

Dipartimento di Impresa e Management

Cattedra di Metodi Statistici per il Marketing

Olio di oliva, una commodity mediterranea

Modelli previsionali

RELATORE

Prof. Pierpaolo D'Urso

CANDIDATO

Gabriele Varia

Matr. 672571

CORRELATRICE

Prof.ssa Livia De Giovanni

ANNO ACCADEMICO

2016/2017



Olive Trees with Yellow Sky and Sun. Vincent Van Gogh nov. 1889

INDICE

INTRODUZIONE.....	6
CAPITOLO I Il settore dell'olio di oliva.....	9
1. Struttura e analisi del mercato dell'olio di oliva.....	9
1.1 Evoluzione storica della coltura.....	11
1.2 Quadro normativo ed enti regolatori del mercato.....	12
1.3 Consumi, importazioni ed esportazioni globali.....	14
1.4 Struttura del mercato e centri produttivi – focus sui principali luoghi di produzione.....	18
1.5 Catena del valore e formazione del prezzo finale.....	20
1.6 Analisi della struttura e andamento dei prezzi all'ingrosso.....	23
CAPITOLO II Agribusiness: ruolo dei modelli previsionali in agricoltura.....	27
2. Agribusiness e mercati delle Commodities agricole.....	27
2.1 Introduzione all'Agribusiness.....	28
2.2 Commodities agricole - focus sugli oli vegetali.....	30
2.3 Ruolo dei modelli predittivi nel settore agricolo	31
2.4 Evoluzione ed introduzione di sistemi di ICT in agricoltura.....	33
CAPITOLO III Modelli previsionali.....	35
3. Modelli Predittivi sulla produzione di olio di oliva nel mercato spagnolo.....	35
3.1 Composizione e analisi primaria dei dati.....	36
3.2 Analisi dei dati metereologici.....	36

3.3	Analisi dei dati sulla produzione.....	38
3.4	Analisi delle correlazioni.....	40
3.5	Alberi decisionali CART.....	47
3.6	Modelli previsionali Xgboost (parametri).....	48
3.7	Dataset definitivo Xgboost.....	49
3.7.1	Modello 00	50
3.7.2	Modello 01.....	53
3.7.3	Modello 02.....	56
3.8	Ruolo critico dei modelli previsionali per le imprese operanti nel mercato dell'olio d'oliva.....	60
3.9	Gestione e riduzione dei rischi legati all'incertezza per le imprese agricole.....	61
3.10	Gestione e riduzione dei rischi legati all'incertezza per i trasformatori e imbottiglieri.....	62
3.11	Commenti finali.....	63
BIBLIOGRAFIA.....		64
SITOGRAFIA.....		66
INDICE DELLE TABELLE		
Tabella 1	“Importazioni globali di olio di oliva e sansa”	15
Tabella 2	“ Esportazioni globali di olio di oliva e sansa”.....	15
Tabella 3	“Variabili meteo”.....	38
Tabella 4	“ Dati primari produzione”	38
Tabella 5	“Copertura campione rispetto alle quantità di olio di oliva prodotto”.....	39
Tabella 6	“variabili sulla produzione e sulle superfici coltivate”	40
Tabella 7	“Coefficienti di correlazione lineare produzione di olive (pvalue 0,001)”.....	41
Tabella 8	“Coefficienti di correlazione lineare resa su ettari irrigati (p value 0,001)”.....	41
Tabella 9	“Superfici irrigate, non irrigate e totali per provincia 2016”.....	42
Tabella 10	“Coefficienti di correlazioni lineare resa su ettari non irrigati (p value 0,01)”	44
Tabella 11	“Indice di correlazione di Pearson resa su ettari non irrigati/Jaén (p value 0,01)”.....	46
Tabella 12	“Indice di correlazione di Pearson resa su ettari irrigati/Jaén (p value 0,01)”.....	46

Tabella 13 “Test modello 00”	51
Tabella 14 “Test modello 01”	54
Tabella 15 “Test modello 02”	57
Tabella 16 “Confronto campagne 2014/15-2015/16”	60

INDICE DELLE FIGURE

Figura 1 “Produzione globale di olio di oliva”	8
Figura 2 “Produzione e consumo globale di olio di oliva”	14
Figura 4 “Reddito per azienda agricola”	18
Figura 3 “Produttività (output totale/input totale)”	18
Figura 5 “Sussidi alla produzione”	18
Figura 6 “Costo di produzione di 1 litro di olio d’oliva per paese”	18
Figura 7 “Catena del valore del prezzo”	20
Figura 8 “Prezzi medi olio extravergine di oliva nei mercati europei più rappresentativi”	23
Figura 9 “Evoluzione dei prezzi dell’olio vergine di oliva”	24-25
Figura 10 “Evoluzione dei prezzi dell’olio lampante”	25-26
Figura 11 “Stadi di evoluzione dell’introduzione di un’innovazione in agricoltura”	33
Figura 12 “Correlazione tra produzione di olive e superfici irrigate per provincia”	43
Figura 13 “Correlazione tra produzioni di olive e superfici non irrigate per provincia”	45
Figura 15 “Overfitting: evoluzione grafica del modello 00”	50
Figura 14 “complessità modello 00”	50
Figura 16 “Importanza features 00”	52
Figura 17 “Complessità modello 01”	53
Figura 18 “Overfitting evoluzione grafica modello 01”	53
Figura 19 “Importanza feature modello 01”	55
Figura 20 “Complessità modello 02”	56
Figura 21 “Overfitting evoluzione grafica del modello 02”	56
Figura 22 “Importanza features 02”	58
Figura 23 “Alberi regressivi modello 02”	59

INTRODUZIONE

Il Mediterraneo è il principale luogo di coltivazione e produzione dell'olio di oliva, tra Spagna, Italia e Grecia viene prodotto il 95% dell'olio di oliva europeo. La Spagna è il paese leader del mercato con il 50% della produzione globale.

“I modelli predittivi sulle quantità di olive prodotte sono indispensabili per l'ottimizzazione delle risorse tecniche e umane impiegate nel raccolto e per la pianificazione globale del marketing e della distribuzione commerciale dell'olio di oliva.”¹ Un buon sistema di *forecasting* fornisce una stima sulle quantità che verranno prodotte nelle campagne future, incrociando questo risultato con le previsioni sulla domanda globale, si ricava una stima sui volumi che verranno scambiati. La domanda globale nell'ultimo quinquennio (2013-2018) è rimasta stabile oscillando intorno ai 3 milioni di tonnellate annue, di contro la produzione globale nell'ultimo quinquennio ha alternato campagne fruttuose a campagne meno proficue, si veda figura 1 p.8.

L'olio di oliva viene scambiato nel mercato *business to business* (B2B) nei pressi dei principali luoghi di produzione e il prezzo di riferimento viene indicato dalle camere di commercio locali. Durante la disastrosa campagna 2014-15 nei due mesi precedenti all'inizio del raccolto, il prezzo dell'olio di oliva a Jaén in Spagna è aumentato del 37% in soli 2 mesi. Nella campagna successiva durante lo stesso periodo, invece, è stata registrata una riduzione del prezzo pari al 36%. Il prezzo dell'olio di oliva viene a formarsi per il 70%² durante la fase di coltivazione delle olive, rendendo i coltivatori i più esposti alla volatilità del prezzo della materia prima. Ad ogni modo, tutte le imprese del settore risentono gli shock di prezzo, per questo stime affidabili sul volume del raccolto sono una crescente necessità per le imprese che sono esposte alla volatilità del prezzo dell'olio.³ Le aziende che hanno previsto il mercato tempestivamente possono immagazzinare prodotto finito e stipulare contratti di fornitura futura in modo da ottimizzare la propria performance.

I modelli predittivi per stimare la resa degli olivi impiegando dati metereologici e sulle superfici coltivate, sono stati sviluppati in collaborazione dalle Università di Cordoba, Siviglia e Perugia, che tra il 1998 e il 2015 hanno pubblicato stime su scala locale e nazionale. Il ciclo biologico dell'olivo, lo stato di salute delle piante e le variabili metereologiche condizionano le quantità prodotte ogni anno.

In questo lavoro di ricerca sono stati realizzati 3 modelli, costruiti sotto forma di alberi decisionali per stimare le produzioni di olio di oliva nel mercato spagnolo. L'obiettivo dei modelli è predire la quantità di olio prodotto

¹ Oteros & al, e. 2014. Testo originale “Early and effective crop forecasting is essential for optimization of technical and human resources for olive harvesting and for planning global marketing and commercial distribution of olive oil.”

² Governo Spagnolo Ministero dell'Ambiente . (2010). *The value chain and price in spanish olive oil industry* .

³ Galan, & al, e. 2004

nelle 13 province spagnole che costituiscono il campione di studio. Campione, che ha una copertura media del 93% rispetto ai volumi prodotti nel mercato spagnolo durante i 20 anni presi in analisi (1995-2015).

I dati relativi alle superfici coltivate e alle produzioni di olio di oliva per singola provincia sono stati ricavati dagli annuari dell'Istituto di Statistica Spagnolo. Le variabili metereologiche che influenzano la produzione sono state scaricate dall'open source dell'ECAD un database finanziato dall'UE che monitora 10.584 stazioni metereologiche nel mondo.

L'indice di correlazione di Pearson mostra un forte legame lineare, pari a 0,80 tra produzione e superfici irrigate coltivate ad olivo, e un indice dello 0,86 tra la produzione e le superfici non irrigate coltivate.

Di contro la correlazione tra la produzione e le variabili meteorologiche risulta debole. Per risolvere il problema della scarsa correlazione tra le variabili metereologiche e la produzione sono stati impiegati strumenti di machine learning a scopo regressivo. Per eseguire l'analisi è stato usato il software R (3.4.3) in cui è stato avviato l'algoritmo xgboost, uno dei metodi più utilizzati per risolvere problemi di classificazione o di regressione in data science. Come in tutti i metodi di machine learning il dataset finale composto da 455 osservazioni di 120 variabili è stato suddiviso in set di apprendimento (1995-2010) e in test set (2011-15).

Il Set di apprendimento viene analizzato in alberi di classificazione e regressione con il metodo di split che minimizza lo scarto quadratico medio dell'errore per predire la funzione obiettivo: ("olive_for_hectar") resa di olive in kg/ha.

I risultati ottenuti nel modello finale indicano un coefficiente di determinazione R^2 (0,98) con un errore assoluto del 2% nel quinquennio complessivo (2011-2015), preso in considerazione per effettuare il test.

L'elaborato è strutturato in 3 capitoli. I primi due capitoli svolgono una funzione introduttiva ai fini dell'analisi dei dati e delle conclusioni che ne derivano. Nel primo capitolo: "Il settore dell'olio di oliva" viene fornita una panoramica sulla struttura del mercato e sulla formazione del prezzo. Nel secondo capitolo vengono introdotti i concetti di Agribusiness, il ruolo delle previsioni e del machine learning per l'ottimizzazione di processi nei mercati delle commodities agricole. Nel terzo capitolo sono stati presentati i tre modelli costruiti per stimare le quantità di olio prodotto, i risultati ottenuti e i commenti finali.

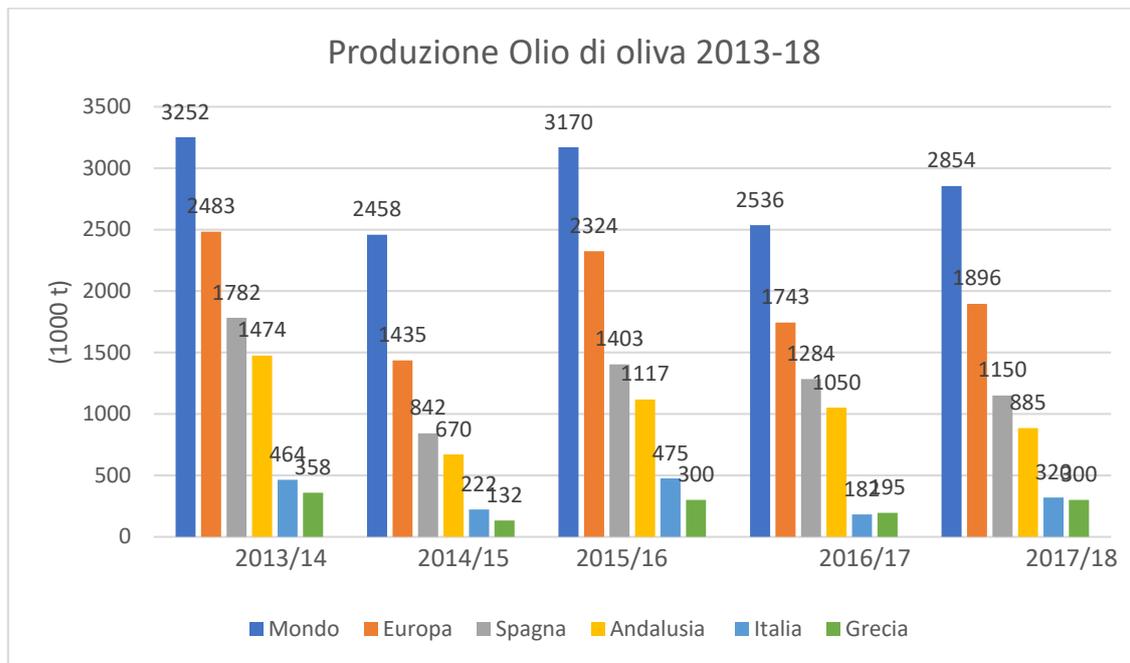


Figura 1 “Produzione globale di olio di oliva”

CAPITOLO I: Il settore dell'olio di oliva

1. Struttura e analisi del mercato dell'olio di oliva

L'olio d'oliva è un olio vegetale ottenuto per mezzo della spremitura delle olive e viene suddiviso in 9 categorie, di cui 4 oli extravergine di oliva.

- L'olio extravergine d'oliva (evo), ottenuto attraverso la spremitura a freddo o meccanica delle olive fresche e ben conservate senza altre operazioni eccetto la decantazione e la centrifuga. L'olio evo mantiene le caratteristiche organolettiche originarie del frutto ed è caratterizzato da una quantità di acido oleico inferiore a 0,8/1 g per 100 g.
- L'olio vergine d'oliva, contiene una soglia inferiore al 2% di acidità.
- l'olio di oliva vergine corrente con una soglia inferiore al 3,3%.
- l'olio vergine lampante contenente una quantità di acido oleico superiore al 3,3% non adatto all'uso alimentare ma utilizzabile per la raffinazione di oli commestibili e per usi non alimentari.⁴

L'olio d'oliva viene scambiato nei mercati locali in cui sono situate le principali colture e i frantoi in cui avviene la trasformazione da materia prima a prodotto finito. I prezzi sono fissati dalle camere di commercio locali o dalle associazioni di categoria e dipendono da diversi fattori che analizzeremo in seguito. L'anno commerciale nel mercato dell'olio inizia il 1° novembre e si conclude il 31 ottobre dato che le olive vengono raccolte, raffinate, confezionate durante la parte finale dell'anno solare, inoltre, vengono immagazzinate e vendute entro 18 mesi di tempo.

Il mercato dell'olio di oliva è una delle principali risorse economiche delle zone rurali del mediterraneo. La filiera produttiva dell'olio è composta da produttori di materia prima, dai trasformatori: frantoi, raffinerie e dalle imprese imbottigliatrici che distribuiscono e commercializzano il prodotto.

La Commissione Europea e il Consiglio Oleico internazionale (COI) sono le principali istituzioni che regolano il mercato globale e forniscono dati aggregati su produzione, consumo, importazioni ed esportazioni.

La campagna olearia globale 2017/2018 dovrebbe registrare un aumento del 16% rispetto alla campagna precedente. Nell'anno commerciale 2016/17 infatti erano stati prodotti 2,5 milioni di tonnellate di olio d'oliva, 300 mila tonnellate in meno rispetto alla media degli ultimi 5 anni (2,86 milioni di tonnellate). Nel quinquennio 2013/14-2017/18, sia in Europa che su scala globale, i massimi sono stati registrati nel primo anno del quinquennio (2013/14), anno in cui la produzione internazionale ha ampiamente superato 3 milioni di tonnellate, mentre il minimo è stato registrato nell'anno successivo. Nella campagna 2014/15 in Europa vi è stato una riduzione pari a 1 milione di tonnellate di olio di oliva che corrispondono al 42% in meno rispetto a

⁴ (Della Porta, 2013)

quanto prodotto nell'anno precedente (2013/14). (European C., 2017) (OIC; International Olive Council, 2017)

L'Unione Europea negli ultimi 5 anni commerciali (2013/14-2017/18) risulta il principale mercato dell'olio di oliva sia per quantità prodotte, il 70 % della produzione, sia per quanto riguarda il consumo che corrisponde al 56% e per le esportazioni che contano il 66% su scala globale. L'olio d'oliva, infatti, è la base della dieta mediterranea, il simbolo e allo stesso tempo una delle principali fonti di sostentamento delle aree rurali. (European C., 2017) Spagna, Italia e Grecia contano il 95% della produzione europea dell'olio di oliva. La Spagna è il paese leader del mercato con il 50% della produzione internazionale e il 73% della produzione all'interno dell'Unione Europea. La domanda d'oliva sta sperimentando una crescita costante, considerando la consapevolezza dei consumatori riguardo le caratteristiche e i benefici generati dall'utilizzo dell'olio evo in campo alimentare (Intl olive Council, 2015).

Diverse ricerche scientifiche hanno evidenziato le proprietà organolettiche dell'olio extravergine di oliva sottolineando le proprietà antiossidanti dei fenoli, i benefici del grasso vegetale sull'epidermide e la riduzione della glicemia nel sangue, associate al consumo di una dose di 10g a pasto. (Violi, Loffredo, & Pignatelli, 2015). L'aumento della domanda di biocombustibili nell'ultimo quinquennio ha comportato l'innalzamento del prezzo della sansa. L'olio di sansa di olive viene infatti ottenuto dai residui della polpa e i frammenti di nocciolino. La sansa e i materiali di scarto della produzione dell'olivo infatti possono essere utilizzati per produrre bio-combustibili rendendo quindi un materiale di scarto una risorsa economica.

Ad ogni modo stando ai report della Food Agricultural Organization (FAO) in collaborazione con l'Organizzazione per la Cooperazione e lo Sviluppo (OECD), la domanda delle commodity dei biocombustibili, dopo un decennio di alti tassi di crescita seguirà un trend costante con margini più ridotti per i rivenditori.

L'ampiezza della distanza tra la domanda potenziale e la domanda primaria del mercato, implica ampi margini di manovra per la riduzione dei diversi gap. Focalizzandoci sull'uso alimentare il gap di *promotion* è quello su cui poter operare in maniera più efficace ed in tempi ridotti, in particolare nei paesi non produttori, in cui il target di consumatori con reddito medio/alto e abitudini alimentari salubri e ricercate ha la massa critica da giustificare gli investimenti pubblicitari. (Mili & Zuniga, 2001)

1.1 Evoluzione storica della coltura

L'origine dell'olio di oliva, coincide con l'espansione della civiltà nel Mediterraneo, la quale ha governato il destino dell'umanità per secoli ed ha segnato profondamente la cultura occidentale. Le origini della pianta *Olea Europea* possono essere rintracciate tra la costa est del Mediterraneo: parte meridionale della penisola anatolica, Siria, Libano, Palestina e Israele stando ai ritrovamenti di tavolette incise, scavi e frammenti di legno trovati nelle tombe antiche. Fossili contenenti tracce di olivo risalenti al Pliocene (1-1,5 milioni di anni fa) furono rinvenute a Mongardino in Piemonte, altri reperti risalenti al paleolitico furono trovati in Nord Africa e in Spagna. L'esistenza dell'albero d'olivo è, quindi, attestata a 12 mila anni fa. (Mattingly, 1988)

L'olivo selvatico in Asia minore è estremamente diffuso e cresce in fitte foreste. Esso si è esteso dalla Siria fino in Grecia attraverso la penisola anatolica. “Nel sedicesimo secolo a. C. i Fenici iniziarono a disseminare olivi nelle isole greche e tra il quattordicesimo e il dodicesimo secolo giunsero nella penisola ellenica.”⁵ La coltivazione dell'olivo è testimoniata a sei mila anni fa da parte delle popolazioni indigene del Mediterraneo e dell'Asia minore. (Luchetti, 2008) In un antico documento rinvenuto in Siria, riferibile al 2000 a.C., l'olio d'oliva aveva un valore cinque volte superiore a quello del vino e due volte e mezzo il valore dell'olio di semi. La diffusione dell'olivo nel commercio è documentato nei suoi primi movimenti verso occidente, i registri indicano l'introduzione dell'*olea europea* in Grecia, in Egitto e nella parte occidentale della Turchia. In queste aree furono ritrovati reperti riconducibili all'olivo in siti archeologici. Nel Palazzo di Cnosso nell'isola di Creta, iscrizioni possono far ricondurre al commercio dell'olio al 1700 a.C. Per quanto riguarda i frantoi vicino Izmir in Turchia un antico processo di trasformazione è datato 600 a.C. Nel quarto secolo a.C. dato che il surplus di produzione rispetto alla domanda interna permise al legislatore Solone di emanare un regolamento per esportare olio. A tale proposito sappiamo che Platone, approfittando di questo permesso, allorché intraprese un viaggio in Egitto, non sdegnò di portare con sé un carico di olio, nella certezza di poterlo vendere a buone condizioni. Lo stesso termine olio deriva dal greco *elaia* conosciuto fin dall'antichità come simbolo di pace e buona volontà. Dal sesto secolo a.C. in poi l'*olea europea* si propagò nei paesi confinanti con il Mediterraneo, gli alberi furono piantati a Tripoli, in Tunisia e in Sicilia prima di arrivare nel sud dello stivale. Secondo la tradizione il primo albero di olivo fu portato durante il regno di Lucio Tarquino Prisco il vecchio (616-578 a.C.) quinto re di Roma probabilmente dal nord Africa. (Vossen, 2007) Come pure è noto che, più tardi, gli agenti dell'imperatore Adriano, convinti che, in ragione della sua abbondanza, l'olio di oliva fosse, come il sale, una delle sostanze che avrebbero potuto rendere maggiormente al fisco, non tardarono a tassarlo. I Romani continuarono l'espansione dell'olivo nei paesi conquistati utilizzando il ramoscello d'olivo in segno di pace, la

⁵ (Luchetti, 2008) testo originale - In the 16th century B.C. the Phoenicians were take the olive tree to the Greek islands: later, between the 14th and the 12th centuries B.C., they introduced it to the Greek mainland where it spread.

coltivazione arrivò in Gallia attorno al 600 a. C., in Spagna invece oggi leader mondiale della produzione di olio d'oliva, la coltura era stata trasportata dai Fenici mille anni prima di Cristo, tuttavia, solo dopo la terza guerra punica fu conosciuta in tutta la penisola iberica. Gli arabi portarono le loro varietà e influenzarono il modo in cui venivano coltivate le olive e anche le parole olivo (aceituna) e olio (aceite), in spagnolo derivano dall'arabo. Le olive venivano raccolte, a seconda dell'uso cui erano destinate, in periodi diversi: ancora acerbe (olive albae o acerbae), non del tutto mature (olive variae o fuscae), mature (olive nigrae). Si raccomandava di staccarle dal ramo con le mani ad una ad una; quelle che non si potevano cogliere salendo sugli alberi, venivano fatte cadere servendosi di lunghi bastoni flessibili (in greco ractriai), sempre ponendo la massima attenzione a non danneggiarle. Alcuni aiutanti raccattavano e riunivano le olive battute che, solitamente venivano macinate il più presto possibile. Con la scoperta dell'America l'olivo fu coltivato anche lontano dal Mediterraneo. I primi olivi furono trasportati da Siviglia alle indie occidentali e quindi raggiunsero il continente americano. Nel 1560 l'olivo era coltivato in Messico, dopo in Perù, California, Cile e Argentina in cui è possibile trovare l'Old Arauco un albero secolare ancora in vita oggi. In tempi recenti gli olivi vengono piantati in paesi distanti dal luogo d'origine come in Sud Africa, Australia, Giappone e Cina. (Vossen P. 2005, 2007)

1.2 Quadro normativo ed enti regolatori

L'olio di oliva come già indicato è regolato dal Consiglio Oleico Internazionale (COI) organismo intergovernativo fondato il 1959 per volontà delle Nazioni Unite. Il COI riunisce produttori, consumatori e gli addetti alla filiera produttiva dell'olio di oliva al livello internazionale. Ha sede a Madrid e si occupa di incrementare il commercio internazionale dell'olio di oliva, fornendo dati, ricerche e studi di settore. Il COI si occupa anche di promuovere il prodotto e di mantenere gli standard qualitativi necessari, oltre a riunire i rappresentanti dei principali paesi aderenti per discutere di eventuali criticità del settore.

