Entrenamientos de desarrollo

El objetivo de esta sección es determinar qué target es el que da mejores resultados para nuestro propósito, calcular el potencial de una estructura atómica de 222 átomos a través del tiempo. Esto debido a que la arquitectura fue desarrollada para funcionar con un dataset de cientos de miles de moléculas pequeñas y en nuestro caso lo estamos aplicando a un dataset con un solo gran arreglo atómico con pasos a través del tiempo.

Para esto tomamos en cuenta tres variaciones del target, primero tomamos la energía total del arreglo tal como fue pensada la arquitectura originalmente, después la energía total del arreglo menos el offset para quedarnos solo con la energía de interacción y por último la energía total del arreglo menos el offset y dividido por el número de átomos para que el cálculo de la energía sea independiente del tamaño del arreglo.

A su vez es importante señalar que la arquitectura es sensible a las unidades del dataset. El código está optimizado para funcionar con la energía en eV, las fuerzas en eV/A y las posiciones en A.

Los entrenamientos se realizaron con 4 GPU’s Nvidia A100

A continuación se muestran los resultados de los entrenamientos.

# Target: Energía total

Este entrenamiento tiene como target la energía total del cluster.

Duración del entrenamiento: 4 horas, 43 minutos y 24 segundos

Configuración relevante:

batch\_size: 32

cutoff\_upper: 6.35

derivative: False

inference\_batch\_size: 4

max\_num\_neighbors: 140

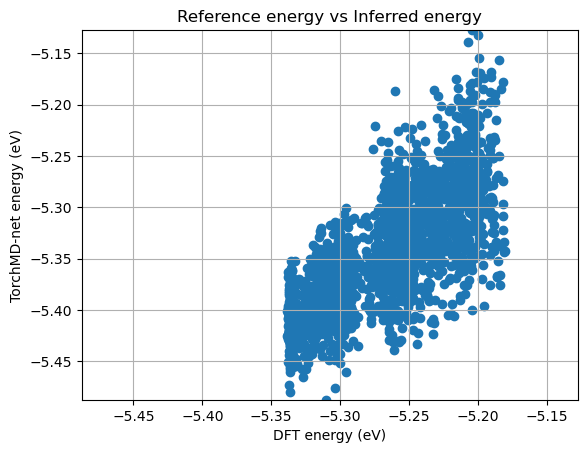
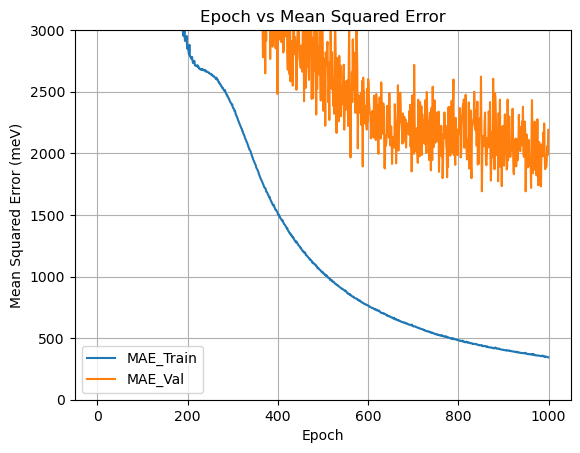
num\_epochs: 1000

seed: 42

test\_size: 0.1

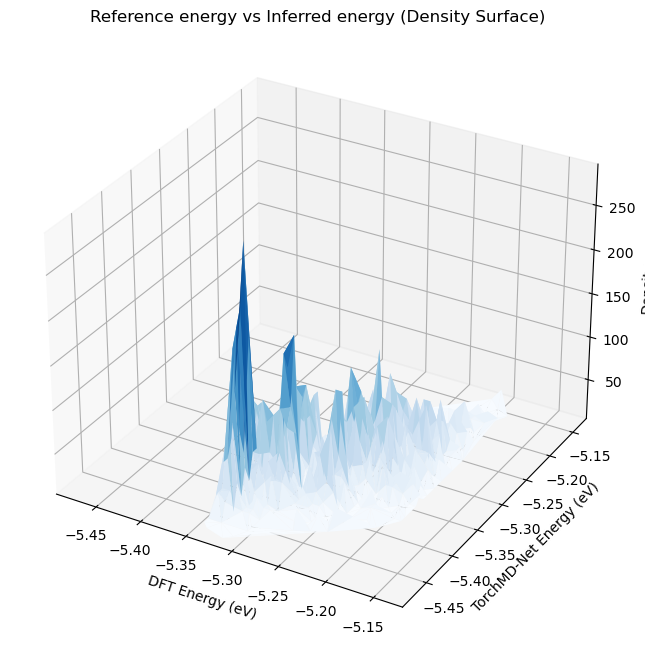
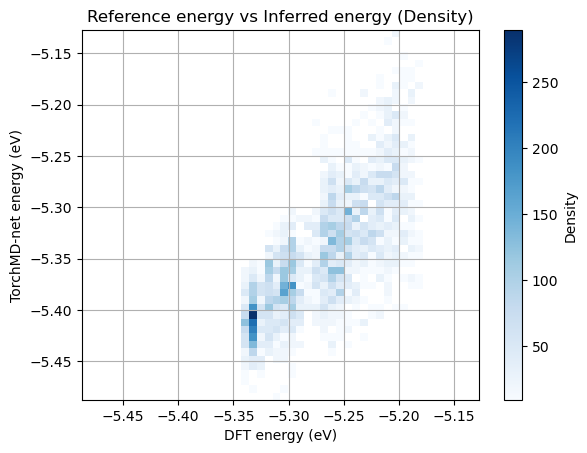
train\_size: 0.81

val\_size: 0.09



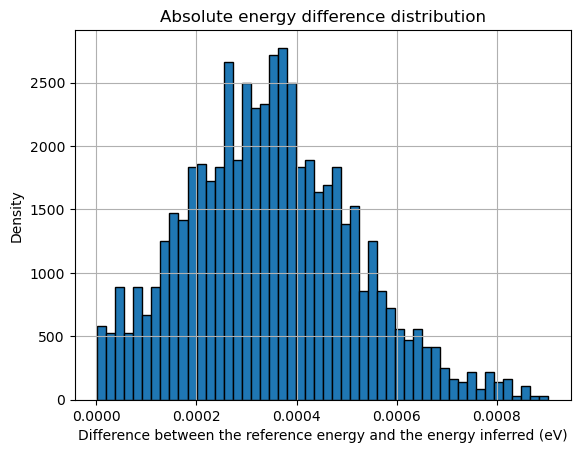
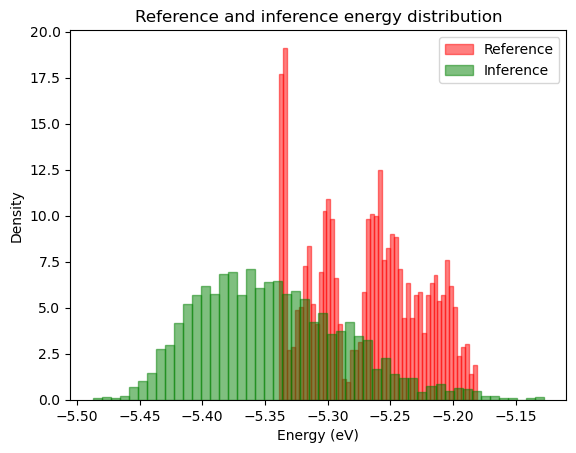
Para obtener la gráfica Epoch vs Mean Squared Error se multiplicó el costo por 1000 para pasar a meV y se dividió por el número de átomos para que esta gráfica sea comparable con las demás del mismo tipo.

Para obtener la gráfica Reference energy vs inferred energy a la predicción se le resto el offset y se dividió por el número de átomos.



