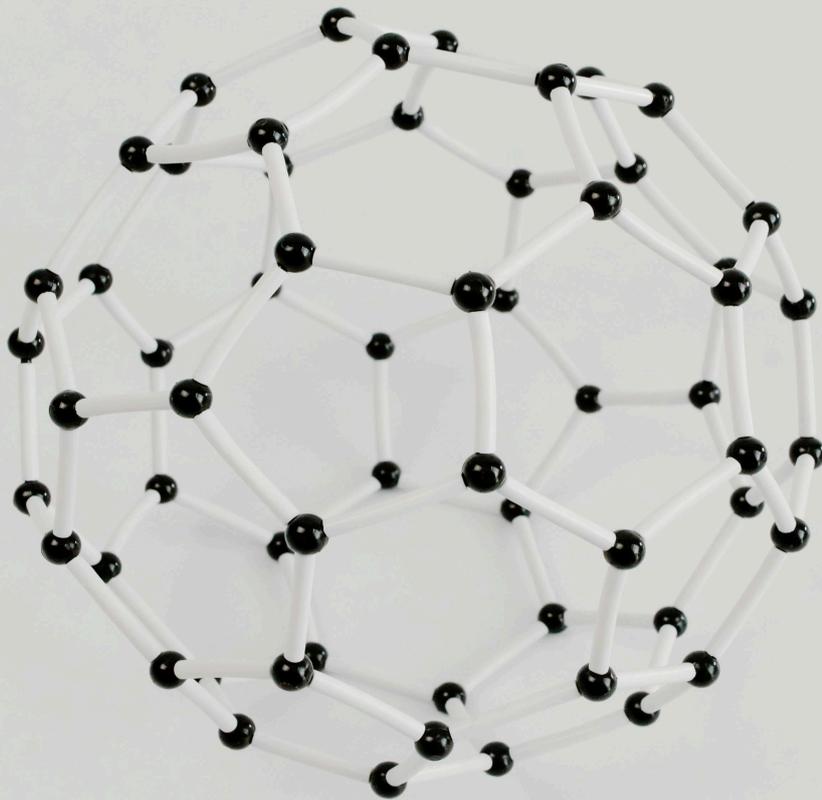


¿La intuición a veces puede traicionarte?



Dr. Francisco J. Valdés Parada

Departamento de Ingeniería de Procesos e Hidráulica.
Universidad Autónoma Metropolitana- Iztapalapa.

M.C. Jessica Sánchez Vargas

Departamento de Biología Molecular y Biotecnología,
Instituto de Investigaciones Biomédicas,
Universidad Nacional Autónoma de México

Abstract

Modeling transport phenomena in systems that involve multiple scales can be performed without ambiguities at the continuum scale level. However, when it is necessary to account for models at higher scale levels and that involve more phases, the same level of clarity and acceptance of the transport equations is not achieved. On the one hand, a rapid solution to this situation consists of using intuitive models founded on (natural or artificial) learning. On the other hand is the rigorous derivation of models at higher scale levels. In this work, both approaches are discussed and the relation that they have. The goal is to show the reader that one approach does not exclude the other and that it is worth dedicating time to discuss both points of view to carry out (rapid or rigorous) solutions to practical problems in science and engineering.

Keywords: Multiphase and multiscale models, upscaling, intuition, neuronal networks

Resumen

El modelado de fenómenos de transporte en sistemas que involucran escalas múltiples puede llevarse a cabo sin ambigüedades al nivel del continuo. Sin embargo, cuando se desea contar con modelos a niveles de escala superiores y que involucran más fases, no se cuenta con el mismo nivel de claridad y aceptación de las ecuaciones de transporte. Por un lado, una solución rápida a esta situación consiste en usar modelos intuitivos fundamentados en el aprendizaje (natural o artificial). Por otro lado, está la deducción rigurosa de los modelos en niveles de escala superiores. En este trabajo se discuten ambos enfoques y la relación que guardan entre sí. El propósito es mostrar al lector que un enfoque no es excluyente

del otro y que vale la pena dedicar tiempo a aprender de ambos puntos de vista para llevar a cabo soluciones (rápidas o rigurosas) de problemas prácticos en ciencias e ingeniería.

Palabras clave: Modelos multifásicos y multiescala, escalamiento, intuición, redes neuronales.