Nel 2005 è stato emanato un regolamento entrato effettivamente in vigore nel 2007 che supera i precedenti accordi internazionali stipulati nel 1956, 1979, 1986. Nell'accordo viene evidenziata l'importanza dell'aspetto qualitativo del prodotto come caratteristica chiave per promuovere il consumo di olio di oliva e olive da tavola, aspetto innovativo rispetto ai precedenti accordi, furono anche introdotte norme riguardanti lo smaltimento dei rifiuti e la sostenibilità dell'olivicoltura.

Il COI è strutturato come una tipica agenzia delle Nazioni Unite sulla sicurezza alimentare. Vi sono tre organi: il consiglio dei membri, il presidente ed il segretariato esecutivo.

Il consiglio costituito da un componente per ogni stato membro approva annualmente il bilancio dell'organizzazione e ne valuta le attività.

Il presidente sovrintende le riunioni ed ha responsabilità legale dell'organizzazione. Ha un ruolo centrale il segretariato esecutivo, un organo collegiale che svolge le attività operative del consiglio, è composta da quattro sezioni operative che svolgono attività diverse ma complementari tra loro: divisione tecnica, divisione studi e valutazioni, divisione promozione, divisione finanziaria.⁶

Dato che l'Unione Europea è il primo produttore, consumatore ed esportatore di olio di oliva, nel quadro della Politica Agricola Comune (PAC), regola la produzione e gli aiuti verso i coltivatori, fornisce gli standard qualitativi, e regola il commercio interno e le esportazioni.

Il PAC detiene due principali finalità riscontrabili nell'intervento attivo delle istituzioni nell'attività produttiva degli agricoltori, che consiste nell'imposizione di un prezzo minimo e nella spinta verso l'utilizzo di tecnologie e sistemi integrati con lo scopo di aumentare la produttività e ridurre l'impatto ambientale.

Nel regolamento (CE) n. 1638/98 furono apportati dei cambiamenti fondamentali nell'erogazione dei sussidi ai coltivatori. Infatti, nel regolamento precedente n 136/1966 venivano forniti aiuti alla produzione fissi per i piccoli produttori. Nella riforma del 1998 vengono ridotti gli strumenti di aiuto, al fine di evitare frodi, ma vengono aumentate le quantità massima garantita da 1,35 milioni a 1,78 milioni di tonnellate di olio di oliva. Inoltre, vengono inseriti i concetti di Quantità Nazionale Garantita (QNG) ovvero il tetto massimo degli aiuti che vengono forniti ai paesi produttori. Entrano a far parte di questi aiuti alla produzione quindi Spagna, Italia, Grecia, Portogallo e in maniera ridotta la Francia. Qualora uno stato membro non dovesse raggiungere la QNG, il 20% della quota non utilizzata potrà essere impiegata a compensazione degli altri stati membri che hanno sfiorato la propria QNG mentre il restante 80% potrà essere riutilizzato per la prossima campagna olivicola.

Regolamento UE n. 61/2011 è l'ultimo aggiornamento delle composizioni chimiche della descrizione e definizione degli oli di oliva e degli oli di sansa di oliva precedentemente inseriti nel CE n. 1513/2001. La molitura di olive in cattivo stato di conservazione infatti provoca un aumento degli alchil esteri informazione supplementare che non era definita precedentemente nella scheda di descrizione e definizione. Quindi nell'interesse dei consumatori è stato introdotto questo nuovo parametro obbligatorio che deve essere analizzato prima che l'olio di oliva possa entrare in commercio.

Al livello nazionale in Italia il Ministero delle politiche agricole ambientali e forestali Mi.P.A.A.F. ha un ruolo strategico nel rappresentare l'Italia nella direzione generale dell'agricoltura e dello sviluppo delle aree rurali della Commissione Europea.⁷

Le questioni riguardanti imballaggi, designazione di origine, registri ed identificazioni delle partite sono di competenze del Mi.P.A.A.F.

⁶ <http://www.internationaloliveoil.org/estaticos/view/102-structure>

⁷ <https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/1019>

Il packaging dell'olio di oliva vendibile al dettaglio deve essere chiuso ermeticamente e in contenitori non eccedenti i 5 litri mentre nel caso di utilizzo in luoghi di ristoro quindi in caso di vendita B2B (business to business) il packaging può arrivare anche a 25 litri. Sul packaging vi è l'obbligo di inserimento della descrizione e provenienza della materia prima utilizzata oltre all'inserimento nel documento di accompagnamento. Le descrizioni possono essere:

- a) miscela di oli di oliva comunitari;
- b) miscela di oli di oliva non comunitari;
- c) miscela di oli di oliva comunitari e non comunitari.

Frantoi e Raffinerie detengono l'obbligo di annotare nei registri movimenti e lavorazioni. Le identificazioni delle partite concernono anche gli stock in magazzino i quali devono annotare il lotto, la quantità di bottiglie/Pet e la capacità di quest'ultimi.

All'interno del sistema SIAN sistema informativo agricolo nazionale in cui le imprese devono indicare data di inizio della lavorazione quantità prodotte e cessazione delle attività così da permettere un controllo da parte delle istituzioni sul mercato.

1.3 Consumo importazioni ed esportazioni

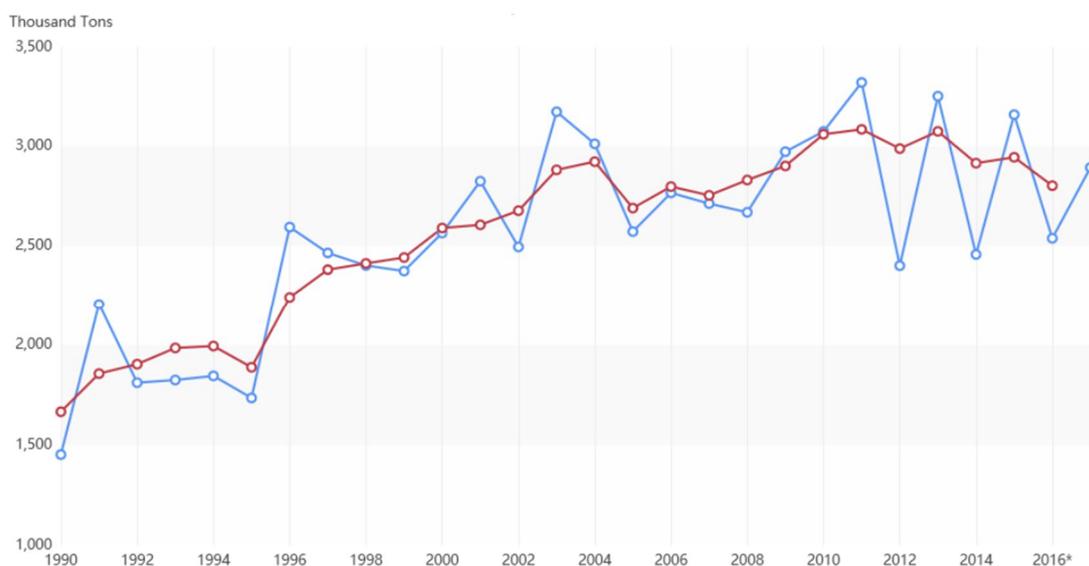


Figura 2 “Produzione e consumo mondiale di olio di oliva” fonte: <https://www.oliveoiltimes.com/olive-oil-business/world-olive-oil-production-figures-2017-show-mixed-results/58946>

-Produzione

-Consumo

Tabella 1 “Importazioni globali di olio di oliva e sansa”

Importazioni	Tonnellate (1000t)		Var. %	Milioni di €		Var. %
	2015	2016		2015	2016	
Anno	2015	2016	Var. %	2015	2016	Var. %
Italia	583	572	-1,9%	1859	1792	-3,6%
Stati Uniti	313	332	6,1%	1100	1193	8,5%
Spagna	251	166	-33,9%	631	358	-43,3%
Francia	115	122	6,1%	434	463	6,7%
Portogallo	106	97	-8,5%	317	289	-8,8%
Germania	71	72	1,4%	292	303	3,8%
Giappone	60	58	-3,3%	274	263	-4,0%
Brasile	58	57	-1,7%	249	255	2,4%
Regno Unito	75	76	1,3%	257	253	-1,6%
Cina	39	45	15,4%	160	178	11,3%
Canada	38	44	15,8%	151	180	19,2%
Svizzera	14	15	7,1%	75	78	4,0%
Australia	20	31	55,0%	73	114	56,2%
Belgio	19	24	26,3%	73	87	19,2%
Messico	18	19	5,6%	65	68	4,6%
Altri	165	244	47,9%	753	0	
Mondo	1945	1974	1,5%	6762	6774	0,2%

Tabella 2 “Esportazioni globali di olio di oliva e sansa”

Esportazioni	Tonnellate (1000t)		Var. %	Milioni di €		Var. %
	2015	2016		2015	2016	
Anno	2015	2016	Var. %	2015	2016	Var. %
Spagna	845	1008	19,3%	2837	3305	16,5%
Italia	362	398	9,9%	1530	1617	5,7%
Tunisia	258	108	-58,1%	763	351	-54,0%
Grecia	203	198	-2,5%	675	630	-6,7%
Portogallo	129	131	1,6%	443	426	-3,8%
Argentina	31	16	-48,4%	101	54	-46,5%
Marocco	34	26	-23,5%	79	59	-25,3%
Turchia	14	19	35,7%	57	66	15,8%
Cile	15	11	-26,7%	54	42	-22,2%
Francia	11	12	9,1%	51	51	0,0%
Stati Uniti	12	13	8,3%	36	37	2,8%
Australia	6	4	-33,3%	24	13	-45,8%
Germania	5	5	0,0%	22	22	0,0%
Belgio	4	6	50,0%	16	22	37,5%
Regno Unito	3	2	-33,3%	12	11	-8,3%
Altri	14	18	28,6%	61	68	11,5%
Mondo	1946	1975	1,5%	6761	6774	0,2%

Il consumo annuale di olio oscilla intorno ai 3 milioni di tonnellate annui prendendo in considerazione l'ultimo quinquennio.⁸ Circa metà del consumo globale avviene all'interno dell'Unione Europea, i paesi produttori sono anche i primi consumatori di olio ma anche Germania, Gran Bretagna e Belgio hanno una domanda interna da non trascurare tra i paesi europei non produttori. (European C., 2017)

Tra gli oli vegetali, l'olio di oliva rappresenta una nicchia del mercato dato che vengono consumati annualmente 60 milioni di tonnellate di olio di palma, 55 milioni di tonnellate di olio di semi di soia e poco meno di 30 milioni di tonnellate di olio di colza.⁹

L'olio di oliva è il grasso vegetale con il prezzo maggiore sia all'ingrosso che al dettaglio e ha un target selezionato, di élite nei paesi lontani dal mediterraneo. L'Italia pur essendo uno dei principali produttori è il primo importatore di olio al mondo. Le importazioni italiane sono principalmente intracomunitarie e provengono dalla Spagna e dalla Grecia, ma anche le importazioni dalla Tunisia, per motivi di vicinanza geografica e di competitività del prezzo hanno procurato l'attenzione del governo italiano che ne ha contingentato le importazioni.¹⁰

La bilancia commerciale italiana nel mercato dell'olio di oliva mostra un saldo negativo di 329 milioni di euro nel 2015 e di 175 milioni di Euro nel 2016. Il saldo delle quantità è anch'esso negativo con 395 mila tonnellate di differenza tra importazioni ed esportazioni nel biennio (2015-16). Questi dati indicano che l'Italia è il primo mercato per consumo.

La Spagna è il primo esportatore al mondo di olio di oliva con 800 mila tonnellate nel 2015 e 1 milione di tonnellate nel 2016. L'Italia è il primo importatore dalla Spagna seguito dagli Stati Uniti.

La domanda primaria di olio d'oliva sia in Europa che nel mondo segue un andamento costante e risente in modo limitato delle oscillazioni delle quantità di olio prodotto annualmente. Tra i paesi europei la Grecia è il paese che ha il consumo pro capite maggiore seguono Italia e Spagna. Nei paesi lontani dal mediterraneo gli Stati Uniti sono i primi importatori e consumatori di olio di oliva con 300 mila tonnellate annue. Altri paesi come Canada, Brasile, Giappone e Australia hanno un consumo annuo in crescita e che si attesta intorno alle 20 mila tonnellate annue.

La stima della domanda primaria è infatti in grandi linee prevedibile basti osservare i dati storici congiuntamente ai trend locali e globali. Risulta più impegnativo stimare l'offerta con mesi di anticipo

⁸ (European C. , 2017) Market situation in the Olive Oil and Tables olives sectors. *Market situation in the Olive Oil and Tables olives sectors*. Bruxelles. p. 11

⁹ <https://www.statista.com/statistics/263937/vegetable-oils-global-consumption/>

¹⁰ <http://www.sviluppoeconomico.gov.it/index.php/it/commercio-internazionale/2021203-settore-olio-importazioni-di-olio-di-oliva-dalla-tunisia-in-regime-di-contingente-tariffario>

rispetto all'inizio del raccolto dato che le colture sono sensibili a shock climatici e scarsità d'acqua che possono ridurre sensibilmente il raccolto delle olive.

Con mercato potenziale si intende il limite teorico a cui la domanda primaria può tendere in una determinata finestra temporale. L'aumento della domanda primaria è strettamente legato all'aumento del consumo nei paesi con alti margini di crescita economica e demografica come India, Cina e Brasile e da paesi come USA e Giappone in cui vi è un'alta concentrazione di clienti target. (Luchetti, 2008)

L'olio evo nei paesi lontani dal mediterraneo è un prodotto destinato all'élite, riservata alla nicchia di individui abituati a condurre uno stile di vita mediterraneo. Considerando un orizzonte temporale più ampio la domanda è destinata a crescere maggiormente dato l'ampliamento dei terreni impiegati per la coltura oggi pari a 10 milioni di ettari. I paesi con la crescita più significativa sono Giappone, USA, Australia, Canada, Francia e i paesi dell'Unione Europea non produttori di olio. In Giappone i 125 milioni di abitanti hanno mediamente redditi alti e una buona conoscenza dei valori nutrizionali e delle diverse proprietà degli alimenti unitamente all'aumento degli scambi agroalimentare con gli altri paesi del Pacifico è considerato il mercato che offre maggiori profitti agli esportatori oltre allo sviluppo dell'utilizzo dell'olio d'oliva in mercati diversi da quello agroalimentare come il settore cosmetico del beauty care e la produzione di biocarburante. Imprese multinazionali hanno lanciato linee di prodotti quali shampoo, bagnoschiuma, saponi, lozioni decantando gli effetti positivi del grasso vegetale. Per produrre i prodotti di cosmetica vengono solitamente utilizzate le oleine delle sostanze che vengono prodotte durante il ciclo produttivo dell'olio. Un altro utilizzo non alimentare che potrebbe condizionare la domanda è l'utilizzo degli scarti della raffinazione per la produzione del biodiesel e di altri carburanti. (Dorado, Ballesteros, Arnal, & Gomez, 2003) La sansa oltre ad essere raffinata per produrre olio di sansa, un olio vegetale commestibile prodotto dagli scarti del ciclo produttivo degli oli di oliva vergini, può essere utilizzato come combustibile per uso industriale o domestico. La sansa è una biomassa umida solitamente conta tra il 30% e il 35% della massa totale adoperata per la trasformazione da materia prima a olio finito. Del totale della sansa è necessario avviare un processo per setacciare il nocciolino dal resto della massa.

1.4 Struttura del mercato e centri produttivi – focus sui principali produttori e luoghi di produzione

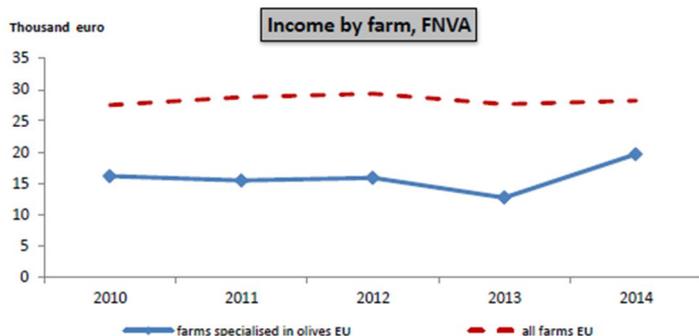


Figura 4 “Reddito per azienda agricola”

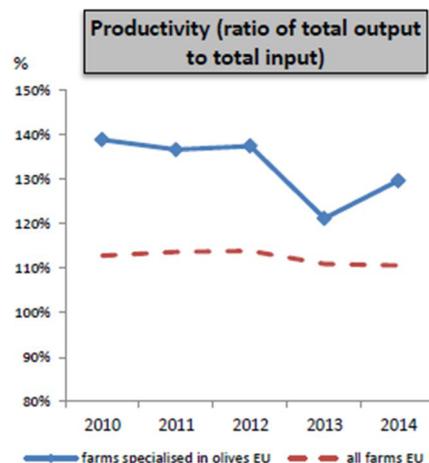


Figura 3 “Produttività (output totale/input totale)”¹

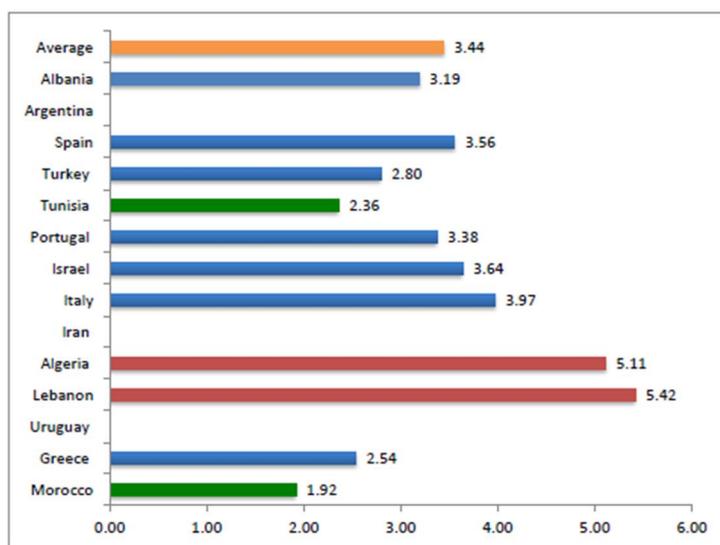


Figura 5 “Costo produzione 1 litro di olio d'oliva per paese”

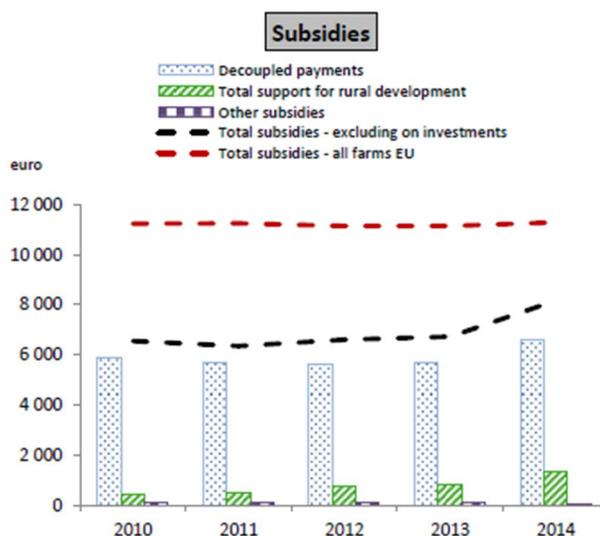


Figura 6 “Sussidi alla produzione”

Fonte (Commissione Europea, 2014) FARMS SPECIALISED IN OLIVES IN THE EU Based on FADN data up to 2014 fig.3 e 4 p. 1, fig.5 p. 2 fig.6 p. 4

Le principali colture di olivi sono situate nell'area Mediterranea tra Spagna, Italia, Grecia, Portogallo, Marocco, Tunisia. In questi 6 paesi, vengono coltivati 8 milioni di ettari dei 10 milioni coltivati su scala globale. (OIC; International Olive Council, 2016)

L'Italia è il leader mondiale per varietà, esistono più di 500 specie di olive, di cui 45 oli d'origine protetta (dop) e una indicazione geografica protetta. (Della Porta, 2013)

Il paese leader per aree adibite alla coltivazione e quantità prodotte è la Spagna, dove i centri produttivi sono in media di dimensioni maggiori e hanno un tasso di produttività per lavoratore nettamente superiore rispetto al tasso degli altri paesi europei. I centri produttivi sono concentrati per il 78% in Andalusia nel sud della Spagna nelle province di Jaén, Cordoba, Granada, Siviglia e Malaga. L'area si distingue per le colture intensive con le caratteristiche delle economie di scale: tasso di produttività per lavoratore maggiore, salari medi più alti e redditività migliore per i produttori rispetto ai competitors di altre zone. (Commissione Europea, 2014)

Le aziende agricole specializzate nella produzione di olive tuttavia hanno redditività inferiore rispetto alla media delle aziende agricole nell'Unione Europea, seppure hanno una produttività (rapporto output/input) di gran lunga superiore rispetto alla media delle altre aziende agricole.

