

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

Análisis de intermitencias hidrodinámicas en flujos bifásicos gas-líquido de alta viscosidad mediante técnicas de Aprendizaje Automático

TESIS

Que para obtener el título de **Ingeniero Petrolero**

PRESENTA

Jesús Alberto González Treviño

DIRECTOR DE TESIS

Dr. José Enrique Guzmán Vázquez



Ciudad Universitaria, Cd. Mx., 2025



PROTESTA UNIVERSITARIA DE INTEGRIDAD Y HONESTIDAD ACADÉMICA Y PROFESIONAL (Titulación con trabajo escrito)



De conformidad con lo dispuesto en los artículos 87, fracción V, del Estatuto General, 68, primer párrafo, del Reglamento General de Estudios Universitarios y 26, fracción I, y 35 del Reglamento General de Exámenes, me comprometo en todo tiempo a honrar a la institución y a cumplir con los principios establecidos en el Código de Ética de la Universidad Nacional Autónoma de México, especialmente con los de integridad y honestidad académica.

De acuerdo con lo anterior, manifiesto que el trabajo escrito titulado <u>ANALISIS DE</u> <u>INTERMITENCIAS HIDRODINAMICAS EN FLUJOS BIFASICOS GAS-LIQUIDO DE ALTA</u> <u>VISCOSIDAD MEDIANTE TECNICAS DE APRENDIZAJE AUTOMATICO</u> que presenté para obtener el titulo de <u>INGENIERO PETROLERO</u> es original, de mi autoría y lo realicé con el rigor metodológico exigido por mi Entidad Académica, citando las fuentes de ideas, textos, imágenes, gráficos u otro tipo de obras empleadas para su desarrollo.

En consecuencia, acepto que la falta de cumplimiento de las disposiciones reglamentarias y normativas de la Universidad, en particular las ya referidas en el Código de Ética, llevará a la nulidad de los actos de carácter académico administrativo del proceso de titulación.

JESUS ALBERTO GONZALEZ TREVIÑO Número de cuenta: 313028306

Jurado Asignado

Presidente:	Dr. José Luis Fernández Zayas
Secretario:	Dra. Ana Paulina Gómora Figueroa
Vocal:	Dr. José Enrique Guzmán Vázquez
Suplente 1:	Dra. Flor Lizeth Torres Ortiz
Suplente 2:	Dr. Jonathan Hernández García

Realizar este trabajo me ha resultado aleccionador: me ha mostrado que no sé nada y más importante aún, me ha enseñado a ser consciente de ello. — JAGT.

Agradecimientos

Quisiera dedicar el presente trabajo a mi hermano y a mis padres. A ellos les agradezco otorgarme, en la medida de sus posibilidades, el privilegio de poder estudiar. De igual manera agradezco a mis tíos Mónica y Manuel por todo el apoyo brindado especialmente en estos últimos años.

Al Dr.Enrique, por integrarme en su equipo de trabajo, por asignarme un tema para poder graduarme y por enseñarme a trabajar siempre con la máxima calidad posible.

A Jonathan, por todos los conocimientos técnicos y extracurriculares que logró transmitirme, por siempre tener espacio para responderme alguna pregunta y por enseñarme el valor de ser integral.

A Matei, por sus comentarios y críticas constructivas precisas y por traer un nuevo ritmo a las reuniones de equipo.

A los profesores de la Facultad de Ciencias por aceptarme en sus asignaturas.

A la Dra. Lizeth por invitarme a los cursos que imparte.

A la Dra. Paulina por su retroalimentación relativa al trabajo escrito.

Al Dr. José Luis por aceptar amablemente ser parte del comité de titulación.

A todos los compañeros y amigos que conocí: José Miguel, Omar, Erik, Sofi, Gloria, Gerado, Silvia, Jeremy, Abi, Juan, Diego y Mariana. Mis mejores deseos a cada uno de ustedes.

A la UNAM, por ser mi alma mater; a la FI por formarme y al IINGEN por permitirme concluir esta estapa de mis estudios.

Resumen

Se analiza un conjunto de series instantáneas de presión en línea y colgamiento de líquido de un flujo tapón glicerina-aire de 0.6 Pa·s en una tubería horizontal de 0.0762 metros de diámetro nominal. Se utiliza un perceptrón multicapa cuyas entradas son las series instantáneas mencionadas. Se descubre que las predicciones de colgamiento de líquido para presiones desconocidas se aproximan al colgamiento de líquido promedio cuando la función de pérdida de la red neuronal es el error cuadrático medio. Para predecir el colgamiento de líquido en las regiones de película, frente y cuerpo del tapón se efectúa una descomposición de dicha serie temporal por medio del modelo de mezcla Gaussiano; se realiza un análisis estadístico del flujo a partir del ajuste del modelo de mezcla y se deducen ecuaciones que permiten estimar para cada una de las fases los volúmenes totales y acumulados en cada una de las regiones de flujo consideradas. Una vez realizada la agrupación, con un conjunto de perceptrones multicapa se predice el colgamiento de líquido por región de flujo para presiones desconocidas.

Abstract

A set of instantaneous series of in-line pressure and liquid holdup for a glycerin-air slug flow of 0.6 Pa \cdot s in a horizontal pipe nominal diameter of 0.0762 meters is analyzed. A multilayer perceptron is used, with the aforementioned instantaneous series as inputs. It is found that liquid holdup predictions for unknown pressures approach the average liquid holdup when the neural network's loss function is the mean squared error. To predict liquid holdup in the stratified region, mixture front and slug's body a decomposition of the time series is performed using the Gaussian Mixture Model; a statistical analysis of the flow is carried out based on the model fit, and equations are derived to estimate the total and accumulated volumes for each phase in the considered flow regions. After grouping, a set of multilayer perceptrons is used to predict liquid holdup by flow region for unknown pressures.

CONTENIDO

Re	lesumen		
A۱	bstra	ct	v
1.	INT	RODUCCIÓN	1
	1.1.	Hipótesis	6
	1.2.	Objetivos	6
	1.3.	Organización del documento	7
2 .	MA	RCO TEÓRICO	8
	2.1.	Conceptos de flujo tapón	8
		2.1.1. Colgamiento de líquido	11
	2.2.	Preliminares de aprendizaje automático	13
	2.3.	Conceptos de aprendizaje probabilístico	14
		2.3.1. Modelos de variables latentes y modelos de mezcla	15
		2.3.2. Máxima verosimilitud para el MMG	18
		2.3.3. Algoritmo E-M para el MMG	19
	2.4.	Conceptos de aprendizaje profundo	21
		2.4.1. Modelo de una neurona artificial	21

		2.4.2.	Modelo de perceptrón multicapa	23
3.	EST	UDIO	PRELIMINAR DEL CONJUNTO DE DATOS	32
	3.1.	Prueba	as bifásicas gas-líquido de alta viscosidad	32
		3.1.1.	Instrumentación empleada	34
		3.1.2.	Adquisición de datos	38
	3.2.	Evalua	ación de la incertidumbre de medición	38
		3.2.1.	Errores aleatorios y sistemáticos	39
		3.2.2.	Evaluación Tipo A de la incertidumbre	40
		3.2.3.	Evaluación Tipo B de la incertidumbre	42
		3.2.4.	Incertidumbre estándar sistemática	43
		3.2.5.	Incertidumbre combinada y expandida	43
		3.2.6.	Consideraciones experimentales para la evaluación de la incertidumbre	44
		3.2.7.	Reporte de incertidumbre	45
	3.3.	Anális	is preliminar del conjunto de datos	48
		3.3.1.	Colgamiento de líquido	48
		3.3.2.	Flujos másicos	50
		3.3.3.	Temperatura	52
		3.3.4.	Investigación de la relación entre la presión y colgamiento de líquido .	53
4.	AN	ÁLISIS	S Y DISCUSIÓN DE RESULTADOS	59
	4.1.	Aplica miento	ción del modelo de mezcla Gaussiano a la serie instantánea de colga- o de líquido	59
		4.1.1.	Identificación de regiones de flujo	63
		4.1.2.	Estimación de volúmenes	71

	4.2. Aplicación del modelo de perceptrón multicapa	82
	4.2.1. Relación entre la presión y el colgamiento de líquido instantáneos por región de flujo	92
5.	CONCLUSIONES	104
	5.1. Contribuciones	106
	5.2. Trabajo futuro	107
	5.3. Recomendaciones	108
А.	. Artículo publicado	109

Lista de Figuras

2.1.	Unidad de tapón de un flujo gas-líquido de alta viscosidad. Fuente: IINGEN- UNAM	10
2.2.	Mecanismo de arrastre de burbujas por plegamiento del frente del tapón. Fuen- te: IINGEN-UNAM.	11
2.3.	Colgamiento de líquido evaluado en cada región de flujo de la unidad de tapón. Fuente: IINGEN-UNAM	12
2.4.	Clasificación de los modelos a utilizar dentro del aprendizaje automático. Fuen- te: Autor	14
2.5.	Relación jerárquica entre las variables de un modelo de mezcla. Fuente: Autor.	16
2.6.	Modelo matemático de una neurona artificial. Adaptado de (Hagan et al., 1997).	21
2.7.	Modelo de una capa de neuronas. Adaptado de (Hagan et al., 1997) \ldots .	23
2.8.	Modelo de perceptrón multicapa. Adaptado de (Hagan et al., 1997) $\ .\ .\ .$.	24
3.1.	Circuito de Flujos Multifásicos IINGEN-UNAM. Adaptado de (Hernández, 2019)	33
3.2.	Tomógrafo en línea. Fuente: Autor	35
3.3.	Medidores de flujos másicos para cada una de las fases	35
3.4.	Transductor de presión. Fuente: Autor	36
3.5.	Distancia entre el segundo transductor de presión y el sensor de tomografía. Fuente: Autor	36

3.6.	Termopares tipo K	37
3.7.	Viscosímetro Brookfield DV2T. Fuente: Fabricante	38
3.8.	Módulos utilizados para la adquisición de datos.	38
3.9.	Series instantáneas de colgamiento de líquido. Los gráficos (a)-(c), (d)-(f) y (g)-(i) corresponden a los experimentos G1L1, G1L2 y G1L3 de la Tabla 3.1, respectivamente. Fuente: Autor	48
3.10.	Flujo másico de la fase gaseosa. Fuente: Autor	50
3.11.	Flujo másico de la fase líquida. Fuente: Autor	51
3.12.	Comparativa entre T_1 , T_2 y $T_{amb.}$. Fuente: Autor.	52
3.13.	Espectro de potencia, EXP G1L1, $\dot{m}_L = 0.6 (kg/s)$. Fuente: Autor	54
3.14.	Espectro de potencia, EXP G1L2, $\dot{m}_L = 1.3 (kg/s)$. Fuente: Autor	54
3.15.	Espectro de potencia, EXP G1L3, $\dot{m}_L = 2.0 (kg/s)$. Fuente: Autor	55
3.16.	Colgamiento de líquido en términos de la presión en línea. Fuente: Autor	57
3.17.	Colgamiento de líquido en términos de la presión en línea (normalizada con el máximo de cada experimento). Fuente: Autor	57
4.1.	Criterio de información de Bayes. Fuente: Autor	60
4.2.	Criterio de información de Akaike. Fuente: Autor	61
4.3.	Medias de cada uno de los componentes del modelo de mezcla para los dife- rentes experimentos. Fuente: Autor	63
4.4.	Desviaciones estándar de cada uno de los componentes del modelo de mezcla para los diferentes experimentos. Fuente: Autor	64
4.5.	Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L1. Fuente: Autor	65
4.6.	Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L2. Fuente: Autor	66
4.7.	Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L3. Fuente: Autor	67

4.8.	Distribuciones de probabilidad. Experimento G1L1. Fuente: Autor	69
4.9.	Distribuciones de probabilidad. Experimento G1L2. Fuente: Autor	70
4.10.	Distribuciones de probabilidad. Experimento G1L3. Fuente: Autor	70
4.11.	Pesos del modelo de mezcla por región de flujo. Fuente: Autor	71
4.12.	Configuración de dos fases con su respectiva fracción de área. Fuente: Autor.	74
4.13.	Configuración de los volúmenes asociados a cada grupo para una determinada fase. Fuente: Autor	75
4.14.	Volumen acumulado de aire en la burbuja de Dumitrescu. Fuente: Autor . $\ .$.	77
4.15.	Volumen acumulado de líquido en la película. Fuente: Autor	78
4.16.	Volumen acumulado de aire en el cuerpo del tapón. Fuente: Autor	79
4.17.	Volumen acumulado de aire en el frente de mezcla. Fuente: Autor	79
4.18.	Volumen acumulado de líquido en el frente de mezcla. Fuente: Autor	80
4.19.	Volumen acumulado de líquido en el cuerpo del tapón. Fuente: Autor	81
4.20.	Comparativa del volumen acumulado de gas en la burbuja de Dumistresco y el volumen acumulado de líquido en la película. Fuente: Autor	82
4.21.	Arquitectura de red para el modelo PCI. Fuente: Autor	84
4.22.	Historia de entrenamiento para el modelo PCI. Fuente: Autor	85
4.23.	Valores de colgamiento de líquido predichos. Modelo PCI. Fuente: Autor	85
4.24.	Historia de entrenamiento. Red G1L1. Fuente: Autor	87
4.25.	Predicciones de colgamiento de líquido. Red G1L1. Fuente: Autor	87
4.26.	Comparativa de $\overline{H_L}$ para las predicciones del modelo PCI. Fuente: Autor	90
4.27.	Predicción del colgamiento de líquido con una red neuronal con 3 neuronas en la capa de salida. Fuente: Autor	91
4.28.	Arquitectura de red para la región de la película de líquido. Fuente: Autor.	94
4.29.	Arquitectura de red para la región del frente de mezcla. Fuente: Autor	95

4.30. Arquitectura de red para la región del cuerpo del tapón. Fuente: Autor	95
4.31. Historia de entrenamiento. Red neuronal para la región de la película de líqui- do. Fuente: Autor.	96
4.32. Historia de entrenamiento. Red neuronal para la región del frente de mezcla. Fuente: Autor	97
4.33. Historia de entrenamiento. Red neuronal para la región del cuerpo del tapón. Fuente: Autor	98
$4.34.$ Predicciones del colgamiento de líquido por región de flujo. Fuente: Autor. $\ .$	99
4.35. Predicciones del colgamiento de líquido promedio y por región de flujo. Fuente: Autor.	101
4.36. Predicciones de colgamiento de líquido para presiones desconocidas por región de flujo. Fuente: Autor	102
4.37. Predicciones de colgamiento de líquido promedio y por región de flujo para presiones desconocidas. Fuente: Autor.	103
A.1. Artículo publicado en Physics of Fluids.	109

Lista de Tablas

2.1.	Funciones de activación.	22
2.2.	Procedimiento utilizado para la construcción de un modelo de perceptrón mul- ticapa	31
3.1.	Conjunto experimental analizado.	33
3.2.	Variables monitoreadas durante la campaña experimental	34
3.3.	Distancia entre transductores de presión en la sección de pruebas	37
3.4.	Evaluación Tipo A y B. Experimento G1L1	45
3.5.	Evaluación Tipo A y B. Experimento G1L2	45
3.6.	Evaluación Tipo A y B. Experimento G1L3	46
3.7.	Incertidumbres. Experimento G1L1	46
3.8.	Incertidumbres. Experimento G1L2	47
3.9.	Incertidumbres. Experimento G1L3	47
4.1.	Número óptimo de componentes. Criterios de Bayes y Akaike	59
4.2.	Parámetros del modelo de mezcla Gaussiano para la serie de colgamiento de líquido	62
4.3.	Identificación de regiones de flujo de la unidad de tapón con los componentes del MMG	68
4.4.	Volumen acumulado de aire por región de flujo.	76

4.5.	Volumen acumulado de glicerina por región de flujo	76
4.6.	Volúmenes totales de glicerina y aire	76
4.7.	Configuración del modelo y del espacio de búsqueda utilizado	83
4.8.	Cantidad de puntos (p_i, H_{L_i}) para el modelo PCI	83
4.9.	Arquitectura, función de activación y tasa de aprendizaje para el modelo PCI.	84
4.10.	Componentes del MMG para las redes neuronales por región de flujo	93
4.11.	Cantidad de puntos (p_i, H_{L_i}) utilizada para cada red neuronal por región de flujo	93
4.12.	Arquitectura, función de activación y tasa de aprendizaje para cada red neu- ronal por región de flujo	94

NOMENCLATURA

Variables

	Símbolo	Unidad (SI)	Unidades (campo)
$\operatorname{\acute{A}rea}$	A	m^2	
Colgamiento	H	-	-
Diámetro externo	D	m	in
Diámetro interno	d	m	in
Flujo másico	\dot{m}	${ m kg} \cdot { m s}^{-1}$	$lb \cdot s^{-1}$
Longitud	L	m	ft
$\mathbf{Presión}$	p	Pa	psi
Presión promedio	\overline{p}	Pa	psi
Presión adimensional	p/p*	-	-
Tiempo	t	S	\mathbf{S}
Temperatura	T	Κ	$^{\circ}\mathrm{C}$
Volumen	V	m^3	bbl

Letras griegas

Flujo multifásico

	Símbolo	Unidad (SI)	Unidades de campo
Densidad	ho	${ m kg} \cdot { m m}^{-3}$	-
Viscosidad dinámica	μ	$Pa \cdot s$	cР
Fracción de vacío	α	-	-

Aprendizaje automático

	\mathbf{S} ímbolo
Función de	¢
activación	φ
Tasa de	n
$\operatorname{aprendizaje}$	'/

Subíndices

Del frente de mezcla	fm
Del modo de Gauss	k
De la película de líquido	plc
De la tubería	tub
De líquido/ de la fase líquida	L
De la fase gaseosa	G
Del tapón/slug	S
De escalado/escalamiento	esc

Conjuntos

De entre	enamiento	\mathcal{T}
De valid	lación	\mathcal{V}
De prue	ba	\mathcal{P}

Abreviaturas

 ${\rm Transductor} \quad Trd$

1. INTRODUCCIÓN

Las actividades relacionadas con la industria petrolera son claves desde el punto de vista económico para nuestro país. Prueba de ello es el 57.37 % que ha aportado Petróleos Mexicanos al ingreso total de la federación en lo referente a la prestación de servicios por parte de empresas públicas del Estado (LIF, 2024) para el 2025. De dicha contribución, un porcentaje significativo recae en la exportación de crudo y gas natural.

Para llegar a los puntos de venta, la mezcla de hidrocarburos que proviene de los yacimientos viaja en primer lugar a través de los pozos y posteriormente en los ductos que se dirigen hacia las instalaciones superficiales. Es en este camino donde las fases que componen dicha mezcla se separan y forman un flujo simultáneo el cual se conoce como flujo multifásico; si bien este tipo de flujo es estudiado en otras industrias, en el sector petrolero gran parte del trabajo ha estado orientado en la predicción del gradiente de presión, con el objetivo de dimensionar de manera óptima las tuberías verticales, horizontales e inclinadas que comprenden el sistema integral de producción.

Específicamente, una de las principales diferencias para predecir el gradiente de presión en el caso multifásico respecto al caso monofásico, es la identificación y cuantificación de las distintas configuraciones geométricas o patrones de flujo que pueden adoptar cada una de las fases dentro de la tubería. Dicha identificación y cuantificación es compleja considerando que los patrones de flujo son función de las densidades, flujos volumétricos, viscosidades, tensiones interfaciales y compresibilidades de cada una de las fases; estas propiedades dependen a su vez de la presión y temperatura en un tiempo y posición específicos. Además, se ha encontrado que los patrones de flujo también son altamente sensibles al diámetro e inclinación del conducto (J. Brill & Arirachakaran, 1992) y (Barnea, 1987). Patrones de flujo comúnes en tuberías verticales son el flujo burbuja y el flujo anular (McQuillan & Whalley, 1985); para tuberías horizontales y con poco ángulo de inclinación configuraciones recurrentes de las fases son el flujo estratificado y el flujo tapón.

Como tal, las primeras aproximaciones al cálculo del gradiente de presión se llevaron a

cabo a mediados del siglo pasado y consistían en relaciones empíricas donde se vinculaban las pérdidas de energía y datos obtenidos en campo (Poettman & Carpenter, 1952), (Gilbert, 1954), (Baxendell & Thomas, 1961). Dichos modelos fueron desarrollados principalmente para pozos de alto flujo volumétrico y consideraban a las fases involucradas como una mezcla homogénea la cual se asumía que viajaba a la misma velocidad. Conceptos como velocidad superficial o velocidad de la mezcla fueron introducidos en este periodo.

No obstante, debido a la gran cantidad de variables involucradas en el flujo gas-líquido, fue necesario investigar grupos adimensionales (Ros, 1961) con el propósito de desarrollar correlaciones para rangos más amplios de propiedades de fluidos y de configuraciones en tubería. Ejemplos de este tipo de correlaciones son las presentadas por (Hagedorn & Brown, 1965) para pozos (tuberías verticales), (Eaton et al., 1967) para tuberías horizontales y (Beggs & Brill, 1973) para tuberías inclinadas. Dichas correlaciones además de utilizar grupos adimensionales, introdujeron para la estimación del gradiente de presión el cálculo de un parámetro (o factor adicional) conocido como colgamiento de líquido¹, que se dene como la fracción entre el área ocupada por líquido y el área total que es transversal al flujo. Como tal, el colgamiento en estas correlaciones se cuantificaba como el líquido atrapado en la tubería después de llevar a cabo un cierre rápido de válvulas. Respecto a las sustancias utilizadas, en la mayoría de los casos se empleaban combinaciones agua-aire o gas natural-crudo ligero.

Por otro lado, con el objetivo de mejorar la precisión, el rango de aplicación así como el entendimiento de los procesos o mecanismos observados en el flujo gas-líquido, se derivaron modelos mecanicistas como el propuesto por (Dukler & Hubbard, 1975), el cual fue desarrollado específicamente para predecir el comportamiento hidrodinámico de un flujo tapón agua-aire. Si bien el modelo es determinista, sus autores mencionan con base en la evidencia experimental, que el flujo tapón presenta características aleatorias. A pesar de esta consideración, el modelo permite estimar la velocidad del tapón, del frente de mezcla así como de la película de líquido. No obstante, dicho modelo requiere como entradas la especificación de parámetros de flujo adicionales, tal como el colgamiento de líquido en el tapón. Correlaciones utilizadas comúnmente para la predicción de este parámetro son las propuestas por (Gregory et al., 1978), (Gomez et al., 2000) y (G. H. Abdul-Majeed, 2000).

De manera paralela al desarrollo de los primeros modelos mecanicistas, mejoras en la instrumentación como el uso de sistemas de rayos X de respuesta rápida permitieron efectuar las primeras mediciones instantáneas de la fracción de vacío. En particular, en el estudio de (Jones Jr & Zuber, 1975) se introduce la aplicación de la función de densidad de probabilidad (FDP) como una forma de identificar los distintos patrones de flujo. Además, a partir del análisis estadístico llevado a cabo, se menciona la idea de que el flujo tapón puede considerarse como una combinación transitoria de los flujos burbuja y anular, de allí el comportamiento bimodal observado en la FDP de la fracción de vacío. Las sustancias utilizadas para este trabajo fueron agua y aire y los experimentos se llevaron a cabo en un canal rectangular; para conductos circulares, en el trabajo de (Vince & Lahey Jr, 1982) se analiza la fracción de

 $^{^{1}}$ Al complemento del colgamiento de líquido se le denomina fracción de vacío.

vacío para un flujo agua-aire mediante el empleo de los cuatro primeros momentos de la FDP para la clasificación de patrones de flujo. En su estudio, se encuentra que la varianza es el mejor indicador para dicha tarea y que la densidad de potencia espectral (DPE²) es sensible a la velocidad superficial de líquido. Otra investigación que incorpora el uso de la DPE para la identificación de patrones de flujo es la de (dos Reis & Goldstein Jr, 2010). En cuanto al análisis del comportamiento bimodal de la serie temporal de colgamiento, (Soto-Cortes et al., 2021) emplea la técnica de estimación de densidad de núcleo para caracterizar al flujo tapón.

Bajo este panorama de modelado mecanicista y de mejoras en la instrumentación así como en la incorporación de técnicas estadísticas, se sumó al estudio del flujo multifásico en tuberías la adopción del modelo de dos fluidos, el cual resultó en la creación del primer simulador comercial para flujos transitorios en condiciones multifásicas (OLGA) (Bendiksen et al., 1991). Una de las principales ventajas de este enfoque es que las configuraciones de las fases se analizan como una parte intrínseca al flujo, a diferencia de las correlaciones de (Eaton et al., 1967) o de (Beggs & Brill, 1973), donde la ecuación a utilizar se determina en función del patrón de flujo. Sin embargo, a pesar de la complejidad matemática introducida en este nuevo enfoque, correlaciones para describir la transición del flujo anular al flujo tapón (Wallis, 1969) o para el factor de fricción en el flujo anular (Wallis, 1970) aún son requeridas.

De toda esta variedad de enfoques mencionada, factores como el incremento en el poder de cómputo así como el acumulamiento de datos disponibles en la literatura, favorecieron que para finales de siglo se incorporara el paradigma de aprendizaje automático o aprendizaje basado en datos (Müller & Guido, 2016), (Bengio et al., 2017) y (A. Zhang et al., 2021) como una herramienta adicional para el estudio del flujo multifásico en tuberías. En el trabajo pionero en este campo, (Ternyik IV et al., 1995), con base en los datos reportados por (Mukherjee, 1979), empleó dos redes neuronales (perceptrones multicapa³) para predecir el colgamiento de líquido y los patrones de flujo en tuberías inclinadas para mezclas gas-líquido de baja viscosidad. Las redes neuronales construidas mostraron una precisión superior en los dos problemas abordados respecto a la propia correlación de (Mukherjee, 1979).

En un estudio posterior, (Osman, 2004) utilizó 199 datos experimentales recabados de los trabajos de (Minami & Brill, 1987) y (G. Abdul-Majeed, 1996) para desarrollar dos redes neuronales de tres capas con el objetivo de estimar el colgamiento de líquido y el patrón de flujo en tuberías horizontales para mezclas agua-aire y aire-queroseno. Para la predicción del colgamiento de líquido, la red neuronal mostró una mejora significativa respecto

²La densidad de potencia espectral o también conocida simplemente como espectro de potencia es una función que describe cómo se distribuye la potencia de una señal en términos de la frecuencia. Esta función puede tomarse como un indicador de la actividad de una serie temporal. Para más información puede consultarse (Koopmans, 1995), (Ziemer et al., 1998), (Oppenheim, 1999), (Proakis, 2001), (Roberts, 2007) y (Howell, 2016).

³El perceptrón multicapa es un tipo de red neuronal en donde el procesamiento de información se efectúa únicamente desde la capa de entrada hasta la capa de salida, i.e., en su arquitectura no se incluyen bucles o algún tipo de retroalimentación. Para más información, consultar (Haykin, 2009), (Bengio et al., 2017) y (A. Zhang et al., 2021).

a las correlaciones de (Guzhov et al., 1967), (Beggs & Brill, 1973), (J. P. Brill et al., 1981), (Gregory et al., 1978), (G. A. Hughmark & Pressburg, 1961) y (G. Hughmark et al., 1962). Para la clasificación de patrones de flujo y tomando como referencia el mapa propuesto por (Mandhane et al., 1974), con la red se determinó adecuadamente más del 94 % de las configuraciones de las fases consideradas. En un estudio similar, (Shippen & Scott, 2004) construyeron una red neuronal de tres capas ocultas para predecir el colgamiento de líquido en tuberías horizontales para líquidos de baja viscosidad a partir de los datos reportados por (Eaton et al., 1967), (Beggs & Brill, 1973), (Mukherjee, 1979), (Minami & Brill, 1987) y (G. Abdul-Majeed, 1996). La red artificial diseñada mostró una predicción y consistencia en todo el rango experimental superior a las correlaciones de esos conjuntos de datos y ligeramente mayor a los modelos mecanicistas de (Taitel & Dukler, 1976a) y (Minami & Brill, 1987). A pesar de las mejoras conseguidas, sus autores resaltan el carácter empírico de las redes neuronales y previenen de su uso general debido a que en su formulación no se incluyen las leyes fundamentales de conservación.

