού. Επιπλέον η μοντελοποίηση και των αροτραίων καλλιεργειών και της αιγοπροβατοτροφίας σε κλίμακα που να καλύπτει το σύνολο της χώρας είναι εξίσου καινοφανής. Τέλος η προσέγγιση των μοντέλων δρώντων υποκειμένων, μια μέθοδος που εισήχθη την τελευταία εικοσαετία στην θεματική της Αγροτικής οικονομίας προσθέτει στις κλασσικές μεθόδους τη διάσταση της επικοινωνίας και διάδρασης μεταξύ των μονάδων λήψης απόφασης, και προσφέρει συμπληρωματικές πληροφορίες, όπως οι τιμές των ενοικίων γης μέσω προσομοίωσης αγοράς ή οι επιπτώσεις μέτρων πολιτικής σε εστιασμένο χωρικά επίπεδο.

Η διατριβή συγκροτείται από τα εξής επιμέρους στοιχεία: (α) συλλογή των δεδομένων από τις διάφορες πρωτογενείς πηγές και διαχείριση τους, (β) ανάπτυξη διεπαφής για μεταφορά των δεδομένων στο υπόδειγμα, (γ) κατασκευή και συγκρότηση του βασικού Υ.ΕΚΜ με βάση τον γραμμικό προγραμματισμό, (δ) παραδειγματική εφαρμογή του για ανάλυση πολιτικής, (ε) επέκταση του υποδείγματος με μετατροπή των μεμονωμένων μονάδων λήψης απόφασης σε εκμεταλλεύσεις/δρώντα υποκείμενα που αλληλεπιδρούν στο χώρο και στο χρόνο, (στ) εφαρμογή του βελτιωμένου υποδείγματος για ανάλυση πολιτικής και συγκριτική ανάλυση και (ζ) δημιουργία συστήματος λήψης απόφασης που θα διευκολύνει τους χρήστες να αλληλεπιδρούν με τα υποδείγματα των προηγούμενων διαδικασιών. Τα κεφάλαια που ακολουθούν είναι δημοσιευμένες ή υπό δημοσίευση εργασίες που περιγράφονται συνοπτικά στη συνέχεια και αντιστοιχούν στα προαναφερόμενα στάδια. Το πλήρες κείμενο τους παρατίθεται κατά περίπτωση στο κυρίως κείμενο και στα παραρτήματα της διατριβής:

- Τα δεδομένα που χρησιμοποιούνται σε ένα Υ.ΕΚΜ για την αξιολόγηση πολιτικής επηρεάζουν σημαντικά τα αποτελέσματα του υποδείγματος και συνεπώς επηρεάζουν ουσιαστικά την αξιοπιστία του. Η δημοσίευση **“Data warehouse technology for agricultural policy data: a Greek case study”** εξετάζει την φύση των δεδομένων που είναι δυνητικά χρήσιμα για την αξιολόγηση της Αγροτικής Πολιτικής και παρουσιάζει μία τυπική εφαρμογή όπου συνδυάζονται δεδομένα από διαφορετικές πηγές (αποδόσεις καλλιεργειών από τις Ετήσιες Έρευνες Γεωργίας της Ελληνικής Στατιστικής Υπηρεσίας και μετεωρολογικές μετρήσεις από το δίκτυο σταθμών του Εθνικού Αστεροσκοπείου Αθηνών).
- Για τις ανάγκες του Ελληνικού υποδείγματος αντλήσαμε από το υπουργείο Αγροτικής Ανάπτυξης δεδομένα της έρευνας του Δικτύου Γεωργικής Λογιστικής Πληροφόρησης (ΔΙΓΕΛΠ / FADN) για τα έτη 2011, 2012 και 2013. Γι αυτό το λόγο το μοντέλο της ελληνικής Γεωργίας που κατασκευάστηκε στα πλαίσια της διδακτορικής διατριβής, έχει την δυνατότητα να χρησιμοποιεί στοιχεία οποιασδήποτε χρονιάς του ΔΙΓΕΛΠ μέσα από δομημένες διαδικασίες. Ο μετασχηματισμός τους σε μορφή τέτοια που να είναι χρησιμοποιήσιμη από το ολοκληρωμένο υπόδειγμα ήταν αρκετά περίπλοκος και απαιτητικός γι αυτό και περιγράφεται ξεχωριστά στο κείμενο **«fadnUtils, An R package for working with FADN data»** του Παραρτήματος.
- Το ολοκληρωμένο υπόδειγμα της Ελληνικής Γεωργίας λειτουργεί σε επίπεδο εκμετάλλευσης (farm model) και περιλαμβάνει στην τρέχουσα έκδοση του τις επιχειρηματικές εκμεταλλεύσεις της φυτικής παραγωγής και την αντίστοιχη αιγοπροβατοτροφία. Οι συμπεριληφθείσες δραστηριότητες περιλαμβάνουν περίπου το 80% των εκτάσεων και το 60% της αξίας παραγωγής σε επίπεδο χώρας.

Στο υπόδειγμα δόθηκε η κωδική ονομασία GREFAM (Greek Representative Farm Model). Η αρχιτεκτονική του υποδείγματος περιγράφεται με λεπτομέρεια στο εγχειρίδιο χρήσης του, το οποίο παρατίθεται στο **«GREFAM Model Reference Manual»** του Παραρτήματος

- Χρησιμοποιήσαμε το μοντέλο GREFAM για να εκτιμήσουμε τις επιπτώσεις διαφορετικών σεναρίων περιφερειοποίησης σε σχέση με το σενάριο της πλήρους εξίσωσης της μοναδιαίας αξίας της ενιαίας ενίσχυσης σε όλη την χώρα. Δείχνουμε ότι με την περιφερειοποίηση μπορούμε να χειριστούμε καλύτερα τους αντικρουόμενους στόχους της ΚΑΠ σε σχέση με το σενάριο της πλήρους εξίσωσης. Το κείμενο (σε διαδικασία κρίσης σε περιοδικό) **«Why to regionalize CAP payments: A farm modeling approach»** πραγματεύεται την παραπάνω προσπάθεια και είναι μία επίδειξη της χρησιμότητας του μοντέλου για την αξιολόγηση ενός επίκαιρου και ρεαλιστικού ερωτήματος αξιολόγησης πολιτικής.
- Στην συνέχεια εξετάσαμε την δυνατότητα συμπλήρωσης της συμβατικής προσέγγισης του μοντέλου φάρμας Υ.ΕΚΜ με χρήση της προσέγγισης *υποδειγμάτων με δρώντα υποκείμενα*. Η εκτεταμένη βιβλιογραφική ανασκόπηση που διενεργήθηκε με στόχο την κριτική καταγραφή της χρήσης τέτοιων υποδειγμάτων στην αξιολόγηση Αγροτικής Πολιτικής βρίσκεται στη δημοσίευση **«A review of Agent Based Models for Policy Evaluation»**.
- Με βάση τα συμπεράσματα από την βιβλιογραφική ανασκόπηση προχωρήσαμε, αρχικά στην δημιουργία ενός μοντέλου δρώντων υποκειμένων/εκμεταλλεύσεων με τη χρήση αντικειμενοστρεφούς προγραμματισμού (Java), όπως περιγράφεται στο κείμενο εργασίας «**Dealing with farm heterogeneity on modeling agricultural policy: an Agent Based Modeling Approach»** στο Παράρτημα. Στην συνέχεια χρησιμοποιήσαμε ένα αντίστοιχο υπόδειγμα, με διαφορετικό τρόπο υλοποίησης, για να διερευνήσουμε τις επιπτώσεις διαφορετικών σεναρίων περιφερειοποίησης σε διαφορετικά παραγωγικά συστήματα ανά την χώρα, όπως περιγράφεται από το κείμενο εργασίας **«Extending a Farm Model by means of Agent Based Model for Evaluating CAP Regionalization Scenarios»**.
- Τέλος, λαμβάνοντας υπόψη το γεγονός πως η απευθείας χρήση υποδειγμάτων δεν είναι δυνατή από τα ενδιαφερόμενα μέρη που επηρεάζονται άμεσα ή/και εμπλέκονται στην διαμόρφωση της αγροτικής πολιτικής, προχωρήσαμε στην δημιουργία ενός Συστήματος Λήψης Απόφασης. Το σύστημα αυτό περιγράφεται στην δημοσίευση **«CAP2020 regionalization design: A Decision Support System»** και η σχεδιάση του επιτρέπει να ενσωματώσει ενδεχόμενες αναβαθμίσεις του GREFAM. Συγκεκριμένα η επέκταση σε όλες τις εκμεταλλεύσεις του δείγματος με σύστημα πρακτόρων θα δώσει στο σύστημα αυτό ακόμα περισσότερες δυνατότητες.

Παρακάτω δίνουμε περισσότερες λεπτομέρειες για τις επιμέρους δημοσιεύσεις που συνθέτουν την διδακτορική διατριβή.

Η τεχνολογία Αποθηκών Δεδομένων στα πλαίσια της Αγροτικής Πολιτικής: Μία μελέτη περίπτωσης για την Ελλάδα (Data warehouse technology for agricultural policy data: a Greek case study)

Τα δεδομένα τα οποία δύνανται να χρησιμοποιηθούν στην αξιολόγηση της αγροτικής πολιτικής παρουσιάζουν τα παρακάτω ιδιαίτερα χαρακτηριστικά:

1. Υπάρχουν πολλές ασύνδετες πηγές πληροφοριών, π.χ. διεθνείς ή εθνικές στατιστικές υπηρεσίες, βάσεις δεδομένων διαχειριστικών φορέων, επιτόπιες έρευνες ή παρελθόντα στοιχεία από πανεπιστήμια κ.λπ., καμία από τις οποίες δεν μπορεί να αγνοηθεί, καθώς αυτά τα δεδομένα είναι στην πραγματικότητα ένας σπάνιος πόρος. Επιπλέον, αυτές οι πηγές έχουν αποθηκευμένα τα δεδομένα τους σε διαφορετικές μορφές ή / και δομές βάσεων δεδομένων.

2. Τα δεδομένα που μας ενδιαφέρουν, αποτυπώνουν πολλές διαστάσεις καθότι συχνά οι υπεύθυνοι χάραξης πολιτικής έχουν πολλαπλούς στόχους. Αυτές οι διαστάσεις μπορούν να ταξινομηθούν ως εξής:

(i) Η βιοφυσική διάσταση. Πχ. δεδομένα μετεωρολογικά/κλιματικά, εδαφολογικά κ.λπ. Για παράδειγμα, στην περίπτωση της διερεύνησης της επίδρασης μιας πολιτικής στη βιοποικιλότητα ή στη διάβρωση του εδάφους, τέτοια δεδομένα μας είναι απαραίτητα

(ii) Η τεχνική διάσταση. Π.χ. εισροές, πρακτικές διαχείρισης, πιθανές αποδόσεις κ.λπ. Οι τεχνικές σχέσεις (δηλαδή ποιες εισροές χρησιμοποιούνται για μια συγκεκριμένη καλλιέργεια σε έναν συγκεκριμένο τομέα) είναι πολύ σημαντικός παράγοντας για τις αποφάσεις παραγωγής των γεωργών και επομένως σχετίζονται άμεσα με τα μοντέλα που στοχεύουν στην αξιολόγηση της αγροτικής πολιτικής.

(iii) Η οικονομική διάσταση. Π.χ. τιμές εισροών και εκροών, παραγωγή, εισόδημα κ.λπ.

(iv) Η κοινωνική διάσταση. Π.χ. πληθυσμός ανά κοινότητα, πυραμίδα ηλικίας κ.λπ. Συχνά η αγροτική πολιτική είτε στοχεύει σε έμμεσες αλλαγές στη στην διάσταση αυτή είτε επηρεάζεται άμεσα από αυτή (π.χ. συρρίκνωση / διατήρηση πληθυσμού, ανάπτυξη δεξιοτήτων κ.λπ.).

3. Είναι επιθυμητό τα δεδομένα να περιέχουν χρονική και χωρική πληροφορία στην μεγαλύτερη δυνατή ανάλυση, π.χ. σε επίπεδο αγροτεμαχίου, ανά ημέρα, κλπ. Ωστόσο συχνά, είτε η χρονική διάσταση δεν καταγράφεται είτε η χωρική διάσταση δίνεται σε επίπεδο διοικητικής περιφέρειας. Έτσι οι δύο αυτές διαστάσεις αποτελούν περιοριστικό παράγοντα για την επιθυμητή λεπτομέρεια.

4. Είναι ιεραρχικά δεδομένα. Για παράδειγμα, η χωρική / διοικητική διάσταση περιλαμβάνει την κοινότητα στο χαμηλότερο επίπεδο και τη χώρα στην κορυφή, η παραγωγική κατεύθυνση αν και είναι συγκεκριμένη (π.χ., παραγωγή γάλακτος από αίγες), συνήθως αθροίζεται σε πιο γενικές κατηγορίες (παραγωγή γάλακτος από αιγοπροβατοτροφία), κλπ. Αυτή η λογική ιεραρχία είναι σημαντική, καθώς μπορεί να διευκολύνει την κατάρτιση βάσεων δεδομένων που περιέχουν πληροφορίες από διαφορετικές πηγές με διαφορετικό επίπεδο λεπτομέρειας. Είναι επίσης χρήσιμη γιατί διαφορετικοί λήπτες αποφάσεων ενδιαφέρονται για διαφορετικά επίπεδα ανάλυσης. Π.χ. ένας δήμος μπορεί να ενδιαφέρεται για πιο εστιασμένη άποψη των δεδομένων, σε αντίθεση με το υπουργείο που επιθυμεί μια πιο συγκεντρωτική εικόνα.

5. Χρησιμοποιούνται από χρήστες με διαφορετικές ανάγκες. Για παράδειγμα, για έναν υπεύθυνο χάραξης πολιτικής αρκεί να μπορεί να περιηγείται και να αναζητά στα δεδομένα, ενώ για κάποιον που φτιάχνει μοντέλα αξιολόγησης πολιτικής τα δεδομένα θα πρέπει ιδανικά να έχουν την δυνατότητα εξαγωγής σε μορφή συμβατή με το μοντέλο του (π.χ. μέσω μιας υπηρεσίας ιστού – web application).

Υπό το πρίσμα των παραπάνω διαπιστώσεων, στην δημοσίευση αυτή κάναμε μία συντομη παρουσίαση της τεχνολογίας της «Αποθήκης Δεδομένων» (Data Warehouse) και συζητήσαμε θέματα σχεδίασης στο πλαίσιο δεδομένων σχετικών με την αξιολόγηση Αγροτικής Πολιτικής. Στην συνέχεια παρουσιάσαμε μία μελέτη περίπτωσης για την κλιματική αλλαγή, όπου συνδυάζοντας δεδομένα από διαφορετικές πηγές προσπαθήσαμε να δούμε το αν η απόδοση κάποιων αροτραίων καλλιεργειών επηρεάζεται από την θερμοκρασία και την βροχόπτωση, για τις περιφέρειες της Κρήτης και της ανατολικής Μακεδονίας & Θράκης.

Τα συμπεράσματα στα οποία καταλήξαμε είναι τα εξής:

- Η χρήση εργαλείων ανοικτού και δωρεάν λογισμικού είναι επαρκής για να καλύψει τις απαιτήσεις σχεδίασης και υλοποίησης αποθηκών δεδομένων για χρήση στην Αγροτική Πολιτική.
- Η διάθεση υπηρεσιών ιστού (web services) από παρόχους τέτοιων δεδομένων (π.χ. ΕΛΣΤΑΤ, ΟΠΕΚΕΠΕ, κλπ) έχει κομβικό ρόλο στην αποτελεσματική υλοποίηση αποθηκών δεδομένων. Η απουσία τέτοιων υπηρεσιών που παρατηρείται στους ελληνικούς δημόσιους οργανισμούς δυσχεραίνει την ανάπτυξη σχετικών εφαρμογών.
- Θα πρέπει να υπάρχει διαδικασία αναζήτησης, επισήμανσης και διόρθωσης λαθών στα δεδομένα,
- Ο σχεδιασμός της ιεραρχίας των διαστάσεων στο σχήμα της αποθήκης δεδομένων θα πρέπει να προβλέπει την μελλοντική προσθήκη και άλλων πηγών δεδομένων.

Εναλλακτικά σενάρια περιφερειοποίησης της ΚΑΠ: Μία προσέγγιση με υπόδειγμα μαθηματικού προγραμματισμού (Why to regionalize CAP payments: A farm modeling approach)

Η ενδιάμεση αναθεώρηση της ΚΑΠ το 2003 εισήγαγε το καθεστώς ενιαίας ενίσχυσης, δηλαδή μια αποσυνδεδεμένη ενίσχυση στους γεωργούς που αντικατέστησε τις μέχρι τότε συνδεδεμένες με την παραγωγή επιδοτήσεις. Η εν λόγω μεταρρύθμιση παρείχε στα κράτη μέλη τη διακριτική ευχέρεια να επιλέξουν μεταξύ τριών διαφορετικών μοντέλων ενίσχυσης: το ιστορικό μοντέλο συνδεδεμένο με τις παρελθούσες ενισχύσεις των επιμέρους εκμεταλλεύσεων, το περιφερειακό μοντέλο όπου οι ενισχύσεις ήταν συνδεδεμένες με συγκεκριμένες περιοχές και το υβριδικό, το οποίο συνδύαζε χαρακτηριστικά των δύο προηγούμενων μοντέλων.

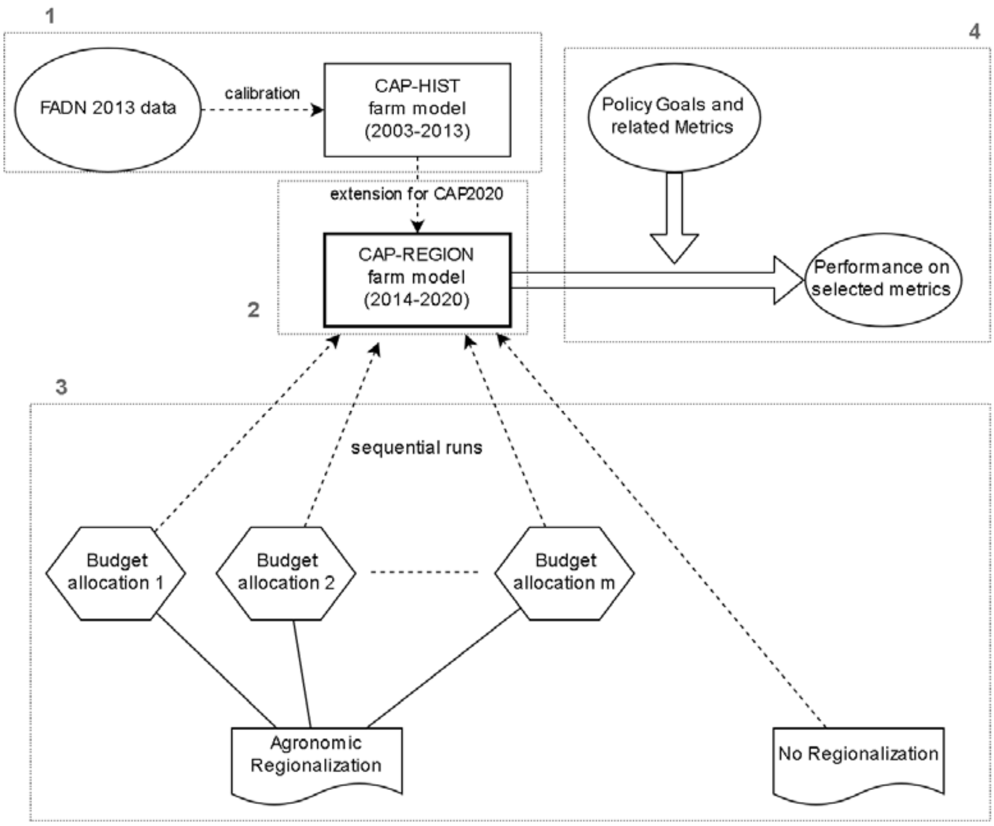
Σύμφωνα με το ιστορικό μοντέλο, η ενίσχυση σε κάθε παραγωγό ισοδυναμούσε με τη χρηματοδοτική στήριξη που έλαβε κατά την «περίοδο αναφοράς» (2000-2002), διατηρώντας στην πράξη άθικτη την προηγούμενη κατανομή των πληρωμών. Αντιθέτως, σύμφωνα με το περιφερειακό μοντέλο, όλες οι γεωργικές εκμεταλλεύσεις μιας συγκεκριμένης περιοχής λαμβάνουν την ίδια πληρωμή ανά εκτάριο (flat rate). Η μεγάλη διαφορά μεταξύ των δύο αυτών μοντέλων βρίσκεται στο πώς διαμορφώνεται η μοναδιαία αξία των δικαιωμάτων (πληρωμή ανά εκτάριο) σε κάθε εκμετάλλευση, παρόλο που το συνολικό ποσό του προϋπολογισμού δεν διαφέρει.

Με τη σειρά της, η μεταρρύθμιση της ΚΑΠ το 2013 αντικατέστησε την ενιαία ενίσχυση με το καθεστώς της βασικής ενίσχυσης. Αυτό συνίσταται σε έναν πυρήνα εισοδηματικής στήριξης στους παραγωγούς (περίπου το 50% του προϋπολογισμού) που συμπληρώνεται από άλλες ενισχύσεις με συγκεκριμένη στοχοθεσία όπως είναι η ενίσχυση για την οικολογική μέριμνα (greening), η αναδιανεμητική ενίσχυση (redistributive payments), οι ενισχύσεις εστιασμένες στις Περιοχές Φυσικών Περιορισμών, το πρόγραμμα νέων γεωργών και οι αποσυνδεδεμένες συμπληρωματικές εθνικές άμεσες ενισχύσεις.

Επιπλέον, η μεταρρύθμιση του 2013 αποσκοπούσε στη μείωση των ανισοτήτων των πληρωμών ανά εκτάριο. Συγκεκριμένα, τα κράτη μέλη που εφάρμοσαν στην προηγούμενη προγραμματική περίοδο το ιστορικό ή το υβριδικό μοντέλο, και άρα διατήρησαν τις ανισότητες στις πληρωμές ανά εκτάριο, θα έπρεπε να εφαρμόσουν μία διαδικασία σύγκλισης αυτών.

Ωστόσο τα κράτη μέλη είχαν τη δυνατότητα να διαφοροποιήσουν τη μοναδιαία αξία μεταξύ διαφορετικών περιοχών καθορισμένων βάσει κοινωνικοοικονομικών ή αγρονομικών κριτηρίων. Η δυνατότητα αυτή ονομάστηκε **περιφερειοποίηση**. Η παρούσα δημοσίευση αξιολογεί την επιλογή της περιφερειοποίησης στην περίπτωση της Ελλάδας σε σχέση με την επιλογή μιας ενιαίας για την χώρα πληρωμής ανά εκτάριο. Στην συνέχεια αξιολογείται μεγάλος αριθμός εναλλακτικών σεναρίων περιφερερειοποίησης, μεταξύ των οποίων και εκείνο που τελικά επιλέχθηκε από το Υπουργείο Αγροτικής Ανάπτυξης για την προγραμματική περίοδο 2014/20.

Το πλαίσιο αξιολόγησης των διαφορετικών σεναρίων πολιτικής δίνεται στο Διάγραμμα 1.



Διάγραμμα 1, Το πλαίσιο αξιολόγησης των σεναρίων περιφερειοποίησης (Διαδικασίες 1-4).

Διαδικασία 1: Ένα Υ.ΕΚΜ με αντικειμενική συνάρτηση την μεγιστοποίηση του ακαθάριστου κέρδους, προσαρμόστηκε στα δεδομένα του FADN για το σύνολο της ελληνικής επικράτειας για το 2013, με βάση το καθεστώς πολιτικής 2007-2013.

Διαδικασία 2: Το παραπάνω μοντέλο επεκτάθηκε ώστε να είναι συμβατό με το καθεστώς 2014-2020, δηλαδή την δυνατότητα της περιφερειοποίησης και κάποιους περιορισμούς σχετικά με το πρασίνισμα.

Διαδικασία 3: Με βάση τον διαχωρισμό των περιφερειών στην Ελληνική περίπτωση (αροτραίες, δενδρώδεις καλλιέργειες και βοσκότοποι), δημιουργούνται 132 πιθανά σενάρια κατανομής του προϋπολογισμού της βασικής και πράσινης ενίσχυσης καθώς επίσης και το σενάριο της ενιαίας πληρωμής ανά εκτάριο.

Διαδικασία 4: Με βάση τους στόχους πολιτικής, όπως περιγράφονται στα επίσημα κείμενα της Ευρωπαϊκής Επιτροπής για την επικείμενη αναθεώρηση (2021-2027), και επιλύοντας το μοντέλο για κάθε σενάριο, καταγράφουμε τις επιδόσεις του κάθε σεναρίου στον κάθε στόχο.

Τα συμπεράσματα στα οποία καταλήξαμε, με βάση την ανάλυση για την ελληνική γεωργία, έχουν ως εξής:

- Η ενιαία πληρωμή ανά εκτάριο στο σύνολο της χώρας (flat rate) είναι σε κάθε περίπτωση χειρότερη από την επιλογή της περιφερειοποίησης. Προσδιορίσαμε τα σενάρια περιφερειοποίησης τα οποία έχουν σε κάποιους από τους στόχους πολιτικής ίση απόδοση με το flat rate και στους υπόλοιπους καλύτερη. Δηλαδή το flat rate δεν συμπεριλαμβάνεται στις αποτελεσματικές κατά Παρέτο επιλογές.
- Η χρήση Υ.ΕΚΜ για την αξιολόγηση των αποσυνδεδεμένων ενισχύσεων μπορεί να παράσχει πληροφορίες σχετικά με την προσαρμογή των εκμεταλλεύσεων σε αντίθεση με την καθαρά λογιστική προσέγγιση που υπολογίζει στατικά τις μεταβιβάσεις πληρωμών. Αυτό είναι ιδιαίτερα σημαντικό στην περίπτωση που το μοντέλο συμπεριλαμβάνει τον μη ουδέτερο χαρακτήρα των αποσυνδεδεμένων ενισχύσεων. Στην περίπτωση μας αυτό συμβαίνει καθώς η αυξομείωση του επιπέδου πληρωμών σε κάθε εκμετάλλευση επηρεάζει το διαθέσιμο κεφάλαιο κίνησης και εν τέλει τις παραγωγικές επιλογές.
- Το εν λόγω Υ.ΕΚΜ μπορεί να παρέχει ποσοτικές κατευθύνσεις στους υπεύθυνους χάραξης της αγροτικής πολιτικής. Για παράδειγμα στην περίπτωση που επιλεγεί το flat rate, το μοντέλο εκτιμάει την απόσταση της απόδοσης του σε κάποιο στόχο πολιτικής από την απόδοση του καλύτερου σεναρίου περιφερειοποίησης σε αυτόν τον στόχο. Έτσι μπορεί να εκτιμηθεί η αντιστάθμιση που απαιτείται (συμπληρωματική ενίσχυση στις μικρές εκμεταλλεύσεις, συνδεδεμένη ενίσχυση για συγκεκριμένους κλάδους, κλπ).

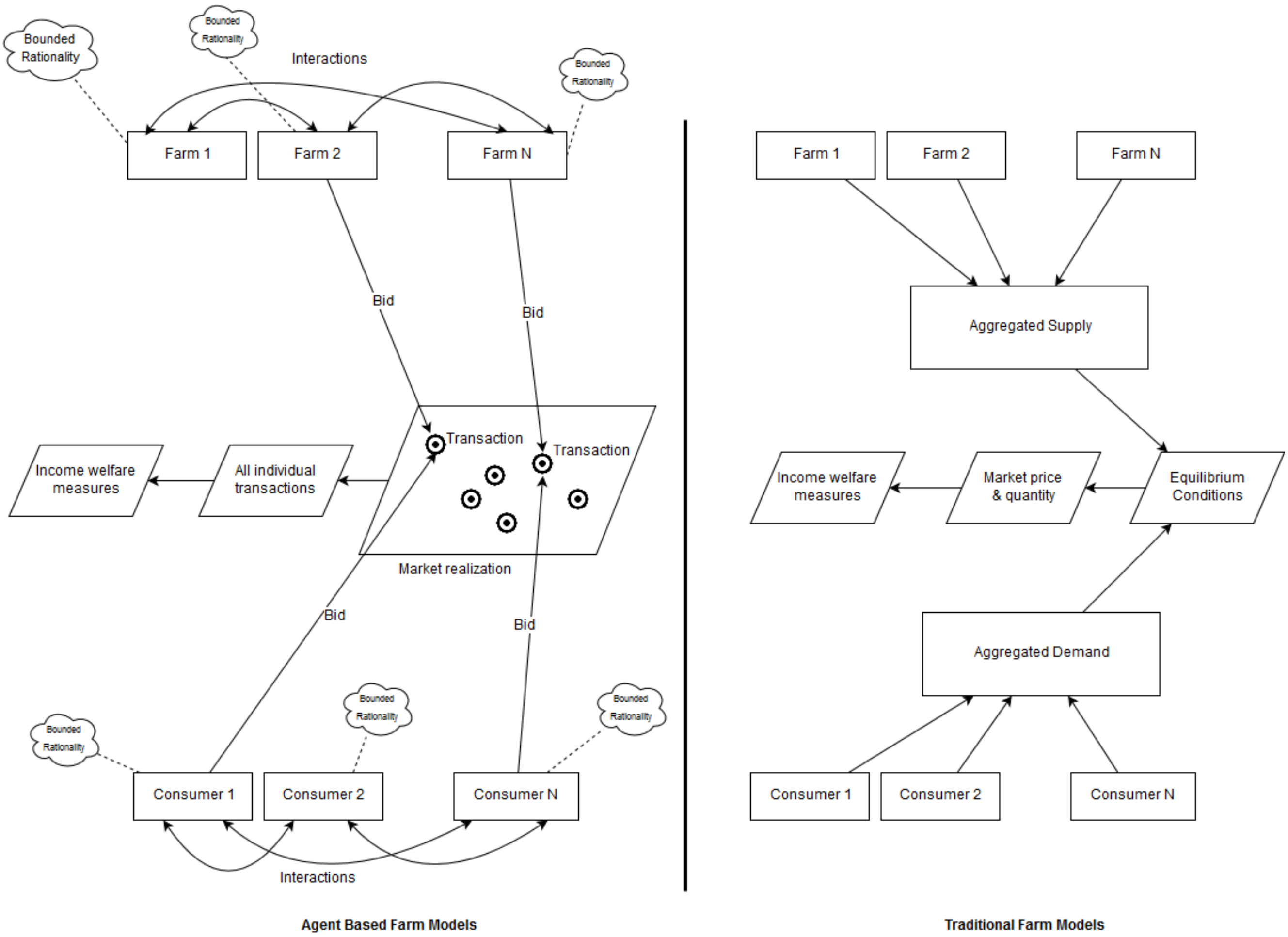
Βιβλιογραφική Ανασκόπηση των μοντέλων δρώντων υποκειμένων για την αξιολόγηση Αγροτικής Πολιτικής (A review of Agent Based Models for Agricultural Policy Evaluation)

Η Αγροτική Πολιτική απομακρύνεται από μέτρα παρέμβασης στην αγορά προς έναν συνδυασμό εθελοντικών και υποχρεωτικών μεταβιβάσεων που συνδέονται με τα χαρακτηριστικά των επιμέρους γεωργικών εκμεταλλεύσεων, τις περιβαλλοντικές τους επιδόσεις και την ικανότητα παροχής ή διατήρησης δημόσιων αγαθών. Κατά συνέπεια, οι επιπτώσεις των μέτρων πολιτικής εντοπίζονται όλο και περισσότερο στο μικρο-επίπεδο και εξαρτώνται από τα ειδικά χαρακτηριστικά της εκάστοτε εκμετάλλευσης. Συνεπώς η ανάλυση πολιτικής σε αυτή την κλίμακα καθίστανται επίκαιρη.

Οι Berger & Troost (2014) συνόψισαν τις προϋποθέσεις που πρέπει να πληρούν τέτοια μικροοικονομικά μοντέλα εκμεταλλεύσεων προκειμένου να είναι χρήσιμα σε αυτό το νέο πλαίσιο πολιτικής:

- επαρκής λεπτομέρεια της χρησιμοποιούμενης τεχνολογίας, διαθέσιμων πόρων και αγρονομικών συνθηκών
- επαρκή κάλυψη της ετερογένειας των εκμεταλλεύσεων
- μοντελοποίηση των αλληλεπιδράσεων μεταξύ των παραγωγών
- ενσωμάτωση της χωρικής διάστασης ώστε να εξετάζονται οι αλληλεπιδράσεις εκμετάλλευσης-περιβάλλοντος

Καταλήγουν στο συμπέρασμα ότι οι τα Μοντέλα Δρώντων Υποκειμένων (ΜΔΥ) έχουν τη δυνατότητα να ανταποκριθούν στις ανωτέρω απαιτήσεις και έτσι μπορούν να συμπληρώσουν τα υφιστάμενα μοντέλα εκμεταλλεύσεων. Μία σύγκριση των δυο τύπων μοντέλων δίνεται από τους Nolan et al. (2009) και την απεικονίζουμε στο Διάγραμμα 2.



Διάγραμμα 2, Αρχετυπικές διαφορές μεταξύ της κλασσικής μικροοικονομικής προσέγγισης μοντέλων φάρμας και αυτής με συστήματα δρώντων υποκειμένων (προσαρμογή από Nolan et al., 2009)

Η ετερογένεια των παραγωγών και των καταναλωτών, η αλληλεπίδραση μεταξύ τους, η χωρική διάσταση (κοινωνικά δίκτυα, αγορές γης, απομιμήσεις κ.λπ.) παρουσιάζονται ως χαρακτηριστικό γνώρισμα των μοντέλων δρώντων υποκειμένων. Επιπλέον, στην περίπτωση των κλασσικών μοντέλων εκμεταλλεύσεων, οι τιμές και οι ποσότητες ισορροπίας προκύπτουν από την ανακατασκευή των καμπυλών προσφοράς και ζήτησης. Αντίθετα στα ΜΔΥ η αγορά προσομοιώνεται με μεμονωμένες συναλλαγές. Επιπρόσθετα, τα ΜΔΥ χρησιμοποιούνται συχνότερα σε περιπτώσεις όπου οι συνθήκες ισορροπίας δεν μπορούν να επιλυθούν αναλυτικά και όταν είναι επιθυμητή η άρση των κλασσικών υποθέσεων της βαλρασιανής αγοράς, όπως στην περίπτωση όπου αναγνωρίζουμε ότι τα υποκείμενα μας έχουν περιορισμένη ικανότητα επεξεργασίας πληροφοριών ή/και πεπερασμένους πόρους για αυτήν. Επιπλέον, τα υποκείμενα μπορούν να μοντελοποιηθούν ώστε να παρουσιάζουν ικανότητες μάθησης στην πάροδο του χρόνου.

Σε αυτή την δημοσίευση εξετάσαμε την σχετική βιβλιογραφία (ΜΔΥ στην αξιολόγηση της Αγροτικής Πολιτικής) από το 2000 εώς και το 2016 με σκοπό (α) να κατηγοριοποιήσουμε τις σχετικές προσεγγίσεις και (β) να εξετάσουμε κρίσιμες παραμέτρους που μπορεί να κάνουν την χρήση των ΜΔΥ ακόμα πιο ευρεία.

Τα συμπεράσματα μας έχουν ως εξής:

- *Για το επίπεδο διαφάνειας των μοντέλων*: Η πλειοψηφία των δημοσιεύσεων ακολουθεί το πρωτόκολλο ODD (Grimm et al., 2010), ωστόσο το γενικό επίπεδο διαφάνειας των μοντέλων πρέπει να βελτιωθεί περαιτέρω. Κατ΄ελάχιστο θα πρέπει να παρέχεται το εκτελέσιμο αρχείο του μοντέλου και τα δεδομένα της δημοσίευσης ώστε να είναι αναπαράξιμα τα αποτελέσματα. Ιδανικά θα πρέπει να υπάρχει πρόσβαση και στον πηγαίο κώδικα.
- *Σε σχέση με την επάρκεια της μοντελοποίησης της εκμετάλλευσης και την χρήση των δυνατοτήτων της προσέγγισης των ΜΔΥ:* Θα πρέπει να γίνουν περισσότερα προς την κατεύθυνση αυτή, ώστε παράμετροι όπως η μάθηση, οι συλλογικές δομές, η προσαρμοστικότητα των υποκειμένων να συνδεθούν καλύτερα σε μοντέλα αξιολόγησης πολιτικής.
- Η μοντελοποίηση της αλληλεπίδρασης μεταξύ των εκμεταλλεύσεων και της χωρικής διάστασης: Τα ΜΔΥ συμπεριλαμβάνουν σε σημαντικό βαθμό τις δύο αυτές διαστάσεις. Ωστόσο περισσότερη προσπάθεια απαιτείται ώστε αυτές να στηρίζονται σε εμπειρικά δεδομένα με την χρήση κατάλληλων στατιστικών μεθόδων.

Επεκτείνοντας ένα μοντέλο εκμετάλλευσης με την προσέγγιση των μοντέλων δρώντων υποκειμένων: Η περίπτωση αξιολόγησης της περιφερειοποίησης της ΚΑΠ (Extending a Farm Model by means of Agent Based Model for Evaluating CAP Regionalization Scenarios)

Η πλειοψηφία των Υ.ΕΚΜ για την αξιολόγηση πολιτικής χρησιμοποιούν αντιπροσωπευτικές εκμεταλλεύσεις, ένα δείγμα δηλαδή του πραγματικού πληθυσμού, βάσει κριτηρίων όπως η παραγωγική δραστηριότητα, το οικονομικό μέγεθος, η διοικητική περιοχή κλπ.

Από την άλλη πλευρά, τα ΜΔΥ βασίζονται στην προσομοίωση ολόκληρου του πληθυσμού των εκμεταλλεύσεων, αν και αυτός συνήθως αναφέρεται σε μικρές σχετικά περιοχές (λίγες κοινότητες). Στο άρθρο αυτό αναλύουμε σχετικά σύντομα τα πλεονεκτήματα και τα μειονεκτήματα της μετάβασης από τις αντιπροσωπευτικές εκμεταλλεύσεις σε μοντελοποίηση με τον πλήρη πληθυσμό εκμεταλλεύσεων.

Τα οφέλη συνοψίζονται ως εξής:

- Δυνατότητα θεώρησης της χωρικής διάστασης με περισσότερη ακρίβεια
- Δυνατότητα να ληφθεί υπόψιν η ετερογένεια των εκμεταλλεύσεων σε μεγαλύτερο βαθμό
- Καλύτερη αναπαράσταση των σχέσεων διάδρασης μεταξύ των εκμεταλλεύσεων
- Καλύτερος συντονισμός με βιοφυσικά μοντέλα

Από την άλλη οι επιπλέον απαιτήσεις που εγείρονται με αυτή την μετάβαση, συνοψίζονται ως εξής:

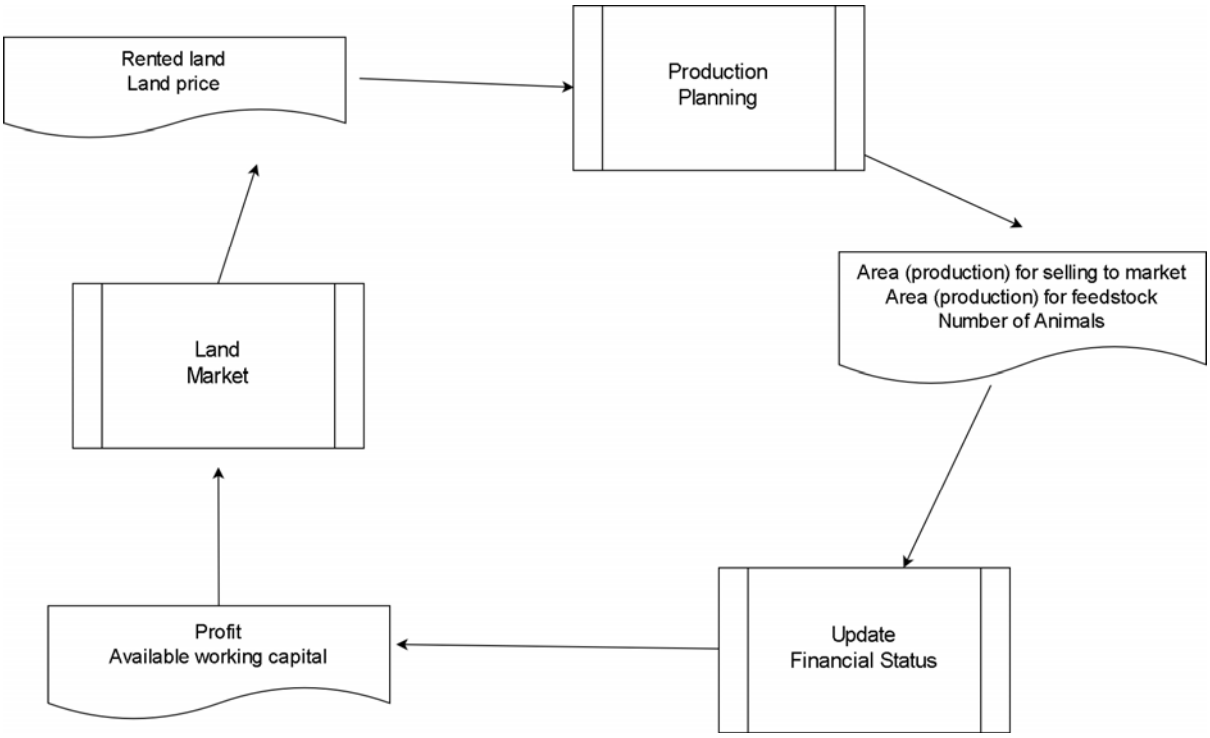
- Αύξηση των απαιτήσεων σε δεδομένα και μειωμένη ακρίβεια των παρατηρήσεων
- Περισσότερες υποθέσεις για κομβικά στοιχεία του μοντέλου
- Τα αποτελέσματα εμπεριέχουν μεγαλύτερη αβεβαιότητα

Στην συνέχεια αναπτύξαμε ένα πρότυπο μοντέλο με δρώντα υποκείμενα, όπως φαίνεται στο Διάγραμμα 3. Τα ορθογώνια σχήματα είναι υπομοντέλα ενώ τα καμπύλα σχήματα συμβολίζουν την ανταλλαγή δεδομένων μεταξύ των διάφορων υπομοντέλων.

Η ροή εργασίας του μοντέλου έχει ως εξής. Μετά την αρχικοποίηση των δεδομένων, ξεκινάει το υπομοντέλο προγραμματισμού παραγωγής των εκμεταλλεύσεων, που βασίζεται στο μοντέλο μαθηματικού προγραμματισμού των αντιπροσωπευτικών εκμεταλλεύσεων. Στην συνέχεια ανανεώνεται η χρηματοοικονομική κατάσταση της εκμετάλλευσης, και όποιες εκμεταλλεύσεις έχουν για κάποιες χρονιές συνεχώς αρνητικό αποτέλεσμα, αποχωρούν από την προσομοίωση. Στην συνέχεια και με βάση τα παραπάνω αποτελέσματα, οι εκμεταλλεύσεις λαμβάνουν μέρος σε μία αγορά ενοικίασης γης, μέσω της οποίας μπορεί να επεκταθούν ή να συρρικνωθούν.

Η εφαρμογή του παραπάνω μοντέλου έγινε στο σενάριο του flat rate και σε τρία επιλεγμένα σενάρια περιφερειοποίησης, τα οποία αναδείχθηκαν από προηγούμενη εργασία (*Why to regionalize CAP payments: A farm modeling approach*), το 0-50-50 (0% του προϋπολογισμού στις αροτραίες, 50% στις δενδρώδεις και 50% στους βοσκότοπους), το 65-20-15 και το 70-0-30. Επιπλέον επιλέχθηκαν μόνο εκμεταλλεύσεις με παραγωγικό προσανατολισμό αροτραίων καλλιεργειών, αιγοπροβατοτροφίας ή μεικτό. Επίσης επιλέχθηκαν 8 χαρακτηριστικοί νομοί.

Τα αποτελέσματα που παρουσιάζονται στην εργασία περιγράφουν για κάθε νομό το ύψος του ενοικίου για τις αρδευόμενες και τις μη αρδευόμενες εκτάσεις καθώς και την κατάσταση στα ακαθάριστα κέρδη ανά παραγωγική κατεύθυνση, σε σύγκριση με το σενάριο της πλήρους ισότητας των ενισχύσεων ανά εκτάριο.



Διάγραμμα 3, Η δομή του μοντέλου δρώντων υποκειμένων

Σύστημα Λήψης Απόφασης για τον σχεδιασμό της περιφερειοποίησης της ΚΑΠ (CAP2020 regionalization design: A Decision Support System)

Ένα Σύστημα Λήψης Απόφασης (ΣΥΑ) είναι ιδιαίτερα χρήσιμο στην διαδικασία της αξιολόγησης σεναρίων αγροτικής πολιτικής εφόσον είναι σε θέση να μοντελοποιεί τις επιπτώσεις της πολιτικής και να τις παρουσιάζει στους αποφασίζοντες. Σχετικά συστήματα αναφέρουμε ενδεικτικά παρακάτω.

Οι Manos et al. (2010) παρουσιάζουν ένα ΣΥΑ σχετικό με την αειφόρο ανάπτυξη και την προστασία του περιβάλλοντος των γεωργικών περιοχών. Το σύστημα στοχεύει στη βελτιστοποίηση του σχεδίου παραγωγής μιας γεωργικής περιοχής λαμβάνοντας υπόψη τους διαθέσιμους πόρους, τις περιβαλλοντικές παραμέτρους και τους κρίσιμους παράγοντες της περιοχής αυτής. Στην εργασία τους, οι Borges et al. (2010) στηρίζουν το ΣΥΑ σε ένα μοντέλο για την πρόβλεψη των επιπτώσεων των αλλαγών της ΚΑΠ στις τιμές της γης στις αγροτικές περιοχές (συμπεριλαμβανομένης της δασικής γης). Οι Louhichi et al. (2010) συνδέουν στο ΣΥΑ ένα βιοοικονομικό μοντέλο γεωργικών εκμεταλλεύσεων. Οι Bournaris and Papathanasiou (2012) παρουσιάζουν ένα ΣΥΑ όπου προσομοιώνονται διαφορετικά σενάρια πολιτικής και υπολογίζονται οι επιπτώσεις στα σχέδια πολιτικής. Τέλος, οι Rovaia et al. (2016) παρουσιάζει ένα ΣΥΑ για τη διαχείριση του αγροτικού τοπίου.

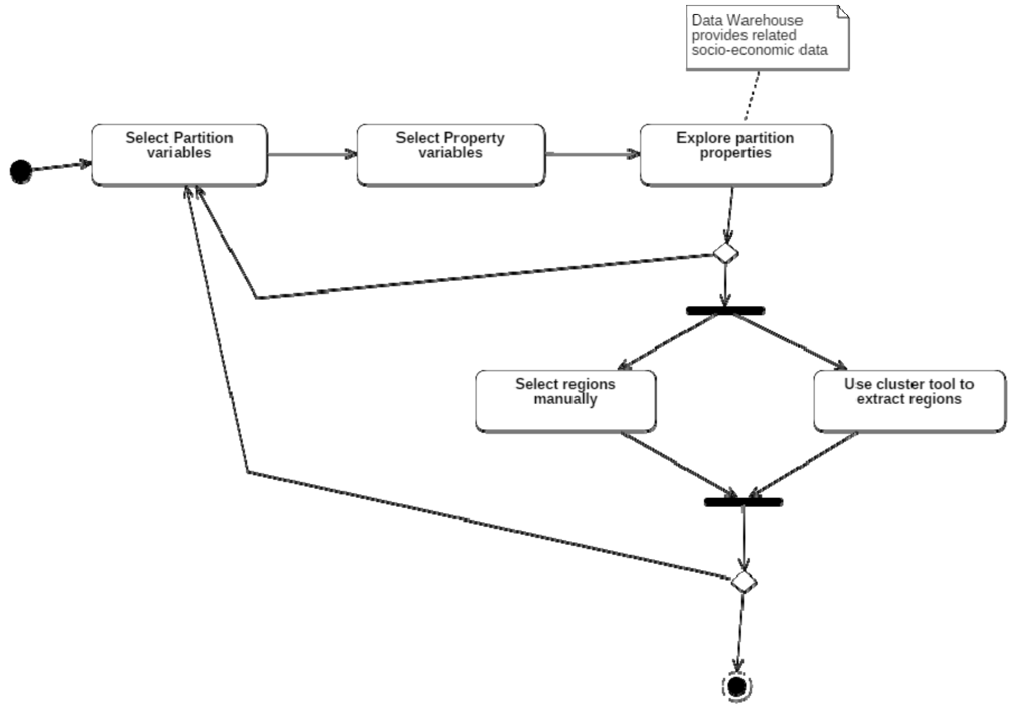
Στα πλαίσια της δυνατότητας της χώρας να εφαρμόσει την περιφερειοποίηση της ΚΑΠ, ένα ΣΥΑ μπορεί να βοηθήσει στην αποτελεσματικότερη και διαφανέστερη σχεδίαση της πολιτικής καθώς και στην άμεση εμπλοκή των ενδιαφερομένων (αγρότες, συνεταιρισμοί, δήμοι, περιβαλλοντικές οργανώσεις, πολίτες κλπ) στην διαδικασία.

Ως εκ τούτου, στην δημοσίευση αυτή υλοποιήσαμε και παρουσιάζουμε ένα τέτοιο ΣΥΑ. Στα Διαγράμματα 4 και 5 παρουσιάζουμε εν συντομία δύο από τις σημαντικότερες λειτουργίες του, την διαδραστική διαμέριση των περιφερειών από τον αποφασίζοντα και την επίσης διαδραστική κατανομή του προϋπολογισμού μεταξύ των περιφερειών.

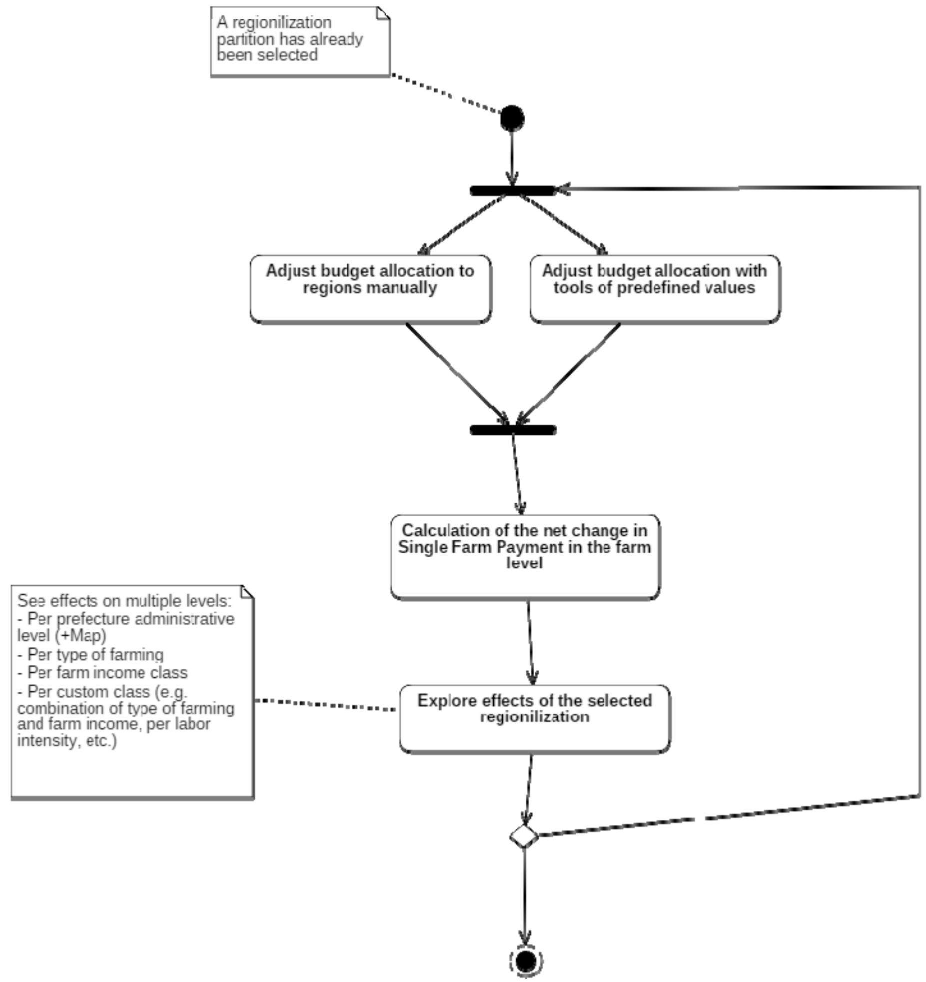
Στην πρώτη περίπτωση ο αποφασίζων επιλέγει τις μεταβλητές με βάση τις οποίες θέλει να γίνει η διαμέριση των περιφερειών (π.χ. νομούς, περιφέρειες, αγρονομικές ζώνες) και τις μεταβλητές που αναδεικνύουν τα χαρακτηριστικά των διαμερίσεων που προκύπτουν (π.χ. σύνολο ενισχύσεων, αριθμός εκμεταλλεύσεων, αξία παραγωγής, κλπ.). Έπειτα, δημιουργώντας περιφέρειες από το σύνολο όλων των πιθανών διαμερίσεων βλέπει τα χαρακτηριστικά της κάθε περιφέρειας (αριθμός εκμεταλλεύσεων, ύψος ενίσχυσης, ενίσχυση ανά στρέμμα, κλπ.), μέχρις ότου καταλήξει στην επιθυμητή δομή της περιφερειοποίησης. Το σύστημα μπορεί να το υποβοηθήσει εκτελώντας ανάλυση ομάδων (cluster analysis).

Στην δεύτερη περίπτωση και αφού ο χρήστης έχει δημιουργήσει τις επιθυμητές περιφέρειες, μπορεί να πειραματιστεί με τις επιπτώσεις διαφορετικών κατανομών του συνολικού προϋπολογισμού στις περιφέρειες (βλέποντας τόσο σε πίνακες όσο και σε χάρτες τις αυξομειώσεις στα ποσά των ενισχύσεων ανά κατηγορία οικονομικού μεγέθους, νομό και είδος καλλιέργειας).

Στο μέλλον σχεδιάζουμε να επεκτείνουμε τη βάση δεδομένων του ΣΥΑ με κοινωνικοοικονομικά δεδομένα. Επιπλέον, θα ενσωματώσουμε τα μοντέλα μαθηματικού προγραμματισμού της διατριβής έτσι ώστε να μπορεί να αξιολογηθεί και η προσαρμογή των εκμεταλλεύσεων στα επιλεγμένα σενάρια πολιτικής.



Διάγραμμα 4. Διαδραστική επιλογή Περιφερειών



Διάγραμμα 5. Διαδραστική εξερεύνηση Κατανομών Προϋπολογισμού

Βιβλιογραφία

Ackrill, R.W., Ramsden, S.J., & Gibbons, J.M, (2001). CAP reform and the rebalancing of support for cereals and oilseeds: a farm-level analysis. European Review of Agricultural Economics, 28(2), 207-226.

Berger, T., & Troost, C. (2014). Agent-based Modelling of Climate Adaptation and Mitigation Options in Agriculture. Journal of Agricultural Economics, 65(2). https://doi.org/10.1111/1477-9552.12045

Borges, P.J., Fragoso, R., Garcia-Gonzalo, J., Borges, J.G., Marques, S., Lucas, M.R. (2010) Assessing impacts of Common Agricultural Policy changes on regional land use patterns with a decision support system: An application in Southern Portugal, Forest Policy and Economics 12 (2010) 111–120

Bournaris T. and Papathanasiou, J. (2012) A DSS for planning the agricultural production, Int. J. Business Innovation and Research, Vol. 6, No. 1, DOI: 10.1504/IJBIR.2012.044259

W. Britz, P. Witzke (2012), CAPRI model documentation, http://www.capri-model.org, accessed at 2/11/2013

E. Galko, P.-A. Jayet (2011), Economic and environmental effects of decoupled agricultural support in the EU, Agricultural Economics 42 (2011) 605–618

Efstratoglou, Sophia & Giannakis, Elias & Psaltopoulos, Dimitrios (2011) Evaluating CAP alternative policy scenarios through a system dynamics approach in rural areas of Greece, 122nd Seminar, February 17-18, 2011, Ancona, Italy 99361, European Association of Agricultural Economists.

Giannakis, E., Efstratoglou, S., & Psaltopoulos, D. (2014). Modelling the impacts of alternative CAP scenarios through a system dynamics approach. Agricultural Economics Review, 15(389-2016-23508), 48.

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: a review and first update. Ecological modelling, 221(23), 2760-2768.

Guindé, L., Millet, G., Rozakis, S., Sourie, J-C., and D. Treguer (2005) The CAP mid-term reform impacts to French cereal farms: a regional supply modelling approach. in F. Arfini (ed.), Modelling Agricultural Policies: State of the Art and new Challenges, Parma: MUP

Katranidis, Stelios D., (2002) Welfare Analysis in a Multi-Market Framework: Implications of the Cap Cotton, Maize and Sugar Beet Regime in Greece, International Congress, August 28-31, 2002, Zaragoza, Spain 24946, European Association of Agricultural Economists.

Klonaris S. & Vlahos G, (2012), The Politics of CAP: The case of Greece, Working Papers 2012-5, Agricultural University of Athens, Department Of Agricultural Economics.

