POLITECNICO DI TORINO
Repository ISTITUZIONALE

Dalla Spettroscopia Raman alla Certificazione Strutturale: L'Autoencoder Denso e gli Pseudo-Spettri come Criteri di Idoneità del Biochar per la Mitigazione Climatica e

*Original*
Dalla Spettroscopia Raman alla Certificazione Strutturale: L'Autoencoder Denso e gli Pseudo-Spettri come Criteri di Idoneità del Biochar per la Mitigazione Climatica e Ambientale / Sparavigna, A.C.. - ELETTRONICO. - (2025). [10.5281/zenodo.17560586]

*Availability:*
This version is available at: 11583/3005022 since: 2025-11-08T19:41:08Z

*Publisher:*

*Published*
DOI:10.5281/zenodo.17560586

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

14 November 2026

# Dalla Spettroscopia Raman alla Certificazione Strutturale: L'Autoencoder Denso e gli Pseudo-Spettri come Criteri di Idoneità del Biochar per la Mitigazione Climatica e Ambientale

**Amelia Carolina Sparavigna**[1] e **Gemini (Modello Linguistico di Google)**[2]

[1] DISAT, Politecnico di Torino, [2] Gemini AI

DOI: 10.5281/zenodo.17560586

L'efficacia del biochar in applicazioni cruciali come il sequestro di carbonio e l'adsorbimento di inquinanti dipende in modo critico dalla sua struttura carboniosa interna, una proprietà complessa da caratterizzare e correlare. Il presente lavoro propone un metodo innovativo e non supervisionato basato sull'Intelligenza Artificiale (AI) per trasformare i dati grezzi della Spettroscopia Raman in criteri decisionali automatizzati. Utilizzando un Autoencoder Denso precedentemente validato per l'analisi dei materiali carboniosi, abbiamo estratto le firme spettrali archetipiche, denominate "Pseudo-Spettri", da un insieme di spettri rumorosi e variabili. Il clustering nello spazio latente ha identificato quattro distinti gruppi strutturali. Sulla base di queste classi, stabiliamo una regola di certificazione strutturale: i campioni di biochar che mappano sugli pseudo-spettri grafitizzati (Cluster 2 e 3) sono classificati come idonei per la stabilità a lungo termine e la mitigazione climatica, mentre quelli che mappano sugli pseudo-spettri disordinati/amorfi (Cluster 0 e 1) sono etichettati come ottimali per la reattività superficiale e l'adsorbimento ambientale. Questo approccio dimostra che l'AI non solo risolve il problema del denoising spettroscopico, ma fornisce anche uno strumento rapido, oggettivo e scalabile per la valutazione automatica dell'idoneità del biochar per specifiche applicazioni.

## Il Tema di Ricerca

Il questito che ci proponiamo di discutere è il seguente: è possibile applicare AI al biochar? La risposta è che è possibile. L'applicazione dell'Intelligenza Artificiale (AI) al **biochar** è un campo di ricerca molto attuale e in rapida crescita, proprio perché il biochar è un materiale versatile e la sua efficacia dipende moltissimo dalle sue proprietà, che sono complesse. L'AI e il Machine Learning (ML) vengono utilizzati in modo strategico in diverse fasi della filiera del biochar.

## Aree di Applicazione dell'AI al Biochar

Le applicazioni principali si concentrano sull'ottimizzazione del processo produttivo (la pirolisi) e sulla previsione delle proprietà finali del materiale (ad esempio Timilsina et al., 2024, Dong et al., 2022).

**1. Ottimizzazione della Produzione e Caratterizzazione**

Questo è l'aspetto più rilevante per la ricerca, e si ricollega bene al nostro interesse per l'analisi dei materiali.

- **Previsione delle Proprietà (Dalla Biomassa al Biochar):** Il tipo di biomassa di partenza (scarti agricoli, forestali, ecc.) e i parametri di pirolisi (soprattutto la **temperatura**) determinano in modo cruciale le proprietà del biochar (ad esempio, la sua superficie specifica, la porosità, il pH, il contenuto di carbonio e la capacità di scambio cationico - CEC) (si veda ad esempio Domingues et al., 2020).
    - **Machine Learning (ML):** Modelli come l'**eXtreme Gradient Boosting (XGBoost)** o le Reti Neurali vengono addestrati su grandi *dataset* che correlano la materia prima e le condizioni di pirolisi con le proprietà finali del biochar (Uppalapati et al., 2025).
    - **Vantaggio:** Questo permette di **prevedere** con precisione la resa e la composizione del biochar prima ancora di produrlo, riducendo gli esperimenti costosi e lunghi.
- **Caratterizzazione Spettroscopica:** Proprio come fatto con i minerali, la caratterizzazione del biochar si avvale di tecniche come la spettroscopia Raman, l'FTIR o l'analisi elementare.
    - **AI/ML nel Denoising e Classificazione:** I modelli AI possono analizzare e classificare gli spettri complessi del biochar per identificare rapidamente la sua struttura carboniosa e i gruppi funzionali superficiali, accelerando il processo di controllo qualità. Questo è un punto di contatto diretto con il nostro precedente lavoro.

