UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY

ESCUELA DE CIENCIAS QUÍMICAS E INGENIERÍA

TÍTULO: USE OF ARTIFICIAL NEURAL NETWORKS FOR ESTIMATING WATER CONTENT OF NATURAL GAS MIXTURES TAKING INTO ACCOUNT HEAVY HYDROCARBONS CONTRIBUTION

Trabajo de integración curricular presentado como requisito para la obtención del título de Petroquímico

Autor:

Montenegro Madroñero Jhon Fabián

Tutor:

Ricaurte Fernández Marvin José, PhD

Co-Tutor:

Viloria Vera Dario Alfredo, PhD

Urcuquí, Julio 2021

SECRETARÍA GENERAL
(Vicerrectorado Académico/Cancillería)
ESCUELA DE CIENCIAS QUÍMICAS E INGENIERÍA
CARRERA DE PETROQUÍMICA
ACTA DE DEFENSA No. UITEY-CHE-2021-00023-AD

A los 7 días del mes de julio de 2021, a las 11:00 horas, de manera virtual mediante videoconferencia, y ante el Tribunal Calificador, integrado por los docentes:

Presidente Tribunal de Defensa	Dr. TAFUR GUISAO, JUAN PABLO , Ph.D.
Miembro No Tutor	Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.
Tutor	Dr. RICAURTE FERNANDEZ, MARVIN JOSE , Ph.D.

El(la) señor(ita) estudiante MONTENEGRO MADROÑERO, JHON FABIAN, con cédula de identidad No. 0401750419, de la ESCUELA DE CIENCIAS QUÍMICAS E INGENIERÍA, de la Carrera de PETROQUÍMICA, aprobada por el Consejo de Educación Superior (CES), mediante Resolución RPC-SO-39-No.456-2014, realiza a través de videoconferencia, la sustentación de su trabajo de titulación denominado: USE OF ARTIFICIAL NEURAL NETWORKS FOR ESTIMATING WATER CONTENT OF NATURAL GAS MIXTURES TAKING INTO ACCOUNT HEAVY HYDROCARBONS CONTRIBUTION, previa a la obtención del título de PETROQUÍMICO/A.

El citado trabajo de titulación, fue debidamente aprobado por el(los) docente(s):

Tutor	Dr. RICAURTE FERNANDEZ, MARVIN JOSE , Ph.D.

Y recibió las observaciones de los otros miembros del Tribunal Calificador, las mismas que han sido incorporadas por el(la) estudiante.

Previamente cumplidos los requisitos legales y reglamentarios, el trabajo de titulación fue sustentado por el(la) estudiante y examinado por los miembros del Tribunal Calificador. Escuchada la sustentación del trabajo de titulación a través de videoconferencia, que integró la exposición de el(la) estudiante sobre el contenido de la misma y las preguntas formuladas por los miembros del Tribunal, se califica la sustentación del trabajo de titulación con las siguientes calificaciones:

Tipo	Docente	Calificación
Miembro Tribunal De Defensa	Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.	10,0
Tutor	Dr. RICAURTE FERNANDEZ, MARVIN JOSE , Ph.D.	10,0
Presidente Tribunal De Defensa	Dr. TAFUR GUISAO, JUAN PABLO , Ph.D.	10,0

Lo que da un promedio de: 10 (Diez punto Cero), sobre 10 (diez), equivalente a: APROBADO

Para constancia de lo actuado, firman los miembros del Tribunal Calificador, el/la estudiante y el/la secretario ad-hoc.

Certifico que *en cumplimiento del Decreto Ejecutivo 1017 de 16 de marzo de 2020, la defensa de trabajo de titulación (o examen de grado modalidad teórico práctica) se realizó vía virtual, por lo que las firmas de los miembros del Tribunal de Defensa de Grado, constan en forma digital.*

MONTENEGRO MADROÑERO, JHON FABIAN
Estudiante

JUAN PABLO TAFUR GUISAO
Firmado digitalmente por JUAN PABLO TAFUR GUISAO
Fecha: 2021.07.13 15:07:52 -05'00'

Dr. TAFUR GUISAO, JUAN PABLO , Ph.D.
Presidente Tribunal de Defensa

MARVIN JOSE
RICAURTE
FERNANDEZ

Dr. RICAURTE FERNANDEZ, MARVIN JOSE , Ph.D.
Tutor

ALICIA ESTELA
SOMMER MARQUEZ

Dr. SOMMER MARQUEZ, ALICIA ESTELA , Ph.D.
Miembro No Tutor

CARLA SOFIA
YASELGA NARANJO

YASELGA NARANJO, CARLA
Secretario Ad-hoc

AUTORÍA

Yo, **JHON FABIÁN MONTENEGRO MADROÑERO**, con cédula de identidad 0401750419, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad de el/la autora (a) del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, mayo 2021.

Jhon Fabián Montenegro Madroñero
CI: 0401750419

AUTORIZACIÓN DE PUBLICACIÓN

Yo, **JHON FABIÁN MONTENEGRO MADROÑERO**, con cédula de identidad 0401750419, cedo a la Universidad de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, mayo 2021.

Jhon Fabián Montenegro Madroñero
CI: 0401750419

ACKNOWLEDGEMENTS

DEDICATION

I dedicate this thesis to my family, who have always supported me. Especially to my parents Fabián and Marta, for always supporting and taking care of me. To my brother and sisters, Carol, Oscar, Deysi, Paola and Melany who are the fundamental basis of my personal growth. To my grandparents, primarily to my grandfather Segundo who taught me the value of work. To my teachers, classmates and friends at Yachay Tech for being fundamental in my academic and personal formation.

Jhon Fabián Montenegro Madroñero

RESUMEN

El gas natural (GN) al ser extraído de los yacimientos trae consigo agua en fase vapor. El agua presente en el gas natural deber ser removida para prevenir consecuencias desastrosas como corrosión, problemas en tuberías, formación de hidratos, dificultades en procesos de compresión y transporte, entre otros. La estimación precisa del contenido de agua en mezclas de gas natural es la base para el diseño de procesos de deshidratación. Para este propósito, se han desarrollado varios métodos de cálculo de contenido de agua (rigurosos y semiempíricos). Los métodos disponibles se basan en datos de presión y temperatura, otros métodos incluyen la concentración de sulfuro de hidrogeno (H_2S), dióxido de carbono (CO_2), y metano. No se han reportado métodos para el cálculo de contenido de agua que consideren la alta concentración de hidrocarburos más pesados que el metano presente en mezclas de gas rico, condensado de gas, o gas licuado de petróleo (GLP). Además, algunos métodos disponibles para la estimación del contenido de agua en el gas natural tienen baja precisión y no son aplicables a las condiciones típicas de operación en el procesamiento de gas. Actualmente, con el desarrollo de la inteligencia artificial, métodos alternativos de estimación como las redes neuronales artificiales (RNA) han probado ser precisas en la estimación de datos para aplicaciones ingenieriles. RNA basan su estructura en el funcionamiento biológico de redes neuronales y pueden aprender desde un conjunto de datos previo a predecir nuevos datos. Por lo tanto, el objetivo de este proyecto de graduación es el desarrollo de un modelo predictivo basado en redes neuronales artificiales para la estimación precisa del contenido de agua en mezclas de gas natural, tomando en cuenta la composición de los hidrocarburos pesados presentes en la mezcla, en condiciones típicas de operación. Para este fin, datos experimentales de contenido de agua en GN fueron requeridos. Los datos experimentales fueron procesados y se diseñó la RNA, después etapas de validación con otros métodos predictivos fueron necesarias para verificar la aplicabilidad del método propuesto. Ecuador al ser un país productor de hidrocarburos, las herramientas para el cálculo de contenido de agua en mezclas de gas natural son de interés para garantizar la continuidad operacional en la producción de hidrocarburos gaseosos e instalaciones de procesamiento.

Palabras claves: gas natural, contenido de agua, red neuronal artificial, hidrocarburos pesados, procesos de deshidratación.

ABSTRACT

When natural gas (NG) is extracted from the reservoir, it brings water in a vapor phase. Water present in natural gas needs to be removed to prevent disastrous consequences such as: corrosion, problems in pipelines, hydrate formation, difficulties in compression and transport processes. Accurate estimation of the water content in natural gas mixtures is the basis of the dehydration process design. In this regard, many water content methods have been developed (rigorous and semi-empirical methods). Some methods are based on pressure and temperature data. Other ones include the concentration of hydrogen sulfide (H_2S), carbon dioxide (CO_2), and methane in natural gas. There are no reported methods for calculating water content that considers the hydrocarbons heavier than methane present in mixtures of rich gas, gas condensate, or liquefied petroleum gas (LPG). Furthermore, some available methods for estimating water content in natural gas have low accuracy, and not all apply to typical operating conditions for natural gas processing. With artificial intelligence development, alternative estimation methods such as artificial neural networks (ANN) have proven to be accurate in estimating data for engineering applications. ANN base their structure on the biological neural networks functioning and can learn from a set of previous data to predict new data. Hence, this graduation project aims to develop a predictive model based on artificial neural networks to precisely estimate water content in natural gas mixtures, taking into account the composition of the heavy hydrocarbons present in the mixture for typical gas processing conditions. For this purpose, experimental data from open literature of water content in NG was required. The experimental data were processed, and the ANN was designed. Then, validation stages with other predictive methods were necessary to check the proposed method's applicability. As Ecuador is a hydrocarbon-producing country, the engineering tools for calculating the water content in natural gas mixtures are of interest to guarantee operational continuity in gaseous hydrocarbon production and processing facilities.

Keywords: natural gas, water content, artificial neural network, heavy hydrocarbons, dehydration process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AAD	Average Absolute Deviation
ADAM	Adaptative Linear Momentum
AE	Absolute Error
ANN	Artificial Neural Network
AI	Artificial Intelligence
C_G	Correction for Gas Gravity
DEA	Diethanolamine
DEG	Diethylene glycol
DGA	Diglylamine
EG	Ethylene glycol
GD	Gradient descent
GH	Gas hydrates
lb/MMSCF	Pounds of Water Per Million Standard Cubic Feet of Gas
LMA	Levenberg-Marquardt Algorithm
LPG	Liquefied Petroleum Gas
MAE	Mean Absolute Error
MDEA	Methyldietamine
MEA	Monoethanolamine
mg/m^3	Milligrams of Water Per Cubic Meter of Gas
ML	Machine Learning
MMbl/yr	Million Barrels per Year
MMSCF/day	Million Standard Cubic Feet per Day
MSE	Mean Square Error
NG	Natural Gas
NGL	Natural Gas Liquids
NSE	Normalized Squared Error
PRM	Peng-Robinson-Modified-Panagiotopoulos-Reid Eos
Relu	Rectified Lineal Unit Function
SRKKD	Soave-Redlich-Kwong-Kabadi-Danner EoS
Tanh	Hyperbolic Tangent Function
TEA	Triethanolamine
TEG	Triethylene Glycol

CHAPTER I

1. INTRODUCTION

Natural gas (NG) is currently seen as one of the primary and most relevant energy sources from domestic to industrial use. So much so that for the year 2019, NG supplied 24.2% of global primary energy demand and at least 25% of the electrical energy generation [1]. The above percentages show the energy importance of natural gas. However, natural gas is not only energy; it is also the main feedstock for the petrochemical industry, used in the manufacture of fertilizers, plastics, nylons, polyesters, etc. [2].

Natural gas extracted from reservoirs contains light hydrocarbons, hydrogen sulfide (H_2S), carbon dioxide (CO_2), water (H_2O), and a variety of unwanted components [3]. Undesirable components like water can cause processing operations problems if not removed or reduced from natural gas streams. Dehydration technologies like molecular sieves and glycol absorbers are typically used to eliminate or reduce the water content in NG [4].

Water content in natural gas refers to the maximum amount of water vapor present in a specific volume of NG at a given pressure and temperature. Usually, water content is reported in pounds of water per million standard cubic feet of gas (lb/MMSCF) in English units, or milligrams of water per cubic meter (mg/m^3) in the International Systems of Units [5]. Original conditions of the reservoirs and NG extraction processes allow sufficient water in contact to saturate the gas [6]. An already saturated gas cannot increase its water content in the vapor phase at a given condition, but dehydration processes are necessary to reach gas industrial required specifications [7].

Estimating water content in natural gas is the critical step for the dehydration processes design. Significant errors in water content estimation in natural gas lead to inappropriate dehydration plant designs. These inadequate designs are reflected in economic losses [4]. Many methods for estimating the water content in NG have been developed. In general, these methods can be divided into two groups: 1) charts, empirical or semiempirical correlations, and 2) rigorous thermodynamic models. Charts and correlations are helpful to get a rapid estimation due to available input data (pressure and temperature), while

1

thermodynamic models present a greater precision, but their application in most cases requires computational procedures [8].

All methods base their calculations on gas pressure and temperature conditions. Some of them have chosen to add the contribution of H_2S, CO_2 and methane concentration. Also, limited ranges for pressure and temperature are proposed for a more accurate calculation of the water content in NG. Thus, accurate methods at low temperatures are not proper at high temperatures and vice versa, same for the pressure [8]. Although recognized models, like the McKetta-Wehe chart, proposed corrections based on natural gas molecular weight [9]. At the moment, methods that take into account the composition of heavy hydrocarbons in NG are not reported in the open literature.

In addition to the methods mentioned above, in recent years, models based on artificial intelligence (AI), specifically machine learning (ML), have emerged for use in solving a large variety of tasks [10]. Artificial neural networks (ANN) are ML models that demonstrate great relevance in prediction and data classification. ANN are based on biological neuron networks functioning, and they can learn to predict values or results established on the knowledge of previous data [11]. ANN have been used in the natural gas industry to predict dew points in dehydration processes with TEG [12]. Also in NG context, Al-Fattah and Startzman [13] developed an ANN to estimate natural gas production.

In Ecuador, around 100 million standard cubic feet per day (MMSCF/day) of natural gas are produced [14]. The natural gas produced in Ecuador is used for several applications that are: (1) for supplying NG to the power plant Termogas Machala located in El Oro province, (2) for supplying liquefied natural gas to Azuay province, specifically to the ceramic sector, (3) for supplying different industrial sectors of the provinces of Pichincha, Guayas, Azuay among others, (4) for supplying domestic natural gas to around 350 families in Bajo Alto province of El Oro [15, 16]. The rest of the natural gas produced is used in energy production plans for supplying EP Petroecuador facilities, and its excess is burned in flares [17]. Besides, Ecuador produces around 1.06 million barrels per year (MMbl/yr) of liquefied petroleum gas (LPG) directly from associated natural gas [18].

This work focuses on developing an alternative and accurate predictive model of water content based on artificial neural networks, taking into account heavy hydrocarbon composition in natural gas. The model developed is intended to be applicable under

typical pressure and temperature conditions in NG processing facilities for various gas types. For this purpose, an intense search of open experimental water content data and available methods was carried out. Then data processing was necessary. After processing the experimental data, the artificial neural network was developed and structured with artificial intelligence software tools. Finally, a group of natural gas mixtures was used to study the developed model's accuracy and applicability, and its results were compared with other available methods for estimating water content in natural gas.

In Ecuador, natural gas and its derivatives (LGP) make it possible to supply part of the domestic market and reduce the importation of petroleum derivatives [15]. The predictive tools for calculating water content in natural gas guarantee operational continuity in gaseous hydrocarbon production and processing facilities in the country.

1.1. Problem Approach

When natural gas is extracted from the reservoir, it is accompanied by water vapor. The amount of water present in natural gas must be removed or decreased during NG processing. If the amount of water in the NG is not reduced or eliminated, its presence can lead to operational processes problems and reduce the gas quality [5]. In Figure 1, the main problems are shown that water can produce during the processing, transportation, and storage of natural gas.
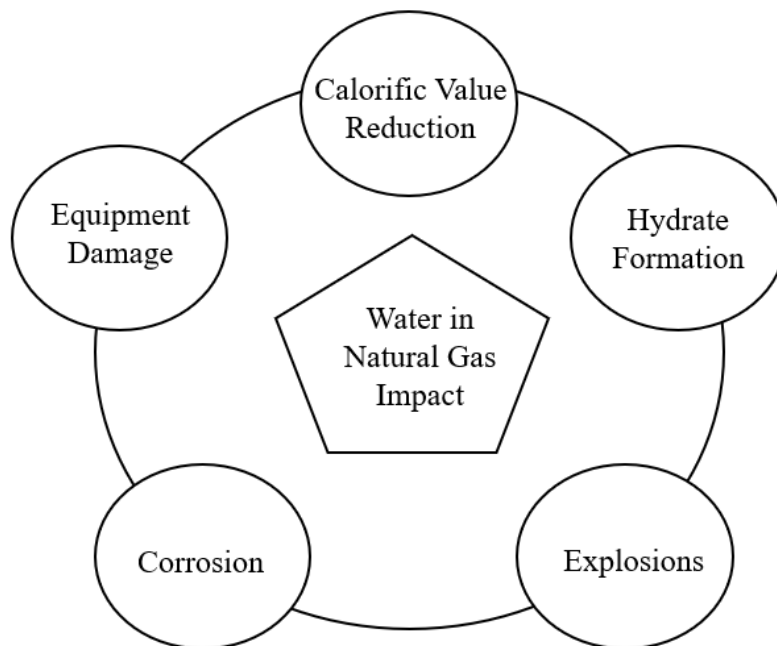
Figure 1. *Emanating problems from the presence of water in natural gas*

- **Corrosion:** variations in pressure and temperature can cause condensation of the water present in natural gas. The liquid water phase can interact with acid gases and may lead to corrosion.

- **Hydrate formation:** gas hydrates (GH) are crystalline solids that are formed when water (or ice) traps small molecules of gases such as methane, ethane, propane, carbon dioxide or hydrogen sulfide under certain conditions of pressure and temperature [19]. Gas hydrates generate plugging of flow pipes.

- **Equipment damage:** water in the liquid phase, even in small quantities, can damage equipment that works specifically with gases and vapor phase, such as compressors. Corrosion and gas hydrates also cause damage to equipment.

- **Calorific value reduction**: water has no calorific value; therefore, its presence in high quantities reduces natural gas's calorific value.

- **Explosions:** the gas used in thermoelectric plants, boilers, and other natural gas combustion processes must strictly comply with a minimum amount of water. The water present in the gas used for combustion can boil violently, increasing the pressure considerably. These pressure increases could damage the structure of the boiler, eventually causing explosions.

In order to avoid the problems mentioned above, a dehydration process must be used. Regardless of which one, the dehydration process is optimally designed to meet the specifications required for processing, using, and commercializing natural gas and its products like LPG [20]. For any dehydration process design, the water content in the natural gas must be known at a given temperature and pressure conditions; in this need, many methods for determining water content in natural gas have been developed [21]. The problem with the available methods is most of them have low precision and have been developed based on scarce experimental data. This has meant some available methods cannot be applied in the entire range of typical pressure and temperature conditions to which natural gas is subjected during its production and processing. Further, no method allows taking into account the compositional classification of natural gas, except for H_2S, CO_2 and methane. This way, available methods could not have a good prediction of water content in mixtures with a large composition of heavy hydrocarbons such as rich gases, LPG and condensed gases.

In this context, this work seeks to develop a predictive and precise model (ANN) for calculating the water content in natural gas, alternative to available methods considering the impact of the heavy hydrocarbons present in the NG. The model must also be applicable in a range of typical pressure and temperature conditions to which natural gas is exposed during its production and processing. Besides, it is sought that the method developed applies to most types of natural gas.

1.2. Objectives

1.2.1. General Objective

To develop an artificial neural network that calculates the water content in natural gas mixtures, taking into account the heavy hydrocarbons contribution, under operating typical natural gas processing conditions.

1.2.2. Specific Objectives

- To search and process experimental data of water content in natural gas mixtures.

- To analyze the effect that heavy hydrocarbons have on natural gas mixtures water content based on available methods.

- To analyze the operating conditions for which the artificial neural network will be applied.

- To compare the developed artificial neural network's performance and applicability with available methods in different types of natural gas mixtures.

CHAPTER II

2. BACKGROUND AND LITERATURE REVIEW

2.1. Natural Gas

Natural gas is a mixture that contains different hydrocarbon and non-hydrocarbon constituents. On the one hand, the hydrocarbon compounds in natural gas are methane (C_1), ethane (C_2), propane (C_3), butanes (nC_4, iC_4), pentanes (nC_5, iC_5), hexane (C_6), heptane and heavier hydrocarbon traces (C_7^+). On the other hand, non-hydrocarbon compounds in natural gas are water (H_2O), hydrogen sulfide (H_2S), carbon dioxide (CO_2), and nitrogen (N_2). The compositional percentage of each of the natural compounds varies according to the type of gas [22].