1. Introducción

Imagina que tienes una dolencia intensa en tu cuerpo; acudes pronto al médico y te da un remedio para el dolor. Además, te recomienda que resuelvas el problema de raíz, te hagas estudios y acudas con un especialista. Sin embargo, como ya no sientes dolor, decides no atender la recomendación y sigues con tu vida. Podrías tener suerte y que se tratase de algún mal pasajero. Si este no es el caso, esta decisión podría traducirse en más problemas en el futuro. Una situación análoga tiene lugar en ocasiones en ingeniería o en la ciencia en general. En este sentido, pueden identificarse dos niveles de profundidad en la búsqueda de una explicación (Liu *et al.*, 2024). Por un lado están las explicaciones más burdas/fáciles/correlacionales, las cuales podrían ser análogas al uso de un tratamiento paliativo. Por otro lado, están las más finas/rigurosas/causales, las cuales pueden asociarse con los tratamientos médicos dados para resolver la razón de tus dolencias. En ingeniería, en muchos casos, se utilizan modelos matemáticos para reproducir resultados experimentales. Sin embargo, no siempre existe un modelo previamente deducido para cada aplicación. Una solución rápida sería usar un modelo que haya probado ser exitoso en un sistema similar o bien proponer intuitivamente un nuevo modelo. Si dicho modelo reproduce, tras ajustes paramétricos, los resultados experimentales, adquiere credibilidad

y popularidad. La desventaja de esta solución es que los modelos suelen estar basados en suposiciones, implícitas o explícitas, que limitan su aplicación a ciertas condiciones y tipos de sistema. Más aún, en el caso de las suposiciones implícitas, se corre el riesgo de promover aplicaciones no fundamentadas que lleven a problemas a futuro. De hecho, no necesariamente porque un modelo matemático reproduzca datos experimentales se debe concluir, de inmediato, que el modelo es físicamente correcto. Más aún, como se mostrará más adelante, cuando este tipo de modelos heurísticos no aportan resultados satisfactorios se suelen proponer correcciones de la misma forma en la que fueron propuestos los modelos originales. La alternativa, al igual que en la analogía inicial, es deducir rigurosamente el modelo matemático correspondiente. Lo cual permite también desarrollar versiones más sofisticadas al relajar algunas de las suposiciones involucradas en la deducción del modelo.

Lo descrito anteriormente ocurre más frecuentemente en sistemas *multifásicos*¹ y *multiescala*². En ingeniería se concibe al nivel de escala del continuo como aquel en el que pueden definirse fases, las cuales están separadas entre sí por interfases. De esta forma, aunque la jerarquía de niveles de escala puede ser muy amplia, en este texto se considera el nivel de escala del continuo (*microescala*) y solo un nivel de escala superior donde hay dos o más fases involucradas (*macroescala*). En este caso, suelen adoptarse modelos con semejanza estructural a sus contrapartes en la microescala o recurrir a ecuaciones empíricas basadas en observaciones experimentales. O incluso, desde años más recientes, se emplean herramientas

derivadas de redes neuronales, para calcular, por ejemplo, los coeficientes del modelo en la macroescala.

¹Sistemas que, a cierto nivel de escala, comprenden más de una fase.

²Sistemas donde existen dos o más niveles de escala interrelacionados.

El objetivo de este trabajo es discutir acerca de los enfoques intuitivo y riguroso para el modelado de fenómenos de transporte en sistemas multiescala. Para ello, el resto de este trabajo está organizado como sigue: en la sección 2 se explica la diferencia entre formular modelos en la microescala (donde hay un alto grado de certidumbre) y en la macroescala (donde los modelos pueden proponerse de forma intuitiva o mediante técnicas rigurosas). Esto se ilustra con dos ejemplos sencillos: flujo de un fluido en un medio poroso y difusión y reacción en partículas catalíticas. Si bien las deducciones intuitivas pueden darse de forma natural con base en el conocimiento aprendido, también pueden darse de manera artificial usando redes neuronales. Esto último se discute a detalle en la sección 3. Por último, se presentan algunas reflexiones finales en la sección 4.