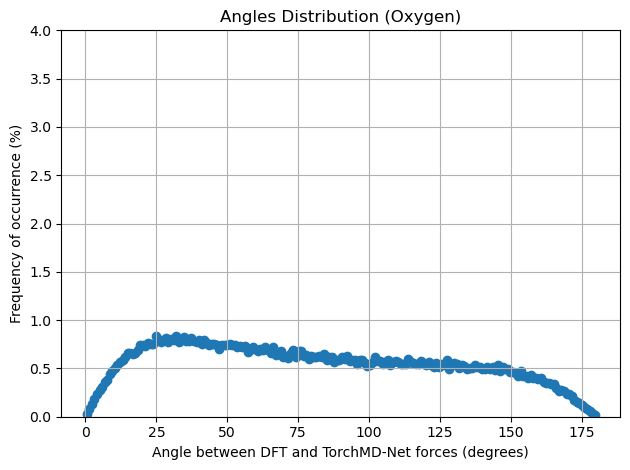
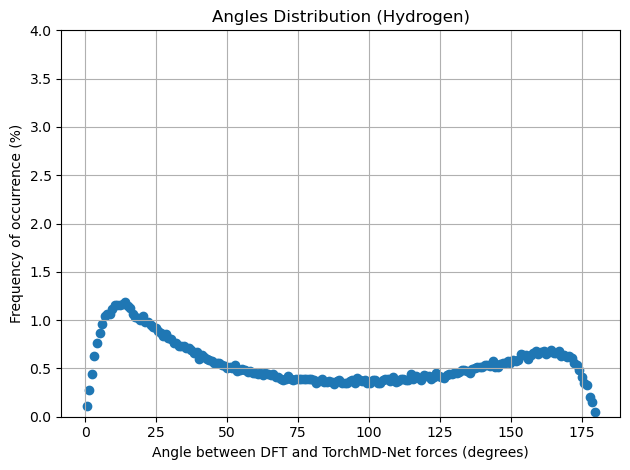
Para obtener la gráfica Reference energy vs Inferred energy (Density) se siguió el mismo procedimiento de la gráfica Reference energy vs inferred energy solamente se cambió la representación de una gráfica de dispersión a una gráfica de densidad.

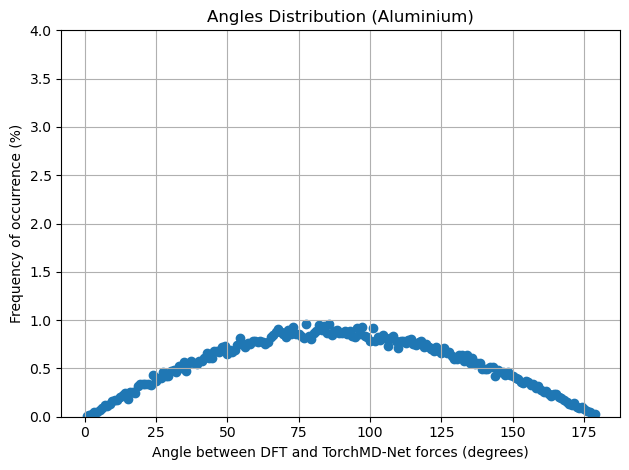
La gráfica Reference energy vs Inferred energy (Density Surface) es representación en tres dimensiones de la gráfica Reference energy vs Inferred energy (Density)



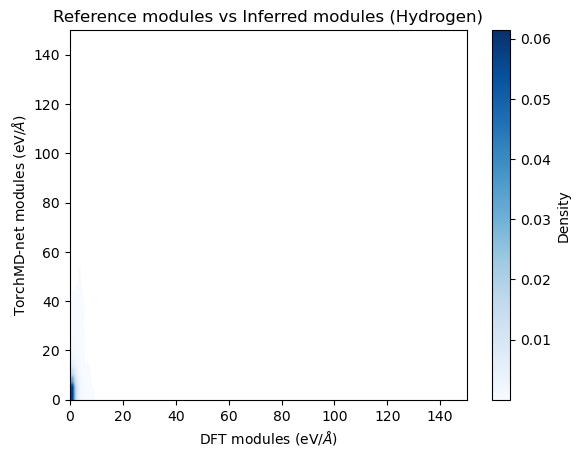
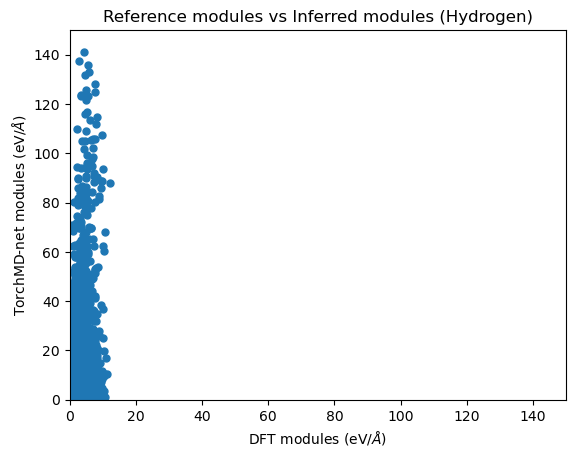
Para tener referencia de las distribuciones de las energías de referencia e inferencia se obtiene la gráfica Reference and inference energy distribution.

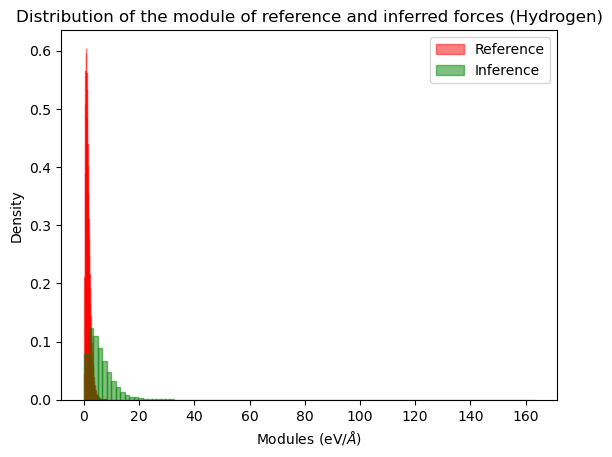
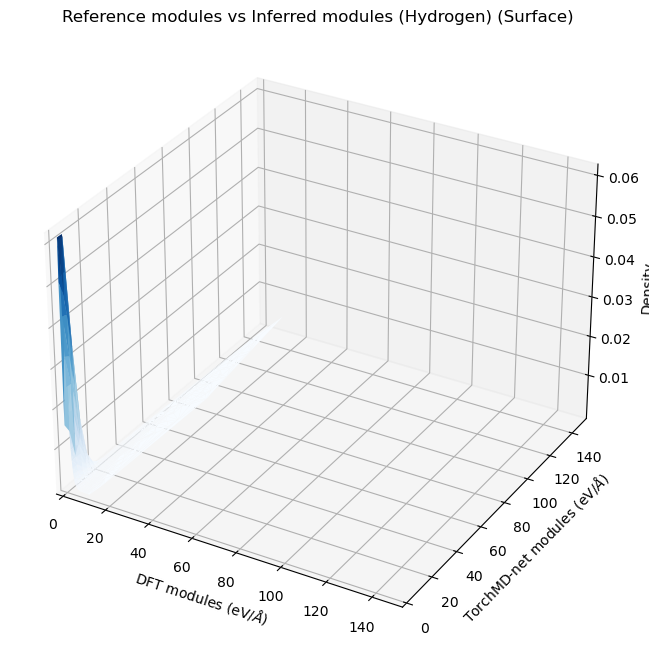
A si mismo con la distribución del valor absoluto de la diferencia entre ambas energías, se obtiene la gráfica Absolute energy difference distribution.



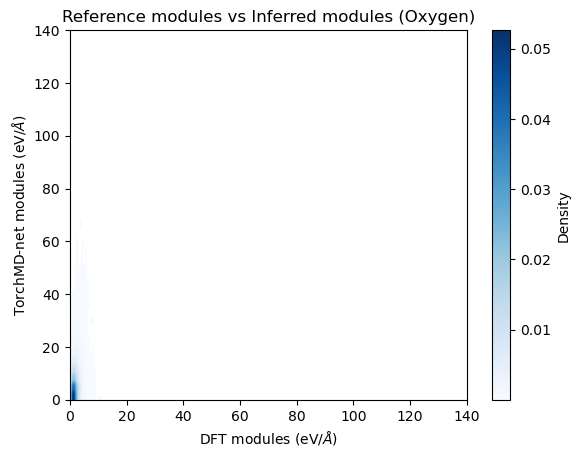
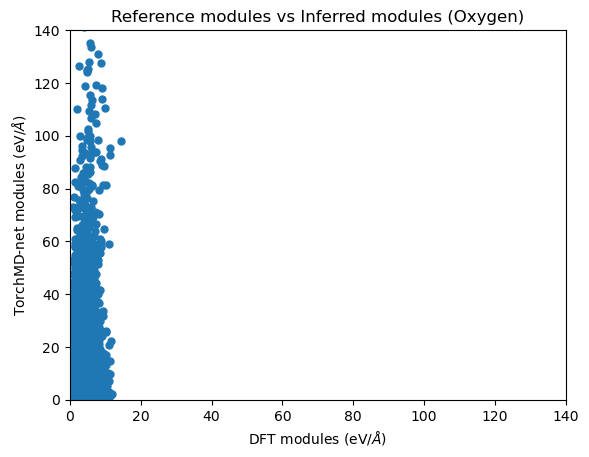