Louhichi, K., Kanellopoulos, A., Guillermo Flichman, S. J., Blanco, M., Hengsdijk, M., Heckelei, T., Berentsen, P., Lansink, A. O., Van Ittersum, M. (2010) FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies, Agricultural Systems 103 (2010) 585–597

Manos B., Bournaris T., Papathanasiou J., Chatzinikolaou P. (2009). Evaluation of tobacco cultivation alternatives under the EU common agricultural policy (CAP), Journal of Policy Modeling, Elsevier, vol. 31(2), pages 225-238.

Manos, B.D., Papathanasiou, J., Bournaris T. and Voudouris, K. (2010) A DSS for sustainable development and environmental protection of agricultural regions, Environ Monit Assess 164:43–52, DOI 10.1007/s10661-009-0873-1

Manos, B., Bournaris, T., Chatzinikolaou, P., Berbel, J., & Nikolov, D. (2013). Effects of CAP policy on farm household behaviour and social sustainability. Land Use Policy, 31, 166-181.

Mantziaris, S., Rozakis, S., & Karanikolas, P. (2017). CAP Implementation under Memorandum of Understanding bail-out conditionality: Impacts on business-oriented arable farms in Greece. New Medit, 16(4), 12-24

Mattas K., E. Loizou, S. Rozakis, E. Tzouvelekas (2006) Agricultural modelling: an input-output approach, in in F. Ferretti (ed.), *Leaves and Cigarettes :* Modelling the tobacco industry, with application to Italy and Greece*,* Milano: Francoangeli Publisher

Nolan, J., Parker, D., Van Kooten, G. C., & Berger, T. (2009). An overview of computational modeling in agricultural and resource economics. Canadian Journal of Agricultural Economics, 57(4), 417–429. http://doi.org/10.1111/j.1744-7976.2009.01163.x

Petsakos A., Rozakis R. and Tsiboukas K (2009). Risk optimal farm plans in the context of decoupled subsidy payments: the case of cotton production in Thessaly, Journal of Farm Management 13(7): 467-483

Petsakos A., S. Rozakis (2009) Critical Review and State-of-the-Art of PMP Models: An Application to Greek Arable Agriculture, in Research Topics in Agricultural and Applied Economics, A. Rezitis (ed.), Bentham Science Publishers (forthcoming)

Pezaros, P. "The C.A.P. in the Pliers of the Multilateral Trading System: Origins, Evolution and Future Challenges" in the S. Bilal – P. Pezaros (Ed.) (2000): Negotiating the Future of Agricultural Policies", Kluwer Law International, The Hague, pp. 51 -80

Rezitis A, Stavropoulos K, (2010) Modeling beef supply response and price volatility under CAP reforms: The case of Greece, Food Policy, Volume 35, Issue 2

Reidsma, P., Janssen, S., Jansen, J., & van Ittersum, M. K. (2018). On the development and use of farm models for policy impact assessment in the European Union – A review. Agricultural Systems, 159(September 2017), 111–125. http://doi.org/10.1016/j.agsy.2017.10.012

Rozakis S, Tsiboukas K and Korasidis M. (2008). CAP Reform Impacts To Greek Cotton Farmers: A Mathematical Programming Approach, Journal of Agricultural and Food Economics 3(1-2): 85-106

Rozakis S. (2010)., Hybrid linear programming to estimate CAP impacts of flatter rates and environmental top-ups, Working Papers 2010-03, Agricultural University of Athens, Department Of Agricultural Economics.

Tzima A., Athanasiadis I., et al. (2007). Agent-based modelling and simulation in the irrigation management sector: applications and potential. Options Mediterraneennes, Series B: Studies and Research, Proceedings of the WASAMED International Conference.

Van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A., Van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS). Agricultural Systems 96, 150-165.

Wilson, P., Gibbons, J.M., & Ramsden, S. J. (2003). The impact of cereal prices and policy on crop rotations and supply response, Journal of Agricultural Economics, 54(2), 313-323.

Πετσάκος Α (2012), Μια βιο-οικονομική προσέγγιση στην κατασκευή υποδειγμάτων προγραμματισμού δραστηριοτήτων των γεωργικών εκμεταλλεύσεων, Διδακτορική διατριβή, τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών

Σιντόρη Α. (2012), Ανάλυση της διαδικασίας λήψεως αποφάσεων με πολλαπλούς στόχους. Η περίπτωση των προβατοτροφικών εκμεταλλεύσεων της Ηπειρωτικής Ελλάδας, Διδακτορική διατριβή, τμήμα Αγροτικής Οικονομίας και Ανάπτυξης, Γεωπονικό Πανεπιστήμιο Αθηνών

Introduction

The Common Agricultural Policy (CAP) was introduced in 1962 and got fully implemented in 1968. It is considered to be the first real EU common policy replacing all relevant national agricultural policies while since then numerous reforms have been applied (Pezaros 2000). For the last 20 years, every year CAP is absorbing more or less about 0,5%-0,6% of the EU GDP and 50%-60% of the EU budget. Therefore CAP evaluation is a persisting issue in the Agricultural Economics field.

The new CAP (2014-2020) design, acknowledging the wide diversity of agronomic production potential and climatic, environmental as well as socio-economic conditions and needs across the EU, offers implementation flexibility to member states. Indicatively, member states may differentiate the basic payment per hectare according to administrative or agronomic criteria; choose from different options for internal convergence for payments per hectare until 2019; opt in for the right to use a redistributive payment for the first hectares; enable the “small farm scheme”, where small farms receive an annual subsidy of 500€ - 1250€ with minimal administrative burden; preserve a limited amount for coupled payments; grant an additional payment for areas with natural constraints (as defined under Rural Development rules) .

Inside this flexible CAP framework, the current agricultural policy evaluation models reach their limits. Econometric models cannot give guidance in such major policy shifts and general or partial equilibrium models cannot reach the necessary microeconomic resolution. Since the impacts of policy measures depend on the specific farm characteristics, getting insights at disaggregated level and spatial scale becomes relevant for both policymakers and researchers. Consequently farm scale policy analysis is becoming very relevant and is suitable for CAP policy analysis.

In Greece there is already a significant number of research work based on farm models for evaluating CAP. However they are mainly focused on crops whose regime have changed drastically, i.e. tobacco and cotton, and do not make projections for the Greek agriculture as a whole. The reader can refer to the papers of Katranidis (2002), Petsakos et al. (2006), Rozakis et al. (2008), Manos et al. (2009), Rozakis (2009), Petsakos et al (2009), Rozakis (2010), Rezitis & Stavropoulos. (2010), Efstratoglou et al. (2011), Sintori (2012), Petsakos (2012).

The main aim of the doctoral dissertation is to develop an integrated model of the Greek agriculture that can be used for the evaluation of agricultural policy. An additional objective is to complement the conventional representative farm model with the Agent Based Modeling approach. For this purpose, the following steps were followed.

Handling farm model input data

Data handling and transformation is a core component of a country wide farm model. . Thus the publication **"Data Case Technology for Agricultural Policy Data: a Greek Case Study"** discusses the agricultural policy related data particularities presents a policy application combining data from different sources using the Data Warehouse technology.

The Greek integrated farm model presented later in the dissertation uses data from the Farm Accountancy Data Network (FADN). The data was obtained for the years 2011, 2012 and 2013 from the Ministry of Rural Development in raw format. The transformation in a form is usable by the farm model was quite laborious and was elaborated by constructin an R package named fadnUtils, as described in the Apendix "fadnUtils, An R package for working with FADN data".

Builiding a farm model that represents the majority of Greek Agriculture

The mathematical programming farm is named Greek REpresentative FArm Model (GREFAM). It covers over 25 activities from the plant and animal sector, including sample farms that represent more than 80% of land and 60% of production value of the Greek agriculture. A detailed description of the model is given in the Appendix **«GREFAM Model Reference Manual»**

Applying GREFAM model to a policy problem

The paper **«Why to regionalize CAP payments: A farm modeling approach»** describes in detail an application of the GREFAM model. The performance of regionalization versus a country uniform payment per hectare is compared for various policy objectives. We show that, at least in the case of Greece, regionalizing CAP payments can cope better with the CAP conflicting objectives that reside simultaneously on economic, social and environmental dimensions. The paper exhibits the usefeulness of the farm model for evaluating a contemporary policy problem.

Augmenting the farm model by means of Agent Based Modeling approach

The paper **«A review of Agent Based Models for Policy Evaluation»** thoroughy examines the advantages and the challenges of complementing the farm modeling approach with Agent Based Modeling.

Based on the above findings we created a relevant ABM based on the Repast object oriented software framework, as described in the working paper «**Dealing with farm heterogeneity on modeling agricultural policy: an Agent Based Modeling Approach»** in the Appendix. We further developed another ABM model with a different implementation approach for examining regionalization scenarios in order to complement the GREFAM findings. This is described in the Working Paper **«Extending a Farm Model by means of Agent Based Model for Evaluating CAP Regionalization Scenarios»** in the main text.

Packing the models into a Decision Support system.

Policy makers and stakeholders, in most cases cannot make a direct use of the above models. A Decision Support System that will act as the interface between non experts and the dissertation's models is a prerequisite for will be of great usefulness for the wider adaptation. agricultural policy evaluation. This is described in the publication "CAP2020 regionalization design: A Decision Support System"

References

Efstratoglou, Sophia & Giannakis, Elias & Psaltopoulos, Dimitrios (2011) Evaluating CAP alternative policy scenarios through a system dynamics approach in rural areas of Greece, 122nd Seminar, February 17-18, 2011, Ancona, Italy 99361, European Association of Agricultural Economists.

Katranidis, Stelios D., (2002) Welfare Analysis in a Multi-Market Framework: Implications of the Cap Cotton, Maize and Sugar Beet Regime in Greece, International Congress, August 28-31, 2002, Zaragoza, Spain 24946, European Association of Agricultural Economists.

Manos B., Bournaris T., Papathanasiou J., Chatzinikolaou P. (2009). Evaluation of tobacco cultivation alternatives under the EU common agricultural policy (CAP), Journal of Policy Modeling, Elsevier, vol. 31(2), pages 225-238.

Mattas K., E. Loizou, S. Rozakis, E. Tzouvelekas (2006) Agricultural modelling: an input-output approach, in in F. Ferretti (ed.), Leaves and Cigarettes : Modelling the tobacco industry, with application to Italy and Greece, Milano: Francoangeli Publisher

Petsakos A., Rozakis R. and Tsiboukas K (2009). Risk optimal farm plans in the context of decoupled subsidy payments: the case of cotton production in Thessaly, Journal of Farm Management 13(7): 467-483

Petsakos A., S. Rozakis (2009) Critical Review and State-of-the-Art of PMP Models: An Application to Greek Arable Agriculture, in Research Topics in Agricultural and Applied Economics, A. Rezitis (ed.), Bentham Science Publishers (forthcoming)

Rezitis A, Stavropoulos K, (2010) Modeling beef supply response and price volatility under CAP reforms: The case of Greece, Food Policy, Volume 35, Issue 2

Rozakis S, Tsiboukas K and Korasidis M. (2008). CAP Reform Impacts To Greek Cotton Farmers: A Mathematical Programming Approach, Journal of Agricultural and Food Economics 3(1-2): 85-106

Rozakis S. (2010)., Hybrid linear programming to estimate CAP impacts of flatter rates and environmental top-ups, Working Papers 2010-03, Agricultural University of Athens, Department Of Agricultural Economics.

Πετσάκος Α (2012), Μια βιο-οικονομική προσέγγιση στην κατασκευή υποδειγμάτων προγραμματισμού δραστηριοτήτων των γεωργικών εκμεταλλεύσεων, Διδακτορική διατριβή που υποβλήθηκε στο τμήμα Αγροτικής Οικονομίας και Ανάπτυξης του Γεωπονικού Πανεπιστημίου Αθηνών

Σιντόρη Α. (2012), Ανάλυση της διαδικασίας λήψεως αποφάσεων με πολλαπλούς στόχους. Η περίπτωση των προβατοτροφικών εκμεταλλεύσεων της Ηπειρωτικής Ελλάδας, Διδακτορική διατριβή που υποβλήθηκε στο τμήμα Αγροτικής Οικονομίας και Ανάπτυξης του Γεωπονικού Πανεπιστημίου Αθηνών

Paper: Data warehouse technology for Agricultural Policy Data: a Greek Case Study

Published in the International Journal of Sustainable Agricultural Management and Informatics, Volume 2, Issue 2-4, 2016, Pages 243-262

Data warehouse technology for Agricultural Policy Data: a Greek Case Study

Michael Maliappis[1], Dimitris Kremmydas[2]

[1]Laboratory of Informatics, Department of Agricultural Economics and Rural Development, Agricultural University of Athens

[2]Laboratory of Agribusiness Management, Department of Agricultural Economics and Rural Development, Agricultural University of Athens

Abstract. Statistical data for agricultural policy analysis has certain unique features: a multitude of sources of very different nature; a variety of dimensional granularity; different end user requirements. The utilization of Data Warehouse (DW) technology is valuable for tackling the above issues and successfully offering data to policy stakeholders and modelers. In this paper, we briefly introduce the DW technology, discuss the DW design issues in the context of policy related data and investigate the several difficulties identified on building and using a DW for monitoring crop responses to climate change for two Greek regions.

Keywords: Agricultural Data, Data warehouse, Online Analytical Processing (OLAP), Agricultural Policy.

1 Introduction

Worldwide, the agricultural sector is receiving a significant amount of state funding through various agricultural policy tools. In a recent report of the Organization for Economic Co-operation and Development, the included countries (50 countries, accounting for the majority of global agricultural value added) provided an annual average of EUR 469 billion of support to their agricultural producers directly in the years 2013-15 (OECD, 2016). Thus the efficient allocation of funding in order to accomplish the strategic goals of the policy makers is essential. Agricultural policy analysis is concerned with evaluating the instruments of providing subsidies to the agricultural sector, ex ante or ex post (Alston & James, 2002).

Although this evaluation is based on theoretical models, most often evidence is sought for empirical validation. In fact, as Runge (2006) notes, the agricultural economics subject itself arose in the late 19th century partly due to the fact that the U.S. Department of Agriculture (USDA) had compiled rich datasets some decades earlier. Consequently data is of prime importance for agricultural policy analysis, since it is utilized by policy makers to make qualitative judgments and by researchers to build quantitative models.

This agricultural policy related data bears certain special features:

1. *There exist many independent sources of information*, e.g. international or national statistical offices, diversified administration databases, field surveys or past data from universities, etc., none of which should be disregarded because agricultural data is actually a scarce resource. Those sources possibly store their data in different formats or/and different database schema definitions.

2. *The related data expands horizontally on various dimensions* since frequently agriculture policy makers pursue multiple goals. Those dimensions may be classified to (i) biophysical (weather or soil related, animal population, etc.). A usage example may be to investigate the effect of a policy to a region's biodiversity or soil erosion; (ii) technical/technological (input-output relationships, management practices, etc). Technical relationships (i.e. what inputs are used for a certain crop in a specific area) are very important factors for farmers' production decisions and thus are directly relevant to agricultural policy; (iii) economic (prices of inputs and outputs, production, income, etc.) containing data that is directly related to policies; (iv) social (population per community, age pyramid, etc) since often agricultural policies target at altering (e.g. shrinking/maintaining population, developing skills, etc) rural societies. See Elizabeth et al. (2005) for a typical policy case study where all of the above dimensions are relevant.
3. *The temporal and spatial dimensions are relevant to their finest available detail*. The first is important to note because temporal dimension might not be always recorded, especially in the case of operational databases. Also, normally policy makers are interested on policy effect estimation to the finest administrative unit and the constraint for doing so is data availability.
4. *Dimensions are mostly of hierarchical kind*. For example the spatial/administrative dimension includes the community at the lowest level and the country at the top; production type can be very specific (e.g. production of milk from goats) concluding to aggregated level (e.g. production from animals); time from daily to yearly; etc. This logical hierarchy is relevant, since it can facilitate the compilation of databases that hold information for different level of detail and can also be useful for presentation purposes to different stakeholders of the policy making process (e.g. a municipality officer may be interested on a more focused view of the data, in contrast with a ministry officer that is interested on an aggregated picture).
5. *Data is utilized by different kind of users, each with diverse needs*. For example, for a high level policy maker or an administration officer it is sufficient to browse the data through a web interface or browse the results of a data mining procedure while for a modeler the data will ideally be directly imported to his / her model (e.g. by means of a web service).

The above specific features of agricultural policy related data designates Data Warehouse (DW) technology to be ideal for usage in agricultural policy monitoring and evaluation. DW can effectively facilitate collection of data from different sources explicitly maintaining temporal information; integrate and present multidimensional data; deal efficiently with hierarchical dimensions; and output data in different ways (Boulil et al., 2014; Rai et al., 2008). The application of DW in agricultural policy evaluation is not a straightforward process since DW technology is a set of processes rather than a ready-to-deliver product and there are specific design requirements that are discussed in the rest of the paper.

In the broad domain of agriculture there are several cases where a DW was introduced to manage statistical data. One of the earliest appearances was that of the US Department of Agriculture's National Agricultural Statistics Service (Yost, 2000). Another attempt was that of the development of a central Data Warehouse at Indian Agricultural Statistics Research Institute (IASRI) at New Delhi (Chaturvedi et al., 2006; Rai et al., 2007).

Abdullah and Hussain (2006) describe an Agriculture extension DW that monitors cotton pests in Pakistan. Van Broekhoven (2007) describes a DW system developed for presenting the Belgian Farm Accountancy Data Network data. Additionally, DW solutions are also applied to agricultural business problems. For instance, Schulze et al. (2007) use a DW in a dairy precision farming context, to collect data from different dairy enterprises and efficiently derive timeline measures for examining disease treatments.

In the rest of the paper, we briefly introduce the DW technology (section 2.1), discuss design issues in the context of policy related data (sections 2.2 and 2.3) and investigate the several difficulties identified on building and using a DW for monitoring crop responses to climate change for two Greek regions (section 3). Finally a summary of conclusions is drawn (section 4).

2 Data Warehouse Technology for Agricultural Policy

2.1 A brief introduction to Data Warehouse technology

As Ballard et al. (1998) notes, a Data Warehouse (DW) is not a product but rather a solution for transforming plain information to knowledge. More specifically they define Data Warehousing as "the design and implementation of processes, tools, and facilities to manage and deliver complete, timely, accurate, and understandable information for decision making. It includes all the activities that make it possible for an organization to create, manage and maintain a data warehouse or data mart".

The current DW process lifecycle includes a wide set of operations (Casters et al., 2010; Kimball and Ross, 2013) as depicted in Figure 1. The first step towards DW development is *the identification of data sources*. Usually a disparate (i.e. in respect to mean of storage, access protocols, logical organization, data quality, etc.) set of sources is used and thus an intermediate procedure called *"Extract, Transform, Load" (ETL)* is required in order to prepare plain data and load it in the DW engine. Finally the DW data is not directly accessible by end-users but accessed by means of *reports, data mining interfaces* and *On-Line Analytical Processing (OLAP)* cubes. We provide some DW technology term definitions in order for not so familiar readers to be able to follow the rest of the paper.

The notions of *fact, dimension* and *measure* are central for designing a DW (Malinowski and Zimányi, 2008). *Facts* are collections of related data items, e.g. the utilized agricultural area taken from a census. *Dimensions* are the structures that categorize facts, e.g. the product, the region or the time that the reported utilized area refers to. *Measures* are facts that are aggregated to dimension tuple, e.g. the utilized agricultural area for a specific region and a specific year. This different dimension aggregation can be depicted as a *cube* (in the case of three dimensions) or a *hypercube* (more than three dimensions), where each cell of the cube is a measure. *Star schema* is a relational database schema where one or more central fact tables are linked to one or more dimension tables. DW development process can be abstracted to the mapping of the selected data source data schemas to the DW star schema. A *Data mart* is a collection of facts, dimensions and measures that are subject specific. A DW is actually a collection of data marts.

In figure 2 we provide a visualization of a *data mart* and the connection between *facts, dimensions, measures, cubes* and *star schemas.* The fact of the utilized agricultural area (UAA) is connected to three dimensions: Date (in month resolution) that UAA was measured, the product and the region that the measurement is referring to. There is a direct correspondence between the star schema and the data mart cube. Each cell of the cube is a measure (what is the UAA of product X in region Y at date Z).

Online Analytical Processing (OLAP) is the multidimensional view of data stored in a DW. *OLAP cubes* are view structures that correspond to DW cubes like in figure 2. OLAP functionality includes basic navigation and browsing, i.e. view a *slice* of the data cube (one or more dimensions constant), *dice* (create a sub-cube), *drill down/up* (from aggregated to more detailed dimensions levels or reverse) and *roll-up* (summarizes data along a dimension). Also statistical analyses, time series creation and more complex modeling can be utilized.

Nevertheless there is a significant literature on DW technology. For a swift introduction we suggest the guides from Ballard et al. (1998) and Lane et al. (2005) and for going deeper into the subject the books of Ponniah (2001), Adamson (2010) and Kimball and Ross (2013).

2.2 Agricultural Policy Data Warehouse Design Issues

As Jukic (2006) notes, in DW design there are two main schools of thought: *a bottom-up approach* (Kimball et al., 2013), where subject-specific data marts are independently created. These are eventually integrated in a common “dimension bus” forming the data warehouse; *a top-down approach* (Inmon, 2002), in which the data warehouse (i.e. the collection of individual data marts) is built after the normalized enterprise data model has been set. To further clarify we will provide the agricultural policy relevant example, where we want to monitor/evaluate the agricultural policy of an EU country.

In the bottom-up approach, we would immediately proceed to creating a data mart for a specific policy measure. Let’s assume that due to data availability we choose to create a data mart for direct payments to farmers, using data from the national payment authority. The star schema contains a fact table about the amount of payments, while the included dimensions are: farm production orientation; farm’s region (in prefecture resolution), time of payment (in month resolution). After some time a request to monitor agro-environmental measures is coming, so we create a new data mart with the same procedure, catering for the alignment of the common dimensions, if possible.

For the case of the top down approach, the agricultural policy is fully analyzed and all desired reports and OLAP functionality are determined from the beginning. The relevant sources shall be established (e.g. see Table 1 for a list of Greek agricultural sources) and all relevant dimensions be included sketched in a *3-normal-form* dimension star schema. Then ETL procedures are applied, data marts are populated and OLAP functionality is provided to end users.

What is the most appropriate approach for building a DW for agricultural policy monitoring and evaluation? As discussed in the introduction, agricultural policy related data is connected with many distinguished dimensions and is derived from many independent sources of information. Thus in the top-down approach all costs are incurred at the beginning of the project. Furthermore, since many independent administrative departments and stakeholders are involved, project requirements may change, unpredicted problems may arise and thus the final outcome is uncertain. Instead a bottom up approach will produce usable results in short time, demonstrating the virtues of using DW technology to the stakeholders increasing the motivation for adoption from other departments too. In resume, the complexity of the stakeholder's structure that is most often found in agricultural policy context can be effectively managed with the bottom-up approach.

2.3 Design and implementation of policy related data marts

Regarding the design and implementation of a single data mart, using the bottom up dimensional model, we adapt the DW design lifecycle proposed by Ballard et al. (1998), as shown in Figure 2, and discuss the various steps in the agricultural policy context.

1) Subject and scope definition: In the beginning a single subject shall be defined. The question that is to be answered here is "what and why do I want to analyze". Example subjects are: "monitoring crop yields in response to climate change" (see following section); "the effect of a specific policy measure on the biodiversity of arable crop fields".

2) Data sources identification: After there is a clear idea on the what-why questions, it is easier to identify the relevant data sources. Regarding agricultural policy related data, apart from the traditional sources (e.g. statistical offices, etc.), there are some emerged data sources worth considering: geo-referenced agricultural and environmental data sources, like meteorological remote sensing systems, satellite images; computer applications where farmers can record their practices (Pinet and Schneider, 2010); data from agricultural institutes and projects (Janssen et al., 2012). In general, there is a variety of data sources that can be proved useful, for example see Table 2 for the case of Greek agriculture. Also the use of web services can provide easy access to that data.

3) Star schema creation: Regarding agricultural policy context, certain dimensions are expected to appear often: Type of activity, temporal and administrative unit classification. Their organization in a hierarchical way will facilitate the data analysis phase and the related operations (drill up/down, slice, etc.). Other dimensions are also expected to be present (farm size, various categorization of farming types, etc.) and should be catered accordingly. Also the logical conformation of the dimensions takes place in this stage and a relevant discussion on dealing with such issues is made in Nilakanta et al. (2008).

A good systematic method for the conformation of a dimension that is common in different data sources is depicted in Figure 3. A directed graph is constructed where each node represents a resolution level of the dimension. There is also a directed hierarchical positioning of the various dimension levels (e.g. municipalities are connected to regions, regions to countries, countries to continents, etc.). If there are dimensions levels of different data sources that are exactly the same, then they are both written within the same node in a way that the information of the source is maintained (e.g. {LAU-1} and [Municipalities], where {} notes that the level name is found in Eurostat data source and [] in Greek statistical office). Any acyclical subgraph that starts from the lowest level node and ends to the highest can be a hierarchical conformed dimension for two or more data sources, as long as they are present in at least one node.

To clarify, in Figure 4, there are three such subgraphs: 1,2,3,5,6,7,9,10 which connects Eurostat and Greek Statistical Office hierarchy with [Prefectures] to be a missing level; 1,2,3,4 ,6,7,9,10 which does the same with {NUTS-3} be missing; and 1,2,3,5,6,8,10 which connects the above data sources and FADN source. A researcher will select the dimension merge that is more suitable for his needs (i.e. missing dimension is not important) and there is also the option of attaching more than one merged dimension to the data.

4) Design and application of ETL processes: Due to the variety of data sources, ETL design is expected to take a significant proportion of the project's workload. If the data was originally produced for operational use (an example can be found in Schulze et al., 2007) it is very possible that the time definition will be missing (i.e. only updated data will be present) and shall be explicitly inserted in this step. Also the application of the ETL process may reveal weaknesses on the data quality of certain organizational units (e.g. maintained in plain excel files, no remote retrieval methods, etc.) and the provided feedback would improve the overall IT infrastructure of the policy structure.

5) Design and creation of OLAP cubes: Due to the existence of different users with different needs, the access to the OLAP cube, ideally should be provided through a web interface with the ability to export in various formats and also through automated retrieval offering web service access.

3 Case Study: A Data Mart for Monitoring Yields and Climate Data

In order to demonstrate the power of using DW technology we apply the previously discussed design workflow in a climate change case study. Our implementation uses free licensed tools: MySQL[2] as DW storage data base, Kettle[3] to facilitate collection, transformation and loading of data and Mondrian[4] to create the OLAP cube and apply data analysis and SpagoBI[5] server to enable the execution of the OLAP cube.

We discuss in detail the implementation of the already presented design workflow:

Step 1, Definition of subject and scope: We want to monitor crop yields, weather conditions and their relation in order for policy makers to anticipate any climate change effects. Due to budget and time constraints we selected to monitor and evaluate two regions: Thrace (8,578 km^2) which is the northernmost Greek region and Crete (8,303 km^2) which is the southernmost one. Both areas cover about 20% of the national utilized agricultural area. We also decided to focus on certain crops: cereals, cotton, tobacco and olives. Our selection diversifies the geographical and product scope so that the results are representative enough.

Step 2, Data sources identification: A major criterion for selecting a data source was, apart from being relevant to our subject definition and scope, to be publicly accessible. See Table 2 for a list of candidate data sources. We selected the "Annual Agricultural Statistical Survey" since it provides annual data in fine grained administrative resolution (after our special request) for production volume and crop areas and thus detailed crop yields can be derived. We also use meteorological data (rain height, temperatures, etc.) recorded from

[2] https://www.mysql.com/

[3] http://community.pentaho.com/projects/data-integration/

[4] http://community.pentaho.com/projects/mondrian/

[5] http://www.spagobi.org

the National Observatory of Athens that maintains a network of 347 scientific meteorological stations[6] all over Greece, since it is the most complete source of this kind of information.

Step 3, Star schema creation: In this step it is crucial to identify and deal with any data peculiarities (e.g. how missing values are treated, is there any implicit dimension hierarchy, etc.), either through studying metadata or contacting the authority that supplied them. A list of the relevant dimensions and measures of the data sources will facilitate the creation of the star schema and the application of the previously described methodology will also be helpful. In our case this list is provided in Table 2. Considering missing values of certain dimensional data, the policy was to ignore the whole row of data. In the case of missing fact data adopted a null value which resulted in exclusion by aggregation calculations.

Dimension conformation was not so difficult for time and administrative dimensions. Regarding time dimension, there is a clear hierarchical relationship (day, month, year). As far as spatial (in meteorological data) and administrative unit (in agricultural statistics) dimensions, again a specific point in space can be clearly attributed to a municipal entity. On the other hand the dimension hierarchy shall be carefully crafted because any future data mart development will overall be more efficient if it is based on the existing dimensions. Thus, before concluding, we investigated other potential future data sources (see Table 1). This is more evident in the production activity dimension, where there are at least three different nomenclatures (Eurostat NACE-2, Eurostat LUCAS, Greek Statistical Office). We based our hierarchy to the Greek Statistical Office but inserted latent levels so that in the future other nomenclatures can be merged where possible.

Finally three star schemas were created: standalone production; standalone meteorological data; and another one for combining them. For the latter (Figure 6) we faced a certain challenges and thus provide more details: Yearly area, production volume and yield can be directly derived from production data. Mean and low temperature, rain mm per day is also directly available from weather data. The challenge was that the time dimension granularity was incompatible between the two data sources (weather on daily and production on yearly basis). An OLAP cube can successfully deal with this by aggregating (averaging, summing or giving the minimum) the finer detailed data (weather) to the least detailed one (production), i.e. present meteorological aggregated data in year resolution. But due to the impact of within year weather conditions to the overall behavior of crops (e.g. a very rainy summer could dramatically decrease/increase yields even if the year average was close to normal), this would result in a significant loss of information. Thus 36 additional measures were calculated: one for each month of a year and for each of the direct measures (mean and low temperature, rain mm), i.e. 12 months times 3 measures. To clarify even further a subset of those 36 measures is: mean temperature of January; mean temperature of February; ...; mean temperature of December; lowest temperature of January; ...; rain mm of December. Consequently in the combined star schema a total of 39 facts are included.

Step 4, Design and application of ETL processes: Regarding the Annual Agricultural Statistical survey the problems we had to deal with was the relatively poor data connection interface (data was provided in excel files) and the fact that administrative coding schemes between 2000-2010 and 2011-2012 periods were different. Regarding meteorological data connectivity was also an issue, as they were provided through plain

---

6 http://meteosearch.meteo.gr/

text files with fixed format but inconsistent across years and recording stations (data on some text files was starting on 11th row while for others on 12th and a relevant problem for data column start). Kettle was a valuable tool and handled efficiently the whole process, although meteorological files were more convenient to be downloaded through an http client and pre-processed with awk scripts. As can be seen in Figure 5, the Kettle transformation takes as input several Excel files with the raw data, combines them with other Excel files containing dimensional data from the previous step, and after numerous transformation steps gives as a result the final dataset which can be uploaded into a database table.

After loading data we observed that aggregated production volume for certain crops was absurd, revealing a false step in transforming the original data. In any case a last validation step is required so as to ensure that loaded data are error-free. Comparing official aggregated data with aggregation from the loaded data is an easy way of validating the ETL process. Overall ETL was designed so as any future data additions (e.g. additional regions or years) to be conveniently imported in the DW engine.

Step 5, Design and creation of OLAP cubes: The final step is the creation of the OLAP cube providing further data analysis capabilities. For each star schema, an OLAP cube has been created using the Mondrian server (a relational OLAP engine) connected to a MySQL database using the JDBC protocol. The OLAP cube contains the necessary metadata so that users can efficiently navigate through the different dimensions and aggregate data at different granularity levels.

The execution of the OLAP cubes has been carried out via the SpagoBI analytical server. For each cube one schema instance has been created containing representative dimensions with characteristic granularities and some of the measures. The analytical server is flexible enough to provide dimensions in any combination and granularity with the available measures. The user can easily select the required granularity of any dimension and the measures related to his/her own view of the data and apply to the cube the analytical capabilities described in section 2.1 (slice, dice, etc.). Moreover, any user, having the appropriate knowledge of the MDX query language, is able to create new calculated measures or to apply specific filters to the data. Furthermore, SpagoBI provides the tools to represent the data graphically or export specific instances of analysis. An architectural overview of the OLAP creation components is provided in Figure 7.

Within the analytical server, there is a tradeoff between the dimensions granularity and the speed of analysis. For each new combination selected, the data should be aggregated and represented after scanning the whole dataset. When the fact table contains huge number of rows, then the query, to aggregate data at finer dimension levels, takes longer time, because of the need to scan full dataset. The SpagoBI analytical server is able to store these aggregations in a memory cache, to facilitate subsequent analyses. These data are useful during one analysis session and for this reason the cache cleared after any restart of the server. The performance issue, with huge fact tables, can be solved by building aggregate tables, which contain pre-calculated summary data. The build of aggregate tables should be a part of the ETL process that populate / refresh the data warehouse.

Figure 8 provides an instance analysis report of several meteorological measures aggregated at the year level of the time dimension. The report is coming from the analytical server and offers a high level view, comparing the measures between two specific regions of Greece.

4 Conclusions

The utilization of Data Warehouse technology is valuable for tackling certain agricultural related data characteristics: a multitude of sources of very different nature and of independent origin; dimensions are mostly hierarchical; temporal and spatial aspects are relevant; fine granularity of data is useful; different end user requirements.

Regarding the design of policy related DW we argue that the bottom-up approach is far more convenient compared to the top-down. In the top-down approach all DW development costs are incurred at the beginning of the project while the bottom up approach will produce usable results in short time, increasing the motivation for adoption from policy stakeholders. Thus for implementing a policy related DW single subject data marts can be incrementally setup and implemented.

For creating a data mart, we propose the following steps: Define a single subject to investigate based on the question “what do I want to analyze and why”; identify all relevant data sources; design and implement a star schema that will cater for connecting the facts and their dimensions dealing with possible hierarchy discrepancies using the proposed technique; design and implement the necessary ETL procedures that will fetch data from the selected sources, transform them accordingly and load them into the DW engine; design and create the necessary OLAP cubes for browsing the data.

In order to demonstrate the power of using DW technology we presented a climate change case study, where we pursue an answer to the question “is the yield of certain crops affected by any climate change effects”. Certain important conclusions can be drawn:

- The use of open source tools for providing a whole data mart solution was adequate in terms of configuration, performance and user experience efficiency.
- The use web services by data providers can facilitate the retrieval of the data. Otherwise an overhead data manipulation cost shall be expected.
- Data validation shall always be part of the ETL process especially if fine detailed granularity data are handled.
- Dimension hierarchy shall be crafted carefully, catering for any future data additions. Thus a good strategy is to review possible future data sources although not used in the current data mart creation.
- There is a tradeoff between performance and dimension granularity and we propose a certain solution for dealing with it

Conclusively in this paper we have successfully crafted a data mart for an agricultural policy related subject (monitoring relation of yields to climatic conditions), although various shortcomings were encountered (data quality and validation, performance, etc.). As far as the future work is concerned, the need for adding more policy related data marts will arise and any additional problems shall be confronted. A consolidation of data from Farm Structural Surveys, FADN micro-data will be very useful to agricultural policy modelers. Also the provided OLAP interface shall cater for automated data retrieval through web services. Finally publicly available spatial data calls for a better integration with quantitative data in the DW context.

References

Abdullah, A. and Hussain, A. (2006) ‘Data Mining a New Pilot Agriculture Extension Data’, *Warehouse Journal of Research and Practice in Information Technology*, Vol. 38 No. 3, pp. 229-249.

Adamson, C. (2010) *Star Schema. The Complete Reference*. McGraw Hill Professional.

Alston, J.M. and James, J.S. (2002) ‘The incidence of agricultural policy’, in Handbook of Agricultural Economics, Elsevier, Vol. 2, Part B, pp. 1689-1749, ISSN 1574-0072, ISBN 9780444510792, http://dx.doi.org/10.1016/S1574-0072(02)10020-X.

Ballard, C., Herreman, D., Schau, D., Bell, R., Kim, E. and Valencic, A. (1998) *Data modeling techniques for data warehousing*, IBM Corporation International Technical Support Organization.

Bédard, Y., Merrett, T. and Han, J. (2001) ‘Fundamentals of spatial data warehousing for geographic knowledge discovery’, *Geographic Data Mining and Knowledge Discovery*, Taylor & Francis, Vol. Research Monographs in GIS, pp. 53–73.

Boulil, K., Le Ber, F., Bimont, S., Grac, C. and Cernesson, F. (2014) ‘Multidimensional modeling and analysis of large and complex watercourse data: an OLAP-based solution’, *Ecological Informatics* Vol. 24, pp.90–106.

Casters, M., Bouman, R. and van Dongen, J. (2010) *Pentaho Kettle Solutions: Building Open Source ETL Solutions with Pentaho Data Integration*, Wiley Publishing, Inc.

Chaturvedi, K.K., Rai, A., Dubey, V.K. and Malhotra, P.K. (2008) ‘On-line Analytical Processing in Agriculture using Multidimensional Cubes’, *Journal of the Indian Society of Agricultural Statistics.* Vol. 62 No.1, pp. 56-64.

Elizabeth H.A. and Mattison, K.N. (2005) ‘Bridging the gaps between agricultural policy, land-use and biodiversity’, *Trends in Ecology and Evolution*, Vol. 20 No. 11, pp.610-616, ISSN 0169-5347, http://dx.doi.org/10.1016/j.tree.2005.08.011.

Inmon, W. (2002) *Building the Data Warehouse*, 3rd ed., Wiley.

Janssen, S., Kraalingen, D. Van, Boogaard, H., Wit, A. De, Franke, J., Porter, C. and Athanasiadis, I. N. (2012) ‘A generic data schema for crop experiment data in food security research’, *In International Environmental Modelling and Software Society (iEMSs) 2012 International Congress on Environmental Modelling and Software*.

Jukic, N. (2006) ‘Modeling strategies and alternatives for data warehousing projects’, *Communications of the ACM*, Vol.49 No.4, pp. 83-88.

Kimball, R. and Ross, M. (2013) *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed., John Wiley & Sons, Inc.

Lane, P., Schupmann, V. and Stuart, I. (2005) Oracle Database Data Warehousing Guide. *10g Release*, *2*(10.2).

Malinowski, E. and Zimányi, E. (2008) *Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications. Springer eBooks*.

Nilakanta, S., Scheibe, K. and Rai, A. (2008) 'Dimensional issues in agricultural data warehouse designs', *Computer Electronics in Agriculture.* Vol. 60, No. 2, pp.263-278. DOI=10.1016/j.compag.2007.09.009 http://dx.doi.org/10.1016/j.compag.2007.09.009

OECD (2016), *Agricultural Policy Monitoring and Evaluation 2016*, OECD Publishing, Paris. DOI: http://dx.doi.org/10.1787/agr_pol-2016-en

Pinet, F. and Schneider, M. (2010) 'Precise design of environmental data warehouses', *Operational Research*, Vol. 10 No. 3, pp. 349–369. http://doi.org/10.1007/s12351-009-0069-z

Pinet, F., Miralles, A. and Bimonte, S. (2010) The use of UML to design agricultural data warehouses. In *AGENG 2020 Conference*. Retrieved from http://hal.cirad.fr/hal-00525347/

Ponniah, P. (2011) *Data warehousing fundamentals for IT professionals,* John Wiley & Sons.

Rai, A., Dubey, V., Chaturvedi, K.K. and Malhotra, P.K. (2008) 'Design and development of data mart for animal resources', *Computer Electronics in Agriculture*, Vol. 64, pp. 111–119.

Rai, A., Malhotra, P.K., Sharma, S.D. and Chaturvedi, K.K. (2007) 'Data Warehousing for Agricultural Research - An Integrated Approach for Decision Making', *Indian Journal of Agricultural Statistics*, Vol. 61 No. 2, pp. 265-274.

Runge, C. F. (2006) *Agricultural economics: a brief intellectual history (No. 13649)*, University of Minnesota, Department of Applied Economics.

Schulze, C., Spilke, J. and Lehner, W. (2007) 'Data modeling for Precision Dairy Farming within the competitive field of operational and analytical tasks', *Computers and Electronics in Agriculture*, Vol. 59 No. 1-2, pp.39–55. http://doi.org/10.1016/j.compag.2007.05.001

Van Broekhoven, E. (2007) 'A new farm accountancy data network for Flanders (Belgium)', in Poppe, K.J. et al. (Eds.), *PACIOLI 15 workshop presentations: integration of farm accounting in research and statistics*. Accessed from http://library.wur.nl/WebQuery/wurpubs/fulltext/21927.

Yost, M. (2000) 'Data warehousing and decision support at the National Agricultural Statistics Service', *Social Sciences Computer Review,* Vol. 18 No. 4, pp. 434-441. DOI=10.1177/089443930001800406 http://dx.doi.org/10.1177/089443930001800406

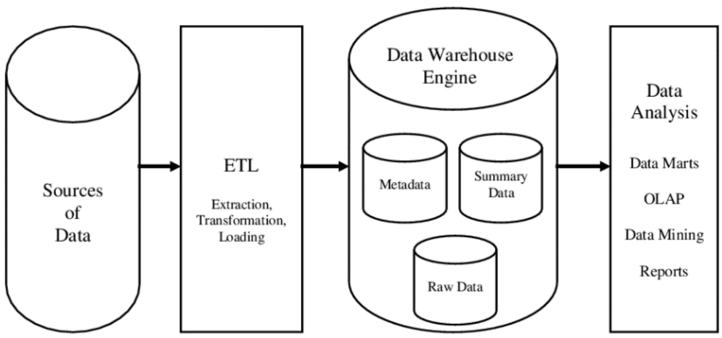


Figure 1 Data Warehouse lifecycle process

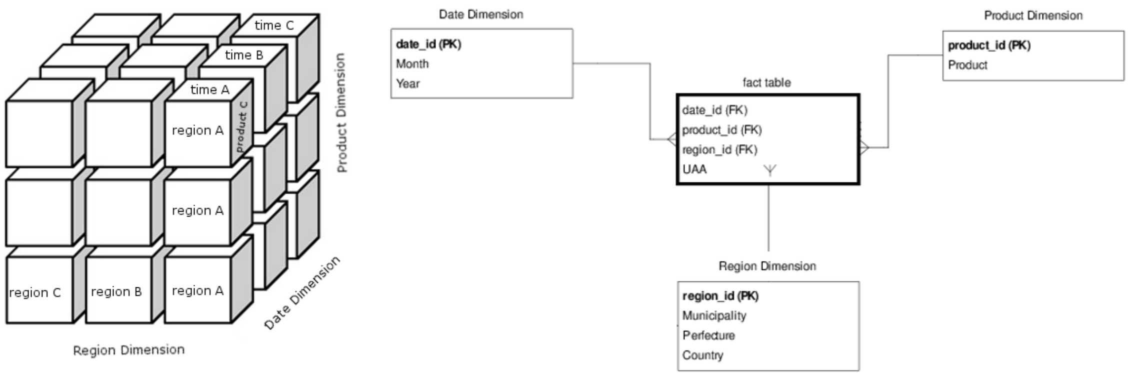


Figure 2, Data Warehouse main notions visualized

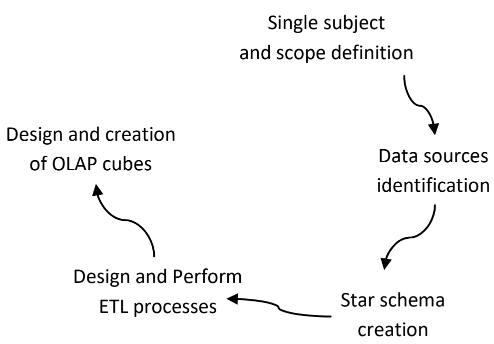


Figure 2, Data Warehouse design workflow (adapted from Ballard et al., 1998)

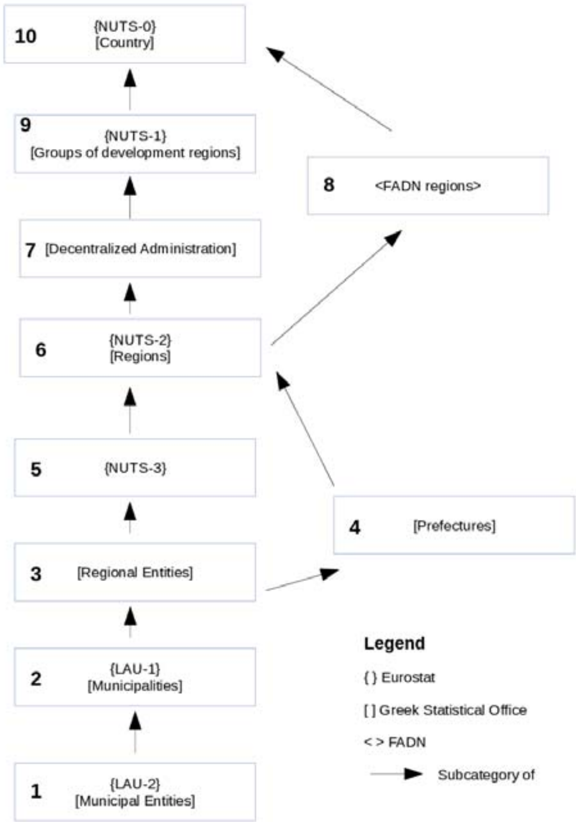


Figure 4, A graph facilitating the conformation of hierarchical dimensions connected to different data sources

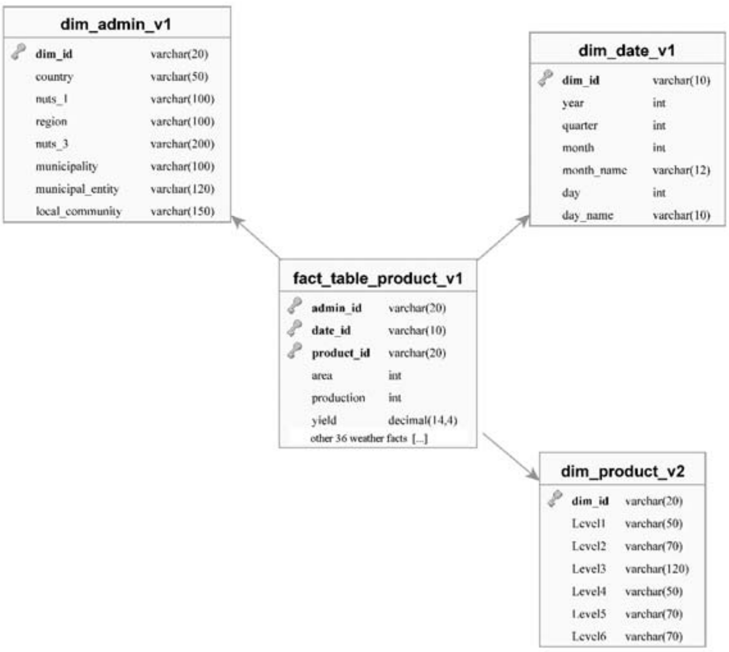


Figure 5, Star schema for the combination of meteorological and production data

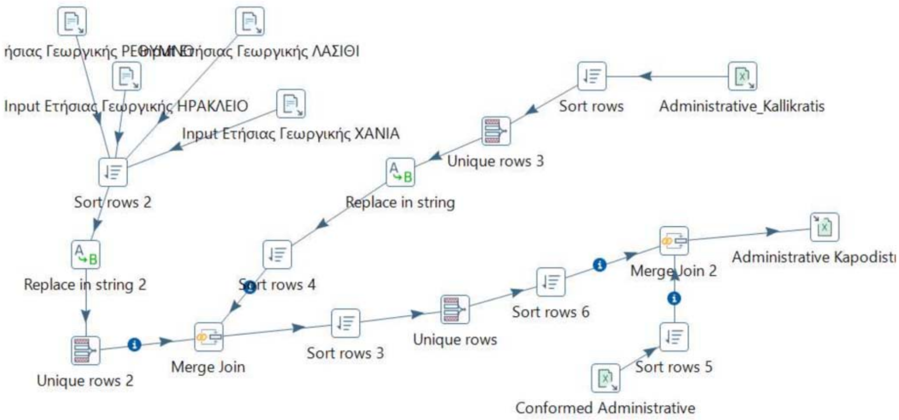


Figure 6, Kettle ETL procedure example

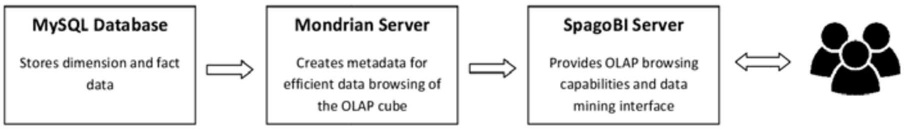
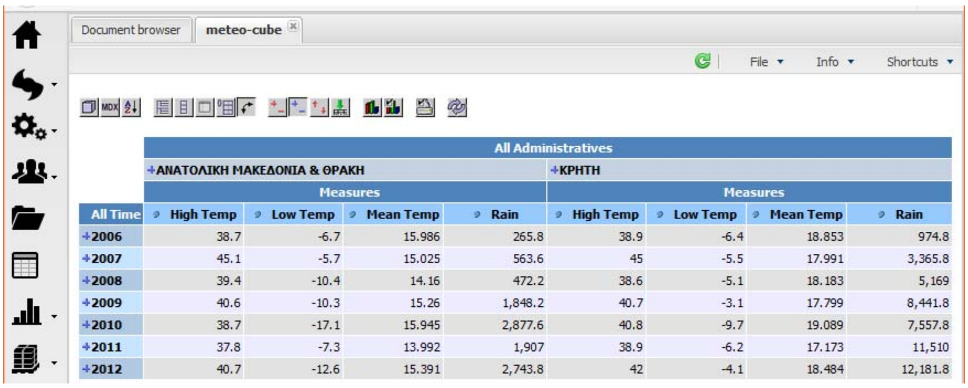


Figure 7, Architectural overview of the OLAP implementation components



	All Administratives							
	ΑΝΑΤΟΛΙΚΗ ΜΑΚΕΔΟΝΙΑ & ΘΡΑΚΗ				ΚΡΗΤΗ			
	Measures				Measures			
All Time	High Temp	Low Temp	Mean Temp	Rain	High Temp	Low Temp	Mean Temp	Rain
2006	38.7	-6.7	15.986	265.8	38.9	-6.4	18.853	974.8
2007	45.1	-5.7	15.025	563.6	45	-5.5	17.991	3,365.8
2008	39.4	-10.4	14.16	472.2	38.6	-5.1	18.183	5,169
2009	40.6	-10.3	15.26	1,848.2	40.7	-3.1	17.799	8,441.8
2010	38.7	-17.1	15.945	2,877.6	40.8	-9.7	19.089	7,557.8
2011	37.8	-7.3	13.992	1,907	38.9	-6.2	17.173	11,510
2012	40.7	-12.6	15.391	2,743.8	42	-4.1	18.484	12,181.8

Figure 8, An instance analysis report of meteorological data OLAP cube

Table 1. Sources of statistical information for Greek Agriculture

Provider	Data Series Name	Type	Starting Year	Frequency	Geographical Coverage	Finest Geographical Resolution	Data Included	Data Availability
EL.STAT.[1]	Census of Agricultural and Livestock Holdings	Census	1961	every 10 years	Whole of Greece	Municipal districts	number of plant and animal agricultural holdings and their properties regarding their legal status, agricultural and tenure status, structural properties (type of crops / animal / activity), production methods	1961,1971,1981,1991 in printed form 2000,2009 in electronic form
EL.STAT.	Annual Agricultural Statistical Survey	Survey	1961	Annual	Whole of Greece	Municipalities (as defined in the "Kapodistrias" law)	agricultural utilized land per type of crop, volume of agricultural (plant and animal) production, utilization of agricultural machineries	Online from 1961 – 2006
EL.STAT.	Farm Structure Survey	Survey	1966	1966, 1977, since 1983 every 2 years (but not 1991 and 2000), since 2010 every 3 years	Whole of Greece	Municipal districts	number of plant and animal agricultural holdings and their properties regarding their legal status, agricultural and tenure status, structural properties (type of crops / animal / activity), production methods	Online since 2003
EL.STAT.	Survey on Crop Production (including permanent cultivations and grape yards)	Survey		Grape yards: Yearly survey, grains and other crops / Basic survey every 10 years for grape yards / research every 5 years for permanent cultivations	Whole of Greece	Prefecture (NUTS-2)	Cultivating area per crop	Online since 2000
EL.STAT.	Agriculture Input and Output Price Index	Index	1967	Monthly	Whole of Greece	760 (output) and 783(input)price-collection-points, from all Greece	Index of output prices (subsidies and transport costs are excluded) for plant and animal products (as classified in European Economic Accounts) Index of input (products and services)	Online since 2001

Provider	Data Series Name	Type	Starting Year	Frequency	Geographical Coverage	Finest Geographical Resolution	Data Included	Data Availability
							prices	
EL.STAT.	Agriculture production factors' index (Cost Index)	Index	1975	Yearly	Whole of Greece	Whole of Greece / 155 points of price collection points	Index of production factor wage. It is comprised of three sub-indexes: labor (payment for one day), land (rent), and capital (loan interests and agricultural machinery rent)	Online since 2005
Greek Ministry of Agriculture	FADN / RICA	Survey	1985	Annual	Whole of Greece	Municipal District / ~ 4000 farms	Accountancy data	Fine detailed data is not freely distributed. Aggregated data is publicly available.
Greek Payment Authority of Common Agricultural Policy Aid Schemes (OPEKEPE)	Registry of Farm Subsidies	Registry	2013	Yearly	Whole of Greece	Plots	Crop type per plot basis	Not publicly available
National Observatory of Athens (Greece)	Public Database of Meteorological Measurements	Real Data	2006	Daily	Network of 347 scientific meteorological stations	Spatial Point	Mean temperature, min-max temperatures, rain, mean wind speed, dominant wind direction	Online as fixed width text files

Provider	Data Series Name	Type	Starting Year	Frequency	Geographical Coverage	Finest Geographical Resolution	Data Included	Data Availability
European Environment Agency	Corine Land Cover	Spatial	1990	1990, 2000, 2006	Whole of Greece	25 hectares	44 classes of land use	Online
European Soil Data Centre (ESDAC)	European Soil Database	Spatial	2001	-	Whole of EU-27	1:1,000,000	Soil related data	Online
EUROSTAT	Land Use and Coverage Area Frame Survey (LUCAS)	Spatial	2006	3-years	Whole of EU-27	270000 points in EU-27	land cover, land use and environmental parameters associated with the individual points surveyed	Online
EUROSTAT	TRADE Database (COMEXT)	Detailed Data	1976	1976 – 1987 is annual, since 1988 is monthly	Whole of EU-27	Intra is from direct collection of information from trade operators / Extra is from custom declarations	Value and quantity of goods traded between EU Member States (intra-EU trade) and between Member States and non-EU countries (extra-EU trade)	Since 2004 are free of charge http://ec.europa.eu/eurostat/web/international-trade/data/database

[1] Hellenic Statistical Authority (EL.STAT.)

Table 2, List of case study dimensions and measures

Annual Agricultural Statistical Survey (Greek Statistical Office)	Meteorological data (National Observatory of Athens)
Dimensions: 1. Time (in year resolution) 2. Administrative unit (in municipal entity resolution. 2000-2010 in “kapodistrias coding scheme”, 2011-2012 in “kallikratis coding scheme” 3. Product code (as defined by Greek Statistical Office) Measures: 1. Area (in 0.1 ha) 2. Production volume (in kgs)	Dimensions: 1. Time (in day resolution) 2. Spatial (point in space) Measures: 1. High Temperature (in C) 2. Low Temperature (in C) 3. Average Temperature (in C) 4. Rain (in mm)

Paper: Why to regionalize CAP payments: A farm modeling approach

Submitted to Land Use Policy, 2019

Why to regionalize CAP payments: A farm modelling approach

Dimitris Kremmydas[1,*], Athanasios Kampas[1], Stelios Rozakis[2,3], Konstantinos Tsiboukas[1], George Vlahos[1]

[1] Dept. Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

[2] School of Environmental Engineering, Technical University of Crete, Greece

[3] Dept. Bioeconomy and Systems Analysis, IUNG, Poland

[*] Corresponding author, kremmydas@aua.gr

Abstract The 2013 CAP reform meant to reduce the variability in the per hectare payments between and within Member States, aiming at converging to an EU-wide uniform payment. However, the Member States were given the option of regionalizing CAP payments, i.e. to differentiate the per hectare payment amongst country regions on the basis of socio-economic or agronomic criteria. By means of a farm model we evaluate the performance of a regionalization (distributive oriented policy setting) versus the case of a country flat rate (procedural oriented policy). We show that in the case of Greece, regionalizing CAP payments can cope better with the CAP conflicting objectives that reside simultaneously on economic, social and environmental dimensions.

Keywords: CAP; regionalization; farm model; Greece; distributive fairness

1. Introduction

The 2003 Mid-term review of the Common Agricultural Policy (CAP) mainly introduced the Single Payment Scheme (SPS), a decoupled direct payment to farmers that replaced the subsidies coupled to production applied at that date. The SPS per farm was calculated on the basis of entitlements distributed to farmers according to the eligible agricultural area they managed. The 2003 reform provided the Member States (MS) the discretion to choose between three different SPS implementation options: a historic (farm specific) model, a regional (flat-rate) model and a hybrid version which combines rationales of the previous two ones (Ciaian et al., 2014).