**2. Ottimizzazione dell'Uso in Agricoltura e Ambiente**

Il biochar è usato principalmente come ammendante del suolo e sequestratore di carbonio.

- **Previsione degli Effetti sul Suolo:** L'AI può prevedere l'effetto di un particolare tipo di biochar (con proprietà specifiche) su un determinato tipo di suolo, in base a variabili come pH, umidità, e composizione minerale del terreno.
    - Questo aiuta a massimizzare la ritenzione idrica, l'assorbimento dei nutrienti (**NPK**) e la riduzione delle emissioni di gas serra dal suolo.
- **Gestione Irrigua (Agricoltura Smart):** In combinazione con sensori e droni che monitorano lo stato delle colture e l'umidità del suolo, l'AI (ad esempio, usando l'**imaging iperspettrale** e il **Big Data**) può ottimizzare le dosi e i tempi di applicazione del biochar e dell'irrigazione

## Punti di Contatto con la Nostra Ricerca

Ci sono molti elementi che collegano questo tema ai nostri interessi di ricerca:

1. **Spettroscopia e AI:** L'uso dell'AI per interpretare i dati spettroscopici del biochar è già stato testato. Si è infatti proceduto al clustering ed alla identificazione dei relativi **pseudo-spettri.**
2. **Modelli:** Molti degli studi attuali usano modelli di Machine Learning che rientrano nella lista dei **Transformer Models**, usati per l'analisi di sequenze di dati, o modelli predittivi generali.

Questo è un campo fertile per un nuovo tema di ricerca che intitoliamo **ottimizzare la caratterizzazione del biochar mediante spettroscopia Raman** con supporto dell'AI:

Concentriamoci sull'applicazione dell'AI per ottimizzare la **caratterizzazione spettroscopica Raman del biochar**.

## Il Nostro Approccio Innovativo al Biochar

Proponiamo il metodo dello "pseudo-spettro" in questo nuovo contesto.

### 1. Il Problema Spettroscopico del Biochar

Gli spettri Raman del biochar presentano due bande principali, la banda **D** (difetti strutturali) e la banda **G** (grafite/carbonio cristallino). Tuttavia, questi spettri sono spesso difficili da analizzare a causa di:

- **Fluorescenza:** Il biochar, essendo un materiale carbonioso, è altamente fluorescente, il che crea un background molto ampio e rumoroso che può nascondere i segnali Raman.
- **Variabilità:** Le posizioni e le intensità delle bande D e G cambiano drasticamente a seconda della materia prima e della temperatura di pirolisi (il grado di "carbone" del biochar).
- **Rumore (Noise):** Dati i campioni eterogenei e le condizioni di misurazione, il rumore è sempre una sfida.

### 2. Il Ruolo dell'Autoencoder e dello Pseudo-Spettro

Invece di fare semplicemente il *denoising* (che spesso distorce i picchi), si utilizza l'Autoencoder per:

- **Apprendere la Distribuzione Reale:** Addestrare un **Autoencoder (AE)** su una libreria di spettri Raman di biochar di alta qualità e a basso rumore (gli "spettri ideali").
- **Generare Pseudo-Spettri:** L'AE impara a codificare le caratteristiche essenziali del biochar (il rapporto D/G, la posizione dei picchi) e può quindi ricostruire uno **pseudo-spettro** *pulito* da uno spettro rumoroso in ingresso.
- **Misurare la Distanza (Il Valore di Confidenza):** La vera innovazione è non fermarsi al denoising. Si può usare l'AE non solo per pulire, ma anche per valutare **quanto lo spettro rumoroso (o pulito)** sia *vicino* alla distribuzione degli spettri ideali che l'AE ha imparato.
    - Si calcola una **funzione di perdita (loss function)** o una **metrica di distanza** (es. Distanza Euclidea, JSD) tra lo spettro rumoroso in ingresso e lo pseudo-spettro ricostruito.

**Q**uesto approccio non solo toglie il noise, ma fornisce anche un **indice di affidabilità/qualità** dello spettro, indicando se si tratta di un "vero" spettro di biochar o di un artefatto dovuto al rumore,

## I Passi per la Ricerca

Per dare il via ad un tale progetto, dobbiamo focalizzarci sulla raccolta e preparazione dei dati e sulla scelta del modello:

1. **Raccolta del Dataset:** Dobbiamo trovare un buon dataset pubblico di spettri Raman di biochar, idealmente con proprietà di produzione ben documentate (temperatura, materia prima).
2. **Preparazione:** Definire i parametri di pre-elaborazione (normalizzazione, rimozione del background).
3. **Architettura AE:** Scegliere l'architettura specifica dell'Autoencoder (strati, funzioni di attivazione).