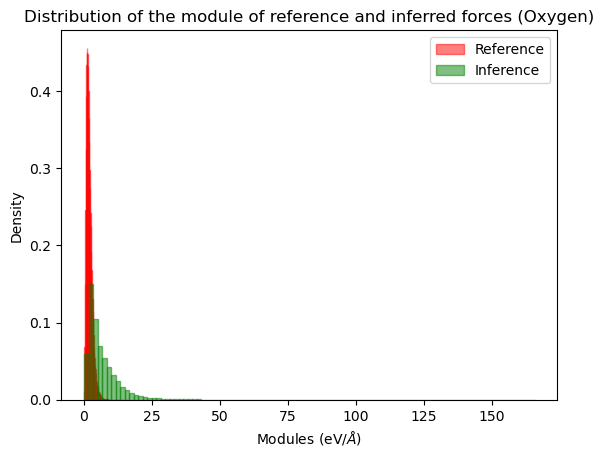
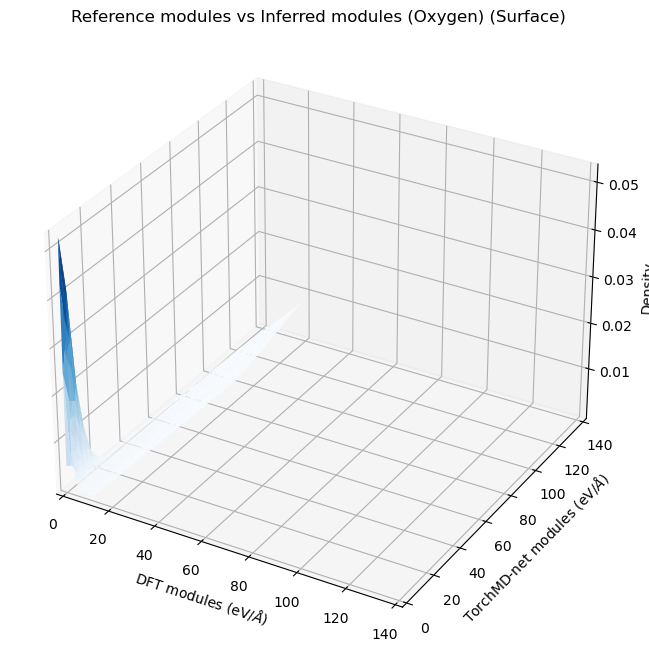
Para analizar las fuerzas nos es útil analizar la diferencia entre el ángulo de la fuerza de referencia y el ángulo de la fuerza inferida. Podemos graficar esta discrepancia para ver que tan acertada es la predicción. Es importante señalar que distintos átomos experimentan distintas fuerzas, por lo que para su correcto análisis se agrupan los átomos por su número atómico.



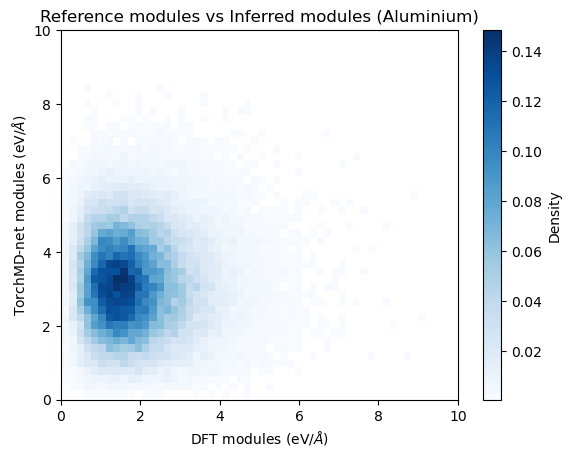
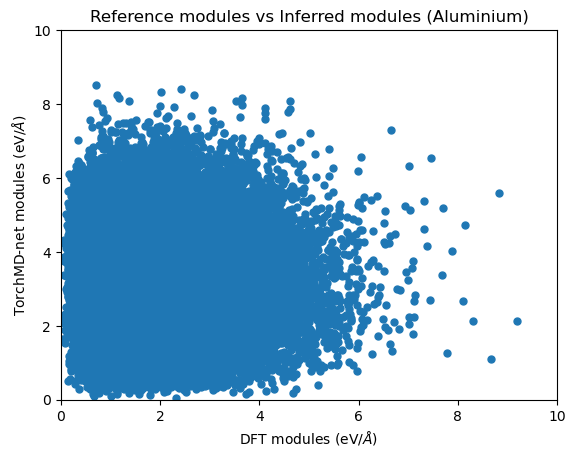


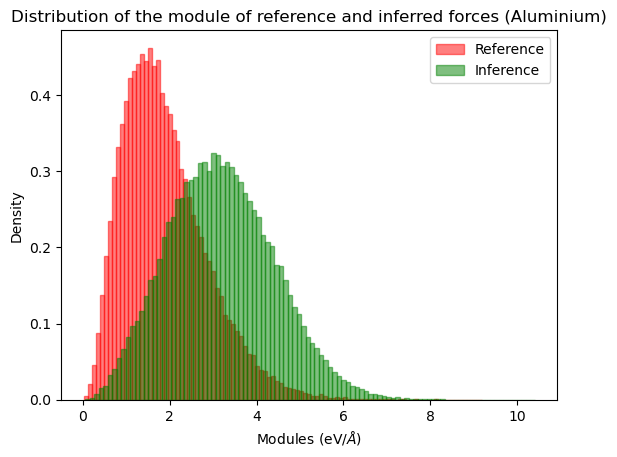
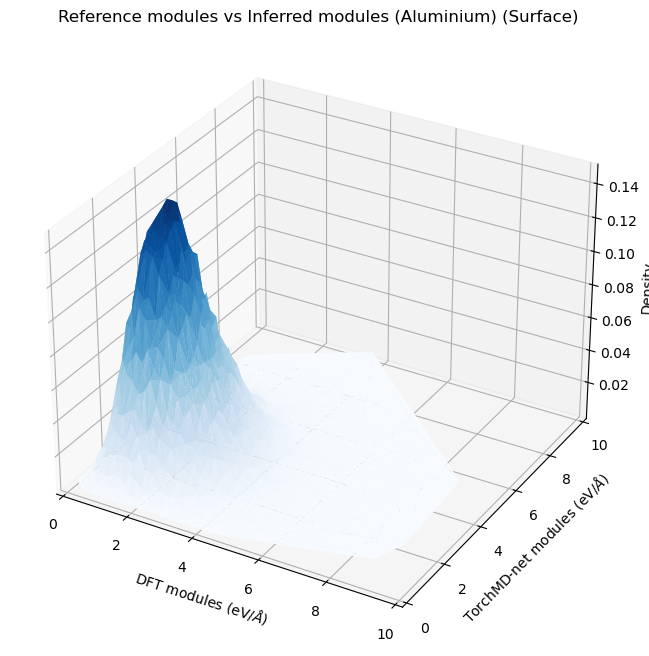
Utilizamos los mismos razonamientos del análisis de la energía para el análisis de los módulos de las fuerzas que experimentan los átomos de Hidrógeno.





Replicamos las gráficas para las fuerzas de los átomos de Oxígeno.





Y las fuerzas de los átomos de Aluminio.

# Target: Energía total menos el offset

Este entrenamiento tiene como target la energía total menos la atómica.

Configuración relevante:

batch\_size: 32

cutoff\_upper: 6.35

derivative: False

inference\_batch\_size: 4

max\_num\_neighbors: 140

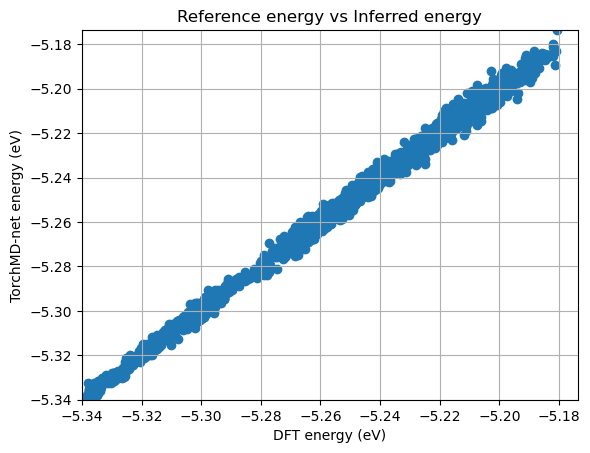
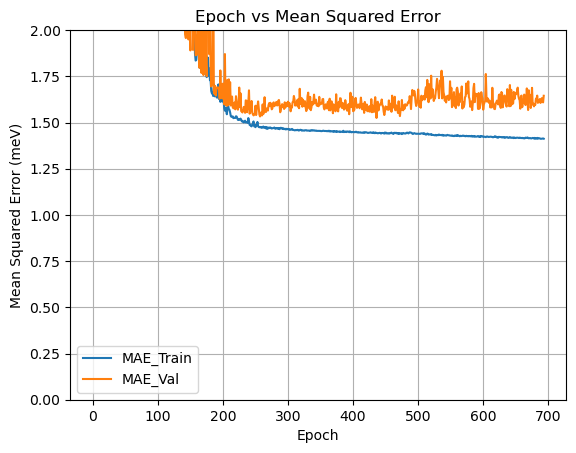
num\_epochs: 1000