Por otro lado, un aspecto relevante y común de los trabajos mencionados que han incorporado el paradigma de aprendizaje automático es lo referente a la naturaleza promedio de los parámetros de flujo de los conjuntos de datos considerados, i.e., un dato de un ensavo o prueba experimental de esos conjuntos contiene la media de los flujos volumétricos (o másicos) para cada fase además de otros valores de propiedades importantes como presión y temperatura. Para esas condiciones, la medición asociada de colgamiento consiste de igual manera en un único punto experimental. Este tipo de recopilación de datos, si bien tiene como objetivo cubrir una gama amplia de propiedades de fluidos así como de configuraciones de tubería, prescinde de la información que aportan las fluctuaciones de esas mismas propiedades de fluidos y de flujo en el dominio del tiempo y del espacio. En ese sentido, para el mejor conocimiento del autor, no se han reportado estudios que aborden el problema de predecir las fluctuaciones instantáneas de colgamiento de líquido para una configuración particular de flujos másicos (o velocidades superficiales) utilizando redes neuronales artificiales. Por su conceptualización y a diferencia de los estudios de (Osman, 2004) y (Shippen & Scott, 2004), el problema de estimación instantánea que se investiga en el presente trabajo no tiene el objetivo de alcanzar el número de combinaciones de flujos másicos de los trabajos publicados, sino en analizar para un rango de pruebas particular, la forma en que se puede predecir el colgamiento de líquido instantáneo en una tubería horizontal con el modelo de perceptrón multicapa. Esta formulación no compromete el trabajo en cuanto al orden datos que requieren los modelos de aprendizaje automático, ya que con el muestreo de las señales efectuado se está en el orden de 10^3 puntos por experimento, que es incluso mayor que el total de puntos que consideraron (Osman, 2004) y (Shippen & Scott, 2004).

De lo anterior, se tiene que el objeto de estudio de la presente investigación es el análisis de un conjunto de familias de intermitencias hidrodinámicas correspondientes a un flujo tapón cuya fase líquida es de alta viscosidad⁴. Como tal, la motivación por estudiar el flujo

⁴A la fecha, debido a la relativa novedad en el estudio de este tipo de crudos, no se dispone de un valor de corte estándar para determinar si un crudo es de "alta viscosidad". En todo caso, para el mejor conocimiento

altamente viscoso se debe a la relevancia que posee el crudo pesado en el panorama energético de nuestro país⁵; no obstante, a pesar de su relevancia, el estudio y entendimiento de este tipo de crudo supone un desafío considerando que la mayor parte de la investigación para flujos líquido-gas, como se ha mostrado en párrafos anteriores⁶, se ha desarrollado principalmente para condiciones de baja viscosidad.

El escenario previamente comentado resulta problemático a causa de que el flujo gas-líquido de alta viscosidad exhibe un comportamiento sustancialmente diferente al de su contraparte menos viscosa, lo cual resulta en que los modelos desarrollados para baja viscosidad no sean directamente aplicables a este nuevo caso (Foletti et al., 2011) y (Farsetti et al., 2014). Por ejemplo, se ha encontrado que a medida que aumenta la viscosidad de la fase líquida, mapas como los de (Taitel & Dukler, 1976b) y (Taitel et al., 1980) predicen inadecuadamente los patrones de flujo; dicha discrepancia es debida a que en condiciones de alta viscosidad existe un reemplazo del flujo estratificado y ondulado por parte del flujo intermitente y anular, respectivamente (Matsubara & Naito, 2011). Así mismo, para el caso del flujo tapón, se ha reportado que la frecuencia de "slugs" aumenta proporcionalmente a la viscosidad de la fase líquida (Gokcal et al., 2008), (Gokcal et al., 2010), (Okezue, 2013) y (Baba et al., 2017). De igual manera, en condiciones de alta viscosidad se ha observado un aumento en el nivel de la película de líquido (E. M. Al-Safran et al., 2013) y en la región del cuerpo del tapón (E. Al-Safran et al., 2015). No obstante, para el flujo tapón altamente viscoso se ha identificado una reducción en la longitud promedio de los bloques de líquido respecto a las longitudes típicas de tapón que se presentan en baja viscosidad (E. M. Al-Safran et al., 2013) y (Baba et al., 2018); otros problemas relativos a este tipo de flujo son las mayores caídas de presión durante su transporte, la precipitación de parafinas y asfaltenos, su contenido de sal y la corrosión inducida en las tuberías (Martínez-Palou et al., 2011), (H. Zhang & Lan, 2017) y (Zhai et al., 2018).

De la descripción anterior, resulta evidente la importancia por desarrollar herramientas que permitan predecir parámetros clave del flujo bifásico como el colgamiento de líquido en condiciones de alta viscosidad, siendo de especial interés la predicción de este parámetro en términos de variables que típicamente se puedan medir en campo. Ahora bien, dado que la presión es la propiedad a la que normalmente se tiene acceso en escenarios reales, el presente trabajo busca aprovechar el poder predictivo del aprendizaje automático para correlacionar las fluctuaciones de presión en línea y colgamiento de líquido de un flujo tapón de alta viscosidad.

Definido el objeto de estudio y la metodología de análisis a utilizar, lo siguiente es tomar como referencia los estudios previos sobre las fluctuaciones de la fracción de vacío (Akagawa,

del autor, el menor valor de viscosidad para la fase líquida, cuyos autores consideran "alta viscosidad" es 0.587 Pa·s (ver (Gokcal et al., 2008), (Kora et al., 2011) y (E. Al-Safran et al., 2015)).

 $^{{}^{5}}$ Ejemplo de ello es el 52.4 % que aportó el crudo pesado en lo que respecta a la producción total nacional de hidrocarburos líquidos en el año 2024 (PEMEX, 2025).

⁶Incluidos los trabajos que han utilizado el paradigma de aprendizaje automático.

1964) y de presión en línea (Montiel Cortés, 2017) para formular la siguiente hipótesis de relación:

1.1. Hipótesis

1. Para una combinación particular y aproximadamente constante de un flujo másico de una fase líquida y un flujo másico de una fase gaseosa en la entrada de un sistema, del flujo bifásico resultante aguas abajo, existe una relación entre el comportamiento hidrodinámico de la serie instantánea de colgamiento de líquido con el comportamiento hidrodinámico de la serie instantánea de presión cuando los puntos de medición de ambas series se encuentran a una distancia de tres metros el uno del otro y entre ambos puntos no existe un cambio en el diámetro interior de la tubería o una modificación en el ángulo de inclinación de la misma.

Esta hipótesis a su vez, produce las siguientes dos subhipótesis:

- 1A. Los principales componentes de frecuencia del espectro de la serie de colgamiento de líquido son significativos en el espectro de la serie de presión.
- 1B. Dado que los perceptrones multicapa se consideran aproximadores universales de funciones (Cybenko, 1989), es posible obtener un perceptrón que relacione los valores de la serie instantánea de presión con los valores de la serie instantánea de colgamiento de líquido. Una vez aprendidos los patrones subyacentes entre ambas variables, se espera reconstruir nuevas series instantáneas de colgamiento de líquido a partir de valores de presión desconocidos.

Una vez enunciadas las hipótesis, los objetivos son los siguientes:

1.2. Objetivos

- 1. Investigar los espectros de frecuencia de las series temporales de presión y colgamiento de líquido.
- 2. Construir un perceptrón multicapa con enfoque regresivo para predecir valores de colgamiento de líquido a partir de valores de presión en línea de un flujo tapón de alta viscosidad.

- 3. Agrupar los valores instantáneos de la serie de colgamiento de líquido utilizando el modelo de mezcla Gaussiano (MMG).
 - 3.1 Estimar los volúmenes totales y acumulados por región de flujo de cada una de las fases con el MMG.
- 4. Diseñar un conjunto de perceptrones multicapa de tipo regresivo para predecir el colgamiento en términos de presión por región de flujo una vez agrupados los datos de colgamiento de líquido.

1.3. Organización del documento

La estructura de este trabajo se enucia a continuación:

En el capítulo dos se describen brevemente los conceptos fundamentales de flujo tapón y de los dos modelos de aprendizaje automático requeridos para abordar el problema de predicción planteado. En el capítulo tres se proporciona la información experimental relativa al conjunto de datos; además, se efectúa un estudio preliminar de las series de tiempo de interés. En el capítulo cuatro se analizan los resultados al aplicar el modelo de mezcla Gaussiano y de perceptrón multicapa al problema de predección de colgamiento de líquido instantáneo en términos de presión. Finalmente en el capítulo cinco, se enlistan las conclusiones, trabajo futuro, recomendaciones y contribuciones obtenidas de la investigación realizada.

2. MARCO TEÓRICO

2.1. Conceptos de flujo tapón

En la producción y transporte de hidrocarburos, es común la ocurrencia de flujo multifásico. Dentro del flujo bifásico gas-líquido en tuberías horizontales y con poco ángulo de inclinación, un patrón de flujo recurrente es el flujo tapón¹, el cual se define como un flujo intercalado de tapones de líquido de alta velocidad con intervalos estratificados de burbujas de gas que se mueven sobre una capa de líquido de menor velocidad (O. Nydal et al., 1992), (Andreussi et al., 1993), (E. Al-Safran, 2009a) y (Hernandez et al., 2018). Dada su naturaleza, el flujo tapón forma parte de la familia de flujo intermitente².

A nivel operativo, el flujo tapón resulta conflictivo a raíz del estrés mecánico que genera en las líneas así como a la paradas de planta asociadas a su comportamiento (Hill & Wood, 1994); otros inconvenientes que acarrea son la corrosión inducida así como los desbordamientos una vez que los bloques de líquido llegan a las instalaciones superficiales (Arabi et al., 2020). No obstante, el flujo tapón es especialmente relevante en el marco del flujo bifásico gas-líquido debido a la relación que presenta con las configuraciones de fases restantes, lo cual se ve reflejado en que el flujo tapón ocupe una posición central en todos los mapas de patrones de flujo (H.-Q. Zhang et al., 2003).

Desde el punto de vista predictivo, el flujo tapón presenta desafíos importantes a causa de su estructura dinámica, la cual conlleva a que parámetros característicos del tren de tapones como frecuencia, velocidad traslacional y colgamiento de líquido exhiban distribuciones esta-

 $^{^1{\}rm Conocido}$ también en español como flujo bache. En el idioma inglés a este flujo se le denomina "slug flow".

²Otra configuración de fases de la familia intermitente es el flujo de burbuja elongada ("plug flow", por su nombre en inglés). La diferencia entre el flujo tapón y el de burbuja elongada es que en este último la cantidad de aire contenida en los tapones es despreciable, de allí que este patrón de flujo se encuentre a menores velocidades superficiales de gas.

9

dísticas distintas en localizaciones separadas de la tubería. Además, la inestabilidad inherente de este patrón de flujo resulta en que no posea periodicidad en el tiempo ni en el espacio (Bendiksen et al., 1996).

A su vez, dependiendo del mecanismo de formación, los flujos tapón se pueden clasificar en hidrodinámicos o por configuración de terreno³ (Ujang et al., 2006). El taponamiento hidrodinámico se genera principalmente por las inestabilidades en la interfase líquido-gas a causa de la diferencia de velocidades entre las fases; por su parte, el taponamiento por configuración de terreno resulta de la acumulación de líquido en las inclinaciones de la tubería con topografía variable (Ujang et al., 2006). En todo caso, en las líneas multifásicas industriales los bloques de líquido de ambos mecanismos pueden interactuar formando un tren de tapones altamente complejo (Issa & Kempf, 2003).

Por otro lado, en el estudio del flujo tapón (principalmente en la vertiente de modelado mecanicista) se hace uso de un concepto denominado como unidad de tapón⁴, la cual se define como la combinación de un tapón de líquido en adición de una región estratificada que comprende la película de esta misma fase y de la burbuja de gas que está por encima de ella (E. Al-Safran, 2009b) y (H.-Q. Zhang et al., 2000). No obstante, en las observaciones de series experimentales las secuencias burbuja-tapón incluyen además del cuerpo y película de líquido, una región conocida como frente de mezcla o de mezclado así como una zona de desprendimiento de líquido ubicada en la parte trasera del tapón⁵. Una descripción de una unidad de tapón de un flujo glicerina-aire considerando las regiones adicionales de flujo mencionadas se proporciona a continuación.

 $^{^{3}}$ El fenómeno de taponamiento severo (o "severe slugging" por su nombre en inglés) es un ejemplo de este tipo de mecanismo de formación.

 $^{^4\}mathrm{Concepto}$ introducido por (Dukler & Hubbard, 1975).

 $^{^5\}mathrm{A}$ esta región de desprendimiento de líquido también es conocida como "cola del tapón."



Figura 2.1: Unidad de tapón de un flujo gas-líquido de alta viscosidad. Fuente: IINGEN-UNAM.

En la Fig.(2.1) se ilustran las distintas regiones de flujo de una unidad de tapón de alta viscosidad. En la figura, el orden alfabético de los incisos coincide con la dirección del flujo; las burbujas pequeñas observadas en las figuras 2.1a y 2.1b están adheridas a la pared de la tubería.

En la Fig.(2.1a) se visualiza la región estratificada de la unidad de tapón, la cual comprende una capa o película de líquido (zona inferior de la tubería) y la región de burbuja principal⁶(parte superior). En esta región estratificada, debido a su menor viscosidad dinámica la fase gaseosa fluye a mayor velocidad que la fase líquida. A medida que la diferencia de velocidades se incrementa, aumentan las inestabilidades en la interfase gas-líquido.

Por su parte, la Fig.(2.1b) ilustra a la región del frente de mezcla. Como tal, las burbujas presentes en esta región de flujo provienen de la burbuja principal y se generan a causa del esfuerzo cortante entre el frente de alta velocidad y la capa de líquido que viaja más lentamente; durante el proceso de mezcla, las nuevas burbujas pueden fragmentarse o bien formar burbujas más grandes (coalescencia). Así mismo, las burbujas generadas pueden recircular a la burbuja principal o viajar hacia el cuerpo del tapón (proceso de aireamiento), este último mecanismo es llevado a cabo por vórtices que a su vez son resultado de la mayor energía cinética que posee el frente del tapón; adicionalmente, para flujos de alta viscosidad, se ha observado otro mecanismo de arrastre donde el frente captura burbujas grandes en un proceso similar a un "tragamiento"⁷ (E. M. Al-Safran et al., 2013). Dicho proceso es causado por un

⁶También denominada como burbuja de Dumitrescu o burbuja de Taylor en honor a sus primeros investigadores: (Dumitrescu, 1943) y (Davies & Taylor, 1950).

⁷Tal mecanismo los autores lo denominan como "swallowing".

doblamiento o plegamiento⁸ de una capa superior del frente del tapón que se mueve a mayor velocidad provocando que se adelante respecto a las capas inferiores del frente. (E. Al-Safran et al., 2015). El proceso mencionado se ilustra en la figura 2.2.



Figura 2.2: Mecanismo de arrastre de burbujas por plegamiento del frente del tapón. Fuente: IINGEN-UNAM.

En cuanto a los bloques de líquido (Fig.2.1c), estos adquieren su volumen y longitud a causa de todo el líquido barrido por el frente de mezcla. Las burbujas provenientes de la zona de mezclado que logran penetrar al cuerpo del tapón pueden ascender a la parte alta de la tubería debido a la flotabilidad que poseen. Dependiendo de la velocidad del tapón, dichas burbujas pueden transicionar en forma de burbujas grandes a burbujas dispersas distribuidas a lo largo de todo el bloque de líquido (Dascher, 1968). Además, tanto las burbujas presentes en el tapón como su líquido contenido son simultáneamente desprendidos en su parte trasera (Fig.2.1d). Estas burbujas salientes se pueden incorporar en la burbuja principal o en la película de líquido de la siguiente unidad de tapón. Respecto al líquido desprendido, este se une por segregación gravitacional a la siguiente capa de líquido.

2.1.1. Colgamiento de líquido

En el estudio del flujo bifásico, uno de los conceptos más importantes es el parámetro que se conoce en la jerga petrolera como colgamiento de líquido. Como tal, este parámetro cuantifica el área ocupada por la fase líquida en un tiempo y posición determinados respecto al área total transversal al flujo correspondiente al área interna de la tubería. Matemáticamente se define como:

$$H_L = \frac{A_L}{A_{tub}} \tag{2.1}$$

En cuanto a sus aplicaciones, el colgamiento de líquido es requerido para estimar el gradiente

⁸Los autores lo nombran "front-folding".

de presión bifásico y para cerrar y resolver los sistemas de ecuaciones planteados por las formulaciones mecanicistas asociadas al flujo tapón; así mismo, la aplicación del concepto de colgamiento de líquido a una unidad específica de tapón produce lo que se conoce como colgamiento de líquido por región de flujo, tal evaluación se ilustra a continuación:



Figura 2.3: Colgamiento de líquido evaluado en cada región de flujo de la unidad de tapón. Fuente: IINGEN-UNAM.

2.2. Preliminares de aprendizaje automático

El aprendizaje automático⁹ es una rama de la Inteligencia Artificial¹⁰ dirigida al desarrollo de modelos y algoritmos computacionales capaces de identificar patrones así como de realizar predicciones y toma de decisiones de un conjunto de datos con la menor intervención humana posible (Ortiz, 2024). El aprendizaje automático se refiere también a la capacidad de un sistema de adquirir su propio conocimiento a partir de los datos que le fueron suministrados (Bengio et al., 2017) y (Brunton, 2021). Dependiendo de la naturaleza en la formulación de un determinado modelo de aprendizaje automático, este puede ser de tipo estadístico, o bien, basado puramente en datos (empírico).

A las aplicaciones donde se proporcionan ejemplos de entrenamiento con sus correspondientes respuestas correctas, se denominan problemas de aprendizaje supervisado (Bishop & Nasrabadi, 2006) y (Marsland, 2011). El objetivo en este tipo de enfoque es minimizar el error entre la respuesta correcta proveniente del conjunto de datos y el valor predicho por el modelo (Liu & Mehta, 2019). Si las predicciones del modelo resultante son precisas respecto a datos no vistos, se dice que el modelo generaliza adecuadamente.

Dependiendo del tipo de variable que se desea predecir, los problemas de aprendizaje supervisado pueden ser de clasificación o regresivos si la variable objetivo se puede agrupar en un número finito de categorías o si la salida deseada es número real, respectivamente (Bishop & Nasrabadi, 2006) y (Müller & Guido, 2016).

Por su parte, a las aplicaciones en donde sólo se proporcionan ejemplos de entrenamiento pero ningún valor objetivo o respuesta correspondiente se denominan problemas de aprendizaje no supervisado. El objetivo en este enfoque es inferir grupos similares en el conjunto de datos (agrupamiento), determinar la distribución de los datos (estimación de densidad) o proyectar los datos desde un espacio de alta dimensión a uno de menor dimensión con el propósito de visualizarlos (Bishop & Nasrabadi, 2006).

Por otro lado, para cubrir los objetivos planteados en la sección 1.2 se requiere hacer uso de modelos de aprendizaje supervisado (perceptrón multicapa) así como de aprendizaje no supervisado (modelo de mezcla Gaussiano). Una panorámica de la clasificación dichos modelos respecto al aprendizaje automático se ilustra en el siguiente diagrama.

⁹O "Machine Learning" por su nombre en inglés.

¹⁰La Inteligencia Artificial (IA) puede definirse como la búsqueda por automatizar tareas que típicamente realizan los humanos (Bellman, 1978) y (Chollet, 2021). Para más información sobre la IA y otras de sus ramas puede consultarse (Brewka, 1996).



Figura 2.4: Clasificación de los modelos a utilizar dentro del aprendizaje automático. Fuente: Autor.

Los detalles del modelo de perceptrón multicapa y del modelo de mezcla Gaussiano se describen en las secciones siguientes.

2.3. Conceptos de aprendizaje probabilístico

El aprendizaje probabilístico se puede entender como el estudio de la aplicabilidad de métodos probabilísticos a tareas que involucran automatización bajo incertidumbre¹¹.

Dentro de los intereses que han motivado el desarrollo del aprendizaje probabilístico, se encuentra la creación de sistemas inteligentes que sean robustos ante cambios inesperados en la distribución de los datos que le son suministrados y el diseño de sistemas que sean capaces de aprender rápidamente utilizando sólo pequeñas cantidades de datos (Murphy, 2023).

Como tal, el campo de estudio del aprendizaje probabilístico abarca desde modelos orientados a la predicción, como los procesos Gaussianos y las redes neuronales Bayesianas, hasta modelos generativos como los autocodificadores variacionales y las redes generativas adversarias. Otro tipo de formulación cubierto por este tipo de aprendizaje son los modelos de

¹¹La definición expuesta está formulada a partir de lo que comenta (Pearl, 1988) sobre el alcance que tiene el campo de estudio que el autor denomina como razonamiento probabilístico.

factores latentes, cuyo caso especial son los modelos de variables latentes.

2.3.1. Modelos de variables latentes y modelos de mezcla

Un modelo de variable latente se define como cualquier modelo probabilístico en el que algunas de sus variables no son directamente observables¹²(Murphy, 2023). Dentro de las formulaciones de variables latentes, un modelo que ha encontrado gran aplicabilidad en diversas disciplinas (McLachlan et al., 2019) es el modelo de mezcla, que en la formulación de variables latentes se representa como sigue (Murphy, 2022):

$$p(\mathbf{x}) = \sum_{k} p(z=k) p(\mathbf{x}|z=k)$$
(2.2)

En la ecuación 2.2, \mathbf{x} es un vector *d*-dimensional de entradas con *n* elementos. Por su parte, *z* es una variable latente que indica cuál componente de mezcla se usa para generar \mathbf{x} ; por su naturaleza, *z* se distribuye categóricamente (Murphy, 2023), es decir:

$$z \sim p(z) \sim \operatorname{Cat}(z)$$
 (2.3)

El término p(z = k) de la ecuación 2.2 se deriva del hecho de como z es desconocida priori, el valor k que puede tomar es incierto.

Por parte, el término $p(\mathbf{x}|z=k)$ representa algún tipo distribución de probabilidad discreta o continua; dependiendo de los requerimientos, dichas distribuciones pueden ser todas discretas o todas continuas o en el caso más general, una combinación de ambas¹³. No obstante, debido a su aplicabilidad, son de especial interés los modelos de mezcla cuyo término $p(\mathbf{x}|z=k)$ involucra a familias de distribuciones exponenciales.

Por otro lado, el modelo de mezcla por definición se expresa como (Bouguila & Fan, 2020) y (Murphy, 2023):

$$p(\mathbf{x}|\boldsymbol{\theta}) = \sum_{k=1}^{K} \lambda_k p_k(\mathbf{x})$$
(2.4)

En la ecuación 2.4, los escalares λ_k se denominan pesos de mezcla (o simplemente pesos) del modelo, los cuales cumplen con $0 \le \lambda_k \le 1$ y $\sum_k \lambda_k = 1$. Por su parte, el término $p(\mathbf{x}|\boldsymbol{\theta})$

 $^{^{12}\}mathrm{De}$ allí la connotación de variable oculta, no observable o latente.

¹³Para más información sobre modelos de mezcla finitos, puede consultarse (Everitt, 1988).

representa la probabilidad de generar \mathbf{x} a partir de $\boldsymbol{\theta}$. Como tal, $\boldsymbol{\theta}$ contiene los pesos del modelo así como cada uno de los parámetros de cada una de las distribuciones p_k .

Ahora bien, comparando los términos equivalentes para las dos formulaciones del modelo de mezcla (ecuaciones 2.2 y 2.4) se tiene que:

$$\begin{cases} p(z=k) = \lambda_k \\ p(\mathbf{x}|z=k) = p_k(\mathbf{x}) = p(\mathbf{x}|\theta_k) \end{cases}$$
(2.5)

En este punto, como el modelo de mezcla contiene tanto variables observables como no observables, resulta conveniente reformularlo como un modelo jerárquico. Tal jerarquía se ilustra en la figura siguiente.



Figura 2.5: Relación jerárquica entre las variables de un modelo de mezcla. Fuente: Autor.

En la figura 2.5, la flecha entre dos variables indica dependencia de la variable representada con más tabulación respecto a la variable con menor tabulación. Por tanto, con la figura en cuestión y con base en el conjunto de ecuaciones de 2.5, se obtiene (Murphy, 2023):

$$\begin{cases} p(z|\boldsymbol{\theta}) = \operatorname{Cat}(z|\lambda) \\ p(\mathbf{x}|z,\boldsymbol{\theta}) = p(\mathbf{x}|\boldsymbol{\theta}_k) \end{cases}$$
(2.6)

Ahora, utilizando el conjunto de expresiones dado por 2.6, las ecuaciones 2.2 y 2.4 resultan en:

$$p(x|\boldsymbol{\theta}) = \sum_{k=1}^{K} p(z=k,\boldsymbol{\theta}) p(\mathbf{x}|z=k,\boldsymbol{\theta})$$
(2.7)

$$p(x|\boldsymbol{\theta}) = \sum_{k=1}^{K} \lambda_k p_k(x|\boldsymbol{\theta}_k)$$
(2.8)

La ecuación 2.8 representa a un modelo de mezcla con k componentes¹⁴, cada uno de ellos con parámetros $\boldsymbol{\theta}_k$.

Adicionalmente, cuando cada una de las distribuciones p_k de la ecuación 2.8 se distribuyen de manera normal, al modelo resultante se le conoce como modelo de mezcla de distribuciones Gaussianas^{15–16} para el vector *d*-dimensional de entradas **x** (Sugiyama, 2015), (Bouguila & Fan, 2020) y (Murphy, 2023), esto es:

$$p(\mathbf{x}|\boldsymbol{\theta}) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(\mathbf{x}|\mu_k, \Sigma_k)$$
(2.9)

Donde $\mathcal{N}(\mathbf{x}|\mu, \Sigma)$ se define como (Bonald, 2019):

$$\mathcal{N}(\mathbf{x}|\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^d det(\Sigma)}} exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$
(2.10)

En la expresión 2.9 el parámetro $\boldsymbol{\theta} = (\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$. A su vez, los vectores $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma}$ contienen el peso, la media y la matriz de covarianza¹⁷ para cada uno de los k componentes del modelo de mezcla. Respecto a la ecuación 2.10, el término *det* hace referencia al determinante de $\boldsymbol{\Sigma}$.

Como tal, el MMG de la ecuación 2.9 es utilizado en el esquema de aprendizaje no supervisado como un método de agrupamiento de puntos $x_i \in \mathbb{R}^d$ (Murphy, 2023). La ventaja que proporciona el MMG respecto a otras técnicas, es que la agrupación que efectúa se lleva a cabo de manera probabilística, esta característica es especialmente útil en problemas donde es incierta la asignación de un punto x_i a un determinado grupo. Otro escenario frecuente de aplicación del MMG es para aproximar distribuciones que exhiben un comportamiento multimodal¹⁸.

¹⁴O también denominados *modos* del modelo de mezcla.

¹⁵O simplemente como modelo de mezcla de Gaussianas, Gaussiano o de Gauss, por su forma más abreviada.

¹⁶Designado en español por las siglas (MMG). En el idioma inglés sus siglas correspondientes son (GMM) las cuales hacen referencia a "Gaussian Mixture Model".

 $^{^{17}}$ Cuando ${\bf x}$ es de dimensión uno, la contraparte de Σ es la varianza y se representa con σ^2 para una mejor distinción respecto al caso de mayor dimensión.

¹⁸De hecho, cuando el número de componentes de un MMG es lo suficientemente grande el modelo es capaz de aproximar cualquier distribución suave sobre \mathbb{R}^d (Murphy, 2023).
Precisada la definición así como los usos del MMG, resta describir la metodología de cálculo de $\boldsymbol{\theta}$. Tal procedimiento se expone en la siguientes dos subsecciones.