2. Modelado en escalas múltiples

En términos de fenómenos de *transporte*³, las ecuaciones que gobiernan lo que ocurre en la microescala son bien conocidas y aceptadas por la comunidad científica, pues tienen un sustento sólido en la termodinámica de los procesos irreversibles (De-Groot y Mazur 2011). Sin embargo, actualmente la solución de las ecuaciones gobernantes en la microescala solo puede llevarse a cabo en sistemas de tamaños que no sean prohibitivos computacionalmente mediante simulaciones numéricas

*directas*⁴. A niveles de escala superiores en donde se encuentra presente más de una fase, las ecuaciones gobernantes del transporte ya no cuentan con la misma claridad, aceptación universal y entendimiento como sus contrapartes en la microescala. Actualmente existen varias técnicas teóricas para deducir modelos de transporte en la macroescala (ver, por ejemplo, Battiato *et al.*, 2019) y aún no existe una técnica universalmente aceptada para deducir modelos a cualquier nivel de escala para cualquier tipo de transporte. Además, existe un uso extenso de modelos empíricos, y sus modificaciones, para el modelado de sistemas multifásicos y multiescala. A continuación se revisan dos ejemplos para ilustrar los puntos expuestos en esta sección que son: el transporte de cantidad de movimiento en medios porosos y el transporte de masa y reacción en partículas catalíticas.

2.1. Transporte de cantidad de movimiento en medios porosos

Imagina un medio poroso constituido por una fase fluida pasando por

los intersticios que se encuentran entre la fase sólida que compone la matriz del medio poroso. Este tipo de configuración puede encontrar aplicación en áreas como la medicina, hidrología subterránea, sistemas celulares, biotecnología, contaminación ambiental, entre muchas otras. Ahora considera que el flujo *incompresible*⁵, isotérmico, *reptante*⁶ y en estado *estacionario*⁷, mientras que el medio poroso es *rígido*⁸ y *homogéneo*⁹ como el mostrado en la figura 1.

³es decir, el estudio de la mecánica de fluidos, así como el transporte de calor y de especies químicas.

⁴de simulación computacional en la que se resuelven las ecuaciones al nivel de la microescala de forma numérica y sin suposiciones de escalamiento.

⁵es decir, que su densidad es constante.

⁶es decir, que las fuerzas de fricción superan a la inercia del fluido.

⁷es decir, que ninguna de las variables del sistema cambia con el tiempo.

⁸es decir, que la fase sólida no se mueve ni se deforma.

⁹es decir, que sus propiedades en la macroescala (como la porosidad) son constantes.

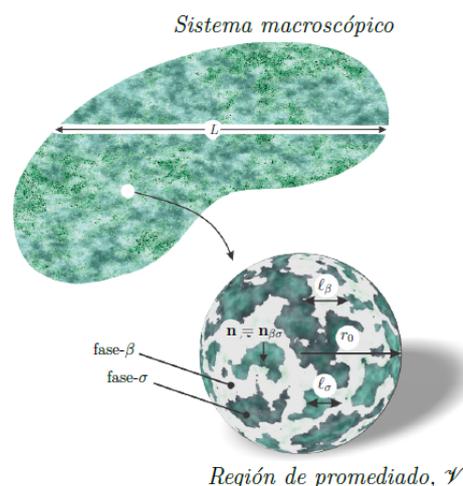


Figura 1: Esquema de un medio poroso compuesto por una fase fluida (fase- β) y una fase sólida (fase- σ). En este esquema se incluye además una región de promediado \mathcal{V} , cuya longitud característica, r_0 , es mucho mayor que las longitudes características en la microescala (ℓ_β , ℓ_σ) y, al mismo tiempo, mucho menor que la longitud característica de la macroescala, L .

Bajo estas condiciones tan particulares, las ecuaciones que gobiernan el transporte de masa total y la cantidad de movimiento en la fase fluida que satura los poros del sistema son las siguientes:

$$\nabla \cdot \mathbf{v} = 0, \quad (1a)$$

$$\mathbf{0} = -\nabla \mathcal{P} + \mu \nabla^2 \mathbf{v}. \quad (1b)$$

En las ecuaciones anteriores, v , \mathcal{P} y μ representan la velocidad, presión (modificada por los esfuerzos volumétricos) y viscosidad del fluido en el medio poroso, respectivamente. Por simplicidad, se supone que la velocidad del fluido en la interfase con la fase sólida en el medio poroso es cero. El modelo macroscópico de cantidad de movimiento correspondiente a esta situación física fue propuesto de manera empírica por Darcy en el siglo XIX al notar que existe una relación lineal entre el gradiente de presión macroscópica aplicado y la velocidad promedio resultante del fluido. Lo anterior puede expresarse de la siguiente forma:

$$\langle v_x \rangle = -\frac{K}{\mu} \frac{\partial \langle \mathcal{P} \rangle^\beta}{\partial x}. \quad (2)$$

En la ecuación anterior, $\langle v_x \rangle$ y $\langle \mathcal{P} \rangle^\beta$ representan respectivamente el promedio superficial de la componente x de la velocidad del fluido y el promedio intrínseco de la presión modificada.