seed: 42

test\_size: 0.1

train\_size: 0.81

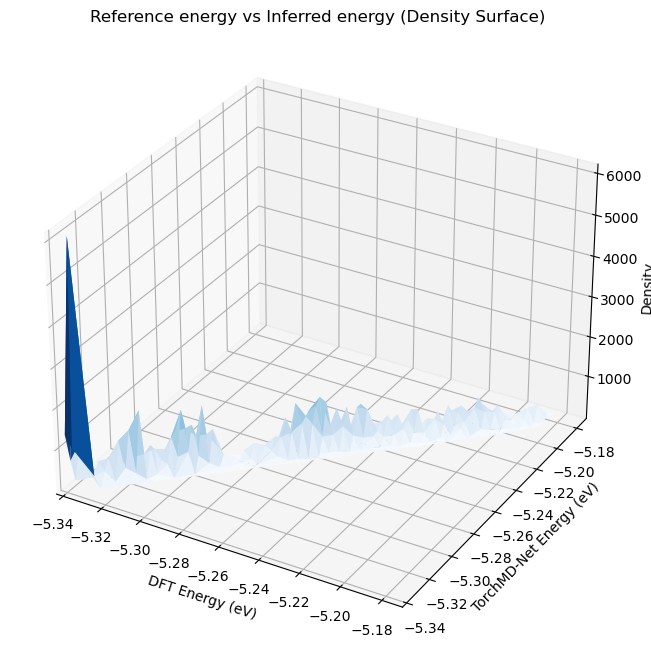
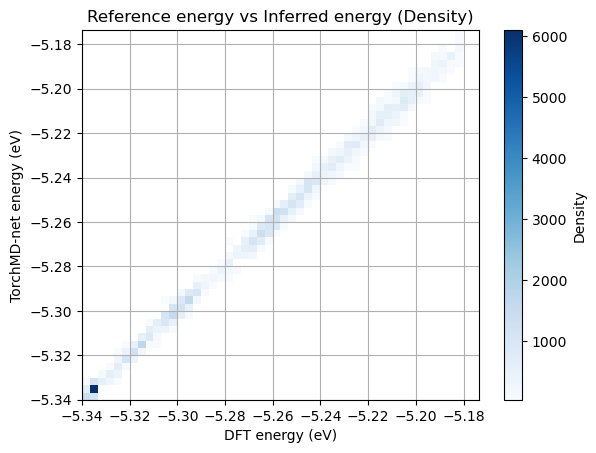
val\_size: 0.09

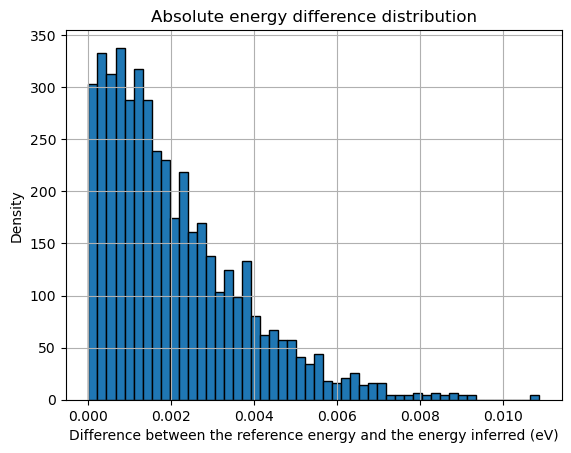
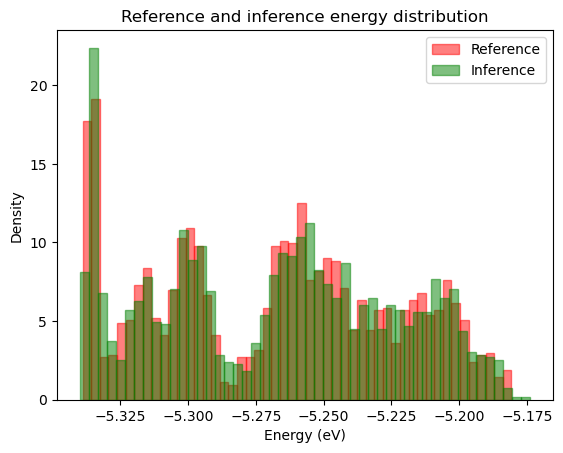


Se sigue el mismo análisis que en el entrenamiento con target Energía Total. Tomando en cuenta las siguientes precisiones:

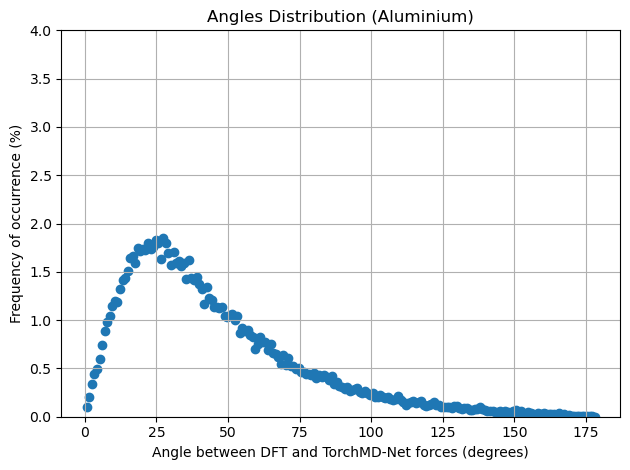
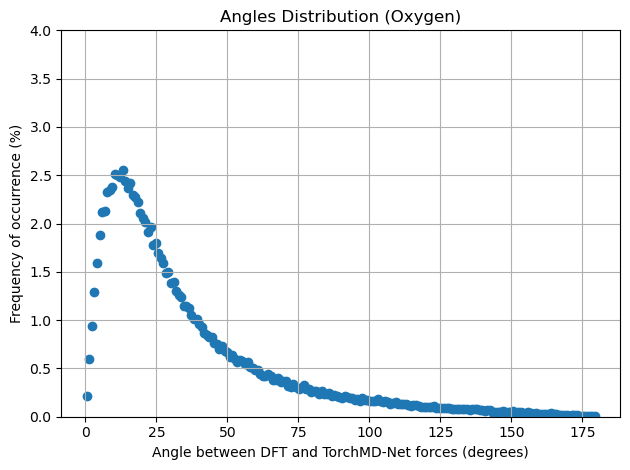
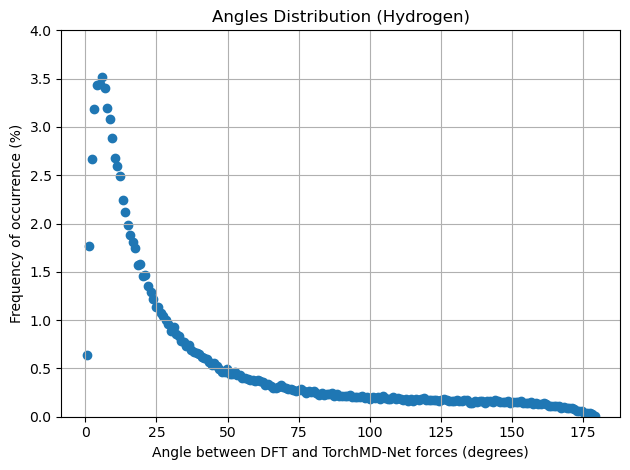
Para obtener la gráfica Epoch vs Mean Squared Error se multiplicó la pérdida por 1000 para pasar a meV y se dividió por el número de átomos.

