

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Statistiche

Corso di Laurea Magistrale in

Scienze Statistiche



**Tesi di Laurea Magistrale**

Crescita delle energie rinnovabili: un'analisi multipaese con modelli di  
diffusione e modelli per dati funzionali

*Relatore:* Prof. Mariangela Guidolin

Dipartimento di Scienze Statistiche

*Correlatore:* Prof. Bruno Scarpa

Dipartimento di Scienze Statistiche

*Laureanda:* Margherita Gaetani

*Matricola N* 2026570

Anno Accademico 2022/2023



# Indice

<b>Introduzione</b>	<b>3</b>
<b>1 Energie rinnovabili: storia e dati</b>	<b>6</b>
1.1 Cenni storici . . . . .	6
1.2 I dati . . . . .	8
1.2.1 Energia eolica . . . . .	9
1.2.2 Energia solare . . . . .	12
1.2.3 Energia idroelettrica . . . . .	16
1.3 Modelli per spiegare e per prevedere . . . . .	20
<b>2 Modelli di diffusione</b>	<b>23</b>
2.1 Modello di Bass . . . . .	23
2.2 Applicazione del modello di Bass . . . . .	25
2.2.1 Analisi del coefficiente relativo al mercato potenziale . . . . .	29
2.2.2 Analisi del coefficiente relativo agli innovatori . . . . .	33
2.2.3 Analisi del coefficiente relativo agli imitatori . . . . .	39
2.3 Analisi dei gruppi . . . . .	44
2.4 Modellazione per spiegare . . . . .	51

---

<b>3</b>	<b>Modelli per dati funzionali</b>	<b>53</b>
3.1	Lisciamento dei dati . . . . .	54
3.1.1	Applicazione del lisciamento ai dati . . . . .	56
3.2	Componenti principali funzionali . . . . .	60
3.2.1	Applicazione della f-PCA ai dati . . . . .	64
3.3	Modello lineare simultaneo . . . . .	71
3.3.1	Applicazione del modello simultaneo ai dati . . . . .	73
3.4	Analisi dei gruppi funzionale . . . . .	80
3.4.1	Applicazione dell'algoritmo funHDDC ai dati . . . . .	82
3.5	Modellazione per prevedere . . . . .	90
	<b>Conclusioni</b>	<b>93</b>
	<b>Bibliografia</b>	<b>97</b>
	<b>Appendice</b>	<b>102</b>



# Introduzione

Con il termine di “*transizione energetica*” si intende il passaggio dalle attuali fonti di produzione energetica, che si basano principalmente sull’uso di petrolio, gas e carbone, a un mix più efficiente e meno inquinante di energie rinnovabili. Si può quindi affermare che la transizione energetica comporta l’adozione su larga scala di fonti di energia rinnovabile.

L’analisi multi-paese presentata in questa tesi mira a comprendere la dinamica temporale del processo di crescita di tre fonti di energia rinnovabile: eolica, solare e idroelettrica. L’obiettivo è, da un lato, spiegare in maniera parsimoniosa l’andamento del trend di crescita e, dall’altro, studiare la relazione tra le curve e alcune variabili esplicative considerate potenzialmente influenti e di interesse. A partire da questi due approcci, viene svolta, in entrambe le situazioni, un’analisi di raggruppamento che consente di identificare gruppi di paesi che presentano caratteristiche simili, consentendo di individuare profili di adozione particolarmente interessanti e rappresentativi.

La tesi è strutturata nel seguente modo. Nel primo Capitolo vengono presentati i dati analizzati riguardanti le serie storiche di produzione pro capite delle tre fonti di energia menzionate, ovvero eolica, solare e idroelettrica. Inoltre, viene fornita una breve panoramica sulla storia delle energie rinnovabili al fine di dare informazioni utili per comprendere come le stesse abbiano acquisito nel corso degli anni una tale importanza.

---

Nel secondo Capitolo, viene proposta una metodologia specificamente concepita per lo studio del processo di diffusione di un'innovazione, ovvero il modello di Bass. Una delle caratteristiche di questo modello è la facile interpretabilità, in quanto i parametri stimati permettono di caratterizzare un paese come innovatore o imitatore e, quindi, di descriverne l'andamento del processo di diffusione. Sulla base dei parametri stimati del modello si propone un'analisi di raggruppamento, con lo scopo di individuare una certa regolarità nei comportamenti innovativi o imitativi dei vari paesi.

Nel terzo Capitolo, i dati vengono trattati da una prospettiva diversa: ogni serie storica di produzione pro capite di energia di ogni paese viene considerata come una singola unità statistica, rappresentata da una curva. I dati sono quindi trattati come osservazioni ripetute nel tempo e assumono una forma funzionale. L'analisi dei dati funzionali prevede una fase di lisciamiento delle curve prima dell'analisi. Per fare ciò, si utilizzano diversi metodi di lisciamiento; quello utilizzato nel presente elaborato è l'espansione in basi. Anche in questo capitolo, si propone un metodo di raggruppamento funzionale. In questo caso, l'obiettivo è individuare fattori comuni nel processo evolutivo della produzione di energia nei vari stati.

A corredo dei risultati ottenuti con i due approcci di analisi vengono presentate delle considerazioni conclusive.





# Capitolo 1

## Energie rinnovabili: storia e dati

### 1.1 Cenni storici

L'uomo, sin dall'antichità, ha saputo sfruttare le energie rinnovabili a proprio vantaggio. Basti pensare all'uso dell'energia geotermica, per riscaldare abitazioni e terme, o all'utilizzo di mulini a vento e ad acqua, per macinare il grano e per irrigare i campi.

L'utilizzo delle energie rinnovabili è divenuto però di minore importanza con l'avvento dei combustibili fossili, tra i quali il carbone, prima, e petrolio e gas successivamente. Ha sicuramente giocato un ruolo fondamentale la prima Rivoluzione Industriale iniziata nel 1760, con l'avvento di una delle più importanti innovazioni tecnologiche, ovvero la prima macchina a vapore. Questa macchina veniva azionata dal carbone e la velocità dello strumento favoriva una maggior estrazione del carbone stesso. Si è assistito in questo periodo storico ad una trasformazione della società, sempre più dipendente dal consumo di materie prime energetiche, utilizzate per alimentare le macchine nelle industrie. Il processo tecnologico ha subito un'accelerazione con l'invenzione del primo motore elettrico nel 1831 e delle prime turbine. Questo ha permesso di utilizzare l'e-

nergia del vapore per generare elettricità nelle prime centrali elettriche di New York e Londra, all fine dell'Ottocento.

Nonostante la centralità dei combustibili fossili, l'interesse verso le fonti rinnovabili non è venuto meno, ma anzi si sono cercate soluzioni che potessero permettere di generare elettricità proprio da queste. Un esempio è la nascita della prima centrale idroelettriche posizionata sulle cascate del Niagara, inaugurata nel 1853. Va comunque detto che per sostenere le condizioni di vita e il ritmo industriale di quegli anni, l'unica fonte che ha permesso di tenere il passo della crescente domanda di elettricità è stato proprio il carbone. Tuttavia, se da una parte il carbone ha consentito un veloce sviluppo industriale, dall'altra ha fortemente inquinato le città e l'atmosfera. Questo continuo sviluppo è stato accelerato dall'utilizzo del petrolio; il primo pozzo petrolifero venne aperto nel 1859 negli Stati Uniti. Ciò che ha favorito la domanda di petrolio è stata l'introduzione del motore a combustione interna, sperimentato a partire dal 1853, ma reso efficiente solo nel 1899, con l'apparizione di vere frizioni che permettevano di azionare un veicolo da fermo. Attualmente, i combustibili fossili sono la fonte di energia più usata al mondo; tra questi, i più comunemente utilizzati sono il petrolio e il carbone, ma anche il gas metano ha progressivamente assunto importanza cruciale per la produzione di energia elettrica. Ad esempio, in un paese come l'Italia il gas rappresenta una fonte di sussistenza importantissima, tanto da posizionarsi al primo posto come paese europeo maggiormente dipendente dal gas naturale, nel 2021 (Murgia, 2022).

Le energie rinnovabili come le conosciamo oggi, fondano le loro basi a partire dal Novecento. Nel 1904 è entrata in funzione la prima centrale geotermica in Italia, negli anni '30 i primi parchi eolici negli Stati Uniti, negli anni '70 sono nati i primi pannelli solari, mentre l'era di sfruttamento delle maree è iniziata in Francia a partire dal 1968. Pur essendoci sempre stato un interesse per le energie rinnovabili, queste hanno rappresentato solo un elemento marginale nella produzione di energia mondiale. Un evento che

ha sicuramente favorito lo sviluppo di queste fonti energetiche è stata la crisi economica del 1973, causata da un rapido aumento dei prezzi del petrolio; a questo si è aggiunta anche la crescente presa di consapevolezza dell'esauribilità dei combustibili fossili e dei danni che questi provocano all'ambiente. Per quanto ci si fosse resi conto che la soluzione che avrebbe garantito un'indipendenza dai combustibili fossili fosse l'utilizzo delle fonti rinnovabili, la messa in atto di questo piano non si è consolidata in pochi anni. Infatti, ci si è trovati di fronte a numerose barriere di tipo tecnico, economico, ambientale, politico e culturale.

Proprio a causa di queste barriere, l'evoluzione delle energie rinnovabili non è stata costante. Tuttavia, a causa del continuo aumento della popolazione e del conseguente aumento della domanda mondiale di energia, la sempre minore disponibilità di combustibili fossili e dei cambiamenti climatici generati da emissioni di CO<sub>2</sub>, l'utilizzo di queste fonti rappresenta una soluzione percorribile. In questi anni, per perseguire questo scopo, la ricerca e lo sviluppo tecnologico sono cresciuti in modo da consentire modalità di produzione di energia sempre più efficienti ed economicamente sostenibili. Le energie rinnovabili che si intende trattare in questo elaborato sono l'energia eolica, solare e idroelettrica, essendo queste tra le più utilizzate. Nella sezione seguente, verranno descritti i dati riguardanti la produzione pro capite delle tre energie rinnovabili.

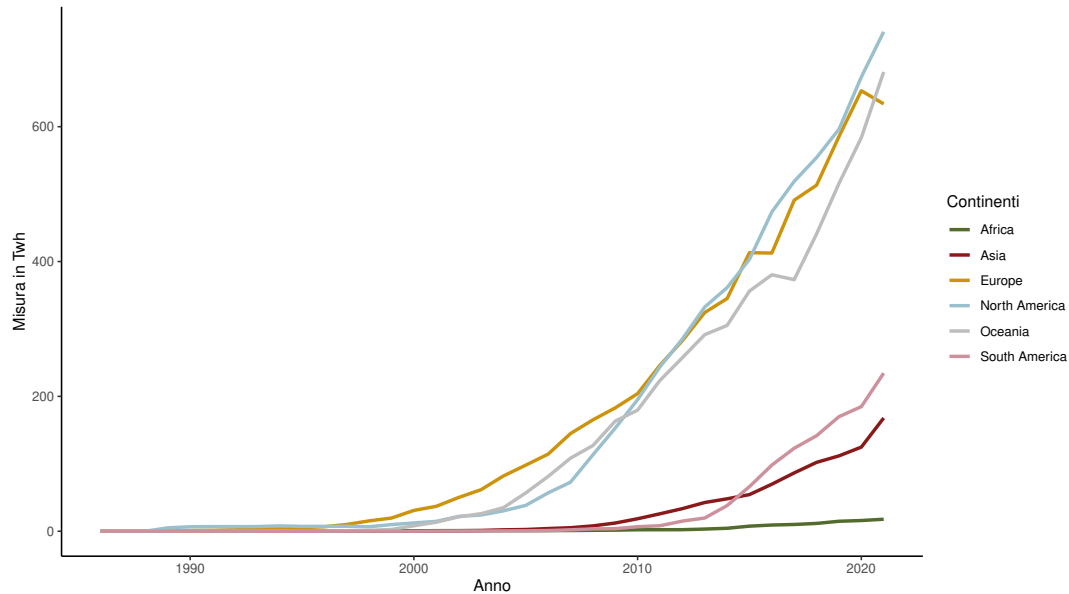
## 1.2 I dati

I dati utilizzati per le analisi riportate nei capitoli successivi sono stati reperiti dal sito Our World in Data (OWD). L'insieme di dati che verrà presentato è stato costruito accorpando diverse informazioni derivanti da fonti quali *BP Statistical Review of World Energy*, *EIA*, l'agenzia statistica e analitica del Dipartimento dell'energia degli Stati Uniti d'America, e *Ember Climate*, un'associazione no-profit senza scopo di lucro, che si

batte per ridurre l'uso del carbone. Dal sito ufficiale di OWD, attraverso un'esplorazione dei grafici riguardanti la produzione pro capite di energie rinnovabili, si è potuto notare fin da subito che l'analisi condotta, si sarebbe dovuta concentrare solo su alcuni stati principalmente per due motivi: il primo perché non tutti gli stati producono tutte e tre le energie analizzate; il secondo perché per certi paesi la produzione è così bassa da non consentire un'analisi ragionevole. Per questa ragione, si sono considerati gruppi di stati e di numerosità differenti a seconda della fonte rinnovabile analizzata. Come anticipato, l'attenzione è centrata su tre fonti rinnovabili: eolica, solare e idroelettrica. Le analisi condotte si basano su intervalli temporali diversi a seconda della fonte analizzata, che per l'energia eolica e idroelettrica va dal 1985 al 2020 e per la solare dal 2000 al 2020.

### 1.2.1 Energia eolica

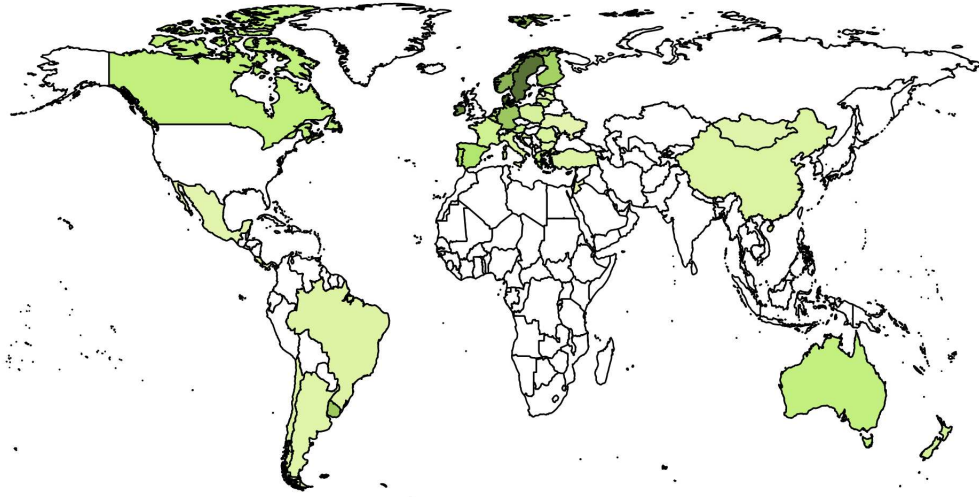
Per il caso dell'energia eolica, in Figura 1.1 sono rappresentate le serie storiche della produzione pro capite della stessa per continente. Si noti come, invece dei cinque continenti, siano state considerate sei regioni, in quanto l'America è stata suddivisa in due macro-aree, una per il nord e una per il sud, poiché si è pensato che l'utilizzo di fonti rinnovabili sia differente nel continente americano considerato nella sua interezza rispetto a una sua suddivisione in Nord e Sud, a causa della più lenta e difficile crescita nei paesi dell'America latina. Analizzando il grafico in Figura 1.1, sembrano esserci due gruppi: uno composto da Oceania, Nord America ed Europa, l'altro composto da Africa, Asia e America del Sud. Questo potrebbe esser giustificato dal fatto che i maggior produttori di energia eolica si trovino proprio nel primo gruppo. Si nota, inoltre, come per il primo gruppo l'uso di energia eolica sia iniziato ben prima rispetto al secondo gruppo. Per condurre le analisi nei capitoli seguenti, si è selezionato un gruppo di 40 Stati, elencati Tabella 1.1 e rappresentanti nella mappa in Figura 1.2; la



**Fig. 1.1.** Produzione di energia eolica pro capite per continente

differenti gradazioni di colore indicano che più questo è scuro, maggiore è la produzione di energia pro capite per quel determinato stato. L'arco temporale analizzato va dal 1985 al 2020.

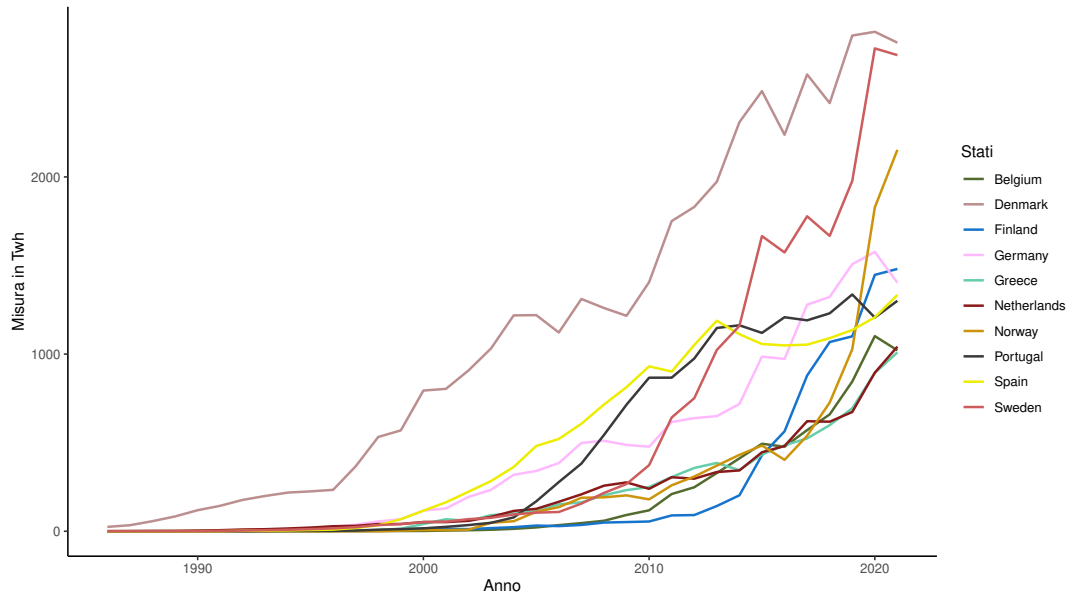
Analizzando nel dettaglio la serie storica dei paesi maggiori produttori di energia eolica pro capite in Figura 1.3, si può vedere come la Danimarca produca una quantità di energia maggiore rispetto agli stati considerati e come l'utilizzo di questa fonte energetica sia iniziato ben prima rispetto alle altre nazioni. La Svezia, invece, è entrata nel mercato dell'energia eolica più tardi rispetto alla Danimarca, ma il crescente interesse nell'investire in questa fonte rinnovabile ha fatto sì che nel 2020 la produzione pro capite si sia di molto avvicinata a quella danese. Lo stesso si può dire della Norvegia e della Finlandia. Rispetto a quanto detto in precedenza, si evince che i maggiori produttori di energia derivante dal vento siano i paesi nordici e quelli scandinavi, grazie soprattutto a condizioni meteorologiche favorevoli.



**Fig. 1.2.** paesi con produzione di energia eolica pro capite maggiore di 100kWh

Continente	Stato
Europa	Austria, Belgio, Bulgaria, Cipro, Danimarca, Finlandia, Francia, Germania, Grecia, Italia, Lituania, paesi Bassi, Polonia, Portogallo, Regno Unito, Repubblica Ceca Romania, Slovacchia, Slovenia, Spagna, Svezia, Svizzera, Ucraina, Ungheria
America del Sud	Argentina, Brasile, Cile, Costa Rica, Panama, Uruguay
America del Nord	Canada, Messico, Stati Uniti d'America
Oceania	Australia, Nuova Zelanda
Asia	Cina, Giordania, Mongolia, Turchia

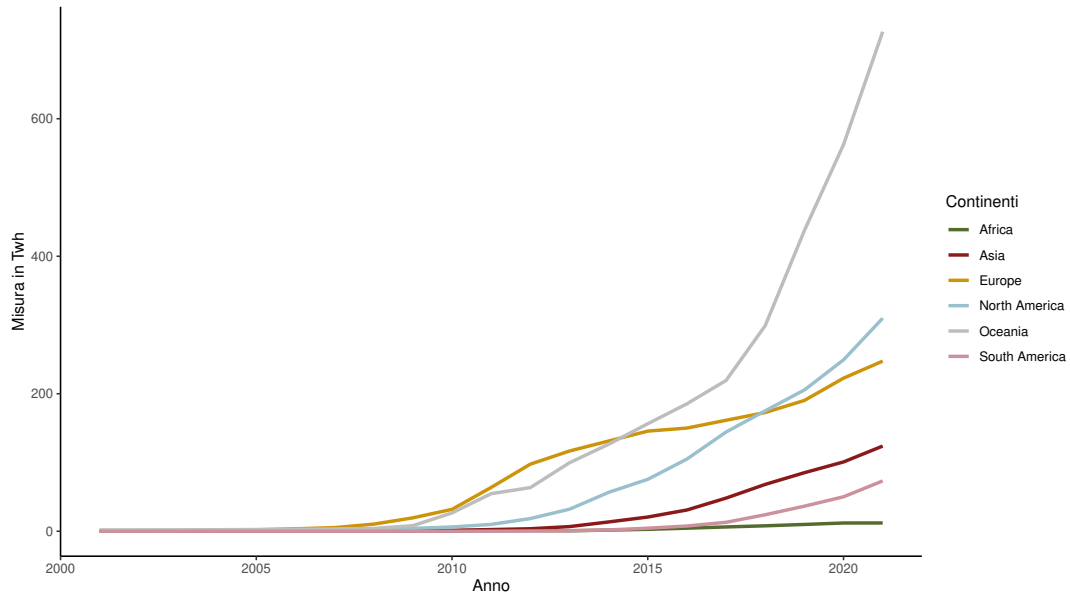
**Tab. 1.1.** paesi appartenenti al campione selezionato per l'energia eolica



**Fig. 1.3.** paesi con maggior produzione di energia eolica pro capite

## 1.2.2 Energia solare

La produzione di energia solare ha storia più breve rispetto alla produzione di energia da altre fonti rinnovabili. In Figura 1.4, il continente che sembra produrne di più è l'Oceania; questo, però, è dovuto anche al fatto che in Australia, il paese con superficie maggiore appartenente a questo continente, la popolazione è piuttosto bassa e, quindi, soddisfare la domanda di energia risulta sicuramente più facile proprio perché ci sono meno abitanti. Situazione opposta si riscontra in Asia, continente che ospita stati tra i più popolati al mondo, come Cina e India, e per i quali sarebbe necessaria una quantità di energia di molto superiore al caso dell'Oceania. Per quanto riguarda l'Europa, invece, si vuole evidenziare come sembrerebbe essere il continente pioniere nell'utilizzo di energia solare. In Figura 1.5 è rappresentato il gruppo di stati selezionato, composto da 46 paesi, maggiori produttori di energia solare e in Tabella 2.5 la suddivisione degli

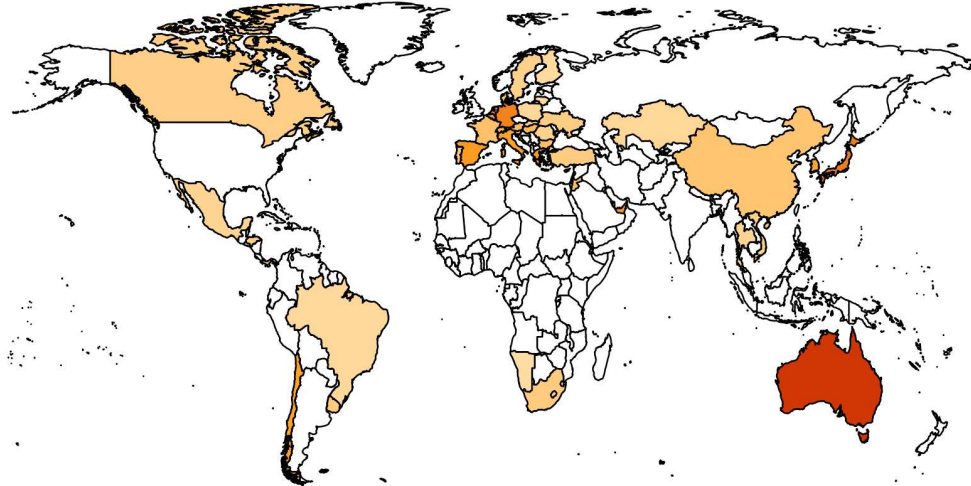


**Fig. 1.4.** Produzione di energia solare pro capite per continente

stessi per continente. Le serie storiche analizzate si estendono per un intervallo che inizia nel 2000 e termina nel 2020.

La Germania, in Figura 1.6 indicata con il color rosa, è stata per più di quindici anni il paese dove si produceva più energia solare pro capite. Nell'ultimo periodo analizzato, invece, è stata superata di gran lunga dall'Australia, seguita da Giappone, paesi Bassi e Israele. Analizzando la serie storica di produzione di energia solare non pro capite in Figura 1.7, si nota però che il paese che produce in assoluto più kWh di energia è il Giappone, seguito dalla Germania e dall'Australia; è quindi uno scenario completamente diverso rispetto a quanto evidenziato precedentemente considerando la produzione pro capite. Infatti, in Australia gli abitanti sono un quinto rispetto al Giappone, quindi soddisfare la domanda di energia solare nel paese oceanico risulta essere più fattibile.

