SOURCE ROCK RECONSTRUCTION AND HYDROCARBON MASS BALANCE ASSESSMENT – APPLIED TO THE WOLFCAMP FORMATION IN THE PERMIAN DELAWARE BASIN

A Dissertation

by

ANINDITO SATRIO BASKORO

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Ethan L. Grossman
Co-Chair of Committee,	Alan Yu
Committee Members,	Mauro Becker
	Hadi Nasrabadi
Head of Department,	Julie Newman

December 2022

Major Subject: Geology

ABSTRACT

Improved efficiency of hydrocarbon resources exploration is directly linked to an accurate petroleum system analysis, including detailed quantification of hydrocarbon (HC) masses involved in a basin from generation to production. The quantified HC masses must be balanced following the fundamental laws of mass conservation. This is particularly important for a basin with unconventional-conventional petroleum systems where the generated HCs from the source rock are either expelled out from or retained within the source rock. Expulsion efficiency becomes even more critical as it largely determines the HCs in place, both for the source rock or out of the source rock accumulations. Before performing the hydrocarbon mass balance calculation, basin-wide source rock (initial quality and quantity and litho- and organo- facies) is reconstructed following the proposed workflow. The workflow is applied to Wolfcamp Formation in the Permian Delaware Basin as it is an excellent example of unconventional-conventional petroleum systems. Firstly, the proposed kinetic-based inversion workflow results show a considerable range of restored initial hydrogen index, ca. 188-594 mgHC/gTOC, and total organic carbon, ca. 1.49-3.17 wt%, values indicating source rock heterogeneity of the Wolfcamp. The workflow restores the measured HI and TOC significantly as Wolfcamp is within the oil to gas generation (transformation ratio ca. 56-84 %) window but initially has a good to excellent source rock potential. Secondly, seven basin-wide lithofacies are identified within the Wolfcamp Play by applying unsupervised machine learning workflows on well-log measurements. Wolfcamp source rock basinal lithofacies is identified as organic-rich siliceous mudstone with marine clastic organofacies dominated to mixed argillaceous mudstone with mixed marine-terrigenous organofacies dominated, interbedded with siltstone and limestone, reflecting a shallow-to-deeper marine to deep marine depositional environments. Finally, the restored source rock property and identified basinal facies are integrated to the 3D Delaware basin model to calculate the generated HC. The calculated generated HC volumes (p90/50/10) are 705/1,332/3,016 Billion Barrels of Oil Equivalent (BBOE). Ultimately, the proposed workflow assesses the mass balance and estimates the mean remaining recoverable hydrocarbon of 33/65/149 BBOE with mean overall expulsion efficiency of 50%, which may be used as a reference in modeling unconventional interbedded play.

DEDICATION

I dedicate this dissertation to my wife, Silvania Dwi Utami, and my son, Byan Arrasyad Baskoro. This one marks our first family project as a legacy for the world.

In memory of my mother, I miss you. Thanks for your countless joys and love that shape me into who I am today.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Ethan Grossman, my committee co-chair, Dr. Alan Yu, and my committee members, Dr. Mauro Becker and Dr. Hadi Nasrabadi, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues in College Station, Texas, my friends and professors at the Chevron CORE Basin Modeling Program, and the department faculty and staff for making my time at Texas A&M University a great experience.

Thanks to my father, mother, and brothers-sisters for their support and encouragement. Finally, thanks to my wife and son for their patience and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Dr. Ethan Grossman (advisor), Dr. Alan Yu (co-advisor), and Dr. Mauro Becker of the Department of Geology and Geophysics, and Dr. Hadi Nasrabadi of the Department of Petroleum Engineering.

Most of the data used in this dissertation were provided by Chevron corporation. The foundation of the 3D basin model was developed by Chevron CORE Basin Modeling Program's professors, including Dr. Alan Yu, Dr. Mauro Becker, and formerly Dr. Andrea Miceli. The development of Chapter 2 was also supervised by Friedemann Baur from Chevron. The python algorithm and workflow for Chapter 3 were developed in collaboration with Eliza Ganguly, MS Student in Petroleum Engineering TAMU (graduated) under the guidance of Dr. Siddharth Misra from the Department of Petroleum Engineering/ Geology and Geophysics TAMU. Chevron's geoscientists have also reviewed the manuscripts, including Barry Katz (Chapters 2 and 3) and Gary Muscio (Chapter 2).

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by Chevron Fellowship from Chevron corporation through the Chevron CORE Basin Modeling Program under the Berg-Hughes Center, Department of Geology and Geophysics, Texas A&M University.

Graduate study was also partially supported by Berg-Hughes Center Fellowship and Research Assistantship, Oswald Graduate Fellowship in Geology and M.T. Halbouty’30/ AAPG Foundation Endowed Fellowship by formerly College of Geoscience, ISS scholarship by International Student Service, and TAMU Energy Institute Fellowship.

NOMENCLATURE

AE	Accumulation Efficiency
BBOE	Billion Barrels of Oil Equivalent
DEN	Bulk Density log
DT	Compressional wave delay time (sonic log)
EE	Expulsion Efficiency
Fm(s)	Formation(s)
GR	Gamma Ray log
HC	Hydrocarbon
HI	Hydrogen Index
OF	Organofacies
PEF	Photo-Electric Factor log
P90/50/10	Percentile 90/50/10 (minimum/median/maximum)
RF	Recovery Factor
SR	Source Rock
TOC	Total Organic Carbon
TR	Transformation Ratio
USGS	United States Geological Survey
VR	Vitrinite Reflectance

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background and Motivation

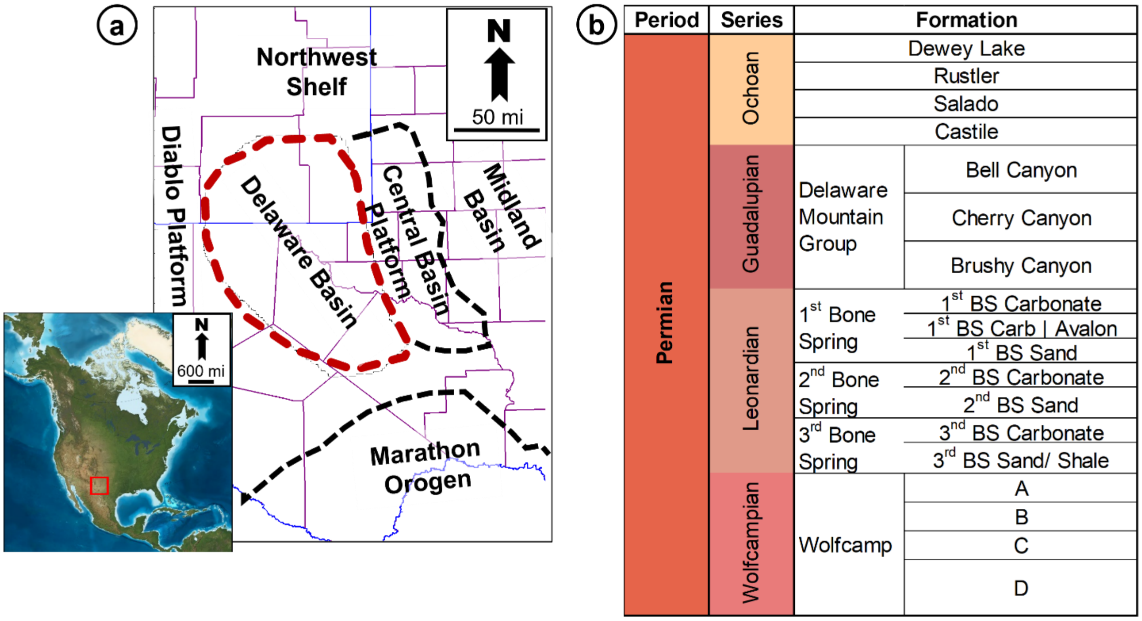
Detailed quantification of hydrocarbon masses involved in a basin from generation to production is necessary for an accurate petroleum system analysis (Hantschel and Kauerauf, 2009). The quantified hydrocarbon masses must be balanced following the fundamental laws of mass conservation. This is particularly important for a basin with combined unconventional-conventional petroleum systems where the generated hydrocarbon from the source rock interval is either expelled out from or retained within the source rock interval. Expulsion efficiency becomes even more critical as it largely determines the hydrocarbon in place, both for the source rock interval or out of the source rock accumulations.

Permian Delaware Basin exhibits an excellent example of combined unconventional-conventional petroleum systems with the Wolfcamp Formation as the important source rock and unconventional reservoir target. According to the USGS resources assessment (Gaswirth et al., 2018), the Wolfcamp and Bone Spring of the Delaware Basin are currently the largest continuous oil and gas resource potential (undiscovered yet technically recoverable) in the United States. To efficiently explore Wolfcamp sourced hydrocarbon resources, it is fundamental in quantifying the total generated hydrocarbon balance and its distribution throughout the basin from expulsion to accumulation and production is necessary. By assessing the mass balance of Wolfcamp's sourced hydrocarbon from generation to production, remaining recoverable resources and insight into expulsion efficiency in unconventional/ hybrid systems can be estimated more accurately.

1.2. Geological Overview

The Permian Delaware Basin is the western major structural subdivision of the Permian Basin in west Texas and southeastern New Mexico (Hills, 1984) (**Figure 1.1a**). The Delaware Basin is bounded to the west by the Diablo Platform, to the north by the North American craton and Northwestern Shelf areas, and to the south by the Marathon orogenic belt. The Delaware Basin is separated from the Midland Basin to the east by the N-S trending Central Basin Platform (Adams, 1965; Hills, 1984). The Delaware Basin occupies a basinal area of 13,000 mi^2 (33,500 km^2) with a maximum depth of 24,000 ft (7,300 m) and is filled by up to 40,000 mi^3 (170,000 km^3) of Phanerozoic sediments (Hills, 1984).

The Wolfcamp Formation (Wolfcampian Series) consists of four main intervals: Wolfcamp A, -B, -C, and -D (Bievenour and Sonnenberg, 2019; Dutton et al., 2005) and is part of the thick Permian strata in the Delaware Basin (**Figure 1.1b**). The deposition took place during the early Permian with a depositional environment of generally a deep, oceanic connected interior continental basin surrounded by carbonate shelf (the Northwestern Shelf, eastern Central Basin Carbonate Platform, western Diablo Platform), the northwestern Pedernal uplift, and the southern siliciclastic shelf in front of the Marathon orogenic belt (Blakey, 2019). Wolfcamp intervals were also deposited during the frequent eustatic fluctuations (Fairhurst et al., 2021). The complex depositional setting caused lithofacies heterogeneity, primarily limestone, silts, and mudstone, within the Wolfcamp intervals in the Delaware Basin.



Period	Series	Formation	
Permian	Ochoan	Dewey Lake	
		Rustler	
		Salado	
		Castile	
	Guadalupian	Delaware Mountain Group	Bell Canyon
			Cherry Canyon
			Brushy Canyon
	Leonardian	1st Bone Spring	1st BS Carbonate
			1st BS Carb | Avalon
			1st BS Sand
		2nd Bone Spring	2nd BS Carbonate
			2nd BS Sand
		3rd Bone Spring	3rd BS Carbonate
			3rd BS Sand/ Shale
	Wolfcampian	Wolfcamp	A
			B
			C
			D

Figure 1.1. (a) Major subdivision and boundaries of the Permian Basin around the Delaware Basin (after Dutton et al. (2015) and Silver and Todd (1969)). (b) Stratigraphic chart of Permian strata, which includes Wolfcamp interval subdivision for the Delaware Basin (after Bievenour and Sonnenberg (2019) and Dutton et al. (2015)).

Hydrocarbon generation from the Wolfcamp source rock (SR) started during the middle Permian (Adams, 1965; Hills, 1984; Kinley et al., 2008), influenced by complex tectonic history and thermal processes, including high thermal conductive evaporites (Adams, 1965; Gardner and Sonnenfeld, 1996; Hills, 1984, 1985; Kinley et al., 2008), intrusive and volcanic activities (Barker and Pawlewicz, 1987), and burial history. Recent studies (Becker et al., 2019; Yu et al., 2020) indicate that burial history is the most dominant factor for source maturity, with the maximum maturity reached during the mid-Permian at maximum burial, then continuing to generate hydrocarbons until the present-day source-rock condition (Hills, 1984; Lew et al., 2013; Schwartz et al., 2015).

There are several studies related to petroleum system analysis for the Wolfcamp and younger Formations in the Delaware Basin (Curtis and Zumberge, 2017, 2018; Fairhurst et al., 2021; Jarvie, 2017; Justman and Broadhead, 2010a, b; Sandvik et al., 1992). Wolfcamp SR is the primary source rock for the Wolfcamp unconventional/ hybrid reservoirs and lower Bone Spring interval with Wolfcamp marine shale and mixed upper Wolfcamp-lower Bone Spring SR facies (Curtis and Zumberge, 2017, 2018; Fairhurst et al., 2021; Jarvie, 2017; Pepper et al., 2020). The Wolfcamp marine carbonate-rich source rock, which is mainly restricted to eastern and northern Delaware Basin margins, is primarily responsible for the hydrocarbon conventional accumulations in the Delaware Basin margins and as far as Northwest Shelf and Central Basin Platform (Curtis and Zumberge, 2017, 2018; Pepper et al., 2020). The signatures of these two different oil families are documented in Curtis and Zumberge (2018); Echegu et al. (2021); and Pepper et al. (2020).

1.3. Related Works and Objectives

Several authors have attempted to perform hydrocarbon mass balance calculations at both basin and field scales (e.g., Baur et al., 2011; Carlton, 2017;

Coutinho, 2008; Duran et al., 2013; Lewan et al., 2002; Muscio et al., 2016). Most of these studies aimed to understand the migration efficiency for conventional hydrocarbon accumulation, calibrated to production data from the conventional field. While expulsion efficiency and retained hydrocarbon portion within the source rock interval were not adequately handled.

Other studies have been performed on modeling or assessing the mass balance of unconventional hydrocarbon accumulations. For example, Rushing et al. (2004) and Kuchinskiy et al. (2012) quantified hydrocarbon volumes in unconventional play with expulsion efficiency calculated using several assumptions such as retaining factor and fractions of hydrocarbons lost from the system. However, the results were not assessed for their mass balance for the whole petroleum system. Expulsion efficiency is calculated using several assumptions, while source rock heterogeneity is generally simplified. Other examples are from basin and petroleum system modeling studies applied to the unconventional system, including Amer et al. (2015); Bryant et al. (2013); Cander (2012); Jarvie et al. (2007); Romero-Sarmiento et al. (2013), among many others. Some of these studies addressed the challenge of modeling unconventional systems, including factors controlling the retained hydrocarbon storage capacity and fluid properties. However, the calculated retained and expelled hydrocarbons were not assessed for their mass balance.

Before the quantity of generated hydrocarbon mass can be calculated as the starting point of hydrocarbon mass balance calculation, it is necessary to:

1) accurately model the thermal maturity (Baur et al., 2011; Coutinho, 2008; Duran et al., 2013; Jarvie et al., 2007; Pepper, 1991; Sandvik et al., 1992),
2) restore the present-day source-rock conditions (hydrogen index (HI) and total organic carbon (TOC)) to their initial state (Jarvie et al., 2007; Kuchinskiy et al., 2012; Lewan et al., 2002; Modica and Lapierre, 2012; Rushing et al., 2004), and
3) understand source rock litho- and organo- facies (Coutinho, 2008; Donovan et al., 2017; Evenick, 2016; Jarvie et al., 2007; Katz et al., 2017; Pepper and Corvi, 1995).

Thermal maturity plays a vital role in kerogen's thermal alteration and cracking process (Jarvie et al., 2007). With increasing thermal maturity, source rock may enter the gas generation window and/or undergo secondary cracking that results in more gas generation, increasing the retained hydrocarbon gas/oil ratio (GOR), and increasing pore pressure and micro-fracturing that enhance the expulsion (Jarvie et al., 2007). At the same time, thermal maturity increases the organic porosity and adsorption capacity, increasing the storage for the retained hydrocarbon within the source rock, and thus decreasing expulsion efficiency and affecting the generated fluid property (Baur et al., 2011; Coutinho, 2008; Duran et al., 2013; Jarvie et al., 2007; Pepper, 1991; Sandvik et al., 1992).

The main goal of restoring the present-day source rock to its initial state is to accurately characterize the initial source rock quality to calculate the total hydrocarbon generation potential. The transformation of organic matter from its initial state to present-day conditions will generate hydrocarbons and form organic porosity within the kerogen, which will also determine the storage capacity for retained hydrocarbon within the source rock (Jarvie et al., 2007; Katz and Arango, 2018; Modica and Lapierre, 2012; Romero-Sarmiento et al., 2013).

Various methods have been proposed to restore the measured source rock to its initial condition aiming to calculate the volume of hydrocarbon generated (e.g., Chen and Jiang (2015); Chen et al. (2016); Jarvie et al. (2007); Justwan and Dahl (2005); Peters et al. (2006); among many others). Most of these methods require measured HI data and assumptions of initial HI (HI_i) to calculate the transformation ratio (TR) and then restore the initial TOC (TOC_i) from measured TOC data. However, estimation of HI_i is challenging for samples with high maturity and for cases in which immature samples are unavailable, or maceral data is lacking. Another challenge is to provide a representative HI_i value for heterogeneous source rocks. Using published methods to restore HI_i and TOC_i typically add considerable uncertainty to the source rock model (e.g., calculated present-day HI and TOC in basin model are not calibrated to measured HI and TOC data) and ultimately to the calculated volume of generated hydrocarbon.

Source rock lithofacies and thickness affect formation overpressure and therefore, hydrocarbon storage mechanism and capacity, which ultimately control the expulsion efficiency of the source rock (Coutinho, 2008; Jarvie et al., 2007; Katz et al., 2017). The organofacies (Donovan et al., 2017; Evenick, 2016; Pepper and Corvi, 1995) directly impact the fluid property and composition (e.g., marine carbonate organofacies have the highest C_{15+} amount) and fluid phase (i.e., mixed marine-terrigenous organofacies are oil and gas prone) of the source rock. The properties of the generated hydrocarbon will ultimately affect the properties of the expelled and retained hydrocarbon (Baur, 2019; di Primio and Horsfield, 2006; Jarvie et al., 2007; Pepper and Corvi, 1995).

Wireline log measurements are often used for lithofacies identification as they are the most common well data and cover broad depth intervals of the formation of interest. With the development of data science and machine learning, numerous works related to lithofacies identification have been performed using various machine learning techniques applied to well-log data (Baldwin et al., 1990; Chang et al., 2002; Raeesi et al., 2012; Wang et al., 2014). They showed how machine learning methods could automate well-log-based lithofacies identification efficiently and comparable to a core description. However, most of these works are based on supervised machine learning methods, well-log and core-derived parameters, or integration with core data to classify lithofacies for the area of field scale. Nevertheless, classifying lithofacies based on well-log data is challenging for datasets lacking core data. Another challenge is consistently classifying lithofacies on a large basin-wide dataset.

Therefore, this study aims to fill these gaps by performing hydrocarbon mass balance calculations using several calculation scenarios, including probabilistic and basin and petroleum system modeling. Before performing the mass balance calculation (**chapter 4**), source rock initial property (**chapter 2**) and litho- and organo- facies (**chapter 2** and **3**) are characterized following the proposed alternative workflow to address the challenges reviewed above. Thermal maturity is a matter of intense discussion, and several studies are investigating it. This study will model thermal maturity under some simplified assumptions calibrated to the available thermal maturity indicator (**chapter 2** and **4**).

The proposed overall workflow of source rock reconstruction and mass balance calculation aims to achieve "bookkeeping" of the generated hydrocarbon from the source rock, considering the SR heterogeneity in unconventional/ hybrid systems. Ultimately, the proposed workflow assesses the mass balance to estimate the remaining recoverable hydrocarbon and expulsion efficiency, which can be used as a comparison or reference in modeling unconventional/ hybrid petroleum systems.

1.4. Summary of Each Three Chapters

Each of the three chapters discusses different parts of the overall workflow of source rock characterization and mass balance calculations. **Chapter 2** discusses the proposed workflow and results in restoring source rock initial HI and TOC as part of the source rock characterization. Organofacies identification and thermal maturity calibration are also discussed in **Chapter 2**. Similar to **Chapter 2**, **Chapter 3** also focuses on characterizing the source rock and emphasizing basin-wide lithofacies identification. Identified organofacies in **Chapter 2** are also tied to identified lithofacies from **Chapter 3** to identify the basin-wide gross depositional environment and organofacies distribution of the source rock intervals. Finally, **Chapter** Error! Reference source not found. i ntegrates the finding of characterized source rock from **Chapter 2** and **3** to calculate the volume of generated hydrocarbon. Finally, **Chapter 4** assesses the Wolfcamp sourced hydrocarbon mass balance using proposed scenarios to estimate remaining recoverable hydrocarbon resources and expulsion efficiency.

As of the time of submission of this dissertation, **Chapter 2** has been reviewed by Marine and Petroleum Geology (Elsevier's journal) and is currently under revision process. **Chapter 3** is currently under review with Unconventional Resources (Elsevier's journal). Both **Chapter 2** and 3 have also been reviewed by Chevron's geoscientists and also presented in the International Meeting for Applied Geoscience and Energy (IMAGE) 2021 by the American Association of Petroleum Geologists (AAPG) and Society of Exploration Geophysicists (SEG) at Denver, Colorado. **Chapter 4** is currently under review with Chevron and will be submitted to Interpretation: The Role of Geochemical Workflows in Understanding Resource Plays (AAPG's journal) once the review is done.

1.4.1. Chapter 2 - Restoring Source Rock Initial Quality and Quantity with Kinetic-Based Inversion - Applied to the Wolfcamp Play in the Permian Delaware Basin

Most of the published methods require measured HI data and assumptions on initial HI (HI_i) to calculate the transformation ratio (TR) and then restore the initial TOC (TOC_i) from measured TOC data. Estimating HI_i can be challenging for samples with high maturity, for cases in which immature samples and maceral data are lacking, or for heterogeneous source rocks. This study presents a workflow using a kinetic-based inversion procedure to restore the source-rock's HI_i and TOC_i from measured HI and TOC data. Unlike the other methods, the kinetic-based inversion method uses simulated TR derived from the basin model and independent of assumed HI_i to restore the HI_i and TOC_i from measured HI and TOC data. Thus, a pre-defined basin model with calibrated thermal model and assigned kinetic (organofacies) for the source rock intervals is required before the kinetic-based inversion procedure is performed. The workflow is applied to the Wolfcamp Play in the Permian Delaware Basin using a proprietary dataset covering the whole basin spatially.

1.4.2. Chapter 3 - Basin-Wide Lithofacies Identification Using Unsupervised Machine Learning: Applied to the Unconventional Wolfcamp Play, Permian Delaware Basin

A basin-wide lithofacies identification based on a large dataset can be challenging when core data are lacking. This study investigates a workflow using unsupervised machine learning (KMeans, Spectral, and Agglomerative Hierarchical clustering) procedure to identify basin-wide lithofacies from a proprietary well-log database with over 1500 wells. The proposed workflow is data-driven and straightforward, based solely on well-log data. The workflow is applied to the Wolfcamp Play in the Permian Delaware Basin.

1.4.3. Chapter 4 - Mass Balance Calculation for the Wolfcamp-Sourced Hydrocarbon in Permian Delaware Basin: Insight on Remaining Recoverable Resource and Expulsion-Migration Efficiency

Detailed quantification of basin-wide hydrocarbon (HC) masses from generation to production is necessary for an accurate petroleum system analysis. Such quantified HC masses must be balanced following the fundamental laws of mass conservation. Mass balance is particularly important for unconventional-conventional petroleum systems in which expulsion efficiency is a critical parameter defining HCs in place, both within the source rock interval (unconventional self-sourced) and outside (conventional or unconventional tight-rock). This study introduces an HC mass balance workflow aimed at “bookkeeping” the HC volumes from generation to production to assess the HC's mass balance, including the remaining recoverable and expulsion efficiency. The mass balance workflow is applied to the Wolfcamp Formation in Permian Delaware Basin.

1.5. References

Adams, J. E., 1965, Stratigraphic-tectonic development of Delaware Basin: AAPG Bulletin, v. 49, p. 2140-2148.

Amer, A., R. di Primio, and R. Ondrak, 2015, 4-D petroleum systems modelling of the Haynesville Shale Play—understanding gas in place: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 935-949.

Baldwin, J. L., R. M. Bateman, and C. L. Wheatley, 1990, Application of a neural network to the problem of mineral identification from well logs: The Log Analyst, v. 31.

Barker, C., and M. Pawlewicz, 1987, The effects of igneous intrusions and higher heat flow on the thermal maturity of Leonardian and younger rocks, western Delaware Basin, Texas: Glass Mountain: SEPM Guidebook, v. 87, p. 27.

Baur, F., 2019, Predicting petroleum gravity with basin modeling: New kinetic models: AAPG Bulletin, v. 103, p. 1811-1837.

Baur, F., R. di Primio, C. Lampe, and R. Littke, 2011, Mass balance calculations for different models of hydrocarbon migration in the Jeanne d'Arc basin, offshore Newfoundland: Journal of Petroleum Geology, v. 34, p. 181-198.

Becker, M., A. M. Romero, and A. Yu, 2019, 3-D basin modeling of the Permian Delaware Basin: tectonic evolution assessment to improve definition of boundary conditions: AAPG Annual Convention and Exhibition.

Bievenour, A., and S. Sonnenberg, 2019, Reservoir characterization of the Bone Spring and Wolfcamp Formations, Delaware Basin, Ward County, West Texas: SPE/AAPG/SEG Unconventional Resources Technology Conference.

Blakey, R., 2019, Greater Permian Basin Series, Colorado Plateau Geosystems. Inc. https://deeptimemaps.com/greater-permian-basin/ *(accessed on May 4th 2022)*

Bryant, I. D., C. B. Stabell, and M. Neumaier, 2013, Evaluation of unconventional resources using a petroleum system modeling approach: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 1394-1401.

Cander, H., 2012, Sweet spots in shale gas and liquids plays: prediction of fluid composition and reservoir pressure: AAPG Annual Convention and Exhibition, p. 12-15.

Chang, H.-C., D. C. Kopaska-Merkel, and H.-C. Chen, 2002, Identification of lithofacies using Kohonen self-organizing maps: Computers & Geosciences, v. 28, p. 223-229.

Chen, Z., and C. Jiang, 2015, A data driven model for studying kerogen kinetics with application examples from Canadian sedimentary basins: Marine and Petroleum Geology, v. 67, p. 795-803.

Chen, Z., C. Jiang, D. Lavoie, and J. Reyes, 2016, Model-assisted Rock-Eval data interpretation for source rock evaluation: Examples from producing and potential shale gas resource plays: International Journal of Coal Geology, v. 165, p. 290-302.

Coutinho, L. F. C., 2008, Analysis of the petroleum mass balance in a mature zone for exploration-Reconcavo Basin, Brazil, Université Pierre et Marie Curie-Paris VI.

Curtis, J. B., and J. E. Zumberge, 2017, An applied geochemical look at the Delaware Basin petroleum systems: AAPG Division of Professional Affairs, Delaware Basin Playmaker’s Forum.

Curtis, J. B., and J. E. Zumberge, 2018, Permian Basin petroleum systems–geochemical insight into hydrocarbon generation, migration and well performance: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 3980-3985.

di Primio, R., and B. Horsfield, 2006, From petroleum-type organofacies to hydrocarbon phase prediction: AAPG Bulletin, v. 90, p. 1031-1058.

Donovan, A. D., J. Evenick, L. Banfield, N. McInnis, and W. Hill, 2017, An organofacies-based mudstone classification for unconventional tight rock

& source rock plays: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 3683-3697.

Duran, E. R., R. di Primio, Z. Anka, D. Stoddart, and B. Horsfield, 2013, 3D-basin modelling of the Hammerfest Basin (southwestern Barents Sea): A quantitative assessment of petroleum generation, migration and leakage: Marine and Petroleum Geology, v. 45, p. 281-303.

Dutton, S. P., E. M. Kim, R. F. Broadhead, W. D. Raatz, C. L. Breton, S. C. Ruppel, and C. Kerans, 2005, Play analysis and leading-edge oil-reservoir development methods in the Permian basin: Increased recovery through advanced technologies: AAPG Bulletin, v. 89, p. 553-576.

Echegu, S., A. K. Bissada, and L. Elrod, 2021, Geochemical characterization and classification of crude oils of the Permian Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 105, p. 223-246.

Evenick, J. C., 2016, Evaluating source rock organofacies and paleodepositional environments using bulk rock compositional data and pristane/phytane ratios: Marine and Petroleum Geology, v. 78, p. 507-515.

Fairhurst, B., T. Ewing, and B. Lindsay, 2021, West Texas (Permian) Super Basin, United States: Tectonics, structural development, sedimentation, petroleum systems, and hydrocarbon reserves: AAPG Bulletin, v. 105, p. 1099-1147.

Gardner, M., and M. Sonnenfeld, 1996, Stratigraphic changes in facies architecture of the Permian Brushy Canyon Formation in Guadalupe

Mountains National Park, west Texas: Society of Economic Palaeontologists and Mineralogists Permian Basin Section, p. 17-40.

Gaswirth, S. B., K. L. French, J. K. Pitman, K. R. Marra, T. J. Mercier, H. M. Leathers-Miller, C. J. Schenk, M. E. Tennyson, C. A. Woodall, and M. E. Brownfield, 2018, Assessment of undiscovered continuous oil and gas resources in the Wolfcamp Shale and Bone Spring Formation of the Delaware Basin, Permian Basin Province, New Mexico and Texas, 2018, U.S. Geological Survey.

Hantschel, T., and A. I. Kauerauf, 2009, Migration and accumulation, Fundamentals of Basin and Petroleum Systems Modeling, Springer, p. 247-340.

Hills, J. M., 1984, Sedimentation, tectonism, and hydrocarbon generation in Delaware Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 68, p. 250-267.

Hills, J. M., 1985, Structural evolution of the Permian basin of west Texas and New Mexico: Structure tectonics of Trans-Pecos Texas: West Texas Geological Society, Publication, v. 85, p. 81.

Jarvie, D. M., 2017, Geochemical assessment and characterization of petroleum source rocks and oils, and petroleum systems, Permian Basin, US: The Houston Geological Society Bulletin, v. 60.

Jarvie, D. M., R. J. Hill, T. E. Ruble, and R. M. Pollastro, 2007, Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central

Texas as one model for thermogenic shale-gas assessment: AAPG Bulletin, v. 91, p. 475-499.

Justman, H., and R. Broadhead, 2010a, Petroleum source rock data for the Bone Spring Formation, Delaware Basin, southeastern New Mexico, Open File Report 529, New Mexico Bureau of Geology and Mineral Resources A Division of New Mexico Institute of Mining and Technology.

Justman, H., and R. Broadhead, 2010b, Petroleum source rock data for the Brushy Canyon Formation, Delaware Basin, southeastern New Mexico, Open File Report 528, New Mexico Bureau of Geology and Mineral Resources A Division of New Mexico Institute of Mining and Technology.

Justwan, H., and B. Dahl, 2005, Quantitative hydrocarbon potential mapping and organofacies study in the Greater Balder Area, Norwegian North Sea: Geological Society, London, Petroleum Geology Conference series, p. 1317-1329.