2.3.2. Máxima verosimilitud para el MMG

Una de las técnicas más utilizadas para la estimación puntual de parámetros es el método de máxima verosimilitud. La función de verosimilitud para una muestra con n elementos de una variable aleatoria $X \sim f(\mathbf{x}|\boldsymbol{\theta})$ donde $\boldsymbol{\theta}$ es un parámetro desconocido se define como (Montgomery & Runger, 2010):

$$\ell(\boldsymbol{\theta}) = \prod_{i=1}^{n} p(x_i | \boldsymbol{\theta})$$
(2.11)

O bien, en su forma logarítmica:

$$\log \ell(\boldsymbol{\theta}) = \sum_{i}^{n} p(x_i | \boldsymbol{\theta})$$
(2.12)

Sustituyendo la definición del MMG (ecuación 2.9) en 2.12 se obtiene:

$$\log \ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left(\sum_{k=1}^{K} \lambda_k \mathcal{N}(x_i | \mu_k, \Sigma_k) \right)$$
(2.13)

De la expresión 2.13, el estimador que maximiza la función de log-verosimilitud es:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} \log \ell(\boldsymbol{\theta}) \tag{2.14}$$

El cálculo de $\hat{\theta}$ para la ecuación 2.13 no es trivial, de allí que su estimación se lleve a cabo de forma numérica. El algoritmo específicamente diseñado para la obtención de $\hat{\theta}$ en problemas de variables latentes se conoce como algoritmo E-M¹⁹ y se explica en la siguiente subsección.

¹⁹Presentado por (Dempster et al., 1977).

2.3.3. Algoritmo E-M para el MMG

El término p(z = k) de la ecuación 2.2 se entiende como una probabilidad *a priori* de que z tome algún valor k. En todo caso, para la tarea de agrupamiento, lo que se desea calcular es la probabilidad de asignar algún grupo k a algún valor $x_i \in \mathbf{x}$; como x_i sí es observale, la probabilidad buscada es $p(z = k | x_i, \boldsymbol{\theta})$. Para el cálculo de esta nueva probabilidad se puede hacer uso de la regla de Bayes (Murphy, 2023), obteniendo:

$$\gamma_{ik} = p(z = k | x_i, \boldsymbol{\theta}) = \frac{p(z = k, \boldsymbol{\theta}) p(x_i | z = k, \boldsymbol{\theta})}{\sum_{k'=1}^{K} p(z = k' | \boldsymbol{\theta}) p(x_i | z = k', \boldsymbol{\theta})}$$
(2.15)

Es decir:

$$\hat{\gamma}_{ik} = \frac{\hat{\lambda}_k \mathcal{N}(x_i | \hat{\mu}_k, \hat{\Sigma}_k)}{\sum_{k'=1}^K \hat{\lambda}_{k'} \mathcal{N}(x_i | \hat{\mu}_{k'}, \hat{\Sigma}_{k'})}$$
(2.16)

Al término $\hat{\gamma}_{ik}$ de la ecuación 2.16 se le conoce como responsabilidad del grupo k sobre el punto x_i . En la ecuación referida, la suma del denominador posee el contador k', donde $k' \neq k$ debido a que k' proviene del cálculo de la probabilidad total de la regla de Bayes.

Por su parte, el algoritmo E-M es una metodología iterativa de estimaciones de máxima verosimilitud para problemas cuyas observaciones pueden considerarse datos incompletos (Dempster et al., 1977). Como tal, el algoritmo comprende dos etapas o pasos²⁰: el paso relativo a la "esperanza" (paso E) y el paso de "maximización" (paso M). En el paso E, se calcula la probabilidad de pertenencia de un punto $x_i \in \mathbf{x}$ a un grupo k; esta probabilidad está determinada por $\hat{\gamma}_{ik}$. Para el paso M, se recalculan las estimaciones de los parámetros del modelo de mezcla para de maximizar la log-verosimilitud. Matemáticamente, el algortimo se expresa como sigue (Sugiyama, 2015).²¹:

Para el paso E:

$$\hat{\gamma}_{ik} = \frac{\hat{\lambda}_k \mathcal{N}(x_i | \hat{\mu}_k, \hat{\Sigma}_k)}{\sum_{k'=1}^K \hat{\lambda}_{k'} \mathcal{N}(x_i | \hat{\mu}_{k'}, \hat{\Sigma}_{k'})}$$
(2.17)

 $^{^{20}}$ Previo a una inicialización aleatoria de los parámetros $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\Sigma}.$

²¹Para mayor información sobre el algoritmo E-M aplicado al MMG puede consultarse (Guzmán et al., 2024). Aproximaciones distintas a la máxima verosimilutud para el MMG pueden revisarse en (Ng, 2000) y (Sahu & Babu, 2020).

A su vez, para el paso M:

$$\hat{\gamma}_{k} = \frac{1}{n} \sum_{i=1}^{n} \hat{\gamma}_{ik} \\
\hat{\mu}_{k} = \frac{\sum_{i=1}^{n} \hat{\gamma}_{ik} x_{i}}{\sum_{k'=1}^{K} \hat{\gamma}_{ik}} \\
\hat{\Sigma}_{k} = \frac{\sum_{i=1}^{n} \hat{\gamma}_{ik} (x_{i} - \hat{\mu}_{k}) (x_{i} - \hat{\mu}_{k})^{T}}{\sum_{k'=1}^{K} \hat{\gamma}_{ik}}$$
(2.18)

Los pasos E y M se repiten hasta alcanzar la convergencia.

Por otro lado, para llevar a cabo el ajuste de un MMG, además de hacer uso del algoritmo E-M se requiere precisar a priori el número k de componentes del modelo de mezcla. Para dicha especificación se recurre a fórmulas que han sido derivadas en el contexto de selección de modelos múltiples o en competencia²². Como tal, la idea detrás de estas fórmulas o criterios es comparar el ajuste proporcionado por un determinado modelo respecto a la cantidad de parámetros que contiene. Dichos criterios se definen como sigue.

Criterio de Akaike²³:

$$AIC = -2\log(\hat{\ell}) + 2k \tag{2.19}$$

Por su parte, el criterio de Bayes se define $como^{24}$:

$$BIC = -2\log(\ell) + k\log(n) \tag{2.20}$$

Para las ecuaciones 2.19 y 2.20, $\hat{\ell}$ representa la estimación de la máxima verosimilitud para un MMG con k componentes; en el criterio de Bayes, n hace referencia al tamaño de la muestra **x**. Para ambas ecuaciones, el término positivo representa una penalización sobre el número k de componentes que posee un modelo; el término negativo hace referencia al ajuste de la log-verosimilitud para ese modelo en cuestión.

Ahora bien, en las ecuaciones referidas observar que para el criterio de Akaike, el término de

²²Para más información sobre la selección de modelos y la extensión del método de máxima verosimilitud, puede consultarse (Kullback & Leibler, 1951) y (Akaike, 1973).

²³Propuesto por (Akaike, 1974).

 $^{^{24}\}ensuremath{\mathrm{Formulado}}$ por (Schwarz, 1978).

penalización es lineal. Por su parte, el término de penalización para el criterio de Bayes toma en cuenta tanto el número de componentes como el tamaño de la muestra. Tal consideración resulta en que para el criterio de Bayes se seleccionen modelos con un menor número de componentes respecto a la selección efectuada con el criterio de Akaike.

Con lo comentado acerca de los criterios para la selección de un MMG óptimo, se cubre lo referente a los conceptos de aprendizaje probabilístico. Los conceptos sobre aprendizaje profundo se describen a continuación.

2.4. Conceptos de aprendizaje profundo

El aprendizaje profundo es una subrama del aprendizaje automático que estudia la representación de características de un conjunto de datos por medio de capas sucesivas que involucran la composición de múltiples transformaciones no lineales (Bengio et al., 2013); en el aprendizaje profundo, las representaciones en cuestión se aprenden a través de modelos conocidos como redes neuronales (Chollet, 2021). Por otro lado, el término *profundo* en este tipo de aprendizaje hace referencia al número de capas ocultas de una red neuronal; en las aplicaciones de aprendizaje profundo, el número de capas ocultas puede ir desde decenas a cientos o miles de ellas (He et al., 2016); dichos modelos, además, pueden incluir miles o millones de hiperparámetros (Krizhevsky et al., 2012).

2.4.1. Modelo de una neurona artificial

El concepto de neurona artificial es el pilar de todo el aprendizaje profundo. Su modelo matemático se ilustra en la Fig.(2.6).



Figura 2.6: Modelo matemático de una neurona artificial. Adaptado de (Hagan et al., 1997).

En la figura 2.6 se representa al modelo de una neurona artificial de una muestra con R entradas o características. Los valores para esa muestra se almacenan en el vector \mathbf{x} . Por su parte, al vector \mathbf{w} que contiene los diferentes pesos $w_{1,1}, w_{1,1}, \ldots, w_{1,R}$ se denomina vector de

pesos; el primer subíndice en **w** indica hacia qué neurona se realiza la conexión, el segundo subíndice se refiere a cuál entrada ha generado la señal²⁵. A su vez, al término *b* se le llama sesgo de la neurona. El símbolo Σ encuadrado es un sumador. A la literal *n* se le conoce como entrada de la red (neurona en este caso). El símbolo ϕ dentro del recuadro denota a la función de activación, la cual limita el rango de amplitud permisible de la señal a un valor finito (Haykin, 2009); la literal *a* indica la salida de la neurona.

El proceso de funcionamiento del modelo de una neurona artificial es el siguiente: de izquierda a derecha, la neurona recibe las señales o estímulos y su contenido se almacena en \mathbf{x} . Los valores de dicho vector se multiplican con los correspondientes elementos de \mathbf{w} ; esta operación hace referencia a "los enlaces de conexión o de sinapsis" (Haykin, 2009). Posteriormente, al sumador llega el producto anterior así como la entrada del sesgo de la neurona. Después, la suma total se asigna a n y esta entra como argumento de la función de activación. Finalmente, la señal modificada se asigna a la salida de la neurona.

De lo anterior descrito, la salida del modelo de una neurona artificial matemáticamente se expresa como:

$$a = \phi \left(\mathbf{w} \cdot \mathbf{x} + b \right) \tag{2.21}$$

La función de activación ϕ se elige con base en los requerimientos de algún problema en específico (Liu & Mehta, 2019). No obstante, en la práctica del aprendizaje profundo moderno la recomendación estándar es utilizar la función ReLU²⁶ (Nair & Hinton, 2009), (Jarrett et al., 2009), (Glorot et al., 2011) y (Bengio et al., 2017). De lo anterior mencionado, el presente estudio considera las funciones de activación siguientes.

ϕ	Definición				
Escalón	$\int a = 0 si n < 0$				
	$\begin{bmatrix} a=1 & si & n \ge 0 \end{bmatrix}$				
Lineal	a = n				
Sigmoide	$a = \frac{1}{1 + e^{-n}}$				
Tanh	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$				
ReLU	a = max(0, n)				

Tabla 2.1: Funciones de activación.

 $^{^{25}}$ Por ejemplo, el término $w_{1,2}$ se lee como el peso hacia la neurona uno proveniente de la entrada uno.

 $^{^{26}}$ En todo caso, a excepción de la función escalón (debido a las restricciones que presenta en su proceso de diferenciación), las funciones de activación restantes de la Tabla 2.1 se incluyen en el espacio de búsqueda de hiperparámetros del presente trabajo.

2.4.2. Modelo de perceptrón multicapa

En los problemas reales se requieren modelos que contengan más de una neurona con el objetivo de aumentar su poder predictivo. Una forma de lograrlo es apilando un serie de neuronas con algún tipo de disposición. Cuando este acomodo se lleva a cabo en paralelo se genera lo que se conoce como una capa de neuronas (Hagan et al., 1997)²⁷. Esquemáticamente se representa como sigue:



Figura 2.7: Modelo de una capa de neuronas. Adaptado de (Hagan et al., 1997)

En la figura 2.7 se ilustra el modelo de una capa de una muestra con R entradas y Q neuronas. Observar que cada neurona posee su propio vector de pesos, sesgo, sumador, función de activación, entrada y salida correspondiente. Lo anterior resulta en que ahora se tenga una matriz de pesos \mathbf{W} a diferencia del vector \mathbf{w} del caso de la sección 2.4.1. De igual manera, los componentes que en el caso anterior eran escalares, ahora pasan a ser vectores. Así mismo, $\boldsymbol{\phi}$ contiene cada una de las funciones de activación de la capa, las cuales en el caso más general pueden ser todas distintas unas de otras.

De lo comentado previamente, se tiene que la salida asociada a la capa de neuronas puede

 $^{^{27}}$ En particular, a la capa cuyas neuronas reciben el vector de entradas se le denomina *capa de entrada*; a las capas internas que son invisibles desde el punto de vista de las entradas de la red y cuyas salidas son las entradas de otras capas se les conoce como *capas ocultas*; por su parte, a la capa cuyas salidas son las salidas del perceptrón se le conoce como *capa de salida*.

ser descrita por la siguiente ecuación.

$$\mathbf{a} = \boldsymbol{\phi} \left(\mathbf{W} \mathbf{x} + \mathbf{b} \right) \tag{2.22}$$

La matriz W de la ecuación 2.22 tiene la forma siguiente (Hagan et al., 1997).

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,R} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,R} \\ \vdots & \vdots & & \vdots \\ w_{Q,1} & w_{Q,2} & \cdots & w_{Q,R} \end{bmatrix}$$
(2.23)

Ahora bien, notar que cada una de las filas de W contiene la información de la conexión de todas las entradas con los pesos correspondientes para alguna neurona en particular. Así mismo, cada una de las columnas de W almacena los pesos de todas las neuronas para una determinada entrada.

Por otro lado, el concepto de perceptrón multicapa surge de añadir capas adicionales al modelo de una capa de neuronas. El arreglo resultante se ilustra en la Fig.(2.8).



Figura 2.8: Modelo de perceptrón multicapa. Adaptado de (Hagan et al., 1997)

En la figura 2.8 se visualiza el esquema de un perceptrón con tres capas ocultas. Observar que cada capa posee su matriz de pesos \mathbf{W} , su vector de sesgos \mathbf{b} , su vector de entradas de cada una de sus neuronas \mathbf{n} , así como sus correspondientes funciones de activación y su vector de salidas \mathbf{a} para esas neuronas. Para diferenciar tales elementos entre una capa y otra se ha optado por la notación de superíndices²⁸ (Hagan et al., 1997). En todo caso, la idea de

²⁸Por ejemplo, el término $w_{1,1}^1$ se lee como el peso hacia la neurona uno de la primera entrada para la

disponer capas sucesivas se logra haciendo que la entrada para alguna capa m corresponda con la salida de la capa anterior. Lo anterior matemáticamente se representa como:

$$\mathbf{a}^{\mathbf{m}} = \boldsymbol{\phi}^{\mathbf{m}} \left(\mathbf{W}^{\mathbf{m}} \mathbf{a}^{\mathbf{m}-1} + \mathbf{b}^{\mathbf{m}} \right) \tag{2.24}$$

En particular, la salida dada por la ecuación 2.24 para el caso de un perceptrón con tres capas ocultas resulta de la siguiente manera.

$$\mathbf{a}^3 = \boldsymbol{\phi}^3 \left(\mathbf{W}^3 \mathbf{a}^2 + \mathbf{b}^3 \right) \tag{2.25}$$

Es decir:

$$a^{3} = \phi^{3} \left(W^{3} \phi^{2} (W^{2} \phi^{1} (W^{1} x + b^{1}) + b^{2}) + b^{3} \right)$$
(2.26)

Definido el modelo, el esquema así como la salida correspondiente a un modelo de perceptrón multicapa, resta describir los conceptos referentes a su entrenamiento. Tal información se proporciona en las subsecciones siguientes.

Etapas de un modelo de aprendizaje supervisado, error de generalización y conjuntos de datos

Independientemente de si el problema de aprendizaje supervisado a resolver es de tipo regresivo o de clasificación, el flujo de trabajo del problema en cuestión engloba secciones bien diferenciadas, las cuales típicamente se identifican en etapa de preprocesamiento de datos, de entrenamiento, de validación y de prueba del modelo.

Como tal, la etapa de preprocesamiento de datos se refiere a la práctica de ajustar la representación de las características de los datos para que estas sean más adecuadas a los algoritmos de entrenamiento (Müller & Guido, 2016). A esta sección también se le conoce como etapa de extracción de características (Bishop & Nasrabadi, 2006). Su importancia radica en que vuelve más eficiente y estable el procesamiento posterior de los datos.

En el presente estudio, para el preprocesamiento del conjunto experimental se utilizó el escalador que toma en cuenta el valor máximo y mínimo de alguna característica en particular. Su expresión se muestra a continuación (adaptada de (Pedregosa et al., 2011)):

primera capa.

$$x_{esc_i} = \left(\frac{x_i - x_{min}}{x_{max} - x_{min}}\right) (max - min) + min$$
(2.27)

En la expresión 2.27, x_i se refiere al valor *i* de la característica **x**. x_{min} y x_{max} son el valor mínimo y máximo de **x**, respectivamente. En cuanto a los términos min y max, estos se refieren al mínimo y máximo deseados en el escalamiento. Además, se menciona que se hizo uso de este escalador debido a que conserva la distribución original de los datos, la cual es una característica útil para algoritmos que dependen de la distancia relativa entre sus puntos, tal como el descenso de gradiente.

Por otro lado, continuando con la descripción del flujo de trabajo de un modelo de aprendizaje supervisado, se señala que la etapa de entrenamiento se refiere al proceso de ajustar los pesos y sesgos del modelo seleccionado haciendo uso de algún algoritmo de optimización. Referente a la etapa de validación, en ella se lleva a cabo la obtención de un conjunto de hiperparámetros óptimos mediante algún algoritmo de búsqueda. Una vez que el modelo ha sido entrenado y que ha sido configurado con sus hiperparámetros óptimos, en la etapa de prueba se efectúa la evaluación de su rendimiento final.

Ahora bien, considerando que el objetivo en el aprendizaje supervisado es producir un modelo que reporte el menor error ante datos no vistos (error de generalización) (Bengio et al., 2017), es imperativo²⁹ que para cada etapa de su desarrollo se disponga de diferentes conjuntos de datos. En particular, si \mathcal{T} se refiere al conjunto de datos para la etapa de entrenamiento y \mathcal{V} y \mathcal{T} a los conjuntos para las etapas de validación y de prueba, respectivamente, se debe cumplir que $\mathcal{T} \cap \mathcal{V} \cap \mathcal{T} = \emptyset$. Adicionalmente, los elementos de cada uno de los conjuntos mencionados deben cumplir el requerimiento de provenir de una repartición aleatoria y balanceada del conjunto experimental original.

Descrita una panorámica de las etapas de trabajo para un modelo de aprendizaje supervisado, en los apartados siguientes se exponen los conceptos relativos al entrenamiento del modelo de perceptrón multicapa.

Función de pérdida

La función de pérdida³⁰ es una medida cuantitativa del rendimiento de un modelo en el conjunto de datos de entrenamiento (Chollet, 2021). Su importancia radica en que permite guiar al algoritmo de descenso de gradiente en la dirección correcta. La función de pérdida más ampliamente utilizada para problemas de regresión es el error cuadrático medio (ECM³¹)

²⁹De hecho el propio concepto de error de generalización hace referencia implícita a un conjunto de datos distinto al utilizado en la etapa de entrenamiento.

 $^{^{30}\}mathrm{También}$ conocida como función de costo.

³¹O "Mean Squared Error (MSE)" por su nombre y siglas en inglés.

(Liu & Mehta, 2019) y (A. Zhang et al., 2021). Matemáticamente se define como:

$$ECM = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
(2.28)

En la ecuación 2.28, n indica el número de datos experimentales o muestras. Por su parte, y_i se refiere al valor experimental de la muestra $i \ge \hat{y}_i$ a la predicción correspondiente efectuada por el modelo seleccionado.

El ECM por definición, amplifica los errores que son más grandes en magnitud, lo cual deriva en que penaliza de forma más severa las desviaciones grandes respecto a las pequeñas. Otra función de pérdida utilizada para problemas de regresión es el error medio absoluto (EMA³²). En todo caso, dado su uso estándar, el presente trabajo sólo contempla el ECM como función de pérdida para el problema de predicción de colgamiento señalado en la sección 1.2.

Entrenamiento del modelo de perceptrón multicapa

El proceso de entrenamiento de un modelo de aprendizaje supervisado es clave dentro de su flujo de trabajo ya que es en esta etapa donde el modelo en caso de ser entrenado adecuadamente, adquiere su poder predictivo. Para el caso del perceptrón multicapa, el proceso de entrenamiento parte de la siguiente ecuación de actualización de parámetros (Adaptado de (Hagan et al., 1997)):

$$\mathbf{W}^{\mathbf{m}}(k+1) = \mathbf{W}^{m}(k) + \eta \,\mathbf{z}_{\mathbf{k}} \; ; \; m = 0, 1, \dots, M-1 \tag{2.29}$$

$$\mathbf{b}^{\mathbf{m}}(k+1) = \mathbf{b}^{m}(k) + \eta \,\mathbf{z}_{\mathbf{k}} \, ; \, m = 0, 1, \dots, M-1$$
(2.30)

Las ecuaciones 2.29 y 2.30 indican la forma de calcular los nuevos valores (representados por k+1) de la matriz de pesos **W** y el vector de sesgos **b** para alguna capa m. Además, para ambas ecuaciones η representa la tasa de aprendizaje³³ y $\mathbf{z}_{\mathbf{k}}$ alguna dirección de búsqueda para la iteración k. Para la iteración k = 0, **W** y **b** se inicializan de manera aleatoria.

Ahora bien, el aspecto clave de las dos ecuaciones anteriores está relacionado en cómo se designa a $\mathbf{z}_{\mathbf{k}}$, ya que a partir de su elección tiene que ser posible la minimización de alguna

³²O "Mean Absolute Error (MAE)" por su nombre y siglas en inglés.

³³La tasa de aprendizaje se vincula con el tamaño del paso en cada una de las iteraciones. Además, su valor afecta la estabilidad y rapidez en el proceso de búsqueda de algún mínimo de **F**. A pesar de su importancia, dado que η es un hiperparámetro (ver subsección 2.4.2), no se dispone de una metodología precisa de su estimación para funciones arbitrarias.

función de pérdida **F**. Como tal, el enfoque usual³⁴ para dicha designación es utilizar el concepto de gradiente de una función con el objetivo de aprovechar la propiedad del descenso más rápido³⁵ de los vectores que este operador proporciona y de esta manera estimar algún mínimo local de **F**. A la técnica mencionada se le conoce como descenso de gradiente y su ecuación de actualización para algún vector de parámetros $\boldsymbol{\theta}$ es la siguiente.

$$\boldsymbol{\theta}(k+1) = \boldsymbol{\theta}(k) - \eta \nabla \mathbf{F}$$
(2.31)

La ecuación 2.31 aplicada específicamente a la matriz de pesos \mathbf{W} y al vector de sesgos \mathbf{b} de alguna capa m resulta en:

$$\mathbf{W}^{\mathbf{m}}(k+1) = \mathbf{W}^{m}(k) - \eta \nabla \mathbf{F} \; ; \; m = 0, 1, \dots, M-1$$
(2.32)

$$\mathbf{b}^{\mathbf{m}}(k+1) = \mathbf{b}^{m}(k) - \eta \nabla \mathbf{F} \; ; \; m = 0, 1, \dots, M-1$$
(2.33)

En los perceptrones multicapa, la no linealidad de las funciones de activación y la relación compleja entre sus pesos y el error correspondiente dificultan un cálculo explícito de $\nabla \mathbf{F}$ en las ecuaciones 2.32 y 2.33. A la técnica específicamente diseñada para estimar numéricamente a $\nabla \mathbf{F}$ se le conoce como algoritmo de retropropagación³⁶. Dicho algoritmo fue desarrollado con base en el principio de la regla de la cadena del Cálculo Diferencial³⁷.

Como tal, el funcionamiento del algoritmo de retropropagación engloba dos etapas (la etapa o paso hacia adelante y el paso hacia atrás per sé). Para la primera de ellas, cada muestra x_i^{38} del conjunto de entrenamiento pasa por todas las capas ocultas del perceptrón hasta que llegar a la capa de salida, donde se obtiene su correspondiente predicción \hat{x}_i . Posteriormente, se calcula el error entre x_i y \hat{x}_i . Después, el error recién calculado se propaga en sentido inverso, lo cual permite el ajuste de los pesos y sesgos de la capa de salida; con estos dos nuevos valores, se actualizan los pesos y sesgos de las capas ocultas. Al estado en donde al menos una vez, para cada muestra x_i se ha completado tanto el paso hacia adelante como hacia atrás del algoritmo se conoce como época.

³⁴Otros enfoques utilizan el gradiente conjugado o bien diferenciación de segundo orden, como los métodos de Newton y de Levenberg-Marquardt. Para más información referente a Optimizacion Numérica puede consultarse (Kelley, 1995), (Dennis Jr & Schnabel, 1996), (Nocedal & Wright, 1999), (Fletcher, 2000), (Sun & Yuan, 2006) y (Bonnans et al., 2006).

³⁵Específicamente, se sabe del cálculo vectorial que el descenso más pronunciado para una función \mathbf{F} está dado por $-\nabla \mathbf{F}$.

³⁶O "backpropagation" por su nombre en inglés.

³⁷El primer desarrollo del algoritmo de retropropagación fue reportado por (Werbos, 1974) y retomado posteriormente en los trabajos de (LeCun, 1985) y (Rumelhart et al., 1986).

³⁸A esta forma de procesar individualmente las muestras x_i se le denomina "descenso de gradiente estocástico", ya que la muestra x_i es elegida al azar. Otras formas de procesar las x_i consisten en analizar todo el conjunto de entrenamiento o bien dividirlo por "lotes".

Con lo anterior mencionado, se tiene que las ecuaciones del algoritmo de retropropagación son las siguientes³⁹ (Adaptado de (Hagan et al., 1997)).

Para el paso hacia adelante:

$$\begin{cases} \mathbf{a}^{0} = \mathbf{x} \\ \mathbf{a}^{m+1} = \boldsymbol{\phi}^{m+1} (\mathbf{W}^{m+1} \mathbf{a}^{\mathbf{m}} + \mathbf{b}^{m+1}) ; \ m = 0, 1, \dots, M - 1 \\ \mathbf{a} = \mathbf{a}^{M} \end{cases}$$
(2.34)

Referente al paso hacia atrás del algoritmo, las ecuaciones son⁴⁰:

$$\begin{cases} \mathbf{s}^{M} = -2\left(\frac{d\mathbf{a}^{M}}{d\mathbf{n}_{M}}\right)\left(x_{i} - \hat{x}_{i}\right) \\ \mathbf{s}^{m} = \left(\frac{d\mathbf{a}^{M}}{d\mathbf{n}_{M}}\right)\left(\mathbf{W}^{m+1}\right)^{T}\mathbf{s}^{m+1} ; \ m = M - 1, \dots, 2, 1. \end{cases}$$

$$(2.35)$$

Por otro lado, se comenta que la ecuación de actualización mostrada en 2.31 es una forma elemental de ajustar los parámetros de θ vía descenso de gradiente. Se conocen como optimizadores a los algoritmos que modifican la ecuación 2.31 con el objetivo de acelerar el proceso de búsqueda de un mínimo de **F** así como para incrementar la estabilidad durante dicho proceso. Se comenta que para el desarrollo de los perceptrones multicapa del presente trabajo se ha elegido el optimizador Adam^{41, 42} debido a la mayor estabilidad y rapida convergencia que ha mostrado en una gran cantidad de aplicaciones.

Otro elemento relevante dentro de la etapa de entrenamiento para algún modelo de aprendizaje profundo es su regularización⁴³ la cual se define como cualquier modificación efectuada

 $^{^{39}}$ El superíndice mse refiere a las capas ocultas y de entrada y el superíndice M hace referencia a la capa de salida del perceptrón.