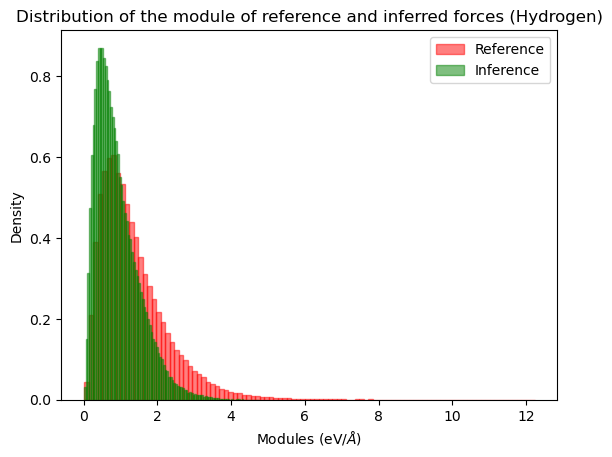
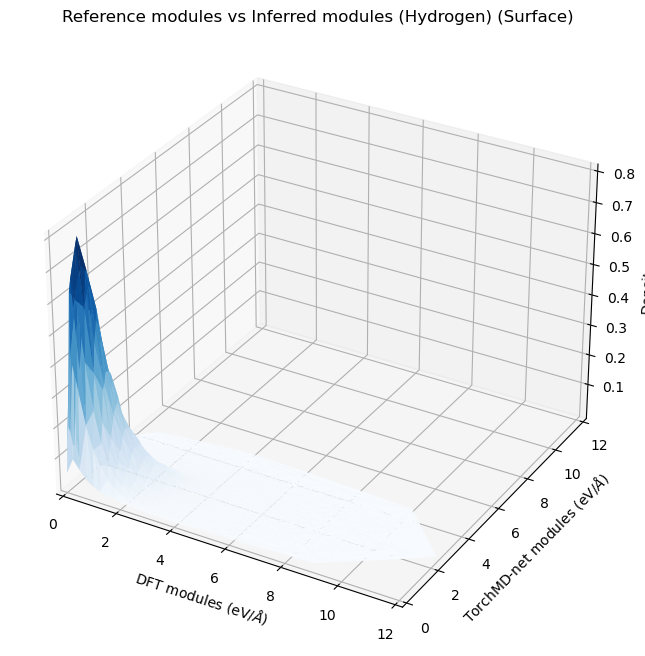
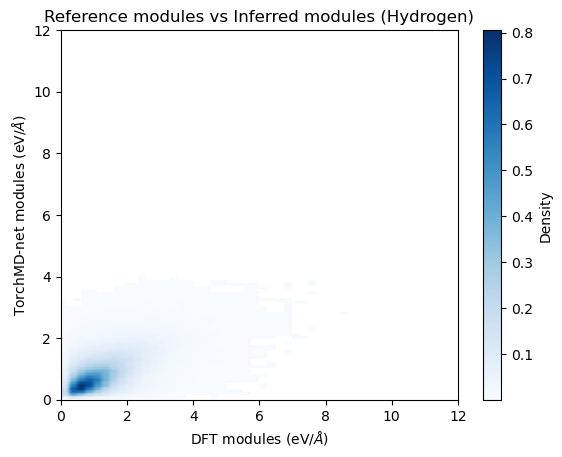
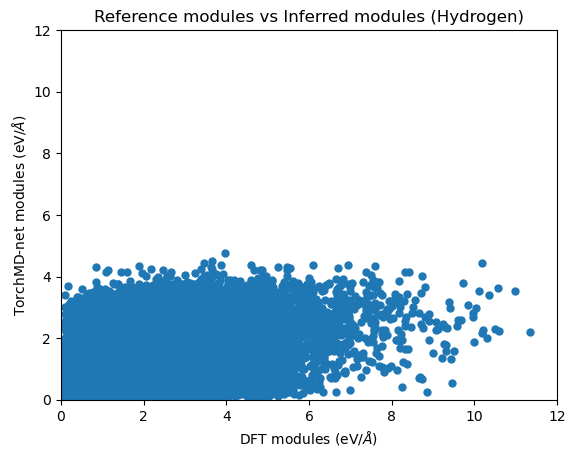
Para obtener la gráfica Reference energy vs Inferred energy de dividió la energía por el número de átomos.

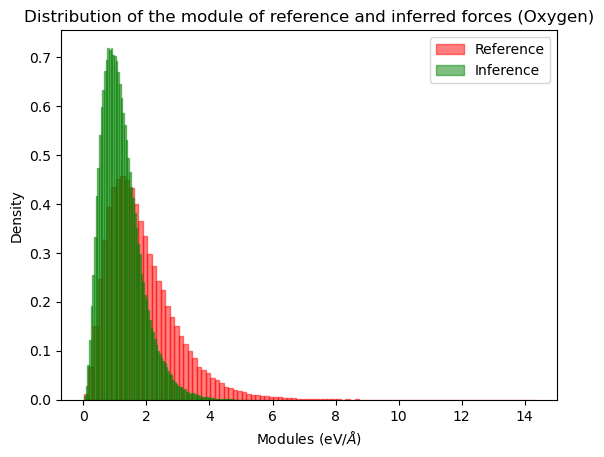
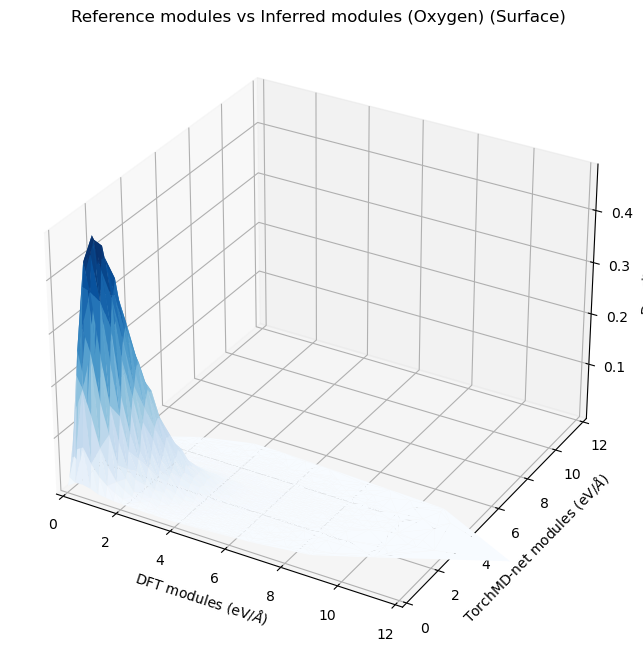
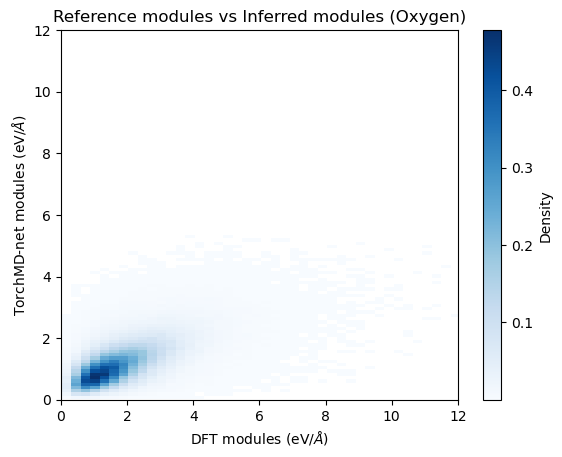
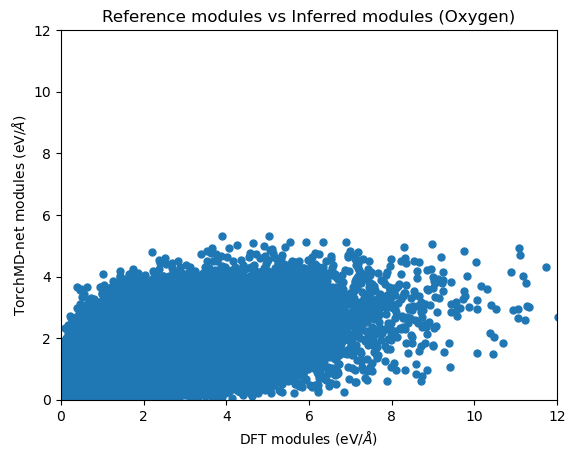