Katz, B. J., and I. Arango, 2018, Organic porosity: a geochemist's view of the current state of understanding: Organic Geochemistry, v. 123, p. 1-16.

Katz, B. J., I. Arango, and F. Frasse, 2017, Expulsion and migration associated with unconventional petroleum systems: The Houston Geological Society Bulletin, v. Vol. 59, No. 5.

Kinley, T. J., L. W. Cook, J. A. Breyer, D. M. Jarvie, and A. B. Busbey, 2008, Hydrocarbon potential of the Barnett Shale (Mississippian), Delaware

Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 92, p. 967-991.

Kuchinskiy, V., K. Gentry, and R. Hill, 2012, Source rock evaluation technique: A probabilistic approach for determining hydrocarbon generation potential and in-place volume for shale plays: AAPG Annual Convention and Exhibition.

Lew, C., S. Baharuddin, and J. D. Pigott, 2013, Deepwater basin model for the Permian: Delaware Basin example: International Petroleum Technology Conference, p. cp-350-00399.

Lewan, M., M. Henry, D. Higley, and J. K. Pitman, 2002, Material-balance assessment of the New Albany-Chesterian petroleum system of the Illinois basin: AAPG Bulletin, v. 86, p. 745-777.

Modica, C. J., and S. G. Lapierre, 2012, Estimation of kerogen porosity in source rocks as a function of thermal transformation: Example from the Mowry Shale in the Powder River Basin of Wyoming: AAPG bulletin, v. 96, p. 87-108.

Muscio, G. P. A., M. A. Everett, L. P. Marpaung, K. T. McDannell, A. A. Miceli Romero, and I. Yurchenko, 2016, Insights into migration efficiency based on stochastic mass balance calculations and full-physics forward basin modeling, AAPG Hedberg Conference, The Future of Basin and Petroleum System Modelling, Santa Barbara, California, AAPG Datapages/ Search and Discovery Article.

Pepper, A., A. Doebbert, J.-M. Laigle, and L. Laigle, 2020, Greater Permian Basin petroleum systems–what are we learning in the transition from conventional to unconventional?: The Houston Geological Society Bulletin.

Pepper, A. S., 1991, Estimating the petroleum expulsion behaviour of source rocks: a novel quantitative approach: Geological Society, London, Special Publications, v. 59, p. 9-31.

Pepper, A. S., and P. J. Corvi, 1995, Simple kinetic models of petroleum formation. Part I: oil and gas generation from kerogen: Marine and Petroleum Geology, v. 12, p. 291-319.

Peters, K., L. Magoon, K. Bird, Z. Valin, and M. Keller, 2006, North Slope, Alaska: Source rock distribution, richness, thermal maturity, and petroleum charge: AAPG Bulletin, v. 90, p. 261-292.

Raeesi, M., A. Moradzadeh, F. D. Ardejani, and M. Rahimi, 2012, Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks: Journal of Petroleum Science and Engineering, v. 82, p. 151-165.

Romero-Sarmiento, M.-F., M. Ducros, B. Carpentier, F. Lorant, M.-C. Cacas, S. Pegaz-Fiornet, S. Wolf, S. Rohais, and I. Moretti, 2013, Quantitative evaluation of TOC, organic porosity and gas retention distribution in a gas shale play using petroleum system modeling: Application to the

Mississippian Barnett Shale: Marine and Petroleum Geology, v. 45, p. 315-330.

Rushing, J., A. Chaouche, and K. Newsham, 2004, A mass balance approach for assessing basin-centred gas prospects: integrating reservoir engineering, geochemistry and petrophysics: Geological Society, London, Special Publications, v. 237, p. 373-390.

Sandvik, E., W. Young, and D. Curry, 1992, Expulsion from hydrocarbon sources: the role of organic absorption: Organic Geochemistry, v. 19, p. 77-87.

Schwartz, K. M., G. Hennenfent, M. Hegmann, M. Hoffnagle, D. Bain, and A. McCallister, 2015, Pay distributions and basin architecture of the Wolfcamp shale in the Delaware Basin: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 1729-1733.

Wang, G., T. R. Carr, Y. Ju, and C. Li, 2014, Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin: Computers & Geosciences, v. 64, p. 52-60.

Yu, A. Z., M. Becker, A. S. Baskoro, and M. Bhatia, 2020, Fluid property variations and its relationship with the geo-history of the Permian Delaware Basin: AAPG Annual Conference and Exhibition.

2. RESTORING SOURCE ROCK INITIAL QUALITY AND QUANTITY WITH KINETIC-BASED INVERSION – APPLIED TO THE WOLFCAMP PLAY IN THE PERMIAN DELAWARE BASIN

2.1. Introduction

Accurate estimation of the hydrocarbon mass generated from the source rock is critical to improving the mass balance assessment between the hydrocarbons that migrated from- and hydrocarbons that remained within- the source rock intervals. Before the quantity of generated hydrocarbon mass can be calculated, it is necessary to restore the present-day source-rock conditions – hydrogen index (HI) and total organic carbon (TOC), to their initial state (Jarvie et al., 2007; Kuchinskiy et al., 2012; Lewan et al., 2002; Modica and Lapierre, 2012; Rushing et al., 2004). Additionally, restoring the present-day source rock to its initial state is essential when assessing unconventional resources as the transformation of organic matter from its initial state to present-day conditions will generate hydrocarbons and form organic porosity within the kerogen, which will also determine the storage capacity for retained hydrocarbon within the source rock (Jarvie et al., 2007; Katz and Arango, 2018; Modica and Lapierre, 2012; Romero-Sarmiento et al., 2013).

Various methods have been proposed to restore the measured source rock to its initial condition aiming to calculate the volume of hydrocarbon generated (e.g., Chen and Jiang, 2015; Chen et al., 2016; Jarvie et al., 2007;

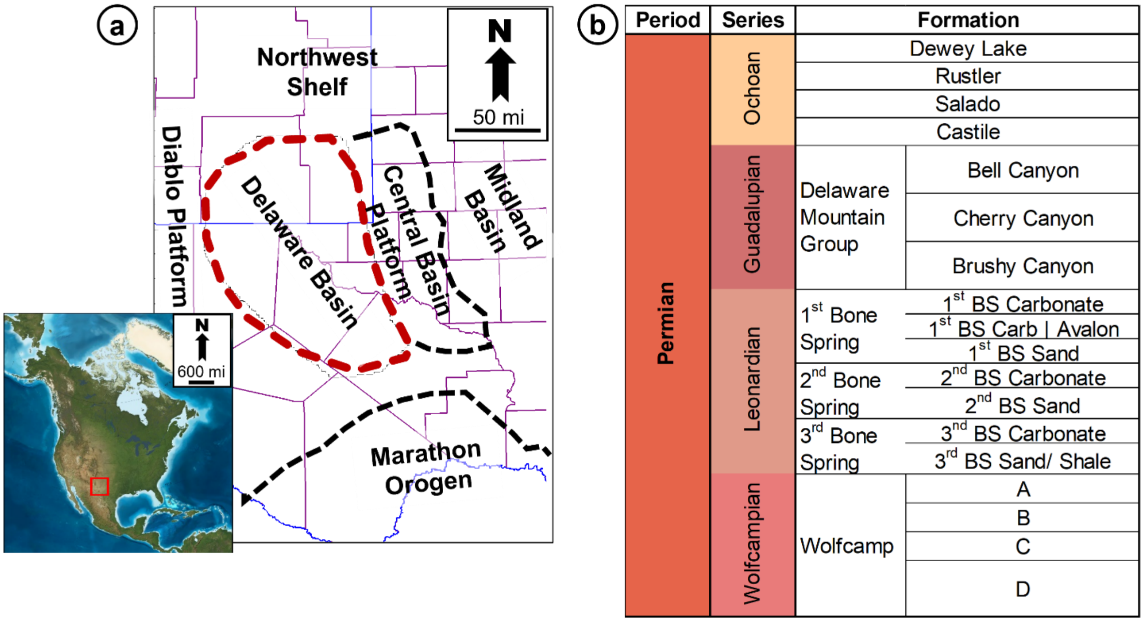
Justwan and Dahl, 2005; Peters et al., 2006; among many others). Most of these methods require measured HI data and assumptions of initial HI (HI_i) to calculate the transformation ratio (TR) and then restore the initial TOC (TOC_i) from measured TOC data (e.g., Jarvie et al., 2007; Justwan and Dahl, 2005; Peters et al., 2006). Transformation ratio is a ratio of the petroleum formed by kerogen to the total amount of petroleum that kerogen can generate (Tissot and Welte, 1984a). For the reconstruction method based on Justwan and Dahl (2005), TR is calculated based on the transformation of assumed HI_i to the present-day (measured) HI data. In contrast, Jarvie et al. (2007) and Peters et al. (2006) calculate TR based on the transformation of both assumed HI_i to measured HI and assumed initial production index (PI_i) to measured PI. Both Justwan and Dahl (2005) and Peters et al. (2006) used measured HI from the immature sample to represent the HI_i, whereas Jarvie et al. (2007) calculated the HI_i based on maceral percentages and assigned average HI_i values for each kerogen-type. Nevertheless, estimation of HI_i is challenging for samples with high maturity, and for cases in which immature samples are unavailable or maceral data is lacking. Another challenge is to provide a representative HI_i value for heterogeneous source rocks.

Another source rock reconstruction method was proposed by Chen and Jiang (2015; Chen et al., 2016). Unlike the other reconstruction methods, Chen and Jiang (2015) proposed restoring HI_i using a statistical approach to fit measured HI and Tmax (the measured temperature at the maximum of the S2 peak of pyrolysis or the remaining hydrocarbon generation potential) data with hydrocarbon generation kinetic. Using their approach, HI_i can be calculated along with its uncertainty range. The restored HI_i and the measured HI data are then used to calculate TR (using the formula from Justwan and Dahl, 2005) and restore TOC_i from measured TOC data (Chen et al., 2016). However, one of the requirements to provide a meaningful empirical model of HI_i transformation vs. Tmax is that the dataset covers the entire Tmax range of hydrocarbon generation (e.g., from immature to high mature) (Chen and Jiang, 2015). In addition, using measured Tmax data brings additional uncertainty as hydrocarbon retardation by specific mineralogy or kerogen and/or organic matter abundance may suppress the measured data (Katz and Lin, 2021). Using published methods to restore HI_i and TOC_i typically add considerable uncertainty to the source rock model (e.g., calculated present-day HI and TOC in basin model are not calibrated to measured HI and TOC data) and ultimately to the calculated volume of generated hydrocarbon.

This paper proposes and evaluates a workflow using a kinetic-based inversion approach to restore the source-rock's HI_i and TOC_i from the measured HI and TOC data. Unlike the previous methods, the kinetic-based inversion method utilizes TR derived from the basin model to restore the HI_i and TOC_i from measured HI and TOC data. In the basin model, calculated TR is not dependent on the assumed HI_i and uses an arbitrary HI_i instead (Hantschel and Kauerauf, 2009). The calculated TR is based on Arrhenius-type kinetic reactions and thus is dependent on thermal maturity and organofacies-dependent kinetic (Hantschel and Kauerauf, 2009), and the model is calibrated to thermal indicators. Sensitivity analysis is performed to address the uncertainty of the restored HI_i and TOC_i based on a kinetic-based inversion procedure. We also illustrate the impact of varying restored source rock HI_i and TOC_i toward the estimated hydrocarbon resources.

2.2. Geological Overview

The case study for this source rock reconstruction is the Wolfcamp Play in the Permian Delaware Basin of West Texas and Southeast New Mexico, USA (**Figure 2.1a**). The Delaware Basin is one of the best examples of a prolific petroleum basin with hybrid petroleum systems. The Wolfcamp Play in the Permian Delaware Basin plays a vital role as one of the primary source rocks, especially for the Middle-Upper Permian strata in the Delaware Basin. It is also one of the main targets for unconventional production along with the Leonardian Bone Spring interval in the basin (**Figure 2.1b**). Performing hydrocarbon mass-balance assessment is necessary to improve the resources assessment, including the distribution, for both conventional and unconventional accumulation sourced from the Wolfcamp interval in Delaware Basin.



Period	Series	Formation	
Permian	Ochoan	Dewey Lake	
		Rustler	
		Salado	
		Castile	
	Guadalupian	Delaware Mountain Group	Bell Canyon
			Cherry Canyon
			Brushy Canyon
	Leonardian	1st Bone Spring	1st BS Carbonate
			1st BS Carb | Avalon
			1st BS Sand
		2nd Bone Spring	2nd BS Carbonate
			2nd BS Sand
		3rd Bone Spring	3rd BS Carbonate
			3rd BS Sand/ Shale
	Wolfcampian	Wolfcamp	A
			B
			C
			D

Figure 2.1. (a) Major subdivision and boundaries of the Permian Basin around the Delaware Basin (after Dutton et al. (2015) and Silver and Todd (1969)). (b) The stratigraphic chart of Permian strata, which includes Wolfcamp interval subdivision for the Delaware Basin.

The Permian Delaware Basin is the western major structural subdivision of the Permian Basin's more massive structure in West Texas and Southeast New Mexico (Hills, 1984) (**Figure 2.1a**). The Delaware Basin is bounded to the west by the Diablo Platform, to the north by the Northern American craton and Northwestern Shelf areas, and to the south by the Marathon orogenic belt. To the east, the Delaware Basin is separated from the Midland Basin by the N-S trending Central Basin Platform (Adams, 1965; Hills, 1984; Silver and Todd, 1969).

The Wolfcamp is part of the thick Permian strata (**Figure 2.1b**) in the Delaware Basin and consists of four main intervals: Wolfcamp A, -B, -C, and -D (Bievenour and Sonnenberg, 2019; Dutton et al., 2005). Wolfcamp intervals were deposited during the early Permian with a depositional environment of generally a deep, oceanic-connected interior continental basin, surrounded by carbonate shelf (the Northwestern Carbonate Platform, eastern Central Basin Carbonate Platform, western Diablo Platform) and the southern siliciclastic shelf in front of the Marathon orogenic belt (Blakey, 2019). This depositional setting resulted in litho- and organo-facies heterogeneity within the Wolfcamp intervals. Hydrocarbon generation from the Wolfcamp source rock (SR) started during the middle Permian (Adams, 1965; Hills, 1984; Kinley et al., 2008), influenced by complex tectonic history and thermal processes, including high thermal conductive evaporites (Gardner and Sonnenfeld, 1996; Hills, 1985; Kinley et al., 2008), intrusive and volcanic activities (Barker and Pawlewicz, 1987), and burial history. Recent studies (Becker et al., 2019; Yu et al., 2020) indicate that burial history is the most dominant factor for source maturity, with the maximum maturity reached during the mid-Permian at maximum burial, then continuing to generate hydrocarbons until the present-day source-rock condition (Hills, 1984; Lew et al., 2013; Schwartz et al., 2015).

2.3. Methods and Procedures

The source rock reconstruction workflow presented in this paper involves two main steps: (1) defining the source rock organofacies; (2) performing a kinetic-based inversion approach to restore the HI_i and TOC_i using measured HI and TOC data as the control for the determined source organofacies.

2.3.1. Determining Organofacies and Kinetic Models

Pepper and Corvi's (1995) source rock organofacies (OF) classification was applied. This divides organofacies into five classes with defined kinetic parameters.

2.3.1.1. Geochemical Data Interpretation for Organofacies

Three methods are used to determine source rock organofacies by interpreting maceral composition, oil biomarkers, and Rock-Eval pyrolysis based on the Wolfcamp geochemical datasets provided in the Delaware Basin.

2.3.1.1.1. Maceral Composition

Maceral composition data are derived from both core and cuttings of five wells covering Wolfcamp A-B-C-D intervals in the eastern part of the basin (**Figure 2.2**). The maceral composition data are plotted on a ternary diagram of amorphous organic matter (AOM) + exinite vs. vitrinite vs. inertinite based on Tissot and Welte (1984b) to define the organofacies type of the source rock. AOM and exinite macerals are assumed to primarily represent marine organic matter.

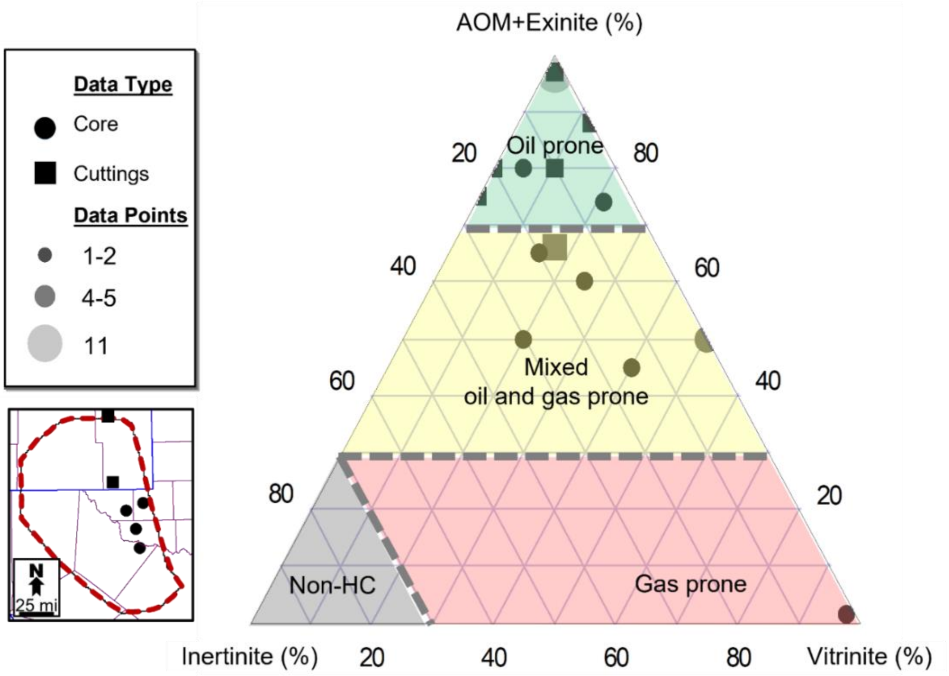


Figure 2.2. Kerogen type interpretation for Wolfcamp interval based on maceral composition data: from marine to mixed marine-terrigenous dominated typed organofacies. Insert map localize the sampled wells. AOM refers to Amorphpus Organic Matter.

Based on **Figure 2.2**, the Wolfcamp SR is interpreted as predominantly oil and mixed oil and gas prone typed organofacies. As no alginite maceral is recorded in the dataset, oil-prone is interpreted as marine typed (carbonate and/or clastic) with no lacustrine typed organofacies. Mixed oil and gas prone is interpreted as mixed marine-terrigenous typed organofacies. One data point also indicates an associated contribution of gas-prone or terrigenous typed organofacies in the Wolfcamp dataset.

2.3.1.1.2. Oil Biomarker Analysis

Biomarker data are from Wolfcamp-sourced (upper and lower intervals) oil wells distributed in the eastern part of the basin (**Figure 2.3**). This study uses oil biomarkers, including $C_{35}S/C_{34}S$ hopanes, C_{29}/C_{30} hopane (H), pristane/nC_{17} (C_{17} normal alkane), and phytane/nC_{18} (C_{18} normal alkane), to determine the organofacies based on general facies (i.e., clastic vs. carbonate) and organic matter type (marine vs. terrigenous).

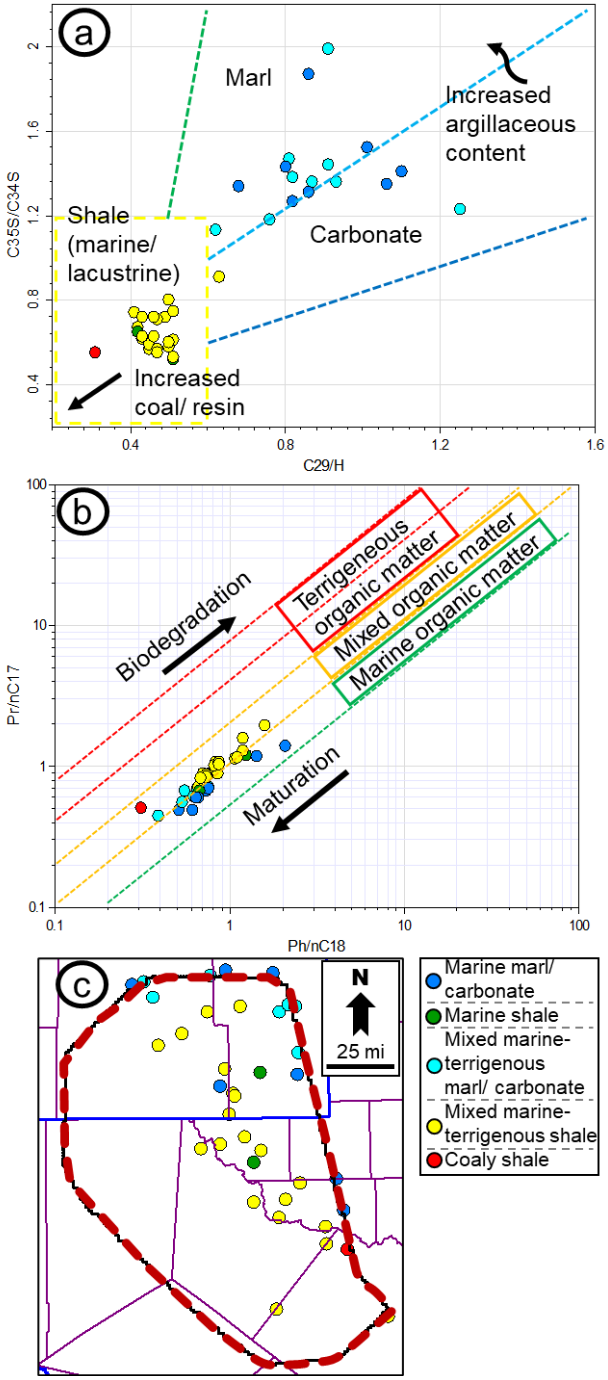


Figure 2.3. Interpretation of Wolfcamp SR sourced oil biomarkers data: source rock's: (a) facies; (b) organic matter type; and (c) inferred source rock groups and distribution in the basin; indicating marine shale and mixed marine-terrigenous shale sourced in the basin depocenter and more carbonate sourced in the Northern and Eastern basin margin.

The homohopanes (C_{31}-C_{35}) are sourced from prokaryotic microorganisms and indicate the redox potential of source rock deposition (Peters and Moldowan, 1991; Peters et al., 2005). High homohopanes values are associated with a highly reducing marine depositional environment. C_{29}-C_{30}/H (hopane) is commonly a major peak on *m/ z* 191 mass chromatograms of saturate fractions of oils and bitumen (Peters et al., 2005). A high C_{29}/H value (>1.0) is associated with anoxic carbonate or marl source rocks. Pristane (C_{19}) and phytane (C_{20}) are primarily sourced from the phytyl side chain of chlorophyll a in phototrophic organisms and bacteriochlorophyll a and b in purple sulfur bacteria, which are a good indicator for redox conditions in the source rock (Peters et al., 2005). Reducing or anoxic conditions promote a reaction of the phytyl side chain to generate phytane eventually. In contrast, the oxic condition promotes the phytyl side chain's conversion to pristane.

A plot of $C_{35}S/C_{34}S$ hopanes vs. C_{29}/C_{30} hopane (H) adapted from Peters et al. (2005) is used to interpret the source rock's facies, including clastic shale vs. marine carbonate and marl (**Figure 2.3a**), while a plot of pristane/nC_{17} (C_{17} normal alkane) and phytane/nC_{18} (C_{18} normal alkane) (**Figure 2.3b**) adapted from Shanmugam (1985) is used to define organic matter type of the source rock and also to compare the interpretation results from the plot of $C_{35}S/C_{34}S$ hopanes vs. C_{29}/H.

Based on **Figure 2.3**, Wolfcamp SR is predominantly marine shale (OF “B”) and mixed marine-terrigenous shale (OF “D/E”) typed organofacies in the basin depocenter with marine- and mixed marine-terrigenous- carbonate/ marl typed organofacies (OF “A”) mostly restricted to the basin’s northern and eastern margin portions (**Figure 2.3c**).

2.3.1.1.3. Rock-Eval Pyrolysis Data

Rock-Eval pyrolysis data are abundant and cover most of the Wolfcamp intervals in the wide spatial distribution in the basin (**Figure 2.4**). The dataset was screened for possible contamination or oxidation (Peters, 1986) and sampled over 75-100 ft. depth intervals before being used for interpretation.

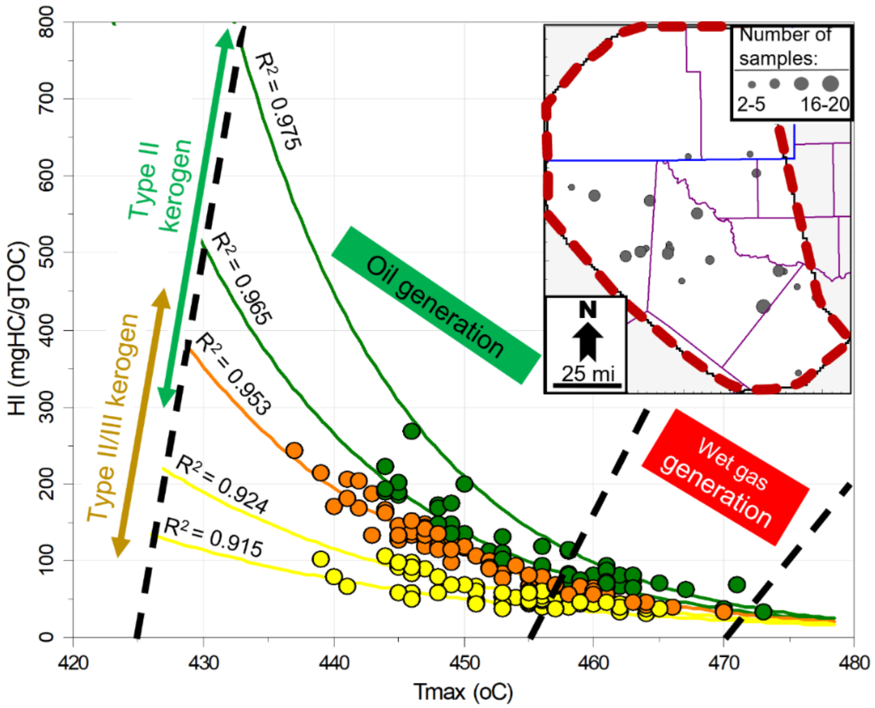


Figure 2.4. A plot of measured Tmax vs. HI for 185 data points from the Wolfcamp Play showing oil to wet gas generation maturity window and predominantly marine and mixed marine-terrigenous typed kerogen.

A plot of HI vs. Tmax (modified after Peters et al. 2006; Tobey and Campbell, 2016) is used to define the kerogen type of the source rock from screened Rock-Eval pyrolysis data based on the roughly projected initial HI following the exponential maturation pathway (Devine, 2014; Jarvie and Lundell, 2001; Montgomery et al., 2005) of the mature data points (**Figure 2.4**). Using this exponential relationship between HI and Tmax, data points are classified into five exponential lines (R^2>0.91), assuming the intercept or flattening low-HI point (<10 mgHC/gTOC) is at Tmax >490^{o}C.

Based on **Figure 2.4**, Wolfcamp SR data are classified into marine (green and orange) and mixed marine-terrigenous (orange and yellow). Further classification for the orange data points is based on the paleo-structure of Wolfcamp intervals (Becker et al., 2019; Yu et al., 2020) and its proximity from the basin margin (distal vs. proximal). For example, the orange data points for Wolfcamp A-B are predominantly marine typed organofacies (distal wells) with few mixed marine-terrigenous typed organofacies for the proximal wells. The orange data points for Wolfcamp C-D are predominantly mixed marine-terrigenous typed organofacies (proximal wells) with few marine typed organofacies for the wells in the basin's distal and deeper portion (main depocenter). This interpretation is based on the assumption that sediments from the proximal facies are rich in terrigenous material due to higher sedimentation rate from the sediment source, diluting the marine organic matter. The mixed marine-terrigenous facies is characterized by clay-rich (argillaceous) mudstone or siltstone source rock with lower HI and TOC (Baskoro et al., 2022 (in prep); Baskoro, 2017; Donovan et al., 2007).

2.3.1.2. Kinetic Model for Determined Organofacies

To conclude, by integrating maceral (**Figure 2.2**), oil biomarker (**Figure 2.3**), and Rock-Eval (**Figure 2.4**) data, the Wolfcamp SR within the Delaware Basin is shown to be predominantly marine clastic shale (OF "B") and mixed marine-terrigenous (OF "D/E") typed organofacies. Marine- and mixed marine-terrigenous carbonate/ marl source rocks (OF "A") may also be present but are mostly restricted to the basin's northern and eastern margin portions (**Figure 2.3**).

The general source rock organofacies type interpretation from this study is supported by others, including Echegu et al. (2021) using Rock-Eval and oil biomarker data and Bievenour and Sonnenberg (2019) using Rock-Eval data. Bievenour and Sonnenberg (2019) also used backscatter electrons and secondary electrons and observed both marine and woody-concave non-porous terrigenous organic matter in siliceous mudstone and siltstone lithofacies groups, indicating a mixture of marine and terrigenous types of kerogen. Supporting the mixture of marine and terrigenous types of kerogen, Kvale et al. (2020) also found prints of a land plant (peltasperm *Germaropteris martinsii*) in the Wolfcamp SR cores they studied.

Therefore, based on the determined organofacies, kinetic models for OF "B" and OF "D/E" from Pepper and Corvi (1995) are adopted for Wolfcamp SR with identified marine or mixed marine-terrigenous typed organofacies.

2.3.2. Kinetic-Based Inversion

The kinetic-based inversion procedure is based on the calibrated thermal history combined with the simulated TR for the source rocks from the basin model and the measured HI and TOC data. The procedure restores HI_i and TOC_i by "inverting" the measured HI and TOC based on their assumed maturation pathway – simulated TR, to their initial (restored) condition for the given source organofacies. Sensitivity analysis is also performed to evaluate the uncertainties involved in the reconstruction of HI_i and TOC_i using this procedure. The following are the procedure steps:

2.3.2.1. Calculation and Validation of Transformation Ratio

In the basin model, the transformation rate of the initial kerogen (HI_i and TOC_i) to present-day kerogen is primarily controlled by thermal history and organofacies type (Hantschel and Kauerauf, 2009). Thus, it is important to have a pre-defined basin model calibrated to thermal maturity indicators, which can be used to calculate the transformation ratio (TR) for the kinetic-based inversion procedure. In this study, multiple 1D basin models with their associated boundary conditions and thermal maturity scenario are extracted from the 3D basin model of the Permian Delaware Basin, which was calibrated against all thermal indicators from wells such as vitrinite reflectance, bottom hole temperature, and fluid property observed (Becker et al., 2019; Yu et al., 2020).

A kinetic model (e.g., Pepper and Corvi, 1995) must be assigned to the intervals or sublayers where TOC and HI calibration data are available or want to be restored (e.g., **Figure 2.4** for this study) to calculate TR. **Figure 2.5** demonstrates the calculated TR from twenty-one (21) 1D basin models, consisting of a total of 185 sublayers of Wolfcamp SR with measured HI and TOC data. **Figure 2.5a** shows that the calculated TR from the basin models generally follows the exponential maturation pathway of measured Tmax data: the higher the measured Tmax, the higher the calculated TR. Higher TR (>70%) values are predominantly distributed in the western portion of the basin (see **Figure 2.4** for the distribution of the wells in the basin) for both OF “B” and “D/E” samples (**Figure 2.5b** and **c**) following the generally high thermal maturity trend (e.g., vitrinite reflectance: Becker et al., 2019; Yu et al., 2020) in the basin. In addition, high TR for OF “D/E” (>80%) is observed only from the Wolfcamp D samples in the western portion of the basin.