Under the historic model, transfers to farmers via the SPS equal the financial support received through various subsidies, in the “reference period” (2000-2002), maintaining thus intact the past distribution of payments across farmers. By contrast, under the regional model all farms in a specific region receive the same flat rate payment per hectare. The major difference between the historic and the regional models concerns the unit value of the entitlements with the aggregate amount of subsidies equal at the national level. In turn, the 2013 CAP reform replaced the SPS with the basic payment scheme (BPS). BPS provides a basic layer of support to farmers which is combined with other payments targeting specific issues such as the greening payments, redistributive payments, payments for Areas of Natural Constraints (ANC), the young farmer scheme, and the Decoupled Complementary National Direct Payments (Caian et al., 2018)).

Furthermore, the 2013 CAP reform meant to reduce the variability in the per hectare payments. In particular, the MS that had applied the historical or the hybrid SPS model, should adjust the unit value of the entitlements towards a more uniform level while the rate of this adjustment was left to national/regional (Caian et al., 2018). Such an internal convergence of the payments per hectare represents one of the core elements of a political agreement in the EU.

Furthermore, the MS had the option to differentiate the unit value of the BPS amongst different groups on the basis of socio-economic or agronomic criteria. Henceforth, we refer to such differentiation as *regionalization.* In fact, six MS (Germany, Spain, France, Finland and United Kingdom and Greece) have regionalized the BPS, with Greece being the only MS to apply purely agronomic criteria according to historical land uses, namely arable land, grassland and permanent crops. On the contrary, the former five MS have used territorial criteria (Henke et al., 2015).

The EU commitment to reach an equitable distribution of farmers' support, as jointly captured by external and internal convergences in the direct payments, represents a plausible choice that epitomizes concerns of procedural fairness (Commission, 2015). Typically, procedural fairness refers to transparent and impartial rules ensuring that each and every member enjoys an equal opportunity to obtain a satisfactory outcome (Krawczyk, 2011). In other words, procedural fairness concerns the distribution of resources, or the income transfers in terms of this paper. Note that procedural fairness along with distributive fairness constitute social justice (Miller, 1999), where distributive fairness primarily concerns the fair allocation of the outcome, e.g. in proportion to agents' claims (Rescher, 2002) Occasionally claims concerning the superiority of procedural fairness appear in the scholarly literature(van den Bos et al., 2001), however there is no consensus that such claims are generally accepted see Lucas et al. (2015).

Against the rationale of procedural fairness, which is conceived as a sign of political acceptability by some EU policy circles see Bureau et al. (2012), this paper proposes that distributive fairness should not be given lesser priority. In fact by means of a farm model we evaluate the performance of a distributive oriented policy setting (regionalization) versus a procedural one (country flat rate). Furthermore we show that regionalizing CAP payments can cope better with the CAP conflicting objectives that reside on different dimensions (economic, social and environmental).

The rest of the paper is organized as follows. Initially we discuss the utilization of farm models for decoupled payments analysis (section 2.1); then provide more details on important aspects of our farm model (sections 2.2 and 2.3); examine how to model regionalization (section 2.4); present and discuss the policy objectives and the metrics that represent them (section 2.5). In section 3 we present the results and discuss the performances of regionalization versus country wide flat rate, concluding in section 4.

2. Materials and methods

2.1 Farm models and decoupled payments

A clear advantage of using a farm model, rather than an accounting approach, stems from the fact farm models are designed to assess both first round effects (income changes) and second round ones (that is the income changes brought about by changes in the cropping pattern induced by wealth effects). Nonetheless, a typical objection against the use of such models for assessing the impacts of policy reforms is that decoupled payments, although having an income effect, they do not distort input choices, i.e. they are allocative neutral (Sinabell et al., 2013). That is to say that decoupled payments do not distort the opportunity cost of resources. As a result, it is the very nature of decoupled payments that renders farm models less suitable for similar policy evaluation since typically they are regarded as a lump-sum transfer to households (Urban et al., 2016). Chambers and Voica (2017) argue that direct payments although being distortionary they primarily affect marginal consumption and leisure choices rather than production choices.

However, against that conventional wisdom, a growing body of literature has emerges which examines how decoupled payments may affect production choices, (Hennessy, 1998) has identified two channels, namely the wealth effects and the insurance effects, through which the distribution of direct (decoupled) payments affect producers' choices , as. Suffice to say that these choices affect the outcome, i.e. they have distributive implications. The main empirical findings of the scholarly literature are twofold. First, decoupled payments affect the relative prices of resources. Ciaian et al. (2014) examine how the implementation details of 2003 CAP reform affect the capitalization of decoupled payments on land prices, while Pavel et al. (2018) examine the same issue for the 2013 CAP reform. Likewise Patton et al. (2008) examined the same issue for Ireland. Notwithstanding, other studies found low and partial capitalization of income support to land prices see Ciaian and Kancs (2012) and O'Neill and Hanrahan (2016). Graubner (2018) examine the conditions under which such a capitalization is possible and who is benefited (farmers or landowners).

Second, decoupled payments affect either the resources' availability (e.g. through credit constraints see Moro and Sckokai (2013)) or the relative risk of income sources (e.g. decoupled payments stabilize farm revenues see Schmid and Sinabell (2007)). O'Toole and Hennessy (2015) found that the higher the proportion of income earned from risk free decoupled transfers, the less pressure credit constraints impose on farmers and therefore their investment decisions are affected by these modified credit constraints. Finally, Martinez Cillero et al. (2017) have found that decoupled payment, through the channel of investments, may have a positive role on farm's technical efficiency. Rizov et al. (2013) provide conditional evidence of such positive link between farms' productivity and decoupled payments. By stark contrast, Hailu and Poon (2017) argue that such a link is mostly negative in the sense that the less efficient farmers may receive a higher level of payment per unit of revenues.

In conclusion, farm models can be used to evaluate decoupled payment scenarios in two modes: (a) as a vehicle to see the accounting effects of altering the current status of transfer; for instance in the case where a detailed representative farm level database exists like FADN (b) to additionally capture the non neutral nature of the decoupled payments in the case where one or more of the aforementioned points are incorporated in the model.

2.2 Evaluation framework overview

This paper compares various options of regionalizing the BPS against the country wide uniform distribution of farmers' support. EU wishes to achieve such a uniformity in the near future. Although various regionalization choices are ranked on the basis of some transparent criteria drawn directly from officially policy documents, their design and implementation add a layer of extra complexity to the CAP decision-making and consequently result in higher (policy) transaction costs. Furthermore, a likely regionalization may provide the pressure groups with the room to exercise their lobbying activities and to exploit rent-seeking possibilities. Thus, a transparent and properly structured process which assesses the different regionalization options on the basis of sound and well defined criteria is an essential prerequisite. The proposed evaluation framework, presented in Figure 1, offers such an opportunity.

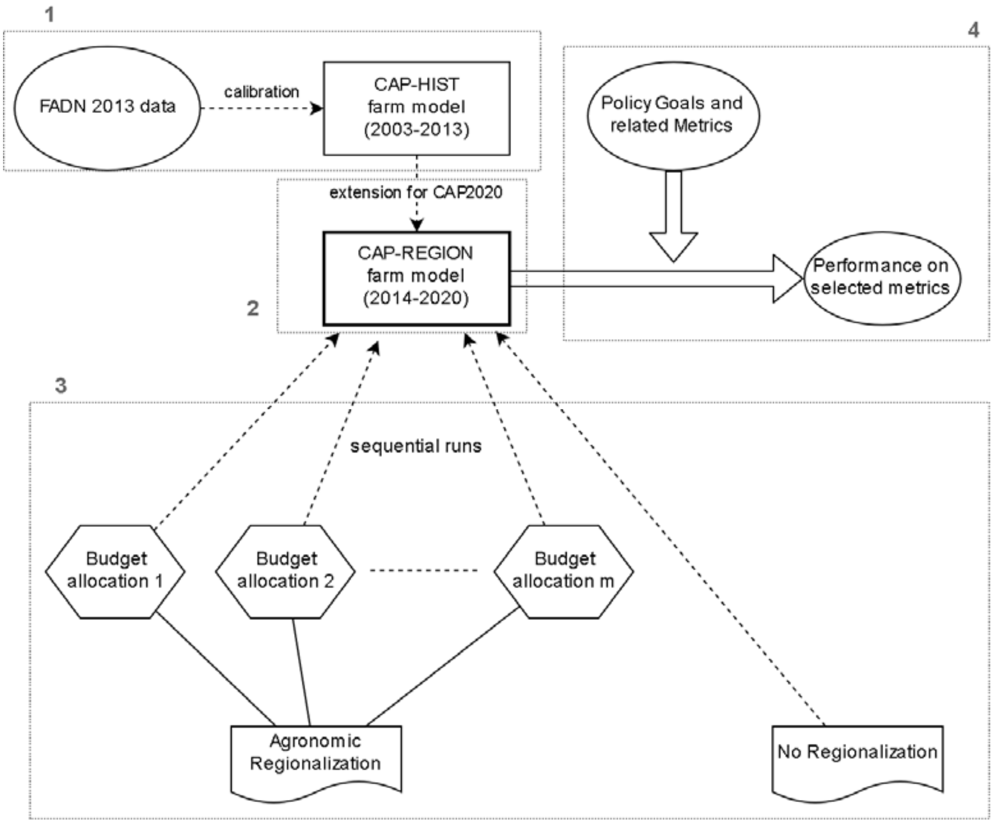


Figure 3, Overview of the evaluation method. For numbering explanation see section 2.

In particular, it comprises the following sequence of procedures (see Figure 3):

Box 1: A mathematical programming model based on the standard assumption of profit maximising farms. The model is calibrated for the 2013 FADN data. At that time the historical model was active, thus the policy constraints are set accordingly. We refer to it as CAP-HIST.

Box 2: In turn, the CAP-HIST is adjusted for the CAP 2014-2020 period adding greening constraints and facilitating regionalization payments. The extended model is termed as CAP-REGION,. Section 2.3 presents in a detailed fashion both CAP-HIST and CAP-REGION.

Box 3: The set of possible budget allocations between the current agronomic regions (arable, permanent crop and grassland) is parametrically created in 5% steps. This concludes to to 132 allocations. We also evaluate the scenario of a uniform country payment. Section 2.4 explains the regionalization rationale applied in this paper.

Box 4: It is self-evident that the evaluation of examined scenarios and its consequent ranking requires one or more assessment criteria. In so doing, the paper carefully selects eight indicative policy goals to serve as guiding criteria, based in the European Commission proposal for the 2021-2027 period (COM (2018) 393). Furthermore, appropriate metrics for each of these criteria are chosen. It should be noted that almost half of these metrics can only be derived by means of the farm modelling approach. Details on the utilized policy goals and metrics are given in section 6.

2.3 Details on the farm model

The model assumes that farms select an activity plan to maximize total gross margin (eq. 1). Vector c^T contains the activities' gross margins and x is the vector of the decision variables (areas of crops in ha, number of livestock heads). Farms are also subject to certain constraints (eq. 2) where the matrix A contains the resources' requirements per unit of an activity and the vector b represents the resource availability. The following constraints were explicitly taken into consideration: total land availability; irrigated land; labor availability; working capital constraint; permanent crop; livestock; crop rotations; as well as flexibility constraints such as existence of contract crops. A detailed account of the model is given in the Appendix B.

$$\max c^T x \tag{1}$$

$$s.t\ Ax \leq b \tag{2}$$

2.3.1 Working capital constraint

It is assumed that decoupled payments are partially or fully channeled to satisfy working capital requirements and the variation of decoupled payments (across different regional policy scenarios) affect the credit constraint of a farm. Working capital demand and supply at the farm level, annotated with FADN codes, is depicted in **Figure 4**

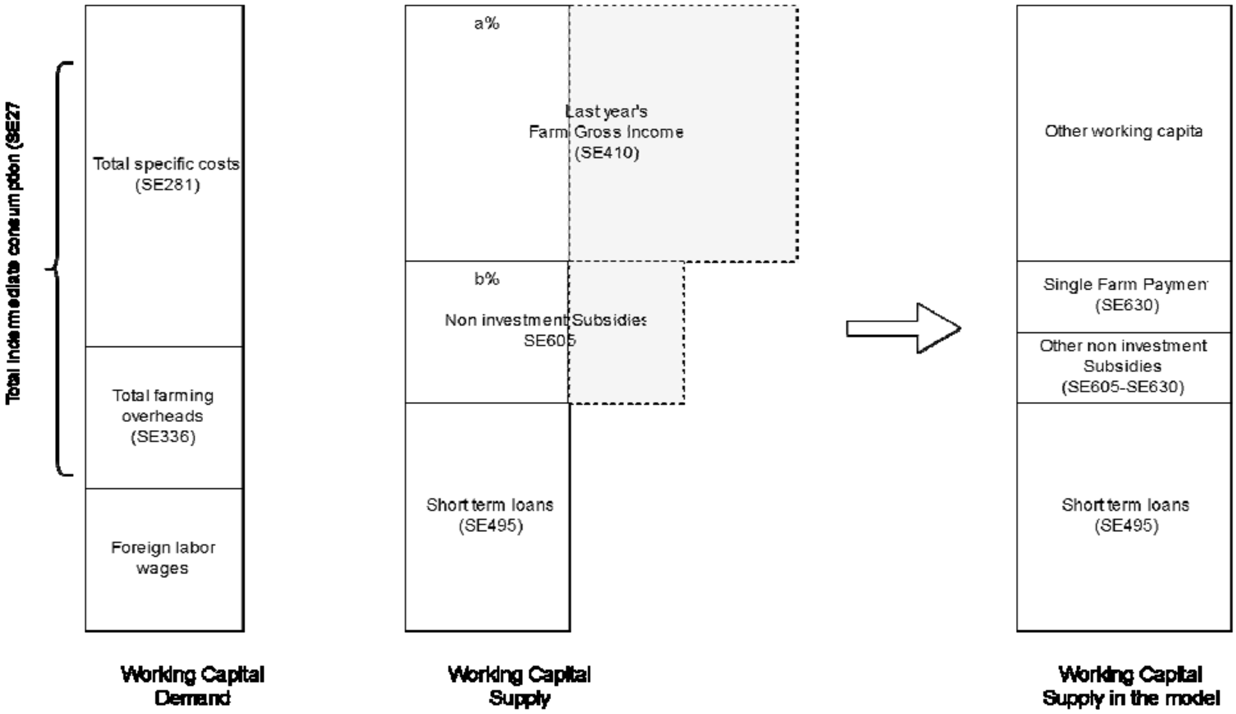


Figure 4, a conceptual diagram of the working capital definition. In parentheses the FADN standard result variable codes[7]

Total specific costs plus any farming overheads plus any hired labour wages are covered by a share of last year's farm gross income plus a share of transfers (including direct payments) plus any new short term loans. Since we cannot impute those shares we assume that direct payments are exclusively used to cover working capital requirements (see Figure 4). Thus, for the observed crop plan of each farm we calculate the demand for working capital and estimate the part of the working capital that is not coming from subsidies (*Other working capital* in Figure 2).

As shown in Table 1, the above assumption provides a fair approximation to the contribution of decoupled payments to the working capital supply. For the 2013 FADN Greek farms, the level of the non-investment subsidies/transfers is more than 50% of the total working capital demand. Thus reducing those transfers, although potentially can be covered by diverting gross margin from family farm consumption demand to working capital demand, will create a short-term credit strain to the farm. Moreover the transfers' contribution to working capital demand is essential for the Greek case, where the banking system and especially agricultural credit is malfunctioning after the 2009 crisis.

[7] See http://ec.europa.eu/agriculture/rica/annex003_en.cfm

Table 1, Percentage of non investment subsidies to total working capital demand in Greek FADN regions for 2013

Working Capital Elements (million €)		Macedonia, Thrace	Epirus, Peloponese, Ionian	Thessalia	Sterea, Aegean, Crete
Supply	Short term Loans (SE495)	8	0	1	5
	Non Investment Subsidies (SE605)	1039	378	382	967
	Previous Year Farm Gross Income (SE410)	2297	1602	805	2663
Demand	Intermediate Consumption (SE275)	1671	649	529	1191
	Foreign Labor Wages (SE370)	217	99	59	137
% of Non-investment subsidies to total working capital demand SE605/(SE275+SE370)		**56%**	**50%**	**65%**	**72%**

2.3.2 Extending CAP-HIST model

Regarding the extension of CAP-HIST model (time-span of 2003-2013) to the CAP-REGION version (the 2014-2020 period) the following steps and assumptions were employed:

a) The Single Payment Scheme (SPS) was replaced by the basic payment scheme (BPS). The SPS is different for each farm on the basis of the value of the historical entitlements allocated. On the other hand, BPS is a uniform payment per hectare for all farms that are within a single region. Section 2.4 provides details on how BPS is connected to regionalization.

b) CAP-REGION incorporates the so-called “greening measures” (i.e. crop diversification and ecological focus areas) and assumes farmers’ full compliance with these requirements.

c) Under the CAP-REGION model, the area of permanent crops remain the same as in CAP-HIST. Thus the required working capital for those activities is fixed. The only part of the working capital supply that can vary across simulations is the single farm payment and in this way can affect the working capital supply. *Other working capital*, *other non investment subsidies* and *short term loans* remain constant (see Figure 2)

2.4 Modeling regionalization

There are primarily three broad categories of regionalization schemes according to how the boundaries of the assigned zones (AZ) are defined (Solazzo et al., 2015). These are:

- An AZ may be defined on the basis of administrative criteria (e.g. NUTS 2 or 3), socio-economic or territorial environmental criteria (e.g. mountainous vs. non-mountainous areas). Each and every farm belongs solely to one AZ and the unit value of the received BPS is determined by the AZ attributes n which a farm belongs to (see Equation 3).

- An AZ may be defined using agronomic criteria (e.g. arable vs. permanent crops). In such a case one farm's land can belong to multiple AZs. Consequently, a farm's BPS unit value equals the weighted (by the area in each zone) average the different unit values of the AZs it belongs (see Equation 4).
- Finally, a AZ may be defined in hybrid way, fusing the above two rationales. Since agronomic criteria insert payment heterogeneity, it is evident that the hybrid case will do the same (see Equation 5).

Equations (3)-(5) define the association between the BPS and the examined regionalization schemes, while an illustrative example is given in appendix A.

administrative-based regionalization
$$BP_f^F = BP_{r(f)}^R \quad \forall f \qquad BP_r^R = BP_r^{BUDGET} \Big/ \sum_{f(r)} TL_f^E \tag{3}$$

agronomic-based regionalization
$$BP_f^F = \frac{\sum_g \sum_{c(g)} \left(BP_g^R \cdot X_{f,c}\right)}{TL_f^E} \qquad BP_g^R = BP_g^{BUDGET} \Big/ \sum_{c(g)} \sum_{f} X_{f,c} \tag{4}$$

Combined regionalization
$$BP_f^F = \frac{\sum_g \sum_{c(g)} \left(BP_{g,r}^R \cdot X_{f,c}\right)}{TL_f^E} \qquad BP_{g,r}^R = BP_{g,r}^{BUDGET} \Big/ \sum_{c(g)} \sum_{f(r)} X_{f,c} \tag{5}$$

where f the set of farms, r the set of regions, c the set of crop activities

BP_f^F: Basic payment unit value applicable to farm-f (euro/ha)

BP_r^R: Basic payment unit value applicable to administrative region-r, where $r(f)$ is the region of farm-f) (euro/ha)

BP_r^{BUDGET}: The Basic Payment budget for region-r, where $f(r)$ is the set of farms that belong to region-r (euro)

$BP_{g,r}^R$: Basic payment unit value applicable to agronomic region-g under administrative region-g, where $c(g)$ is the crop-set related to g (euro/ha)

$BP_{g,r}^{BUDGET}$: The Basic Payment budget for agronomic region-g under administrative region-g,(euro)

TL_f^E: Total eligible land for farm-f (ha)

$X_{f,c}^B$: Area of crop-c in farm-f in the reference period (ha))

2.5. Policy goals and metrics

A typical truism in the scholarly literature is that the efficacy of a policy is judged against its stated objectives. This is often referred to it as result-oriented evaluation (Burton and Schwarz, 2013) or outcome based judgment (Heinrich Carolyn, 2003). Therefore, this section explicitly presents the set of policy objectives against which the examined regionalization schemes are assessed. These policy objectives are primarily originated from an official EU document (COM (2018) 393 final). Suffice to say that for each policy objective an appropriate metric is carefully selected through which the derived results are evaluated. Table 2 presents these objectives and metrics.

Table 2, Policy Objectives

Type	Objective	Indicator/Metric	Units	Abbreviation	direction
Economic	*Enhance a smart, resilient and diversified agricultural sector ensuring food security*	Change in the farm gross profit for livestock farms	10^6 €	GP-LVST	max
		Change in the farm gross profit for specialist field farms	10^6 €	GP-FIELD	max
		A productivity proxy related to output-input value	n/a	PROD	max
Environmental	*Focus on environmental care and climate action*	Fallow land in non-LFA areas	10^3 ha	non-LFA FALLOW	max
		Area of irrigated crops	10^3 ha	IRRIGATED	min
Social	*Focus on the socio-economic fabric of rural areas*	The Gross Farm Profit of small farms	10^6 €	SM-FARM-GP	max
		Fallow land in areas facing environmental constraints	10^3 ha	LFA FALLOW	min
Political	Political Acceptability	% of farms that receive less than 75% of the baseline payments	%	MIN ENVY	min
		Gini coefficient of the distribution of the unit value of BPS	n/a	UNIFORM BP	min

The PROD metric is related to the output to input ratio, i.e. the value of production for one unit of input. Each regionalization scenario results in a distribution of output to input ratios, one for each farm. To calculate the metric we sort the distribution and get the 75% quartile. Thus PROD is a measure of the productivity of most productive farms. A lower value means that the most productive farm group has a lower productivity.

The nonLFA FALLOW objective captures the potential environmental beneficial role that fallow land may have. It primary refers to the beneficial contribution of fallow-land to biodiversity conservation (see (Banks-Leite et al., 2014; Herzon et al., 2011; Kovács-Hostyánszki and Báldi, 2012)) , it may also have a beneficial impact on climate through carbon sequestration (Piñeiro et al., 2009).

On the other hand, the LFA-FALLOW objective measures the maintenance of agricultural activity in areas with natural constraints. This coincides to a large degree with High Nature Value (HNV) farming systems and is considered essential not only in order to maintain the social vitality but also in order to enhance biodiversity (Terres et al., 2013) It is worth noting that in Greece, 92% of HNV pastures and 84% of cultivated land are found in areas with natural constraints. .

The MIN-ENVY objective, a political one with national significance, refers to maintaining the status quo of historical rights. The rationale of such an objective echoes the third proposition in (Zajac, 2001) where the status quo is considered as a right per se, where the divergence of which is considered unjust. In particular, there is a very diverse situation across farms regarding the basic payment value per hectare in Greece. Thus the immediate abolition of those differences would possibly lead to a politically troublesome situation in Greece. The proposed metric is the percentage of farms that receive less than 75% of the baseline gross pillar I payments.

The UNIFORM-BP objective which is also a political one but at the EU level, refers to officially declared goal of achieving a convergence of direct payments per hectare in the long-run[8]. It is estimated using the Gini inequality index of the payment per hectare distribution under different regionalization schemes. Suffice to say that the distribution with the lower inequality represents a better adjustment of the EU commitment towards facing the challenge of territorial imbalance across the Union (EC, 2017).

2.6. Data

The model utilizes all arable, vegetable and permanent crops and a major part of the grazing livestock (sheep and goats). By doing so, the model represents almost 2.8×10^5 farms (83% of FADN dataset); 2.5×10^6 hectares of utilized agricultural land (85% of FADN dataset) and 60% of the output value. It should be stressed, however, that while this sample is about the 50% of total farms which receive direct payments in Greece, it represents the vast majority of commercial farms since the rest are excluded from the FADN dataset as being too small.

The CAP-HIST version is calibrated for the 2013 FADN data. The calibration consists mainly on disaggregating farm variable costs to individual activities, eliciting the labour requirements of crop and livestock activities and adjusting global values like the nutrient content of produced and purchased feedstock. This results in a Mean Absolute Deviation (MAD) index for the cropping pattern

[8] Article 23 of Regulation 1307/2013

of 0.82. More details on the calibration of the model can be found in the model's online manual (Kremmydas, 2018).

Regarding the Pillar I budget, the FADN sample farms have received 32 million EUR representing payments of 1.4 billion EUR. We keep the represented payments budget for both the baseline model (CAP-HIST) and for the flat rate and regionalization scenario models (CAP-REGION). The unit value of each region is calculated as the total budget divided by the represented area of each region.

Concerning regionalization scenarios, we created them, based on the following grounds. First, the actual regionalization scheme for the period 2014-2020 in Greece that included three agronomic regions: arable crops, permanent crops and grazing areas, and second all the valid combinations of budget allocation to those regions with a 5% resolution. That resulted in creating 132 scenarios. For instance one scenario was the case where 45% of budget was allocated to arable regions, 15% to permanent crops and 40% to grazing land, etc.. The complete list of scenarios is given in appendix C.

Regarding the implementation of the model and the analysis of the results, the model input data were prepared in R using among others the fadnUtil package; the model was built and run in GAMS; and the analysis was performed in R using the rPref package for eliciting Pareto optimal scenarios (Roocks, 2016). For each scenario the optimization runs deliver the crop mix and associated values of indicators.

3. Results and Discussion

Presumably, the ranking of the examined scenarios requires an aggregation of their performance over the selected policy objectives. However, since the objective of this paper is to shed some light in the dilemma between procedural fairness (country wide flat rate) versus distributive fairness (regional allocation), the analysis did not apply any multi-criteria methods. The adopted procedure comprises the following steps.

First, for each policy objective, the performance range of the examined scenarios was divided into three equal drawn zones. Second, we consider all scenarios that fall within the same policy objective zone as being equivalent. The latter is probably a heroic assumption but is justified, however, on the basis that it does not bear the extra transactions costs to trace the differentiated benefits of alternative scenarios within the same zone. Put it differently, the extra transaction cost of targeting a specific policy within a specific zone may be higher than the missed efficiency losses of using horizontal policy choices. Third, each zone was assigned an ordinal score; 1 for the lower stripe, 2 for the middle and 3 for the higher one.

Figure 5**Error! Reference source not found.** summarizes and provides visual insights of the above steps. It is a 3x3 grid where each grid cell refers to a specific policy objective. In each grid cell, scenario performances are drawn as grey and black jittered points[9] and the overall distribution is given with a violin plot[10]; the red square is the flat rate scenario performance. The graph was created using the ggplot2 R package (Wickham, 2016).

For instance in the upper left corner of the grid the performances of all 132 scenarios in the policy goal of GP-FIELD (gross profit of field cropping farms) are drawn. The y-axis is the performance scale of GP-FIELD (in mil EUR) with a range from 1000 to 2000. The violin plot consists of a box plot (median approximately 1400, 25% and 75% quartiles are approximately 1250 and 1600) and a density plot that is symmetrical in the x-axis (the performance below the median gathers more probability density). Furthermore each regionalization scenario is represented as a point while te country wide flat rate is the red square. The scenario/point performance is valued along the y-axis, however close points are scattered across the x-axis so they do not overlap.

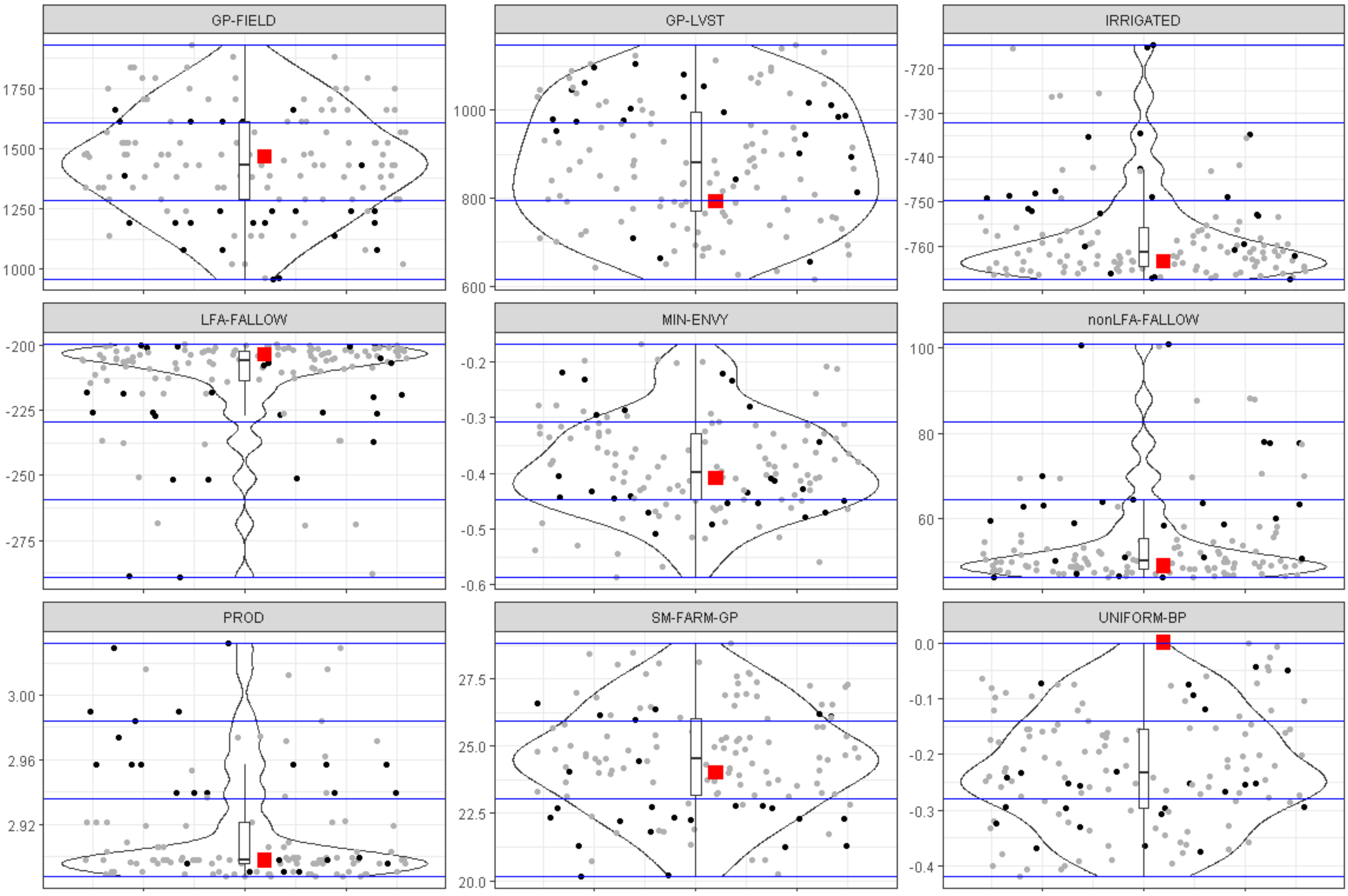


Figure 5, Scenario performance on individual policy objectives (see text for detailed description)

[9] Jittered points are scattered across the x-axis so that they are visually distinguishable, i.e.if two points have very similar performance they are placed in different x-axis positions without altering their y-axis position

[10] A violin plot is an extension of box-plot where the local density information is added thus displaying the distributional characteristics of data more clearly (Hintzee & Nelson, 1998)

If we consider goals independently, we see that flat rate is performing relatively moderate in four out of nine objectives (GP-FIELD, GP-LVST, MIN-ENVY, SM-FARM-GP); it excels in two (LFA-FALLOW, UNIFORM-BP) and has a poor performance in the rest three (IRRIGATED, nonLFA-FALLOW, PROD).

However since the selection of a specific policy setting has a simultaneous effect in all policy objectives, the above approach is not sufficient. Thus, using the ordinal scores assigned in step three, the non-dominated (Pareto optimal) scenarios across all objectives were identified. A non-dominated Pareto scenario is one where no other scenario has a better performance on all objectives while for a dominated (Pareto non optimal) scenario there exists at least another one that is performing better in one or more objectives, all the rest having equal performance. As a result a rational decision maker will only consider Pareto optimal scenarios. Thus we reduce the 133 scenarios (132 agronomic plus the flat rate) to 25 Pareto optimal scenarios, as given in Table 3.

Table 3, Non-dominated scenarios

	Scenario[1]	**GP-LVST**	**GP-FIELD**	**nonLFA-FALLOW**	**LFA-FALLOW**	**IRRIGATED**	**MIN-ENVY**	**UNIFORM-BP**	**SM-FARM-GP**	**PROD**
	flat rate	+	++	+	+++	+	++	+++	++	+
	Pareto optimal									
SC-1	AR:65/TR:30/GR:5, AR:70/TR:30/GR:0, AR:65/TR:35/GR:0	+	+++	+	+++	+	+++	+++	+++	+
SC-2	AR:25/TR:35/GR:40, AR:25/TR:30/GR:45, AR:20/TR:35/GR:45, AR:25/TR:25/GR:50, AR:25/TR:45/GR:30, AR:25/TR:40/GR:35, AR:20/TR:45/GR:35	+++	+	+	+++	+	+	++	+	++
SC-3	AR:20/TR:40/GR:40, AR:20/TR:30/GR:50	+++	+	+	+++	++	+	+	+	++
SC-4	AR:15/TR:45/GR:40, AR:10/TR:50/GR:40	+++	+	++	++	++	++	+	+	++
SC-5	AR:10/TR:45/GR:45, AR:10/TR:40/GR:50	+++	+	++	++	++	+	+	+	+++
SC-6	AR:0/TR:55/GR:45, AR:0/TR:50/GR:50	+++	+	+++	+	+++	+++	++	+	+++
SC-7	AR:65/TR:20/GR:15	++	+++	+	+++	+	++	+++	+++	+
SC-8	AR:20/TR:55/GR:25, AR:20/TR:50/GR:30	++	+	+	+++	++	++	+	+	++
SC-9	AR:70/TR:0/GR:30, AR:65/TR:0/GR:35	+++	+++	+	+++	+	+++	++	+++	+
SC-10	AR:45/TR:20/GR:35, AR:40/TR:25/GR:35	+++	++	+	+++	+	++	+++	++	+

[1] budget allocations refer to ARABLE/TREES/PASTURES. For example AR:40%/TR:55%/GR:5% refers to the scenario where 40% of the budget goes to arable crops, 55% to trees and 5% to Pastures

A straightforward result is that the flat rate scenario is not Pareto optimal. That means that there is one other scenario that outperforms it in one policy objective at least without being outperformed by it in the other objectives. Indeed, as seen in **Error! Reference source not found.** where the ordinal scores of the non-dominated scenarios are displayed, the SC-1 group of scenarios has similar performance to the flat rate scenario but outperform it in the GP-FIELD, MIN-ENVY and SM-FARM-GP objectives. Similarly the SC-7 scenario group outperforms flat rate to GP-FIELD, GP-LVST and SM-FARM-GP objectives and SC-10 is better in GP-FIELD, GP-LVST.

In the case where the flat rate scenario is fr seome reason preferred the above results can provide directions to policy makers for compensating the underperformance of specific objectives. In our case the flat rate scenario need to be complemented by coupled payments to field crops and additional support for small farms (e.g. redistributive payments).

4. Conclusions

Overall, the contribution of this paper is threefold. First, it provides a structured conceptual approach for policy analysis of regionalizing CAP. Second, by so doing, the policy design and its assessment is transparent and comprehensive, hence appealing for a broad consultation procedure of the involved agents.
The main conclusion of this paper is that the flat rate option, at least in the Greek case, is outperformed by regionalization policy settings. Nevertheless, the flat rate basic payment scheme does not address all objectives equally and shall be complemented by additional measures/schemes that will compensate (young farmers, etc.).

Also the use of a farm model to evaluate decoupled payments is a key demonstration of the paper. A pure accounting approach estimates statically the budget transfers and it does not consider the farms' production adaptation. Farm models can provide insights on the farm reactions to policy adapting crop mix. We went beyond that as we captured the non neutral nature of the decoupled payments. As a matter of fact we introduced the effects of the decoupled payment level in the credit constraint bound so that the adaptation of the farm's production choices were illustrated.

Considering that the new CAP design is pursuing multi-dimensional goals. extending to the economic, social and environmental domain the need for the employment of a multi-criteria multi-stakeholder evaluation environment becomes apparent. We aspire to further develop the current modeling framework towards this promising applied research field. The use of a DSS to assist such decision making process becomes a necessity, however it goes beyond the scope of the current work shaping the plan for future research (Kremmydas et al., 2019).

5. References

Banks-Leite, C., Pardini, R., Tambosi, L.R., Pearse, W.D., Bueno, A.A., Bruscagin, R.T., Condez, T.H., Dixo, M., Igari, A.T., Martensen, A.C., Metzger, J.P., 2014. Using ecological thresholds to evaluate the costs and benefits of set-asides in a biodiversity hotspot. Science 345, 1041.

Bureau, J.-C., Tangermann, S., Matthews, A., Viaggi, D., Crombez, C., Knops, L., Swinnen, J., 2012. The Common Agricultural Policy after 2013. Intereconomics 47, 316-342.

Burton, R.J.F., Schwarz, G., 2013. Result-oriented agri-environmental schemes in Europe and their potential for promoting behavioural change. Land Use Policy 30, 628-641.

Chambers, R.G., Voica, D.C., 2017. "Decoupled" Farm Program Payments are Really Decoupled: The Theory. American Journal of Agricultural Economics 99, 773-782.

Ciaian, P., Kancs, d.A., 2012. The Capitalization of Area Payments into Farmland Rents: Micro Evidence from the New EU Member States. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 60, 517-540.

Ciaian, P., Kancs, D.A., Swinnen, J., 2014. The Impact of the 2013 Reform of the Common Agricultural Policy on Land Capitalization in the European Union. Applied Economic Perspectives and Policy 36, 643-673.

Ciaian, P., d'Artis, K., Espinosa M., 2018. The Impact of the 2013 CAP Reform on the Decoupled Payments' Capitalisation into Land Values. Journal of Agricultural Economics 69, 306-337.

Commission, E., 2015. Direct Payments: The Basic Payment Scheme from 2015 - Convergence of the value of payment entitlements ('Internal Convergence'), in: Development, D.f.A.a.R. (Ed.), Brussels.

Graubner, M., 2018. Lost in space? The effect of direct payments on land rental prices. European Review of Agricultural Economics 45, 143-171.

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Hailu, G., Poon, K., 2017. Do Farm Support Programs Reward Production Inefficiency? Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie 65, 567-589.

Heinrich Carolyn, J., 2003. Outcomes–Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness. Public Administration Review 62, 712-725.

Henke, R., D'Andrea, M.R., Benos, T., Castellotti, T., Pierangeli, F., Lironcurti, S.R., De Filippis, F., Giua, M., Rosatelli, L., Resl, T., Heinschink, K., 2015. Implementation of the First Pillar of the CAP 2014 – 2020 in the EU Member States, in: Policies, P.D.B.S.a.C. (Ed.). European Parliament, Brussels.

Hennessy, D.A., 1998. The Production Effects of Agricultural Income Support Policies under Uncertainty. American Journal of Agricultural Economics 80, 46-57.

Herzon, I., Ekroos, J., Rintala, J., Tiainen, J., Seimola, T., Vepsäläinen, V., 2011. Importance of set-aside for breeding birds of open farmland in Finland. Agriculture, Ecosystems & Environment 143, 3-7.

Jerry L. Hintze & Ray D. Nelson (1998) Violin Plots: A Box Plot-Density Trace Synergism, The American Statistician, 52:2, 181-184, DOI: 10.1080/00031305.1998.10480559

Kovács-Hostyánszki, A., Báldi, A., 2012. Set-aside fields in agri-environment schemes can replace the market-driven abolishment of fallows. Biological Conservation 152, 196-203.

Krawczyk, M.W., 2011. A model of procedural and distributive fairness. Theory Decis 70, 111-128.

Kremmydas Dimitris (2018), GREFAM model manual, content available from https://www.researchgate.net/publication/329840868_GREFAM_model_Reference_manual, DOI: 10.13140/RG.2.2.13592.65285

Kremmydas D., Malliapis M., Nellas L., Polymeros A., Rozakis S., Tsiboukas K., (Forthcoming 2019) CAP 2020 regionalization design: A Decision Support System In Communications in Computer and Information Science, Springer

Lucas, T., Kamble, S.V., Wu, M.S., Zhdanova, L., Wendorf, C.A., 2015. Distributive and Procedural Justice for Self and Others: Measurement Invariance and Links to Life Satisfaction in Four Cultures. Journal of Cross-Cultural Psychology 47, 234-248.

Martinez Cillero, M., Thorne, F., Wallace, M., Breen, J., Hennessy, T., 2017. The Effects of Direct Payments on Technical Efficiency of Irish Beef Farms: A Stochastic Frontier Analysis. Journal of Agricultural Economics 0.

Miller, D., 1999. Principles of Social Justice. Harvard University Press, Cambridge.

Moro, D., Sckokai, P., 2013. The impact of decoupled payments on farm choices: Conceptual and methodological challenges. Food Policy 41, 28-38.

O'Toole, C., Hennessy, T., 2015. Do decoupled payments affect investment financing constraints? Evidence from Irish agriculture. Food Policy 56, 67-75.

O'Neill, S., Hanrahan, K., 2016. The capitalization of coupled and decoupled CAP payments into land rental rates. Agricultural Economics 47, 285-294.

Patrick Roocks (2016). rPref: Database Preferences and Skyline Computation. R package version 1.2. https://CRAN.R-project.org/package=rPref

Patton, M., Kostov, P., McErlean, S., Moss, J., 2008. Assessing the influence of direct payments on the rental value of agricultural land. Food Policy 33, 397-405.

Piñeiro, G., Jobbágy Esteban, G., Baker, J., Murray Brian, C., Jackson Robert, B., 2009. Set-asides can be better climate investment than corn ethanol. Ecological Applications 19, 277-282.

Rescher, N., 2002. Fairness : theory & practice of distributive justice. Transaction Publishers, New Brunswick.

Rizov, M., Pokrivcak, J., Ciaian, P., 2013. CAP Subsidies and Productivity of the EU Farms. Journal of Agricultural Economics 64, 537-557.

Schmid, E., Sinabell, F., 2007. On the choice of farm management practices after the reform of the Common Agricultural Policy in 2003. Journal of Environmental Management 82, 332-340.

Sinabell, F., Schmid, E., Hofreither, M.F., 2013. Exploring the distribution of direct payments of the Common Agricultural Policy. Empirica 40, 325-341.

Solazzo, R., Donati, M., Arfini, F., 2015. Cap towards 2020 and the cost of political choices: The case of Emilia-romagna region. Land Use Policy 48, 575-587.

Terres, J., Nisini, L., Anguiano, E., 2013. Assessing the risk of farmland abandonment in the EU in: Institute for Environment and Sustainability, J.R.C. (Ed.). European Commission, Ispra.

Urban, K., Jensen, H.G., Brockmeier, M., 2016. How decoupled is the Single Farm Payment and does it matter for international trade? Food Policy 59, 126-138.

van den Bos, K., Lind, E.A., Wilke, H.A.M., 2001. The psychology of procedural and distributive justice viewed from the perspective of fairness heuristic theory, in: Cropanzano, R. (Ed.), Justice in the workplace: From theory to practice, Vol. 2. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US, pp. 49-66.

Zajac, E., 2001. Political Economy of fairness. MIT Press, Cambridge.

Appendix A

We present a regionalization illustrative example. A hypothetical country with three farms and a total direct payments budget of 1000 euro is selected. The details for each farm (prefecture it belongs to, if the region is LFA or not, and direct payment rights) are given in the table below.

			Rights (ha)			
	Prefecture	LFA	Arable Crops	Tree Crops	Grazing Areas	Sum
Farm 1	A	Yes	10	2	1	13
Farm 2	B	No	3	1	5	9
Farm 3	C	No	1	10	-	11
		Sum	14	13	6	

We apply an administrative type of regionalization scenario, where Prefecture A is the region #1 with 30% of the budget and prefectures B and C are region #2 with 70% of the budget.

Regionalization Option I, Administrative based			
	Region 1	Region 2	
Definition	Prefecture A	Prefectures B+C	
Budget Allocation	30%	70%	
Budget (euros)	0.3*1000=300	0.7*1000=700	
Unit Value (euro/ha)	300/13=23.07	700/(9+11)=35	
	Farm 1	Farm 2	Farm 3
Farms SFP	13*23.07=	9*35=	11*35=

We apply an agronomic based scenario where arable crops, trees and grazing land are the three distinct regions.

Regionalization Option II, Agronomic based			
	Region 1	Region 2	Region 3
Definition	Arable Crops	Tree crops	Grazing Land
Budget Allocation	50%	30%	20%
Budget (euros)	0.5*1000=500	0.3*1000=300	0.2*1000=200
Unit Value (euro/ha)	500/14=35.7	300/13=23.07	200/6=33.3
Farms SFP	Farm 1 10*35.7+2*23.07+1*33.3=	Farm 2 3*35.7+1*23.07+5*33.3=	Farm 3 1*35.7+10*23.07=

Finally we apply a mixed regionalization type of scenario, where budget allocation is differentiated between arable crops, trees and grazing land and also whether the region is LFA or not.

Regionalization Option III, Mixed			
	Region 1	Region 2	Region 3
Definition	Arable Crops at LFAs	Tree crops and Grazing Areas at LFAs	Non-LFAs (all crops)
Budget Allocation	30%	20%	50%
Budget (euros)	0.3*1000=300	0.2*1000=200	0.5*1000=500
Unit Value (euro/ha)	300/10=30	200/(2+1)=66.66	500/(9+11)=25
Farms SFP	Farm 1 10*30+(2+1)*66.66=	Farm 2 9*25=	Farm 3 11*25=

Appendix B

Objective Function

We assume that farm selects a crop plan so as to maximize their gross income

$$TGP = GP^{market} + GP^{lvst} - WAGES + SUBS \tag{1.1}$$

Total farm Gross Profit equals Gross Profit (without wages paid) from crops sold at the market plus the Gross Profit (without wages paid) from the livestock activities minus wages paid for all activities plus subsidies received by the farm

$$GP^{market} = \sum_{c^m} \left(\left(PR_{c^m}^{market} \cdot YI_{c^m} - VC_{c^m} \right) \cdot X_{c^m}^{market} \right) \tag{1.1.1}$$

Gross Profits from market crops equal, for each crop, the gross margin per hectare (price times yield minus variable costs plus coupled payments).

$$GP^{lvst} = INC^{lvst} - FEED^{prod} - FEED^{purch} - EXP^{lvst} \tag{1.1.2}$$

Gross Profits from livestock activities equals the income from livestock activities minus the cost of in-farm produced feedstock (excluding wages) minus the cost of purchased feedstock minus any non-feedstock costs

$$INC^{lvst} = M_{a\_l}^{quant} \cdot M_{a\_l}^{price} + INC^{SELL} \tag{1.1.2.1}$$

Income from livestock equals the income from milk plus that from selling young animals

$$M_{a\_l}^{quant} = X_{ta^{a\_l}}^{animals} \cdot M_{ta^{a\_l}}^{YIELD} \cdot M_{a\_l}^{price} \tag{1.1.2.2}$$

The quantity of milk of type a_l (sheep or goat milk) equals the number of animals of type ta that produce milk a_l (the subset ta^{a_l}) times their yield times the price of milk a_l

$$INC^{SELL} = \sum_{ta} \left(X_{ta}^{animals} \cdot S_{ta}^{share} \cdot S_{ta}^{price} \right) \tag{1.1.2.3}$$

The income from selling young animals equals the sum of the number of animals of type ta times the observed percentage of young sold animals times their observed price

$$FEED^{prod} = \sum_{cf} \left(VC_{cf} \cdot X_{cf}^{feed} \right) \tag{1.1.2.4}$$

Produced feedstock cost equal, for each feedstock crop, the variable costs (excluding wages) per hectare minus any coupled payment per hectare times the area cultivated

$$FEED^{purch} = \sum_{fs\_purch} \left(PR_{fs\_purch}^{feed} \cdot F_{fs\_purch} \right) \tag{1.1.2.5}$$

Purchased feedstock cost equals the price of the feedstock (euro per tn) times the quantity purchased

$$WAGES = L^{foreign} \cdot WAGE \tag{1.1.3}$$

The wages paid are equal to the required foreign labor (in hours) multiplied by the wage. The required foreign labor equals the labor required for crop (market and feedstock) cultivation plus the required labor for breeding any animals minus the available family labor.

$$LAB^{crops} = \sum_{cm} \left(LB_{cm}^{CR} \cdot X_{cm}^{market} \right) + \sum_{cf} \left(LB_{cf}^{CR} \cdot X_{cf}^{feed} \right) \tag{1.1.3.1}$$

The labor required for crops equals the labor required for one hectare of a market crop (LB_c^R) times the cultivated area plus the labor required for one hectare of a feedstock crop times the corresponding areas

$$LAB^{animl} = \sum_{ta} \left(LB_{ta}^{AN} \cdot X_{ta}^{animals} \right) + \sum_{l\_a} \left(LB_{l\_a}^{ML} \cdot M_{a\_l}^{quant} \right) \tag{1.1.3.2}$$

The labor required for animal breeding equals the sum over all animals types of the labor per animal times the number of animals in this system

$$SUBS = SP \cdot LE + SO + \sum_{cm} \left(CP_c \cdot X_{cm}^{market} \right) + \sum_{cf} \left(CP_{cf} \cdot X_{cf}^{feed} \right) \tag{1.1.4}$$

The value of subsidies equal the unit value of the decoupled payment (Single Payment of Single Farm Payment) times the eligible land plus the value of pillar II subsidies plus the value of coupled payments

Constraints

Land, Labor and Capital

$$\sum_{cm}\left(X_{cm}^{market}\right)+\sum_{cf}\left(X_{cf}^{feed}\right)\leq LA^{T} \tag{1.2}$$

The total cropping area cannot exceed total land

$$\sum_{cm}\left(X_{cm\_i}^{market}\right)+\sum_{cf}\left(X_{cf\_i}^{feed}\right)\leq LA^{I} \tag{1.3}$$

The total irrigated cropping area cannot exceed total irrigated land

$$LAB^{crops}+LAB^{animl}\leq L^{family}+L^{foreign}\leq TOTAL\_LABOUR \tag{1.4}$$

The required labor shall be equal or more than the available family labor and less than the total available labor

$$\sum_{c}\left(VC_{c}\cdot X_{c}^{market}\right)+FEED^{prod}+FEED^{purch}+WAGES\leq SUBS+WC^{E}$$

The required working capital (left hand side) equals to the non-labor variable costs plus any foreign labor expenditure. This cannot exceed the sum of subsidies plus a farm excess working capital.

Livestock Constraints

$$SUPPLY_{n}^{CF}+SUPPLY_{n}^{FS\_PURCH}\geq NEED_{n}^{MAINT}+NEED_{n}^{MILK}\qquad\forall n \tag{1.6}$$

For each nutrient type, the nutrient supplied by purchased and produced feedstock shall cover the needs for maintaining the weight of the farm animals (kg)

$$SUPPLY_{n}^{CF}=\sum\nolimits_{cf}\left(CONT_{n,cf}^{C\_F}\cdot YI_{cf}\cdot X_{cf}^{feed}\right)\qquad\forall n \tag{1.6.1}$$

The nutrient supplied by produced equals the area of those crops times the yield times the content per ton of that crops

$$SUPPLY_{n}^{FS\_PURCH}=\sum_{fs\_purch}\left(CONT_{fs\_purch}^{FS\_PURCH}\cdot F_{cf\_purch}\right)\qquad\forall n, \tag{1.6.2}$$

The nutrient supplied by purchased feedstock equals the quantity purchased times the content per ton of that feedstock (kg)

$$NEED_{n}^{MAINT}=\sum_{ta}\left(NEED_{n}^{AN}\cdot X_{ta}^{animals}\right)\qquad\forall n \tag{1.6.3}$$

The maintenance demand for a nutrient equal for each animal type the per-animal maintenance needs times the number of animals of that type

$$NEED_n^{MILK} = \sum_{a\_l} \left(NEED_{a\_l}^{LT} \cdot M_{a\_l}^{quant} \right) \qquad \forall n \tag{1.6.4}$$

The milk production demand for a nutrient equal for each livestock product the needs of nutrient per unit of product (liter) times the quantity of production

$$SUPPLY_{"DM"}^{CF} + SUPPLY_{"DM"}^{FS\_PURCH} \geq \cdot \sum_{ta} \left(WEIGHT_{ta} \cdot X_{ta}^{animals} \right) \tag{1.7}$$

The Dry Matter Supply (DM) shall exceed the 1% of the total weight of all animals

$$X_{ta}^{animals} \leq CAP_{ta} \qquad (1.8$$

The number of animals of type ta cannot exceed the observed number of animals in the farm (capacity)

Flexibility Constraints

$$LB_{c^m} \cdot LA^T \leq X_{c^m}^{market} \leq UB_{c^m} \cdot LA^T \qquad \forall c^m \in flex\_m \tag{1.2}$$

Certain market crops (flex_m: sugar beet and sunflower) are bounded to be within a range of their observed area

$$LB_{c^f} \cdot LA^T \leq X_{c^f}^{feed} \leq UB_{c^f} \cdot LA^T \qquad \forall c^f \in flex\_f \tag{1.3}$$

Certain feedstock crops (flex_f: irrigated nd non-irrigated fodder other) are bounded to be within a range of their observed area

Appendix C,
List of
regionalization
scenarios

	Arable	Trees	Grazing		Arable	Trees	Grazing		Arable	Trees	Grazing
c1	100%	0%	0%	c34	40%	50%	10%	c67	50%	25%	25%
c2	95%	5%	0%	c35	35%	55%	10%	c68	45%	30%	25%
c3	90%	10%	0%	c36	85%	0%	15%	c69	40%	35%	25%
c4	85%	15%	0%	c37	80%	5%	15%	c70	35%	40%	25%
c5	80%	20%	0%	c38	75%	10%	15%	c71	30%	45%	25%
c6	75%	25%	0%	c39	70%	15%	15%	c72	25%	50%	25%
c7	70%	30%	0%	c40	65%	20%	15%	c73	20%	55%	25%
c8	65%	35%	0%	c41	60%	25%	15%	c74	15%	60%	25%
c9	60%	40%	0%	c42	55%	30%	15%	c75	70%	0%	30%
c10	55%	45%	0%	c43	50%	35%	15%	c76	65%	5%	30%
c11	50%	50%	0%	c44	45%	40%	15%	c77	55%	15%	30%
c12	45%	55%	0%	c45	40%	45%	15%	c78	50%	20%	30%
c13	40%	60%	0%	c46	35%	50%	15%	c79	45%	25%	30%
c14	95%	0%	5%	c47	30%	55%	15%	c80	40%	30%	30%
c15	90%	5%	5%	c48	25%	60%	15%	c81	35%	35%	30%
c16	85%	10%	5%	c49	80%	0%	20%	c82	30%	40%	30%
c17	80%	15%	5%	c50	75%	5%	20%	c83	25%	45%	30%
c18	75%	20%	5%	c51	70%	10%	20%	c84	20%	50%	30%
c19	70%	25%	5%	c52	65%	15%	20%	c85	15%	55%	30%
c20	65%	30%	5%	c53	60%	20%	20%	c86	65%	0%	35%
c21	55%	40%	5%	c54	55%	25%	20%	c87	55%	10%	35%
c22	50%	45%	5%	c55	50%	30%	20%	c88	50%	15%	35%
c23	45%	50%	5%	c56	45%	35%	20%	c89	45%	20%	35%
c24	40%	55%	5%	c57	40%	40%	20%	c90	40%	25%	35%
c25	90%	0%	10%	c58	35%	45%	20%	c91	35%	30%	35%
c26	85%	5%	10%	c59	30%	50%	20%	c92	30%	35%	35%
c27	80%	10%	10%	c60	25%	55%	20%	c93	25%	40%	35%
c28	75%	15%	10%	c61	20%	60%	20%	c94	20%	45%	35%
c29	70%	20%	10%	c62	75%	0%	25%	c95	15%	50%	35%
c30	65%	25%	10%	c63	70%	5%	25%	c96	10%	55%	35%
c31	55%	35%	10%	c64	65%	10%	25%	c97	60%	0%	40%
c32	50%	40%	10%	c65	60%	15%	25%	c98	55%	5%	40%
c33	45%	45%	10%	c66	55%	20%	25%	c99	50%	10%	40%

Paper: A review of Agent Based Modeling for agricultural policy evaluation

Published in the Agricultural Systems Volume 164, July 2018, Pages 95-106

A review of Agent Based Modeling for agricultural policy evaluation

Dimitris Kremmydas[1†], Ioannis N. Athanasiadis[2], Stelios Rozakis[3, 4]

[1] Dept. Agricultural Economics & Rural Development, Agricultural University of Athens, Greece

[2] Information Technology, Wageningen University, Netherlands

[3] School of Environmental Engineering, Technical University of Crete, Greece

[4] Dept. Bioeconomy and Systems Analysis, IUNG, Poland

[†]Corresponding author: email: kremmydas@aua.gr, phone+30 2105294783

Abstract: Farm level scale policy analysis is receiving increased attention due to a changing agricultural policy orientation. Agent based models (ABM) are farm level models that have appeared in the end of 1990's, having several differences from traditional farm level models, like the consideration of interactions between farms, the way markets are simulated, the inclusion of agents' bounded rationality, behavioral heterogeneity, etc. Considering the potential of ABMs to complement existing farm level models and that they are a relatively recent approach with a growing demand for new models and modelers, we perform a systematic literature review to (a) consolidate in a consistent and transparent way the literature status on policy evaluation ABMs; (b) examine the status of the literature regarding model transparency; the modeling of the agents' decision processes; and the creation of the initial.