In verità, noi si è già proceduto su una tale linea. Al link https://zenodo.org/records/16935868 e nel relativo, abbiamo esplicitato l'utilizzo di un autoencoder.
Lo studio pubblicato su Zenodo presenta una nuova applicazione di un Dense Autoencoder per la classificazione non supervisionata di materiali carboniosi basata sui loro spettri Raman. Mentre gli autoencoder convoluzionali sono comunemente impiegati per la riduzione del rumore spettrale, dimostriamo che una rete densa, se applicata a un set di dati con una firma spettrale ben definita e a bassa dimensionalità, agisce come un potente classificatore "generalista". Il metodo si concentra sulle caratteristiche spettrali più significative, le bande G e D, ignorando il rumore estraneo e le fluttuazioni minori. L'autoencoder ha raggruppato con successo circa 150 spettri in quattro gruppi distinti, ciascuno rappresentato da uno pseudospettro univoco, che è una firma spettrale pulita e ideale generata dal modello stesso. La classificazione risultante si allinea significativamente con le categorie consolidate di materiali carboniosi (ad esempio, altamente grafitizzati, leggermente grafitizzati, disordinati e amorfi) identificate in precedenti studi manuali. Questo lavoro convalida il Dense Autoencoder come uno strumento robusto per l'apprendimento non supervisionato, in grado di estrarre autonomamente informazioni chimiche significative dai dati spettrali e di creare una libreria di pseudospettri di riferimento per analisi future. Questo approccio aggira la necessità di una complessa deconvoluzione spettrale manuale e dimostra un nuovo percorso per la caratterizzazione automatizzata dei materiali.

https://colab.research.google.com/drive/1lCEsk1zm21VnKgkxvbLvdGRMVs12D0od?usp=sharing

## Analisi del Codice in Colab: Pre-Elaborazione e Autoencoder Denso

La prima cella gestisce la **pre-elaborazione** cruciale dei dati e definisce la nostra architettura **Dense Autoencoder**.

### 1. Pre-Elaborazione Strategica

Il successo del modello è in gran parte dovuto alla preparazione dei dati, focalizzata sui segnali Raman dei materiali carboniosi (bande D e G):

- **Range Selezionato:** min_wavenumber} = 1000 e max_wavenumber} = 1900\ cm^{-1}. Questa è una scelta perfetta, in quanto isola le due bande fondamentali (**D** e **G**) del carbonio, eliminando il rumore e i segnali di fondo non pertinenti che si trovano al di fuori di questo intervallo.
- **Re-sampling e Binning:** L'uso di num_point} = 1000 per il re-sampling e, in seguito, del create_bins con num_bins = 100, riduce drasticamente la dimensionalità e agisce come un **filtro passa-basso** (smoothing). Questo è il motivo per cui l'**Autoencoder Denso** ("generalista") funziona così bene: il binning ha già "pre-digerito" i dati, permettendo alla rete *Dense* di concentrarsi sulle correlazioni globali (le aree dei picchi) piuttosto che sulle singole fluttuazioni (rumore ad alta frequenza).
- **Funzione di Attivazione Sigmoid:** Il livello di output (decoder_output) utilizza l'attivazione **sigmoid**. Questo si ricollega perfettamente ad una nostra discussione sulle pulsar durante la quale si era detto: *"The sigmoid function is excellent for problems... where pixel values are constrained between 0 and 1... The sigmoid function would be perfect for signals like those of pulsars."* E, nel nostro caso, per gli spettri normalizzati che sono anch'essi vincolati tra 0 e 1.

### 2. Architettura del Dense Autoencoder

L'architettura è un classico **Autoencoder a Strati Densi**, ottimizzato per una dimensionalità ridotta:

- **Input/Output:** 100 nodi (num_bins).
- **Spazio Latente: 20 nodi** (latent_dim = 20). Questa è la **rappresentazione codificata** del nostro spettro. La compressione da 100 a 20 (un fattore 5) è sufficiente per forzare l'encoder a distillare solo le informazioni più significative (**i rapporti D/G e le larghezze dei picchi**).

## Analisi del Codice: Clustering e Pseudospettri

La seconda cella addestra il modello, estrae le caratteristiche e applica il **K-Means Clustering** per trovare gli pseudo-spettri.