La motivazione principale di questo squilibrio tra redditività e produttività, può essere spiegato dalle dimensioni medie delle aziende agricole oleari, inferiori rispetto alla media delle aziende agricole nell'Unione Europea. L'input di lavoro familiare in Grecia è considerato il 92% degli input totali con medie decisamente inferiori in Italia e Spagna dove si attesta al 50%. Le tenute olearie solitamente sono a conduzione familiare, hanno dimensioni medio/piccole, in Grecia la media di terreno coltivata per azienda agricola è di 3 ha contro i 4 ha dell'Italia e i 12 ha in Spagna. Le aziende agricole specializzate in olivi tra i paesi sopra menzionati contano un terzo delle colture totali per paese, tuttavia in Spagna il raccolto è superiore alla somma congiunta di Italia e Grecia. È indicativo che gli asset delle aziende agricole in generale siano mediamente del 50% più alti delle aziende specializzate in produzione di olive. Le dimensioni inferiori e conseguentemente minor valore in immobilizzazioni sono da computare al fatto che l'olea europea ha un periodo d'incubazione superiore alle altre colture, infatti, la pianta entra in piena produzione solo dopo 12-15 anni dalla installazione. Ampliare le aree specializzate, quindi, è un investimento a medio-lungo termine che solo grandi produttori possono sopportare, unitamente ai sussidi istituzionali forniti dal fondo europeo per lo sviluppo delle zone agricole e rurali. Senza i sussidi e la detassazione ad hoc per le aziende agricole, le produzioni europee infatti si ritroverebbero fuori mercato considerando i costi di produzione inferiori sia nel nord Africa che in Asia e nelle Americhe. Nel grafico a barre redatto dall'OIC viene mostrato il costo di produzione per paese in cui spiccano i costi in Marocco e Tunisia in cui il costo di produzione è pari a € 1,98/lt e €2,36/lt, circa la metà del costo di produzione italiano € 3,97/lt. La riforma della UE introdotta nel 1998 ha avuto lo scopo di ridurre il numero degli strumenti di assistenza e di concedere un aiuto alla produzione effettiva abolendo gli aiuti in quota fissa verso i piccoli produttori che erano stati previsti dal regolamento 136/66 della CEE. Beneficiano di tali sussidi più di 2 milioni di coltivatori registrati nell'UE e l'aiuto viene fornito in base alle quantità effettivamente prodotte. I paesi che beneficiano di tali sussidi sono Spagna, Italia, Grecia, Portogallo e in misura ridotta la Francia. Vengono inserite le Quantità Nazionali Garantite (QNG) di cui il 42% del totale è assegnato alla Spagna, il 31% all'Italia e il 23% alla Grecia. In caso in cui il 20% della quota non venga

utilizzata potrà essere adoperata nel caso in cui un altro stato membro eccedesse il proprio QNG. (CE, n. 1638/98)

1.5 Catena del valore e formazione del prezzo focus sul mercato spagnolo

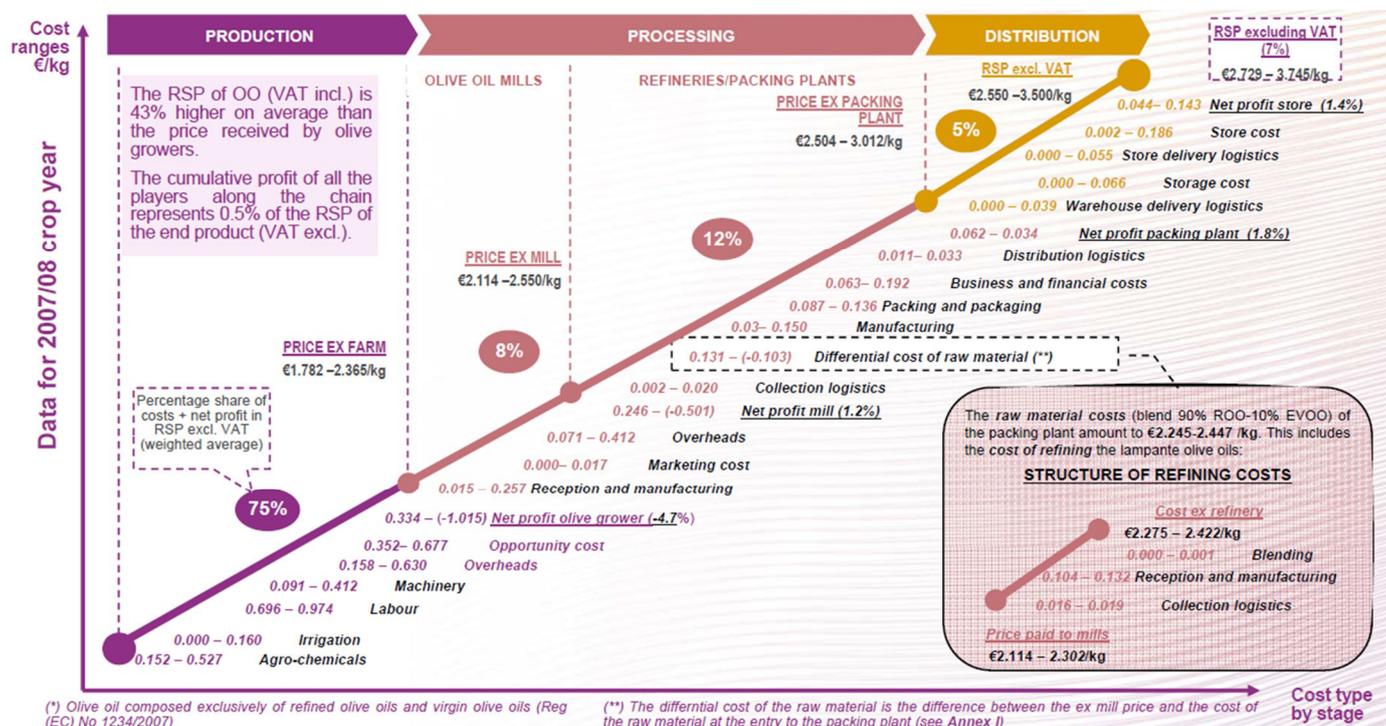


Figura 7 “Catena del valore del prezzo” Fonte: Governo Spagnolo Ministero dell'Ambiente . (2010). The value chain and price in spanish olive oil industry . Madrid p. 48.

Il prezzo dell’olio d’oliva a scaffale nel supermercato è influenzato da tutti i passaggi precedenti che avvengono dalla coglitura della materia prima fino al trasporto del bene finale al luogo di vendita al dettaglio. La filiera produttiva dell’olio d’oliva è composta da 3 fasi:

- Produzione delle olive
- Fase di trasformazione
- Distribuzione

La fase di produzione delle olive avviene nelle aziende agricole e vanno distinte le imprese che operano in maniera indipendente da chi invece opera in cooperative o partnership commerciali. Le dimensioni medie delle imprese agricole variano per singolo paese e la voce di costo che pesa maggiormente nella costituzione del prezzo dell’olio è il costo del personale. In Spagna vengono distinte 3 tipi di imprese in base alle dimensioni ed ai metodi di coltivazione utilizzati: tradizionali o coltivazioni estensive, coltivazione intensive e super intensive.

Le imprese tradizionali coltivano da 80 a 120 alberi per ettaro e utilizzano metodi di raccolta tradizionali o meccanizzati.

Le imprese che utilizzano sistemi di coltivazione intensivi utilizzano suolo con migliori caratteristiche vengono irrigati e hanno una densità media tra 200 e 400 piante per ettaro e nel caso delle colture super intensive possono contare anche 800 piante per ettaro. Questi sistemi risultano il nuovo modo per far crescere le olive riducendo i costi e massimizzando la resa. Questo tipo di imprese agricole sono in espansione in Andalusia dove sono adibiti alla coltivazione dell'olivo tanti ettari quanto in tutto il nostro paese.

Le principali attività svolte in questa fase sono il management del suolo che si occupa di come disporre le piante sul suolo definendo le distanze e quindi il tipo di coltivazione se tradizionale o intensiva in base alle caratteristiche e altre scelte di stretta competenza tecnica. I sistemi d'irrigazione che in Spagna coprono il 25% del totale delle terre irrigate AE (2015-16). L'irrigazione è una strategia che va attuata in base alle condizioni climatiche, alle condizioni della coltura e del suolo ed è molto variabile tra una regione ed un'altra. Nelle colture intensive dell'Andalusia in cui le piante ricevono migliaia di ore di luce all'anno e ricevono poca acqua piovana, è necessario l'installazione di impianti d'irrigazione.

Fertilizzanti, e trattamenti per preservare la salute degli olivi: il costo di questa attività dipende dalla frequenza e dal tipo di fertilizzanti adoperati, dallo stato di salute delle piante e dai macchinari utilizzati. I costi dei fertilizzanti aumenta all'aumentare della quantità di olive prodotte.

Potatura: Attività che regola la forma, la struttura e incide sulle quantità prodotte durante il ciclo biologico biennale dell'olivo. Può avvenire manualmente o per mezzo dell'utilizzo di macchinari specialmente nelle imprese intensive.

La fase del raccolto è la fase più costosa del processo produttivo delle olive. Questa fase conta dal 40% al 60% dei costi totali sostenuti durante la produzione delle olive. Le imprese tendono ad utilizzare la raccolta meccanica delle olive considerando l'alto costo del personale impiegato stagionalmente nella coglitura delle olive.

Infine, le olive vengono trasportate al frantoio dove inizierà la fase di trasformazione, il costo di questa fase dipende dalla distanza dal luogo di coltivazione al frantoio che solitamente sono situati in zone limitrofe alle principali zone coltivate ad olivo.

La fase di trasformazione

La fase inizia nei frantoi, in Spagna i frantoi attivi sono più di 1700 divisi nelle 13 comunità autonome, simili alle nostre regioni per dimensioni. L' Andalusia è chiaramente la comunità autonoma con il maggior numero di frantoi, ne sono presenti infatti più di 800. I frantoi solitamente vendono l'olio all'ingrosso agli intermediari che forniranno il prodotto finito alle imprese imbottigliatrici e alle raffinerie in cambio del loro servizio ottengono una percentuale come commissione sulla vendita. Le raffinerie lavorano l'olio lampante, l'olio non commestibile ottenuto dagli scarti della produzione di olio d'oliva e dell'olio extra vergine d'oliva acquistato dalle cooperative o dai frantoi. In Spagna ci sono 15 raffinerie di cui 9 in Andalusia. Le raffinerie lavorano l'olio lampante e lo mescolano insieme ad oli vergini per ottenere olio d'oliva commestibile che potranno successivamente imbottigliare o vendere all'ingrosso.

Le imprese imbottigliatrici utilizzano PET o vetro come packaging per l'olio e svolgono l'importante attività del quality management e della tracciabilità necessaria per assicurare la qualità e la tracciabilità lungo la filiera produttiva. Ci sono imprese produttrici integrate con le raffinerie che producono tipi di olio di diversa qualità dall'olio EVO fino all'olio lampante. Impianti di imbottigliamento di proprietà delle cooperative o di proprietà di un frantoio di grandi dimensioni che imbottigliano solo olio evo e imprese indipendenti che possono imbottigliare tutti i tipi di olio e alcune delle quali sono leader del settore in Europa.

Nel sistema distributivo operano piattaforme di distribuzione e servizi di acquisto, ipermercati e supermercati e negozi di piccola dimensione specializzati e non.

Le piattaforme distributive occupano di immagazzinare l'olio e controllarne la qualità organizzare la logistica in entrata ed in uscita e trasportarli al luogo di vendita finale.

La fase che influenza maggiormente e pesa di più sul valore finale del prezzo dell'olio è la fase di produzione delle olive. Il 68% del prezzo finale viene costituito in questa fase e la voce maggiore di costo è il fattore lavoro e le spese generali condotte dalle imprese agricole.

Nella fase di trasformazione nel suo complesso viene costituito il 27% del prezzo finale ripartito rispettivamente al 7% e al 20% tra i frantoi e le imprese imbottigliatrici. Sia per i Frantoi che per le imprese imbottigliatrici le voci di costo più rilevanti sono il costo dei macchinari e dell'impiego di quest'ultimi.

Il sistema distributivo invece conta il 5% della catena del valore del prezzo e non vi è una voce di costo rilevante in generale dato che tra i diversi sistemi distributivi vi sono differenze sostanziali, tuttavia il costo di immagazzinamento, trasporto e gestione dei negozi fisici sono caratteristiche comuni a tutti i sistemi distributivi ma pesano in maniera diverse in base alle dimensioni ed ai volumi scambiati.

1.6 Analisi della struttura e andamento dei prezzi

La struttura e l'andamento dei prezzi necessitano una premessa. L'olio d'oliva viene scambiato nel mercato all'ingrosso *business to business* (B2B) nei mercati locali ed il prezzo viene fissato in €/100 kg.

Prenderemo in analisi i dati forniti settimanalmente dalla Commissione Europea DG per l'Agricoltura e per lo sviluppo delle aree rurali, Unita G.4 Colture Arabe e olio d'oliva, in base al prezzo di vendita dell'olio extra vergine d'oliva, dell'olio vergine d'oliva e dell'olio vergine lampante a Jaén in Andalusia, a Bari in Puglia e a Canea (*Xavia*) sull'isola di Creta.. (European Commission DG Agriculture and Rural Development, 2016/17)

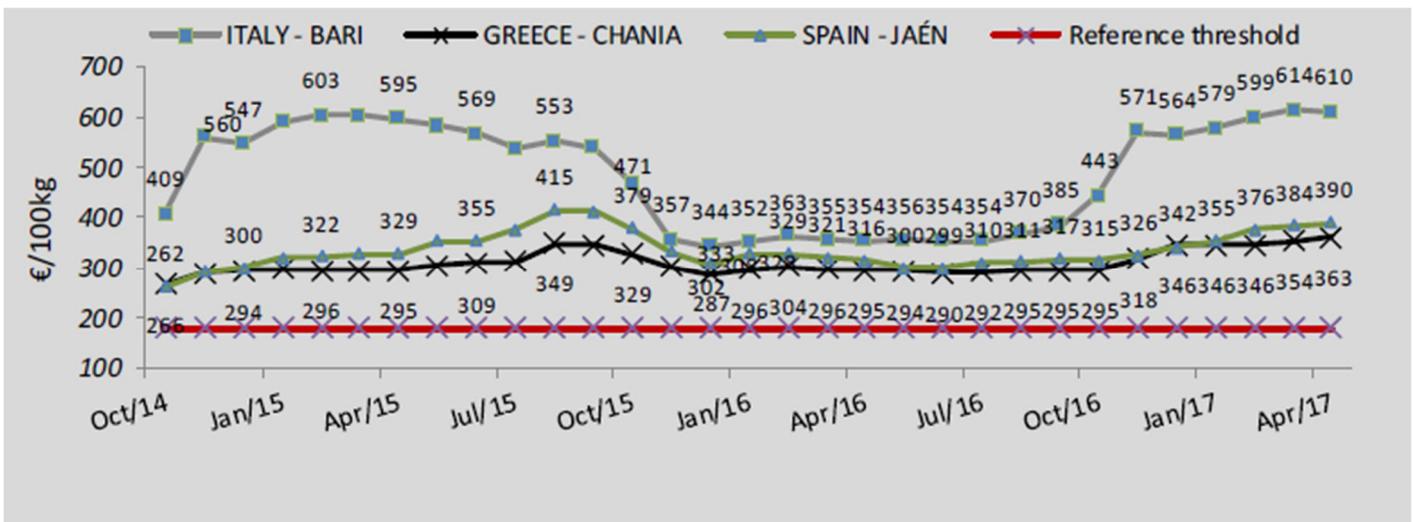


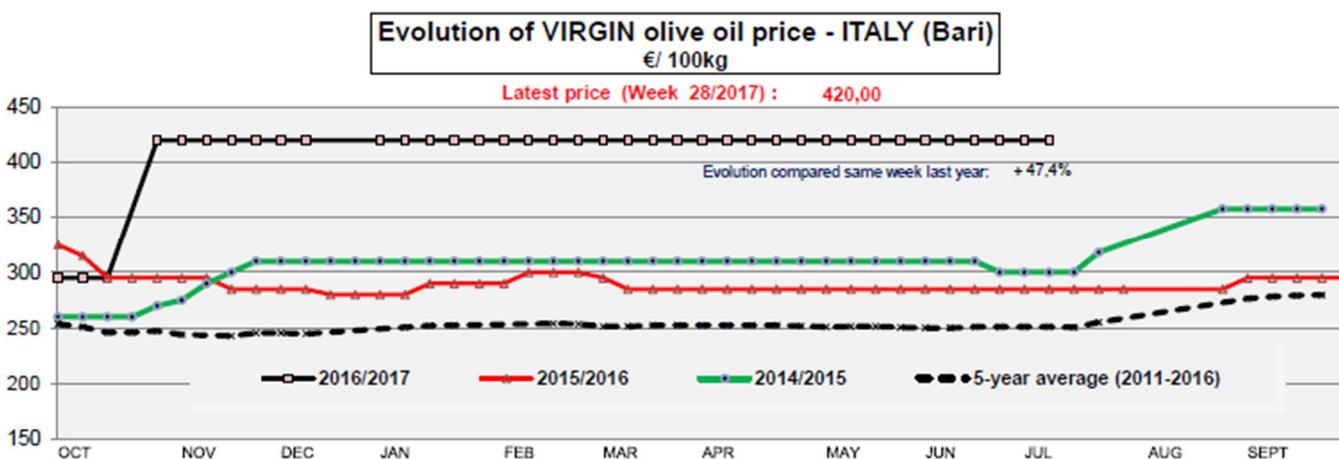
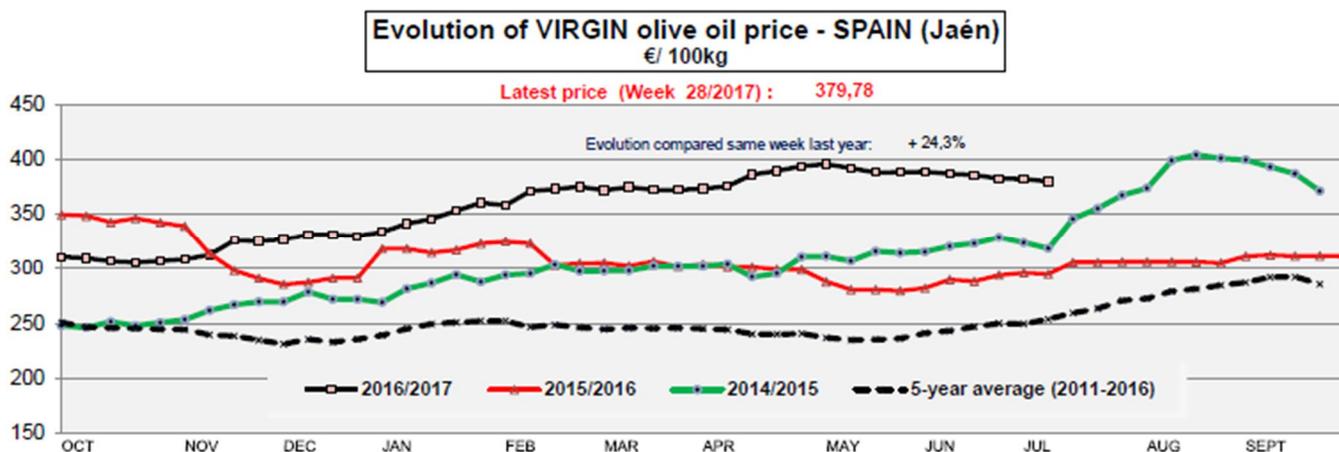
Figura 8 “Prezzi medi olio extravergine di oliva nei mercati europei più rappresentativi”

Fonte: https://ec.europa.eu/agriculture/olive-oil/prices_en

Confrontando l'andamento del prezzo nello spazio a Jaén il prezzo € 379,78/100 kg risulta il 32% inferiore rispetto al prezzo scambiato a Bari mentre a Canea il prezzo di € 365/100 kg varia rispetto a Jaén di soli € 14,78, (3%). Confrontando i prezzi di Canea e di Bari la variazione del prezzo è pari a € 197 e 35% in meno rispetto a Bari (€ 562/100 kg). Nella Figura 8 “Prezzi medi olio extravergine di oliva nei mercati europei più rappresentativi” vengono mostrati l'andamento congiunto dei prezzi da ottobre 2014 ad aprile 2017 la curva del prezzo di Bari sottende Canea e Jaén, che invece si ritrovano poco distanti specialmente tra gennaio del 2016 e gennaio 2017 in cui il pattern è sovrapponibile e i prezzi coincidono. La differenza di prezzo dipende da diversi fattori quali le quantità prodotte nella zona, la qualità e i costi di produzione. Se in Grecia il costo di produzione è particolarmente basso € 2,50 per litro, in Spagna il prezzo è di € 3,50/lt e in Italia è superiore a €4/lt. La Spagna essendo il leader per quantità prodotte riesce a mantenere un prezzo allo stesso livello della Grecia. In Spagna le olive vengono raccolte 3 settimane dopo rispetto all'Italia il che fa rendere dallo stesso peso di materie prime un 25% in più.

L'olio extravergine d'oliva a differenza dell'olio vergine e dell'olio lampante ha un grado di differenziazione maggiore in quanto è un prodotto di qualità superiore e viene ottenuto dalla sola spremitura meccanica delle olive e nessun altro processo esclusi la centrifuga e la decantazione. Olio vergine e lampante invece sono effettivamente commodities dato che il loro grado di differenziazione è inferiore e dato che questi possono essere utilizzato per ottenere oli raffinati.