Keywords: Agent Based Modeling; ABM; Agricultural policy; Systematic review

Introduction

Agricultural policies are moving away from market intervention measures toward a combination of voluntary and compulsory aids on top of basic flat rate support measures related to farm features, its environmental performance and capacity to provide ecosystem services. Consequently impacts of policy measures depend on the specific farm characteristics. So getting insights at disaggregated level and spatial scale becomes relevant for both policymakers and researchers; consequently farm scale policy analysis is receiving increased attention (Langrell et al., 2013).

Berger & Troost (2014) summarized the requirements that farm-scale models need to fulfill in order to provide useful insights within this new policy context: sufficient detail of farm management and agronomic conditions; model the heterogeneity in behavioral constraints and behaviors; include farm interactions; incorporate spatial dimension; consider farm-environment interactions and feedback; move from a comparative-static to a comparative-dynamic analysis; moderate data requirements connected to existing data sources; employ comprehensive sensitivity and uncertainty analysis. They conclude that ABMs have the potential to meet the above requirements and thus can complement existing simulation approaches.

Also, in a recent review paper, Reidsma et al. (2018) examined the development and use of farm models for policy impact assessment. Agent Based models (ABM), about 15% of all 184 papers considered, were found to have the potential to provide important additions to farm level mathematical programming models.

Agent based models in agricultural economics have appeared in the end of 1990’s. Some of the early adopters were the CORMAS group which employed a multi-agent approach to study renewable source management within an agricultural systems context (Bousquet et al., 1998). Balmann (1997) used a cellular automata approach for modeling structural change of agricultural production systems; and Berger (2001) used a spatial multi-agent programming model to assess policy options in the diffusion of innovations and resource use changes.The latter two approaches, which were policy evaluation oriented, can be considered descendants of the recursive mathematical programming (MP) approach, as the initial ABMs included a typical MP production/investment problem coupled with a land market module that was solved iteratively. The innovative elements were: the ability to include farms’ interaction and in this way to evaluate the direction of the structural change (farm growth/shrinking, farm entry/exit) and the explicit consideration of the spatial dimension.

The additions of ABMs to traditional farm level microeconomic models[11], in the conceptual level, are well summarized in Nolan et al. (2009) and are shown in Figure 6. Farm and consumer heterogeneity, spatial location and the consideration of interactions between farms and/or consumers (social networks, land markets, imitation, etc.) are presented as a distinctive feature of ABMs. Moreover in the case of traditional farm models, market outcome is the combination of the aggregate supply and demand functions while in the ABM case, market is simulated by means of individual transactions. Additionally, although traditional farm level models can potentially do so, Nolan et al. (2009) note that since ABM is most often used in cases where equilibrium conditions either cannot be identified or analytically solved, they generally relax the assumption of full rationality. This allows the assumption that economic agents facing limited information and/or information processing capacity and finite resources. Furthermore they can be endowed with adaptive mechanisms and learning capabilities.

[11] Farm type models are originally built by means of mathematical programming, econometric modeling or simulation techniques. Due to suitability to investigate novel policy instruments (advantage over econometric models) and their time and cost efficiency (comparing with simulation models) mathematical programming in various forms (LP. NLP, MILP) prevailed to the others. When we mention throughout the text the term “traditional models” for agricultural policy analysis, we refer to the above three categories, most often though in MP models. On the other hand, combined econometric-mathematical programming models as well as ABMs or ABMs combined with mathematical programming modules are novel approaches still in the making.

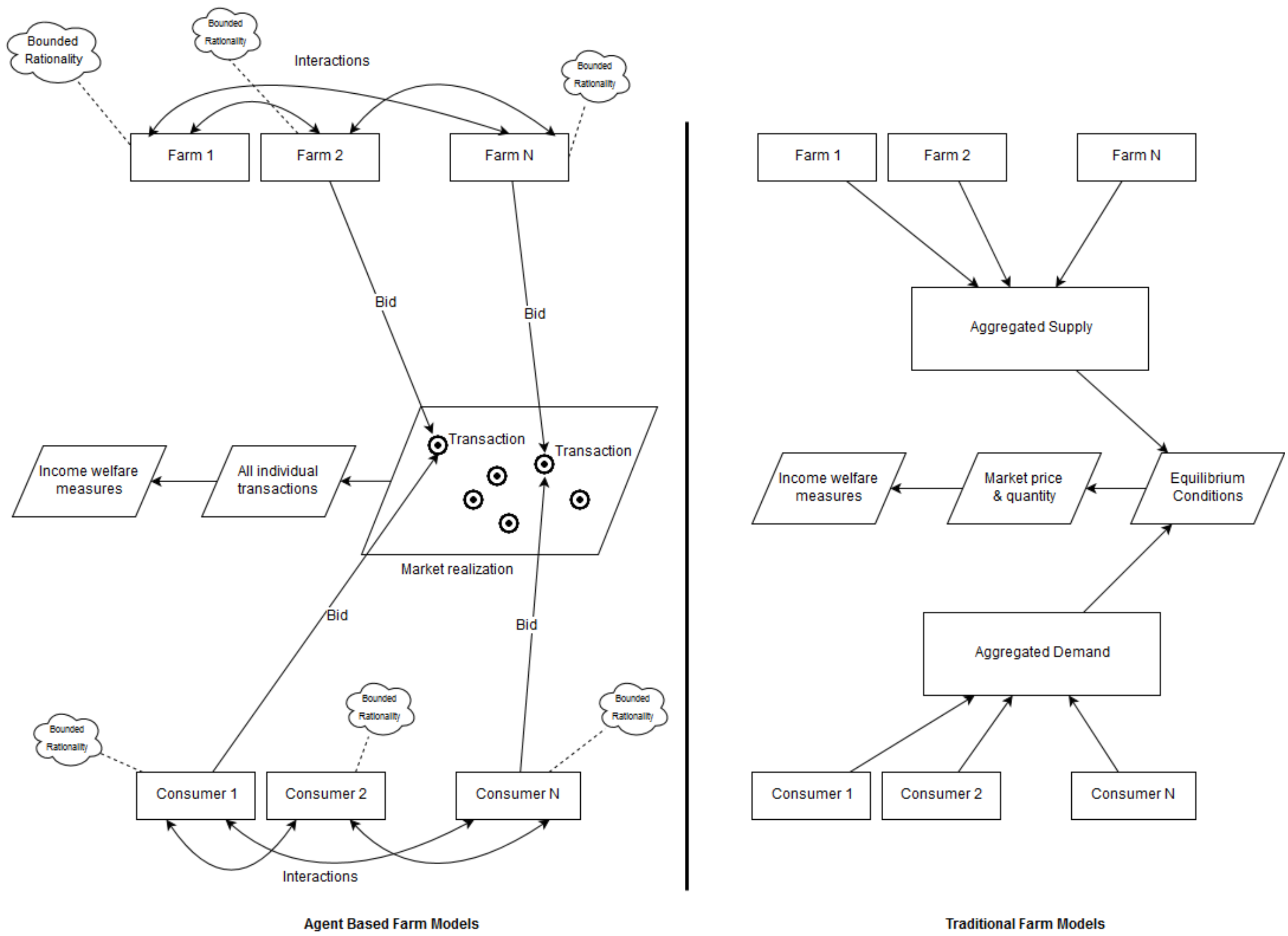


Figure 6, Conceptual difference between Agent Based Modeling approach and traditional microeconomic farm models (adapted from Nolan et al., 2009)

In a 2007 review, Matthews et al. note that "there is an increasing pressure from funding agencies to develop (Agent Based Land Use Models) tools that are of practical use by end-users and other stakeholders". Later in a methodological overview of agricultural and farm level modeling development and implementation, Langrell et al. (2013) found that although there is a substantial increase of ABMs models over time, "a large number of existing farm level models are developed for specific purposes and locations and are not easily adaptable and reusable (for policy evaluation)".

Thus, considering the potential of ABMs to complement existing farm level models and that they are a relatively recent approach with a growing demand for new models and modelers, the aims of the paper are twofold: (a) to consolidate in a consistent and transparent way the literature status on ex-ante policy evaluation ABMs; (b) to examine the critical aspects to gain more acceptance from the wider farm modeling community.

Both targets of the paper are pursued by employing a systematic literature review (SLR) approach, for related publications since 2000. The remainder of the paper is organized as follows. Section 2 describes the

SLR method used in this study. Section 3 presents the results of the SLR and the discussion of the findings; section 4 concludes the paper.

Literature Review Design

Review Protocol

The first step of the review protocol is to develop a transparent search strategy for discovering papers that are potentially related to ABM applications in the agricultural policy evaluation domain. Selection criteria are used to classify papers in groups. This addresses the first target of the paper, i.e. a consolidation of the existing ABM policy literature.

Then we clearly and explicitly specify research questions related to the second aim of the paper; an examination of the most critical aspects for further adoption of empirical ABMs from farm modelers. We use a structured process to extract all information needed to address the review questions in a meaningful way.

Search strategy and Selection criteria

Search is confined to papers written in English and published in peer-reviewed journals between 2000 and 2016 and either in title, abstract or keywords include one or more of “agent-based” ,“agent based”, “abm”, “multi-agent” or “multi agent” and any word beginning from “polic” and in title any word beginning from “farm”, “agricul”, “biodivers” or “crop”. This is equivalent to the following SCOPUS search command:

SRCTYPE (j) AND (TITLE-ABS-KEY ("agent-based" OR "agent based" OR "abm" OR "multi-agent" OR "multi agent") AND (TITLE-ABS-KEY (polic*) OR INDEXTERMS(polic*)) AND (TITLE-ABS-KEY (farm*) OR TITLE-ABS-KEY (agricul*) OR TITLE-ABS-KEY (biodivers*) OR TITLE-ABS-KEY (crop*))) AND (PUBYEAR > 1999) AND (PUBYEAR < 2017) AND LANGUAGE (english)

The search produced 176 documents that were further refined based on the criteria detailed below:

Criterion 1: the relevance to the Agent Based Modeling (criterion 1a) and Agriculture domain (criterion 1b). Based on abstract inspection and on full text inspection when necessary we removed 11 papers that were not agent based models but rather were just mentioning the term (NOT ABM). We removed 5 papers where ABM was a fraction of a larger model and thus there were not many details on the ABM implementation (PARTIALLY ABM). We removed 29 papers that were dealing with marine or coastal areas, urban areas, etc., and thus were irrelevant to agriculture (NOT AGRICULTURE).

Criterion 2: *the focus to agricultural policy evaluation subject*. We consider a paper to be relevant if the agricultural policy is a key component of the model that directly affects the model outcome and consequently the paper focuses on the relation of the policy to the model outcome. We included papers which attempted an ex-ante evaluation of a specific policy or evaluated at two or more alternative agricultural policies or different components of a single policy. Based on abstract inspection and on full text inspection when necessary, we removed 72 items and came down to 59 papers that were ABM for agricultural policy evaluation.

Criterion 3: the granularity of the agent. We identified two distinct categories, with different methodological issues. The first uses agents to represent individual farms and the second assigns them to aggregated entities, e.g. representative farms, regions, etc , or non-farm entities like landscape cells, animal or plant agents, etc. We selected to deal only with individual farm models. Based on full text inspection, we removed 8 papers.

Criterion 4: Regarding the questions that are addressed. We distinguish between data-driven models and theory-driven models, following Barlas (1996) and Polhill et al. (2013). Data-driven models focus on reproducing real world situations and thus are driven and validated by collected data and evidence. In the second category the models are based on qualitative information and second order data (stylized facts) and are used for exploring questions in principle, e.g. looking for emerging properties like resilience, etc. Ex-ante policy evaluation is pursued by means of farm models that simulate an actual farming system (Reidsma et al., 2018, Langrell et al., 2013). Due to the empirical policy orientation of the paper, we focus on data-driven ABM. We thus proceed with the data-driven (empirical) individual-farm ABM excluding 19 papers that were individual farm theory driven ABM policy evaluation papers.

An overview of the refinement process is in Figure 7 and a detailed correspondence of criteria to publications, can be found in the excel supplement.

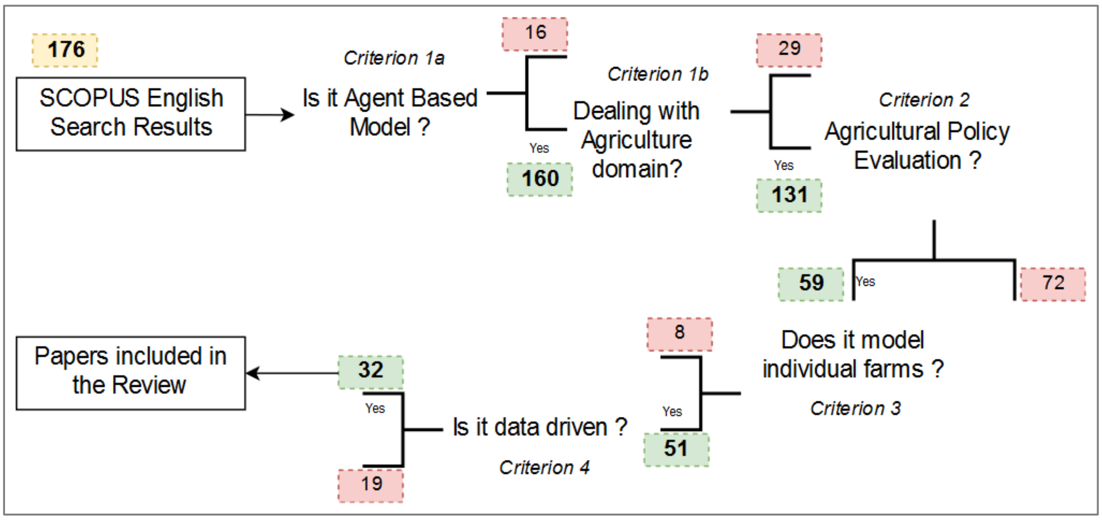


Figure 7, Overview of search results filtering process

Thus we conclude to 32 empirical-based and individual-farm relevant papers published between 2000 and 2016 as in Table 1. In Figure 8 we depict the temporal evolution of the various recognized categories. The agriculture-related ABMs (greens) are constantly increasing from 2005 and onwards and the same happens for agricultural policy evaluation ABMs (dark greens).

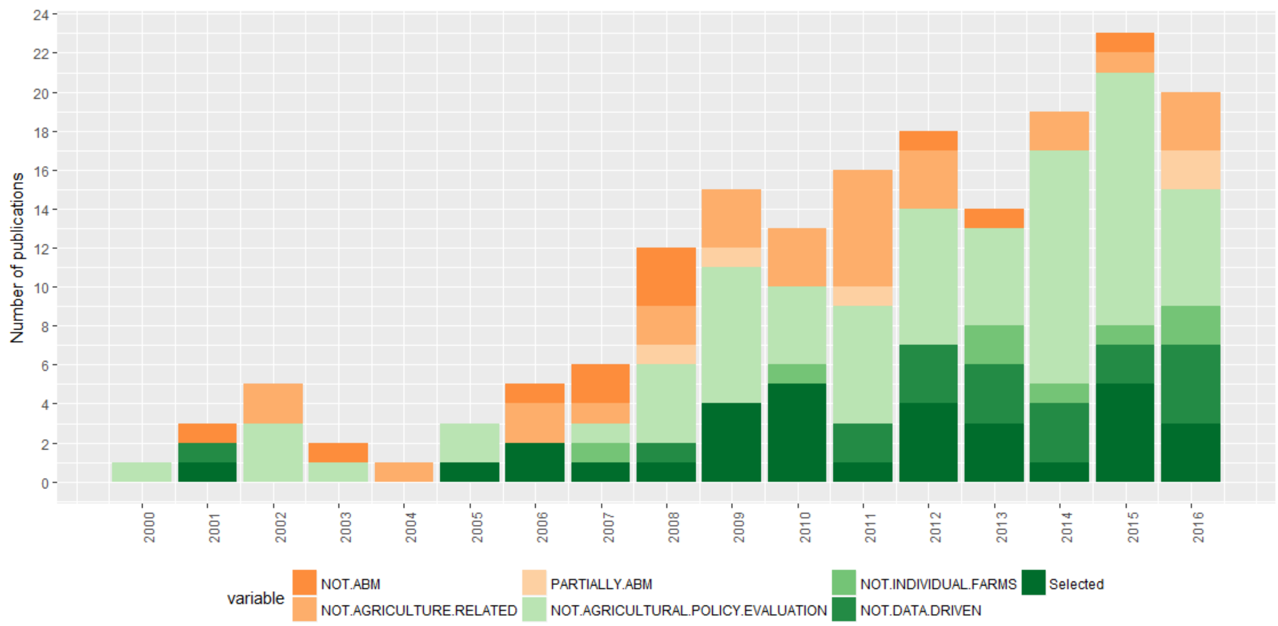


Figure 8, temporal distribution of search filtering process

Table 4, List of reviewed papers

Authors	Year	Title	Source title	Short Name
Berger T.	2001	Agent-based spatial models applied to agriculture: A simulation tool for technology diffusion, resource use changes and policy analysis	Agricultural Economics	Berger (2001)
Sengupta R., Lant C., Kraft S., Beaulieu J., Peterson W., Loftus T.	2005	Modeling enrollment in the Conservation Reserve Program by using agents within spatial decision support systems: An example from southern Illinois	Environment and Planning B: Planning and Design	Sengupta et al. (2005)
Happe K., Kellermann K., Balmann A.	2006	Agent-based analysis of agricultural policies: An illustration of the agricultural policy simulator AgriPolis, its adaptation and behavior	Ecology and Society	Happe et al.(2006)
Berger T., Schreinemachers P., Woelcke J.	2006	Multi-agent simulation for the targeting of development policies in less-favored areas	Agricultural Systems	Berger et al. (2006)
Schreinemachers P., Berger T., Aune J.B.	2007	Simulating soil fertility and poverty dynamics in Uganda: A bio-economic multi-agent systems approach	Ecological Economics	Schreinemachers et al. (2007)

Authors	Year	Title	Source title	Short Name
Happe K., Balmann A., Kellermann K., Sahrbacher C.	2008	Does structure matter? The impact of switching the agricultural policy regime on farm structures	Journal of Economic Behavior and Organization	Happe et al. (2008)
Brady M., Kellermann K., Sahrbacher C., Jelinek L.	2009	Impacts of decoupled agricultural support on farm structure, biodiversity and landscape mosaic: Some EU results	Journal of Agricultural Economics	Brady et al. (2009)
Freeman T., Nolan J., Schoney R.	2009	An agent-based simulation model of structural change in canadian prairie agriculture, 1960-2000	Canadian Journal of Agricultural Economics	Freeman et al. (2009)
Happe K., Schnicke H., Sahrbacher C., Kellermann K.	2009	Will they stay or will they go? simulating the dynamics of single-holder farms in a dualistic farm structure in Slovakia	Canadian Journal of Agricultural Economics	Happe et al. (2009)
Sahrbacher C., Jelinek L., Kellermann K., Medonos T.	2009	Past and future effects of the common agricultural policy in the Czech Republic	Post-Communist Economies	Sahrbacher et al. (2009)
Le Q.B., Park S.J., Vlek P.L.G.	2010	Land Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human-landscape system. 2. Scenario-based application for impact assessment of land-use policies	Ecological Informatics	Le et al.(2010)
Gibon A., Sheeren D., Monteil C., Ladet S., Balent G.	2010	Modelling and simulating change in reforesting mountain landscapes using a social-ecological framework	Landscape Ecology	Gibon et al. (2010)
Lobianco A., Esposti R.	2010	The Regional Multi-Agent Simulator (RegMAS): An open-source spatially explicit model to assess the impact of agricultural policies	Computers and Electronics in Agriculture	Lobianco & Esposti(2010)
van der Straeten B., Buysse J., Nolte S., Lauwers L., Claeys D., van Huylenbroeck G.	2010	A multi-agent simulation model for spatial optimisation of manure allocation	Journal of Environmental Planning and Management	van der Straeten et al. (2010)
Roeder N., Lederbogen D., Trautner J., Bergamini A., Stofer S.,	2010	The impact of changing agricultural policies on jointly used rough pastures in the Bavarian Pre-Alps: An economic and ecological scenario	Ecological Economics	Roeder et al.(2010)

Authors	Year	Title	Source title	Short Name
Scheidegger C.		approach		
Happe K., Hutchings N.J., Dalgaard T., Kellerman K.	2011	Modelling the interactions between regional farming structure, nitrogen losses and environmental regulation	Agricultural Systems	Happe et al.(2011)
Chen X., Lupi F., An L., Sheely R., Viña A., Liu J.	2012	Agent-based modeling of the effects of social norms on enrollment in payments for ecosystem services	Ecological Modelling	Chen et al. (2012)
Brady M., Sahrbacher C., Kellermann K., Happe K.	2012	An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services	Landscape Ecology	Brady et al.(2012)
Bakam I., Balana B.B., Matthews R.	2012	Cost-effectiveness analysis of policy instruments for greenhouse gas emission mitigation in the agricultural sector	Journal of Environmental Management	Bakam et al.(2012)
Nainggolan D., Termansen M., Fleskens L., Hubacek K., Reed M.S., de Vente J., Boix-Fayos C.	2012	What does the future hold for semi-arid Mediterranean agro-ecosystems? - Exploring cellular automata and agent-based trajectories of future land-use change	Applied Geography	Nainggolan et al.(2012)
Schouten M., Opdam P., Polman N., Westerhof E.	2013	Resilience-based governance in rural landscapes: Experiments with agri-environment schemes using a spatially explicit agent-based model	Land Use Policy	Schouten al.(2013)
Huber R., Briner S., Peringer A., Lauber S., Seidl R., Widmer A., Gillet F., Buttler A., Le Q.B., Hirschi C.	2013	Modeling social-ecological feedback effects in the implementation of payments for environmental services in pasture-woodlands	Ecology and Society	Huber et al.(2013)
Widener M.J., Bar-Yam Y., Gros A., Metcalf S.S., Bar-Yam Y.	2013	Modeling policy and agricultural decisions in Afghanistan	GeoJournal	Widener et al.(2013)
Daloğlu I., Nassauer J.I., Riolo R., Scavia D.	2014	An integrated social and ecological modeling framework—impacts of agricultural conservation practices on water quality	Ecology and Society	Daloğlu et al.(2014)

Authors	Year	Title	Source title	Short Name
Smajgl A., Xu J., Egan S., Yi Z.-F., Ward J., Su Y.	2015	Assessing the effectiveness of payments for ecosystem services fordiversifying rubber in Yunnan, China	Environmental Modelling and Software	Smajgl et al.(2015)
Wossen T., Berger T.	2015	Climate variability, food security and poverty: Agent-based assessment of policy options for farm households in Northern Ghana	Environmental Science and Policy	Wossen & Berger(2015)
Troost C., Walter T., Berger T.	2015	Climate, energy and environmental policies in agriculture: Simulating likely farmer responses in Southwest Germany	Land Use Policy	Troost et al.(2015)
Guillem E.E., Murray-Rust D., Robinson D.T., Barnes A., Rounsevell M.D.A.	2015	Modelling farmer decision-making to anticipate tradeoffs between provisioning ecosystem services and biodiversity	Agricultural Systems	Guillem et al.(2015)
Morgan F.J., Daigneault A.J.	2015	Estimating impacts of climate change policy on land use: An agent-based modelling approach	PLoS ONE	Morgan & Daigneault(2015)
Appel F., Ostermeyer-Wiethaup A., Balmann A.	2016	Effects of the German Renewable Energy Act on structural change in agriculture – The case of biogas	Utilities Policy	Appel et al.(2016)
Baillie S., Kaye-Blake W., Smale P., Dennis S.	2016	Simulation modelling to investigate nutrient loss mitigation practices	Agricultural Water Management	Baillie et al. (2016)
Wossen T., Berger T., Haile M.G., Troost C.	2016	Impacts of climate variability and food price volatility on household income and food security of farm households in East and West Africa	Agricultural Systems	Wossen et al. (2016)

Research Questions

To define research questions in relation to the critical aspects of wider scientific community acceptance, we consider four ABM related review papers in the agricultural related fields (land use change, socio-environmental issues, etc.) published so far:

Parker et al. (2003) reviewed multi-agent systems for the simulation of land-use change. Regarding empirical modeling they conclude than ABMs greatest advantage and at the same time shortcoming is their flexibility of specification and design that calls for focusing on verification and validation procedures. Furthermore, among others, they recognize the following challenges: the consolidation of the different individual decision making approaches and the communication of the models.

Bousquet & Le Page (2004) reviewed the development of multi-agent systems for ecosystem management. They find that the greatest advantage of ABMs is the combination of their spatial nature and the ability to represent networks. Among others, they raise the questions of whether individual decision making rules shall be based on theory or elicited from observation; and of the credibility of the model, i.e. the presentation of its structure and assumptions and their validity.

Matthews et al. (2007), list as distinct advantages of ABMs the ability to couple social and environmental models; the capacity to study the emergence of collective responses to environmental management policies; and the ability to model individual decision making entities incorporating the interactions among them. They find that the prime challenge of ABM is to show that they can provide new insights into complex natural resource systems and their management.

Kaye-Blake et al. (2009), provides a more technical overview of the various approaches of different existing models regarding the modeling of markets (land, water, labor, etc); the incorporation of risk preferences and other personality traits in the agent decision making; and the issues of information transfer and opinion transfer between agents.

Based on the advantages and challenges listed by the aforementioned review papers, and also on the requirements of farm-level models sketched by Berger & Troost (2014) mentioned already in the introduction, we shaped the following research questions (RQ):

RQ1: What is the status of the published corpus regarding model transparency? Transparency is crucial for empirical policy modeling. End users of ABM shall be able to easily identify the assumptions, relationships, and data used in a model. Since ABMs are loosely implemented in software, even when object oriented paradigm is adopted, transparency is a difficult issue to tackle with and thus we classify the reviewed papers in order to provide an overall evaluation of the transparency status. Furthermore, this is a longstanding problem that the ABM community has recognized, e.g. see the OpenABM computational model library in https://www.comses.net.

RQ2: What is the approach of the published papers regarding the modeling of agent behavior? In past review papers the ABMs flexibility to model individual behavior is considered a major advantage and in Reidsma et al. (2018) ABMs are found to be promising for modeling farmer interactions and farm structural change. However the high degree of modeling freedom results in a loose family of models very diverse between them and difficult to compare, reuse and summarize. Thus we attempt a structured classification of the various behavior modeling approaches, in order to identify potential strengths and weaknesses.

RQ3: What methods are used for initializing agent population? Agricultural policy ABM is used so as to represent an existing farming system in fine-grain detail, e.g. in plot level or/and farm population level of a certain area. However, available datasets are usually not sufficient due to aggregated or incomplete data. Consequently it is necessary to initialize/synthesize the farm population and allocate it in space. The validity of the initial virtual population has important implications for the validity of the model itself, since any significant diversion of the properties of the virtual population from the real one renders the model results disputable.

There are also other important challenges that we do not examine here, mainly because they are of a more general farm modeling interest and discussing them would require significant space and would rather distract the focus from empirical ABM. However, we understand to be important and thus provide key references that came up during the review process: *the model's process validity*, where the papers of Robinson et al. (2007) and the book edited by Smajgl & Barreteau (2014) highlight how to use empirical methods to accurately represent human behavior; *how to deal with model uncertainty*, where Troost et al. (2014) use a systematic approach based on Design of Experiments (DOE); Parry et al. (2013) uses a Bayesian sensitivity analysis approach; and Ligmann-Zielinska et al. (2014) propose a simulation framework based on quantitative uncertainty and sensitivity analyses to build parsimonious ABMs.

Data extraction and synthesis

In order to address the research questions, we read the full texts of the 32 primary studies and used a data extraction form to record our findings. The data extraction form is given in Table 5 while the extracted data can be found in detail in the excel file in the supplementary material.

Table 5, Data extraction form

	Data Extracted		Comments
RQ1 (Model transparency)	1.1	Does the paper follow the well established Overview, Design concepts, Details (ODD, Grimm et al., 2010) documentation protocol and/or its extension ODD+D (Muller et al., 2013)?	An indicator of the documentation quality
	1.2	What is the level of the results' reproducibility?	In detail: (a) Is executable or source code available? (b) Is a source dataset available?
	1.3	Does the paper explicitly report the simulation	how the modeler ascertains

		verification process?	that the model is credibly coded and run in the simulator
RQ2 (Agent behavior)	2.1 to 2.29	We adjusted the *Overview, Design concepts, Details + human Decision making* (ODD+D) of Muller et al. (2013) so as to categorize the reviewed papers on several aspects of agents' decision making. More specifically, we took the ODD+D *Design concepts* section and converted most of the guiding questions to classification questions.	The agent behavior aspects include: Individual Decision Making, Learning, Individual Sensing, Individual Prediction, Interaction, Collectives, Heterogeneity, Stochasticity and Observation. See the Appendix for a detailed description of the 29 elementary data extracted
RQ3 (Population synthesis)	3.1	What is the data source used to create the initial population?	
	3.2	What is the method to create the initial population?	
	3.3	What is the method to position agents in space?	

An important note regarding the data extraction process is that we abstain from concluding that a certain property or feature is not existent in a paper. Due to the complex model structure (for almost half of the papers we had to consider an additional source like another paper or a manual) ABMs most often have, it is possible that a feature was not stated clearly or not reported at all; thus a Type II error (false negative) is probable.

Finally we followed up with a synthesis by collating and summarizing the extracted data in a manner that is suitable to answer our research questions. We employed descriptive and qualitative analysis on our data, while statistical meta-analysis was not possible due to the Type II error and the relatively small number of observations.

Results and Discussion

Literature consolidation

More than 65% of the papers mention that they use a modeling framework[12]. Agripolis is used in eight papers while MP-MAS in six, while the rest used by one paper are Aporia, ALUAM-AB, ARLUNZ, CORMAS,

[12] "A modeling framework is a collection of building blocks (i.e., coded methods) and a generic system structure (i.e., abstract classes representing actors in the system, how they can interact and behave, as well as scheduling actions) that enable researchers to focus on conceptual representations of the study system; justification of model parameterization; and calibration rather than

LUDAS, RegMAS, RF-MAS, ERA. Regarding modeling toolkits, RepastJ or Repast Simphony is used in three papers, while Netlogo in two.

Land use change and *environmental impact assessment* is within the subject of about one third of the papers while *structural change* and *income, production or market projections* of one quarter of the papers. On average, the study area is approximately 1000 km2, including around 1600 agents with a time span of 20 years.

The journal with the most reviewed publications is *Agricultural Systems*, an indicator of the multi-disciplinary nature of the ABM approach. Also many papers are published in journals directly related to environmental management and some to journals related to geography, another indicator of the spatial nature of ABM.

In Figure 4, Agricultural economics (Ag.Econ) journals appear in deep and marine blue that is prior to 2010 and they are located mainly in the south west quartile of the map, which means that they cite similar references, in other words they drill from the same sources. Policy, systems and environmental analysis journals appear after 2010, they cite both Ag.Econ (the seminal papers) and others. A possible explanation is that first publications concern the methodology and theory so they fulfilled requirements of Ag.Econ journals whereas the latter ones focus on implementing the methodology with emphasis in the environment. Another explanation could be that after succeeding to the rigorous scrutiny of Ag.Econ journals, teams who developed such ABMs were solicited in research projects undertaken by multidisciplinary consortia. The output of these projects had a broader scope beyond disciplinary journals in agricultural economics, notwithstanding higher impact factors.

developing a model from scratch. Frameworks are significantly more refined than general ABM toolkits, as they integrate domain knowledge and preassemble building blocks that facilitate domain-specific research questions (e.g., land-use change, production decisions)". (Murray-Rust et al., 2014)

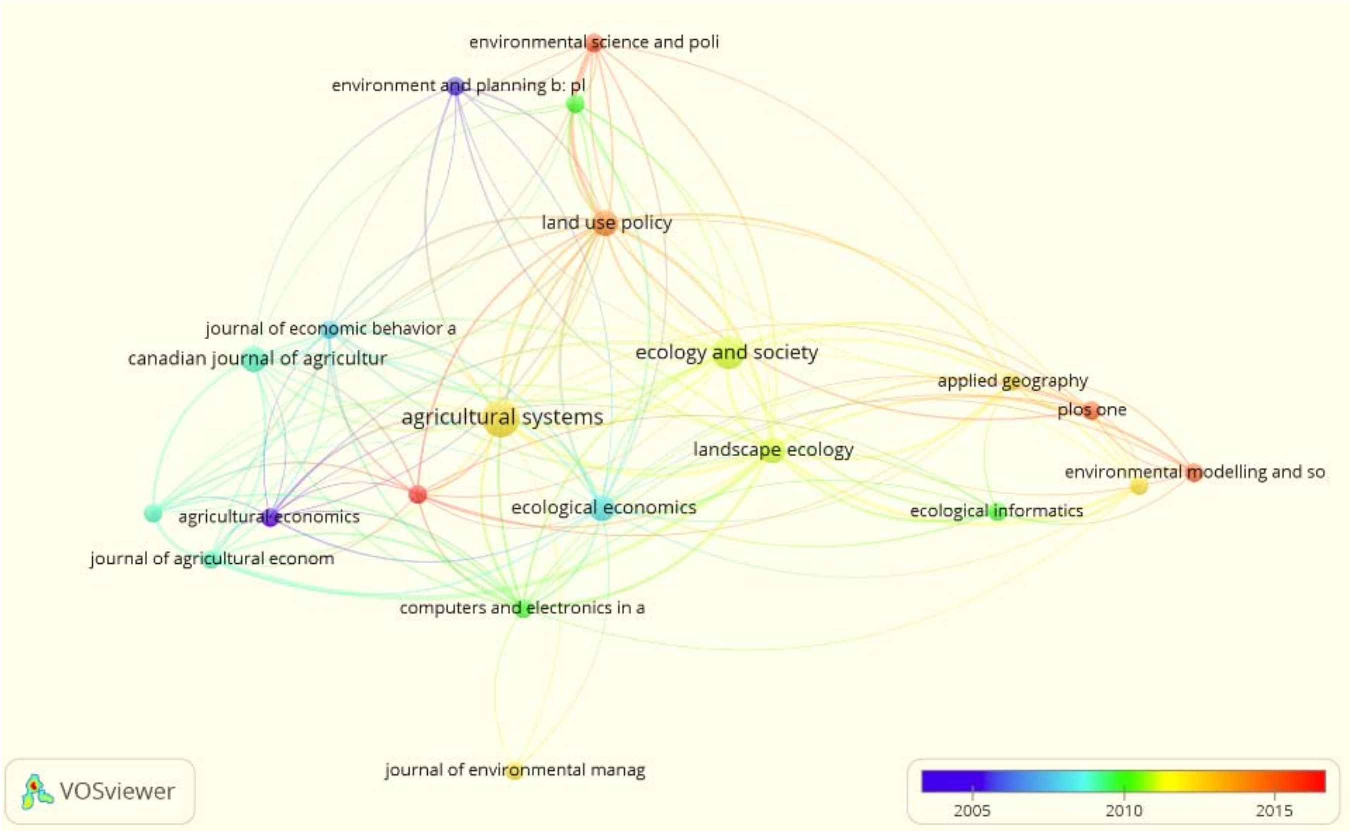


Figure 4, Network of journals using bibliographic coupling analysis in VOSviewer. The positioning of items is determined based on the number of references they share; edges between two nodes denote that there is at least one common reference between them. Two source journals are omitted from the graph as outliers to optimize visibility: Agricultural Water Management that is only linked to Land Use Policy and Geojournal not linked to any other journal

Model Transparency

Over 60% of the reviewed papers followed the ODD or the ODD+D (Muller et al., 2013) documentation protocol. This clearly enhances the readability of the models by other researchers. But still, since ODD is originally targeting ecology ABMs, the ODD+D (Muller et al., 2013) seems a promising extension that covers several human decision making aspects and it should be more widely adopted.

Another effort towards improving documentation quality is to prototype the creation process of the empirically based ABMs itself. The paper of Smajgl et al. (2011) is moving towards this direction. They propose a parameterization procedure for empirical ABMs, composed by three steps: Extracting different agent classes and corresponding behaviors; eliciting each agent class behavior parameters or rules; and assigning each individual member of the simulation population to some kind of behavior. This framework can be potentially transformed to a documentation protocol, like ODD, with relevant questions common to all empirically based ABMs that will clarify to a great extent each model's approach.

On most papers (22 out of 32) of the reviewed papers we did not recognize any possibility of *reproducing the results*. In two papers the source code was available, in another two the source and the model's dataset was provided, and in another six the executable files and data was available to reproduce the results.

Reproducibility provides credibility to empirical models and more attention shall be given by authors and by journals publishing related work. We believe that at minimum, an executable and a related dataset shall be available to model users.

Regarding *model verification*, for the models that provided source code, this is partially fulfilled since the end users can check themsevles the model verification, although practically this may not hold, e.g. the end user does not have command of the model's programming language. In any case, in two reviewed papers the verification process is explicitly stated to be performed by means of unit testing. *Unit tests* are a powerful tool for doing so: As Daloglu et al. (2014) is describing, the software development is happening in small steps, and for each step code test units are written that are fed with a predefined input followed by a comparison of the expected and observed output of the test unit. This testing process could also act as the public verification of the model when unit tests are given alongside with the executable.

Overall, we propose a four-level incremental scale to characterize the quality of model transparency: access to model documentation; following a documentation protocol; dataset and executable; dataset and source code (Table 4). In Figure 4 we give an assessment of the model transparency quality of the reviewed papers.

Table 6, Proposed verification stages

Documentation	Level 1	Explanation of the simulation model by self-means (as discussed in Muller et al., 2014)
	Level 2	Follow a broadly recognized and well structured (initial conditions, timing, interaction, unit tests, exposition of the mechanics of the simulation) documentation procedure, e.g. ODD+D
Reproducibility	Level 3	Ability to reproduce the results (take the simulation executable and run it with only one dataset and reach the same results)
	Level 4	Ability to change the assumptions of the simulation, run the model and test the sensitivity of the results (source code is provided)

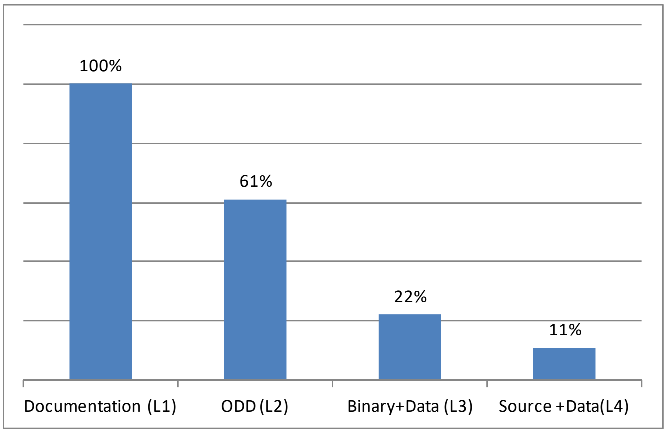


Figure 9, Quality of Model Transparency

Agent Behavior

Results

Regarding the decision making entity (the agent), almost 70% of the papers refer to a *farmer/farm* where the decision making (DM) process is revolved around production or/and investment while the rest to a *farm household* where DM also includes consumption. Other DM objects found, although less frequent, regard the land use or a conversion to a management practice.

We did not notice papers to include agents in lower or higher scales. By *agents* we mean entities that display autonomous and proactive agency in contrast to passive entities, e.g. “agents” that serve as database for other real agents to retrieve info from. This latter type of so-called agency is present in many papers, but since it is a merely technical software construct, it does not affect the dynamics of the simulation and we are not interested on reviewing and reporting on this. In the existing literature, decision making was studied only in a single scale (that of the farmers agents), and the effects of decision making at different scales are largely unexplored.

Regarding the DM algorithm, about 60% of the reviewed papers are considering rational agents using *explicit mathematical programming optimization* (MP), about 20% employ reflective agents using *simple rules* (SR), e.g. if neighbor is in state A, then do B. The rest employ some type of *behavioral heuristics* (BH), e.g. calculate the utility of the alternatives and select the maximum. If we regard papers that use the same modeling framework as a single paper (it is plausible to do so, since a modeling framework uses the same DM approach across all related publications), then MP is used 45%, SR by 30% and BH by 25% of items.

In almost 20% of the papers the agent DM process is itself a stochastic process, e.g. the agent maps a probability of selection to the alternatives and the simulator select randomly using those probabilities. Also we did not notice any paper to explicitly consider the variability of any parameter of the DM algorithm, e.g. the variance of price is a parameter of the agent’s decision model.

In over 85% of the papers the agents were adaptive. We considered an agent to be adaptive if he is capable of responding to other agents and/or its environment change of state; this is a very broad definition of adaptiveness where even simple reactiveness is included. On all papers that included adaptive agents, spatial aspects were incorporated in the DM (e.g. an agent is located in space and thus holds specific endowments). In almost 70% of the reviewed papers, temporal dimension was also affecting DM (e.g. data from past events or prospects of future outcomes). On the other hand we did not notice a paper that incorporated social norms or cultural values in DM.

We identified learning in two papers. By *learning* we mean the improvement of the agent's performance in the course of time by gaining more information/knowledge of the environment.

In 85% of the papers agents were sensing their environment and the nature of that sensing was rather global, e.g. all agents read a product global price; than local, e.g. read the neighbor's price (4 papers). We did not notice any paper to explicitly model the sensing process but rather information was directly provided to agents. We also did not notice any paper to model errors in sensing, e.g. stochastic sensing could serve as such, or costs for sensing.

In about half of the reviewed papers the agents make predictions, i.e. the estimation of future conditions the agents will experience, like the use of expected prices or yields; however if we group papers by modelling framework, only in one quarter of approaches agents make predictions.. In three papers the projection to the future was endogenously modeled.

As far as agents interaction is concerned, we identified it in 60% of the reviewed papers. Over 70% of those was referring to a land market and the rest to some kind of information exchange through a network. Land market was primarily implemented as non-direct type of interaction, e.g. agents were submitting bids to a database and they were globally cleared, while information exchange in most cases were modeled as a direct agent to agent interaction.

In two papers we identified collectives, i.e. emerging aggregations of agents that affect individual agents. In all reviewed papers agents were heterogeneous regarding their state variables, e.g. resource endowments, but only in five, agents were exhibiting diverse behavior, e.g. different goals and thus a diversified DM process.

Regarding simulation stochasticity, in one paper a global parameter was itself a stochastic element that was updated in each simulation turn. In less than 30% of the reviewed papers it was reported that many runs were performed to account for the randomness in simulation parameters. In some of those papers it was stated that "*multiple runs with different initial random seeds were performed*". However since for pseudo-number generators, the series of two different seeds are correlated, the correct way to perform multiple runs is to use a single seed across all runs, using the first n numbers for the first run, the second n numbers for the next run, etc.

Regarding the presentation of the results, in all papers aggregated results were shown. In 25% of the papers a distribution of an observation variable was also given and in almost 30% a GIS map was provided.

In Figure 5, we provide a graphical overview of the above results that provide. In the horizontal axis is the specific dimension we examine (e.g. Adaptive agents in model?, DM with spatial aspects in model?, etc.) and in the bar we show the percentage of papers we have positively recognized to do so (e.g. in ~80% of papers we recognized to contain adaptive agents).

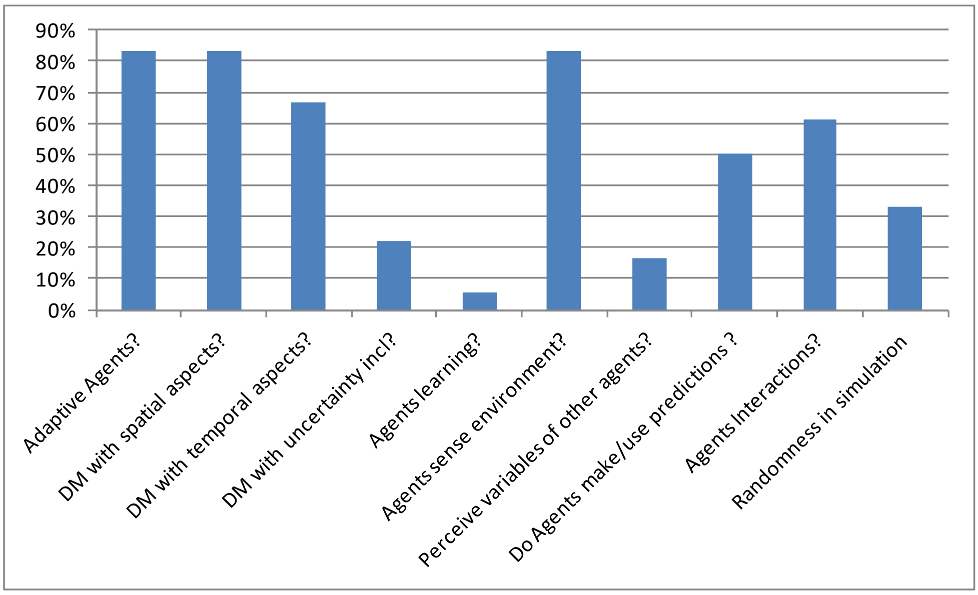


Figure 10.Various aspects of Agent Decision Making

Discussion

Agents' reactiveness can be considered to be the minimum requirement for a model to be classified to the ABM discipline. From a modeling perspective, the agent's decision making algorithm shall contain a parameter, representing another agent's or an environmental stimulus, which potentially varies during the simulation. For more complex ABM settings, one or more of agents' sensing, interaction, prediction, learning and collectives shall be explicitly modeled. For the vast majority of the reviewed papers agents' reactiveness was easily identifiable, but the rest ABM elements, with the exception of interaction, do not seem to be frequently modeled.

We also find that emergent phenomena are not highlighted in the majority of the reviewed papers. By *emergent phenomena*, as Grimm & Railsback (2005) note, we consider output properties that are not simply the sum of the properties of the individuals and cannot easily be predicted by a priori consideration of the individual agents. For example the existence of path dependence on the distribution of farm sizes or a skewed distribution of the land uses can be considered emergent properties; they are properties of the system and not of the individual agents and cannot be derived by examining agents in isolation. ABMs are very suitable for highlighting emergent properties. The fact that most reviewed paper are not focusing on those properties can be attributed to their empirical orientation and that highlighting emergent phenomena might distract their scope. An exception is the paper of Happe et al. (2008) that examines the evolution of structural change in relation to different policy regimes. In any case, in the majority of the reviewed papers, we recognize spatial explicit models with heterogeneous agents' that nevertheless is a good argument to use the ABM approach, but we do not see the modeling of *complex adaptive systems* as discussed in Xepapadeas (2010)[13].

Thus a future research direction is to answer whether it is feasible that complex adaptive systems modeling, using ABMs, can provide useful insights for empirically based questions and another is how to do this credibly without increasing the model uncertainty and loosing focus from the policy question. One possible direction is to include more frequently currently overlooked elements (agents' sensing, prediction, learning and collective); another path may be the incorporation of heterodox theories on economic agent decision making, containing components on human bounded rationality, evolutionary decision making and interaction. For instance as discussed in Foley (1994), Day (2008) and Elsner (2012), could serve toward this end.

Regarding agent interaction, it is included in most of the reviewed papers and is modeled mostly in an indirect way (e.g. a third party clears the collected bids of all agents). We believe that more empirical research should be conducted for modeling explicitly the mechanisms and the parameters of the agent interactions. Good examples of empirical investigations about farmers' interaction are found in Mertens et al. (2016) and in Manson et al. (2016).

In a few papers, agents are interviewed about their reactions to various scenarios and then those are inserted in the ABM model. A promising extension of this approach is that of Delmotte et al. (2016). The farmers are iteratively providing decision choices through software that then feed the ABM model. A remote (e.g. web based) gamification framework, where farmers will participate in a *business game*, providing their decisions online, can potentially replace the one-shot interview that elicits agents attitudes.

Regarding the stochasticity of the models, most papers do not report how they deal with the randomness in the simulation. It is not mentioned explicitly that multiple runs were performed and furthermore result-variables are reported without statistical measures (mean, standard deviation, etc.). ABM may be considered as stochastic computer experiments, since agents' properties are usually random distributions (e.g. positioning of agents, multivariate distributions of agent properties, etc.); and also agents' interactions can be modeled only as stochastic processes, e.g. agents are randomly selecting another agent from a set of neighboring-agents to commit a transaction. Thus the stochastic nature of ABMs dictates that the results should be given in the form of appropriate statistical distribution parameters, something that is not common among reviewed papers. Furthermore advanced data analysis techniques, like time series analysis, spatiotemporal methods and data mining algorithms could be incorporated in ABM software packages as discussed and exhibited in Lee et al. (2015). Also we find very interesting, although not popular

[13] "Economic, social, and ecological systems are examples of Complex Adaptive Systems. Economic systems are comprised of individual agents that pursue their own objectives and interact among themselves. These interactions lead to the emergence of macro behaviors that ultimately may feed back to influence the actions of individual agents, but typically on different time and spatial scales. The actions of individual agents and the emerging macroscopic outcomes may also be influenced by actions taken by regulatory institutions in their attempt to mitigate externalities associated with individual actions."

among the reviewed papers, the insertion of a stochastic element in the DM process. For empirical models, that could compensate for the lack of exact knowledge on agent behavior.

Overall, since there is homogeneity of the subject and the object of decision making, we believe that it is feasible to develop a unifying decision making and interaction framework. A common object oriented programming framework would help towards this direction as Bell et al. (2015) propose. In fact *Aporia framework* (Murray et al., 2014) uses such an object-oriented approach for modeling the agent decision making for agricultural land use and may be a good step toward this end.

Population synthesis

Regarding the *data source used to create the initial population*, in 18 papers a microeconomic database was used, in 9 interviews with all or a sample of the agent population and in two papers GIS data was used. For 3 papers we could not identify the data source. We find that data scarcity is not a major barrier since detailed geographical data (e.g. cadastral maps, land use maps, etc.) and disaggregated data of farm surveys, are often used by the reviewed papers. Interviews may also prove cost effective when models deal with relatively small areas with a few agents.

Regarding *the method used to create the population*, in 8 papers *cloning* was used. By *cloning* we mean that a limited number of agents, less than the number of the simulation agents, were replicated in order to reach the final agent population. In eight papers a *monte carlo* method was used, where the agents' population is randomly drawn from an empirical joint distribution of the farm properties; the latter is created from the available data for a limited number of agents. Finally for three papers the agent population was a one-to-one correspondence of real data and for the rest 13 we could not identify how the initial population was created.

The problem with the cloning approach is that, it reduces the variability of the model data compared to real population, multiplying the sampling error and possibly affecting the validity of the model dynamics. Furthermore, no sensitivity analysis regarding the random effect of the population generation process can be conducted, since only one population can be generated, i.e. the clones of the sample farms. Monte Carlo methods, as discussed in Berger & Schreinemachers (2006) hold better statistical properties.

Regarding *the method used to position agents in space* in 12 papers we could not identify this method, in 18 papers it was randomly positioned and in two the plots of the farmers were corresponding to real data. . Random positioning ignores the likely spatial autocorrelation of their properties but can be overlooked if the simulation is dealing with a spatially homogeneous farming system. Otherwise, provisions should be made to spatially allocate the agents based on at least some plausible evidence. In any case, spatial location can potentially be included in the population synthesis process; spatial location being a farm property. Mack et al. (2013) is closer to this approach.

From a software engineering point of view, incorporating population synthesis as a distinct module with a special user interface may provide to end users the ability to experiment on the impact of data downscaling assumptions to the model output.

Overall, regarding population synthesis and spatial allocation, there seems to be a rigorous research interest, not directly related to agricultural policy ABM, but with potentially applicable results to empirical models for agricultural policy evaluation. For instance the paper of Harland et al. (2012) reviews and compares three state-of-the-art spatial population generation techniques (deterministic reweighting, conditional probabilities and simulated annealing) and Hamada et al. (2015) present a novel kernel estimator for reconstructing an entire population from a small sample survey.

Conclusions

ABMs can complement conventional farm models for policy analysis, as pointed by Berger (2001): heterogeneity of behavior can easily be modeled; a wide range of farm to farm interaction can be included like information exchange, markets of locally available resources with endogenous price formation, etc.; dynamic comparative analysis can be undertaken as opposed to the comparative static approach of equilibrium based farm models; spatial element is inherently included and that allows to investigate the spatial dynamics of various properties, e.g. the land rents. Another key strength is the ability to link human and environmental elements using space as the common element, a very important feature considering the pro-environmental orientation of contemporary agricultural policy.

In this review we examined the ABM literature on policy evaluation from 2000 to 2016 in order to (a) consolidate it in a consistent and transparent way; (b) to examine the critical aspects of empirical based individual farm policy evaluation ABMs that will expand their use.

Regarding the literature status on policy evaluation ABMs, there is a significant increase in the number of publications after 2008 at a large extent due to the potential of early seminal papers published in the previous period. We distinguished between individual-farm ABMs and not-individual or non-farm ABMs, and between data-driven and theory-driven approaches. Figure 8 provides an illustrative summary of their evolution. In this respect, researchers can carry over from the detailed literature classification, either for examining the groups of papers that we are not focusing into, or for a future review on the same subject.

We examined several critical aspects of empirical-based farm ABMs in relation to wider adaptation for policy analysis. Those aspects are based on past reviews and on generic farm model requirements sketched by Berger & Troost (2014). A summary of our findings is given below:

- *Modeling transparency*: We find that the majority of the papers follow the ODD protocol (Grimm et al., 2010), however the overall level of modeling transparency has potential to be further improved. At a minimum an executable and related data shall be available to end users. When for privacy or copyright reasons data cannot be shared it is advised to make available synthetic sample data together with the model. Last but not least, unit testing is a good practice to be employed for public model verification.

- *The sufficient detail of farm management and agronomic conditions and the heterogeneity in behavioral constraints and behaviors*: ABMs can be as analytic as the traditional microeconomic models regarding the details on those aspects. Moreover, they can incorporate behavioral parameters that other type of models cannot; Learning, collective structures, modeling complex adaptive systems. We propose that more rigorous research is needed primarily on whether incorporating those can provide useful insights for empirically based questions and how to do this without increasing the model uncertainty and loosing focus from the policy question.
- *Farm interaction and incorporation of spatial dimension*: ABMs exhibit those two features to a satisfactory degree. However, more work shall be done so that interactions are modeled in a direct way and established on empirical data. More information shall be provided on the population initialization that includes positioning in space and statistically sound methods shall be established for doing so; the above two additions will improve the spatial dimension.

Overall, although ABMs clearly outperform mainstream modeling approaches in certain aspects, they face difficulties to be widely adopted by modelers and applicable for large scale assessment. By means of literature review, the present work attempted to identify some of them and provide insights for enhancement which along with advances in computing and standardization of parameterization and calibration processes can spread their use by policy analysts and decision makers.

References

Appel, F., Ostermeyer-Wiethaup, A., & Balmann, A. (2016). Effects of the German Renewable Energy Act on structural change in Agriculture – The case of biogas. Utilities Policy, 1–11. http://doi.org/10.1016/j.jup.2016.02.013

Baillie, S., Kaye-Blake, W., Smale, P., & Dennis, S. (2016). Simulation modelling to investigate nutrient loss mitigation practices. Agricultural Water Management, 177, 221–228. http://doi.org/10.1016/j.agwat.2016.07.018

Balmann, A. (1997). "Farm-Based Modelling of Regional Structural Change: A Cellular Automata Approach." European Review of Agricultural Economics 24(1): 85-108.

Barlas, Y. (1996). Formal aspects of model validity and validation in system dynamics. System Dynamics Review, 12(3), 183–210. http://doi.org/10.1002/(SICI)1099-1727(199623)12:3<183::AID-SDR103>3.0.CO;2-4

Bell, A. R., Robinson, D. T., Malik, A., & Dewal, S. (2015). Modular ABM development for improved dissemination and training. Environmental Modelling & Software, 73, 189–200. https://doi.org/10.1016/j.envsoft.2015.07.016

Berger, T., 2001. Agent-based models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics 25(2/3), 245-260.

Berger, T., & Schreinemachers, P. (2006). Creating agents and landscapes for multiagent systems from random samples. *Ecology and Society*, *11*(2), 19.

Berger, T., Schreinemachers, P., & Woelcke, J. (2006). Multi-agent simulation for the targeting of development policies in less-favored areas. Agricultural Systems, 88(1), 28–43. http://doi.org/10.1016/j.agsy.2005.06.002

Berger, T., & Troost, C. (2014). Agent-based Modelling of Climate Adaptation and Mitigation Options in Agriculture. Journal of Agricultural Economics, 65(2). https://doi.org/10.1111/1477-9552.12045

Bousquet, F., Bakam, I., Proton, H., & C. Le Page. (1998). Cormas: common-pool resources and multi-agent systems. In Lecture Notes in Artificial Intelligence (Vol. 1416). https://doi.org/10.1007/3-540-64574-8

Bousquet, F., & Le Page, C. (2004). Multi-agent simulations and ecosystem management: A review. Ecological Modelling, 176(3-4), 313–332. http://doi.org/10.1016/j.ecolmodel.2004.01.011

Brady, M., Kellermann, K., Sahrbacher, C., & Jelinek, L. (2009). Impacts of decoupled agricultural support on farm structure, biodiversity and landscape mosaic: Some EU results. Journal of Agricultural Economics, 60(3), 563–585. http://doi.org/10.1111/j.1477-9552.2009.00216.x

Brady, M., Sahrbacher, C., Kellermann, K., & Happe, K. (2012). An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services. Landscape Ecology, 27(9), 1363–1381. http://doi.org/10.1007/s10980-012-9787-3

Chen, X., Lupi, F., An, L., Sheely, R., Viña, A., & Liu, J. (2012). Agent-based modeling of the effects of social norms on enrollment in payments for ecosystem services. Ecological Modelling, 229, 16–24. http://doi.org/10.1016/j.ecolmodel.2011.06.007

Daloğlu, I., Nassauer, J. I., Riolo, R., & Scavia, D. (2014). An integrated social and ecological modeling framework - Impacts of agricultural conservation practices on water quality. Ecology and Society, 19(3), 12. http://doi.org/10.5751/ES-06597-190312

Day, R. H. (2008). Micro foundations for meso and macro economic theory. Journal of Evolutionary Economics, 18(2), 261–273. https://doi.org/10.1007/s00191-007-0084-2

Delmotte, S., Barbier, J., Mouret, J., Le, C., Wery, J., Chauvelon, P., … Lopez, S. (2016). Participatory integrated assessment of scenarios for organic farming at different scales in Camargue , France. Agsy, 143, 147–158. http://doi.org/10.1016/j.agsy.2015.12.009

Elsner, Wolfram. Microeconomics of Interactive Economies: Evolutionary, Institutional, and Complexity Perspectives. a'non-Toxic'Intermediate Textbook. Edward Elgar Publishing, 2012.