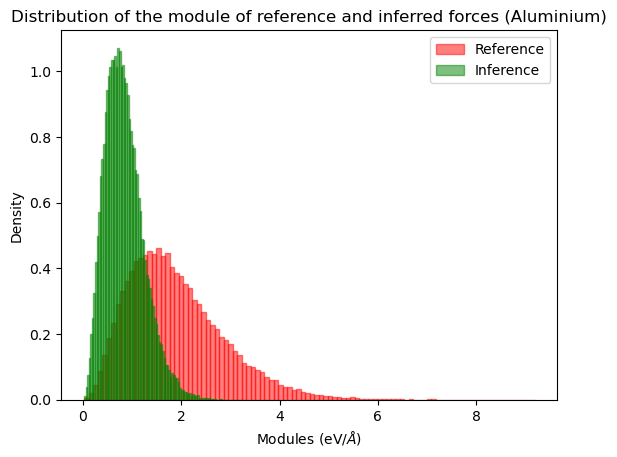
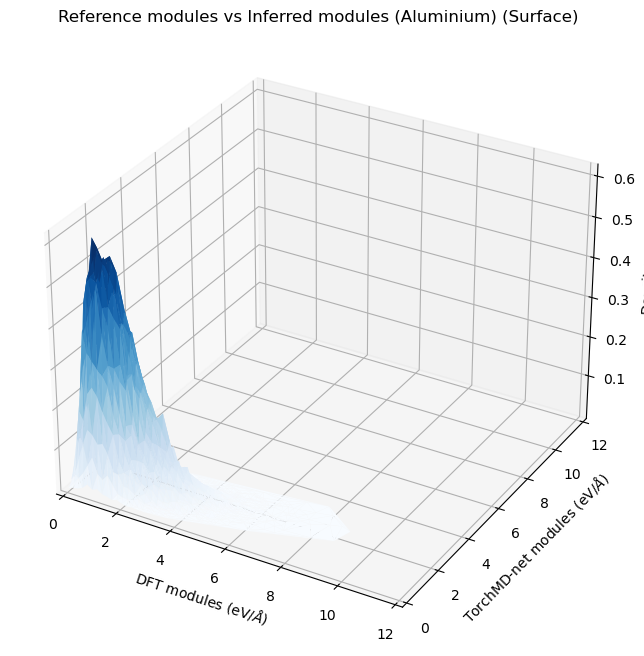
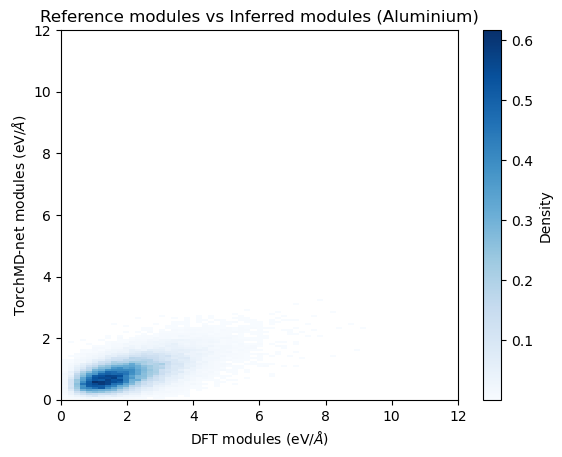
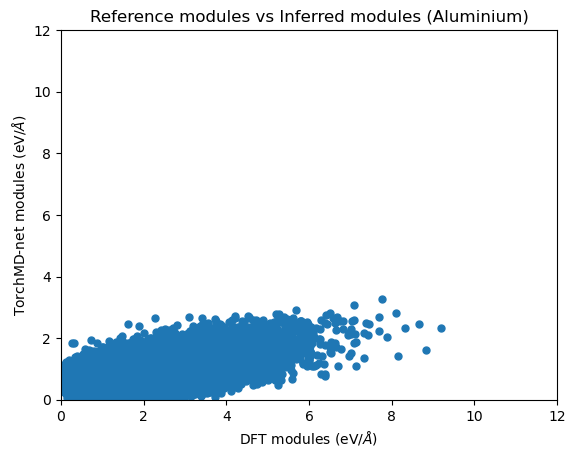


Estas gráficas toman como base la gráfica Reference energy vs Inferred energy.









# Target: Energía total menos el offset entre en número de átomos

Este entrenamiento tiene como target la energía total menos la atómica entre el número de átomos.

Duración del entrenamiento: 2 horas 42 minutos y 1 segundos

Configuración relevante:

batch\_size: 32

cutoff\_upper: 6.35

derivative: False

inference\_batch\_size: 4

max\_num\_neighbors: 140

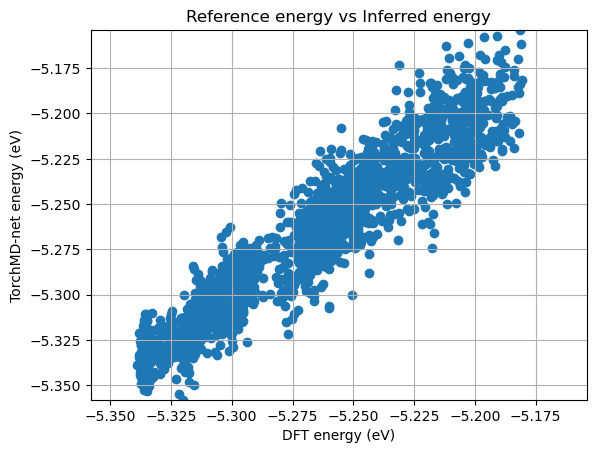
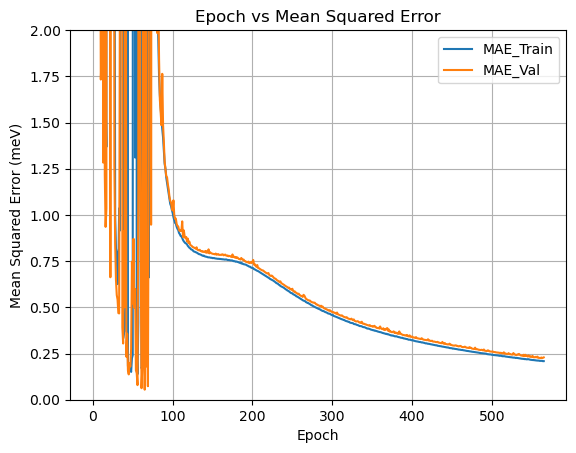
num\_epochs: 1000

seed: 42

test\_size: 0.1

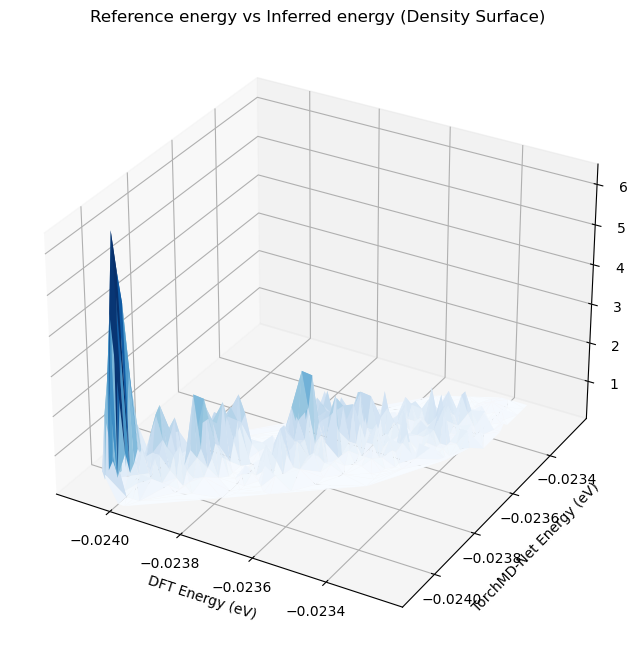
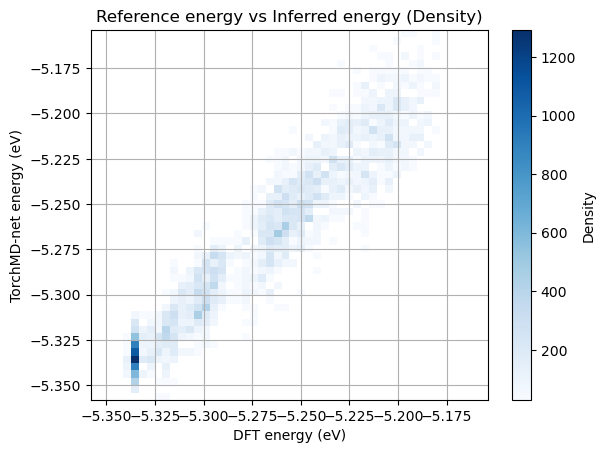
train\_size: 0.81