Los operadores de promedio superficial e intrínseco de una variable

en la microescala ψ , se definen como sigue (Whitaker 1999):

$$\langle \psi \rangle = \frac{1}{V} \int_{\mathcal{V}} \psi dV, \quad (3a)$$

$$\langle \psi \rangle^\beta = \frac{1}{V_\beta} \int_{\mathcal{V}_\beta} \psi dV. \quad (3b)$$

Aquí, \mathcal{V} representa una región de promediado (cuyo volumen es V) que contiene porciones de las fases sólida y fluida del medio poroso (ver figura 1). De hecho, el espacio ocupado por la fase fluida en \mathcal{V} se denota como \mathcal{V}_β y su volumen es V_β . Por último, el coeficiente K en la ecuación (2) representa la permeabilidad del medio poroso, la cual sólo depende de la microestructura y no de las propiedades del fluido. Se verificó de manera experimental que el uso de este modelo está limitado por las suposiciones mencionadas al inicio de esta sección. Por lo anterior. En 1901 Forchheimer propuso la siguiente modificación a la ley de Darcy para estudiar el flujo en un medio poroso en condiciones de flujo laminar

$$-\frac{\partial \langle \mathcal{P} \rangle^\beta}{\partial x} = \frac{\mu}{K} \langle v_x \rangle + \rho \beta \langle v_x \rangle^2. \quad (4)$$

En esta ecuación ρ es la densidad del fluido y β es el coeficiente de resistencia inercial y debe determinarse experimentalmente. Como es de esperarse, la corrección de Forchheimer es sólo una de muchas modificaciones empíricas que se han hecho a la ley de Darcy. Fue hasta la segunda

mitad del siglo pasado que se hizo una deducción rigurosa de este modelo y se encontró que, a partir de las ecuaciones gobernantes en la microescala, se obtienen los siguientes modelos matemáticos en la macroescala (ver por ejemplo, el capítulo 3 de Whitaker 1999):

$$\nabla \cdot \langle \mathbf{v} \rangle = 0, \quad (5a)$$

$$\langle \mathbf{v} \rangle = -\frac{\mathbf{K}}{\mu} \cdot \nabla \langle \mathcal{P} \rangle^\beta. \quad (5b)$$

En las ecuaciones anteriores puede notarse que la ecuación de conservación de masa (ecuación (5a)) guarda una forma exactamente igual que su contraparte en la microescala (ecuación (1a)) y que la ecuación de balance de cantidad de movimiento (ecuación (5b)) es una generalización en 3D de la ley de Darcy en la forma dada en la ecuación (2) y que no guarda un símil con su contraparte en la microescala (ecuación (1b)). Vale la pena destacar que los métodos de escalamiento (como la homogeneización o el *promedio volumétrico*¹⁰) permiten no sólo deducir el modelo en la macroescala, sino también predecir los coeficientes involucrados en el modelo, en este caso el tensor de permeabilidad \mathbf{K} a partir de la solución de un problema de *cerradura*¹¹ asociado en un dominio representativo de la microescala. Esto constituye un esquema predictivo que no requiere de intuición o datos *experimentales*¹² para ser formulado o para predecir sus coeficientes. Más aún,

al relajar algunas de las suposiciones de partida en la formulación del modelo en la microescala, o bien durante el proceso de escalamiento, es posible deducir diversas modificaciones a la ley de Darcy y esto ha sido estudiado ampliamente en la literatura.

¹⁰Esta técnica consiste brevemente en los siguientes pasos: 1) aplicación de un operador de promediado (intrínseco o superficial) a las ecuaciones gobernantes en la microescala, 2) Intercambiar las operaciones de integración y diferenciación mediante los teoremas general del transporte y del promediado espacial, 3) Deducción, simplificación y solución formal de problemas de cerradura que permitan estimar las desviaciones locales respecto a las variables promedio y 4) Sustitución de la solución formal de los problemas de cerradura en el modelo obtenido en el paso 2) y, de ser pertinente, llevar a cabo simplificaciones adicionales al modelo macroscópico.