Freeman, T., Nolan, J., & Schoney, R. (2009). An Agent-Based Simulation Model of Structural Change in Canadian Prairie Agriculture, 1960-2000. Canadian Journal of Agricultural Economics/Revue Canadienne D'agroeconomie, 57(4), 537–554. http://doi.org/10.1111/j.1744-7976.2009.01169.x

Foley, D. K. (1994). A Statistical Equilibrium Theory of Markets. *Journal of Economic Theory*, (62), 321–345.

Gibon, A., Sheeren, D., Monteil, C., Ladet, S., & Balent, G. (2010). Modelling and simulating change in reforesting mountain landscapes using a social-ecological framework. Landscape Ecology, 25(2), 267–285. http://doi.org/10.1007/s10980-009-9438-5

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: a review and first update. *Ecological modelling*, *221*(23), 2760-2768.

Grimm, V., & Railsback, S. F. (2005). Individual-based modeling and ecology (Vol. 2005). Princeton: Princeton university press.

Guillem, E. E., Murray-Rust, D., Robinson, D. T., Barnes, A., & Rounsevell, M. D. A. (2015). Modelling farmer decision-making to anticipate tradeoffs between provisioning ecosystem services and biodiversity. Agricultural Systems, 137, 12–23. http://doi.org/10.1016/j.agsy.2015.03.006

Murray-Rust, D., Robinson, D. T., Guillem, E., Karali, E., & Rounsevell, M. (2014). An open framework for agent based modelling of agricultural land use change. Environmental Modelling and Software, 61, 19–38. https://doi.org/10.1016/j.envsoft.2014.06.027

Happe, K., Kellermann, K., & Balmann, A. (2006). Agent-based Analysis of Agricultural Policies: an Illustration of the Agricultural Policy Simulator AgriPoliS, its Adaptation and Behavior. Ecology and Society, 11(1). http://doi.org/49

Happe, K., Balmann, A., Kellermann, K., & Sahrbacher, C. (2008). Does structure matter? The impact of switching the agricultural policy regime on farm structures. Journal of Economic Behavior and Organization, 67(2), 431–444. http://doi.org/10.1016/j.jebo.2006.10.009

Happe, K., Schnicke, H., Sahrbacher, C., & Kellermann, K. (2009). Will they stay or will they go? simulating the dynamics of single-holder farms in a dualistic farm structure in Slovakia. Canadian Journal of Agricultural Economics, 57(4), 497–511. http://doi.org/10.1111/j.1744-7976.2009.01167.x

Happe, K., Hutchings, N. J., Dalgaard, T., & Kellerman, K. (2011). Modelling the interactions between regional farming structure, nitrogen losses and environmental regulation. Agricultural Systems, 104(3), 281–291. http://doi.org/10.1016/j.agsy.2010.09.008

Harland, K., Heppenstall, A., Smith, D., & Birkin, M. (2012). Creating realistic synthetic populations at varying spatial scales: A comparative critique of population synthesis techniques. JASSS, 15(1). http://doi.org/10.18564/jasss.1909

Kaye-Blake, W., Li, F. Y., Martin, A. M., McDermott, A., Neil, H., & Rains, S. (2009). A review of multi-agent simulation models in agriculture.

Kellermann, K., Happe, K., Sahrbacher, C., Balmann, A., Brady, M., Schnicke, H., & Osuch, A. (2008). AgriPoliS 2.1 – Model documentation.

Langrell, S., Ciaian, P., Espinosa, M., Paloma, S. G., Heckelei, T., Louhichi, K., ... & Vard, T. (2013). Farm level modelling of CAP: a methodological overview. *JRC Scientific and Policy Report, EUR*, *25873*.. Retrieved from http://publications.jrc.ec.europa.eu/repository/handle/JRC79969

Le, Q. B., Park, S. J., Vlek, P. L., & Cremers, A. B. (2008). Land-Use Dynamic Simulator (LUDAS): A multi-agent system model for simulating spatio-temporal dynamics of coupled human–landscape system. I. Structure and theoretical specification. Ecological Informatics, 3(2), 135-153.

Lee, J. S., Filatova, T., Ligmann-Zielinska, A., Hassani-Mahmooei, B., Stonedahl, F., Lorscheid, I., … Parker, D. C. (2015). The complexities of agent-based modeling output analysis. *JASSS*, 18(4). http://doi.org/10.18564/jasss.2897

Ligmann-Zielinska, A., Kramer, D. B., Cheruvelil, K. S., & Soranno, P. A. (2014). Using uncertainty and sensitivity analyses in socioecological agent-based models to improve their analytical performance and policy relevance. *PloS one*, *9*(10), e109779.

Lobianco, A., & Esposti, R. (2010). The Regional Multi-Agent Simulator (RegMAS): An open-source spatially explicit model to assess the impact of agricultural policies. Computers and Electronics in Agriculture, 72(1), 14–26. http://doi.org/10.1016/j.compag.2010.02.006

Louhichi, K., Ciaian, P., & Gomez y Paloma, S. (2013.). Farm-level models for EU policy analysis: Review of recent literature and comparison of most relevant models. In S. Langrell (Ed.), Farm level modelling of CAP: A methodological overview,. EC-JRC-IPTS.

Mack, G., Möhring, A., Ferjani, A., Zimmermann, A., & Mann, S. (2013). Transfer of single farm payment entitlements to farm successors : impact on structural change and rental prices in Switzerland. Bio-Based and Applied Economics, 2(2), 113–130.

Manson, S. M., Jordan, N. R., Nelson, K. C., & Brummel, R. F. (2016). Modeling the effect of social networks on adoption of multifunctional agriculture. *Environmental Modelling and Software*, 75, 388–401. http://doi.org/10.1016/j.envsoft.2014.09.015

Matthews, R. B., Gilbert, N. G., Roach, A., Polhill, J. G., & Gotts, N. M. (2007). Agent-based land-use models: A review of applications. Landscape Ecology, 22(10), 1447–1459. http://doi.org/10.1007/s10980-007-9135-1

Mertens, A., Van Meensel, J., Mondelaers, K., Lauwers, L., & Buysse, J. (2016). Context Matters—Using an Agent-Based Model to Investigate the Influence of Market Context on the Supply of Local Biomass for Anaerobic Digestion. *BioEnergy Research*, 9(1), 132–145. http://doi.org/10.1007/s12155-015-9668-0

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., ... & Schwarz, N. (2013). Describing human decisions in agent-based models–ODD+ D, an extension of the ODD protocol. Environmental Modelling & Software, 48, 37-48.

Müller, B., Balbi, S., Buchmann, C. M., de Sousa, L., Dressler, G., Groeneveld, J., Weise, H. (2014). Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling and Software*, 55, 156–163. http://doi.org/10.1016/j.envsoft.2014.01.029

Nainggolan, D., Termansen, M., Fleskens, L., Hubacek, K., Reed, M. S., de Vente, J., & Boix-Fayos, C. (2012). What does the future hold for semi-arid Mediterranean agro-ecosystems? - Exploring cellular automata and agent-based trajectories of future land-use change. Applied Geography, 35(1-2), 474–490. http://doi.org/10.1016/j.apgeog.2012.09.007

Nolan, J., Parker, D., Van Kooten, G. C., & Berger, T. (2009). An overview of computational modeling in agricultural and resource economics. Canadian Journal of Agricultural Economics, 57(4), 417–429. http://doi.org/10.1111/j.1744-7976.2009.01163.x

Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: A review. Annals of the Association of American Geographers, 93(2), 314–337. http://doi.org/10.1111/1467-8306.9302004

Parry, H. R., Topping, C. J., Kennedy, M. C., Boatman, N. D., & Murray, A. W. A. (2013). A Bayesian sensitivity analysis applied to an Agent-based model of bird population response to landscape change. Environmental Modelling and Software, 45, 104–115. https://doi.org/10.1016/j.envsoft.2012.08.006

Polhill, J. G., Gimona, A., & Gotts, N. M. (2013). Nonlinearities in biodiversity incentive schemes: A study using an integrated agent-based and metacommunity model. Environmental Modelling and Software, 45, 74–91. http://doi.org/10.1016/j.envsoft.2012.11.011

Reidsma, P., Janssen, S., Jansen, J., & van Ittersum, M. K. (2018). On the development and use of farm models for policy impact assessment in the European Union – A review. Agricultural Systems, 159 (September 2017), 111–125. http://doi.org/10.1016/j.agsy.2017.10.012

Robinson, D. T., Brown, D. G., Parker, D. C., & others. (2007). Comparison of empirical methods for building agent-based models in land use science. Journal of Land Use Science, 2(1), 31–55. https://doi.org/10.1080/17474230701201349

Sahrbacher, C., Jelinek, L., Kellermann, K., & Medonos, T. (2009). Past and future effects of the Common Agricultural Policy in the Czech Republic. Post-Communist Economies, 21(4), 495–511. http://doi.org/10.1080/14631370903339872

Sengupta, R., Lant, C., Kraft, S., Beaulieu, J., Peterson, W., & Loftus, T. (2005). Modeling enrollment in the Conservation Reserve Program by using agents within spatial decision support systems: An example from southern Illinois. Environment and Planning B: Planning and Design, 32(6), 821–834. http://doi.org/10.1068/b31193

Schouten, M., Opdam, P., Polman, N., & Westerhof, E. (2013). Resilience-based governance in rural landscapes: Experiments with agri-environment schemes using a spatially explicit agent-based model. *Land Use Policy*, 30(1), 934–943. http://doi.org/10.1016/j.landusepol.2012.06.008

Schouten, M., Polman, N., Westerhof, E., & Kuhlman, T. (2012). Rural landscapes in turbulent times: a spatially explicit agent-based model for assessing the impact of agricultural policies. In *Managing Market Complexity: The Approach of Artificial Economics* (No. 662, pp. 195-208). Springer.

Schreinemachers, P., & Berger, T. (2011). An agent-based simulation model of human–environment interactions in agricultural systems. *Environmental Modelling & Software*, *26*(7), 845-859.

Schreinemachers, P., Berger, T., & Aune, J. B. (2007). Simulating soil fertility and poverty dynamics in Uganda: A bio-economic multi-agent systems approach. Ecological Economics, 64(2), 387–401. http://doi.org/10.1016/j.ecolecon.2007.07.018

Smajgl, A., Brown, D. G., Valbuena, D., & Huigen, M. G. A. (2011). Empirical characterisation of agent behaviours in socio-ecological systems. Environmental Modelling and Software, 26(7), 837–844. https://doi.org/10.1016/j.envsoft.2011.02.011

Smajgl, A., Egan, S., Kirby, M., Mainuddin, M., Ward, J., Kroon, F., 2013. The Mekong Region Simulation (Mersim) Model - Design Document. CSIRO Climate Adaptation Flagship (Townsville).

Smajgl, A., Barreteau, O. (eds.) (2014). Empirical Agent-Based Modelling - Challenges and Solutions, Vol. 1, The Characterisation and Parameterisation of Empirical Agent-Based Models, Springer

Troost, C., & Berger, T. (2014a). Dealing with uncertainty in agent-based simulation: Farm-level modeling of adaptation to climate change in southwest Germany. *American Journal of Agricultural Economics*, 97(3), 833–854. http://doi.org/10.1093/ajae/aau076

Troost, C., Walter, T., & Berger, T. (2015). Climate, energy and environmental policies in agriculture: Simulating likely farmer responses in Southwest Germany. *Land Use Policy*, 46, 50–64. http://doi.org/10.1016/j.landusepol.2015.01.028

Troost, C., 2014. MPMAS Central Swabian Jura (Version 3.1) – Model Documentation. University of Hohenheim, Stuttgart https://www.uni-hohenheim.de/fileadmin/einrichtungen/mp-mas/msa_model_v3_1_documentation.pdf

van der Straeten, B., Buysse, J., Nolte, S., Lauwers, L., Claeys, D., & van Huylenbroeck, G. (2010). A multi-agent simulation model for spatial optimisation of manure allocation. Journal of Environmental Planning and Management, 53(8), 1011–1030. http://doi.org/10.1080/09640568.2010.495546

Wossen, T., Berger, T., Haile, M. G., & Troost, C. (2016). Impacts of climate variability and food price volatility on household income and food security of farm households in East and West Africa. Agricultural Systems. http://doi.org/10.1016/j.agsy.2017.02.006

Xepapadeas, A. (2010). Modeling complex systems. Agricultural Economics, 41, 181–191. https://doi.org/10.1111/j.1574-0862.2010.00499.x

Appendix

ODD+D (Muller et al., 2013)			Review	
Design Category	Examples	Guiding questions	Review Questions	Comments
Individual Decision Making	Name subjects (individuals agents / households, on communal level, top down decision maker) and objects of decisions, e.g.: Form of land use, distribution of labor, choices of buying and selling	II.ii.a What are the subjects and objects of decision-making? On which level of aggregation is decision-making modeled? Are multiple levels of decision making included?	1. Subject of decision making 2. Objects of decision making 3. Are there any agents in a higher/lower scale other than farm?	"Agents in lower/higher scale": We consider only entities that display autonomous agency. For instance we do not consider an "agent" that is used merely as a information database for other simulation agent to retrieve info from (a "ministry agent" that responds to queries on subsidy prices, is not consider an actual agent and thus it is not "an agent in higher scale than farm").
	Rational choice (classical optimization approach, utility maximization), bounded rationality (satisficing approach), no objectives (routine based, trial and error)	II.ii.b What is the basic rationality behind agents' decision-making in the model? Do agents pursue an explicit objective or have other success criteria?		
	Decision tree, utility function, random choice	II.ii.c How do agents make their decisions?	4. How do agents make their decisions? [SRB, BH, MP]	SRB: Simple rule based, e.g. if neighbor is in state A, then do B BH: Behavioral Heuristics, e.g. calculate the

				utility of the alternatives and select the maximum MP: explicit mathematical programming optimization
	Adaption of resource extraction level in dependence of ecological state of resource	II.ii.d Do the agents adapt their behavior to changing endogenous and exogenous state variables? And if yes, how?	5. Are agents adaptive? [Yes, No]	Definition of *adaptive agent*: An agent is considered adaptive if he is capable of responding to other agents and/or its environment. We consider a very broad definition of adaptiveness, i.e. even simple reactiveness is included.
	Cultural norms, trust	II.ii.e Do social norms or cultural values play a role in the decision-making process?	6. Do social norms or cultural values play a role in the decision-making process? [Yes, No]	
	Space-theory based models	II.ii.f Do spatial aspects play a role in the decision process?	7. Do spatial aspects play a role in the decision process? [Yes, No]	Is the Decision making shaped within a spatial dimension (e.g. an agent is located in space and thus holds specific endowments that affect the DM process as paramaters)
	Discounting, memory	II.ii.g Do temporal aspects play a role in the decision process?	8. Do temporal aspects play a role in the decision process? [Yes, No]	Does the agent decision process includes data from past events (e.g. memory) or prospects of future outcomes (e.g. expected price)
	Not at all / stochastic elements mimic uncertainties in agents' behavior / agents explicitly	II.ii.h To which extent and how is uncertainty included in the agents' decision rules?	9. Is uncertainty included in the agent's decision rules? [VAR, STOCH, No]	VAR: if agent consider the variability of estimated values that are used in the decision process (e.g. the variance of price is a parameter of the agent's decision

	consider uncertain situations or risk			model) STOCH: if the agent DM process is itself a stochastic process (e.g. the agent maps a probability of selection to the alternatives and the simulator select them randomly using those probabilities)
Learning	Change of aspiration levels depending on past experiences	II.iii.a Is individual learning included in the decision process? How do individuals change their decision rules over time as consequence of their experience?	10. Are agents learning? [Yes, No]	Definition of Learning: The agent is improving its performance in the course of time by gaining more information/knowledge of the environment.
	Evolution, genetic algorithms	II.iii.b Is collective learning implemented in the model?		
Individual Sensing		II.iv.a What endogenous and exogenous state variables are individuals assumed to sense and consider in their decisions? Is the sensing process erroneous?	11. Do agents sense their environment ? [Yes, No] 12. Does sensing includes erroneous elements? [Yes, No]	
	(Multiple) resources (including working power, monetary resources, other income resources) and behavior of other agents	II.iv.b What state variables of which other individuals can an individual perceive? Is the sensing process erroneous?	13. Do agents perceive state variables of other individuals? [Yes, No]	
	Local, network, global (whole model space)	II.iv.c What is the spatial scale of sensing?	14. What is the spatial scale of sensing? [Not applicable, Local, Global]	Local sensing: If agent can perceive data or events in his local neighborhood (e.g. yields

				of neighborhood) Global sensing: If the agent can perceive data or events in any part of the environment (e.g. market price)
	Sensing is often assumed to be local, but can happen through networks or can even be assumed to be global.	II.iv.d Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?	15. Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables? [Not applicable, Explicitly modeled, information is given directly]	
		II.iv.e Are costs for cognition and costs for gathering information included in the model?	16. Are costs for cognition and/or costs for gathering information included in the model? [Yes, No]	
Individual Prediction	Extrapolation from experience, from spatial observations	II.v.a Which data uses the agent to predict future conditions?	17. Do agents make/use predictions? [Yes, No]	Definition of prediction: The estimation of future conditions the agents will experience (e.g. expected prices or yields)
		II.v.b What internal models are agents assumed to use to estimate future conditions or consequences of their decisions?	18. Is the agent prediction process modeled explicitly? [Yes, No]	
	(External) uncertainty, (internal) capability of the agent	II.v.c Might agents be erroneous in the prediction process, and how is it implemented?	19. Might agents be erroneous in the prediction process? [Not applicable, Yes, No]	
Interaction	Direct interactions, indirect interactions (mediated by	II.vi.a Are interactions among agents and entities assumed as direct or indirect?	20. Are there interactions between agents? [Yes, No] 21. Are interactions among agents	Indirect interactions are those that there is a mediator between agents. For example a

	the environment / the market, auction)		and entities direct or indirect? [Not applicable, Direct, Indirect]	land market auctioneer that gathers all bids and clears the market. Instead in direct interactions, agents are “talking” each other. For example a land market auction mechanism, where two picked agents are exchanging bids that will result or not to a transaction
	Spatial distances (neighborhood), networks, type of agent	II.vi.b On what do the interactions depend?	22. What types of interactions are present ? [Not applicable, Land Market, Information networks, ….]	
	Explicit messages (Matthews et al., 2007)	II.vi.c If the interactions involve communication, how are such communications represented?		
	Centralized vs. decentralized, group based tasks	II.vi.d If a coordination network exists, how does it affect the agent modeler? Is the structure of the network imposed or emergent?		
Collectives	Social groups, human networks and organizations	II.vii.a Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Are these aggregations imposed by the 103odeler or do they emerge during the simulation?	23. Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? [Yes, No] 24. Are these aggregations imposed by the modeler or do they emerge during the simulation? [Not applicable, imposed, emerged]	
	Collective as emergent property vs. as a definition by the modeler (separate	II.vii.b How are collectives represented?		

	kind of entity with its own state variables and traits)			
Heterogeneity	Would an exchange of one agent with another at the beginning have an effect on the simulation?	II.viii.a Are the agents heterogeneous? If yes, which state variables and/or processes differ between the agents?	25. Are agents heterogeneous ? [Yes, no]	
		II.viii.b Are the agents heterogeneous in their decision-making? If yes, which decision models or decision objects differ between the agents?	26. What kind of heterogeneity exists? [Not applicable, State variables, Behavioral, Both]	
Stochasticity		II.ix.a What processes (including initialization) are modeled by assuming they are random or partly random?	27. Where is randomness involved? [no randomness, initialization, decision making, communication, multiple runs]	By *randomness* we mean the use of a probabilistic element DM: Stochastic element in agent DM algorithm Runs: Many runs in order to measure the randomness effect Within-runs: A global parameter is a stochastic element updated each run
Observation		II.x.a What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected?	28. In what relative-to-farm-scale are results presented? [farm level, aggregated level]	
		II.x.b What key results, outputs or characteristics of the model are emerging	29. Are there any emergent phenomena identified in the output/results? [Yes, No]	Definition of emergent phenomena: (as in Grimm & Railsback, 2005) • Emergent properties are not simply the sum of the

		from the individuals? (Emergence)		properties of the individuals, • Emergent properties are of a different type than the properties of the individuals (e.g., the spatial distribution of individuals is a system property of a type that none of the system's individuals has), and • Emergent properties cannot easily be predicted by looking only at the individuals.

Working Paper: Extending a Farm Model by means of Agent Based Model for Evaluating CAP Regionalization Scenarios

Extending a Farm Model by means of Agent Based Model for Evaluating CAP Regionalization Scenarios

Dimitris Kremmydas
Researcher, Agricultural University of Athens

Introduction

Environmental concerns in contemporary agricultural policy make the modeling at a disaggregated spatial level very relevant (Ewert et al. 2011, Kampas et al. 2012, Donati et al. 2015). Furthermore it is widely acknowledged that agricultural systems should be apprehended as coupled socio-ecological systems (Ewert et al. 2009, Filatova et al. 2016) and thus it is necessary to the link of farm decisions with agri-environmental status, and consider any related feedbacks. In that context, farm scale policy analysis is receiving increased attention (Langrell et al., 2013). Louhichi et al. (2013) identify five types of models with the farm as the decision-making unit: mathematical programming, econometric, econometric-mathematical programming, simulation approach and agent based models.

Among these types, agent based models (ABM) have appeared in the end of 1990's (Balmann 1997; Berger 2001; Happe 2004) for modeling structural change of agricultural production systems. ABMs can be considered descendants of the recursive mathematical programming (MP) approach, as the initial ABMs included a typical MP production/investment problem coupled with a land market module that was solved recursively. The innovative elements were: the ability to include farms' interaction and in this way to evaluate the direction of the structural change (farm growth, farm entry/exit); and the implicit consideration of spatial dimension which potentially could lead to operational integration with biophysical models. Since then, ABMs have been used for a wide range of policy evaluation goals: land use change (Brady et al. 2012, Nainggolan et al 2012, Troost et al. 2014), environmental impact assessment (Daloglu et al. 2014, Gimona et al. 2011, Schouten et al. 2013), structural change (Appel et al. 2016, Mack et al. 2013, Troost et al. 2015, Lobianco et al. 2010), production and market projections (Appel et al. 2016, Lobianco et al. 2010) and technology diffusion (Alexander et al. 2015, Rebaudo et al. 2013).

Due to the growing body of ABM literature, we identify a distinct class of ABMs, where the decision-making unit is the individual farm, naming it as "individual farm agent based modeling" (IF-ABM). The individual farm scale raises special modeling issues and several methodological challenges which are discussed in the rest of the paper. As an example, consider the works of Troost et al. (2015) and Sorda et al. (2013), which both deleloped agent based models for modeling biogas production in an agricultural system. The first paper considers individual farms as the agents of the simulation while the latter defines agents in a more aggregate level, such as federal government, bank, electric utility, municipality representative investor and municipality substrate supplier. The methodological and data requirements are obviously very different. Troost et al. (2015) shall be classified as IF-ABM while Sorda et al. (2013) does not fit in this subcategory.

The IF-ABM approach offers several advantages regarding the evaluation of contemporary agricultural policies, considering that environmental aspects are in the foreground. Such models can be far more easily coupled with existing environmental and ecological models than the mainstream representative farm models. Some examples of IF-ABM coupled with a water quality model is found in Daloglu et al. (2014), with a nitrogen loss module in Happe et al. (2011), with an individual-based model of skylark breeding population to assess biodiversity in Guillem et al. (2015) and with a species metacommunity model in Gimmona et al. (2011) and Polhill et al. (2013). Other advantages are their ability to include space dimension, to better represent farms' heterogeneity and to thoroughly model the interaction dynamics. The latter is very important for exploring the complexity of socio-ecological systems (Xepapadeas 2010, Levin et al. 2013, Polhill et al. 2016).

In this paper, after discussing the pros and cons of switching from representative to individual farm modeling, we present a policy evaluation case study modeled in the ABM approach. We describe the model and data setup and present the results, closing with a discussion on how the ABM and the representative farm modeling can complement each other.

The transition from Representative to Individual Farms

Most often, current farm-level approaches are most commonly using *representative farms* (Reidsma et al., 2018). The actual population of individual farms has been reduced through sampling to a number of representative farms, based on relevant criteria such as production activity, economic size, administrative region, etc. On the other hand, as discussed in Kremmydas et al. (2018), Agent Based Models for empirical policy evaluation use data on individual farms, trying to reproduce the behavior and the effects on the actual farm population in a study area.

A question arises: What are the additional benefits and the additional costs of modeling with individual-farms compared to representative-farms? We summarize them in Table 1 and discuss them in more detail underneath.

Table 1, Summary of additional cost/benefit analysis of switching from "representative farms" to "individual/agent farm" scale regime

Additional Benefits	Additional Costs
More accurate representation of spatial dimension –	Data availability and accuracy is diminished
Can represent farm heterogeneity better	Higher number of assumptions
Better representation of interaction dynamics	Conclusions less robust
Better alignment with biophysical models	

The transition from representative to individual farm level modeling comes with the following benefits:

1. Spatial dimension is more accurately represented. Disaggregating the agents means that each agent can be positioned in an individual location. For instance, 5 representative groups of agents can be located to maximum five locations while decomposing the group to 500 agents, means that now at maximum 500 locations can be attributed. This is particularly important when including environmental aspects in the policy model, where space is essential (e.g. land use change, diffusion effects, etc.).

2. Representing Farm heterogeneity is enhanced though ABM. This is self-evident since the number of agents are increased and potentially individual property values can be assigned.

3. From (1) and (2) it follows that the accuracy of the interacting dynamics is clearly improved. Despite that, modelling agents; interactions, this is only possible at inter-group level. By contrast, , the transition to individual agent regime, combined with the spatial allocations of the agents allows modelling of fairly detailed interaction dynamics, drilling down to modelling interactions among individuals. This enables the potential to identify intra-group effects that emerge from those diverse interactions.

We illustrate this in Figure 1, where we have 3 interacting representative groups and we break the groups, simply by cloning the group agents (G1 has 3 g1 agents, G2 has 2 g2, G3 has 3 g3) and positioning them in a social network. The interaction dynamics representation is more elaborate, independently of the heterogeneity of the agents (even if agents are clones of the representative agent).

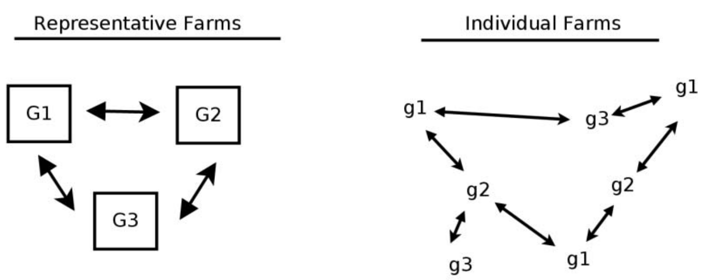


Figure 11, Interaction dynamics comparison between "representative agents" and "agents" scale regime

However the transition from representative to individual farm is accomplished at the expense of additional costs:

1. Data availability and data accuracy is limited. The finest level of the current agricultural databases is usually representative farms through the use of stratified sampling (e.g. EU Farm Accountancy Data Network -FADN or US Agricultural Resource Management Survey - ARM). Thus the use of “representative farms” modeling has a sound statistical base, which is not the case for the “individual farm” level. Either field surveys have to be conducted or statistical techniques to be employed in order to include in the model the farm population.

2. Large number of assumptions. As previously stressed, the majority of low-level data sources focuses on representative agents and most probably not all the required data variables of the agent level are available. For example the managerial capacity of each farm can only be deduced by data not collected. In the case of the individual farm's future price expectation formation, no related datasets are available and assumptions are rather heroic Thus for the "missing variables" new assumptions have to be made.

3. Due to points (1) and (2), the results of the model may no be robust. This drawback calls for using experimental design and sensitivity analysis techniques to assess the output reliability.

In particular, the evaluation of regionalization with a representative farm model, a representative farm model cannot provide insights into the impacts of different policy options to farms through the effects on local resource markets like land market and animal feeds that are produced and consumed locally. Thus the purpose of the study is to complement the existing representative farm model with an Agent Based version that will provide insights on the above facets.

The Agent Based Farm model

An overview of the components and the functionality of the ABM model are given in Figure 2. The rectangular shapes are submodels while the curved shapes contain the data that is transferred from one submodel to the next. The ABM model, after some initialization procedures, starts from the *Plan Production* submodel and finishes in the *Land Market* submodel. Each completion of the *Plan Production – Update Financial Status – Land Market* path constitutes a model round and corresponds to a single production year.

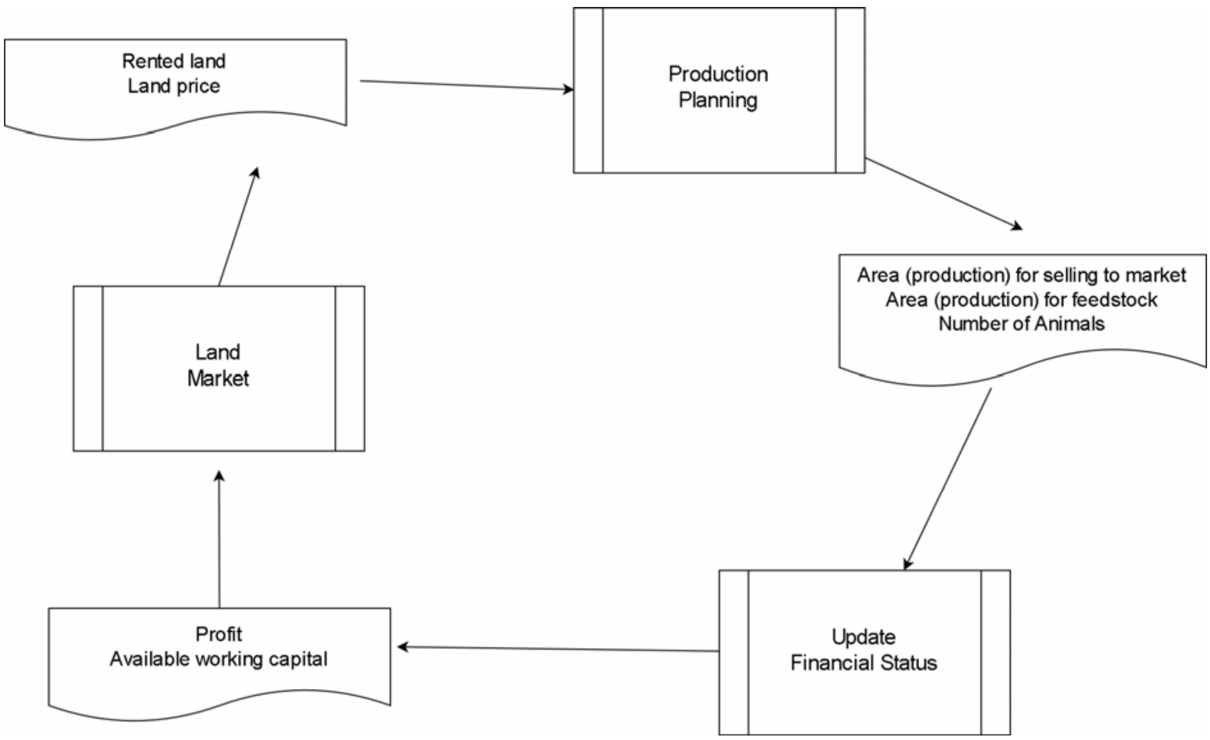


Figure 12, Overview diagram of the ABM model

Regarding the *Production Planning* phase*,* we maintain the GREFAM model and input data structure and thus all the details found in the GREFAM manual hold for the ABM farm mathematical programming model. The output of the model is the level of crop activity (in hectares) for selling to the market and for feeding the farm's animals. Each farm also decides on the number of animals it will hold for the current production year. This combined with the milk yield is the base for calculating the livestock income.

In the *Update Financial* Status submodel we combine the production decisions of the production planning submodel with the prices of the activities' output and the land rent paid by each farm to update their financial status. We keep the current round's gross profit but we also have the accumulated gross profit for all previous rounds. Then the available working capital is updated as a percentage of the current round profit. In the current version of the model we keep this fixed across all rounds but it is differentiated from farm to farm based on the observed value of the base year. In this step, farms that their accumulated profit is negative are considered to exit farm activity and their owned land is considered abandoned wile their rental contracts are immediately expired.

The *Land Market* module performs the following operations:

- Calculate the land market supply. This is the sum of abandoned land plus any expired rental contracts during the current round.
- Calculate the farms' Willingness-To-Pay for land. It equals the farms' shadow price for irrigated and non-irrigated land. We calculate the shadow prices by parametrically solving the model for additional land units using the updated working capital from the *Update Financial Model* submodel.
- Market Clearing. For each land supply unit a random procedure sets the new tenant and the price. The probability for a farm renting a land unit is proportional to its shadow price, thus farms with higher land shadow price will end up with a higher proportion of the newly rented land. For each rented and unit, the price is a random number between the current average land price in the area and the shadow price of the new tenant.

The *Land Market* module updates the rented areas and a new round begins from the *Production Planning* submodel.

In representative farm models, as explained in Kremmydas et al. (2018), market is cleared in an aggregated way. The willingness to accept and the willingness to pay for each farm will be calculated, not differing from the current ABM approach, and the demand and supply functions will be constructed in order to find their intersection that defines the land equilibrium price. In contrast, in the ABM approach, the market is cleared through individual transactions. That inevitably drives ABM to include randomness in the clearing result, since the probability that two agents (a buyer and a seller) will meet to negotiate a transaction is a random variable itself.

Data

We consider only farms that belong to TYPE-OF-FARMING codes 1 (field crops), 4 (livestock farms) and 8 (mixed field and livestock farms) of the FADN nomenclature. We do so because in the current GREFAM version permanent crop cannot vary and thus those farms cannot grow or shrink.

We also assume that farms that are located in the same NUTS-3 unit can be considered. This is indeed a heroic assumption and will be settled in a subsequent version of the model where methods for recreating the farm population in the municipality level will be used.

We selected a limited number of NUTS-3 areas based on the following criteria

- To contain more than 30 FADN sample farms. This assures that results will be of a minimum level of statistical confidence in the NUTS-3 level.
- The farms that belong to the considered TYPE-OF-FARMING (i.e., field, livestock, mixed) to be more than 60% of all farms in the region. Thus the conclusions drawn for the selected farms will be representative for the NUTS-3 area.
- The permanent crops, vegetable and grazing areas not to exceed a 20% threshold of total utilized agricultural area in the NUTS-3 level.

The selected NUTS-3 units are shown in Table 2.

Table 2, Selected NUTS-3 properties

	Number of farms		% land use of permanent crops	Type-of-farming for selected farms		
NOMOS	**All FADN farms**	**Selected farms[1]**		**Field**	**Livestock**	**Mixed**
KARDITSA	139	135	9%	93%	4%	3%
PREVEZA	66	58	27%	10%	62%	28%
THESSALONIKI	295	256	9%	82%	16%	3%
SERRES	286	248	6%	89%	7%	4%
AIT/NIA	108	91	31%	13%	53%	34%
KASTORIA	69	53	12%	66%	32%	2%
BOIOTIA	86	66	9%	100%	0%	0%
LARISA	198	133	5%	82%	13%	5%
	[1] farms that belong to type-of-farming 1,4 and 8					

An additional selection criterion was to include regions with a diversified production profile. SERRES, BOIOTIA and KARDITSA are predominantly arable areas while PREVEZA and AIT/NIA are mainly livestock areas.

Regarding the selected policy scenarios, we examine the flat rate versus a set of regionalization scenarios. This set is derived from the Pareto efficient set acquired from the evaluation of a greater number of scenarios using a representative farm model. The selected regionalization scenarios are shown in Table 3.

Table 3, Selected Regionalization scenarios

Scenario				
Name	**Arable**	**Trees**	**Grazing**	**Function**
agron.0-50-50.2013	0%	50%	50%	pro-environmental, Support livestock
agron.70-0-30.2013	70%	0%	30%	Support field crops
agron.65-20-15.2013	65%	20%	15%	Resembles flat rate

Results

We run the ABM extension model for a 10 year period for all selected NUTS-3 areas, for each of the flat rate scenario and the chosen regionalization scenarios. We recorded data on land use, animal heads, land prices, shadow prices for irrigated and non-irrigated areas and farm accounting elements like working capital and gross margin. The overall size of the output data file amounts to over 300 MB.

Land Rent

In Table 4 we present the median ratio of the rent price of regionalization scenarios to the flat rate scenario. For instance in the first cell (Irrigated Land > AIT/NIA > agron.0-50-50.2013), the median rent of the regionalization scenario is the 0.94 of the flat rate scenario.

Table 4, Land rent price index

		agron.0-50-50.2013	agron.65-20-15.2013	agron.70-0-30.2013
Irrigated Land	AIT/NIA	0.94	1.00	0.98
	BOIOTIA	0.70	1.03	1.03
	KARDITSA	0.94	1.05	1.08
	KASTORIA	1.55	1.13	1.04
	LARISA	0.66	1.07	1.08
	PREVEZA	0.95	1.08	1.09
	SERRES	0.70	1.06	1.09
	THESSALONIKI	1.02	1.05	1.08
		agron.0-50-50.2013	agron.65-20-15.2013	agron.70-0-30.2013
non Irrigated	AIT/NIA	0.98	1.11	1.14
	BOIOTIA	0.49	1.07	1.07

Land	KARDITSA	0.52	1.09	1.15
	KASTORIA	0.33	1.14	1.18
	LARISA	0.52	1.10	1.15
	PREVEZA	1.04	1.09	1.21
	SERRES	0.43	1.11	1.17
	THESSALONIKI	0.52	1.11	1.20

The three regionalization scenarios have different impact to rent of irrigated and non-irrigated land, even in the same area. For instance, in *THESSALONIKI*, the *agron.0-50-50.2013* scenario drops significantly the land rent of the non irrigated land (0.52 of the flat rate scenario) while does not affect that of the irrigated areas (1.02).

Additionally, the three regionalization scenarios have a different impact across different prefecture. For instance, for irrigated land, the *agron.0-50-50.2013* scenario rises land rents for *KASTORIA* while lowers them for LARISA, SERRES and BOIOTIA.

Gross Margin

In Table 5 we present the ratio of the sum of gross margins over all farms of regionalization scenarios to that of the flat rate scenario. Those ratios are given for each NOMOS and type of farming. For instance in the first cell (AIT/NIA > ARABLE > agron.0-50-50.2013), the sum of gross margin of AIT/NIA of ARABLE farms is 0.75 of the flat rate equivalent.

Table 5, Gross Margin index

PREFECTURE	FARMING TYPE	agron.0-50-50.2013	agron.65-20-15.2013	agron.70-0-30.2013
AIT/NIA	ARABLE	0.75	1.02	1.03
	LIVESTOCK	0.93	1.01	1.02
	MIXED	0.94	1.01	0.99
BOIOTIA	ARABLE	0.74	1.05	1.06
KARDITS	ARABLE	0.63	1.06	1.09
	LIVESTOCK	1.04	1.01	1.08
	MIXED	0.81	1.02	1.08
KASTORIA	ARABLE	0.82	1.01	1.07
	LIVESTOCK	0.85	0.99	1.02
	MIXED	0.66	1.05	1.08
LARISA	ARABLE	0.65	1.05	1.07
	LIVESTOCK	0.99	0.96	1.01
	MIXED	1.16	0.99	1.01

PREFECTURE	FARMING TYPE	agron.0-50-50.2013	agron.65-20-15.2013	agron.70-0-30.2013
PREVEZ	ARABLE	0.24	1.03	1.16
	LIVESTOCK	0.79	1.00	0.99
	MIXED	0.97	1.03	1.05
SERRES	ARABLE	0.58	1.06	1.08
	LIVESTOCK	1.12	0.99	1.08
	MIXED	1.02	1.07	0.90
THESSALON	ARABLE	0.43	1.09	1.13
	LIVESTOCK	0.89	1.00	0.98
	MIXED	0.56	1.10	1.15

We observe that the *agron.0-50-50.2013* scenario affects negatively the gross margin of farms specialized in arable farming. However the impact is much differentiated among prefectures.

Discussion and Conclusions

In this paper we use an Agent Based Model to complement the policy analysis performed in a previous paper by means of a representative farm model. The latter provided adequate guidelines in the country level, however effects due to farm interaction (e.g. structural change, land value, etc.) or in finer spatial detail were not observable.

Based on the results fo the representative farm model, we select the most interesting policy scenarios and regions with different production structure (only arable, mostly livestock, mixed). We then evaluate the effects of the different policy implementations to the different regions.

The working paper is planned to be further improved with the following additions:

- In the current setting individual farms are the original FADN dataset farms. However advanced population synthesis techniques can be used in order to recreate the farm population more accurately. Beyond the agricultural modelling domain there is active research on synthesizing population based on sample data as in Harland et al. (2012)
- We will represent the farm spatial distribution even more accurately by using GIS and cadastral data to represent biophysical parameters (soil quality, weather, etc.) and other spatial specific properties like the number of farms' different plots.
- The current model and data setting can easily be extended to a simulation metamodel using Desing of Experiments techniques as discussed in Kleijnen et al. (2005).

References

Alexander, P., Moran, D., & Rounsevell, M. D. A. (2014). Evaluating potential policies for the UK perennial energy crop market to achieve carbon abatement and deliver a source of low carbon electricity. Biomass and Bioenergy, 82, 3–12. http://doi.org/10.1016/j.biombioe.2015.04.025

Appel, F., Ostermeyer-Wiethaup, A., & Balmann, A. (2016). Effects of the German Renewable Energy Act on structural change in Agriculture – The case of biogas. Utilities Policy, 1–11. http://doi.org/10.1016/j.jup.2016.02.013

Balmann, A. (1997). "Farm-Based Modelling of Regional Structural Change: A Cellular Automata Approach." European Review of Agricultural Economics 24(1): 85-108.

Berger, T., 2001. Agent-based models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics 25(2/3), 245-260.

Brady, M., Sahrbacher, C., Kellermann, K., & Happe, K. (2012). An agent-based approach to modeling impacts of agricultural policy on land use, biodiversity and ecosystem services. Landscape Ecology, 27(9), 1363–1381. http://doi.org/10.1007/s10980-012-9787-3

Daloğlu, I., Nassauer, J. I., Riolo, R., & Scavia, D. (2014). An integrated social and ecological modeling framework - Impacts of agricultural conservation practices on water quality. Ecology and Society, 19(3), 12. http://doi.org/10.5751/ES-06597-190312

Donati, M., Menozzi, D., Fioravanzi, M. (2015). Understanding farmers' responses to CAP reform, New Medit, 14 (3), pp. 29-39.

Ewert, F., van Ittersum, M. K., Bezlepkina, I., Therond, O., Andersen, E., Belhouchette, H., … Wolf, J. (2009). A methodology for enhanced flexibility of integrated assessment in agriculture. Environmental Science and Policy, 12(5), 546–561. http://doi.org/10.1016/j.envsci.2009.02.005

Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I., Andersen, E. (2011). Scale changes and model linking methods for integrated assessment of agri-environmental systems, Agriculture, Ecosystems and Environment, 142 (1-2), pp. 6-17.

Filatova, T., Polhill, J. G., & van Ewijk, S. (2016). Regime shifts in coupled socio-environmental systems: Review of modelling challenges and approaches. Environmental Modelling and Software, 75, 333–347. http://doi.org/10.1016/j.envsoft.2015.04.003

Gimona, A., & Polhill, J. G. (2011). Exploring robustness of biodiversity policy with a coupled metacommunity and agent-based model. Journal of Land Use Science, 6(2-3), 175–193. http://doi.org/10.1080/1747423X.2011.558601

Gimona, A., & Polhill, J. G. (2011). Exploring robustness of biodiversity policy with a coupled metacommunity and agent-based model. Journal of Land Use Science, 6(2-3), 175–193. http://doi.org/10.1080/1747423X.2011.558601

Guillem, E. E., Murray-Rust, D., Robinson, D. T., Barnes, A., & Rounsevell, M. D. A. (2015). Modelling farmer decision-making to anticipate tradeoffs between provisioning ecosystem services and biodiversity. Agricultural Systems, 137, 12–23. http://doi.org/10.1016/j.agsy.2015.03.006

Happe, K., Balmann, A., & Kellermann, K. (2004). The agricultural policy simulator (AgriPoliS): an agent-based model to study structural change in agriculture (Version 1.0) (No. 71). Discussion paper//Institute of Agricultural Development in Central and Eastern Europe.

Happe, K., Hutchings, N. J., Dalgaard, T., & Kellerman, K. (2011). Modelling the interactions between regional farming structure, nitrogen losses and environmental regulation. Agricultural Systems, 104(3), 281–291. http://doi.org/10.1016/j.agsy.2010.09.008

Harland, K., Heppenstall, A., Smith, D., & Birkin, M. (2012). Creating realistic synthetic populations at varying spatial scales: A comparative critique of population synthesis techniques. JASSS, 15(1). http://doi.org/10.18564/jasss.1909

Kampas, A., Petsakos, A., Rozakis, S. (2012) . Price induced irrigation water saving: Unraveling conflicts and synergies between European agricultural and water policies for a Greek Water District, Agricultural Systems, 113, pp. 28-38.

Kremmydas, D., Athanasiadis, I. N., & Rozakis, S. (2018). A review of Agent Based Modeling for agricultural policy evaluation. *Agricultural Systems*, *164*, 95-106.

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M. (2005). State-of-the-Art Review: A User's Guide to the Brave New World of Designing Simulation Experiments. INFORMS Journal on Computing, 17(3), 263–289. http://doi.org/10.1287/ijoc.1050.0136

Langrell, S., Ciaian, P., Espinosa Goded Maria, Gomez Y Paloma, S., Heckelei, T., Sckokai, P., … Vard, T. (2013). Farm level modelling of CAP: a methodological overview. Publications Office of the European Union. Retrieved from http://publications.jrc.ec.europa.eu/repository/handle/JRC79969

Levin, S., Xepapadeas, T., Crépin, A. S., Norberg, J., De Zeeuw, A., Folke, C., ... & Ehrlich, P. (2013). Social-ecological systems as complex adaptive systems: modeling and policy implications. Environment and Development Economics, 18(02), 111-132.

Louhichi, K., Ciaian, P., & Gomez y Paloma, S. (2013.). Farm-level models for EU policy analysis: Review of recent literature and comparison of most relevant models. In S. Langrell (Ed.), Farm level modelling of CAP: A methodological overview,. EC-JRC-IPTS.

Mack, G., Möhring, A., Ferjani, A., Zimmermann, A., & Mann, S. (2013). Transfer of single farm payment entitlements to farm successors : impact on structural change and rental prices in Switzerland. Bio-Based and Applied Economics, 2(2), 113–130.

Nainggolan, D., Termansen, M., Fleskens, L., Hubacek, K., Reed, M. S., de Vente, J., & Boix-Fayos, C. (2012). What does the future hold for semi-arid Mediterranean agro-ecosystems? - Exploring cellular automata and agent-based trajectories of future land-use change. Applied Geography, 35(1-2), 474–490. http://doi.org/10.1016/j.apgeog.2012.09.007

Polhill, J. G., Filatova, T., Schlüter, M., & Voinov, A. (2016). Modelling systemic change in coupled socio-environmental systems. Environmental Modelling & Software, 75, 318-332.

Polhill, J. G., Gimona, A., & Gotts, N. M. (2013). Nonlinearities in biodiversity incentive schemes: A study using an integrated agent-based and metacommunity model. Environmental Modelling and Software, 45, 74–91. http://doi.org/10.1016/j.envsoft.2012.11.011

Rebaudo, F., & Dangles, O. (2013). An agent-based modeling framework for integrated pest management dissemination programs. Environmental Modelling & Software, 45, 141–149. http://doi.org/10.1016/j.envsoft.2012.06.014

Reidsma, P. *et al.* (2018) 'On the development and use of farm models for policy impact assessment in the European Union – A review', *Agricultural Systems*. Elsevier, 159(September 2017), pp. 111–125. doi: 10.1016/j.agsy.2017.10.012.

Sorda, G., Sunak, Y., & Madlener, R. (2013). An agent-based spatial simulation to evaluate the promotion of electricity from agricultural biogas plants in Germany. Ecological Economics, 89, 43–60. http://doi.org/10.1016/j.ecolecon.2013.01.022

Troost, C., & Berger, T. (2014). Dealing with uncertainty in agent-based simulation: Farm-level modeling of adaptation to climate change in southwest Germany. American Journal of Agricultural Economics, 97(3), 833–854. http://doi.org/10.1093/ajae/aau076

Troost, C., Walter, T., & Berger, T. (2015). Climate, energy and environmental policies in agriculture: Simulating likely farmer responses in Southwest Germany. Land Use Policy, 46, 50–64. http://doi.org/10.1016/j.landusepol.2015.01.028

Xepapadeas, A. (2010). Modeling complex systems. Agricultural Economics, 41, 181–191. http://doi.org/10.1111/j.1574-0862.2010.00499.x

Appendix - Code

Code Description

The simulation is controlled through a batch script file, *simulation.full.bat*. This controls the looping over all scenarios and NUTS-3 regions and some other configuration aspects like the number of rounds (years) and whether a debug mode is on or off.

Data initialization is done using the *init_simuation.R* script. We use fadnUtils package to read data for the selected NUTS-3 region and filter the field, livestock and mixed farms. It creates the initial farm monitoring files like the farm accounting books, the land and rental register. It also calculates the initial market prices and the average rental price for irrigated and non-irrigated land.

The *keep_history.R* script, as its names implies, records the value of an extensive number of variables for each round and each farm.

The Production Planning phase is realized by calling a gams model file that is very similar to the GREFAM model.

The *update_accounts.R* script loads the simulation current status (land and rental register, production plan, farm accounts) and updates them after the proper calculations. It also determines the farms that exit the simulation and reallocate their land.

For realizing the Land Market module, the shadow price for each additional unit of land is calculated. The gams model file is used here too, called by the *simulation_full.bat* where a special external parameter is passed so that the gams file iteratively calculates those shadow prices.

The *land_market.R* script reads the shadow prices and the abandoned land estimated in previous steps, clears the market and update the appropriate registers.