2.1.1. Types of Natural Gas

Natural gas can mainly be classified by its source and composition [23].

- **Classification by source**
  - *Associated natural gas:* it is the gas that is extracted together with oil and contains large amounts of hydrocarbons, such as ethane, propane, butane, and others.

  - *Non-associated natural gas:* it is the gas found in deposits that do not contain crude oil. The most abundant component in non-associated natural gas is methane.

- **Classification by composition**
  - *Sour gas:* contains considerable amounts of carbon dioxide and sulfur derivatives (hydrogen sulfide, mercaptans, sulfides and disulfides).
  - *Sweet gas:* it is considered sweet gas when it does not have sulfur compounds or its composition is below the regulations' values (< 4 ppmv).
  - *Rich gas:* it is the gas that contains significant amounts of hydrocarbons heavier than methane; generally, it is the associated natural gas.

6

o *Lean gas:* gas that contains smaller amounts of hydrocarbons heavier than methane, usually is non-associated natural gas.

The mole composition in natural gas of hydrocarbon and non-hydrocarbon compounds is enormously varied and will depend on the gas source. In Ecuador, varieties of natural gas are produced. The most representative mixtures for their usefulness are the natural gas from the Shushufindi industrial field, which feeds the same field's gas processing plant and natural gas from the Amistad offshore field [24]. Table 1 shows the composition of different natural gas in some regions worldwide and representative samples from Ecuador.

***Table 1.** Typical gas composition in some regions*

Natural Gas Composition (mole %)					
	Canada (Alberta)	Bach Ho (Field Vietnam)	Miskar (Field Tunisia)	Rio Arriba (New Mexico)	Ecuador (Field Shushufindi)	Ecuador (Field Amistad)
He	-	-	-	-	-	-
N_2	3.20	0.21	16.90	0.68	3.80	0.61
CO_2	1.70	0.07	13.59	0.82	12.66	0.12
H_2S	3.30	-	0.09	-	-	-
C_1	77.10	71.79	63.95	96.91	45.20	98.81
C_2	6.60	13.59	3.34	1.33	7.45	0.37
C_3	3.10	7.60	0.96	0.19	22.59	0.09
iC_4	1.00	2.03	-	-	2.74	-
nC_4	1.00	2.05	0.54	0.05	5.56	-
iC_5	1.50	1.66	-	-	-	-
nC_5	1.50	1.00	0.63	0.02	-	-

Source: Adapted from Kidnay and Parrish [2]; Vicuña and Parreño [24].

The "richness" or commercial appreciation of natural gas is evaluated by the GPM [25]. GPM is a unit of measurement and represents the volume of liquid that can be obtained per thousand standard cubic feet of natural gas, generally expressing the amount of the compounds, such as propane, butanes and other heavy components, which in practice can be obtained as liquids (generally GPM greater than two (GPM > 2) are considered rich gases) [26]. Table 2 summarizes the procedure for calculating GPM from the composition of the natural gas mixture.

Table 2. *GPM calculation*

Component	Mole fraction	Gal/mole	GPM per Component
He	0.000	-	-
N_2	0.010	-	-
CO_2	0.030	-	-
H_2S	0.000	-	-
C_1	0.850	-	-
C_2	0.058	10.126	1.548
C_3	0.030	10.433	0.825
iC_4	0.007	12.386	0.228
nC_4	0.008	11.937	0.252
iC_5	0.003	13.860	0.110
nC_5	0.002	13.713	0.072
C_6	0.002	15.566	0.082
C_7^+	0.000	17.464	0.000
GPM Mixture			3.116

Source: Adapted from Engineering Data Book [25].
Note: To obtain the GPM per component, multiply Gal/mol by the mole fraction of each component and divide by 379.49 SCF/mol to get gallons per standard cubic foot of the gas; then multiply by 1000. GPM mixture will be the sum of all GPM per component.

With a GPM of 10.80, the Shushufindi industrial field's natural gas is a rich gas with significantly heavier hydrocarbons than methane present in its composition. On the other hand, the GPM of natural gas from field Amistad is 0.12, being a lean gas where its composition is mostly methane.

2.1.2. Natural Gas Processing

Natural gas processing is essentially comparable to crude oil refining. The objective of gas processing is to condition the crude natural gas from the well into sales gas for delivering to customers [27]. Generally, natural gas processing begins with the removal of impurities, which are mainly non-hydrocarbon compounds. Fractionation stages are also necessary to separate hydrocarbons from obtaining compounds with greater added value [9]. In Figure 2, a generalized block diagram shows the main stages of natural gas processing.
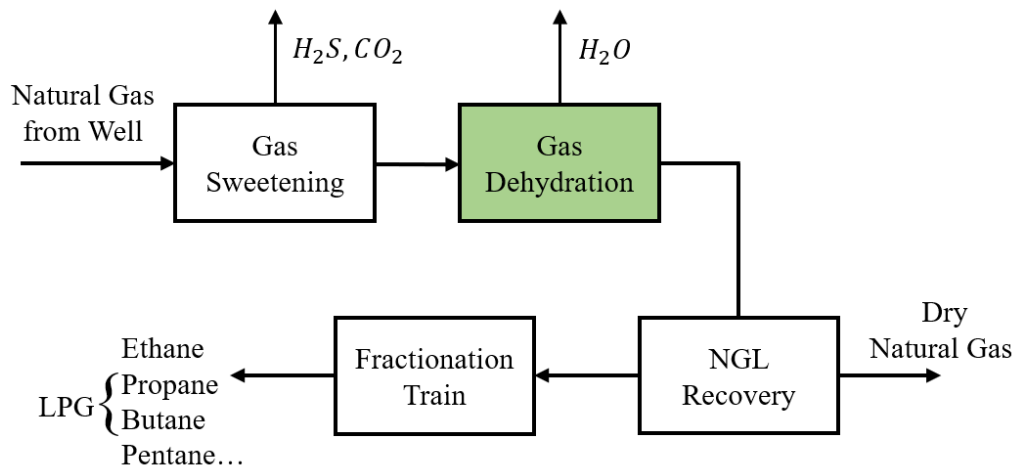
Figure 2. *Generalized natural gas processing block diagram*

- **Gas Sweetening**: natural gas sweetening is a process by which the content of acid gases such as CO_2 or H_2S, is reduced. Acid gases in natural gas can cause problems in the handling and processing stages, such as corrosion, poisoning, harmful odors, and emissions of compounds that cause acid rain. Normally, quality standards are $CO_2 < 2$ vol.%, and $H_2S < 4$ ppmv [2]. There are several methods for the removal of acid gases from a natural gas mixture. These include physical and chemical methods. In physical processes, a substance is used as a solvent and absorbs acid gases from natural gas as gas in solution without chemical reactions [28]. Usually, the solvent used is regenerated with pressure reduction and the application of heat. In chemical processes, the natural gas to be treated is contacted countercurrently with a solution of an active component that reacts with acid gases to form complexes, soluble in the solvent. The solvents most used in the gas industry for chemical sweetening natural gas are aqueous solutions of amines and water (30 - 70 vol% respectively). The amines used commercially for gas treating are: monoethanolamine (MEA), diethanolamine (DEA), triethanolamine (TEA), diglylamine (DGA), methyldietamine (MDEA) [25].

- **Gas Dehydration**: refers to the processes for eliminating or reducing water content in natural gas. Section 2.2. explains in detail the natural gas dehydration process.

- **Natural Gas Liquids Recovery**: the recovery of natural gas liquids (NGL) involves processes to separate ethane and heavier hydrocarbons from methane; this process produces a quality lean natural gas to transport through pipelines [26]. The two main techniques for recovering NGL are absorption and mechanical refrigeration. On the one hand, in the absorption method, absorbent oil is used,

which traps NGL. On the other hand, the mechanical refrigeration method is the simplest and most straightforward process for NGL recovery on a large scale. This method is provided by a refrigeration cycle, which usually uses propane as the refrigerant [29]. Once the NGL has been recovered from the natural gas stream, they must be separated or fractionated.

- **Fractionation Train:** the fractionation of natural gas liquids is used in gas processing plants to separate heavy hydrocarbons mixed in natural gas and obtain individual products. These hydrocarbons are ethane, propane, n-butane, isobutane, and pentanes with heavier compounds, called natural gasoline. The separation of these compounds is carried out by relative volatility [3]. The compounds' facilities can be separated depending on the relative volatility within the mixture and the products' required quality. A fractionation train (Figure 3) is made up of many columns or separators. The columns' number depends on components to be separated from the hydrocarbon mixture in the feed [2].
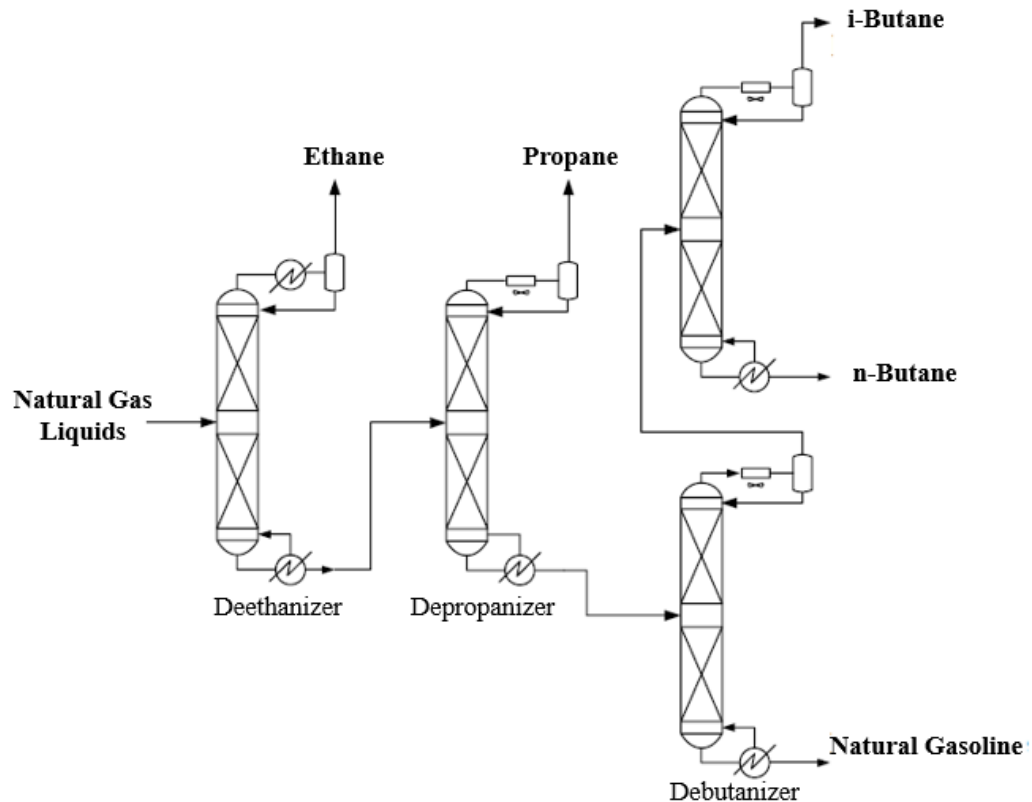
Figure 3. *Typical fractionation train in natural gas processing*

2.2. Dehydration Process

Water present in natural gas at temperatures below its dew point tends to condense. In natural gas mixtures, two types of dew points can be distinguished: the hydrocarbon dew point and the water dew point [20]. Usually, for a saturated rich natural gas, the water dew point can be present at temperatures lower than hydrocarbons' dew point. On the contrary, in saturated lean natural gas mixtures, the dew point occurs at higher temperatures (Figure 4) [24]. In terms used in engineering, dehydrating a gas reduces its water dew point, so the lower the water's dew point, the lower the water content in the gas. The main reasons for removing water vapor from natural gas are the following:

- o Liquid water and natural gas can form hydrates that plug equipment and pipes,
- o Liquid water from natural gas is corrosive, mainly if it contains CO_2 and H_2S,
- o Water vapor from natural gas can condense in lines, causing plugging,
- o For optimizing the operation of the compressors,
- o For achieving the quality required for its transport in pipes and commercialization [4].

The water content specification in natural gas is commonly 4 to 7 lb/MMSCF [2]. The Ecuadorian technical regulation INEN 2 489:2009 establishes a maximum limit of 4 lb/MMSCF for processed natural gas [30].
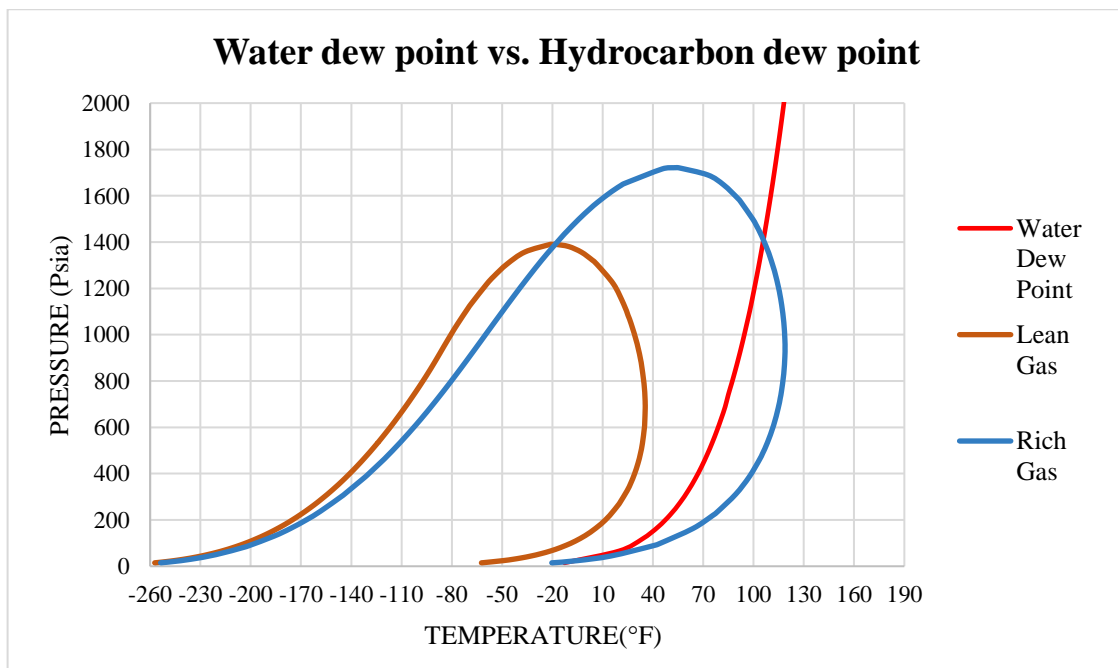
Figure 4. *Phase envelope - rich gas and lean gas*

In order to meet the quality standards of water content in natural gas, natural gas must be subjected to a dehydration process. The dehydration processes most used in the natural gas industry are absorption and adsorption processes [31].

2.2.1. Absorption Processes

Absorption is a mass transfer phenomenon from a gas phase to a liquid phase; this is possible by adding a liquid with high hygroscopicity or ease of removing water. Absorption dehydration consists of removing water vapor from the gas through intimate contact with a liquid desiccant [25]. The contact takes place in a packed or dishes tower. Glycols are the most effective liquid desiccants; the most used for the dehydration of natural gas are: ethylene glycol (EG), diethylene glycol (DEG), and triethylene glycol (TEG) [4].

Triethylene glycol has gained universal acceptance as the most effective glycol because it achieves greater moisture remotion, has low evaporation losses and operating costs. It is a nonvolatile, odorless, colorless and hygroscopic liquid [32]. TEG has been used to dehydrate sweet and sour gases in operating ranges of pressure (25-2500 psia) and temperature (40-160 ºF) [25]. The disadvantages of using TEG are that it is not suitable for cryogenic processes, and it is corrosive if contaminated with H_2S [33].

A dehydration plant that uses TEG (Figure 5) comprises two zones: the dehydration zone, which is favored by high pressures and low temperatures, and the regeneration zone, which is favored by low pressures and high temperatures. There are two complementary operations. The first refers to cleaning the wet natural gas that enters the absorption tower, and the second corresponds to the decontamination of the glycol to prevent impurities from reaching the reboiler [2]. It should be noted that the liquid-vapor phase equilibrium governs the absorption of water vapor and the concentration of glycol.
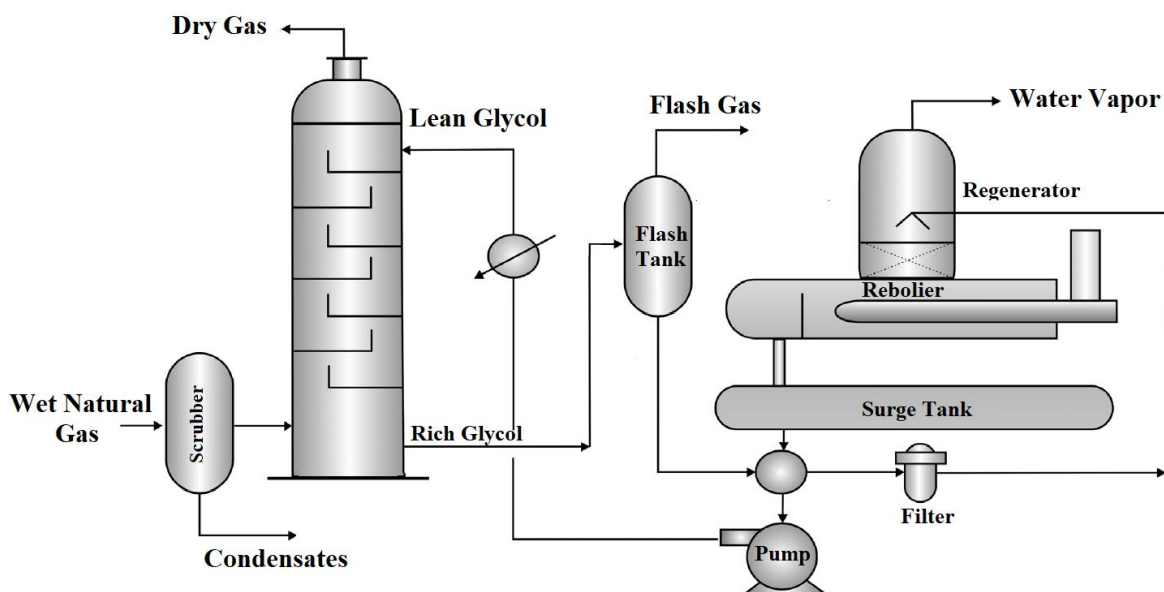
Figure 5. *Diagram of a glycol dehydration plant*
Source: Adapted from Engineering Data Book [25].

The regenerated or lean glycol is continuously pumped to the absorber tower's upper plate, where it flows through the liquid downcomers that connect each plate or contact tray. The natural gas rises with water vapor content and flows through the bubble capsules or the valves placed in each plate [32]. This process is repeated at each contact plate: the descending liquid absorbs the water vapor, and the ascending natural gas is dehydrated.

The following information must be available to design glycol dehydration plants: gas inlet flow, gas specific gravity, operating pressure, maximum contact tower working pressure, gas inlet temperature, "water content of inlet natural gas," and outlet gas specification [3]. Similarly, two fundamental design criteria must also be taken into account for an adequate plant performance: glycol/water ratio and lean TEG concentration. A value of 2 to 6 gallons of TEG per pound of water removed (gal TEG/lb H_2O) is considered appropriate for most glycol dehydration process requirements. A value of 3 gal TEG/lb H_2O is most often used in dehydrators [25]. Normally, the concentration of lean TEG is 98 to 99.5% by weight. To obtain a TEG purity greater than 98% by weight, heating to 400 °F and atmospheric pressure are required in the regenerator [31]. To avoid the presence of liquid (water, heavy hydrocarbons) in the natural gas feed, a stripper separator (Scrubber) must be placed before the contactor or absorber [3].

In the operation of dehydration plants by glycol may occur operational problems such as:

o Glycol contamination due to the presence of suspended impurities, which can generate foam in the contact tower. Thus, glycol carryover by gas can occur.

o Pumping problems when handling low temperatures which increase the viscosity of highly concentrated glycol solutions.

o Glycol losses should be controlled to less than 0.1 gallons for each MMSCF of natural gas treated, avoiding temperatures above 160 °F and foam formation in the contactor [25].

Despite the problems associated with its operation, glycol absorption is the traditional gas dehydration method since it has proven to be efficient in achieving the required quality specifications, and its operation is well known [32]. In Ecuador, the Bajo Alto dehydration plant located in El Oro province is an absorption plant that uses TEG to dehydrate the natural gas from the Amistad field. Once natural gas is dehydrated goes to the thermoelectric plant Termogas Machala [34].