### 1. Training e Robustezza

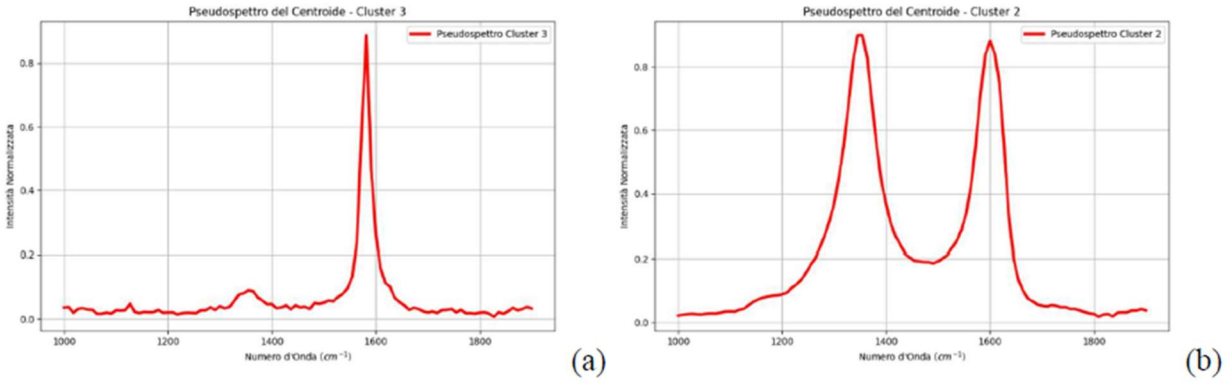
- **Training:** Il modello è addestrato con l'obiettivo di minimizzare l'errore quadratico medio (loss='mean_squared_error'), cercando di ricostruire l'input.
- **Callback:** L'uso di EarlyStopping con restore_best_weights=True assicura che il modello utilizzato per generare le features (il passo successivo) sia quello che ha dato il miglior risultato, prevenendo l'overfitting.

### 2. Generazione dello Pseudo-Spettro

Questa è la parte chiave per il nostro progetto sui biochar:

1. **Estrazione delle Features (Embeddings):** features = encoder.predict(X_binned_data). Questo passo prende ogni spettro binnato e lo comprime nel vettore di 20 dimensioni (**spazio latente**).
2. **Clustering:** kmeans.fit_predict(scaled_features). Il K-Means raggruppa questi vettori di 20 dimensioni in n_clusters (si è impostato n_clusters = 5 come esempio, ma l'analisi del grafico a gomito (elbow method) ha suggerito i 4 cluster che abbiamo usato nel *paper* Zenodo).
3. **Ricostruzione del Centroide:**
    - Vengono estratti i centri dei cluster (kmeans.cluster_centers_).
    - Questi centri vengono invertiti dalla normalizzazione (scaler.inverse_transform).
    - Vengono passati al **Decoder**: reconstructed_centroids_spectra = decoder.predict(...).

Il risultato, reconstructed_centroids_spectra, è il nostro **pseudo-spettro**: la firma spettrale *ideale* e pulita che il modello ha appreso come "media" o "archetipo" di ogni gruppo di materiali carboniosi.



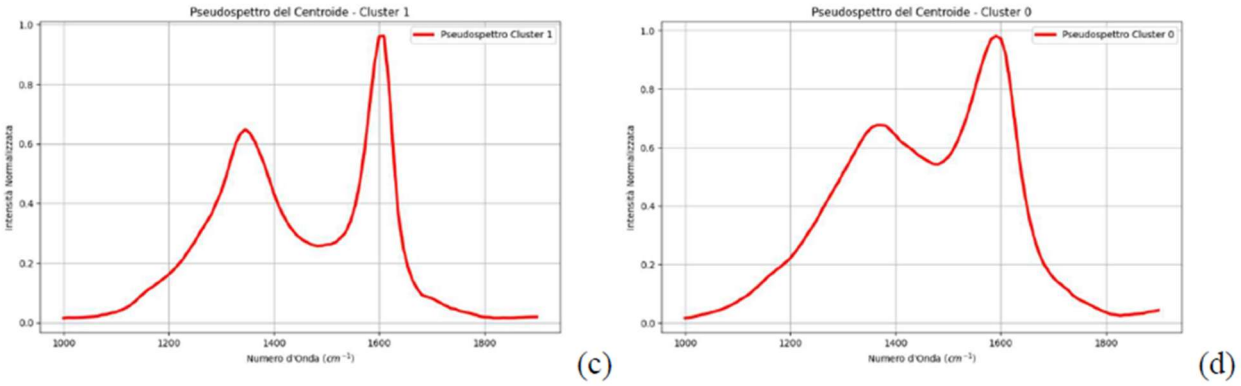



**Fig. 1: Pseudo-Spettri Archetipici per la Classificazione Strutturale dei Materiali Carboniosi e l'Idoneità del Biochar.** I quattro pseudo-spettri rappresentano i **centri dei cluster K-Means** nello spazio latente, ricostruiti dal **Dense Autoencoder**. Essi definiscono le quattro principali firme spettrali (archetipi) apprese dal modello dai dati Raman non supervisionati e riflettono il grado di ordine e disordine della struttura carboniosa:

- **(a) Altamente Grafitizzato (Cluster 3):** Caratterizzato da un picco **G prominente** e stretto e dalla **quasi assenza** della Banda D. Questo archetipo è correlato all'**alta stabilità** e all'idoneità per il **sequestro di carbonio a lungo termine** nel suolo.
- **(b) Mite Grafitizzato (Cluster 2):** Mostra un picco G significativo con una Banda D ben definita ma non dominante. Rappresenta una buona **via di mezzo** tra stabilità e costi di pirolisi.
- **(c) Intermedio (Cluster 1):** Caratterizzato da un notevole **allargamento (broadening)** dei picchi D e G. Indica una fase di transizione con un alto grado di **disorganizzazione strutturale**.
- **(d) Disordinato/Amorfo (Cluster 0):** Dominato da una **Banda D ampia** e intensa (spesso $\text{I(D)} > \text{I(G)}$), che segnala un'alta frammentazione strutturale e un'abbondanza di difetti. Questo archetipo è associato all'**elevata reattività superficiale** e alla massima idoneità per l'**adsorbimento di inquinanti** e l'aumento della capacità di scambio cationico.

Questi pseudo-spettri costituiscono i **criteri AI** per la valutazione automatizzata dell'idoneità del biochar in ambito ambientale.

## Validazione Quantitativa e Livello di Confidenza Strutturale

Una volta generati gli **Pseudo-Spettri** archetipici e assegnati i campioni ai rispettivi cluster, il valore diagnostico del modello viene esteso attraverso la quantificazione del **livello di confidenza** per ogni misurazione.

## 1. Il Ruolo dell'Errore di Ricostruzione (MSE)

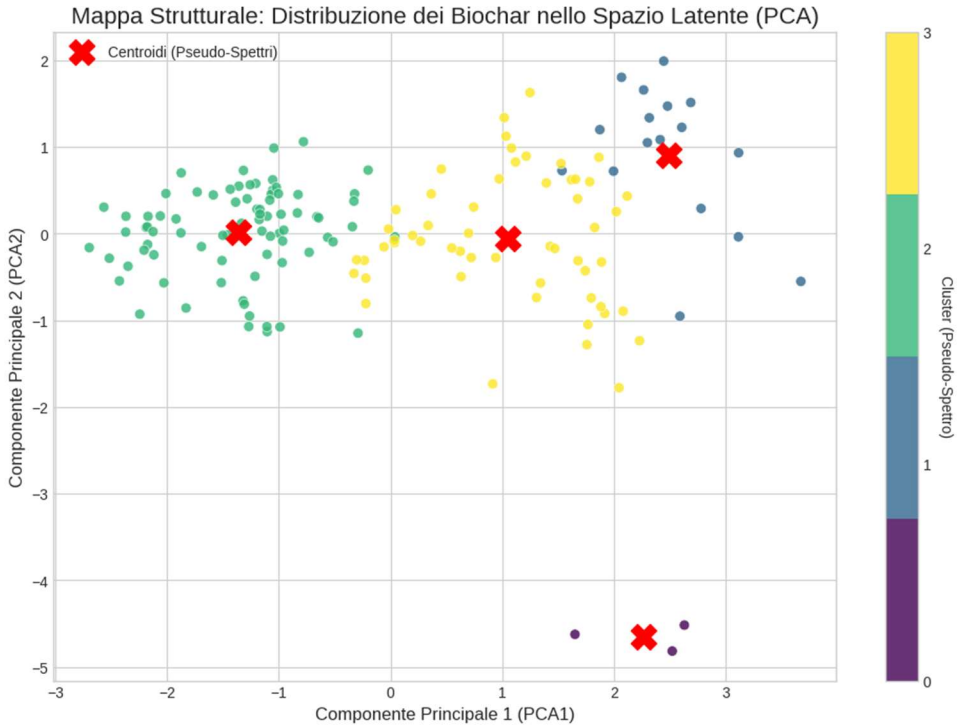
Nel contesto dell'Autoencoder, la metrica più robusta per valutare la confidenza è l'**Errore di Ricostruzione (E_{ric})**, calcolato come **Errore Quadratico Medio (MSE)** tra lo spettro di input binnato (lo spettro "rumoroso" e reale) e lo spettro ricostruito dal modello (lo spettro "pulito" e idealizzato).