We provide the source code of the above scripts in the following pages.

runSimulation.full.bat

```
@ECHO OFF

REM .\runSimulation.full.bat | tee simulation.log

cls

REM ------------------------
REM  configuration
call ./configuration.bat

REM  Number of years running simulation
SET YEARS=10

REM Shal gams output lst file
set DEBUG=FALSE

for %%N in (
SERRES
"AIT/NIA"
KASTORIA
LARISA
PREVEZ
THESSALON
KARDITS
BOIOTIA
    ) do (

    REM START nomos LOOP
    for %%S in (
flat.rate.2013
agron.0-50-50.2013
agron.70-0-30.2013
agron.65-20-15.2013
        ) do (

        REM START scenario LOOP
echo --------------------------------------------------------------------
echo ----                                  ----------------------------------
echo ---- NOMOS: %%N,    SCENARIO: %%S
echo ----                                  ----------------------------------
echo --------------------------------------------------------------------
echo Initializing Data
REM delete previous data data
del .\gams\scenarios\%%S\%RESULTS_FOLDER%\* /Q
del .\gams\scenarios\%%S\%RESULTS_FOLDER%\history\* /Q
del .\gams\*.lst /Q
del .\gams\*.log /Q

%Rscript%    --vanilla    R_code\init_simulation.R    "%%S"    "%%N"    "%RESULTS_FOLDER%"
%TRANSPORT_COST_OVER% %LEASE_DURATION%
IF %ERRORLEVEL% NEQ 0 (
  echo PROBLEM with R. Return code: %ERRORLEVEL%
  exit
)

   echo.
   echo ----------------------Keep History
%Rscript% --vanilla  R_code\keep_history.R %%S 0 1 %RESULTS_FOLDER%
IF %ERRORLEVEL% NEQ 0 (
  echo PROBLEM with R. Return code: %ERRORLEVEL%
  exit
)

REM  Loop simulation years
for /l %%t in (1, 1, %YEARS%) do (
```

```
    echo.
    echo ---------------------------------------------------------------------------------
----------
    echo Round %%t
    echo ---------------------------------------------------------------------------------
----------

   echo.
   echo ----------------------Crop Plan Phase -----------------------
   @echo echo GAMS did not run > ./gams/cmd_interface.bat

IF "%DEBUG%"=="TRUE" (

%GAMS% "model1.fadn2012.scenario.LP.withLivestock.agronomic.model.gms"  --scen_input="%%S"
--sf_file="sf.txt"          --round_num=%%t     --mode="normal"     --debug=%DEBUG%     --
result_folder=%RESULTS_FOLDER%      workDir="%ccd%/gams"      -o=crop.plan.%%t.lst      -
lf=crop.plan.%%t.log -lo=2

) ELSE (

%GAMS% "model1.fadn2012.scenario.LP.withLivestock.agronomic.model.gms"  --scen_input="%%S"
--sf_file="sf.txt"          --round_num=%%t     --mode="normal"     --debug=%DEBUG%     --
result_folder=%RESULTS_FOLDER% workDir="%ccd%/gams" -o=crop.plan.%%t.lst -lo=0

)

IF %ERRORLEVEL% NEQ 0 (
  echo PROBLEM with GAMS. Return code: %ERRORLEVEL%
  exit
)

call ./gams/cmd_interface.bat

   echo.
   echo ----------------------Update Accounts Phase R-%%t------------------

    %Rscript% --vanilla R_code\update_accounts.R %%S %%t %RESULTS_FOLDER%
    IF %ERRORLEVEL% NEQ 0 (
      echo PROBLEM with R. Return code: %ERRORLEVEL%
      exit
    )


   echo.
   echo ----------------------LandMarket Phase  R-%%t-----------------------

   echo -----find shadow prices
   @echo echo GAMS did not run > ./gams/cmd_interface.bat

    IF "%DEBUG%"=="TRUE" (

        %GAMS%    "model1.fadn2012.scenario.LP.withLivestock.agronomic.model.gms"       --
scen_input="%%S"   --sf_file="sf.txt"     --round_num=%%t   --mode="seek_SHADOW_land"   --
land_range=%LAND_RANGE%    --land_step=%LAND_STEP%    --result_folder=%RESULTS_FOLDER%   --
debug=%DEBUG% workDir="%ccd%/gams" -o=shadow.land.%%t.lst -lf=shadow.land.%%t.log -lo=2

    ) ELSE (

        %GAMS%    "model1.fadn2012.scenario.LP.withLivestock.agronomic.model.gms"       --
scen_input="%%S"   --sf_file="sf.txt"     --round_num=%%t   --mode="seek_SHADOW_land"   --
land_range=%LAND_RANGE%    --land_step=%LAND_STEP%    --result_folder=%RESULTS_FOLDER%   --
debug=%DEBUG% workDir="%ccd%/gams" -o=shadow.land.%%t.lst -lo=0
    )

    IF %ERRORLEVEL% NEQ 0 (
      echo PROBLEM with GAMS. Return code: %ERRORLEVEL%
      exit
    )
```

```
    call ./gams/cmd_interface.bat

   echo -----clear market R-%%t
REM run land market
%Rscript% --vanilla R_code\land_market.R %%S %%t %RESULTS_FOLDER% %DEBUG% %LEASE_DURATION%
IF %ERRORLEVEL% NEQ 0 (
  echo PROBLEM with R. Return code: %ERRORLEVEL%
  exit
)

   echo.
   echo ----------------------Keep History  R-%%t
%Rscript% --vanilla  R_code\keep_history.R %%S %%t 1 %RESULTS_FOLDER%
IF %ERRORLEVEL% NEQ 0 (
  echo PROBLEM with R. Return code: %ERRORLEVEL%
  exit
)

REM end loop simulation years
)

REM copy history
copy                                       ".\gams\scenarios\%%S\%RESULTS_FOLDER%\history.txt"
".\gams\scenarios\history.%%S.%%N.txt"

        REM END scenario LOOP
        )

    REM END nomos LOOP
    )

:END
```

init_simulation.R

```
# init_simulation.R
#
# DESCRIPTION ----
# ..........................................................................
# Initialize input data
#
#
# INPUT FILES----
# ..........................................................................
# - info.data.incl.rds (info from prepare_data process)
# - prod.data.incl.rds (prod from prepare_data process)
# - farms.txt (the FID.f of farms that are used in scenario)
# - other_working_capital.original.txt (working capital calculated from prepare_data for all greece)
# - land.currentAllocation.txt (the farm-activity pairs calculated from the prepare_data for all greece)
#
#
# OUTPUT FILES ----
# ..........................................................................
# - land_register.txt (land register for each farm)
# - rental_register.txt
# - farm_accounting.txt (the accounting status of the farm: FID, FID.f, WC.avail, PROFIT.current, PROFIT.past, PROFIT.accum)
# - %sf_file%.inc  (sf of farms in the simulation)
# - market_prices.inc (the price of crops sold in market. Applied to all farms)
# - feed_prices.inc (prices of purchased feedstock)
#
#
# PARAMETERS ----
# ..........................................................................
# - args[1]: scenario name
# - args[2]: nomos
# - args[3]: result folder name
# - args[4]: transport overhead
# - args[5]: lease_duration

#...........................................................................................................
# load libraries and parameters ----

#options(warn=-1)

library(data.table)
library(fadnUtils)
source("R_code/variables.R")

scenarios.basepath = "gams/scenarios/"

args = commandArgs(trailingOnly=TRUE)
#scenario = "flat.rate.2013"
scenario = args[1]
#nomos="SERRES"
nomos=args[2]
#result.folder="results"
result.folder=args[3]
#transport.overhead=.25
transport.overhead=as.numeric(args[4])
#lease_duration=4
lease_duration=as.numeric(args[5])

save.folder = paste0(scenarios.basepath,scenario,"/",result.folder);
print(paste0("saving to ",save.folder))

#...........................................................................................................
```

```
# functions ----
filter.nomos = function(NOMOS.text, TF.c=c("1","4","8"),

info.data.sel.data=info.data.sel,prod.data.sel.data=prod.data.sel,lvst.number.data.sel.dat
a=lvst.number.data.sel) {
  info.data.sel.nomos          =          info.data.sel.data[NOMOS==NOMOS.text          &
substr(TYPE_OF_FARM_real,1,1)%in%TF.c]
  prod.data.sel.nomos = prod.data.sel.data[FID%in%info.data.sel.nomos$FID]
  lvst.number.sel.nomos                                                                  =
merge(lvst.number.data.sel.data[FID%in%info.data.sel.nomos$FID][,FID:=as.numeric(FID)],
                                info.data.sel.nomos[,list(FID,WEIGHT)],
                                by="FID")
  fids.nomos=info.data.sel.nomos[,FID]

  #select certain fids
  #fids.nomos=c(3103);

  return(list(
    info=info.data.sel.nomos,
    prod=prod.data.sel.nomos,
    lvst.number=lvst.number.sel.nomos,
    fids=fids.nomos
  ))
}

#.........................................................................................
......................
# loading fadnUtils ----
print("..reading fadnUtils")
info.data = readRDS(paste0(scenarios.basepath,scenario,"/info.data.incl.rds"))
prod.data = readRDS(paste0(scenarios.basepath,scenario,"/prod.data.incl.rds"))
lvst.number.data=ELL_2013$livestock$livestock.number

#.........................................................................................
......................
# reading eligible farms ----
print(paste0("Reading farms from ",paste0(scenarios.basepath,scenario,"/farms.txt")))
fids.selected                                                                           =
data.table(read.table(file=paste0(scenarios.basepath,scenario,"/farms.txt"),
                                      sep=" ",
                                      header = F,
                                      col.names = c("FID.f"))
)[,list(FID.f,FID=as.numeric(gsub("f","",FID.f)))]
print(paste0("      .... ", nrow(fids.selected)," eligible farms from all Greece found"))

info.data.sel = info.data[FID%in%fids.selected$FID]
prod.data.sel = prod.data[FID%in%fids.selected$FID]
lvst.number.data.sel=lvst.number.data[FID%in%fids.selected$FID]

nomos.data=filter.nomos(nomos)

#.........................................................................................
......................
# Write %sf_file%.inc ----
save.sf.file(save.folder,nomos.data$fids)

#.........................................................................................
......................
# create farm_accounting.txt   ----
#FID, FID.f, WC.avail, PROFIT.current, PROFIT.past,PROFIT.accum
print("Creating farm_accounting")

farm_accounting=merge(
  data.table(
    FID=nomos.data$fids,  PROFIT.current=0.0, PROFIT.past=0.0,PROFIT.accum=0.0
  ),
```

```
data.table(read.table(file=paste0(scenarios.basepath,scenario,"/other_working_capital.orig
inal.txt"),
                        col.names                                                    =
c("FID.f","WC.avail")))[,FID:=as.numeric(gsub("f","",FID.f))][FID%in%nomos.data$fids],
  all.x=T,by="FID"
)
farm_accounting[is.na(farm_accounting)]=0;


setcolorder(farm_accounting,c("FID","FID.f","WC.avail","PROFIT.current","PROFIT.past","PRO
FIT.accum"))
save.farm.accounting(save.folder,farm_accounting)

#.........................................................................................
......................
# create and save land_abandoned.txt  ----
# land_rented_total(sf),land_rented_irr(sf)

print("Creating land_abandoned")
land_abandoned = data.table(TYPE=c("IRR","nIRR"),AREA=as.numeric(c(0.0,0.0)))

save.land.abandoned(save.folder,land_abandoned)


#.........................................................................................
......................
# create and save land_register.txt  ----
# land_rented_total(sf),land_rented_irr(sf)

print("Creating land_register")
#load and calculate irrigated land
# TOTAL=all land (OWNED+RENTED)
# OWNED=land that is owned by the farm (MUTATED+FIXED)
# RENTED=land that is rented (only MUTATED)
# MUTATED=land that farms can select crops
# FIXED=land with permanent crops (we consider it to be OWNED by the farm)
#
simulation.fids = read.simulation.fids(paste0(save.folder,"/sf.txt"))
land.current                                                                             =
read.current.allocation(paste0(save.folder,"/../land.currentAllocation.txt"))[FID%in%simul
ation.fids$FID]

activities.incl = unique(land.current$ACTIVITY)
activities.incl.irr = c(activities.incl[grep("\\.irr",activities.incl)],
                        activities.incl[grep("\\.garden",activities.incl)],
                        "tobacco")
#find total and rented land
land.status = dcast(
  rbind(
    land.current[,list(AREA=sum(AREA),TYPE="TOTAL"),by=FID],
    info.data.sel[FID%in%simulation.fids$FID,list(FID,AREA=RENTED_UAA,TYPE="RENTED")]
  ),
  FID~TYPE,value.var = "AREA"
)
land.status[is.na(land.status)]=0

#if rented>total (not normal) then set total=rented
land.status[RENTED>TOTAL,RENTED:=TOTAL]

#calculate OWNED
land.status[,OWNED:=TOTAL-RENTED]

#add mutated land
land.status=merge(
  land.status,
  dcast(
```

```
land.current[ACTIVITY%in%mutated.activities,list(AREA=sum(AREA),TYPE="MUTATED"),by=FID],
    FID~TYPE,value.var = "AREA"
  ),
  all.x=T,by="FID"
)
land.status[is.na(land.status)]=0

#calculate OWNED.FIXED
land.status[,":="(OWNED.FIXED=TOTAL-MUTATED)]
land.status[OWNED.FIXED>OWNED,RENTED:=round(RENTED-(OWNED.FIXED-OWNED),2)]
land.status[OWNED.FIXED>OWNED,OWNED:=OWNED.FIXED]

#allocate MUTATED to RENTED and OWNED
land.status[,OWNED.MUTATED:=OWNED-OWNED.FIXED]
land.status[,RENTED.MUTATED:=RENTED]

#add IRR land
land.status=merge(
  land.status,
  land.current[ACTIVITY%in%activities.incl.irr                                         &
ACTIVITY%in%mutated.activities,list(IRR.MUTATED=sum(AREA)),by=FID],
  all.x=T,by="FID"
)
land.status[is.na(land.status)]=0

land.status=merge(
  land.status,
  land.current[ACTIVITY%in%activities.incl.irr                                         &
!(ACTIVITY%in%mutated.activities),list(IRR.FIXED=sum(AREA)),by=FID],
  all.x=T,by="FID"
)
land.status[is.na(land.status)]=0

land.status[,OWNED.FIXED.IRR:=IRR.FIXED]
land.status[,OWNED.FIXED.nIRR:=round(OWNED.FIXED-OWNED.FIXED.IRR,2)]

land.status[(OWNED+RENTED)>0,OWNED.MUTATED.IRR:=round(IRR.MUTATED*(OWNED/(OWNED+RENTED)),2
)]
land.status[(OWNED+RENTED)>0,RENTED.MUTATED.IRR:=round(IRR.MUTATED*(RENTED/(OWNED+RENTED))
,2)]
land.status[is.na(land.status)]=0

land.status[OWNED.MUTATED>0,OWNED.MUTATED.nIRR:=round(OWNED.MUTATED-OWNED.MUTATED.IRR,2)]
land.status[OWNED.MUTATED.nIRR<0,OWNED.MUTATED.nIRR:=0]
land.status[is.na(land.status)]=0

land.status[,RENTED.MUTATED.nIRR:=round(RENTED.MUTATED-RENTED.MUTATED.IRR,2)]

print("Simulation land (ha):")
print(colSums(land.status))

land_register=land.status[,list(
  FID,FID.f=paste0("f",FID),
  OWNED.all=round(OWNED.MUTATED,2),                 OWNED.irr=round(OWNED.MUTATED.IRR,2),
OWNED.nirr=round(OWNED.MUTATED.nIRR,2),

RENTED.all=round(RENTED.MUTATED,2),RENTED.irr=round(RENTED.MUTATED.IRR,2),RENTED.nirr=roun
d(RENTED.MUTATED.nIRR,2),

FIXED.all=round(OWNED.FIXED,2),FIXED.irr=round(OWNED.FIXED.IRR,2),FIXED.nirr=round(OWNED.F
IXED.nIRR,2)
)]

save.land.register(save.folder,land_register)

#.........................................................................................
......................
```

```r
# create and save rental_register.txt  ----

print("..calculating initial land rental prices ...")

#the means of the rent in areas both for irrigated and non irrigated land
rent = info.data.sel[FID%in%nomos.data$fids & RENTED_UAA>0,V285/RENTED_UAA]

if(length(rent)>2) {
  #do a kmeans
  rent.cluster = kmeans(rent,2)

  rent.irr=round(max(rent.cluster$centers),2)
  rent.nirr=round(min(rent.cluster$centers),2)
} else {
  rent.irr=round(mean(rent),2)
  rent.nirr=round(mean(rent),2)
}

print("..creating land_register")
rental_register                                                              =
data.table(FID=integer(),TYPE=character(),ROUND.expire=integer(),AREA=numeric(),PRICE=nume
ric())

#write few area first a 1-year lease
rental_register = rbind(
  rental_register,
  land_register[RENTED.irr>0 & RENTED.irr<lease_duration,
               list(
                 FID,TYPE="IRR",ROUND.expire=1,AREA=RENTED.irr,PRICE=rent.irr)]
)

rental_register = rbind(
  rental_register,
  land_register[RENTED.nirr>0 & RENTED.nirr<lease_duration,
               list(
                 FID,TYPE="nIRR",ROUND.expire=1,AREA=RENTED.nirr,PRICE=rent.nirr)]
)

#write farms with more area
for(r in seq(from=1,to = lease_duration,by = 1)) {
  rental_register = rbind(
    rental_register,
    land_register[RENTED.irr>lease_duration,
                 list(

FID,TYPE="IRR",ROUND.expire=r,AREA=round(RENTED.irr/lease_duration,2),PRICE=rent.irr)]
  )

  rental_register = rbind(
    rental_register,
    land_register[RENTED.nirr>lease_duration,
                 list(

FID,TYPE="nIRR",ROUND.expire=r,AREA=round(RENTED.nirr/lease_duration,2),PRICE=rent.nirr)]
  )

}

save.rental.register(save.folder,rental_register)

#.....................................................................................
......................
# write market_prices.inc ----
#  prices(c,f)
print("writing market_prices.inc ")
prod.market = nomos.data$prod[!is.na(PRICE),list(FID,ACTIVITY,PRICE)]
```

```
cat(
  file=paste0(save.folder,"/market_prices.inc"),

paste0("prices('",prod.market$ACTIVITY,"','f",prod.market$FID,"')=",round(prod.market$PRIC
E,2),";"),
  sep="\n"
)

#.......................................................................................
......................
# write feed_prices.inc ----
#  PR_fs_purch(fs_purch)
print("writing feed_prices.inc ")
prod.feed.local                                                                        =
nomos.data$prod[ACTIVITY.FAMILY%in%unique(activity.map.dt$ACTIVITY.FAMILY)]
prod.feed.local.prices                                                                 =
prod.feed.local[,list(PRICE=mean(PRICE,na.rm=T)),by=ACTIVITY.FAMILY]

#transport.overhead
prod.feed.global.prices                                                                =
prod.data.sel[ACTIVITY.FAMILY%in%unique(activity.map.dt$ACTIVITY.FAMILY),list(PRICE=mean(P
RICE,na.rm=T)*(1+transport.overhead)),by=ACTIVITY.FAMILY]

prod.feed.local.prices=rbind(
  prod.feed.local.prices,
  prod.feed.global.prices[!ACTIVITY.FAMILY%in%prod.feed.local.prices$ACTIVITY.FAMILY]
)

cat(
  file=paste0(save.folder,"/feed_prices.inc"),

paste0("PR_fs_purch('",prod.feed.local.prices$ACTIVITY.FAMILY,"')=",round(prod.feed.local.
prices$PRICE,2),";"),
  sep="\n"
)
```

keep_history.R

```
# keep_history.R
#
# DESCRIPTION ----
# ........................................................................
# Save input data of each round to a different file (for statsitical analysis purposes)
#
#
# OUTPUT FILES ----
# ........................................................................
# - history.txt
#
#
# PARAMETERS ----
# ........................................................................
# - args[1]: scenario name
# - args[2]: cur.round
# - args[3]: copy.files {1: copy the original input files, 0: do not copy the original
files}
# - args[4]: result folder
#

#.........................................................................................
......................
# load libraries and parameters ----
#options(warn=-1)

library(data.table)
source("R_code/variables.R")

args = commandArgs(trailingOnly=TRUE)
#scenario = "flat.rate.2013"
scenario = args[1]

#cur.round=1
cur.round = args[2]

#copy.files=1
copy.files = args[3]

#result.folder="results"
result.folder=args[4]

scenarios.basepath = "gams/scenarios/"
history.new = data.table(FID.f=character(),ROUND=numeric(),variable=character(),
                         SPECIFIC=character(),SPECIFIC2=character(),value=numeric())
save.folder = paste0(scenarios.basepath,scenario,"/",result.folder);
print(paste0("reading and writing to ",save.folder))

#.........................................................................................
......................
# writing files ----
# other_working_capital.inc ----
print("writing other_working_capital.inc....")
history.new = add.to.history(history.new,

read.wc(paste0(save.folder,"/other_working_capital.inc"))[,list(FID.f,    ROUND=cur.round,
SPECIFIC=NA, SPECIFIC2=NA, WC)])

# land_register.txt----
print("writing land_register....")
history.new = add.to.history(history.new,

read.land.register(save.folder)[,LAND.TOTAL.all:=RENTED.all+OWNED.all+FIXED.all][,-
c("FID")][,":="(ROUND=cur.round, SPECIFIC="land_register", SPECIFIC2=NA)]
                              )
```

```
file.copy(from=paste0(save.folder,"/land_register.txt"),to=paste0(save.folder,"/history/la
nd_register.",cur.round,".txt"))

# farm_accounting.txt ----
print("writing farm_accounting.txt....")
farm_accounting=read.farm.accounting(save.folder)[,-c("FID")]
history.new = add.to.history(history.new,

farm_accounting[,":="(SPECIFIC="farm_accounting",SPECIFIC2=NA, ROUND=cur.round)])
file.copy(from=paste0(save.folder,"/farm_accounting.txt"),to=paste0(save.folder,"/history/
farm_accounting.",cur.round,".txt"))
# rental_register.txt ----
#
data.table(FID=integer(),TYPE=character(),ROUND.expire=integer(),AREA=numeric(),PRICE=nume
ric())
#FID.f=character(),ROUND=numeric(),variable=character(),
#SPECIFIC=character(),SPECIFIC2=character(),value=numeric())

print("writing rental_register.txt....")
rental_register=read.rental.register(save.folder)
history.new = add.to.history(history.new,
                    rental_register[,list(
                      FID.f=paste0("f",FID),ROUND=cur.round,
                      SPECIFIC=TYPE,SPECIFIC2=ROUND.expire,
                      AREA.rented=AREA,RENT.price=PRICE)]
)
file.copy(from=paste0(save.folder,"/rental_register.txt"),to=paste0(save.folder,"/history/
rental_register.",cur.round,".txt"))
# land_abandoned.txt ----
# data.table(TYPE=character(),AREA=numeric())

print("writing land_abandoned.txt ....")
land_abandoned=read.land.abandoned(save.folder)
history.new = add.to.history(history.new,
                             land_abandoned[,list(
                               FID.f="ALL",ROUND=cur.round,
                               SPECIFIC=TYPE,SPECIFIC2=NA,
                               AREA.abandoned=AREA)]
)
file.copy(from=paste0(save.folder,"/rental_register.txt"),to=paste0(save.folder,"/history/
rental_register.",cur.round,".txt"))

# market_prices.inc ----
#//TODO
# feed_prices.inc ----
#//TODO
if(cur.round>0) {


# farm_plan.txt ----
  print("writing farm_plan.txt....")

farm.plan.data=data.table(read.table(file=paste0(save.folder,"/farm_plan.txt"),sep="\t",st
rip.white = T,header = F,col.names = c("FID.f","variable","SPECIFIC","value")))
  history.new = rbind(history.new,farm.plan.data[,":="(SPECIFIC2=NA, ROUND=cur.round)])


# farm_plan_outcome.txt ----
  print("writing farm_plan_outcome.txt....")
  farm.outcome.data=read.farm.plan.outcome(save.folder)
  history.new                                                                           =
add.to.history(history.new,farm.outcome.data[,":="(SPECIFIC="farm_plan_outcome",SPECIFIC2=
NA, ROUND=cur.round)])

file.copy(from=paste0(save.folder,"/farm_plan_outcome.txt"),to=paste0(save.folder,"/histor
y/farm_plan_outcome.",cur.round,".txt"))
```

```
# land.shadow.txt ----
  print("writing land.shadow.txt....")
  history.new = add.to.history(history.new,
                               read.land.shadow.prices(save.folder)[,list(FID.f,
ROUND=cur.round, SPECIFIC=TYPE, SPECIFIC2=EXTRA_LAND, LAND.SHADOW.PRICE=EXTRA_PROFIT)])



}

#Write file ----
print(".. saving history.txt")
cat(file=paste0(save.folder,"/history.txt"),append = T,

paste(history.new$FID.f,history.new$ROUND,history.new$variable,history.new$SPECIFIC,histor
y.new$SPECIFIC2,history.new$value,sep="\t"),
    sep="\n"
    )
```

update_accounts.R

```
# update_accounts.R
#
# DESCRIPTION ----
# ...............................................................................
# Update the accounts (availble working capital) of farms
#
#
# INPUT FILES----
# ...............................................................................
# - history.txt (history of variables)
# - farm_plan_outcome.txt (Economic results of the planned production for each farm)
# - land_register.txt
# - farm_accounting.txt
#
#
# OUTPUT FILES ----
# ...............................................................................
# - other_working_capital.inc (Available working capitalfrom previous year's results for
each farm)
# - history.txt (history of variables)
# - land_register.txt
# - farm_accounting.txt
#
#
# PARAMETERS ----
# ...............................................................................
# - args[1]: scenario name
# - args[2]: current round
# - args[3]: result folder name
#
#
#...........................................................................................
......................
# load libraries and parameters ----
#options(warn=-1)

library(data.table)
source("R_code/variables.R")

args = commandArgs(trailingOnly=TRUE)
#scenario = "flat.rate.2013"
scenario = args[1]
#round.cur=3
round.cur = as.numeric(args[2])
#result.folder="results"
result.folder=args[3]

scenarios.basepath = "gams/scenarios/"

save.folder = paste0(scenarios.basepath,scenario,"/",result.folder);
print(paste0("reading and writing to ",save.folder))
#...........................................................................................
......................
# load farm_plan_outcome.txt----
farm_plan_outcome=read.farm.plan.outcome(save.folder)

#...........................................................................................
......................
# load farm_accounting----
farm_accounting=read.farm.accounting(save.folder)

#...........................................................................................
......................
# load rental_register----
rental_register=read.rental.register(save.folder)
```

```
#...........................................................................................
......................
# Find Profit (GTP) and  ----
print("Calculating GTP ...")
farm_accounting = farm_accounting[,list(
  FID,FID.f,WC.avail,PROFIT.past=PROFIT.current, PROFIT.accum
)]

#Calculate PROFIT without rental expenditure
farm_accounting = merge(
  farm_accounting,
  farm_plan_outcome[,list(FID.f,PROFIT.current=             INC_lvst+INC_market_crop+SUBS-
COST_feed_prod-COST_feed_purch-COST_market_crop)],
  all.x=T,by="FID.f"
)
farm_accounting[is.na(farm_accounting)]=0

#substract from PROFIT the rental expenditure
rent.paid = rental_register[,list(RENT.paid=sum(AREA*PRICE)),by=FID]

farm_accounting = merge(
  farm_accounting,rent.paid,all.x=T,by="FID"
)
farm_accounting[is.na(farm_accounting)]=0
farm_accounting = farm_accounting[,list(
  FID,FID.f,WC.avail,PROFIT.past,PROFIT.current=PROFIT.current-RENT.paid,
PROFIT.accum=PROFIT.accum+PROFIT.current-RENT.paid
)]
if(round.cur>1) {
  print("R>1 ...")


#...........................................................................................
......................
  # (R>1) Find new working capital  ----

  print(" .... calculating new working capital ...")

  farm_accounting[,PROFIT.change:=round((PROFIT.current-PROFIT.past+1)/(PROFIT.past+1),2)]

  farm_accounting[,WC.avail:=round(WC.avail*(1+PROFIT.change),2)]
  farm_accounting[WC.avail<0,WC.avail:=0]
  farm_accounting[is.na(WC.avail),WC.avail:=0]
  farm_accounting[,PROFIT.change:=NULL]



#...........................................................................................
......................
  # (R>1)  Calculate abandoned land  ----
  print(".... calculating abandoned land")

  land_register=read.land.register(save.folder)
  print.land.register(land_register)

  fids.negative.profit=farm_accounting[PROFIT.accum<0,FID]
  print(paste0("......number of farms with PROFIT<0: ", length(fids.negative.profit)))

  if(length(fids.negative.profit)>0) {

    land_abandoned = read.land.abandoned(save.folder)

    print(paste0("fid negative: ",paste0(fids.negative.profit,collapse = " ")))

    #update land_Register
    abandoned.irr=land_register[FID%in%fids.negative.profit,sum(OWNED.irr+RENTED.irr)]
    abandoned.nirr=land_register[FID%in%fids.negative.profit,sum(OWNED.nirr+RENTED.nirr)]
```

```
    land_abandoned[TYPE=="IRR",AREA:=AREA+abandoned.irr]
    land_abandoned[TYPE=="nIRR",AREA:=AREA+abandoned.nirr]
    save.land.abandoned(save.folder,land_abandoned)


    land_register[FID%in%fids.negative.profit,
                  ":="(
                    OWNED.all=0.0,OWNED.irr=0.0,OWNED.nirr=0.0,
                    RENTED.all=0.0,RENTED.irr=0.0,RENTED.nirr=0.0
                  )]

    #update rental_Register (remove rental contracts for bankrupt farms)
    rental_register=rental_register[!FID%in%fids.negative.profit]
    save.rental.register(save.folder,rental_register)


    print("New Land status after considering farms with negative profit")
    print.land.register(land_register)
  }

  #write land abandoned
  save.land.register(save.folder,land_register)


  #remove bankrupt farms from sf.txt
  #save.sf.file(save.folder,land_register[!FID%in%fids.negative.profit,FID])


} else {

  print("round==1, doing nothing ...")
}


#.........................................................................................
......................
# Save farm_accounting ----
save.farm.accounting(save.folder,farm_accounting)
```

land_market.R

```
# land_market.R
#
# DESCRIPTION ----
# ................................................................................
# Clear the land market
#
# INPUT FILES----
# ................................................................................
# - land_register.txt
# - land.shadow.txt (land shadow prices for each farm)
# - farm_accounting.txt
#
#
# OUTPUT FILES ----
# ................................................................................
# - land_register.txt
# - rental_register.txt
#
#
# PARAMETERS ----
# ................................................................................
# - args[1]: scenario name
# - args[2]: current round
# - args[3]: result folder name
# - args[4]: debug status {TRUE, FALSE}
# - args[5]: lease_duration
#
#.........................................................................................
......................
# load libraries and parameters ----

#options(warn=-1)

library(data.table)
source("R_code/variables.R")

args = commandArgs(trailingOnly=TRUE)
#scenario = "flat.rate.2013"
scenario = args[1]

#round.cur = 13
round.cur = as.numeric(args[2]);

#result.folder="results"
result.folder=args[3]

#debug.status="TRUE"
debug.status=args[4]

#lease_duration=4
lease_duration=as.numeric(args[5])

scenarios.basepath = "gams/scenarios/"

save.folder = paste0(scenarios.basepath,scenario,"/",result.folder);
print(paste0("saving to ",save.folder))

#.........................................................................................
......................
# functions ----

#rent.granularity: granularity of land transactions in ha (the average size of plot)
clear.land.market = function(data.land.supply,
                             data.shadow.prices,
                             median.market.price,
                             rent.granularity = .25
```

```
) {

  market.results = data.table(FID=character(),AREA=numeric(),PRICE=numeric())

  if(data.land.supply<=0) {
    print("....... no land supply")
    return(market.results)
  }


  data.shadow.prices[,EXTRA_LAND:=1]

  if(nrow(data.shadow.prices)<=0) {
    print("....... no shadow prices")
    return(market.results)
  }

  market.fids = unique(data.shadow.prices$FID)

  round.count=round(data.land.supply/rent.granularity,0)
  print(paste0(".... land market will be cleared in ",round.count, " rounds"))

  round.fid.comb = data.table(t(replicate(round.count,sample(market.fids,3,replace = F))))

  data.random = runif(round.count)

  round.cur.count=0

  for(plot in 1:round.count) {

    round.cur.count=round.cur.count+1;
    if(round.cur.count>(round.count+1000)) {
      break
    }

    if( (plot%%100)==0) {
      print(paste0(" ....... reached round ", plot));
      #print(paste0("      .......       land      remaining      for      rental:      ",
data.shadow.prices[,sum(EXTRA_LAND)]))
    }

    #get 3 random FID
    round.fids = as.numeric(round.fid.comb[plot,])
    round.shadow                     =                   data.shadow.prices[FID%in%round.fids,
list(FID,EXTRA_LAND,EXTRA_PROFIT)]

    #find max EXTRA_PROFIT (shadow price) and min EXTRA_LAND
    round.shadow   =    round.shadow[,   .SD[EXTRA_PROFIT==max(EXTRA_PROFIT)],by=FID    ][,
.SD[which.min(EXTRA_LAND)],by=FID ]

    #the land is rented to the one with highest EXTRA_PROFIT for a random price between
EXTRA_PROFIT and median.market.price
    round.winner=round.shadow[EXTRA_PROFIT==max(EXTRA_PROFIT)]

    if(nrow(round.winner)>1) {
      round.winner=round.winner[1]
    }

    if(nrow(round.winner)==0) {
      cat("|");
      plot--
      next
    }

    #print(round.winner)
    market.results=rbind(market.results,
                         round.winner[,list(FID,AREA=rent.granularity,
```

```
PRICE=median.market.price+data.random[plot]*(EXTRA_PROFIT-median.market.price)*.5
                         )]
    )

    #update shadow prices so that the winner can rent less area
    data.shadow.prices[FID%in%round.winner$FID  &  EXTRA_LAND%in%round.winner$EXTRA_LAND  &
EXTRA_PROFIT%in%round.winner$EXTRA_PROFIT,
                       EXTRA_LAND:=EXTRA_LAND-rent.granularity]

    #remove shadow prices with EXTRA_LAND<0
    data.shadow.prices=data.shadow.prices[EXTRA_LAND>0]


  } #end for loop over plots

  print("")

  market.results[,FID:=as.numeric(FID)]

  return(market.results)

}
#.....................................................................................
......................
# finding land supply (land_rented.inc ) ----
land_register=read.land.register(save.folder)
farm_accounting=read.farm.accounting(save.folder)
rental_register=read.rental.register(save.folder)
land_abandoned=read.land.abandoned(save.folder)

# ... land supply from bankrupt farms (update accounts has already settled ABANDONED land-
---
print("..finding land supply from abandoned land")
print(paste0("Land abandoned: ",land_abandoned[,sum(AREA)]))
print(paste0("Land abandoned IRR: ",land_abandoned[TYPE=="IRR",sum(AREA)]))
print(paste0("Land abandoned nIRR: ",land_abandoned[TYPE=="nIRR",sum(AREA)]))
print("")
#update abandoned land
supply.land.IRR.total = land_abandoned[TYPE=="IRR",sum(AREA)]
supply.land.nIRR.total = land_abandoned[TYPE=="nIRR",sum(AREA)]
# ... land supply from expired rental contracts ----
print("..finding land supply from expiring rental contracts")
#irr
current.irr.rental.area=rental_register[ROUND.expire==round.cur                          &
TYPE=="IRR",list(RENTED.irr.expired=sum(AREA)),by=FID]
land_register=merge(land_register,current.irr.rental.area,all.x=T,by="FID")
land_register[is.na(land_register)]=0
land_register[,RENTED.irr:=round(RENTED.irr-RENTED.irr.expired,2)]
land_register[,RENTED.all:=round(RENTED.all-RENTED.irr.expired,2)]
land_register[,RENTED.irr.expired:=NULL]
rental_register=rental_register[! (ROUND.expire<=round.cur & TYPE=="IRR")]
supply.land.IRR.total=supply.land.IRR.total+current.irr.rental.area[,sum(RENTED.irr.expire
d)]
print(paste0("Land            IRR            from            expiring            rental:
",current.irr.rental.area[,sum(RENTED.irr.expired)]))

#nirr
current.nirr.rental.area=rental_register[ROUND.expire==round.cur                         &
TYPE=="nIRR",list(RENTED.nirr.expired=sum(AREA)),by=FID]
land_register=merge(land_register,current.nirr.rental.area,all.x=T,by="FID")
land_register[is.na(land_register)]=0
land_register[,RENTED.nirr:=round(RENTED.nirr-RENTED.nirr.expired,2)]
land_register[,RENTED.all:=round(RENTED.all-RENTED.nirr.expired,2)]
land_register[,RENTED.nirr.expired:=NULL]
rental_register=rental_register[! (ROUND.expire==round.cur & TYPE=="nIRR")]
supply.land.nIRR.total=supply.land.nIRR.total+current.nirr.rental.area[,sum(RENTED.nirr.ex
pired)]
```

```
print(paste0("Land nIRR from expiring rental: ",current.nirr.rental.area[,sum(RENTED.nirr.expired)]))

#find supply
print(paste0("supply.land.total: ", supply.land.IRR.total+supply.land.nIRR.total))
print(paste0("supply.land.nIRR.total: ", supply.land.nIRR.total))
print(paste0("supply.land.IRR.total: ", supply.land.IRR.total))
#read shadow prices and ranges (land.shadow.txt) ----
print("..read land shadow prices")
shadow.prices = read.land.shadow.prices(save.folder)



shadow.prices.irr = shadow.prices[TYPE=="IRR"]
if(debug.status=="TRUE") {
  print(".......IRR shadow prices:")
  print(shadow.prices.irr[,as.list(quantile(EXTRA_PROFIT)),by=EXTRA_LAND])
}

shadow.prices.nirr = shadow.prices[TYPE=="nIRR"]
if(debug.status=="TRUE") {
  print(".......nIRR shadow prices:")
  print(shadow.prices.nirr[,as.list(quantile(EXTRA_PROFIT)),by=EXTRA_LAND])
}

#do market ----

print("..clearing market for nIRR land")
#clear market for non-irrigted
if(nrow(shadow.prices.nirr[EXTRA_PROFIT>0])>0) {
  land.market.nIRR.results = clear.land.market(
    data.land.supply = supply.land.nIRR.total,
    data.shadow.prices= shadow.prices.nirr[
EXTRA_PROFIT>0,list(FID,EXTRA_LAND,EXTRA_PROFIT)],
    median.market.price = median(rental_register[TYPE=="nIRR",PRICE])
  )

} else {
  land.market.nIRR.results=data.table(FID=character(),
                                       AREA=numeric(),
                                       PRICE=numeric())
}

print("..clearing market for IRR land")
#clear market for irrigted
if(nrow(shadow.prices.irr[EXTRA_PROFIT>0])>0) {
  land.market.IRR.results = clear.land.market(
    data.land.supply = supply.land.IRR.total,
    data.shadow.prices= shadow.prices.irr[
EXTRA_PROFIT>0,list(FID,EXTRA_LAND,EXTRA_PROFIT)],
    median.market.price = median(rental_register[TYPE=="IRR",PRICE])
  )
} else {
  land.market.IRR.results=data.table(FID=character(),
                                      AREA=numeric(),
                                      PRICE=numeric())
}

print("..processing market clearing results")
land.market.results  = rbind(
  land.market.nIRR.results[,TYPE:="nIRR"],
  land.market.IRR.results[,TYPE:="IRR"]
)

# ...............................................................
#Update land_register ----

land.rented.total.IRR = land.market.results[TYPE=="IRR",sum(AREA)]
```

```
land.rented.total.nIRR = land.market.results[TYPE=="nIRR",sum(AREA)]

print(paste0(".... total land rented: ", land.rented.total.nIRR+land.rented.total.IRR))
print(paste0(".... total IRR land rented: ", land.rented.total.IRR))
print(paste0(".... total nIRR land rented: ", land.rented.total.nIRR))
#...update rental_register with newly rented land ----
#data.table(FID=integer(),TYPE=character(),ROUND.expire=integer(),AREA=numeric(),PRICE=num
eric())
if(nrow(land.market.results)>0) {

  print("..updating rental_register with newly rented land")

  rental_register=rbind(rental_register,
                        land.market.results[,

list(FID,TYPE,ROUND.expire=round.cur+lease_duration,
                                                 AREA,PRICE)
                                             ])

  #...update land_register  with newly rented land ----
  print("..updating land_register with newly rented land")

  #remove rented land
  land_register=land_register[,-c("RENTED.all","RENTED.irr","RENTED.nirr")]

  #update with rental_Register
  land_register=merge(
    land_register,
    rental_register[TYPE=="IRR",list(RENTED.irr=sum(AREA)),by=FID],
    all.x=T,by="FID"
  )

  land_register=merge(
    land_register,
    rental_register[TYPE=="nIRR",list(RENTED.nirr=sum(AREA)),by=FID],
    all.x=T,by="FID"
  )

  land_register[is.na(land_register)]=0;

  #put columns in correct order
  land_register=land_register[,list(
    FID,FID.f,
    OWNED.all, OWNED.irr, OWNED.nirr,
    RENTED.all=RENTED.irr+RENTED.nirr,RENTED.irr,RENTED.nirr,
    FIXED.all,FIXED.irr,FIXED.nirr
  )]


} else {
  print("land market had no transactions")
  land.status.new =land.status
}
#...still abandoned land ----
if(nrow(land.market.results)>0) {

  print("..updating land_register with still abandoned land")

  land.abandoned.still.IRR=max(0,round(supply.land.IRR.total-land.rented.total.IRR,2))
  land.abandoned.still.nIRR=max(0,round(supply.land.nIRR.total-land.rented.total.nIRR,2))
  print(paste0(".... total IRR land still abandoned: ", land.abandoned.still.IRR))
  print(paste0(".... total nIRR land still abandoned: ", land.abandoned.still.nIRR))

  land_abandoned[TYPE=="IRR",AREA:=land.abandoned.still.IRR]
  land_abandoned[TYPE=="nIRR",AREA:=land.abandoned.still.nIRR]

  save.land.abandoned(save.folder,land_abandoned)
```

```
}

print("Land status after clearing market")
print.land.register(land_register)
#write land_register----
save.land.register(save.folder,land_register)

#write rental_register----
save.rental.register(save.folder,rental_register)
```

Paper: CAP 2020 regionalization design: A Decision Support System

Forthcoming in CCIS Springer, "Communications in Computer and Information Science" series, Volume number CCIS 953

CAP 2020 regionalization design: A Decision Support System

Dimitris Kremmydas[1,†], Michael Malliapis[1], Leyteris Nellas[1], Apostolos Polymeros[3], Stelios Rozakis[2], Kostas Tsiboukas[1]

[1] Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Ieara Odos 75 118 55, {kremmydas,michael,enellas,tsiboukas}@aua.gr

[2]Department of Environmental Engineering, Technical University of Crete,

srozakis@isc.tuc.gr

[3] Directorate of Agricultural Extension, Greek Ministry of Agriculture,

apolymeros@minagric.gr

Abstract. The latest Common Agricultural Policy reform provides national authorities with several implementation options for fine tuning individual goals. Among other, member states can opt for regionalization, i.e. vary the basic payment unit value between national agronomic or administrative regions that have been defined at the beginning of the programming period. We present a Decision support System that support national authorities to implement regionalization in a transparent way facilitating collaboration with different shareholders.

Keywords: Common Agricultural Policy, Decision Support System; Basic Payment Scheme

1 Introduction

Common Agricultural Policy (CAP) is the agricultural policy of the European Union (EU), introduced in 1962 and fully implemented in 1968. It is considered to be the first real EU common policy replacing all relevant national agricultural policies while since then numerous reforms have been applied (table 1). For the last 20 years, CAP is absorbing more or less about 0,5%-0,6%[14] of the EU GDP and 50%-60%[15] of the EU budget annualy. Therefore CAP evaluation is a persisting issue in the Agricultural Economics field.

Table 7, EU-CAP reform milestones

Year	Short Description
1979	Overproduction problems. Measures are put in place to align production with market needs. Introduction of market quotas and expenditure ceiling.
1992	“McSharry reform”. The CAP shifts from market support to producer support. Introduction of direct aid payments, “set-aside” payments, measures to encourage retirement.
1999	“Agenda 2000”.Introduction of two Pillars, production support and rural development. Agri-environment schemes became compulsory.
2003	“Midterm CAP reform”. The link between subsidies and production is cut. Introduction of “Decoupled

[14] “CAP expenditure and CAP reform path”, accessed from http://ec.europa.eu/agriculture/cap-post-2013/graphs/graph2_en.pdf

[15] “CAP expenditure in the total EU expenditure”, accessed from http://ec.europa.eu/agriculture/cap-post-2013/graphs/graph1_en.pdf

	payments" and "Cross-compliance". "Multifunctionality of agriculture" notion
2006	Sugar regime reform.
2008	"Health Check". Enforcement of the 2003 reform.
2013	"2014-2020 CAP reform". Introduction of "national envelopes" for members states, i.e. flexibility in the budgeting and implementation of first pillar measurements. Introduction of "Basic Payment Scheme", "Green Payment", "Young Farmers Scheme" and "Redistributive Payment". Gradual abolition of "Historical model".
Compiled from: a. "The Common Agricultural Policy: A story to be continued", European Commission, accessed from http://ec.europa.eu/agriculture/50-years-of-cap/files/history/history_book_lr_en.pdf b. Pezaros (2000) c. "Overview of CAP Reform 2014-2020", European Commission, accessed from http://ec.europa.eu/agriculture/policy-perspectives/policy-briefs/05_en.pdf	

The new CAP design, acknowledging the wide diversity of agronomic production potential and climatic, environmental as well as socio-economic conditions and needs across the EU, offers implementation flexibility to member states. Indicatively, member states may: differentiate the basic payment per hectare according to administrative or agronomic criteria; choose from different options for internal convergence for payments per hectare until 2019; opt in for the right to use a redistributive payment for the first hectares; enable the "small farm scheme", where small farms receive an annual subsidy of 500€ - 1250€ with minimal administrative burden; preserve a limited amount for coupled payments; grant an additional payment for areas with natural constraints (as defined under Rural Development rules)[16].

The latest Common Agricultural Policy reform (CAP2020) provides national authorities with several implementation options for fine tuning individual goals. 30% of the national CAP funding is connected to the farmers' compliance to a predefined set of pro-environmental practices. Up to 5% can be devoted to farms of areas with natural constraints, up to 13% to coupled payments, up to 10% to small farmers' scheme, up to 2% to new farmers' scheme and up to 3% to the national rights stock. The rest, called basic payment scheme (BPS) is the main layer of income support (over 50% of the national budget), based on payment entitlements activated on eligible land and decoupled from production.

Within this scheme, among other options, Member States (MS) can opt to apply BPS in finer scale than the national level, termed hereafter as *regionalization*

In the Direct Payments regulation (1307/2013), Article 23(1) notes

> *Member States may decide, by 1 August 2014, to apply the basic payment scheme at regional level. In such cases, they shall define the regions in accordance with objective and non-*

[16] Compiled from European Commission MEMO, "CAP Reform – an explanation of the main elements", accessed from http://europa.eu/rapid/press-release_MEMO-13-621_en.htm

discriminatory criteria such as their agronomic and socio- economic characteristics, their regional agricultural potential, or their institutional or administrative structure.

Thus, MS can differentiate the unit value of the basic payment (BP) on the basis of national, agronomic or administrative regions that have been defined at the beginning of the programming period. Policy assigned regionalization regions (RR) can coincide with administrative or geographic regions but can also be not related to them, such as the case of agronomic criteria where a region is defined on the basis of specific crop areas (e.g. arable or permanent crops). Hence, regionalization regions may represent a broader category than administrative or geographic regions and shall not be confused with them.

The only regulation guideline regarding regionalization is that it should be in accordance with objective and non-discriminatory criteria. Practically MS are totally free to draw the regions and allocate the BPS budget.

This flexibility provides to policy makers numerous different alternatives on how to form regions and allocate the budget. Additionally the fact that different stakeholders are affected in a distinct way, call for a transparent design process. Towards this end we propose a Decision support System (DSS) that will support national authorities to implement regionalization in a transparent way facilitating collaboration with different shareholders. In this paper we present its design overview and give a proof-of-concept implementation.

In section 2, we provide details on the mathematical representation of modeling regionalization; in section 3 we give a small review of the use of decision support systems in agricultural policy evaluation; in sections 4 and 5 we present the design and the implementation of the employed regionalization DSS.

2 Modeling CAP2020 Regionalization

There are three regionalization types, based on how regionalization regions (RR) are defined.

- RRs are administrative-based partitions (e.g. prefecture-based) or socio-economic related partitions (e.g. mountainous vs. non-mountainous areas). The distinctive feature in this case is that each farm is related with only one RR. The farm's basic payment unit value (BPUV) equals to the RR basic payment unit value that the farm belongs to (Eq. 1).
- RRs are agronomic based partitions (e.g. Arable vs. Tree crops). In this case farms can be related to more than one RR, e.g. half of farm area is connected to arable RR and the other half to tree RR. The farm's BPUV equals the average of each agronomic region (agronomic=crop) basic payment unit value weighted by the share of each crop area to total farm area in a reference year, as in Eq. 2.
- RR definition is a hybrid case of the previous two cases. For example when the RRs are mountainous vs. non-mountainous arable crops vs. non-mountainous permanent crops. Then the farm's BPUV is like the second case but the agronomic basic payment unit value can differ from one farm to another, as in Eq. 3.

administrative-based regionalization	$BP_f^F = BP_{r(f)}^R \quad \forall f$	$BP_r^R = BP_r^{BUDGET} \Big/ \sum_{f(r)} TL_f^E$	(1)
agronomic-based regionalization	$BP_f^F = \frac{\sum_g \sum_{c(g)} \left(BP_g^R \cdot X_{f,c}\right)}{TL_f^E}$	$BP_g^R = BP_g^{BUDGET} \Big/ \sum_{c(g)} \sum_f X_{f,c}$	(2)
Combined regionalization	$BP_f^F = \frac{\sum_g \sum_{c(g)} \left(BP_{g,r}^R \cdot X_{f,c}\right)}{TL_f^E}$	$BP_{g,r}^R = BP_{g,r}^{BUDGET} \Big/ \sum_{c(g)} \sum_{f(r)} X_{f,c}$	(3)

where

BP_f^F: Basic payment unit value applicable to farm-f (euro/ha)

BP_r^R: Basic payment unit value applicable to administrative region-r, where $r(f)$ is the region of farm-f) (euro/ha)

BP_r^{BUDGET}: The Basic Payment budget for region-r, where $f(r)$ is the set of farms that belong to region-r (euro)

$BP_{g,r}^R$: Basic payment unit value applicable to agronomic region-g under administrative region-g, where $c(g)$ is the crop-set related to g (euro/ha)

$BP_{g,r}^{BUDGET}$: The Basic Payment budget for agronomic region-g under administrative region-g,(euro)

TL_f^E: Total eligible land for farm-f (ha)

$X_{f,c}^B$: Area of crop-c in farm-f in the reference period (ha)

For an illustrative example regarding those three regionalization types, see Kremmydas et al. (2018).

Therefore the policy-makers options regarding regionalization can be decomposed to the following sequential decisions:

a) the regionalization type, i.e. administrative, agronomic or hybrid
b) the allocation of farms and/or crops to the corresponding RRs (defining $f(r)$ and $c(g)$ sets)
c) the allocation of the total budget to the defined RRs (defining BP_r^{BUDGET}, BP_g^{BUDGET}, $BP_{g,r}^{BUDGET}$)

The DSS addresses those three phases, as described in section 4.

3 Decision Support Systems and Agricultural Policy Evaluation

Decision Support System (DSS) is any kind of a computer program facilitating decision making process. It is an umbrella term that covers any computerized system that supports decision making in an organization. DSS enhances the capability of decision makers to take more accurate and on-time decisions. It has been acknowledged that decisions utilizing DSS can be made more quickly and accurately than unaided decisions (Djamasbi and Loiacono, 2008; Todd and Benbasat, 2000; Chan et al., 2017).

A properly designed DSS is an interactive software-based system intended to help decision makers compile useful information from a combination of raw data, documents, and personal knowledge, or business models to identify and solve problems and make decisions (Sprague, 1980).

According to Power and Sharda (2007), there are five categories of DSS which can be recognized by identifying the dominant architectural component that provides the functionality for supporting decision-making (Power, 2002). The five categories include model-driven DSS, as well as communications-driven, data-driven, document driven and knowledge-driven DSS. The architecture is most often comprised of three fundamental components: the database or knowledge base, the model (i.e., the decision context and user criteria) and the user interface. The database or knowledge base holds the data used by the model to derive its conclusions. The user interface is the way through which the user interacts with the DSS to provide the necessary inputs and pick the results.

Agricultural policy needs strategic decisions and requires DSS to evaluate and understand the outcome of each alternative for optimal decision-making. So, the domain is a privileged area for the technology of DSS. There are a lot of DSS covering several aspects of this area and some distinguished papers are mentioned below.

Manos et al. (2010) present a DSS for sustainable development and environmental protection of agricultural regions. The system aims at the optimization of the production plan of an agricultural region taking in account the available resources, the environmental parameters, and the vulnerability map of the region. In their paper, Borges et al. (2010) demonstrate the use of a model base approach to anticipate the impacts of changes in CAP and/or in prices on land use in rural areas (including forest land). In Louhichi et al. (2010) is presented a bio-economic farm model for different bio-physical and socio-economic contexts, facilitating the linking of micro and macro analysis. Model use is illustrated with an analysis of the impacts of the CAP reform of 2003 for arable and livestock farms in a context of market liberalization. Bournaris and Papathanasiou (2012), present a DSS for the planning of agricultural production in agricultural holdings or in agricultural areas. The system simulates different scenarios and policies and proposes alternative production plans. Finally, a paper of Rovaia et al. (2016) presents a comprehensive model for the governance of rural landscape and a first simplified application to a cultural landscape.

4 The CAP2020 Regionalization Decision Support System

National authorities have a great flexibility on how to draw regionalization regions and allocate budget. Consequently they can potentially pursue a wide range of objectives. This means that the required data in order to evaluate the objectives can only approximately be determined during the development of the DSS and very probably new data will be requested during the consultation with other stakeholders.

The DSS knowledge base currently contains data from the Greek Payments Authority on previous CAP regime; the current direct payment allocation per farm size and prefecture. However the database can easily be extended to contain other socio-economic data like the income indicators per farm size and farm activity from the national FADN database; the regional GDP per sector from the national statistical authorities; etc.

The DSS models the effects of the policy makers decisions regarding the three regionalization options (type, region definition, budget allocation) to the Single Farm Payment value for each farm, as described in section 2. Output is provided in the basis of farm grouping of farm economic size and type of farming. Spatial output is also present, providing a visual representation of the effects n each prefecture. In any case the model can be extended to provide output for other measures that represent individual policy goals.

Overall, given an established strategic goal, the DSS provides a clear picture of how that goal is affected for any selected scenarios. The DSS usage is expected to be in an iterative mode: policy makers and stakeholders draw regions, try some budget allocations and observe the effects and then restart the process to fine tune policy results.

We distinguish two DSS use cases that correspond to the regionalization design decisions that are described in section 2 and another one that extends the DSS with collaboration features.

4.1 Select regionalization type and define regions

Policy makers form a regionalization scenario, i.e. select regionalization type and define regions, by means of exploratory analysis. The definition of regions is based on some partition variables, e.g. the NUTS nomenclature, the altitude or some crop classification like arable vs. permanent crops. Thus the user selects the partition variables which identify the different regions. The user very probably will further consolidate those regions to more homogeneous ones. In order to do so, he will examine certain regions’ property variables, e.g. the prevailed crop pattern, the importance of agriculture, the current single payment unit value, etc. He can thus refine initial region creation either manually or through a clustering tool that will suggest him the regions that are as homogeneous as possible. The activity diagram of this use case is provided in Figure 1.

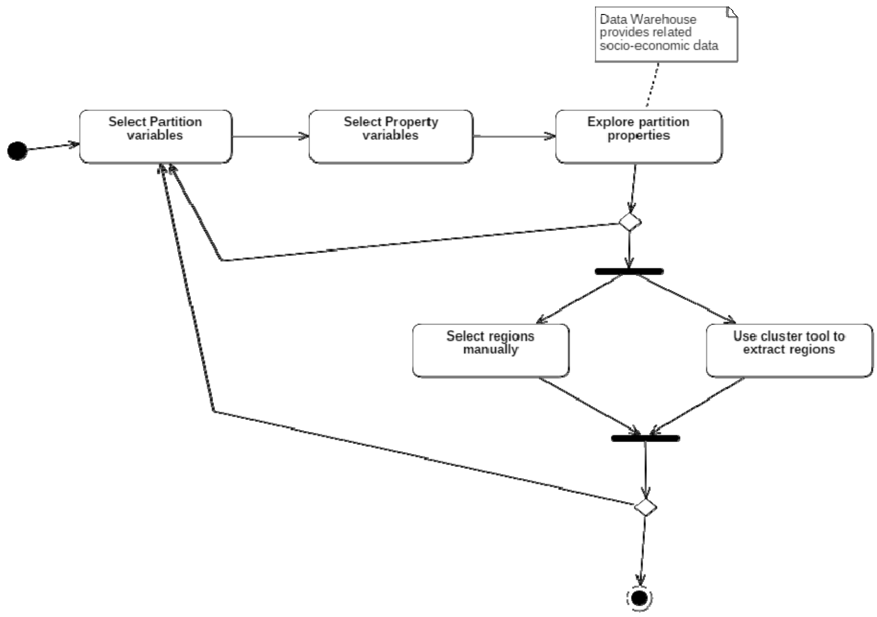


Fig. 1, Activity Diagram for exploratory extraction of regions

4.2 Budget allocation across regionalization regions

When the user has concluded to a region formation scenario he is ready to set budget allocation. This is expected to be a trial and error exercise where policy makers investigate the effects across different stakeholders for different allocations. Users can manually set the budget share or can use tools of predefined allocations, e.g. budget share proportional to the number of entitlements or to the gross value of direct payments in each region. Then the DSS engine will calculate the indicators and present them to the user. Based on the results the user can save the regionalization scenario and restart the process. The indicators of the scenario effect will span to different stakeholder classes, e.g. farms per NUTS administrative level or per type of farming or per farm income class. The activity diagram of this use case is provided in Figure 2.

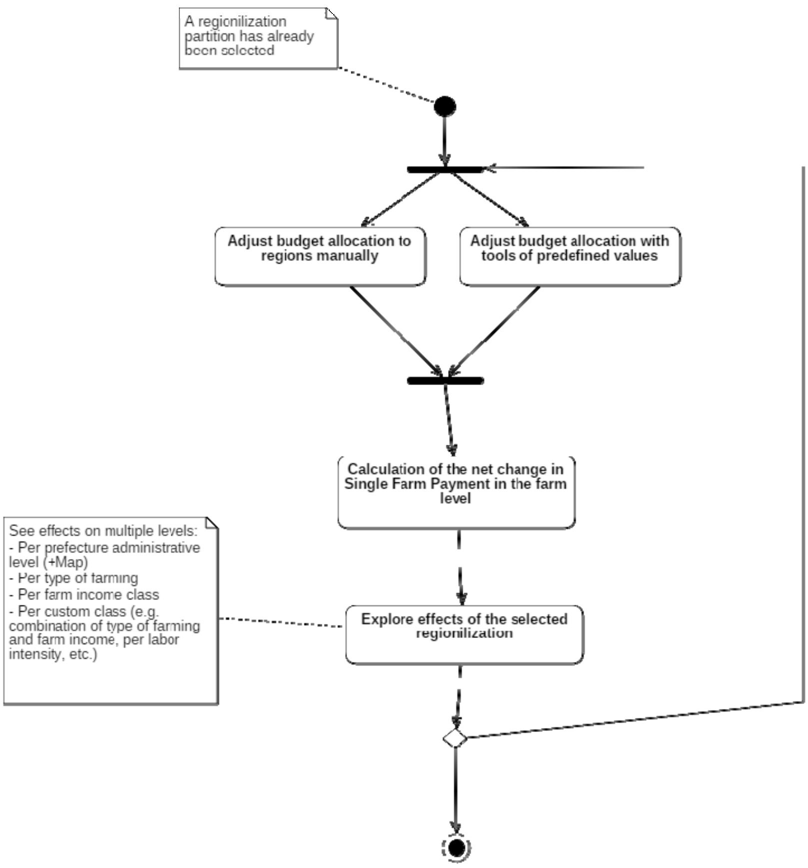


Fig. 2, Activity Diagram for exploring the effects of various budget allocations between regions

4.3 Dissemination and collaboration

Dissemination and collaboration use case: Due to the different interests of stakeholders, collaboration is a very important aspect of the regionalization decision process and thus is incorporated into the DSS. When a user is satisfied with a scenario (regionalization region definition + budget allocation) he can save it and choose to share it, either with other users of the system or in public. A discussion channel, e.g. a forum thread, will be automatically created so that other users can comment and discuss scenarios. Users will also be able to load the scenarios of other users in order to adjust them to their point of view.

5 A swift exhibition of the system

We used the R-Shiny web application framework (Chang et al., 2017) for agile development.

For the region formation stage, we used the following partition variables: NUTS-3, Altitude, Less-Favored-Area, current regionalization regions, municipalities.

For deciding on the region homogeneity we provided the following property variables: number of farms, sum of utilized agricultural area, mean Single Farm Payment unit value, sum of Single Farm Payment value.

In Figures 3, one can see that the user has selected to partition regionalization regions (RR) by Prefectures (Fig. 3a) and to examine the homogeneity of the formulated RRs using the variables of *number of farms, sum of utilized agricultural rea, sum of SFP value* and *mean unit* value.

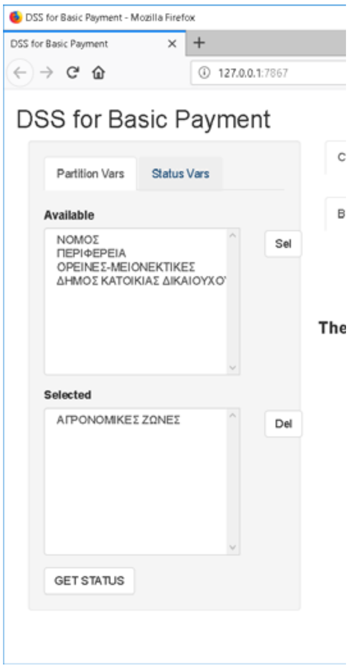


Fig. 3a, Partition variables

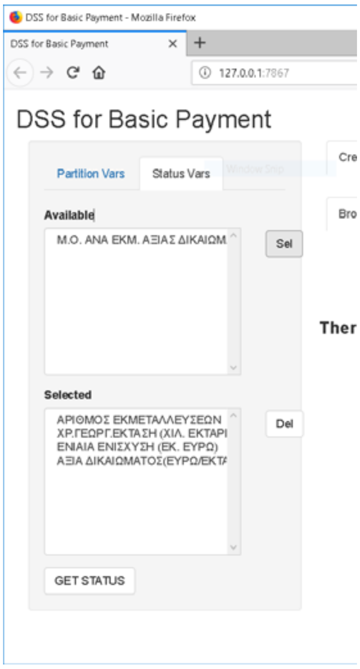
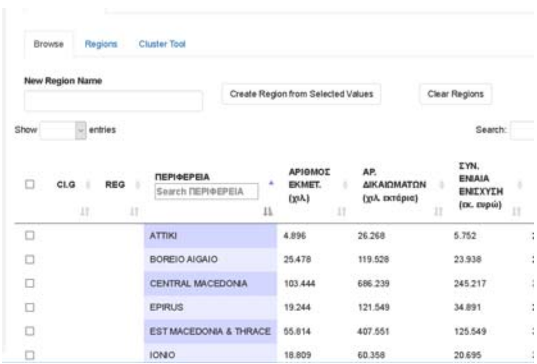


Fig. 3b, homogeneity status variables

In Figure 4a we show the status of the DSS when the user has clicked the STATUS button. The values of the status variables are shown grouped by the partition variables. The user can sort prefectures based on status variables and/or filter values. Furthermore on the clusters tab the user can perform a hierarchical or a k-means clustering in order effectively see the homogeneous RRs. In Figure 4b, the user has selected to perform k-means cluster, eliciting 3 regions based on the number of farms contained in an RR and the mean SFP unit value. The DSS outputs the mean values of the clusters and also adds to Cl.G column the number of the cluster that each RR belongs to (see Fig. 4c).