2.2.2. Adsorption Processes

Dehydration with solid desiccants is a process that works under the principle of adsorption. Adsorption involves a form of adhesion between the solid desiccant particles and the water vapor in the gas. Dehydration with solids is much more efficient than dehydration with glycol; with this technique, water content of 0.05 lb/MMSCF is reached [3]. However, to reduce the adsorber's size, a glycol absorber is often used to perform initial dehydration, thereby reducing the mass of solid desiccant required for final dehydration [35].

Solid bed dehydration is a good alternative to glycol dehydration in applications such as:
o Dehydration to achieve dew points of water less than -40 °F to -58 °F,
o Hydrocarbon dew point control units where the simultaneous extraction of water and hydrocarbon are required to achieve both sales specifications. Solid desiccant is frequently used to control the hydrocarbon dew point in high-pressure lean gas streams [9],
o Simultaneous dehydration and H_2S removal from natural gas is required,
o Dehydration of gases containing H_2S, where the solubility of H_2S in glycol can cause corrosion problems,

   o Dehydration and sulfide components (H_2S, COS, CS_2, mercaptan) removal from NGL or LPG streams [36].

Desiccants must possess a large affinity for water, a large surface area per unit volume, high mechanical strength, abrasion resistance, be chemically inert, and have reasonable costs. The most used solid desiccants are silica gel, alumina, and molecular sieve [37]. A comparison of the physical properties of each desiccant is shown in Table 3.

In natural gas dehydration processes by adsorption systems, the most widely used desiccants are molecular sieves, which are crystalline forms of aluminosilicates that exhibit a high degree of water adsorption. They provide even a dew point of -150 °F and can be used simultaneously to sweet and dehydrate natural gas. Molecular sieves in their structure form cavities connected by uniform pores with diameters from 3 to 10 Å, depending on the type of sieve [4]. Since they are made according to specific pore size, molecular sieves allow adsorption to be selective; only molecules whose diameter is smaller than the molecular sieve's pore size are adsorbed [25].

Table 3. Comparison of the physical properties of desiccants used for dehydration of NG

Properties	Silica gel	Alumina	Mol. sieves
Specific area [m^2/g]	750-830	210	650-800
Pore volume [cm^3/g]	0.4-0.45	0.21	0.27
Pore diameter [Å]	22	26	3-10
Design capacity [kg H_2O/100 kg desiccant]	7-9	4-7	9-12
Density [kg/m^3]	721	800-880	690-720
Heat capacity [J/kg/°C]	920	240	200
Regeneration temperature [°C]	230	240	290
Heat of desorption [J]	3256	4183	3718

  Source: Adapted from Netusil and Ditl [4].

Figure 6 shows the typical gas dehydration process with solid desiccant. When the wet gas enters the plant, it passes through a separator (Scrubber) to remove all solids and liquids. Subsequently, the gas flows to the top of the adsorber containing a desiccant bed. While one adsorber tower is dehydrating, the other regenerates by a stream of hot gas [25]. The gas to be processed passes through the adsorbent bed during the adsorption stage, where the water is selectively retained. When the bed becomes saturated, a stream of hot gas passes in counterflow to the adsorbent bed for regeneration. After regeneration and before adsorption, the bed must be cooled, and this is achieved by circulating cold gas

through the adsorption bed in the same flow direction; later, the same gas can be used for the regeneration process [3]. The change of beds is carried out using a time controller, which executes the operations' changes at certain times depending on the cycle; however, the different stages' duration varies considerably [38].

In Ecuador, the Shushufindi Gas Plant has two adsorbers of molecular sieves formed by synthetic zeolites. These molecular sieves are used to dehydration the associated natural gas that comes from the field with the same name (Shushufindi). Regeneration of the molecular sieves in this plant is carried out using residual gas, which is preheated at 500 °F for 6 hours [18].
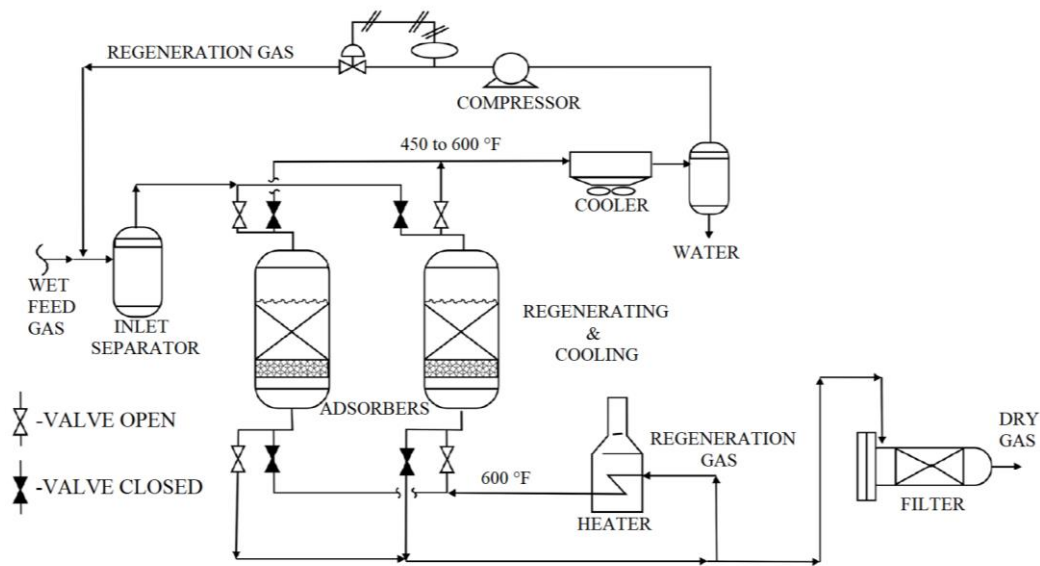
Figure 6. *Diagram of the dehydration process with solid desiccants*
Source: Adapted from Engineering Data Book [25].

Design of plants for the dehydration of natural gas by adsorption necessarily requires the precise estimation of the "water content" in the gas to be dehydrated. In molecular sieves, the water content in saturated natural gas is one of the most important factors for sizing the adsorber, as is the flow to be treated and the expected lifetime of the molecular sieve [21].

2.3. Methods for Calculating Water Content in Natural Gas

Natural gas extracted from the well can contain water in vapor form and free water associated with condensates. Besides, the sweetening processes with aqueous amines

16

solutions provide the natural gas's necessary conditions to become saturated in water at a given pressure and temperature [39]. The water vapor content of natural gas increases with increasing temperature and decreases with increasing pressure. A saturated natural gas's water content is on the order of a hundred or thousand pounds of water per million standard cubic feet of gas (lb/MMSCF). On the contrary, the specification of the maximum water content for transportation through gas pipelines and the sale of natural gas is in the order of 4 to 7 lb/MMSCF [25]. The dew point of water in natural gas is an indirect measure of natural gas's water content. The dew point is the temperature at which natural gas is saturated with water vapor at a certain pressure. At the dew point, the gas is in equilibrium with liquid water. Any decrease in temperature or increase in pressure would cause the water vapor in natural gas condensation [40].

In natural gas treatment processes, the so-called dew point depression is defined as the difference between the dew temperature of the gas saturated with water vapor and the dew point of the same gas after the dehydration process. To evaluate any natural gas dehydration process, by convention, the start point is to determine the saturated natural gas's water content [33]. Experimental determination of the water content in natural gas is the most accurate way to calculate it. However, the experimental calculation of water content is very expensive and time-consuming. Thus, the role of methods for predicting water content in natural gas becomes of utmost importance [41]. The methods for estimating water content can be divided into empirical such as charts and correlations; and theoretical, which are rigorous thermodynamic models [4]. Some suitable methods for estimating water content in natural gas mixtures are showed below.

2.3.1. McKetta and Wehe Chart

The most widely used manual method to quantify the water content in natural gas is the McKetta and Wehe chart. McKetta and Wehe published in 1958 the chart for estimating the water content in sweet natural gas. This chart for its importance has been reproduced in many publications, but especially in the GPSA Engineering Data Book [40]. The chart was generated from empirical data, limited to sweet gases and should not be used for sour gas compositions greater than 5% mol of $H_2S + CO_2$ [25]. Figure 7 shows the pressure-temperature chart performed by McKetta and Wehe. The pressure and temperature

conditions are needed to use this chart, and the point where they intersect, taken to the sides, indicates the water content in lb/MMSCF.

Figure 7. McKetta and Wehe chart

Source: Adapted from Engineering Data Book [25].

19
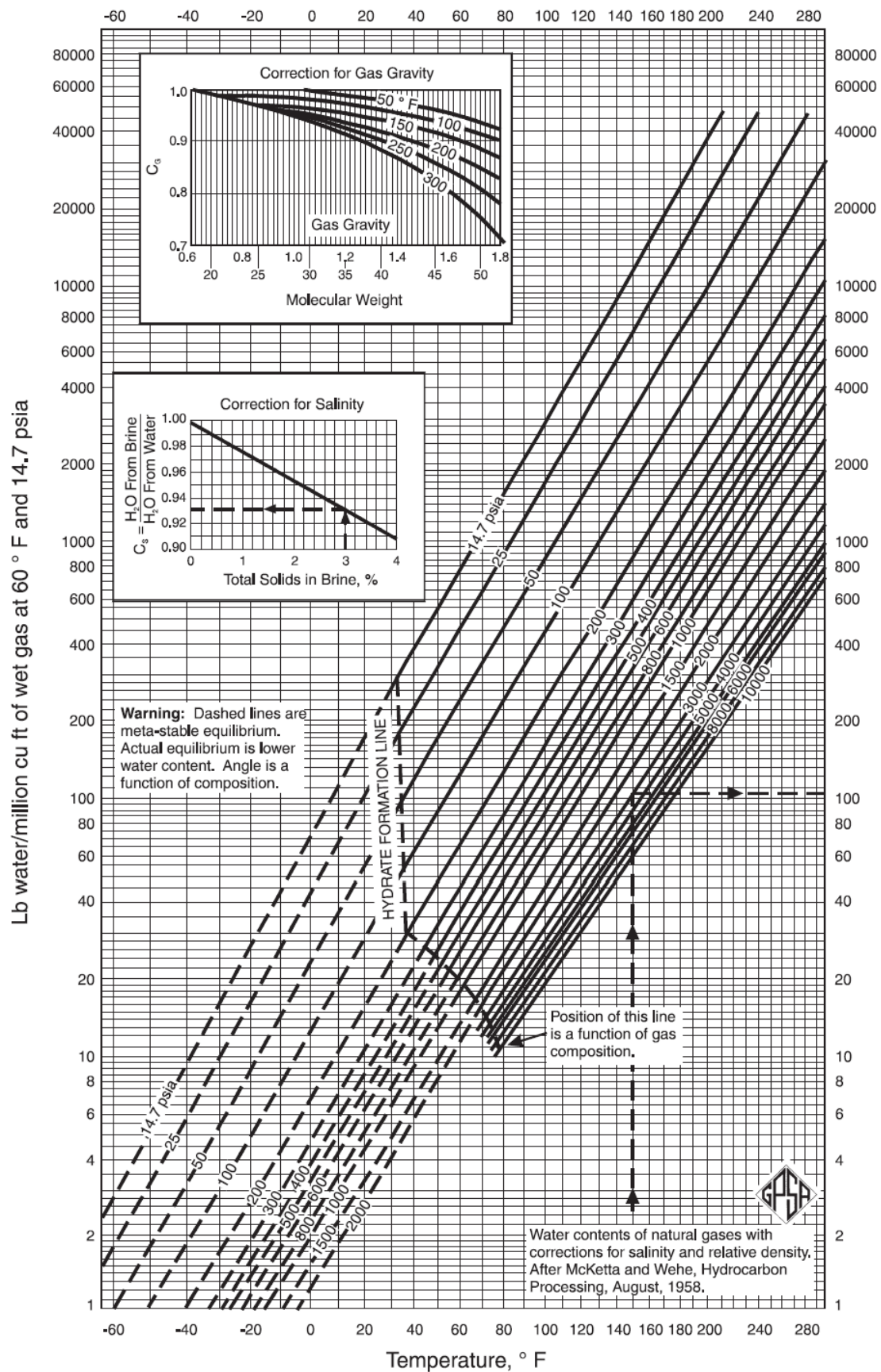
The McKetta and Wehe chart also presents two corrections to the final value of the calculated water content; the first correction (C_G) is a function of the gas specific gravity. The second is a function of the salts content of the brine.

2.3.2. Bukacek Correlation

Several empirical correlations have been proposed from the McKetta and Wehe chart publication; one is the Bukacek correlation. This correlation uses two mathematical expressions (Eqs. (1) and (2)).

$$WC_{H_2O} = 0.016 \left(47484 \frac{P_{sw}}{P} + B \right) \tag{1}$$

$$logB = \frac{-3083.87}{459.6 + 1.8(t + 32)} + 6.69449 \tag{2}$$

Where WC_{H_2O} is the water content in natural gas (in g/m³); P_{sw} is the water saturated vapor pressure (in kPa); P is the pressure of the natural gas (in kPa); B is the correction term and, t is the temperature of natural gas (in °C). Naturally, these units can be extended to any system of units [42]. Correlation is valid only for sweet gases and for temperatures in a range 15.6 to 237.8 °C (60 to 460 °F) and pressures from 103.4 to 68947.7 kPa (15 to 10,000 psia). The apparent simplicity equation takes complexity and precision in the correct calculation of the water saturated vapor pressure [40].

2.3.3. Maddox Correction

Maddox et al., 1998 [43] presented a correction for calculating water content in sour natural gas mixtures. In their work, the water content in sour natural gas is assumed to be the sum of the contribution of sweet natural gas, H_2S, and CO_2 [44]. Eq. (3) shows the mathematical expression of this correction.

$$WC_{H_2O,sour} = y_{Hc} \cdot WC_{H_2O,sweet} + y_{CO_2} \cdot WC_{CO_2} + y_{H_2S} \cdot WC_{H_2S} \tag{3}$$

Where $WC_{H_2O,sour}$ is the water content in sour natural gas; y_{Hc}, y_{CO_2} and y_{H_2S} are the mol fraction of hydrocarbons, carbon dioxide, and hydrogen sulfide, respectively;

WC_{CO_2} and WC_{H_2S} are the water content of carbon dioxide and hydrogen sulfide [43]. $WC_{H_2O,sweet}$ can be obtained with McKetta-Wehe chart or the Bukacek correlation. While WC_{CO_2} and WC_{H_2S} are calculated by Eq. (4):

$$logWC_{NonHC} = a_0 + a_1 \cdot log\ P + a_2 \cdot (logP)^2 \tag{4}$$

NonHc stands for H_2S or CO_2, P is the pressure of the natural gas (in psia). The coefficients a_i are temperature dependent and are tabulated in Table 4.

Table 4. *Maddox a_i coefficients*

CO₂				H₂S			
T(K)	a_0	a_1	a_2	T(K)	a_0	a_1	a_2
299.82	2.202797	−1.058120906	0.3427	299.82	2.314489	−0.678582201	0.3004
310.93	2.496936	−1.036484877	0.3103	310.93	2.544338	−0.70422576	0.3046
327.60	2.930297	−0.990489109	0.24	327.60	2.890856	−0.799408896	0.3319
344.26	3.18733	−0.924518418	0.2139	344.26	3.241253	−0.896948037	0.3646
				377.60	3.773303	−1.030222461	0.4232
				410.93	4.277016	−1.255345485	0.4897

2.3.4. Wichert Chart

Wichert Chart is an empirical method for calculating water content in sour natural gas mixtures. It is necessary to know the molar concentration of sour gases (H_2S, CO_2) and the pressure and temperature of the gas to apply the model. This method adds a variable called %H_2S Equivalent, which integrates the molar concentrations of sour gases into a single concentration. Eq. (5) is the mathematical expression of the %H_2S Equivalent.

$$\%H_2S\ Equivalent = mol\%\ H_2S + 0.7mol\%\ CO_2 \tag{5}$$

This chart is applicable up to values of 350 °F and up to a pressure of 10,000 psia, for sour gas mixtures with an %H_2S Equivalent up to 50% [25]. The temperature is intersected with the % H_2S Equivalent; then a vertical line is drawn up until it intersects with the gas pressure; drawing a horizontal line to the left, the water content ratio value will be obtained. Finally, the acid gas's water content is cleared from the water content ratio, for which the water content of the sweet natural gas is required by the McKetta and Wehe chart. Figure 8 shows the model developed by Wichert.
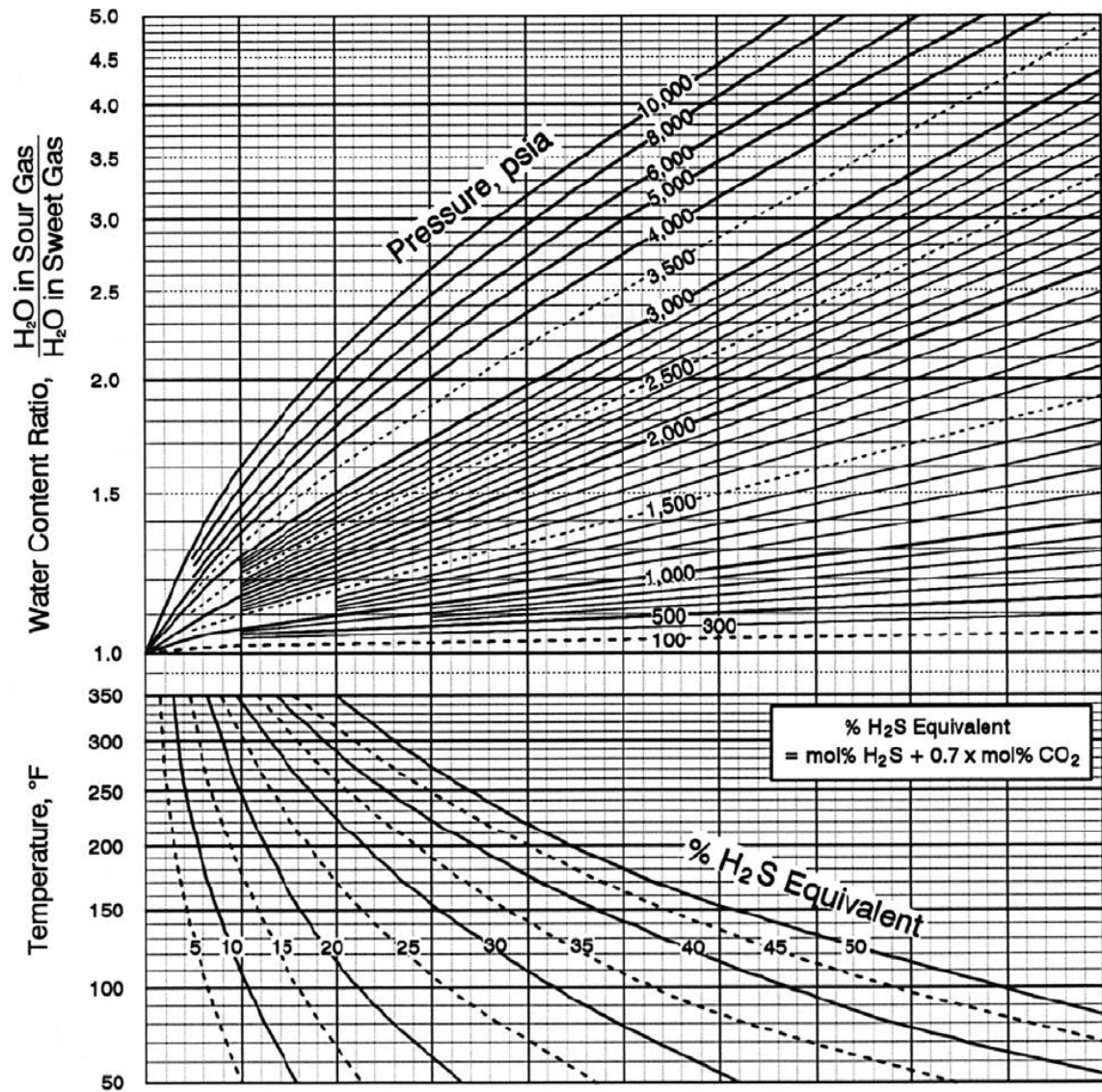
Figure 8. *Wichert Chart for calculating water content in sour gas*

Source: Adapted from Engineering Data Book [25].