¹¹Este es un problema que se deduce para estimar las desviaciones de las propiedades bajo estudio respecto a sus valores promedio

¹²El que los modelos deducidos por métodos rigurosos puedan ser resultados sin la necesidad de datos experimentales no quiere decir que no haga uso de estos para fines de validación.

¹³Se le denomina como pseudo reacción a este coeficiente porque contiene los efectos de los coeficientes de adsorción, desorción y reacción.

2.2 Transferencia de masa y reacción en catalizadores porosos

Considera ahora un medio poroso con las mismas propiedades del ejemplo de la sección anterior (rígido, homogéneo, saturado por una sola fase fluida, etc.) solo que ahora la fase fluida lleva consigo a una especie química en solución diluida. Suponga que esta especie se transporta por difusión en la fase fluida y que tiene lugar una reacción química heterogénea de primer orden en los sitios activos localizados en la superficie de los sólidos que componen el medio poroso. La ecuación diferencial y la condición de

frontera interfacial que describen este caso de estudio en la microescala son las siguientes:

$$\frac{\partial c}{\partial t} = \nabla \cdot (\mathcal{D} \nabla c), \quad \text{en la fase fluida,} \quad (6a)$$

$$-\mathbf{n} \cdot \mathcal{D} \nabla c = kc, \quad \text{en la interfase sólido-fluido.} \quad (6b)$$

En las ecuaciones anteriores c representa la concentración molar de la especie química, \mathbf{n} es el vector unitario normal dirigido de la fase fluida a la sólida (ver figura 1), mientras que los coeficientes \mathcal{D} y k representan la difusión molecular y la pseudo reacción¹³ heterogénea. Usando la intuición y con algo de entrenamiento en el área de transferencia de masa en catalizadores porosos, resulta lógico proponer que el modelo en la macroescala sea el siguiente:

$$\phi \frac{\partial \langle c \rangle^\beta}{\partial t} = \phi \nabla \cdot (D_{ef} \nabla \langle c \rangle^\beta) - a_v k \langle c \rangle^\beta. \quad (7)$$

En la ecuación anterior, ϕ es la porosidad de la partícula catalítica, D_{ef} es el coeficiente de difusión efectiva, mientras que a_v es el área interfacial por unidad de volumen. El coeficiente de difusión efectiva debe depender de la microestructura del medio poroso y, además, ha habido un debate en la literatura sobre si este coeficiente depende o no de la reacción química. Esta pregunta es difícil de resolver con un modelo generado meramente de forma intuitiva, pues queda a consideración del autor y de los datos experimentales contra los que se valide el modelo. La anterior divergencia en opiniones no tendrá cabida en un modelo deducido de forma rigurosa, donde se identifican claramente las razones y suposiciones que llevan al modelo macroscópico. Retomando la ecuación (7), la incorporación de un término de reacción va de la mano con la idea de concebir al medio poroso como un

pseudo-continuo en donde se combina lo que ocurre en la fase fluida (y por ello se multiplican los términos de acumulación y difusión por la porosidad) y la reacción en la interfase sólido-fluido (la cual se multiplica por a_v).

A pesar de lo lógico y fácilmente aceptable que aparente ser el modelo dado en la ecuación (7), hay varios puntos débiles que pueden llevar a serios errores en su aplicación. Si el medio poroso no es *isótropo*¹⁴ la difusividad efectiva no puede ser un escalar, sino que debe ser un tensor de segundo orden. Lo anterior surge de forma similar al caso arriba expuesto para la permeabilidad del modelo de Darcy, que era originalmente un escalar, y se dedujo más tarde que debía ser, en general, un tensor de segundo orden. Más aún, la dependencia del coeficiente de difusión efectiva con la tasa de reacción debe ser clara. Esto último merece una explicación más detallada: ya que el modelo dado en la ecuación (7) tiene dos coeficientes de medio efectivo (la difusividad efectiva y la reacción efectiva, dada por $a_v k$). Si el coeficiente de difusión no dependiera de la reacción, entonces podría llevarse a cabo un experimento sin reacción para determinarlo y posteriormente un experimento con reacción para calcular el valor de $a_v k$. Sin embargo, si D_{ef} dependiera de la reacción química, la estrategia descrita arriba resulta inadecuada.