 $^{^{40}}$ El término **s** se conoce como sensibilidad y se refiere al cambio de la función de pérdida respecto a la activación de una neurona en una determinada capa del perceptrón. En la literatura se puede encontrar también con el símbolo δ .

⁴¹Cuyas siglas provienen de su nombre en inglés: "Adaptative Moment Estimation."

⁴²Ideado por (Kingma & Ba, 2014).Otros optimizadores utilizados ampliamente son AdaDelta, propuesto por (Zeiler, 2012); RMSprop dado a conocer por (Tieleman & Hinton, 2012) y AdaGrad, formulado por (Duchi et al., 2011).

⁴³La noción de regularización se ha generalizado a lo largo del tiempo (Michelucci, 2018). Inicialmente, la se definió como un término de penalización a la función de pérdida (Bishop, 1995).

en un algoritmo de aprendizaje para lograr que el modelo seleccionado generalice de mejor manera (Kukačka et al., 2017) reduciendo el error en el conjunto de prueba (error de generalización), pero no en el conjunto de entrenamiento (Bengio et al., 2017). Como tal, la idea de regularización se originó para lidiar con el problema de sobreajuste⁴⁴, el cual se refiere a la condición cuando un modelo construido se ajusta demasiado a las particularidades del conjunto de entrenamiento pero no generaliza adecuadamente ante nuevos datos (Müller & Guido, 2016). Una forma práctica de investigar la condición de sobreajuste es revisar si $\mathbf{F}|_{\mathcal{T}} \ll \mathbf{F}|_{\mathcal{P}}$ para el modelo entrenado. No obstante, es altamente recomendable monitorear el comportamiento de la función de pérdida en los conjuntos de entrenamiento y de validación durante todo el proceso de entrenamiento.

Por otra parte, dentro de la variedad de técnicas empleadas para evitar el problema de sobreajuste⁴⁵, el presente trabajo considera el uso de la parada temprana, la cual consiste en detener el entrenamiento cuando la función de pérdida en el conjunto de validación ha alcanzado un valor mínimo (Michelucci, 2018).

Ahora bien, con la etapa de entrenamiento descrita, a continuación se cubre el tópico referente a optimización de hiperparámetros, que es esencial dentro de la etapa de validación de un modelo de aprendizaje supervisado.

Optimización de hiperparámetros

La tarea de buscar conjuntos óptimos de hiperpárametros posee gran relevancia en todo el aprendizaje automático supervisado, debido a la influencia que tienen en el rendimiento del modelo de aprendizaje, es decir, en su capacidad predictiva ante datos no vistos. (Pinto et al., 2009), (Bergstra et al., 2011) y (Hutter et al., 2019).

Específicamente, un hiperparámetro se define como cualquier variable de un modelo o de un algoritmo de aprendizaje cuyo valor no está determinado directamente durante el proceso de entrenamiento⁴⁶. Por otro lado, se denomina optimización de hiperparámetros al problema de encontrar valores adecuados para cada uno de los hiperparámetros de un modelo seleccionado. (Bergstra et al., 2011).

Como tal, la búsqueda de hiperparámetros inicialmente se llevaba a cabo de manera manual. Como es de esperar, esta forma sólo cubría una porción muy pequeña del espacio de búsqueda además de no ser susceptible de reproducibilidad en los resultados.

Posteriormente, una de las primeras alternativas a la búsqueda manual fue la búsqueda por

⁴⁴Para más información sobre el concepto de sobreajuste, puede consultarse (Ying, 2019).

⁴⁵Las técnicas de regularización más utilizadas son L_1 y L_2 y descarte aleatorio (o "Dropout" por su nombre en inglés). Esta última técnica fue presentada por (Srivastava et al., 2014).

⁴⁶El prefijo *hiper* en la palabra hiperparámetro puede interpretarse como una referencia a que el valor de ese parámetro está "más allá de los datos", en este caso, de los datos de entrenamiento.

malla, que resultó ser altamente ineficiente (Aggarwal et al., 2018). No obstante, se ha reportado que a la fecha, la búsqueda aleatoria por malla proporciona una estimación aceptable en los conjuntos de datos en los que se ha utilizado (Bergstra et al., 2011). Más recientemente, debido al incremento en la complejidad así como en la cantidad de modelos de aprendizaje automático, se ha visto necesario disponer de técnicas que sean computacionalmente eficientes y que además cubran un espacio de búsqueda lo suficientemente grande. En esa dirección, un procedimiento específicamente diseñado para la optimización de hiperparámetros es el algoritmo de hiperbanda (Li et al., 2018), el cual es una extensión del algoritmo de reducción sucesiva presentado por (Jamieson & Talwalkar, 2016). Otra técnica comúnmente utilizada es la búsqueda Bayesiana (Snoek et al., 2012), (Eggensperger et al., 2013), (Thornton et al., 2013) y (Garnett, 2023). En todo caso, se comenta que para los objetivos y alcances del presente trabajo, el método seleccionado para la optimización de hiperparámetros de los perceptrones multicapa considerados es la búsqueda aleatoria por malla.

A continuación, se resume la metodología que se seguirá en el presente estudio para el desarrollo de alguna red neuronal.

Tabla 2.2: Procedimiento utilizado para la construcción de un modelo de perceptrón multicapa.

Paso	Descripción
1	Definir un modelo de RNA
2	Definir un espacio de búsqueda de hiperparámetros $ ightarrow \mathcal{S}$
3	Dividir $\{(x_i, y_i)\}$ para obtener \mathcal{T}, \mathcal{V} y \mathcal{P}
4	Realizar el escalamiento respectivo para cada conjunto de datos del paso anterior
5	Seleccionar un algoritmo de búsqueda de hiperparámetros $ ightarrow \mathbf{ALG}$
6	Encontrar un conjunto óptimo de hiperparámetros en ${\cal S}$ utilizando ${\cal V}$
7	Construir un hipermodelo a partir del conjunto óptimo de hiperparámetros encon-
	$\mathrm{trado} \to \mathbf{H}\mathbf{M}$
8	Entrenar ${f HM}$ utilizando ${\cal T}$
9	Evaluar rendimiento final de ${f HM}$ utilizando ${\cal P}$

3. ESTUDIO PRELIMINAR DEL CONJUNTO DE DATOS

El presente capítulo introduce al conjunto de datos a estudiar: su procedencia, las combinaciones de flujos másicos analizadas así como el remuestreo efectuado en las señales; se incluye también una visualización de las variables de flujo de interés. Además, se expone una descripción de la instrumentación utilizada y el análisis de incertidumbre de las variables monitoreadas durante los experimentos.

3.1. Pruebas bifásicas gas-líquido de alta viscosidad

El conjunto de datos proporcionado¹ formó parte de una campaña experimental de pruebas bifásicas gas-líquido de alta viscosidad llevada a cabo en el circuito de flujos multifásicos del Instituto de Ingeniería de la UNAM.

Actualmente, la sección de pruebas² tiene una extensión de 50 metros medidos desde el punto de inyección (d) hasta la desembocadura en las tolvas de almacenamiento (f). (ver figura 3.1)

 $^{^{1}}$ El conjunto experimental se generó de manera previa a la incorporación del autor al equipo de trabajo.

²Para más detalles sobre el circuito experimental consultar (Hernández, 2019).



Figura 3.1: Circuito de Flujos Multifásicos IINGEN-UNAM. Adaptado de (Hernández, 2019)

Como tal, la campaña contempló la realización de tres experimentos, en los cuales el flujo másico promedio de la fase líquida se estableció con valores de 0.6, 1.3 y 2.0 kg/s, respectivamente. El flujo másico promedio de la fase gaseosa se estableció, para todos los experimentos, con un valor de 5×10^{-3} kg/s. Las combinaciones de flujos másicos mencionadas se resumen en la tabla siguiente:

Tabla 3.1: Conjunto experimental analizado.

	η	$\dot{n}_L (kg/s)$	s)
$\dot{m}_G \left(kg/s \right)$	0.6	1.3	2.0
0.005	G1L1	G1L2	G1L3

Cada experimento de la Tabla 3.1 se muestreó en tres tramos de datos de 180 segundos resultando en un total de 540 segundos por experimento. La tasa de adquisición fue de 7 Hz. Relativo a la configuración de la sección de pruebas, la campaña se llevó a cabo en una tubería horizontal de tres pulgadas (0.0762 metros) de diámetro nominal.³

Respecto a las sustancias utilizadas, se empleó aire para la fase gaseosa ($\mu_G = 1.815 \times 10^{-5} \text{ Pa} \cdot \text{s}$); las densidades del aire medidas a 18°C y a la presión del transductor uno (ver figura 3.1) fueron de 1.409, 1.887 y 2.368 kg/m³ para los experimentos G1L1, G1L2 y G1L3, respectivamente. Por su parte, para la fase líquida se empleó glicerina ($\rho_L = 1260 \text{ kg/m}^3$, $\mu_L = 0.6 \text{ Pa} \cdot \text{s}$ medida a 18°C).

 $^{^{3}}$ Dichas especificaciones corresponden a una tubería de acero al carbón sin costura (cédula 80).

Las variables monitoreadas durante la campaña de pruebas fueron las siguientes:

No.	Variable
1	Colgamiento de líquido.
2	Flujo másico de la fase líquida y de la fase gaseosa.
3	Presión en cuatro puntos distintos de la sección de pruebas.
4	Temperatura de los fluidos y ambiental.
5	Viscosidad dinámica de la fase líquida.

A excepción de la viscosidad de líquido, todas las variables mencionadas fueron medidas en tiempo real. Lo anterior resultó en la generación de diez series de tiempo para el análisis posterior.

La instrumentación requerida para evaluar cada parámetro de la Tabla 3.2 se describe a continuación.

3.1.1. Instrumentación empleada

Colgamiento de líquido

El equipo de medición seleccionado para monitorear al colgamiento de líquido en la presente campaña experimental fue un sistema de tomografía modelo m3c, marca ITS. El sistema está compuesto de un sensor en línea (tomógrafo) (ver figura 3.2), un módulo de adquisición de datos y un software para configurar a los dispositivos según se requiera. Se optó por este tipo de instrumentación debido a que ofrece ventajas como ausencia de radiación, respuesta rápida, funcionamiento para altas presiones y temperaturas (Ismail et al., 2005), además de contar con la característica de que no es invasiva al flujo.⁴

La forma de operación del tomógrafo consiste en medir, en píxeles, las permitividades de los fluidos de la sección transversal al flujo en un plano de la tubería. Posteriormente, mediante un algoritmo de reconstrucción⁵, se obtienen las imágenes que representan cada una de las fracciones de ocupación de las fases.

⁴Esto es particularmente importante en flujos de naturaleza inestable como el flujo tapón (Taitel & Dukler, 1977). Cualquier perturbación externa puede alterar considerablemente la dinámica del flujo y con ello las lecturas de medición posteriores.

⁵Típicamente el algoritmo LBP (Hansen et al., 2019).





Figura 3.2: Tomógrafo en línea. Fuente: Autor.

Flujos másicos

Los flujos másicos de cada una de las fases se cuantifican con medidores tipo Coriólis marca Endress + Houser, modelo Promass 83 F, cada uno en su respectiva línea de flujo (ver figura 3.3). Ambos se encuentran colocados en la sección de inyección del circuito.





(a) Coriólis, fase gaseosa. Fuente: Autor. (b) Coriólis, fase líquida. Fuente: Autor.

Figura 3.3: Medidores de flujos másicos para cada una de las fases.

Presión

La medición de la presión fue efectuada utilizando transductores industriales de tipo piezométrico serie U5300 de alta precisión de la marca TE Connectivity de 15 psi (ver figura 3.4). Durante la campaña experimental, se contó con cuatro transductores dispuestos a lo largo de la sección de pruebas (ver figura 3.1). Todos los transductores están colocados en la parte superior de la tubería.



Figura 3.4: Transductor de presión. Fuente: Autor.

La disposición de los transductores fue diseñada de tal manera que fuera posible mesurar la presión en zonas clave de la sección de pruebas, por ejemplo, el primer transductor se utiliza para monitorear la presión próxima al punto de inyección; por su parte, el tercero evalúa la presión posterior al tramo curvo de la sección; el cuarto, reporta la presión medida cerca de la salida hacia las tolvas de almacenamiento. Caso especial es el del segundo transductor, el cual fue colocado estratégicamente cercano al tomógrafo. La distancia entre el transductor dos y el tomógrado es de tres metros (ver figura 3.5).



Figura 3.5: Distancia entre el segundo transductor de presión y el sensor de tomografía. Fuente: Autor.

Además, tomando como referencia la ubicación del primer transductor, la distancia de los transductores restantes es la siguiente:

Tabla 3.3: Distancia entre transductores de presión en la sección de pruebas.

L	(m)
$Trd_1 - Trd_2$	17
$Trd_2 - Trd_3$	4
$Trd_3 - Trd_4$	20
Total	41

Temperatura

Para la medición de temperatura se utilizaron termopares tipo K (ver figura 3.6); dos de ellos se encuentran instalados en la sección de pruebas (ver figura 3.1) e inspeccionan la temperatura de los fluidos. El termopar restante está colocado en la sección de adquisición de datos y reporta la temperatura ambiente.





 (a) Termopar colocado en la sección de pruebas.
 (b) Termopar colocado en la sección de ad-quisición de datos. Fuente: Autor.

Figura 3.6: Termopares tipo K.

Viscosidad de líquido

La viscosidad de la fase líquida fue evaluada con un viscosímetro marca Brookfield, modelo DV2T (ver figura 3.7). Dadas las condiciones ambientales en el laboratorio durante las pruebas, la viscosidad dinámica de la fase líquida resultó en un valor promedio de 0.6 Pa · s medido a $18^{\circ}C^{6}$.

 $^{^{6}}$ Dado que la glicerina es un fluido newtoniano, se prescindió en el reporte de medición de la información sobre la configuración de las revoluciones por minuto.



Figura 3.7: Viscosímetro Brookfield DV2T. Fuente: Fabricante.

3.1.2. Adquisición de datos

Las variables de la Tabla 3.2 monitoreadas con la instrumentación descrita en la subsección 3.1.1 son digitalizadas por medio de dos módulos de adquisición de datos (ver figura 3.8). El módulo que recopila las señales de presión, temperatura y flujos másicos para cada una de las fases es un adquisidor marca IMC modelo Cronos-flex que incluye su respectivo software (imc STUDIO), este tipo de adquisidor es de uso general. El módulo que digitaliza la señal de colgamiento de líquido es propio del sistema m3c de tomografía descrito anteriormente y su uso es específico para esta tarea.



(a) Adquidor IMC Cronos-flex. Fuente: Fabricante.



(b) Adquisidor ITS m3c. Fuente: Fabricante.

Figura 3.8: Módulos utilizados para la adquisición de datos.

3.2. Evaluación de la incertidumbre de medición

Todo trabajo que involucre el uso de datos experimentales requiere una estimación de la incertidumbre de medición. De hecho, como lo menciona (Hughes & Hase, 2010), un resultado experimental no se considera completo hasta que no se haya reportado un análisis de incertidumbre correspondiente.

La incertidumbre de medición según el (JCGM, 2008a), se refiere al parámetro no negativo

que caracteriza la dispersión de los valores atribuidos a un mensurando⁷. Por otro lado, el análisis de incertidumbre se puede llevar a cabo de manera general o detallada. El procedimiento de evaluación puede ser utilizando el método de Monte Carlo (MMC) o el método de series de Taylor (MST) (Coleman & Steele, 2018). Con base en las condiciones en las que se efectuó la presente campaña experimental, el análisis de incertidumbre realizado a cabo es de tipo general y el procedimiento utilizado es el de series de Taylor.

3.2.1. Errores aleatorios y sistemáticos

Todo sistema de medición está sujeto a diferentes fuentes de error las cuales imposibilitan conocer el valor verdadero de la magnitud de interés. Si X es un determinado mensurando, las mediciones sucesivas $\{x_1, x_2, \ldots, x_m\}$ hechas por el sistema de medición se pueden modelar como sigue (Coleman y Steele, 2018):

$$x_{1} = x_{v} + (F_{1})_{1} + (F_{2})_{1} + \dots + (F_{k})_{1}$$

$$x_{2} = x_{v} + (F_{1})_{2} + (F_{2})_{2} + \dots + (F_{k})_{2}$$

$$\vdots$$

$$x_{m} = x_{v} + (F_{1})_{m} + (F_{2})_{m} + \dots + (F_{k})_{m}$$
(3.1)

Donde x_v representa el valor verdadero del mensurando X; $(F_k)_m$ es el valor del error asociado a la fuente de error k en la medición m.

Ahora bien, (ISO, 1993), (JCGM, 2008b), (COTNNMET, 2002), (Coleman & Steele, 2018) y (Hughes & Hase, 2010) sugieren agrupar las diferentes fuentes de error con base en su variabilidad: a los errores que varían durante el proceso de medición se denominan errores aleatorios y a los errores que permanecen constantes durante las mediciones sucesivas se les conoce como errores sistemáticos. En particular, si se denota como b a la suma de las diferentes fuentes de error aleatorias⁸, la expresión para el valor de la medición m es:

$$x_m = x_v + r_m + \beta$$

$$r = \sum_{w=1}^W r_w$$

$$b = \sum_{q=1}^Q b_q$$
(3.2)

⁷El término mensurando se refiere a la magnitud que se desea medir. (JCGM, 2008a)

 $^{^{8}}b$ deriva históricamente de la palabra en inglés "bias" y r se refiere a "random".

De las expresiones en (3.2) se puede observar que el número de fuentes de error de tipo sistemático Q puede diferir de la cantidad de fuentes de error aleatorio W. Notar también que para la expresión de x_m , el parámetro β no posee el subíndice m debido a que este tipo de error, por definición, permanece constante para esa medición; caso contrario para r que sí posee el subíndice ya que los errores aleatorios varían en función de m.

Dado que las diferentes fuentes de error se agrupan dependiendo de si la fuente es sistemática o aleatoria, la evaluación de la incertidumbre contempla a su vez esta diferenciación: para las fuentes de error aleatorio la estimación de la incertidumbre se denomina **Evaluación Tipo A** y se lleva a cabo mediante análisis estadístico. Para las fuentes de error sistemático la estimación de la incertidumbre se conoce como **Evaluación Tipo B** y puede ser determinada a partir del conocimiento previo del mensurando o asumiendo distribuciones de probabilidad⁹.

3.2.2. Evaluación Tipo A de la incertidumbre

Cuando se realizan N mediciones de un mensurando X se tiene, estadísticamente hablando, una distribución muestral de los valores de X. El hecho de que se tenga un conjunto de valores posibles para un único mensurando es debido a las fuentes de error aleatorio. Dado que el valor de cada una de las fuentes de dicho error cambia de una medición a otra, la variación total se ve reflejada en forma de dispersión en los valores muestreados. Al parámetro que se utiliza para cuantificar la dispersión en una distribución muestral se conoce como desviación estándar de la muestra, y es denotado por S_X .

$$S_X \to r = \sum_{w=1}^W r_w \tag{3.3}$$

Normalidad en las fuentes de error aleatorio

Si se representan gráficamente los valores muestreados de un proceso de medición, se observará que una gran parte de las lecturas se encuentra alrededor de un valor *medio* mientras que una porción significativamente menor se encontrará alejada de dicho valor de referencia. Se ha identificado en numerosas aplicaciones (Hughes y Hase, 2010) que este comportamiento es modelado adecuadamente por una distribución normal o Gaussiana, de allí la razón de considerar a los errores aleatorios como *normalmente distribuidos*.¹⁰

⁹Funciones comunes en este tipo de evaluación son la distribución uniforme y triangular.

¹⁰Esta consideración, cuyo origen es esencialmente práctico, está sustentada por el Teorema del Límite Central (TLC), el cual establece que conforme el número de mediciones aumente durante un determinado proceso de medición, la distribución muestral se aproximará a una curva normal.

Expresiones. Evaluación Tipo A

Sabiendo que los errores aleatorios se distribuyen normalmente, las expresiones a utilizar corresponden a los parámetros de dicha distribución:

$$\overline{X} = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$S_X = \left[\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{X})\right]^{1/2}$$
(3.4)

En la expresión 3.4, \overline{X} es la media muestral y N es el tamaño de la muestra. Por su parte, para los experimentos de la Tabla (3.1) con base en lo mencionado en dicha sección, se tiene que $n_{datos \, por \, tramo} = 1260$. Dado que hay tres tramos de datos por experimento, se deduce que $N = N_{exp} = n_{tramos} \times n_{datos \, por \, tramo} = 3,780$ datos para cada experimento de algún mensurando X.

Por otro lado, debido a la dispersión en los valores muestreados, en los reportes de incertidumbre el valor de medición que se asocia al mensurando corresponde a la media de la distribución muestral. Ahora bien, dado que S_X cuantifica la dispersión muestral y no la dispersión de la media, no es posible asociarla directamente a \overline{X} . Para este cálculo se utiliza la desviación estándar de la media $S_{\overline{X}}$ que se define como:

$$S_{\overline{X}} = \frac{S_X}{\sqrt{N}} \tag{3.5}$$

Cuando la distribución muestral analizada es una composición de otras distribuciones muestrales, la ecuación 3.5 se transforma en:

$$S_{\overline{X}} = \frac{1}{N} \sqrt{\sum_{k=1}^{K} S_{X_k}^2}$$

$$(3.6)$$

Observar en la expresión 3.6 que si k = 1 se recupera la ecuación 3.5.

Una vez obtendida $S_{\overline{X}}$, el rango de valores asociado a la media \overline{X} de un mensurando X es:

$$\overline{X} \pm S_{\overline{X}} \tag{3.7}$$

La formulación dada en 3.7 es una expresión de la incertidumbre aleatoria estándar.

Para completar el reporte asociado a las fuentes de error aleatorio, es recomendable calcular un intervalo de confianza para \overline{X} . La función utilizada para estimar dicho intervalo es la distribución t (Walpole et al., 1993). Para obtener un valor de la distribución t es necesario calcular dos parámetros: α , que está asociado al nivel de confianza que se desea tener en el intervalo y ν , que se refiere a los grados de libertad de la distribución; en particular $\alpha = (1 - confianza)/2$ y $\nu = N - 1$. Además, un valor de referencia para $\alpha = 0.025$ el cual corresponde a una confianza del intervalo de 95%.

Con lo mencionado en el párrafo anterior se obtiene:

$$t_{\nu,95} S_{\overline{X}} \tag{3.8}$$

Para el conjunto experimental de la Tabla 3.1, dado que $N = N_{exp} \gg 30$, por el Teorema del Límite Central es prácticamente indistinto utilizar el valor de la distribución t o el valor z de la distribución normal estándar. De hecho, como N es constante para todos los experimentos se tiene que $t \approx z = 1.959963$.

De lo anterior, el intervalo de confianza resulta en (Coleman & Steele, 2018):

$$\operatorname{Prob}\left(\overline{X} - t_{\nu,95} S_{\overline{X}} \le \mu \le \overline{X} + t_{\nu,95} S_{\overline{X}}\right) = 0.95 \tag{3.9}$$

O bien, para el presente conjunto de datos:

$$\operatorname{Prob}\left(\overline{X} - 1.959963 \, S_{\overline{X}} \le \mu \le \overline{X} + 1.959963 \, S_{\overline{X}}\right) = 0.95 \tag{3.10}$$

3.2.3. Evaluación Tipo B de la incertidumbre

Flujos másicos

Para la evaluación del errores sistemáticos de los medidores Coriolis, se utilizó el error máximo de medición que proporciona el fabricante, el cual es de 0.10% F.S¹¹; el rango de medición de los medidores es de 0 - 5 (kg/s).

¹¹F.S. se refiere a escala completa (Full Scale).

Presión

Para los transductores de presión, a partir de las especificaciones del fabricante, la identificación de las fuentes de error sistemático fue la siguiente:

- Fuente debida a la no linealidad, histéresis y repetibilidad. Con un valor de 0.1 % F.S.
- Fuente debida a la estabilidad en el largo plazo. Con un valor de 0.1% F.S.
- Fuente debida al offset. Con un valor de $0.25\,\%$ F.S.
- Fuente debida al span. Con un valor de 0.25 % F.S.

El rango de los transductores de presión es de 0 - 15 psi.

Temperatura

Los valores de los errores asociados a fuentes sistemáticas para los termopares tipo K se tomaron como los límites de los errores estándar con base en lo que reporta el ("ASTM E230", 2017), los cuales tienen un valor de 0.75 % F.S. El intervalo de operación de este tipo de termopares es de -200 a 1250 °C.

La temperatura de trabajo durante la campaña experimental fue de 18°C.

3.2.4. Incertidumbre estándar sistemática

Una vez que las fuentes de error sistemático han sido identificadas y estimadas. La incertidumbre estándar sistemática se calcula como (Coleman & Steele, 2018):

$$\beta = \left(b_1^2 + b_2^2 + \dots b_W^2\right)^{1/2} = \left(\sum_{w=1}^W b_w^2\right)^{1/2}$$
(3.11)

3.2.5. Incertidumbre combinada y expandida

Cuando son conocidas las incertidumbres resultantes a cada tipo de fuente de error (ISO, 1993) sugiere proporcionar una incertidumbre total que considere los dos tipos de incertidumbre. A este nuevo valor se le conoce como *incertidumbre combinada* u_c y se calcula como:

$$u_c = (S_X^2 + \beta^2)^{1/2} \tag{3.12}$$

Finalmente, (ISO, 1993) recomienda determinar el valor de la incertidumbre combinada con un nivel de confianza en particular, lo cual se lograr multiplicando u_c por un factor de cobertura k. Cuando el nivel de confianza es de 95%, el factor de cobertura con base en la distribución t es aproximadamente igual a 2, por lo que la expresión para calcular la incertidumbre expandida resulta en:

$$U_{95} = 2 u_c \tag{3.13}$$

3.2.6. Consideraciones experimentales para la evaluación de la incertidumbre

Con la información generada durante la campaña experimental y con las especificaciones disponibles de los fabricantes, se evalúo la incertidumbre Tipo A y Tipo B de todas las variables de la Tabla 3.2 excepto para el colgamiento de líquido, del cual sólo se conoce su incertidumbre asociada a errores aleatorios; su incertidumbre sistemática se deja indicada.

Respecto a las unidades de los mensurandos, se reportan de la manera en que fueron configuradas originalmente en los diferentes adquisidores de datos. Las cantidades que aparecen de la Tabla 3.4 a la Tabla 3.9 se presentan, si el número resultante del cálculo computacional lo permite, con un redondeo final de cuatro cifras significativas.

3.2.7. Reporte de incertidumbre

El resumen de incertidumbre realizado para los experimentos de la Tabla (3.1) se expone a continuación:

	Errores aleatorios			Errores sistemáticos	
X	\overline{X}	S_X	$S_{\overline{X}}$	$t_{95}S_{\overline{X}}$	β
H_L	0.6486	1.308e-01	3.686e-03	7.226e-03	β_{H_L}
$\dot{m}_G \left(kg/s ight)$	4.996e-03	4.549e-04	1.281e-02	2.512 e- 02	1.249 e-05
$\dot{m}_L \left(kg/s ight)$	0.6181	1.413e-02	3.669e-04	7.194 e-04	5.000e-03
$p_1 (psi)$	5.804	6.347 e-02	1.699e-03	3.331e-03	5.511 e-02
$p_2 (psi)$	3.829	6.025 e- 02	1.666e-03	3.266e-03	5.511 e-02
$p_3 \left(psi \right)$	3.288	5.749e-02	1.588e-03	3.113 e-03	5.511 e-02
$p_4 \ (psi)$	0.9612	1.756e-02	4.909e-04	9.625 e- 04	5.511 e-02
$T_1 (^{\circ}C)$	17.88	1.414e-02	2.837 e-04	5.562 e- 04	1.350 e-01
$T_1 (°C)$	18.02	2.744e-02	5.478e-04	1.074 e- 03	1.350e-01
T_{amb} (° C)	18.06	4.476e-02	1.249e-03	2.448e-03	1.350e-01

Tabla 3.4: Evaluación Tipo A y B. Experimento G1L1.