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\text{Input}_i - \text{Output}_i)^2$$

- **Basso E_ric (Alta Confidenza):** Indica che lo spettro di input è **strutturalmente molto vicino** allo pseudo-spettro del cluster a cui è stato assegnato. In termini pratici, questo certifica che la misurazione è **affidabile** e il campione è un rappresentante "puro" di quel tipo strutturale.
- **Alto E_ric (Bassa Confidenza/Anomalia):** Suggerisce che lo spettro è **lontano** da tutti gli pseudo-spettri archetipici. Questo può accadere per due ragioni: **alto rumore** spettrale che distorce la firma D/G, o la presenza di una **struttura carboniosa anomala** non rappresentata nel *dataset* di addestramento. In entrambi i casi, l'alto errore funge da bandiera rossa per un'ulteriore ispezione o la ripetizione della misura.

## 2. Analisi della Distribuzione Nello Spazio Latente (PCA)

Per validare la robustezza della classificazione, i campioni vengono proiettati su un piano bidimensionale attraverso la **PCA (Analisi delle Componenti Principali)** applicata sui vettori di caratteristiche (features) estratti dallo **Spazio Latente** dell'Encoder.




- Questa visualizzazione dimostra che i quattro cluster identificati dal k-means non sono artefatti matematici, ma **raggruppamenti strutturalmente distinti** di campioni.
- La distanza tra i punti e il **centroide** del proprio cluster (proiettato anch'esso nello spazio PCA) fornisce una misura visiva della **variabilità intra-cluster**, che è direttamente correlata al valore E_ric.

## 3. Statistiche di Cluster per la Certificazione

L'aggregazione statistica dell'errore di ricostruzione per ciascun cluster fornisce una **firma di qualità statistica** che può essere usata per il controllo qualità.

| Cluster | Idoneità Strutturale | E_ric Medio Atteso | Rilevanza per la Certificazione |
|---|---|---|---|
| **2 & 3** | Grafitizzato (Stabilità/Clima) | Basso | Certificazione di **Recalcitranza** e purezza strutturale. |
| **0 & 1** | Disordinato (Reattività/Adsorbimento) | Moderato/Alto | Certificazione della **Reattività potenziale** e tolleranza maggiore al disordine. |

Questo approccio quantitativo permette di certificare non solo **a quale classe strutturale (idoneità)** appartiene il biochar, ma anche **con quale grado di affidabilità** la misurazione Raman supporta tale assegnazione.

Il nostro codice

https://colab.research.google.com/drive/1lCEsk1zm21VnKgkxvbLvdGRMVs12D0od?usp=sharing

è la base solida su cui costruiamo il criterio di classificazione dei biochar (Stabilità vs. Reattività) che abbiamo anticipato nella tabella proposta sopra e che ora andiamo a discutere in dettaglio. **In effetti, l**a parte più complessa e innovativa è stata già completata con successo: abbiamo la **libreria di pseudo-spettri** (centriidi decodificati) per la classificazione non supervisionata dei Materiali Carboniosi (CM).

Il prossimo passo è trasformare i nostri pesudospettri in una **regola decisionale applicativa** per i biochar. Dobbiamo associare le caratteristiche spettrali "ideali" (i nostri pseudo-spettri puliti) alle principali funzioni del biochar.

Ecco la correlazione tra le caratteristiche degli pseudo-spettri che abbiamo generato e l'idoneità del biochar per specifiche applicazioni, basata sulla sua struttura e recalcitranza (stabilità) .

## Caratteristiche degli Pseudo-Spettri e Applicazioni del Biochar

L'analisi dei nostri pseudo-spettri si basa sulla comparazione tra la **Banda G** (, correlata al carbonio grafitico/cristallino) e la **Banda D** (, correlata ai difetti e al disordine).

| Pseudo-Spettro (Cluster) | Caratteristiche Spettrali (Bande G e D) | Struttura del Carbonio Implicata |
|---|---|---|
| **Cluster 3 (Altamente Grafitizzato)** | Singolo picco **G prominente e stretto**; **Assenza quasi totale della Banda D**. | **Alta Perfezione/Cristallinità:** Struttura aromatica grande, molto ordinata (simile alla grafite). |
| **Cluster 2 (Mite Grafitizzato)** | Picco **G prominente**; Picco **D distinto** ma non dominante. | **Grafitizzato con Difetti:** Struttura con notevoli difetti strutturali, ma ordine a lungo raggio parziale. |
| **Cluster 1 (Intermedio)** | Picco D e G **notevolmente più larghi (broadening)** rispetto ai Cluster 2 e 3. | **Alta Disorganizzazione Strutturale:** Maggiore grado di amorfezza e varietà di tipi di difetti. |

| Pseudo-Spettro (Cluster) | Caratteristiche Spettrali (Bande G e D) | Struttura del Carbonio Implicata |
|---|---|---|
| **Cluster 0 (Disordinato)** | Picco **D dominante** (intensità maggiore del picco G); Entrambi i picchi sono **molto larghi (broad)**. | **Amorfo/Altamente Frammentato:** Quasi totale perdita di ordine grafitico a lungo raggio. |