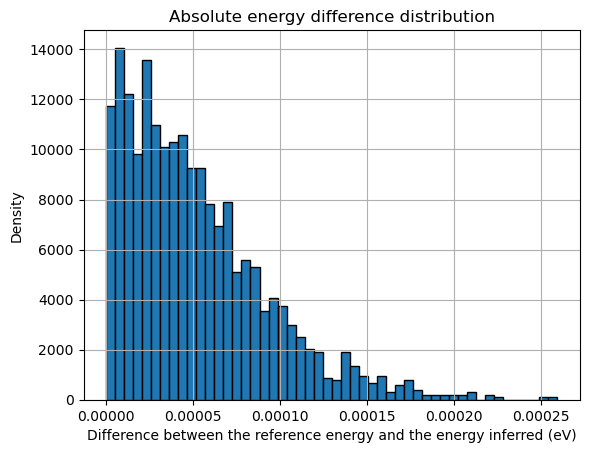
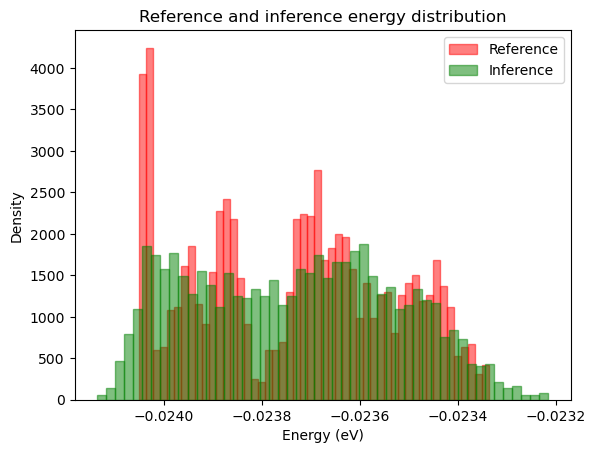
val\_size: 0.09



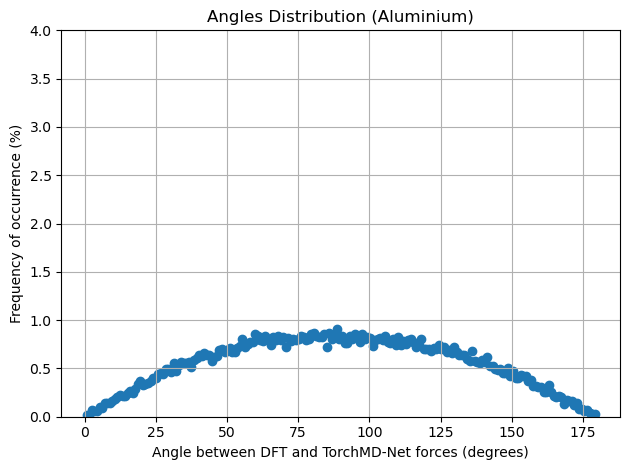
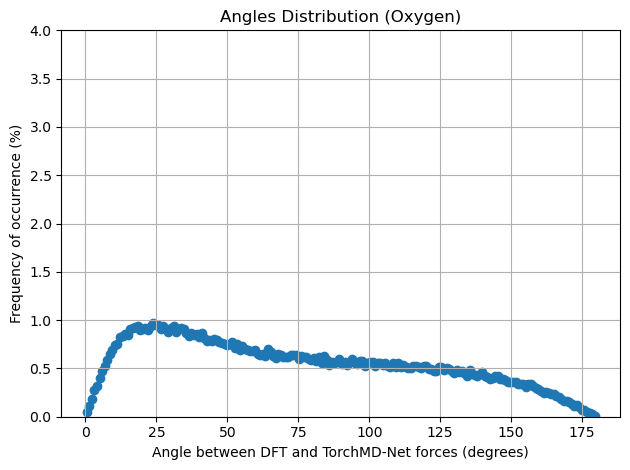
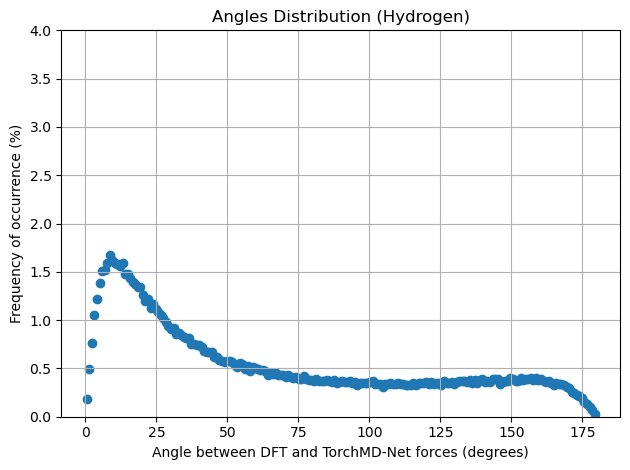
Se sigue el mismo análisis que en el entrenamiento con target Energía Total. Tomando en cuenta las siguientes precisiones:

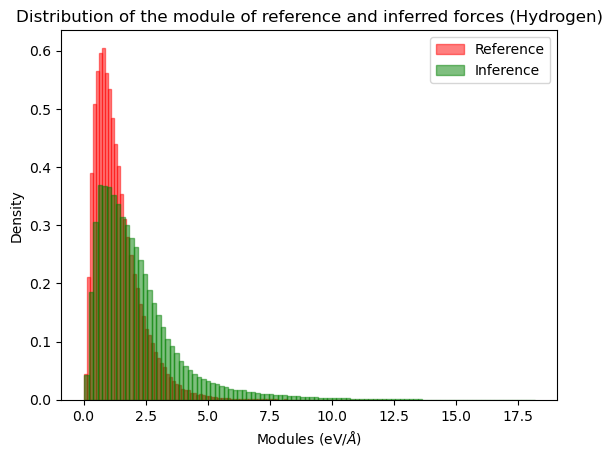
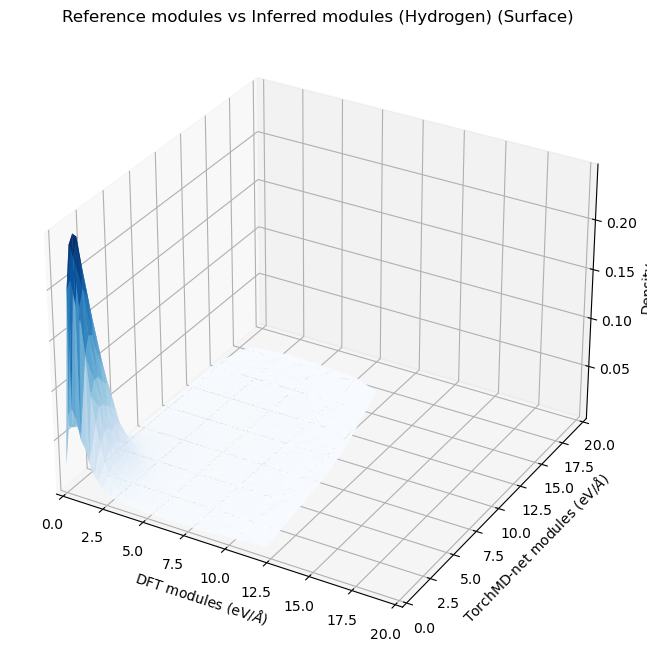
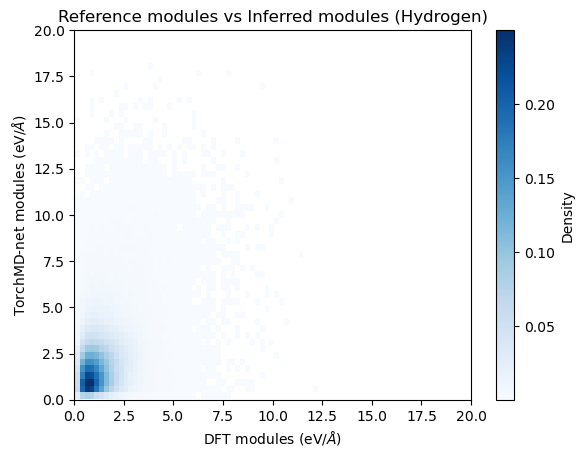
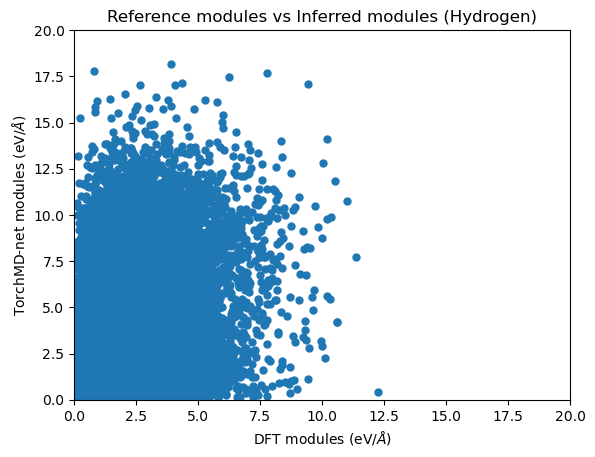
Para obtener la gráfica Epoch vs Mean Squared Error se multiplicó la pérdida por 1000 para pasar a meV