Tabla 3.5: Evaluación Tipo A y B. Experimento G1L2.

	Errores aleatorios			Errores sistemáticos	
X	\overline{X}	S_X	$S_{\overline{X}}$	$t_{95}S_{\overline{X}}$	β
H_L	0.7167	1.662 e-01	4.683 e-03	9.181e-03	β_{H_L}
$\dot{m}_G \left(kg/s ight)$	5.037 e-03	4.120e-04	1.160e-02	2.274e-02	1.249 e-05
$\dot{m}_L \left(kg/s ight)$	1.294	1.464 e-02	3.496e-04	6.855e-04	5.000e-03
$p_1 (psi)$	11.6	6.632 e- 02	1.768e-03	3.466e-03	5.511e-02
$p_2 (psi)$	7.654	6.405 e-02	1.763 e-03	3.457 e-03	5.511 e-02
$p_3 (psi)$	6.603	6.112e-02	1.683 e-03	3.299e-03	5.511e-02
$p_4 \ (psi)$	1.995	1.877 e-02	5.155e-04	1.011e-03	5.511 e-02
$T_1 (°C)$	17.84	7.939e-03	1.372 e- 04	2.690e-04	$1.350\mathrm{e}{-01}$
$T_1 (°C)$	17.98	6.981 e- 03	1.730e-04	3.391e-04	1.350e-01
$T_{amb}(°C)$	18.0	5.716e-02	1.056e-03	2.071e-03	1.350 e-01

		Errores a	aleatorios		Errores sistemáticos
X	\overline{X}	S_X	$S_{\overline{X}}$	$t_{95}S_{\overline{X}}$	β
H_L	0.7637	1.638e-01	4.615 e-03	9.048 e-03	β_{H_L}
$\dot{m}_G \left(kg/s ight)$	5.058 e-03	4.762 e- 04	1.335e-02	2.617 e-02	1.249 e-05
$\dot{m}_L \left(kg/s ight)$	1.966	3.642 e- 02	5.707 e-04	1.119e-03	5.000 e-03
$p_1 (psi)$	17.34	7.516e-02	1.509e-03	2.958e-03	$5.511\mathrm{e}{-02}$
$p_2 (psi)$	11.49	6.351 e- 02	1.490 e-03	2.920e-03	5.511e-02
$p_3 (psi)$	9.936	6.193 e-02	1.490 e-03	2.922e-03	$5.511\mathrm{e}{-02}$
$p_4~(psi)$	3.046	1.637 e-02	3.565 e-04	6.989e-04	$5.511\mathrm{e}{-02}$
$T_1 (^{\circ}C)$	17.89	2.393e-02	1.644 e-04	3.222e-04	$1.350\mathrm{e} ext{-}01$
$T_1 (^{\circ}C)$	18.05	2.702e-02	$2.197\mathrm{e}{\text{-}}04$	4.308e-04	$1.350\mathrm{e}{ ext{-}01}$
T_{amb} (° C)	18.02	3.471 e- 02	8.968e-04	1.758e-03	$1.350\mathrm{e}{ ext{-}01}$

Tabla 3.6: Evaluación Tipo A y B. Experimento G1L3.

Tabla 3.7: Incertidumbres. Experimento G1L1.

	Incertidumbres		
X		U_{95}	
H_L	$(1.711e-02+\beta_{H_L}^2)^{1/2}$	$2(1.711e-02+\beta_{H_L}^2)^{1/2}$	
$\dot{m}_G \left(kg/s \right)$	4.551e-04	9.101e-04	
$\dot{m}_L \left(kg/s ight)$	1.499e-02	2.998e-02	
$p_1 (psi)$	8.406e-02	1.681 e-01	
$p_2 (psi)$	8.165 e-02	1.633 e-01	
$p_3 (psi)$	7.964 e-02	$1.593 \mathrm{e}{-01}$	
$p_4 (psi)$	5.784 e-02	1.157 e-01	
$T_1 (°C)$	1.357 e-01	2.715e-01	
$T_1 (°C)$	1.378e-01	2.755 e- 01	
$T_{amb}(^{\circ}C)$	1.422 e-01	2.845 e-01	

	Incertidumbres		
X	u_c	U_{95}	
H_L	$(2.761 \text{e-} 02 + \beta_{H_L}^2)^{1/2}$	$2(2.761e-02+\beta_{H_L}^2)^{1/2}$	
$\dot{m}_G \left(kg/s ight)$	4.122e-04	8.244e-04	
$\dot{m}_L \left(kg/s ight)$	1.547 e-02	3.095 e- 02	
$p_1 (psi)$	8.623 e-02	1.725e-01	
$p_2 \left(psi \right)$	8.450e-02	1.690e-01	
$p_{3}\left(psi ight)$	8.230e-02	1.646e-01	
$p_4~(psi)$	5.822 e-02	1.164 e-01	
$T_1 (^{\circ}C)$	1.352 e- 01	2.705 e-01	
$T_1 (^{\circ}C)$	1.352 e- 01	2.704 e-01	
$T_{amb} (^{\circ}C)$	1.466e-01	2.932e-01	

Tabla 3.8:Incertidumbres.Experimento G1L2.

Tabla 3.9: Incertidumbres. Experimento G1L3.

	${f Incertidumbres}$		
X		U_{95}	
H_L	$(2.682e-02+\beta_{H_L}^2)^{1/2}$	$2(2.682e-02+\beta_{H_L}^2)^{1/2}$	
$\dot{m}_G \left(kg/s ight)$	4.763e-04	$9.527\mathrm{e}{-04}$	
$\dot{m}_L \left(kg/s ight)$	3.677 e-02	7.353e-02	
$p_1 (psi)$	9.320e-02	1.864 e-01	
$p_2 \left(psi \right)$	8.409e-02	1.682 e-01	
$p_3 \left(psi \right)$	8.291 e-02	1.658e-01	
$p_4 (psi)$	5.749e-02	1.150e-01	
$T_1 (°C)$	1.371e-01	2.742e-01	
$T_1 (°C)$	1.377e-01	2.754e-01	
$T_{amb}(^{\circ}C)$	1.394 e-01	2.788e-01	

Efectuada la evaluación de incertidumbre de las pruebas de la Tabla 3.1, lo siguiente es realizar un análisis preliminar de las series de tiempo monitoreadas durante la campaña experimental.

3.3. Análisis preliminar del conjunto de datos



3.3.1. Colgamiento de líquido

Figura 3.9: Series instantáneas de colgamiento de líquido. Los gráficos (a)-(c), (d)-(f) y (g)-(i) corresponden a los experimentos G1L1, G1L2 y G1L3 de la Tabla 3.1, respectivamente. Fuente: Autor.

En la figura 3.9 se exponen las series instantáneas de colgamiento de líquido de los experimentos de la Tabla 3.1 para diferentes intervalos de tiempo. De la figura, las crestas en la serie indican el paso de los tapones de glicerina; el espacio comprendido entre cada una de esas crestas está asociado a la longitud de las burbujas de Dumitrescu; por su parte, para los tres experimentos, el colgamiento alrededor de 0.6 corresponde a la región de la película de líquido. Los valores intermedios de colgamiento hacen referencia a la región del frente de mezcla y a la cola del tapón.

Ahora bien, una comparativa del gráfico 3.9 permite visualizar que la frecuencia de los tapones de glicerina es proporcional al aumento de flujo másico de líquido; además, dicho incremento en la cantidad de líquido reduce, a su vez, la longitud de las burbujas de Dumitrescu cuya duración transiciona de alrededor de tres segundos para el experimento G1L1 a menos de un segundo para el experimento G1L3; de igual manera, se observa que a medida que aumenta la cantidad de secuencias burbuja-tapón, sus formas adquieren configuraciones más irregulares dando como resultado que los valores de colgamiento en la región de la película de líquido así como en el cuerpo del tapón presenten mayor dispersión en el experimento G1L3 respecto a los observados en la prueba G1L1.

Por otro lado, el hecho de que para alguna región de flujo se tenga una variedad¹² de valores de colgamiento asociada, abre la discusión sobre qué valor de dicha serie debe emplearse para determinar el inicio/fin de alguna región de flujo en particular. Ahora bien, como es incierto establecer *a priori* ese valor de corte, queda claro que si se desea efectuar un agrupamiento de los valores de la serie de colgamiento de líquido por región de flujo, este se tendrá que llevar a cabo necesariamente de manera probabilística.^{13,14}

 $^{^{12}}$ El término "variedad", en este contexto, se refiere a que una determinada región de flujo comprende una distribución de valores de colgamiento y no un único valor constante.

¹³Ante la complejidad de esta problemática se han propuesto valores de corte deterministas, como 0.7 (O. J. Nydal, 1991). Sin embargo, un valor de corte fijo puede llevar a identificaciones imprecisas de las zonas de tapón, especialmente para casos con alta velocidad superficial de gas.

¹⁴Notar que el agrupamiento probabilístico que se requiere es precisamente una de las aplicaciones del modelo de mezcla Gaussiano, tal como se comentó en la sección 2.3.1.

3.3.2. Flujos másicos

Flujo másico de gas



Figura 3.10: Flujo másico de la fase gaseosa. Fuente: Autor.

En la figura 3.10 se visualizan las series de tiempo de flujo másico de la fase gaseosa para los tres experimentos¹⁵. Se nota que las curvas están superpuestas debido a que el flujo másico de aire se mantiene aproximadamente constante para todo el conjunto de datos. Lo anterior se confirma observando los promedios (líneas punteadas en el gráfico) los cuales son bastante cercanos entre sí.

Por otro lado, las fluctuaciones observadas en el gráfico corresponden a la diferencia de presión que existe entre la descarga de aire en la salida de la sección de pruebas hacia las tolvas y la presión debida a la carga de aire del compresor desde su tanque de almacenamiento. Tal diferencia está presente en toda la serie de tiempo a causa de que la descarga de secuencias burbuja-tapón es continua.

¹⁵Cada experimento de la Tabla 3.1 se ilustra con un código de colores particular: azul para el caso G1L1; naranja para el caso G1L2 y verde para el caso G1L3. Este código de colores se mantiene a lo largo de todo el documento.

Flujo másico de líquido



Figura 3.11: Flujo másico de la fase líquida. Fuente: Autor.

En la figura 3.11 se presentan los valores instantáneos de flujo másico para la fase líquida así como sus respectivos valores promedio (líneas puntedas) para los tres experimentos.

En este caso, las fluctuaciones observadas son debidas al movimiento helicoidal de las cavidades generadas entre el impulsor y el estator de la bomba durante el proceso de inyección de líquido en la línea monofásica, de allí que a medida que se inyecta mayor cantidad de líquido, se incremente la amplitud de las fluctuaciones. No obstante, a pesar de dichas fluctuaciones, notar que no existen valores atípicos en la series de tiempo, lo cual confirma que el flujo másico de líquido fue aproximadamente constante para los experimentos realizados.

3.3.3. Temperatura



Figura 3.12: Comparativa entre T_1 , T_2 y $T_{amb.}$. Fuente: Autor.

La figura 3.12 ilustra una comparativa de las temperaturas de los fluidos T_1 y T_2 y de la temperatura ambiental. De la figura, se observa que la temperatura del primer termopar es ligeramente menor que la temperatura del segundo termopar y de la temperatura ambiente para un determinado experimento. En todo caso, lo relevante del gráfico es notar que las fluctuaciones de las dos mediciones de temperatura en línea son lo suficientemente pequeñas y cercanas a la temperatura ambiente (que es aproximadamente 291.15(K), 18°C). Lo anterior permite considerar que el flujo se desarrolla bajo condiciones isotérmicas.^{16, 17}

3.3.4. Investigación de la relación entre la presión y colgamiento de líquido

La hipótesis 1A expuesta en la sección 1.1, conlleva a la investigación de la densidad de potencia espectral de potencia (DPE) para la serie de presión del segundo transductor y para la serie de colgamiento de líquido. Dicha investigación está motivada también con el objetivo de evaluar una posible reconstrucción de la serie de tiempo de colgamiento de líquido a partir de una serie de Fourier con los principales componentes de frecuencia de la serie de presión. A continuación, se realiza una comparativa gráfica de la DPE de ambas series de tiempo para los tres experimentos.

 $^{^{16}}$ La razón de utilizar la temperatura ambiente como referencia es que la medición de viscosidad de la fase líquida no se efectuó en condiciones en línea sino en las condiciones de la sección de adqusición de datos, que es donde está instalado el termopar que monitorea la temperatura ambiental (ver figura 3.6).

¹⁷La consideración de flujo isotérmico sustenta la representatividad del valor constante de 0.6 Pa · s para la viscosidad dinámica de la fase líquida para toda la campaña experimental.




Figura 3.13: Espectro de potencia, EXP G1L1, $\dot{m}_L = 0.6 (kg/s)$. Fuente: Autor.



Figura 3.14: Espectro de potencia, EXP G1L2, $\dot{m}_L = 1.3 (kg/s)$. Fuente: Autor.



Figura 3.15: Espectro de potencia, EXP G1L3, $\dot{m}_L = 2.0 (kg/s)$. Fuente: Autor.

En las figuras 3.13, 3.14 y 3.15 se grafica, respectivamente, la densidad de potencia espectral para los tres experimentos de la Tabla 3.1. Para cada gráfico, el eje vertical a la izquierda (color azul) ilustra los valores de la DPE para la presión del segundo transductor; por su parte, el eje vertical a la derecha (color verde) expone los valores de la DPE para el colgamiento de líquido. Para los valores de la DPE mostrados se prescinde del componente de frecuencia cero.

Relativo al espectro de la presión del segundo transductor, se observa que para los tres experimentos, los cuatro primeros componentes de frecuencia están localizados en un intervalo menor a 0.5 Hz; de igual manera, se visualiza que el aspecto general de dicho espectro está concentrado hacia la izquierda del gráfico.

Sin embargo, para el espectro de colgamiento de líquido, se observa que su primer componente de frecuencia es mayor a 0.5 Hz para todos los experimentos; así mismo, se visualiza que sus cuatro primeros armónicos se trasladan a la derecha a medida que aumenta el flujo másico de líquido.

El aspecto relevante de esta comparativa es apreciar que un incremento aproximadamente lineal en el flujo másico de líquido no implica que el acomodo de los principales componentes de frecuencia del experimento G1L1 se mantenga y que únicamente se traslade hacia un intervalo de frecuencia más alta, sino que a medida que aumenta la cantidad de líquido en el sistema, el acomodo previo se pierde dando como resultado que los cuatro primeros armónicos presenten mayor dispersión en los experimentos siguientes. Esta mayor dispersión en los principales componentes de frecuencia está asociada por su parte, a una mayor variabilidad en la morfología de las familias de intermitencias, de allí que las secuencias burbuja-tapón del experimento G1L3 sean más disímiles entre sí que las observadas en la serie de tiempo del experimento G1L1 (ver Fig.(3.9)).

Ahora bien, dado que los principales componentes de frecuencia de la serie de colgamiento de líquido no son significativos en el espectro de la serie de presión, se descarta una posible reconstrucción de H_L en términos de los armónicos de p. De este hecho, se sigue que el análisis posterior se llevará a cabo en amplitud.





Figura 3.16: Colgamiento de líquido en términos de la presión en línea. Fuente: Autor.



Figura 3.17: Colgamiento de líquido en términos de la presión en línea (normalizada con el máximo de cada experimento). Fuente: Autor.

En la figura 3.16 se ilustra la relación, en amplitud, del colgamiento de líquido en términos de la presión de la presión en línea del segundo transductor. Para una mejor visualización, dicha relación, se grafica en la figura 3.17 en términos de la presión normalizada con el máximo de cada experimento.

Ahora bien, respecto a la figura 3.16 se observa que los valores de colgamiento de líquido cubren prácticamente todo el intervalo comprendido entre 0.57 y 1.0, siendo el experimento G1L3 el que mayor cantidad de puntos posee en la región de valores medios de colgamiento así como en la región del cuerpo del tapón.

Por su parte, en la figura 3.17 se observa con mayor detalle el acomodo característico que presenta el flujo. De dicha disposición, es posible apreciar que la mayoría de puntos para los tres experimentos caen en los valores bajos y altos de colgamiento, los cuales corresponden a la película de líquido y al cuerpo del tapón, respectivamente. De igual manera, se visualiza que el intervalo de valores de presión se restringe a medida que aumenta el flujo másico de líquido. Adicionalmente, del gráfico en cuestión se visualiza que en conjunto, los valores extremos de colgamiento presentan una estratificación proporcional al flujo másico de líquido para la región de la película e inversamente proporcional para la región del cuerpo del tapón; son estos patrones de proporcionalidad los que se pretenden capturar por medio del modelo de perceptrón multicapa.. No obstante, notar de ambos gráficos, que para algún valor p_i de presión existen múltiples valores de colgamiento asociados, i.e., como la relación entre p y H_L no cumple con la condición de unicidad¹⁸, se deduce que la relación en cuestión no es una función.¹⁹

Complementario a lo anterior, se menciona que queda descartado algún tipo de manipulación en los valores (p_i, H_{L_i}) para forzar la condición de unicidad ya que el objetivo es, precisamente, correlacionar las fluctuaciones entre ambas variables. Por último, la multiplicidad en los valores de colgamiento en términos de la presión se debe que dichas variables fluctúan a frecuencias fundamentales diferentes, tal como lo muestra la comparativa de sus espectros de la subsección 3.3.4, i.e., la relación entre p y H_L sería una función si sus frecuencias fundamentales fueran muy próximas la una a la otra; sin embargo, como lo exponen las figuras 3.13, 3.14, 3.15, 3.16 y 3.17 no es el caso.

 $^{^{18}}$ La multiplicidad de los valores de colgamiento para algún valor p_i de presión puede corroborarse fácilmente empleando la prueba de la línea recta (ver (Bartle & Sherbert, 1964)) en cualquiera de los dos gráficos anteriores.

¹⁹Para mayor información sobre la definición de función, puede consultrase (Bartle & Sherbert, 1964), (Spivak, 1967) y (Apostol, 1973).

4. ANÁLISIS Y DISCUSIÓN DE RESULTADOS

El presente capítulo expone la aplicación de los modelos de mezcla Gaussiano y de perceptrón multicapa descritos en el capítulo 2 al conjunto de datos estudiado en el capítulo 3 para abordar el problema de predicción de colgamiento de líquido instantáneo en términos de la presión en línea planteado en la sección 1.2.

4.1. Aplicación del modelo de mezcla Gaussiano a la serie instantánea de colgamiento de líquido

Para ajustar un modelo de mezcla Gaussiano a alguna serie de colgamiento se requiere estimar, como se comentó en la sección 2.3.3, el número óptimo¹ de componentes del modelo a partir de los criterios de información de Bayes y de Akaike. El número óptimo de componentes para cada experimento de la Tabla 3.1 se resume en la Tabla 4.1.

	k		
Experimento	Bayes	Akaike	
G1L1	6	13	
G1L2	7	14	
G1L3	8	12	

Tabla 4.1: Número óptimo de componentes. Criterios de Bayes y Akaike.

¹El término "óptimo" en este contexto se refiere a un número de componentes del modelo de mezcla tal que maximice la log-verosimilitud durante el proceso de ajuste y que dicho modelo además cumpla la restricción de no sobreajustarse a los datos.

En las figuras 4.1 y 4.2, se expone respectivamente, el ajuste para los criterios de Bayes y de Akaike para diversos números de componentes del modelo de mezcla. En ambos gráficos se señala, además, el número óptimo de componentes para algún experimento en particular.



Figura 4.1: Criterio de información de Bayes. Fuente: Autor.



Figura 4.2: Criterio de información de Akaike. Fuente: Autor.

Específicamente, la Fig.(4.1) muestra los valores correspondientes al criterio de información de Bayes (CIB) para diferentes modelos de mezcla de cada uno de los experimentos. Del gráfico, se visualiza que cada una de las tres curvas presentan una caída significativa en los valores del criterio para los primeros tres componentes y posteriormente, a partir del cuarto las curvas mencionadas comienzan a estabilizarse. De igual manera, se apunta que el valor k óptimo del CIB aumenta unitariamente a medida que el flujo másico de líquido se incrementa. Tal aumento en dicho valor de k refleja el hecho de que se requieren modelos de mezcla más complejos para aproximar la densidad del colgamiento cuando la cantidad de líquido en el sistema aumenta. Esta respuesta del CIB tiene sentido considerando que las secuencias burbuja-tapón son más irregulares para el experimento G1L3 que las observadas en la prueba G1L1 (ver Fig.(3.9)).

Respecto al criterio de información de Akaike (CIA), se observa de la Fig.(4.2) que al igual que en el caso anterior, las tres curvas disminuyen significativamente para los primeros tres componentes y posteriormente su comportamiento se estabiliza. No obstante, observar que el CIA alcanza el k óptimo en los últimos valores del intervalo considerado. Así mismo, visualizar que entre los k óptimos de los distintos experimentos no se distingue alguna tendencia en función del flujo másico de líquido.

Por otro lado, notar comparativamente de la Tabla 4.1 así como de las figuras 4.1 y 4.2 que el valor óptimo de k estimado por el CIA supone una diferencia de siete componentes

para los experimentos G1L1 y G1L2 y de cuatro para la prueba G1L3. Tal valor óptimo de k relativo al CIA resulta excesivo desde el punto de vista de estudio del flujo, de allí que el análisis posterior tome en cuenta el número óptimo de componentes calculado con el criterio de Bayes. Con base en la ecuación 2.9, el modelo de mezcla² para el colgamiento de líquido de cada uno de los experimentos de la Tabla 3.1 resulta en:

$$q(\boldsymbol{H}_{\boldsymbol{L}}|\boldsymbol{\theta}) = \sum_{k=1}^{K} \lambda_k \, \mathcal{N}(\boldsymbol{H}_{\boldsymbol{L}}|\,\mu_k,\sigma_k^2)$$
(4.1)

Los valores de λ_k , μ_k y σ_k para los diferentes experimentos se proporcionan a continuación:

		G1L1			G1L2			G1L3	
k	λ_k	μ_k	σ_k	λ_k	μ_k	σ_k	λ_k	μ_k	σ_k
1	0.5421	0.5858	0.0065	0.3113	0.5940	0.0062	0.2225	0.6074	0.0082
2	0.2270	0.6049	0.0116	0.2266	0.6092	0.0088	0.2161	0.6272	0.0104
3	0.0736	0.6370	0.0217	0.1177	0.6334	0.0138	0.0970	0.6614	0.0171
4	0.0314	0.7390	0.0467	0.0466	0.6881	0.0284	0.0561	0.7093	0.0218
5	0.0267	0.9331	0.0458	0.0271	0.7884	0.0401	0.0302	0.7806	0.0270
6	0.0992	0.9949	0.0022	0.0420	0.9374	0.0387	0.0333	0.8688	0.0336
7	-	-	-	0.2288	0.9898	0.0031	0.0546	0.9516	0.0201
8	-	-	-	-	-	-	0.2902	0.9809	0.0040

Tabla 4.2: Parámetros del modelo de mezcla Gaussiano para la serie de colgamiento de líquido.

En la Tabla 4.2 se exponen para cada uno de los experimentos, los pesos, medias y desviaciones estándar de cada uno de los componentes de su correspondiente modelo de mezcla Gaussiano. La tabla en cuestión está organizada³ de tal manera que los diferentes valores de kaumenten en proporción a μ . Se optó por este diseño para facilitar la lectura e interpretación de los valores que contiene así como para el futuro proceso de visualización de los parámetros del modelo de mezcla.

Ahora bien, ajustado un MMG a cada una de las series instantáneas de colgamiento de líquido, lo siguiente es llevar a cabo una identificación de los componentes del modelo de mezcla en términos de las distintas regiones de flujo de la unidad de tapón.

²En la ecuación 4.1 para evitar equívocos relativos a la literal p, se optó por denotar el término $p(\mathbf{x}|\boldsymbol{\theta})$ de la ecuación 2.9 por $q(\mathbf{x}|\boldsymbol{\theta})$.

³Por ejemplo, el último componente del modelo de mezcla para un determinado experimento corresponde con el valor de μ más grande de ese experimento.

4.1.1. Identificación de regiones de flujo

La presente sección aborda la cuestión de asociar un determinado componente del modelo de mezcla a una región de flujo en particular. Esto es especialmente importante para llevar a cabo un análisis de los resultados del ajuste del modelo en términos de variables de flujo. No obstante, dicha identificación no es trivial considerando que el número óptimo de componentes calculado con el criterio de Bayes discrepa del número de regiones de flujo que se consideran en la unidad de tapón.

Por otro lado, desde el punto de vista hidrodinámico los parámetros μ_k y σ_k del modelo de mezcla son claves para llevar a cabo la asociación anteriormente descrita. Por tal motivo, resulta conveniente analizar gráficamente el colgamiento de líquido en combinación con los valores de dichos parámetros. El proceso se describe como sigue.



Figura 4.3: Medias de cada uno de los componentes del modelo de mezcla para los diferentes experimentos. Fuente: Autor.



Figura 4.4: Desviaciones estándar de cada uno de los componentes del modelo de mezcla para los diferentes experimentos. Fuente: Autor.

En las figuras 4.3 y 4.4 se exponen, respectivamente, las medias y desviaciones estándar de los tres modelos de mezcla. De la Fig.(4.4) se observa que el último componente de cada modelo presenta el menor valor de desviación estándar para los tres experimentos, caso contrario a lo que se observa en la Fig.(4.3) en donde el último componente presenta el mayor valor de colgamiento para todos los experimentos. Regresando a la Fig.(4.4), se visualiza que el penúltimo componente para todos los modelos de mezcla presenta un aumento significativo en su valor σ correspondiente. Tal contraste entre σ_K y σ_{K-1} así como en los valores de μ_k proporcionan la pauta para establecer que el último componente para el modelo de mezcla de un determinado experimento está asociado a la región del cuerpo del tapón y se deduce a su vez, que los penúltimos componentes agrupan valores de colgamiento de la región del frente de mezcla. Esta última asignación tiene sentido físico considerando que en el frente del tapón se presentan vórtices de mezcla debido al resbalamiento entre las fases⁴, i.e., los remolinos de mezcla en el frente son capturados por el MMG con una mayor desviación estándar. Observar además, que en conjunto, las medias de la figura 4.3 y las desviaciones estándar de la figura 4.4 forman una unidad estadística de tapón.

Una vez que se han asociado el último y el penúltimo componente de los modelos de mezcla a la región del tapón y frente respectivamente, el siguiente paso es identificar los componentes restantes para las regiones de película y frente. Sin embargo, en este nuevo caso

⁴Ver figura 2.2a.

la forma de proceder es incierta considerando que en la figura 4.4 el patrón exhibido por las desviaciones estándar (en forma de "torre") para los experimentos G1L1 Y G1L2 discrepa del comportamiento triangular del experimento G1L3. Por tal motivo, se propone realizar y comparar, además de las dos figuras anteriores, un gráfico adicional donde se ilustren en el plano (p, H_L) cada una de las medias de los modelos de mezcla. Tales gráficos se muestran en las figuras (4.5), (4.6) y (4.7).

Para el experimento G1L1, se visualiza en la figura 4.5 que las dos primeros medias del modelo de mezcla están superpuestas a los valores de la película de líquido; este hecho confirma que el último componente (línea azul punteada) está asociado al cuerpo del tapón. Relativo a la media del tercer componente notar que su valor es cercano al colgamiento de líquido promedio (línea negra continua). Como tal, la dificultad que presenta el componente tres se puede superar observando que el valor de σ_3 en la figura 4.4 es más cercano en orden de magnitud a los dos primeros valores de σ que al del cuarto componente, lo cual se corrobora con la curva azul de la figura 4.3, ya que la pendiente de los componentes 1-2 es similar a la pendiente de los componentes 2-3, y dichas pendientes son considerablemente distintas de la pendiente 3-4. De lo anterior, se establece que el tercer componente del modelo de mezcla agrupa en mayor medida valores de la película de líquido. Adicionalmente, dado que el tercer componente es de frontera y el último componente está restringido al tapón, los componentes restantes intermedios quedan asociados al frente de mezcla.