2.3.5. Thermodynamic Models

Thermodynamic models for calculating water content in natural gas mixtures are based on the equality of chemical potentials. Although most thermodynamic methods have successfully predicted phase equilibrium for various systems, many of them have poor predictions for water-rich hydrocarbon systems [45]. Several authors have modified traditional Equations-of-State to successfully predict the phase equilibrium of water-rich hydrocarbon systems to solve this bad prediction. Two relevant models are cited below.

- SRK-Kabadi-Danner Equation-of-State (SRKKD EoS): Kabadi and Danner [45] presented a model based on the Soave-Redlich-Kwong Equation-of-State, with modifications in the mixing rules and the binary interaction parameters, to more adequately represent the liquid-vapor equilibrium of systems composed of hydrocarbons and water. They also developed a means to estimate water-hydrocarbon equilibrium when no data are available [46].

- PR-Modified-Panagiotopoulos-Reid Equation-of-State (PRM EoS): PRM Eos is a modification of the mixing rule for PR-Panagiotopoulos-Reid, which was initially developed from the Peng-Robinson Equation-of-State [47]. This thermodynamic model also fits the binary iteration parameters of the work developed by Panagiotopoulos and Reid. Software companies developed PRM for better prediction for Three-phase separators for water-hydrocarbon systems in commercial simulators (AVEVA PRO/II and Aspen HYSYS).

Improvements in phase equilibrium water-hydrocarbon to the traditional thermodynamic models have made better and accurate water content predictions in natural gas mixtures. Most of these thermodynamic are programmable in computational tools; however, this requires a certain degree of complexity and time-consuming [5]. Commercial software facilitates the use of thermodynamic models, but this requires an economical expense.

2.3.6. Other Methods

Many methods have been proposed to calculate water content in natural gas; there are differences in their development and applicability. Thus, some have been developed strictly for sweet natural gases and others for sour natural gases. The pressure and temperature ranges are also different for each method [5]. Table 5 summarizes some available methods for calculating water content.

Accurate water content determination requires careful study of existing literature and methods, if possible, compared with available experimental data [25]. Previously described methods estimate the water content as a function of pressure and temperature mainly. Unique sour gas methods add H_2S and CO_2 concentration for water content calculation. Although in the industry, there are many types of natural gas mixtures rich in heavy hydrocarbon components such as associated natural gas and its derivatives, among them

natural gas liquids and liquefied petroleum gas. No method for calculating the water content in NG considers the concentration of heavy hydrocarbons present in the mixture.

Table 5. Some available methods for estimating water content in NG

Method	Natural Gas Type	Operating Temperature (°F)	Operating Pressure (psia)
Sharma–Campbell Method [48]	Sweet	80 to 160	< 2000
Robinson et al., [49]	Sour	-50 to 350	300 to 10000
Guo and Ghalambor's Chart [50]	Sour	-60 to 280	14.7 to 10000
Ning Yingnan's formula [40]	Sweet	NA	14.7 to 10000
Sloan correlation [5]	NA	-40 to 120	200 to 1000
Wichert and Wichert chart [25]	Sour	-50 to 350	100 to 10000
Gordon's Chart [20]	Sweet	-60 to 400	14.7 to 10000
Torbjørn's Chart [20]	Sweet	-50 to 212	14 to 7352
Mohammadi and Richon [11]	Sweet	-38 to 40	< 1500

Note: NA stands no available information.

2.4. Fundamentals of Artificial Neural Network

Science has been inspired by nature in a wide variety of environments. For example, airplanes use aerodynamics principles taken from the way birds stay in the air. Nature also provides a powerful tool to learn from mistakes through natural selection and develop designs that generate excellent efficiency and adaptation to its environment. In Artificial Intelligence (AI), examples of how natural models have been adapted to create algorithms that allow them to learn from their mistakes and become more efficient in their tasks. An example of the above are artificial neural networks (ANN).

Artificial neural networks are software algorithms whose foundation is inspired by the behavior of neurons in the human brain [51]. Millions of chemical reactions occur in the brain that neurons communicate with each other. Neurons connect to each other through connections called synapses, which are electrical exchanges between the synaptic button (sending neuron) and the dendritic button (receiving neuron). These connections allow a neuron to send electric shocks to the cells around it. When a neuron receives a stimulus, its electrochemistry is altered, accumulating energy that, in a certain amount, the neuron discharges it to the neighboring neurons, forming a synaptic connection [52].

Artificial neural networks base their structure on neuronal biological networks, where a simple unit receives information from various sources to generate and transmit its data in a network of simple units. The simple ANN unit acquires the name of a neuron or node [53].

2.4.1. Node

The node or neuron is the smallest unit of processing in an artificial neural network. Similar to biological neurons, the node receives external stimuli through connections. External stimuli are input values (x_i) connected to the node by parameters called weights (w_i). Inside the node, a weighted sum of the input values multiplied by their weight is performed. A parameter called bias (b) is used to adjust the output and the weighted sum of the inputs to the neuron [54]. The weighted sum of the input values plus the bias are based on the linear regression model expressed in Eq. (6).

$$z = \sum_{i=0}^{n} (x_i \cdot wi) + b \tag{6}$$

Where z represents the weighted sum plus the bias.

Artificial neural network applications require that the output values are not always linear. Thus, to break with linearity, z is passed through an activation function ($\emptyset$) [51]. Therefore, the output value (y) of a neuron is represented by the mathematical expression in Eq. (7).

$$y = \emptyset(z) \tag{7}$$

Figure 9 shows a node's diagram with three input values denoted by X_1, X_2, X_3, and one output value y_1. The influence of each input value in y_1 is determined by the value of its respective weight W_1, W_2, W_3.
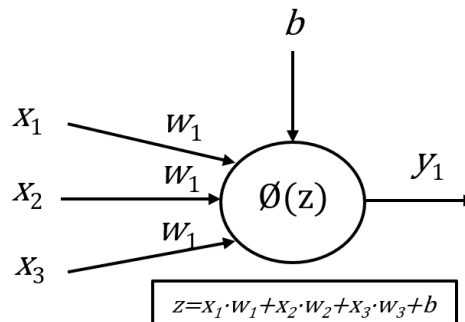
Figure 9. *Diagram of a node with three input values and one output value*

The activation function is the reason why artificial neural networks have the ability to model countless problems [54]. Activation functions are mathematical equations that determine the output of a node. There are linear and nonlinear activation functions; nonlinear functions are the most widely used because they allow artificial neural networks to expand their applicability and modeling power. Normally the set of output values that an activation function grants are in a range determined as (0,1) or (-1,1) [51]. Frequently used activation functions are listed below:

- Linear Function: also called identity function, it is the most used function (Eq. (8)) for modeling linear problems where the output is a real value [51].

$$\emptyset(x) = x \qquad (8)$$

- Sigmoid Function: the S-shaped term "sigmoid"; this function (Eq. (9)), also called logistic, maps the entire real range of x into (0,1). This simple function has two useful properties: 1) it can model a conditional probability distribution, and 2) its derivative has a simple form [55].

$$\emptyset(x) = \frac{1}{1 + e^{-x}} \qquad (9)$$

- Hyperbolic Tangent Function (*tanh*): this function (Eq. (10)) is analogous to the sigmoid function. Its interval widens to (-1,1), where high values asymptotically tend to 1, and very low values tend to -1. The tanh function has two important characteristics: 1) it is centered at the value of 0; 2) it enables the obtaining of negative and positive outputs [55].

$$\emptyset(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad (10)$$

- Rectified Lineal Unit Function (Relu): Relu function transforms the entered values, canceling the negative values and leaving the positive ones as they are entered. The Relu function (Eq. (11)) is frequently used in models where negative values are inadmissible [51].

$$\emptyset(x) = \begin{cases} 0 \ for \ x < 0 \\ x \ for \ x \geq 0 \end{cases} \qquad (11)$$

Figure 10 presents the representative graphs of the outputs of the activation functions explained.
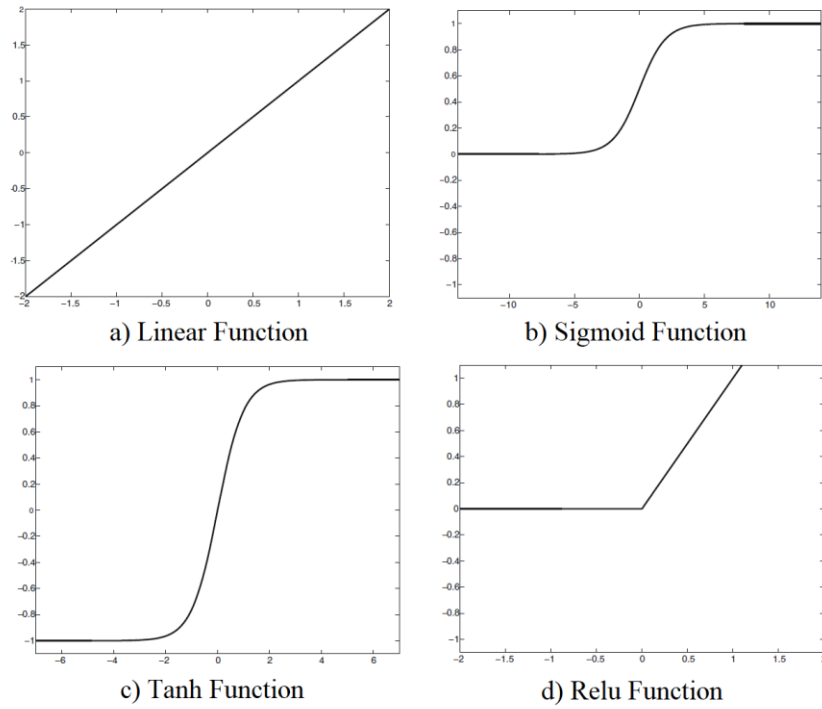
Figure 10. *Various activation functions*
Source: Adapted from Aggarwal [51].

The node and activation functions are the core of artificial neural networks. However, a single node cannot cover the solution of complex problems. Several nodes must be grouped, which together and how they have organized form the artificial neural networks topology [52].

2.4.2. Artificial Neural Networks Topology

In an artificial neural network, the nodes are connected to each other and grouped at different levels called layers [52]. A layer is a set of nodes; if the layer receives the input values are called "input layers", while the "output layers" export the data that the artificial neural network has processed. The layers between the output layer and the input layer are called "hidden layers" [53]. Hidden layers represent the processing structure of artificial neural networks. How layers are connected, the number of nodes per layer, and the number of layers present in a neural network give the name to what is known as artificial neural network topology [54].

Whit variation in the number of nodes per layer and number of layers in the network, complex ANN can be structured [51]. Artificial neural network models by the number of layers are classified into three groups: 1) Perceptron with the simplest structure which only has one input layer and one output layer; 2) Monolayer Neural Network which has one input layer, one hidden layer and one output layer; 3) Multilayer Neural Network which has one input layer, one output layer and as many layers as required by the network system [53]. The connections between neurons can be the feed-forward networks with only forward connections or the feedback networks that can have forward and backward connections. Figure 11 shows a multilayer feedforward neural network diagram, where the direction of the arrows represents the forward connections between layers.

Input Layer   Hidden Layers   Output Layer
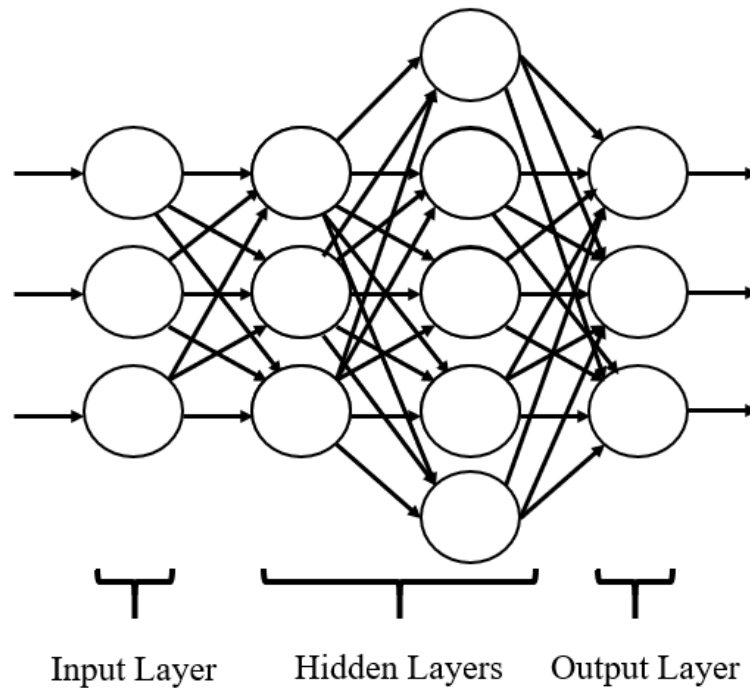
Figure 11. *Feed-forward multilayer neural network*

2.4.3. Artificial Neural Networks Training

Like humans, an artificial neural network can learn through experience. In other words, an artificial neural network can solve problems after training with other similar problems [10]. ANN training consists of giving the ANN a set of data (training data) with known inputs and outputs. Thus, from the training data, an artificial neural network will be able to predict outputs for new input data sets. Once an ANN can solve problems that have never been seen before, it is said that the ANN meets generalization [55]. In the training

stage, ANN parameters (weights and biases) are adjusted so that the output values of the ANN are the closest to the target value of the training data. This learning through experience in machine learning is called supervised learning [51]. It is necessary to use training algorithms to find the appropriate ANN parameters. Figure 12 shows the general steps that the training algorithms take for supervised learning.
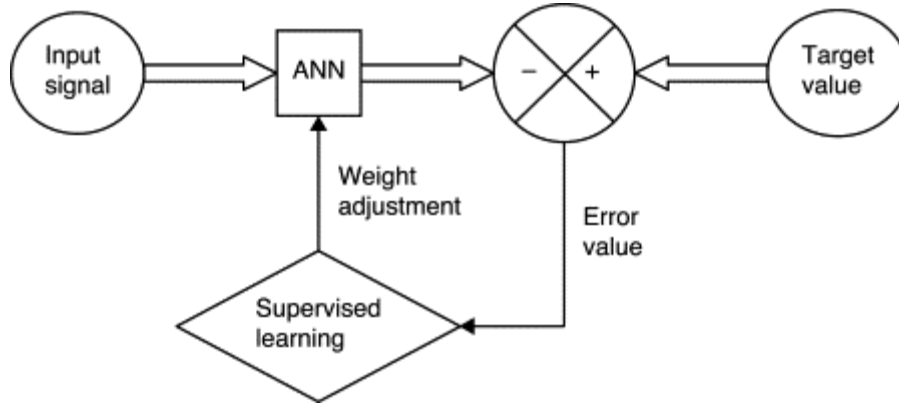
Figure 12. Supervised learning algorithm

During the training stage, the ANN parameters are adjusted as often as necessary to obtain a minimal difference between the neural network's output and the training data set's target values. This difference between the ANN values and the target values is called objective function and is usually represented by statistical tools that express the deviation [54]. Each time all the weights are adjusted, and the neural network output is obtained, called an epoch or iteration. Each epoch or iteration reduces the objective function [53].

2.4.3.1. Objective Function

The objective function plays a fundamental role in the learning of ANN. This objective function defines the course of learning by evaluating the error between the ANN output values concerning the training data's target values. The choice of the objective function depends on the particular application [51].

When establishing an objective function (Eq. (12)), two different terms must be chosen: error and regularization.

$$Objective\ function = error\ term + regularization\ term \qquad (12)$$

- **Error term**: error is the most important term in the objective function; It measures how the ANN fits into the training data set. All of these errors can be measured with a statistical tool [56]. The most important errors used in the field of ANN are described below.

-Mean Square Error (MSE): calculates the mean of square differences between the ANN outputs and the training data set targets (Eq. (13)).

$$MSE = \frac{\sum (ANNoutput - target\ value\ )^2}{training\ instances\ number} \tag{13}$$

MSE is the most commonly chosen error term to form part of the cost function because squaring differences remove their negative signs and influence larger differences [54].

-Mean Absolute Error (MAE): calculates the mean of the absolute differences between the ANN outputs and training data set targets (Eq. (14)) [56].

$$MAE = \frac{\sum |ANNoutput - target\ value|}{training\ instances\ number} \tag{14}$$

-Normalized Squared Error (NSE): divides the squared difference between the ANN outputs and the targets by a normalization coefficient (Eq. (15)). If the NSE has a value of less or equal to one, then the neural network predicts within the allowed range, while a zero value means that the ANN makes a perfect prediction [56].

$$NSE = \frac{\sum (ANNoutput - target\ value\ )^2}{normalization\ coefficient} \tag{15}$$

NSE can be considered the default error term when solving classification and prediction problems.

- **Regularization Term**

Regularization consists of adding a penalty value to the cost function. This penalty produces simpler models that generalize better [57]. The regularization term can take any value, or it is also considered according to the ANN parameters (weights and biases).

2.4.3.2. Optimization Supervised Algorithm

As mentioned before, a learning artificial neural network consists of obtaining weights and biases that minimize the objective function. The objective function is, in general, a function of the ANN parameters. These parameters, "weights and the biases" are initially random, and by changing them only once, it would be impossible to reach an immediate solution. Instead, a search through the infinite parameter space consists of a succession of iterations or epochs. In each epoch, the objective function will decrease by adjusting ANN parameters [55]. Optimization supervised algorithms generate a sequence of parameters to reduce the objective function at each iteration. The question is, when do you have to stop training? The optimization algorithm stops the ANN training when the following specified condition is satisfied [51]:

   o   The objective function improvement in one epoch is less than a set value.
   o   The objective function has been minimized to a goal value.
   o   A maximum number of iterations or epochs is reached.

The supervised optimization algorithm or training algorithm determines how ANN parameters change. The most commonly used optimization algorithms are described below [58].

- Gradient descent (GD): GD is the simplest optimization algorithm. Here, the parameters (weights and biases) are updated in each epoch going in the gradient objective function's negative direction [59]. From modifications to the GD algorithm, several other algorithms have been developed, such as AdaGrad and ADAM. Figure 13 shows a diagram of how the descending gradient works to find new parameters; in this example, the objective function that has been chosen is the MSE.
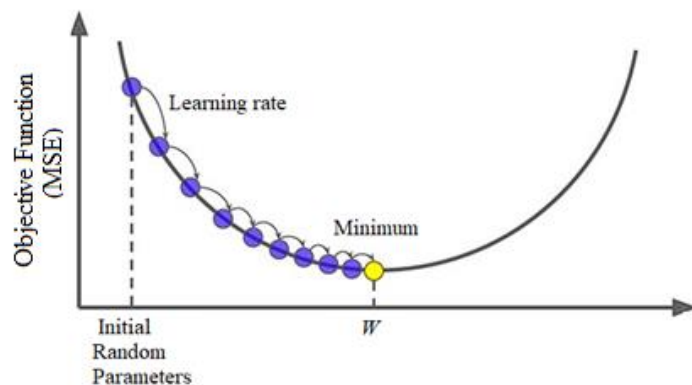
Figure 13. *Representation of gradient descent*

- Levenberg-Marquardt algorithm (LMA): LMA proposes a damped Gauss-Newton scheme. Like Gauss-Newton, it takes the linear approximation of the model and a certain direction of descent. LMA is highly used in the optimization environment, especially in the least-squares minimization [60].

- Adaptative linear momentum (ADAM): the ADAM is similar to gradient descent but implements a more sophisticated method of calculating the training direction that generally produces faster convergence [59].

Optimization Supervised Algorithm programming requires a deep knowledge of calculation tools. Generally, these algorithms are already implemented in computational tools for ANN design.