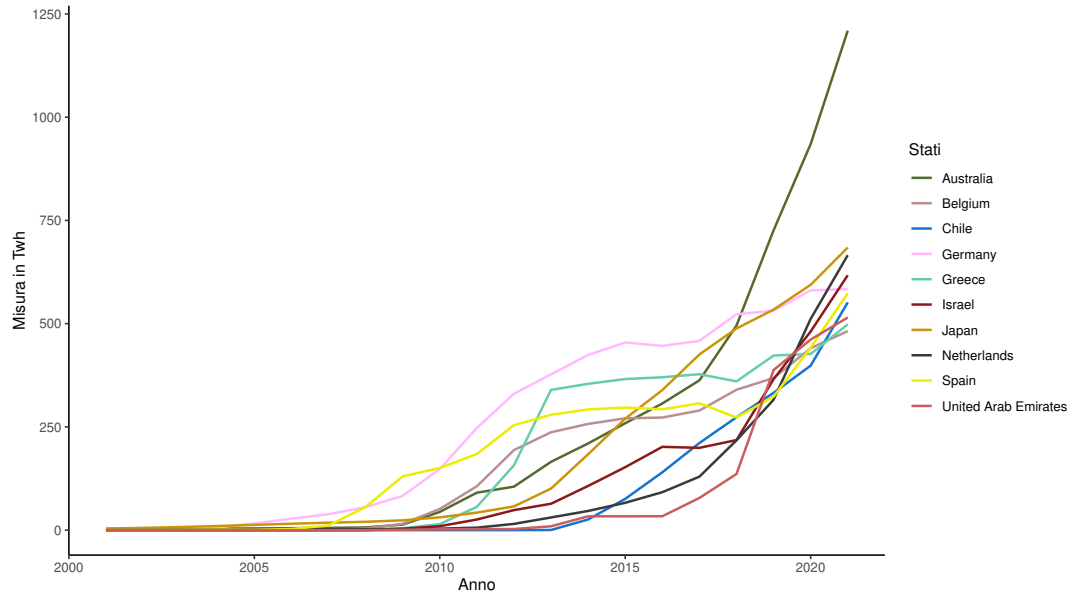




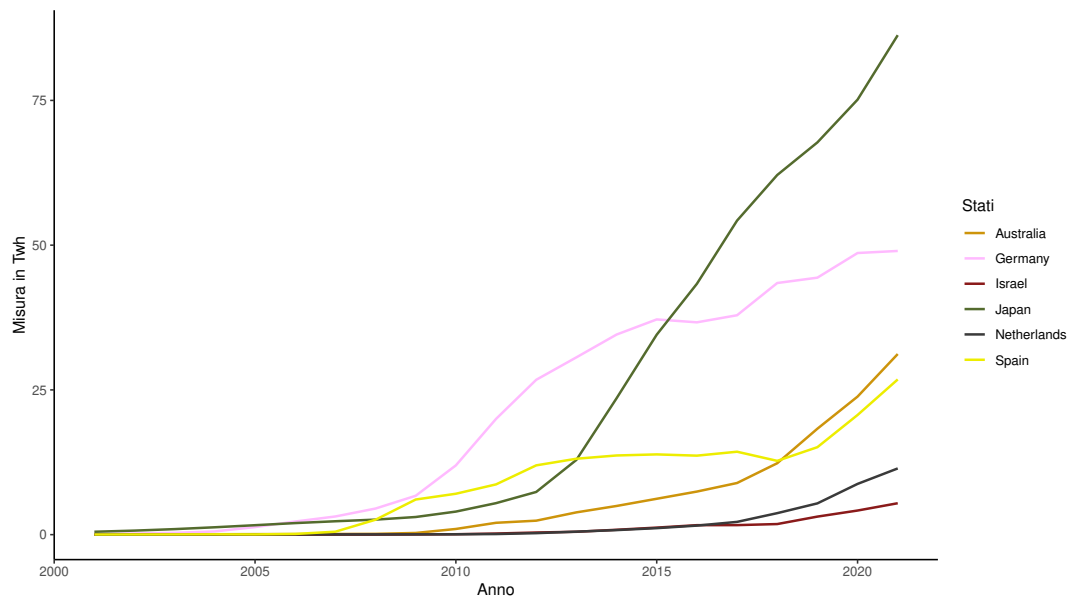
**Fig. 1.5.** paesi con produzione di energia solare pro capite maggiore di 50kWh

Continente	Stato
Europa	Austria, Belgio, Bulgaria, Cipro, Croazia, Danimarca, Estonia, Finlandia, Francia, Germania, Grecia, Irlanda, Italia, Lettonia, Lituania, Lussemburgo, Norvegia, paesi Bassi, Polonia, Portogallo, Regno Unito, Romania, Spagna, Svezia, Ucraina
America del Sud	Brasile, Cile, El Salvador, Honduras, Panama, Uruguay
America del Nord	Canada, Messico, Stati Uniti d'America
Oceania	Australia
Asia	Cina, Corea del Sud, Emirati Arabi Uniti, Giappone, Giordania, Israele, Kazakistan, Thailandia, Turchia, Vietnam
Africa	Namibia, Sud Africa

**Tab. 1.2.** paesi appartenenti al campione selezionato per l'energia solare



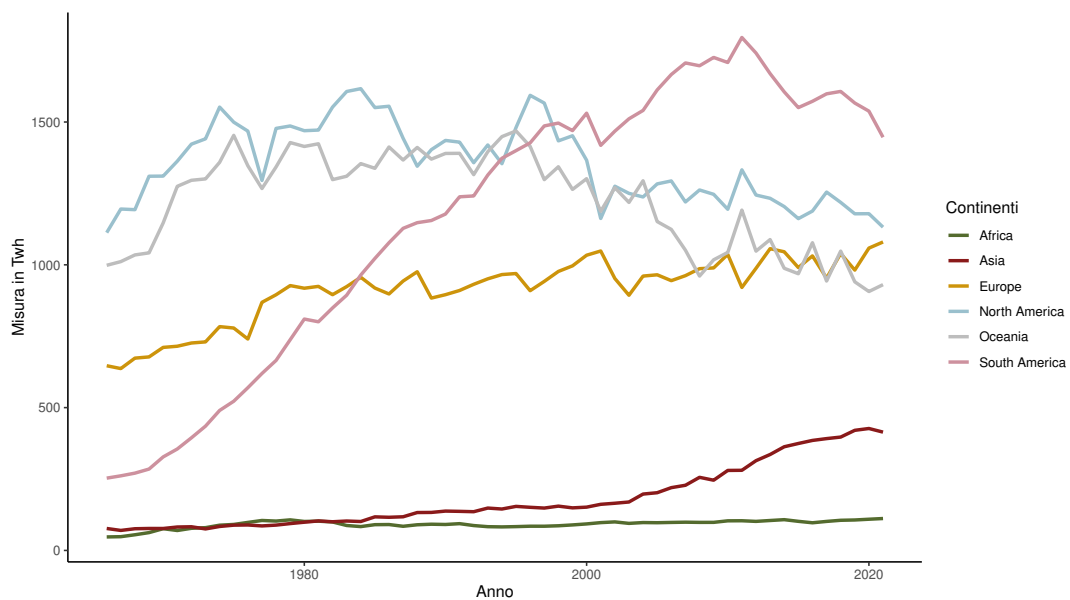
**Fig. 1.6.** paesi con maggior produzione di energia eolica pro capite



**Fig. 1.7.** paesi con maggior produzione di energia eolica non pro capite

### 1.2.3 Energia idroelettrica

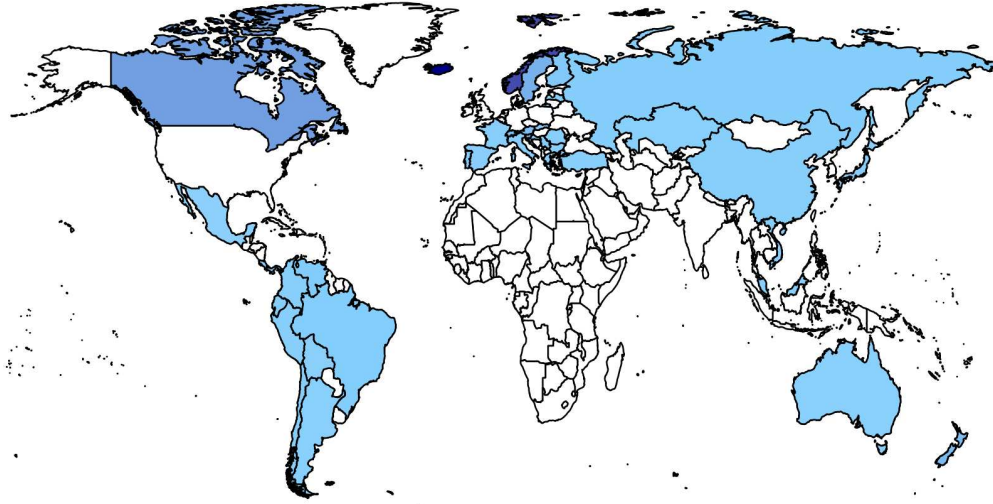
La storia dell'energia idroelettrica fonda le sue radici ben prima rispetto alle altre due fonti rinnovabili analizzate. Grazie ad una maggiore disponibilità di dati, si riportano in Figura 1.8 le serie storiche dei continenti a partire dal 1965. Si noti come Europa, America del Sud e America del Nord abbiano un andamento abbastanza simile tra loro, che si differenzia solo per la quantità di energia prodotta. Ciò che invece salta subito all'occhio è l'America del Sud, per la quale la produzione di energia è caratterizzata da un trend crescente. Del tutto diversa è la dinamica del processo di produzione di energia pro capite per Africa e Asia, per le quali, nonostante abbiano iniziato ad utilizzare l'idroelettrico a partire almeno dal 1965, i volumi di produzione rimangono piuttosto limitati.



**Fig. 1.8.** Produzione di energia idroelettrica pro capite per continente

Per l'energia idroelettrica sono stati selezionati 40 Stati, rappresentati in Figura 1.9

ed elencati in Tabella 2.6. L'intervallo temporale analizzato va dal 1985 al 2020.



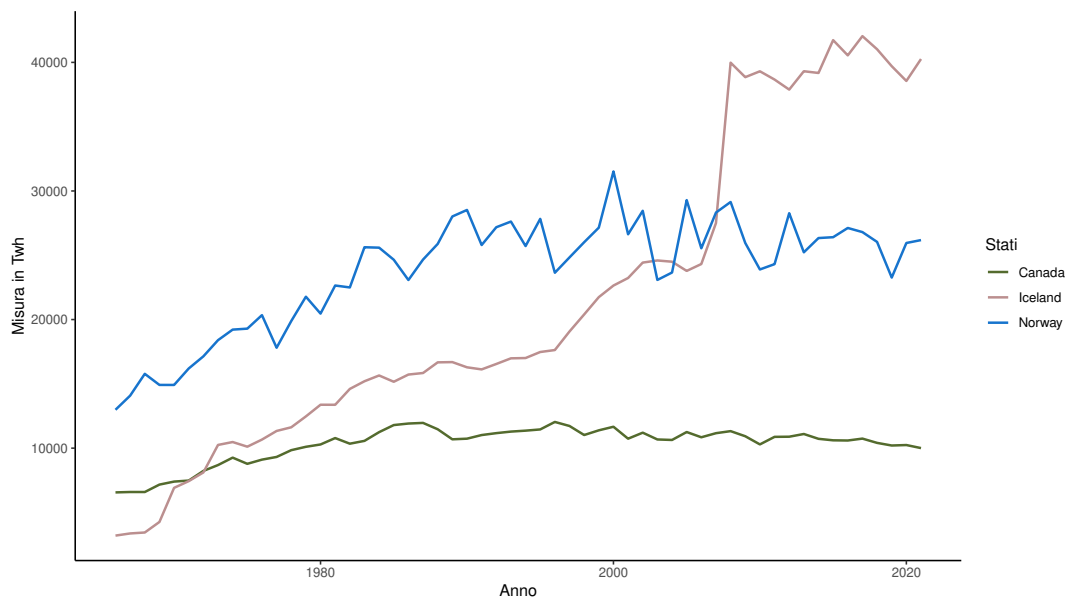
**Fig. 1.9.** paesi con produzione di energia idroelettrica pro capite maggiore di 250kWh

In Figura 1.10 sono rappresentati i maggiori produttori di energia idroelettrica pro capite, ovvero Canada, Islanda e Norvegia. L'Islanda ha registrato negli anni un trend crescente, con una crescita esponenziale soprattutto intorno al 2005. A differenza delle serie storiche delle altre due energie analizzate, si noti come quella dell'energia idroelettrica sia visivamente differente, questo perché è influenzata da condizioni meteorologiche diverse rispetto al solare e all'eolico, tra le quali la siccità. Se, ad esempio, per produrre più energia solare è sufficiente installare più pannelli fotovoltaici in una zona ad alta insolazione, lo stesso non è possibile per l'energia idroelettrica, in quanto, se la portata di un fiume si riduce drasticamente, questo fenomeno porterà a una diminuzione della produzione di energia.

Visto il trend crescente di produzione di energia idroelettrica in America Latina, in Figura 1.11 vengono rappresentate le nazioni appartenenti a questa macro-area, alle quali sono state aggiunte anche Costa Rica e Messico. Tra queste regioni, il Venezuela è

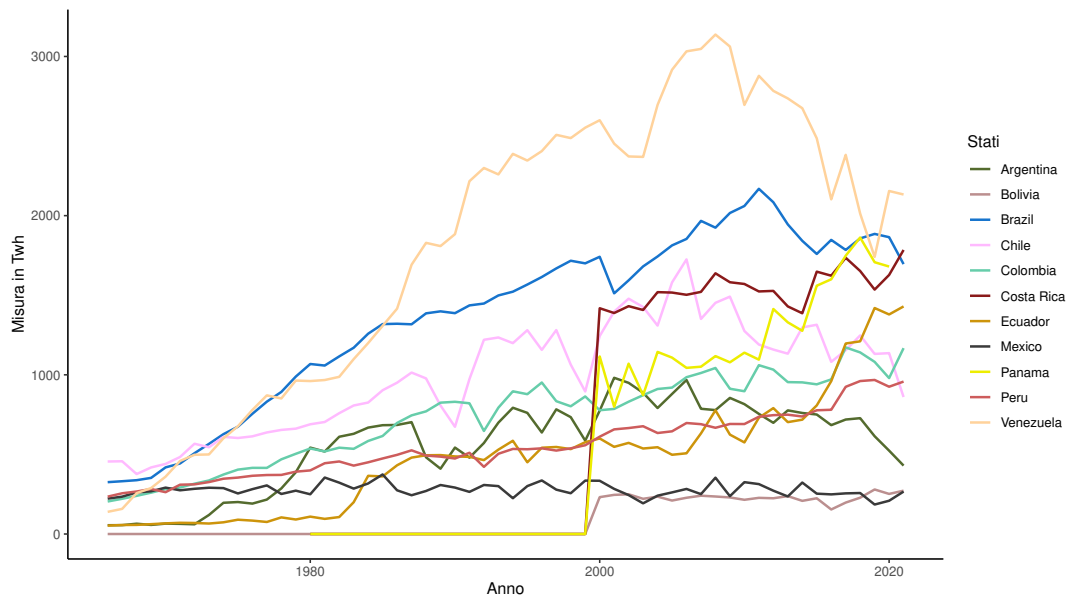
Continente	Stato
Europa	Bulgaria, Croazia, Finlandia, Francia, Grecia, Islanda, Italia, Lettonia, Norvegia, Portogallo, Romania, Russia, Serbia, Slovacchia, Slovenia, Spagna, Svezia, Svizzera, Ucraina
America del Sud	Argentina, Bolivia, Brasile, Cile, Colombia, Costa Rica, Ecuador, Panama, Perù
America del Nord	Canada, Messico, Stati Uniti d'America
Oceania	Australia, Nuova Zelanda
Asia	Cina, Giappone, Kazakistan, Malesia, Thailandia, Turchia, Vietnam

**Tab. 1.3.** paesi appartenenti al campione selezionato per l'energia idroelettrica



**Fig. 1.10.** paesi con maggior produzione di energia idroelettrica pro capite

il maggior produttore, ma un paese che sicuramente vale la pena citare è la Costa Rica; la sua entrata nel mercato idroelettrico è arrivata ben più tardi rispetto ad altri stati, ma in pochi anni è riuscito a sfruttare al meglio le risorse idriche tanto da arrivare ad un livello di produzione così alto da soddisfare quasi interamente la domanda di energia del paese. Si noti, però, come molte delle serie storiche rappresentate in Figura 1.11 siano caratterizzate da un trend crescente, per alcuni stati più rapido rispetto ad altri.



**Fig. 1.11.** paesi dell'America Latina con maggior produzione di energia idroelettrica pro capite

### 1.3 Modelli per spiegare e per prevedere

L'idea che sta alla base di questa tesi è quella di *spiegare* la crescita della produzione pro capite di energia rinnovabile e di *prevedere* il suo comportamento nel tempo. L'idea di questo dublice obiettivo, deriva dall'articolo *To explain or to predict?* (Shmueli, 2010), dove viene chiarita la differenza tra modelli per spiegare un fenomeno e modelli per prevedere. Per questo motivo, si ritiene utile riassumere brevemente il contributo di questo articolo.

In generale, la modellazione statistica è uno strumento utilizzato per sviluppare e testare teorie attraverso la previsione, la descrizione e la spiegazione della causalità. In particolare, la modellazione esplicativa e quella predittiva sono spesso fuse tra loro; infatti, entrambe sono necessarie per generare e testare ipotesi, anche se ognuna gioca un ruolo diverso. Nello specifico, per modellazione esplicativa si intende l'uso di modelli statistici per testare la causalità dell'evento oggetto di studio. In questo caso, le ipotesi di ricerca sono date in termini di costrutti teorici piuttosto che di variabili misurabili. A differenza di queste ultime, i costrutti sono astrazioni che descrivono un fenomeno di interesse teorico e possono essere osservabili o non osservabili. La modellazione predittiva, invece, è il processo di applicazione dei modelli statistici o algoritmi di *data mining* ai dati, con l'obiettivo di prevedere nuove o future osservazioni. Lo scopo della previsione non stocastica è prevedere i valori di output  $Y$  per delle nuove osservazioni, dati i loro valori di input  $X$ . La differenza sostanziale tra questi due approcci risiede nel fatto che i dati misurabili non sono rappresentazioni accurate dei costrutti a loro sottostanti. Per spiegare questa disparità formalmente, si consideri che il costrutto  $\mathcal{X}$  causi il costrutto  $\mathcal{Y}$  attraverso la funzione  $F$ , ovvero  $\mathcal{Y} = \mathcal{F}(\mathcal{X})$ . Siano  $\mathbf{X}$  e  $\mathbf{Y}$  operazionalizzazioni di  $\mathcal{X}$  e  $\mathcal{Y}$  rispettivamente. L'operazionalizzazione di  $F$  in un modello statistico  $f$ , come ad esempio  $E(Y) = f(X)$ , è attuata considerando le dinamiche delle variabili

oggetto di studio. In un contesto predittivo, si considerano solo  $X, Y$  e  $f$ . La disparità tra i due approcci cresce poiché l'obiettivo della modellazione esplicativa è di ugualgiare  $f$  con  $\mathcal{F}$  il più fedelmente possibile anche l'inferenza statistica applichi le ipotesi teoriche. I dati  $\mathbf{X}$ ,  $\mathbf{Y}$  sono strumenti necessari per stimare  $f$ , la quale è utilizzata per testare le ipotesi di causalità. Invece, in un contesto predittivo, le entità di interesse sono  $\mathbf{X}$  e  $\mathbf{Y}$ , e la funzione  $f$  è utilizzata come strumento per generare previsioni dei nuovi valori di  $\mathbf{Y}$ . Queste differenze si manifestano principalmente in quattro aspetti:

- *causa-associazione*: nei modelli esplicativi  $f$  rappresenta la funzione causale sottostante e si assume che  $\mathbf{X}$  causi  $\mathbf{Y}$ ; nei modelli predittivi  $f$  cattura l'associazione tra  $\mathbf{X}$  e  $\mathbf{Y}$ .
- *teoria-dati*: nei modelli esplicativi  $f$  è costruita basandosi su  $\mathcal{F}$  in modo tale da supportare l'interpretazione della relazione stimata tra  $\mathbf{X}$  e  $\mathbf{Y}$ . Nei modelli predittivi  $f$  è spesso costruita a partire dai dati.
- *retrospettivo-prospettivo*: i modelli predittivi sono prospettivi, mentre quelli esplicativi sono retrospettivi.
- *distorsione-varianza*: il focus nei modelli esplicativi è minimizzare la distorsione, per ottenere la rappresentazione più accurata della teoria sottostante; al contrario, i modelli predittivi cercano di minimizzare la combinazione di varianza e distorsione.

Dopo aver testato diversi modelli, sia nel caso predittivo che in quello esplicativo, è necessario scegliere quello che più è in grado di spiegare il fenomeno di interesse. A questo proposito, subentrano altre disparità tra i due approcci considerati, ovvero la convalida dei modelli e la valutazione degli stessi. Per i modelli esplicativi, la convalida



si articola in due fasi: la convalida del modello, che accerta che  $f$  rappresenti adeguatamente  $F$ , e l'adattamento del modello, che convalida che  $\hat{f}$  si adatti ai dati  $X, Y$ . Al contrario, nei modelli predittivi la convalida si focalizza sulla generalizzazione, ovvero la capacità di  $\hat{f}$  di prevedere nuovi dati  $\{X_{new}, Y_{new}\}$ .

Per quanto riguarda la valutazione del modello, la massima priorità in termini di performance del modello per il caso dei modelli esplicativi è valutare il potere esplicativo, il quale misura la forza della relazione indicata da  $\hat{f}$ . Si è soliti utilizzare il valore di  $R^2$  e la statistica test  $F$ . L'obiettivo dei modelli predittivi è, invece, l'accuratezza predittiva, ovvero quanto bene si adatta  $\hat{f}$  ai nuovi dati. Normalmente, si usano metriche *out-of-sample* o alcune loro approssimazioni.

È importante sottolineare che se l'accuratezza predittiva può essere utilizzata sia per scopi predittivi che esplicativi, lo stesso non si può dire dell'accuratezza esplicativa, la quale non può essere impiegata per valutare modelli predittivi. Questo è dovuto al fatto che misure come l' $R^2$  e la statistica test  $F$  forniscono indicazioni circa il livello di associazione, ma non di casualità.

Per raggiungere il duplice obiettivo descritto in questa sezione, si è deciso di utilizzare in questa tesi due metodologie: per la parte esplicativa i modelli di diffusione, in particolare il modello di Bass (Bass, 1969), trattato nel Capitolo 2; per la parte predittiva i modelli per dati funzionali, in particolare il modello simultaneo, trattato nel Capitolo 3 (Ramsay and Silverman, 2005).

# Capitolo 2

## Modelli di diffusione

Il crescente interesse per le energie rinnovabili registrato negli ultimi anni e la continua crescita nell'adozione delle stesse, rende particolarmente interessante cercare di spiegarne l'evoluzione nel tempo tramite modelli di diffusione. Alla base di questa scelta c'è l'intenzione di caratterizzare il processo di diffusione attraverso alcuni parametri, i quali permettano di ricavare informazioni riguardo a come e con quale velocità si sta evolvendo questo processo nell'insieme di paesi scelti. In questo senso, l'approccio scelto ha valore esplicativo.

### 2.1 Modello di Bass

Dopo aver introdotto i dati nel Capitolo 1, si passa ad una loro analisi tramite modelli di diffusione. A tale scopo, si farà riferimento ad un specifico modello molto noto in letteratura, ovvero il modello di Bass (Bass, 1969). Si tratta di un modello la cui struttura matematica è collegata ai modelli impiegati per la descrizione di eventi epidemiologici; infatti, ciò che viene indicata come un'innovazione si propaga in un

sistema sociale attraverso un comportamento imitativo tra persone, proprio come una malattia si propaga nella popolazione attraverso il contagio (Muller et al., 2009).

Il modello di Bass (BM) è definito dalla seguente equazione differenziale di primo ordine:

$$z'(t) = \left[ p + q \frac{z(t)}{m} \right] [m - z(t)] \quad t > 0 \quad (2.1)$$

dove, nel caso in esame,  $z'(t)$  rappresenta la variazione della produzione pro capite di energie rinnovabili al tempo  $t$  ed è data da:

- $p$  coefficiente di innovazione;
- $q$  coefficiente di imitazione;
- $m$  mercato potenziale raggiungibile;
- $z(t)$  produzione cumulata al tempo  $t$ ;
- $m - z(t)$  mercato residuo.

La soluzione in forma chiusa del modello di Bass, poiché  $z(t)/m = 0$ , è definita come:

$$y(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \quad t > 0. \quad (2.2)$$

Il termine  $y(t)$  può essere visto come una funzione di ripartizione,  $0 < y(t) < 1$ , e dipende dai parametri  $p$  e  $q$ , la cui grandezza determina la velocità di crescita, fino alla saturazione. Visto che  $z(t) = y(t)m$  allora la soluzione in forma chiusa diventa:

$$z(t) = m \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \quad t > 0 \quad (2.3)$$

Si noti come il mercato potenziale  $m$  agisca come parametro di scala del processo di diffusione e venga assunto costante.

Si vuole sottolineare che il modello di Bass è un modello parsimonioso, infatti, richiede la stima di soli tre parametri,  $m$ ,  $p$  e  $q$ , i quali permettono di descrivere efficacemente il processo di diffusione di un prodotto o una tecnologia, prevedendo le varie fasi del suo ciclo di vita, ovvero la crescita, la maturità e il declino. Inoltre, è facilmente applicabile ai dati, visto che sono necessarie serie storiche aggregate (Guidolin, 2023).

## 2.2 Applicazione del modello di Bass

Nelle sottosezioni di seguito si presenta un'analisi separatamente per i coefficienti  $m$ ,  $p$  e  $q$  dei modelli stimati, derivanti dall'applicazione del BM, per le tre fonti di energia rinnovabile scelte, ovvero eolica, solare e idroelettrica. Per ogni coefficiente e per ogni energia rinnovabile, si propone un grafico a barre dei primi dieci paesi per i quali i valori di  $m$ ,  $p$  e  $q$  sono più elevati e una possibile interpretazione dei valori assunti dai coefficienti stessi. A tale scopo, si ricorda che sono state analizzate le serie storiche della produzione pro capite per ogni fonte di energia, non la produzione totale di ognuna di esse. Nelle Tabelle 2.1, 2.2 e 2.3 vengono riportati i valori associati ai coefficienti  $m$ ,  $p$  e  $q$ . L'asterisco riportato a fianco ad alcuni dei valori indica la significatività dello stesso ad almeno il 10%

Capitolo 2: Modelli di diffusione

---

Stato	$m$	$p$	$q$	Stato	$m$	$p$	$q$
Argentina	40138	0	0.48394*	Italia	4482*	0.000034*	0.27643*
Australia	10851*	0.000032*	0.22363*	Lettonia	1349*	0.000105*	0.22591*
Austria	13618*	0.00008*	0.25829*	Lituania	5719*	0.000001*	0.36123*
Belgio	11529*	0.000005*	0.30731*	Lussemburgo	52914*	0.000076	0.13723*
Brasile	2098	0	0.51549*	Messico	1297*	0.0000003*	0.40169*
Bulgaria	2212*	0.0000005*	0.43308*	Mongolia	1601	0	0.47002*
Canada	12521*	0.000008*	0.30258*	Norvegia	301303*	0.000005	0.19274*
Cile	2299	0	0.46397*	N.Zelanda	7399*	0.000041*	0.29346*
Cina	3238*	0.000009*	0.36356*	paesi Bassi	17561*	0.000212*	0.17284*
Cipro	1956*	0	0.47867*	Panama	1031	0	0.70969*
Costa Rica	25484	0.000063	0.16068*	Polonia	3875*	0.0000003*	0.41892*
Croazia	3532*	0.000002*	0.41645*	Portogallo	17383*	0.000019*	0.31056*
Danimarca	69081*	0.001003*	0.14273*	Regno Unito	15471*	0.000015	0.26692*
Estonia	6460*	0.000004*	0.35434*	Romania	3139	0	0.52798*
Finlandia	23741*	0.0000004*	0.35094*	Spagna	19275*	0.000133*	0.25109*
Francia	6540*	0.000013*	0.29066*	Svezia	29740*	0.000011*	0.28902*
Germania	31601*	0.000223*	0.17913*	Turchia	2805*	0.0000008*	0.37913*
Giordania	484	0	0.76499*	Usa	14071*	0.000036*	0.25445*
Grecia	12602*	0.000127*	0.20562*	Ucraina	907*	0.000001*	0.32793*
Irlanda	32980*	0.000071*	0.22841*	Uruguay	9359	0	0.66756*

**Tab. 2.1.** Valori dei coefficienti  $m$ ,  $p$  e  $q$  derivanti dall'applicazione del BM per l'energia eolica

Capitolo 2: Modelli di diffusione

Stato	$m$	$p$	$q$	Stato	$m$	$p$	$q$
Australia	15290*	0.000091*	0.33769*	Lituania	289*	0.000035	0.50736*
Austria	1889*	0.000123*	0.40976*	Messico	321	0	0.98164*
Belgio	3559*	0.000215*	0.41912*	Namibia	152*	0.000063	0.52532*
Brasile	167	0	1.24259*	paesi Bassi	79373*	0.0000011	0.42374*
Bulgaria	1841*	0.000057*	0.52942*	Panama	322	0	0.94364*
Canada	957*	0.000059*	0.49429*	Polonia	4493	0	0.83211*
Cile	2081*	0.0000004	0.72636*	Portogallo	1493*	0.000282*	0.35174*
Cina	1077*	0.0000009	0.66117*	Regno Unito	1371*	0.000006*	0.61774*
Cipro	1687*	0.000039*	0.47888*	Rep.Ceca	2300*	0.000269*	0.45422*
Corea del Sud	6473*	0.000032*	0.38338*	Romania	695*	0.000039	0.67757*
Danimarca	1440*	0.000042*	0.50348*	Slovacchia	1076*	0.000169*	0.47379*
El Salvador	368	0	0.88923*	Slovenia	1349*	0.000099*	0.48662*
EAU	5829*	0	65327*	Spagna	3805*	0.000693*	0.37214*
Finlandia	15181*	0	0.43548*	Sud Africa	826*	0.000002*	0.63476*
Francia	1632*	0.000115*	0.43675*	Svezia	14462*	0	0.48225*
Germania	5909*	0.000551*	0.37137*	Svizzera	2516*	0.000081*	0.42899*
Giappone	5764*	0.000103*	0.40374*	Thailandia	556*	0.000016*	0.54831*
Giordania	934	0	0.99771*	Turchia	501	0	1.09133*
Grecia	3624*	0.000065*	0.51625*	Usa	3291*	0.000036*	0.45667*
Honduras	653*	0	0.81216*	Ucraina	43484*	0	0.46516*
Israele	4097*	0.000054*	0.41664*	Ungheria	10129*	0	0.64902*
Italia	4482*	0.000034*	0.0.27643*	Uruguay	628	0	0.91429*
Kazakistan	9375	0	0.84293*	Vietnam	899	0	1.37196*