Figura 4.5: Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L1. Fuente: Autor.

En lo que respecta al experimento G1L2. En la figura 4.6 se muestra que el colgamiento de líquido promedio se ha desplazado significativamente hacia arriba respecto al caso anterior; además, se observa que hay un nuevo componente (línea verde claro discontinua) y es este nuevo valor el más cercano al colgamiento de líquido promedio. Por su parte, notar que la aparición del cuarto componente desplaza hacia abajo μ_3 lo cual da pauta para establecer que el tercer componente esté asociado a la película de líquido; esta suposición se confirma analizando en la figura 4.4 las pendientes entre los componentes 1-2 y 2-3, las cuales son similares entre sí y discrepan, a su vez, de la inclinación entre los componentes 3-4, siendo esta última pendiente un indicador de que el componente 3 es el último grupo asociado a la película de líquido y que, por tanto, los componentes cuatro a seis abarcan los valores de colgamiento de líquido pertenecientes al frente de mezcla al igual que en el caso anterior.



Figura 4.6: Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L2. Fuente: Autor.



Figura 4.7: Medias del modelo de mezcla y valores $(p/p^*, H_L)$. Experimento G1L3. Fuente: Autor.

Por último, para el experimento G1L3 mostrado en la figura 4.7 se visualiza que nuevamente, el componente adicional (línea dorada punteada) del modelo de mezcla es el más cercano al colgamiento de líquido promedio. Se observa además, que el tercer componente (línea magenta punteada) está localizado en el límite entre la región de película y frente del tapón. Para la identificación de los componentes restantes, notar de derecha a izquierda en la figura 4.3 que la pendiente a partir del componente 7 y hasta el componente 4 es similar; en cambio, la inclinación entre los componentes 3-1 de ese mismo gráfico discrepa de la pendiente 7-4 previamente mencionada. Con este apunte, se deduce que el tercer componente se asocia con la película de líquido; los componentes 4 a 7 quedan por tanto asociados al frente de mezcla.

De la discusión anterior, la identificación de las regiones de la unidad de tapón con los componentes del modelo de mezcla es la siguiente:

	Compon	entes del N	AMG asociados
Región de flujo	G1L1	G1L2	G1L3
Película de líquido-burbuja de Dumitrescu Frente del tapón Cuerpo del tapón	$ \begin{array}{c} 1^{o}, 2^{o}, 3^{o} \\ 4^{o}, 5^{o} \\ 6^{o} \end{array} $	$ \begin{array}{c} 1^{o}, 2^{o}, 3^{o} \\ 4^{o}, 5^{o}, 6^{o} \\ 7^{o} \end{array} $	$ \begin{array}{c} 1^{o}, 2^{o}, 3^{o} \\ 4^{o}, 5^{o}, 6^{o}, 7^{o} \\ 8^{o} \end{array} $

Tabla 4.3: Identificación de regiones de flujo de la unidad de tapón con
los componentes del MMG.

Ahora bien, la identificación mostrada en la Tabla 4.3 debe entenderse como una clasificación posterior al propio agrupamiento que efectúa el modelo de mezcla con base en la probabilidad, i.e., una vez que un determinado punto H_{L_i} se ha asignado a alguno de los componentes del modelo, dicho valor se reetiqueta en términos de alguna región de flujo. Como tal, este último paso no forma parte del ajuste del modelo de mezcla en sí mismo, de allí el análisis gráfico llevado a cabo. No obstante, este reetiquetado al ser implementado después de la asignación que realiza el modelo de mezcla, hereda su carácter probabilístico; de otra manera la identificación de regiones de flujo se estaría efectuando de una forma puramente determinista. Esta distinción puede entenderse observando las distribuciones de probabilidad del modelo de mezcla.

En las figuras 4.8. 4.9 y 4.10 se grafica la función de densidad de probabilidad (FDP) total o de la mezcla (línea gris continua) así como de cada una de las distribuciones latentes (líneas puntedas de colores) para las series instantáneas de colgamiento de los experimentos de la Tabla 3.1. Se incluye además, un histograma para los valores de dicha serie.

Al comparar las tres figuras en cuestión se observa que la densidad de la primera curva normal (línea azul punteada) disminuye a medida que el flujo másico de líquido se incrementa; notar que dicho decremento coincide a su vez con un ascenso en la densidad de la última curva normal correspondiente. Esta relación de descenso-ascenso sugiere la idea de que cuando aumenta el flujo másico de líquido, la masa de dicha fase se redistribuye⁵ en las distintas regiones de flujo de la unidad de tapón, por lo que se deduce que la FDP de la última curva normal de los tres experimentos es sensible al aumento de la frecuencia de "slugs".

Por otra parte, se observa que la forma de la curva de la FDP total es más compleja para valores de colgamiento menores a 0.7; dicha complejidad se ve reflejada en una mayor superposición y cantidad de distribuciones normales. Notar que tal intervalo de colgamiento coincide con los valores frontera de la región estratificada. Este último hecho, permite entender que el componente incremental estimado con el criterio de Bayes es debido al mayor número de inestabilidades en la interfase gas-líquido a medida que aumenta la diferencia de velocidades entre las fases.

 $^{^{5}}$ Dicha redistribución de la cantidad de líquido se expresa en términos de densidad de puntos, de allí que se refleje en la FDP total.

Adicionalmente, se visualiza que las mayores desviaciones estándar corresponden a la familia de distribuciones que comprende desde la cuarta hasta la penúltima curva normal de los tres experimentos; por su parte, las medias de dichas distribuciones, como se había comentado, caen en la región del frente del mezcla. Este hecho verifica la identificación de componentes realizada anteriormente.



Figura 4.8: Distribuciones de probabilidad. Experimento G1L1. Fuente: Autor.



Figura 4.9: Distribuciones de probabilidad. Experimento G1L2. Fuente: Autor.



Figura 4.10: Distribuciones de probabilidad. Experimento G1L3. Fuente: Autor.

Por último, considerando que el comportamiento del flujo exhibe tendencias claras y aprovechando que cada componente del modelo de mezcla se ha asociado a una región de flujo de la unidad de tapón, es posible cuantificar a partir de los pesos del modelo, el peso o fracción de ocupación de alguna región de flujo en particular. Este nuevo peso se ha denominado λ_{rf} y su gráfico permite visualizar que el frente de mezclado, a pesar de tener asociadas las mayores desviaciones estándar, es la región de flujo con menor cantidad de puntos para un determinado experimento; más importante aún, las tendencias mostradas en el gráfico de λ_{rf} son la base de un prospectivo diseño de perceptrones multicapa por región de flujo.



Figura 4.11: Pesos del modelo de mezcla por región de flujo. Fuente: Autor.

4.1.2. Estimación de volúmenes

Una vez que se ha ajustado un modelo de mezcla Gaussiano a la serie instantánea de colgamiento de líquido y que se ha asociado cada componente a una región de flujo en particular, es posible obtener una aproximación a los volúmenes totales y por región de flujo de cada una de las fases.

Para efectuar dicho cálculo es necesario hacer un uso combinado de ecuaciones de la Mecánica de Fluidos, de ecuaciones deducidas del modelo de mezcla y de ecuaciones de flujo bifásico.

El procedimiento se describe a continuación:

Ecuaciones de la Mecánica de Fluidos

Partiendo de la ecuación de flujo másico para una determinada fase f:

$$m_f = \int_0^{t_{exp}} \dot{m}_f \, dt \tag{4.2}$$

Dado que el flujo másico a la entrada de la sección de pruebas (ver figura 3.1) se mantiene aproximadamente constante para un determinado experimento, la masa promedio de dicha fase se estima como:

$$m_f \approx \overline{\dot{m}_f} \int_0^{t_{exp}} dt = \overline{\dot{m}_f} t_{exp} \tag{4.3}$$

Por su parte, se sabe que el volumen se calcula como:

$$V = \frac{m_f}{\rho_f} \tag{4.4}$$

Sustituyendo la expresión (4.3) en (4.4) se obtiene:

$$V = \frac{\overline{\dot{m}_f} t_{exp}}{\rho_f} \tag{4.5}$$

Ecuaciones derivadas del modelo de mezcla

El ajuste del modelo de mezcla Gaussiano resulta en un agrupamiento probabilístico de los datos procesados. Una vez finalizado el entrenamiento con el algoritmo E-M, los pesos resultantes del modelo se pueden estimar como la cantidad de puntos del grupo k respecto al número total de datos:

$$\frac{n_k}{n}$$
 (4.6)

Dado que el número total de datos está relacionado con la duración total del experimento, la fracción dada por la ecuación 4.6 representa la *fracción de ocupación* para un componente k del MMG:

$$t_k^* = \frac{n_k}{n} \tag{4.7}$$

Por su parte, el tiempo asociado a un componente del MMG o *tiempo de ocupación* se calcula volviendo dimensional la fracción de ocupación para ese componente:

$$(\Delta t)_k = t_k^* t_{exp} \tag{4.8}$$

Por otro lado, la fracción de tiempo cumple:

$$\sum_{k} (\Delta t)_k = t_{exp} \tag{4.9}$$

Sustituyendo la ec.(4.9) en (4.5) se obtiene:

$$V = \frac{\overline{\dot{m}_f} \sum_k (\Delta t)_k}{\rho_f} \tag{4.10}$$

Desarrollando:

$$V = \frac{\overline{\dot{m}_f} \left(\Delta t_1 + \Delta t_2 + \dots \Delta t_k \right)}{\rho_f} \tag{4.11}$$

$$V = \frac{\overline{\dot{m}_f} \,\Delta t_1}{\rho_f} + \frac{\overline{\dot{m}_f} \,\Delta t_2}{\rho_f} + \dots + \frac{\overline{\dot{m}_f} \,\Delta t_k}{\rho_f} \tag{4.12}$$

$$V = V_1 + V_1 + \dots + V_k \tag{4.13}$$

De las expresiones (4.12) y (4.13) se observa que la ecuación para calcular el volumen asociado al componente k del MMG es la siguiente:

$$V_k = \frac{\overline{\dot{m}_f}(\Delta t)_k}{\rho_f} \tag{4.14}$$

La ecuación 4.13 proporciona la forma de estimar el volumen total para una determinada fase. Notar que dicha ecuación al no estar ponderada por algún factor que relacione la fracción que ocupa una determinada fase en un espacio, se trata por tanto, de una ecuación monofásica.

Ecuaciones de flujo bifásico

Considerar una tubería de longitud L, área A_{tub} en donde se tiene la distribución de fases siguiente:



Figura 4.12: Configuración de dos fases con su respectiva fracción de área. Fuente: Autor.

Para el caso multifásico, la forma de evaluar el área que ocupa una determinada fase es ponderar el área de dicha fase respecto al área total transversal al flujo, que en este caso es el área interna de la tubería.

$$c_f = \frac{A_f}{A_{tub}} \tag{4.15}$$

El área de una determinada fase se calcula como sigue:

$$A_f = c_f A_{tub} \tag{4.16}$$

Para calcular el volumen, se multiplica a ambos lados de la ecuación 4.16 por la longitud de la tubería, obteniendo:

$$A_f L = c_f A_{tub} L \tag{4.17}$$

Lo cual resulta en:

$$V_f = c_f V \tag{4.18}$$

Observar que en la ecuación 4.18 $V = V_{tub}$.

Ahora bien, considerar el diagrama siguiente:



Figura 4.13: Configuración de los volúmenes asociados a cada grupo para una determinada fase. Fuente: Autor.

En la figura 4.13, V_{f_k} representa el volumen asociado al componente k del modelo de mezcla o bien, el volumen correspondiente a la fracción que ocupan los puntos que contiene ese componente. En todo caso, considerando que los datos están agrupados, V_{f_k} se puede entender como el volumen acumulado para ese grupo en particular. No obstante, dado que el flujo no es monofásico, cada bloque V_{f_k} está ponderado con su correspondiente valor de c_{f_k} ; c_{f_k} es el valor medio de los puntos que almacena ese grupo del modelo de mezcla.

Además, con base en la figura 4.13, se deduce que el volumen V de la ecuación 4.18 corresponde con V_{tub} debido a que existe un único bloque de esa fase. Cuando existe más de un bloque de puntos para una fase, el volumen V pasa a ser la porción de volumen de un determinado grupo, a esta nueva región de volumen no ponderada se denomona V_k . Por su parte, V_f que originalmente era el volumen del único bloque, ahora es el volumen efectivo o real de cada uno de los componentes del modelo de mezcla; este último volumen se etiqueta como V_{f_k} . Lo anterior mencionado deriva en la expresión siguiente:

$$V_{f_k} = c_{f_k} V_k \tag{4.19}$$

Sustituyendo la ecuación 4.18 en 4.14 se obtiene:

$$V_{f_k} = c_{f_k} \frac{\overline{\dot{m}_f}(\Delta t)_k}{\rho_f} \tag{4.20}$$

Con la ecuación 4.20 es posible calcular el volumen acumulado de un algún componente del modelo de mezcla para una determinada fase.

En particular, para la ecuación 4.20 si la fase en cuestión corresponde a la fase líquida, se tiene entonces que $f = L \rightarrow c_L \rightarrow H_L$, por lo que la ecuación 4.20 se transforma en:

$$V_{L_k} = \frac{\dot{m}_L(\Delta t)_k}{\rho_L} H_{L_k} \tag{4.21}$$

Análogamente, para el caso donde la fase es gaseosa se obtiene:

$$V_{G_k} = \frac{\overline{\dot{m}_G}(\Delta t)_k}{\rho_G} \alpha_{G_k} \tag{4.22}$$

Los resultados de aplicar las ecuaciones 4.21 y 4.22 al conjunto de datos de la Tabla 3.1 se resumen a continuación:

	$V_G (m^3)$			$V_G(ft^3)$		
Experimento	Burbuja	Frente	Tapón	Burbuja	Frente	Tapón
G1L1	0.7254	0.0201	0.0011	25.6172	0.7099	0.0381
G1L2	0.4382	0.0353	0.0040	15.4733	1.2472	0.1397
G1L3	0.2819	0.0399	0.0079	9.9569	1.4098	0.2795

Tabla 4.4: Volumen acumulado de aire por región de flujo.

Tabla 4.5: Volumen acumulado de glicerina por región de flujo.

	$V_L \ (m^3)$		$V_L \ (bbl)$			
Experimento	Película	Frente	Tapón	Película	Frente	Tapón
G1L1	0.1329	0.0121	0.0264	0.8356	0.0763	0.1662
G1L2	0.2219	0.0481	0.1271	1.3956	0.3022	0.7991
G1L3	0.2836	0.1139	0.2458	1.7837	0.7160	1.5458

Tabla 4.6: Volúmenes totales de glicerina y aire.

	V_L		I	V_G
$\mathbf{Experimento}$	m^3	bbl	m^3	ft^3
G1L1	0.1714	1.0781	0.7466	26.3652
G1L2	0.3970	2.4970	0.4774	16.8601
G1L3	0.6432	4.0455	0.3298	11.6461

El análisis de los volúmenes calculados se detalla a continuación.

En las figuras 4.14 y 4.15 se grafican respectivamente, el volumen acumulado de la fase gaseosa en la región de burbuja y el volumen acumulado de la fase líquida en la película. Se incluyen además, para cada volumen sus respectivas bandas de incertidumbre.

En a la figura 4.14, se visualiza el volumen acumulado de aire disminuyendo conforme el flujo másico de líquido aumenta. Ahora bien, dado que el flujo másico de aire es aproximadamente constante para los tres experimentos, el comportamiento observado puede tomarse como un indicador de la compresibilidad del aire por la fase líquida, como se ingresa más cantidad de glicerina al sistema en el mismo espacio, el líquido necesariamente comprimirá en mayor medida a la masa de aire; esta compresibilidad puede ser vista en la serie instantánea de colgamiento de líquido en forma de burbujas de Dumitrescu más pequeñas en los experimentos G1L2 Y G1L3, respecto a las burbujas observadas del experimento G1L1. Dicho aumento en la compresibilidad del aire está relacionado con la mayor irregularidad que exhiben las familias de intermitencias a medida que se incrementa el flujo másico de líquido, tal como se comentó en la sección 3.3.4.

Para la figura 4.15 se observa que el volumen acumulado de glicerina que transporta la película de líquido es proporcional al flujo másico de dicha fase; esta tendencia es acorde con lo revisado en la figura 3.17, en la cual se visualiza que el nivel promedio de la película de líquido aumenta a medida que su flujo másico también lo hace.



Figura 4.14: Volumen acumulado de aire en la burbuja de Dumitrescu. Fuente: Autor.



Figura 4.15: Volumen acumulado de líquido en la película. Fuente: Autor.

En las figuras (4.16) y (4.17) se ilustran los volúmenes acumulados de aire en el cuerpo del tapón y frente de mezcla; se muestran además, las respectivas bandas de incertidumbre para cada uno de esos volúmenes.

De la figura 4.17 se puede observar que el volumen de aire crece con un comportamiento aproximadamente no lineal conforme aumenta el flujo másico de líquido; por su parte para la figura 4.17 se visualiza que el volumen promedio de aire en el cuerpo del tapón exhibe un comportamiento aproximadamente lineal a medida que se incrementa la cantidad de líquido. En todo caso, el comportamiento de ambos está vinculado debido a que el volumen acumulado de aire que captura el cuerpo del tapón es función del volumen acumulado de aire atrapado previamente por el frente de mezcla. En particular, para un determinado experimento, el volumen de aire que captura el frente es significativamente mayor que su correspondiente volumen en el cuerpo del tapón. Tal diferencia de volúmenes sugiere que las burbujas alojadas en el frente se mantienen en dicha región o bien se fragmentan o recirculan hacia la burbuja principal. En cualquier caso, se apunta que estos dos volúmenes son una medida cuantitativa⁶ del proceso de aireamiento de los tapones; se comenta además, que la estimación de los volúmenes en cuestión puede ser extendida para comparar el proceso de aireamiento en condiciones de alta y baja viscosidad.

⁶El proceso de aireamiento de los tapones típicamente se estudia de manera visual, consultar e.g. (E. M. Al-Safran et al., 2013) y (E. Al-Safran et al., 2015).



Figura 4.16: Volumen acumulado de aire en el cuerpo del tapón. Fuente: Autor.



Figura 4.17: Volumen acumulado de aire en el frente de mezcla. Fuente: Autor.

En las figuras 4.18 y 4.19 se comparan, respectivamente, los volumenes acumulados de glicerina en la región del frente y cuerpo del tapón; se incluyen además sus correspondientes bandas de incertidumbre.

Para la figura 4.19 se observa que el volumen acumulado de líquido en el tapón posee un comportamiento aproximadamente lineal respecto al flujo másico de líquido; de igual manera se visualiza en la figura 4.18 que el volumen acumulado en el frente de mezcla es proporcional respecto al flujo másico de dicha fase; esta tendencia de ambos volúmenes y en conjunto con la tendencia ascendente del volumen acumulado de glicerina en la película (ver figura 4.15) permite entender que la cantidad de líquido incremental se distribuye en todas las regiones de flujo de la unidad de tapón. En todo caso, como el intervalo experimental es pequeño no es posible observar si a mayores flujos másicos de líquido el volumen acumulado en el tapón superará a su contraparte en la región de la película, i.e., para este intervalo experimental la región que transporta mayor volumen acumulado de glicerina es la película de líquido.



Figura 4.18: Volumen acumulado de líquido en el frente de mezcla. Fuente: Autor.



Figura 4.19: Volumen acumulado de líquido en el cuerpo del tapón. Fuente: Autor.

La figura 4.20 muestra simultáneamente el volumen acumulado de líquido y aire en la región de la película y burbuja, respectivamente; se incluyen además las correspondientes bandas de incertidumbre para cada uno de los volúmenes.

Del gráfico referido se observa que para el experimento G1L1 el volumen acumulado de aire en la burbuja es considerablemente mayor al volumen acumulado de líquido en la película, esto es explicable debido a que para este caso, las burbujas de Dumitrescu tienen una longitud (o tiempo de ocupación) mayor que los tapones.

Respecto al experimento G1L2, se observa que el volumen acumulado de líquido en la película ha aumentado respecto al caso anterior calculado en la misma región de flujo. No obstante, dicho volumen acumulado de líquido aún no supera al volumen acumulado de aire en la región de burbuja.

Relativo al experimento G1L3 se visualiza que el volumen acumulado de líquido en la película es ya muy cercano al volumen acumulado de aire para la región de burbuja. Este comportamiento confirma lo mecionado anteriormente: un mayor volumen de líquido se traduce en una mayor compresibilidad de la fase gaseosa. Por otra parte, a partir de las tendencias mostradas en la Fig.(4.20) se deduce que el volumen de aire en la burbuja, si bien inicialmente es mayor que el volumen de líquido en la película, puede ser superado por este último a medida que se incremente el flujo másico de líquido. Además, analizando comparativamente la Tabla 4.6 y las figuras 4.16, 4.17 y 4.20, se observa que para el intervalo experimental considerado, la región de burbuja es la que posee mayor volumen acumulado de aire.



Figura 4.20: Comparativa del volumen acumulado de gas en la burbuja de Dumistresco y el volumen acumulado de líquido en la película. Fuente: Autor.

Ahora bien, realizado el análisis de los volúmenes acumulados se da por finalizado el apartado relativo al modelo de mezcla Gaussiano. La aplicación del modelo de perceptrón multicapa se describe en la siguiente sección.

4.2. Aplicación del modelo de perceptrón multicapa

En el presente apartado se aplica el modelo de perceptrón multicapa descrito en la sección 2.4.2 para abordar el problema de predicción del colgamiento de líquido instantáneo a partir de valores de presión en línea formulado en la sección 1.2.

El paso inicial para abordar el problema predictivo en cuestión es establecer, con base en el procedimiento expuesto en la Tabla 2.2 la configuración del modelo así como el espacio de búsqueda de hiperparámetro correspondiente. Los resultados de la aplicación de dicha metodología se resumen en la Tabla 4.7. Al modelo considerado se le denomina como PCI (que hace referencia a presión-colgamiento instantáneo). Por otra parte, para garantizar la consistencia metodológica durante el presente análisis, cualquier otra red neuronal que se construya en el estudio actual seguirá el procedimeinto según se indica en la Tabla 4.7.

Nombre	PRESIÓN_COLGAMIENTO_INSTANTÁNEO (PCI)		
Configuración del modelo			
Tipo de RNA	Perceptrón multicapa		
Capa de entrada	Número de neuronas: 1; Función de activación: lineal		
Capa de salida	Número de neuronas: 1; Función de activación: lineal		
Tamaño de lote	30		
Regularizadores	Parada temprana		
Optimizador	ADAM		
Función de pérdida	ECM		
Espacio de búsqueda			
Número de capas ocultas	[1, 3], paso 1		
Número de neuronas por capa oculta	[2, 32], paso 2		
Funciones de activación	lineal, sigmoide, tanh, relu		
Tasa de aprendizaje	$[10^{-4}, 10^{-2}]$, muestreo logarítmico		

Tabla 4.7: Configuración del modelo y del espacio de búsqueda utilizado.

Respecto a la división del conjunto experimental, con base en lo mencionado en la sección 2.4.2, la proporción de datos a utilizar para el modelo PCI es la siguiente⁷: 70 % de las muestras para la etapa de entrenamiento del modelo; 15 % para su validación y 15 % para la etapa de prueba del mismo⁸. La partición resultante se resume en la Tabla 4.8.

Tabla 4.8: Cantidad de puntos (p_i, H_{L_i}) para el modelo PCI.

\mathcal{T}	\mathcal{V}	\mathcal{P}	Total
7938	1701	1701	11340

Con la configuración y espacio de búsqueda mostrados en la Tabla 4.7 y los datos por conjunto de la Tabla 4.8, se obtiene para el modelo PCI la siguiente arquitectura, función de activación y tasa de aprendizaje:

 $^{^{7}}$ Como menciona (Ajbar et al., 2024), dado que no existe una regla sistemática para determinar la proporción entre los distintos conjuntos de datos, la cantidad destinada a cada uno de ellos puede cambiar de un problema a otro.

⁸Recordar que \mathcal{T} , \mathcal{V} y \mathcal{P} hacen referencia a los conjuntos de entrenamiento, validación y prueba, respectivamente.

Tabla 4.9: Arquitectura, función de activación y tasa de aprendizaje para el modelo PCI.

Arquitectura	$\phi~$ en capas ocultas	η
1-8-1	sigmoide	2.0951e-04

La arquitectura proporcionada en la Tabla 4.9 se ilustra a continuación.



Figura 4.21: Arquitectura de red para el modelo PCI. Fuente: Autor.

Ahora bien, una vez que se ha conseguido una arquitectura e hiperparámetros óptimos lo siguiente es analizar el comportamiento del modelo durante el proceso de entrenamiento. Tal evaluación se presenta en la figura 4.22.



Figura 4.22: Historia de entrenamiento para el modelo PCI. Fuente: Autor.



Figura 4.23: Valores de colgamiento de líquido predichos. Modelo PCI. Fuente: Autor.

En la figura 4.22 se presenta la historia de entrenamiento del perceptrón multicapa PCI (presión-colgamiento instantáneo) para los conjuntos de entrenamiento y validación. Del gráfico, se observa una caída significativa en los valores de la función de pérdida durante las primeras dos épocas⁹; para dicha función en el conjunto de validación se observa una ligera caída para la primera época y a partir de la segunda se visualiza un comportamiento constante. Comparativamente, del gráfico se puede notar que los valores de la función de pérdida en el conjunto de validación son mayores que los valores que presenta dicha función en el conjunto de entrenamiento; de lo anterior se deduce que el modelo considerado no presenta sobrejuste y que por tanto, es apto para generalizar ante valores de presión no vistos. En la Fig.(4.23) se muestran las series instantáneas de presión y colgamiento de líquido así como las predicciones efectuadas por el modelo PCI para los tres experimentos.

Del gráfico (4.23) se observa que el colgamiento predicho exhibe una tendencia ascendente respecto al flujo másico de líquido (y por lo tanto de la presión en línea). Notar que el conjunto de valores predichos es más cercano a los valores de la película de líquido para el experimento de 0.6 kg/s y progresivamente se aleja de dicha región de flujo; para el experimento de 2.0 kg/s al igual que el de 1.3, los valores predichos se encuentran en la región del frente de mezcla.

Por otro lado, resulta interesante que la predicción de la red neuronal consista únicamente en valores intermedios de colgamiento de líquido para un determinado experimento. A priori, según lo establecido por la subhipótesis 1B de la sección 1.1, la predicción esperada sería haber obtenido una aproximación a los distintos valores instantáneos de la serie de H_L , a partir de los valores instantáneos de la serie de presión, es decir, una estimación del colgamiento con valores en todas las regiones de flujo. Otro posible escenario sería haber predicho al menos una región de flujo en específico (la que contuviera más datos, en todo caso) para distintos valores de presión. Sin embargo, como es observado en la figura 4.23 las predicciones de la red neuronal no cubren la totalidad de alguna región particular del flujo. Además, con base en la figura 4.22 y considerando que el perceptrón PCI posee una arquitectura óptima, se descarta la posibilidad de que las épocas con las cuales se entrenó el modelo hayan sido insuficientes (subaprendizaje del modelo) para captar la relación subyacente entre p y H_L ya que la curva de la función de pérdida muestra un comportamiento claramente convergente.

Para entender e interpretar de mejor manera el comportamiento observado en la figura 4.23 resulta conveniente considerar un caso diferente y simplificado del problema original de predicción. Para este nuevo caso se propone examinar las predicciones de un perceptrón multicapa entrenado sólo con los datos de un experimento, poniendo especial atención en el error cuadrático medio, que fue la función de pérdida seleccionada para evaluar el rendimiento del modelo. Para esta versión hipotética se consideran los valores (p_i, H_{L_i}) del experimento G1L1. Los gráficos relativos a este nuevo caso se exponen a continuación.