Lo anterior, claramente es una motivación para llevar a cabo deducciones rigurosas del modelo de transporte y reacción en la macroescala. Usando más de un tipo de técnica de escalamiento, se ha reportado en la literatura que, cuando el número de *Damköhler*¹⁵ es menor a la unidad, la ecuación (7) (reemplazando D_{ef} por

el tensor de segundo orden D_{ef}) es el resultado del escalamiento de sus contrapartes en la microescala (ver, por ejemplo, el capítulo 1 de Whitaker 1999). Sin embargo, cuando la tasa de reacción es superior a la difusión molecular, el coeficiente de difusión efectiva depende del número de Damköhler, en general. Por supuesto, el modelo discutido en esta sección es un caso particular y su uso está acotado por las suposiciones involucradas en su formulación. Por ejemplo, si el transporte de masa involucrara convección, si hubiera acumulación en la interfase sólido-fluido o bien si hubieran más especies químicas presentes, el modelo de medio efectivo de la ecuación (7) se modificaría sustancialmente y resultaría cada vez más difícil e incierto el modelo que podría surgir de postulaciones basadas en la intuición. Es por ello que la deducción rigurosa de modelos en la macroescala sigue siendo un campo de investigación activa en la actualidad.

¹⁴es decir, si el transporte en la microescala no se da igual en todas las direcciones.

¹⁵Este número adimensional relaciona la tasa de transporte por difusión respecto a la tasa de consumo por reacción.

Si bien ya se discutió sobre las ventajas de enfoques intuitivos y rigurosos para la deducción de modelos en la macroescala. Estos dos acercamientos pueden converger en el día a día. Más aún, hay ciertas aplicaciones, como la de redes neuronales, donde puede resultar difusa la línea que divide un modelo intuitivo de uno riguroso. Esto último se abordará en la siguiente sección.

3. Intuición artificial

Las redes neuronales artificiales, inspiradas en las redes neuronales biológicas, consisten en una red grande de varias capas de neuronas artificiales interconectadas. Se caracterizan por la propagación de señales desde una

capa de entrada, pasando a través de múltiples capas intermedias (ocultas), hasta llegar a una capa de salida. Toda red debe ser entrenada para efectuar de forma satisfactoria la operación para la que fue creada. Una red entrenada posee un conjunto de parámetros específicos que son ajustados de forma iterativa al tratar de minimizar una función de pérdida (localizada fuera de la red neuronal) durante el entrenamiento. A pesar de la variedad de redes neuronales existentes, aquí nos enfocaremos en las *PINNs*¹⁶, o Redes Neuronales Informadas por la Física. Éstas se caracterizan por contener dos partes importantes en su función de pérdida: Por un lado, se incorpora un término que calcula el error entre datos experimentales usados para entrenar la red y los datos de salida de la red neuronal con un determinado conjunto de parámetros. Por otro lado, se utilizan leyes físicas que pueden incluir ecuaciones diferenciales y condiciones de frontera. La combinación de ambas partes de la función de pérdida guía el proceso de aprendizaje hacia soluciones que no sólo sean acordes con los datos experimentales disponibles, sino que también conserven sentido físico. De esta forma, el papel que juegan las ecuaciones diferenciales y condiciones de frontera es fijar límites físicos que la red neuronal no debe violar en su búsqueda por encontrar una relación entre las variables de entrada y salida.

Es importante reconocer que las redes neuronales usadas actualmente se consideran, en general, modelos de caja negra. Lo anterior se debe a que, principalmente en redes complejas (como aquellas de aprendizaje *profundo*¹⁷), no hay completa claridad de los patrones específicos que reconoce cada una de las múltiples capas de transformaciones, ni de qué forma

contribuyen a la decisión final. A pesar de lo anterior, los modelos actuales de redes neuronales son funcionales en aplicaciones concretas en las ingenierías. Sin embargo, no suelen ser asociadas con la generación de conocimiento nuevo de un sistema (Liu *et al.*, 2024). Por lo anterior, podría decirse que las redes neuronales actuales, por lo general, asemejan a ese remedio paliativo que da el médico, como es también el caso de los modelos empíricos o correlaciones basadas en datos experimentales.

Para entrenar una red neuronal se requiere una gran cantidad de datos (del orden de miles a decenas de miles), cuyo número depende de la complejidad de la tarea (o número de parámetros que involucra), el tipo de red y la calidad de los datos (Bailly *et al.*, 2022). En muchos casos, la cantidad de datos disponibles en aplicaciones biológicas o ingenieriles no es suficiente para el uso de una red neuronal *convencional*¹⁸.