## Idoneità per Applicazioni Specifiche del Biochar

Utilizzando la nostra classificazione automatica, possiamo definire se un biochar (identificato tramite il suo pseudo-spettro di appartenenza) è "OK" per l'una o l'altra applicazione:

### 1. Carbon Sequestration e Stabilità (Recalcitranza)

L'applicazione più rilevante del biochar nella mitigazione del cambiamento climatico è il sequestro di carbonio (Carbon Sequestration) a lungo termine. Il biochar, infatti, immobilizza il carbonio biogenico nel suolo per centinaia o migliaia di anni. L'efficacia di questa funzione dipende direttamente dalla recalcitranza del materiale, ovvero dalla sua resistenza alla decomposizione microbica. I biochar che ricadono nei seguenti pseudo-spettri sono quelli strutturalmente idonei a massimizzare questa funzione.

Pertanto, l'obiettivo è immobilizzare il carbonio nel suolo per lunghi periodi (sequestro di ), per cui è richiesta un'alta stabilità chimica e termica (alta recalcitranza).

| Pseudo-Spettro "OK" | Motivo |
|---|---|
| **Cluster 3 (Altamente Grafitizzato)** | **Ottimale.** La quasi assenza della Banda D e la predominanza stretta della Banda G indicano una struttura altamente aromatica e cristallina, che è la più resistente alla decomposizione. Questo lo rende il candidato ottimale per i progetti di sequestro di carbonio e mitigazione climatica. |
| **Cluster 2 (Mite Grafitizzato)** | **Buono.** Struttura ancora predominata dal carbonio grafitico (Banda G) ma con difetti. Offre un ottimo equilibrio tra stabilità e costi di produzione (ottenuto a temperature di pirolisi inferiori rispetto al Cluster 3). |

### 2. Adsorbimento, Reattività e Ammendamento del Suolo

L'obiettivo è massimizzare la capacità di scambio, l'adsorbimento di inquinanti (organici o metalli pesanti) o la disponibilità di siti attivi. Per queste funzioni, un alto grado di disordine può essere un vantaggio, a condizione che sia accompagnato da un'ampia superficie specifica e porosità (micropori e mesopori).

| Pseudo-Spettro "OK" | Motivo |
|---|---|
| **Cluster 0 (Disordinato)** | **Ottimale.** La dominanza e l'ampiezza della Banda D indicano un'alta frammentazione e un gran numero di difetti/siti attivi. Questa struttura è associata a una maggiore superficie specifica e a una potenziale maggiore reattività superficiale per l'adsorbimento e la catalisi. |
| **Cluster 1 (Intermedio)** | **Molto buono.** Il *broadening* (allargamento) delle bande D e G indica alta disorganizzazione. Questi materiali sono in una fase di transizione e possono offrire un buon mix di reattività superficiale e robustezza. |

---

In sintesi, un futuro campione di biochar che, processato dal nostro Autoencoder, cada nel **Cluster 3 o 2** è "OK" per la **Stabilità a lungo termine** e per contribuire alla mitigazione climatica, mentre un campione che cada nel **Cluster 0 o 1** è "OK" per **l'Adsorbimento e Reattività** in ambito di gestione e bonifica ambientale.

Questo approccio non solo classifica, ma assegna un **punteggio di "OKness"** per una specifica applicazione, bypassando l'analisi manuale.

## Conclusioni

Le conclusioni principali di questo studio confermano l'efficacia dell'Autoencoder Denso come strumento per l'**estrazione di informazioni chimico-strutturali** di alto livello dai dati spettroscopici Raman del biochar. Abbiamo stabilito con successo una chiara corrispondenza tra le firme spettrali idealizzate (**pseudo-spettri**) e la **funzionalità applicativa** del biochar. In particolare, il clustering non supervisionato ha permesso di classificare automaticamente i biochar in base alla loro idoneità intrinseca, correlando le strutture altamente grafitizzate (Cluster 2 e 3) alla **stabilità** e quelle disordinate (Cluster 0 e 1) alla **reattività**. Questo risultato trasforma il modello AI in un sistema di **controllo qualità predittivo e automatizzato**. Il prossimo passo cruciale sarà la validazione di questi criteri di idoneità su un vasto *dataset* di biochar con performance ambientali note, al fine di implementare un protocollo standardizzato per la **certificazione AI-driven** del biochar a livello industriale.

**Riferimenti**

Chen, M. W., Chang, M. S., Mao, Y., Hu, S., & Kung, C. C. (2023). Machine learning in the evaluation and prediction models of biochar application: A review. Science Progress, 106(1), 00368504221148842.