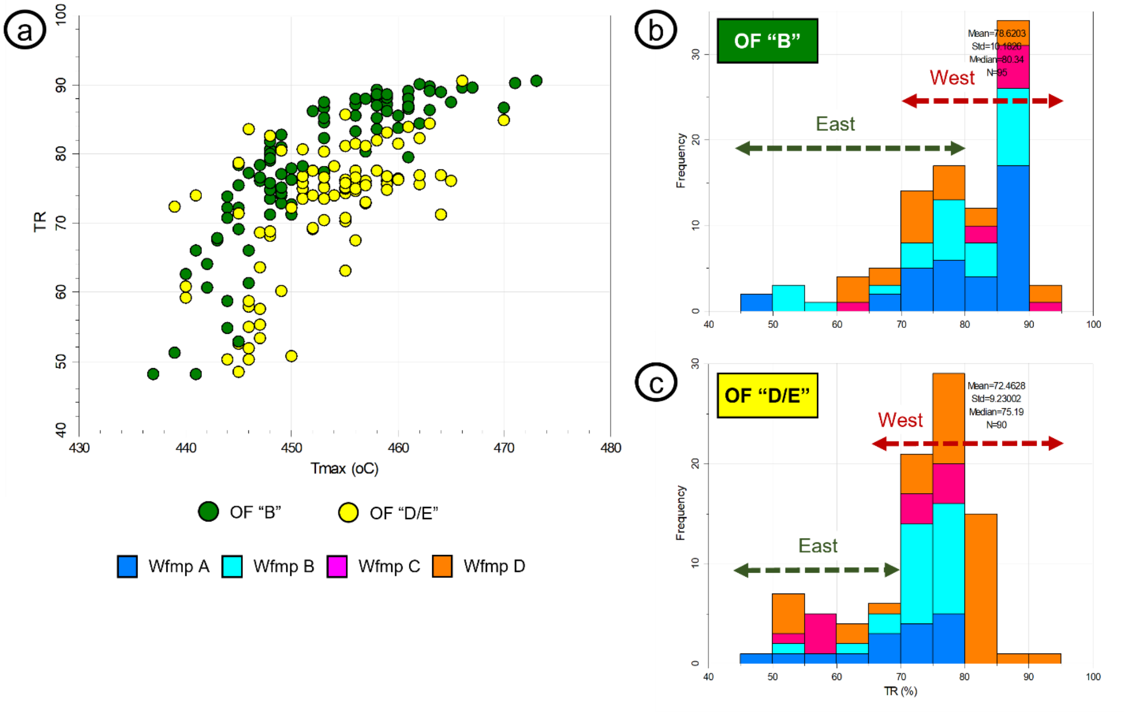


Figure 2.5. Simulated TR distribution of 185 source units (model sublayers) considered from the Wolfcamp Play: (a) TR plot against the measured Tmax; a histogram with Wolfcamp interval for source units with (b) OF "B"; and (c) OF "D/E". East: eastern part; West: western part of Delaware Basin.

2.3.2.2. Calculation of Restored HI_i and TOC_i

By definition, TR (transformation ratio) is:

$$\mathrm{TR} = \frac{\mathrm{HI}_i - \mathrm{HI}}{\mathrm{HI}_i} \qquad (1)$$

In the kinetic-based inversion procedure, HI_i is calculated by rearranging **equation (1)** to solve HI_i using **equation (2)**. TOC_i is calculated using **equation (3)**. These equations are provided in the basin model (T. Hantschel, T. Fuchs, and M. de Lind, 2021, personal communication):

$$\mathrm{HI}_i = \frac{\mathrm{HI}}{(1-\mathrm{TR})} \qquad (2)$$

$$\mathrm{TOC}_i = \frac{\mathrm{TOC}}{(1-0.001*\mathrm{W}*(\mathrm{HI_i}-\mathrm{HI}))} \qquad (3)$$

HI and TOC are measured hydrogen index (mgHC/gTOC) and total organic carbon (wt%) values, while TR is the transformation ratio (fraction) for the source rock at a given layer, derived from the basin model. A unit conversion factor of the HI, from mg/gTOC to g/gTOC, is denoted by 0.001. W is the weight factor, which represents the mass fraction of carbon in the primary generated hydrocarbon and varies by the type of organofacies (e.g., 0.75 for OF "B"; and 0.68 for OF "D/E").

2.4. Results and Interpretation of Restored HI_i and TOC_i for Wolfcamp Play

HI_i and TOC_i are calculated on 185 HI and TOC measurements. These HI and TOC data are from the screened Rock-Eval pyrolysis dataset and correspond to sublayers of various Wolfcamp SR intervals (A, B, C, and D) from multiple wells (1D basin model) across the basin. **Figure 2.6** shows the inversion procedure applied to a representative well for Wolfcamp A and B intervals. Since the inversion procedure restores HI_i and TOC_i using the measured HI and TOC data and simulated TR, the present-day simulated HI and TOC from the basin model will match measured HI and TOC data when using these restored HI_i and TOC_i values. Sensitivity analysis is also performed to illustrate the uncertainties in the inversion procedure and their impact on restored HI_i and TOC_i.

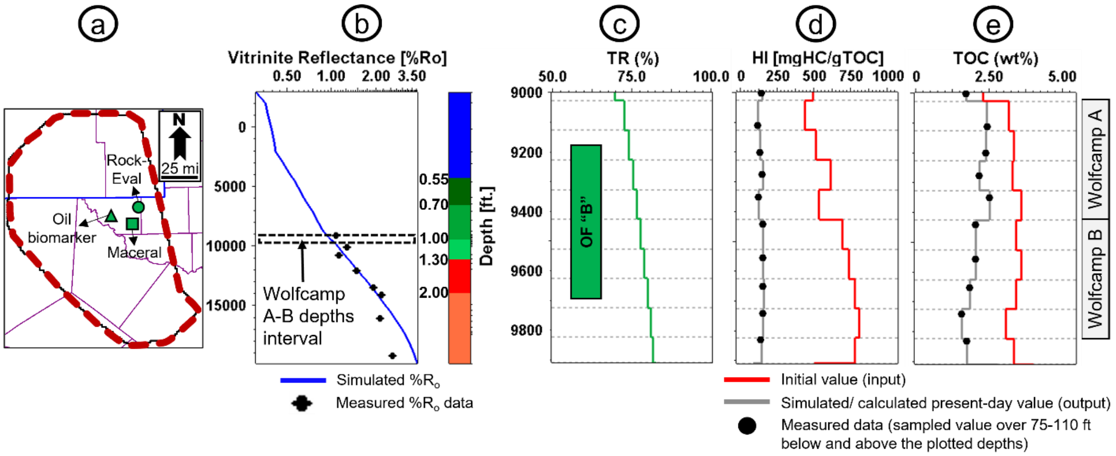


Figure 2.6. Example of the kinetic-based inversion procedure at a well. (a) Well location; (b) thermal maturity calibration; (c) determined organofacies (OF) and its simulated TR; (d) restored (initial value), measured, and calculated HI; (e) restored (initial value), measured, and calculated TOC.

2.4.1. Restored Source Rock Distribution – Wolfcamp Play

Figure 2.7 demonstrates measured data vs. restored HI_i and TOC_i for the Wolfcamp Play dataset. As expected from a source rock sample with high thermal maturity (**Figure 2.4** and **Figure** 2.5), there is a significant increase between measured and restored initial HI and TOC values (**Figure 2.7**). The increase from HI to HI_i is significantly higher for samples with identified OF “B” (an increase from 191 up to 749 mgHC/gTOC) than for samples with OF “D/E” (an increase from 60 up to 370 mgHC/gTOC). Similarly, the increase from TOC to TOC_i is significantly higher for samples with identified OF “B”, from 0.26 to 2.83 wt%, than for samples with OF “D/E”, from 0.06 to 0.74 wt%. The considerable variation of both restored HI_i and TOC_i indicate source rock heterogeneity associated with the preservation and mixing of marine and terrigenous organic matter during Wolfcamp deposition.

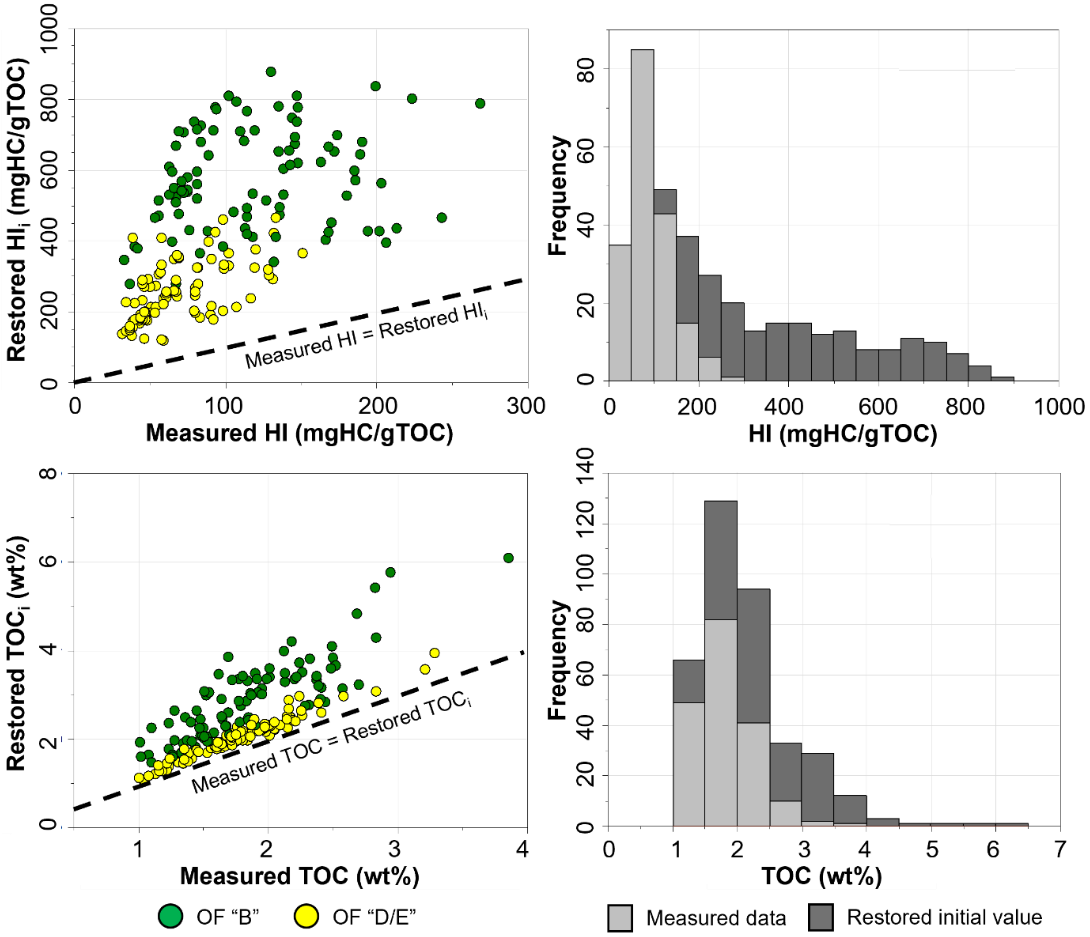


Figure 2.7. Plots (left) and histograms (right) of measured data vs. restored initial value for both HI (upper) and TOC (lower) for the Wolfcamp Play indicate the significance of restored initial value from measured data.

Figure 2.8 demonstrates the distribution of restored source rock (HI$_i$ and TOC$_i$) within Wolfcamp Play intervals. In general, significant variations are observed in Wolfcamp A, B, and C intervals, while lower HI$_i$ and TOC$_i$ dominate Wolfcamp D. Based on **Table 2.1**, each Wolfcamp interval has a similar range of simulated TR of 60-90%, but the HI_i varies. Better quality and quantity source rocks primarily are Wolfcamp A and B, indicated by the P90/50/10 values of HI_i (**Table 2.1**) and TOC_i (**Table 2.2**), showing P50 HI_i of 530 and 430 mgHC/gTOC in Wolfcamp A and B, respectively, while P50 HI_i is about 300 mgHC/gTOC for Wolfcamp C and D intervals. These contrast with lower P90/50/10 of HI_i and TOC_i values for Wolfcamp C and D, indicating lower quality and quantity of the source rock.

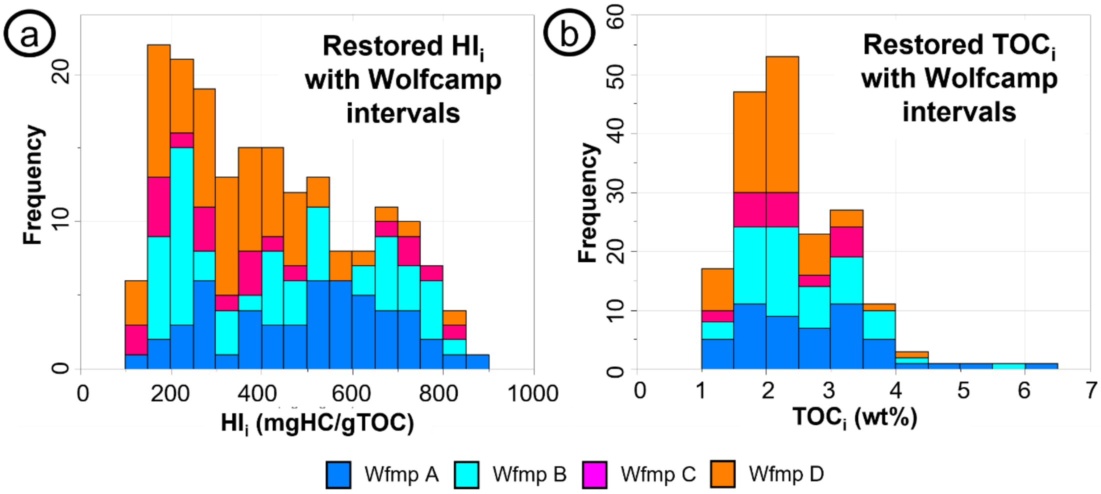


Figure 2.8. Wolfcamp Play restored source rock distribution by Wolfcamp intervals based on kinetic-based inversion procedure (a) HI_i, and (b) TOC_i.

Table 2.1. Wolfcamp Play simulated TR (%) and measured and restored HI($_i$) (mgHC/gTOC) distribution (P90/P50/P10) based on the kinetic-based inversion procedure.

		P90	P50	P10
Wolfcamp A N = 52	TR (%)	63	77	87
	Measured HI	65	92	186
	Restored HI$_i$	246	530	713
Wolfcamp B N = 53	TR (%)	61	76	87
	Measured HI	47	81	168
	Restored HI$_i$	197	429	736
Wolfcamp C N = 21	TR (%)	58	77	90
	Measured HI	38	72	147
	Restored HI$_i$	179	303	715
Wolfcamp D N = 59	TR (%)	61	77	88
	Measured HI	37	64	138
	Restored HI$_i$	159	325	550

Table 2.2. Wolfcamp Play measured and restored TOC($_i$) (wt%) distribution (P90/P50/P10) based on the kinetic-based inversion procedure.

		P90	P50	P10
Wolfcamp A N = 52	Measured TOC	1.23	1.78	2.68
	Restored TOC$_i$	1.50	2.64	3.86
Wolfcamp B N = 53	Measured TOC	1.31	1.80	2.40
	Restored TOC$_i$	1.64	2.27	3.59
Wolfcamp C N = 21	Measured TOC	1.18	1.70	2.26
	Restored TOC$_i$	1.56	2.21	3.33
Wolfcamp D N = 59	Measured TOC	1.21	1.71	2.21
	Restored TOC$_i$	1.47	2.09	2.97

The source rock potential based on TOC$_i$ vs. S2$_i$ (initial hydrocarbon generation potential) is also illustrated in **Figure 2.9**. Wolfcamp A and B show to be better source rock than Wolfcamp C and D. Overall, the Wolfcamp SR are within good to excellent source rock potential range. **Figure 2.9** also demonstrates the positive relationship between TOC$_i$ and S2$_i$.

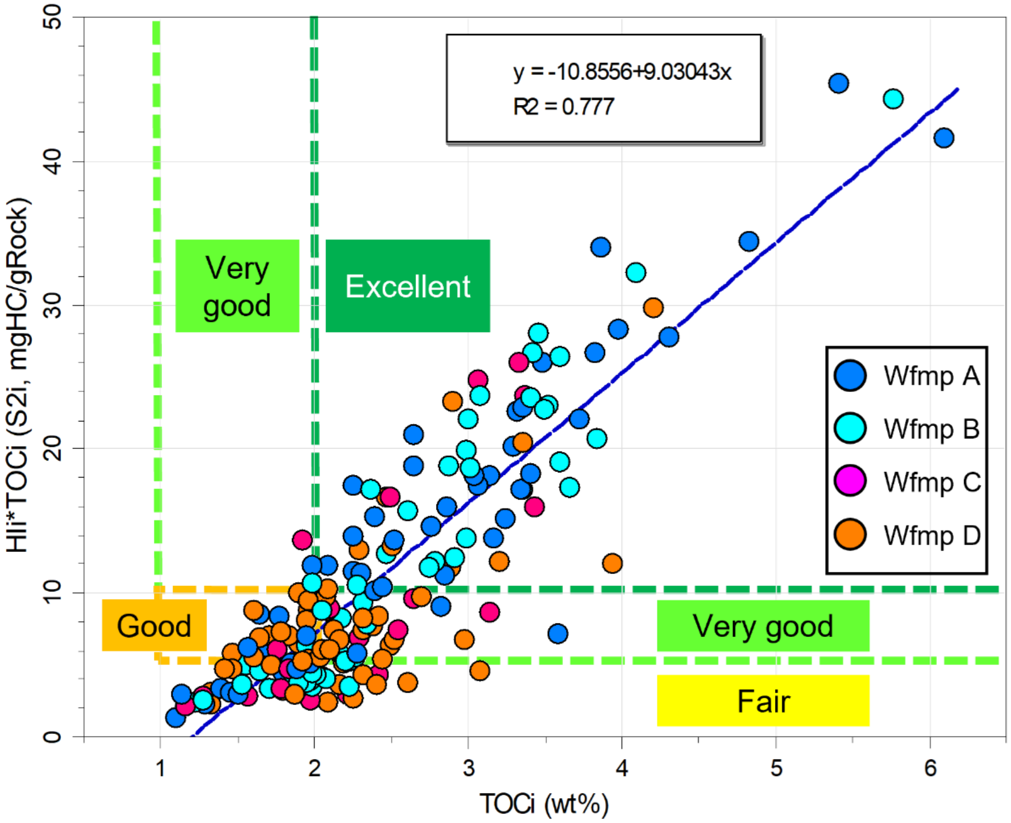


Figure 2.9. A plot of TOC$_i$ vs. S2$_i$ for the Wolfcamp Play showing excellent source rock potential is mostly from Wolfcamp A and B intervals.

2.4.2. Sensitivity Analysis and Uncertainty Ranges

As the kinetic-based inversion procedure is TR-dependent, a sensitivity analysis is necessary to address the impacts of various scenarios related to thermal maturity due to burial and organofacies (e.g., kinetic model) toward the reconstruction of HI_i and TOC_i. The scenarios of varying surface erosion, which were determined to be the essential control of source maturity and organofacies (kinetic model), are assessed. This sensitivity analysis is performed at one well only to demonstrate the uncertainties involved.

2.4.2.1. Sensitivity due to Thermal Maturity

The thermal maturity of the source rock is calibrated with vitrinite reflectance and any available thermal indicators. However, the uncertainties of the calibration data - such as sample type (e.g., whole-rock vs. kerogen isolated), geological context, maceral abundance and identification, and vitrinite chemistry - lead to uncertainties in the observed vitrinite reflectance value (Katz and Lin, 2021) and thus to the model calibration. Delaware Basin is highly over-pressured in the Permian strata (Lee and Williams, 2000; Sinclair, 2007), and vitrinite reflectance enhancement and retardation could occur (Qiu et al., 2007). Vitrinite-rich organofacies (e.g., mixed marine-terrigenous and/or terrigenous-rich organic matters) and/or oxygen-rich environment (oxic) during the deposition (Fang and Jianyu, 1992; Fermont, 1988) can also enhance the observed vitrinite reflectance. Therefore, thermal maturity uncertainty must be considered when restoring the source rock's initial quality and quantity.

Additionally, other thermal maturity indicators (e.g., Raman spectroscopy (Manos, 2018) and carbonate clumped isotope thermometry (Naylor et al., 2020), among many others) may also help to constrain the thermal maturity.

Figure 2.10 gives the sensitivity analysis example from the same well as in **Figure 2.6** but only for Wolfcamp A interval. The base case scenario is assumed as OF "B" typed source rock and surface erosion of 900 ft. for the burial history. From this base case scenario, Wolfcamp A interval is at maturity of TR 74-78%, and the restored initial source quality and quantity are in the range of 436-614 mgHC/gTOC for HI_i and 3.17-3.60 wt% for TOC_i, respectively.

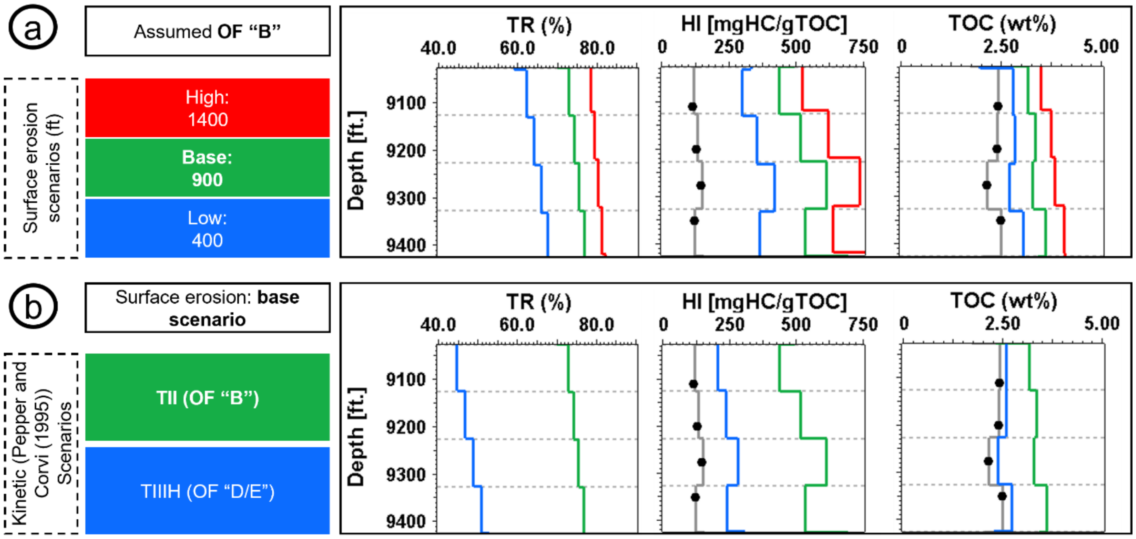


Figure 2.10. Sensitivity analysis examples of the assumptions (TR) involved in this SR restoration workflow. (a) Thermal maturity (surface erosion); (b) organofacies (kinetic). From left to right: simulation scenarios affecting TR; simulated TR; restored HI_i and TOC_i (colored lines) and simulated HI and TOC value (grey line) with measured HI and TOC data (dots).

Two additional surface erosion scenarios (± 500 ft. from the base case) are included due to thermal maturity uncertainty in the basin's central-eastern part (**Figure 2.10a**). For the Wolfcamp A interval, the predicted TR is 78-82% and 62-68% for the high and low surface erosion scenarios, respectively. The restored HI$_i$ is 525-737 mgHC/gTOC for the higher erosion case and 299-420 mgHC/gTOC for the lower erosion case, and the corresponding restored TOC$_i$ is 3.48-4.07 wt% for the higher erosion case and 2.69-3.04 wt% for the lower erosion case.

2.4.2.2. Sensitivity due to Organofacies

Although the organofacies are defined by integrating macerals, oil biomarkers, and Rock-Eval, due to data limitations, most of the organofacies determined in the inversion procedure, especially in the western part of the basin, are with uncertainty. It is essential to incorporate maceral and biomarkers to define the organofacies for each data point used for the inversion rather than Rock-Eval data alone. This is because OF classification from Pepper and Corvi (1995) is based on similarity of organic precursor, depositional environment, and early diagenetic history and has no direct relationship to HI value (Baur, 2019).

Figure 2.10b illustrates the impact of organofacies uncertainty on the restored HI$_i$ and TOC$_i$ by considering two different source types, OF “B” vs. “D/E”. With the base case of erosion, the OF “D/E” type yields a TR maturity of 45-51%, compared with a TR maturity of 74-78% with OF “B”. The restored HI$_i$ are in the range of 205-280 mgHC/gTOC compared with HI$_i$ of 436-614 mgHC/gTOC for OF “B”, and the corresponding restored TOC_i is 2.37-2.73 wt% compared with TOC_i of 3.17-3.60 wt% for OF “B”. These variations indicate the importance of determining source rock organofacies appropriately.

2.4.2.3. Sensitivity due to Kinetic Models

Another uncertainty relates to the variation in kinetic models. There are at least ten kinetic models published for the OF "B", including Abu-Ali et al. (1999), Baur (2019), Behar et al. (1997), diPrimio and Horsfield (2006), Espitalié et al. (1988), Pepper and Corvi (1995), Ungerer (1990), and Vandenbroucke et al. (1999). A complete list of the kinetic models can be found in Baur (2019). To demonstrate the impact of kinetic models on TR, **Figure 2.11** shows the TR vs. depth and thermal maturity (%R_o) from various kinetic models for source rock with OF "B". Within the same depth interval of Wolfcamp A (see **Figure 2.6** for the well location), TR derived from multiple kinetic models of "B" varies from 30% to 90%, indicating the importance of selecting an appropriate kinetic model.

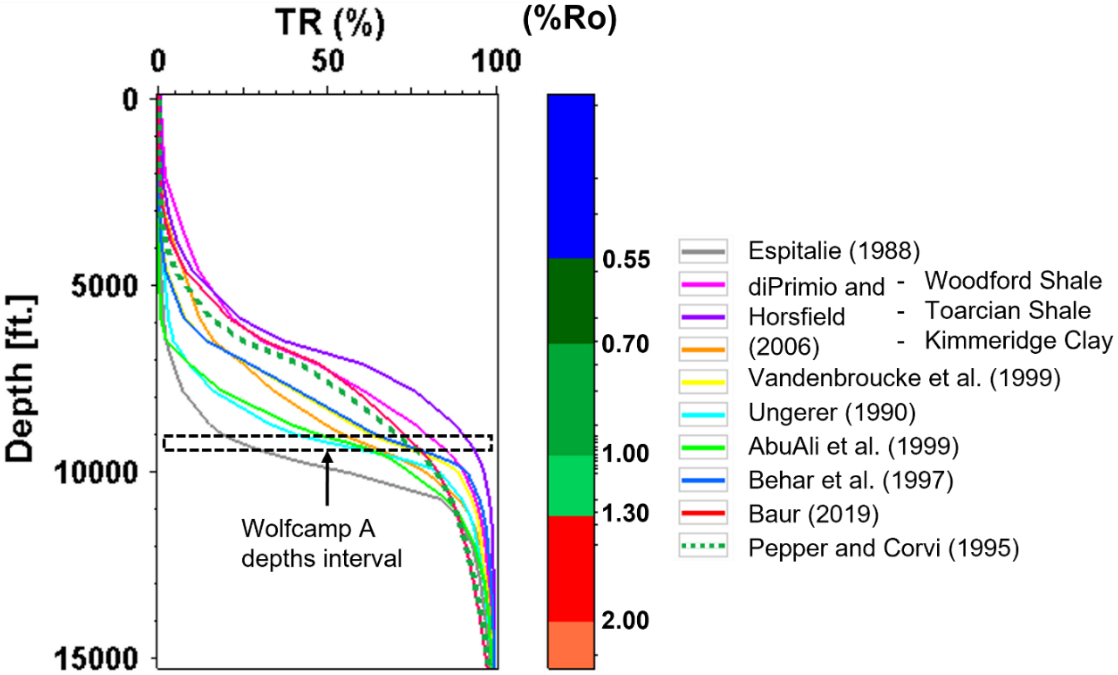


Figure 2.11. A TR vs. depth vs. %Ro plot from various kinetic models for source rock with OF "B“ at a well showing a wide range of TR within the Wolfcamp A depths interval.

2.5. Discussion

2.5.1. Limitation of the Inversion Procedure

We have evaluated the inversion procedure on wide-range TR maturity levels assuming the same HI and TOC values. For example, if assuming HI = 100 mgHC/gTOC, the restored HI_i is >1,000 mgHC/gTOC when TR >90% (see **equation (2)**). Therefore, this procedure only applies to a TR and HI domain shown in **Figure 2.12**. The dashed line can screen the HI data against TR to avoid the area domain that is physically not applicable in calculating the restored HI_i and TOC_i using the inversion procedure. For example, all Wolfcamp Play datasets we applied here are within the applicable domain (HI_i <1,000 mgHC/gTOC).

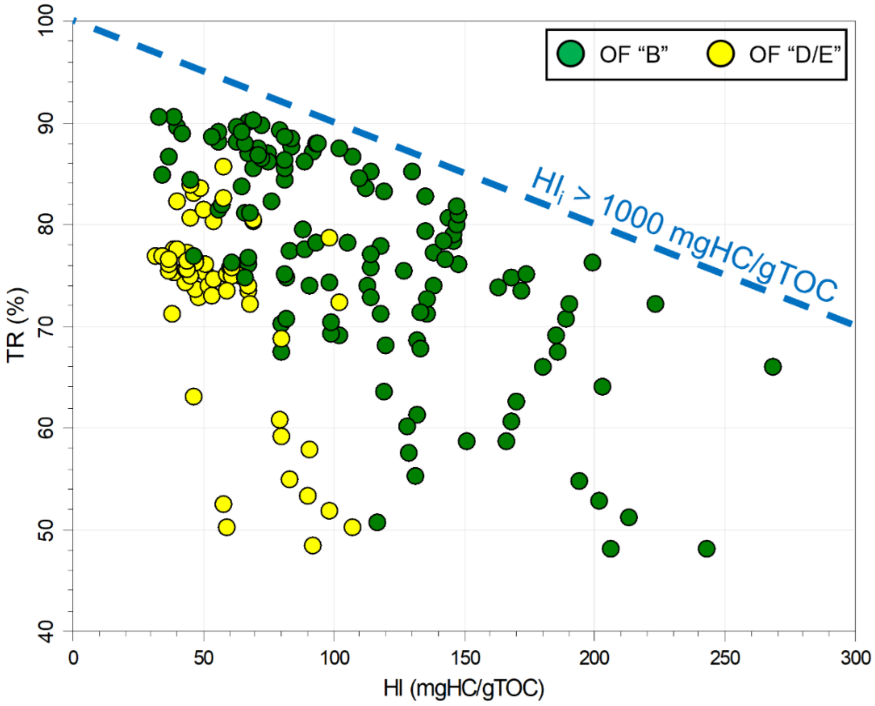


Figure 2.12. HI vs. TR plot with a dashed line to screen for projected restored HI_i >1,000 mgHC/gTOC using the kinetic-based inversion procedure.