¹⁶del inglés Physics-Informed Neural Networks

¹⁷Es decir, aquellas que capturan patrones complejos mediante la interrelación y operación entre (al menos tres) capas neuronales ocultas.

¹⁸Al no tener suficientes datos, se promueve el sobreajuste durante el entrenamiento y esto evita que la red reconozca los patrones generales requeridos para llevar a cabo predicciones satisfactorias.

Como se mencionó arriba, el incorporar leyes físicas en las PINNs, permite reducir el número de datos necesarios para el entrenamiento al restringir el espacio de soluciones admitibles de la red. De esta forma, el algoritmo de aprendizaje alcanza más fácilmente la solución correcta y encuentra relaciones generales aún con pocos datos de entrenamiento (Kim y Lee 2024). Más aún, las PINNs subsanan también escenarios donde no se conocen (o se conocen parcialmente) las ecuaciones gobernantes mediante el uso de datos

experimentales. Esto último hace posible la aplicación de PINNs a problemas multiescala donde no se conoce el modelo en la macroescala pero probablemente sí se conozca el modelo en la microescala; es decir, donde no se cuenta con un modelo escalado.

A continuación se describirán dos ejemplos prácticos donde se desconoce el modelo escalado o la cerradura, y donde se pueden usar PINNs para subsanarlo. Para la primera, se plantea el uso de la red neuronal con el fin de evitar la necesidad de un problema de cerradura, el cual genera la relación entre la microescala y las variables en la macroescala. Para esto, se debe recordar que el escalamiento de las ecuaciones gobernantes en la microescala, usando, por ejemplo, el método del promedio volumétrico, implica un primer paso de promediado, seguido del uso de ciertos teoremas integrales. De esta forma, el modelo se expresa en términos de variables promedio e integrales de variables puntuales. Una desventaja que tienen las técnicas de escalamiento actuales es que para establecer una relación formal entre las variables en la microescala y en la macroescala, es necesario contar con un modelo lineal. Una alternativa a esto, es recurrir a una red neuronal para predecir directamente el resultado de la integral de las variables en la microescala. Esto se logra al entrenar la red para predecir el campo de la variable puntual y, posteriormente, llevar a cabo la integración. Con lo anterior se puede proceder a cerrar el modelo macroscópico. Para ilustrar la aplicación descrita en el párrafo anterior, conviene retomar la ecuación (1a). Tras la aplicación del operador de promediado superficial, definido en la ecuación (3a), se obtiene el siguiente resultado

$$\frac{1}{V} \int_{\mathcal{V}} \nabla \cdot \mathbf{v} dV = 0. \quad (8)$$

Si ahora se aplica el teorema del promediado espacial (Whitaker1999)¹⁹

$$\langle \nabla \cdot \mathbf{a} \rangle = \nabla \cdot \langle \mathbf{a} \rangle + \frac{1}{V} \int_{\mathcal{A}_{\beta\sigma}} \mathbf{n} \cdot \mathbf{a} dA, \quad (9)$$

al resultado anterior, se obtiene que

$$\nabla \cdot \langle \mathbf{v} \rangle + \frac{1}{V} \int_{\mathcal{A}} \mathbf{n} \cdot \mathbf{v} dA = 0. \quad (10)$$

A diferencia del problema clásico de Darcy, donde no hay transporte de masa hacia la interfase fluido-sólido, aquí se permitió este escenario.

¹⁹En la ecuación (9), \mathbf{a} es una variable definida en la fase fluida en la microescala y es suave por tramos.

Lo anterior implica que el segundo término del lado izquierdo de la ecuación (10) no es cero, y se requiere calcular la integral de $\mathbf{n} \cdot \mathbf{v}$ entre el fluido y el sólido en la microescala. La estimación de este último término sería el objetivo de la generación de una red neuronal que, partiendo de los datos experimentales disponibles (como por ejemplo mediciones globales del transporte de masa total, elementos de la microestructura, entre otros) y rigiéndose por los modelos en la microescala y promedio (ecuación (10)), prediga el valor de esa integral. Así, el resultado podría sustituirse en la expresión (10) y, de esta forma, cerrar el modelo promedio. Un segundo caso consiste en utilizar una PINN para sustituir por completo el proceso de escalamiento. En esta aplicación, la diferencia reside en que el objetivo es que la red aprenda a predecir una variable promedio a partir de cierta información en la microescala la cual puede obtenerse mediante simulaciones numéricas directas. Es decir, no intenta escalar el modelo explícitamente, sino que aprende a obtener el resultado correcto sin que en algún momento se