Olio vergine d'oliva



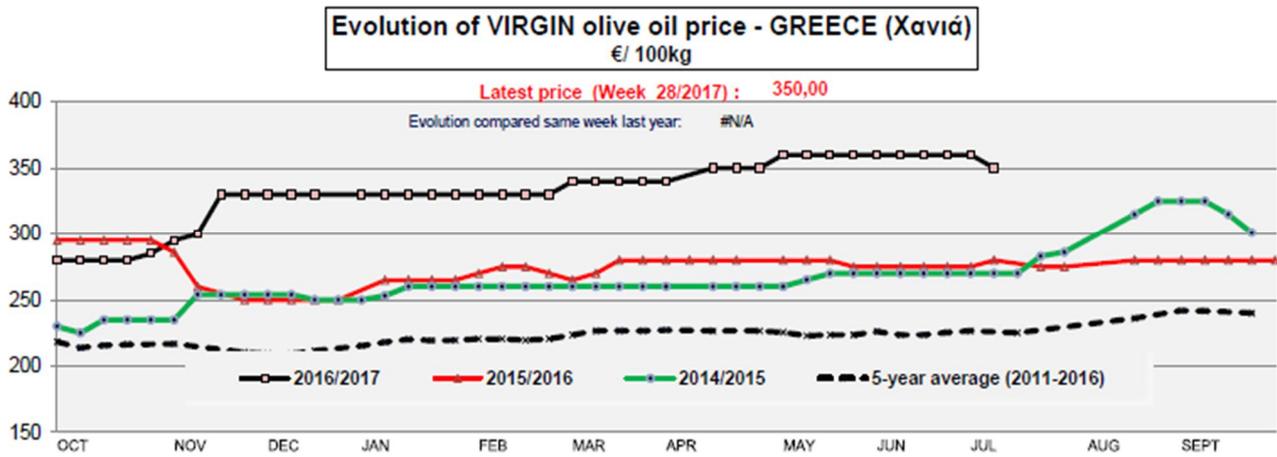


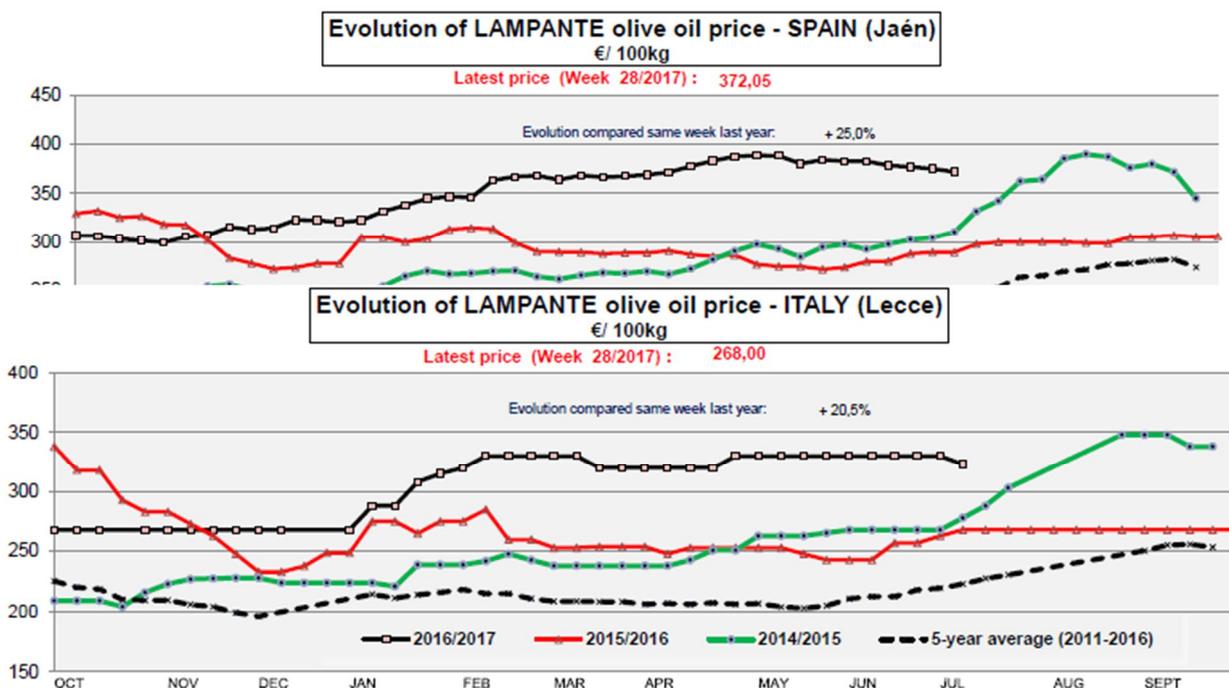
Figura 9 “Evoluzione prezzi dell’olio vergine di oliva”

fonte: https://ec.europa.eu/agriculture/olive-oil/prices_en

L’olio vergine d’oliva contiene una soglia di acido oleico in un range compreso tra l’1-2% è quindi di qualità e gusto inferiore rispetto all’evo, ad ogni modo è adatto all’uso alimentare ed è facilmente confondibile con un olio evo da parte di un consumatore medio.

Il prezzo dell’olio vergine rispetto all’olio extravergine a Bari è inferiore di € 142/100 kg. In Andalusia la differenza tra l’olio evo e l’olio vergine è molto ridotta inferiore a € 7/100 kg il che indica probabilmente una differenza in qualità meno significativa rispetto a quella di Bari. In Grecia la variazione tra i due prodotti è anch’essa molto bassa pari a € 15/100 kg. Il prezzo dell’olio vergine ha minore volatilità rispetto all’evo e il prezzo medio nei tre mercati negli ultimi 5 anni è stato pari a € 225 oggi invece ha un prezzo medio compreso tra € 350 - €420 dato che segue pari passo l’andamento dell’evo.

Olio Lampante



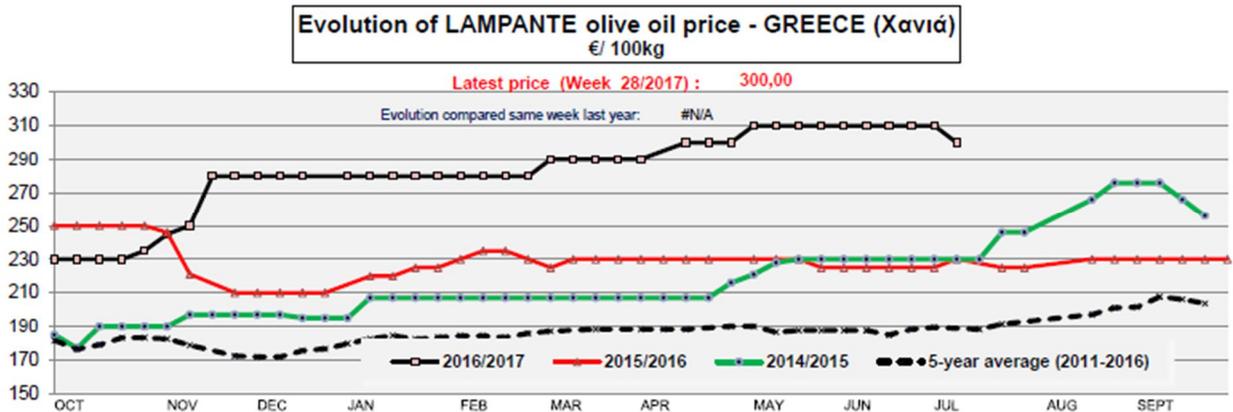


Figura 10 “Evoluzione dei prezzi dell’olio lampante

fonte: https://ec.europa.eu/agriculture/olive-oil/prices_en”

L’olio vergine lampante è caratterizzato da un tasso di acido oleico compreso tra il 2% - 3,3% per questo non è considerato adatto all’uso alimentare. Il nome lampante nasce durante l’Impero Romano dall’utilizzo di tale olio come combustibile per le lampade ad olio. Il prezzo è chiaramente inferiore e segue i pattern delle categorie di olio superiori negli stessi mercati in grandi linee. Viene indicato in Italia dal mercato di Lecce e il prezzo nella sett. 28 del 2017 è pari a € 320/100 kg contro i € 372,05/100 kg di Jaén e € 300/100 kg di Canea. La variazione rispettivamente rispetto all’olio vergine sono di - € 100 in Italia, - € 7,73 a Jaén e - € 50 a Creta. Il prezzo medio degli ultimi 5 anni ha particolare volatilità e risulta in aumento annualmente da giugno a settembre in tutti e tre i mercati.

L’olio lampante viene utilizzato per usi non alimentari e può essere adoperato nel processo di raffinamento per ottenere olio commestibile.

¹¹ https://ec.europa.eu/agriculture/olive-oil/prices_en

CAPITOLO II Agribusiness e ruolo dei modelli predittivi in agricoltura

2. Commodities alimentari e agribusiness

Il termine Commodity indica una sostanza o una materia prima standard che può essere scambiata, venduta o comprata sul mercato. La domanda primaria delle commodities e l'assenza di diversificazione nelle materie prime fanno sì che i produttori, a meno che non posseggano una quota di mercato significativa, non possano condizionare il prezzo di mercato. Gli operatori nella filiera produttiva delle commodities, quindi, possono essere definiti *price taker*. Il ruolo delle previsioni incrociate su domanda e offerta delle commodities sono di cruciale importanza per poter assicurare la stabilità dei singoli operatori e del mercato nel suo complesso.

L'agribusiness è il settore che combina le competenze in ambito agrario con quelle di management del settore agricolo. Le aree di interesse di queste imprese infatti includono: fertilizzanti e prodotti chimici, produzioni agricole, allevamento, macchinari e altre immobilizzazioni materiali e immateriali, management del raccolto (*harvest management*), waste management, risk management, marketing e vendite al dettaglio. L'integrazione verticale della filiera produttiva è la strategia che le principali corporation del settore agroalimentare hanno attuato con successo minimizzando così i rapporti di interdipendenza con eventuali fornitori e distributori.

I produttori di commodities subendo il prezzo di mercato sono soggetti alla volatilità del prezzo. Si possono verificare infatti scenari in cui il prezzo di mercato unitario sia inferiore al costo di produzione unitario del bene. I produttori infatti stipulano contratti futures a scopo cautelativo. I futures sono degli strumenti finanziari derivati che svolgono l'obiettivo di fissare in t_0 quantità e prezzo di scambio futuri di un determinato bene, così da minimizzare il rischio di prezzo. Gran parte dei contratti di future, hanno delle opzioni correlate che servono come mossa di copertura (*hedging*), infatti se il prezzo del sottostante non dovesse seguire l'andamento previsto il sottoscrittore ha limitato la sua perdita al solo costo del contratto di opzione.

Nell'ultimo decennio vi è stato un aumento senza precedenti nella domanda di prodotti agricoli, tra il 2004-06 e il 2014-16 la domanda di cereali è cresciuta del 25% passando da 2 miliardi di tonnellate a 2.5 miliardi. Secondo il rapporto stilato in cooperazione tra la FAO e l'OECD le ragioni di questa crescita sono la conseguenza di due fenomeni distinti tra loro: l'aumento della ricchezza diffusa in Cina composta congiuntamente dall'aumento demografico e dall'aumento del reddito pro-capite, il secondo motivo invece concerne l'utilizzo dei carburanti di origine vegetale come l'etanolo, il propano e il green diesel che hanno avuto un grande sviluppo nell'ultimo decennio ma secondo il rapporto il settore sarà meno attrattivo e vivrà una contrazione notevole nel decennio successivo. (OECD-FAO, 2017)

Nel primo paragrafo quindi verrà introdotto l'approccio dell'agribusiness così da avere un quadro della gestione dei processi aziendali delle imprese agricole e nello specifico di quelle olivicole. Nel secondo paragrafo considereremo la costituzione e i principali mercati di commodities alimentari con un focus sull'olio

olio di palma l'olio vegetale con la quota di mercato maggiore rispetto ai competitors. Nel terzo paragrafo è stato svolto una review sul ruolo dei modelli predittivi in agricoltura in generale e nel quarto ed ultimo paragrafo il concetto di innovazione in agricoltura.

2.1 Introduzione all'agribusiness

Il termine Agribusiness fu coniato nel 1957 da J, due economisti di Harvard nel libro *A concept of Agribusiness* in cui definiscono l'Agribusiness come: *“la somma totale delle operazioni coinvolte nella produzione e distribuzione di beni agricoli; funzionamento produttivo dell'impresa agricola; immagazzinamento, trattamento e distribuzione di beni agricoli e prodotti derivati da essi.”*¹² (Davis & A., 1957)

L'opera dei due economisti di Harvard si inquadra l'approccio e le strategie attuate dai grandi colossi del settore agroalimentare oltreoceano considerando tutte le operazioni svolte da quest'ultime dalla produzione degli input agricoli alla distribuzione al consumatore finale passando per le attività di promozione del prodotto, dalla logistica in entrata ed in uscita, del ciclo di vita del prodotto. Nel testo vengono individuate 5 principali aree che compongono l'intero sistema dell'agribusiness.

- I. Input: primo sottosistema che viene costituito dall'attività necessaria per poter avviare la produzione della commodity. Quindi in questo sottosistema fanno parte i fertilizzanti, i semi, i macchinari che possono essere prodotti importati o distribuiti. Nella coltivazione dell'olivo ad esempio bisogna considerare in questo sottoinsieme gli ettari di terra utilizzati o da utilizzare, disposizione e irrigazione delle piante, fertilizzanti ed eventuali protezioni da porre alle piante reti protettive e pesticidi.
- II. Produzione: Gli input sono direttamente utilizzati per la produzione di una commodity agricola che può essere coltivata e venduta senza subire alcun tipo di processo come le olive da tavola ad esempio oppure utilizzate come materie prime per costituire un prodotto finale come nel caso delle olive utilizzate per essere trasformate in olio d'oliva.
- III. Trattamento o lavorazione: Le commodities in generale possono essere trasformate in diversi prodotti. I livelli di trasformazione dipendono dal tipo di procedimento che deve essere attuato, che può essere semplice come pulire o sbucciare un frutto o essere un procedimento complesso che implica alterazioni chimiche come nel caso del cotone della gomma o dello stesso olio d'oliva.

¹² (Davis & A., 1957) Testo originale *“the sum total of all operations involved in the manufacture and distribution of farm supplies; production operations on the farm; and the storage, processing and distribution of farm commodities and items made from them.”*

- IV. Marketing: Le attività di marketing concernono tutte le attività di trasferimento della merce all'utente. Vengono pressoché individuati 4 possibili *pattern*
1. Trasferimento degli input dall'impresa produttrice all'impresa che necessita di quest'ultimi per avviare il ciclo produttivo. In questo caso quindi si tratta di attività di Marketing business to business
 2. Trasferimento dal produttore al trasformatore.
 3. Trasferimento dal trasformatore al distributore
 4. Trasferimento dal trasformatore al consumatore finale
- V. Attività di supporto: Consistono in tutti i soggetti che forniscono attività cruciali a sostegno dell'attività di impresa. Fornitori di servizi, intermediari, istituti di credito, associazioni di categoria, cooperative o organizzazioni di ricerca e regolatori del mercato.

L'Agribusiness quindi opera con commodities che possono avere una *short shelf-life* letteralmente breve durata sullo scaffale, o comunque una durata limitata rispetto ad altre commodities quali i gas fossili o i metalli. Nel settore dell'agribusiness infatti svolge un ruolo fondamentale l'attività di marketing sia nel B2B che nel B2C per poter stimolare la domanda e chiamarla all'azione onde evitare che il prodotto non possa più essere utilizzato perché scaduto o comunque non più adatto ad essere utilizzato per un determinato scopo.

L'Agribusiness integra i principi di management aziendale con i principi tecnici e i know how del settore agricolo, dalla produzione agricola alla consegna al consumatore finale.

Come tutte le materie interdisciplinari non è ben chiaro quali siano i limiti dell'agribusiness dato che mette insieme diverse competenze individuali e di organizzazione nella gestione di un'attività profittevole. È un settore che deve saper cogliere le opportunità, valutare i rischi. La valutazione dei rischi richiede l'analisi dei dati e delle informazioni che possono indicare la miglior strategia da percorrere per l'impresa. Trend correnti, questioni relative ai canali distributivi, gestione delle risorse umane e utilizzo dei macchinari, attività finanziarie e di contabilità devono essere pianificate coerentemente con la mission e la vision dell'impresa.

I due autori avevano attenzionato i principali stakeholder dell'agribusiness, le popolazioni rurali che vivono dell'economia creata da quest'ultime. Oggi la metà della popolazione globale vive in metropoli e in zone urbane, l'abbandono delle zone rurali ha comportato la distruzione dell'economie rurali in buona parte di Spagna e Italia che sono i primi mercati dell'olio d'oliva al mondo. Sono infatti stati avviati progetti per tentare di invertire questo trend, tuttavia come riscontrabile in report della commissione dell'UE per l'agricoltura e lo sviluppo delle zone rurali, senza sussidi istituzionali la maggior parte di queste imprese non potrebbe operare. Il costo del personale nelle imprese agricole europee è la voce principale nella costituzione della catena del valore delle commodity, riescono a competere con i costi di produzione di altre aree geografiche in cui i costi del personale sono competitivi e i costi di produzione in generale.

2.2 Commodities agricole - focus oli vegetali

Il settore agroalimentare nel mondo sta affrontando una sfida globale: garantire l'accesso a materie prime sicure e salubri in modo da poter nutrire la popolazione mondiale in crescita. La sfida più impegnativa delle istituzioni internazionali è quello di ottenere questo risultato incentivando un utilizzo efficiente delle risorse naturali, in modo da produrre le commodities agricole in maniera sostenibile, riducendo le emissioni di gas serra². (FAO OECD, 2017)

Il mercato internazionale delle commodities e le principali istituzioni hanno il ruolo di immettere liquidità nel settore, assicurando un ritorno certo ai produttori ed evitando shock di sistema, la banca mondiale in risposta all'instabilità dei prezzi delle commodities sta cercando di perfezionare l'accesso ai mercati dei futures esistenti per i paesi sottosviluppati, introducendo anche strumenti di scambio appropriati per coprire i rischi. (World Bank, 1999)

“È noto che mercati incompleti (mercati in cui non è possibile coprire tutti i rischi) conducano all'inefficienza degli stessi. Quindi, il perfezionamento del mercato dovrebbe essere il prossimo passo della *world-trade agenda* se l'efficienza dovesse essere pienamente assicurata e tutti i benefici del libero mercato realizzati.”¹³

Il rischio di prezzo nel mercato delle materie prime dipende dal movimento stesso del prezzo che segue una distribuzione non normale e quindi determina la volatilità. Volatilità che può essere calcolata tramite una variabile spread come la varianza o la deviazione standard. Costituire un modello con lo scopo di perfetta copertura dai rischi è impossibile, tuttavia ridurre il rischio di prezzo al costo della sola opzione migliora l'efficienza del mercato.

“Nell'ultimo decennio la domanda di materie prime ha visto un aumento senza precedenti. Tra il biennio 2004-2006 e il 2014-2016, il consumo di cereali (frumento, mais, riso e altri tipi di grano) è aumentato da 2 miliardi a 2,5 miliardi di tonnellate, implementando quindi 500 milioni di tonnellate nella domanda. [...] Questo aumento è stato guidato da due fattori principali: la crescita cinese e l'aumento della produzione di biocarburanti.”¹⁴

Ad ogni modo le previsioni indicano una riduzione nei tassi di crescita della domanda delle commodities agricole per usi alimentari e non nel periodo di riferimento. In particolare, oli vegetali e zucchero di canna i cui prezzi erano aumentati sensibilmente a causa dei loro utilizzi per la produzione di biocarburanti dovrebbero

¹³ Voituriez et al. 2001, What explains price volatility changes in commodity markets? Answers from the world palm-oil market, Nogent sur Marne, Cedex, France testo originale” It is known that incomplete markets (markets where not all risks can be hedged) lead to inefficiency. Hence, market completion should be the next logical step on the world-trade agenda if efficient is to be fully assured and all benefits of free trade realized.”

¹⁴ FAO OECD 2017, Agricultural outlook 2017-2026, p. 23 testo originale This increase was driven by two main factors: the rise of China and the growth in biofuel production. Between 2004-06 and 2014-16, the total consumption of cereals (wheat, maize, rice, and other coarse grains) increased from 2.0 bln t to 2.5 bln t, adding almost 500 Mt of additional demand This increase was driven by two main factors: the rise of China and the growth in biofuel production

ridurre i tassi di crescita. Gli oli vegetali sono una delle commodities agricole più scambiate nel mercato contano il 42% del totale e il mercato continuerà a crescere ma a tassi di crescita moderati. L'Indonesia stando alle stime della FAO dovrebbe registrare una contrazione nella crescita passando dal 6,1% dell'ultimo decennio a una media dell'1,5% durante il medio periodo 2017-2026. Malesia e Indonesia continueranno a dominare il mercato degli oli vegetali dato che la loro produzione è orientata alle esportazioni. Nello specifico in questo paragrafo focalizzeremo l'attenzione sull'olio di palma, l'olio vegetale che conta il 60% delle esportazioni globali e il 30% della produzione di oli vegetali globali. (Carter et al., 2007)

“L'output dell'olio di palma crudo è cresciuto da 3 milioni di tonnellate nel 1974/75 a 40 milioni nel 2007 registrando un aumento annuale superiore all'8% annui”¹⁵ In termini di volumi quindi risulta l'olio vegetale più scambiato al mondo e per questo motivo può risultare utile introdurre alcuni aspetti e criticità in modo da avere un termine di paragone per il mercato dell'olio di oliva.

Il mercato dell'olio di palma risulta essere il principale olio vegetale per volumi di scambio al mondo i suoi usi infatti variano ma è molto utilizzato su scala industriale per prodotti fritti e inscatolati come riso, noodles e patatine. Lo stato solido dell'olio di palma ne favorisce l'utilizzo anche per l'utilizzo industriale. I motivi del successo dell'olio di palma sono riscontrabili nel costo di produzione inferiore rispetto alle principali alternative. A livello internazionale è stata mossa una campagna mediatica contro l'utilizzo del prodotto considerato non ecosostenibile e nocivo per la salute, tuttavia non è stata registrata una diminuzione dello scambio dell'olio vegetale sebbene in diversi prodotti in particolare nel nostro paese è stato sostituito da altri oli vegetali come l'olio di semi di girasole o di soia. Riguardo l'utilizzo non alimentare dell'olio di palma è stato posto un tetto massimo per la produzione di biodiesel Malesia e Indonesia infatti i governi hanno posto un tetto di 6 milioni di tonnellate annui per la produzione di biocarburanti. (OECD FAO, 2017)

2.3 Ruolo dei modelli predittivi nel settore agricolo

Data l'importanza strategica del settore alimentare per il mantenimento della stabilità sociale e della sicurezza nazionale, governi e organizzazione intergovernative regolano ed intervengono nel mercato fornendo sussidi, incentivi, e limitazioni nei confronti dei produttori. Quest'ultimi sono diventati i principali realizzatori di previsioni sulle quantità agricole che verranno prodotte nel breve e nel medio-lungo termine. Stime affidabili sulle quantità prodotte e sui relativi prezzi sono di fondamentale importanza per gli addetti alla filiera produttiva che devono pianificare con mesi d'anticipo le strategie aziendali, la maniera in cui impiegare le risorse e finanziare la propria attività economica.

¹⁵ Carter et al. 2007, Palm oil markets and future supply, Oxford UK. Testo originale Global crude palm oil (CPO) has increased from less than 3 million tonnes in 1974/75 to almost 40 million tonnes in 2006/07. This represents an average annual growth rate of more than 8%

L'*international journal of forecasting* (IJF) rivista scientifica internazionale pubblicata da Elsevier, si occupa di riunire gli operatori delle diverse aree d'interesse a cui vengono applicati i modelli previsionali. I principali settori dei modelli predittivi chiamati *forecasting* in inglese sono marketing e vendite, finanza, clima, decision making, energia, problemi ed applicazioni legali. Gli obiettivi che si propone l'IJF è di chiudere i gap tra teoria e pratica, proponendo metodi utili e rilevanti per orientare politiche e decisioni. Le previsioni in ambito agricolo condividono alcuni punti comuni con le previsioni macroeconomiche:

- Orizzonte Temporale

Le previsioni possono essere divise in previsioni a breve termine che corrispondono a previsioni con un orizzonte temporale inferiore ai 12 mesi, le cui ripercussioni e raggio di azione non va oltre l'anno commerciale corrente. Medio termine indica un periodo compreso tra 2-5 anni e quindi coerenti con una pianificazione che non sfora 5 anni commerciali dal momento presente.