2.4.3.3. Overfitting and Underfitting

The goal of a good machine learning model, in this case, artificial neural networks, is to generalize the training data. It allows future predictions to be made on data that the ANN has never seen. Overfitting occurs when the model has been trained too much for the training data, it does not allow good predictions for new input data [61]. Underfitting occurs when the model cannot identify patterns, and the ANN will always get terrible results [51]. In Figure 14, a simplification to an example of regression allows visualizing the problem of underfitting and overfitting. Finding a balance for a correct generalization becomes necessary. The sweet spot is the middle ground that must be found in ANNs learning to ensure avoiding overfitting and underfitting, and an ideal fit can sometimes be a difficult task [53].
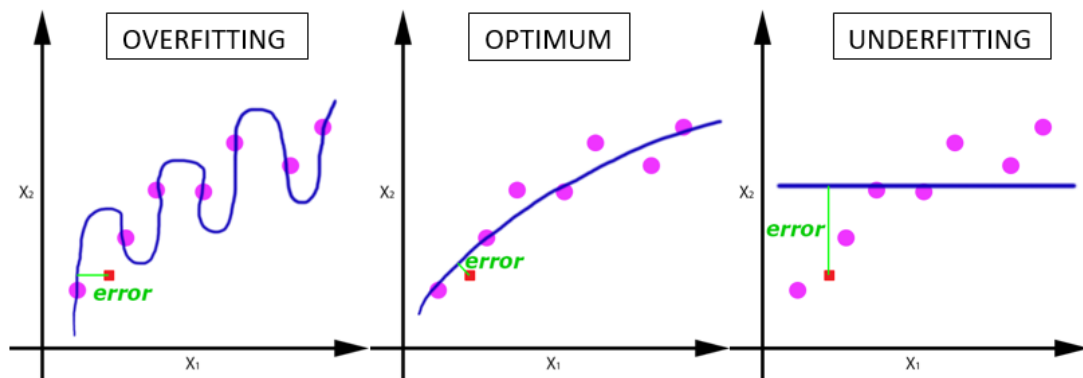
Figure 14. Comparison Overfitting-Underfitting

There are several strategies to avoid overfitting and underfitting. A strategy has already been discussed, which is Regularization. When regularization is used, the complexity of the model is minimized while minimizing the objective function. The results in simpler models that tend to generalize better than excessively complex [61]. Second, data must be sufficient to represent and generalize the desired model; an analysis of the training data is important when implementing any artificial neural network model [54].

The third and most important solution to overfitting and underfitting is validation. Validation is the data set that does not participate in the ANN training group. The validation group and the ANN output must be compared with statistical tools, taking the validation group as real data [10]. From a data universe, 80% is usually taken for the training set and 20% for the validation set; these percentages in machine learning make mention of the Pareto principle, but they are not mandatory [62].

CHAPTER III

3. METHODOLOGY

The artificial neural network's development for calculating the water content in natural gas, taking heavy hydrocarbons into account, required a series of steps summarized in Figure 15. First, an intensive bibliographic search of the available methods for calculating water content in NG and experimental data was carried out. The most influential variables in the water content of NG and the collected experimental data were analyzed from the available methods. Second, for modeling, the data bank was processed, and two data sets were obtained. Data set 1 was used to design the ANN. An artificial intelligence software was used to find the artificial neural network structure that best fits the training data. Finally, data set 2 was used as a validation group of the artificial neural network. The results obtained by the ANN were compared with available methods for calculating the water content in NG. It was to verify the neural network's applicability developed for different natural gas types and in an adequate operating range according to the gas industry's needs.
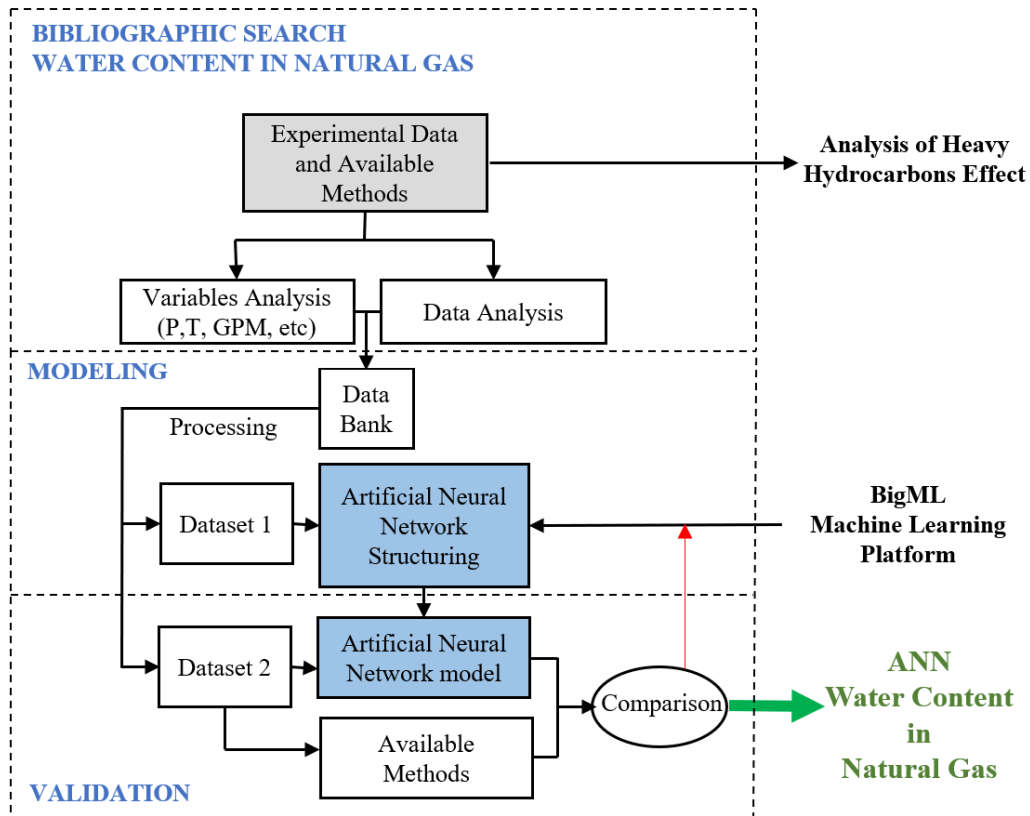
Figure 15. *Methodology diagram*

3.1. Modeling Bases

For developing the artificial neural network for calculating the water content in natural gas, taking heavy hydrocarbons into account, the following bases are taken:

1. The collected experimental data must comply with the following information to be considered in this work: pressure, temperature, the total composition of the natural gas mixture, and experimental water content.

2. The experimental data of all types of natural gas available in the open literature are considered in this work; this includes sour gases, sweet gases, rich gases, and lean gases.

3. Since no method for calculating water content reports errors greater than 25%, an absolute error (AE) of $< 25\%$ was established to limit experimental data. Eq. (16) shows the absolute error formula to compare the experimental data for water content with available methods, which were called theoretical values.

$$AE = \frac{|[theoretical\ value - experimetal\ value]|}{experimental\ value} \tag{16}$$

4. For experimental data, possible measurement errors are not taken into account in this work since the data was collected from different sources.

5. The influence of heavy hydrocarbons on the water content in natural gas was made based on available methods.

3.2. Bibliographic Search Water Content in Natural Gas

The bibliography search was carried out in various open sources of reliable information such as books, online scientific information services, scientific and engineering journals, and natural gas handbooks. The first step was an in-depth study of available methods for calculating water content in natural gas mixtures. This study analyzed the typical influencing variables and heavy hydrocarbons on the water content in natural gas. Experimental water content data were also analyzed, avoiding extreme deviations from the available methods.

3.2.1. Variables Analysis

Variables analysis is essential for the development of any model or calculation method. For modeling, knowing the most important variables related to the water content in a natural gas mixture was necessary to develop the artificial neural network properly. On the one hand, many variables used as input values hinder the learning process, decreasing the accuracy of the model's predictions. On the other hand, creating a model with few input variables facilitates the neural network training process, but neglecting influencing variables can cause the model to not comply with generalization [61]. Naturally, there are variables more influential than others, and the number of total variables of the ANN was determined based on available models. A total of 4 variables are considered in available methods for calculating water content in natural gas. To these variables in this work is added the contribution that the concentration of hydrocarbons heavier than methane can influence in water content. The typical variables are: temperature, pressure, acid gas content, and molecular weight.

3.2.2. Data Analysis

From the bibliographic search, experimental data of the water content in natural gas mixtures were also obtained. For the appropriate analysis of the experimental data and future application of the ANN, pressure and temperature operating range was established to which natural gas is usually subjected during its processing. For this purpose, a minimum and maximum limit of pressure and temperature were established. Pressure and temperature limits were set with natural gas processing experts and natural gas treatment handbooks.

The experimental data were also compared with available methods for calculating water content in NG to eliminate data extremely far from the pre-established models. The experimental data come from different research works, making it impossible to obtain a general error for pressure, temperature, and water content measurements. This work assumes that all the experimental data obtained from the different research works are real data. Thus, with all the experimental data that complied with the operating range and modeling bases 1 and 3, the final data bank was formed.

3.3. ANN Modeling

The design of an artificial neural network involves a sequence of steps that go from data processing to the final step, which is the validation of the ANN [10]. The steps followed in this work for modeling the artificial neural network were: data bank processing and ANN structuring.

3.3.1. Data Processing

Data processing was carried out from the data bank compiled in the bibliographic search. Since artificial neural networks' activation functions are a range (-1,1), having an adequate scale for the database facilitates the neural network's convergence and training [63]. In this work, all variables were linearly normalized in a range (-1,1). The mathematical expression used for the normalization method is defined in Eq. (17).

$$x_{norm} = \frac{x - x_{max}}{x_{max} - x_{min}}(r_{max} - r_{min}) + r_{min} \qquad (17)$$

where x is the original value; x_{norm} is the normalized value; x_{max} and x_{min} are the maximum and minimum values of the concerned variable, respectively; r_{min} and r_{max} are the desired values of the normalized variable range, in this case (-1,1).

Once all the values of all the variables were normalized, including the water content, the data bank normalized was divided into two groups; data set 1 and data set 2. Data set 1 formed part of the artificial neural network structuring, and Data set 2 was used as a validation group for the ANN. The normalized data bank was divided 74% for data set 1 and 26% for the validation group.

3.3.2. Artificial Neural Network Structuring

Structuring an artificial neural network involves finding the topology with the training algorithm best adapted to the training data set and complies with generalization to other data. The process of finding a good ANN topology does not have a defined sequence of steps. Several authors assure that monolayer structures should be tested first, progressively increasing the number of nodes one by one in the hidden layer. If monolayer

neural networks do not have good results, multilayer neural networks allow better results to be obtained [51]. However, finding a suitable ANN multilayer topology becomes tedious given many combinations in the number of hidden layers and the number of neurons per layer. Many machine learning software, programming languages, and AI platforms allow access to instruments that make it easy to find good ANN topologies.

To find the artificial neural network topology that best fits this work, the BigML machine learning software has been used. BigML is a consumable, programmable, and scalable Machine Learning software. In this work, the tools for artificial neural networks of BigML are employed to obtain the best performance of the ANN. About 5000 multilayer artificial neural network topologies were designed and evaluated in BigML. Doing it manually took too much time, so the BigMLs Automatic Network Search tool was also used. This tool allows evaluation batches of 200 neural networks in a maximum period of 5 hours. During each batch, several training algorithms are tested. This search tool also tests the different activation functions described above. After the evaluation and training of any batch, the software yields the ANN structure with the best adjustment performance for the training data set [64]. BigMl also allows hosting the ANN in their platform for later use. The objective function for all searches was established as an MSE of 0.01 for the ANN's normalized outputs compared to the training data. Finally, the validation data set was used to evaluate the ANN with the best performance.

3.4. Artificial Neural Network Validation.

Dataset 2 was used as a validation group for the neural networks. To validate the ANNs designed, the following statistical tools were used to measure the error or variation of the ANN output values with the values of the validation group:

- **Mean Square Error (MSE$_v$):** analogous to Eq. (13), the MSE and being used as a cost function are also used to validate the ANN. The following nomenclature MSE$_v$ was used for the validation section since the mathematical formula is the same, but the compared values change as shown in Eq. (18).

$$MSE_v = \frac{\sum(WC_{exp} - WC_{ANN})^2}{validation\ instances} \tag{18}$$

where WC_{exp} is the experimental water content of the validation group; WC_{ANN} is the water content estimated for the ANN. This calculation is made for all validation group instances.

Note: this work refers to each natural gas sample with its respective variables pressure, temperature, etc.

- **Coefficient of Determination (R^2):** it is a statistical tool that measures how well a prediction fits linearly to the general linear model [65]. Eq. (19) describes the coefficient of determination.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(WC_{i,exp} - WC_{i,ANN})^2}{\sum_{i=1}^{N}(WC_{i,exp} - \overline{Wc})^2} \tag{19}$$

where $WC_{i,exp}$ is the experimental water content of the validation group; $WC_{i,ANN}$ is the water content predicted for the ANN; $\overline{Wc}$ is the mean water content of observed validation data being evaluated. The values of R^2 range from 0 to 1. Values close to 1 represent that the values produced by the ANN have a good fit to the experimental ones, while values of R^2 that tend to 0 mean that the values predicted by the ANN do not have a good fit to the experimental data.

MSE and R^2 values are typically used to validate the normalized outputs of the artificial neural network. However, the normalized ANN outputs do not make any physical sense. To validate the ANN outputs, the average absolute deviation (AAD) was used as a parameter. The ANN's normalized output data was transformed into water content values in lb/MMSCF and compared with the data from the validation group.

The mathematical expression of AAD is shown in Eq. (20).

$$\text{AAD} = \frac{1}{N}\sum_{i=1}^{N}\left|\left(\frac{WC_{exp} - WC_{ANN}}{WC_{exp}}\right)\right| \tag{20}$$

Usually, the AAD is compared in percentage; therefore, Eq. (20) should be multiplied by 100%. High AAD% values mean that ANN outputs data is vastly different from the experimental data. In contrast, small AAD% shows that the ANN gives accurate results.

Once the best performing ANN was obtained, the ANN results were compared with available water content methods. For comparison with charts and correlations, Bukacek Correlation and McKetta-Wehe Chart were taken. On the thermodynamic methods side, ANN was compared with SRKKD EoS and PRM EoS. Finally, to observe the ANN's applicability, the validation group was divided by gas types, and the AAD% was evaluated concerning the other available calculation methods. Because Mcketta-Wehe chart and Bukacek correlation are applicable only for sweet gases, for the sour gas samples, Maddox correction is applied to these methods for calculating water content in sour mixtures.

CHAPTER IV

4. RESULTS AND DISCUSSION

4.1. Analysis of Influencing Variables on Water Content

The four influencing variables that the available methods for calculating water content in natural gas mixtures usually report are temperature, pressure, acid gas content, and molecular weight.

- **Temperature and pressure:** these two variables are the most important for estimating water content in natural gas. For the hydrocarbons-water system, pressure and temperature define the vapor-liquid phase equilibrium. The temperature is directly proportional to the water content; the higher the temperature, the higher the natural gas's water content. The pressure, on the other hand, is inversely proportional to water content. Methods such as the Bukacek correlation have been developed only taking into account these two variables.

- **Acid gas content:** the presence of acid gases such as H_2S and CO_2 increases the natural gas's water content. H_2S has an asymmetric electron charge arrangement; as such, H_2S has a higher polarity than hydrocarbons. Water, a polar substance, has higher solubility in polar materials [66]. Instead, CO_2 has strong dipole-induced dipole attractions to water. For most of the methods developed for calculating water content in natural gas, the presence of acid gases determines whether a method is applicable or not. This is the case of the Mcketta-Wehe chart that is not applicable for natural gas mixtures with considerable concentrations of H_2S and CO_2. For the acid gas content, the contribution of H_2S and CO_2 was taken according to Eq. (21).

$$y' = yH_2S + 0.75yCO_2 \tag{21}$$

  where y is the mol fraction, H_2S and CO_2 refer to the gases. In order to group in a single variable the influence of acid gases and reduce the input values to the ANN, y' parameter received the name of Equivalent mol H_2S [67].

- **Molecular weight:** molecular weight as influential on water content has several

appreciations: methods applied for sweet gases report that the higher the molecular weight, the water content tends to be lower [25]. On the contrary, sour gases methods report that the higher the molecular weight, the water content increases [42]. Remarkably, the molecular weight is considered influential in the water content according to the type of gas used to develop the model.

4.1.1. Heavy Hydrocarbons Effect on Water Content

There are no studies or methods that show the influence of hydrocarbons heavier than methane on natural gas's water content in the open literature. McKetta-Wehe chart is the first model developed that reports a correction to natural gas's water content considering specific gravity and molecular weight. McKetta-Wehe chart was established only to be used in sweet gases, where acid gases' compositions are negligible [6]. Therefore, if acid gases' composition is negligible, the components that increase the molecular weight can only be heavier hydrocarbons than methane.

4.1.1.2. Correction for Gas Gravity

The correction factor for gas gravity (C_G), Figure 16, presented in the McKetta-Wehe chart is established in function of temperature and specific gravity (instead of molecular weight). By analyzing the correction for gas gravity, it can be determined that with good content of heavy hydrocarbons and the appropriate temperature condition, there is up to a 30% variation concerning the calculation of water content in rich natural gases compared to lean gases. Therefore, not considering the content of heavy hydrocarbons in the available water content methods could be why some available methods have low accuracy.

In the gas industry, natural gas compositions are reported with a composition of heavy hydrocarbons from ethane to heptane or sometimes represented as C_7^+ (used to represent heptane composition in conjunction with small amounts of heavier hydrocarbons). Considering the hydrocarbons ethane, propane, butanes, pentanes, hexanes, heptane+, at least six variables should be added to any water content calculation model taking into account the presence of heavy hydrocarbons. In addition to pressure and temperature, mandatory conditions for calculating water content require at least eight variables to add the water content's heavy hydrocarbons. Reducing the number of variables in the model

is an important step to optimize the system's functioning and, in the artificial neural network, to optimize the learning process [55]. The variable reduction should be made, trying not to omit relevant information for the model. Thus, in this work, it is proposed to group all the heavy hydrocarbon compositions in a single variable: the GPM.
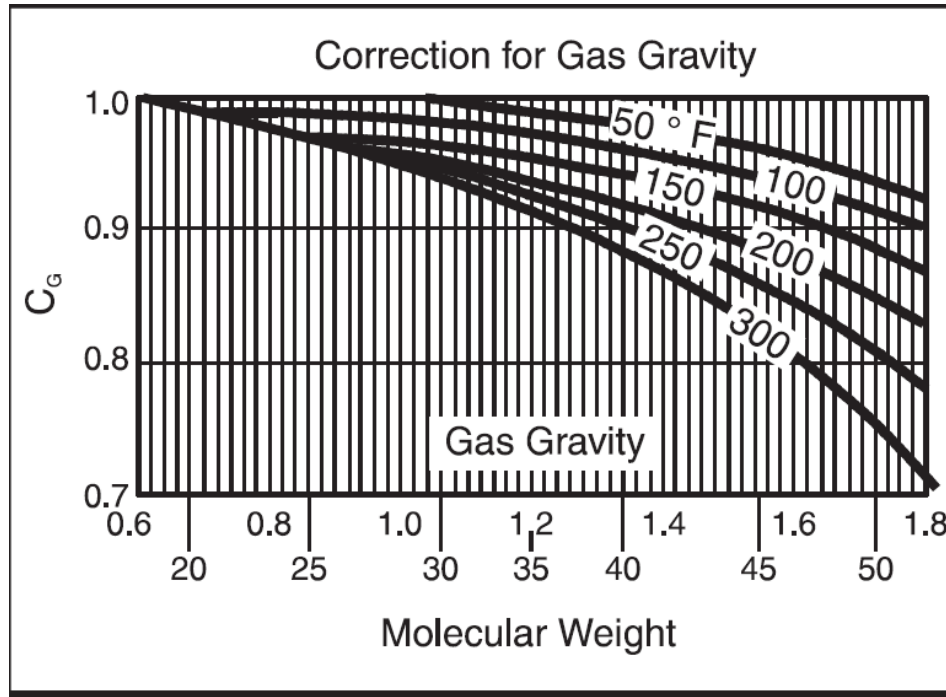
Figure 16. *Gravity correction factor*

Source: Adapted from Engineering Data Book [25].

The present work seeks that the developed model was applicable for both sour and sweet gas mixtures; therefore, the molecular weight's duality of influence will conflict with the ANN training process. Thus, the molecular weight is discarded as a variable, and the remaining 3 variables are added to the contribution variable of heavy hydrocarbons. Finally, for the development of the artificial neural network, the following variables were considered as input values: temperature (°F), pressure (Psia), equivalent mol H_2S (y′) and GPM.