**Tab. 2.2.** Valori dei coefficienti  $m$ ,  $p$  e  $q$  derivanti dall'applicazione del BM per l'energia solare

Stato	$m$	$p$	$q$	Stato	$m$	$p$	$q$
Argentina	37325*	0.0131*	0.0605*	Lettonia	776906*	0.00151	0.00774*
Australia	43819*	0.0203*	0.0217*	Malesia	660148*	0.00031	0.00234*
Austria	373242*	0.0123*	0.0225*	Messico	39857*	0.00739	0.00511*
Bolivia	4867*	0.00057*	0.23388*	Norvegia	2402941*	0.01051*	0.01716*
Brasile	143270*	0.00864*	0.03487*	N.Zelanda	410787*	0.01648*	0.01522*
Bulgaria	287226*	0.00075	0.03354*	Panama	32899*	0.00061*	0.19599*
Canada	1472400*	0.00781*	0.00697*	Perù	421918*	0.00104	0.00213*
Cile	62037*	0.01267*	0.06131*	Portogallo	105583*	0.00854*	0.01823*
Cina	433736*	0.00013	0.08368*	Romania	48691*	0.00982*	0.04546*
Colombia	329670*	0.00218	0.01489*	Russia	550735*	0.00199	0.00556*
Costa Rica	4867*	0.00057*	0.23388*	Serbia	39949*	0.00713*	0.11213*
Croazia	54242*	0.00618*	0.11634*	Slovacchia	31735*	0.00993*	0.08921*
Ecuador	749753*	0.00053	0.02491*	Slovenia	78582*	0.00689*	0.09774*
Finlandia	1112344*	0.00228	0.00272*	Spagna	45810*	0.01496*	0.02257*
Francia	64457*	0.01799*	0.01853*	Svizzera	367303*	0.01357*	0.01564*
Giappone	172822*	0.00398	0.00076*	Svezia	753413*	0.01089*	0.00733*
Grecia	47833*	0.00454*	0.03795*	Turchia	184547*	0.00212	0.02483*
Islanda	20125402*	0.00063	0.03969*	Venezuela	121177*	0.01956*	0.06739*
Italia	238131*	0.00287	0.00655*	Vietnam	482515*	0.00009	0.09111*
Kazakistan	35367*	0.01062*	0.03421*	Usa	67633*	0.01713*	0.01096*

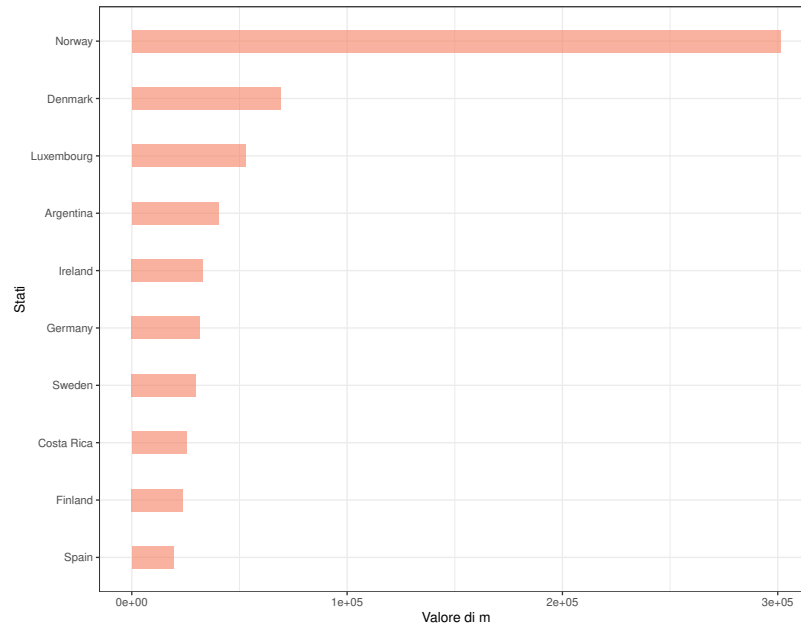
**Tab. 2.3.** Valori dei coefficienti  $m$ ,  $p$  e  $q$  derivanti dall'applicazione del BM per l'energia idroelettrica

### 2.2.1 Analisi del coefficiente relativo al mercato potenziale

La prima analisi proposta fa riferimento alla stima del parametro  $m$  nel caso dell'energia eolica. In Figura 2.1 si nota al primo posto la Norvegia, seguita dal Lussemburgo e dalla Danimarca. Il risultato che desta più sorpresa è quello relativo al valore del coefficiente  $m$  attribuito al Lussemburgo; ci si poteva aspettare che questo valore fosse ben al di sotto di quello di altri paesi, non essendo la nazione più grande a disposizione dal punto di vista della dimensione del mercato. Approfondendo i risultati ottenuti, si è notato come ad  $m$  sia associato un p-value al di sopra del 10%, il quale indica la non significatività del coefficiente. Le conclusioni che se ne possono ricavare devono quindi essere tratte con molta accortezza. Si vuole comunque sottolineare che il Lussemburgo si è prefissato come obiettivo il raggiungimento di un'autonomia energetica pari al 25% dalle rinnovabili entro il 2030 e, per questo, ha deciso di investire 30 miliardi di euro nell'energia eolica offshore, creando un'isola energetica nel mar del Nord. Troviamo una situazione simile anche per la Norvegia, la quale ha investito anch'essa sull'eolico galleggiante; si stima che, non appena le installazioni saranno a pieno regime, permetteranno di coprire almeno il 35% del fabbisogno elettrico del paese.

Infine, la Danimarca ha iniziato a investire nell'eolico a metà degli anni Ottanta, un periodo storico caratterizzato da una crisi petrolifera, da timori per il nucleare e dalla necessità di trovare una soluzione alle elevate emissioni di CO<sub>2</sub> delle centrali elettriche a carbone. Fortunatamente, la sensibilità dei cittadini ha sposato incentivi e sovvenzioni statali e il crescente interesse per questa realtà rinnovabile ha fatto sì che, ad oggi, la Danimarca sia il primo paese europeo per consumo di energia eolica.



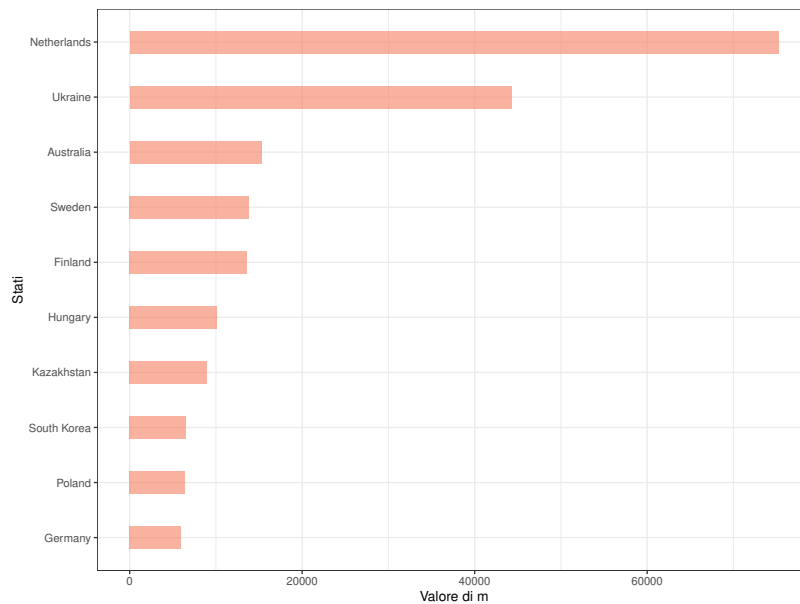


**Fig. 2.1.** Grafico a barre del coefficiente  $m$  per l'energia eolica

I risultati ottenuti per l'energia solare sono riportati in Figura 2.2; in prima posizione sono rappresentati i paesi Bassi, seguiti dall'Ucraina e dall'Australia. Considerando i paesi Bassi, la radiazione solare dell'estate 2020 ha permesso di generare tra giugno e agosto il 40% di energia in più con i pannelli solari rispetto allo stesso periodo dell'anno scorso. Questo risultato è frutto sia dell'aumento dell'esposizione al sole sia della crescita della capacità fotovoltaica installata. Per quanto riguarda l'Australia, sulla costa occidentale è stato avviato un progetto per la costruzione del più grande impianto di energia rinnovabile del mondo, tra eolico e solare. Infatti, in un paese come questo, è possibile installare pannelli solari e fotovoltaici più facilmente rispetto ad altri paesi, essendo uno degli stati meno densamente popolati e avendo così a disposizione più terre libere.

Per il caso dell'Ucraina, invece, non ci si sarebbe aspettati un valore così alto del

coefficiente relativo al mercato potenziale; analizzando più dettagliatamente i risultati ottenuti dall'applicazione del modello, si può affermare che questo coefficiente non sia utile ai fini dell'analisi, dato che l'intervallo di confidenza dello stesso comprende lo 0. Si vuole ricordare, inoltre, che i risultati riportati nel grafico in Figura 2.2 fanno riferimento all'intervallo temporale [2000, 2020], escludendo quindi l'ultimo biennio e, di conseguenza, la situazione attuale che sta vivendo l'Ucraina a causa della guerra. Nonostante questo, si può affermare che ciò che gioca a favore della produzione di energia solare in Ucraina è la sua latitudine, che permette molti giorni solari tra l'inizio della primavera e fino al tardo autunno. Infatti, considerando la quantità di radiazione solare per unità di superficie, l'Ucraina supera di gran lunga la Germania, riconosciuta come nazione leader in Europa per l'energia solare.

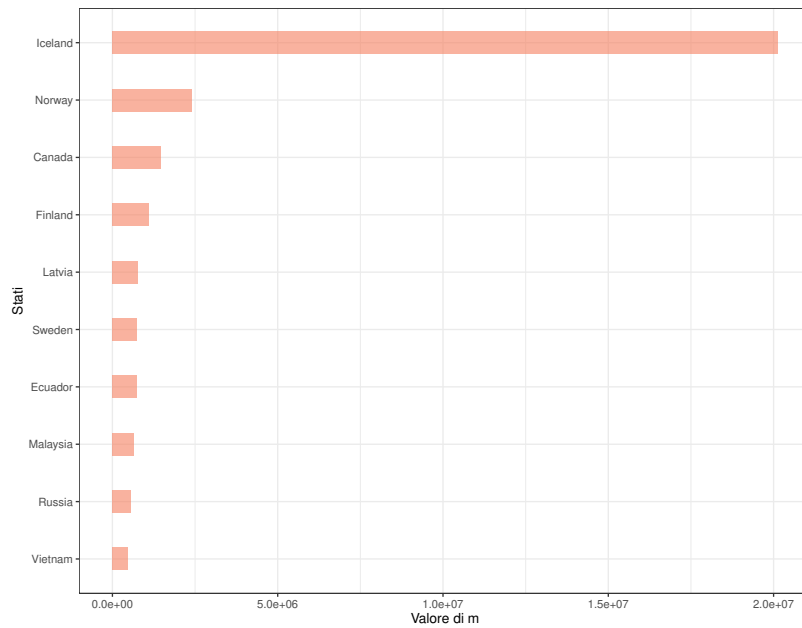


**Fig. 2.2.** Grafico a barre del coefficiente  $m$  per l'energia solare

Infine, in Figura 2.3 è rappresentato il coefficiente relativo al mercato potenziale dell'energia idroelettrica dei primi dieci paesi tra quelli considerati. Tra le prime posizioni

troviamo tre paesi nordici, ovvero Islanda, Norvegia e Finlandia.

Come per il caso del Lussemburgo citato prima, anche il coefficiente relativo al mercato potenziale dell'Islanda potrebbe considerarsi non propriamente adeguato a descrivere la situazione attuale. Analizzando nel dettaglio i risultati derivanti dall'applicazione del modello di Bass, si è notato come l'intervallo di confidenza del coefficiente in esame comprendesse lo 0 e, quindi, come questo non sia utile a descrivere il processo di diffusione, proprio come per il caso dell'Ucraina per l'energia solare. Si tratta perciò, anche in questa occasione, di un caso evidentemente molto particolare che va interpretato con estrema cautela.



**Fig. 2.3.** Grafico a barre del coefficiente  $m$  per l'energia idroelettrica

### 2.2.2 Analisi del coefficiente relativo agli innovatori

In questa sottosezione vengono riportati nei grafici a barre i coefficienti dei primi dieci paesi innovatori per tutte e tre le energie analizzate. Inoltre, si è deciso di rappresentare anche la serie storica di questi Stati per evidenziare il comportamento tipico degli innovatori. Ci si aspetta, infatti, che i processi di diffusione di questi siano iniziati ben prima rispetto a quelli degli imitatori e che siano, inoltre, in una fase piuttosto avanzata.

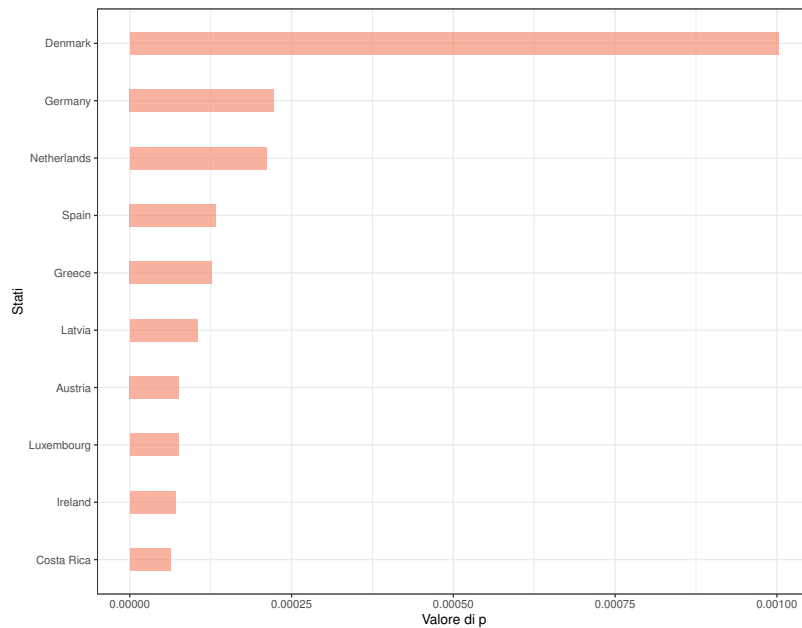
Per quanto riguarda l'energia eolica, la Danimarca, in Figura 2.4, mostra di essere lo stato più innovatore tra quelli considerati, seguita dalla Germania e dai paesi Bassi. La terra danese è sede di una brezza climatica costante, che consente l'utilizzo di grandi pale eoliche; proprio in Danimarca, infatti, la maggior parte di energia, pari al 48% di energia totale, deriva da fonte eolica.

I paesi Bassi sono un paese che sta investendo molto nell'eolico, tanto da volersi rendere autonoma dalle importazioni energetiche sfruttando la forza del vento. Proprio per questo prevede, entro il 2030, di triplicare la capacità di produzione energetica proveniente dall'eolico offshore. In Germania, invece, la presenza di vento, soprattutto durante i mesi autunnali e invernali, contribuisce alla sempre più crescente produzione di energia eolica; come per altri paesi citati in precedenza, anche la Germania ha l'obiettivo di quadruplicare la capacità eolica offshore totale entro il 2030.

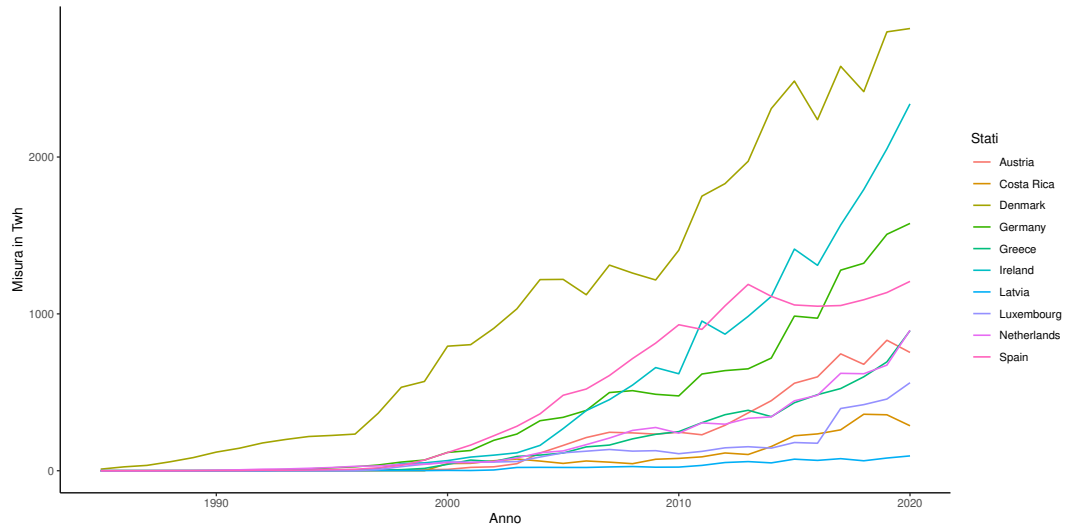
Si noti come il Costa Rica, pur essendo uno stato molto piccolo a livello di superficie, sia compreso tra i primi dieci stati innovatori. Questo paese ha dimostrato negli ultimi anni un forte orientamento verso il tema della sostenibilità e verso una quasi totale autonomia a livello di energie rinnovabili; a conferma di ciò, nel 2021 il 99.05% di energia è stato prodotto interamente da fonte rinnovabile. In particolare, per quanto riguarda l'energia eolica, nel 1996 venne installato il primo parco eolico e, fino ai giorni d'oggi,

la produzione di questa energia ha registrato una costante, seppure lenta, crescita. Attualmente, il 12.44% di energia è derivante da fonte eolica.

Nel grafico in Figura 2.5 è evidente come i gli stati considerati innovatori si siano mossi già a partire dagli anni novanta del Novecento verso il mercato dell'eolico. Il caso della Danimarca è sicuramente il più rilevante; emerge infatti come, negli anni considerati, abbia prodotto e continui a produrre una quantità di energia eolica pro capite ben al di sopra degli altri nove paesi e come si sia rivolta a questo mercato prima del 1990. Queste considerazioni chiariscono la prima posizione di tale stato nel grafico in Figura 2.4.



**Fig. 2.4.** Energia eolica: primi dieci paesi innovatori



**Fig. 2.5.** Serie storica dei primi dieci paesi innovatori per l'energia eolica

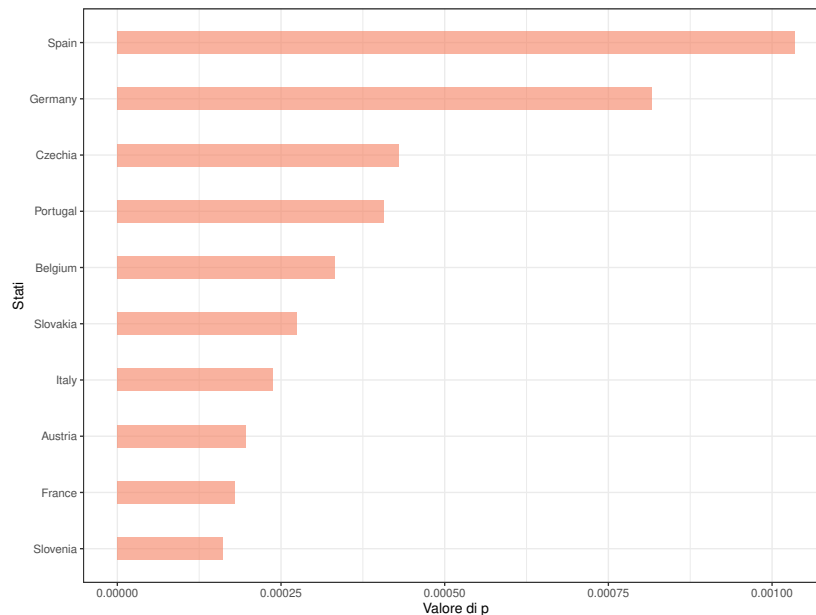
Sempre restando tra i paesi innovatori nell'utilizzo di energie derivanti da fonti rinnovabili, si passa ad analizzare il caso dell'energia solare, riportata in Figura 2.6. I primi 10 paesi sono tutti in Europa, a prova del fatto che in Europa ci sia un forte orientamento verso l'utilizzo esclusivo di fonti rinnovabili.

La Spagna sembra essere il paese più innovatore tra quelli considerati, dove il 9.85% di energia deriva proprio dal fotovoltaico. L'istituto di ricerca di mercato *IHS Markit* considera al quinto posto nel mondo la Spagna tra i mercati più interessanti nelle rinnovabili, grazie alla stabilità giuridica in questo campo. Oltre a questo, sono da tenere in considerazione l'alta insolazione e la grande superficie di terreni liberi.

Se dalla Spagna ci si poteva attendere un risultato simile, diverso è il caso della Germania, al secondo posto; questo paese, infatti, non gode delle stesse condizioni meteorologiche della terra spagnola, ma è riuscito negli anni a produrre una così elevata quantità di energia solare tanto da classificarsi al secondo posto tra gli innovatori. Ad oggi, la Germania si conferma il più grande mercato solare d'Europa, con una produzione di

circa 59 TWh nel 2022.

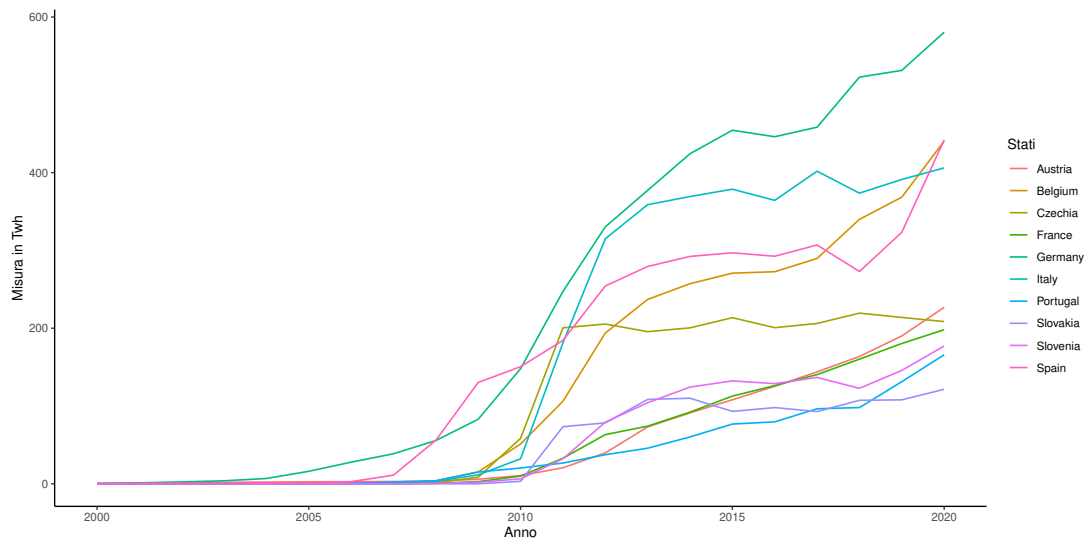
In terza posizione si trova la Repubblica Ceca, la quale, a partire dal 2005, ha incentivato, attraverso una serie di leggi, la produzione di energia solare. Infatti, nel 2010 era, dopo Germania e Italia, il terzo produttore in Europa. Ad oggi, però, sembrerebbe che la crescita del fotovoltaico si sia del tutto arrestata, a causa di leggi retroattive che hanno cambiato le tariffe agevolate sulla produzione di questa energia. L'effetto di queste leggi è evidente nel grafico in Figura 2.7, dal quale emerge che la produzione pro capite di energia solare sembra essersi stabilizzata proprio a partire dal 2010, a differenza delle altre serie storiche le quali hanno subito una crescita più o meno lenta nel corso degli anni.



**Fig. 2.6.** Energia solare: primi dieci paesi innovatori

Il comportamento innovativo della Germania trova conferma nella serie storica riportata in Figura 2.7, essendo tra le prime ad esser entrata nel mercato del solare e

avendo da sempre prodotto, ad eccezione di qualche anno tra il 2005 e il 2010, una quantità di energia pro capite superiore alle altre nove nazioni rappresentate.

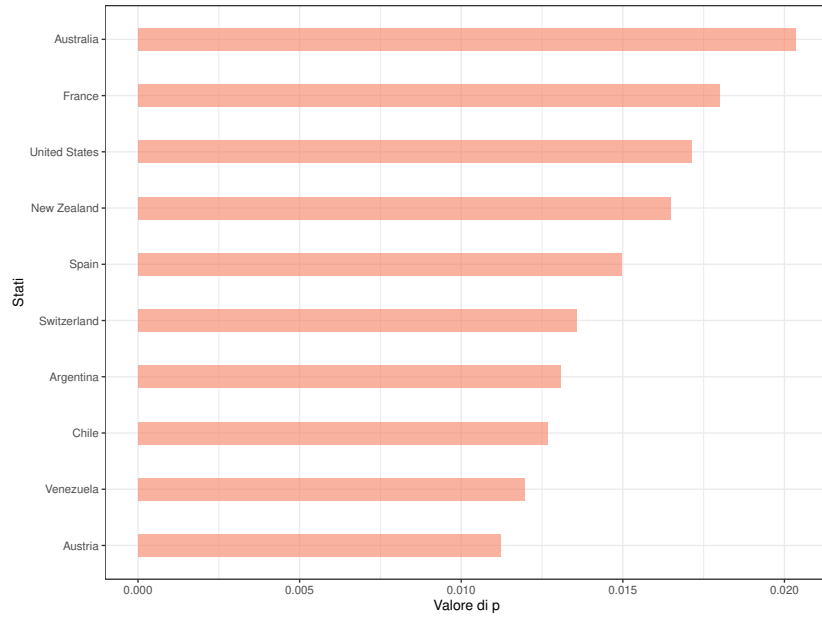


**Fig. 2.7.** Serie storica dei primi dieci paesi innovatori per l'energia solare

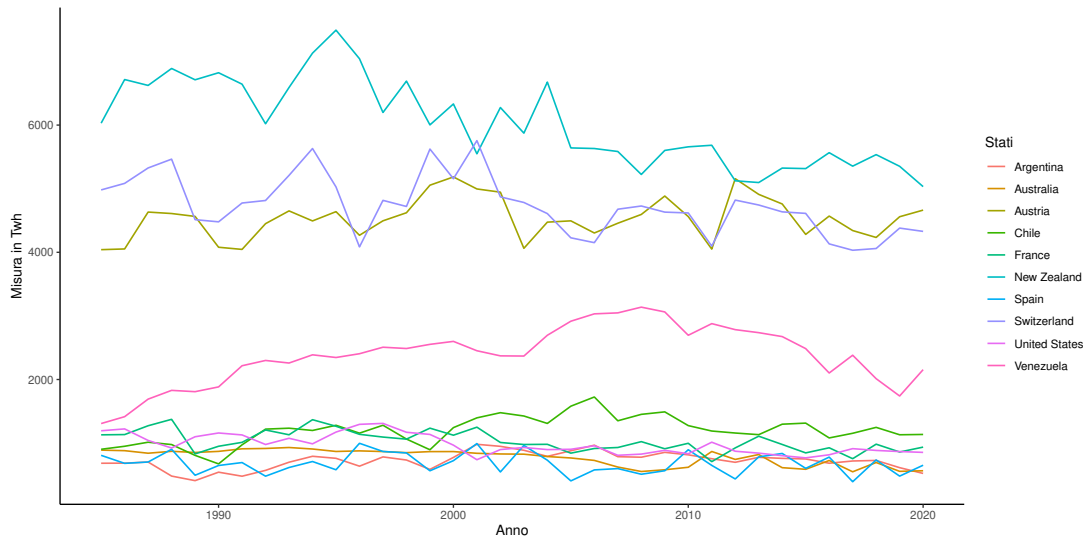
Infine, in Figura 2.8 sono rappresentati i valori assunti dal coefficiente  $p$  in ordine decrescente dei primi dieci paesi, derivante dall'applicazione del modello di Bass per l'energia idroelettrica. Tra i primi stati, troviamo la Francia dove, attualmente l'energia idroelettrica è la principale fonte rinnovabile. Lo stesso si può dire per l'Australia, per la quale l'idroelettrico è il settore che pesa di più tra le energie rinnovabili. Gli Stati Uniti, invece, sono considerati tra i maggiori produttori di questa energia, grazie soprattutto all'abbondante presenza di risorse idriche naturali, come fiumi, laghi e ghiacciai.