Lungo termine quando l'orizzonte temporale è superiore ad un lustro, e in questo caso le previsioni svolgono un ruolo importante per il settore di ricerca e sviluppo, per le innovazioni tecnologiche e sul cambiamento di un determinato settore.

- Performance delle stime e comparazione

Le performance dei modelli predittivi in generale possono essere valutati a posteriori e questo rende difficile la comparazione tra diversi modelli utilizzati per lo stesso scopo. Se a livello tecnico i modelli sono comparabili considerando misure sintetiche quale l'R *adjusted* non è detto che poi il modello rispecchi l'andamento reale e che i valori osservati siano relativamente simili ad i valori stimati.

In agricoltura tuttavia ci sono delle dovute differenze con altre attività economiche, in particolare nel mercato delle commodities agricole queste sono indifferenziate o comunque simili e gli operatori subiscono il prezzo di mercato. L'attività degli operatori ad ogni modo è strettamente influenzano dall'attività dei governi che svolgono un ruolo chiave nel controllo di un mercato d'interesse per la sicurezza nazionale. Infatti, gli stati forniscono sussidi, pongono limiti o in caso contrario forniscono incentivi per ampliare le zone adibite ad una determinata coltura. La politica di sviluppo delle aree rurali nell'Unione Europea viene finanziata tramite il Fondo europeo agricolo per lo sviluppo rurale che ha disposto 100 miliardi di Euro per il periodo 2014-2020 a cui si aggiungono altri fondi nazionali che gli stati membri dovranno stanziare per lo sviluppo dell'agricoltura nell'Unione Europea. Nel regolamento 1306/2013 approvato dal parlamento europeo viene inserito il CAP Politica Agricola Comune composta da 121 articoli con lo scopo di chiarire le modalità di finanziamento dell'unione europea verso le imprese agricole, quindi il ruolo degli organi di controllo e degli stati membri oltre al chiarimento dei ruoli e degli obiettivi di sviluppo dell'intero settore. Negli articoli 6 e nello specifico nell'articolo 22 del CAP l'UE si impegna a essere il principale realizzatore e fornitore di materiale tecnico per fornire lo stato delle colture e delle previsioni sulla resa agricola alle imprese agricole. In particolare, nel regolamento vengono evidenziati le immagini satellitari delle aree coltivate ed il

sistema di monitoraggio agro-meteorologico per potere fornire stime sulle quantità di raccolto. Nel prossimo capitolo verrà avviata un modello di forecasting locale per stimare le quantità di olive prodotte utilizzando dati agro-meteorologici. L'obiettivo infatti dell'analisi sarà quello di cercare di aumentare il valore predittivo e di analizzare le implicazioni che derivano da un sistema integrato di innovazione e monitoraggio dei risultati che risultano la base per la pianificazione aziendale.

2.4 Evoluzione e introduzione di sistemi di ICT in agricoltura

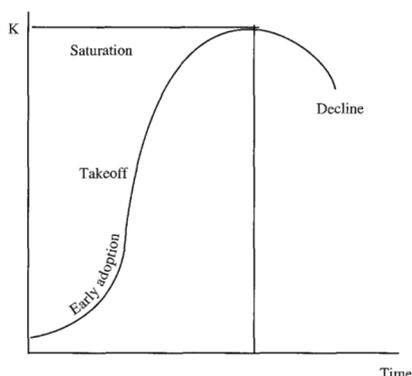


Figura 11 “stadi di evoluzione di una innovazione in agricoltura

Fonte: Sunding D., Zilbermann. (2001) Cap. 4 p.24 figura 3

L'adozione di nuove tecnologie in agricoltura è raramente un processo immediato. Le imprese agricole, specialmente in Italia e nei paesi mediterranei, in cui la maggioranza delle imprese sono a carattere familiare, non risultano propensi ad adottare strumenti di Information Communication Technology (ICT) come strumenti integrati che richiedono solitamente investimenti con un'ottica di medio-lungo termine. Le istituzioni hanno un ruolo chiave nell'incentivare l'adozione di nuove tecnologie che possono essere adoperate per molteplici scopi. La crescita demografica su scala globale richiede un conseguente aumento delle produzioni agricole e per avviare questo processo in maniera sostenibile, il settore non può prescindere dall'utilizzo di tecnologie che abbiano lo scopo di massimizzare la produttività minimizzando l'erosione del suolo.

Innovazioni possono essere definite nuove metodologie, consuetudini o strumenti utilizzati per raggiungere nuovi obiettivi. In letteratura vengono distinte le innovazioni che incorporano l'utilizzo di nuovi strumenti o beni utilizzati direttamente nelle produzioni agricole e le innovazioni di processo come l'utilizzo di strumenti per predire la resa futura delle coltivazioni.

Possiamo suddividere le innovazioni considerando la classificazione e la natura:

- innovazioni di tipo meccanico, utilizzo di strumenti meccanici come trattori o macchinari utilizzati per la coglitura dei frutti

- innovazioni biologiche come l'introduzione di nuove tipologie di semi o di nuovi metodi per coltivare una determinata coltura
- innovazioni chimiche introduzione di nuovi fertilizzanti e pesticidi
- innovazioni agronomiche utilizzo di strumenti quali il monitoraggio delle colture e di soluzioni di management del raccolto.
- Innovazioni biotecnologiche che riguardano la salute e lo stato delle colture quindi la possibilità che le colture contraggono malattie che possano segnare la coltura
- Innovazioni legate all'ICT e all'analisi di dati che possono avere molteplici usi ed implicazioni dall'installazione, al monitoraggio fino alla distribuzione del prodotto finito.

Le innovazioni possono anche essere distinte per il loro impatto economico per le imprese agricole e per il mercato. Queste categorie includono:

- Aumento della resa delle colture
- Aumento della qualità
- Riduzione dei rischi
- Protezione ambientale
- Aumento della durata dei prodotti agricoli

Come già accennato precedentemente vi è un intervallo di tempo tra quando una tecnologia viene sviluppata e quando quest'ultima viene largamente diffusa solitamente come confermato da [Roger 1962] che introdusse un nuovo tipo di mais ibrido in Iowa (midwest degli Stati Uniti d'America) e ne studiò lo sviluppo in tutto il paese denotando che l'introduzione delle tecnologie in agricoltura segua lo stesso pattern dell'innovazione di prodotto ovvero la diffusione ad S. La diffusione ad S indica che inizialmente l'innovazione viene utilizzata da pochi utilizzatori chiamati *early adopters* fino a che la tecnologia non venga resa nota dalla massa critica necessaria per raggiungere lo stadio del *take off* decollo, momento in cui l'innovazione raggiunge il suo momento di diffusione maggiore. Il momento successivo al decollo è la saturazione ovvero quando l'innovazione ha raggiunto la base di utenti massima e quindi non vi sono alti margini di crescita la curva quindi si appiattisce e precede la fase in cui quest'ultima viene superata e quindi la fase finale del declino dell'innovazione. Per quanto riguarda l'analisi dei dati aerobiologici e meteorologici per predire in maniera affidabile la resa degli olivi l'innovazione è stata apportata come spesso accade dall'università di Cordoba in Spagna all'inizio degli anni Ottanta. Gli *early adopters* è stato la giunta regionale dell'Andalusia e le cooperative di olio che nelle province di Cordoba, Jaén, Granada, Siviglia e Malaga. Questa metodologia poi è arrivata in ambito accademico alle università italiane ma non ha suscitato l'interesse delle imprese. In Andalusia la estensione in ettari coltivati ad olivo è aumentata di 170.000 ettari dal 1996/97 e le superfici irrigate sono passate da 144.047 ha, a 478.152 un aumento di 334.000 ettari in vent'anni.

CAPITOLO III Modelli predittivi

3. Modelli Predittivi sulla produzione di olio di oliva nel mercato spagnolo

Dato che il 70% del prezzo dell'olio di oliva viene a formarsi durante la fase di coltivazione e di raccolta delle olive, per le imprese operanti nel mercato è una crescente necessità disporre di una previsione accurata sulle quantità che verranno prodotte. Le imprese che dispongono dell'informazione possono indirizzare le proprie strategie operative di rotazione del magazzino e dell'impiego delle risorse umane in modo da ottimizzare i processi, determinando una riduzione dei costi operativi per l'impresa.

Le previsioni sono uno strumento che il management può consultare per scegliere tempi e modalità più favorevoli nella stipulazione dei contratti di fornitura, di trasporto, permettendo di ottimizzare i processi di logistica e le strategie di marketing e vendita.

Come area di studio è stato preso in considerazione il mercato spagnolo, paese leader nel mercato per quantità prodotte ed esportate. In Spagna, infatti, viene prodotto il 50% della produzione globale di olio di oliva. Nello specifico l'Andalusia è la regione più produttiva dove viene prodotto in media l'80% dell'olio spagnolo.

Come metodo previsionale è stato adoperato un algoritmo di data mining a scopo regressivo. Il gradient boosting che utilizza tecniche iterative nella costruzione di alberi di regressione.

Il campione è composto da 13 province spagnole, ed è stato selezionato in base alla produttività totale. Nelle province prese in esame viene prodotto il 93% dell'olio totale in Spagna. Le province escluse, quindi, non modificherebbero i risultati dello studio.

Province inserite nel campione per comunità autonoma (regione):

- Andalusia: Almeria, Cadice, Ciudad Real, Cordoba, Granada, Huelva, Jaén, Malaga, Siviglia.
- Estremadura: Badajoz, Cáceres
- Castiglia: Ciudad Real, Toledo
- Catalogna: Tarragona

Periodo di interesse (1995-2016) 20 campagne olivicole.

L'obiettivo dei modelli predittivi è stimare le quantità di olio di oliva prodotte annualmente nelle 13 province, adoperando variabili causali sulle superfici adibite alla coltivazione e sulle variabili metereologici che influenzano la produzione di olive.

Verrà quindi composto un *training set* per l'intervallo temporale (1998-2010) per addestrare il sistema supervisionato, e un *test set* sulle campagne (2011-2015) così da poter verificare le capacità predittive del modello.

3.1 Composizione e analisi primaria dei dati

I dati possono essere suddivisi in 2 categorie:

1. Dati metereologici
2. Dati riguardanti la produzione di olio di oliva e le superfici coltivate per provincia.

I dati metereologici sono stati scaricati dall'Open Source *European Climate Assessment and Datasets* (ECA&D) fondata dallo *European Climate Support Network* (ECSN) nel 1998 ricevendo aiuti finanziari dall'*European National Meteorological Services* (EUMETNET) e dalla Commissione Europea. Nei datasets sono presenti 10.584 stazioni metereologiche in tutto il mondo, il 78% di questi dati possono essere scaricati dal sito per ricerche non commerciali e/o scientifiche.¹⁶ Nella sezione utilizzata per costituire i modelli sono presenti: 31.551.768.000 di osservazioni e possono essere scaricate per un massimo di 500.000 osservazioni per query. I 12 elementi registrati giornalmente dalle stazioni sono: nuvolosità, umidità, temperatura media, temperatura massima, temperatura minima, piovosità, pressione sul livello del mare, altezza della neve, radiazioni solari, direzione del vento, raffiche di vento, velocità del vento.

I dati sulla produzione e sulle superfici coltivate per provincia sono stati scaricati dagli annuari di statistica pubblicati dall'Istituto di Statistica Spagnolo (capitolo 12).

3.2 Analisi dei dati metereologici

Premessa

Qualsiasi coltura è sensibile alle condizioni climatiche circostanti. Diversi studi sulla coltivazione dell'Olea Europea hanno dimostrato come il cambiamento climatico in atto alteri i cicli biologici della pianta. Ad esempio, le temperature miti durante l'inverno potrebbero interrompere la fase di riposo che richiede la gemma con conseguenze negative sulla fioritura e sulla produzione finale.¹⁷

Composizione iniziale dei dati metereologici

Sono state eseguite un totale di 52 *query*¹⁸, 4 *query* per singola provincia di interesse per l'intervallo di tempo disponibile (1979-2017).

¹⁶ <http://www.ecad.eu/>

¹⁷ Garcia-Mozo, Orlandi, & Galan, Olive flowering phenology variation between different cultivars in Spain and Italy: modelling analysis, 2009

¹⁸ <http://www.ecad.eu/dailydata/customquery.php>

Sono state scaricate per ogni singola provincia:

- Temperature massima (“TX”)
- Temperatura minima (“TN”)
- Temperatura media (“TG”)
- Quantità di pioggia (“RR”)

Le osservazioni vengono costituite giornalmente, le temperature vengono indicate in (0,1° C) mentre le piogge in (0,1 mm). Ogni singola query è scaricata in formato di testo (.txt) è composta da 4 colonne e da 13880 righe

Le osservazioni (righe) indicano i giorni dall’1 gennaio 1979 al 31 dicembre 2017.

Le variabili (colonne) invece: codice identificativo della query: (“SOUID”),Data: (anno, mese, giorno), Informazione: (TX, TN,TG,RR) in 0,1° C o 0,1 mm, Qualità della registrazione: (Q_TX,Q_TN,Q_TG,Q_RR) in forma binaria 0 per informazione registrata correttamente, 9 in caso di registrazione anomala

Manipolazione dei dati.

Obiettivo:

Raggruppare i dati metereologici in variabili mensili e stagionali inserendole in un’unica osservazione annuale per provincia, così da poter studiare l’incidenza delle condizioni meteo sulla quantità di olive prodotte in ogni campagna.

Procedimento

I dati metereologici scaricabili dal database dell’ECAD sotto forma di osservazioni giornaliere sono stati raggruppati così da formare un’unica osservazione mensile per ogni variabile (temperatura minima, massima, media e pioggia).

Inoltre, è stata creata una nuova variabile causale (“rain_days”) giorni effettivi di pioggia. Giorni in cui la quantità di pioggia registrata è \geq a 1 mm di pioggia.

Le variabili metereologiche dei mesi di novembre e dicembre sono stati inseriti nell’anno successivo tramite l’utilizzo della funzione *lag*. La campagna olearia infatti inizia l’1 novembre e si conclude il 31 ottobre e quindi i fenomeni atmosferici di novembre e dicembre influiscono sulla produzione dell’anno successivo.

Sono state quindi raggruppate altre 2 variabili per il periodo estivo ed invernale così da poter studiare come incidono le temperature stagionali sulla produzione. Le due stagioni sono state costituite da 4 mesi ognuno

Estate = giugno + luglio + agosto + settembre

Inverno = novembre + dicembre + gennaio + febbraio

Ogni osservazione del dataset “meteo” è composta quindi da 14 intervalli temporali costituiti da 12 mesi e 2 stagioni per 5 variabili climatiche: temperatura massima (“TX”), temperatura minima (“TN”), temperatura media (“TG”), quantità di pioggia in 0,1 mm (“RR”) e giorni di pioggia (“rain_days”). Ogni osservazione, quindi, è composta da anno (“year”), provincia (“region”) e da 70 variabili metereologiche.

Il dataset “meteorologia” è quindi formato da 488 osservazioni di 72 variabili.

Tabella 3 “Variabili meteo”

region	year	TX_s	TG_s	TN_s	RR_s	rain_days_s
--------	------	------	------	------	------	-------------

3.3 Analisi dei dati sulla produzione di olio d’oliva e superfici coltivate

I dati sulla produzione di olive e sulle variabili che ne influenzano la produzione sono stati ricavati dagli annuari di statistica, pubblicati dall’Istituto di Statistica spagnolo capitolo 12 (superficie delle coltivazioni e produzioni agricole sezione Olivi).

Dati primari

Le singole osservazioni sono composte da 18 variabili e si dividono in dati sulle superfici coltivate ad olivo, quantità di olive e tipi di olio prodotti.

Tabella 4 “ Dati primari produzione”

region	year	olive_for_hectar_w		olive_for_new_trees		olive_production	evoo	virgin_oil	lampante	oil_tot
hectar_nw	hectar_w	hectar_tot	hectar_nw_active	hectar_w_active	new_trees	olive_for_hectar_nw				

Provincia (“region”), Anno (“year”), Ettari non irrigati (“hectar_nw”), Ettari irrigati (“hectar_w”), Ettari totali (“hectar_tot”), Ettari non irrigati in produzione (“hectar_nw_active”), Ettari irrigati in produzione (“hectar_w_active”), Alberi seminati (“new_trees”), Resa ettari non irrigati kg/ha (“olive_for_hectar_nw”), Resa ettari irrigati in kg/ha (“olive_for_hectar_w”), olive per albero seminato in kg/ha (“olive_for_new_trees”), quantità olive da pressa in t (“olive_production”), quantità olio extra vergine di oliva (“evoo”), quantità olio vergine (“virgin_oil”), quantità olio lampante (“lampante”), quantità totale olio (“oil_tot”).

Obiettivo

Costituire un dataset contenente le quantità di olive, di olio prodotti e le superfici coltivate annualmente (1995-2016) per singola provincia.

Modalità di campionamento

I dati su produzioni e superfici coltivate sono state scaricate dagli annuari di statistica¹⁹ capitolo 12. Sono state inserite solo le province in cui vi fosse una stazione metereologica monitorata dall'ECAD per motivi di semplicità e di coesione del formato dei dati.

Tra le 27 province totali in cui viene prodotto olio in Spagna sono state selezionate 13 province così da ottenere una copertura media rispetto al totale prodotto del 93%. Nel campione sono state inserite tutte le 8 province dell'Andalusia, regione in cui vengono prodotte in media nei 20 anni di interesse l'80% del totale di tutta la Spagna e le 2 province dell'Estremadura, seconda regione più produttiva. Sono state inserite altre 3 province, 2 della Castiglia e 1 della Catalogna per rappresentare anche aree geografiche distanti l'una dall'altra ed estese su tutto il territorio.

Tabella 5 “Copertura campione rispetto alle quantità di olio di oliva prodotto”

1997	93,69%
1998	92,89%
1999	94,00%
2000	89,81%
2001	95,82%
2002	93,91%
2003	76,07%
2004	98,06%
2005	95,31%
2006	92,45%
2007	94,17%
2008	93,66%
2009	93,51%
2010	97,72%
2011	87,71%
2012	95,13%
2013	89,49%
2014	94,92%
2015	91,96%
2016	93,32%
media	93,18%

- Andalusia: Almeria, Cadice, Ciudad Real, Cordoba, Granada, Huelva, Jaén, Malaga, Siviglia.
- Estremadura: Badajoz, Cáceres
- Castiglia: Ciudad Real, Toledo

¹⁹ <http://www.mapama.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica/>

- Catalogna: Tarragona

Manipolazione dei dati

I dati per ogni singolo anno sono stati riportati dagli annuari su un singolo foglio di calcolo per ogni anno in formato (csv) dato che l'intestazione degli annuari risultava illeggibile dal software. Sono stati riuniti in un unico dataset tutti i file sui dati della produzione delle 13 province, creando un dataset contenente 260 osservazioni 20 (anni) per 13 (province). È stato quindi verificato che non ci fossero valori NA (non disponibili) e che le quantità prodotte fossero corrette.

Per la costituzione del dataset “spain” sono state esclusi i dati sugli alberi seminati, perché contenenti celle vuote o illeggibili, ed è stato preferito utilizzare gli ettari in produzione nelle modalità (“hectar_w_active”) ettari irrigati in produzione ed (“hectar_nw_active”) ettari non irrigati in produzione rispetto agli ettari totali per tipologie. Per semplificare sono stati unite le tipologie di olio (“evoo”, “virgin”, “lampante”) in un’unica variabile (“oil_tot”).

Le variabili incluse nel dataset “spain” per ogni osservazione sono:

Tabella 6 “variabili sulla produzione e sulle superfici coltivate”

region	year	hectar_nw_active	hectar_w_active	olive_for_hectar_nw	olive_for_hectar_w	olive_production	oil_tot
--------	------	------------------	-----------------	---------------------	--------------------	------------------	---------

region: provincia

year: anni (1996-2015)

hectar_nw_active: sono gli ettari coltivati ad olivo in produzione irrigati

hectar_w_active: sono gli ettari coltivati ad olivo in produzione NON irrigati

olive_production: olive prodotte in 1000 t

oil_tot: olio totale in 1000 t

3.4 Analisi delle correlazioni

Premessa

Per avviare l’analisi delle correlazioni sono stati uniti i dataset “meteo” e “spain” in un unico dataset contenente 260 osservazioni per 70 variabili.

Tabella 7 “Coefficients di correlazione lineare produzione di olive (pvalue 0,001)”

Variabile	Coeff. di correlazione lineare
olive_production	1
oil_tot	0,993
hectar_nw_active	0,86
hectar_w_active	0,808
olive_for_hectar_nw	0,739
olive_for_hectar_nw_2	0,628
olive_for_hectar_w	0,513
olive_for_hectar_nw_1	0,505
olive_for_hectar_w_2	0,441
olive_for_hectar_w_1	0,322
TN_august	0,134
rain_days_s	-0,07
TG_may	0,049
TG_april	0,031

Dalla matrice delle correlazioni tra la produzione di olive (“olive_production”) possiamo affermare che la produzione di olio (“oil_tot”) ha un indice di correlazione (0,9935) prossimo ad una correlazione perfetta positiva. Le variabili ettari attivi non irrigati (“hectar_nw_active”), ettari attivi irrigati (“hectar_w_active”) e la resa (kg/ha) per ettari non irrigati (“olive_for_hectar_nw”) sono fortemente correlate, rispettivamente con indici della correlazione di Pearson (0,86), (0,80) e (0,73). Gli altri valori sulla resa, quantità di olive prodotte per ettaro hanno un indice di correlazione moderata compresa tra (0,32 e 0,62) indice di correlazione di Pearson. I dati meteorologici significativi sono le temperature medie di agosto (0.133), i giorni di pioggia in estate (-0,07) e le temperature medie di maggio e aprile (0,049) (-0,03) variabili che sono correlate debolmente con la variabile dipendente.

Tabella 8 “Coefficients di correlazione lineare resa su ettari irrigati (p value 0,001)”

Variabile	Coeff. di correlazione lineare
olive_for_hectar_w	1
olive_for_hectar_w_2	0,666
olive_for_hectar_nw	0,623
olive_for_hectar_w_1	0,569
olive_production	0,512
oil_tot	0,494
olive_for_hectar_nw_2	0,484
hectar_nw_active	0,424
olive_for_hectar_nw_1	0,307
hectar_w_active	0,277
TN_august	0,212
TG_april	0,205
TG_may	0,113
rain_days_s	-0,157

La matrice delle correlazioni resa ettari irrigati (“olive_for_hectar_w”) indica un rapporto lineare dello (0,52) con la produzione di olive totale (“olive_production”) e dello (0,49) con l’olio totale (“oil_tot”).

Vi è una correlazione positiva moderata dello (0,623) con la resa delle superfici coltivate non irrigate.