4.2. Artificial Neural Network Model

4.2.1. Data Analysis and Operating Conditions

As mentioned before, a pressure and temperature range must be established to develop and apply the ANN. According to the recommendations of gas processing experts and manuals, the minimum pressure condition was set at 60 psia and the maximum condition at 3200 psia. The lower temperature condition was established at 32 °F because hydrate formation is favored at low temperatures, and hydrate formation conditions are not desired in the gas industry [19]. Later, for the maximum temperature, a temperature around 300 °F was established. The defined pressure and temperature operating range and modelling bases were used to analyze the data found in the bibliographic search.
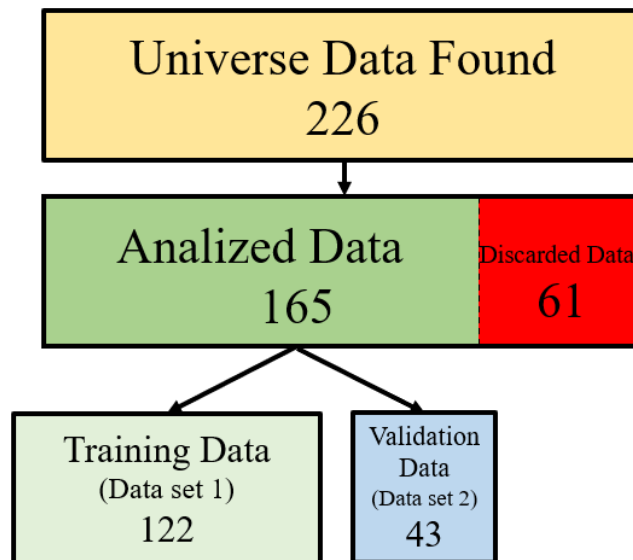
Figure 17. Database distribution

From the bibliographic search, a universe data of 226 instances of natural gas was obtained. The modeling bases and the operating range established for this work allowed to discard instances; finally, 165 instances of natural gas mixtures formed the analyzed data bank. For the development of the ANN, data set 1 with 74 % of the analyzed data bank (122 instances) formed the training group, while data set 2, the 26% (43 instances) the validation group (Figure 17). Satisfactory studies Tatar et al., 2016 [68] recommend an 80% distribution for the training stage and 20% for validation. The data bank's distribution in this work is due to many instances of the data bank, which is why 6% more data is offered for the validation group, giving a wider margin for validation. The data bank division was carried out randomly, seeking to distribute both the training and

validation data throughout the established operational range of pressure and temperature. Figure 18 shows the final distribution of the training and validation group in this work.

There are methods for predicting water content in natural gas that consider pressure conditions up to 15,000 psia [25] in the open literature. However, under typical gas processing conditions, the pre-dehydration process stages rarely exceed 3000 psia [2]. The condition range established in this work for pressure is sufficient for typical natural gas processing conditions. Furthermore, as shown in Figure 18, most of the open literature data are below 1500 psia; this shows that the defined pressure range for the ANN is optimal for its application in dehydration processes in the gas industry. Concerning the temperature, as mentioned before, low temperatures favor the formation of gas hydrates. The water content in natural gas in hydrate equilibrium tends to be lower than in liquid-vapor equilibrium [3]. This adds to the lack of experimental data that describes the water content in this phase equilibrium due to the difficulty of measurement under these conditions [11].

Figure 18. Database distribution under typical natural gas processing conditions

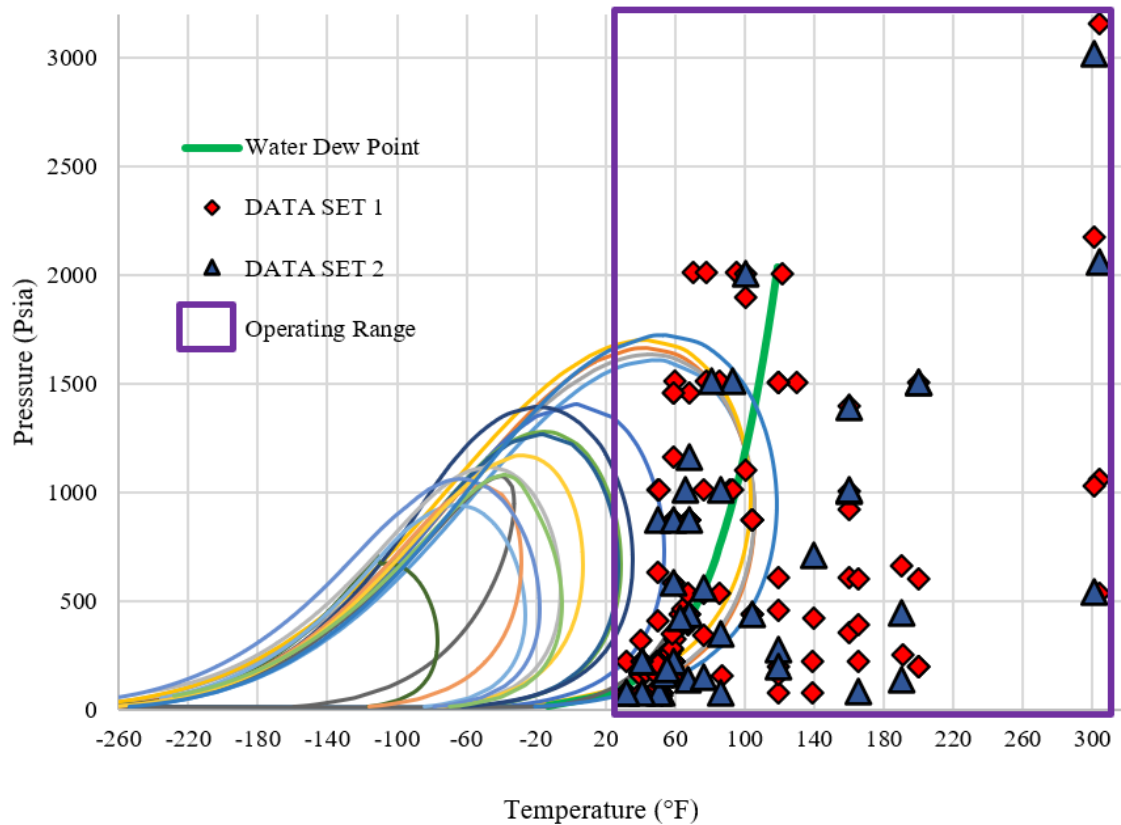Therefore, using experimental data measured for water content in liquid-vapor equilibrium should not be used for modeling natural gas systems in conditions of gas hydrates equilibrium. Thus, in this work, the lower temperature limit was set at 32 °F and

for the maximum temperature limit at 300 °F. Very high temperatures are also avoided since in the natural gas industry there are cooling processes (gas liquefaction), increasing the gas temperature a lot and then cooling requires a considerable energy cost. Finally, the operational range established for temperature in this work is considered optimal for applying the developed artificial neural network.

4.2.2. Artificial Neural Network Structure

Artificial neural network training was performed on the BigML software, various artificial neural network structures were tested with many training algorithms, and the activation functions such as Linear, Tanh, Sigmoid and Relu. The validation group was used in comparison with the ANN outputs. Then, statistical indicators MSE_v and R^2 were considered for the water content's normalized data, and the %AAD for the ANN outputs converted to lb/MMSCF. Values ≤ 20 of AAD% were taken as a satisfactory result. Models that do not comply with the AAD% are not reported in this work. Several authors state that a monolayer ANN is sufficient to obtain good results for nonlinear regressions. However, multilayer ANN is more effective, especially for predicting structures with many input values [51]. Since monolayer ANN models did not comply with a satisfactory AAD%, the BigML platform was used to find multilayer ANN structures with the best performance. About 5000 multilayer ANN structures were tested in batches of 200 structures. The best ANN structure of each batch and the comparison of their results with the validation group are reported in Table 6.

Table 6. *Best ANN structures developed*

Notation	Hidden layers	Training Algoritm	Normalized output		Not normalized
			MSE_v	R^2	AAD%
ANN(A)	**Relu**(533),**Relu**(301)	AdaGrad	0.036	0.95	18.75
ANN(B)	**Sigmoid**(54),**Sigmoid**(54),**Simoid**(54)	ADAM	0.020	0.89	19.26
ANN(C)	**Linear**(998),**Linear**(524),**Linear**(302)	ADAM	0.030	0.96	17.23
ANN(D)	**Sigmoid**(16),**Sigmoid**(36),**Sigmoid**(67)	ADAM	0.010	0.99	13.90
ANN(E)	**Relu**(524),**Relu**(524)	ADAM	0.015	0.93	11.80
ANN(F)	**Relu**(342),**Relu**(890),**Relu**(40)	AdaGrad	0.220	0.92	15.80
ANN(G)	**Tanh**(120),**Tanh**(38)	AdaGrad	0.017	0.91	9.17
ANN(H)	**Sigmoid**(46),**Sigmoid**(23),**Sigmoid**(11)	ADAM	0.010	0.97	7.00
ANN(I)	**Tanh**(120),**Tanh**(38)	ADAM	0.030	0.95	11.50
ANN(J)	**Relu**(528),**Relu**(258),**Relu**(122)	AdaGrad	0.020	0.98	12.96

Note: The hidden layers are ordered from left to right, each activation function is applied to the entire layer, and the number in parentheses represents the total node per layer.

For all ANN structures, the MSE_v and R^2 values are acceptable. The AAD%, however, presents a lower value in the **ANN(H)** structure; this means this structure has more accurate output predictions than the others when comparing with the experimental water content of the validation group in lb/MMSCF. Although the R^2 is higher for the ANN(D) and ANN(J) structures than the **ANN(H)** structure, Draper and Smith [69] indicate that R^2 should be evaluated with other statistical indicators that allow validating a model since the R^2 does not strictly indicate whether a regression model is or not suitable. Thus, the MSE_v and the AAD% take on greater relevance to validate the ANN developed in this work. MSE_v values denote the differences between the ANN output values and the experimental validation group values [54]. **ANN(H)** and ANN(D) structure present the smallest MSE_v values, ensuring a minimum difference between the normalized values of the ANN and the validation group. Several models for the calculation of water content in natural gas base their validation on the AAD%. A small AAD% reflects more accurate predictions for the developed ANN. Therefore, the AAD% defines **ANN(H)** as the structure with the best performance.

Figure 19 shows a developed multilayer artificial neural network ANN(H) scheme that presents the best performance. It is a feedforward multilayer ANN; its structure seems a bit complex compared to other works related to water content calculation in natural gas mixtures. Mohammad et al., 2016 [39] propose a monolayer ANN with 8 nodes. On the other side, Mohammadi and Richon 2007 [11] propose a monolayer ANN with 6 nodes. However, monolayer ANN structures in this work did not present accurate predictions. Increasing the number of nodes was one solution so that monolayer ANNs can generate more accurate predictions. However, an extremely large number of neurons in a single layer limits the learning process [53]. Multilayer ANN structure is justified in this work because 4 input variables are presented. Previous works present 2 and 3 input variables; also, the instances used for training are smaller than those used in this work. Besides, sometimes, as in this work, increasing the number of hidden layers is more effective, although it should be clarified that for non-linear regressions, more than 3 hidden layers do not provide better results [51].
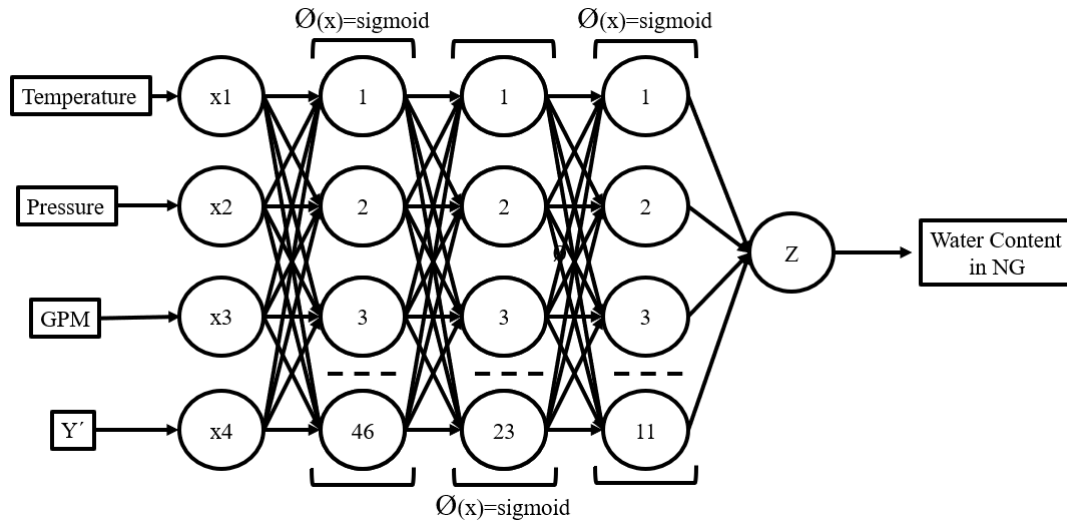
Figure 19. *Artificial neural network structure developed for calculating water content in natural gas*

4.3. Artificial Neural Network Applicability

The experimental data which ANN was developed comes from 7 different sources. Many studies for the experimental calculation of water content were carried out with synthetic gas (SM) mixtures that emulated natural gas mixtures depending on the study. Natural gas mixtures (NG) from wells were also obtained. SM and NG in all studies were subjected to artificial water content saturation in their respective works. The validation group is conformed of NG and SM mixtures; its properties are detailed in the Appendix section. The samples that were part of the artificial neural network development can be classified either by their richness or content of acid gases, as shown in Figure 20.
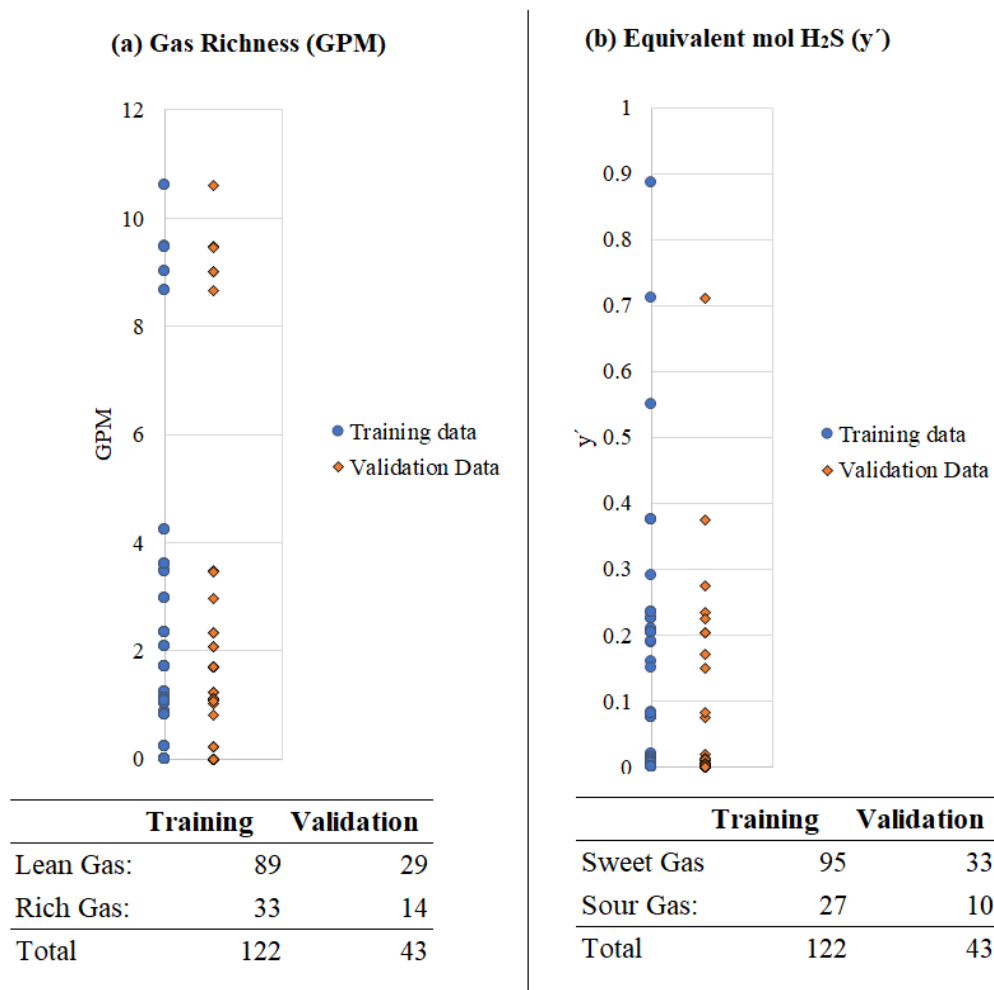
Figure 20. *Dataset classification:*
(a) by heavy hydrocarbons content, (b) by acid gases content

The applicability of the developed artificial neural network must be defined by evaluating its performance compared to other predictive models of water content in natural gas mixtures. For the evaluation, Bukacek correlation, McKetta-Wehe chart, SRKKD EoS and PRM EoS were selected. Bukacek correlation and McKetta-Wehe chart must not be used for sour gases, so the Maddox correction was used to adjust the water content values calculated by these methods in sour gases. For their part, SRKKD EoS and PRM EoS thermodynamic methods do not report limitations to work for both sour and sweet gases [45]. First, a comparison of the entire validation group was made with the ANN developed and the available water content methods.

Figure 21 shows the AAD% comparison of this work's results and other available methods for calculating water content. The AAD% obtained for the ANN is 7.00%, the lower value than the other available methods. This means that for this data set (validation group) the model developed in this work has more accurate predictions than other predictive methods. Bukacek, being a correlation that depends only on temperature and pressure, shows the largest deviation with an AAD% of 14.54%. For their part, thermodynamic methods present adequate AAD% values, especially since the SRKKD EoS model has similar values to the model developed in this work. The difference between the two methods is the facility of their application. For its part, the SRKKD EoS thermodynamic method requires a computational tool for its implementation and is time-consuming. The use of thermodynamic models is greatly facilitated with commercial software acquisition, but this requires an economic cost. On the other hand, the ANN developed in this work is hosted on the BigML software with free access, and only the four normalized input variables are required to estimate the water content in natural gas. The ANN developed in this work also presents an opportunity for improvement as more experimental data is added to the database. A detailed manual for using the ANN of this work can be found in the Appendix section.
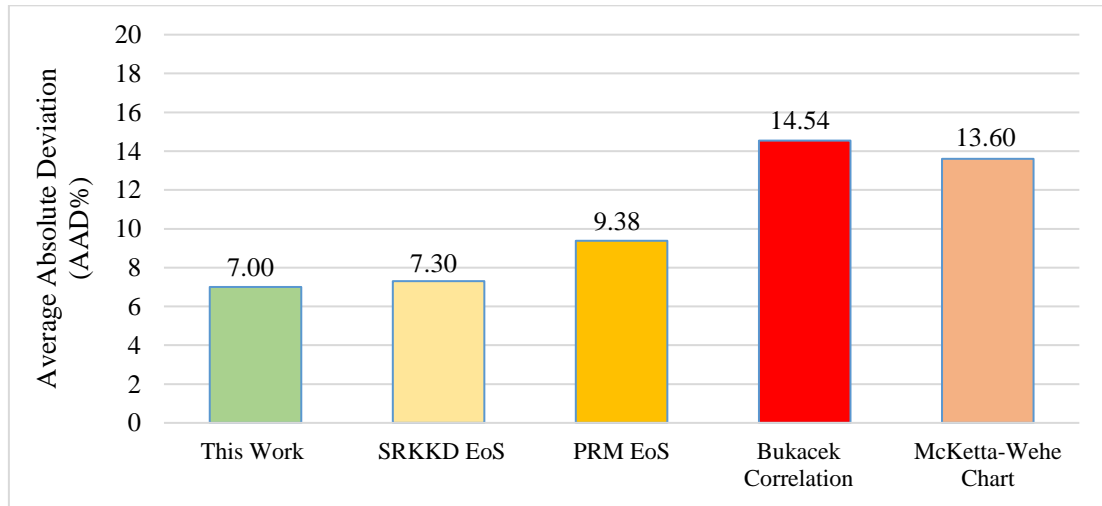
Figure 21. *Methods for calculating water content in NG performance for the validation group*