2.5.2. Advantages of the Inversion Procedure

Despite the limitation and uncertainty, advantages of using kinetic-based inversion procedures to restore HI_i and TOC_i include:

1. This inversion procedure does not require the assumption of HI_i, which is usually determined from the measurements of the immature sample, but is often lacking.

2. This workflow uses TR derived from the thermal indicator calibrated basin model (3D or multiple 1D/2D) to restore HI_i and TOC_i from the measured HI and TOC data.
3. The method uses the measurements of HI and TOC of the source rock as controls to restore HI_i and TOC_i rather than assuming HI_i and TOC_i as the input for the source rock in the basin model. Therefore, the calculated present-day HI and TOC from the basin model match the measured HI and TOC.
4. The inversion procedure captures the heterogeneity of the restored HI_i and TOC_i of the source rock, as demonstrated by the variability of the restored HI_i and TOC_i values of the Wolfcamp Play dataset in the Delaware Basin.

The procedure provides a way to restore the HI_i and TOC_i based on our understanding of the thermal history, organofacies, and proven kinetic models and measured HI and TOC data of the source rock.

2.5.3. Implications to Calculated Ultimate Expellable Hydrocarbon Potential

To illustrate the impact of restoring the measured HI and TOC data in hydrocarbon resources estimation, we calculate the "ultimate expellable hydrocarbon potential (UEP)" (Pepper and Roller, 2021). UEP is a mass fraction of the generative and expellable hydrocarbons from the organic matter, which can accumulate as a reservoir fluid within-, and/or migrated out from-, the source rock. The calculated mass is expressed in liquid and gas volumetric units at surface conditions per unit area (Pepper and Roller, 2021).

The calculated UEP based on the P90/50/10 values of the restored HI_i and TOC_i in barrels of oil equivalent (boe) per acre-ft. (1 acre = 4046.86 m^2; 1 foot (ft.) = 0.3048 m) of source rocks are shown in **Table 2.3**. High-quality Wolfcamp A and B intervals have the UEP of 97/335/648 and 86/238/613 boe/acre-ft., respectively, based on their P90/50/10 values HI_i and TOC_i. The relative lean intervals of Wolfcamp C and D have UEP of 74/168/545 and 62/170/389, respectively, based on their P90/50/10 values HI_i and TOC_i. This calculation shows that the UEP has an extensive range of values based on P90/50/10, signifying the importance of considering the different HI_i and TOC_i for hydrocarbon resource evaluation. For example, a variation of 35% and 46% of restored HI_i and TOC_i (**Table 2.1** and **Table** 2.2), respectively, of P10 values relative to P50 values for Wolfcamp A interval yields a variation of 93% calculated UEP (**Table 2.3**). This variation is critical to perform the whole-cycle hydrocarbon mass balance from generation potential, trapped and retained volume, to hydrocarbon production in the unconventional/ hybrid plays like Delaware Basin to address the UEP volume uncertainties associated with restored HI_i and TOC_i.

Table 2.3. Wolfcamp Play calculated UEP in boe/acre.ft based on P90/P50/P10 values of restored HI_i and TOC_i.

	P90	P50	P10
Wolfcamp A	97	335	648
Wolfcamp B	86	238	613
Wolfcamp C	74	168	545
Wolfcamp D	62	170	389

2.6. Conclusion

Wolfcamp SR is predominantly the marine clastic (OF “B”) and the mixed marine-terrigenous (OF “D/E”) organofacies. The P90/50/10 values of TR maturity are about 60/77/90 % indicating the peak oil generation window. Based on this kinetic-based inversion result, the restored HI_i and TOC_i are shown to be excellent source rock for Wolfcamp A and B intervals. Generally, Wolfcamp A and B have initial source rock potential (P50) of 530 and 429 mgHC/gTOC of HI_i with 2.64 and 2.27 wt% of TOC_i, respectively. Wolfcamp C and D are relatively lean but are still considered good to very good source rocks with P50 of restored HI_i of 303 and 325 mgHC/gTOC and TOC_i of 2.21 and 2.09 wt%, respectively.

Considering uncertainties, e.g., burial history and simulated TR, the restored HI_i and TOC_i range could be even higher. Therefore, thermal maturity, organofacies, and kinetic models must be constrained before using the kinetic-based inversion procedure.

This proposed kinetic-based inversion method has several advantages compared with other methods. The preliminary calculation of the restored HI_i and TOC_i and the UEP support the massive hydrocarbon volume predicted and produced for the Wolfcamp Play in the Delaware Basin.

2.7. References

Abu-Ali, M. A., J.-L. L. Rudkiewicz, J. G. McGillivray, and F. Behar, 1999, Paleozoic petroleum system of central Saudi Arabia: GeoArabia, v. 4, p. 321-336.

Adams, J. E., 1965, Stratigraphic-tectonic development of Delaware Basin: AAPG bulletin, v. 49, p. 2140-2148.

Barker, C., and M. Pawlewicz, 1987, The effects of igneous intrusions and higher heat flow on the thermal maturity of Leonardian and younger rocks, western Delaware Basin, Texas: Glass mountains: SEPM Guidebook, v. 87, p. 27.

Baskoro, A. S., E. Ganguly, A. Z. Yu, and S. Misra, 2021, End-to-end machine learning workflow for lithofacies classification and source rock organo-facies identification in the Permian Delaware Basin, southwest USA: AAPG/SEG International Meeting for Applied Geoscience & Energy.

Baur, F., 2019, Predicting petroleum gravity with basin modeling: New kinetic models: AAPG Bulletin, v. 103, p. 1811-1837.

Becker, M., A. M. Romero, and A. Yu, 2019, 3-D basin modeling of the Permian Delaware Basin: tectonic evolution assessment to improve definition of boundary conditions: AAPG Annual Convention and Exhibition.

Behar, F., M. Vandenbroucke, Y. Tang, F. Marquis, and J. Espitalie, 1997, Thermal cracking of kerogen in open and closed systems: determination of kinetic parameters and stoichiometric coefficients for oil and gas generation: Organic geochemistry, v. 26, p. 321-339.

Bievenour, A., and S. Sonnenberg, 2019, Reservoir characterization of the Bone Spring and Wolfcamp Formations, Delaware Basin, Ward County, West

Texas: Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July 2019, p. 891-906.

Blakey, R., 2019, Colorado Plateau Geosystems. Inc., Greater Permian Basin Series: https://deeptimemaps.com/greater-permian-basin/ (accessed 4 May 2022).

Brooks, J., C. Cornford, and R. Archer, 1987, The role of hydrocarbon source rocks in petroleum exploration: Geological Society, London, Special Publications, v. 26, p. 17-46.

Chen, Z., and C. Jiang, 2015, A data driven model for studying kerogen kinetics with application examples from Canadian sedimentary basins: Marine and Petroleum Geology, v. 67, p. 795-803.

Chen, Z., C. Jiang, D. Lavoie, and J. Reyes, 2016, Model-assisted Rock-Eval data interpretation for source rock evaluation: Examples from producing and potential shale gas resource plays: International Journal of Coal Geology, v. 165, p. 290-302.

Chen, Z., X. Liu, Q. Guo, C. Jiang, and A. Mort, 2017, Inversion of source rock hydrocarbon generation kinetics from Rock-Eval data: Fuel, v. 194, p. 91-101.

Echegu, S., A. K. Bissada, and L. Elrod, 2021, Geochemical characterization and classification of crude oils of the Permian Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 105, p. 223-246.

Espitalié, J., P. Ungerer, I. Irwin, and F. Marquis, 1988, Primary cracking of kerogens. Experimenting and modelling C1, C2–C5, C6–C15 and C15+ classes of hydrocarbons formed: Organic Geochemistry, v. 13, p. 893-899.

Fang, H., and C. Jianyu, 1992, The cause and mechanism of vitrinite reflectance anomalies: Journal of Petroleum Geology, v. 15, p. 419-434.

Fermont, W. J., 1988, Possible causes of abnormal vitrinite reflectance values in paralic deposits of the Carboniferous in the Achterhoek area, The Netherlands: Organic geochemistry, v. 12, p. 401-411.

Gardner, M., and M. Sonnenfeld, 1996, Stratigraphic changes in facies architecture of the Permian Brushy Canyon Formation in Guadalupe Mountains National Park, west Texas: Society of Economic Palaeontologists and Mineralogists Permian Basin Section, p. 17-40.

Hantschel, T., and A. I. Kauerauf, 2009, Petroleum generation, Fundamentals of Basin and Petroleum Systems Modeling, Springer, p. 151-198.

Hills, J. M., 1984, Sedimentation, tectonism, and hydrocarbon generation in Delaware Basin, west Texas and southeastern New Mexico: AAPG bulletin, v. 68, p. 250-267.

Hills, J. M., 1985, Structural evolution of the Permian basin of west Texas and New Mexico: Structure tectonics of Trans-Pecos Texas: West Texas Geological Society, Publication, v. 85, p. 81.

Jarvie, D. M., R. J. Hill, T. E. Ruble, and R. M. Pollastro, 2007, Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment: AAPG bulletin, v. 91, p. 475-499.

Justwan, H., and B. Dahl, 2005, Quantitative hydrocarbon potential mapping and organofacies study in the Greater Balder Area, Norwegian North Sea: Geological Society, London, Petroleum Geology Conference series, p. 1317-1329.

Katz, B. J., and I. Arango, 2018, Organic porosity: a geochemist's view of the current state of understanding: Organic Geochemistry, v. 123, p. 1-16.

Katz, B. J., and F. Lin, 2021, Consideration of the limitations of thermal maturity with respect to vitrinite reflectance, Tmax, and other proxies: AAPG Bulletin, v. 105, p. 695-720.

Kinley, T. J., L. W. Cook, J. A. Breyer, D. M. Jarvie, and A. B. Busbey, 2008, Hydrocarbon potential of the Barnett Shale (Mississippian), Delaware Basin, west Texas and southeastern New Mexico: AAPG bulletin, v. 92, p. 967-991.

Kuchinskiy, V., K. Gentry, and R. Hill, 2012, Source rock evaluation technique: a probabilistic approach for determining hydrocarbon generation potential and in-place volume for shale plays: AAPG Annual Convention and Exhibition, Long Beach, California.

Kvale, E. P., C. M. Bowie, C. Flenthrope, C. Mace, J. M. Parrish, B. Price, S. Anderson, and W. A. DiMichele, 2020, Facies variability within a mixed carbonate–siliciclastic sea-floor fan (upper Wolfcamp Formation, Permian, Delaware Basin, New Mexico): AAPG Bulletin, v. 104, p. 525-563.

Lee, M.-K., and D. D. Williams, 2000, Paleohydrology of the Delaware basin, western Texas: overpressure development, hydrocarbon migration, and ore genesis: AAPG bulletin, v. 84, p. 961-974.

Lew, C., S. Baharuddin, and J. D. Pigott, 2013, Deepwater basin model for the Permian: Delaware Basin example: IPTC 2013: International Petroleum Technology Conference, p. cp-350-00399.

Lewan, M., M. Henry, D. Higley, and J. K. Pitman, 2002, Material-balance assessment of the New Albany-Chesterian petroleum system of the Illinois basin: AAPG bulletin, v. 86, p. 745-777.

Manos, T. A., 2018, Thermal maturity modeling of organic-rich mudrocks in the Delaware Basin using Raman spectroscopy of carbonaceous material, Texas A&M University.

Modica, C. J., and S. G. Lapierre, 2012, Estimation of kerogen porosity in source rocks as a function of thermal transformation: Example from the Mowry Shale in the Powder River Basin of WyomingEstimation of Kerogen Porosity as a Function of Thermal Transformation: AAPG bulletin, v. 96, p. 87-108.

Naylor, H. N., W. F. Defliese, E. L. Grossman, and C. R. Maupin, 2020, Investigation of the thermal history of the Delaware Basin (West Texas, USA) using carbonate clumped isotope thermometry: Basin Research, v. 32, p. 1150-1165.

Pacton, M., G. E. Gorin, and C. Vasconcelos, 2011, Amorphous organic matter—Experimental data on formation and the role of microbes: Review of Palaeobotany Palynology, v. 166, p. 253-267.

Pepper, A., and E. Roller, 2021, Ultimate expellable potentials of source rocks from selected super basins: What does "world class" look like?: AAPG Bulletin, v. 105, p. 1069-1097.

Pepper, A. S., and P. J. Corvi, 1995, Simple kinetic models of petroleum formation. Part I: oil and gas generation from kerogen: Marine and petroleum geology, v. 12, p. 291-319.

Peters, K., L. Magoon, K. Bird, Z. Valin, and M. Keller, 2006, North Slope, Alaska: Source rock distribution, richness, thermal maturity, and petroleum charge: AAPG bulletin, v. 90, p. 261-292.

Peters, K., and J. Moldowan, 1991, Effects of source, thermal maturity, and biodegradation on the distribution and isomerization of homohopanes in petroleum: Organic geochemistry, v. 17, p. 47-61.

Peters, K., C. Walters, and J. Moldowan, 2005, Source-and age-related biomarker parameters, The biomarker guide, v. 2, Cambridge University Press, p. 483-607.

Peters, K. E., 1986, Guidelines for evaluating petroleum source rock using programmed pyrolysis: AAPG bulletin, v. 70, p. 318-329.

Qiu, N., W. Wang, and M. Xie, 2007, Physical and chemical environments of abnormal vitrinite reflectance evolution in the sedimentary basins: Frontiers of Earth Science in China, v. 1, p. 341-350.

Romero-Sarmiento, M.-F., M. Ducros, B. Carpentier, F. Lorant, M.-C. Cacas, S. Pegaz-Fiornet, S. Wolf, S. Rohais, I. J. M. Moretti, and P. Geology, 2013, Quantitative evaluation of TOC, organic porosity and gas retention distribution in a gas shale play using petroleum system modeling: Application to the Mississippian Barnett Shale, v. 45, p. 315-330.

Rushing, J., A. Chaouche, and K. Newsham, 2004, A mass balance approach for assessing basin-centred gas prospects: integrating reservoir engineering, geochemistry and petrophysics: Geological Society, London, Special Publications, v. 237, p. 373-390.

Schwartz, K. M., G. Hennenfent, M. Hegmann, M. Hoffnagle, D. Bain, and A. McCallister, 2015, Pay distributions and basin architecture of the Wolfcamp shale in the Delaware Basin: Unconventional Resources Technology Conference, San Antonio, Texas, 20-22 July 2015, p. 1729-1733.

Shanmugam, G., 1985, Significance of coniferous rain forests and related organic matter in generating commercial quantities of oil, Gippsland Basin, Australia1: AAPG bulletin, v. 69, p. 1241-1254.

Sinclair, T. D., 2007, The generation and continued existence of overpressure in the Delaware Basin, Texas, Durham University.

Silver, B. A., and R. G. Todd, 1969, Permian cyclic strata, northern Midland and Delaware basins, west Texas and southeastern New Mexico: AAPG Bulletin, v. 53, p. 2223-2251.

Tissot, B. P., and D. H. Welte, 1984a, From kerogen to petroleum, Petroleum formation and occurrence, Springer, p. 160-198.

Tissot, B. P., and D. H. Welte, 1984b, kerogen: composition and classification, Petroleum Formation and Occurrence, Springer, p. 131-159.

Tobey, M. H., and C. E. Campbell, 2016, Hydrogen index as a maturity proxy–some pitfalls and how to overcome them: AAPG Pacific Section and Rocky Mountain Section Joint Meeting.

Ungerer, P., 1990, State of the art of research in kinetic modelling of oil formation and expulsion: Organic Geochemistry, v. 16, p. 1-25.

Vandenbroucke, M., F. Behar, and J. Rudkiewicz, 1999, Kinetic modelling of petroleum formation and cracking: implications from the high pressure/high temperature Elgin Field (UK, North Sea): Organic Geochemistry, v. 30, p. 1105-1125.

Yu, A. Z., M. Becker, A. S. Baskoro, and M. Bhatia, 2020, Fluid property variations and its relationship with the geo-history of the Permian Delaware Basin: AAPG Annual Conference and Exhibition.

3. BASIN-WIDE LITHOFACIES IDENTIFICATION USING UNSUPERVISED MACHINE LEARNING: APPLIED TO THE UNCONVENTIONAL WOLFCAMP PLAY, PERMIAN DELAWARE BASIN

3.1. Introduction

Accurate estimation of basin-wide hydrocarbon resources, both conventional and unconventional, requires several inputs, including the source rock lithofacies (facies with similar lithological characteristics). Source rock lithofacies and thickness affect formation overpressure and therefore, hydrocarbon storage mechanism and capacity, which ultimately control the expulsion efficiency of the source rock (Coutinho et al., 2009; Jarvie et al., 2007; Katz et al., 2017). Lithofacies is also associated with organofacies, facies of kerogens with common organic material from similar depositional environment and early diagenetic history (Pepper and Corvi, 1995), controlled by the depositional environment (Donovan et al., 2017; Evenick, 2016; Pepper and Corvi, 1995). The organofacies directly impact the fluid property and composition (e.g., marine carbonate organofacies have the highest C_{15+} amount) and fluid phase (i.e., mixed marine-terrigenous organofacies are oil and gas prone) of the source rock. The properties of the generated hydrocarbon will ultimately affect the properties of the expelled and retained hydrocarbon (Baur, 2019; di Primio and Horsfield, 2006; Jarvie et al., 2007; Pepper and Corvi, 1995). Therefore, identifying the basin-wide lithofacies of the source rock interval is critical for hydrocarbon resource estimation.

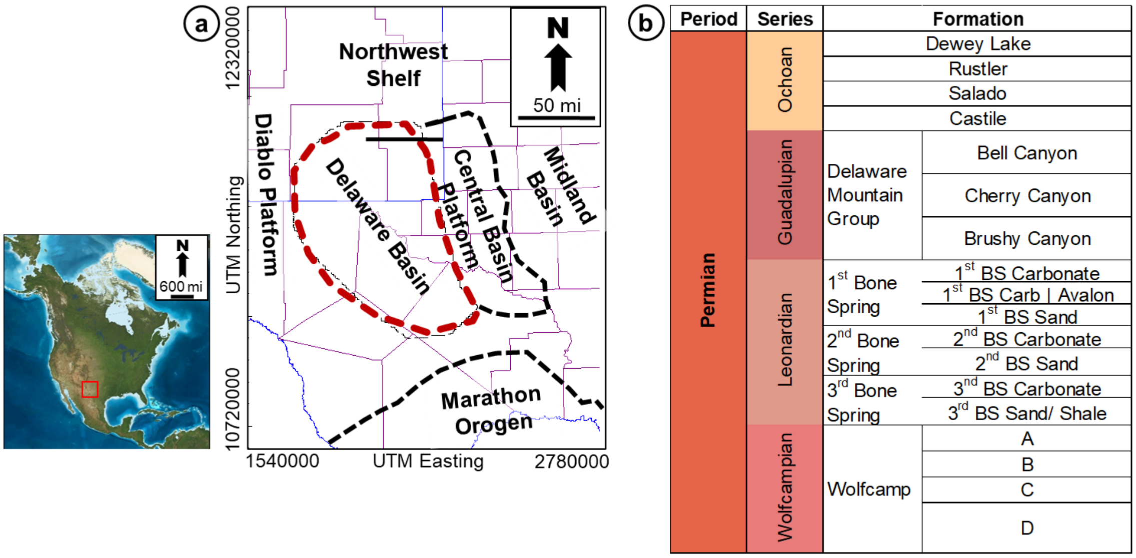
Wireline log measurements are often used for lithofacies identification as they are the most common well data and cover broad depth intervals of the formation of interest. With the development of data science and machine learning, numerous works related to lithofacies identification have been performed using various machine learning techniques applied to well-log data (Chang et al., 2002; He et al., 2019; Li and Misra, 2018; Raeesi et al., 2012; Wang et al., 2014). They showed how machine learning methods could automate well-log-based lithofacies identification efficiently and comparable to core description. However, most of these works are based on supervised machine learning methods, well-log and core-derived parameters, or integration with core data to classify lithofacies for the area of field scale. He et al. (2019) and Li and Misra (2018) have used unsupervised learning on well-log data for well-to-well correlations. Nevertheless, classifying lithofacies based on well-log data is challenging for datasets lacking core data. Another challenge is consistently classifying lithofacies on a large basin-wide dataset.

This paper examines a workflow using unsupervised machine learning methods to identify lithofacies from a proprietary basin-wide well dataset. Unlike other methods, the workflow is data-driven and straightforward, based solely on well-log data. The interpreted lithofacies from the machine learning methods are compared with other studies based on core interpretation and spectral gamma-ray logs from the dataset. We also illustrate the impact of the identified lithofacies distribution on the interpretation of the basin-wide gross depositional environment and organofacies distribution along with their laterally and vertically variability.

3.2. Geological Overview

The case study for this unsupervised machine learning application is the Wolfcamp Play in the Permian Delaware Basin of West Texas and New Mexico, USA (**Figure 3.1**). The Permian Delaware Basin is the western major structural subdivision of the Permian Basin's more massive structure in west Texas and southeastern New Mexico (Hills, 1984) (**Figure 3.1a**). The Delaware Basin is bounded to the west by the Diablo Platform, to the north by the North American craton and Northwestern Shelf areas, and to the south by the Marathon orogenic belt. The Delaware Basin is separated from the Midland Basin to the east by the N-S trending Central Basin Platform (Adams, 1965; Hills, 1984). The Delaware Basin occupies a basinal area of 13,000 mi^2 (33,500 km^2) with a maximum depth of 24,000 ft (7,300 m) and is filled by up to 40,000 mi^3 (170,000 km^3) of Phanerozoic sediments (Hills, 1984).

The Wolfcamp Formation consists of four main intervals: Wolfcamp A, -B, -C, and -D (Bievenour and Sonnenberg, 2019; Dutton et al., 2005); and is part of the thick Permian strata in the Delaware Basin (**Figure 3.1b**). The deposition took place during the early Permian with a depositional environment of generally a deep, oceanic connected interior continental basin surrounded by carbonate shelf (the Northwestern Shelf, eastern Central Basin Carbonate Platform, western Diablo Platform), the northwestern Pedernal uplift, and the southern siliciclastic shelf in front of the Marathon orogenic belt (Blakey, 2019). Wolfcamp intervals were also deposited during the frequent eustatic fluctuations (Fairhurst et al., 2021). The complex depositional setting caused lithofacies heterogeneity, primarily limestone, silts, and mudstone, within the Wolfcamp intervals in the Delaware Basin.



Period	Series	Formation	
Permian	Ochoan	Dewey Lake	
		Rustler	
		Salado	
		Castile	
	Guadalupian	Delaware Mountain Group	Bell Canyon
			Cherry Canyon
			Brushy Canyon
	Leonardian	1st Bone Spring	1st BS Carbonate
			1st BS Carb | Avalon
			1st BS Sand
		2nd Bone Spring	2nd BS Carbonate
			2nd BS Sand
		3rd Bone Spring	3nd BS Carbonate
			3rd BS Sand/ Shale
	Wolfcampian	Wolfcamp	A
			B
			C
			D

Figure 3.1. (a) Major subdivision and boundaries of the Permian Basin around the Delaware Basin (after Dutton et al. (2015) and Silver and Todd (1969)). (b) Stratigraphic chart of Permian strata, which includes Wolfcamp interval subdivision for the Delaware Basin (after Bievenour and Sonnenberg (2019) and Dutton et al. (2015)).

The Wolfcamp Play is one of the main targets for unconventional production, along with the Leonardian Bone Spring interval in the Permian Delaware Basin (Gaswirth et al., 2018). Among the four intervals, Wolfcamp A is the most drilled target zone for unconventional play in the Delaware Basin (EIA, 2019; Fairhurst et al., 2021). Wolfcamp source rocks, especially Wolfcamp A and B, are the source for the Wolfcamp unconventional play on the Delaware Basin and mixed oil sources on the Central Basin Platform and Northwest Shelf areas (Fairhurst et al., 2021). The primary source rocks of Wolfcamp A and B are organic-rich, silica-rich facies. Identifying basin-scale lithofacies is necessary to improve hydrocarbon resources assessment, including the distribution, for both conventional and unconventional accumulation sourced from the Wolfcamp interval in the Delaware Basin.

3.3. Methods

The unsupervised machine learning methods used to identify the lithofacies are based on four sets of well-log measurements. A large amount of well-log data from over 1500 wells are considered, but only a small subset of wells was selected and used in this study. Two main steps are required to identify the basin-wide lithofacies from the well-log dataset in this workflow: (1) pre-processing of the well-data; and (2) clustering the well-data using unsupervised machine learning methods.

3.3.1. Pre-Processing

The main goals of the pre-processing step is to prepare the dataset and select the well data for the clustering. Pre-processing is essential in automating the data preparation from 1500 wells available for time and computational resource usage efficiency.

3.3.1.1. Preparing the Dataset

The first step is to select the well dataset with the log type of measurement and formation of interest out of 1500 wells available for the clustering (**Figure 3.2a**). Out of 1500 wells available, we implement four filtering criteria:

1. wells must be located within the Delaware Basin,
2. wells must have four types of well-log measurements: photoelectric factor (PEF), gamma-ray (GR), sonic (DT), and bulk density (DEN)). These four types of well-log measurements represent lithology indicators and complement each other in lithofacies interpretation,
3. wells must cover >60% Wolfcamp depth interval,
4. wells must have well-log measurements with consistent and appropriate formats.

Following the four filtering criteria, 31 wells were selected out of 1500 wells (**Figure 3.2b**).

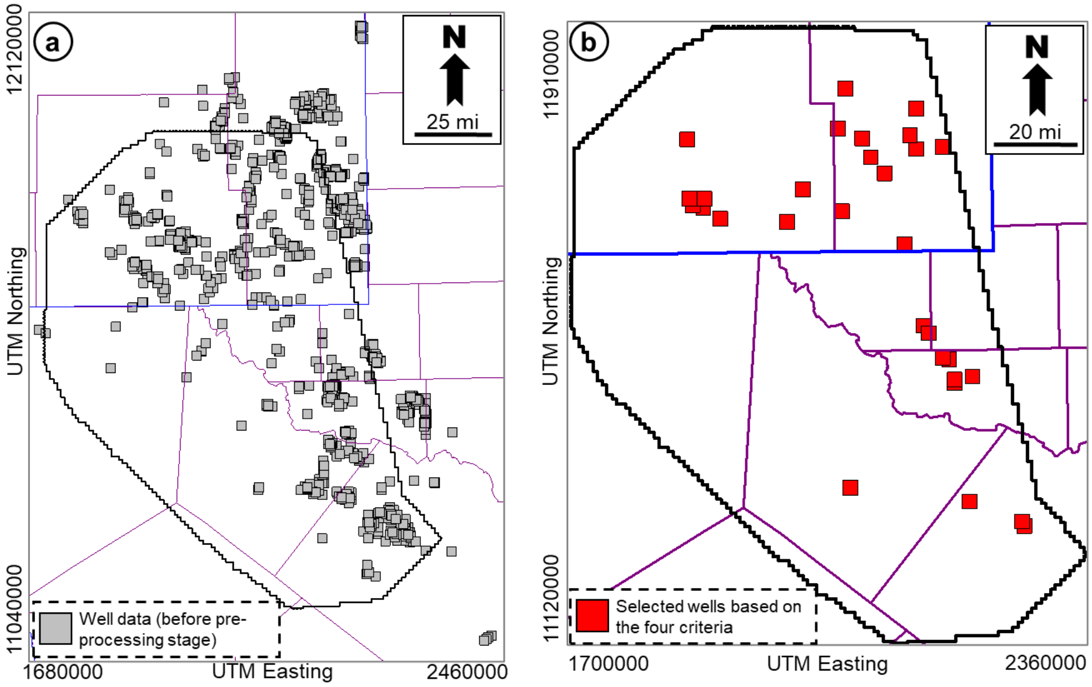


Figure 3.2. Before (a) and after (b) pre-processing the well dataset to select well data of Wolfcamp interval for the clustering. Note that "Others" represent types of well-log measurement other than the four types of well-log measurement used for the clustering in this workflow including GR (gamma-ray), DEN (bulk density), DT (sonic), and PEF (photoelectric factor).

3.3.1.2. Data Cleaning

The next step is to create the DataFrame from the dataset as the input for the clustering. The DataFrame is then scaled using a Standard Scaler and cleaned from missing values and outliers, accounting for 63,311 data points (**Figure 3.3**).

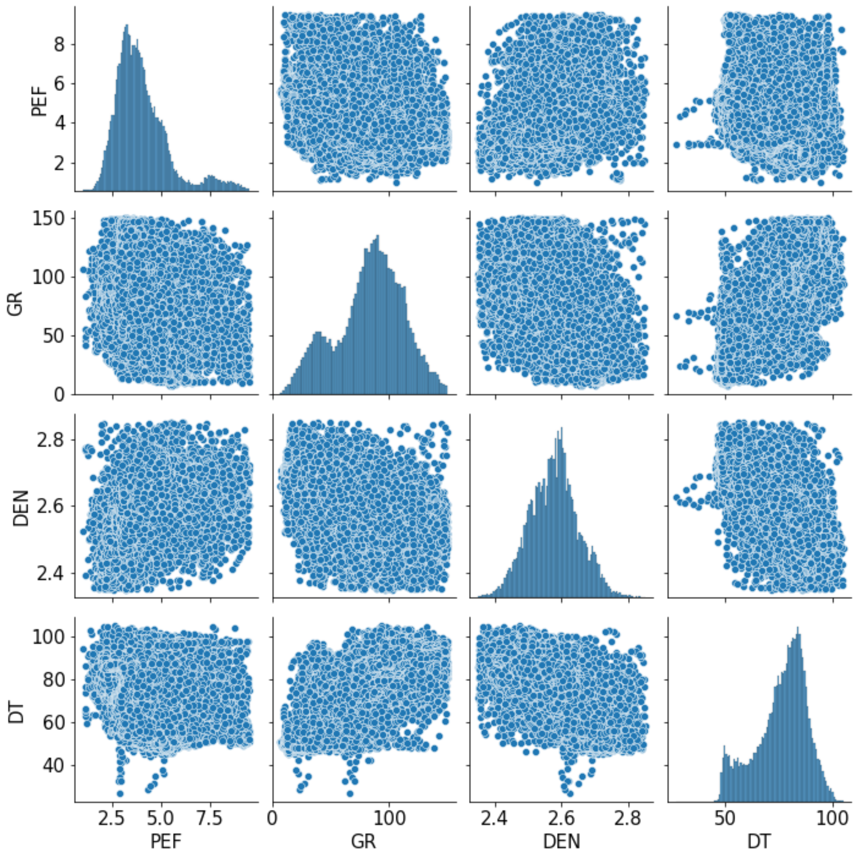


Figure 3.3. Matrix scatter-plot of the four types of well-log measurement for the 31 well data of Wolfcamp DataFrame (63,311 data points). PEF: photoelectric factor; GR: gamma-ray; DEN: bulk density; DT: sonic.

3.3.2. Clustering

The main goal of the clustering step is to classify the data points into several compact and well-separated facies that can be interpreted as consistent lithofacies throughout the basin. For clustering, unsupervised, or data-driven machine learning, clustering methods (Alloghani et al., 2020; Gentleman and Carey, 2008) were applied for the well-log data (**Figure 3.3**) following the workflow.