In Figura 2.9 sono rappresentate le serie storiche dei primi dieci paesi innovatori, le quali sembrano essere sostanzialmente stazionarie. Questo andamento conferma la fase di maturità del mercato idroelettrico per questi stati.





**Fig. 2.8.** Energia idroelettrica: primi dieci paesi innovatori



**Fig. 2.9.** Serie storica dei primi dieci paesi innovatori per l'energia idroelettrica

### 2.2.3 Analisi del coefficiente relativo agli imitatori

In questa sottosezione, si presentano i risultati, similmente a quanto riportato nella sottosezione precedente, relativi ora al caso dei paesi imitatori.

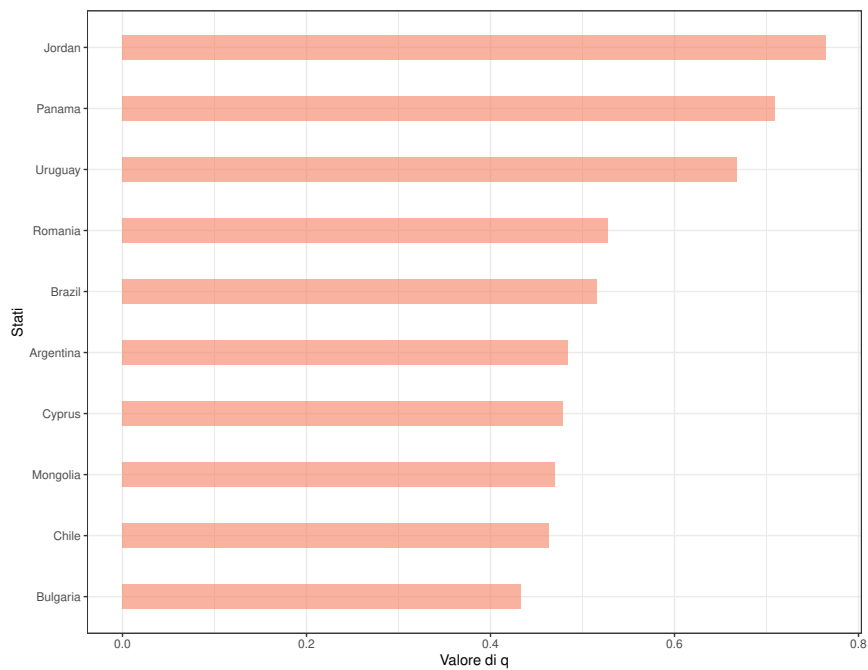
Nello specifico, in Figura 2.10 sono rappresentati i primi dieci paesi imitatori; considerando le prime quattro posizioni, tre di queste sono occupate da stati collocati in Sud America. In particolare, per lo stato di Panama, al secondo posto, a partire dal 2014, gli investimenti in energia eolica sono cresciuti notevolmente e questo sottolinea il grande interesse verso il completo utilizzo di energia derivante da fonti rinnovabili.

Al terzo posto, invece, l'Uruguay, pur essendo un piccolo paese dell'America Latina, è tra le prime cinque posizioni dei produttori di energia eolica in tutto il mondo, investendo proprio nella creazione di parchi eolici.

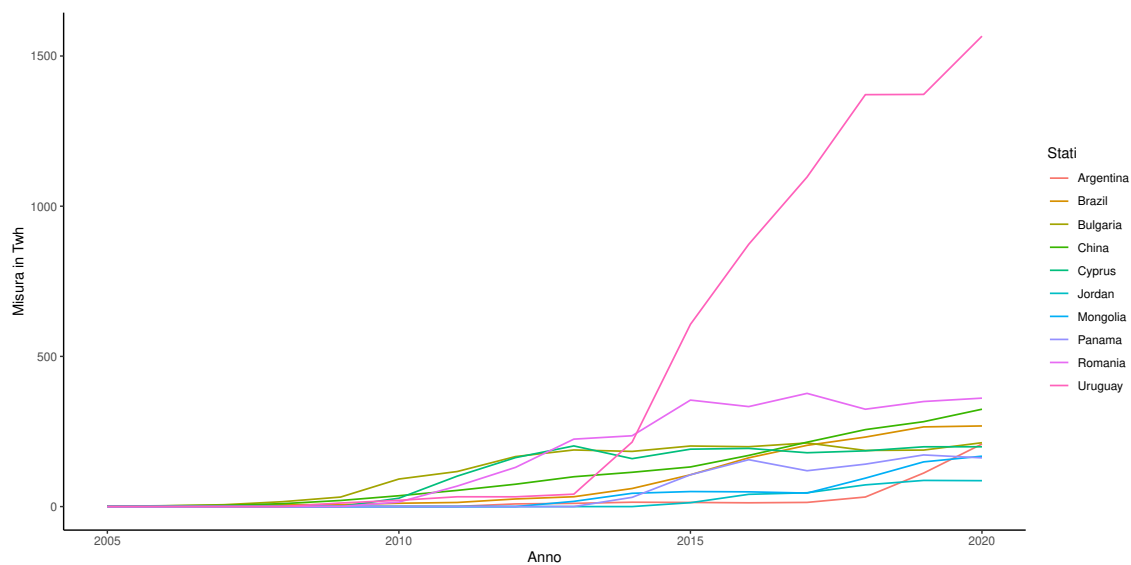
Infine, al primo posto troviamo la Giordania, dove la prima centrale elettrica eolica è stata costruita nel 2015. Questo progetto, per un paese povero di risorse naturali come è la Giordania, rappresenta un passo significativo verso l'indipendenza dalle fonti di energia convenzionali.

A differenza delle serie storiche rappresentate nei grafici precedenti per il caso degli stati innovatori, per i paesi imitatori si può notare come, in Figura 2.11, sia raffigurato il processo di diffusione tipico di questi comportamenti, il quale si trova in una fase di crescita, soprattutto per l'Uruguay, piuttosto che di maturità.

Passando ora al caso del solare e analizzando la Figura 2.12, si riscontra che il Vietnam è il paese a cui è associato il maggior valore del coefficiente  $q$ ; infatti, è uno stato dall'enorme potenziale che ha registrato una rapida espansione in termini di energia eolica e solare negli ultimi anni, oltre a essere uno dei mercati in maggior crescita per quanto riguarda la domanda di elettricità. La svolta verso il solare è iniziata a partire dal 2014 e, nel corso degli anni, è diventato il più grande mercato solare del sud-est



**Fig. 2.10.** Energia eolica: primi dieci paesi imitatori

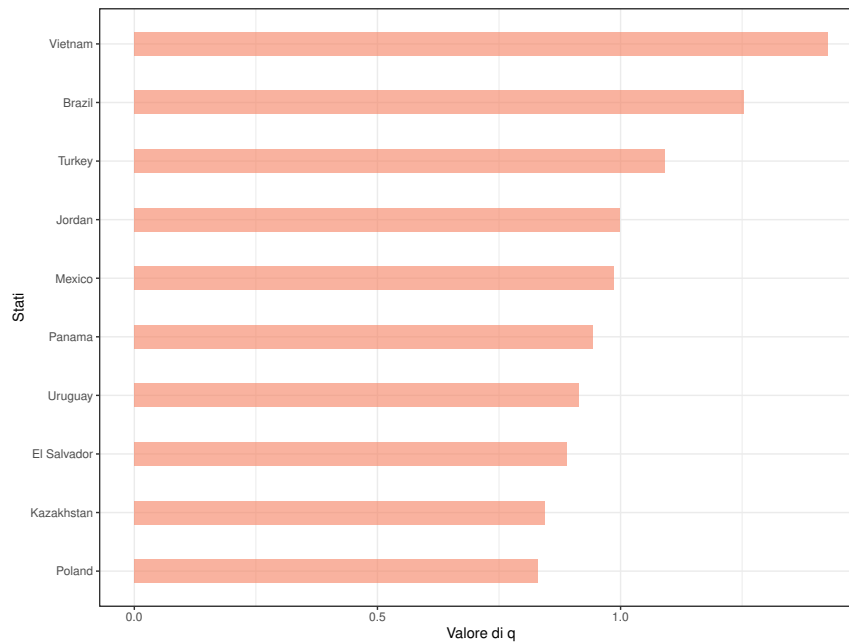


**Fig. 2.11.** Serie storica dei primi dieci paesi imitatori per l'energia eolica

asiatico.

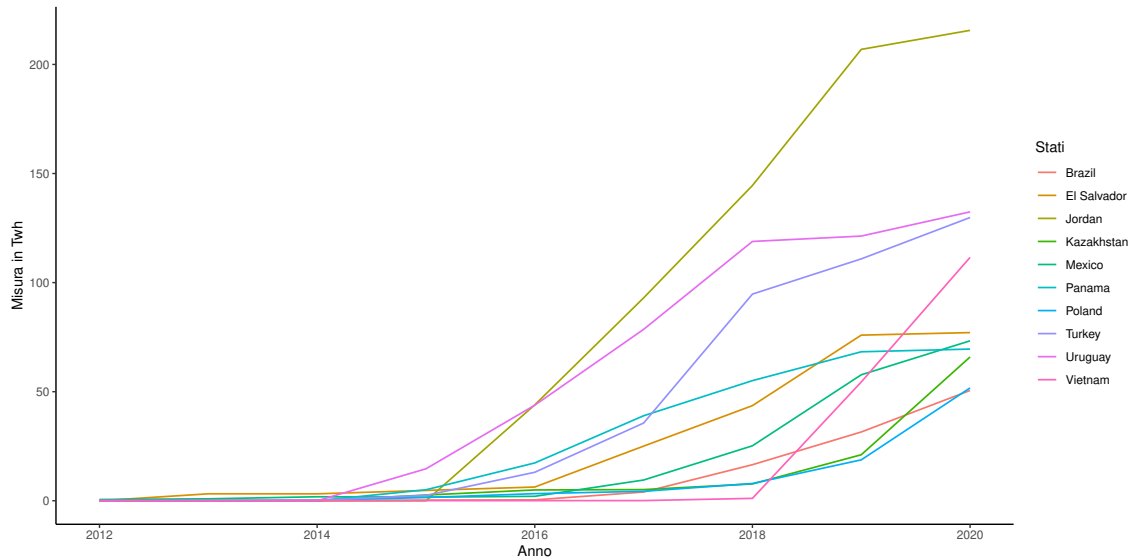
L'avvio del il mercato dell'energia solare del Brasile è avvenuto nel 2013; nei tre anni successivi la crescita della capacità installata è stata piuttosto lenta, ma a partire dal 2017 questa capacità è raddoppiata anno dopo anno.

Infine, la Turchia, al terzo posto, ha avuto nel corso degli ultimi dieci anni, un comportamento analogo al Brasile, penetrando nel mercato del solare nel 2013 e incrementando notevolmente la capacità solare istallata solo negli ultimi anni.



**Fig. 2.12.** Energia solare: primi dieci paesi imitatori

Le serie storiche riportate in Figura 2.13, sottolineano nuovamente le dinamiche del processo di diffusione dei paesi imitatori; a differenza degli innovatori, per i quali l'entrata nel mercato solare è avvenuta entro il 2010, per gli imitatori questo avvenimento si colloca dopo il 2010.

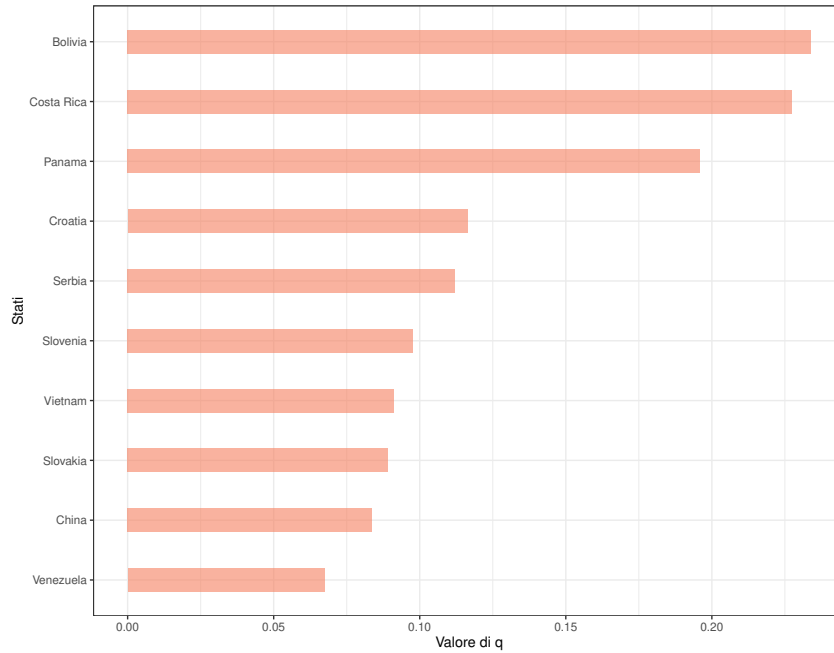


**Fig. 2.13.** Serie storica dei primi dieci paesi imitatori per l'energia solare

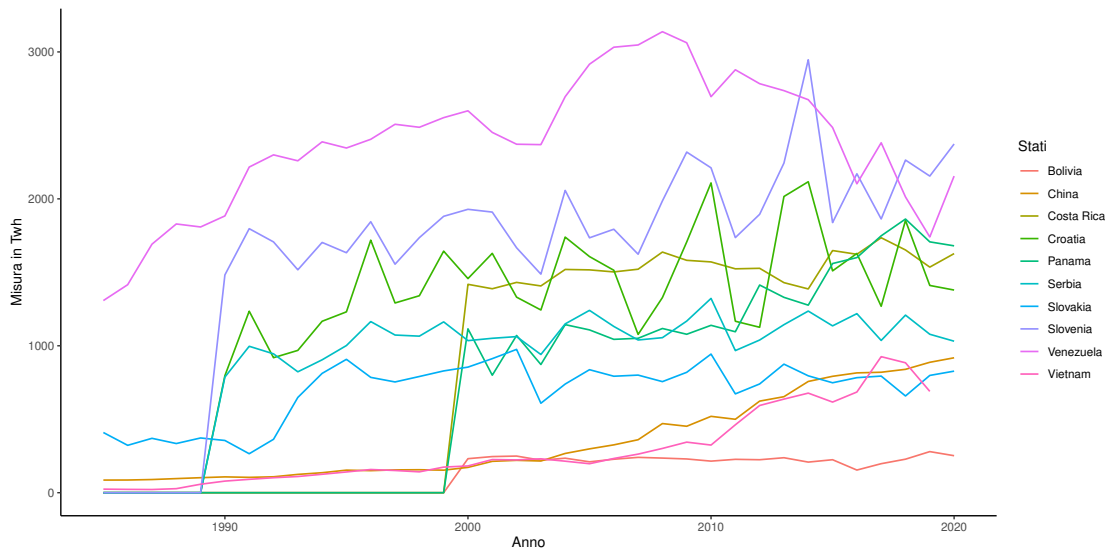
Considerando, infine, l'energia idroelettrica, dalla Figura 2.14, si evince che il paese considerato maggior imitatore di questa energia è la Bolivia; questa nazione sta cercando negli ultimi anni di rendere la produzione di energia il più sostenibile possibile, promuovendo proprio la generazione della stessa dall'idroelettrico.

Al secondo posto troviamo il Costa Rica, per il quale l'energia idroelettrica rappresenta la prima fonte di energia rinnovabile, in quanto fornisce fino al 73% del fabbisogno nazionale. Questa, viene prodotta principalmente grazie alla diga posta sul fiume Reventazon. Lo stesso si può dire di Panama, dove la percentuale si aggira intorno al 65%. Si tratta di una nazione dal grande potenziale idrico, il quale, negli anni, è stato sfruttato per soddisfare la crescente domanda di energia.

Come detto precedentemente per l'energia eolica, si riscontra, in Figura 2.15, l'andamento tipico del processo di diffusione dei paesi imitatori, per i quali l'entrata nel mercato idroelettrico è avvenuta più tardi e i volumi di produzione risultano inferiori.



**Fig. 2.14.** Energia idroelettrica: primi dieci paesi imitatori

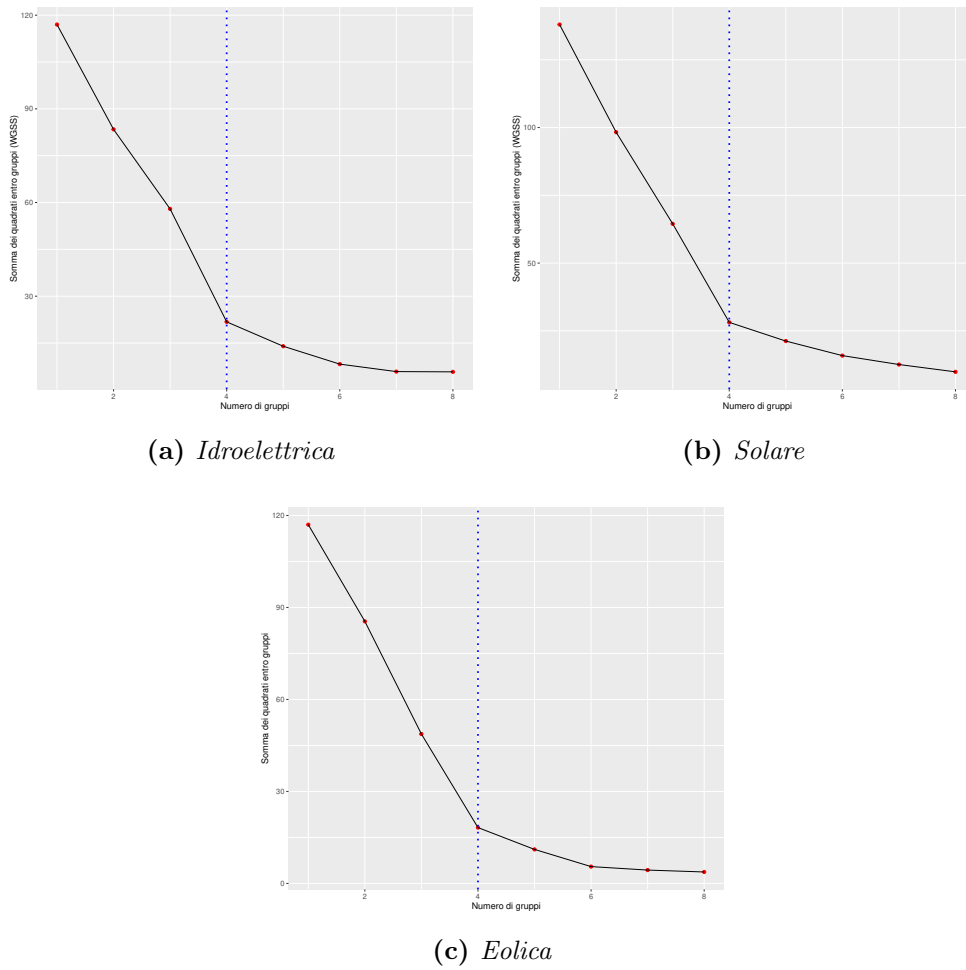


**Fig. 2.15.** Serie storica dei primi dieci paesi imitatori per l'energia idroelettrica

## 2.3 Analisi dei gruppi

Per cercare di valutare se tra gli stati considerati ce ne siano alcuni con caratteristiche comuni, si è deciso di condurre un'analisi di raggruppamento, considerando i valori standardizzati di  $m$ ,  $p$  e  $q$  derivanti dall'applicazione del modello di Bass alle serie storiche della produzione pro capite di energia rinnovabile. Lo scopo è condurre un'analisi che consenta di identificare gruppi di paesi tra loro simili per caratteristiche comuni relativamente al processo di diffusione. A tale scopo, il metodo di raggruppamento prescelto è basato sull'algoritmo delle  $k$ -medie (Azzalini e Scarpa, 2012).

Per determinare il numero di gruppi con cui suddividere il campione, si è soliti utilizzare il cosiddetto "metodo del gomito", da cui deriva il grafico riportato in Figura 2.16. Sull'asse delle  $x$  troviamo il numero di gruppi, mentre su quello delle  $y$  la somma dei quadrati entro in gruppi. Per il caso in esame, si noti come la scelta del numero dei gruppi sia stata immediata, dato il risultato ottenuto, rappresentato nel grafico in Figura 2.16. Infatti, il gomito è evidente in corrispondenza di un numero di gruppi pari a quattro. Si è scelta pertanto questa suddivisione per tutte e tre le energie rinnovabili analizzate.

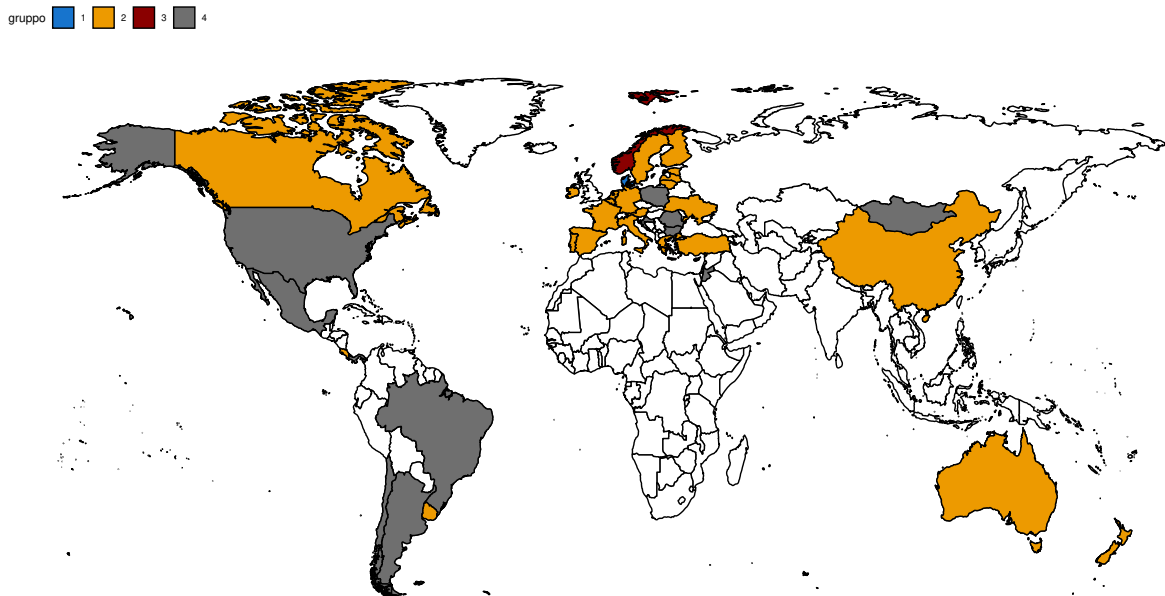


**Fig. 2.16.** Numero di gruppi ottimale per le tre energie rinnovabili



Di seguito, in Figura 2.17 è riportata la mappa e i gruppi per il caso dell'energia eolica e in Tabella 2.4 la formazione degli stessi:

- il gruppo blu è rappresentato da un singolo stato, la Danimarca, considerato il primo paese innovatore, per il quale il coefficiente  $p$  è molto più elevato rispetto agli altri stati;
- il gruppo arancione è formato dai paesi innovatori;
- il gruppo rosso, con solo la Norvegia, rappresenta il paese con il più grande mercato potenziale; anche in questo caso, analizzando più nel dettaglio i risultati ottenuti dall'applicazione del modello di Bass, ci si rende conto che l'intervallo di confidenza di  $m$  comprende lo zero;
- il gruppo grigio rappresenta i paesi imitatori.



**Fig. 2.17.** Gruppi individuati dall'algoritmo k-medie per l'energia eolica

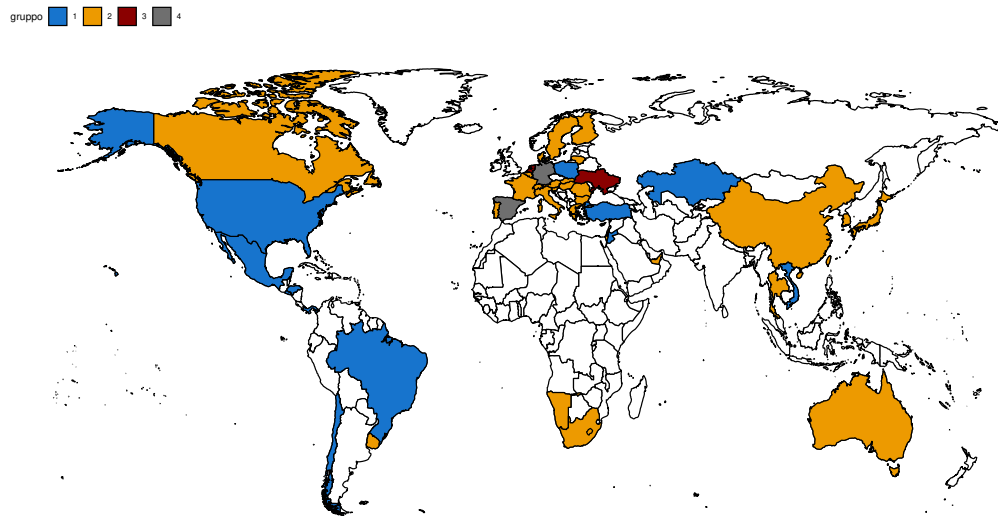
Gruppo	Colore identificativo	paese
1	Blu	Danimarca
2	Arancione	Australia, Austria, Belgio, Canada, Cina, Costa Rica, Estonia, Finlandia, Francia, Germania, Grecia, Irlanda, Italia, Lettonia, Lussemburgo, paesi Bassi, Portogallo, Regno Unito, Spagna, Svezia, Turchia, Ucraina, Usa
3	Rosso	Norvegia
4	Grigio	Argentina, Brasile, Bulgaria, Cile, Cipro, Croazia, Giordania , Messico, Mongolia, Panama, Polonia, Romania, Ungheria

**Tab. 2.4.** Gruppi derivanti dall'algoritmo delle k-medie per l'energia eolica

In Figura 2.18, si riporta il risultato dell'applicazione dell'algoritmo per il caso dell'energia solare. I gruppi individuati si possono così descrivere:

- il gruppo blu è costituito dai primi dieci stati considerati imitatori;
- il gruppo arancione è un gruppo misto: alcuni dei paesi appartenenti a questo gruppo sono caratterizzati da valori di  $p$  elevati, altri invece da un alto valore di  $m$ ;
- il gruppo rosso considera Ucraina e paesi Bassi, due paesi il cui coefficiente relativo al mercato potenziale risulta posizionarsi tra le prime tre posizioni; per entrambi questi stati, l'intervallo di confidenza di  $m$  comprende lo zero e ciò suggerisce che l'utilizzo di tale coefficiente per descrivere il processo di diffusione potrebbe essere non affidabile;

- il gruppo grigio è costituito dai paesi considerati innovatori per eccellenza, ovvero Spagna e Germania.



**Fig. 2.18.** Gruppi individuati dall'algoritmo k-medie per l'energia solare

La composizione dei gruppi evidenziati in Figura 2.18 è descritta più dettagliatamente in Tabella 2.5.