The applicability of the model can be extended by evaluating more experimental data. However, more experimental data are not available in the open literature. To evaluate the ANN's performance for various natural gas types, the validation group was divided, and the% AAD was calculated for gas types. Figure 22 shows the AAD% values by gas type obtained for the ANN developed and the available methods for predicting water content

in NG. The ANN developed in this work presents AAD% lower than the other methods for rich gases, lean gases, and sweet gases. However, sour gases present the second-highest AAD% below the McKetta-Wehe Chart. The AAD% of 8.7% value for sour gases in this work does not represent a bad prediction considering that methods such as the Wichert chart and Maddox correction (models developed strictly for sour gases) report deviations of 10% concerning experimental data of sour gases [25]. Then, the possibility of using more experimental data is open for better evaluation and comparison with other available methods.
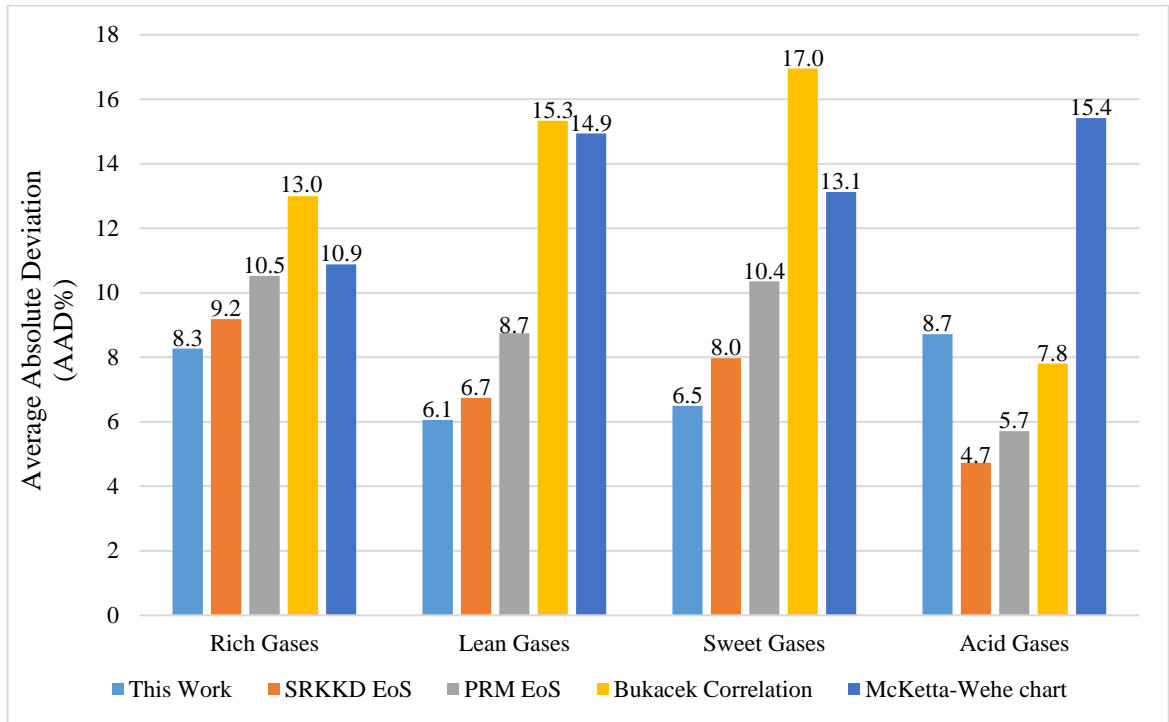
Figure 22. *AAD% by Gas Type to compare methods for calculating water content*

McKetta-Wehe chart is the most commonly used method in calculating the water content of natural gas because it gives a quick estimate, and if used carefully, the error for sweet gases is reported as less than 5% [40]. The problem with McKetta-Wehe chart is that its facility of use is overshadowed by its difficulty in reading the water content, leaving the final value of the water content at the discretion of the person who used the chart [39]. "A picture is worth a thousand words", Figure 23 shows the correlation of the predictions of the McKetta-Wehe chart and the ANN developed in comparison with the experimental water content in natural gas, which ordinate gives a straight line of 45°. It can be seen that the values generated by both models spread around the 45° line, which does not imply an overestimation or an underestimation [69]. However, visually the McKetta-Wehe chart

values protrude more in the deviation lines 10% than the ANN values, which are closer to the 45° line.
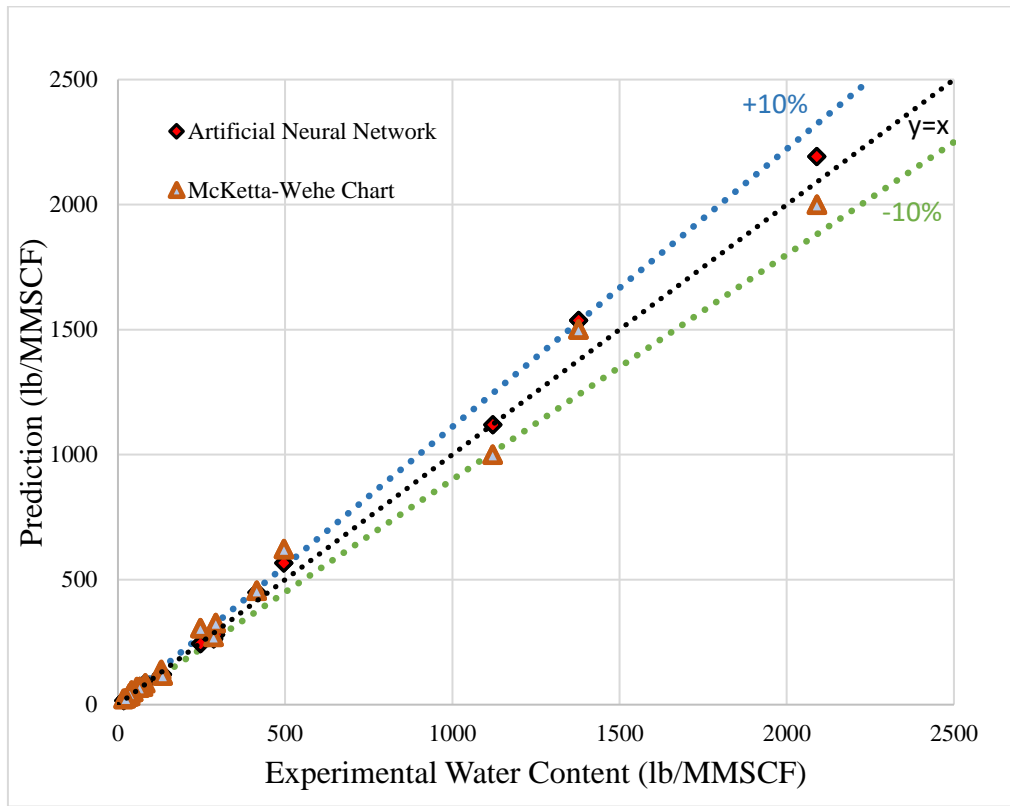
Figure 23. *Correlation between experimental data and model predictions*

The contribution of heavy hydrocarbons in this work has allowed the development of a model with applicability to calculate water content in natural gas of various types. Considering the analysis of the influence on the water content that heavy hydrocarbons contribute, it is deduced that a large quantity of these heavy hydrocarbons in the mixture could decrease the water content calculated with other available methods by up to 30%. The composition of heavy hydrocarbons could then be taken as a variable that reduces the amount of water in natural gas. The composition of heavy hydrocarbons contrasts with the variable of acid gas content, which increases the amount of water in natural gas. Also, with the variable pressure and temperature fundamental in the calculation of water content. This model makes it possible to precisely calculate the water content for various types of gas and in a suitable operating range for the typical pressure and temperature conditions to which the gas is subjected in its processing.

CONCLUSIONS AND RECOMMENDATIONS

o The artificial neural network model developed for calculating water content in natural gas, taking into account the GPM as heavy hydrocarbon contribution, obtained satisfactory and accurate results.

o The analysis of heavy hydrocarbons' contribution in calculating the water content in natural gas was carried out based on the available methods due to the lack of studies on the importance of heavy compounds in water content.

o The artificial neural network model developed was shown to predict the water content efficiently for sour and sweet gas mixtures, in the same way for rich and lean gases. This represents a huge advantage in the applicability of the model for natural gas dehydration plants' design.

o Given the great variety of natural gas mixtures present in Ecuador and this being a producer of gaseous hydrocarbons, an accurate tool for estimating the water content in NG is of utmost importance to ensure the operational continuity of gas processing and satisfy the internal demand for natural gas and its derivatives such as LPG and NGL.

o It is recommended to use this neural network only for natural gas mixtures in the operating pressure and temperature range defined in this work.

o For future studies are necessary experimental measurements of water content at the same pressure and temperature intervals for different gas types to better observe the influence of heavy hydrocarbons on natural gas's water content.

REFERENCES

[1]   BP pcl, Statisical Review of World Energy, 69 ed., 2020. [Online]. Available: https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html.

[2]   A. J. Kidnay and W. R. Parrish, Fundamentals of Natural Gas Processing, Boca Raton: Taylor and Francis Group, 2006. ISBN: 978-0-8493-3406-1.

[3]   S. Mokhatab, W. A. Poe and J. Y. Mak, Handbook of Natural Gas Transmission and Processing, 3 ed., Cambridge: Elsevier, 2015. ISBN: 978-0-12-801499-8.

[4]   M. Netusil and P. Ditl, «Natural Gas Dehydration», 2012. ISBN: 978-953-51-0820-7. doi: 10.5772/45802.

[5]   V. Aimikhe and M. Adeyemi, «A Critical Evaluation of Natural Gas-water Formula Correlations», *Journal of Scientific Research & Reports,* vol. 25, pp. 1-20, 2019. doi: 10.9734/JSRR/2019/v25i630206.

[6]   A. Bahadori, H. B. Vuthaluru and S. Mokhatab, «Rapid Estimation of Water Content of Sour Natural Gases», *Journal of the Japan Petroleum Istitute,* vol. 52, n° 5, pp. 270-274, 2009. doi: 10.1627/jpi.52.270.

[7]   V. Aimikhe, O. Joel, S. Ikiensikimama and S. Iyuke, «Experimental Determination of Water Content of Sweet Natural Gas with Methane Component Below 70%», *International Journal of Scientific and Engineering Research,* vol. 8, 2017.

[8]   A. Mohammadi, V. Samieyan and B. Tohidi, «Estimation of Water Content in Sour Gases», *SPE,* vol. 94133, 2005. doi: 10.2118/94133-MS.

[9]   A. Bahadori, Natural Gas Processing, Houston: Gulf Professional Publishing, 2014. doi: ISBN:9780080999715.

[10]   K. L. Du and M. N. S. Swamy, Neural Networks in a Softcomputing Framework, Germany: British Library Cataloguing, 2016. ISBN-13: 9781846283024.

[11]   A. H. Mohammadi and D. Richon, «Use of Artificial Neural Networks for Estimating Water Content of Natural Gases», *Ind. Eng. Chem. Res.,* vol. 46, pp. 1431-1438, 2007.

[12]   M. A. Ahmadi, R. Soleiman and A. Bahadori, «A Computational Intelligence Scheme for Prediction of Equilibrium Water Dew Point of Natural Gas in TEG Dehydration Systems», *Fuel,* vol. 137, pp. 145-154, 2014. doi: 10.1016/j.fuel.2014.07.072.

[13]   S. Al-Fattah and R. Startzman, «Neural Network Approach Predicts U.S. Natural Gas Production», *SPE Production & Facilities,* vol. 82411, 2003. doi: 10.2118/82411-PA.

[14]   Petroamazonas EP, «Resumen Ejecutivo de Operaciones», 2020. [Online]. Available: https://www.petroamazonas.gob.ec/wp-content/uploads/downloads/2020/12/reporte_diario-36.pdf.

[15]   M. G., Gas Natural, Latacunga: Universidad de las Fuerzas Armadas, ESPE-Extensión Latacunga, Ingenieria Petroquimica, 2019.

[16]   EU Redacción, En Bajo Alto, 350 familias esperan que no aumente gas por tuberías, *EL Universo,* 28 April 2019. [Online]. Available: https://www.eluniverso.com/noticias/2019/04/28/nota/7305626/bajo-alto-350-familias-esperan-que-no-aumente-gas-tuberias/.

[17]   Petroamazonas EP, «Reporte Gerencial», 2018. [Online]. Available: https://www.petroamazonas.gob.ec/?p=9880.

[18]   M. Ricaurte, D. Viloria and M. Echeverria, Operational Improvements in the Shushufindi Gas Plant as an Alternative to Increasing the Availability of Liquefied Petroleum Gas (LPG) in the Ecuadorian Market, Yachay Tech, Urcuquí, Escuela de Ciencias Químicas e Ingeniería, 2020.

[19]   V. Nardone, M. MÜller, A. Bouza, Y. Sánchez and J. Montbrun, «Evaluación de modelos empíricos para la predicción de hidratos de gas natural», *Revista de la Facultad de Ingeniería Universidad Central de Venezuela,* vol. 24, nº 3, 2019. ISSN: 0798-4065.

[20]   V. J. Aimikhe, O. F. Joel, S. S. Ikiensikimama and S. Iyuke, «Evaluation of Charts in Determining Water Content of Sweet Natural Gas with Methane Component of 60 - 67%», *European Journal of Engineering and Technology,* vol. 7, nº 1, pp. 1-15, 2019. ISSN: 2056-5860.

[21]   A. Terrigeol, C. Coquelet and A. Chapoy, Water Content Assessment in Acid Gas, Italy, Florence: Gas Processors Association – Europe, 2015.

[22]   S. Faramawy, A.-. Zaki and E. Sakr, «Natural Gas Origin, Composition, and Processing: A review», *Journal of Natural Gas Science and Engineering,* vol. 34, pp. 34-54, 2016. doi: 10.1016/j.jngse.2016.06.030.

[23]   Quadrennial Energy Review, «Energy Transmission, Storage, and Distribution Infrastructure», QER, report, 2015. [Online]. Available: https://www.energy.gov/sites/prod/files/2015/07/f24/QER_NaturalGasAppendix.pdf-.

[24]   S. Vicuña and M. Parreño, Evaluación de la Factibilidad de Industrializar Helio a partir de Gas Natural y Gases de Pozos Petroleros Producidos en Ecuador, Escuela Politécnica Nacional, Facultad de Ingeniería Química y Agroindustria, 2011.

[25]   G. P. S. A. (U.S.) and G. P. Association, Engineering data book, 12th ed., vol. I & II, Tulsa, Oklahoma: Gas Processors Suppliers Association, ©2012, 2004.

[26]   Congressional Research Service, «Natural Gas Liquids: The Unknown Hydrocarbons», CRS Reports, 2018. [Online], Available: shorturl.at/ptvRS.

[27]   R. J. Clews, Project Finance for the International Petroleum Industry, London: Elsevier, 2016. ISBN: 978-0-12-800158-5.

[28]   E. Erdmann, A. Ruiz, J. Martínez, J. Gutierrez and E. Tarifa, «Natural Gas Sweetening with Amines. Process Simulation and Parametric Sensitivity Analysis», *Avances en Ciencias e Ingeniería,* vol. 3, nº 4, pp. 89-101, 2012. ISSN: 0718-8706.

[29]   C. Olsen, T. Kozman, J. Lee and K. Yuvamitra, «A Comparative Study of Natural Gas Liquids Recovery Methods», *Distributed Generation & Alternative Energy Journal,* vol. 27, 2012. doi: 10.1080/21563306.2012.10505411.

[30]   Instituto Ecuatoriano de Normalización, Norma Técnica Ecuatoriana, NTE INEN 2 489, Quito: INEN, 2009. [Online]. Available: https://www.normalizacion.gob.ec/buzon/normas/2489.pdf.

[31]   D. L. Christensen, Thermodynamic Simulation of Water/glycol Mixture, Denmark: Aalborg University Esbjerg, 2009. Available: https://projekter.aau.dk/projekter/files/17059482/Gas_Dehydration.pdf.

[32]   A. Bahadori and H. B. Vuthaluru, «Rapid estimation of equilibrium water dew point of natural gas in TEG», *Journal of Natural Gas Science and Engineering,* vol. 1, p. 68–71, 2019. doi: 10.1016/j.jngse.2009.08.001.

[33]   F. Gironi, M. Maschietti and V. Piemonte, «Natural Gas Dehydration: A Triethylene Glycol-Water System Analysis», *Energy Sources Part A-recovery Utilization and Environmental Effects,* vol. 32, pp. 861-1868, 2010. doi: 10.1080/15567030902804756.

[34]   R. Sandoval and G. Sevilla, Integración de la Planta Deshidratadora de Gas Natural de Bajo Alto en Machala a la Plataforma Scada del Centro de Monitoreo y Control Hidrocarburífero (CMCH) para el Control y la fiscalización que ejerce la Agencia de Regulación y Control, Universidad Central del Ecuador, Facultad de Ingeniería de Geología, Minas y Petróleos, y Ambiental., 2013.

[35]   J. M. Campbell, «Gas Processing in 1982», *Society of petroleum Engineers of AIME,* pp. 465-467, 1982). doi: 10.2118/10895-PA.

[36]   T. Yan and C. J. H. Honggang, «Formation of Carbonyl Sulfide During Molecular Sieve Dehydration Process from High Sour Natural Gas in Chin,» *Society of Petroleum Engineers,* vol. 131908, pp. 1-4, 2010. doi: 10.2118/131908-MS.

[37]   J. Campbell, «Chapter 9» in *Technical Assistance Service for the Design, Operation, and Maintenance of Gas Plants*, Campbell Petroleum Series, 2003, pp. 5-10. doi: 10.9707/1944-5660.1342.

[38]   M. Stewart, «Surface Production Operations» in *Design of Gas-Handling Systems and Facilities*, 3 ed., vol. 2, Houston, Texas, Gulf Publishing, 2014, pp. 229-232.

[39]   A. A. Mohammad, A. Zainal, K. P. Le Thi, K. Tomoaki and B. Alireza, «Estimation of Water Content of Natural Gases Using Particle Swarm Optimization Method», *Petroleum Science and Technology,* vol. 34, nº 7, pp. 595-600. doi: 10.1080/10916466.2016.1153655.

[40]   J. Carroll, «Chapter 10, Water Content of Natural Gas», in *Natural Gas Hydrates A Guide for Engineers*, Gulf Professional Publishing, 2014, pp. 263-299. doi: ISBN: 9780128005750.

[41]   K. Nasrifar, F. Alavi and J. Javanmardi, «Prediction of Water Content of Natural Gases Using the PC-SAFT Equation of State», *Fluid Phase Equilibria An International Journal,* nº 453, pp. 40-45, 2017. doi: 10.1016/j.fluid.2017.08.023.

[42]   L. Changjun, X. Wu, W. Jia and O. Sun, «New correlations for water content of sweet and sour natural gases», *International Journal of Oil Gas and Coal Technology,* vol. 6, pp. 392-407, 2013. doi: 10.1504/IJOGCT.2013.054853.

[43]   R. Maddox, L. Lilly, M. Moshfeghian and E. Elizondo, «Estimating Water Content of sour Natural Gas Mixtures», Norman, Laurance Reid Gas Conditioning Conference, 1998.

[44]   L. Zhu, L. Luling, J. Zhua, L. Qin and F. Junming, «Analytical Methods to Calculate Water Content in Natural Gas», *Chemical Engineering Research and Design,* nº 93, p. 148–162, 2015. doi: 10.1016/j.cherd.2014.05.021.

[45]   V. N. Kabadi and R. P. Danner, «A Modified Soave-Redlich-Kwong Equation of State for Water-Hydrocarbon Phase Equilibria», *Ind. fng. Chem. Process Des. Dev,* vol. 24, pp. 537-541, 1985. doi: 01 96-4305/85/1124-0537.

[46]   SimSci, PRO/II 10.0, Concise Help Contents, Schneider Electric Software LLC, 1994-2016.

[47]   A. Hassan, Phase Behavior of Multicomponent Mixtures of Complex Molecules in Supercritical Fluids, Oregon State University, Department of Chemical Engineering, 1997.

[48]   S. Sharma and J. Campbell, «Predict Natural-gas Water Content with Total Gas Usage», *Oil & Gas Journal,* vol. 67, nº 31, p. 136–137, 1969.

[49]   J. H. Robinson, E. Wichert, R. G. Moore and R. A. Heidemann, «Estimation of the water content of sour natural gases», *Oil and Gas Journal,* p. 76–78, 1978. doi: 10.2118/6098-PA.

[50]   B. Guo and A. Ghalambor, Natural Gas Engineering Handbook, Houston: Gulf Publishing Company, 2012. ISBN: 9780127999951.

[51]   C. C. Aggarwal, Neural Networks and Deep Learning, Gewerbestrasse: Springer Nature Switzerland AG, 2018. ISBN 978-3-319-94462-3.

[52]   M. C. Nwadiugwu, Neural Networks, Artificial Intelligence and the Computational Brain, Nigeria: University of Ilorin, Department of Anatomy, 2020.

[53]   G. Crescenzio, Artificial Neural Networks: Tutorial, Italy: University of Foggia, Department of Clinical and Experimental Medicine, 2015.

[54]   J. Hilera and H. V. Martinez, Redes Neuronales Artificiales : Fundamentos, Modelos y Aplicaciones, Madrid: SERBIULA (sistema Librum 2.0), 1995.

[55]   K. Gurney, An Introduction to Neural Networks, London: Taylor & Francis Group, 1997 ISBN 0-203-45151-1.