3.3.2.1. Finding the Number of Clusters Present in the Data

In this workflow, we primarily deploy clustering methods that require a specified number of clusters as user input. An essential requirement for robust clustering is determining the optimal number of clusters in the dataset. The clusters should be consistent and robust. The optimal number of clusters present in the data (*k*) is determined using several intrinsic metrics, including the Elbow curve method (Edwards and Cavalli-Sforza, 1965; Kodinariya and Makwana, 2013), Silhouette score (Rousseeuw, 1987), Calinski-Harabasz index (Caliński and Harabasz, 1974), and Davies-Bouldin index (Davies and Bouldin, 1979). These metrics assess the goodness of the data partitioning without the use of any external information. These scores serve as heuristic tools to evaluate the clustering performance (as shown by Gonzalez and Misra (2022)). These scores are based on the compactness (similarity) of each cluster and the separation (distinction) between the clusters. A higher value of Calinski-Harabasz score and lower value of Davies-Bouldin score indicate a robust clustering. For the Silhouette score, the range of performance is set from -1 to 1, where 1 represents the best clustering performance.

3.3.2.2. Determining the Best Clustering Method

Three clustering methods were applied to the well-log data: KMeans (Hartigan and Wong, 1979), Agglomerative Hierarchical (Johnson, 1967), and Spectral (Ng et al., 2001). KMeans is the distance-based clustering where each centroid sets itself in the center of one of the *k* clusters in the data.

Agglomerative Hierarchical clustering considers each data point as an individual cluster and merges the most similar data points until *k* clusters are formed. On the other hand, Spectral clustering roots in graph theory, where communities of nodes are grouped in a graph based on the edges connecting them.

Best clustering methods for the given dataset are determined with the help of three intrinsic cluster validation metrics: Silhouette score, Calinski-Harabasz index, and Davies-Bouldin index. An Adjusted Rand Score is also used to determine whether the clusters from the two clustering methods are similar. The closer to 100% the score is, the more similar the clusters are, indicating consistent results, even from different clustering methods. Similar clustering comparisons have been performed by Gonzalez and Misra (2022) and Chakravarty and Misra (2022) to find the best clustering method.

3.3.2.3. Clusters Interpretation and Validation

Clusters from the best-unsupervised learning methods are then interpreted into lithofacies based on the feature set derived from the well-log measurements. The interpretation is compared with other studies based on core interpretation and spectral gamma-ray logs from the dataset. The interpreted lithofacies are compared at individual wells nearby to see whether the clustering workflow can classify the various well-log features following the basin-wide geological variation. The interpreted lithofacies are validated for Wolfcamp A samples due to the availability of other geological and geochemical data (literature and dataset). The identified basin-wide lithofacies are then used to construct basin-wide gross depositional environment (GDE) and associated organofacies.

3.4. Results and Interpretation

3.4.1. Clustering Evaluation

3.4.1.1. First-Level Clustering

All four intrinsic metrics indicate that the optimal number of clusters present is 2 (**Figure 3.4**). It is evident from **Figure 3.5a** that KMeans and Spectral clustering perform better than Hierarchical clustering, with KMeans having the best intrinsic cluster validation metric scores. This is also supported by a high Adjusted Rand Score of 94.35% between KMeans and Spectral clustering (**Table 3.1**). With two as the number of clusters from KMeans clustering, 46,466 data points are assigned as Facies A, while 16,845 data points are assigned as Facies B. Based on the matrix scatter plot (**Figure 3.5b**), it is not easy to interpret the classified facies as each cluster has a wide range and overlapping features of the well-log measurements. Generally, Facies A is predominantly silt and mudstone lithofacies, while Facies B is predominantly carbonate lithofacies. Thus, the second-level clustering workflow is necessary to further subdivide Facies A and Facies B into more refined lithofacies on each new subset of data from the first-level clustering: DataFrame for Facies A and DataFrame for Facies B.

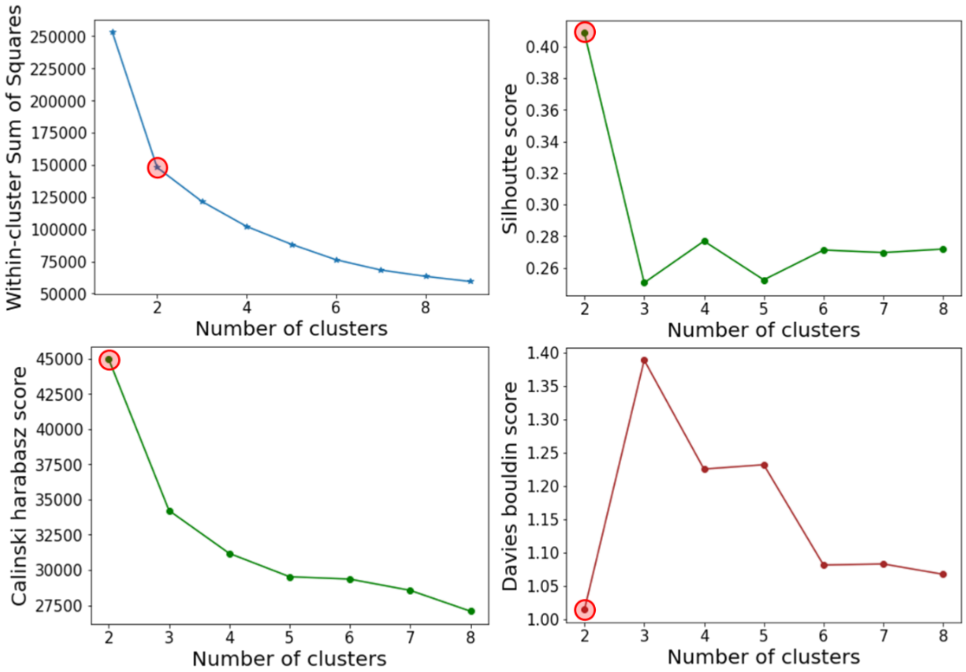


Figure 3.4. First-level clustering: all four intrinsic metrics indicate 2 as the optimal number of clusters present in the Wolfcamp data.

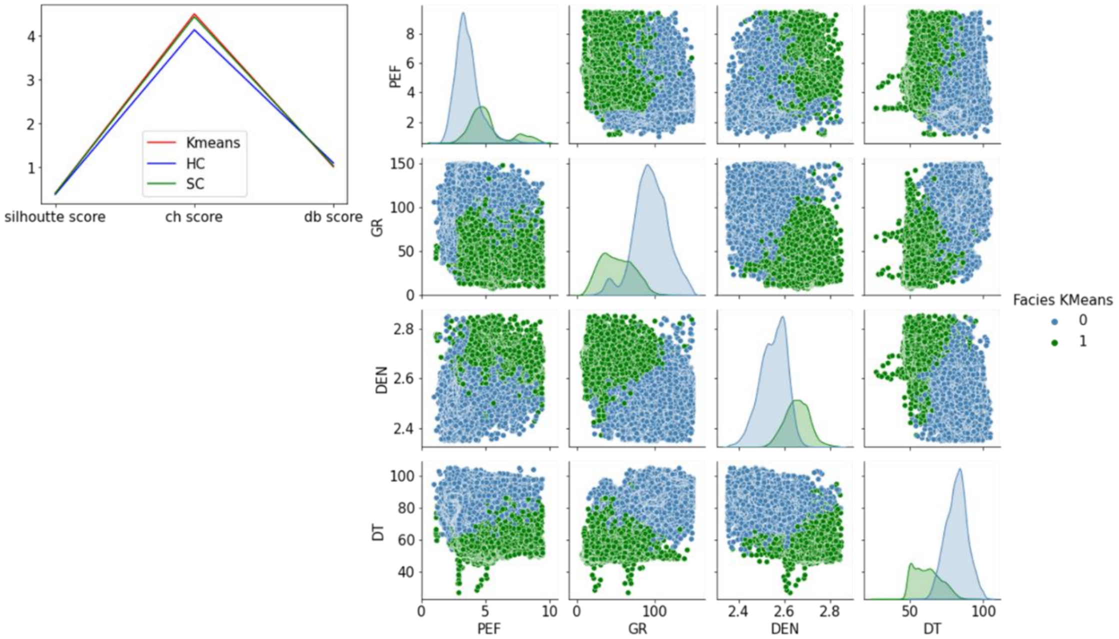


Figure 3.5. Clustering results for the first level-clustering of the Wolfcamp data: (a) intrinsic cluster validation metrics; (b) matrix scatter-plot of the identified KMeans facies.

Table 3.1. Comparison of Adjusted Rand Score of the facies similarity among different clustering methods.

			Adjusted Rand Score (%)		
Clustering	**Samples**	**N of Facies**	**KMeans-HC**	**KMeans-SC**	**HC-SC**
1st-level	63311	2	80.81	94.35	83.03
2nd-level: DF Facies A	46466	4	45.41	66.15	45.73
2nd-level: DF Facies B	16845	3	63.53	88.38	59.87
Combined 2nd-level	63311	7	50.23	72.07	49.49
1st-level (7)	63311	7	52.61	53.81	43.26

3.4.1.2. Second-Level Clustering

The same procedure as the first-level clustering is applied for the second-level clustering to further subdivide the mudstone and carbonate lithofacies into more detailed lithofacies. Two-level clustering has been shown to be useful and informative in identifying physically consistent groups by Falola et al. (2022). For the Facies A dataset, three intrinsic metrics indicate that the optimal number of clusters present is 4 (**Figure 3.6**), and thus the workflow proceeds with four as the number of clusters. Based on **Figure 3.7a**, both KMeans and Spectral clustering perform significantly better than Hierarchical clustering. This is also supported by a higher Adjusted Rand Score of 66.15% between KMeans and Spectral clustering than Hierarchical clustering (**Table 3.1**). Using four as the number of clusters with KMeans clustering, 46,466 data points are subdivided into 4 Facies with data points ranging from 9,288 to 14,089 data points (**Figure 3.7b**).

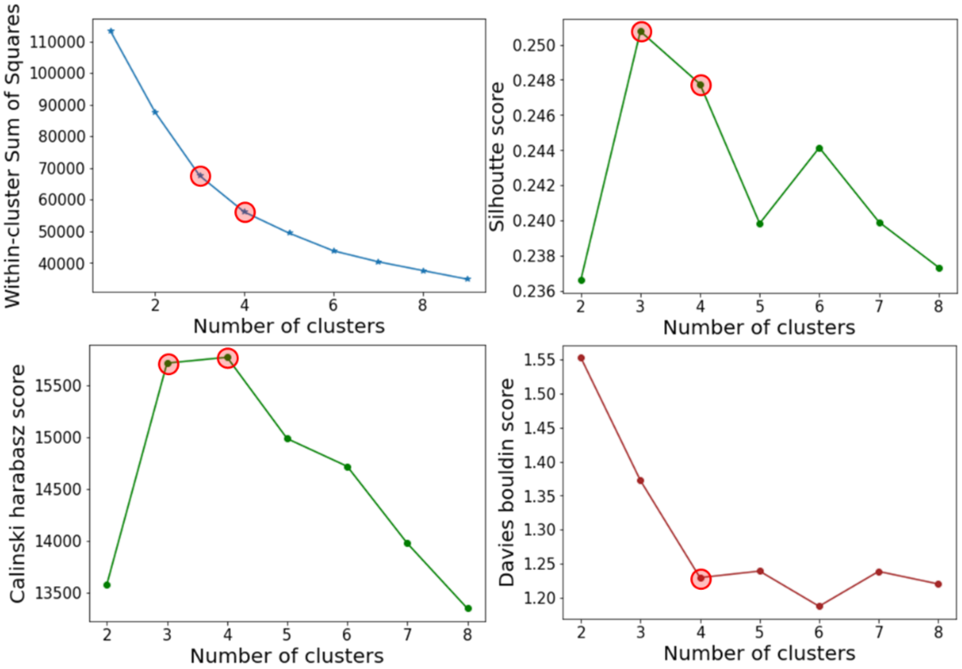


Figure 3.6. Second-level clustering of Facies A - Wolfcamp: All four intrinsic metrics indicate 4 as the optimal number of clusters present in the Wolfcamp data - Facies A.

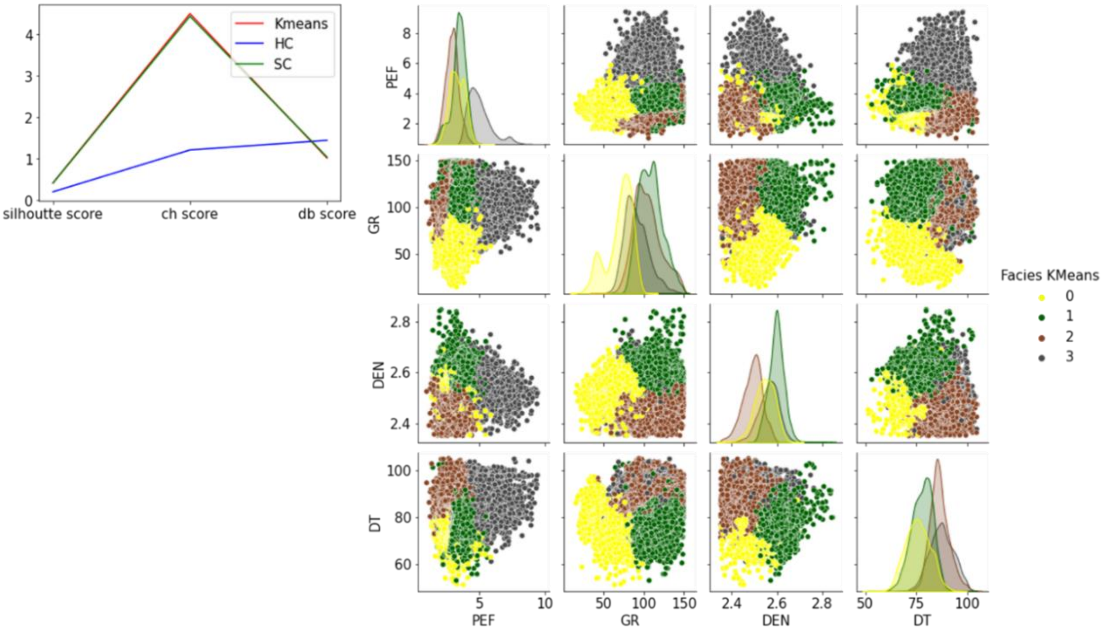


Figure 3.7. Clustering results for the second level-clustering of the Wolfcamp data – Facies A: (a) intrinsic cluster validation metrics; (b) matrix scatter-plot of the identified KMeans facies.

For the Facies B data, three intrinsic metrics indicate that the optimal number of clusters present is 3 (**Figure 3.8**), and thus the workflow proceeds with three as the number of clusters. Based on **Figure 3.9a**, both KMeans and Spectral clustering perform significantly better than Hierarchical clustering. This is also supported by a high Adjusted Rand Score of 88.38% between KMeans and Spectral clustering (**Table 3.1**). Using three as the number of clusters with KMeans clustering, 16845 data points are subdivided into 3 Facies with data points ranging from 3200 to 7417 data points (**Figure 3.9b**). The combined clusters, with seven identified basin-wide lithofacies, have a combined Adjusted Rand Score of 72.07% between KMeans and Spectral clustering, compared to

53.81% if only first-level clustering with seven as the number of clusters is applied.

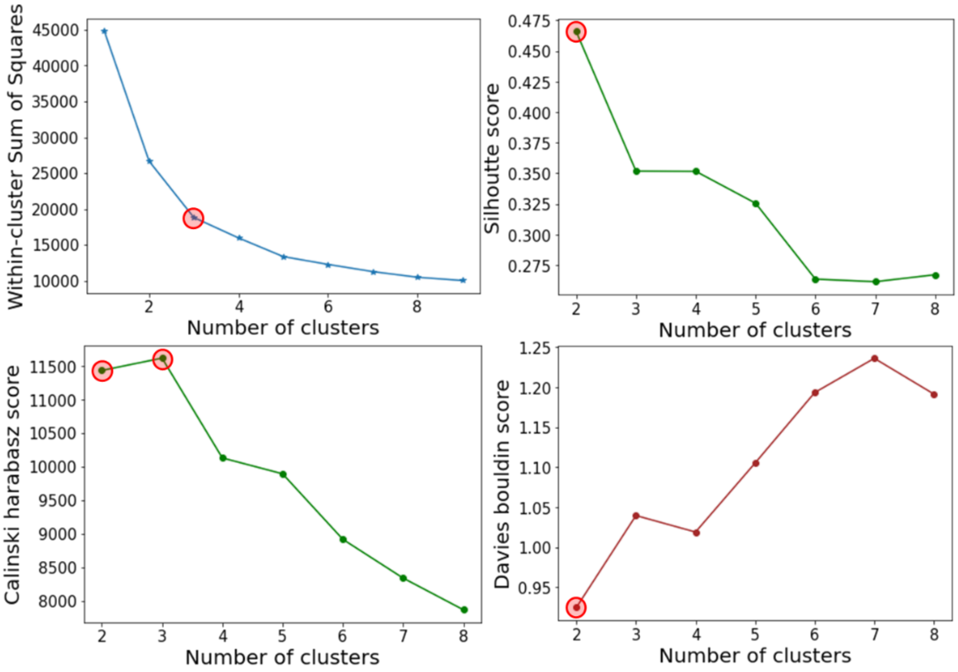


Figure 3.8. Second-level clustering of Facies B - Wolfcamp: intrinsic metrics indicate 2 or 3 as the optimal number of clusters present.

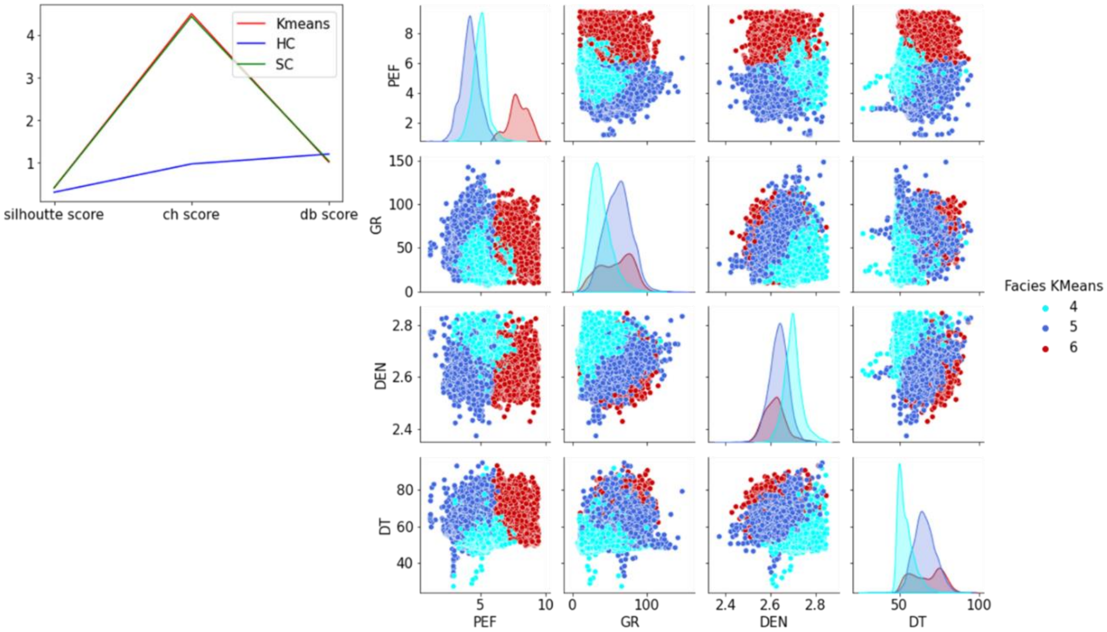


Figure 3.9. Clustering results for the second level-clustering of the Wolfcamp data – Facies B: (a) intrinsic cluster validation metrics; (b) matrix scatter-plot of the identified KMeans facies.

3.4.2. Lithofacies Interpretation and Validation

3.4.2.1. Facies Interpretation Based on the Well-Log Measurement

Seven lithofacies from the combined data (**Figure 3.10**) are interpreted based on their well-log measurements. Generally, Facies 0-3 are interpreted as clastic source-rock facies, while Facies 4-6 are interpreted as carbonate/ marl with non-source-rock facies (see **Table 3.2** for the summary).

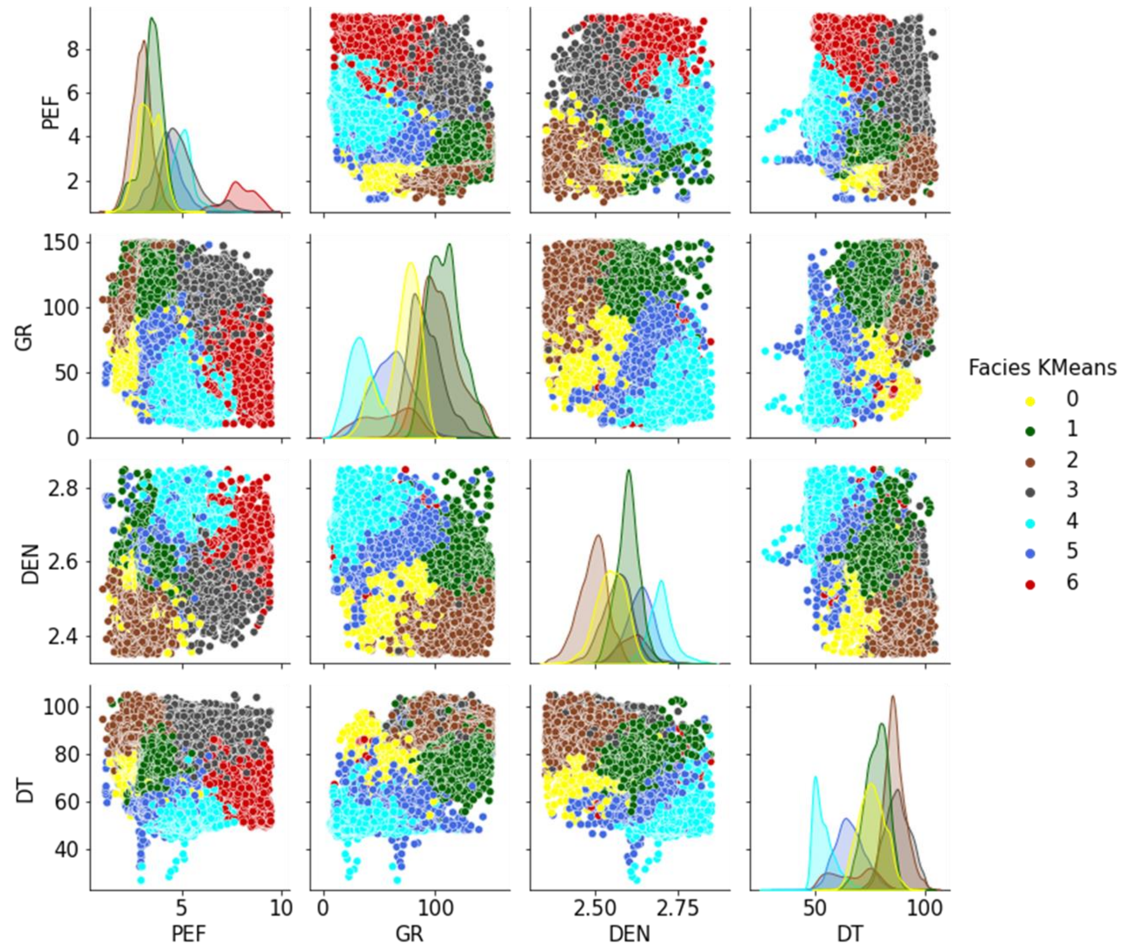


Figure 3.10. Matrix scatter-plot of the identified KMeans facies from the combined data of Facies A and B (second-level clustering).

Table 3.2. Interpreted lithofacies and their associated well-log data distribution.

Facies KMeans	Interpreted Lithofacies	N of samples		PEF barns/electron)	GR (API)	DT (µs/ft.)	DEN (g/cm^3)
0	Siltstone	10729	Mean ± SD:	3.31 ± 0.59	70 ± 16	76 ± 6	2.54 ± 0.04
			p10/50/90:	2.56/3.29/4.09	42/74/87	68/76/84	2.49/2.55/2.60
1	Mixed argillaceous mudstone	14089	Mean ± SD:	3.44 ± 0.56	108 ± 15	78 ± 5	2.60 ± 0.03
			p10/50/90:	2.66/3.48/4.09	90/108/129	71/78/84	2.55/2.60/2.64
2	Siliceous mudstone (organic-rich)	12360	Mean ± SD:	2.92 ± 0.50	104 ± 17	86 ± 5	2.49 ± 0.04
			p10/50/90:	2.27/2.93/3.56	84/102/129	81/86/93	2.43/2.50/2.54
3	Calcareous-argillaceous mudstone (organic-rich)	9288	Mean ± SD:	5.03 ± 0.94	91 ± 15	87 ± 6	2.56 ± 0.04
			p10/50/90:	4.05/4.82/6.29	74/89/111	80/87/95	2.51/2.57/2.62
4	Limestone (dolomitic)	6228	Mean ± SD:	4.97 ± 0.60	36 ± 15	54 ± 5	2.70 ± 0.04
			p10/50/90:	4.23/4.98/5.61	20/34/56	49/53/61	2.65/2.70/2.75
5	Mixed carbonate-siliciclastic (shaly)	7417	Mean ± SD:	4.17 ± 0.69	61 ± 18	66 ± 7	2.63 ± 0.04
			p10/50/90:	3.26/4.17/5.07	38/61/84	57/65/75	2.58/2.64/2.69
6	Heavy mineral bearing (ankerite (?)) interval	3200	Mean ± SD:	7.91 ± 0.76	58 ± 22	67 ± 10	2.62 ± 0.05
			p10/50/90:	6.90/7.89/8.91	27/60/85	54/68/79	2.56/2.62/2.68

Facies 0-3 have relatively high gamma-ray (GR), high sonic (DT), and low bulk density (DEN), indicative of clastic and possible organic-rich interval or hydrocarbon-bearing zone, with Facies 2 and 3 being richer in organic content than Facies 0 and 1. Compared with Facies 1-3, Facies 0 is likely coarser grain siltstone, indicated by relatively lower GR values. Facies 0, 1, 2, and 3 are interpreted as predominantly siltstone, mixed argillaceous mudstone, siliceous mudstone, and calcareous-argillaceous mudstone, respectively, based on their photoelectric factor (PEF) values. This interpretation is supported by DEN and DT values, as siliceous mudstone has lower DEN and higher DT values than calcareous-argillaceous mudstone.

Facies 4, 5, and 6 are dominated by low GR, low DT, high DEN, and high PEF, indicating carbonate-dominated lithofacies. Facies 4 is interpreted as limestone-dominated lithofacies, while Facies 5 is interpreted as mixed carbonate-siliciclastic as the log values are between typical limestone log characteristics (e.g., Facies 4) and siliciclastic lithofacies (e.g., Facies 0). Facies

6 is unique due to its high PEF value, indicating heavy minerals associated with carbonate (e.g., ankerite).

3.4.2.2. Comparison with Other Studies and Dataset

The interpreted lithofacies from this study are compared with the interpretation from other studies based on core and spectral GR data (dataset).

3.4.2.2.1. Well-to-Well Lithofacies Comparison with Another Study: Wolfcamp A as an Example

Figure 11 shows KMeans clustering-derived lithofacies compared with core data derived from other studies (Bievenour and Sonnenberg, 2019). Both interpreted facies indicate that Wolfcamp A at this well-location is predominantly siliceous mudstone with some interbedded siltstone and minor carbonate. Note that the core well location is from a nearby well. Therefore, there might be vertical heterogeneity variability between wells. Another vertical variability difference is possibly caused by differences in depth measurement scale between well-log (every 0.5 ft) and core data (centimeters).

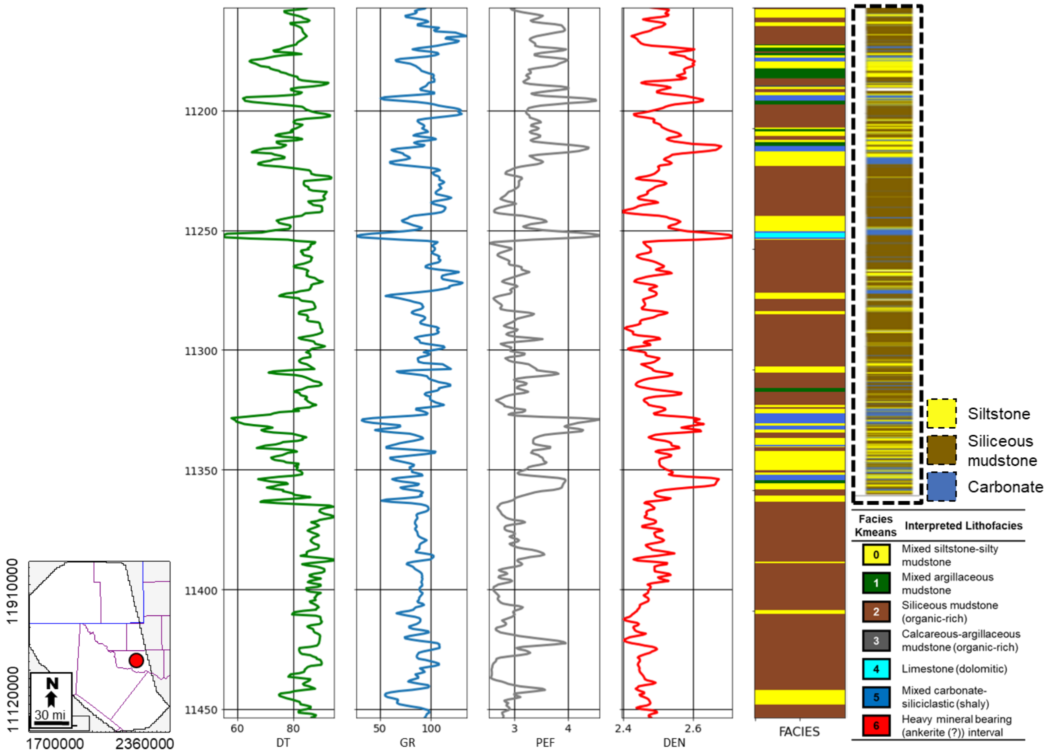


Figure 3.11. Comparison of interpreted lithofacies (Wolfcamp A) from the KMeans facies vs. from the core data from nearby well (dashed-box) derived from Bievenour and Sonnenberg (2019).

3.4.2.2.2. Comparison with Spectral GR Log (Dataset): Wolfcamp A and B as Examples

Figure 3.12 compares the interpreted lithofacies from the KMeans vs. measured spectral gamma-ray from the nearby wells. From the depth plots, mixed argillaceous mudstone is differentiated from siliceous mudstone by the uranium content, with the latter having high uranium content, primarily associated with a reducing environment. This supports the interpretation of siliceous mudstone being richer in organic matter than the mixed argillaceous mudstone.