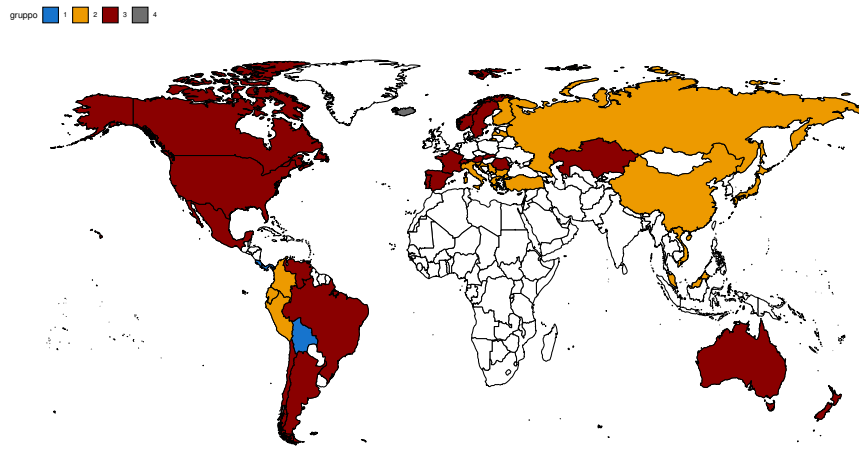
Infine, in Figura 2.19 è riportata una mappa nella quale sono evidenziati i quattro gruppi derivanti dall'applicazione del metodo di raggruppamento per il caso dell'energia idroelettrica, che possiamo così descrivere:

- il gruppo blu rappresenta i primi tre stati con il più alto coefficiente di imitazione, per i quali il valore di  $q$  è ben più alto rispetto a quello degli altri paesi considerati; inoltre, il coefficiente  $m$  è minore rispetto a quello dei paesi innovatori;
- il gruppo arancione è composto dai paesi imitatori, ad esclusione dei primi tre appartenenti al gruppo blu;
- il gruppo rosso è costituito interamente dai paesi innovatori;

Gruppo	Colore identificativo	paese
1	Blu	Brasile, Cile, El Salvador, Giordania, Honduras, Kazakistan, Messico, Panama, Polonia, Turchia, Uruguay, Vietnam
2	Arancione	Australia, Austria, Belgio, Bulgaria, Canada, Cina, Cipro, Corea del Sud, Danimarca, Emirati Arabi Uniti, Finlandia, Francia, Giappone, Grecia, Israele, Italia, Lituania, Namibia, Portogallo, Regno Unito, Repubblica Ceca, Romania, Slovacchia, Slovenia, Sud Africa, Svezia, Svizzera, Thailandia, Ungheria, Usa
3	Rosso	paesi Bassi, Ucraina
4	Grigio	Germania, Spagna

**Tab. 2.5.** Gruppi individuati dell’algoritmo delle k-medie per l’energia solare

- il gruppo grigio è composto solo dall'Islanda, paese a cui è associato il più alto valore del coefficiente  $m$ ; si ricordi, in questo caso, che l'intervallo di confidenza per questo coefficiente comprende lo zero, ad indicare la non utilità del coefficiente ai fini dell'analisi.



**Fig. 2.19.** Gruppi individuati dall'algoritmo delle k-medie per l'energia idroelettrica

In Tabella 2.6 vengono riportati gli stati suddivisi per gruppo di appartenenza, identificati dai colori blu, arancione, rosso e grigio.

Gruppo	Colore identificativo	paese
1	Blu	Bolivia, Costa Rica, Panama
2	Arancione	Bulgaria, Cina, Colombia, Croazia, Ecuador, Finlandia, Giappone, Grecia, Italia, Lettonia, Malesia, Perù, Russia, Serbia, Slovenia, Turchia, Vietnam
3	Rosso	Argentina, Australia, Austria, Brasile, Canada, Cile, Francia, Messico, Norvegia, Nuova Zelanda, Portogallo, Romania, Slovacchia, Spagna, Svezia, Svizzera, Usa, Venezuela
4	Grigio	Islanda

**Tab. 2.6.** Gruppi derivanti dall’algoritmo delle k-medie per l’energia idroelettrica

## 2.4 Modellazione per spiegare

Le analisi condotte in questo Capitolo hanno permesso di identificare i paesi innovatori e imitatori, nel percorso di adozione delle energie rinnovabili, attraverso la valenza esplicativa del modello di Bass, che per mezzo dell’interpretazione di tre parametri, consente di comprendere la dinamica sottostante a un processo di diffusione.

L’aspetto che qualifica quest’analisi è il fatto di aver applicato un solo modello ad un ampio numero di paesi, trovando che questo riusciva a catturare in maniera perlopiù efficace la crescita nelle adozioni delle energie rinnovabili. A partire da questa analisi, grazie alla classificazione del comportamento degli stati imitatori e innovatori, si è potuto mettere in evidenza similitudini e differenze tra questi. I paesi innovatori pur essendo entrati nel mercato delle fonti rinnovabili ben prima rispetto ai paesi imitatori,

si trovano ancora in una fase di crescita del processo di diffusione, comportamento che ci si aspetta, e che viene confermato con l'applicazione del modello, dai paesi imitatori. Questo andamento crescente naturalmente segnala un trend sicuramente positivo verso la transazione energetica, in quanto rappresenta un impegno mondiale verso una più totale indipendenza dalle fonti non rinnovabili. Per quanto riguarda, invece, le possibili differenze tra gli stati innovatori e imitatori si sottolineano principalmente i differenti volumi di quantità di energia prodotta, minori per i paesi imitatori.

La scelta di operare un'analisi di raggruppamento non sulle singole serie storiche, bensì sui parametri stimati dal modello, è stata motivata dall'interesse a capire se ci fosse una regolarità nei comportamenti dei vari stati. I risultati ottenuti hanno permesso di identificare quattro gruppi per ciascuna delle energie rinnovabili; tra questi gruppi, due sono rappresentati interamente da paesi innovatori o imitatori, mentre gli altri due gruppi sono in grado di discriminare quegli stati considerati innovatori o imitatori per eccellenza; esempi paradigmatici, considerando gli innovatori, sono la Danimarca per l'eolico oppure la Spagna e la Germania per il solare, mentre tra gli imitatori troviamo la Bolivia, Costa Rica e Panama per il caso dell'idroelettrico.

# Capitolo 3

## Modelli per dati funzionali

In questo capitolo, si presenta la modellazione scelta per perseguire il secondo obiettivo che ci si è posti, ovvero prevedere la produzione pro capite di energia rinnovabile.

Per il caso in esame, si considera come unità statistica la singola serie storica di produzione pro capite di energia e uno degli approcci più adatti per trattare questo tipo di dato è l'analisi dei dati funzionali (*Functional Data Analysis - FDA*). La variabile casuale appartiene a uno spazio funzionale, e non a uno spazio scalare o vettoriale, di cui si osserva una porzione finita, definita da istanti di tempo  $t_1, \dots, t_T$ . Ogni curva, quindi ogni osservazione, è generata da un processo del tipo:

$$x_i(t) = f_i(t) + \varepsilon_i(t) \quad i = 1, \dots, n \quad t = 1, \dots, T \quad (3.1)$$

dove  $x_i(t)$  rappresenta l'osservazione soggetta a un errore  $\varepsilon_i(t)$  della funzione casuale  $f_i(t)$ , mentre  $n$  è il numero di curve considerate e  $t$  gli istanti temporali. Per quanto riguarda l'analisi condotta in questa tesi, il numero di unità statistiche varia a seconda dell'energia rinnovabile; in particolare si ha  $n_e = 40$ ,  $n_s = 46$  e  $n_i = 40$ , dove  $e$ ,  $s$  e  $i$  indicano rispettivamente le energie eolica, solare e idroelettrica.



Inoltre, gli istanti temporali per i quali si osserva l'unità statistica sono, in questo caso, equispaziati: le serie storiche sono annuali, perciò gli istanti vengono misurati a distanza di un anno e ogni curva è osservata nello stesso intervallo di tempo  $\mathcal{T}$ . Anche per quanto concerne gli intervalli temporali, si ricorda che questi sono differenti a seconda dell'energia rinnovabile analizzata; in particolare per l'energia eolica e idroelettrica le misurazioni vanno dal 1985 al 2020, mentre per l'energia solare dal 2000 al 2020.

### 3.1 Lisciamento dei dati

Considerando che in ciascuna osservazione è presente una componente d'errore che si intende rimuovere, è possibile rendere i dati più regolari attraverso diversi approcci; uno dei più utilizzati nella FDA è l'espansione in funzioni di base. Attraverso questo metodo si riesce ad approssimare una generica funzione  $f(t)$  tramite la seguente combinazione lineare:

$$f(t) = \sum_{k=1}^K c_k \phi_k(t) = \phi(t)^\top c \quad (3.2)$$

A seconda dei dati che si intendono analizzare, è possibile utilizzare diversi approcci, che si distinguono in base alla natura del dato stesso: dato periodico e dato aperiodico. Nel primo caso, si utilizza la serie di Fourier, mentre nel secondo caso, analizzato nel presente elaborato, si utilizzano le splines. Le splines sono funzioni polinomiali a tratti, definite dall'intervallo di validità, dal numero di nodi e dal grado del polinomio stesso. Il numero delle  $K$  funzioni di base è dato dalla somma tra il numero di nodi e l'ordine. Generalmente si utilizzano le B-splines cubiche e, una volta fissate, è possibile procedere in due modi differenti, utilizzando le splines di regressione o le splines di lisciamento. Nel primo caso, la stima della matrice di coefficienti  $\mathbf{C} = [c_1, \dots, c_n]$  di dimensione  $K \times n$  diventa un problema di regressione lineare.

Sia  $\mathbf{x} = [x_1, \dots, x_n]$  la matrice di dimensione  $T \times n$  delle osservazioni e  $\Phi = [\phi_1, \dots, \phi_K]$  la matrice  $T \times K$  delle funzioni di base valutate nei tempi osservati  $t_1, \dots, t_T$ ; allora l'Equazione 3.2 può essere riscritta in forma matriciale come segue

$$\mathbf{x} = \Phi \mathbf{C} + \varepsilon \quad (3.3)$$

Utilizzando il criterio dei minimi quadrati, si ottiene, per ciascuna osservazione  $x_i$ , la stima dei vettori dei coefficienti  $\hat{c}_i$ , ovvero

$$\hat{c}_i = (\Phi^\top \Phi)^{-1} \Phi^\top x_i \quad (3.4)$$

Nel secondo caso, ovvero utilizzando le splines di lisciamiento, è possibile lisciare i dati introducendo una misura di penalità. Generalmente, si considera come penalità l'integrale della derivata seconda di  $f$  valutata al tempo  $t$ :

$$PEN(x) = \int \{D^2 x(t)\}^2 dt \quad (3.5)$$

In questo caso, per stimare il vettore dei coefficienti  $\hat{c}_i$  si usa il criterio dei minimi quadrati penalizzati, ovvero

$$SSE = \sum_{i=1}^n \|x_i - \Phi c_i\|^2 + \lambda \int \{D^2 x(t)\}^2 dt \quad (3.6)$$

dove  $\lambda$  è un parametro di regolarizzazione che funge da parametro di lisciamiento, il quale può essere scelto via convalida incrociata.

Dall'Equazione 3.6 si ricava la stima dei coefficienti  $c_i$ , ossia

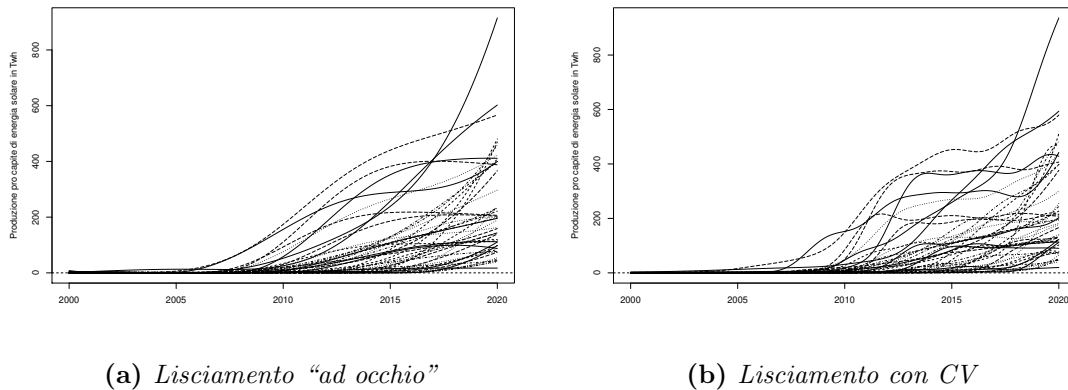
$$\hat{c}_i = (\Phi^\top \Phi + \lambda R_\phi)^{-1} \Phi^\top x_i \quad (3.7)$$

dove la matrice di penalizzazione  $R_\phi$  è tale che  $R_\phi = \int D^2 \Phi(t) D^2 \Phi^\top(t) dt$ .

### 3.1.1 Applicazione del lisciamiento ai dati

Considerando che il processo di diffusione dei Paesi imitatori e innovatori inizia in anni differenti, nel momento in cui si lisciano le curve, queste potrebbero assumere valori minori di zero in alcuni punti dell'intervallo considerato. Per evitare questa possibile situazione, si sono imposti a zero tutti quei coefficienti  $c_k$  minori di zero.

Per la scelta del parametro di lisciamiento, si è inizialmente utilizzata la convalida incrociata, ma non fornito delle curve sufficientemente lisce, si è scelto  $\lambda$  manualmente, provando una serie di valori differenti e selezionando quello che più lisciasse le curve. Si riporta in Figura 3.1 un esempio che evidenzia la differenza tra le curve lisciate con il parametro di lisciamiento scelto manualmente e con la convalida incrociata.



**Fig. 3.1.** Differenza tra dati liscati con metodo “ad occhio” e con CV

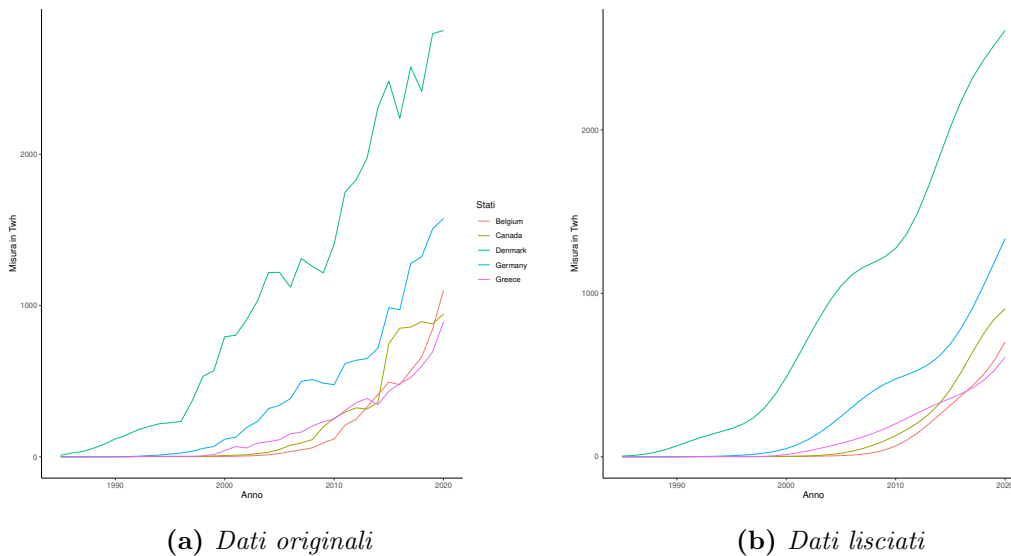
In Tabella 3.1 vengono riportati i valori di  $\lambda$  e il numero di basi utilizzate per le tre energie rinnovabili analizzate. Si ricorda che il numero di funzioni di base è pari alla somma tra il numero di nodi e l'ordine della spline, pari a 3.

Si riporta in Figura 3.2 un esempio di cinque curve, a destra originali e a sinistra lisciate. Si nota come le curve lisciate riescano ad eliminare il rumore che caratterizza

Energia	$\lambda$	Numero di basi	Intervallo temporale
Eolica	1.75	40	1990-2020
Solare	2.5	26	2000-2020
Idroelettrica	1	40	1990-2020

**Tab. 3.1.** Numero di basi e valore di  $\lambda$  per il lisciamiento dei dati dell'energia

le serie dei dati originali, rappresentando perfettamente il trend di questi ultimi. Per cercare di capire quale sia l'andamento medio della produzione pro capite di energia e la varianza della stessa, si presentano, in questa sottosezione, queste due statistiche di sintesi per dati funzionali.



**Fig. 3.2.** Differenza tra dati originali e dati lisciati per cinque Paesi

La media funzionale è definita come la media valutata ad ogni tempo osservato delle

$n$  funzioni:

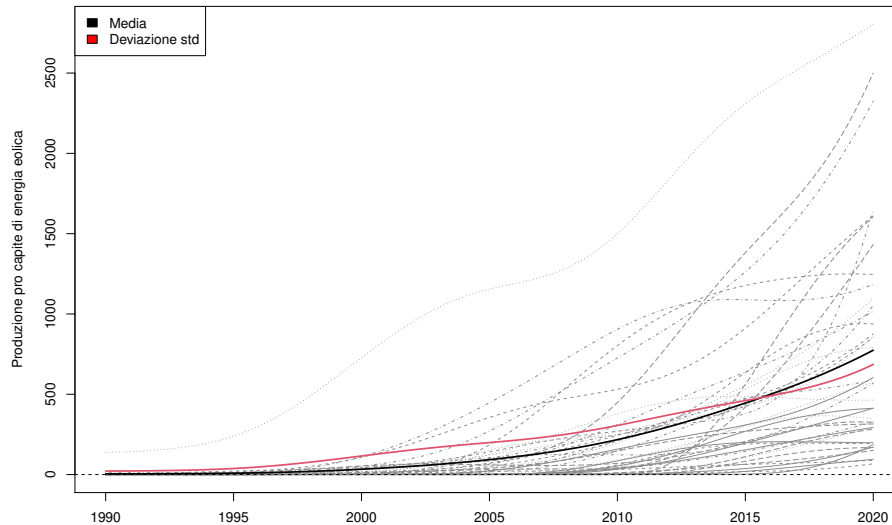
$$\bar{x}(t) = \frac{1}{n} \sum_{i=1}^n x_i(t) \quad (3.8)$$

Analogamente, la varianza è definita come

$$\text{Var}_X(t) = \frac{1}{n-1} \sum_{i=1}^n [x_i(t) - \bar{x}(t)]^2 \quad (3.9)$$

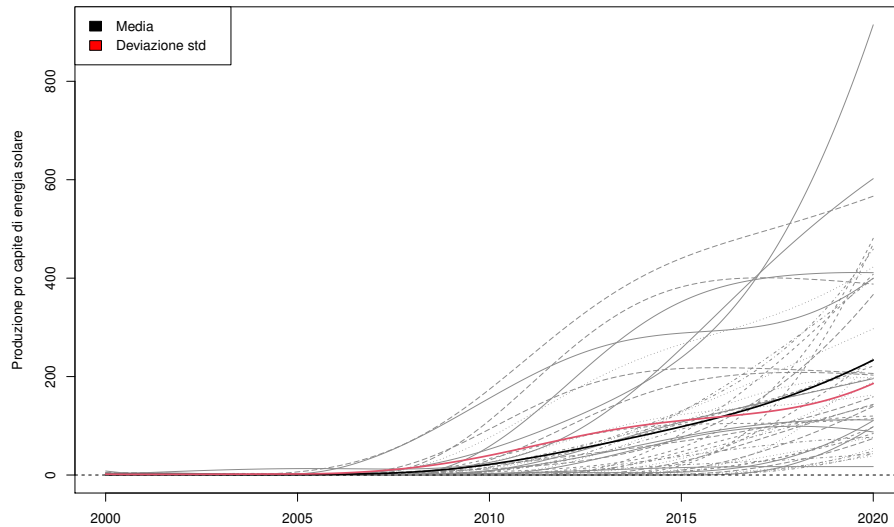
Nei grafici di seguito si rappresentano media e deviazione standard congiuntamente per ogni energia rinnovabile.

In particolare, in Figura 3.3 sono raffigurate in grigio le curve lisce dei Paesi analizzati, mentre, in nero, la produzione media di energia eolica pro capite, la quale è caratterizzata da una crescita più o meno costante nell'intervallo considerato; lo stesso si può dire della deviazione standard, in rosso, la quale aumenta via via ci si avvicina al 2020.



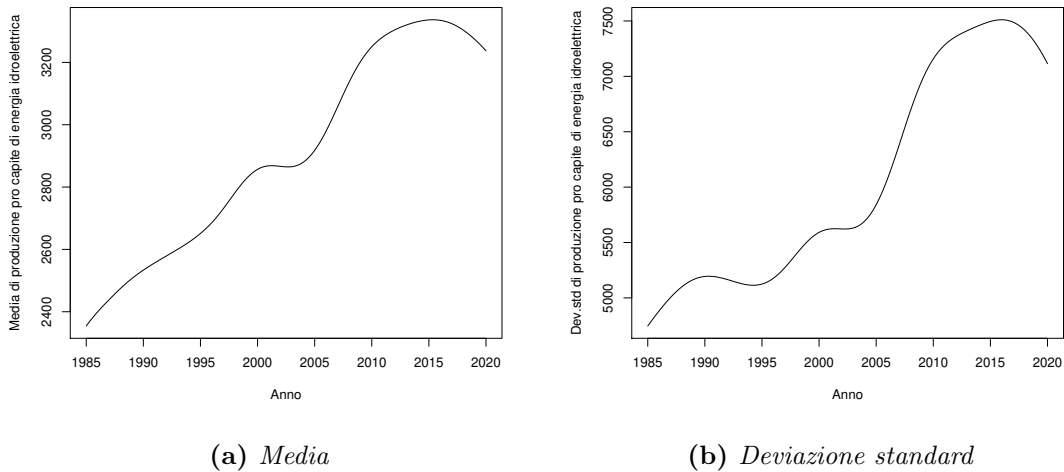
**Fig. 3.3.** Media e deviazione standard per l'energia eolica

La situazione per quanto riguarda l'energia solare, in Figura 3.4, è la medesima rispetto al caso dell'energia eolica, con una crescita più o meno costante sia di produzione media che di deviazione standard lungo tutto l'intervallo considerato.



**Fig. 3.4.** Media e deviazione standard per l'energia solare

Infine, per l'energia idroelettrica si è deciso di rappresentare media e deviazione standard in due grafici separati rispetto alle curve lisce, per aver una maggior chiarezza dei risultati. Infatti, diversamente dagli altri due casi di energia, in Figura 3.5 si nota come la produzione media di energia idroelettrica, a sinistra, abbia subito una crescita quasi costante fino al 2015 circa e poi una riduzione negli ultimi cinque anni; lo stesso si può dire della deviazione standard, raffigurata a destra, la quale, a differenza degli altri due casi di energia, è circa il doppio della media.



**Fig. 3.5.** Media e deviazione standard per l'energia idroelettrica

## 3.2 Componenti principali funzionali

L'analisi delle componenti principali (PCA) è uno dei metodi più utilizzati per le principali fonti di variabilità nei dati. Anche nel contesto dei dati funzionali, la f-PCA (*functional principal component analysis*) è un metodo molto diffuso. Se nel contesto multivariato, attraverso la decomposizione in valori singolari si ottengono gli autovalori a cui sono associati gli autovettori, in ambito funzionale ad ogni autovalore è associata un'autofunzione. Le componenti principali sono rappresentate dagli autovettori, nel caso multivariato, o dalle autofunzioni, nel caso funzionale, mentre gli autovalori sono indicatori dell'importanza delle stesse componenti.

Diversamente da quanto accade nel contesto multivariato, nell'ambito dell'analisi dei dati funzionali vengono esaminate le funzioni  $x_i(s)$ , dove  $s$  rappresenta l'indice continuo e non discreto.

Il modo più appropriato per combinare le funzioni di pesi  $\beta(s)$  con le funzioni  $x(s)$

è attraverso il prodotto interno

$$\int \beta x = \int \beta(s)x(s)ds. \quad (3.10)$$

Si vuole sottolineare che nel caso multivariato l'integrale è sostituito dalla sommatoria. Utilizzando la formulazione in Equazione 3.10, i punteggi della componente funzionale con peso  $\beta$  sono definiti come

$$f_i = \int \beta x_i = \int \beta(s)x_i(s)ds. \quad (3.11)$$

La determinazione delle componenti principali avviene in due passi:

1. Al primo passo, il vettore di pesi  $\xi_1(s)$  è tale da massimizzare

$$\frac{1}{N} \sum_i f_{i1}^2 = \frac{1}{N} \sum_i \left( \int \xi_1 x_i \right)^2 \quad s.v. \quad \int \xi_1(s)^2 ds = \|\xi_1\|^2 = 1 \quad (3.12)$$

si impone il vincolo di norma unitaria affinché il problema sia ben definito e garantisca una soluzione;

2. All' $m$ -esimo passo, si calcola un nuovo vettore di pesi  $\xi_i$  il quale massimizza la media dei quadrati dei punteggi sotto due vincoli: uno di norma unitaria,  $\|\xi_1\|^2 = 1$ , e l'altro di ortogonalità,  $\int \xi_k(s)\xi_m(s)ds = 0$ ,  $k < m$ .

Con il primo passo si individua la più grande fonte di variabilità e con i passi successivi si trovano nuovi contributi alla variabilità totale, i quali, grazie ai vincoli di ortogonalità, spiegano aspetti diversi della variabilità stessa.

Le autofunzioni rappresentano un sistema di basi ortonormale che può essere utilizzato per rappresentare i dati di partenza, in modo tale da poter scrivere ogni curva come combinazione lineare di queste basi, ovvero

$$\hat{x}_i(t) = \sum_{k=1}^K f_{ik} \xi_k(t). \quad (3.13)$$



Si può affermare che le autofunzioni ottenute tramite f-PCA rappresentano la base più efficiente, nel senso che l'errore totale della somma dei quadrati

$$PCASSE = \sum_{i=1}^N \int [x_i(t) - \hat{x}_i(t)]^2 dt \quad (3.14)$$

è il minimo tra tutti gli errori possibili.