 $^{{}^{9}}$ Recordar del capítulo 2 que una época se refiere a una iteración en donde todo el conjunto de datos de entrenamiento ha sido procesado a través de la red neuronal.



Figura 4.24: Historia de entrenamiento. Red G1L1. Fuente: Autor.



Figura 4.25: Predicciones de colgamiento de líquido. Red G1L1. Fuente: Autor.

En las figuras 4.24 y 4.25 se presentan la historia de entrenamiento así como los valores predichos de colgamiento de líquido para el caso hipotético considerado, respectivamente.

Específicamente, la Fig.(4.24) exhibe una caída significativa de la función de pérdida en las primeras dos épocas, para dicha curva se visualiza además que alcanza la convergencia alrededor de la época tres; en cuanto a la curva de validación, esta exhibe de igual manera una caída para las dos primeras épocas; no obstante, comienza a estabilizarse a partir de la época cuatro. Se aprecia también que la función de pérdida en el conjunto de validación es significativamente mayor respecto a la curva del conjunto de entrenamiento. Notar además, que para ambas curvas no se observan fluctuaciones considerables. De lo anterior se descarta que el modelo presente la condición de sobreajuste.

Por otro lado, en la figura 4.25 se exponen los valores (p_i, H_{L_i}) del experimento G1L1 así como las predicciones del colgamiento realizadas por la red neuronal hipotética considerada. Observar que los nuevos valores de H_L son aproximadamente constantes para todo el intervalo de presión del conjunto de prueba. Notar que en conjunto, los valores predichos dan el aspecto de formar una línea horizontal a diferencia de la apariencia de recta con un ángulo de inclinación de la Fig.(4.23) del caso original.¹⁰

Adicionalmente, el hecho de que las predicciones de colgamiento para un determinado experimento sean bastante cercanas unas a otras en los dos casos de análisis considerados, sugiere la idea de la existencia de un valor al cuál se aproximan dichan predicciones una vez que la función de pérdida se minimiza durante el proceso de entrenamiento. El cálculo del valor de aproximación se requiere para explicar e interpretar el significado de las predicciones de colgamiento de líquido del caso hipotético y para el modelo PCI considerados. Como tal, la pauta para llevar a cabo la estimación mencionada se encuentra en la función de pérdida de ECM. Si bien es sabido que la función de pérdida no se minimiza explícitamente en el conjunto de validación (por definición), observar en las figuras 4.22 y 4.24 que la curva de validación sí tiende hacia un valor mínimo. Con ese último estado de la curva de validación posteriormente se evalúan las predicciones observadas en los gráficos 4.23 y 4.25. En otras palabras, cuando el ECM en el conjunto de validación se acerca a un mínimo una vez que dicha función de pérdida se ha minimizado en el conjunto de entrenamiento, los valores predichos \hat{y}_i se aproximan a una constante k. En el límite, lo anterior resulta en:

$$\lim_{\hat{y}_i \to k} \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \frac{1}{n} \sum_{i=1}^n (y_i - k)^2$$
(4.23)

 $^{^{10}}$ Se deduce, por tanto, que la tendencia de los valores predichos ilustrados en la Fig. (4.23) corresponde a un patrón de proporcionalidad que captura el perceptrón en el colgamiento a medida que los valores de presión aumentan.

$$ECM = \frac{1}{n} \sum_{i=1}^{n} (y_i - k)^2$$
(4.24)

De la ecuación previa, es posible obtener el valor k deseado, calculando el mínimo de ese ECM.

$$\frac{d(ECM)}{dk} = \frac{1}{n} [-2(y_1 - k) - 2(y_2 - k) + \dots + (-2)(y_n - k)]$$
$$\frac{d(ECM)}{dk} = \frac{1}{n} [-2y_1 + 2k - 2y_2 + 2k + \dots + -2y_n + 2k]$$

Agrupando términos correspondientes:

$$\frac{d(ECM)}{dk} = \frac{1}{n} \left[-2(y_1 + y_2 + \dots + y_n) + 2\sum_{i=1}^n k \right]$$
$$\frac{d(ECM)}{dk} = \frac{1}{n} \left[-2\sum_{i=1}^n y_i + 2\sum_{i=1}^n k \right]$$
$$\frac{d(ECM)}{dk} = \frac{1}{n} \left[-2\sum_{i=1}^n y_i + 2nk \right]$$
(4.25)

Igualando la ecuación 4.25 a cero y despejando k:

$$-2\sum_{i=1}^{n} y_i + 2nk = 0$$

$$k = \frac{1}{n}\sum_{i=1}^{n} y_i$$
(4.26)

Adicionalmente, derivando (4.25):

$$\frac{d^2(ECM)}{dk^2} = 2k > 0 \tag{4.27}$$
Como la segunda derivada del ECM es positiva, se concluye que el ECM es mínimo cuando k es el promedio de $\{y\}$.

Ahora bien, una vez conocido el valor de k, es posible entender que las predicciones de H_L mostradas en la gráfica (4.25) son una aproximación al colgamiento de líquido promedio de algún experimento en particular.

Por otro lado, se podría suponer que el valor de k sólo es válido para el caso donde la red neuronal se entrena con un único experimento. Para comparar el alcance del valor de k en el caso real considerar la imagen siguiente.



Figura 4.26: Comparativa de $\overline{H_L}$ para las predicciones del modelo PCI. Fuente: Autor.

En la figura 4.26 se exponen las predicciones del modelo PCI para los diferentes experimentos así como su colgamiento promedio experimental (ilustrado con el marcador de estrella). Del gráfico en cuestión, se observa que a pesar de la tendencia ascendente que exhiben los valores predichos de un experimento a otro, las predicciones aún permanecen cercanas al colgamiento de líquido promedio de ese experimento. De este hecho, es posible deducir que el valor de k proporcionado en la expresión (4.26) permite entender e interpretar adecuadamente las predicciones de colgamiento para el caso original considerado.

Por otro lado, el comportamiento promedio observado en las predicciones de colgamiento para el modelo PCI así como para el caso hipotético formulado, podría argumentarse que es debido a que en la configuración de dichos modelos se utilizó sólo una neurona en sus capas de salida. Para evaluar esta suposición se construyó un perceptrón multicapa siguiendo la metodología de la Tabla 2.2 pero con la modificación de que este perceptrón tenga tres neuronas¹¹ en su capa de salida. Las predicciones correspondientes a este nuevo modelo se exponen a continuación:



Figura 4.27: Predicción del colgamiento de líquido con una red neuronal con 3 neuronas en la capa de salida. Fuente: Autor.

En la figura 4.27 se exponen los valores de colgamiento predichos para cada una de las tres neuronas de la capa de salida del perceptrón (triángulos en colores rojo, negro y azul). Se grafica también el colgamiento promedio de cada experimento.

Como tal, del gráfico referido se visualiza que las predicciones del perceptrón con tres neuronas en la capa de salida son, de igual manera que en los casos anteriores, muy próximas al colgamiento de líquido promedio de un determinado experimento. Este hecho permite descartar que los valores predichos efectuados por el modelo PCI mostrados en la Fig.(4.23) sean particulares a que dicho perceptrón tenga sólo una neurona en su capa de salida.

Con base en los resultados mostrados en las figuras 4.23 y 4.25, resulta evidente que si el modelo de perceptrón multicapa es entrenado con parejas de valores instantáneos (p_i, H_{L_i})

 $^{^{11}}$ Asumiendo que cada neurona está asociada a la predicción de colgamiento en cada una de las regiones de flujo (película, frente y tapón).

y dicho modelo se configura además, según lo establecido por la Tabla 4.7, las correspondientes predicciones de ese perceptrón se aproximarán al colgamiento de líquido promedio de un determinado experimento; esta aproximación promedio a la serie de colgamiento difiere sustancialmente con la idea de utilizar el modelo de perceptrón multicapa como una forma de poder reconstruir nuevas series instantáneas de colgamiento; esta discrepancia implica necesariamente una refutación de la hipótesis 1B formulada en la sección 1.1. En todo caso, tener presente que la aproximación promedio de colgamiento obtenida por el modelo PCI es consecuencia de que la relación entre presión y colgamiento de líquido no es una función, tal como se comentó en la sección 3.3.4.

Por último, el hecho de que las predicciones del modelo PCI se aproximen únicamente al promedio de la serie instantánea de colgamiento de líquido condiciona, indirectamente, una prospectiva predicción de dicho parámetro en las distintas regiones de flujo de la unidad de tapón, i.e., si se aspira a tal pronóstico se requerirá efectuar una descomposición en las parejas de valores (p_i, H_{L_i}) o bien, utilizar un modelo de red neuronal que sea robusto al problema de no unicidad mencionado anteriormente. No obstante, dado que hacer uso de una red neuronal de distinta naturaleza supone un nuevo trabajo por sí mismo, en el presente estudio se elige la opción de agrupar los valores (p_i, H_{L_i}) por medio del modelo de mezcla Gaussiano explicado en la sección anterior. Una vez logrado el agrupamiento, se construye un perceptrón multicapa para predecir el colgamiento en cada región de flujo. Dicho procedimiento se detalla a continuación:

4.2.1. Relación entre la presión y el colgamiento de líquido instantáneos por región de flujo

En el presente apartado se aborda la problemática de predecir el colgamiento de líquido por región de flujo. Como tal, el primer obstáculo en esta nueva tarea es el comportamiento variable observado en el número componentes óptimo del modelo de mezcla (Ver Tabla 4.1), lo cual deriva en que el significado físico que se asigna a cada componente difiera de un experimento a otro. Otro aspecto a tomar en cuenta es que desde el punto de vista predictivo, tiene que existir alguna tendencia entre p y H_L que la red neuronal pueda capturar y que tenga sentido respecto de lo que se observa experimentalmente del flujo.

Las consideraciones previas resultan en la siguiente reasignación que toma como referencia la Tabla (4.3):

	Componentes del MMG por región de flujo			
Experimento	Película	Frente de mezcla	Tapón	
G1L1	$1^{o}, 2^{o}, 3^{o}$	4^o	6^o	
G1L2	$1^{o}, 2^{o}, 3^{o}$	5^o	7^{o}	
G1L3	$1^{o}, 2^{o}, 3^{o}$	6^o	8^o	

Tabla 4.10: Componentes del MMG para las redes neuronales por región de flujo.

Una vez efectuada la reasociación de componentes del modelo de mezcla, lo siguiente es determinar la cantidad de valores (p_i, H_{L_i}) que tendrá cada red neuronal para los procesos de búsqueda de hiperparámetros óptimos así como del entrenamiento del hipermodelo correspondiente. El número de puntos en cuestión se expone en la tabla siguiente:

Tabla 4.11: Cantidad de puntos (p_i, H_{L_i}) utilizada para cada red neuronal por región de flujo.

	F	Puntos (p_i, H_{L_i})		
Región de flujo	\mathcal{T}	\mathcal{V}	\mathcal{P}	Total
Película	5409	1159	1159	7727
Frente de mezcla	232	50	50	332
Tapón	1665	356	357	2378

De la Tabla 4.11 notar que la región del tapón así como la región del frente poseen sólo un componente; no obstante, se observa que la cantidad de datos para la región del cuerpo del tapón es significativamente mayor; esta situación es explicable debido a que la región del frente de mezcla a pesar de abarcar un intervalo más amplio de valores de colgamiento es también la región de flujo con menor densidad de puntos en los tres experimentos (ver figura 4.11). En cuanto al tapón y la película de líquido, la tabla muestra que dichas regiones contienen más datos aún cuando su intervalo de valores es más restringido.

Una vez generados los conjuntos de datos necesarios, el siguiente paso consiste en aplicar la metodología descrita en la Tabla 2.2 para obtener la red neuronal correspondiente a cada región de flujo. Por consistencia metodológica y para evitar sesgos en la configuración de los modelos, el espacio de búsqueda es el mismo para los tres ensayos considerados. Además, dado que no existe una mejora sustancial en el uso de más de una neurona en la capa de salida, la actual configuración para cada uno de los modelos contempla sólo una neurona para la capa mencionada. Las configuraciones de los hipermodelos resultantes se expone a continuación.

Región de flujo	Arquitectura	ϕ en capas ocultas	η
Película	1-2-14-1	relu	1.0415e-03
Frente de mezcla	1-4-1	anh	2.0072 e-03
Tapón	1 - 16 - 4 - 1	$\operatorname{sigmoide}$	3.9978e-03

Tabla 4.12: Arquitectura, función de activación y tasa de aprendizaje para cada red neuronal por región de flujo.

Para la Tabla 4.12 se visualiza que η está en el mismo orden de magnitud para las tres regiones de flujo. En este caso las tasas de aprendizaje son mayores a la calculada para el modelo PCI (ver Tabla 4.9). De la tabla 4.12 también se puede apreciar que los modelos para la película y el tapón son considerablemente más complejos que el modelo relativo al frente de mezcla, siendo éste último más parecido a la red neuronal de la sección anterior en cuanto al número de capas ocultas y neuronas se refiere. Además, se comenta que por facilidad de lectura se ha optado por la notación de arquitectura abreviada; la visualización de cada uno de los modelos construidos se ilustra a continuación.



Figura 4.28: Arquitectura de red para la región de la película de líquido. Fuente: Autor.



Figura 4.29: Arquitectura de red para la región del frente de mezcla. Fuente: Autor.



Figura 4.30: Arquitectura de red para la región del cuerpo del tapón. Fuente: Autor.

Encontradas las arquitecturas óptimas para cada región de flujo, se evalúa el desempeño de los hipermodelos durante el proceso de entrenamiento. Los gráficos de la función de pérdida en los conjuntos de entrenamiento y validación se exponen a continuación:



Figura 4.31: Historia de entrenamiento. Red neuronal para la región de la película de líquido. Fuente: Autor.

La figura 4.31 ilustra el comportamiento de la función de pérdida en los conjuntos de entrenamiento y validación correspondiente al perceptrón multicapa asociado a la película de líquido. Del gráfico, se observa que el modelo converge aproximadamente en la época cuatro. También se visualiza que la función de pérdida durante el entrenamiento decrece alrededor el triple desde la época uno hasta su estabilización; por su parte, para la función de pérdida en el conjunto de validación se aprecia que desde el comienzo sus valores son cercanos a 0.04; además, para esta misma curva se distingue la presencia de ligeras fluctuaciones a partir de la época diez. Por otra parte, el sobreajuste del modelo se descarta debido al hecho de que la curva de validación a pesar de las fluctuaciones mencionadas, es siempre mayor que la curva de entrenamiento.



Figura 4.32: Historia de entrenamiento. Red neuronal para la región del frente de mezcla. Fuente: Autor.

En cuanto al desempeño de la red neuronal para el frente de mezcla, se visualiza en la figura 4.32 que la función de pérdida tiene una caída abrupta alrededor de la época doce, sin embargo, continúa decayendo hasta estabilizarse por completo para la época cuarenta aproximadamente. Además, se observa que para el conjunto de validación su curva es bastante cercana a la del conjunto de entrenamiento en las primeras veinte épocas y no es sino hasta que la función de pérdida se logra minimizar, que la curva de validación se separa de la curva de entrenamiento. Por otro lado, el presente gráfico permite ejemplificar la importancia del uso adecuado de la parada temprana: si bien la curva de entrenamiento presenta una caída importante al inicio, la parada temprana identifica que a pesar de ello la función de pérdida sigue disminuyendo de una manera que considera significativa; épocas después una vez que se alcanza la convergencia, detiene el modelo. Ahora bien, en lo que respecta al poder predictivo de la red, dado que ambas curvas se separan al final del entrenamiento, se descarta la presencia de sobreajuste.



Figura 4.33: Historia de entrenamiento. Red neuronal para la región del cuerpo del tapón. Fuente: Autor.

Continuando con el análisis del desempeño de hipermodelos, en la figura 4.33 se ilustra la historia de entrenamiento del perceptrón multicapa relativo al cuerpo del tapón. Del gráfico considerado, se observa que en este caso el ECM en el conjunto de entrenamiento decrece alrededor de una séptima parte de la primera a la segunda época y que alcanza la estabilización aproximadamente en la época ocho. Respecto a la curva de validación, se puede visualizar que inicia con valores próximos a 0.05 y posteriormente presenta una trayectoria similar a la curva de entrenamiento. Adicionalmente, se distingue que la mayor proximidad entre ambas curvas ocurre justo antes de que se alcance la convergencia. Después de dicha convergencia, las curvas guardan una distancia suficiente como para descartar la presencia de sobreajuste.

Una revisión comparativa del valor del ECM en las figuras (4.31), (4.32) y (4.33) respecto al valor del ECM en la figura 4.22 revela que la curva en el conjunto de validación alcanza valores menores de alrededor del triple para los modelos particulares a cada región de flujo. Este hecho es un indicador de que el poder predictivo de un modelo de perceptrón multicapa es susceptible a la variabilidad que poseen los valores que le fueron suministrados, i.e., a mayor dispersión en las entradas del modelo, mayor será el error de las predicciones. Tal apunte sustenta la descomposición efectuada en los valores instantáneos (p_i, H_{L_i}) .

Ahora bien, explicada la historia de entrenamiento de cada modelo de red neuronal por región de flujo, resta analizar las predicciones del colgamiento de líquido correspondientes. Dicha información se presenta en el gráfico siguiente.



Figura 4.34: Predicciones del colgamiento de líquido por región de flujo. Fuente: Autor.

En la figura 4.34 se exponen los diferentes valores predichos de colgamiento de líquido de cada una de las redes neuronales construidas. De arriba hacia abajo, los valores corresponden a las predicciones de colgamiento de líquido en el tapón, frente de mezcla y película de líquido. Del gráfico, se visualiza en primer lugar, la tendencia descendente en los valores predichos para H_{LS} . Tal tendencia a la baja es debida a que los tapones capturan mayor cantidad de aire (ver figuras 4.16 y 4.17) a causa del aumento de su velocidad al incrementarse el flujo másico de líquido. De igual manera, el aumento en la cantidad de aire atrapada por los "slugs" está relacionado con la mayor frecuencia de tapón derivada del incremento de la cantidad de masa de líquido; recordar que el aumento de dicha frecuencia se vincula a su vez, con la mayor irregularidad en las secuencias burbuja-tapón (ver figura 3.9) la cual resulta en un espectro de colgamiento con una dispersión significativa en sus principales componentes de frecuencia (ver figura 3.15).

Respecto a las predicciones del modelo para el frente de mezcla, se observa que tienen una tendencia ascendente de una pendiente considerablemente más pronunciada respecto a las predicciones del colgamiento para la región de la película de líquido. Lo anterior resulta en que los valores predichos para el frente sean más cercanos al cuerpo del tapón para el experimento G1L3 y más próximos a la película de líquido para el experimento G1L1. Además, se precisa que a pesar de que el intervalo de valores de colgamiento que abarca el frente de mezcla es el más amplio de las tres regiones de flujo consideradas, la alta dispersión que posee dificulta

la predicción en otras zonas de esta región de flujo debido a que los componentes restantes del MMG asociados a ella no exhiben algún tipo de tendencia. Lo anterior resulta en que el colgamiento para el frente de mezclado sea el valor de más difícil predicción.¹²

Respecto a la región de la película, de la figura 4.34 se visualiza que los valores predichos de colgamiento aumentan de forma muy controlada en proporción al incremento del flujo másico de líquido; este patrón está en acuerdo con lo observado en la Fig.(3.17) para la región de flujo en cuestión. Se identifica además, que todas las predicciones de colgamiento son menores a 0.65.

A su vez, una comparativa de la figura 4.34 muestra que las tendencias de las predicciones de colgamiento de cada perceptrón multicapa son todas diferentes entre sí. Dicha discrepancia es un reflejo de la complejidad subyacente en las estructuras de las regiones de este patrón de flujo.

Por otro lado, la idea del valor k calculado para el modelo PCI de la sección anterior permite entender, para el presente caso, que las predicciones de colgamiento mostradas en la figura 4.34 para las regiones del cuerpo del tapón, frente y película de líquido son una aproximación al \overline{H}_{LS} , $\overline{H}_{L_{fm}}$ y $\overline{H}_{L_{plc}}$, respectivamente.

Complementario a las predicciones de colgamiento por región de flujo de la figura 4.34 se grafican dichos valores acompañados de las predicciones de colgamiento efectuadas por el modelo PCI de la sección 4.2.

¹²De allí que diversos autores prescindan de incluirlo en sus análisis. No obstante, el estudio específico del frente de mezcla proporciona información valiosa sobre el comportamiento del flujo además de que dicha región es un elemento esencial de la unidad de tapón.



Figura 4.35: Predicciones del colgamiento de líquido promedio y por región de flujo. Fuente: Autor.

En la figura 4.35 se grafican comparativamente, las predicciones de colgamiento de líquido en el cuerpo, frente de mezcla y película así como los valores correspondientes al colgamiento promedio efectuado por el modelo PCI. En la figura, se puede visualizar con claridad la menor proporción de datos para el frente de mezcla respecto a las predicciones de los modelos restantes.

Por otro lado, una revisión de la figura 4.35 permite entender de mejor manera que el aumento del colgamiento promedio conforme se incrementa el flujo másico de líquido se debe a un crecimiento en el nivel de líquido de la película así como a una frecuencia mayor de tapones. Además, notar que la tendencia de \overline{H}_L presenta un comportamiento similar al de la tendencia de los valores predichos de colgamiento para la región del frente de mezcla, siendo este parecido más evidente para los dos primeros experimentos; tal familiaridad resulta interesante considerando que el modelo PCI y el del frente de mezcla poseen una arquitectura semejante. No obstante, se precisa que para las predicciones del frente de mezcla, específicamente en el experimento G1L3, el colgamiento crece en mayor proporción respecto al incremento correspondiente que presenta el colgamiento de líquido promedio. Este crecimiento ligeramente mayor en el frente de mezcla para el experimento G1L3 permite apreciar que el modelo de perceptrón multicapa es capaz de aprender patrones sutiles en los datos sin tener que asumir *a priori*, un cierto tipo de comportamiento (lineal, cuadrático, etcétera). Esta característica de aprendizaje autónomo es especialmente valiosa en conjuntos de datos que son difíciles de modelar, como es el caso de $H_{L_{plc}}$, para el cual no se dispone de fórmulas predictivas explícitas o como H_{LS} , cuya predicción está escasamente reportada para flujos de alta viscosidad.

Ahora bien, capturados los patrones de colgamiento en las distintas regiones de flujo, se está en condiciones de poder predecir con los correspondientes modelos de perceptrón multicapa, nuevos valores de colgamiento para presiones desconocidas¹³. En esta nueva predicción, se comenta que se tomaron en cuenta valores de presión más allá del rango experimental. Tal inclusión es lícita considerando la cercanía de los nuevos valores respecto a los datos experimentales y porque se sabe, de otras campañas de prueba, que para ese intervalo de presión el patrón de flujo continúa siendo intermitente. Dicho lo anterior, las predicciones mencionadas se ilustran en la siguiente figura.



Figura 4.36: Predicciones de colgamiento de líquido para presiones desconocidas por región de flujo. Fuente: Autor.

En la figura 4.36 se exponen las predicciones de colgamiento de líquido por región de flujo para presiones desconocidas correspondientes a los flujos másicos de líquido de 0.9, 1.6 y 2.3 (kg/s).

Del gráfico en cuestión, se muestra que el colgamiento de líquido en el tapón, incluso para el experimento de 2.3 (kg/s) no es menor a 0.95. Por su parte, para la película de líquido

 $^{^{13}\}mathrm{Los}$ cuales se generan a partir de nuevos valores de flujo másico de líquido.

se observa que todas las predicciones de colgamiento para nuevas presiones se mantienen en un intervalo menor a 0.65. Relativo a la región del frente de mezcla, se distingue el rápido crecimiento de las predicciones, las cuales pasan de un valor de 0.74 aproximadamente para el experimento de 0.9 (kg/s) a alrededor de 0.90 para la prueba de 2.3 (kg/s).

Finalmente, resta comparar las predicciones de la Fig.(4.36) añaniendo los valores correspondientes de colgamiento de líquido promedio del modelo PCI para los nuevos experimentos considerados. Tal gráfico resume la solución encontrada al problema de predicción de colgamiento planteado en la sección 1.2 y permite dar por terminada la presente investigación.



Figura 4.37: Predicciones de colgamiento de líquido promedio y por región de flujo para presiones desconocidas. Fuente: Autor.

5. CONCLUSIONES

En este apartado se describen las conclusiones obtenidas de la actual investigación. Para una mejor claridad se enlistan dependiendo de la técnica de análisis utilizada. Se incluye además, el trabajo futuro correspondiente así como las recomendaciones y contribuciones relativas al estudio realizado.

Para la comparativa de densidad de potencia espectral, se concluye que:

- 1. El primer armónico de los tapones de glicerina es proporcional al aumento de flujo másico de líquido y es mayor a 0.5 Hz en todos los experimentos considerados; por su parte, el primer armónico de presión resultó menor a 0.5 Hz para todos los experimentos.
- 2. A medida que aumenta el flujo másico de líquido, aumenta la dispersión de los primeros cuatro armónicos en el espectro de colgamiento de líquido.
- 3. Para un experimento en particular, las frecuencias de los primeros cuatro armónicos del colgamiento de líquido no mostraron significancia en el correspondiente espectro de presión.

A su vez, para el modelo de mezcla Gaussiano (MMG) las conclusiones son:

- 1. La función de densidad de probabilidad (FDP) de la serie de colgamiento de líquido se puede entender como una distribución compuesta cuyas distribuciones latentes están asociadas a cada una de las regiones de flujo de la unidad de tapón.
- 2. El número de componentes empleando el criterio de información de Bayes muestra ser directamente proporcional al aumento de flujo másico de líquido. Para el criterio de información de Akaike, no fue posible observar una tendencia clara respecto a la misma variable de flujo.

- 3. El componente incremental estimado con el criterio de información de Bayes se debe a las mayores inestabilidades en la interfase gas-líquido a medida que aumenta la diferencia de velocidad entre las fases.
- 4. La no linealidad en la penalización que efectúa el criterio de Bayes sobre el número óptimo de componentes resulta en modelos de mezcla más interpretables desde el punto de vista del flujo que el criterio de Akaike.
- 5. En conjunto, los valores de la media y de desviación estándar del modelo de mezcla permiten definir el concepto de unidad estadística de tapón.
- 6. Con base en los pesos del modelo de mezcla, es posible estimar el tiempo de ocupación de cada región de flujo para un determinado experimento.
- 7. Los componentes de mayor desviación estándar del modelo de mezcla están asociados al frente del tapón debido a la presencia de vórtices de mezcla. A pesar de su alta dispersión, dicha región de flujo es la que posee menor cantidad de puntos en todos los experimentos considerados.
- 8. Para los volúmenes estimados con base en el MMG se menciona que:
 - a) El volumen total de aire se comprime a medida que aumenta el flujo másico de glicerina.
 - b) Para el intervalo experimental considerado, la región que transporta mayor volumen acumulado de aire y de glicerina es la región estratificada.
 - c) El volumen acumulado de aire que atrapa el cuerpo del tapón es proporcional al flujo másico de glicerina.
 - d) A medida que se incrementa el flujo másico de líquido, el volumen acumulado de glicerina asociado se distribuye en todas las regiones de flujo de la unidad de tapón.
 - e) Los volúmenes acumulados de aire en el frente de mezcla y cuerpo del tapón pueden tomarse como una medida cuantitativa del proceso de aireamiento de los tapones.

Relativo a las conclusiones sobre el modelo de perceptrón multicapa, se tiene que:

- 1. Dado que la relación entre las series instantáneas de presión y colgamiento de líquido no cumple con la definición de función, para el espacio de búsqueda de hiperparámetros y configuración del modelo considerados, se encontró que el menor error de generalización una vez que la función de pérdida fue minimizada se alcanza cuando las predicciones se aproximan al colgamiento de líquido promedio de algún experimento en particular.
- 2. Para el perceptrón multicapa considerado, no se presenta una mejora significativa en las predicciones de colgamiento al aumentar el número de neuronas para la capa de salida.