Clay minerals identified in mixed argillaceous mudstone and siliceous mudstone lithofacies are predominantly mixed-layer clay and illite (**Figure 3.12** - thorium vs. potassium plots). This clay mineral identification supports the interpretation of a silica-rich lithology for the low PEF lithofacies (e.g., siliceous mudstone), as minor montmorillonite (PEF ~2.0) or kaolinite (PEF ~1.8) is identified in the plot which may confuse the interpretation with silica-rich minerals (PEF ~1.8).

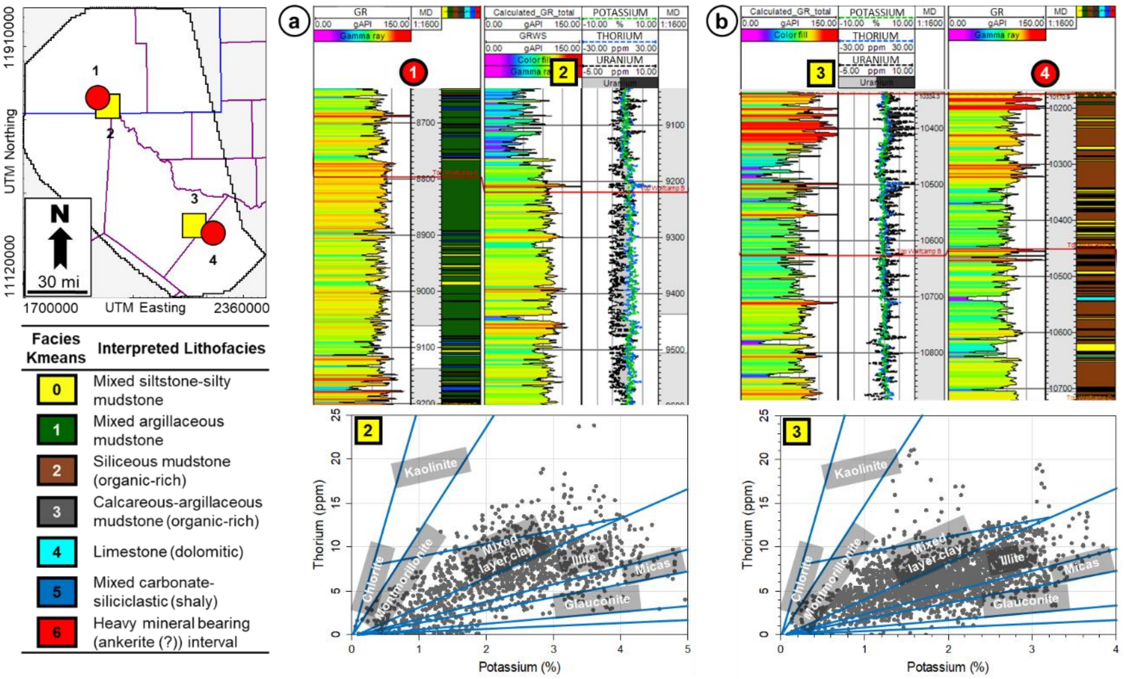


Figure 3.12. Comparison of the interpreted lithofacies (Wolfcamp A & B) from the KMeans vs. measured spectral gamma-ray from the nearby wells. (a) mixed argillaceous mudstone dominated well in the northwest area; (b) siliceous mudstone dominated well in the southeast area; Upper: depth plot highlighting the high uranium content in (b); Lower: clay mineral identification plot showing predominantly mixed-layer clay and illite with minor montmorillonite and kaolinite, indicating low PEF values (PEF~2.0) are associated to silica minerals.

3.4.2.2.3. Basin-Wide Distribution Comparison with Interpreted Lithofacies from Other Studies

Figure 3.13 shows the distribution of the interpreted lithofacies at different well locations in the basin from both KMeans and other studies based on core data of the Wolfcamp A interval (Bievenour and Sonnenberg, 2019; Colborne and Sonnenberg, 2019; Jones, 2019; Kvale et al., 2020; Thompson et al., 2018). Both interpretations indicate that the northern basin margin is limestone-dominated, while more siliceous mudstone is encountered in the central to southern parts of the basin. However, uncertainty exists in lithofacies heterogeneity due to different definitions of “Wolfcamp A” among various studies and this study.

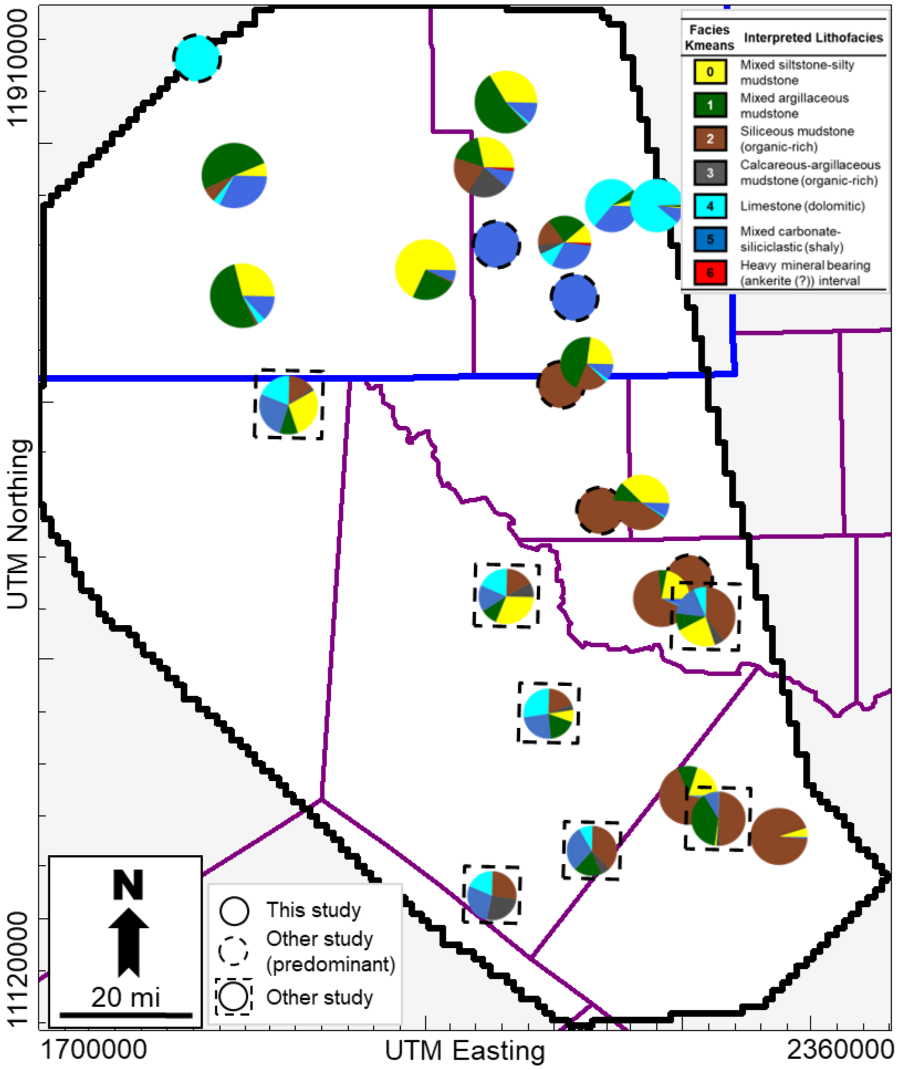


Figure 3.13. Basin-wide lithofacies distribution of Wolfcamp A, both from this study (KMeans) and other studies based on core-data (dashed circle and dashed square).

3.4.2.3. From Basin-Wide Lithofacies to Depositional Environments and Organofacies

3.4.2.3.1. Wolfcamp A Gross-Depositional Environments (GDE) and Organofacies

Figure 3.14 is the interpreted lithofacies and depositional environment map based on this workflow and other studies and the interpreted organofacies from the geochemical data (Baskoro et al., 2021) for the Wolfcamp intervals. With the formation thickness as a reference, the interpreted lithofacies-organofacies (or GDE) of the Wolfcamp Play is generally divided into the shallow marine, shallow-to-deeper marine transition, and deep marine facies. In reference to the facies maps by Blakey (2019), the shallow marine facies is equivalent to the restricted shelf and open marine shelf facies, while the shallow-to-deeper marine transition and deep marine facies are equivalent to the deep marine slope to shallow-to-deeper marine transition and deep basin marine facies, respectively.

The shallow marine facies is dominated by limestone lithofacies. The shallow-to-deeper marine transition facies is dominated by mixed argillaceous mudstone lithofacies with some contribution from other lithofacies, especially non-shale, depending on the geographic location in the basin. The deeper marine facies is dominated by siliceous mudstone lithofacies with some contribution from other lithofacies depending on the geographic location in the basin. The variability of the lithofacies can be explained by multiple sediment sources (**Figure 3.14**) surrounding the Delaware Basin during Wolfcamp deposition (Blakey, 2019). **Figure 3.15** shows the variability of the lithofacies in the W-E cross-section in the northern part of the basin, from shallow-to-deeper marine transition facies to shallow marine facies in the east margins. **Figure 3.16** is the N-S cross-section in the eastern part of the basin, showing a depositional environment from shallow-to- deeper marine transition to deep marine facies and lithofacies from mixed argillaceous mudstone to siliceous mudstone dominated.

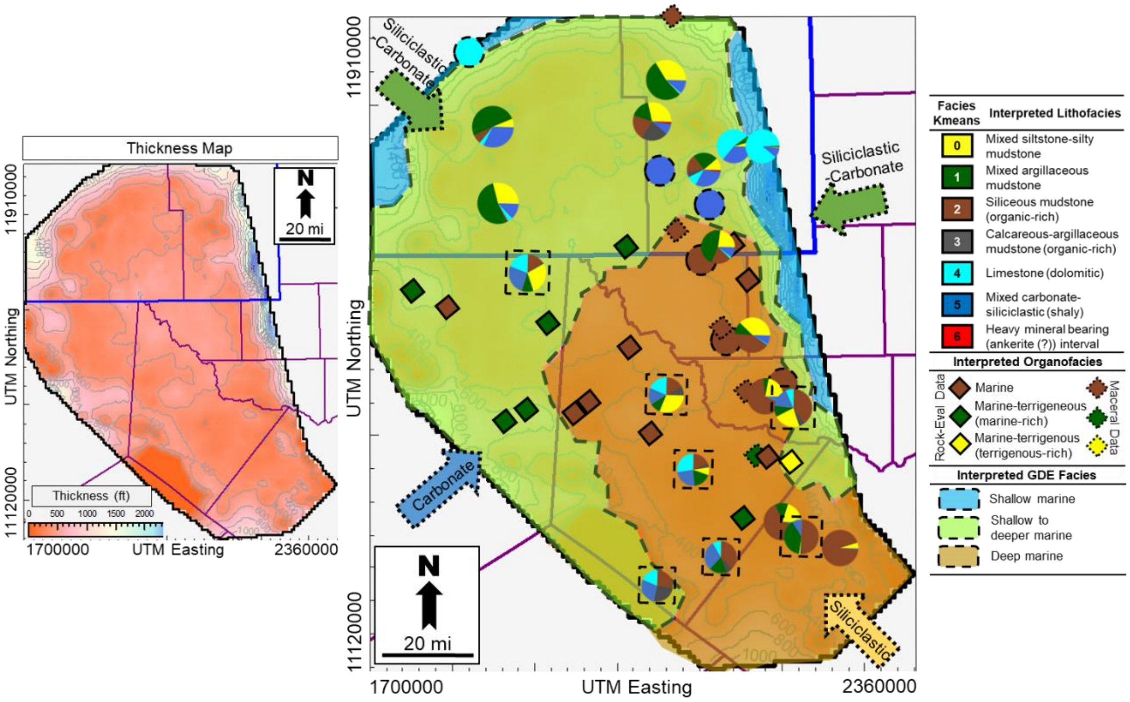


Figure 3.14. Interpreted gross-depositional environments based on litho- and organo- facies distribution and formation thickness of Wolfcamp A interval: marine shelf, transitional shallow to deep marine, and deep marine.

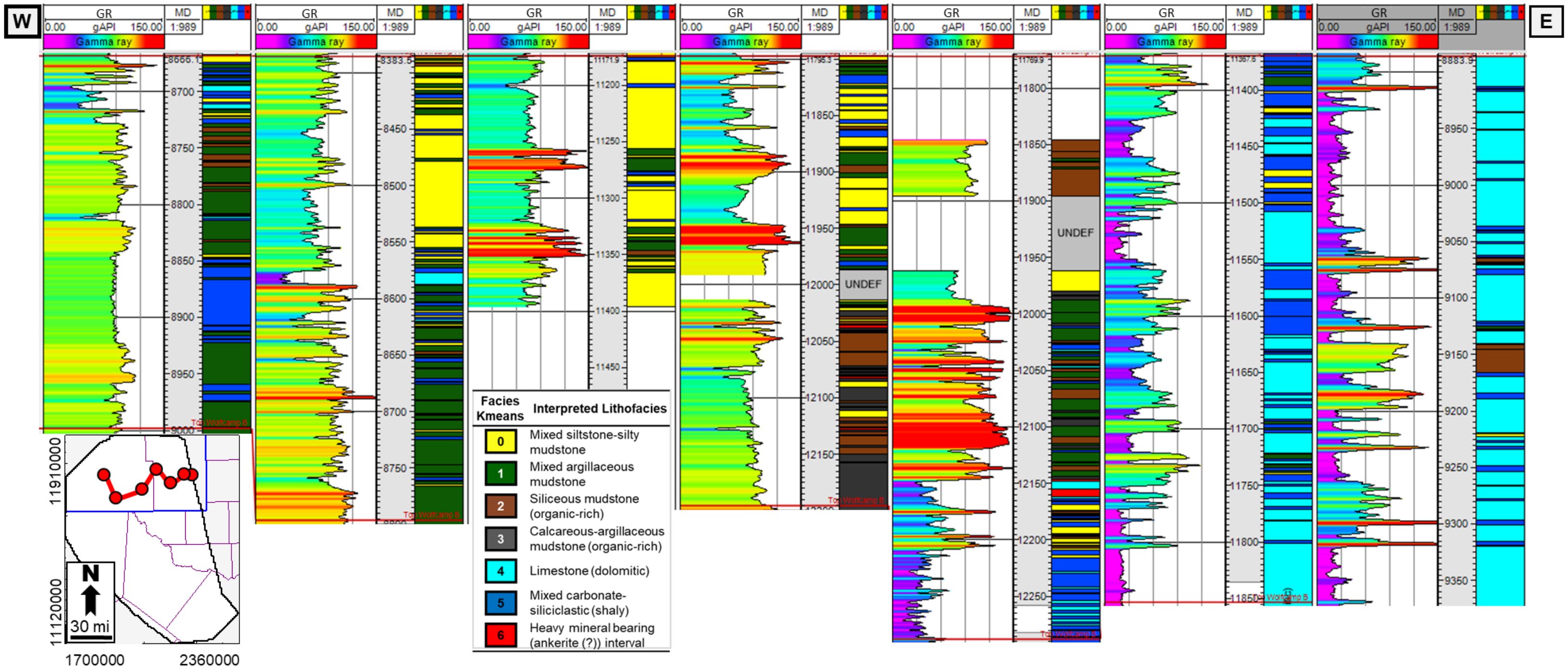


Figure 3.15. West-East northern cross-section of interpreted lithofacies distribution of Wolfcamp A interval, showing the carbonate dominated marine shelf in the east margin.

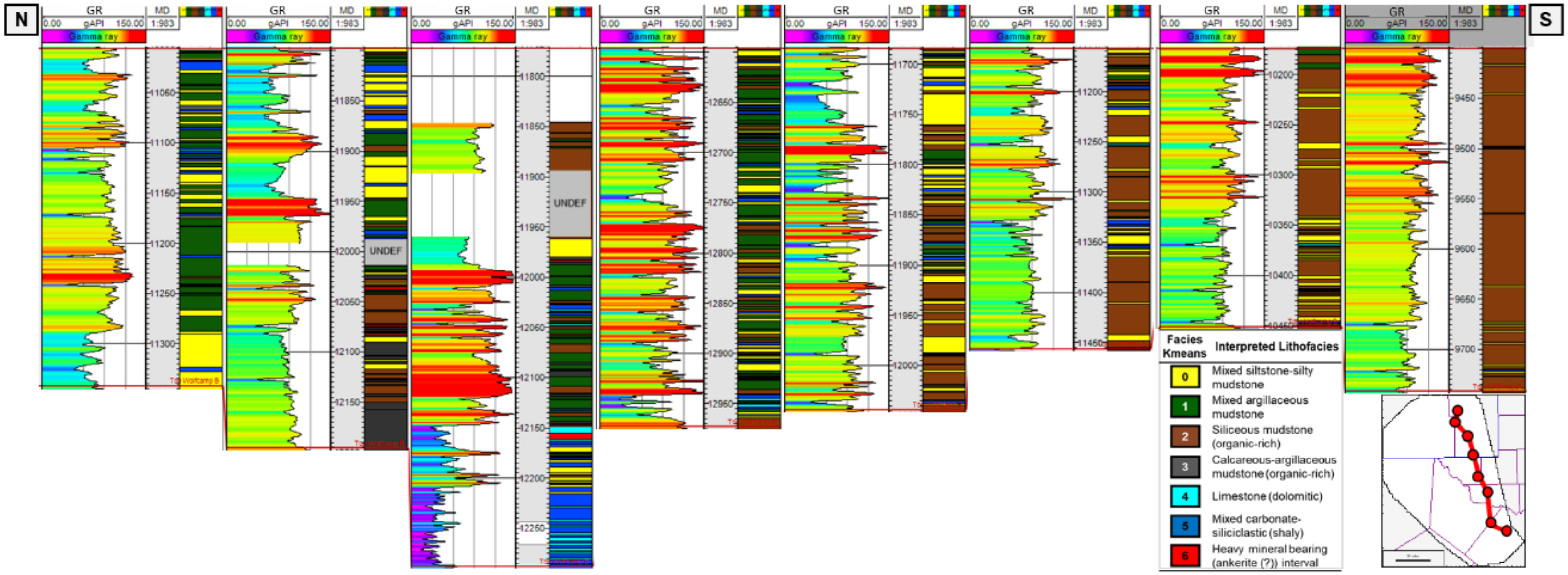


Figure 3.16. North-South eastern cross-section of interpreted lithofacies distribution of Wolfcamp A interval, showing the GDE change from shallower to deeper marine and lithofacies change from predominantly mixed argillaceus mudstone to siliceous mudstone.

Each GDE facies (**Figure 3.14**) is generally associated with organofacies. A detailed fundamental concept of GDE, lithofacies, and their association to organofacies are illustrated in Donovan et al. (2017), Evenick (2016), and Pepper and Corvi (1995). The shallow-to-deeper marine transition facies is predominantly mixed marine-terrigenous organofacies (OF “D/E’-“B”) with mixed argillaceous mudstone and coarser grain lithofacies. Marine carbonate-marl (OF “A”) or mixed marine carbonate-clastic (OF “A/B”) may also be encountered adjacent to shallow marine facies associated with limestone or calcareous mudstone lithofacies. The deep marine facies is dominated by marine clastic organofacies associated with siliceous mudstone (OF “B”). Mixed marine-terrigenous organofacies (OF “D/E”) may also be found, especially adjacent to the eastern basin margin transported through sedimentation from the Central Basin Platform (Kvale et al., 2020). Additionally, Bievenour and Sonnenberg (2019) concluded that siliceous mudstone lithofacies are richer in marine organic matter than the siltstone facies based on their backscatter and secondary electron observations. Even though no Rock-Eval or maceral data are available for the shallow marine facies, most oil biomarkers (upper Wolfcamp sourced-oil) indicated a marine carbonate-marl source rock or OF “A” (Baskoro et al., 2021), associated with limestone and mixed carbonate-siliciclastic lithofacies.

3.4.2.3.2. Wolfcamp B-C-D Intervals GDE and Organofacies

The GDE maps for Wolfcamp B, C, and D intervals are reconstructed in **Figure 3.17**. In general, each Wolfcamp interval has a similar overall GDE and litho- and organo- facies trend but with a lateral variability reflecting lateral changes of the depositional environment. The lateral variability (e.g., from north to south) includes deepening in relative water depth, changing lithofacies from mixed argillaceous to siliceous mudstone lithofacies dominated, and changing organic matter richness from relatively poorer to richer.

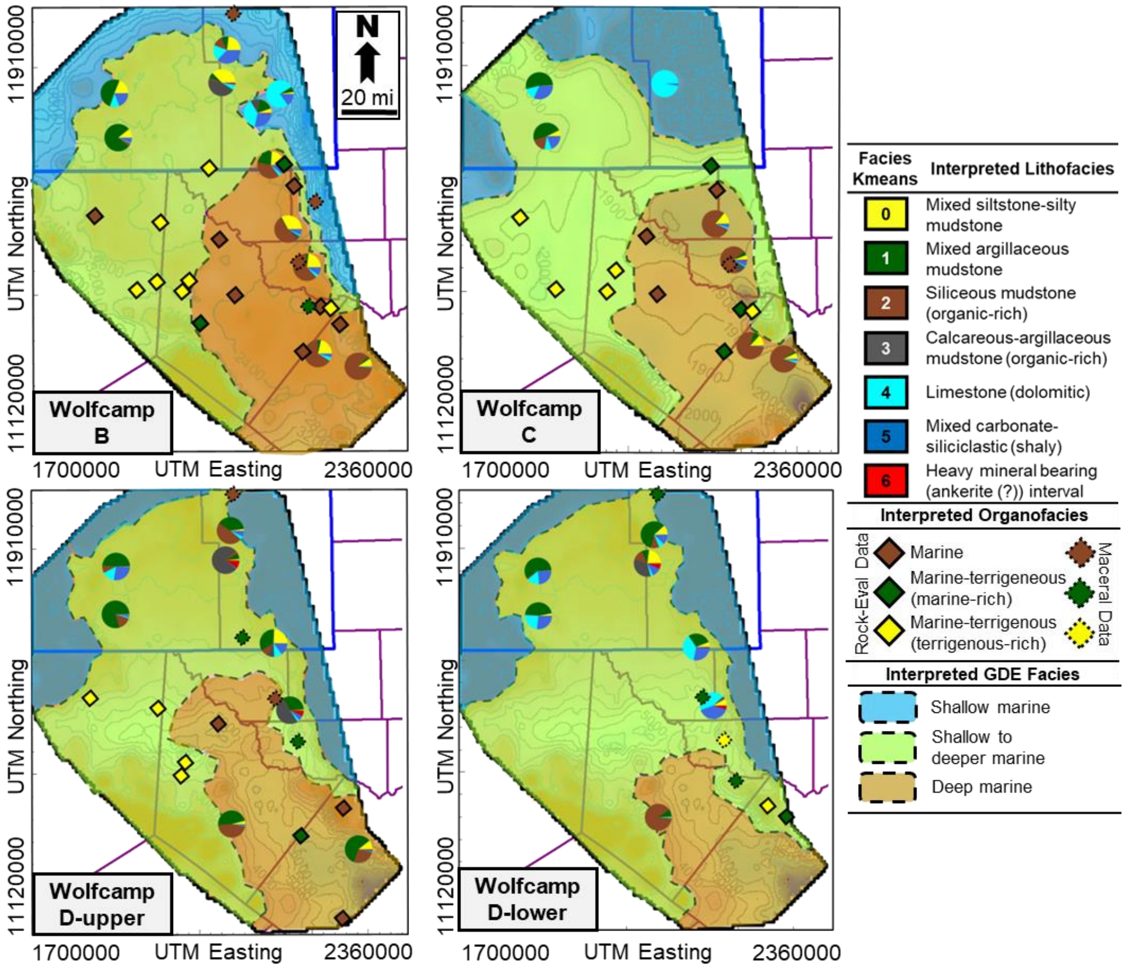


Figure 3.17. Interpreted gross-depositional environment and litho- and organo- facies distribution of Wolfcamp B-C-D (upper and lower) intervals.

Vertical variability is also spotted among the different Wolfcamp intervals (**Figure 3.18**). One of the significant variabilities is that the mixed marine-terrigenous organofacies of the shallow-to-deeper marine transition facies of Wolfcamp B-C-D intervals are richer in terrigenous organic matter than Wolfcamp A interval's shallow-to-deeper marine transition facies. Another vertical variability is observed, including the changes in the predominant litho- and organo-facies from older to younger intervals (e.g., from Wolfcamp D to Wolfcamp A), reflecting the changes in relative water depth and overall depositional environments, from relatively shallower to deeper marine.

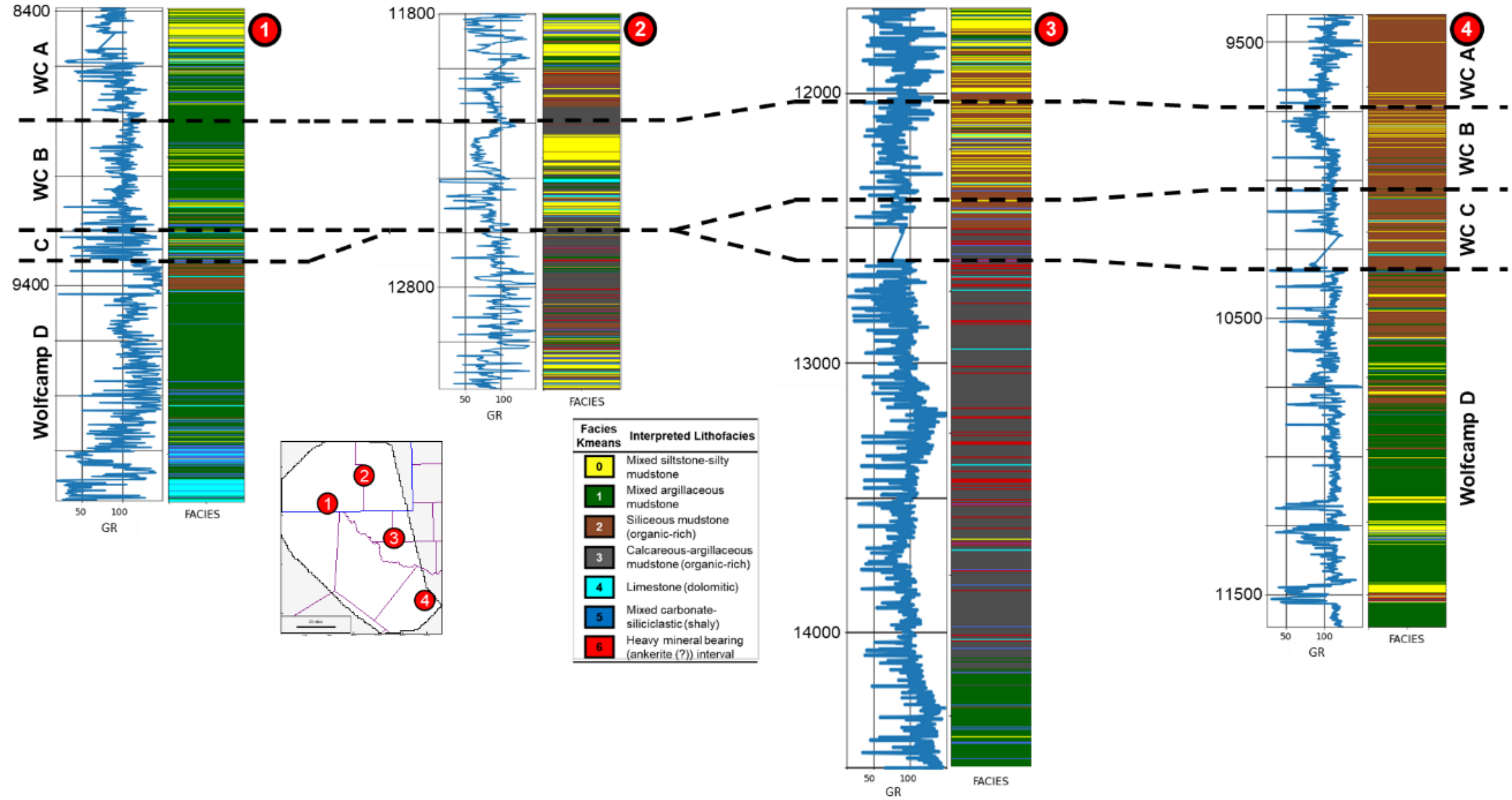


Figure 3.18. Wolfcamp A-B-C-D intervals heterogeneity across different well-location in the Delaware Basin.

3.5. Discussion

3.5.1. Data Challenge and Pre-Processing Procedure

There are challenges in developing the pre-processing workflow. Firstly, the large number of well-data requires automation for each pre-processing stage. Secondly, well-log aliases in the dataset are inconsistent due to different operators (e.g., gamma-ray in numerous aliases such as GR, GRWS, GR_1, GR1, HCGR, GAMMA (>250 aliases). It is critical to QC the data and filter the best subset of data to carry out the proposed workflow.

3.5.2. Advantages of the Unsupervised Machine Learning

As demonstrated in **Figure 3.10** and **Table 3.2**, even though each facies can be distinguished from each other by at least one type of well-log measurement, the other types of well-log measurement are sometimes blurry or partially overlapping (e.g., PEF and GR values between Facies 0 and 5), which complicates the interpretation. Additionally, the lithofacies interpretation is based on a broad range of well-log measurements (e.g., between P90 and P10 values or 80% of data), which is assumed to represent the facies signature. The narrower the cluster is, e.g., the smaller the range between P90 and P10 values, the easier and more accurate the interpretation.

Despite the limitations and challenges, there are several advantages of using unsupervised machine learning methods to classify lithofacies from well-log data. Firstly, the workflow is data-driven and straightforward; no human interference is necessary. The data-driven workflow suggests the number of clusters once the appropriate input data is compiled as a DataFrame. Only four types of well-log measurements (PEF, GR, DT, and DEN) are necessary to yield meaningful clusters of lithofacies in this geological case study. The dataset must cover the spatial distribution of the study area to yield representative clusters. Only 31 out of 1500 wells are used in this study to draw the basin-wide GDE. Secondly, the workflow consistently captures lateral and vertical lithofacies variability, as demonstrated with the Wolfcamp Play dataset in the Delaware Basin. Distinct basin-wide lithofacies trends are recognized geographically across the basin, even though only a subset of data is used in this workflow.

3.6. Conclusion

Seven basin-wide lithofacies are identified within the Wolfcamp Play dataset by applying two levels of unsupervised machine learning workflow. The lithofacies interpretation concludes that there are three predominant facies on the gross depositional environments for the Wolfcamp intervals. The shallow marine facies is dominated by limestone lithofacies with marine carbonate/ marl organofacies ”A” and “D/E” sources. The shallow-to-deeper marine transition facies is generally dominated by mixed argillaceous mudstone and other lithofacies with mixed marine-terrigenous organofacies “D/E” sources. The deep marine facies is dominated by siliceous mudstone and other lithofacies with marine clastic organofacies “B” sources. The vertical and lateral variability of depositional environments and lithofacies are recognized throughout the different Wolfcamp intervals, e.g., from both north to south and older to younger are deepening in relative water depth, changing in lithofacies from mixed argillaceous to siliceous mudstone lithofacies, and changing in organic matter richness from relatively poorer to richer.

Because the procedure is a data-driven workflow, proper pre-processing is necessary to handle the large basin-wide dataset, well-log aliases, well-log scale, and computational resources. Even though most clusters are compact and well-separated, few clusters have an overlapping well-log signature, making the interpretation challenging.