Applicare la PCA nel caso multivariato, equivale a trovare autovettori e autovalori della matrice di correlazione o di varianza-covarianza. Sia  $X$  la matrice  $N \times p$  con valori  $x_{ij}$ . La prima componente principale si ottiene come:

$$\max_{\xi_1^\top \xi_1 = 1} \frac{1}{N} \xi_1^\top X^\top X \xi_1. \quad (3.15)$$

Utilizzando la matrice di varianza  $V = \frac{1}{N} X^\top X$  si ottiene

$$\max_{\xi_1^\top \xi_1 = 1} \xi_1^\top V \xi_1. \quad (3.16)$$

Sia  $V = UDU^\top$ , dove  $D$  è la matrice diagonale con gli autovalori e  $U$  è la matrice ortogonale di autovettori. Sia  $\theta = U^\top \xi$  e, quindi,  $\xi = U\theta$ . Il vincolo  $\|\xi_1\|^2 = 1$  diventa  $\|U\theta\|^2 = \theta U^\top U \theta = \|\theta\|^2 = 1$ . Il problema diventa allora

$$\max_{\theta^\top \theta = 1} \theta^\top D \theta \quad (3.17)$$

che è equivalente a prendere il più grande autovalore di  $D$  e il relativo autovettore. È possibile applicare la decomposizione di  $V$  in  $UDU^\top$  solo se la matrice ha rango pieno, ovvero  $p > n$ , e ci sono al massimo  $\min\{p, n - 1\}$  autovalori non nulli. Nel contesto dei dati funzionali,  $p$  non è una quantità fissa, ma è infinito-dimensionale. Infatti, non ci sono infinite componenti come sono infiniti i punti su cui posso valutare una funzione, ma ce ne sono al più  $\min\{p, N\}$ , ma visto che  $p = \infty$  allora ce ne sono al massimo  $N - 1$ . Analizzando allora il caso funzionale, le funzioni dei pesi delle componenti principali  $\xi_i(s)$  soddisfa

$$\int v(s, t) \xi(t) dt = \rho \xi(s) \quad (3.18)$$

per un appropriato autovalore  $\rho$ . La parte a sinistra dell'Equazione 3.18 è una trasformata integrale di  $V$ , detta operatore di covarianza, dalla funzione di peso  $\xi$  definita come

$$V\xi = \int v(\cdot, t)\xi(t)dt. \quad (3.19)$$

L'autoequazione è definita come

$$V\xi = \rho\xi \quad (3.20)$$

la quale, per essere risolta, sfrutta la rappresentazione delle funzioni attraverso le basi, in modo da ottenere dei coefficienti. Come per il caso multivariato, anche per i dati funzionali si calcolano le autofunzioni  $\xi_j$  di  $v(s, t)$  come soluzione dell'autoequazione funzionale. Si noti che nel caso multivariato il numero di variabili  $p$  è solitamente minore di  $N$ , ma ciò non è sempre vero nel contesto funzionale: il numero di funzioni osservate è più grande di  $N$ . Questo implica che il massimo numero di autovalori non nulli è  $\min\{N - 1, K, n\}$ , dove  $N$  rappresenta la numerosità campionaria,  $K$  il numero di basi e  $n$  il numero di punti in cui è stata valutata la curva. Tipicamente il minimo è  $K$  se si ha una divisione in basi parsimoniosa. Le componenti funzionali risolvono il seguente autosistema

$$\int v(s, t)\xi(t)dt = \rho\xi(s) \quad (3.21)$$

in due possibili modi:

1. Il primo metodo considera una griglia finita  $t = [t_1, \dots, t_K]$  e trova poi il valore singolare della matrice di varianze e covarianze  $\Sigma(t, t)$ , valutata sulla nuova griglia. Infine, gli autovalori si ottengono considerando un metodo di interpolazione.
2. Il secondo metodo considera l'espansione in funzioni di base utilizzata per rappresentare le osservazioni  $x_i$  anche per le autofunzioni e in questo modo è possibile scrivere l'Equazione 3.18 in termini dei coefficienti dell'espansione in funzioni di base.

### 3.2.1 Applicazione della f-PCA ai dati

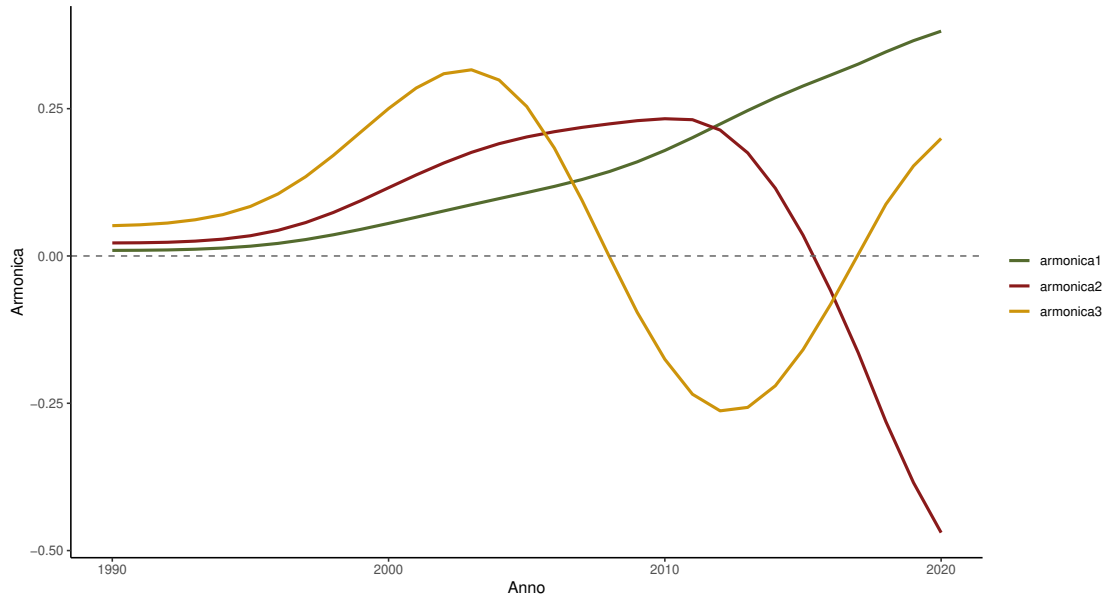
Applicando il metodo delle f-PCA ai dati sulla produzione di energia pro capite, si è deciso di considerare tre componenti principali per le energie eolica e solare e due per l'energia idroelettrica. Nelle Figure 3.7, 3.10 e 3.13 sono rappresentate le armoniche individuate per ogni tipo di energia; la linea continua rappresenta l'andamento medio delle osservazioni, mentre i segni + e - corrispondono agli effetti conseguenti, rispettivamente, all'aggiunta o alla sottrazione di un opportuno multiplo di ciascuna armonica. In ogni grafico è indicata anche la percentuale di varianza spiegata da ogni armonica. In particolare, si nota che per i casi dell'energia eolica e solare, l'andamento delle armoniche e l'interpretazione che se ne ricava è pressoché uguale.

Per l'energia eolica la prima componente principale (PC) spiega il 91% della variabilità totale; è positiva e crescente su tutto l'intervallo considerato e i Paesi a cui corrispondono punteggi positivi sono quelli con una produzione pro capite di energia eolica maggiore, come si può vedere in Figura 3.8.

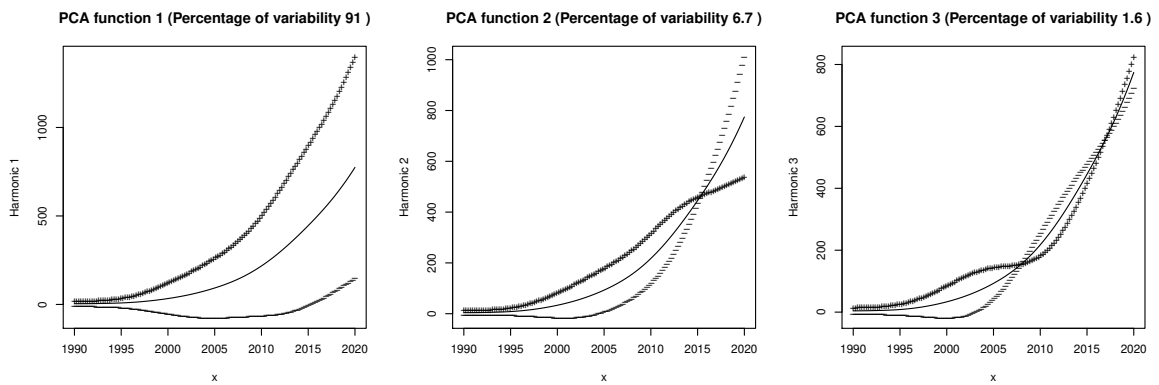
La seconda PC, invece, spiega il 6.7% di variabilità ed è positiva fino al 2015 e negativa nell'ultima parte dell'intervallo considerato; in questo caso, i punteggi positivi sono associati a quegli stati che hanno registrato una crescita esponenziale di produzione di energia a ridosso del 2015, ma pur sempre al di sotto della media.

Infine, la terza PC, spiega una piccola percentuale di variabilità, ovvero l'1,6%, e corrisponde ad un effetto di accelerazione nella partenza della crescita, e un ritardo nella fase intermedia. Questo lo si nota dal grafico a destra in Figura 3.7, dove si vede che i Paesi che hanno valori positivi per questa componente hanno una crescita accelerata all'inizio, per poi rallentare, anche rispetto agli altri stati, nella fase centrale (2008-2018).

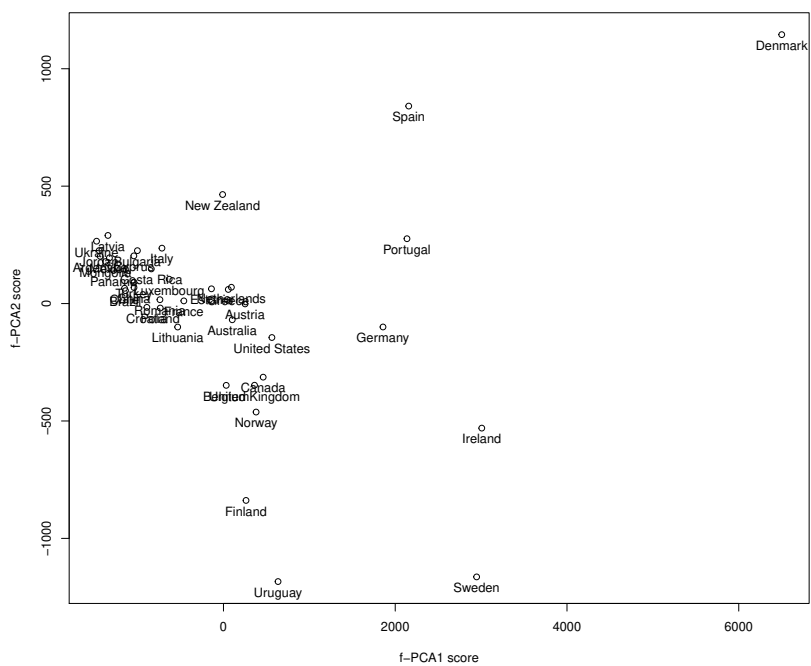
Come anticipato, la dinamica nel caso dell'energia solare è pressoché analoga a quella



**Fig. 3.6.** Andamento delle armoniche per l'energia eolica

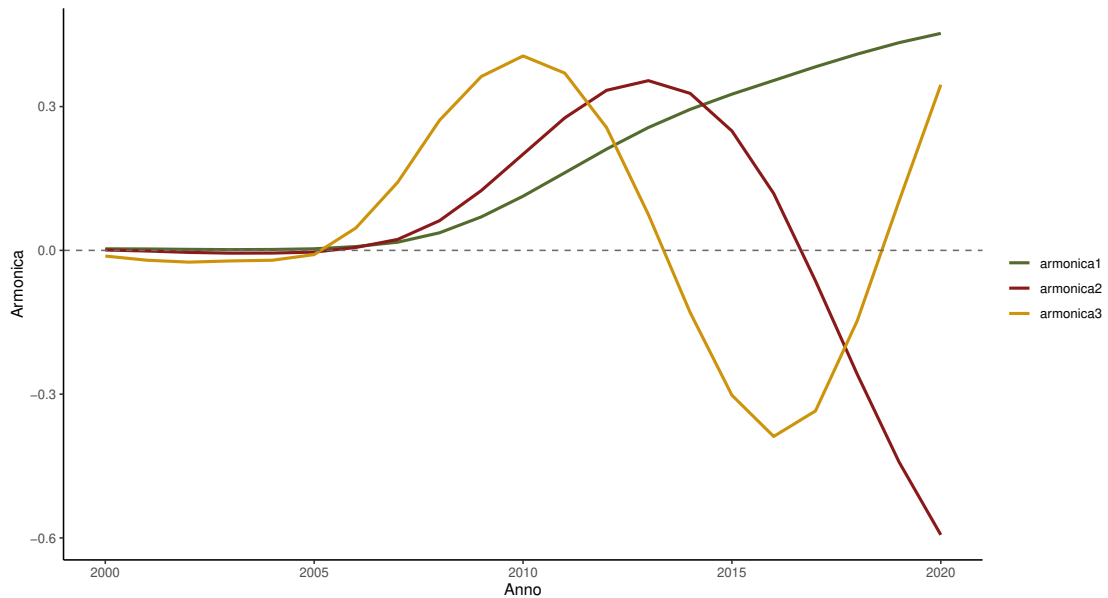


**Fig. 3.7.** Componenti principali per l'energia eolica

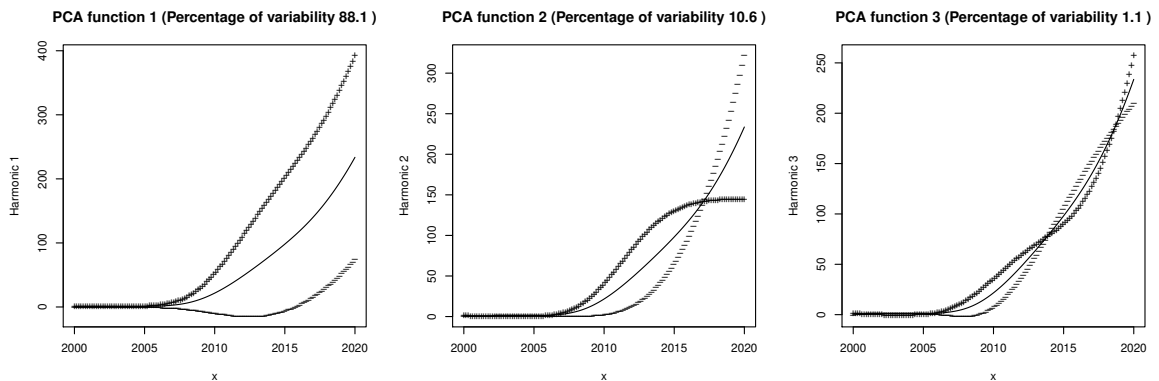


**Fig. 3.8.** Punteggi delle prime due armoniche per l'energia eolica

dell'eolico: l'andamento delle armoniche è molto simile ed è possibile fornire le stesse interpretazioni riguardo le componenti principali. Cambiano ovviamente le percentuali di variabilità spiegata da ogni componente principale, che in questo caso corrispondono all'88.1% per la prima, 10.6% per la seconda e 1.1% per la terza.

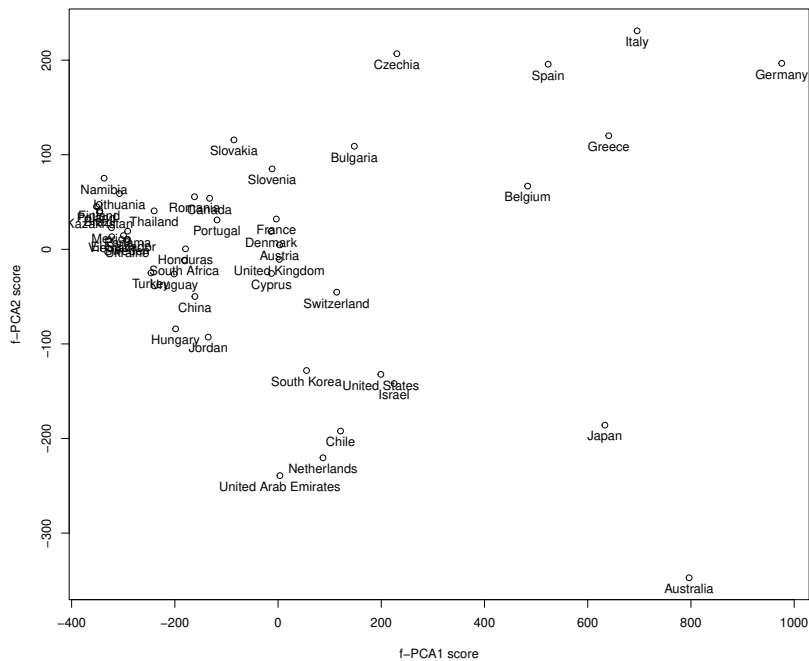


**Fig. 3.9.** Andamento delle armoniche per l'energia solare



**Fig. 3.10.** Componenti principali per l'energia solare

In Figura 3.11 vengono rappresentati i punteggi delle prime due PC per meglio comprendere i risultati ottenuti. Si nota come per i maggiori produttori di energia solare corrispondano punteggi positivi ed elevati sia della prima che della seconda armonica, come, ad esempio, per la Germania e la Spagna. A Paesi come Australia e Giappone sono associati valori positivi della prima armonica e negativi della seconda, pur essendo tra i maggiori produttori di energia solare; una possibile motivazione potrebbe essere che stiano ancora vivendo una fase di crescita di produzione, a differenza di Paesi come la Germania o l'Italia che sono più in una fase di maturazione.

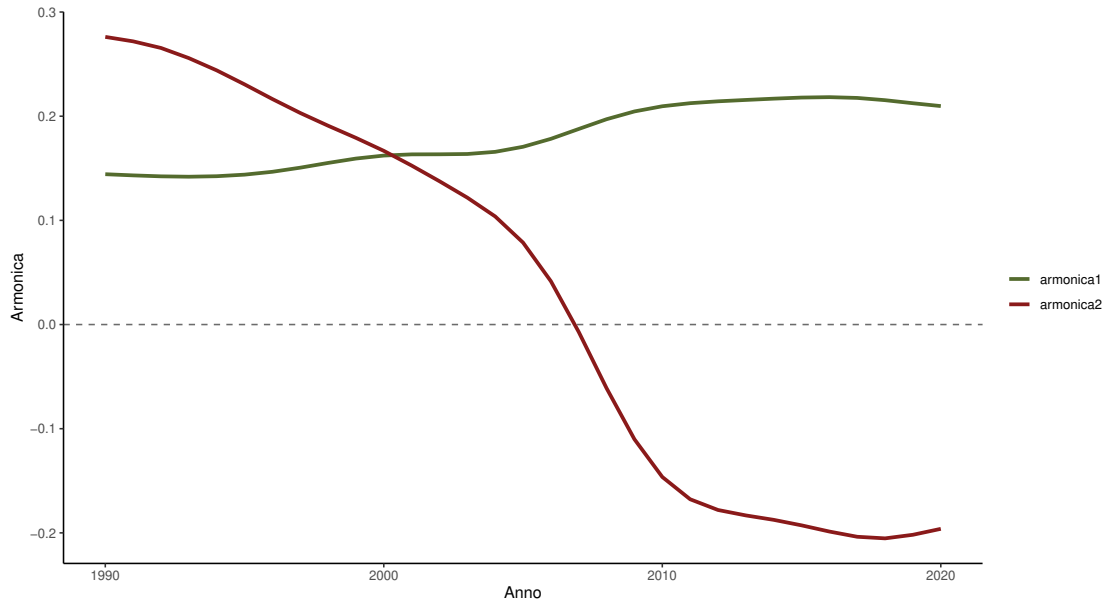


**Fig. 3.11.** Punteggi delle prime due armoniche per l'energia solare

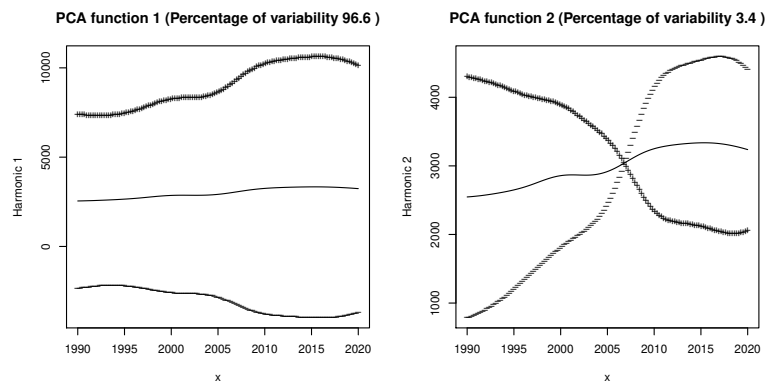
Anche per il caso dell'energia idroelettrica, la prima armonica è sempre positiva e rappresenta quegli stati che producono una quantità di energia di molto al di sopra

della media e spiega il 96.6% di variabilità.

La seconda armonica, invece, è positiva fino al 2007 e negativa nell'ultima parte dell'intervallo considerato.



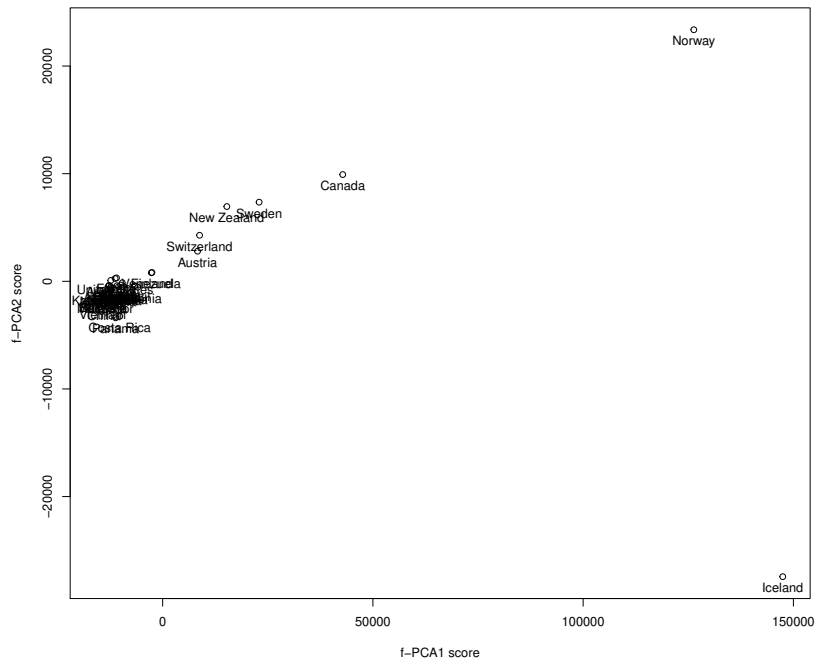
**Fig. 3.12.** Andamento delle armoniche per l'energia idroelettrica



**Fig. 3.13.** Componenti principali per l'energia idroelettrica



Analizzando i punteggi delle due armoniche in Figura 3.11 si nota come al maggior produttore di energia idroelettrica siano associati valori elevanti sia della prima che della seconda armonica.



**Fig. 3.14.** Punteggi delle prime due armoniche per l'energia eolica

### 3.3 Modello lineare simultaneo

Nell'ambito dell'analisi dei dati funzionali, è possibile applicare diversi modelli statistici che si differenziano in base alla natura delle variabili risposta e delle esplicative. Nel caso in esame, l'idea è prevedere la risposta funzionale relativa alla produzione pro capite di energia rinnovabile attraverso alcune covariate funzionali. In questo senso, allora, il modello più opportuno è il modello di regressione simultaneo: sia la  $y$  che le varie covariate  $z_j$  sono funzioni misurate allo stesso istante temporale  $t$  e l'influenza delle  $z_j$  è simultanea, ovvero  $z_j$  influenza  $y(t)$  solo attraverso il suo valore  $z_j(t)$  al tempo  $t$ , quindi non ci sono effetti ritardati o posticipati.

Siano  $z_{ij}$  le  $q$  covariate, ognuna delle quali è moltiplicata per il suo regressore  $\beta_j$ . Il modello di regressione simultaneo è

$$y_i(t) = \sum_{j=1}^q z_{ij}(t)\beta_j(t) + \varepsilon_i(t) \quad (3.22)$$

La notazione matriciale è

$$\mathbf{y}(t) = \mathbf{Z}(t)\beta(t) + \epsilon(t) \quad (3.23)$$

dove la matrice  $\mathbf{Z}$  di dimensione  $N \times q$  contiene le covariate funzionali  $z_{ij}$ , il vettore  $\beta$  di dimensione  $q$  contiene i coefficienti di regressione e  $\mathbf{y}$  è il vettore di dimensione  $N$  della risposta funzionale. Si stima un'espansione in funzioni di base sia per ogni osservazione, come in Equazione 3.2, sia i regressori  $\beta_j$ . Per questi ultimi l'espansione in funzioni di base può essere espressa come segue

$$\beta_j(t) = \sum_k^{K_j} b_{kj}\theta_{kj}(t) = \theta_j(t)^\top b_j(t). \quad (3.24)$$

È possibile lisciare anche i coefficienti  $\beta$ , penalizzando la somma dei quadrati dei residui

$$PENSSSE_\lambda(\beta) = \sum_{i=1}^n \int (y_i(t) - x_i(t)\beta_j(t))^2 dt + \sum_j \lambda_j \int [L\beta_j(t)]^2 dt \quad (3.25)$$

dove  $L$  è l'operatore differenziale, diverso a seconda del tipo di curva analizzata; se il fenomeno non è periodico si utilizza  $D^2$ .

Per esprimere il modello in Equazione 3.23 tenendo in considerazione la penalizzazione dei coefficienti  $\beta$ , è necessario definire alcune matrici a blocchi. Si definisca

$$K_\beta = \sum_j^q K_j \quad (3.26)$$

e il vettore  $\mathbf{b}$  di dimensione  $K_\beta$  strutturato inserendo i singoli  $\mathbf{b}_q$  verticalmente, ovvero

$$\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_q^\top) \quad (3.27)$$

Allora, in forma matriciale il modello in Equazione 3.23 può essere definito come

$$\mathbf{y} = \mathbf{Z}(t)\Theta(t)\mathbf{b} + \varepsilon(t) \quad (3.28)$$

dove la matrice  $\Theta$  di dimensione  $q \times q$  è

$$\Theta = \begin{bmatrix} \theta'_1 & 0 & \dots & 0 \\ 0 & \theta'_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \theta'_q \end{bmatrix}$$

Per calcolare gli intervalli di confidenza dei coefficienti  $\beta$  stimati, si definisca il residuo dell' $i$ -esima osservazione per le  $j$ -esima curva come segue

$$r_{ij} = y_{ij} - \mathbf{Z}_j(t_i)\beta(t_i) \quad (3.29)$$

e si stimi la matrice di varianza

$$\Sigma_\epsilon^* = \frac{1}{N} \mathbf{r}'\mathbf{r} \quad (3.30)$$

dove  $\mathbf{r}$  è la matrice dei residui. Sia  $\mathbf{C}$  la matrice di regressione dei coefficienti dell'espansione in base della variabile risposta, quindi  $\mathbf{y}(t) = \mathbf{C}\phi(t)$ . Allora la stima del

vettore  $\mathbf{b}$  si ottiene nel modo seguente

$$\begin{aligned}\hat{\mathbf{b}} &= \mathbf{A}^{-1} \left[ \int \Theta(t)' \mathbf{Z}(t) \mathbf{C} \phi(t) dt \right] \\ &= \mathbf{A}^{-1} \left[ \int \phi(t)' \otimes (\Theta(t)' \mathbf{Z}(t)') dt \right] \text{vec}(\mathbf{C}) \\ &= \mathbf{c2bMap} \text{vec}(\mathbf{C})\end{aligned}\tag{3.31}$$

dove  $\mathbf{c2bMap}$  è una matrice che converte i coefficienti di liscio della variabile risposta nel coefficiente di regressione. Utilizzando anche la matrice  $\mathbf{y2cMap}$ , la quale viene utilizzata per calcolare la matrice dei coefficienti di regressione  $\mathbf{C}$  delle osservazioni originali, la varianza di  $\mathbf{b}$  è ottenibile come

$$\text{Var}[\hat{\mathbf{b}}] = \mathbf{c2bMap} \mathbf{y2cMap} \Sigma_{\epsilon}^* \mathbf{y2cMap}' \mathbf{c2bMap}'\tag{3.32}$$

Per valutare la bontà di adattamento di un modello, è possibile utilizzare l' $R^2$  funzionale

$$R^2(t) = 1 - \frac{\sum (y_i(t) - \hat{y}_i(t))^2}{\sum (y_i(t) - \bar{y}_i(t))^2}\tag{3.33}$$

Per ottenere invece la significatività, si definisce la statistica  $F$  punto a punto come segue

$$F(t) = \frac{\text{var}\{\hat{y}(t)\}}{\sum (y_i - \hat{y}_i)^2 / (n - p)}\tag{3.34}$$

che indica dove è presente la maggior quantità di segnale relativamente alla varianza.

### 3.3.1 Applicazione del modello simultaneo ai dati

In questa sottosezione, vengono riportati i risultati derivanti dall'applicazione del modello simultaneo ai dati. Come variabile risposta si considera la produzione pro capite di energia e come esplicative il PIL pro capite, l'ISU (indice di sviluppo umano), la percentuale di popolazione che ha accesso a internet, la densità e il livello di emissioni di CO<sub>2</sub>. Ogni esplicativa, così come la variabile risposta, è stata lisciata, utilizzando il metodo descritto nel Paragrafo 3.1.