Gli indici di Pearson sopramenzionati indicano che le rese tra i diversi tipi di superficie non sono correlate fortemente.

Nel grafico è evidenziata la correlazione tra gli ettari irrigati e la produzione. Vediamo come la provincia di Jaén sia l’area con la maggior distensione di ettari irrigati, nonché l’area più produttiva. A Jaén nel (2016) erano presenti 270.680 ha irrigati che hanno prodotto 4,5 milioni di tonnellate di olive. È possibile distinguere anche i pattern di Cordoba e di Granada in cui la superficie coltivata è superiore a 50.000 ha. Le altre aree, avendo una superficie irrigata ristretta, confermano il rapporto lineare tra quantità prodotte ed ettari coltivati.

Tabella 9 “Superfici irrigate, non irrigate e totali per provincia 2016”

Province	Sup irrigate in (ha)	superfici non irrigate in (ha)	Totale (ha)
Almeria	13.888	6.872	20.760
Cadice	2.277	21.392	23.669
Cordoba	53.280	291.499	344.779
Granada	69.287	123.372	192.659
Huelva	4.392	24.544	28.936
Jaén	271.916	314.158	586.074
Malaga	20.282	103.938	124.220
Siviglia	42.380	98.529	140.909
Badajoz	11.523	139.465	150.988
Cáceres	1.151	51.032	52.183
Tarragona	11.831	56.184	68.015
Ciudad Real	4.193	143.407	147.600
Toledo	3.460	117.163	120.623
totale	509.860	1.491.555	2.001.415

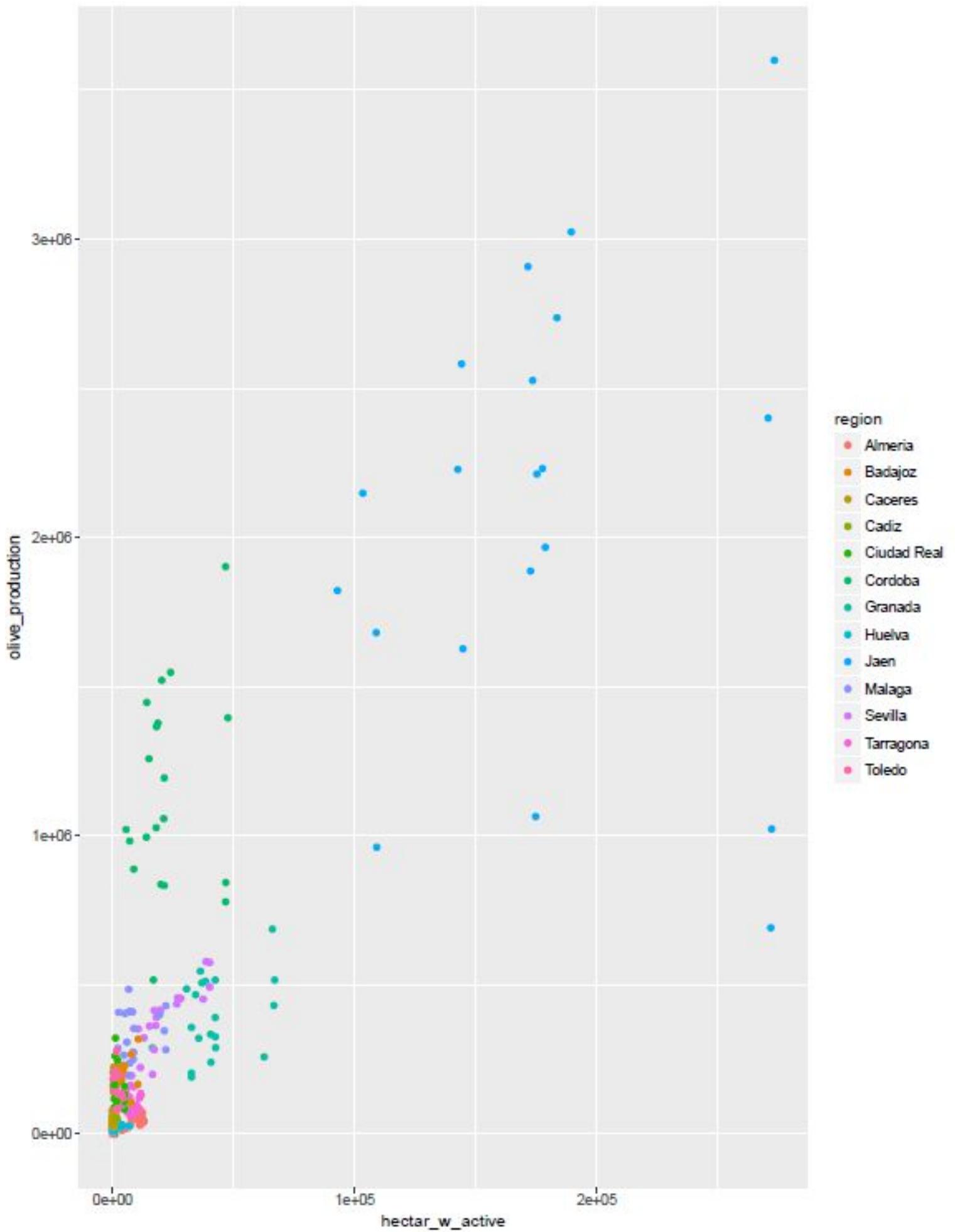


Figura 12 “Correlazione tra produzione di olive e superfici irrigate per provincia”

Tabella 10 “Indice di correlazione di Pearson resa su terreni non irrigati (p value 0,01)”

Variabile	indice di correlazione di Pearson
olive_for_hectar_nw	1
olive_production	0,739
oil_tot	0,716
olive_for_hectar_nw_2	0,715
olive_for_hectar_w	0,623
hectar_nw_active	0,619
olive_for_hectar_nw_1	0,559
olive_for_hectar_w_2	0,498
hectar_w_active	0,406
olive_for_hectar_w_1	0,372
TG_april	0,205
TN_august	0,189
TG_may	0,177
rain_days_s	-0,167

Nelle due matrici possiamo confrontare le diverse variabili che incidono sulla resa dei terreni irrigati e dei terreni non irrigati nell’area di Jaén. Sono state incluse solo le correlazioni significative al livello pvalue 0,01.

Notiamo come i due terreni siano correlati in maniera differente con le variabili metereologiche. Nella resa dei terreni non irrigati vi è infatti una correlazione positiva del 65% con le precipitazioni invernali ed in particolare con le precipitazioni del mese di dicembre. Al contrario, i terreni irrigati sono positivamente correlati al 57% con le piovosità del mese di marzo. È evidente che la resa dei terreni irrigati non necessita delle piogge invernali allo stesso modo dei terreni non irrigati. Le piogge nel mese di marzo sono positivamente correlate con i terreni irrigati perché favoriscono l’impollinazione delle piante.

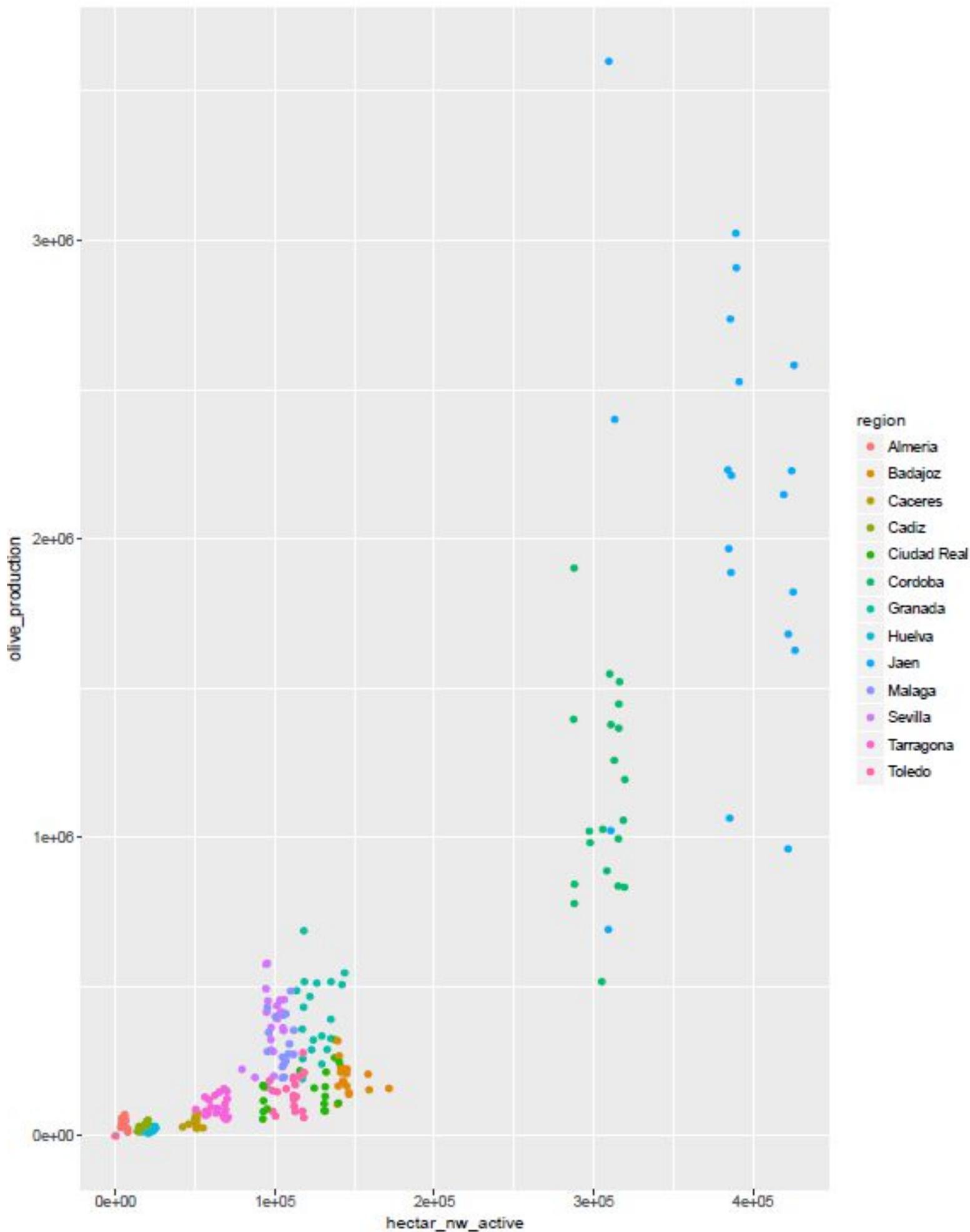


Figura 13 “correlazione tra produzioni di olive ed ettari non irrigati in produzione per provincia”

Tabella 11 “Indice di correlazione di Pearson resa su ettari non irrigati/ Jaén (p value 0,01)”

Variabile	Indice di correlazione
olive_for_hectar_nw	1
olive_production	0,916
oil_tot	0,897
rain_days_dec	0,652
RR_w	0,649
rain_days_w	0,652

Tabella 12 “Indice della correlazione di Pearson resa su ettari irrigati/Jaén (pvalue 0,01)”

Variabile	Indice di correlazione
olive_for_hectar_w	1
olive_production	0,945
oil_tot	0,904
RR_marc	0,578
rain_days_march	0,574

Nelle due matrici possiamo confrontare le diverse variabili che incidono sulla resa dei terreni irrigati e dei terreni non irrigati nell’area di Jaén. Sono state incluse solo le correlazioni significative al livello pvalue 0,01.

Notiamo come i due terreni siano correlati in maniera differente con le variabili metereologiche. Nella resa dei terreni non irrigati infatti vi è una correlazione positiva del 65% con le precipitazioni invernali ed in particolare con le precipitazioni del mese di dicembre. Di contro, i terreni irrigati sono positivamente correlati al 57% con le piovosità del mese di marzo. le piogge nel mese di marzo sono positivamente correlate con i terreni irrigati perché favoriscono l’impollinazione delle piante che avviene solitamente tra aprile e maggio.

È chiaro che le piogge invernali non influenzano la resa dei terreni irrigati allo stesso modo in cui influenzano la resa delle superfici che non ricevono irrigazione.

3.5 Alberi decisionali di classificazione e regressione

Gli alberi decisionali, *classification and regression trees* (CART) sono una procedura binaria ricorsiva capace di processare attributi continui e nominali come target e come predittori. Iniziando dal nodo radice i dati sono divisi in due bambini (children), e ognuno dei bambini a sua volta viene suddiviso in nipoti (grandchildren). Gli alberi crescono fino alla misura massima, quando non viene selezionata una regola di stop. L'albero di dimensioni massime viene quindi potato di nuovo alla radice tramite il nuovo metodo di potatura che minimizza la funzione di perdita. Il meccanismo CART è inteso per produrre non un albero, ma una sequenza di alberi potati annidati, ognuno dei quali è candidato per essere l'albero ottimale. L'albero "di giusta dimensione" o "onesto" è identificato per valutare la performance predittiva di ogni albero nella sequenza di potatura su indipendente dati di test. Le prestazioni dell'albero vengono sempre misurate su dati di test indipendenti (o tramite convalida incrociata) e procedono alla selezione dell'albero solo dopo che la valutazione basata sui test venga conclusa.

L'obiettivo di questa analisi è quello di predire la funzione obiettivo $f(y)$ minimizzando l'errore quadratico medio dell'errore.

“Gli alberi di classificazione e regressione non sono i primi alberi decisionali introdotti nel machine learning ma sono i primi ad essere stati descritti con rigore analitico e supportato da tecniche di statistiche sofisticate e dalla teoria delle probabilità”²⁰

Nei meccanismi di gradient boosting, la procedura di apprendimento si adatta consecutivamente a nuovi modelli, al fine di fornire una stima più accurata della variabile di risposta

$$Obj(\Theta) = L(\Theta) + \Omega(\Theta)$$

Funzione Obiettivo: Training loss + regularization

Training loss: misura l'adattamento del modello sui dati del set di addestramento

Regularization: misura la complessità del modello

L2 normalizza gli score delle foglie per la regressione lineare

Ottimizzando la training loss aumenta la predittività dei modelli, mentre la regularization riduce la complessità del modello.

²⁰Kumar V. et. Al (2009) p. 180. testo originale “CART was not the first decision tree to be introduced to machine learning, although it is the first to be described with analytical rigor and supported by sophisticated statistics and probability theory”

3.6 Modelli previsionali Xgboost (parametri)

Premessa

L'algoritmo processa solo variabili numeriche, quindi, prima di impostare i parametri del modello è necessario rendere le variabili di testo come le province ("region") in formato numerico.

PARAMETRI Xgboost regressione lineare²¹

Per avviare l'analisi è necessario impostare 4 tipologie di parametri.

- A) Parametri generali
- B) Parametri di boosting
- C) Parametri di learning task
- D) Command line parameters (parametri di linea di comando)

I parametri generali e di boosting vengono impostati di default dal programma e possono essere modificati in base alle esigenze dell'analisi.

I parametri di learning task e di linea di comando invece devono essere modificati manualmente, per fornire il set di apprendimento e l'obiettivo che si vuole raggiungere con il modello.

A)

Booster: di default viene utilizzato (gbtree).

Nthread: numero dei threads adoperati.

Num_pbuffer: buffer di predizione, normalmente impostato sul numero di istanze di addestramento.

Num_feature: determina le dimensioni del boosting.

B)

Eta: (default 0.3) tasso di apprendimento.

Gamma: (default 0) riduzione minime perdite per costruire la partizione di una foglia dell'albero regressivo.

Massima profondità:(default 6) Massima profondità di un albero, aumentando questo valore il modello diventa più complesso e può creare overfitting.

Min child weight: (default 0) Nella regressione lineare, corrisponde al numero minimo di casi necessari per raggiungere ciascun modo. Più è largo più conservativo renderà l'algoritmo.

Alpha: (default è 0) (L1) regolarizza i termini sui pesi, aumentando questo valore il modello diviene più conservativo.

²¹ <http://xgboost.readthedocs.io/en/latest/R-package/xgboostPresentation.html>

Lambda: (default=1) (L2) regolarizza i termini sui pesi, aumentando il valore renderà il modello più conservativo.

Lambda_bias: (L2) regolarizza i termini dell'errore.

Metodo di costruzione degli alberi automatico: utilizzato nel modello *exact greedy algorithm*.

C)

Obiettivo: Regressione lineare.

Base score: (default 0.5) il punteggio di predizione iniziale di tutte le istanze, bias globale.

“Eval metric”: metriche di validazione “rmse” (scarto quadratico medio dell'errore). “mae” (errore assoluto medio).

D)

Use_buffer: (default 1) Parametri di learning task (apprendimento).

Numero round.

Data: Train_matrix

Test: data.

Task: quattro tipi di compiti: train, pred, eval, dump.

train: impostare i vettori che devono essere addestrati dalla matrice di addestramento.

pred: predizione test, impostare i vettori su cui vengono effettuate le predizioni.

eval: per valutare il dataset, cross validation k-fold .

dump: per riportare il Modello in forma di testo.

3.7 Dataset definitivo Xgboost

I risultati della precedente analisi hanno mostrato come le variabili meteorologiche influiscano in maniera differente sulle superfici irrigate e su quelle non irrigate.

Per questo il dataset utilizzato per creare le matrici delle correlazioni è stato successivamente diviso in 2 osservazioni per ogni anno e provincia in base al tipo di superficie.

Esempio: sono state incluse due osservazioni per la campagna 2010 nella provincia di Almeria tipo 1 e tipo 2. Nel tipo 1 verranno inseriti i dati relativi ai terreni irrigati mentre nel tipo 2 i dati sui terreni non irrigati.

Questo passaggio è stato eseguito per studiare come le condizioni meteo influiscano univocamente sulle superfici irrigate e sulle superfici non irrigate.

Il dataset finale quindi è costituito da 455 osservazioni di 120 variabili.

Il dataset successivamente è stato diviso in sezione training contenente il periodo (1998-2010) e nella sezione test (2011-2015).

(“olive_for_hectar”) resa in kg/ha, (“prod_olive”) produzione di olive in t, (“hectar_active”) ettari in produzione.

$$\text{("olive_for_hectar")} = \text{("prod_olive")}/\text{("hectar_active")}$$

$$\text{("prod_olive")} = \text{("olive_for_hectar")} * \text{("hectar_active")}$$

$$\text{("prod_oil_tot")} \text{ per convenzione è stata costituita } = \text{("prod_olive")} * (0,2)$$

Solitamente per ogni kilogrammo di olive in Spagna viene prodotto dal 18% al 24% di olio, nel modello per convenzione utilizzeremo il 20%.

la funzione obiettivo: resa (kg/ha) è determinata dalla minimizzazione della funzione di perdita e dalla regolarizzazione (cv).

$$f(\text{"olive_for_hectar"}) = \min(\text{rmse}, \text{mae}) + \text{regularization}$$

3.7.1 Modello 00

Overfitting : Grafical model evolution

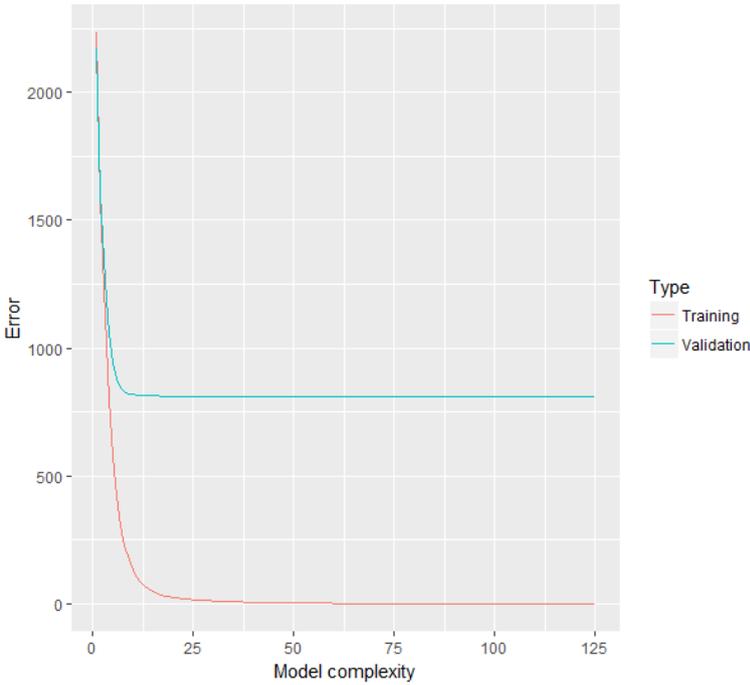


Figura 14 “Overfitting: evoluzione grafica del modello 00”

Model complexity

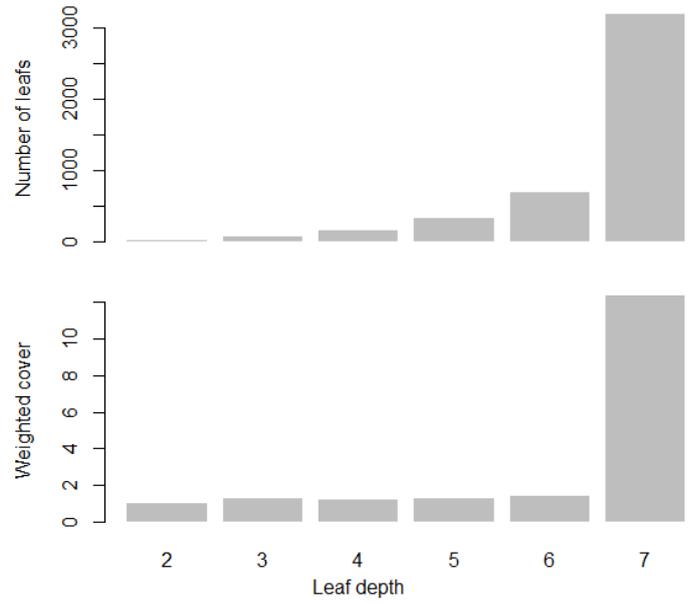


Figura 15 “complessità modello 00”

Cross validation Modello 00 `xgb_00.cv <- xgb.cv(booster = "gbtree", nthread = 8, alpha = 0,003, lambda = 0,004, lambda_bias = 0,04, objective = "reg:linear", base_score = 1, eval_metric = c("rmse", "mae"), use_buffer = T, nround = nround, data = trainMatrix, task = "train", nfold = 100)`

Per rendere il Modello più conservativo sono stati aumentati i parametri lineari di default pari a (0). Alpha (0,003), lambda (0,004) e lambda bias (0,04). Sono stati impostati gli obiettivi: regressione lineare con metrica di valutazione: scarto quadratico medio dell'errore (rmse) ed errore assoluto medio (mae). La complessità del modello, dato dal numero di iterazioni per minimizzare la funzione di perdita è 125 ovvero il numero massimo di round.