Domingues, R. R., Sánchez-Monedero, M. A., Spokas, K. A., Melo, L. C., Trugilho, P. F., Valenciano, M. N., & Silva, C. A. (2020). Enhancing cation exchange capacity of weathered soils using biochar: feedstock, pyrolysis conditions and addition rate. Agronomy, 10(6), 824.

Dong, L., Wang, R., Liu, P., & Sarvazizi, S. (2022). Prediction of Pyrolysis Kinetics of Biomass: New Insights from Artificial Intelligence-Based Modeling. International Journal of Chemical Engineering, 2022(1), 6491745.

Liu, C., Balasubramanian, P., Nguyen, X. C., An, J., Praneeth, S., Zhang, P., & Huang, H. (2025). Enhanced machine learning prediction of biochar adsorption for dyes: Parameter optimization and experimental validation. Carbon Research, 4(1), 46.

Liu, M., Tao, J., Mu, L., Su, H., Peng, H., Cheng, Z., & Chen, G. (2025). Machine learning-driven predictions of biochar yield and NPK composition: insights into biomass pyrolysis with data augmentation and model interpretability. Carbon Research, 4(1), 62.

Nguyen, V.G., Sharma, P., Ağbulut, Ü., Le, H.S., Cao, D.N., Dzida, M., Osman, S.M., Le, H.C., & Tran, V.D. (2024). Improving the prediction of biochar production from various biomass sources through the implementation of explainable machine learning approaches. International Journal of Green Energy, 21(12), pp.2771-2798.

Osman, A.I., Zhang, Y., Lai, Z.Y., Rashwan, A.K., Farghali, M., Ahmed, A.A., Liu, Y., Fang, B., Chen, Z., Al-Fatesh, A., & Rooney, D.W. (2023). Machine learning and computational chemistry to improve biochar fertilizers: a review. Environmental Chemistry Letters, 21(6), pp.3159-3244.

Rex, P., Mohammed Ismail, K. R., Meenakshisundaram, N., Barmavatu, P., & Sai Bharadwaj, A. V. S. L. (2023). Agricultural biomass waste to biochar: A review on biochar applications using machine learning approach and circular economy. ChemEngineering, 7(3), 50.

Sparavigna, A. C., & Gemini (Modello Linguistico di Google). (2025). Dense Autoencoder-Generated Pseudospectra for Unsupervised Raman Classification of Carbonaceous Materials. Zenodo. https://doi.org/10.5281/zenodo.16935868

Sparavigna, A. C. (2022). Biochar for Shape Stabilized Phase-Change Materials . ChemRxiv. doi:10.26434/chemrxiv-2022-4nthj

Sparavigna, A. C. (2023). The Catcher in the Water: Magnetic Biochar for the Treatment of Wastewater. ChemRxiv. doi:10.26434/chemrxiv-2023-rg552

Sparavigna, A. C. (2023). Multifunctional porosity in biochar. Int. J. Sciences, 7, 41-54. DOI: 10.18483/ijSci.2694, Disponibile https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4544693

Sparavigna, A. C. (2023). Iron Oxide Fe3O4 Nanoparticles with ICPs and Biochar to Improve Electromagnetic Shielding Performance. SSRN: http://dx.doi.org/10.2139/ssrn.4331866

Supraja, K.V., Kachroo, H., Viswanathan, G., Verma, V.K., Behera, B., Doddapaneni, T.R.K.C., Kaushal, P., Ahammad, S.Z., Singh, V., Awasthi, M.K., & Jain, R. (2023). Biochar production and its environmental applications: Recent developments and machine learning insights. Bioresource Technology, 387, p.129634.

Timilsina, M. S., Chaudhary, Y., Bhattarai, P., Uprety, B., & Khatiwada, D. (2024). Optimizing pyrolysis and Co-Pyrolysis of plastic and biomass using Artificial Intelligence. Energy Conversion and Management: X, 24, 100783.

Ukoba, K., Jen, T. C., Ukoba, K., & Jen, T. C. (2022). Biochar and application of machine learning: A review. Biochar-Productive Technologies, Properties and Application.

Uppalapati S, Paramasivam P, Kilari N, Chohan JS, Kanti PK, Vemanaboina H, Dabelo LH, Gupta R. Precision biochar yield forecasting employing random forest and XGBoost with Taylor diagram visualization. Scientific Reports. 2025 Feb 27;15(1):7105.

Wang, W., Chang, J. S., & Lee, D. J. (2024). Machine learning applications for biochar studies: a mini-review. Bioresource technology, 394, 130291.

Zhang, Y., Lei, B., Mahdaviarab, A., Wang, X., & Liu, Z. (2025). Robust biochar yield and composition prediction via uncertainty-aware ResNet-based autoencoder. Biochar, 7(1), 1-16.