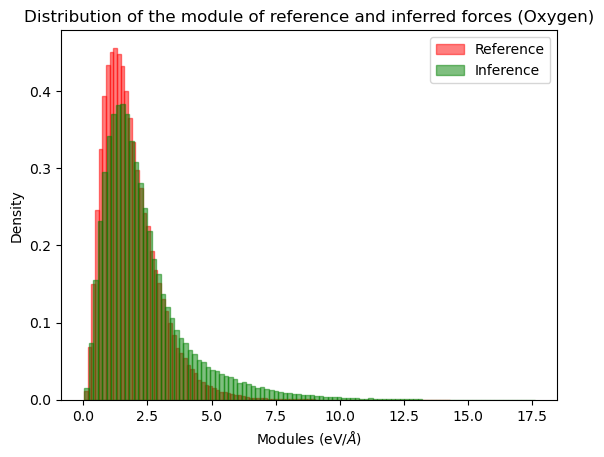
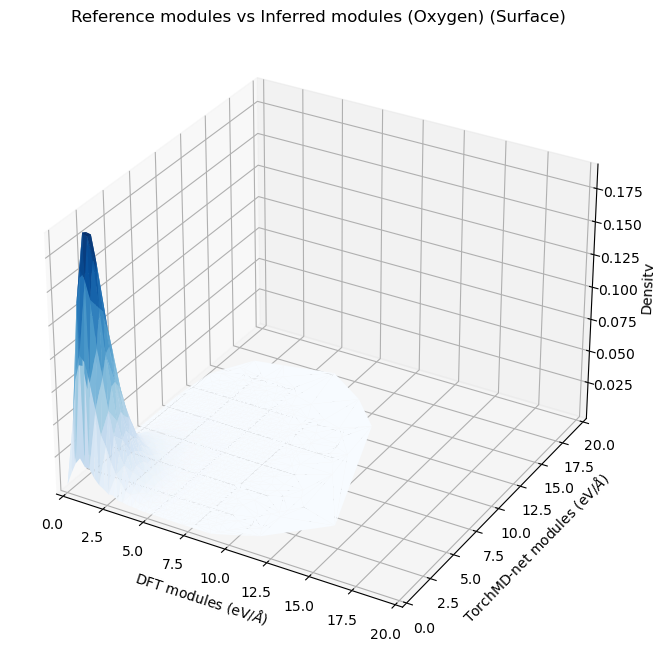
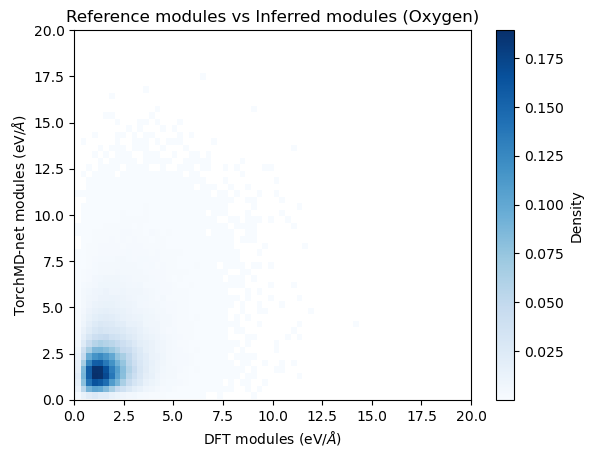
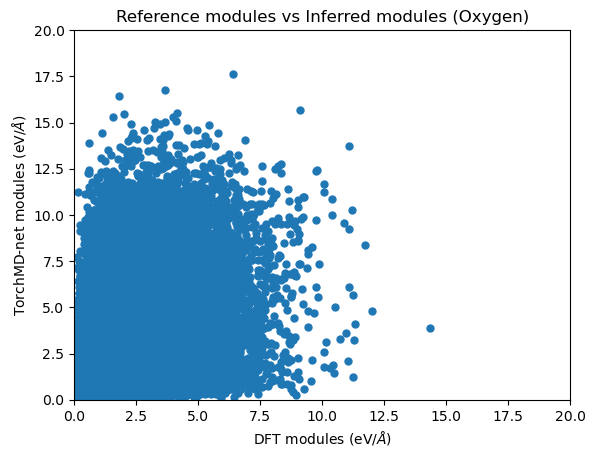


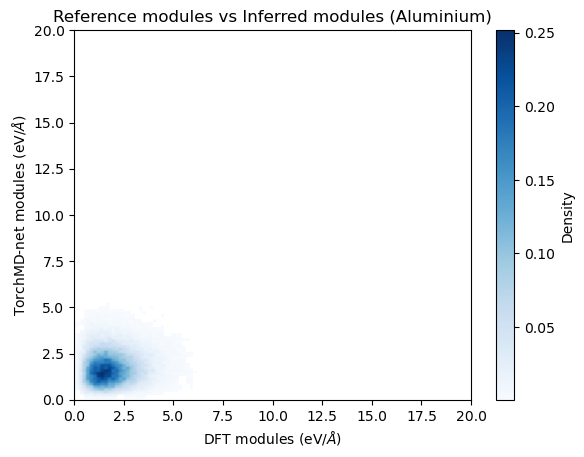
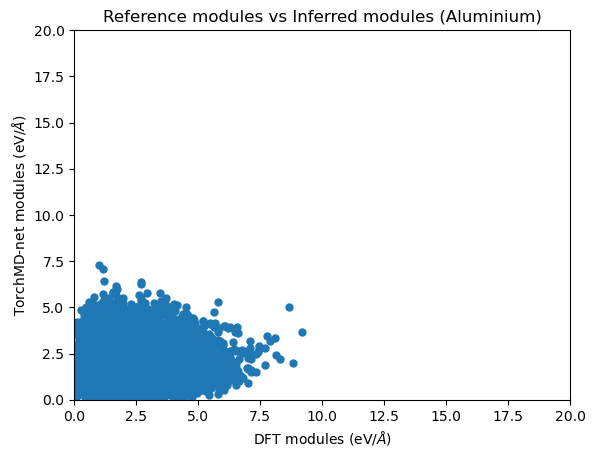


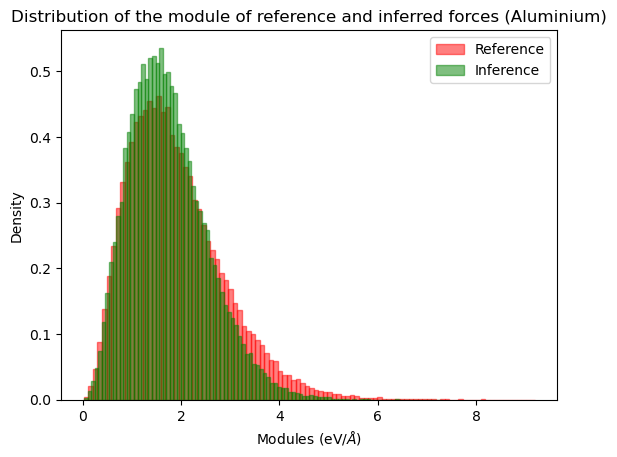
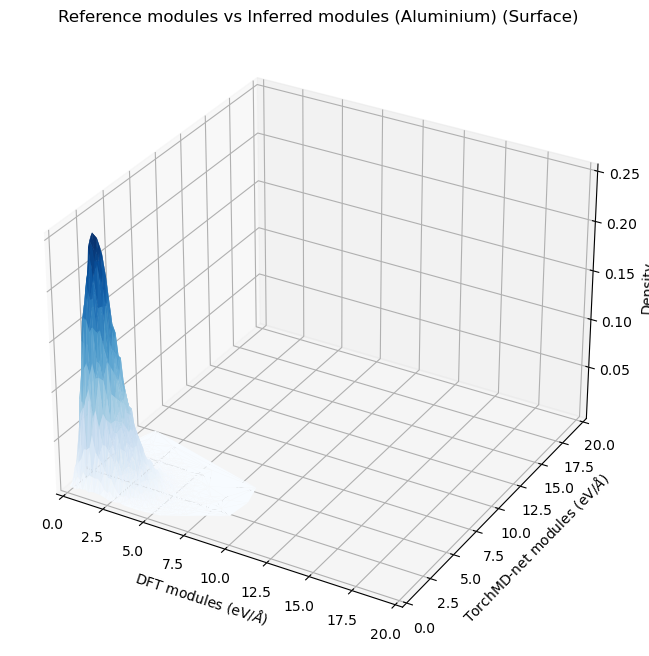
Estas gráficas toman como base la gráfica Reference energy vs Inferred energy.











# Conclusiones

Podemos observar que la arquitectura es altamente sensible al target, si bien el código está diseñado para trabajar con la energía total de la molécula el offset de arreglos grandes afecta considerablemente la precisión de las inferencias. El resultado mejora bastante si se resta el offset. Otro arreglo intuitivo es dividir por el número de átomos para que las inferencias sean independientes del tamaño del arreglo, sin embargo la arquitectura no soporta esta corrección.

El entrenamiento con mejores resultados fue el que tiene como target La Energía total menos el offset.