- 3. Los perceptrones multicapa por región de flujo capturan adecuadamente las tendencias exhibidas de su región de flujo correspondiente.
- 4. Las predicciones de colgamiento de los perceptrones multicapa por región de flujo aproximan el colgamiento de líquido promedio de esa región de flujo.
- 5. Los perceptrones multicapa específicos a cada región de flujo de la unidad de tapón tienen un poder predictivo mayor respecto al perceptrón multicapa cuyas entradas son las series instantáneas de presión y colgamiento de líquido.
- 6. El frente de mezclado debido a su alta dispersión y menor densidad de puntos, resulta en la región de flujo de más difícil predicción.
- 7. Cuando la viscosidad de la fase líquida es de 0.6 Pa·s y el flujo másico de la fase gaseosa se mantiene aproximadamente constante, las predicciones de los perceptrones multicapa muestran que el colgamiento de líquido en la región de la película es directamente proporcional al aumento de flujo másico de líquido y que el colgamiento en el cuerpo del tapón es inversamente proporcional respecto al aumento de esa misma variable.

5.1. Contribuciones

- 1. Se presenta al modelo de mezcla Gaussiano (MMG) como una herramienta para analizar la serie de colgamiento de líquido.
- 2. Se presenta una forma sistemática de estimar la densidad multimodal del colgamiento de líquido mediante el uso del MMG y de los criterios de información de Bayes y de Akaike.
- 3. Se presenta una metodología de estimación de volúmenes totales y por región de flujo utilizando el MMG.
- 4. Se proporciona una forma cuantitativa de estudiar el proceso de aireamiento de los tapones por medio de los volúmenes promedio de aire en el frente y cuerpo del tapón.
- 5. Se presenta una forma de calcular el colgamiento de líquido promedio en el tapón a partir del ajuste de un MMG.
- 6. Con base en los parámetros del MMG, se presenta el concepto de unidad estadística de tapón.

5.2. Trabajo futuro

Del análisis efectuado en el presente estudio, las actividades futuras correspondientes se describen a continuación:

- 1. Aplicar la metodología de análisis propuesta a matrices experimentales con mayores combinaciones de flujos másicos y comparar los resultados para los casos de alta y baja viscosidad.
- 2. Analizar el comportamiento instantáneo del colgamiento de líquido en condiciones de flujo no isotérmico; para dicho estudio se requeriría instalar en la sección de pruebas instrumentación para efectuar mediciones de viscosidad en línea.

Por su parte, con el objetivo de continuar con la investigación de la relación entre las series de tiempo de presión y colgamiento de líquido, se tienen los siguientes puntos por desarrollar respecto al análisis espectral de dichas señales:

- 1. Comparar los espectros de presión y colgamiento de líquido para presiones medidas en la parte inferior de la tubería.
- 2. Comparar los espectros de la presión hidrostática en línea y de colgamiento de líquido.

Referente a la temática de aprendizaje profundo, se tienen los siguientes puntos a desarrollar:

- 1. Realizar un análisis del tiempo de cómputo y de la cantidad de modelos considerados comparando las técnicas de búsqueda aleatoria por malla, hiperbanda y búsqueda Bayesiana.
- 2. Evaluar el comportamiento de las predicciones realizadas por el perceptrón multicapa utilizando funciones de pérdida distintas al error cuadrático medio.
- 3. Incorporar además de la parada temprana, técnicas de regularización como L_1, L_2 y descarte aleatorio.
- 4. Diseñar una función de activación ad hoc al comportamiento que exhibe la serie de colgamiento de líquido.

A su vez, para el modelo de mezcla Gaussiano, el trabajo futuro comprende los siguientes puntos:

- 1. Con dicho modelo, estudiar estadísticamente la evolución del flujo tapón en distintos puntos de la tubería.
- 2. Investigar el ajuste de dicho modelo en espacios de dimensión superior.
- 3. Comparar la densidad resultante del colgamiento de líquido obtenida con el MMG con otras metodologías como la estimación de densidad de núcleo.

5.3. Recomendaciones

Finalmente, con la intención de mejorar la planeación de campañas experimentales futuras se hace mención de las observaciones siguientes:

- 1. Dado que la banda de frecuencias de los tapones de glicerina resultó cercana a la frecuencia de Nyquist, se recomienda aumentar la tasa de muestreo del adquisidor de datos del tomógrafo a por lo menos 10 Hz.
- 2. De ser posible, se recomienda que para campañas experimentales futuras las series de tiempo muestreadas con diferentes adquisidores de datos estén sincronizadas.
- 3. Con el objetivo de analizar el impacto que tiene el cambio de geometría en el comportamiento hidrodinámico de los tapones, se recomienda colocar un tomógrafo antes y después de la curva de la sección de pruebas y comparar los resultados obtenidos para los casos de alta y baja viscosidad.
- 4. Con el propósito de disponer de una cantidad de datos mayor, se recomienda aumentar por experimento, el tiempo de duración por tramo así como el número de tramos. Lo anterior con la intención de que el trabajo estadístico obtenido sea lo más representativo posible.
- 5. Mientras sea experimentalmente factible, se recomienda instalar tramos de tubería transparente cercanos al punto de inyección del circuito con la finalidad de estudiar el mecanismo de formación del flujo tapón en alta viscosidad.

Apéndice A

Artículo publicado

Una versión de las ideas fundacionales expuestas en el presente estudio se encuentra en la publicación adjunta llevada a cabo por parte del equipo de trabajo.



Figura A.1: Artículo publicado en Physics of Fluids.

Bibliografía

- Abdul-Majeed, G. (1996). Liquid holdup in horizontal two-phase gas—liquid flow. *Journal* of Petroleum Science and Engineering, 15(2-4), 271-280.
- Abdul-Majeed, G. H. (2000). Liquid slug holdup in horizontal and slightly inclined two-phase slug flow. Journal of Petroleum Science and Engineering, 27(1-2), 27-32.
- Aggarwal, C. C., et al. (2018). Neural networks and deep learning (Vol. 10). Springer.
- Ajbar, W., Torres, L., Guzmán, J., Hernández-García, J., & Palacio-Pérez, A. (2024). Development of artificial neural networks for the prediction of the pressure field along a horizontal pipe conveying high-viscosity two-phase flow. *Flow Measurement and Instrumentation*, 96, 102541.
- Akagawa, K. (1964). Fluctuation of void ratio in two-phase flow: 1st report, the properties in a vertical upward flow. Bulletin of JSME, 7(25), 122-128.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. En Selected papers of hirotugu akaike (pp. 199-213). Springer.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716-723.
- Al-Safran, E. (2009a). Investigation and prediction of slug frequency in gas/liquid horizontal pipe flow. Journal of petroleum science and engineering, 69(1-2), 143-155.
- Al-Safran, E. (2009b). Prediction of slug liquid holdup in horizontal pipes.
- Al-Safran, E., Kora, C., & Sarica, C. (2015). Prediction of slug liquid holdup in high viscosity liquid and gas two-phase flow in horizontal pipes. *Journal of Petroleum Science and Engineering*, 133, 566-575.
- Al-Safran, E. M., Gokcal, B., & Sarica, C. (2013). Investigation and prediction of highviscosity liquid effect on two-phase slug length in horizontal pipelines. SPE production & operations, 28(03), 296-305.
- Andreussi, P., Bendiksen, K., & Nydal, O. (1993). Void distribution in slug flow. International Journal of Multiphase Flow, 19(5), 817-828.
- Apostol, T. M. (1973). Mathematical analysis. Narosa Publishing House Pvt. Ltd.
- Arabi, A., Ragui, K., Salhi, Y., & Filali, A. (2020). Slug frequency for a gas-liquid plug flow: Review and development of a new correlation. International Communications in Heat and Mass Transfer, 118, 104841.

- ASTM E230 [DOI: 10.1520/E0230_E0230M-17]. (2017). ASTM International. https://www.astm.org/Standards/E230.htm
- Baba, Y. D., Aliyu, A. M., Archibong, A. E., Abdulkadir, M., Lao, L., & Yeung, H. (2018). Slug length for high viscosity oil-gas flow in horizontal pipes: Experiments and prediction. Journal of Petroleum Science and Engineering, 165, 397-411.
- Baba, Y. D., Archibong, A. E., Aliyu, A. M., & Ameen, A. I. (2017). Slug frequency in high viscosity oil-gas two-phase flow: Experiment and prediction. *Flow measurement and instrumentation*, 54, 109-123.
- Barnea, D. (1987). A unified model for predicting flow-pattern transitions for the whole range of pipe inclinations. *International journal of multiphase flow*, 13(1), 1-12.
- Bartle, R. G., & Sherbert, D. R. (1964). Introduction to real analysis. John Wiley & Sons, Inc.
- Baxendell, P., & Thomas, R. (1961). The calculation of pressure gradients in high-rate flowing wells. Journal of Petroleum Technology, 13(10), 1023-1028.
- Beggs, D. H., & Brill, J. P. (1973). A study of two-phase flow in inclined pipes. *Journal of Petroleum technology*, 25(05), 607-617.
- Bellman, R. E. (1978). Artificial intelligence: can computers think? Boyd Fraser publishing Company.
- Bendiksen, K. H., Malnes, D., Moe, R., & Nuland, S. (1991). The dynamic two-fluid model OLGA: Theory and application. SPE production engineering, 6(02), 171-180.
- Bendiksen, K. H., Malnes, D., & Nydal, O. J. (1996). On the modelling of slug flow. Chemical Engineering Communications, 141(1), 71-102.
- Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1798-1828.
- Bengio, Y., Goodfellow, I., & Courville, A. (2017). Deep learning (Vol. 1). MIT press Cambridge, MA, USA.
- Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. Advances in neural information processing systems, 24.
- Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford university press.
- Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer.
- Bonald, T. (2019). Expectation-maximization for the gaussian mixture model.
- Bonnans, J.-F., Gilbert, J. C., Lemaréchal, C., & Sagastizábal, C. A. (2006). Numerical optimization: theoretical and practical aspects. Springer Science & Business Media.
- Bouguila, N., & Fan, W. (2020). Mixture models and applications (Vol. 530). Springer.
- Brewka, G. (1996). Artificial intelligence—a modern approach by Stuart Russell and Peter Norvig, Prentice Hall. Series in Artificial Intelligence, Englewood Cliffs, NJ. The Knowledge Engineering Review, 11(1), 78-79.
- Brill, J. P., Schmidt, Z., Coberly, W. A., Herring, J. D., & Moore, D. W. (1981). Analysis of two-phase tests in large-diameter flow lines in Prudhoe Bay field. Society of Petroleum Engineers Journal, 21 (03), 363-378.

- Brill, J., & Arirachakaran, S. (1992). State of the art in multiphase flow. Journal of Petroleum Technology, 44 (05), 538-541.
- Brunton, S. L. (2021). Applying machine learning to study fluid mechanics. Acta Mechanica Sinica, 37(12), 1718-1726.
- Chollet, F. (2021). Deep learning with Python. simon; schuster.
- Coleman, H. W., & Steele, W. G. (2018). Experimentation, validation, and uncertainty analysis for engineers. John Wiley & Sons.
- COTNNMET. (2002). Guía para la expresión de la incertidumbre en las mediciones. NMX-CH-140-IMNC-2002. IMNC.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics* of control, signals and systems, 2(4), 303-314.
- Dascher, R. E. (1968). The measurement of liquid holdup in two-phase horizontal slug flow.
- Davies, R., & Taylor, G. I. (1950). The mechanics of large bubbles rising through extended liquids and through liquids in tubes. Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 200(1062), 375-390.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. Journal of the royal statistical society: series B (methodological), 39(1), 1-22.
- Dennis Jr, J. E., & Schnabel, R. B. (1996). Numerical methods for unconstrained optimization and nonlinear equations. SIAM.
- dos Reis, E., & Goldstein Jr, L. (2010). Characterization of slug flows in horizontal piping by signal analysis from a capacitive probe. *Flow Measurement and Instrumentation*, 21(3), 347-355.
- Duchi, J., Hazan, E., & Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, 12(7).
- Dukler, A. E., & Hubbard, M. G. (1975). A model for gas-liquid slug flow in horizontal and near horizontal tubes. *Industrial & Engineering Chemistry Fundamentals*, 14(4), 337-347.
- Dumitrescu, D. T. (1943). Strömung an einer Luftblase im senkrechten Rohr. ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik, 23(3), 139-149.
- Eaton, B. A., Knowles, C. R., & Silberbrg, I. (1967). The prediction of flow patterns, liquid holdup and pressure losses occurring during continuous two-phase flow in horizontal pipelines. *Journal of Petroleum technology*, 19(06), 815-828.
- Eggensperger, K., Feurer, M., Hutter, F., Bergstra, J., Snoek, J., Hoos, H., Leyton-Brown, K., et al. (2013). Towards an empirical foundation for assessing bayesian optimization of hyperparameters. NIPS workshop on Bayesian Optimization in Theory and Practice, 10(3), 1-5.
- Everitt, B. (1988). Finite mixture distributions. Springer Science & Business Media.
- Farsetti, S., Farisè, S., & Poesio, P. (2014). Experimental investigation of high viscosity oil-air intermittent flow. Experimental thermal and fluid science, 57, 285-292.
- Fletcher, R. (2000). Practical methods of optimization. John Wiley & Sons.

- Foletti, C., Farisè, S., Grassi, B., Strazza, D., Lancini, M., & Poesio, P. (2011). Experimental investigation on two-phase air/high-viscosity-oil flow in a horizontal pipe. *Chemical* engineering science, 66(23), 5968-5975.
- Garnett, R. (2023). Bayesian optimization. Cambridge University Press.
- Gilbert, W. (1954). Flowing and gas-lift well performance. Drilling and production practice.
- Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. Proceedings of the fourteenth international conference on artificial intelligence and statistics, 315-323.
- Gokcal, B., Al-Sarkhi, A., Sarica, C., & Al-Safran, E. M. (2010). Prediction of slug frequency for high-viscosity oils in horizontal pipes. SPE projects, facilities & construction, 5(03), 136-144.
- Gokcal, B., Wang, Q., Zhang, H.-Q., & Sarica, C. (2008). Effects of high oil viscosity on oil/gas flow behavior in horizontal pipes. SPE Projects, Facilities & Construction, 3(02), 1-11.
- Gomez, L., Shoham, O., & Taitel, Y. (2000). Prediction of slug liquid holdup: horizontal to upward vertical flow. *International journal of multiphase flow*, 26(3), 517-521.
- Gregory, G., Nicholson, M., & Aziz, K. (1978). Correlation of the liquid volume fraction in the slug for horizontal gas-liquid slug flow. International journal of multiphase flow, 4(1), 33-39.
- Guzhov, A., Mamayev, V., & Odishariya, G. (1967). A study of transportation in gas-liquid systems. International Gas Union.
- Guzmán, J., González-Treviño, J., Torres, L., Aragón-Rivera, F., Hernández-García, J., Palacio-Pérez, A., & Klapp, J. (2024). Probabilistic learning approach for the liquid holdup analysis of high-viscosity intermittent flows. *Physics of Fluids*, 36(3).
- Hagan, M. T., Demuth, H. B., & Beale, M. (1997). Neural network design. PWS Publishing Co.
- Hagedorn, A. R., & Brown, K. E. (1965). Experimental study of pressure gradients occurring during continuous two-phase flow in small-diameter vertical conduits. *Journal* of Petroleum technology, 17(04), 475-484.
- Hansen, L. S., Pedersen, S., & Durdevic, P. (2019). Multi-phase flow metering in offshore oil and gas transportation pipelines: Trends and perspectives. *Sensors*, 19(9), 2184.
- Haykin, S. S. (2009). Neural networks and learning machines: International Version. Pearson Education (US).
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.
- Hernandez, J., Montiel, J., Palacio-Perez, A., Rodríguez-Valdés, A., & Guzmán, J. (2018).
 Horizontal Evolution of Intermittent Gas-Liquid Flows With Highly Viscous Phases.
 Multiphase Flow: Theory and Applications; Witpress: Southampton, UK, 157.
- Hernández, J. (2019). Estudio del mecanismo del transporte en oleoductos horizontales de crudo pesado con presencia de inhibidores de arrastre [Tesis doctoral, UNAM].
- Hill, T., & Wood, D. (1994). Slug flow: Occurrence, consequences, and prediction. SPE University of Tulsa Centennial Petroleum Engineering Symposium, SPE-27960.

- Howell, K. B. (2016). Principles of Fourier analysis. CRC Press.
- Hughes, I., & Hase, T. (2010). Measurements and their uncertainties: a practical guide to modern error analysis. OUP Oxford.
- Hughmark, G., et al. (1962). Holdup in gas-liquid flow. Chemical Engineering Progress, 58(4), 62-65.
- Hughmark, G. A., & Pressburg, B. (1961). Holdup and pressure drop with gas-liquid flow in a vertical pipe. *AIChE Journal*, 7(4), 677-682.
- Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). Automated machine learning: methods, systems, challenges. Springer Nature.
- Ismail, I., Gamio, J., Bukhari, S. A., & Yang, W. (2005). Tomography for multi-phase flow measurement in the oil industry. *Flow measurement and instrumentation*, 16(2-3), 145-155.
- ISO. (1993). Guide to the Expression of Uncertainty in Measurement. (Corrected and reprinted 1995).
- Issa, R., & Kempf, M. (2003). Simulation of slug flow in horizontal and nearly horizontal pipes with the two-fluid model. *International journal of multiphase flow*, 29(1), 69-95.
- Jamieson, K., & Talwalkar, A. (2016). Non-stochastic best arm identification and hyperparameter optimization. *Artificial intelligence and statistics*, 240-248.
- Jarrett, K., Kavukcuoglu, K., Ranzato, M., & LeCun, Y. (2009). What is the best multistage architecture for object recognition? 2009 IEEE 12th international conference on computer vision, 2146-2153.
- JCGM. (2008a). 200:2008(E). International vocabulary of basic and general terms in metrology (VIM). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML.
- JCGM. (2008b). Evaluation of measurement data-Guide to the expression of uncertainty in measurement. (First edition). BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML.
- Jones Jr, O. C., & Zuber, N. (1975). The interrelation between void fraction fluctuations and flow patterns in two-phase flow. *International Journal of Multiphase Flow*, 2(3), 273-306.
- Kelley, C. T. (1995). Iterative methods for linear and nonlinear equations. SIAM.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Koopmans, L. H. (1995). The spectral analysis of time series. Elsevier.
- Kora, C., Sarica, C., Zhang, H.-Q., Al-Sarkhi, A., & Alsafran, E. (2011). Effects of high oil viscosity on slug liquid holdup in horizontal pipes. SPE Canada Unconventional Resources Conference, SPE-146954.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.
- Kukačka, J., Golkov, V., & Cremers, D. (2017). Regularization for deep learning: A taxonomy. arXiv preprint arXiv:1710.10686.
- Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. The annals of mathematical statistics, 22(1), 79-86.

- LeCun, Y. (1985). Une procedure d'apprentissage ponr reseau a seuil asymetrique. Proceedings of cognitiva 85, 599-604.
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2018). Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185), 1-52.
- LIF. (2024). Ley de Ingresos de la Federación para el Ejercicio Fiscal de 2025.
- Liu, Y. H., & Mehta, S. (2019). Hands-On Deep Learning Architectures with Python: Create deep neural networks to solve computational problems using TensorFlow and Keras. Packt Publishing Ltd.
- Mandhane, J., Gregory, G., & Aziz, K. (1974). A flow pattern map for gas—liquid flow in horizontal pipes. *International journal of multiphase flow*, 1(4), 537-553.
- Marsland, S. (2011). Machine learning: an algorithmic perspective. Chapman; Hall/CRC.
- Martínez-Palou, R., de Lourdes Mosqueira, M., Zapata-Rendón, B., Mar-Juárez, E., Bernal-Huicochea, C., de la Cruz Clavel-López, J., & Aburto, J. (2011). Transportation of heavy and extra-heavy crude oil by pipeline: A review. *Journal of petroleum science* and engineering, 75(3-4), 274-282.
- Matsubara, H., & Naito, K. (2011). Effect of liquid viscosity on flow patterns of gas-liquid two-phase flow in a horizontal pipe. *International Journal of Multiphase Flow*, 37(10), 1277-1281.
- McLachlan, G. J., Lee, S. X., & Rathnayake, S. I. (2019). Finite mixture models. Annual review of statistics and its application, 6(1), 355-378.
- McQuillan, K., & Whalley, P. (1985). Flow patterns in vertical two-phase flow. *International Journal of Multiphase Flow*, 11(2), 161-175.
- Michelucci, U. (2018). Applied deep learning. A case-based approach to understanding Deep neural networks.
- Minami, K., & Brill, J. (1987). Liquid holdup in wet-gas pipelines. SPE Production Engineering, 2(01), 36-44.
- Montgomery, D. C., & Runger, G. C. (2010). Applied statistics and probability for engineers. John wiley & sons.
- Montiel Cortés, J. C. (2017). Estudio de flujos bifásicos horizontales de mezclas bifásicas gas-líquido con fases de alta viscosidad. [Tesis de licenciatura]. UNAM.
- Mukherjee, H. (1979). An experimental study of inclined two-phase flow. The University of Tulsa.
- Müller, A. C., & Guido, S. (2016). Introduction to machine learning with Python: a guide for data scientists. .°'Reilly Media, Inc.".
- Murphy, K. P. (2022). Probabilistic machine learning: an introduction. MIT press.
- Murphy, K. P. (2023). Probabilistic machine learning: Advanced topics. MIT press.
- Nair, V., & Hinton, G. E. (2009). 3D object recognition with deep belief nets. Advances in neural information processing systems, 22.
- Ng, A. (2000). CS229 Lecture notes. CS229 Lecture notes, 1(1), 1-3.
- Nocedal, J., & Wright, S. J. (1999). Numerical optimization. Springer.
- Nydal, O., Pintus, S., & Andreussi, P. (1992). Statistical characterization of slug flow in horizontal pipes. *International journal of multiphase flow*, 18(3), 439-453.

- Nydal, O. J. (1991). An experimental investigation on slug flow [Tesis doctoral, University of Oslo].
- Okezue, C. (2013). Application of the gamma radiation method in analysing the effect of liquid viscosity and flow variables on slug frequency in high viscosity oil-gas horizontal flow. WIT Trans. Eng. Sci, 79, 447-461.
- Oppenheim, A. V. (1999). Discrete-time signal processing. Pearson Education India.
- Ortiz, F. L. T. (2024). Descifrando la Dinámica de Fluidos: El Papel del Aprendizaje Automático. *TIES, Revista de Tecnología e Innovación en Educación Superior*, (10), 1-11.
- Osman, E.-S. A. (2004). Artificial neural network models for identifying flow regimes and predicting liquid holdup in horizontal multiphase flow. *SPE production & facilities*, 19(01), 33-40.
- Pearl, J. (1988). Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann Publishers, Inc.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
- PEMEX. (2025). Estadísticas petroleras. https://www.pemex.com/ri/Publicaciones/ Paginas/IndicadoresPetroleros.aspx
- Pinto, N., Doukhan, D., DiCarlo, J. J., & Cox, D. D. (2009). A high-throughput screening approach to discovering good forms of biologically inspired visual representation. *PLoS* computational biology, 5(11), e1000579.
- Poettman, F. H., & Carpenter, P. G. (1952). The multiphase flow of gas, oil, and water through vertical flow strings with application to the design of gas-lift installations. *Drilling and Production practice*.
- Proakis, J. G. (2001). Digital signal processing: principles algorithms and applications. Pearson Education India.
- Roberts, M. J. (2007). Fundamentals of signals and systems. McGraw-Hill Science/Engineering/Math.
- Ros, N. (1961). Simultaneous flow of gas and liquid as encountered in well tubing. *Journal* of Petroleum Technology, 13(10), 1037-1049.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Sahu, N., & Babu, P. (2020). New derivation for Gaussian mixture model parameter estimation: MM based approach. arXiv preprint arXiv:2001.02923.
- Schwarz, G. (1978). Estimating the dimension of a model. The annals of statistics, 461-464.
- Shippen, M. E., & Scott, S. L. (2004). A neural network model for prediction of liquid holdup in two-phase horizontal flow. SPE Production & Facilities, 19(02), 67-76.
- Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems, 25.
- Soto-Cortes, G., Pereyra, E., Sarica, C., Torres, C., & Soedarmo, A. (2021). Signal processing for slug flow analysis via a voltage or instantaneous liquid holdup time-series. *Flow Measurement and Instrumentation*, 79, 101968.

- Spivak, M. (1967). Calculus. Cambridge University Press.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.
- Sugiyama, M. (2015). Introduction to statistical machine learning. Morgan Kaufmann.
- Sun, W., & Yuan, Y.-X. (2006). Optimization theory and methods: nonlinear programming (Vol. 1). Springer Science & Business Media.
- Taitel, Y., & Dukler, A. (1977). A model for slug frequency during gas-liquid flow in horizontal and near horizontal pipes. *International journal of multiphase flow*, 3(6), 585-596.
- Taitel, Y., Barnea, D., & Dukler, A. (1980). Modelling flow pattern transitions for steady upward gas-liquid flow in vertical tubes. AIChE journal, 26(3), 345-354.
- Taitel, Y., & Dukler, A. E. (1976a). A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow. AIChE journal, 22(1), 47-55.
- Taitel, Y., & Dukler, A. E. (1976b). A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow. AIChE journal, 22(1), 47-55.
- Ternyik IV, J., Bilgesu, H., & Mohaghegh, S. (1995). Virtual measurement in pipes: part 2liquid holdup and flow pattern correlations. SPE Eastern Regional Meeting, SPE-30976.
- Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013). Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 847-855.
- Tieleman, T., & Hinton, G. (2012). Lecture 6.5 RMSProp, COURSERA: Neural Networks for Machine Learning (Technical report) (Available online at http://www.cs.toronto. edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf). University of Toronto.
- Ujang, P. M., Lawrence, C. J., Hale, C. P., & Hewitt, G. F. (2006). Slug initiation and evolution in two-phase horizontal flow. *International Journal of Multiphase Flow*, 32(5), 527-552.
- Vince, M., & Lahey Jr, R. (1982). On the development of an objective flow regime indicator. International Journal of Multiphase Flow, 8(2), 93-124.
- Wallis, G. B. (1969). One-dimensional two-phase flow. McGraw Hill Book Co. Inc.
- Wallis, G. B. (1970). Annular two-phase flow—Part 1: a simple theory.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (1993). Probability and statistics for engineers and scientists (Vol. 5). Macmillan New York.
- Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. PhD thesis, Committee on Applied Mathematics, Harvard University, Cambridge, MA.
- Ying, X. (2019). An overview of overfitting and its solutions. Journal of physics: Conference series, 1168, 022022.
- Zeiler, M. D. (2012). Adadelta: an adaptive learning rate method. arXiv preprint arXiv:1212.5701.
- Zhai, S., Pang, S., Shao, J., Zhu, H., Zhang, X., & Liu, P. (2018). The FAC experimental research on vapor-water two-phase flow piping system of CAP1400 secondary loop. *MATEC Web of Conferences*, 207, 04001.

- Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2021). Dive into deep learning. arXiv preprint arXiv:2106.11342.
- Zhang, H., & Lan, H.-q. (2017). A review of internal corrosion mechanism and experimental study for pipelines based on multiphase flow. *Corrosion Reviews*, 35(6), 425-444.
- Zhang, H.-Q., Jayawardena, S. S., Redus, C. L., & Brill, J. P. (2000). Slug dynamics in gas-liquid pipe flow. J. Energy Resour. Technol., 122(1), 14-21.
- Zhang, H.-Q., Wang, Q., Sarica, C., & Brill, J. P. (2003). Unified model for gas-liquid pipe flow via slug dynamics—part 1: model development. J. Energy Resour. Technol., 125(4), 266-273.
- Ziemer, R. E., Tranter, W. H., & Fannin, D. R. (1998). Signals and systems: continuous and discrete. Prentice Hall.