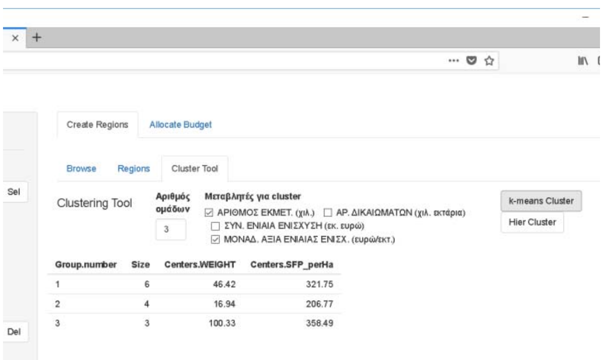


Fig. 4a, Values of status variables grouped by

partition variables (STATUS button) **Fig. 4b, Clustering tool**

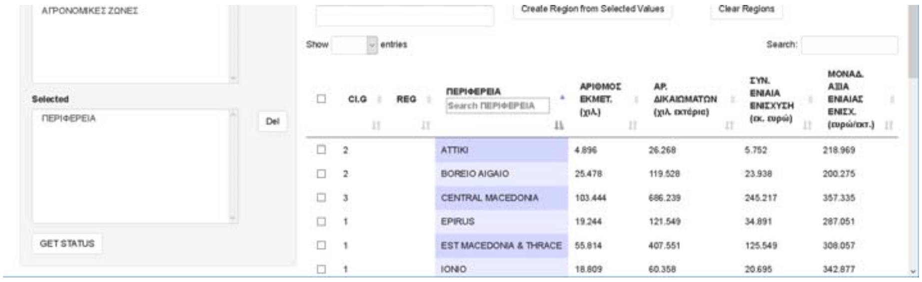


Fig. 4c, Transfer of cluster analysis results

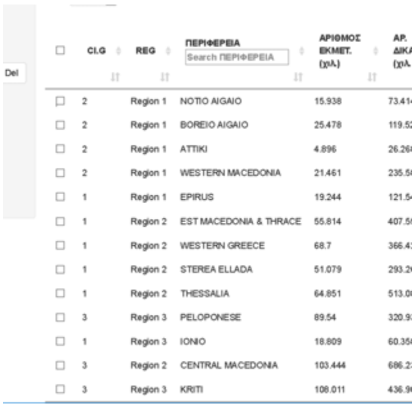
After the user has a clear picture of the homogeneity of the potential RRs, he can further proceed on creating the real RRs. In Figure 5a, the user has created three regions (Region 1, 2 and 3) after sorting by the CI.G column (the cluster number the region belongs to). Those RRs are a furter grouping of the potential RRs. For instance Region 1 is composed of *NOTIO AGAIO, VOREIO AGAIO, ATTIKI, WESTERN MACEDONIA, EPIRUS*. The user can explore many further groupings since the DSS provides instantly their properties regarding the status variables. This is shown in Figure 5b.

When the user has concluded to a certain RR formation he proceeds to the Budget allocation phase, as shown in Figure 5c. There, apart from defining the total SFP budget (in this case 1159 mil. Euro), he allocates it to the different regions. It will be also be possible to define Convergence and Redistributive Payment scheme parameters; however this is still under development.

When the user click the *See Effects* button the DSS model is run and all relevant results are given back. They are given grouped by NUTS-3, by Prefecture, by economic size and by type of farming, so that the user can explore the effects on several dimensions. Also there is a *Complete* option that provides a detailed grouping of all above dimensions. For instance, the user can explore the effects on KRITI prefecture for Arable type of farming for economic size of *2000 – 8000* economic size group.

Furthermore, for each of the above dimensions, information is also presented by means of maps. In Figure 7a the net effect (mil. Euro) of the current Regionalization scenario for each NUTS-3 region is presented in a map. Users can also see this map for a certain type of farming or for a certain economic size.

Finally the distribution of the SFP unit value is given in a table and in a chart. In Figure 7b, the user sees the distribution of the SFP unit value in the current situation (blue line) vs. the user created scenario (red line). Those distributions are also provided for the economic size and type of farming dimensions.



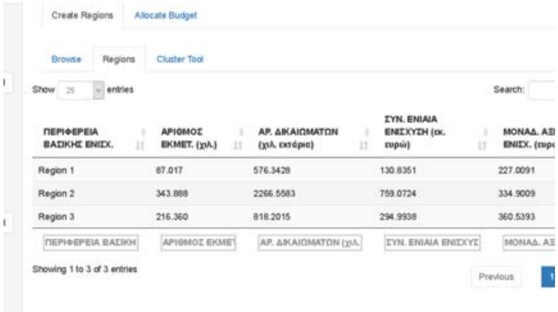


Fig. 5b, Status variables for the created RRs

Fig. 5a, Three regions are created

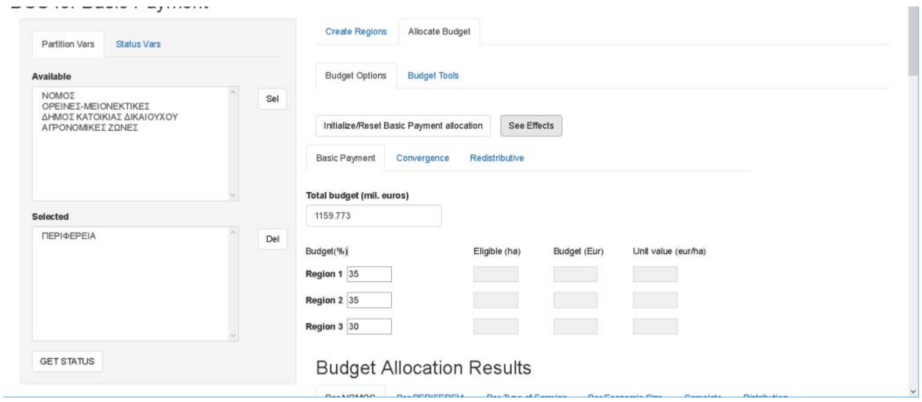


Fig. 5c, Budget allocation panel

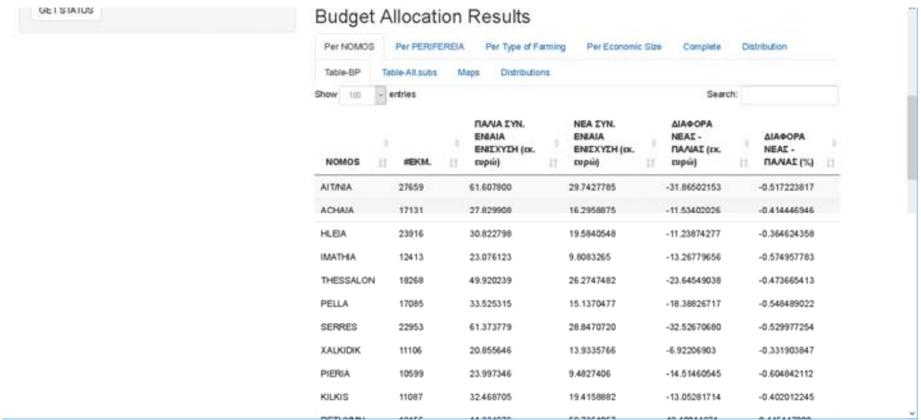


Fig 6, Budget allocation results

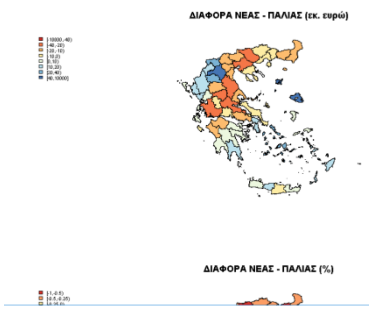


Fig. 7a, Three regions are created

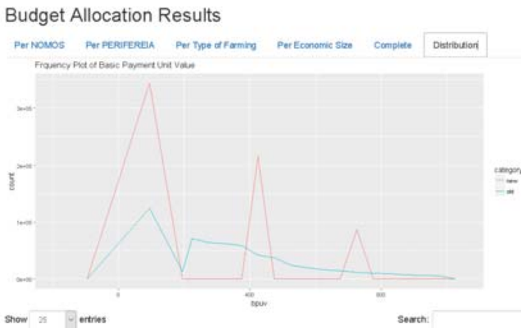


Fig. 7b, Status variables for the created RRs

6 Conclusions

The latest Common Agricultural Policy reform (CAP2020) provides national authorities with several implementation options for fine tuning individual goals. This creates the need for a transparent policy design process. Towards this end we presented a DSS that support regionalization design in a transparent way facilitating collaboration with different shareholders, in three distinct steps: selection of regionalization type and definition of regions; budget allocation across regionalization regions; dissemination and collaboration between stakeholders.

In the future we plan to extend the DSS database with socio-economic data. Furthermore a mathematical programming farm model will be incorporated so that the adaptation of farms to selected policy scenarios can be evaluated. Finally a pilot implementation with selected stakeholders is also planned.

References

1. Borges, P.J., Fragoso, R., Garcia-Gonzalo, J., Borges, J.G., Marques, S., Lucas, M.R. (2010) Assessing impacts of Common Agricultural Policy changes on regional land use patterns with a decision support system: An application in Southern Portugal, Forest Policy and Economics 12 (2010) 111–120
2. Bournaris T. and Papathanasiou, J. (2012) A DSS for planning the agricultural production, Int. J. Business Innovation and Research, Vol. 6, No. 1, DOI: 10.1504/IJBIR.2012.044259
3. Chan, S.H., Songb, Q., Sarkerc, S., and Plumleed, R. D. (2017) Decision support system (DSS) use and decision performance: DSS motivation and its antecedents, Information & Management 54 934–947.
4. Chang, W., Cheng, J., Allaire, J. J., Xie, Y., & McPherson, J. (2017). shiny: Web Application Framework for R. Retrieved from https://cran.r-project.org/package=shin
5. Djamasbi, S. and Loiacono, E. (2008) Do men and women use feedback provided by their Decision Support Systems (DSS) differently? Decision Support Systems, 44, 854–869.
6. European Parliament, C. of the E. U. (2013). EU 1307/2013: Regulation (EU) No 1307/2013 of the European Parliament and of the Council of 17 december 2013 establishing rules for direct payments to farmers under support schemes within the framework of the common agricultural policy and repealing

Counc. Official Journal of the European Union, 2008(347), 608–670. https://doi.org/10.1017/CBO9781107415324.004

7. Henke, R., D'Andrea, M. P., & Benos, T. (2015). Implementation of the first pillar of the CAP 2014 – 2020 in the EU Member States, http://doi.org/10.2861/662407
8. Kremmydas D., Kampas A., Rozakis S., Tsiboukas K., Vlahos G. (2018) CAP regionalization scheme in Greece: A rapid policy assessment method, 162nd EAAE Seminar, April 26-27,2018, Budapest, Hungary
9. Louhichi, K., Kanellopoulos, A., Guillermo Flichman, S. J., Blanco, M., Hengsdijk, M., Heckelei, T., Berentsen, P., Lansink, A. O., Van Ittersum, M. (2010) FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies, Agricultural Systems 103 (2010) 585–597
10. Manos, B.D., Papathanasiou, J., Bournaris T. and Voudouris, K. (2010) A DSS for sustainable development and environmental protection of agricultural regions, Environ Monit Assess 164:43–52, DOI 10.1007/s10661-009-0873-1
11. Pezaros, P. "The C.A.P. in the Pliers of the Multilateral Trading System: Origins, Evolution and Future Challenges" in the S. Bilal – P. Pezaros (Ed.) (2000): Negotiating the Future of Agricultural Policies", Kluwer Law International, The Hague, pp. 51 -80
12. Power, D.J. (2002) Decision Support Systems: Concepts and Resources for Managers, Greenwood/Quorum Books, Westport, CT.
13. Power, D.J. and Sharda, R. (2007) Model-driven decision support systems: Concepts and research directions, Decision Support Systems 43, 1044–1061
14. Rovaia, M., Andreolia, M., Gorellib, S., Jussila, H. (2016) A DSS model for the governance of sustainable rural landscape: A first application to the cultural landscape of Orcia Valley (Tuscany, Italy), Land Use Policy 56 (2016) 217–237
15. Sprague, R (1980). A Framework for the Development of Decision Support Systems. MIS Quarterly. Vol. 4, No. 4, pp.1-25

APPENDIX

G.RE.FA.M manual

Greek Representative Farm Model

G.RE.FA.M

for Agricultural Policy Analysis

Reference manual

version 0.3

Athens

November 2018

CONTRIBUTORS

	Contribution
Dimitris Kremmydas Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Iera Odos 75, 118 55, Athens, Greece kremmydas@aua.gr	Coordination, Documentation, Processing and preparation of FADN raw data, Mathematical Programming formulation, Livestock model
Stelios Rozakis Institute of Soil Science and Plant Cultivation, Pulaw, Poland rozakis@aua.gr	Mathematical Programming formulation
Konstantinos Tsiboukas Department of Agricultural Economics & Rural Development, Agricultural University of Athens, Iera Odos 75, 118 55, Athens, Greece tsiboukas@aua.gr	Consultancy on Livestock model

FUTURE WORK

Title	Description
Handle *Olive groves* activity better	Currently Olive groves (FADN code = 154) is entered as a single activity under the GAMs code olivefr However, we can further split the activity to “Olive grove for olive oil” and “Olive oil for Table Olives” and “Olive grove for mixed production”. We can do this using PRODUCT 281 and 282 where the Production of each type of product is recorded
Run model with data on three consecutive years	Currently the production choices of farms are limited among the observed production choices of the calibration year. If based on three consecutive years the set of production choices for each farm is extended
Calculation of CO2 and Nitrogen emissions based on FAO methodology	
Include Risk on the objective function	

CHANGELOG

v 0.1

Copy Model specification from an early paper on regionalization

Abstraction of the modeling process so that data and model are distinguished

GRIFAM logo added

v 0.2

minor

Added a TODO List / GRIFAM renamed to GREFAM / Housekeeping of Activities

major

Added Livestock submodel

v 0.3

minor

major

Variable Cost Estimation per yield, with quantiles

Table of Contents

Introduction

The scope of the GREFAM model is to use the Greek FADN representative farm data for policy impact assessment. Since the model has a narrow focus it includes in depth details tailored to Greek agriculture.

The purpose of this manual is to give details on the model specification and the input data computation, so as to make it transparent to end users.

Conceptually the model consists of various definition or data files (static components) and of various processes (dynamic components), as shown in Figure 13. The model specification describes the mathematical programming problem (objective function and constraints) and the corresponding data needed in a GAMS file. This file is combined with the Model Input data (in txt or gdx files) and the model is solved so as one or more Result files are produced. In order to create the Model Input Data, several Raw FADN data (through the fadnUtils R package) and other auxiliary raw data are processed.

In the following sections we describe in more detail those model components.

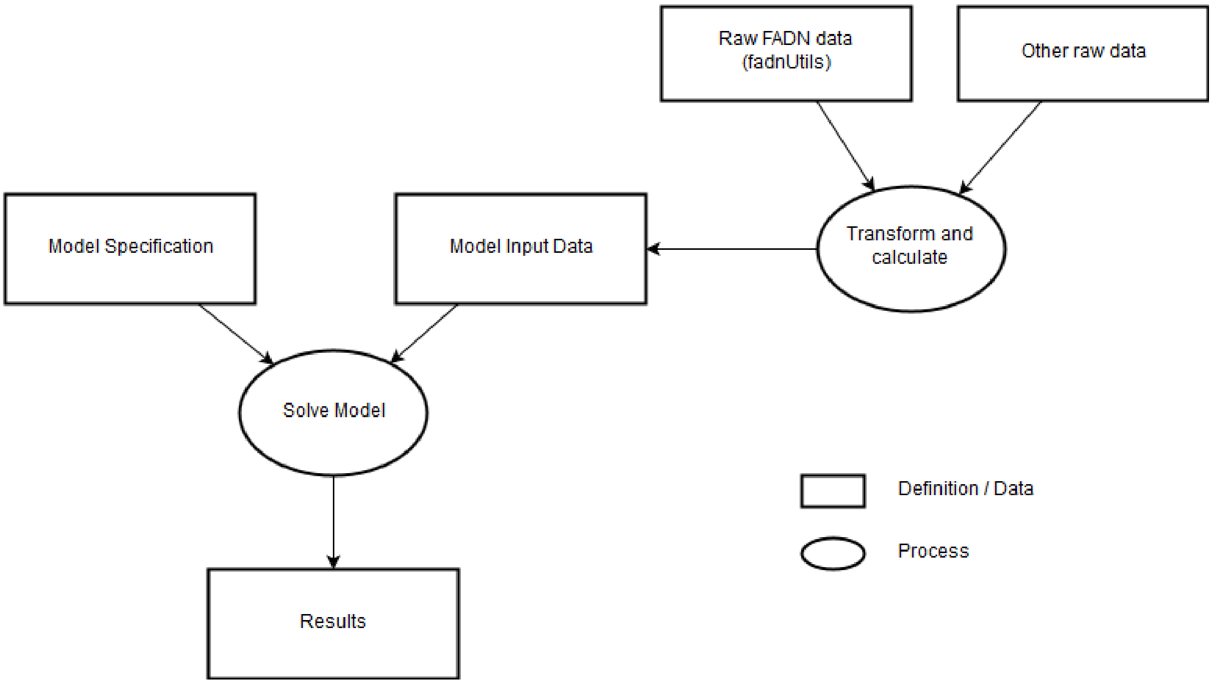


Figure 13, Model Components

Model Specification

Important Sets

Activities

In the model we consider a broad number of agricultural activities. In Table 8 we provide the detailed list along with the corresponding FADN code. The *short name* column provides the GAMS set name of the activity.

Activities are divided to Crop and Livestock activities. For Crop activities, we differentiate between irrigated and non-irrigated activities. For example maize for grain activity is split into *mzegrn* (non irrigated) and *mzegrn.irr.cmb* (irrigated) activities. Thus objective function coefficients (gross margin) are calculated separately in those cases.

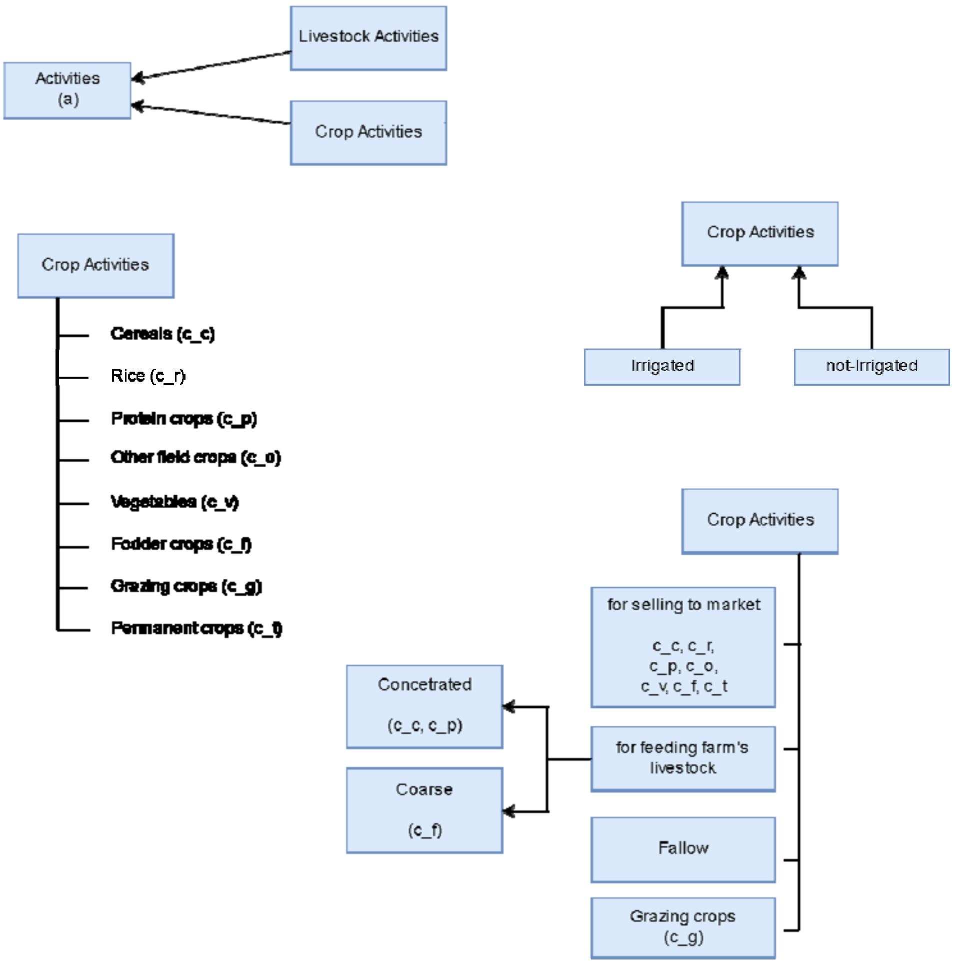


Figure 14, Hierarchy of the Activities Set (the GAMS set name in parentheses)

Table 8, Model Agricultural Activities

	Activity	FADN code	Short Name (GAMS code)			Comments
			Non-irrigated	Irrigated / Combined	Other	
Cereals	Common wheat and spelt	120	cwheat	-		
	Durum wheat	121	dwheat	-		
	Rye (including meslin)	122	rye	-		
	Barley	123	barley	barley.irr.cmb		
	Oats	124	oats	-		
	Grain maize	126	mzegrn	mzegrn.irr.cmb		
	Rice	127	rice.irr.cmb			
	Other cereals	128	ceroth			Includes also millet, triticale, buckwheat and sorghum
	Summer cereal Mixes	125	sumCereal			
Other Field Crops	Cotton	347	cotton	cotton.irr.cmb		
	Potatoes	130		potat.irr.cmb		Includes early and seed potatoes.
	Sunflower	332	sunflr	sunflr.irr.cmb	sunflr.energ	
	Tobacco - Basmas	370	tob.basmas	tob.basmas.irr.cmb		
	Tobacco - Katerini	371		tob.kater.irr.cmb		
	Peas		peas	peas.irr.cmb		
	Lentils		lentils	lentils.irr.cmb		
	Protein Crops		protein	protein.oth.cmb		
Vegetables	Tomatoes	337	veg.tom	veg.tom.irr.cmb	veg.tom.shelter	
	Leeks, spinach, lettuce	336		veg.leeks.irr.cmb	veg.leeks.shelter,	

					veg.leeks.garden	
	Vegetables grown for fruit or flowers (not tomatoes)	338	veg.flowers	veg.fruits.irr.cmb		Marrows, courgettes, aubergines, gherkins, globe artichokes, sweet peppers
	Vegetables grown for roots, bulbs or tubers (except potatoes)	339		veg.roots.irr.cmb		
	Legume vegetables	340		veg.legumes.irr.cmb		Includes peas and beans. Excludes lentils, chick peas
Fodder and Fallow	Fodder maize	326		mze.fod.irr.cmb		
	Other fodder plants	328	fod.oth	fod.oth.irr.cmb		Lucerne and other fodder plants
	Rough grazing	151	grazing.rgh			Generally uncultivated and not fertilized land, including scrub, used as poor quality pasture.
	Permanent pasture	150	pasture.perm			Grassland grown for 5 years or more on cultivated land.
	Temporary grass	147	grass.temp			Grassland grown for less than 5 years on arable land. Includes areas grown for less than one year and the production of hay and/or silage from these areas (
	Fallows and set aside	146	fallow			Includes all arable land included in the crop rotation system, whether worked or not, but with no intention to produce a harvest for the duration of a crop year.
Permanent Crops	Olive groves	154	olivefr	olivefr.irr		Currently there is no distinction as whether production is olive oil or table olives

	Vines	155	vinesfr	vinesfr.irr		Currently there is no distinction as whether production is grapes or wine
	Nuts	351	nutsfr	nutsfr.irr		Includes walnuts, hazelnuts, almonds, chestnuts
	Stone fruit	350	stonefr	stonefr.irr		Includes plums, peaches, apricots, cherries.
	Oranges	354		orangefr.irr		
	Lemons	356		lemonfr.irr		
	Tangerines, mandarines, clementines	355		tangerfr.irr		
	Small fruit and berries	352	berriesfr			Includes red currants, black currants, white currants, gooseberries, raspberries, figs. Excludes strawberries, melons and pineapples (to be recorded in Vegetables)
	Pome fruit	349	pomefr	pomefr.irr		Includes apples, pears, and quinces
	Tropical and sub-tropical fruit	353		tropicalfr.irr		Includes bananas, avocados, mangoes, papayas

Notes:

For Vegetables

- *.shelter refers to Crops grown under shelter (greenhouses, permanent frames, accessible plastic tunnels) during the whole or greater part of the growing season.
- *.garden refers to Crops grown under short rotation with other horticultural crops, with almost continuous occupation of the land and

several harvests per year.

- By default (if no shelter or garden suffix is present), we refer to Crops grown in rotation with field scale crops.

For Permanent Pasture and Rough grazing, The value of hay and/or grass used as feedingstuffs for livestock can be indicated when marketable under 'Farm use'

Variables

Variables	Description
X_{cm}^{market}	The area of crop-c, for selling to market (hectare)
X_{cf}^{feed}	The area of feedstock crop-c^f, for feeding animals in the farm (hectare)
X^{fallow}	The fallow land (hectare)
$X_{ta}^{animals}$	The number of animals of type-ta the farm decides to breed (heads)
$F_{c^f\_purch}$	Quantity of feedstock of type c^{f_purch} purchased (tones)
L^{family}	The family labour used (hours)
$L^{foreign}$	The foreign labour used (hours)

Objective Function

We assume that farm selects a crop plan so as to maximize their gross income (1.1) subject to certain constraints as in (1.2)-(1.11). We now provide more details on the farm model formulation.

The gross income (1.1) equals to non-labor gross margin (Price times Yield plus Coupled payments minus variable costs) minus the foreign labor expenditure (Wage times Foreign labor requirements, i.e. the selected crop plan labor requirements minus the available family labor) plus Single payment plus any Pillar II payments.

$$TGP = GP^{market} + GP^{lvst} - WAGES + SUBS \tag{1.1}$$

Total farm Gross Profit equals Gross Profit (without wages paid) from crops sold at the market plus the Gross Profit (without wages paid) from the livestock activities minus wages paid for all activities plus subsidies received by the farm

$$GP^{market} = \sum_{c^m} \left(\left(PR_{c^m}^{market} \cdot YI_{c^m} - VC_{c^m} \right) \cdot X_{c^m}^{market} \right) \tag{1.1.1}$$

Gross Profits from market crops equal,
for each crop, the gross margin per hectare (price times yield
minus variable costs plus coupled payments).

$$GP^{lvst} = INC^{lvst} - FEED^{prod} - FEED^{purch} - EXP^{lvst} \quad (1.1.2)$$

Gross Profits from livestock activities equals the income from livestock activities minus the cost of in-farm produced feedstock (excluding wages) minus the cost of purchased feedstock minus any non-feedstock costs

$$INC^{lvst} = M_{a\_l}^{quant} \cdot M_{a\_l}^{price} + INC^{SELL} \quad (1.1.2.1)$$

Income from livestock equals the income from milk plus that from selling young animals

$$M_{a\_l}^{quant} = X_{ta^{a\_l}}^{animals} \cdot M_{ta^{a\_l}}^{YIELD} \cdot M_{a\_l}^{price} \quad (1.1.2.2)$$

The quantity of milk of type a_l (sheep or goat milk) equals the number of animals of type ta that produce milk a_l (the subset ta^{a_l}) times their yield times the price of milk a_l

$$INC^{SELL} = \sum_{ta} \left(X_{ta}^{animals} \cdot S_{ta}^{share} \cdot S_{ta}^{price} \right) \quad (1.1.2.3)$$

The income from selling young animals equals the sum of the number of animals of type ta times the observed percentage of young sold animals times their observed price

$$FEED^{prod} = \sum_{cf} \left(VC_{cf} \cdot X_{cf}^{feed} \right) \quad (1.1.2.4)$$

Produced feedstock cost equal, for each feedstock crop, the variable costs (excluding wages) per hectare minus any coupled payment per hectare times the area cultivated

$$FEED^{purch} = \sum_{fs\_purch} \left(PR_{fs\_purch}^{feed} \cdot F_{fs\_purch} \right) \quad (1.1.2.5)$$

Purchased feedstock cost equals the price of the feedstock (euro per tn) times the quantity purchased

$$WAGES = L^{foreign} \cdot WAGE \quad (1.1.3)$$

The wages paid are equal to the required foreign labor (in hours) multiplied by the wage. The required foreign labor equals the labor required for crop (market and feedstock) cultivation plus the required labor for breeding any animals minus the available family labor.

$$LAB^{crops} = \sum_{c^m} \left(LB_{c^m}^{CR} \cdot X_{c^m}^{market} \right) + \sum_{c^f} \left(LB_{c^f}^{CR} \cdot X_{c^f}^{feed} \right) \tag{1.1.3.1}$$

The labor required for crops equals the labor required for one hectare of a market crop (LB_c^R) times the cultivated area plus the labor required for one hectare of a feedstock crop times the corresponding areas

$$LAB^{animl} = \sum_{ta} \left(LB_{ta}^{AN} \cdot ANIM_{ta} \right) + \sum_{l\_a} \left(LB_{l\_a}^{ML} \cdot M_{a\_l}^{quant} \right) \tag{1.1.3.2}$$

The labor required for animal breeding equals the sum over all animals types of the labor per animal times the number of animals in this system

$$SUBS = SP \cdot LE + SO + \sum_{c^m} \left(CP_c \cdot X_{c^m}^{market} \right) + \sum_{c^f} \left(CP_{c^f} \cdot X_{c^f}^{feed} \right) \tag{1.1.4}$$

The value of subsidies equal the unit value of the decoupled payment (Single Payment of Single Farm Payment) times the eligible land plus the value of pillar II subsidies plus the value of coupled payments

Total farm Gross Income equals Gross Profit (without wages paid) from crops sold at the market plus the Gross Profit (without wages paid) from the livestock activities minus wages paid for all activities plus subsidies received by the farm

$$GI = GP^{market} + GP^{lvst} - WAGES + SUBS$$

$$SUBS = SP \cdot LA^{E} + SO$$

The value of subsidies equal the unit value of the decoupled payment (Single Payment of Single Farm Payment) times the eligible land plus the value of pillar II subsidies

$$WAGES = (LAB^{crops} + LAB^{animl} - FLB) \cdot WAGE$$

The wages paid are equal to the required foreign labor (in hours) multiplied by the wage. The required foreign labor equals the labor required for crop (market and feedstock) cultivation plus the required labor for breeding any animals minus the available family labor.

$$LAB^{crops} = \sum_{c} \left(LB_{c}^{CR} \cdot X_{c}^{market}\right) + \sum_{cf} \left(LB_{cf}^{CR} \cdot X_{cf}^{feed}\right)$$

The labor required for crops equals the labor required for one hectare of a market crop (LB_c^R) times the cultivated area plus the labor required for one hectare of a feedstock crop times the corresponding areas

$$LAB^{animl} = \sum_{lsys} \left(LB_{lsys}^{AN} \cdot A_{lsys}\right)$$

The labor required for animal breeding equals the sum over all livestock breeding systems of the labor required for one ewe (for this system) times the number of animals in this system

Gross Profits from market crops equal, for each crop, the gross margin per hectare (price x yield minus variable costs plus coupled payments).

$$GP^{market} = \sum_{c} \left((PR_c \cdot YI_c - VC_c + CP_c) \cdot X_c^{market}\right)$$

Gross Profits from livestock activities equals the income from livestock activities minus the cost of in-farm produced feedstock minus the cost of purchased feedstock

$$GP^{lvst} = INC^{lvst} - FEED^{prod} - FEED^{purch}$$

$$INC^{lvst} = \sum_{lsys} \left(M^{yield} \cdot M^{price} \cdot A_{lsys}\right)$$

Income from livestock equals, for each livestock system (intensive, semi-intensive, extensive), the number of ewe times the milk yield times the milk price

$$FEED^{prod} = \sum_{cf} \left((VC_{cf} - CP_{cf}) \cdot X_{cf}^{feed}\right)$$

Produced feedstock cost equal, for each feedstock crop, the variable costs per hectare minus any coupled payment per hectare times the area cultivated

$$FEED^{purch} = \sum_{cf\_crs\_purch} \left(PR_{cf\_crs\_purch}^{feed} \cdot F_{cf\_crs\_purch}^{crs}\right) + \sum_{cf\_cntr\_purch} \left(PR_{cf\_cntr\_purch}^{feed} \cdot F_{cf\_cntr\_purch}^{cntr}\right)$$

Purchased feedstock cost equal, for each coarse and concentrated feedstock crop that can be purchased ($c^{f\_cntr\_purch}$, $c^{f\_crs\_purch}$), the price of the feedstock (euro per tn) times the quantity purchased

Figure 15, Objective function graphical decomposition

Constraints

Land and Labor

In (1.2) total cropping area cannot exceed total land, in (1.3) the irrigated area is bounded by the current irrigated area and in (1.4) a labor availability constraint is set.

$$\sum_{cm}\left(X_{cm}^{market}\right) + \sum_{cf}\left(X_{cf}^{feed}\right) \leq LA^{T} \tag{1.2}$$

The total cropping area cannot exceed total land

$$\sum_{cm}\left(X_{cm\_i}^{market}\right) + \sum_{cf}\left(X_{cf\_i}^{feed}\right) \leq LA^{I} \tag{1.3}$$

The total irrigated cropping area cannot exceed total irrigated land

$$LAB^{crops} + LAB^{animl} \leq L^{family} + L^{foreign} \leq TOTAL\_LABOUR \tag{1.4}$$

The required labor shall be equal or more than the available family labor and less than the total available labor

Working Capital

In (1.5) a working capital constraint is set for each farm. The required working capital (left hand side) equals to the non-labor variable costs plus any foreign labor expenditure. This cannot exceed the sum of subsidies plus a farm *excess working capital*.

$$\sum_{c}\left(VC_c \cdot X_c^{market}\right) + FEED^{prod} + FEED^{purch} + WAGES \leq SUBS + WC^{E} \tag{1.5}$$

Livestock Constraints

$$SUPPLY_n^{CF} + SUPPLY_n^{FS\_PURCH} \geq NEED_n^{MAINT} + NEED_n^{MILK} \qquad \forall n \tag{1.6}$$

For each nutrient type, the nutrient supplied by purchased and produced feedstock shall cover the needs for maintaining the weight of the farm animals (kg)

$$SUPPLY_n^{CF} = \sum_{cf} \left(CONT_{n,cf}^{C\_F} \cdot YI_{cf} \cdot X_{cf}^{feed} \right) \tag{1.6.1}$$

The nutrient supplied by produced equals the area of those crops times the yield times the content per ton of that crops

$$SUPPLY_n^{FS\_PURCH} = \sum_{fs\_purch} \left(CONT_{"DM",fs_{purch}}^{FS\_PURCH} \cdot F_{cf\_purch} \right) \tag{1.6.2}$$

The nutrient supplied by purchased feedstock equals the quantity purchased times the content per ton of that feedstock (kg)

$$NEED_n^{MAINT} = \sum_{ta} \left(NEED_n^{AN} \cdot X_{ta}^{animals} \right) \tag{1.6.3}$$

The maintenance demand for a nutrient equal for each animal type the per-animal maintenance needs times the number of animals of that type

$$NEED_n^{MILK} = \sum_{a\_l} \left(NEED_{a\_l}^{LT} \cdot M_{a\_l}^{quant} \right) \tag{1.6.4}$$

The milk production demand for a nutrient equal for each livestock product the needs of nutrient per unit of product (liter) times the quantity of production

$$SUPPLY_{"DM"}^{CF} + SUPPLY_{"DM"}^{FS\_PURCH} \geq 0.01 \cdot \sum_{ta} \left(WEIGHT_{ta} \cdot X_{ta}^{animals} \right) \tag{1.7}$$

The Dry Matter Supply (DM) shall exceed the 1% of the total weight of all animals

$$X_{cf\_g}^{feed} = LA_{cf\_g}^{O} \quad \forall c^{f\_g} \tag{1.8}$$

The area devoted to grazing feedstock is fixed and equal to the observed one

Decision space

We confine the farm decision space to specific crops. In this version of the model we confine the available decision space crops to those observed in the baseline year. In future version of the model, the calibration data will be based on three consecutive years, thus the decision base will be more realistic.

$$X_{c^{IMP}} = 0 \qquad \forall \ c^{IMP} \tag{1.9}$$

Sets:: c^{IMP}: The crops that are impossible for farm to cultivate (subset of c)

Permanent crops

In (1.10) we fix the permanent crop area to be equal to the observed ones, i.e. farms cannot expand or contract permanent crops. We assume this since starting a permanent crop activity is an investment decision since they have a production life of 30-50 years and thus bind farm capital in the long term. However in the model we consider income and variable costs connected to permanent crops activity, since they affect the short-term status of a farm. For a similar reason we choose to exclude depreciation from our considerations. Depreciation affects farm in the long run. If the net income (gross income minus depreciation) is negative, it is a sign that the farm is not viable in the long run within its current structure. The long-term viability is nevertheless an interesting and important aspect of examining policy impacts but in this paper we are occupied with the short-term production effects.

$$X_{c^T} = LA^O_{c^T} \qquad \forall c^T \tag{1.10}$$

Sets:: c^T: The permanent crops (subset of c)

Parameters:: $LA^O_{c^T}$: Land observed for crop-c (ha)

Crop rotations

In (1.11) we set crop rotations. The rotation coefficient ($ROT^{coef}_{c(rot),rot}$) denotes how many hectares of a crop participate in the rotation.

$$\sum_{c(rot)} X_{c(rot)} \cdot ROT^{coef}_{c(rot),rot} \geq 0 \qquad \forall rot \tag{1.11}$$

Sets:: rot: the rotation set / $c(rot)$: the crops that are included in a rotation

Parameters:: $ROT^{coef}_{c(rot),rot}$: The rotation coefficient

More specifically we set the rotation constraints as in Table 9.

Table 9, Rotation constraints

dwheat + cwheat + mzegrnIR + barley + fallow	≥	0.25	cottonIR
dwheat + cwheat + barley + fallow	≥	0.25	cotton
cwheat + barley + fallow + cotton	≥	0.25	dwheat

dwheat + cwheat + barley + fallow + cotton	≥	0.25	mzegrnIR
barley	≥	0.25	cwheat

Thus, the first rotation constraint implies the sum of dwheat, cwheat, mzegrnIR,barley,fallow areas shall be greater that 0.25 of the cottonIR area; in other words for every one hectare of cottonIR there shall be at least

Flexibility constraints

In order to remedy the fact that a linear programming problem is an approximation of the farm decision process, we include certain flexibility constraints:

In (1.12) we define some lower and upper limits for certain contract crops. Contract farming is treated with flexibility constraints in the model. Farms that grow one of the contract crops are imposed to decide between 80% and 120% of the current area. All other farms cannot select those contract crops.

$$LOW_{c^{cc}} \cdot LA^{O}_{c^{cc}} \leq X_{c^{cc}} \leq 1.2 \cdot UPP_{c^{cc}} \tag{1.12}$$

Sets:: c^{CC}: contract crops (sugar beet, sunflower) /

Parameters:: $LOW_{c^{cc}}, UPP_{c^{cc}}$: The lower/upper allowable share of the main crop in the total farm area (%)

CAP 2020 Greening constraints

Next we augmented the baseline model with greening constraints, as in (2.4)-(2.7).

Crop diversification 1, 95%

$$X_{f^{30},C_1} + X_{f^{30},C_2} \leq 0.95 \cdot TL^{E}_{f^{30}} \quad \forall f_{30}, c_1 \in c, c_2 \in \{c - c_1\} \tag{2.5}$$

Crop diversification 2, 75%

$$X_{f^{10},c} \leq 0.75 \cdot TL^{E}_{f^{10}} \quad \forall f_{10}, c \tag{2.6}$$

Ecologic Focus Area

$$0.7 \cdot \sum_{c_l} \left(X_{f^{15},c^L}\right) + X_{f^{15},\text{"fallow"}} \geq 0.05 \cdot TL^{E}_{f^{15}} \quad \forall f_{15} \tag{2.7}$$

where f^{10} , f^{15}, f^{30}: the set of farms with more than 10,15 and 30 ha of arable land respectively

Parameters

We provide in details the data used in GREFAM model in Table 10.

Table 10, Parameters (data) used in GREFAM model

GAMS name	Description
General information on farm	
w(f)	Weight of farm-f (number of represented farms)
SPc(f)	Single Farm Payment in baseline/historical regime (euro/ha)
So(f)	Other subsidies, i.e. agri-environmental, NATURA, mountain, etc. (euro)
Gross Margins	
yields(c,f)	Yields for a crop on a farm (tn/ha)
prices(c,f)	Expected prices for a crop on a farm (euro/tn)
totalVarCost(a,f)	Variable cost requirements (intermediate consumption) for a crop on a farm, excluding wages (euro/ha)
CP(a)	Coupled payment for activity a (euro/ha)
Land	
Lir(f)	Irrigated Land owned by farm (ha)
Le(f)	Land eligible for single payments, i.e. the rights in 2013 before the CAP2014 (ha)
Xc(a,f)	Area of activity-a observed for farm-f (ha)
farm_use(c,f)	Area cultivated with crops for feeding in-farm animals (farm use) (ha)
IMP_X(c,f)	Impossible combinations for farm f of crop c (0=impossible, >0 possible)
Labor	
lb_cr(f,c)	Labor requirements for a crop activity (hours/ha)
famL(f)	Family Labor availability (hours)
forL(f)	Foreign Labor availability (hours)
total_labor_obs(f)	Total Labor observed in the farm, according to estimtion of labor per hour

GAMS name	Description
	(hours)"
wage(f)	Wage (euro/hour human labor)
Livestock	
lvst_purch_feeds_obs(f)	Observed value of purchased feedstock for livestock
anim(f,ta)	Number of animals per type-of-animals (ta) (number)
M_quant(f,a_l)	The milk quantity observed produced for each livestock activity in one year (kgr)
M_price(f,a_l)	The milk price for each livestock activity (euro/lt)
INC_sell(f)	income from animal sales (euro)
LB_an(ta)	Labor requirements per animal type (hours/animal)
LB_ml(a_l)	Labor requirements per kg of livestock activity (milk) (hours/kg)
WEIGHT_an(ta)	The average weight (over 1 year) for each animal type (kg)
CNT_fs_purch(fs_purch,n)	The content of the purchased feedstocks in nutrients (dm pr and fib in g, enrg in MJ)
PR_fs_purch(fs_purch)	The price of purchasable global feedstocks including transportation costs (euro/tn)
CNT_c_f(c,n)	The nutrient content of the crops that can be used for feedstock (dm pr and fib in g, enrg in MJ)
NEED_an(ta,n_con)	The need of one animal for maintainance purposes for the whole year (dm pr and fib in gr, enrg in MJ)
NEED_lt(a_l,n_con)	The need of nutrient for producing 1 kg of product-a_l (sheep or goat milk) (dm pr and fib in gr, enrg in MJ)
lvst_purch_feeds_obs_coarse(f)	The observed value of purchased coarse feedstock (euro)
lvst_purch_feeds_obs_conc(f)	The observed value of concetrated concetrated feedstock (euro)
lvst_prod_feeds_obs(f)	The observed value of produced feedstock (euro)
lvst_specific_costs_obs(f)	The observed specific livestock costs (euro)
lvst_costs_paid_obs(f)	Observed value of paid costs for livestock (euro)
M_yield(f,ta)	milk yield per animal for farm-f (tn/animal)
animal_to_milk(ta,a_l)	Type of animal that produces type of milk (0-1 data)

GAMS name	Description
INC_sell_per_animal(f)	Income from selling animals per animal for farm f (euro/animal)
anim_sheep_ratio	The ratio of other to female sheeps (ratio)
anim_goat_ratio	The ratio of other to female goats (ratio)
out_grazing_days(f)	The outside-farm LU grazing days reported
out_grazing_cont(n)	The content of one kgr of outside-farm grazing matter (dm pr and fib in g, enrg in MJ)
out_grazing_yield	The yield of one day of grazing (tn/day)
out_grazing_day_conv	The conversion of 1 LU-grazing day to sheep/goat grazing day (number)
cnt_fs_purch(fs_purch,n)	The content of the purchased feedstocks in nutrients (dm pr and fib in g, enrg in MJ
SETS: f: farms a: activities (crop+livestock) c: crop activities a_l: livestock activities ta: type of animal fs_purch: feedstocks that are purchased n: nutrient categories (protein, energy, etc.)	

Model Input Data

Data Coverage

We are currently calibrating data to a single year, namely 2013. This is planned to be changed to a three year calibration. This will be more accurate since more data will be available for estimating the farm production plan.

We do not use the whole farm dataset but exclude certain farms that may affect the credibility of the results. More specifically we exclude:

1. Farms that they have at least one crop activity that is not in the included activities (see Table 8). This results in 1054 sample farms excluded from the original 4779 sample farms
2. We also exclude any farms that present at least one yield outlier (based on the IQR method) in any farm activity. Another 146 sample farms were removed.

The data coverage after the above filtering is

	Original FADN	Filtered dataset	%
Farms	334.713	230.787	69.0%
Area (th. Ha)	3075.8	2048.7	66.6%
Crop Output Value (mil €)	4787.6	3231.4	67.5%
Livestock Output Value (mil €)	1313.5	672.3	51.2%

Furthermore, in the filtered dataset we run certain data transformations like calculating the production in grazing land; distinguishing irrigation crops and eliciting the areas allocated for in-farm livestock feeding. More details on those can be found in the *Preparation of GREFAM Input Data* supplementary material.

Gross Margin Calculation

In GREFAM we define Gross Margin (GM) as GM=PRICE·YIELD-VARCOST, where VARCOST is the variable cost per hectare, without including labour costs.

Activities' gross margin is the most important data element in a farm model. The relative gross margins between activities determine the optimum crop mix and thus are the main drivers of the model outcome. Furthermore it is important to maintain the farms gross margin heterogeneity found in FADN data, rather than employing homogeneous gross margins from econometric estimation.

However deriving individual farm gross margins for each activity from FADN data is not a straightforward exercise. There are two primary problems

1. Prices and Yields, although reported on per farm ad per activity basis, may contain outliers and thus a data filtering process must be applied
2. Variable costs are reported for the whole farm and not for individual activities. Thus we need to derive variable costs for each activity in each farm

The steps we take to calculate individual farm gross margins are shown in Figure 16.

Initially we isolate from FADN dataset of years 2011, 2012 and 2013 the farms with a single activity crop. Also any farm with livestock activity is removed. Based on this monocrop farm subset, we detect and remove any farm with price or yield outliers using the Inter-Quantile-Range method. Then we proceed with calculating variable costs reference values needed for the goal programming variable costs allocation goal programming model. Reference data consists of minimum, median and maximum values for each variable cost category (fertilizers, water, etc.) for each type of activity (common wheat, maize, etc.). The solution of the goal programming problem provides estimates of the variable costs per hectare for each farm and each activity. Those estimates are combined with the filtered yields and prices to obtain gross margins for each activity in each farm.

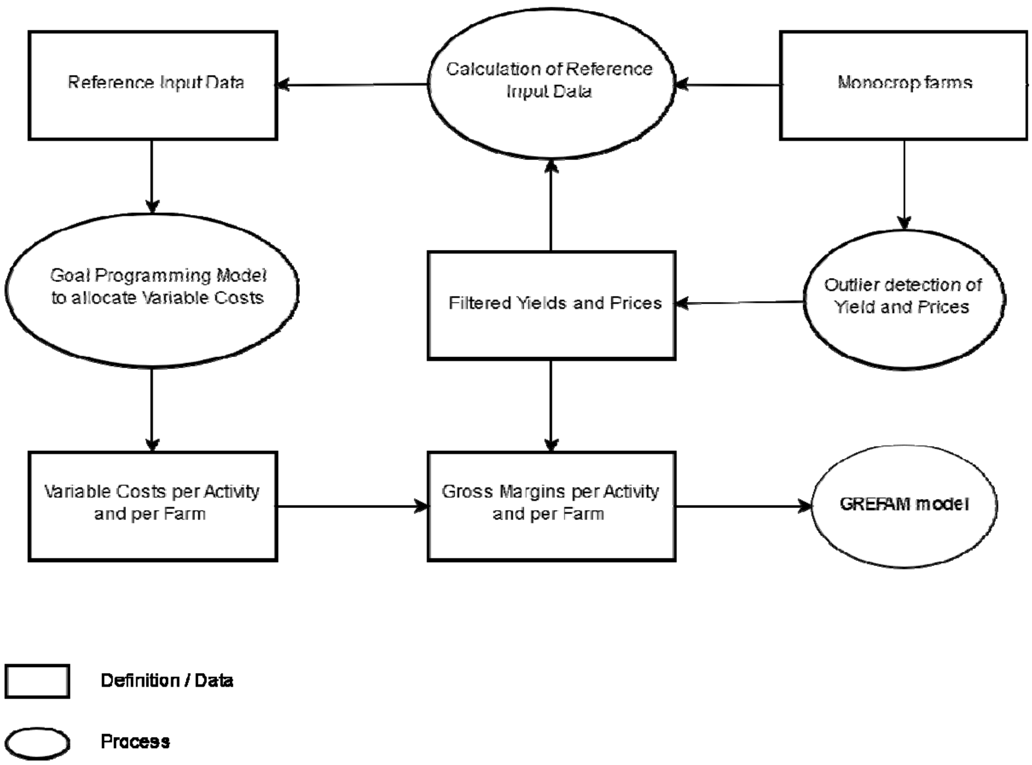


Figure 16, Overview of the Gross Margin Calculation

On the next subsections we provide more information on the various elements of the gross margin estimation process.

FADN Variable cost categories

We calculate the individual variable cost categories that are recorded in the FADN dataset. Those are:

- **SPECIAL CROP COSTS** that includes costs that we assume to be strongly dependent on crop yield. It is the sum of the following table-F FADN variables:
  - Water (V281)
  - Fertilizers and soil improvers (V274)
  - Crop protection products (V275)
  - Contract Work (V260).
  - Motor fuels and lubricants (V262)
  - Electricity (V279)
  - Heating Fuels (V280)
  - Wages (V259)
- **SPECIAL LIVESTOCK COSTS** is the sum of the following table-F FADN variables:
  - Purchased Concentrated feeding stuffs for grazing stock (V264)
  - Purchased Coarse fodder for grazing stock (V265)
  - Produced Feeding stuffs for grazing stock (V268)
- **OTHER COSTS**, that include more overhead costs, that are not so dependent on the farm's targeted yield. This is the sum of following FADN table F variables:
  - Car expenses (V263).
  - Seeds and seedlings purchased (V272)
  - Seeds and seedlings produced and used on the farm (V273)
  - Other specific crop costs (V276). According to FADN definitions, they are more generic costs, e.g. packaging, supplies for the preservation and processing, short-term buildings rent, etc.
  - Other specific livestock costs (V271)

Eliciting single-activity farms

We use the Greek FADN 2012, 2013 and 2014 datasets. The notation and the variable definition is based on the RI/CC 1256 rev. 7 EU document (February 2011) that describes those datasets in detail. Also for each year, we remove livestock farms (Any farm that had at least one animal of any kind). Thus the rest analysis refers to only-crop-farms.

We then calculate the number of activities each farm is involved into. In this case we consider ACTIVITY.FAMILY, that may include irrigated and non-irrigated variants or even a set of similar activities (e.g. winter cereals or other vegetables).

Below we show the quantiles of the number of activities that the FADN dataset are engaged into:

```
##    fadnYEAR 0% 25% 50% 75% 100%
## 1:     2012  1   2   2   3    7
## 2:     2013  1   2   2   3    8
## 3:     2014  1   2   2   3    7
```

We see that half of the farms have up to 2 activities, also 50% of them have exactly 2 activities and 25% of farms have 3 or more.

For the single activity farms, all reported costs are referring to the activity that the farm is engaged into. However certain activities are occur more frequently than others. Below we report the activity occurrence in sinfle activity farm subset.

```
##     ACTIVITY.FAMILY 2012 2013 2014
##  1:         olivefr  143  175  142
##  2:         vinesfr  104  108   29
##  3:          cotton   47   52   60
##  4:        wcereals   30   35   53
##  5:          mzegrn   28   29   28
##  6:         stonefr   28   22   33
##  7:          pomefr   12   11   12
##  8:            rice    9    6    6
##  9:         fod.oth    8   16   12
## 10:   protein.crops    8    8    6
## 11:         veg.oth    7   13    1
## 12:   citrus.fruits    6    6    7
## 13:         veg.tom    5    8   15
## 14:          nutsfr    4    1    1
## 15:      tropicalfr    4    5    7
## 16:           potat    1    3    4
## 17:          sunflr    1    2    4
## 18:         mze.fod    1    0    1
## 19:       berriesfr    1    1    1
## 20:         tobacco    1    2    0
## 21:         grazing    0    0    1
##     ACTIVITY.FAMILY 2012 2013 2014
```

We see that 175 single activity farms are engaged in olive trees and 108 in vine cultivation for 2013. However for other activities we have fewer observations. Nevertheless, for each activity we merge the observations of all three years; for instance for rice we have 9+6+6=21 single activity observations.

Calculation of Variable Cost reference data

The variable cost reference data is an essential element of the allocation procedure. It provides the lower and upper limits that individual farm variable costs can take and also the most probable value. They are farm-wide data.

That reference data consists, for each cost category (e.g. fertilizer, contract work, etc.) and each activity:

1. a MIN and MAX expenditure value per hectare
2. The MEDIAN expenditure value per hectare

Furthermore, since it is reasonable to expect that variable costs per hectare are connected to the output value per hectare, we calculate reference values for different ranges of output value for each activity type. More specifically, for each different type of activity, we split the OUTPUT VALUE (PRICE times YIELD) distribution to quantiles and assign each observation to the appropriate one. Yet for activities with few single farm observations, we may calculate reference values for the whole range of output values per ha.

In Figure 17 we present the results of the reference values for maize grain. In the x-axis we show the two output value ranges (up to 2000 €/ha and >2000 €/ha). In the y-axis is the expenditure per hectare. In each cell of the figure grid a different cost category is presented; we additionally calculate the number of work hours. For instance in the SP.CONTRACT.WORK cost category we see that the reference value (median) is 80 €/ha for output values up to 2000 €/ha and 120 €/ha for output values > 2000€/ha.

In the supplement *Variable Reference Estimates* we provide in details the reference values for each activity.

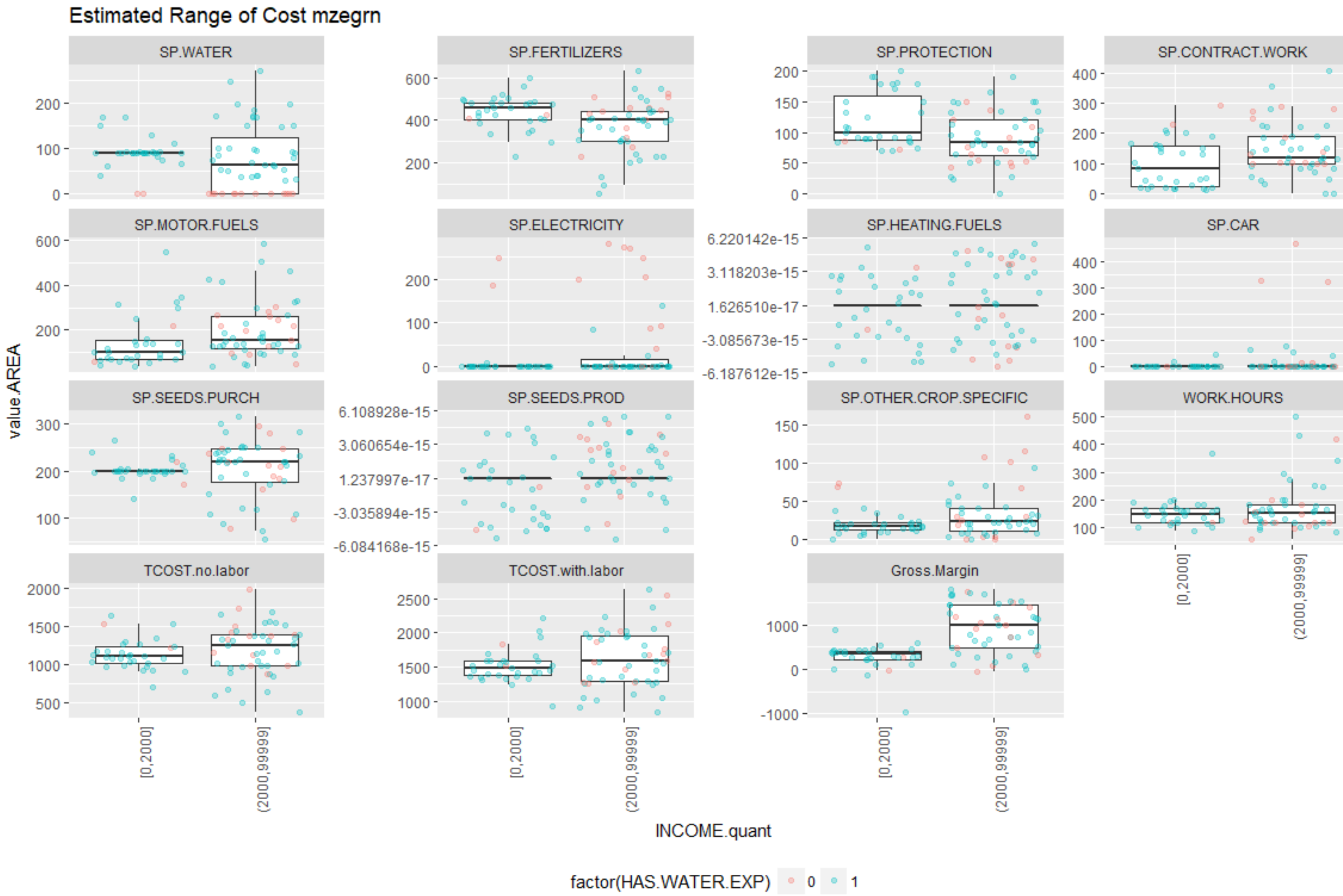


Figure 17, Reference values for Maize Grain

Goal Programming Model to estimate variable costs for each activity on each farm

The goal programming model is based on that of (Guindé, Millet, Rozakis, Sourie, & Treguer, 2005) and (Kampas, Petsakos, & Rozakis, 2012) [17]. A core difference is on how reference values were obtained (in the aforementioned papers this was through an OLS) and the usage of different estimates for different output values per hetare.