[56]   K. Gil, Neural Network Supervision: Notes on Loss Functions, Labels and Confidence Estimation, University of Passau, Faculty of Computer Science and Mathematics, 2019.

[57]   T. Hinz, P. Barros and S. Wermter, «The Effects of Regularization on Learning Facial Expressions with Convolutional Neural Networks», *Springer: Lecture Notes in Computer Science,* pp. 80-87, 2016. ISBN: 978-3-319-44780-3. doi: 10.1007/978-3-319-44781-0_10.

[58]   I. Livieris and P. Pintelas, Evaluation of Neural Network Training Algorithms Using Biomedical Data, University of Patras, Department of Mathematics, 2008.

[59]   A. Zhu, Y. Meng and C. Zhang, «An Improved Adam Algorithm Using look-ahead», *Natural science fundation,* pp. 19-22, 2017. doi: 10.1145/3094243.3094249.

[60]   Y.-C. Du and A. Stephanus, «Levenberg-Marquardt Neural Network Algorithm for Degree of Arteriovenous Fistula Stenosis Classification Using a Dual Optical Photoplethysmography Sensor», *Sensors,* vol. 18, p. 2322, 2018. doi: 10.3390/s18072322.

[61]   X. Ying, «An Overview of Overfitting and its Solutions», *Journal of Physics: Conference Series,* vol. 1168, p. 022022, 2019. doi: 10.1088/1742-6596/1168/2/022022.

[62]   H. Abbass, «A Memetic Pareto Evolutionary Approach to Artificial Neural Networks», *Lecture Notes in Artificial Intelligence,* vol. 2256, pp. 113-152, 2001. doi: 10.1007/3-540-45656-2_1.

[63]   M. Ricaurte and C. Veloz, Predictive Modeling of the Primary Settling Tank in the Ibarra Wastewater Treatment Plant based on Artificial Neural Networks, Yachay Tech, Urcuquí, Escuela de Ciencia Químicas e Ingeniería , 2020.

[64]   [BigML], https://bigml.com. BigML, Inc. Corvallis, Oregon, USA, 2011.

[65]   D. Zhang, «A Coefficient of Determination for Generalized Linear Models», *The American Statistician,* vol. 71, 2016. doi: 10.1080/00031305.2016.1256839.

[66]   M. A. Clark, Experimentally Obtained Saturated Water Content, University of Calgary, Faculty of Graduate Studies, 1999.

[67]   A. H. Mohammadi, A. Chapoy, B. Tohidi and D. Richon, «A Semiempirical Approach for Estimating the Water Content of Natural Gases», *Ind. Eng. Chem. Res.,* vol. 43, pp. 7137-7147, 2004. doi: 10.1021/ie049867m.

[68]   A. Tatar, S. Nasery, M. Bahadori, A. Bahadori, M. Bahadori, A. Barati-Harooni and A. Najafi-Marghmaleki, «Prediction of Water Removal Rate in a Natural Gas Dehydration System Using Radial Basis Function Neural Network», *Petroleum Sciense and Technology,* vol. 34, nº 10, pp. 951-960, 2016. doi: 10.1080/10916466.2016.1166131.

[69]   N. Draper and H. Smith, Applied Regression Analysis, 3 ed., John Wiley & Sons, Inc., 1998. ISBN: 9780471170822. doi: 10.1002/9781118625590.ch15.

APPENDIX

APPENDIX A:

Validation data set and comparison

.

The data group's composition to validate the ANN developed in this work is found in Table 7 for natural gas mixtures and Table 8 for synthetic mixtures. Table 9 shows the comparison of the ANN developed results with the available methods for estimating water content.

Table 7. *Validation Group Natural Gas Mixtures*

Composition (mol%)	Natural Gas Mixtures (NG)										
	NG1	NG2	NG3	NG4	NG5	NG6	NG7	NG8	NG9	NG10	NG11
C_1	63.73	62.16	98.20	93.22	88.21	86.35	94.36	83.85	66.12	86.90	61.25
C_2	13.51	12.19	0.56	2.92	8.36	6.19	2.64	3.46	11.86	7.36	15.86
C_3	8.63	7.58	0.19	0.72	1.76	1.55	0.96	0.66	7.35	3.17	8.87
iC_4	3.40	3.67	0.03	0.09	0.29	0.21	0.00	0.09	1.89	0.64	1.77
nC_4	4.60	4.73	0.04	0.14	0.44	0.31	0.44	0.13	3.36	0.77	2.36
iC_5	1.90	1.82	0.00	0.00	0.00	0.00	0.00	0.00	1.90	0.25	1.94
nC_5	1.50	1.88	0.01	0.06	0.03	0.13	0.00	0.07	0.93	0.16	0.89
C_6	2.10	0.19	0.01	0.05	0.00	0.06	0.00	0.07	1.87	0.04	1.10
C_7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00
N_2	0.14	4.15	0.84	1.94	0.91	4.86	1.00	10.35	2.14	0.03	4.14
CO_2	0.85	1.63	0.11	0.85	0.00	0.17	0.60	1.29	2.58	0.65	1.82
H_2S	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
He	0.00	0.00	0.02	0.03	0.00	0.15	0.00	0.04	0.00	0.00	0.00

Table 8. *Validation Group Synthetic Mitures*

Composition (mol%)	Synthetic Gas Mixtures (SM)														
	SM1	SM2	SM3	SM4	SM5	SM6	SM7	SM8	SM9	SM10	SM11	SM12	SM13	SM14	SM15
C1	70.0	72.5	89.0	100.0	94.0	67.5	96.0	70.0	80.0	50.0	86.9	70.0	83.0	71.3	96.9
C2	0.0	0.0	0.0	0.0	4.0	4.5	3.0	11.0	0.0	0.0	7.3	0.0	0.0	0.0	3.1
C3	0.0	0.0	0.0	0.0	0.0	1.9	1.0	6.0	0.0	0.0	3.2	0.0	0.0	3.8	0.0
iC4	0.0	0.0	0.0	0.0	0.0	0.5	0.0	4.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0
nC4	0.0	0.0	0.0	0.0	0.0	0.7	0.0	4.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0
iC5	0.0	0.0	0.0	0.0	2.0	0.0	0.0	2.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0
nC5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
C6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CO2	30.0	0.0	11.0	0.0	0.0	18.8	0.0	0.0	20.0	50.0	0.6	30.0	0.0	6.3	0.0
H2S	0.0	27.5	0.0	0.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0	17.0	18.8	0.0
He	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 9. *Evaluation of ANN vs Available methods for calculating Water Content in NG*

Samples	Pressure (psia)	Temperature (°F)	Wcexp (lb/MMSCF)	This work (lb/MMSCF)	SRKKD EoS (lb/MMSCF)	PRM EoS (lb/MMSCF)	Bukacek Correlation (lb/MMSCF)	Mckketa-Wehe Chart (lb/MMSCF)
NG1	73.00	51.80	130.00	119.07	122.58	121.20	133.67	140.00
NG2	76.00	50.90	128.00	115.84	116.85	116.75	124.22	135.00
NG3	72.52	32.00	55.57	53.64	59.64	58.64	NA	70.00
	217.56	41.00	28.74	29.63	29.40	28.27	NA	34.00
	870.23	50.00	11.92	9.68	12.67	10.64	18.26	17.00
NG4	217.56	41.00	28.45	28.07	29.57	28.63	NA	34.00
	870.23	59.00	16.91	16.47	17.42	15.16	24.13	24.00
NG5	72.52	41.00	81.69	86.07	84.64	83.84	NA	86.00
NG6	1160.30	68.00	17.19	15.41	19.29	18.13	27.03	22.00
	217.56	41.00	28.07	26.68	29.43	28.45	NA	34.00
NG7	1514.10	93.00	36.80	36.92	34.90	32.94	47.11	34.00
	1014.10	66.00	23.10	18.93	19.73	16.97	27.48	28.00
	1014.10	86.00	38.30	38.52	37.00	33.68	48.54	40.00
	1514.10	81.00	25.50	24.76	24.50	20.44	34.25	24.00
NG8	72.52	32.00	57.95	55.24	59.04	58.29	NA	70.00
	217.56	59.00	55.10	61.35	56.55	55.70	64.58	64.00
NG9	179.00	55.40	65.00	73.41	57.47	57.09	67.97	70.00
NG10	166.00	54.50	75.00	74.16	62.14	61.29	70.34	72.00
NG11	184.00	55.40	65.00	70.78	56.75	56.39	66.39	68.00
SM1	435.11	68.00	45.55	41.79	41.80	40.76	50.15	51.47
SM2	1392.00	160.00	247.00	243.86	206.84	232.46	244.40	306.27
SM3	1000.00	160.00	286.00	263.14	267.23	278.72	308.03	271.57
SM4	141.56	67.19	119.69	123.41	113.59	112.83	123.82	126.00
	146.20	76.01	147.24	156.02	151.84	158.33	160.66	160.00
	560.57	76.01	41.61	43.32	50.79	55.20	53.59	50.00
SM5	442.37	190.67	1120.91	1119.64	1079.71	1115.06	1155.79	1000.00
	74.11	85.91	402.29	403.37	372.83	379.83	412.87	400.00
	279.05	119.93	305.88	298.95	289.32	294.32	328.47	325.00
	710.97	139.55	215.16	210.86	211.90	214.99	248.23	235.00
	82.24	165.47	3272.48	2598.50	2400.75	2503.89	3213.51	3000.00
	138.37	190.49	3315.23	2974.58	2560.28	2667.08	3385.34	3200.00
SM6	1499.55	199.99	477.48	572.93	470.20	548.09	522.08	506.39
	200.01	120.00	415.12	448.18	402.55	414.70	443.45	455.58
SM7	2059.54	304.27	2089.83	2192.69	2148.88	1889.86	2260.46	2000.00
SM8	3015.33	301.01	1377.39	1536.62	1486.12	1833.13	1662.04	1500.00
	542.44	301.01	6364.48	5212.01	4937.51	5086.29	6545.05	6500.00
SM9	2000.00	100.00	40.60	45.71	42.54	40.40	49.28	55.60
SM10	870.23	68.00	30.21	30.34	30.37	33.82	31.66	32.28
SM11	425.00	62.60	35.00	34.97	35.11	33.49	43.31	44.11
SM12	435.11	104.00	133.27	120.25	127.71	128.67	145.60	115.99
SM13	1010.00	160.00	292.00	279.32	278.04	286.11	305.79	326.75
SM14	1499.55	199.99	496.24	566.10	500.57	546.43	522.08	621.77
SM15	580.15	59.00	23.46	22.95	23.74	21.71	30.87	26.00
AAD%				**7.00**	**7.30**	**9.38**	**14.54**	**13.60**

Note: WCexp stands experimental water content; AAD stands absolute average deviation; NA stands no applicable.

APPENDIX B:

Manual for the use of the ANN in BigML for

estimating water content in natural gas

Next, the steps for using the artificial neural network to estimate water content in natural gas mixtures developed in this work are listed.

1. Register on the BigML platform at the following link:

https://bigml.com/accounts/register/

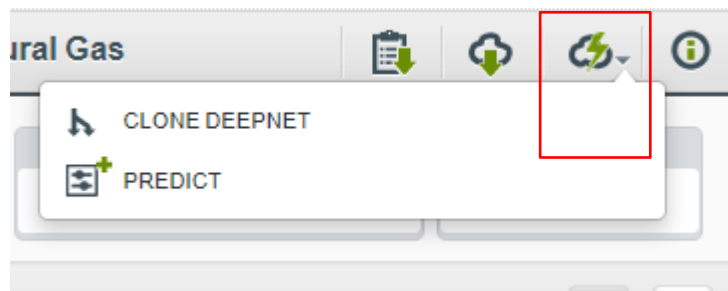-Opening a BigML account is free. Once the account is created, the account must be entered. Inside the platform, you should see this image.
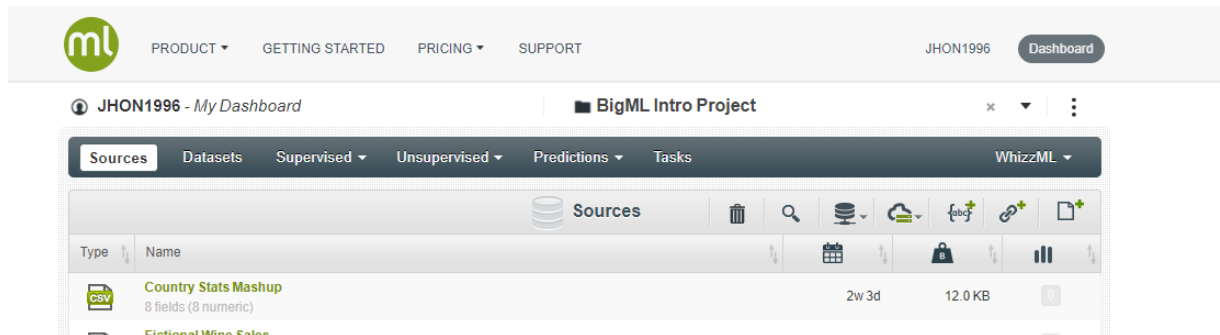
2. Once inside the platform, click on the following link

https://bigml.com/shared/deepnet/fO60BGP51o1njg9J9FrOL24bRDB
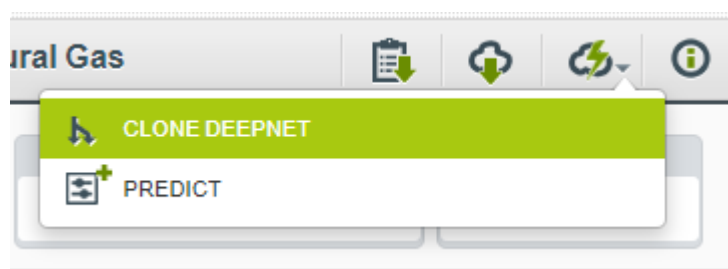
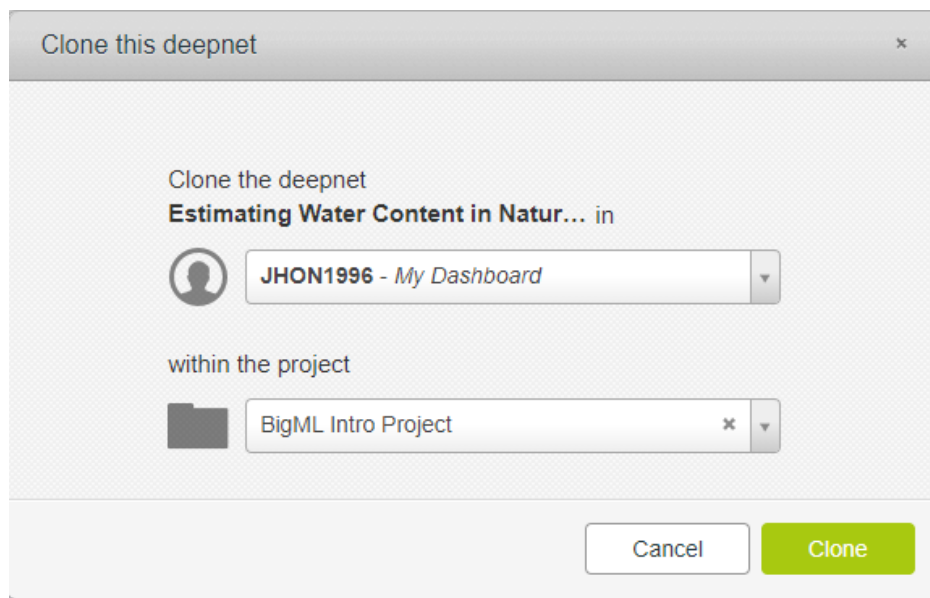This image should appear on the screen
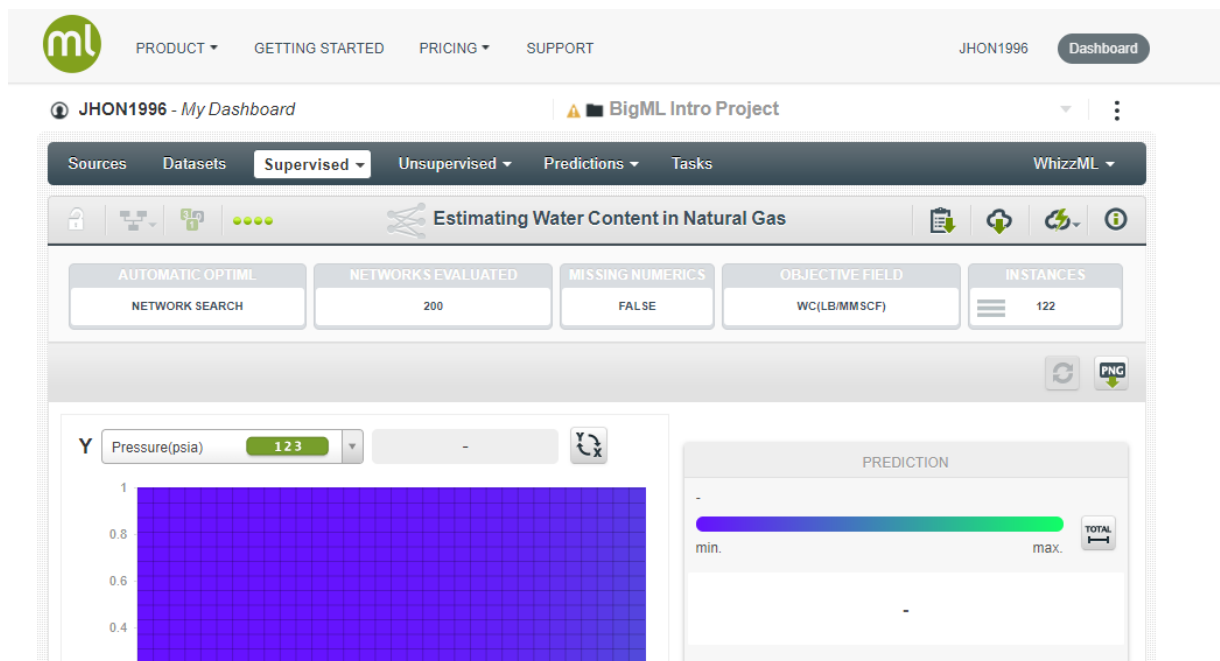
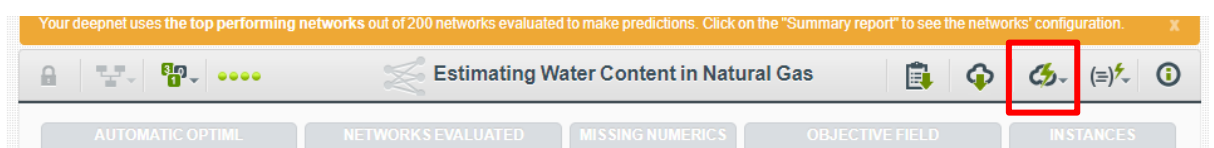3. Click on the symbol in the image that is framed in the red box

4. A box will appear in which you must click on the option "Clone Deepnet"
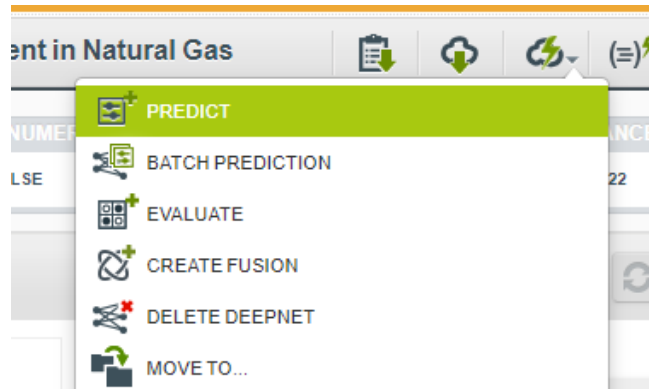

-Followed a box would appear where you must confirm in "Clone"

5. Once the artificial neural network is cloned, you can use it to make predictions in the box marked with red in the following image.

-A box appears where you can see the option "Predict"

6. Finally, the prediction interface appears where the normalized values of pressure, temperature, mol equivalent H_2S (Y´), and GPM of the gas mixture that you want to predict the water content must be located.

-Remember to have the data in the English system of units and then apply Eq. (17) to normalize them with the maximum and minimum values for each variable used in the ANN training tabulated in Table 10.

Table 10. *Values to normalize data in ANN*

	GPM	Pressure(psia)	Temperature(°F)	Y´	Water content
Min	0.00	64.00	32.00	0.00	11.78
Max	10.60	3156.02	304.27	0.89	7124.42

-Remember that the resulting water content is normalized, so the normalization must be reversed with the water content values in Table 10.