$R^2 = 0.996$ RSS (grado di incertezza residua) 0.7017751

Tabella 13 “Test model 00”

campagne	olive_pred.00	olive_prod	oil_prod	oil_pred.00	errore	errore %
2011-12	6.742.883	6.973.371	1.394.674	1.348.577	46.097	-3%
2012-13	5.288.704	3.124.198	624.839	1.057.741	432.902	+41%
2013-14	7.432.300	8.326.845	1.665.368	1.486.460	178.908	-12%
2014-15	5.814.929	3.795.069	759.013	1.162.986	403.973	+35%
2015-16	6.859.960	6.336.790	1.267.358	1.371.992	104.634	+8%
2011-16	32.138.776	28.556.273	5.711.252	6.427.756	716.504	+13%

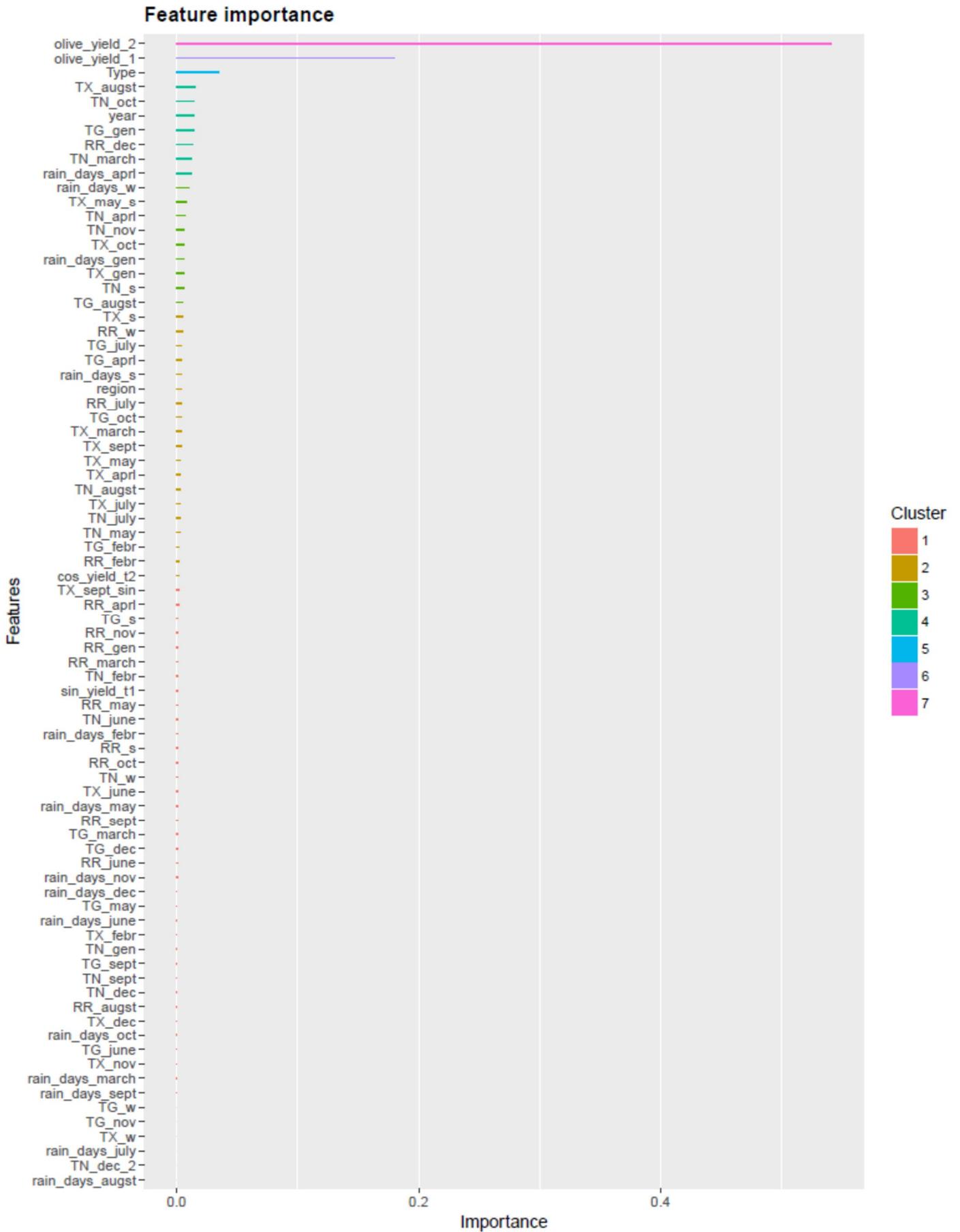


Figura 16 “Importanza features 00”

Commenti Modello 00

I risultati del modello 00 indicano una adattabilità dei dati con un $R^2 = 0,996$ e grado di incertezza residua $RSS = 0,7017751$.

I risultati del test dimostrano un errore assoluto del 13% nelle 5 campagne testate complessivamente.

Il modello stima in maniera accurata la resa delle campagne 2011/12, 2013/14 e 2015/16 con un errore assoluto rispettivamente del -3%, -12%, +8%.

Al contrario, sovrastima del 35% e del 42% la resa nelle campagne 2012/13, 2014/15.

La feature importance indica la quantità con cui ogni attributo migliora la performance misurata, pesata dal numero di osservazioni di cui ogni nodo è responsabile.

Le variabili causali vengono divise in 7 cluster in base al loro grado di importanza.

La variabile con il grado di importanza maggiore è la resa dei terreni non irrigati seguito dalla resa dei terreni irrigati. Tra le variabili meteorologiche quelle che influenzano maggiormente la resa nel modello sono: temperatura massima di agosto, temperatura media di ottobre, temperatura media di marzo, piogge invernali.

3.7.2 Model.01

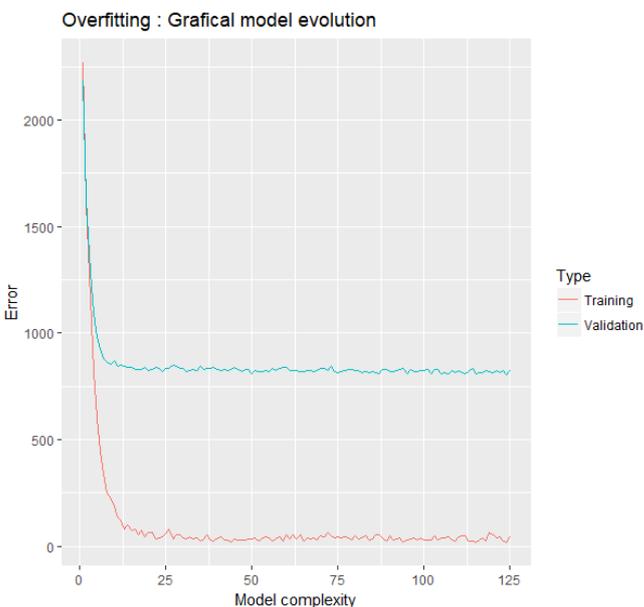


Figura 18 “Overfitting evoluzione grafica modello 01”

$R^2 = 0.997$ $RSS = 0,689$

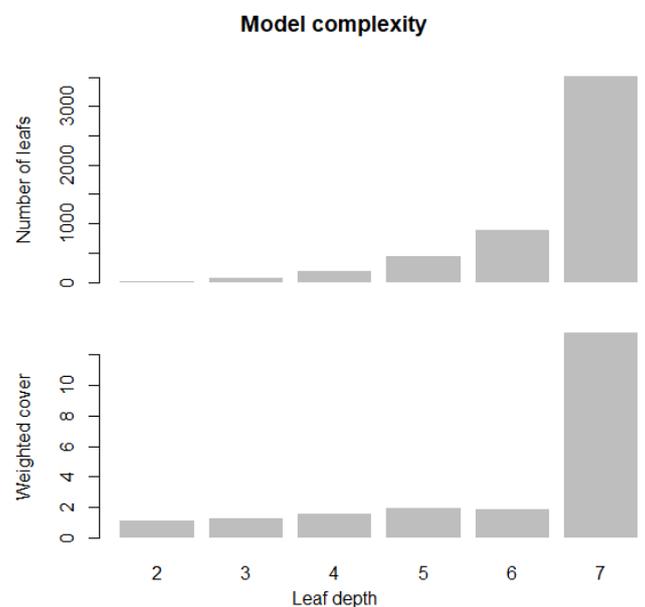


Figura 17 “Complessità modello 01”

Cross validation

```
xgb_01.cv <- xgb.cv(booster = "dart",nthread = 8, sample_type = "weighted",normalize_type = "forest",rate_drop = 0.01,one_drop = T,skip_drop = 0.001,alpha = 0.003,lambda_bias = 0.04,# objective = "reg:linear",base_score = 1, eval_metric = c("rmse", "mae"), use_buffer = T, nround = nround, data = trainMatrix, task = "train",nfold = 100,xgb_model = xgb_00
```

Nel Modello 01 è stato utilizzato il dart booster un algoritmo differente dal gbtree. Con questo boosting infatti vengono pesati e normalizzati i pesi. I parametri lineari sono stati aumentati per rendere il modello più conservativo). Sono stati impostati gli obiettivi: regressione lineare con metrica di valutazione: scarto quadratico medio dell'errore (rmse) ed errore assoluto medio (mae). La complessità del modello, dato dal numero di iterazioni per minimizzare la funzione di perdita è 125 ovvero il numero massimo di round.

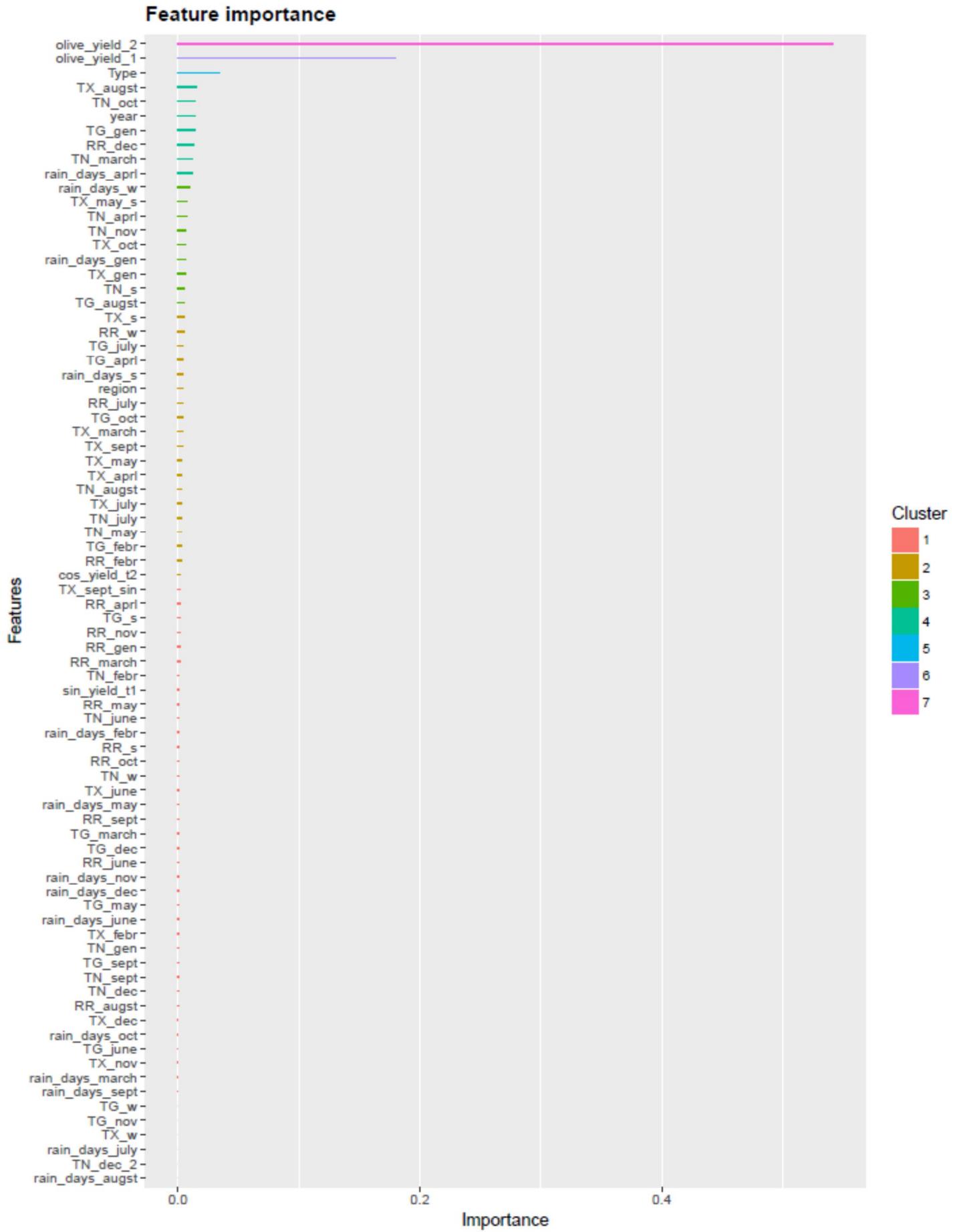
Tabella 14 "Test modello 01"

anno	olive_pred.01	olive_prod	oil_prod	oil_pred.01	errore	errore %
2011-12	6.742.883	6.973.371	1.394.674	1.348.577	-46.097	-3%
2012-13	5.288.704	3.124.198	624.839	1.057.741	432.902	+41%
2013-14	7.432.300	8.326.845	1.665.368	1.486.460	-178.908	-12%
2014-15	5.814.929	3.795.069	759.013	1.162.986	403.973	+35%
2015-16	6.859.960	6.336.790	1.267.358	1.371.992	104.634	+8%
2011-16	32.138.776	28.556.273	5.711.252	6.427.756	716.504	+13%

Commenti modello 01

Il modello si adatta ai dati con un R^2 (0,997) tuttavia resta un grado di incertezza residua RSS del (0,689). I risultati indicano un errore complessivo nel quinquennio del +13%. Il modello risulta altamente predittivo per le campagne 2011/12, 2013/14 e 2015/16 in cui viene commesso un errore assoluto del -3%, -12% e +8% di contro non riesce ad interpretare la riduzione in termine di quantità prodotta nelle campagne 2012/13 e 2014/15 registrando un errore assoluto del 42% e del 35%. I risultati sono corrispondenti al modello 00 anche se sono stati adoperati diversi algoritmi nel primo infatti è stato adoperato gbtree mentre nel modello 01 è stato utilizzato dart. Il grafico sulle feature importance raggruppa le variabili causali in 7 cluster in base al loro grado di importanza. La resa nei terreni non irrigati è la variabile più importante seguita dalla resa dei terreni irrigati. La variabile meteorologiche che incidono maggiormente sulla produzione sono le temperature massime del mese di agosto, temperature medie nel mese di ottobre, temperature medie e piovosità di gennaio, temperature medie di marzo, piovosità di aprile agosto, temperature medie nel mese di ottobre, temperature medie e piovosità di gennaio, temperature medie di marzo, piovosità di aprile.

Figura 19 “Importanza feature modello 01”



3.7.3 Modello 02

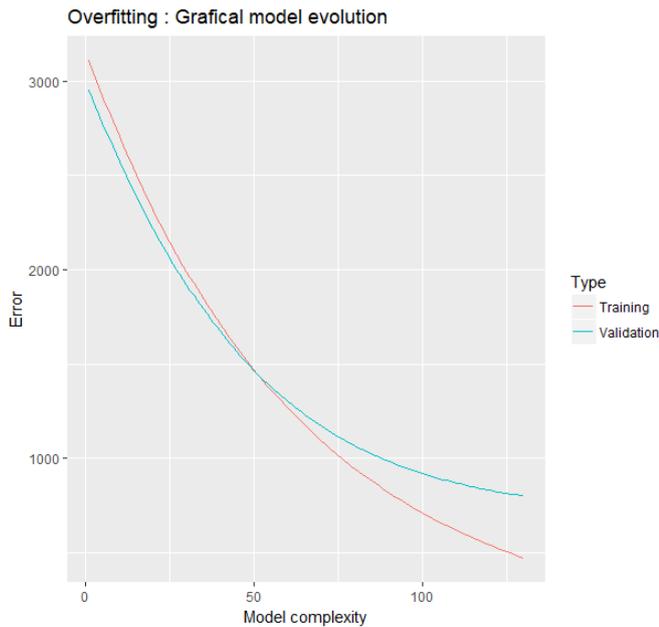


Figura 21 “Overfitting evoluzione grafica del modello 02”

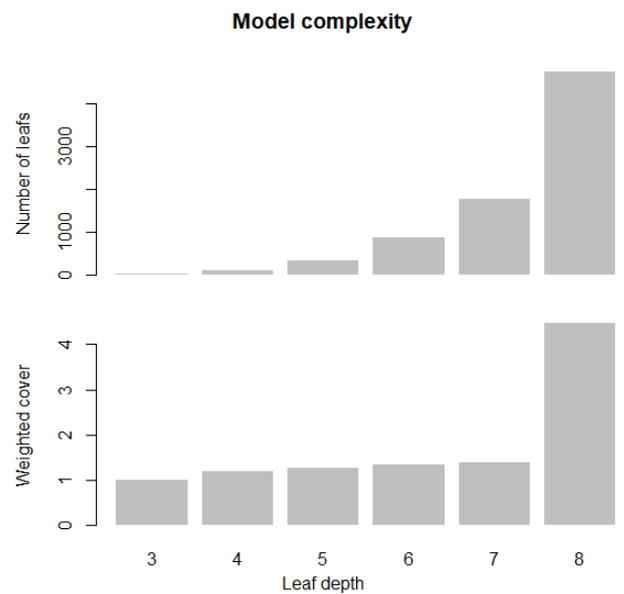


Figura 20 “Complessità modello 02”

$$R^2 = 0,997$$

$$RSS=0,689744$$

Cross validation

```
xgb_02 <- xgb.cv (booster = "gbtree", nthread = 8 eta = 0.016 gamma = 0,0 max_depth = 7,
min_child_weight = 0.2, subsample = 0.6, colsample_bytree = 0.8, colsample_bylevel = 0.8, sample_type =
"weighted", normalize_type = "forest", rate_drop = 0.01, # one_drop = T, # skip_drop = 0.01, # linear
parameter, alpha = 0.003, lambda = 0.004, lambda_bias = 0. objective = "reg:linear",base_score = 1,
eval_metric = c("rmse", "mae"), use_buffer = T, nround = 130, data = trainMatrix, task = "train", nfold = 100
```

Per il Modello 02 è stato impiegato il boosting gbtree impostando i parametri di dart rate_drop e skip_drop 0.01. I parametri lineari sono stati aumentati per rendere il modello più conservativo. Sono stati impostati gli obiettivi: regressione lineare con metrica di valutazione: scarto quadratico medio dell'errore (rmse) ed errore assoluto medio (mae). La complessità del modello, dato dal numero di iterazioni per minimizzare la funzione di perdita è 130.

Tabella 15 “Test modello 02”

campagna	olive prod.	olio prod.	olive stima	olio stima	Errore	Errore in %
2011-12	6.973.371	1.394.674	6.408.555	1.281.710	-112.964	-8%
2012-13	3.124.198	624.839	4.897.743	979.548	354.709	+57%
2013-14	8.326.845	1.665.368	6.625.849	1.325.169	-340.199	-20%
2014-15	3.795.079	759.013	5.222.246	1.044.449	285.436	+38%
2015-16	6.336.790	1.267.358	6.096.058	1.219.211	-48.147	-4%
2011-16	28.556.283	5.711.252	29.250.451	5.850.087	138.835	+2%

Commento Modello.02

L'R² 0,997 indica un buon adattamento del modello ai dati, con un grado di incertezza residua pari a 0,6968.

I risultati indicano un errore complessivo del -2%, analogamente agli altri 2 modelli vengono registrati errori del +57% e del +38% nelle campagne 2012/13 e 2014/15. Nelle campagne 2011/12, 2013/14 e 2015/16 gli errori invece sono rispettivamente dell' -8%, -20% e -4%.

Questi risultati dimostrano che questo modello sia il più affidabile rispetto ai due costruiti precedentemente.

Tuttavia, anch'esso ha delle limitazioni nello stimare correttamente il risultato di ogni singola campagna.

La feature importance indica la quantità con cui ogni attributo migliora la performance misurata, pesata dal numero di osservazioni di cui ogni nodo è responsabile.

Il grafico delle feature importance (figura 22) indica come le variabili meteorologiche che influenzano la produzione possono essere suddivise in 3 cluster in base al loro grado di importanza.

Cluster 3: TGs TXs Tns e RRw TGaugust Le temperature estive (medie, min. e max) sono le feature con il maggior grado di importanza del modello seguite dalla piogge invernali e dalla temperatura media di agosto.

Cluster 2: Le temperature di luglio, agosto e settembre seguite da giorni di pioggia invernali e da quantità di pioggia nei mesi da dicembre a marzo costituiscono il secondo cluster di feature per importanza.

Cluster 1: vengono inserite le feature che hanno una minor importanza le 5 feature con il miglior score sono: temperatura massima di agosto, temperatura minima di febbraio, piogge estive, temperature minime del mese di marzo e piogge di novembre.