[17] Kampas, A., Petsakos, A., & Rozakis, S. (2012). Price induced irrigation water saving: Unraveling conflicts and synergies between European agricultural and water policies for a Greek Water District. *Agricultural Systems*, *113*, 28-38.

The objective is to minimize absolute deviation of the estimated variable cost per hectare from both data-derived variable reference values and of the observed cost total expenditure. The mathematical formulation of the goal programming approach is given in equations (1.1) – (1.7).

$$\min \; w_\delta \sum_i \left(x_i \cdot (\delta_i^+ + \delta_i^-)\right) + w_\varphi \sum_c (\varphi_c^+ + \varphi_c^-) \tag{1.1}$$

Subject to

$$vc_{i,c} \leq VC_{i,c}^{max} \quad \forall i,c \tag{1.2}$$

$$vc_{i,c} \geq VC_{i,c}^{min} \quad \forall i,c \tag{1.3}$$

$$\sum_c (vc_{i,c}) - \delta_i^+ + \delta_i^- = VC_i^{avg} \quad \forall i \tag{1.4}$$

$$\sum_i (x_i \cdot vc_{i,c}) - \varphi_c^+ + \varphi_c^- = VC_c^{obs} \quad \forall c \tag{1.5}$$

$$\delta_i^+, \delta_i^-, \varphi_c^+, \varphi_c^-, vc_{i,n} \geq 0 \tag{1.6}$$

Sets

i for the activities / crops

c for different cost categories

Variables

δ_i^+, δ_i^-: the deviation of the estimation from the regression estimate for crop i (euro/ha)

φ_c^+, φ_c^-: the deviation of the estimation from the observed (aggregate) value on cost type c (euro)

$vc_{i,c}$: the model estimation of variable cost type c of crop i (euro/ha)

Parameters

$VC_{i,c}^{max}, VC_{i,c}^{min}$: The maximum/minimum allowed aggregate variable cost for crop i (euro/ha)

VC_i^{avg}: A reference aggregate variable cost for crop i (euro/ha)

VC_c^{obs}: The observed value of cost type c (euro)

x_i: The area of crop i (ha)

In the objective function (1.1) we minimize the deviation from a regression estimated total variable cost (weighted with the area of each crop) and the deviation from the observed cost.

In (1.2) and (1.3) the estimated variable costs (VC) per cost-type and crop cannot exceed the min and max values that were obtained from the reference data procedure described above. In (1.4), for each crop, we define the deviation between the GP-estimated aggregated (over cost types) VC and the regression-estimated aggregated (total) cost. In (1.5), for each cost type, we define the deviation of all (over crops) VC from the observer VC.

APPENDIX

FADN.utils manual

FADN.utils

R package

version 0.3

Dimitris Kremmydas

Athens, 2017

Purpose of the package

FADN data is used by most farm modelers. Usually a model-ready dataset is provided by an official FADN national office. However working with the raw FADN data is desired since more information is available.

The FADN raw dataset is difficult to handle since many transformations are required in order to get model-ready data. A detailed description of the raw FADN dataset contents are given in RI/CC 1256 rev. 7.

In any case this package will provide functionality for handling FADN raw data within the R language framework.

Conventions

DT: data.table class. See documentation here and a swift introduction for the data.table class

Data structures

fadn.container

A list containing the FADN raw and processed dataset for some year.

It consists of:

- *$tableAI* a DT containing the static information, i.e. columns 1 to 407. It is a fadn.info class.
- *$tableK* a DT with the full farm activities information (table K). It is a fadn.prod class
- *$tableJ* a DT with subsidies received by farms (table J)
- *$tableLMN* a DT containing information on quotas, selected direct payments and purchase/sales of livestock (tables L, M and N)
- *$prod* a data.table containing only the leaf-activities, i.e. not higher hierarchy activities like cereals, trees, etc. that are also named according to some pre-defined rules

It has the following attributes:

- *year* the year that the dataset refers to
- *originalFile* the full path to the file that the raw dataset was stored
- *numOfFarms* the number of farms it contains

fadn.info

A data.table containing columns 1 to 407 of the raw data, i.e. tables A – I.

One can access the original heading number by using numeric indexes. For example in order to access the *national weight of the farm calculated by the Member State* that is given in heading 20 for all farms, one can request ELL_2012$tableAI[,20]

fadn.prod

Functions

getFormulaResult

Aggregates farm columns given a specific formula format.

Input Parameters:

- **data** a fadn.container, containing all tables
- **SEdata** a data.table of already calculated SE
- **formulaString** The formula String to use for aggregation

The formula string is in the form {x}{+,-}{x}. x can be any of:

- "#Y" where Y is the heading number of a fixed column (1 to 405). For example "#20" returns the sampling weights
- "**K**xxx**(**z**)**" where xxx is a product code of table K and z is a column number, e.g. for area equals to 4. For example "K120(4)" returns the areas of common wheat.
- "**K**xxx..yyy**(**z**)**" that returns the sum of column z for all product codes between xxx and yyy (including them)
- "**K**xxx..yyy**(**z..w**)**" that returns the sum of all columns between z and w for all product codes between xxx and yyy
- **SE**xxx where xxx is the column name of a SE dataset

Examples

```
#Total Utilised Agricultural Area (UAA)
  seData$SE025=getFormulaResult(data,
                                seData,
                                "#48+#49+#50");

  #Total crop output
```

```
  seData$SE135=getFormulaResult(data,
                                   seData,
                                  "K120..148(7..10)-120..148(6)+
                                     K150..161(7..10)-K150..161(6)");

  #Total livestock output
  seData$SE206=getFormulaResult(data,
                                   seData,
                                    "#231+  #232+  #234+  #235+  #237+
                                    #238+  #240+  #241+  #243+  #244+
                                    #246+#247+ #249+ #250+ #252+ #253
                                    -#233-#236-#239-#242-#245-#248-
                                    #251-#254
                                   +K162..171(7..10)-
                                    K162..171(6)+K313(7..10)-K313(6)")
```

Use cases

Get farms engaged in a limited number of activities

```
# Load fadn.prod objet of ELL_2013 to a variable
prod.2013= ELL_2013$.prod

# for each farm get number of activities
farms.activity.n= prod.2013 [,.N, by=list(FID)]

# Find frequency of mono-crops (N=1) farms for each activity
table(droplevels(prod.2013[FID%in%farms.activity.n[N==1,FID],ACTIVITY]))
```

See the relationship of variable costs and output value for single-activity farms

```
# See previous Use Case
prod.2013= ELL_2013$.prod
farms.activity.n= prod.2013 [,.N, by=list(FID)]

#Create the various cost categories and load them in a new variable
costs.2013=info.2013[,.(FID,
                        fadnREGION,
                        ECONOMIC_SIZE_real,
                        TYPE_OF_FARM_real,
                        MACHINERY=V260+V261+V262+V263,
                        LVST.GRAZ=V264+V265+V268,
                        LVST.OTH=V271,
                        CROP.SEEDS=V272+V273,
                        CROP.FERTIL=V274,
                        CROP.PROTECT=V275,
                        CROP.OTHER=V276,
                        CROP.WOOD=V277,
                        OVERH.UTIL=V278+V279+V280+V281,
                        TCOST=V260+V261+V262+V263+
                        V264+V265+V268+V271+V272+V273+
                        V274+V275+V276+V277+V279+V280+V281
                        )]

#Create a tabular form of the farms' activities. The table looks like:
#"FID"   "PERIFEREIA"     "cwheat" "cwheat.cmb"     "cwheat.irr.cmb"
#1       "WESTERN MACEDONIA"       6000     0         0
#2       "EPIRUS" 3000     2000     0
prod.2013.outval.table=dcast(prod.2013[],FID+PERIFEREIA~ACTIVITY,value.var  =
                                    "OUTPUT.VALUE",sum)
```

```
#Create a table where costs and activity information are in tabular form
regress.data.outval=merge(costs.2013,prod.2013.outval.table,by="FID")
```

Merge activity information from two distinct year

Compute the annual worked hours for each farm

```
labor.avail=ELL_2013$tableAI[,
                             list(FID,
                                  time.worked=HOURS_OWNER+V58+V62+V66+V70+HO
                                  URS_WIFE+HOURS_FAMILY_OTHERS+V77+HOURS_HIR
                                  ED_WORKERS+HOURS_TEMP_WORKERS
                                  )
                             ]
```

Compute an ordered distribution of the activities by area

```
#Get information on the weight of each farm
prod.2013=merge(prod.2013,
                ELL_2013$tableAI[,list(FID,WEIGHT)],
                all.x=T,
                by="FID")

#Compute weighted area for each activity

prod.2013$wAREA=prod.2013$WEIGHT*prod.2013$AREA

#Aggregate by activity
prod.2013.byAREA=prod.2013[,
                           list(N=.N,
                                wN=sum(WEIGHT),
                                wAREA=sum(wAREA),

                                      wOUTVAL=sum(as.numeric(WEIGHT*OUTPUT.V
                                      ALUE))),
                           by=ACTIVITY]

#Compute total area and output value
tarea.2013=sum(prod.2013.byAREA$wAREA)
tvalue.2013=sum(prod.2013.byAREA$wOUTVAL)

#Compute percentage of area and value in total
prod.2013.byAREA[,':='(pwAREA=wAREA/tarea.2013,
                       pwOUTVAL=wOUTVAL/tvalue.2013)]

#Sort by weighted area
setorder(prod.2013.byAREA,-wAREA)

#Compute cumulative sum
prod.2013.byAREA$CpwAREA=cumsum(prod.2013.byAREA$pwAREA)
prod.2013.byAREA$CpwOUTVAL=cumsum(prod.2013.byAREA$pwOUTVAL)

#Write results in clipboard in exce-format (paste to excel to see them)
write.excel(prod.2013.byAREA,dec=",")
```

Select farms that are engaged in certain activities

```
prod.2013=ELL_2013$prod

#This is the vector of activities we are interested in (cereals)
```

```
activities.incl=c("dwheat","mzegrn.irr.cmb",
                  "barley","cwheat","oats","rye","rice.irr.cmb","ceroth",
                  "mze.fod.irr.cmb","cwheat.irr.cmb",
                  "barley.irr.cmb","dwheat.irr.cmb","sumCereal","mzegrn")

#See what percentage of area on each farm is coverd by activities.incl
farms.prod.2013=merge(
  prod.2013[,list(total.area=sum(AREA)),by=FID],
  prod.2013[ACTIVITY%in%activities.incl,list(incl.area=sum(AREA)),by=FID],
  all.x=T,
  by="FID")

#compute percentage
farms.prod.2013$inclPerc=farms.prod.2013$incl.area/farms.prod.2013$total.area

#for farms that have incl.area=0, set inclPerc=0 (in the merge it was NA)
farms.prod.2013[is.na(incl.area),
                ':='(incl.area=0,inclPerc=0)]

#see quantile distribution
quantile(farms.prod.2013$inclPerc,seq(0,1,.1),na.rm=T)

setorder(farms.prod.2013,"inclPerc")

write.excel(farms.prod.2013)
```

Plot yields for a certain activity for a certain region

Calculate total variable costs per area for all farms

Load and prepare raw files

In order to load and prepare a raw file, the following additional files have to be prepared:

activities.definitions.txt

A definition of the activity defined by the combination of PRODUCT, TYPE and MISSING columns. Activity is the name given to the activity and activity.family is the crop-family that this activity belongs to.

Example:

```
PRODUCT     TYPE  MISSING     ACTIVITY  ACTIVITY.FAMILY
120    1      0      cwheat   cwheat
120    2      0      cwheat.cmb   cwheat
120    3      0      cwheat.folup   cwheat
120    6      0      cwheat.irr.cmb   cwheat
120    7      0      cwheat.irr.folup   cwheat
120    10     0      cwheat.energ   cwheat
120 {0,1,2,3,4,5,6,7,10}      {1,2,3,4,5,6}      cwheat.miss cwheat
```

APPENDIX

Conference Paper

Dealing with farm heterogeneity on modeling agricultural policy: An Agent Based Modeling Approach

Presented in in the 20th Annual Workshop on the Economic Science with Heterogeneous Interacting Agents (WEHIA), in Sophia Antipolis, in the Côte d'Azur region of France, on May 2015

Dealing with farm heterogeneity on modeling agricultural policy: An Agent Based Modeling Approach

Dimitris Kremmydas[1†], Stelios Rozakis[1], Ioannis Athanasiadis[2]

[1] Department of Agricultural Economics and Rural Development, Agricultural University of Athens
[2] Electrical and Computer Engineering Dept, Democritus University of Thrace
[†] Corresponding author

Abstract

The Common Agricultural Policy (CAP) has been gradually transformed from directly supporting prices and production to a decoupled scheme, where farmers receive a payment per hectare regardless of their production decisions. Within this framework and given the multitude of CAP's objectives, ranging from market competiveness to multifunctionality of agriculture, the inclusion of the heterogeneity of farms on modeling CAP will continuously arise in the immediate future as a key research question. In this paper we make a brief discussion of the aspects of farm heterogeneity within the agricultural policy modeling context and we show that Agent Based Modeling approach, coupled with Object Oriented System Analysis, is a very good alternative for considering all aspects of diversity of farms. Finally we present Agroscape, a flexible ABM Agricultural Policy Framework that can easily incorporates both behavioral and capacity heterogeneity presenting a proof-of-concept case study.

Keywords: Agricultural Policy, Agent Based Modeling, Object Oriented System analysis, Farm heterogeneity
JEL Codes: Q12, Q18, C63

1 Introduction

"Heterogeneity" (from Greek ετερογένεια / heterogeneia, i.e. of different family/gene) is defined as "the quality of being consisted of dissimilar or diverse ingredients or constituents". The term can refer to two things: (a) the fact that farms, within a specific geographical area, differ from each other in almost all of their aspects (local heterogeneity) and (b) that between two distant regions the distributions of farm characteristics are different (inter-regional heterogeneity). In this paper we focus on local heterogeneity.

The need for employing a discussion focusing on farm heterogeneity emerges from the last (European) Common Agricultural Policy (CAP) reform. More specifically, the CAP-2020 reform introduces a number of new objectives, such as advancing agricultural productivity, securing fair farmers' income, even in less advantaged areas, maintaining food security through the promotion of the respect of certain standards, safeguarding the sustainable management of natural resources and the protection of the viability of rural economy.

The proposed means to achieve the above goals are also innovative. Indicatively, there are measures to facilitate collective investment; assist small farms to develop; foster knowledge transfer and innovation; enhance competitiveness; promote food chain organization & risk management; restore, preserve & enhance ecosystem services; promote resource efficiency and transition to a low-carbon economy; promote social inclusion, poverty reduction and economic development in rural areas. All of those measures have a local and non-aggregate dimension and so the individual characteristics of farms should be taken into account by policy modelers.

Furthermore the new CAP departs from a status where strict rules were applied for all member states, introducing implementation flexibility. The member states are equipped with a toolbox of measures that can fine tune in order to cater their specific needs. Consequently, national policy makers are required to make more focused decisions. In this manner considering farm heterogeneity is applicable in the design of the CAP policy.

In this paper we discuss some aspects of farm heterogeneity within the agricultural policy modeling context arguing that Agent Based Modeling (ABM) approach is a very good alternative for dealing with it. Finally we present an ABM framework, capable of incorporating the various forms of farm heterogeneity and provide some proof-of-concept results.

2 Theoretical Considerations

2.1 Aspects of farmer's heterogeneity

Given that the scope of the paper is focused on agricultural policy the discussion will be confined to aspects of farm heterogeneity that concern CAP. This is already a wide scope, because as mentioned above, this policy is concerned with many facets of an agricultural system (technical, economic, environmental, and social).

There is an evident first layer of farm heterogeneity which could be termed as "**endowment heterogeneity"**. Farms own land of different soil quality and have different initial land and labor endowments. This kind of heterogeneity is most often included in farm models mainly by differentiating farm yields and variable costs.

A second layer could be the **"managerial ability heterogeneity"** referring to both technical and economic efficiency of the farm management. In this layer only operational (short-term) and tactical (mid-term) managerial decisions are included. As usually observed in farm efficiency analysis case studies, there is a significant variation between farms in managerial ability (Nuthall, 2001) and it is of interest to policy makers since it is an essential factor in successfully explaining varying agricultural output and supply relationships. There exist several studies providing explanations of this variation. As Nuthall (2009) argues, managerial ability is related to education, training and intelligence, age and experience and also "to social capital which involves the networks a manager may have, as well as the relevant components of the current culture".

A third layer of heterogeneity is about the strategic orientation of farms, i.e. theirs long-term goals and how the take decisions, which could be termed as **"decision making heterogeneity"**. In the literature it is widely recognized that not all farms are profit maximizers. Apart from the "satisficing" principle and the notion of bounded rationality that Simon (1957) introduced, there is the case of the multiplicity of goals –potentially modeled in a multi-objective problem, for example in Karanikolas et al. (2013). There is evidence that the age of the farm manager is a significant factor shaping his behavior.

2.2 Including farmer's heterogeneity in Agricultural Policy models

Farm heterogeneity could be included in the agricultural policy discussion in two modes. Firstly as a factor that differentiates the impact of a policy from one farm to another, or from one group of farms to another, and thus is used in the result analysis stage in order to get a more detailed view of the final state of the system (result analysis mode).

Secondly as a factor that has an endogenous effect on the results of the policy, and thus is included directly into the model (modeling mode). An example of the latter can be found at Liu et al. (2007): "the socioeconomic differences among people in a relevant case study lead to different choices and behaviors, which in turn result in very different ecological outcomes than one would find were everyone to have the same preferences for ecosystem services". Another case of heterogeneity "modeling mode" would be in the case of Agricultural Value Chains (Nolan, 2009). As noted, "modern agricultural systems (production, distribution, marketing) are in a state of transition, with increasingly numerous and heterogeneous agents interacting in the value chain". Comprehension of the function of this value chain should include the endogenous modeling of the heterogeneity of the relevant agents.

For short to mid-term models it is reasonable to assume that ignoring endogenous effects of heterogeneity will be a good approximation of reality. On the other hand, for long term modeling, like human – nature interaction is, heterogeneity becomes a crucial element of the modeling process.

Nevertheless, representing heterogeneity in policy models can have some intrinsic limitations. As Heckelei (2013) notes, current agricultural databases have a limited level of detail. Thus heterogeneity ends, at the good scenario, being inferred through statistical methods . Also the spatial heterogeneity should be specifically relevant for a particular policy impact otherwise the benefits are questionable. This argument apparently holds for any kind of heterogeneity.

3 The Agent Based Modeling approach

3.1 The Agroscape ABM Framework

3.1.1 Design Principles

The Agroscape ABM is a modeling framework with the aim of facilitating the modeling of a variety of agricultural production systems. Its design is based on the following principles:

(a) The agricultural production system, being a coupled human – nature system, is a complex adaptive system. It contains reciprocal and feedback loops exhibiting nonlinearities like thresholds (Liu et al., 2007). Since the agricultural policies are increasingly focusing on environmental goals, the modeling process should incorporate this complexity.

(b) A good approach for modeling complex systems is Object Oriented Analysis and Design (OOAD). Briefly, OOAD is about structurally and functionally decomposing a system into smaller units with less complexity and less responsibilities (Booch, 2007). The collaboration of those simpler components is considered to provide the functionality of the system-as-a-whole (Solms, 2014) and thus the question of modeling the complex system is transformed in the questions of modeling many smaller and simpler (non-complex) components and their interactions. More technically OOAD is about applying the principles of abstraction, encapsulation, modularity and hierarchy to the system under consideration. Although OOAD is very closely related to software design, the last two decades is applied to other disciplines as well.

(c) The Agent Based Modeling is an adequate approach for modeling complex adaptive systems using OOAD. It is evident that ABM fits very well with the concepts of Object Oriented Programming, representing agents as software objects and focusing on their interaction.

3.1.2 Main Modeling Elements

Agroscape is programmed in Repast Simphony[18], which is a Java ABM programming framework providing an ABM scheduling namespace and many visual enhancements. Since our framework is actually build upon the Repast Simphony Object model we present briefly its essential elements:

1. *The Context* interface is actually a *Collection* that holds simulation objects. *Contexts* can include other *contexts*, thus providing the ability to the modeller for creating hierarchies of collections of agents. *Contexts* support *Projection* and *Data Layers* classes. All objects in Repast start their life in a root *Context*.
2. *The Projection* interface is a collection of relations between simulation objects. Spatial or network relationships are represented by corresponding projections. A *Projection* is always attached to a specific *context*, imposing a structure upon the contained agents. Apart from Continuous and grid space implementations there are also GIS and Network implementations of this interface
3. *The Data Layer* interface allow the efficient handling of the interaction between agents and data. A *Data Layer* can be either an abstract matrix attached to a *context* or a matrix attached to a Grid Projection thus storing one value for each grid's cell (*GridValueLayer*).

One can see in detail the class hierarchy of Repast in its API documentation[19].

Regarding Agroscape, the skeleton of the framework is shown in the UML 2.0 class diagram on Figure 18. The two core classes are *Farmer* and *Space*, the former being a POJO (Plain Old Java

[18] http://repast.sourceforge.net/repast_simphony.php

[19] http://repast.sourceforge.net/docs/api/repast_simphony/

Object) and the latter a *Grid Projection*, i.e. a pixeled surface where all activity is taking place. The simulation can also contain many *PropertyGridValueLayer* (a *Data Layer* class) objects, i.e. spatial properties (e.g. soil quality, crop suitability, nitrates concentration etc.). *Farmer* is related to *Space* indirectly through *Plot*, which is a logical grouping of space and this is realized through a *LandRegistryAuthority* class that is responsible for the bookkeeping of the ownership of the Plots. The actual ownership relation is between a *HumanAgent* and a *Plot*, since it is expected that also non-Farmers might own land. *Farmer* is also related to *Space* by the *residentIn* association. Additionally, all *Farmer* objects exist within a *FarmerContext* object. The latter can contain an arbitrary number of *Network* objects, representing various kinds of relationships between farmers, for example a social network, an information exchange network, etc. The activity of the agents is realized through attached behaviors, contained in *FarmerContext* and explained in more details right after. The *behaviors* scheme provides modeling extensibility, since new behaviors are easy to be added.

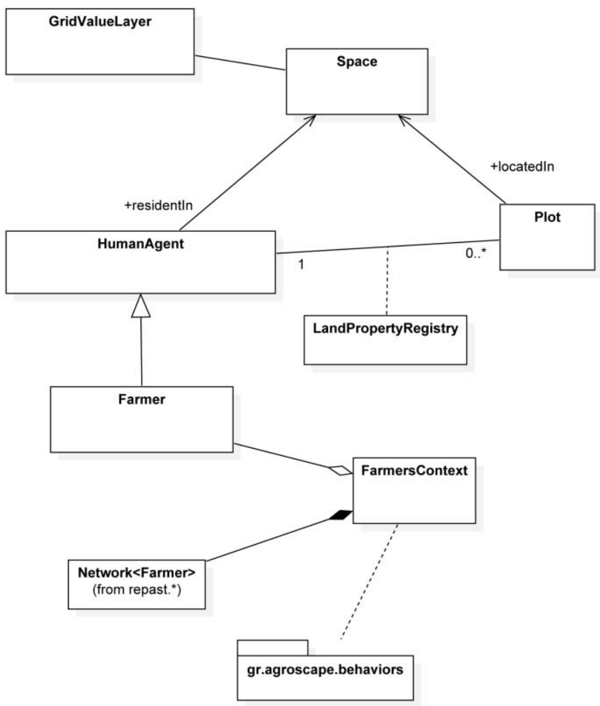


Figure 18, UML class diagram of the Agroscape skeleton

3.1.3 The Behavior Package

The aspired flexibility of Agroscape is founded on the idea of the "behavior package". The agents that are defined in the skeleton of the framework are idle by default. In order to act they need to be attached to one or more behavior classes. A behavior class is an implementation of an action, like taking production decisions, realizing production, making transactions in a land market, deciding for the adoption of a new technology, etc. It is an OOAD way of programming agents in the simulation's timeline.

If ones tries to implement many kind of behaviors for a specific agent with the conventional way of hard-coding them directly into its namespace, the maintenance of the code becomes cumbersome and very possibly conflicts appear between them. On the other hand, the “behavior package” is an innovative and pluggable approach to implement and addi new behaviors to various agents of the simulation is easy and also keeps the modeling complexity in manageable levels.

That is because the behaviors are independent of each other and use only the core classes of the simulation, as described above, without being affected by their properties. The modeler of a certain behavior is also absolutely responsible for scheduling its operations and implementing the logic, without being affected by other already implemented behaviors. Furthermore one behavior that is attached to an agent can also use objects from another behavior of the same agent, since they are all connected to the same agent object. Finally modelers have the possibility to attach different set of behaviors to different agents.

In the implementation side, the idea is based on *IScheduledBehavior* interface which defines a single *getAnnotatedClass* method, returning an object containing *ScheduledMethodAnnotation* annotations. This annotation class defines several properties of a method’s scheduling, like the start and interval ticks of the simulation clock. In this way the modeler have full flexibility on the timing of behaviors. What is left is to model the behavior of the agent.

In order to do so, he has to extend the abstract *AFarmerBehavior* class that contains a reference to the *Farmer* object that exhibits the behavior and also implements *IScheduledBehavior*. Also, since frequently all agents that are attached to a specific behavior will need to either access common classes, or communicate with each other, all behaving objects are contained in an extension of a *ABehaviorContext*. This context also contains a *IscheduledBehaviorDataLoader* object that is responsible for loading all the behaving objects into the context and also adding their behavior to the simulation’s schedule.

A simple illustrative example of a behavior is given in Appendix.

Figure 19, Class diagramm of Behavior package

3.2 Catering for Farm Heterogeneity with the Agroscape framework

Our approach for representing heterogeneity deviates from searching for a suitable typology, where a classification according to certain macro-indicators is performed and certain farm types are derived, e.g. Amico et al. (2013). Rather than performing a statistical analysis of the observed outputs we facilitate the analysis of the production system in terms of system components and theirs relations and embed the heterogeneity in agents' state, actions and interactions.

Following the "behavior package" approach, the various forms of heterogeneity is not modeled directly into a Farmer object but rather is embodied in the individual behavior classes. This approach, although does not seem very natural, is not limiting at all, because any behavior class can exchange information with the skeleton classes of the model.

More specifically, as far as land endowment heterogeneity is regarded, any behavior can have access to the plots that a farmer use and so has access to spatial diversity through the *PropertyGridValueLayer* class. For capital and labor endowment heterogeneity, a behavior can introduce attributes that express this fact, as shown in the case study.

As far as "managerial ability heterogeneity" is concerned, one approach would be to model managerial ability as a [0,1] coefficient that is multiplied with expected yield to give the actual yield, differentiating efficient farmers. Another approach would be to make a low-level modeling of the technical or economic decisions that a farmer is following, embedding managerial details. A relevant modeling of the production realization is shown in Daydé et al. (2014). Although that approach seems to be much more complex, the Agroscape framework could easily facilitate it, showing that the OOAD benefits.

Also decision making heterogeneity is very easily modeled within the proposed framework. Farmers can easily be attached to different specific decision making behaviors interacting with other simulation elements very easily. Furthermore already implemented decision making models can be incorporated to an Agroscape model, utilizing the flexibility and the power of the Java programming language. For instance, since there are many farm models written in GAMS mathematical programming software, a special adapter class could be crafted to use the already written code for a production behavior.

In order to illustrate the above arguments, a simple proof-of-concept case study has been implemented, where farmers own land of different crop suitability and exhibit varying behavior. One can first examine the appendix for an even simpler "hello-world" example.

3.3 A proof-of-concept case study: The *arableCropProduction* Behavior

Arable Crop yearly decisions have been extensively used in agricultural policy modeling over the last decades. An elementary model can be represented as a linear programming problem where an individual farmer f is supposed to choose a cropping plan $\vec{x}^f$ and input use among technically feasible activity plans $A^f \cdot \vec{x}^f \leq \vec{b}^f$ so as to maximize gross margin gm^f. The optimization problem for the farmer f can be expressed as follows (Kremmydas et al., 2012):

$$\begin{cases} \max\limits_{x^f} \quad gm^f\left(\mathbf{x}^f, \boldsymbol{\theta}^f, \boldsymbol{\kappa}\right) \equiv \sum\limits_{i=1}^{n} \left\{ \left[\left(p_i \mid p_i^f + ps_i \right) y_i^f + ls_i - c_i^f \right] x_i^f \right\} \\ \text{s.t.} \quad \mathbf{A}^f\left(\boldsymbol{\theta}^f\right)\mathbf{x}^f \leq \mathbf{b}^f\left(\boldsymbol{\theta}^f\right) \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \\ \qquad\qquad \mathbf{x}^f \geq 0 \qquad\qquad \mathbf{x} \in \mathbb{R}^{n} \end{cases} \tag{1}$$

Where

The $m \times n$ matrix A^f and the $m \times 1$ vector $\vec{b}^f$ represent respectively the technical coefficients and the capacities of the m constraints on production. The vector of parameters θ^f includes yields for crop i (y_i^f), variable costs (c_i^f), prices dependent on quality (p_i^f) and subsidies linked to crop quantity (ls_i^f). Symbol stands for the vector of general economic parameters which includes prices not dependent on farm (p_i) and subsidies specific to crop cultivated area (ps_i).

Below we describe the transformation of such an elementary model to an agent based model, following the "Behavior package" approach.

The ArableCropProductionBhv

Following the OOAD principles we decompose the above linear programming problem domain to its constituent elements, representing each as a different class. We also provide additional elements in order to represent the spatial dimension, not currently available to the above formulation, and we finally introduce classes for different decision making strategies (linear programming being one member of the set of strategies). All of the derived classes are related to the skeleton classes of the Agroscape framework that has already been presented in 3.1.2.

Thus the derived classes of the domain are:

1. *ArableCropCultivation* represents the various crop cultivations that are available, e.g. maize, durum wheat, barley, etc.
2. *ArableCropProductionDecision* is a map from a *Plot* (see 3.1.2) to an *ArableCropCultivation*, denoting the fact that the farmer's decision is actually the assignment of an arable crop to each of the owned plots.

   ExpectedCropPrices, *ExpectedPlotCropVarCost* and *ExpectedPlotCropYield* are the corresponding elements of the farmer's objective function. So a farmer has certain expectations about the next year's prices of a crop output (*ExpectedCropPrices*) and he has an expectation regarding the variable cost and the yield of the "Crop x Plot" combinations. In the current paper exercise we implemented the formation of those expectations to be really simple (taken from pre-defined values plus/minus a random number). In the future we could implement a more realistic but complex modeling of how those expectations are formed (e.g. prospect theory, evolutionary algorithms, through networking with other farmers, etc.) and here the power of the OOAD approach emerges: Even if we would insert into the model such complex procedures, we need only to change the respective *Expectation* class while the rest of the model would left intact. That is because we first tackled the complexity of the relationships between the various classes of the domain, modeling the points of contact between them, and encapsulated the complexity of the classes into themselves. In this way OOAD releases the modeling process from the burden of dealing with the complexity of each element of the system at the same time.

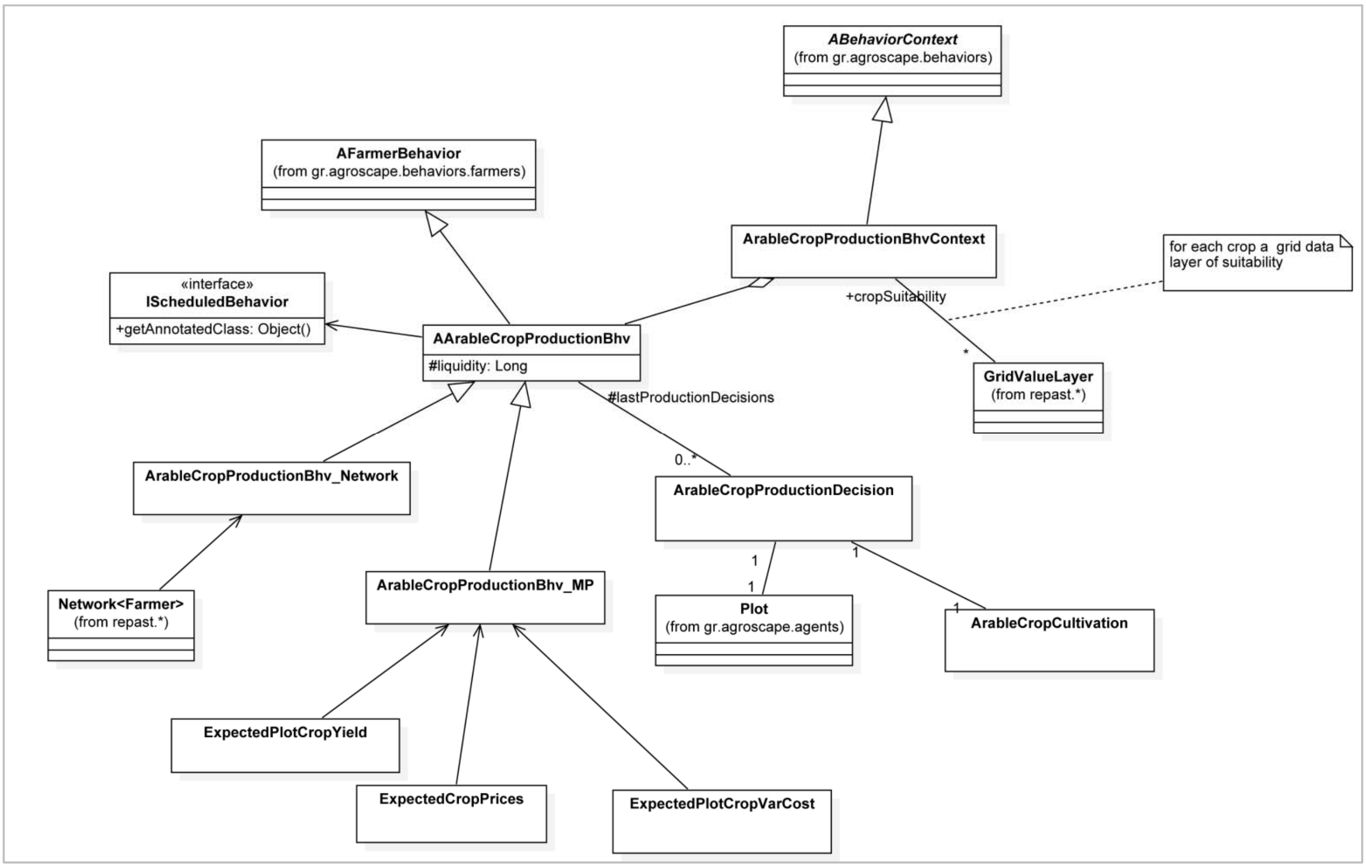


Figure 20, An overview of the arableCropProductionBehavior

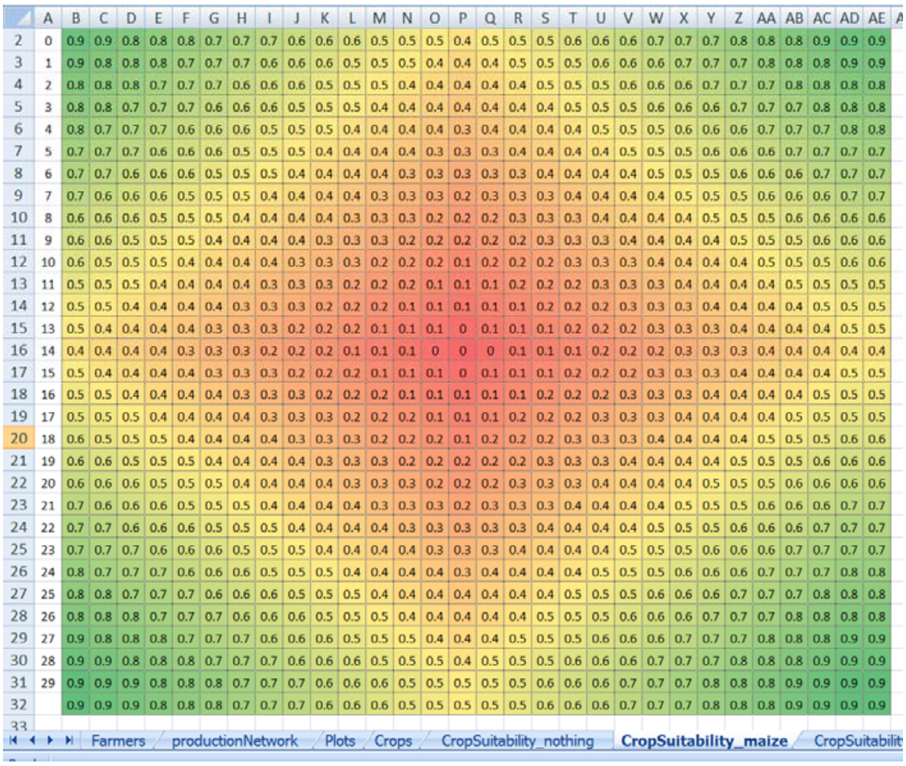
Data and Results

The data management issue on modeling agricultural policy cannot be overlooked. The transformation and loading of data for such models is usually cumbersome, especially if ones goes to plot-level detail or include spatial data, since modeling software (like GAMS) does not provide explicit data handling mechanisms.

Although the data used in this exercise was fictitious, we followed the OOAD approach in the data management aspect, unbinding the mechanics of the model with the data loading process and hopefully giving more flexibility to potential modelers.

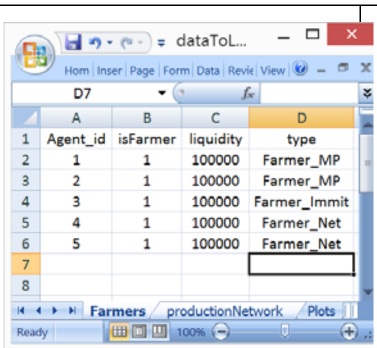
In order for loading data into the Agroscape framework, the *IScheduledBehaviorDataLoader<T>* shall be implemented. This interface defines the *setup(ABehaviorContext<T> container)* method that is called for each behavior loaded during the initialization of the simulation and the *ABehaviorContext<T>* top context is passed. The implementation should be done so as all *AArableCropProductionBhv<T>* objects are loaded in the container.

In our example case we kept all data in an excel sheet, as shown in figures 4 - 7. In order to load the data we created an *ExcelDataLoader* class implementing the required interface. In Figure 25 we show the essential part of the code. The class creation takes the excel workbook and the setup method, that is called from the simulatin initation procedure provides the container that the data loader class loads *AArableCropProductionBhv* objects with private and unexposed functions. If one wants to load data through a different excel structure, then he has to change the internal functionality of *ExcelDataLoader* (or write a new implementation) without being concerned about the stability of the overall model.



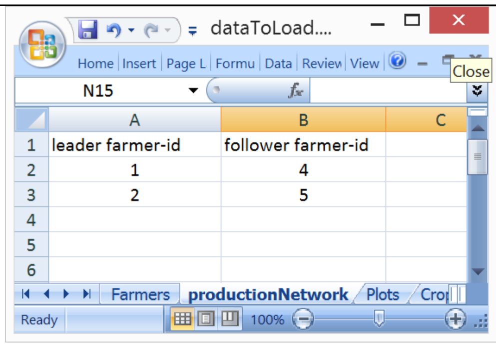
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
2	0	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9	0.9
3	1	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9
4	2	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.8
5	3	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8
6	4	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8
7	5	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.7
8	6	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7
9	7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7
10	8	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.6
11	9	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6
12	10	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6
13	11	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5
14	12	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5
15	13	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5
16	14	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0	0	0	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.4
17	15	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5
18	16	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5
19	17	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5
20	18	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6
21	19	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6
22	20	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.6
23	21	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7
24	22	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7
25	23	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.7
26	24	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.3	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8
27	25	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8
28	26	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.8
29	27	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.4	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9
30	28	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9	0.9
31	29	0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9	0.9	0.9
32		0.9	0.9	0.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.6	0.6	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9	0.9	0.9

Figure 21, Excel Data for the Crop Suitability of maize



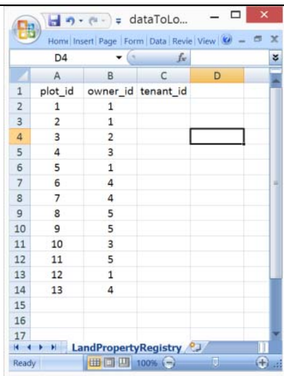
	A	B	C	D
1	Agent_id	isFarmer	liquidity	type
2	1	1	100000	Farmer_MP
3	2	1	100000	Farmer_MP
4	3	1	100000	Farmer_Immit
5	4	1	100000	Farmer_Net
6	5	1	100000	Farmer_Net

Figure 22, Data for assigning decision strategies to farmers



	A	B
1	leader farmer-id	follower farmer-id
2	1	4
3	2	5

Figure 23, Data for creating the network



	A	B	C
1	plot_id	owner_id	tenant_id
2	1	1	
3	2	1	
4	3	2	
5	4	3	
6	5	1	
7	6	4	
8	7	4	
9	8	5	
10	9	5	
11	10	3	
12	11	5	
13	12	1	
14	13	4	

Figure 24, Data for Land Property Registry

```java
public class ExcelDataLoader implements IScheduledBehaviorDataLoader<AArableCropProductionBhv> {

    private Workbook excelWB;

    private ArrayList<ArableCropCultivation> crops;
    private SimulationContext simulationContext;

    public ExcelDataLoader(Workbook excelWB) {
        super();
        this.excelWB = excelWB;
        simulationContext=SimulationContext.getInstance();
    }



    @Override
    public void setup(ABehaviorContext<AArableCropProductionBhv> container) {
        this.loadCrops(container);
        this.loadCropSuitabilities(container);
        this.setupPaymentAuthority(container);
        this.addAgents(container);
    }
```

Figure 25, ExcelDataLoader essential part

After loading the data the simulation was run for a certain number of iterations, recording the decisions of the farmers. We recorded data through the relevant time-saving Repast Simphony mechanism. For example the allocation of the crop to plots was recorded in a video and a graph of the total land per crop was also easily configured to output. The vibration of the surface allocated to the crops is to the feedback mechanism of the deterioration and restoration of the cropSuitability feature of the model.

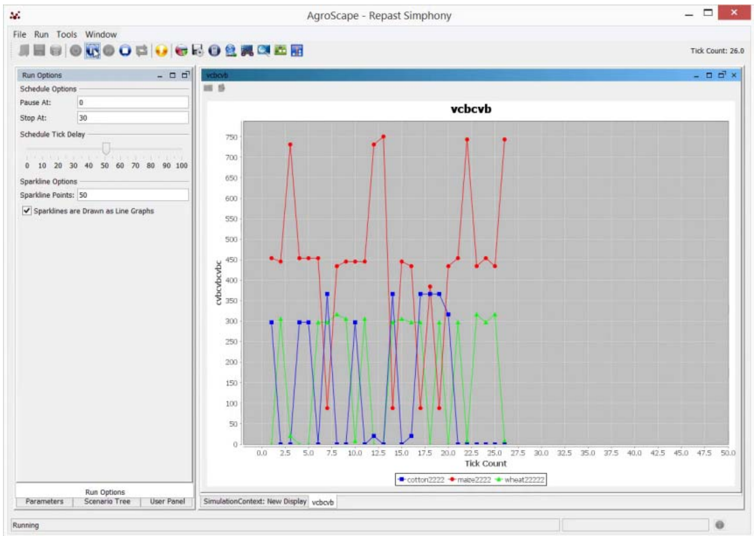


Figure 26, The time series of total land per arable crop

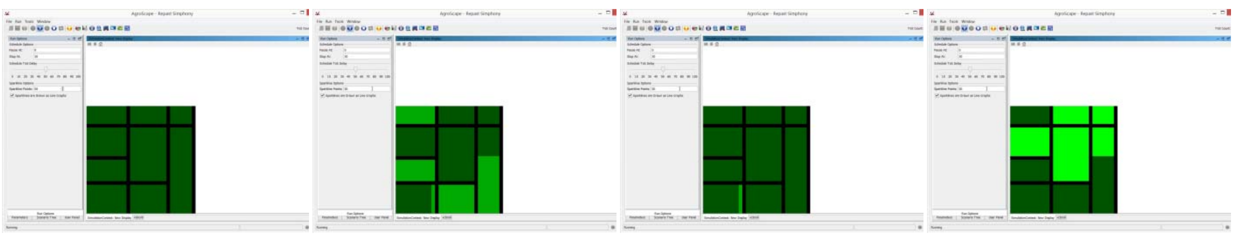


Figure 27, Crops to Plots allocation evolving through time

3.4 Conclusions and Future Research

In this paper we attempted, firstly to prove that agent based modeling should be considered as a well suited modeling approach for dealing with farm heterogeneity in agricultural policy modeling and secondly to propose an agent based modeling framework (Agroscape) that relies heavily on the OOAD principles

The advantages of this approach are:

- Endowment and Managerial farm heterogeneity can be represented as easily as in other approaches. Furthermore space is inherently represented in ABM simulation systems whereas this in not the case in general.
- Managerial heterogeneity can be modeled more efficiently compared to other approaches
- If the OOAD principles are followed, the managerial heterogeneity modeling can be carried out without the complexity "explosion" of the modeling process. The latter is present when one tries to build models that deal with many and different aspects of a system at the same time.

We implemented a proof-of-concept case study with just 5 farmers following two different decision making strategies for selecting an arable crop to cultivate in one of their owned plots (30X30 grid containing 13 plots).

Future work could include:

- Incorporate a Land market througth the behavior package mechanism
- Incorporate otherproduction decision behaviors, like Animal Husbandry production decisions (independently or jointly with arable crop decisions), permanent crop installation and handling decisions, etc
- Model other key players of the Agricultural value chain and investigate on the interaction with farmers (e.g. an information exchange network).
- Implement a real case policy evaluation case

4 References

Booch Grady, et al., 2007. Object-Oriented Analysis and Design with Applications (3nd Ed.).Pearson Education, Inc., Boston, MA, USA.

D. Kremmydas, A. Petsakos, and S. Rozakis. 2012. Parametric Optimization of Linear and Non-Linear Models via Parallel Computing to Enhance Web-Spatial DSS Interactivity. *Int. J. Decis Support Syst. Technol.* 4, 1 (January 2012), 14-29. DOI=http://dx.doi.org/10.4018/jdsst.2012010102

D'Amico Mario, et al., 2013. Agricultural systems in the European Union: an analysis of regional differences. New Medit, N. 4/2013.

Daydé, C., Couture, S., Garcia, F., Martin-Clouaire, R. (2014). . Investigating operational decision-making in agriculture. In: Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. Presented at 7th Congress on Environmental Modelling and Software (IEMSS 2014), San Diego, USA (2014-06-15 – 2014-06-19). Downloaded from http://prodinra.inra.fr/record/258496

Daydé, C., Couture, S., Garcia, F., Martin-Clouaire, R., (2014). . Investigating operational decision-making in agriculture. In: Proceedings of the 7th International Congress on Environmental Modelling and Software, June 15-19, San Diego, California, USA. Presented at 7th Congress on Environmental Modelling and Software (IEMSS 2014), San Diego, USA (2014-06-15 – 2014-06-19). Downloaded from http://prodinra.inra.fr/record/258496

European Commision, 2013. MEMO: The common agricultural policy (CAP) and agriculture in Europe – Frequently asked questions, Brussels 26 June 2013, accessed from http://europa.eu/rapid/press-release_MEMO-13-631_en.htm.

Heckelei Thomas, 2013. General methodological issues on farm level modelling. Chapter in Farm level modelling of CAP: a methodological overview, Ed. Langrell Stephen, JRC Scientific and Policy Reports,

Herbert A. Simon, 1957. Models of Man: Social and Rational. New York: John Wiley and Sons, Inc

Jianguo Liu , et al., 2007, Complexity of Coupled Human and Natural Systems, Science 317 (5844), 1513-1516. [DOI:10.1126/science.1144004]

Jianguo Liu , Thomas Dietz, Stephen R. Carpenter, Marina Alberti, Carl Folke , et al., 2007. Complexity of Coupled Human and Natural Systems, Science 14 September 2007: 317 (5844), 1513-1516. [DOI:10.1126/science.1144004]

Juliano J. Assunção, Maitreesh Ghatak, 2003. Can unobserved heterogeneity in farmer ability explain the inverse relationship between farm size and productivity, Economics Letters, Volume 80, Issue 2, August 2003, Pages 189-194, ISSN 0165-1765, http://dx.doi.org/10.1016/S0165-1765(03)00091-0.

McConnell, D. J. & Dillon, John L. & Food and Agriculture Organization of the United Nations., (1997). Farm management for Asia : a systems approach. Rome : Food and Agriculture Organization of the United Nations

Nolan, J., Parker, D., Van Kooten, G. C. and Berger, T., 2009. An Overview of Computational Modeling in Agricultural and Resource Economics. Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, 57: 417–429. doi: 10.1111/j.1744-7976.2009.01163.x

Nuthall, P., 2009. Modelling the origins of managerial ability in agricultural production. Australian Journal of Agricultural and Resource Economics, 53: 413–436. doi: 10.1111/j.1467-8489.2009.00459.x

Nuthall, P.L., 2001. Managerial ability — a review of its basis and potential improvement using psychological concepts. Agricultural Economics, 24: 247–262. doi: 10.1111/j.1574-0862.2001.tb00028.x

Ondersteijn, C. J. M. & Giesen, G. W. J. & Huirne, R. B. M., 2003. "Identification of farmer characteristics and farm strategies explaining changes in environmental management and environmental and economic performance of dairy farms," Agricultural Systems, Elsevier, vol. 78(1), pages 31-55, October.

Pavlos Karanikolas, Stelios Rozakis, Dimitris Kremmydas, 2013. "The multiplicity of goals in tree-cultivating farms in Greece," Working Papers 2013-2, Agricultural University of Athens, Department Of Agricultural Economics.

Rob J.F. Burton, The influence of farmer demographic characteristics on environmental behaviour: A review, Journal of Environmental Management, Volume 135, 15 March 2014, Pages 19-26, ISSN 0301-4797, http://dx.doi.org/10.1016/j.jenvman.2013.12.005

Solms Fritz, 2014, Object-Oriented Analysis and Design using UML, Book, accessed from http://www.fritzsolms.net/sites/default/files/documents/ObjectOrientedAnalysisAndDesignUsingUML.pdf

Appendix, A “hello world” example: The Stupido Behavior

The implementation of a “hello-world” behavior will now be analyzed. The farmer that is attached to this behavior prints the value of an internal *stupidoProperty* every tick and updates this property every two ticks. The fact that the required *stupidoProperty* attribute is contained within a *StupidoBhv* object and not within a *Farmer* object enables the controlling of complexity to a manageable level for an arbitrarily large number of behaviors. That is because every new behavior can be developed without being affected by other behaviors, since their namespace can be absolutely independent and thus farmers can be attached to any behavior without programming conflicts. One can see the structure of the behavior’s files in Figure 28.

StupidoBhv (Figure 30) is the actual behavior object, extending the *AFarmerBehavior<StupidoBhv>* class. The *AFarmerBehavior<T>* class (Figure 31) is actually enforcing the connection between the behavior object and the farmer object that should be contained there. Since there is an association between the behavior and the farmer objects, one behavior object can use the other behavior objects attached to the same farmer. Finally, in order for the *StupidoBhv* to take action in the simulation’s timeline, it has to implement the *IScheduledBehavior<T>* interface. This is actually realized in two steps. First the *getAnnotatedClass* (lines 34-36, Figure 30) is implemented, returning the behaving object itself. Second, in the returned class (in this case any *StupidoBhv* object), *@ScheduledMethod* annotations have to be inserted accordingly. One can see that this is done in lines 23 and 28 (Figure 30) scheduling the *setRandom* (every 2 ticks) and *print* (every 1 tick) methods respectively.

The *StupidoBhvContext* (Figure 29) is the class that contains all the *StupidoBhv* objects and acts as a container of common functionality. The access of the individual behavior objects is facilitated through the container attribute in *StupidoBhv* (line 13, Figure 30). In our case we need a common random generator which is defined in line 18 of *StupidoBhvContext* (Figure 29). This generator is initialized and used by all contained behaving objects.

Also one should note that agents in the behavior context are loaded using a IScheduledBehaviorDataLoader<T> interface. The implementing classes take a collection of farmers, create the behaving objects and add them to the context and to the simulation timeline.

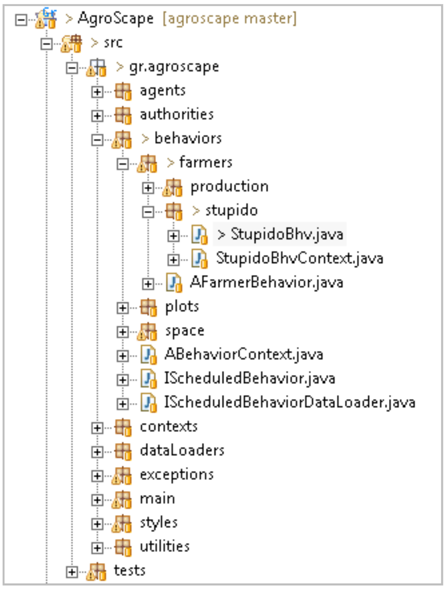


Figure 28, The structure of the files in the Stupido behavior

StupidoBhvContext.java

```java
package gr.agroscape.behaviors.farmers.stupido;

import gr.agroscape.agents.Farmer;

/**
 * The context for the StupidoBhv
 * @author Dimitris Kremmydas
 */
public class StupidoBhvContext extends ABehaviorContext<StupidoBhv> {

    protected Random randomGenerator ;

    public StupidoBhvContext(Collection<? super Farmer> owners) {
        super("stupidoBehavior", new DefaultStupidoDataLoader(owners));
        this.randomGenerator = new Random(System.currentTimeMillis());
        this.loadBehavingObjects();
    }


    public int getRandom() {
        return this.randomGenerator.nextInt();
    }

}


/**
 * Inner class to load stupidoBhv ojects
 * @author Dimitris Kremmydas
 */
class DefaultStupidoDataLoader implements IScheduledBehaviorDataLoader<StupidoBhv> {

    private Collection<? super Farmer> owners;

    public DefaultStupidoDataLoader(Collection<? super Farmer> owners) {
        super();
        this.owners = owners;
    }

    @Override
    public void setup( ABehaviorContext<StupidoBhv> container) {
        Collection<IScheduledBehavior<StupidoBhv>> r = new ArrayList<IScheduledBehavior<StupidoBhv>>();
            for (Object f : owners) {
                StupidoBhv toadd = new StupidoBhv((Farmer)f,(StupidoBhvContext)container);
                r.add(toadd);
            }
            container.addAll(r);
    }

} //end class
```

Figure 29, The StupidoBhvContext.java code

*StupidoBhv.java

```java
package gr.agroscape.behaviors.farmers.stupido;

import gr.agroscape.agents.Farmer;

public class StupidoBhv extends AFarmerBehavior<StupidoBhv> implements IScheduledBehavior<StupidoBhv> {

    /**
     * A reference to the container context
     */
    protected StupidoBhvContext container;

    public StupidoBhv(Farmer owner, StupidoBhvContext c) {
        super(owner);
        this.container=c;
    }

    private int stupidoProperty;


    @ScheduledMethod (start=2,interval = 2)
    public void setRandom() {
        this.stupidoProperty =  this.container.getRandom();
    }

    @ScheduledMethod (start=2,interval = 1)
    public void print() {
        System.err.println("Farmer, id="+this.owner.getID() + ", stupido random=" + this.stupidoProperty);
    }

    @Override
    public Object getAnnotatedClass() {
        return this;
    }

}
```

Figure 30, The StupidoBhv.java code

AFarmerBehavior.java

```java
package gr.agroscape.behaviors.farmers;

import gr.agroscape.agents.Farmer;

 * Every "Farmer Behavior" object contains a reference to the "Farmer" object that cont
public abstract class AFarmerBehavior<T> implements IScheduledBehavior<T> {

    /**
     * A reference to the owner Farmer
     */
    protected Farmer owner;


    /**
     * Constructor
     * @param owner
     */
    public AFarmerBehavior(Farmer owner) {
        this.owner = owner;
    }


    /**
     * Getter
     * @return
     */
    public Farmer getOwner() {
        return owner;
    }

}
```

Figure 31, The AFarmerBehavior source code