The proposed two-level unsupervised machine learning workflow is data-driven and straightforward. Using only four types of well-log measurements (PEF, GR, DT, and DEN) and selecting wells from a large dataset, the workflow can consistently capture the variability of the lithofacies that outlines the basin-wide geological variations.

3.7. References

Adams, J. E., 1965, Stratigraphic-tectonic development of Delaware Basin: AAPG bulletin, v. 49, p. 2140-2148.

Alloghani, M., D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf, 2020, A systematic review on supervised and unsupervised machine learning algorithms for data science: Supervised unsupervised learning for data science, p. 3-21.

Baskoro, A. S., F. Baur, A. Z. Yu, and E. L. Grossman, 2021, Source rock restoration using a multi-well inversion approach for the Wolfcamp Play in the Permian Delaware Basin of west Texas and New Mexico: AAPG/SEG International Meeting for Applied Geoscience & Energy.

Baur, F., 2019, Predicting petroleum gravity with basin modeling: New kinetic models: AAPG Bulletin, v. 103, p. 1811-1837.

Bievenour, A., and S. Sonnenberg, 2019, Reservoir characterization of the Bone Spring and Wolfcamp Formations, Delaware Basin, Ward County, West Texas: Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July 2019, p. 891-906.

Blakey, R., 2019, Colorado Plateau Geosystems. Inc., Greater Permian Basin Series: https://deeptimemaps.com/greater-permian-basin/ *(accessed on May 4th 2022)*

Caliński, T., and J. Harabasz, 1974, A dendrite method for cluster analysis: Communications in Statistics-theory and Methods v. 3, p. 1-27.

Chakravarty, A., and S. Misra, 2022, Unsupervised learning from three-component accelerometer data to monitor the spatiotemporal evolution of meso-scale hydraulic fractures: International Journal of Rock Mechanics and Mining Sciences, 151, 105046.

Chang, H.-C., D. C. Kopaska-Merkel, and H.-C. Chen, 2002, Identification of lithofacies using Kohonen self-organizing maps: Computers & Geosciences, v. 28, p. 223-229.

Colborne, J., and S. Sonnenberg, 2019, Reservoir facies, depositional processes, and the implications on reservoir characterization of the Wolfcamp A, Texas Delaware Basin: Unconventional Resources Technology Conference, Denver, Colorado, 22-24 July 2019, p. 880-890.

Coutinho, L., H. d. B. Penteado, F. Lorant, and J. Rudkiewicz, 2009, A compositional petroleum mass balance in a mature zone for exploration-Reconcavo Basin, Brazil.

Davies, D. L., and D. W. Bouldin, 1979, A cluster separation measure: IEEE Transactions on pattern analysis and machine intelligence, p. 224-227.

di Primio, R., and B. Horsfield, 2006, From petroleum-type organofacies to hydrocarbon phase prediction: AAPG bulletin, v. 90, p. 1031-1058.

Donovan, A. D., J. Evenick, L. Banfield, N. McInnis, and W. Hill, 2017, An organofacies-based mudstone classification for unconventional tight rock & source rock plays: SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, 24-26 July 2017, p. 3683-3697.

Dutton, S. P., Kim, E. M., Broadhead, R. F., Raatz, W. D., Breton, C. L., Ruppel, S. C., & Kerans, C. 2005, Play analysis and leading-edge oil-reservoir development methods in the Permian basin: Increased recovery through advanced technologies: AAPG Bulletin, v. 89 (5), p. 553-576.

Edwards, A. W., and L. L. Cavalli-Sforza, 1965, A method for cluster analysis: Biometrics, p. 362-375.

EIA, U. S., 2019, Permian Basin Wolfcamp and Bone Spring shale plays geology review: U.S. Energy Information Administration Report.

Evenick, J. C., 2016, Evaluating source rock organofacies and paleodepositional environments using bulk rock compositional data and pristane/phytane ratios: Marine and Petroleum Geology, v. 78, p. 507-515.

Fairhurst, B., T. Ewing, and B. Lindsay, 2021, West Texas (Permian) Super Basin, United States: Tectonics, structural development, sedimentation, petroleum systems, and hydrocarbon reserves: AAPG Bulletin, v. 105, p. 1099-1147.

Falola, Y., S. Misra, J. Foster, and M. Bhatia, 2022, Data-driven workflow for the preemptive detection of excess water producing wells drilled in unconventional shales: Journal of Natural Gas Science and Engineering, 104609.

Gaswirth, S. B., K. L. French, J. K. Pitman, K. R. Marra, T. J. Mercier, H. M. Leathers-Miller, C. J. Schenk, M. E. Tennyson, C. A. Woodall, and M. E. Brownfield, 2018, Assessment of undiscovered continuous oil and gas resources in the Wolfcamp Shale and Bone Spring Formation of the Delaware Basin, Permian Basin Province, New Mexico and Texas, 2018, US Geological Survey.

Gentleman, R., and V. J. Carey, 2008, Unsupervised machine learning, Bioconductor case studies, Springer, p. 137-157.

Gonzalez, K., and S. Misra, 2022, Unsupervised learning monitors the carbon-dioxide plume in the subsurface carbon storage reservoir: Expert Systems with Applications, 201, 117216.

Hartigan, J. A., and M. A. Wong, 1979, Algorithm AS 136: A k-means clustering algorithm: Journal of the royal statistical society. series c (applied statistics), v. 28, p. 100-108.

He, J., H. Li, and S. Misra, 2019, Data-driven in-situ sonic-log synthesis in shale reservoirs for geomechanical characterization: SPE Reservoir Evaluation & Engineering, 22(04), 1225-1239.

Hills, J. M., 1984, Sedimentation, tectonism, and hydrocarbon generation in Delaware Basin, west Texas and southeastern New Mexico: AAPG bulletin, v. 68, p. 250-267.

Jarvie, D. M., R. J. Hill, T. E. Ruble, and R. M. Pollastro, 2007, Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment: AAPG bulletin, v. 91, p. 475-499.

Johnson, S. C., 1967, Hierarchical clustering schemes: Psychometrika, v. 32, p. 241-254.

Jones, R., 2019, Nanopetrophysical characterization of the Wolfcamp A shale Formation in the Permian Basin of southeastern New Mexico, USA, The University of Texas at Arlington.

Katz, B. J., I. Arango, and F. Frasse, 2017, Expulsion and migration associated with unconventional petroleum systems: The Houston Geological Society Bulletin, v. Vol. 59, No. 5.

Kodinariya, T. M., and P. R. Makwana, 2013, Review on determining number of Cluster in K-Means Clustering: International Journal of Advance Research in Computer Science and Management Studies v. 1, p. 90-95.

Kvale, E. P., C. M. Bowie, C. Flenthrope, C. Mace, J. M. Parrish, B. Price, S. Anderson, and W. A. DiMichele, 2020, Facies variability within a mixed carbonate–siliciclastic sea-floor fan (upper Wolfcamp Formation, Permian, Delaware Basin, New Mexico): AAPG Bulletin, v. 104, p. 525-563.

Li, H., and S. Misra, 2018, Assessment of miscible light-hydrocarbon-injection recovery efficiency in Bakken shale formation using wireline-log-derived indices: Marine and Petroleum Geology, 89, 585-593.

Ng, A., M. Jordan, and Y. Weiss, 2001, On spectral clustering: Analysis and an algorithm: Advances in neural information processing systems, v. 14.

Pepper, A. S., and P. J. Corvi, 1995, Simple kinetic models of petroleum formation. Part I: oil and gas generation from kerogen: Marine and petroleum geology, v. 12, p. 291-319.

Raeesi, M., A. Moradzadeh, F. D. Ardejani, and M. Rahimi, 2012, Classification and identification of hydrocarbon reservoir lithofacies and their heterogeneity using seismic attributes, logs data and artificial neural networks: Journal of Petroleum Science & Engineering, v. 82, p. 151-165.

Rousseeuw, P. J., 1987, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis: Journal of computational applied mathematics, v. 20, p. 53-65.

Silver, B. A., and R. G. Todd, 1969, Permian cyclic strata, northern Midland and Delaware basins, west Texas and southeastern New Mexico: AAPG Bulletin, v. 53, p. 2223-2251.

Thompson, M., P. Desjardins, J. Pickering, and B. Driskill, 2018, An integrated view of the petrology, sedimentology, and sequence stratigraphy of the Wolfcamp Formation, Delaware Basin, Texas: SPE/AAPG/SEG Unconventional Resources Technology Conference.

Wang, G., T. R. Carr, Y. Ju, and C. Li, 2014, Identifying organic-rich Marcellus Shale lithofacies by support vector machine classifier in the Appalachian basin: Computers & Geosciences, v. 64, p. 52-60

4. MASS BALANCE OF GENERATION, RETENTION, AND PRODUCTION FOR THE WOLFCAMP-SOURCED HYDROCARBON IN PERMIAN DELAWARE BASIN: INSIGHT ON REMAINING RECOVERABLE RESOURCE AND EXPULSION EFFICIENCY

4.1. Introduction

Improved efficiency of hydrocarbon resources exploration is directly linked to an accurate analysis of the petroleum system in a basin. An accurate petroleum system analysis requires detailed quantification of hydrocarbon (HC) masses involved in a basin from generation to production (Hantschel and Kauerauf, 2009a). The quantified HC masses must be balanced following the fundamental laws of mass conservation. This is particularly important for a basin with unconventional-conventional petroleum systems where the generated HCs from the source rock interval are either expelled out from or retained within the source rock interval. Expulsion efficiency becomes even more critical as it largely determines the HCs in place, both for the source rock interval or out of the source rock accumulations (Arango and Katz, 2019; Bai et al., 2017).

Several authors have attempted to perform HC mass balance calculations at both basin and field scales (e.g., (Baur et al., 2011; Coutinho, 2008; Duran et al., 2013; Lewan et al., 2002; Muscio et al., 2016). They calculated the HC mass balance using basin and petroleum system modeling (Baur et al., 2011; Coutinho, 2008; Duran et al., 2013), stochastic/ probabilistic methods (Lewan et al., 2002), or comparison of both methods (Muscio et al., 2016). Most of these studies aimed to understand the accumulation efficiency for conventional HC accumulation, calibrated to production data from the conventional field. Expulsion efficiency and retained HC portion within the source rock interval were not discussed in depth.

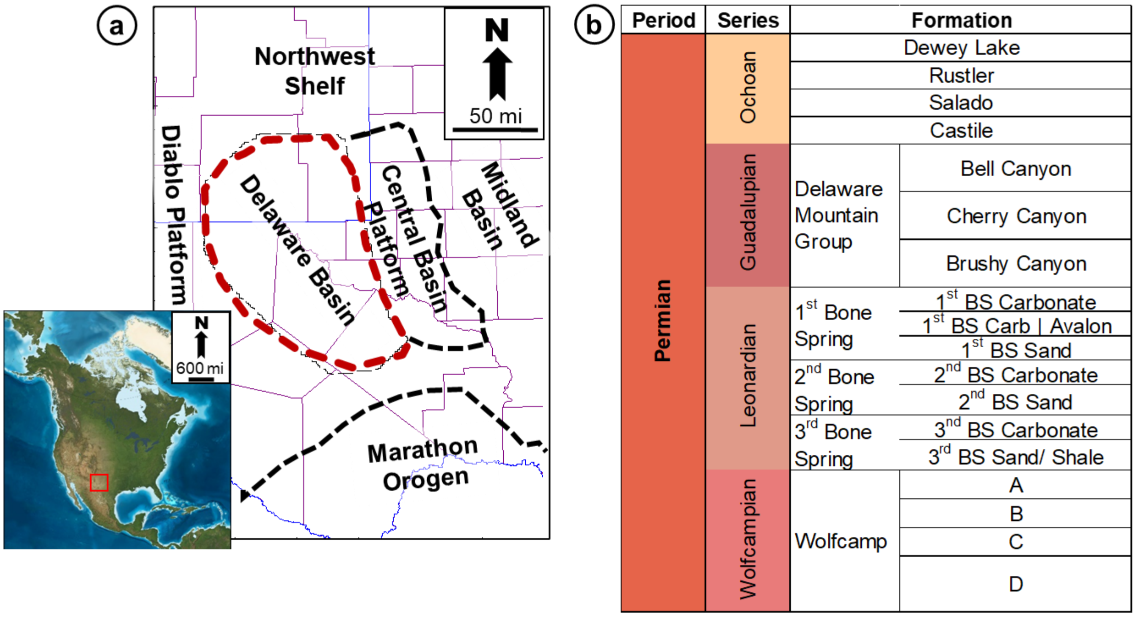
Numerous studies have been performed on modeling or assessing the mass balance of unconventional HC accumulations. For example, Rushing et al. (2004) and Kuchinskiy et al. (2012) calculated the HC volumes of tight-rock and source rock plays, respectively, by employing a stochastic/ probabilistic approach using a modified HC generation potential formula from Schmoker (1994). HC volumes quantification focused on the source rock, with expulsion efficiency calculated using several assumptions regarding retaining factor and fractions of HCs lost from the system, while source rock heterogeneity was generally simplified. However, the results were not assessed for their mass balance for the whole petroleum system.

Other examples are from basin and petroleum system modeling studies applied to the unconventional system, including Amer et al. (2015); Bryant et al. (2013); Cander (2012); Jarvie et al. (2007); Romero-Sarmiento et al. (2013), among many others. Some of these studies addressed the challenge of modeling unconventional systems, including factors controlling the retained HC storage capacity and fluid properties. However, the calculated retained and expelled HCs were not assessed for their mass balance.

Thus, this study aims to fill these gaps by performing HC mass balance calculations using several calculation scenarios, including basin and petroleum system modeling. The proposed workflow of mass balance calculation aims to achieve the distribution of the generated HC from the source rock, considering the SR heterogeneity in an unconventional system. This calculation assesses the mass balance to estimate the remaining recoverable and derive expulsion efficiency. The multiple calculation scenarios of mass balance also aim to establish a reference or comparison for modeling the unconventional system.

4.2. Geological Overview

The case study for this HC mass balance calculation is the Wolfcamp Play in the Permian Delaware Basin of West Texas and New Mexico, USA (**Figure 4.1a**). The Delaware Basin is bounded to the west by the Diablo Platform, to the north by the North American craton and Northwest Shelf areas, and to the south by the Marathon orogenic belt. The Delaware Basin is separated from the Midland Basin to the east by the N-S trending Central Basin Platform (Adams, 1965; Dutton et al., 2005; Hills, 1984; Silver and Todd, 1969). Delaware Basin occupies a basinal area of 13,000 mi^2 (33,500 km^2) with a maximum depth of 24,000 ft (7,300 m) and is filled by up to 40,000 mi^3 (170,000 km^3) of Phanerozoic sediments (Hills, 1984).



Period	Series	Formation	
Permian	Ochoan	Dewey Lake	
		Rustler	
		Salado	
		Castile	
	Guadalupian	Delaware Mountain Group	Bell Canyon
			Cherry Canyon
			Brushy Canyon
	Leonardian	1st Bone Spring	1st BS Carbonate
			1st BS Carb | Avalon
			1st BS Sand
		2nd Bone Spring	2nd BS Carbonate
			2nd BS Sand
		3rd Bone Spring	3nd BS Carbonate
			3rd BS Sand/ Shale
	Wolfcampian	Wolfcamp	A
			B
			C
			D

Figure 4.1. (a) Major subdivision and boundaries of the Permian Basin around the Delaware Basin (after Dutton et al. (2015) and Silver and Todd (1969)). (b) The stratigraphic chart of Permian strata, which includes Wolfcamp interval subdivision for the Delaware Basin.

The Wolfcamp Formation (Fm) is part of the thick Permian strata in the Delaware Basin and consists of four main intervals: Wolfcamp A, -B, -C, and -D (**Figure 4.1b**). Wolfcamp sediments were deposited during the early Permian in a generally deep, oceanic-connected interior continental basin surrounded by carbonate shelf (Northwest Shelf, eastern Central Basin Platform, western Diablo Platform), the northwestern Pedernal uplift, and the southern siliciclastic shelf in front of the Marathon orogenic belt (Blakey, 2019). This depositional setting resulted in litho- and organo-facies heterogeneity within the Wolfcamp intervals in the Delaware Basin. Wolfcamp sediments were deposited during the frequent eustatic fluctuations that cause lithofacies heterogeneity, primarily limestone, silts, and shales (Fairhurst et al., 2021).

The Wolfcamp Play is one of the main targets for unconventional production in the Permian Delaware Basin, along with the Leonardian Bone Spring interval (Gaswirth et al., 2018). Among the four intervals, Wolfcamp A is the most drilled target zone for unconventional play in the Delaware Basin (Fairhurst et al., 2021; Popova, 2019).

Several studies have performed petroleum system analysis for the Wolfcamp and younger Permian intervals in the Delaware Basin (Curtis and Zumberge, 2017, 2018; Fairhurst et al., 2021; Jarvie, 2017; Justman and Broadhead, 2010a, b). Wolfcamp marine shale facies is the primary source rock for the Wolfcamp unconventional/ hybrid reservoirs and the lower Bone Spring interval (Curtis and Zumberge, 2017, 2018; Fairhurst et al., 2021; Jarvie, 2017; Pepper et al., 2020). The marine carbonate-rich source rock, which is mainly restricted to eastern and northern Delaware Basin margins, is primarily responsible for the HC accumulations in the Delaware Basin margins and as far as the Northwest Shelf and Central Basin Platform (Curtis and Zumberge, 2017, 2018; Pepper et al., 2020). The signatures of these two different oil families are documented in Baskoro et al. (2021a); Curtis and Zumberge (2018); Echegu et al. (2021), and Pepper et al. (2020). Wolfcamp Fm is suitable for mass balance calculation due to its importance as a source rock and unconventional system.

4.3. Workflow and Methods

The HC mass balance procedure involves "bookkeeping" of the calculated HC from generation to production. A 3D basin and petroleum system model (BPSM) is utilized to calculate the HC masses from generation (Hantschel and Kauerauf, 2009b) to expulsion-retention. The HC mass balance calculation is then performed on the calculated generated HC as the starting point with the produced HC data as the endpoint.

4.3.1. 3D Basin and Petroleum System Modeling

The 3D BPSM covers the Delaware Basin area within the basin outline (**Figure 4.1**). The main focus of the mass balance calculation is the Wolfcamp-sourced HC within the Permian strata in the Delaware Basin. It is essential to calibrate the 3D BPSM with thermal maturity indicators (e.g., vitrinite reflectance) and then to calibrate the source rock (Wolfcamp) with measured source rock property (HI and TOC) before calculating the HC mass.

4.3.1.1. Model Building

4.3.1.1.1. 3D Delaware Basin Model

The foundation of the 3D BPSM, including the associated boundary conditions, layer subdivisions, age and facies assignment, and thermal maturity scenario, are documented in Becker et al. (2019) and Yu et al. (2020). **Figure 4.2a** the thermal maturity calibration to measured vitrinite reflectance in different areas of the basin. The thermal maturity at the Wolfcamp intervals is from the oil to gas generation window, with the oil dominating in the Wolfcamp A interval and the gas generation dominating in the Wolfcamp D interval (**Figure 4.2b**). For the reliable HC mass calculation from the Wolfcamp source, the 3D BPSM is built with restored source rock initial TOC, HI, and litho- and organo- facies for the Wolfcamp A, B, C, and D intervals.

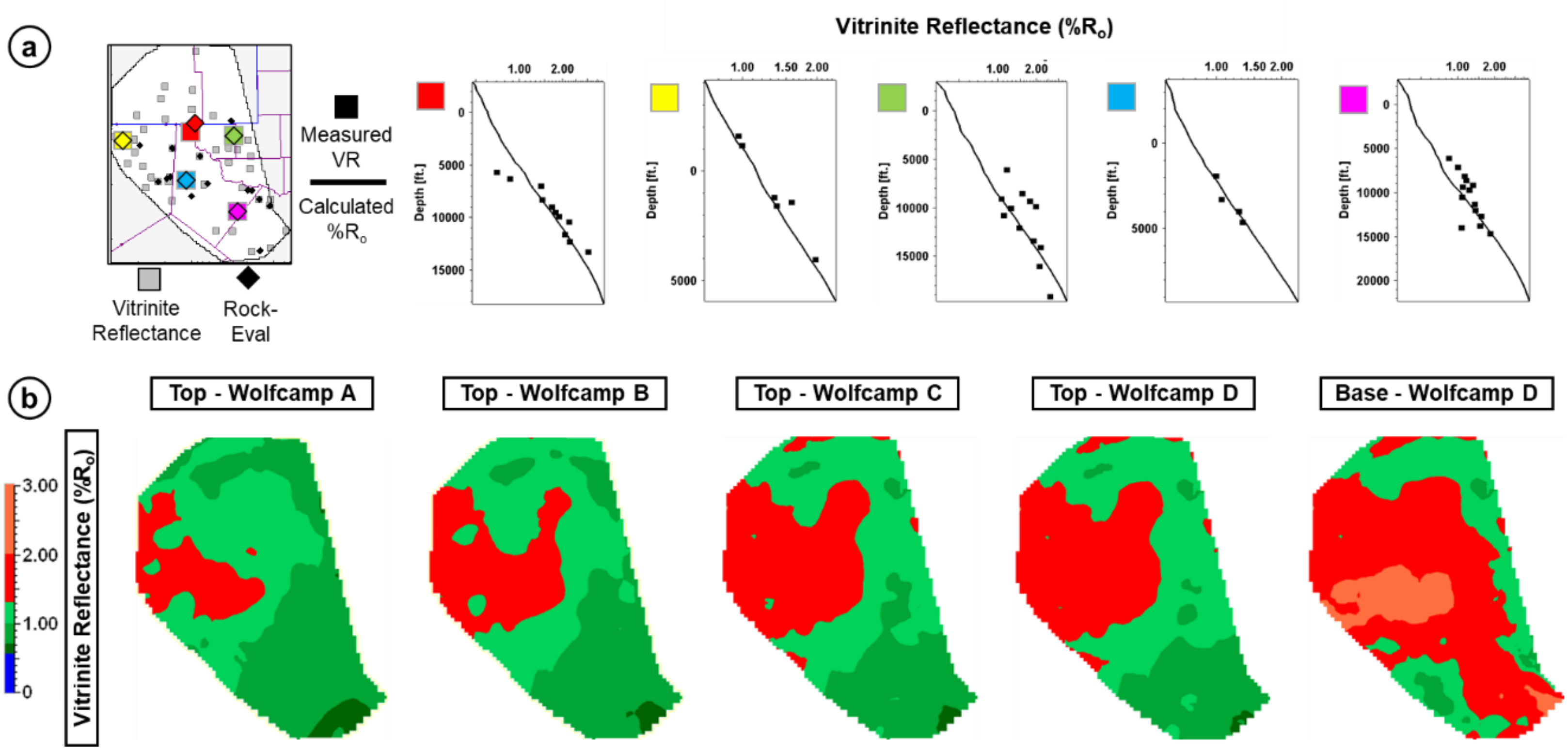


Figure 4.2. Thermal maturity (a) calibration to measured vitrinite reflectance data and (b) model (easy %Ro) shown at top and base of each Wolfcamp interval, indicates peak oil to gas generation window.

4.3.1.1.2. Restored Source Rock Initial HI-TOC and Litho- and Organo-Facies Parameters for the Wolfcamp Formation

The lithofacies variability (**Figure 4.3**) is based on the basin-wide lithofacies study in Baskoro et al. (2021b). Generally, the predominant source rock lithofacies is organic-rich mudstone (more siliceous in the south and more argillaceous in the north) with interbedded siltstone and limestone, reflecting depositional environments from shallow to deep marine (Bievenour and Sonnenberg, 2019; Blakey, 2019; Fairhurst et al., 2021). In the model, siltstone and limestone layers are determined as non-SR, while the mudstone layers are designated SR intervals.

Deep marine or siliceous mudstone facies is predominantly associated with marine clastic OF “B” type kinetic model, while the shallow-to-deeper marine or mixed argillaceous mudstone facies is predominantly associated with mixed marine-terrigenous OF “D/E” type, depending on the initial HI and TOC values (Baskoro et al., 2021a; Baskoro et al., 2021b). Baur (2019) kinetic model is chosen as it is based on the Pepper and Corvi (1995) kinetic model and has updated transformation ratio and fluid properties prediction.

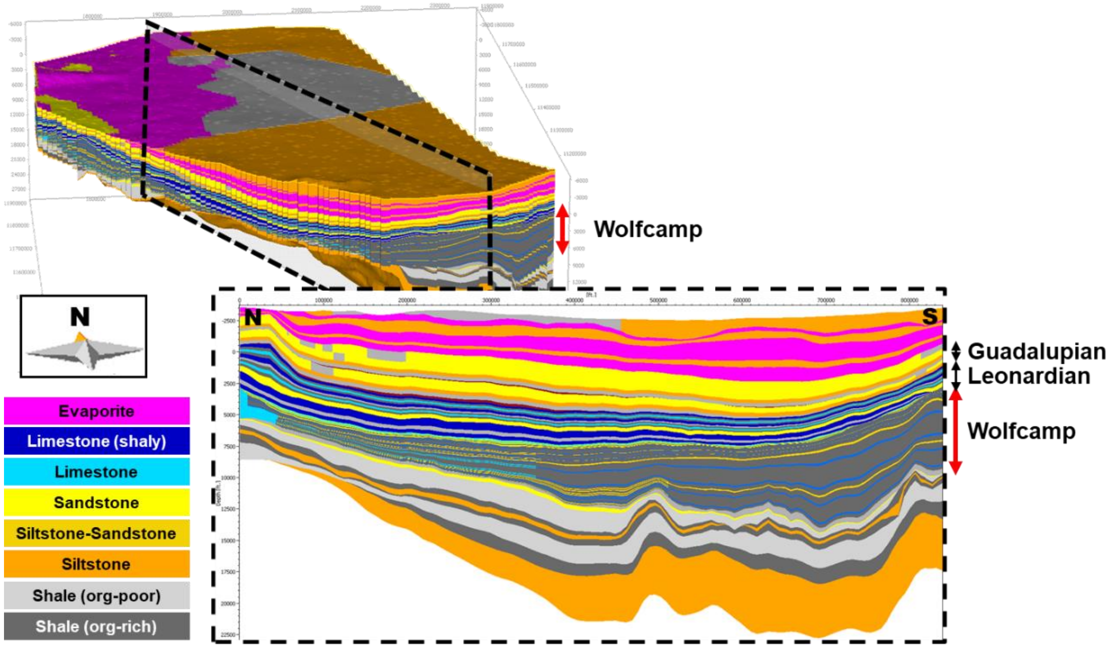


Figure 4.3. 3D basin and petroleum system model of Delaware Basin with enhanced input of variability of Wolfcamp lithofacies.

Table 4.1 shows the p90/50/10 restored initial HI and TOC values adopted in the model, based on the SR restoration study in Baskoro et al. (2021a). For more accurate calculations, each Wolfcamp interval (A, B, C, and D) is subdivided into 2 to 3 layers, where each layer is assigned two initial HI and TOC values for each p90/50/10 model: one value for the shallow-to-deeper marine facies and one value for the deep marine facies (see **Figure 4.4** for the example from the p50 model). Wolfcamp A and B generally have a better initial source rock potential, from very good to excellent, than Wolfcamp C and D.

Table 4.1. P90/50/10 restored initial: HI (upper) and TOC (lower), of Wolfcamp SR intervals adopted in the 3D model.

Initial HI (mgHC/gTOC)		Shallow-to-Deeper Marine			Deep Marine		
		p90	p50	p10	p90	p50	p10
Wolfcamp A	Upper	217	313	579	318	473	647
	Middle	266	288	499	413	552	700
	Lower	200	320	576	423	475	765
Wolfcamp B	Upper	179	207	426	420	444	691
	Middle	213	242	503	407	445	660
	Lower	192	225	348	382	455	532
Wolfcamp C	Upper	167	186	218	403	454	562
	Lower	151	187	202	403	423	573
Wolfcamp D	Upper	150	281	398	307	373	464
	Lower	129	215	407	129	215	407

Initial TOC (mgHC/gTOC)		Shallow-to-Deeper Marine			Deep Marine		
		p90	p50	p10	p90	p50	p10
Wolfcamp A	Upper	1.27	1.66	3.00	1.54	1.93	3.05
	Middle	1.50	1.88	3.24	2.28	2.61	4.19
	Lower	1.66	2.36	3.17	1.64	3.21	5.03
Wolfcamp B	Upper	1.37	1.94	2.15	1.99	2.72	3.97
	Middle	1.51	2.23	3.30	2.10	2.57	3.16
	Lower	1.80	1.95	2.63	1.94	2.48	2.77
Wolfcamp C	Upper	1.44	1.79	2.66	1.94	2.22	3.27
	Lower	1.38	2.09	2.45	1.45	2.00	2.92
Wolfcamp D	Upper	1.85	2.28	2.82	1.48	1.87	2.40
	Lower	1.24	1.67	2.18	1.24	1.67	2.18

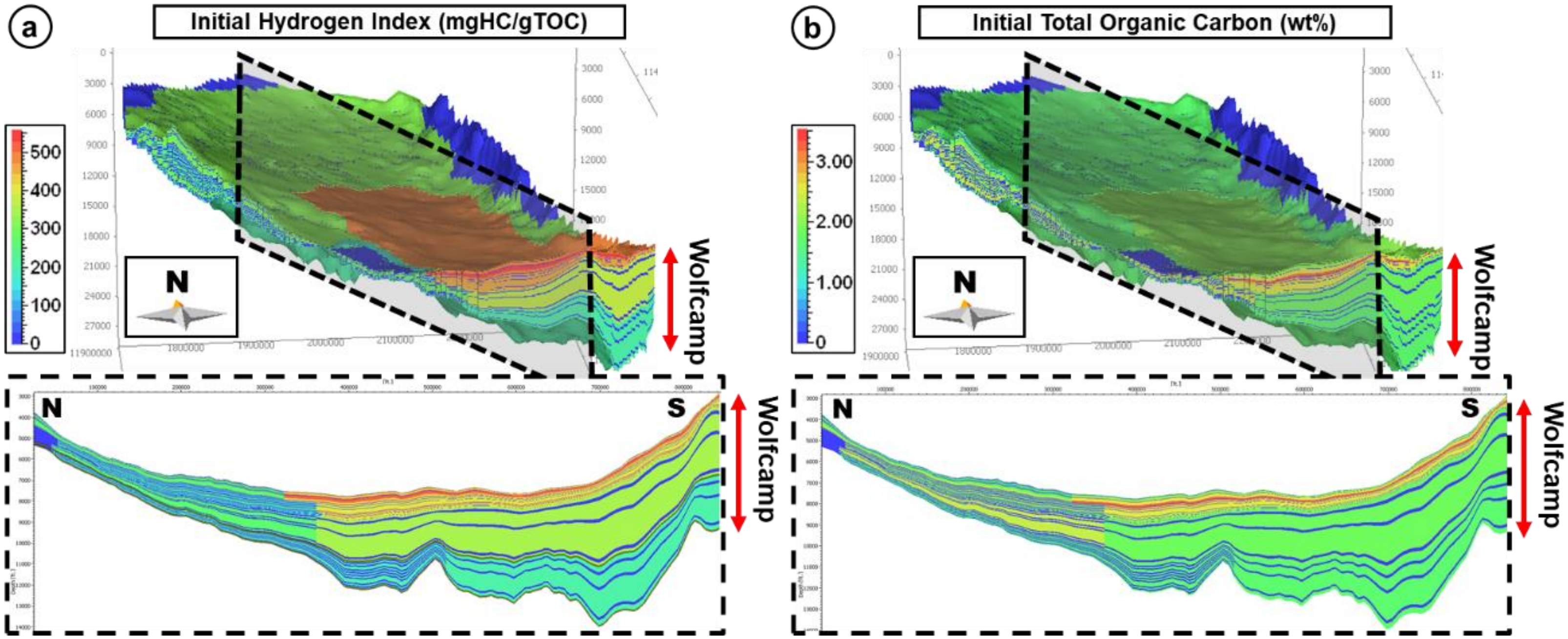


Figure 4.4. 3D basin and petroleum system model of Delaware Basin with enhanced input of variability of Wolfcamp (p50 model) restored initial (a) HI and (b) TOC. Note that only Wolfcamp is displayed in this figure.