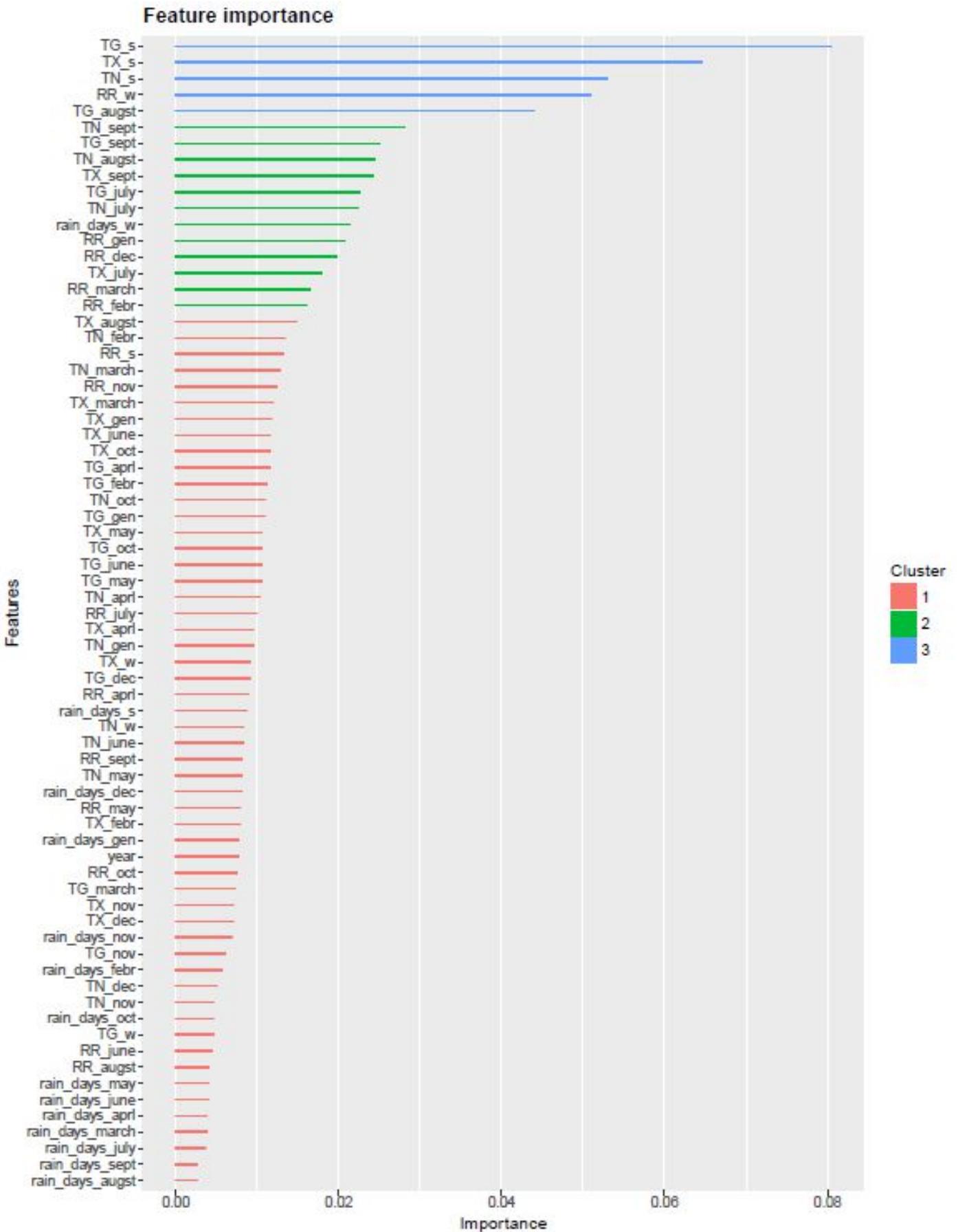


Figura 22 “Feature importance model02”

Rappresentazione grafica del Modello 02



Figura 23 “Alberi regressivi modello 02”

3.8 Ruolo critico dei modelli previsionali per le Imprese del settore

Tabella 16 “Confronto campagne 2014/15-2015/16

campagna	Prod. Olio Spagna in t	prezzo €/t Jaén	variaz %	variaz % produzione	stima model.02 in t
			prezzo set- nov		
2014/15	750.000	settembre € 409/t			1.044.449
		novembre €560/t	37%		
2015/16	1.267.358	settembre €553/t		69%	1.219.211
		novembre €357/t	-36%		

I modelli predittivi sulle quantità di olio prodotte sono indispensabili per il miglior impiego delle risorse a tutti i livelli della filiera produttiva e distributiva. Un buon sistema di *forecasting* fornisce un'indicazione sui volumi scambiati nel futuro e di conseguenza è possibile comprendere in quale momento può risultare più conveniente acquistare, immagazzinare e vendere materia prima, dato che il prezzo dell'olio è determinato dalle quantità prodotte in una determinata area. Nella disastrosa campagna 2014/15 in Spagna sono state prodotte 750.000 t di olio. A Jaén il prezzo è aumentato nei soli due mesi di inizio della campagna del 37%, passando da 409 €/t a 560 €/t, superando i 600€/t nel mese di febbraio.

L'anno successivo sono state prodotte 1.267.358 t di olio di oliva, il 69% in più rispetto alla precedente campagna. Il prezzo a settembre 2015 era 553€/t, di contro a dicembre dopo la fine della raccolta era crollato a 357€/t registrando un calo del 36% in solo 2 mesi.

Nel modello 02 le previsioni della campagna 2014/15 avrebbero previsto una produzione di 1.044.449 registrando un errore in termini assoluti del +38%. Nella campagna 2015/16 invece aveva registrato un errore di appena il 4%. Per ridurre l'errore assoluto è importante considerare ulteriori variabili che determinano la variabilità della produzione. Per le imprese del mercato agroalimentare è estremamente importante raccogliere i dati meteorologici nei pressi delle aree coltivate. Essi infatti influenzano quantità e qualità del raccolto e possono indurre il management a preservare ove possibile le colture, onde evitare riduzioni significative del raccolto. Installare stazioni meteorologiche nei pressi delle aree d'interesse ha costi facilmente ammortizzabili anche per imprese di piccole dimensioni, e con i giusti sistemi di prevenzione (*warning system*) si possono affrontare problematiche legate a fenomeni estremi che possono ridurre sensibilmente la resa. In Andalusia da più di trent'anni la giunta regionale di agricoltura e pesca con la collaborazione delle Università di Cordoba e di Siviglia hanno fatto sistema per incentivare lo sviluppo delle aree rurali della regione. L'Andalusia rappresenta un modello di sostenibilità in termini di produzioni agricole e nello specifico nella coltura degli olivi. Istituzioni, Università e cooperative agricole hanno lavorato congiuntamente nell'interesse della comunità e dello sviluppo del territorio, in cui solo tra le province di Jaén, Cordoba e Granada più di 1 milione di ettari sono destinati alla coltivazione degli olivi, rendendo la regione il luogo con la maggiore densità di olivi al mondo. (Instituto de estadística 2006)

In Andalusia viene prodotto circa il 25% della produzione di olio d'oliva su scala globale, e le colture intensive rendono possibile l'aumento dei margini sia per gli agricoltori che per i distributori. Le variazioni in termini di quantità e qualità, preventivabili con il giusto anticipo grazie all'attività di forecasting, indirizzano i processi primari dell'impresa quali: logistica in entrata, attività operative, logistica in uscita, marketing, vendite e risorse umane.

3.9 Gestione e riduzione dei rischi legati all'incertezza per i produttori d'olio

Il 70% del costo finale dell'olio evoo viene a formarsi al primo step della filiera produttiva, nella gestione delle imprese agricole. I dati sui costi dei processi sono forniti dall' *Observatorio de Precios de los Alimentos* MARM per conto del Ministero dell'ambiente Spagnolo e dal report dell' International Olive Oil Council sulle aziende agricole spagnole, greche ed italiane. Le voci di costo che le imprese agricole devono sopportare sono: i prodotti chimici utilizzati per preservare le piante, i costi di irrigazione, manodopera, macchinari utilizzati per la raccolta, costi generali d'impresa, costi opportunità. La catena del valore del prezzo contiene anche il profitto netto dell'agricoltore, profitto che può anche risultare negativo specialmente nelle annate in cui si registra produzione scarsa. È ormai un obiettivo noto delle istituzioni, specie dell'UE, fornire incentivi alle imprese agricole che devono fronteggiare la competizione di paesi in cui i costi di manodopera sono notevolmente inferiori.

La voce di maggior rilievo nella costituzione del prezzo è come già anticipato il costo di manodopera .In Spagna il costo del personale agricolo incide tra € 0,696 - 0,974 per ogni kg di evoo prodotto, che rappresenta il 38% e il 39% dei costi delle imprese agricole € 1,782 – 2,552. Il prezzo finale medio € 3,076 - € 4,118 per litro, mostra come il rapporto tra costo della manodopera agricola e prezzo finale sia del 23%. La catena del valore del prezzo, quindi, dipende per più di 1/5 dalla manodopera agricola. Le imprese agricole utilizzando dei sistemi di previsione della produzione efficaci ed efficienti, non possono ridurre il costo della manodopera ma possono aiutare ad utilizzare la forza lavoro nel modo più efficiente. Il costo d'irrigazione, che si attesta tra € 0 e € 0,16 per litro, 6% in rapporto al prezzo totale delle imprese agricole, 3% del prezzo totale. I costi d'irrigazione possono anch'essi essere ottimizzati calcolando la quantità e il periodo di concentrazione delle piogge e quindi minimizzando gli sprechi ed evitando shock dovuti alla siccità come accaduto nelle ultime due campagne olivicole 2016-17 e 2017-18. Il costo dei macchinari utilizzati varia tra € 0,091 - € 0,412 per kg di evoo prodotto e anch'essa dipende dai volumi prodotti e con una stima effettiva le imprese agricole possono ammortizzare quest'ultima voce in modo tale da pesare il meno possibile sul bilancio d'esercizio.

3.10 Gestione e riduzione dei rischi legati all'incertezza per le aziende confezionatrici

Le imprese confezionatrici si trovano a fronteggiare un mercato ad alta intensità competitiva, in cui disporre di informazioni maggiormente accurate sulle quantità di olive prodotte pone l'impresa in una posizione dominante rispetto ai competitors. L'asimmetria informativa causata da una efficace previsione del trend di produzione permette di poter stimare il costo della materia prima in maniera più accurata, consentendo al management di acquisire nel momento più adatto, a condizioni più favorevoli. La differenziazione, coerentemente con il posizionamento competitivo, è il processo chiave per creare preferenze di marca. L'olio extravergine d'oliva, essendo tra gli oli vegetali con prezzo indicativo maggiore e caratteristiche superiori rispetto ad altri oli vegetali, viene venduto ad un prezzo, anche 10 volte superiore rispetto ad altri oli vegetali raffinati.

Se un rischio esiste, non affrontarlo non elimina affatto gli eventuali costi, i processi di risk management infatti mirano a prevenire le perdite finanziarie con specifiche attività protettive.

La logistica in entrata è l'attività di gestione e regolamento dei flussi delle materie prime. Questo processo viene ottimizzato con l'utilizzo di tecniche predittive efficienti nelle varie zone di interesse. Infatti, l'impresa confezionatrice disponendo di informazioni relative alle quantità prodotte e alla localizzazione può stimarne il prezzo d'acquisto e di trasporto, comparandolo tra i possibili fornitori di materie prime, effettuando delle scelte operative in anticipo rispetto ai competitors.

Le scelte operative possono riguardare tempi, quantità, provenienza e qualità della materia prime al fine di costituire linee di prodotti che soddisfino i bisogni della base clienti dell'impresa

Il prodotto finito dopo 1-3 mesi viene imbottigliato per essere distribuito. Le imprese confezionatrici che hanno previsto con maggior accuratezza quantità e qualità prodotte hanno la possibilità di avviare la propria campagna marketing anticipando i competitors, nel caso in cui l'azienda si occupi di confezionare per conto di terzi non vi sono differenze, saranno solo quest'ultimi ad occuparsi dei processi di marketing e vendite oltre che i mercati in cui operare con maggior efficacia .

La catena del valore nella costituzione del prezzo finale dell'olio d'oliva e nell'olio extravergine in Spagna mostra come il 20% del valore totale dipende dalle attività di Marketing e il 4% dal sistema distributivo.

In Spagna sono presenti 1471 impianti di confezionamento, di cui i top 20 contano il 65% del totale. (Gobierno de Espana Ministero de Ambiente, 2010) Questo dato dimostra che in Spagna come in Italia vi sia una forte concentrazione, infatti solo 5 imprese producono il 35% del totale secondo le stime del MARM. Questo dato indica che pochi player hanno la possibilità di condizionare il mercato.

Nel sistema distributivo spagnolo hanno un ruolo chiave le piattaforme distributive e gli intermediari indipendenti che operano principalmente verso l'estero.

Nel mercato interno i principali canali distributivi sono ipermercati, supermercati e discount. I consumatori spagnoli stando allo studio del MARM Observatorio de precios de los alimentos (2008) acquistano il 40% di olio extravergine d'oliva negli ipermercati mentre acquistano il 39% nei supermercati meno del 20% nei discount stores.

3.11 Commenti finali

L'utilizzo di modelli previsionali per stimare la resa degli olivi in maniera affidabile e con anticipo rispetto all'inizio della campagna olearia, ha implicazioni su tutte le imprese che operano nel mercato. Il primo obiettivo da raggiungere per le imprese è la riduzione dei costi e dei rischi legati alla volatilità del prezzo della materia prima. Il 70% del prezzo dell'olio di oliva viene a formarsi durante la fase di coltivazione delle olive. Sebbene il 98% delle imprese agricole europee siano iscritte e ricevano gli aiuti finalizzati alla produzione da parte della Commissione Europea per l'Agricoltura, talvolta, essi non sono sufficienti per permettere di chiudere in attivo l'esercizio. Sia nella raccolta, sia nell'effettiva fornitura delle stime agli addetti alla filiera produttiva, le istituzioni hanno un ruolo chiave al fine di mantenere l'efficienza del mercato e consentire di ridurre l'utilizzo di fattori produttivi quali le risorse umane, che gravano sul costo finale del prodotto. Ridurre l'utilizzo di fattori produttivi è infatti l'obiettivo del concetto di sviluppo sostenibile, che in ambito agricolo è più che mai uno scopo necessario, considerato l'aumento della popolazione e le criticità segnalate nell'Outlook 2017-2026, che è stato redatto dalla collaborazione dell'OECD e della FAO. Il modello 02 fornisce un'ottima previsione nell'orizzonte temporale dei 5 anni delle campagne di riferimento, con un R^2 dello (0.98) e un errore assoluto del 2%. Gli errori assoluti nelle campagne 2014/15 e 2012/13 implicano che il modello possa essere migliorato inserendo delle variabili che influenzano la produzione, come l'indice di polline nell'aria suggerito dalla letteratura sull'argomento. Il metodo utilizzato può essere adoperato per fornire una stima complessiva di tutto il mercato dell'olio di oliva globale, dato che il 95% della produzione globale avviene nell'area mediterranea tra: Spagna, Tunisia, Italia, Grecia, Portogallo e Turchia. Per migliorare la predittività dei modelli può risultare utile installare stazioni meteorologiche nelle aree di studio, al fine di avere informazioni più dettagliate sull'incidenza delle variabili meteorologiche ed aerobiologiche, che influenzano le produzioni.

Bibliografia

- Allen P. G. (1994) Economic forecasting in agriculture. *International Journal of Forecasting* n 10. Amherst, University of Massachusetts.
- Abzar, N. (2010). A Review of Waste Management Options in Olive Oil Production. Turkey.
- Aguilera, F., Fornaciari, M., Valenzuela, L., Galàn, C., & Msallem, M. (2014). Phenological models to predict the main flowering phases of olive (*Olea europea* L.) along a latitudinal and longitudinal gradient across the Mediterranean region. *Int J Biometeorol* .
- Anania, G., & Pupo D'Andrea, M. (2007). The global market for olive oil: actors, trends, prospects and research needs Paper prepared for presentation at the. *I Mediterranean Conference of Agro-Food Social Scientists*.
- Avolio, E., Pasqualoni, Fornaciari, M., Bonofiglio, T., Orlandi, F., Bellecci, C., et al. (2009). Correlation between large-scale atmospheric fields and the olive pollen season in Central Italy. *International Journal of Biometeorology* .
- Candau. (1998). Forecasting olive (*Olea europaea*) crop production by monitoring airborne pollen. *Aerobiologia* , 185-190.
- Carter C. et al. (2007) Palm oil markets and future supply. LMC International Ltd. Oxford UK.
- Chen, T., Guestin C.(2016). Xgboost: A Scalable Tree Boosting System. University of Washington.
- Commissione Europea . (2014). *Farms specialised in olives in the EU based on FADN data*.
- council, o. I. (2007). *Production techniques in olive growing*. Madrid.
- de Graaff, J. (1999). Olive oil production and soil conservation in southern Spain.
- De Sousa, A., & Vines, R. (2017). Olive Oil Prices Are Going Through the Roof, Hoarding oil of a new kind after terrible harvests in Italy, Spain and Greece. *Bloomberg*.
- Della Porta, G. (2013). *Tutto sull'olio extravergine d'oliva*.
- Dhiab, A., Mimoun, M., Oteros, J., Garcia-Mozo, H., Dominguez-Vilches, E., Galàn, C., et al. (2016). Modeling olive-crop forecasting in Tunisia.
- D'Arpa, Uricchio, V. F., & Barca, E. (2010). Il data mining nello sviluppo di modelli previsionali per la gestione integrata di modelli previsionali per la gestione integrata dell'agricoltura sostenibile.
- Dorado, M., Ballesteros, E., Arnal, J., & Gomez, J. (2003). Exhaust emissions from a Diesel engine fueled with transesterified waste olive oil. *Fuel* , 1311-1315.
- European Commission DG Agriculture and Rural Development . (2016/17). *Olive oil weekly prices* .
- European, C. (2017). Market situation in the Olive Oil and Tables olives sectors. *Market situation in the Olive Oil and Tables olives sectors*. Bruxelles.
- European, U. (2012). *EU olive oil farms report*.
- Fantozzi, P. (1994). Introduction to the Problem of Evaluating Olive Oil Quality . *Acta* .
- Fornaciari, M., & Pieroni L., O. (2001). A new approach to consider the pollen variable in forecasting yield models.

- Fornaciari, M., & Pieroni, L. (1998). A regression model for the start of the pollen. *grana* .
- Fornaciari, M., Orlandi, Romano, B., & Agronomy, R. f. (2005). Yield Forecasting for Olive Trees: A New Approach in a Historical Series.
- Friedman, J.H.(1999). Greedy function approximation: A Gradient Boosting Machine. IMS 1999 Reitz Lecture.
- Galàn, C., Garcia-Mozo, H., Vazquez, L., Ruiz, L., Diaz de la Guardia, C., & Dominguez-Vilches, E. (2008). Modeling Olive Crop Yield in Andalusia, Spain. *Agronomy Journal* .
- Galan, & al, e. (2004). Forecasting olive (*Olea europaea*) crop yield based on pollen emission.
- Garcia-Mozo, H., Mestre, A., & Galàn, C. (2010). Phenological trends in Southern Spain: A response to climate change . *Elsevier* .
- Garcia-Mozo, H., Orlandi, F., & Galan, C. (2009). Olive flowering phenology variation between different cultivars in Spain and Italy: modelling analysis.
- Gennaro, D., Roselli, & Medicamento. (2009). Evoluzione degli scambi commerciali dell'Italia nel mercato mondiale degli oli di oliva.
- Gonzalez, M. F., Candau, P., Morales, J., & Tomas, C. (1998). Forecasting olive crop production based on ten consecutive years of monitoring airborne pollen in Andalusia (Southern Spain). *Agriculture Ecosystem and environment* .
- Governo spagnolo ministero dell'ambiente . (2010). *The value chain and price in spanish olive oil industry* . Madrid .
- Haylock, M. R., & . Hofstra, A. M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006.
- Hirst, J. (1952). An automatic volumetric spore trap. *Annals of applied Biology - Wiley Online Library* .
- Intl olive council. (2015). *International olive oil production costs* .
- Luchetti, F. (2008). Importance and future of olive oil in the world market — an introduction to olive oil.
- Martínez, M. G. (2002). A Repositioning Strategy for Olive Oil. *agribusiness* .
- Mason, J. et. Al (1999) Boosting Algorithms as Gradient Descent. University of Canberra.
- Mattingly, D. J. (1988). Oil for export? A comparison of Libyan, Spanish and Tunisian olive oil production in the Roman empire.
- Menapace, L., & Colson, G. (2011). Consumers' preferences for geographical origin labels: evidence from the Canadian olive oil market. *European Review of Agricultural Economics* .
- Mili, S., & Zuniga, M. (2001). Exploring Future Developments in International Olive Oil Trade and Marketing: A Spanish Perspective. *Agribusiness, Vol. 17 (3)* , 397-415.
- OIC; International Olive Council. (2016). *Market data*. Madrid.
- Orlandi. (2013). Potential shifts in olive flowering according to climate variations. *Meteorological application*, 497 – 503.

- Orlandi, F., Vazquez, L., Ruga, L., Bonofiglio, T., Fornaciari, M., & Garcia-Mozo, H. (2005). Bioclimatic requirements for initial flowering in two Mediterranean regions located at the same latitude (Andalusia, Spain, and Sicily, Italy) . *AAEM* .
- Oteros, Orlandi, Garcia-Mozo, Aguilera, Dhiab, & al., e. (2014). Better prediction of Mediterranean olive production. *Agronomy for Sustainable Development* , 685-694.
- Pasqualoni, L., Federico, S., Bonofiglio, T., Orlandi, F., Avolio, E., Bellecci, C., et al. (2007). Il rilascio del polline come indicatore dei cambiamenti climatici: La fioritura dell'olivo nel centro Italia. *Italian Journal of Agrometeorology*.
- Ronchin, C. (2016). *oli e frantoi d'Italia 2016*. Treviso: Chartesia srl.
- Spain, o. o. (n.d.). da <http://www.oliveoilfromspain.com/oofs/everything/regions.asp>
- Sunding D., Zilbermann. (2001) The agricultural innovation process research and technology adoption in a changing agricultural sector. Handbook of Agricultural Economics. University of California Berkeley
- Tasgodan, C. (2005). Country market power in EU olive oil trade. *South-Eastern Europe Journal of Economics* , 211-219 .
- Trichopoulou, A., & P., L. (1997). Healthy Traditional Mediterranean Diet: An Expression of Culture, History, and Lifestyle.
- Tura, Failla, Pedò, Gigliotti, Serraiocco, & Bassi. (2008). Italy Effects of Seasonal Weather Variability on Olive Oil Composition in Northern Italy. . *Olive and Olive Oil Research* .
- Violi, F., Loffredo, L., & Pignatelli, P. (2015). Extra virgin olive oil use is associated with improved post-prandial blood glucose and LDL cholesterol in healthy subjects. *Nutrition and Diabetes* .
- Visioli, F., & Galli, C. (1998). Olive Oil Phenols and Their Potential Effects on Human Health. *American Chemical Society* .
- Vossen. (2007). Olive Oil: History, Production, and Characteristics of the World's Classic Oils. *HortScience*.
- Vossen, P. (2005). Olive oil production .
- West, J., & Kimber, V. (2015). Innovations in air sampling to detect plant pathogens. *Wiley*.

Sitografia

<https://data.oecd.org/agroutput/crop-production.htm>

<http://www.ecad.eu/dailydata/customquery.php>

<http://www.frantoionline.it/normativa-olio/>

<http://www.internationaloliveoil.org/estaticos/view/102-structure>

<http://www.ismeamercati.it/flex/FixedPages/IT/Frantoi.php/L/IT#MenuV>

<http://www.juntadeandalucia.es/agriculturaypesca/observatorio/servlet/FrontController?action=Subsector&table=3940&ec=subsector&subsector=33>

<http://www.mapama.gob.es/gl/estadistica/temas/publicaciones/anuario-de-estadistica/default.aspx>

<https://www.olimerca.com/precios/tipoInforme/1>

<http://www.oliveoilfromspain.com/oofs/everything/regions.asp>

<https://www.politicheagricole.it/flex/cm/pages/ServeBLOB.php/L/IT/IDPagina/1019>

<http://www.rivistadiagraria.org/articoli/anno-2016/olio-extravergine-oliva-sintesi-del-quadro-legislativo-del-sistema-sanzionatorio/>

<https://www.statista.com/statistics/263937/vegetable-oils-global-consumption/>

<http://www.sviluppoeconomico.gov.it/index.php/it/commercio-internazionale/2021203-settore-olio-importazioni-di-olio-di-oliva-dalla-tunisia-in-regime-di-contingente-tariffario>

<http://xgboost.readthedocs.io/en/latest/R-package/xgboostPresentation.html>