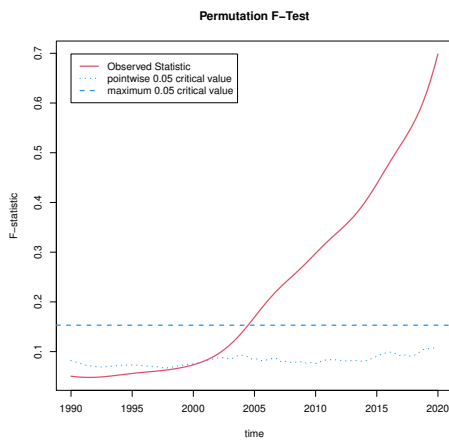
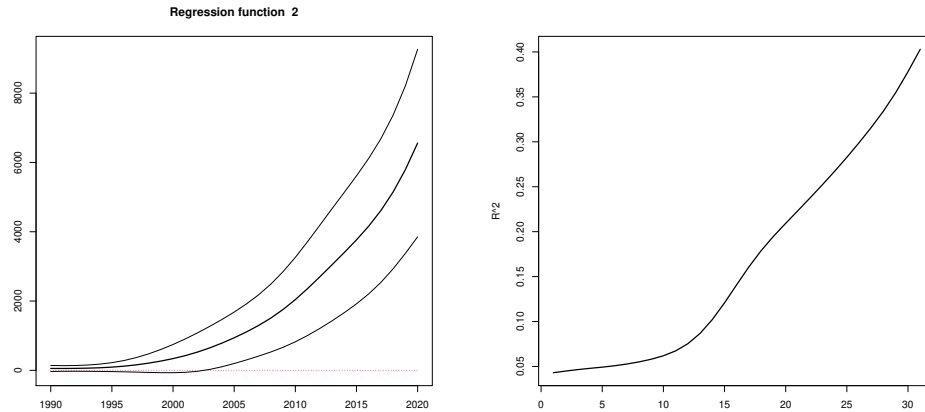
## Energia eolica

Per quanto riguarda l'energia eolica, si è adattato il modello simultaneo considerando prima tutte e cinque le esplicative. Nessuna delle covariate, però, è risultata significativa, in quanto gli intervalli di confidenza dei regressori stimati comprendevano lo 0; inoltre, l' $R^2$  assumeva un valore massimo di 0.18, a conferma della non accuratezza del modello nello spiegare il fenomeno di interesse. Per questa ragione, si è optato per una selezione *stepwise forward*, partendo prima con l'esplicativa maggiormente significativa, ovvero l'indice ISU, e aggiungendo ad ogni passo, successivo al primo, una delle covariate a disposizione. Nessuna delle variabili aggiunte ha portato ad un miglioramento della bontà di adattamento del modello e alla significatività dello stesso. Il miglior modello è quindi risultato essere quello con la sola covariata relativa all'indice ISU, il cui effetto sulla risposta viene riportato nel grafico in alto a sinistra in Figura 3.15. Dal grafico si evince che la variabile esplicativa ha un effetto positivo sulla variabile risposta e crescente via via che ci si avvicina al 2020. Infatti, variabili esplicative come quella analizzata apportano un contributo favorevole ad un processo di diffusione, come quello delle energie rinnovabili, essendo delle variabili di carattere socio-economico e di sviluppo di un Paese.

Per valutare la bontà di adattamento del modello, si riporta in Figura 3.15 nel grafico in alto a destra, la curva dell' $R^2$ ; l'andamento della curva è crescente e il valore massimo raggiunto è circa 0.40.

Infine, per cercare di capire la significatività del modello, si riporta il test  $F$  di permutazione in Figura 3.15 nel grafico in basso; si nota come ci sia poca evidenza di significatività del modello per i primi anni e come, invece, l'effetto sembri maggiormente rilevante negli ultimi anni.

Per il caso appena esaminato, si nota come il contributo della covariata, la signifi-



**Fig. 3.15.** Coefficiente di regressione,  $R^2$  e test  $F$  per l'energia solare

catività del modello e la bontà di adattamento abbiano un andamento simile, via via crescente man mano che ci si avvicina al 2020; questo è probabilmente dovuto al fatto che la maggior parte delle curve iniziano ad assumere valori positivi solo dopo il 2000, essendo l'energia eolica una fonte rinnovabile sfruttata soprattutto negli ultimi anni, per gli stati considerati.

### **Energia solare**

I risultati ottenuti dall'applicazione del modello simultaneo alla produzione pro capite di energia solare sono molto simili al caso dell'eolico. Anche in questo caso si è proceduto utilizzando una selezione delle variabili di tipo *stepwise forward*. Tale procedura ha portato alla determinazione di un modello con una sola variabile esplicativa, ovvero l'indice di sviluppo umano, in quanto le altre non sono risultate significative e/o non sono state in grado di apportare una maggiore significatività al modello.

Nei grafici in Figura 3.16 si riportano le curve relative all'andamento del coefficiente di regressione della covariata, all' $R^2$  e al test  $F$  di permutazione. Dal primo grafico a sinistra, si nota come il coefficiente di regressione relativo all'indice di sviluppo umano abbia un andamento crescente e positivo, soprattutto dopo il 2010, andamento simile al caso dell'eolico. Per quanto riguarda la curva relativa alla bontà di adattamento del modello si nota come questa abbia un andamento decrescente nella prima parte dell'intervallo e crescente nell'ultima parte, raggiungendo un valore massimo di 0.25. Lo stesso si può dire sulla significatività del modello, il quale risulta essere significativo sempre dopo il 2010.

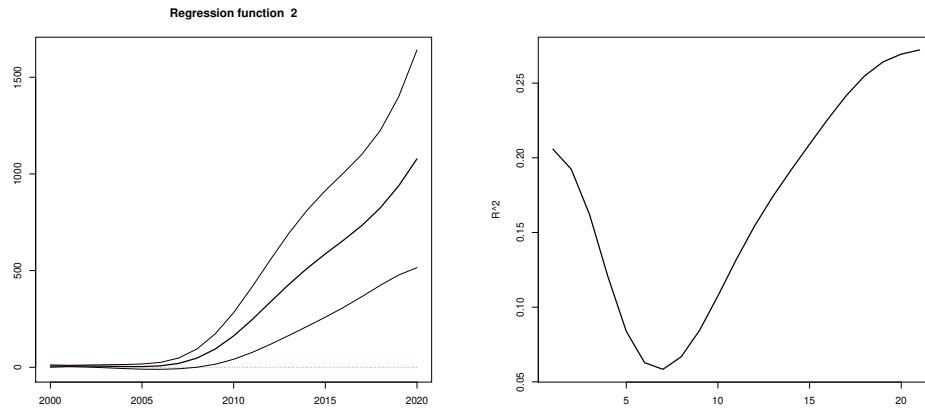
Si vuole ricordare che l'energia solare è la più recente tra le tre energie analizzate e il suo utilizzo inizia principalmente dopo la prima decade del 2000; è per questo che i risultati che si ottengono dall'applicazione del modello simultaneo non sono in generale

molto soddisfacenti, visto che la serie storica è molto breve, e al più risultano essere significativi solo dopo il 2010.

### **Energia idroelettrica**

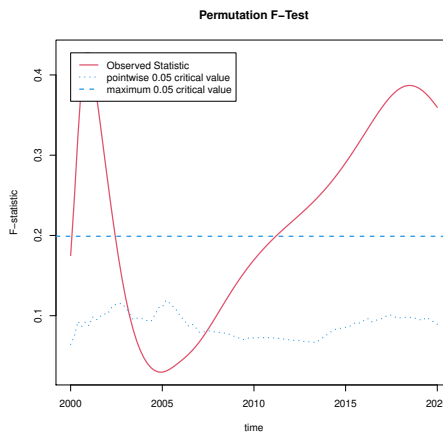
La situazione per l'energia idroelettrica è simile ai casi precedenti; dopo l'applicazione del modello con tutte e cinque le covariate, attraverso una selezione di tipo *stepwise forward*, l'unica esplicativa che è risultata essere significativa è il PIL pro capite. Si riportano di seguito, in Figura 3.17, i risultati ottenuti dall'applicazione del modello simultaneo. A differenza di quanto è emerso nei primi due casi, per l'energia idroelettrica non assistiamo a curve via via crescenti man mano che ci si avvicina all'estremo superiore dell'intervallo temporale. Infatti, questa fonte rinnovabile è tra le più sviluppate e il suo utilizzo è iniziato ben prima del 1990. Il contributo della covariata relativa al PIL pro capite è positivo e pressoché costante lungo tutto l'intervallo analizzato. La curva relativa all'indice di bontà di adattamento raggiunge il valore massimo di 0.45 in corrispondenza del 2005, proprio come il test  $F$  di permutazione e i loro andamenti risultano essere pressoché uguali.





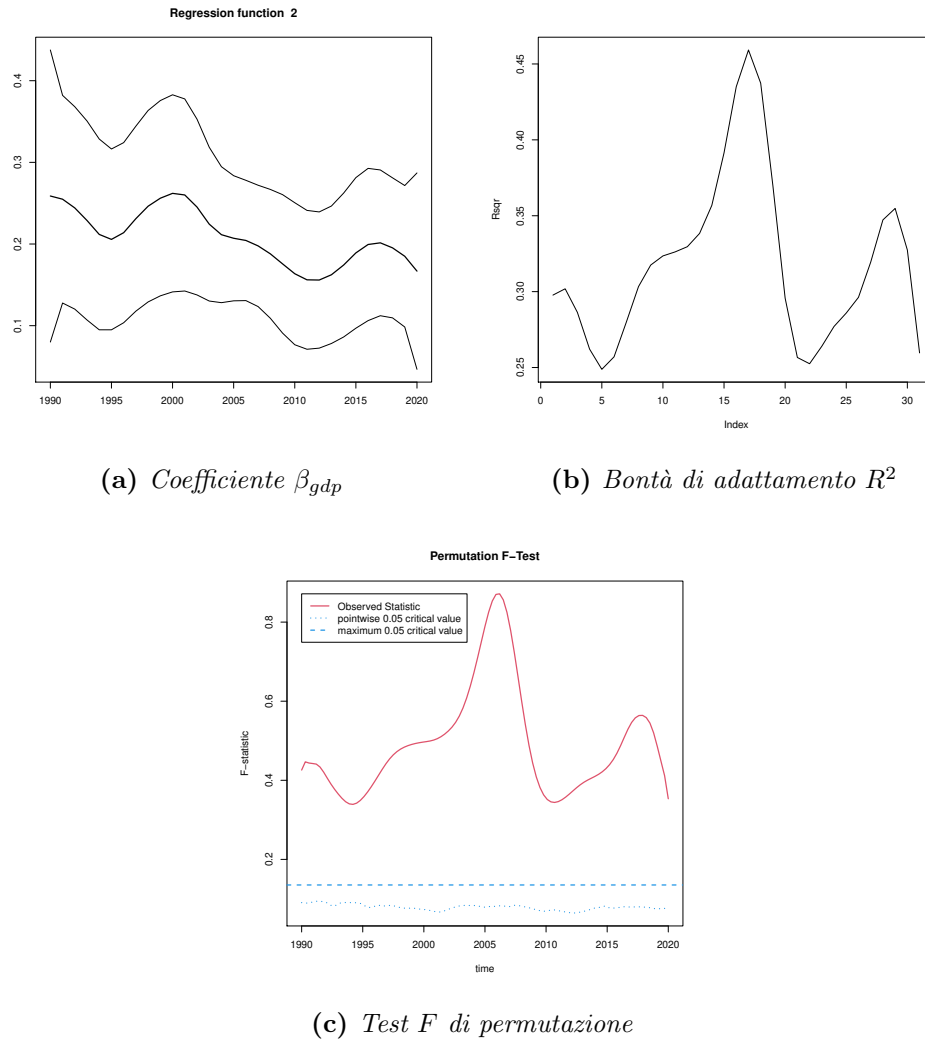
(a) Coefficiente  $\beta_{ISU}$

(b) Bontà di adattamento  $R^2$



(c) Test  $F$  di permutazione

Fig. 3.16. Coefficiente di regressione,  $R^2$  e test  $F$  per l'energia solare



**Fig. 3.17.** Coefficiente di regressione,  $R^2$  e test  $F$  per l'energia idroelettrica

### 3.4 Analisi dei gruppi funzionale

Per raggruppare le unità statistiche nell'ambito dei dati funzionali, sono disponibili diversi approcci. In questa Sezione si è scelto di trattare una metodologia che si basa su modelli mistura progettati per dati funzionali, la quale adatta i modelli di Bouveyron et al. (2007) proposti nel contesto multivariato, chiamati *High-dimensional data clustering* (HDDC). L'idea è quella di modellare e raggruppare le unità statistiche funzionali in sottospazi che abbiano una dimensione ridotta rispetto a quella di origine, individuando  $K$  gruppi omogenei. Il metodo di clustering risultante, detto funHDDC, è un metodo di raggruppamento model-based, il quale raggruppa i dati funzionali attraverso la loro proiezione su sottospazi funzionali, specifici del gruppo. Si consideri un insieme di  $n_k$  curve osservate appartenenti al gruppo  $k$ , ciascuna delle quali viene rappresentata attraverso una specifica espansione in funzioni di base, con vettore dei coefficienti  $\gamma_i \in \mathbb{R}^p$ ,  $i = 1, \dots, n_k$ . Si assuma che  $\{\gamma_1, \dots, \gamma_{n_k}\}$  siano realizzazioni indipendenti di un vettore casuale  $\Gamma \in \mathbb{R}^p$ . Inoltre, si assuma che il processo stocastico associato al  $k$ -esimo gruppo possa essere descritto in un sottospazio funzionale latente,  $L_2[0, T]$ , a bassa dimensionalità, di dimensione  $d_k \leq p$ . Si indichi con  $\{\psi_1, \dots, \psi_p\}$  il sistema di funzioni di base scelto per rappresentare i dati originali su uno spazio finito-dimensionale. Il nuovo sistema di funzioni di base sarà  $\{\varphi_1, \dots, \varphi_{d_k}\}$ , ottenuto mediante trasformazione lineare  $\varphi_{kj} = \sum_{l=1}^p q_{k,jl} \psi_l$ , con  $Q = (q_{k,jl}) = [U_k, V_k]$  matrice ortogonale  $p \times p$ , la quale può essere scomposta in due matrici:  $U_k$  di dimensione  $p \times d_k$  e  $V_k$  di dimensione  $p \times (p - d_k)$ , con  $U_k^\top U_k = I_{d_k}$ ,  $V_k^\top V_k = I_{p-d_k}$  e  $U_k^\top V_k = 0$ . Sia  $\{\lambda_1, \dots, \lambda_{n_k}\}$  il vettore latente dei coefficienti della nuova espansione in basi  $\{\varphi_{kj}\}_{j=1, \dots, d_k}$ , che si assumono essere realizzazioni indipendenti di un vettore casuale latente  $\Lambda \in \mathbb{R}^{d_k}$ . I vettori casuali  $\Lambda$  e  $\Gamma$  sono legati dalla seguente relazione:

$$\Gamma = U_k \Lambda + \varepsilon \tag{3.35}$$

dove  $\varepsilon \in \mathbb{R}^p$  è il termine d'errore indipendente e casuale. Si assuma, inoltre, per  $\Lambda$  la seguente distribuzione normale multivariata:

$$\Lambda \sim \mathcal{N}(m_k, S_k) \quad (3.36)$$

con  $S_k = \text{diag}(a_{k1}, \dots, a_{kd_k})$ . Ugualmente per  $\varepsilon$  si ha

$$\varepsilon \sim \mathcal{N}(0, \Xi_k) \quad (3.37)$$

Dalle assunzioni in Equazione 3.36 e 3.37, si ricava che la distribuzione di  $\Gamma$  è la seguente:

$$\Gamma \sim \mathcal{N}(\mu_k, \Sigma_k) \quad (3.38)$$

dove  $\mu_k = U_k m_k$  e  $\Sigma_k = U_k S_k U_k^\top + \Xi_k$ . La matrice  $\Xi_k$  è tale che  $\Delta_k = \text{cov}(Q_k^\top \Gamma) = Q_k^\top \Sigma_k Q_k$ , da cui si ricava che  $\Sigma = Q_k \Delta_k Q_k^\top$ . La matrice diagonale a blocchi  $\Lambda_k$  di dimensione  $p \times p$  ha la seguente forma:

$$\Lambda_k = \begin{bmatrix} S_k & \mathbf{0} \\ \mathbf{0} & b_k I_{p-d_d} \end{bmatrix}$$

Rispetto a quanto appena descritto, si nota come gli elementi diagonali di  $S_k$  modellano la varianza dei dati osservati, mentre il parametro  $b_k$  modella la varianza dell'errore che si osserva al di fuori dei sottospazi.

Si assuma l'esistenza di una variabile casuale latente  $Z = (Z_1, \dots, Z_K) \in \{0, 1\}^K$  la quale indica l'appartenenza della curva osservata  $x_i$  a uno dei  $K$  gruppi. L'obiettivo del raggruppamento è prevedere il valore  $z_i = (z_{i1}, \dots, z_{iK})$  di  $Z$  per ogni curva osservata  $x_i$ . Come in precedenza, sia  $\gamma$  il vettore dei coefficienti dell'espansione in base di  $x_i$ , la cui distribuzione è una mistura di variabili casuali normali:

$$p(\gamma) = \sum_{k=1}^K \pi_k \phi(\gamma; \mu_k, \Sigma_k) \quad (3.39)$$

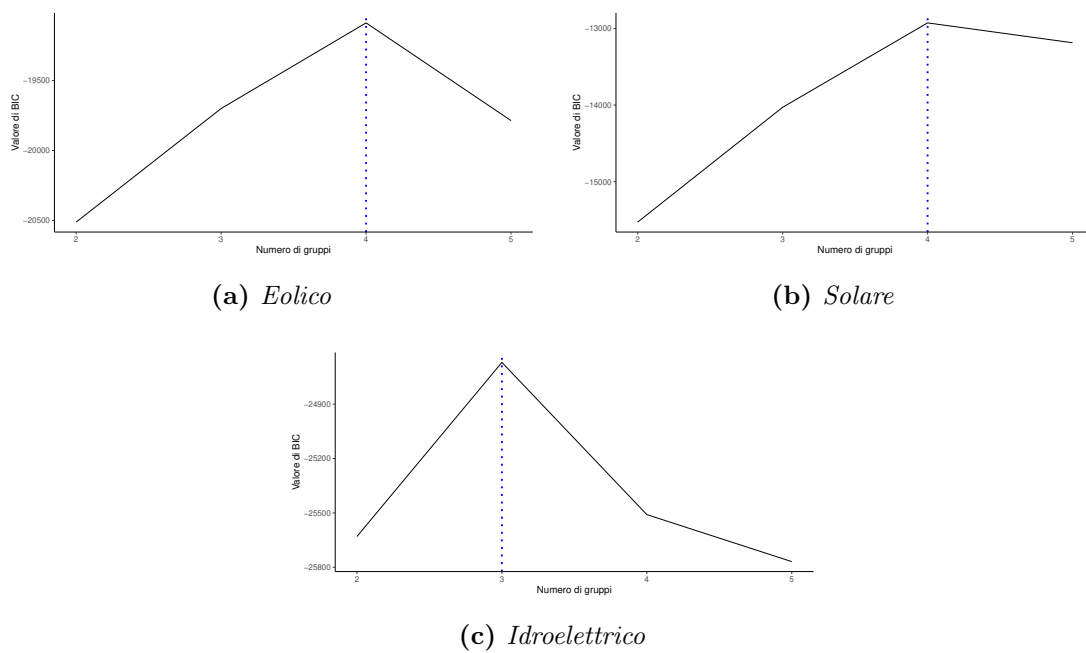
dove  $\pi_k = P(Z_k = 1)$  è la probabilità a priori del  $k$ -esimo gruppo e  $\phi$  è la densità gaussiana valutata in  $\gamma$ . Comprimendo i parametri  $a_{kj}, b_k, Q_k, d_k$  del modello è possibile ottenere rappresentazioni più o meno parsimoniose del modello. Il metodo più utilizzato per la stima dei parametri del modello appena definito è l'algoritmo iterativo EM. Attraverso questo algoritmo, si stimano tutti i parametri del modello, eccetto  $d_k$  e  $K$  i quali controllano la complessità del modello stesso. In particolare, gli autori propongono lo scree-test di Cattell (1966) per la stima di  $d_k$ , mentre per il numero di gruppi  $K$  suggeriscono il criterio BIC.

Dal momento che nell'algoritmo funHDDC si è interessati ad ottenere una partizione dei dati a disposizione, è necessario aggiungere un ulteriore step di classificazione, per ogni osservazione. In questo senso, si utilizza la regola MAP (*maximum a posteriori*), la quale assegna ogni osservazione  $\gamma_i \in \mathbb{R}^p$  al gruppo per il quale  $\gamma_i$  ha la più alta probabilità a posteriori  $P(Z_{ik} = 1 | \gamma_i)$ .

Si vuole sottolineare che, grazie alle proprietà di cui gode l'algoritmo EM, l'algoritmo funHDDC converge a un massimo locale, purché venga definito il punto di partenza dell'algoritmo. Inoltre, non è necessario invertire la matrice  $\Sigma_k$  grazie alla forma assunta da  $\Lambda_k$  e ciò rende applicabile questo metodo anche quando si trattano spazi ad alta dimensionalità, per i quali la matrice di covarianza è solitamente quasi singolare.

### 3.4.1 Applicazione dell'algoritmo funHDDC ai dati

Come descritto precedentemente, è necessario scegliere il numero  $K$  di gruppi, in corrispondenza del quale il BIC è più prossimo allo zero. Di seguito, in Figura 3.18, si rappresentano i tre grafici, uno per ciascun tipo energia rinnovabile, i quali evidenziano il numero di gruppi con cui suddividere il campione.



**Fig. 3.18.** Numero di gruppi derivanti dall'applicazione dell'algoritmo funHDDC

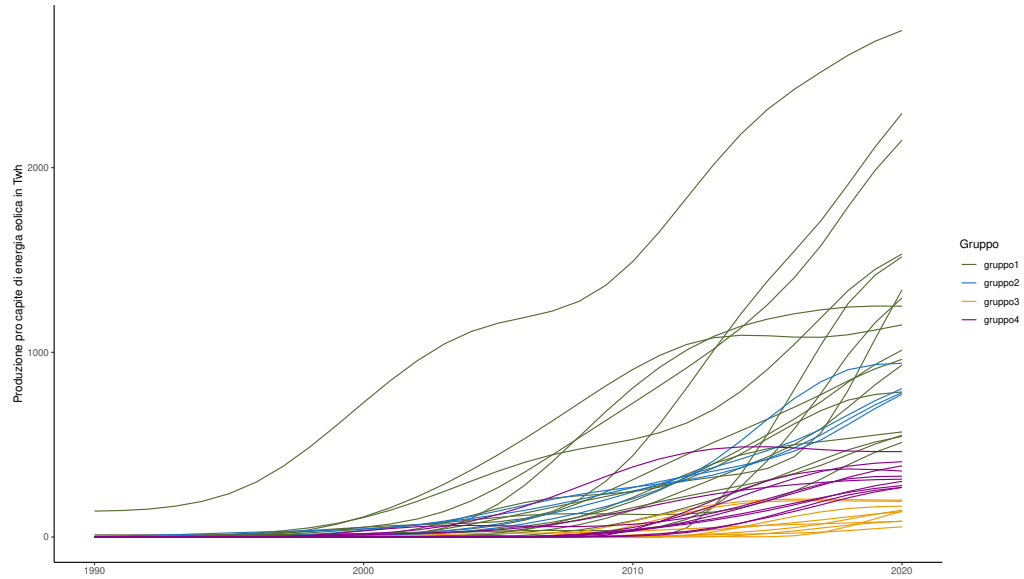
Si riporta nelle Tabelle 3.2, 3.3 e 3.4 la composizione dei gruppi evidenziati dall'algoritmo funHDDC, per i quali si fornisce una possibile interpretazione

In particolare, per l'energia eolica, in Figura 3.19 vengono evidenziati quattro gruppi; in generale, la caratteristica che accomuna gli stati all'interno di ciascun gruppo è in termini di quantità prodotta. Infatti, Paesi come Australia, Canada, Grecia e Paesi Bassi, appartenenti al secondo gruppo e riportati in Figura 3.20, producono circa la stessa quantità di energia eolica pro capite.

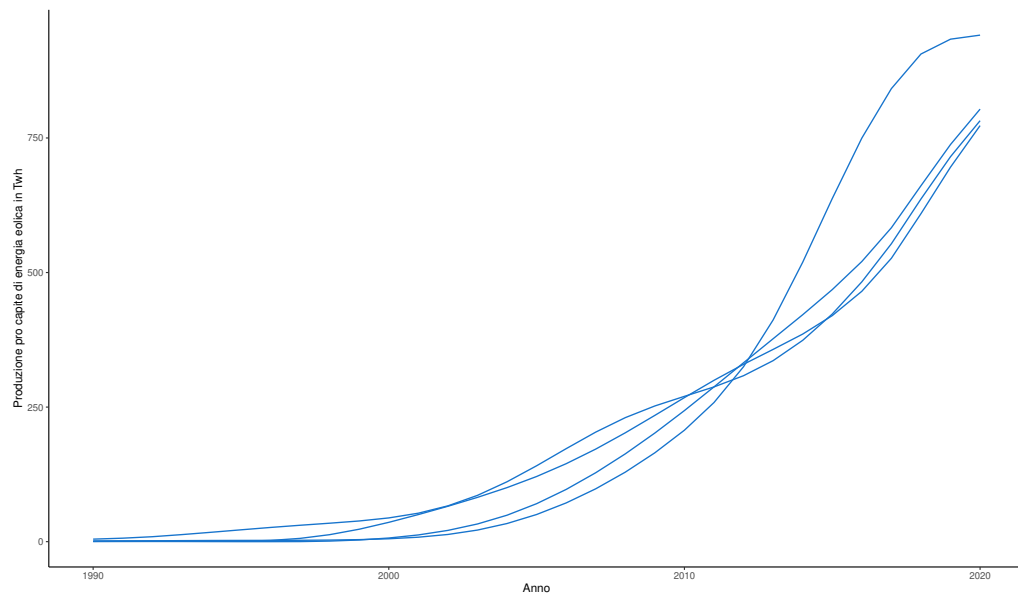
L'unico gruppo i cui stati sembrano essere accomunati da un fattore geografico, oltre alla quantità prodotta, è il quarto gruppo, il più numeroso; infatti, è costituito principalmente da Paesi europei, ai quali si aggiungono l'Uruguay e gli Stati Uniti.