deduzca un modelo en la macroescala. Por supuesto, es posible mejorar las predicciones al incorporar información adicional (por ejemplo, datos en la macroescala, ecuaciones diferenciales, condiciones de frontera, etc.) a la función de pérdida. Debe mencionarse que existen también aplicaciones de las redes neuronales a sistemas jerárquicos donde se requiere mayor grado de dominio de los métodos de escalamiento (Kim y Lee 2024). Un ejemplo son la generación de redes neuronales informadas por dichos métodos para la predicción de los esquemas de cerradura o de la forma promedio de las ecuaciones de transporte. Así también, existen aplicaciones de las redes para la solución de los problemas de cerradura, previamente deducidos.

4. Reflexiones finales

Como se detalló en las secciones anteriores, puede buscarse nuevo conocimiento, en general, por dos vías: La primera está impulsada por la aplicación, dentro de la cual se pueden incluir todo tipo de técnicas que surjan de la intuición humana o artificial. La intuición es útil como una solución pronta, pero no como un sustituto a una solución a largo plazo. La intuición (natural o artificial) sigue siendo intuición y debe tomarse con precaución, pues no pueden obtenerse respuestas ampliamente generalizables. La segunda vía de generación de conocimiento es aquella impulsada por la comprensión, generalmente basada en metodologías concretas, principios fundamentales y haciendo uso de herramientas matemáticas. La deducción rigurosa permite crear formas de pensar que no se han descrito antes y son difíciles de obtener de manera intuitiva (humana o artificial) actualmente. Aunque ambos enfoques son de gran relevancia,

plantean preguntas diferentes y generan respuestas que no pueden ser reemplazadas por el otro enfoque. Por lo anterior, es completamente válido hacer uso de soluciones temporales o del uso de la tecnología para resolver nuestros problemas, siempre y cuando se reconozca aún la necesidad de aprender los fundamentos y hacer deducciones rigurosas, a modo de no depender irremediabilmente de estas “soluciones prácticas” que son específicas para algunos casos concretos. En síntesis, la respuesta que proponemos a la pregunta sobre si la intuición puede traicionarte, es: no se puede generalizar, pero en muchos casos se puede confiar en la intuición a corto plazo para obtener soluciones rápidas. Sin embargo, lo anterior no sustituye la necesidad de continuar haciendo deducciones rigurosas para avanzar en la ciencia. De hecho, la relación entre soluciones rigurosas e intuitivas es simbiótica. Esto se debe a que la deducción de muchos modelos rigurosos ha sido motivada por modelos intuitivos previos. Además, la deducción y cerradura de algunos modelos intuitivos se benefician de conocimientos físicos previos. Incluso, las herramientas actuales que emplean redes neuronales pueden ser mejoradas si se cuenta con principios matemáticos bien establecidos. Es por ello que vale la pena aprender sobre ambos enfoques para atender (rápida o rigurosamente) una amplia variedad de problemas en ciencias e ingeniería.

Referencias

- [1] Bailly, A., Blanc, C., Francis, E., Guillotin, T., Jamal, F., Wakim, B., & Roy, P. (2022). Effects of dataset size and interactions on the prediction performance of logistic regression and deep learning models. *Computer Methods and Programs in Biomedicine*, 213, 106504.
- [2] Battiato, I., V, P. T. F., Malley, D. O., Miller, C. T., Takhar, P. S., Valdes-Parada, F. J., & Wood, B. D. (2019). Theory and Applications of Macroscale Models in Porous Media. *Transport in Porous Media*, 130 (1), 5-76.
- [3] De-Groot, S. R., & Mazur, P. (2011). *Non-Equilibrium Thermodynamics*. Dover Publications.
- [4] Kim, D., & Lee, J. (2024). A Review of Physics Informed Neural Networks for Multiscale Analysis and Inverse Problems. *Multiscale Science and Engineering*, 6 (1), 1-11.
- [5] Liu, Z., Ma, P., Wang, Y., Matusik, W., & Tegmark, M. (2024). KAN2.0: Kolmogorov-Arnold Networks Meet Science.
- [6] Whitaker, S. (1999). *The Method of Volume Averaging*. Springer Netherlands.