Gruppo	Paese
1	Austria, Belgio, Danimarca, Estonia, Finlandia, Francia, Germania, Irlanda, Lituania, Lussemburgo, Norvegia, Portogallo, Spagna, Svezia, Regno Unito, Uruguay, Usa
2	Australia, Canada, Grecia, Paesi Bassi
3	Argentina, Bulgaria, Cipro, Giordania, Lettonia, Messico, Mongolia, Panama, Ucraina
4	Brasile, Cile, Cina, Costa Rica, Croazia, Italia, Nuova Zelanda, Polonia, Romania, Turchia

**Tab. 3.2.** Gruppi individuati dall'algoritmo funHDDC per l'energia eolica



**Fig. 3.19.** Energia eolica: curve dei gruppi derivanti dall'algoritmo funHDDC



**Fig. 3.20.** Curve dei Paesi appartenenti al Gruppo2 per l'energia eolica

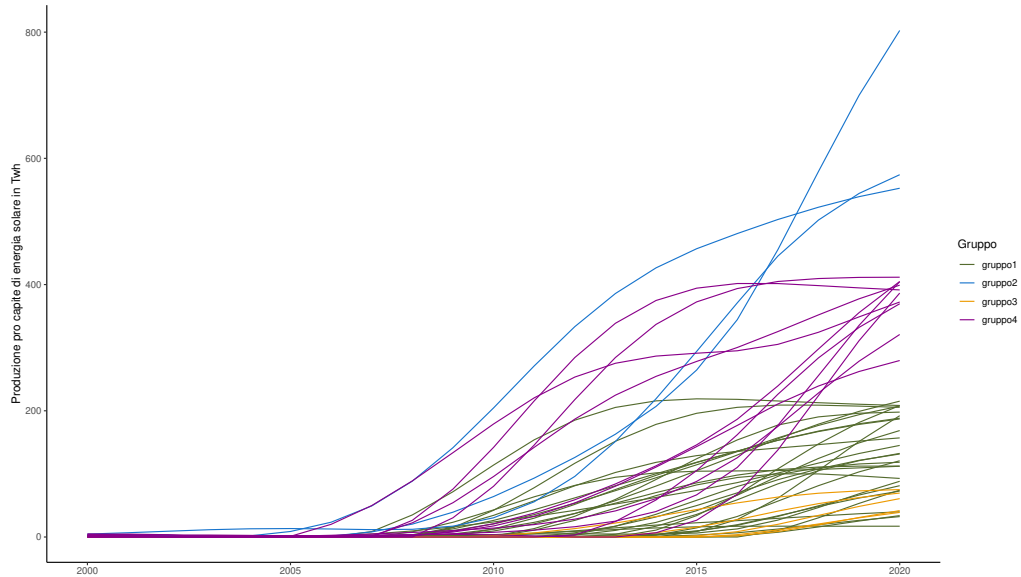


Per quanto riguarda l'energia solare sono stati selezionati quattro gruppi, elencati in Tabella 3.3 e raffigurati in Figura 3.21. Come per il caso dell'eolico, in generale, le curve sembrano essere suddivise in base alla quantità di energia pro capite prodotta; se ne riporta un esempio nel grafico a sinistra in Figura 3.22, rappresentante il terzo gruppo. Questa interpretazione può essere estesa anche ai gruppi uno e quattro.

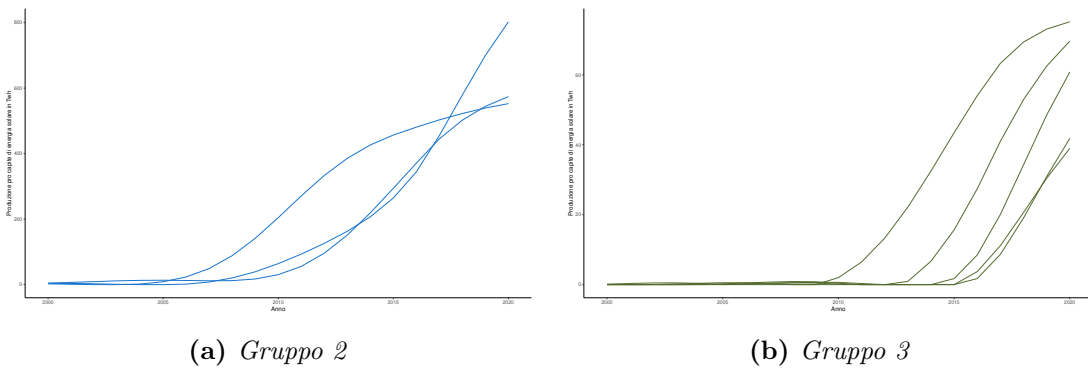
Invece, uno dei gruppi a cui può essere associata un'interpretazione diversa è il terzo, rappresentato a destra nella medesima figura del precedente, in quanto è costituito da stati tra i maggiori produttori al mondo di energia solare, ovvero Australia, Germania e Giappone.

Gruppo	Paese
1	Austria, Bulgaria, Canada, Cina, Cipro, Danimarca, El Salvador, Finlandia, Francia, Giordania, Honduras, Lituania, Namibia, Polonia, Portogallo, Regno Unito, Repubblica Ceca, Romania, Slovacchia, Slovenia, Sud Africa, Svezia, Turchia, Ucraina, Ungheria, Usa, Uruguay
2	Australia, Germania, Giappone
3	Brasile, Kazakistan, Messico, Panama, Thailandia
4	Belgio, Corea del Sud, Cile, Emirati Arabi Uniti, Grecia, Israele, Italia, Paesi Bassi, Spagna, Svizzera, Usa

**Tab. 3.3.** Gruppi individuati dall'algoritmo funHDDC per l'energia solare



**Fig. 3.21.** Energia solare: curve dei gruppi derivanti dall'algoritmo funHDDC



**Fig. 3.22.** Curve dei Paesi del Gruppo2 e Gruppo3 per l'energia solare

Infine, i gruppi evidenziati dall'algoritmo per l'energia idroelettrica rappresentano un caso apparentemente diverso. Sono stati selezionati tre gruppi, rappresentati in Figura 3.23 ed elencati in Tabella 3.4, e l'andamento delle curve in ciascun gruppo è diverso; se per i casi precedenti la suddivisione sembrava esser avvenuta principalmente

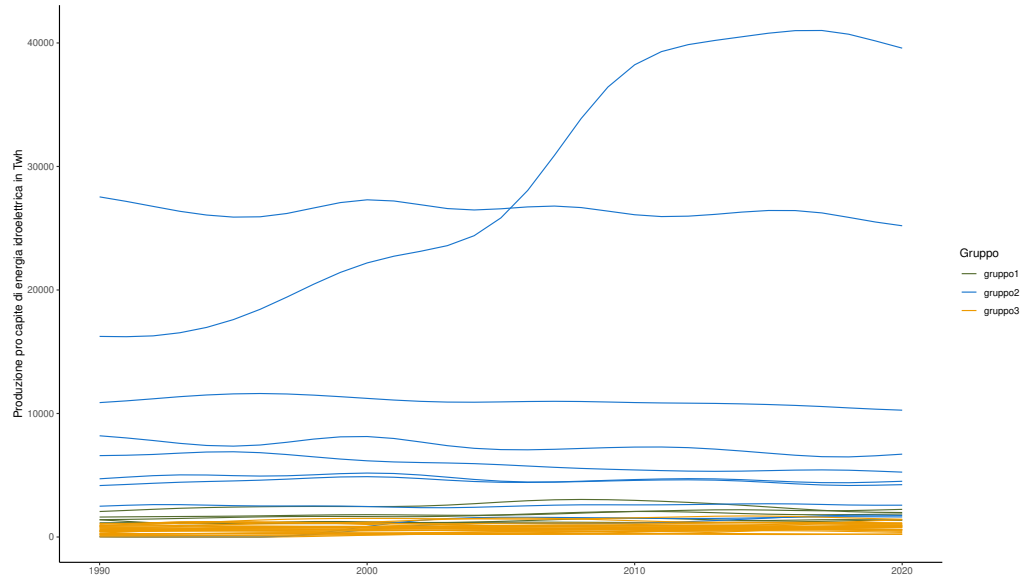
in termini di quantità, ora sembra che ciò che accomuna gli stati all'interno di ciascun gruppo è l'evolversi della curva di produzione pro capite di energia.

Gruppo	Paese
1	Brasile, Finlandia, Lettonia, Russia, Slovenia
2	Austria, Canada, Costa Rica, Islanda, Norvegia, Nuova Zelanda, Panama, Svezia, Svizzera, Venezuela
3	Argentina, Australia, Bolivia, Bulgaria, Cile, Cina, Colombia, Croazia, Ecuador, Francia, Giappone, Grecia, Italia, Kazakistan, Malesia, Messico, Perù, Portogallo, Romania, Serbia, Slovacchia, Spagna, Turchia, Usa, Vietnam

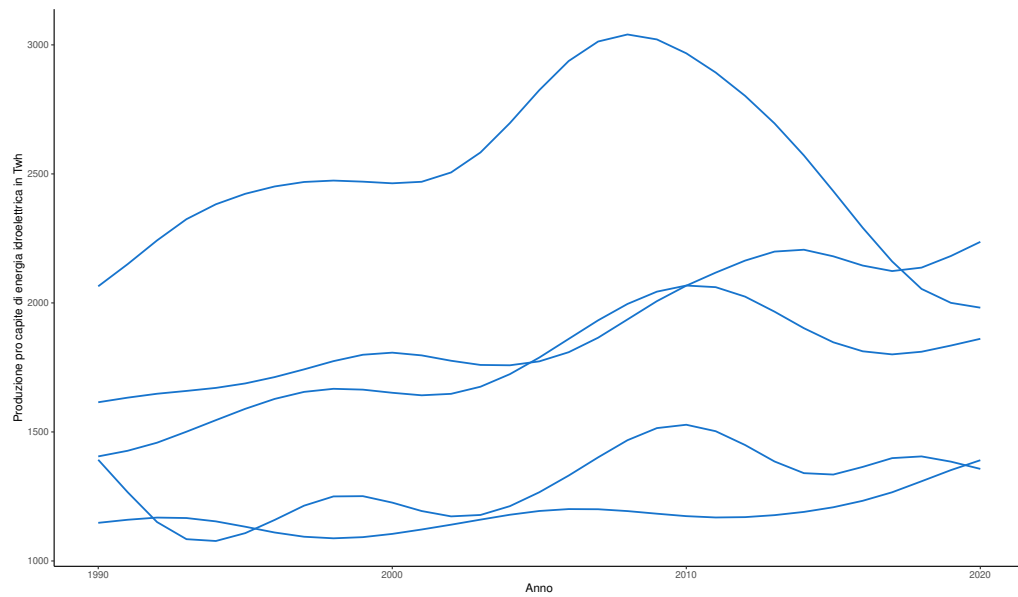
**Tab. 3.4.** Gruppi individuati dall'algoritmo funHDDC per l'energia idroelettrica

In particolare, il secondo gruppo, in Figura 3.24 è composto da stati principalmente nordici o dove l'energia idroelettrica rappresenta una delle principali fonti di energia rinnovabile e per i quali l'andamento della curva sembra essere quasi sempre costante nel tempo.

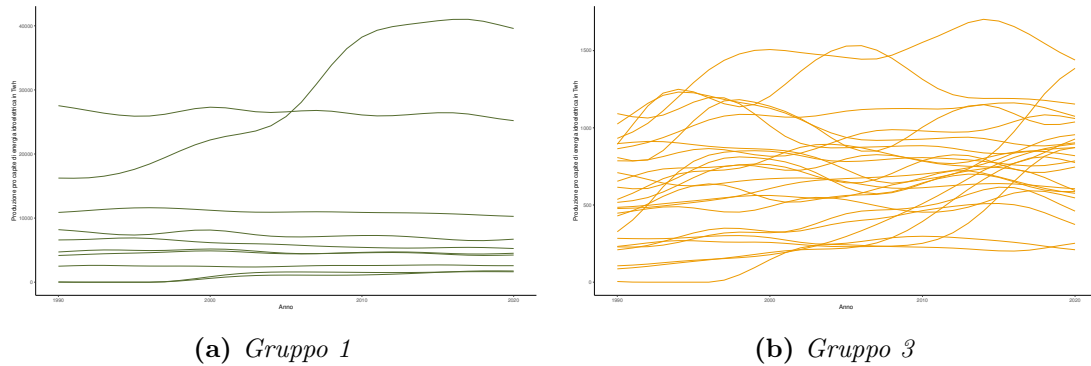
Le curve delle nazioni appartenenti agli altri due gruppi, in Figura 3.25, hanno un andamento meno regolare nel corso degli anni, con curve che oscillano maggiormente nell'intervallo considerato; la differenza sostanziale tra questi due gruppi, oltre alla diversa numerosità, è la quantità prodotta.



**Fig. 3.23.** Energia idroelettrica: curve dei gruppi derivanti dall’algoritmo funHDDC



**Fig. 3.24.** Curve dei Paesi appartenenti al Gruppo2 per l’energia idroelettrica



**Fig. 3.25.** Curve dei Paesi del Gruppo2 e Gruppo3 per l'energia idroelettrica

### 3.5 Modellazione per prevedere

Le analisi condotte in questo Capitolo hanno permesso di comprendere, in primo luogo, quali variabili hanno un impatto significativo sulla produzione pro capite di energia rinnovabile. Tra queste variabili, l'indice di sviluppo umano e il PIL pro capite risultano essere significative. Essendo variabili di carattere socio-economico, indicano che la crescita delle energie rinnovabili analizzate è influenzata da un fattore culturale.

Inoltre, si osserva che l'evoluzione della produzione di energia idroelettrica presenta differenze rispetto a quella delle energie solare ed eolica, essendo la fonte più sviluppata tra le tre. Infatti, i risultati evidenziano un comportamento simile tra eolico e solare: in entrambi i casi, l'indice ISU risulta essere la variabile significativa. Inoltre, nell'analisi di raggruppamento si sono selezionati quattro gruppi e nell'analisi delle componenti principali sono state individuate tre armoniche per descrivere il fenomeno. Nel caso dell'energia idroelettrica, invece, il PIL pro capite è la variabile più significativa; dall'analisi di raggruppamento sono emersi tre possibili gruppi, mentre la f-PCA ha evidenziato al presenza di due armoniche. Questa distinzione sottolinea il fatto che quanto più un paese è culturalmente avanzato, tanto maggiore è la crescita dell'ado-

zione di determinate fonti di energia, soprattutto per l'energia eolica e solare, che sono le più recenti tra le tre fonti analizzate. Per quanto riguarda l'energia idroelettrica, invece, il PIL pro capite ha un'influenza positiva, il che non sorprende poiché è la fonte energetica più sviluppata tra le tre, suggerendo che maggiore è la produzione, maggiore è anche la spesa e il consumo energetico della popolazione.

L'analisi di raggruppamento ha permesso di identificare gruppi di paesi che condividono simili quantità di produzione di energia. Ciò ha consentito di discriminare tra i maggiori produttori di energia, considerati pionieri nell'adozione delle energie rinnovabili, e coloro che possono essere definiti *"late adopters"*, ovvero paesi che si sono inseriti più recentemente nel mercato delle energie rinnovabili.

Complessivamente, le analisi condotte in questo capitolo evidenziano l'importanza delle variabili socio-economiche e culturali nello sviluppo e nell'adozione delle energie rinnovabili. I risultati sottolineano l'importanza di un alto sviluppo umano e di una cultura orientata all'innovazione e alla sostenibilità per promuovere l'adozione di fonti energetiche rinnovabili.



# Conclusioni

Questo elaborato di tesi è stato svolto con l'intento di perseguire un duplice obiettivo: *spiegare* il processo di diffusione della produzione pro capite di energia rinnovabile e *prevedere* il suo comportamento.

L'analisi basata sul modello di Bass ha consentito di comprendere e spiegare il processo di diffusione delle energie rinnovabili nei paesi considerati, classificandoli come innovatori o imitatori. I paesi innovatori sono quelli che hanno adottato le energie rinnovabili inizialmente, avviando il processo di diffusione, mentre i paesi imitatori sono entrati nel mercato in un secondo momento, seguendo l'esempio degli innovatori. È interessante notare che entrambe le categorie, innovatori e imitatori, stanno ancora vivendo una fase di crescita nella diffusione delle energie rinnovabili. Ciò indica che c'è ancora un ampio margine di sviluppo e adozione di queste fonti rinnovabili in tutto il mondo. Inoltre, l'aumento della produzione di energia rinnovabile evidenzia un crescente interesse a livello internazionale per ridurre la dipendenza dai combustibili fossili e promuovere soluzioni energetiche sostenibili. Questo indica una consapevolezza globale dell'importanza della transizione verso fonti energetiche più pulite per affrontare sfide come il cambiamento climatico.

Su questa base, è stata condotta un'analisi di raggruppamento utilizzando i valori dei coefficienti  $m$ ,  $p$  e  $q$  derivanti dal modello di Bass, al fine di identificare eventua-



li fattori comuni nel comportamento dei diversi paesi considerati. Da questa analisi, sono emersi gruppi significativi e coerenti. Alcuni gruppi sono composti interamente da stati innovatori o imitatori, mentre altri includono paesi innovatori o imitatori che si distinguono per valori molto elevati dei coefficienti  $p$  e  $q$ . Questa classificazione, ha permesso di evidenziare quei paesi considerati pionieri nell'adozione delle energie rinnovabili, come Spagna e Germania per l'energia solare e Danimarca per l'energia eolica. Il comportamento di questi paesi è stato e continua ad essere di estrema importanza, in quanto rappresenta un modello virtuoso per gli altri paesi che si stanno avvicinando al mercato delle fonti rinnovabili.

Per raggiungere il secondo obiettivo di prevedere l'evoluzione futura della produzione di energia rinnovabile, è stata utilizzata un'analisi che considera la serie storica come un'unità statistica, nota come analisi dei dati funzionali. A tale scopo, si sono dapprima lisciati i dati per rimuovere il rumore che spesso li caratterizza. Successivamente, è stato applicato un modello simultaneo che ha preso in considerazione alcune variabili esplicative di natura economica e di sviluppo tecnologico di un paese. Tra le variabili esplicative utilizzate vi sono il PIL pro capite, l'indice di sviluppo umano (ISU), la percentuale di popolazione con accesso a Internet, la densità di popolazione e le emissioni di CO<sub>2</sub>. Queste variabili sono state considerate al fine di comprendere meglio i fattori che influenzano la produzione di energia rinnovabile e di poter effettuare previsioni sull'andamento futuro. L'analisi dei dati funzionali ha consentito di analizzare in modo più approfondito le dinamiche temporali e identificare le relazioni tra le variabili esplicative e la produzione di energia rinnovabile.

È interessante notare che le variabili significative nel modello differiscono a seconda del tipo di energia rinnovabile considerata. Per l'energia solare ed eolica, l'unico fattore significativo è l'indice di sviluppo umano (ISU), il quale ha un effetto positivo e crescente sulla produzione di energia. D'altra parte, per l'energia idroelettrica, la variabile signi-

ficativa è il PIL pro capite, il quale influisce positivamente sulla produzione di energia. Il fatto che le uniche variabili significative siano correlate allo sviluppo socio-economico e culturale di un paese sottolinea l'importanza del reddito, dell'aspettativa di vita e del livello di istruzione nel promuovere la produzione di energia rinnovabile. Questi risultati indicano che i paesi con un maggiore sviluppo socio-economico e una migliore qualità della vita tendono ad adottare e promuovere in modo più efficace le energie rinnovabili. Ciò suggerisce l'importanza di politiche e strategie volte a migliorare questi indicatori socio-economici per favorire la transizione verso un sistema energetico più sostenibile. Anche in questo caso, a completamento del modello di regressione simultaneo, è stata svolta un'analisi di raggruppamento. In questo caso, l'obiettivo era capire se le curve di produzione presentassero delle caratteristiche comuni che potessero accomunare gli stati. I risultati ottenuti hanno evidenziato la presenza di gruppi in cui le unità statistiche sono accomunate dalla quantità di energia prodotta. Ciò significa che alcuni paesi mostrano comportamenti simili nella produzione di energia rinnovabile, indipendentemente da fattori geografici.

Pur avendo utilizzato due approcci diversi, il modello di Bass da una parte e i modelli per dati funzionali dall'altra, le analisi di raggruppamento condotte hanno portato a risultati simili in termini di raggruppamento degli stati. Ciò suggerisce che entrambi gli approcci sono stati in grado di individuare similitudini e pattern comuni nella produzione di energia rinnovabile tra gli stati.



# Bibliografia

- [1] A. Azzalini, B. Scarpa (2012), *Data Analysis and Data Mining: an introduction*, Oxford University Press.
- [2] C. Bouveyron, S. Girard, C. Schmid (2007), High-Dimensional Data Clustering, *Computational Statistics and Data Analysis* **52**(1), 502–519.
- [3] C. Bouveyron, J. Jacques (2011), Model-Based Clustering of Time Series in Group-Specific Functional Subspaces, *Advances in Data Analysis and Classification* **5**(4), 281-300.
- [4] M. Guidolin (2023) *Innovation Diffusion Models: Theory and Practice*, First Edition. John Wiley and Sons.
- [5] E. Muller, R. Peres, Y. Mahajan (2009), *Innovation Diffusion and New Product Growth*, Marketing Science Institute.
- [6] J. Ramsay, G. Hooker, S. Graves (2009), *Functional data analysis with R and MATLAB*, Springer.
- [7] J. Ramsay, B. W. Silverman (2005), *Functional Data Analysis*, Springer.
- [8] G. Shmueli (2010), To Explain or to Predict?, *Statistical Science*, **25**(3), 289-310.



# Sitografia

- <https://github.com/owid/energy-data>
- <https://www.key4biz.it/solare-ed-eolico-ecco-la-top-ten-mondiale-dei-paesi-piu-green/402475/>
- <https://www.studiopitagora.net/it/blog/i-paesi-piu-rinnovabili-del-mondo-i-passi-fatti-e-gli-obiettivi-da-raggiungere-256>
- <https://www.hivepower.tech/it/blog/energia-rinnovabile-in-lituania-cosa-dovresti-sapere>
- <https://www.rinnovabili.it/energia/idroelettrico/costa-rica-300-giorni-rinnovabili/>
- <https://www.vpsolar.com/il-mercato-fotovoltaico-in-spagna-si-preannuncia-uno-dei-piu-interessanti-nei-prossimi-anni/>
- <https://www.enelgreenpower.com/it/media/news/2022/08/vietnam-nuova-compagnia>
- <https://www.edison.it/it/onoff/vietnam-impianto-fotovoltaico-galleggiante>
- <https://www.assocamerestero.it/notizie/espansione-dellimpianto-solare-fotovoltaico-sao-goncalo-brasile>

## Sitografia

---

- <https://www.amistades.info/post/emirati-arabi-uniti-giordania-israele-baratto-energetico>
- [https://news.ihsmarkit.com/prviewer/release\\_only/slug/bizwire-2021-5-5-ihsmarkit-rankings-show-united-states-already-the-worlds-most-attractive-market-for-renewables-investment](https://news.ihsmarkit.com/prviewer/release_only/slug/bizwire-2021-5-5-ihsmarkit-rankings-show-united-states-already-the-worlds-most-attractive-market-for-renewables-investment)
- <https://www.edison.it/it/onoff/energia-eolica-danimarca>





# Appendice

In questa sezione, si presentano i risultati ottenuti applicando il modello Guseo-Guidolin (Guseo e Guidolin, 2009), per i quali non si è ritenuto opportuno dedicare uno dei principali capitoli in quanto l'applicazione del modello in vari paesi non ha portato esiti soddisfacenti.

Se consideriamo un mercato potenziale dinamico, allora il modello di Bass diventa:

$$z'(t) = m(t) \left[ p + q \frac{z(t)}{m(t)} \right] \left[ 1 - \frac{z(t)}{m(t)} \right] + m'(t) \frac{z(t)}{m(t)} \quad t > 0 \quad (3.40)$$

dove la variazione del mercato potenziale ha un effetto sulle adozioni istantanee, la quale può essere positiva e rinforzante, se  $m(t)$  è crescente, o negativa, se  $m(t)$  è decrescente. La soluzione in forma chiusa dell'equazione è

$$z(t) = m(t) \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p} e^{-(p+q)t}} \quad (3.41)$$

dove  $m(t)$  è una funzione libera. Nel modello Guseo-Guidolin, si considera che lo sviluppo del mercato potenziale dipenda da un processo di comunicazione del nuovo prodotto, il quale precede la fase di adozione. Il mercato potenziale prende allora la seguente forma

$$m(t) = K \sqrt{\frac{1 - e^{-(p_c+q_c)t}}{1 + \frac{q_c}{p_c} e^{-(p_c+q_c)t}}} \quad t > 0 \quad (3.42)$$

dove i parametri  $p_c$  e  $q_c$  governano il processo di comunicazione:  $p_c$  descrive il comportamento degli innovatori che iniziano a parlare del nuovo prodotto, mentre  $q_c$  rappresenta

la forza di diffusione della parola una volta che si è ricevuta l'informazione, rendendola virale. Il parametro  $K$  indica il comportamento asintotico di  $m(t)$  quando tutti i consumatori informati diventeranno eventualmente utenti. La struttura cumulata è:

$$z(t) = K \sqrt{\frac{1 - e^{-(p_c+q_c)t}}{1 + \frac{q_c}{p_c} e^{-(p_c+q_c)t}} \cdot \frac{1 - e^{-(p_s+q_s)t}}{1 + \frac{q_s}{p_s} e^{-(p_s+q_s)t}}} \quad t > 0 \quad (3.43)$$

dove le adozioni cumulate sono descritte come il prodotto di due fasi separate: comunicazione ( $p_c$  e  $q_c$ ) e adozione ( $p_s$  e  $q_s$ ). Si noti che il modello di Bass può essere visto come un caso speciale del modello Guseo-Guidolin, dove il propagarsi dell'informazione è così veloce che un insieme potenziale di utenti sono pronti ad acquistare non appena il prodotto entra nel mercato, ovvero  $m(t) = K$ .

Per capire quale modello fosse preferibile tra il modello di Bass e il modello Guseo-Guidolin, si è utilizzato il coefficiente di correlazione parziale multiplo al quadrato,  $\tilde{R}^2$ , che assume valori nell'intervallo  $[0; 1]$ . Questo indice permette il confronto tra due modelli annidati, uno completo,  $m_2$ , e uno ridotto,  $m_1$ , la cui formulazione è la seguente:

$$\tilde{R}^2 = (R_{m_2}^2 - R_{m_1}^2)/(1 - R_{m_1}^2) \quad (3.44)$$

Un modello completo si può considerare migliore rispetto ad uno ridotto se  $\tilde{R}^2 \geq 0.3$ . Per valutare quale modello fosse preferibile per il caso in esame, si è pensato di considerare anche la significatività dei coefficienti del modello; quindi, il modello Guseo-Guidolin sarà preferibile se  $\tilde{R}^2 \geq 0.3$  e tutti i coefficienti risultano significativi almeno al 10%. Date queste considerazioni, nelle tabelle di seguito è riportato un elenco dei Paesi per i quali il modello Guseo-Guidolin è risultato essere il migliore, per tutte e tre le energie rinnovabili analizzate.

Al variare dei valori dei parametri  $p_c$  e  $q_c$ , fissati  $p_s$  e  $q_s$ , si possono presentare diverse situazioni; infatti, se nel mercato considerato è presente una buona comunicazione, ovvero ad esempio  $p_c = 0.01$  e  $q_s = 0.9$ , allora il mercato potenziale subirà una crescita

Stato	$\tilde{R}^2$	Tipo di energia
Danimarca	0.8364	Eolica
Nuova Zelanda	0.9558	Eolica
Portogallo	0.9816	Eolica
Spagna	0.9062	Eolica
Croazia	0.9218	Idroelettrica
Serbia	0.9380	Idroelettrica
Slovenia	0.9254	Idroelettrica
Belgio	0.9880	Solare
Germania	0.9880	Solare
Regno Unito	0.9807	Solare

**Tab. 3.5.** Paesi per i quali il modello Guseo-Guidolin è preferibile al modello di Bass

veloce. Al contrario, se il processo di comunicazione non è efficace, con  $p_c = 0.001$  e  $q_c = 0.05$ , allora la crescita del mercato potenziale sarà lenta. Nei tre casi esaminati, ci troviamo davanti a una situazione di mezzo, in cui i valori di  $p_c$  e  $q_c$  sono rispettivamente circa 0.001 e 0.5, quindi sembrerebbe che il processo di comunicazione non sia né troppo convincente, ma nemmeno non efficiente. Analizzando nello specifico i Paesi per i quali il modello Guseo-Guidolin è risultato essere preferibile al modello di Bass, per il caso dell'eolico, i coefficienti  $p_c$  e  $p_s$  associati alla Danimarca sono i più elevati tra i quattro Paesi considerati, stato che il modello di Bass valuta come innovatore per eccellenza. Invece, la nazione con i coefficienti  $q_c$  e  $q_s$  maggiori è il Portogallo, considerato dal modello di Bass al ventesimo posto tra gli imitatori. Per l'idroelettrico, alla Croazia è associato il valore di  $p_c$  maggiore, alla Slovenia  $q_c$  e alla Serbia quelli di  $p_s$  e  $q_s$ . Infine, per il solare, la Germania ha i più alti valori dei coefficienti  $p_c$  e  $p_s$ , Paese che nel modello

di Bass è stato valutato secondo tra gli innovatori; il valore di  $q_c$  maggiore appartiene al Belgio, mentre quello di  $q_s$  al Regno Unito.