In the current model, the carbonate-dominated lithofacies with marine carbonate OF “A” type in the basin margin is excluded from the calculation, as it is assumed that this source rock primarily charged the Central Basin Platform and Northwest Shelf outside the basinal portion of the Delaware Basin (Baskoro et al., 2021a; Curtis and Zumberge, 2018; Pepper et al., 2020).

4.3.1.2. Hydrocarbon Generation Simulation and Calibration

Three model cases of p90/50/10 are built with parameters of probable initial HI-TOC for each lithofacies. Each probable model is then simulated and matches the measured HI (**Figure 4.5a**) and TOC (**Figure 4.5b**) for the Wolfcamp Fm in most parts of the basin.

The siliceous mudstone facies is predominantly modeled with the kinetic model of OF “B”, while the mixed argillaceous mudstone is predominantly modeled with the kinetic model of OF “D/E” (**Figure 4.6a**). Calculated transformation ratio (TR) generally varies from 15 to 90 % (**Figure 4.6b**), with high TR dominating the center part of the basin in response to OF “B” dominated facies with higher thermal maturity levels than the southernmost part of the basin.

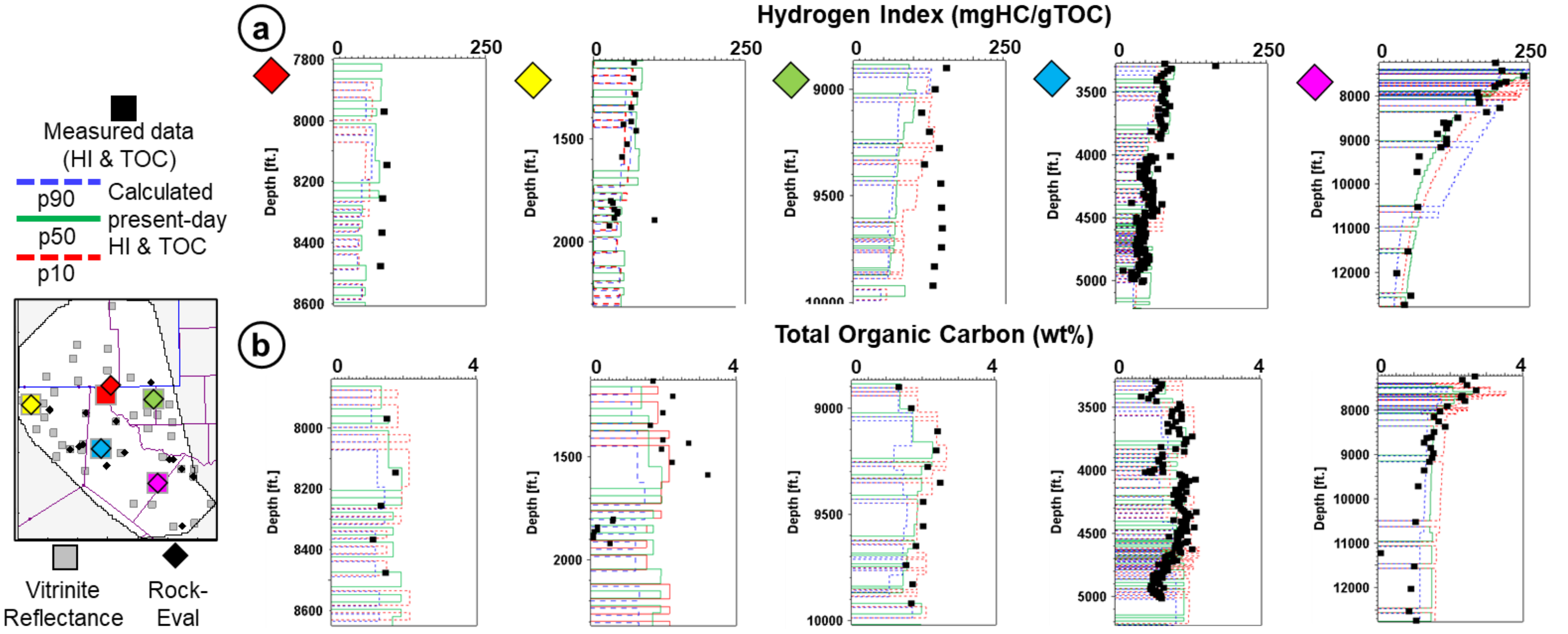


Figure 4.5. Calibration of the present-day (a) HI and (b) TOC based on p90/50/10 initial HI and TOC input for the Wolfcamp source rock.

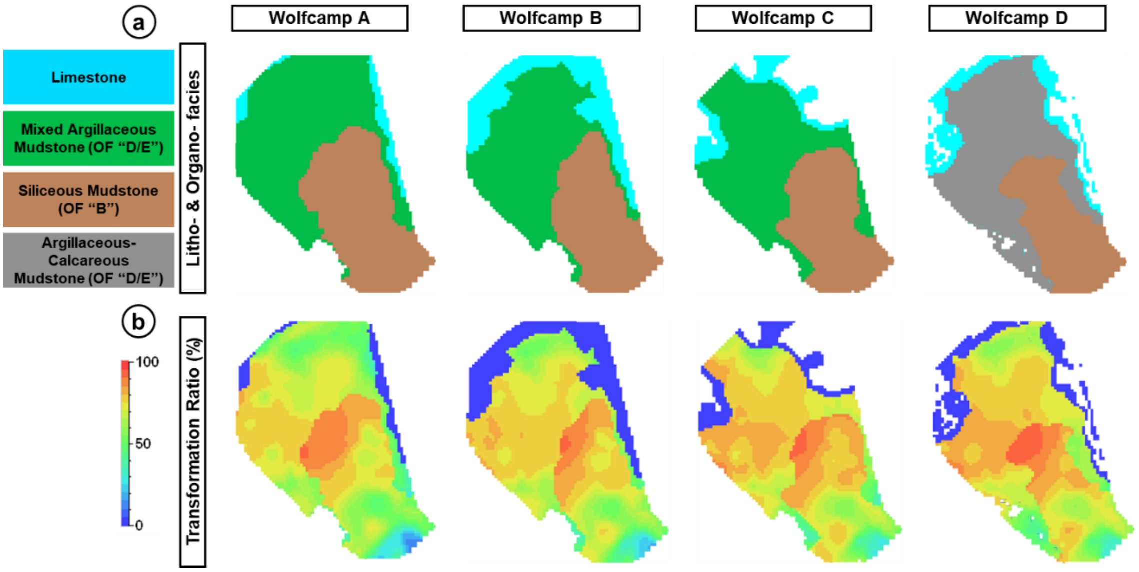


Figure 4.6. (a) Predominant basin-wide litho- and organo- facies distribution of the source rock unit for each Wolfcamp interval and (b) calculated transformation ratio on top of each Wolfcamp intervals.

Figure 4.7 illustrates the calculated present-day HI and TOC distribution for different Wolfcamp intervals based on the initial HI and TOC input from the p50 model. Generally, Wolfcamp intervals in the northern-western part of the basin are dominated by lower present-day HI and TOC as a result of lower initial HI-TOC (**Figure 4.4**), OF “D/E” dominated organofacies (**Figure 4.6**), and higher thermal maturity level (**Figure 4.2**). In contrast, Wolfcamp intervals in the southern-eastern part of the basin have higher initial HI-TOC, OF “B” dominated organofacies, and lower thermal maturity. Since the 3D BPSM is calibrated to the thermal maturity indicator and measured HI and TOC, the workflow proceeds with HC generation simulation.

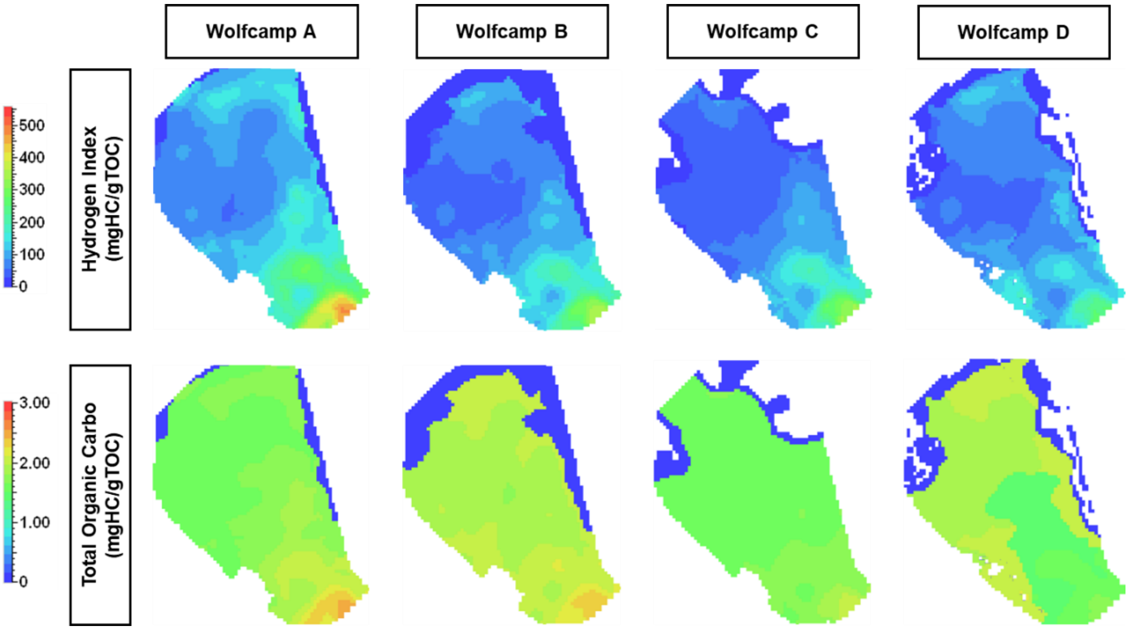


Figure 4.7. Calculated present-day HI & TOC on main source rock unit for each Wolfcamp intervals (p50 model).

HC generation simulation yields three different calculated masses of generated HC in oil and gas phases in Giga tons (GTons) unit: p90/50/10 models. **Figure 4.8** shows the cumulative oil and gas volume (BBOE) of generated HC over geological time from the whole Wolfcamp SR intervals for each probable case. The cumulative oil and gas volumes are 271/519/1,693 BBO and 434/814/1,324 BBOE, respectively, for p90/50/10 models. The total oil equivalent volumes of generated HC for each p90/50/10 model are 705/1,332/3,016 BBOE. These calculated volumes are the starting point for the main calculation of HC mass balance.

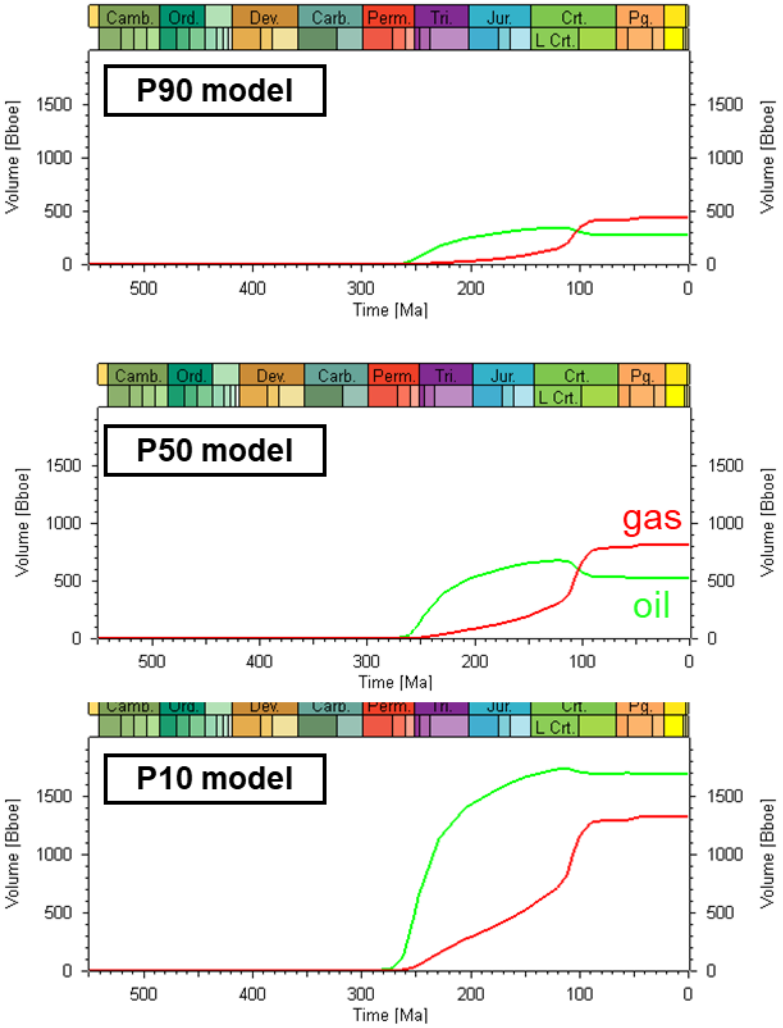


Figure 4.8. Calculated cumulative generated oil and gas volumes over geological time from the whole Wolfcamp SR intervals for each p90/50/10 model.

4.3.2. Hydrocarbon Mass Balance Calculation

4.3.2.1. Mass Balance Calculation Scheme

4.3.2.1.1. Assumptions Used in the Calculation

Figure 4.9 illustrates the “bookkeeping” scheme of the calculated HC volume from generation to production. All the calculated masses from generation to production must balance, following the fundamental law of mass conservation (Hantschel and Kauerauf, 2009a).

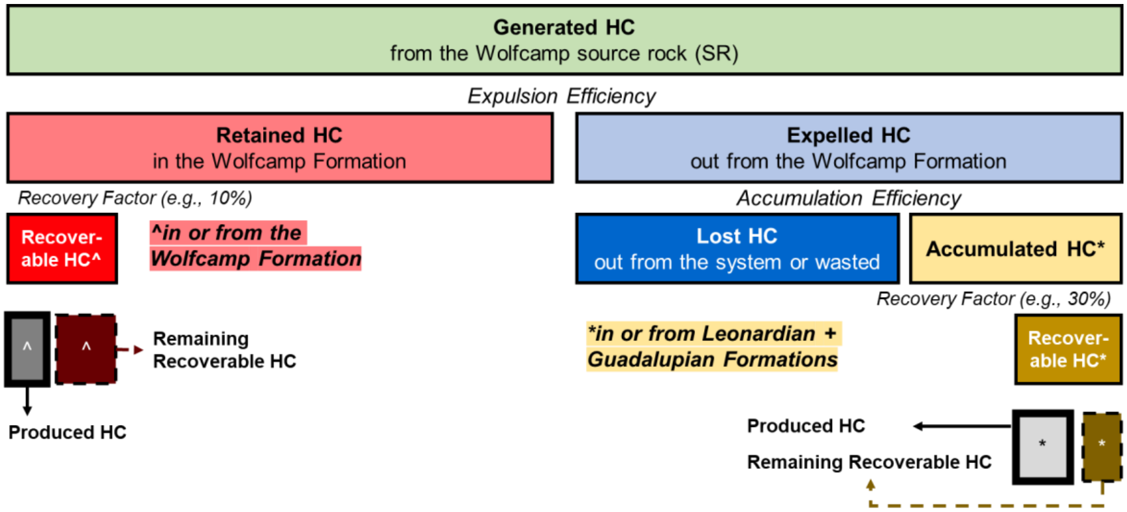


Figure 4.9. "Bookkeeping" scheme of calculated hydrocarbon masses from generation to production applied to Wolfcamp sourced hydrocarbon.

Several assumptions are involved in this mass balance calculation scheme. Firstly, the "produced HC" is based on production data compiled from Dutton et al. (2005) and Enverus (2019) data. Secondly, the recoverable HC is estimated based on the recovery factor (RF) derived from the literature study. This calculation uses typical RF for unconventional play of 10% (Fryklund and Stark, 2020; Liu et al., 2018; Qian et al., 2020) to calculate recoverable HC (unconventional accumulation) from the Wolfcamp Fm. RF of 30%, higher than the Wolfcamp Fm's RF, accounts for combined conventional (Dutton et al., 2005; Fryklund and Stark, 2020) and unconventional accumulations in Leonardian and Guadalupian Fms. Thirdly, Wolfcamp SR is assumed to contribute hydrocarbon charge for the whole Permian strata, including the Leonardian and Guadalupian Fms.

4.3.2.1.2. Defining the Terms Used in the Calculation

In this mass balance scheme (**Figure 4.9**), “Wolfcamp Formation” represents one geologic unit or formation of whole Wolfcamp interbedded source and non-source (reservoir and carrier) intervals. While “Wolfcamp SR” refers to source rock interval (organic-rich shale) parts of the “Wolfcamp Formation”. For example, the term “generated HC” represents the total HC volume generated from the whole “Wolfcamp SR” intervals of the “Wolfcamp Formation” (**Figure 4.9**).

Most of the terms used are derived from Hantschel and Kauerauf (2009a; b) and Peters et al. (2012). One of the important concepts for this calculation is “expulsion”. As described in Hantschel and Kauerauf (2009a) and Katz et al. (2017), expulsion has been defined as either the movement of HCs out of the kerogen (e.g., (Ertas et al., 2006; Pepper, 1991; Rushing et al., 2004) or out of the source rock (e.g., (Katz et al., 2017; Leythaeuser et al., 1984; Sandvik et al., 1992). Since one of the main purposes of the mass balance calculation is HC resource assessment, the term expulsion herein is closely tied to the latter definition.

Rather than detailed expulsion from each source rock interval only, expulsion herein refers to the movement of HCs out of the Wolfcamp Fm as one source rock-dominated geological unit. Thus, the term “expelled HC” represents the proportion of “generated HC” expelled out from the Wolfcamp Fm, which may further migrate to the carrier or reservoir outside the Wolfcamp Fm (**Figure 4.9**). While the term “retained HC” represents the remaining proportion of the “generated HC,” which remains in the Wolfcamp Fm, both in the shale (source rock play) or non-shale (tight-rock play) intervals. The proportions of the retained HC and expelled HC are defined by “expulsion efficiency - EE”. The term “EE” herein represents the overall EE of the Wolfcamp Fm. This simplification is relevant, considering Wolfcamp Play is a self-sourced play.

Using an assumed RF of 10% for the Wolfcamp Fm, “recoverable HC” in the Wolfcamp Fm can be calculated by multiplying the RF with the retained HC (**Figure 4.9**). From this point, the “remaining recoverable HC” in the Wolfcamp Fm can be estimated by subtracting the calculated recoverable HC with the “produced HC” from the Wolfcamp Fm based on production data.

The following important concept in this calculation is “migration,” which herein refers to “secondary migration” or migration from the source rock to the reservoir or trap through the carrier bed or fault (Hantschel and Kauerauf, 2009a; Baur et al., 2011; Duran et al., 2013). Secondary migration or “migration” represents the movement of HC sourced from the Wolfcamp Fm or the “expelled HC” in the “Leonardian and Guadalupian Fms” carrier bed. Portions of the expelled HC are either accumulated or trapped in Leonardian and Guadalupian reservoirs as “accumulated HC,” or migrated out from the system or wasted in the system as “lost HC” (**Figure 4.9**). The proportion of accumulated and lost HC is defined by “accumulation efficiency - AE”, equivalent to accumulation or trapping efficiency in other studies (e.g., Baur et al., 2011; Duran et al., 2013; (Rushing et al., 2004). In this calculation, the assumed AE is 10%.

Using an assumed combined RF of 30% for the combined Leonardian and Guadalupian Fms, “recoverable HC” from the combined Leonardian and Guadalupian Fms can be calculated by multiplying the RF with the accumulated HC. From this point, the “remaining recoverable HC” from the combined Leonardian and Guadalupian Fms can be estimated by subtracting the calculated recoverable HC with the “produced HC” from the combined Leonardian and Guadalupian Fms based on production data.

4.3.2.2. Mass Balance Calculation Scenarios

Three calculation scenarios are performed for the mass balance calculation: (1) backward calculation to derive expulsion-accumulation efficiencies to match USGS-estimated remaining recoverable HCs, (2) forward calculation with assumed expulsion-accumulation efficiencies, and (3) forward BPSM simulation with assumed geological model. The mass balance is then assessed using these three different calculation scenarios, focusing on the retained HC portion. The calculation scheme and assumptions involved in each scenario are integrated into the section below to make it easier for the reader to follow each calculation scenario and its impact on the estimated remaining recoverable HC in and EE of the Wolfcamp Fm.

4.4. Hydrocarbon Mass Balance Assessment from Multiple Calculation Scenarios

4.4.1. Calculation 1 – Inverse Calculation of Required Expulsion-Accumulation Efficiency to Accommodate USGS Estimated Remaining Recoverable HC

4.4.1.1. Calculation Scheme and Assumptions Involved

Calculation 1 is performed backward from recoverable HC to p90/50/10 generated HC (**Figure 4.9**). In calculation 1, the remaining recoverable HC portion is represented by the remaining recoverable resources based on the USGS estimates, both for Wolfcamp Fm (Gaswirth et al., 2018) and combined Leonardian-Guadalupian Fms (Gaswirth et al., 2018; Schenk et al., 2008). Since recoverable HCs are already estimated, retained-expelled-accumulated HC can be back-calculated by putting it into the mass balance calculation scheme along with the generated HC. The “required” expulsion and accumulation efficiencies can be derived to balance the estimated remaining recoverable based on the USGS estimates.

4.4.1.2. Results

Table 4.2 illustrates the results of calculation 1. The estimated expulsion and accumulation efficiency required to accommodate the USGS estimated recoverable HC in the Wolfcamp (35/78/140 BBOE) and combined Leonardian and Guadalupian Fms (18/31/51 BBOE) are 48/40/53 % and 21/22/11 %, respectively. It can also be observed that the calculated AE is 23/23/12 %, higher than the assumed minimum AE (10%) in this mass balance calculation scheme. The calculated EE from calculation 1 and USGS estimates serve as a comparison or reference model for the following two calculations.

Table 4.2. Hydrocarbon mass balance inverse calculation of required expulsion-accumulation efficiency to accommodate USGS estimated remaining recoverable HC.

HC Volumes are in BBOE (Billion Barrels of Oil Equivalent)	**Generated HC** from the Wolfcamp SR	**Retained HC** in the Wolfcamp Fm			**Expelled HC** out from the Wolfcamp Fm	**Expulsion Efficiency (%)** (derived from Mass-Balance Calculation with USGS Estimates and Production Data)	**Accumulated HC** in the Leonardian + Guadalupian Fms (from Wolfcamp expelled HC)			**Accumulation Efficiency (%)** (derived from Mass-Balance Calculation with USGS Estimates and Production Data)
		Recoverable HC					**Recoverable HC**			
		Produced HC from the Wolfcamp Fm	**Remaining Recoverable HC** (USGS Estimates - Wolfcamp Fm (2018))	**Retained HC** (without RF)			**Produced HC** from the Leonardian & Guadalupian Fms	**Remaining Recoverable HC** (USGS Estimates - Leonardian & Guadalupian Fms (2007; 2018))	**Accumulated HC** (without RF)	
P90	705.28	2.11	34.74	368.50	336.78	**48%**	5.74	17.55	77.63	**23%**
P50	1,332.59	2.11	77.66	797.70	534.89	**40%**	5.74	30.79	121.77	**23%**
P10	3,016.52	2.11	139.91	1,420.20	1,596.32	**53%**	5.74	50.59	187.77	**12%**

4.4.2. Calculation 2 – Multiple Expulsion Efficiency Scenarios to Estimate Remaining Recoverable HC

4.4.2.1. Calculation Scheme and Assumptions Involved

Calculation 2 aims to calculate the remaining recoverable HC by applying multiple scenarios of assumed possible minimum and maximum EE of 20 and 80%, respectively, on each calculated p90/50/10 generated HC through forward calculation, following the mass balance scheme from generated HC to recoverable HC (**Figure 4.9**). From these calculations, scenarios and the range of minimum-maximum EE for each p90/50/10 generated HC, which are possible according to mass balance, can be inferred. Then, the mean remaining recoverable HC from the possible minimum and maximum possible EE scenarios for each p90/50/10 generated HC is calculated and compared to USGS estimated remaining recoverable HC in Wolfcamp Fm (Gaswirth et al., 2018). The calculated mean EE is compared to the calculated EE from calculation 1. Additionally, by assuming a minimum AE of 10%, the expelled HC portion is further assessed to infer whether the generated HC from the Wolfcamp SR was sufficient to charge Leonardian and Guadalupian Fms or not.

4.4.2.2. Results

Table 4.3 demonstrates the results of calculation 2. The calculated recoverable HC for the Wolfcamp Fm is “larger” than the produced HC, using all p90/50/10 generated HC and EE scenarios. The calculated mean EE from all p90/50/10 generated HC is 50%, comparable to the EE calculated from calculation 2, with the largest difference of EE of 10% from the p50 case. With the mean EE of 50%, the calculated remaining recoverable HC is 33/65/149 BBOE, comparable to the USGS estimated remaining recoverable HC in the Wolfcamp Fm (Gaswirth et al., 2018).

Calculation 2 also indicates that the Wolfcamp generation balance (except for the p90 case with EE of 20%) is sufficient to charge both Wolfcamp Fm and Leonardian-Guadalupian Fms accumulation using a wide range of EE scenarios, even when assuming a minimum AE of 10% - lower than the required AE from the calculation 1, and minimum generated HC (p90 model) scenarios (**Table 3b**). The calculated accumulated HC significantly underestimated the USGS estimated remaining recoverable HC in the Leonardian-Guadalupian Fms (Gaswirth et al., 2018; Schenk et al., 2008).

Table 4.3. Hydrocarbon mass balance calculation with multiple assumed expulsion efficiency scenarios to estimate mean expulsion efficiency and mean remaining recoverable HC.

HC Volumes are in BBOE (Billion Barrels of Oil Equivalent)	**Generated HC** from the Wolfcamp SR	Assumed **Expulsion Efficiency** (%)	**Retained HC** in the Wolfcamp Fm		**Produced HC** from the Wolfcamp Fm	**Remaining Recoverable HC** in the Wolfcamp Fm		**USGS Estimates -** Wolfcamp Fm (2018)	**Expelled HC** out from the Wolfcamp Fm	Assumed **Accumulation Efficiency = 10%**		**Produced HC** from the Leonardian + Guadalupian Fms	**Remaining Recoverable HC** in the Leonardian + Guadalupian Fms	**USGS Estimates -** Leonardian & Guadalupian Fms (2007; 2018)
				Recoverable HC			**Mean**			**Accumulated HC** in the Leonardian + Guadalupian Fms (from Wolfcamp expelled HC)	**Recoverable HC**			
P90	705.28	**80%**	141.06	14.11	2.11	12.00	**33.15**	34.74	564.22	56.42	16.93	5.74	11.19	17.55
	705.28	**20%**	564.22	56.42	2.11	54.31			141.06	14.11	4.23	5.74	-1.51	

HC Volumes are in BBOE (Billion Barrels of Oil Equivalent)	**Generated HC** from the Wolfcamp SR	Assumed **Expulsion Efficiency** (%)	**Retained HC** in the Wolfcamp Fm		**Produced HC** from the Wolfcamp Fm	**Remaining Recoverable HC** in the Wolfcamp Fm		**USGS Estimates -** Wolfcamp Fm (2018)	**Expelled HC** out from the Wolfcamp Fm	Assumed **Accumulation Efficiency = 10%**		**Produced HC** from the Leonardian + Guadalupian Fms	**Remaining Recoverable HC** in the Leonardian + Guadalupian Fms	**USGS Estimates -** Leonardian & Guadalupian Fms (2007; 2018)
				Recoverable HC			**Mean**			**Accumulated HC** in the Leonardian + Guadalupian Fms (from Wolfcamp expelled HC)	**Recoverable HC**			
P50	1,332.59	**80%**	266.52	26.65	2.11	24.54	**64.52**	77.66	1,066.07	106.61	31.98	5.74	26.24	30.79
	1,332.59	**20%**	1,066.07	106.61	2.11	104.50			266.52	26.65	8.00	5.74	2.26	

HC Volumes are in BBOE (Billion Barrels of Oil Equivalent)	**Generated HC** from the Wolfcamp SR	Assumed **Expulsion Efficiency** (%)	**Retained HC** in the Wolfcamp Fm		**Produced HC** from the Wolfcamp Fm	**Remaining Recoverable HC** in the Wolfcamp Fm		**USGS Estimates -** Wolfcamp Fm (2018)	**Expelled HC** out from the Wolfcamp Fm	Assumed **Accumulation Efficiency = 10%**		**Produced HC** from the Leonardian + Guadalupian Fms	**Remaining Recoverable HC** in the Leonardian + Guadalupian Fms	**USGS Estimates -** Leonardian & Guadalupian Fms (2007; 2018)
				Recoverable HC			**Mean**			**Accumulated HC** in the Leonardian + Guadalupian Fms (from Wolfcamp expelled HC)	**Recoverable HC**			
P10	3,016.52	**80%**	603.30	60.33	2.11	58.22	**148.72**	139.91	2,413.22	241.32	72.40	5.74	66.66	50.59
	3,016.52	**20%**	2,413.22	241.32	2.11	239.21			603.30	60.33	18.10	5.74	12.36	

4.4.3. Calculation 3 – 3D HC Expulsion Simulation Prediction of Remaining Recoverable HC and Expulsion Efficiency

4.4.3.1. Calculation Scheme and Assumptions Involved

Similar to calculation 2, calculation 3 is also performed forward from generated HC to recoverable HC (**Figure 4.9**). The main difference is that the expelled and retained HC are calculated from the expulsion simulation with the 3D geological basin model. Calculation 3 aims to test modeling source rock as unconventional play using default parameters built in PetroMod® expulsion by performing mass balance calculation. The concept of expulsion in this 3D expulsion simulation by PetroMod® is similar to the concept applied in this mass balance calculation. And thus, expelled HC portion represents the proportion of the generated HC which is expelled out from the Wolfcamp Fm, which may further migrate to the reservoir or trap through a carrier or fault to a reservoir outside the Wolfcamp Fm. Since HC migration simulation is not part of this study, the expelled HC portion is not further assessed for its AE.

In PetroMod®, the expulsion simulation requires an “expulsion factor” parameter, which represents a portion of the generated mass, excluding the residual saturation, that is expelled from the “source rock” (e.g., organic-rich Wolfcamp SR layer). However, after testing a wide range of expulsion factors (20-80 %), the difference in the calculated retained HC volume is insignificant. This is because the amount of retained HC is determined by the source rock's adsorption capacity, controlled by the source rock's property (initial HI-TOC and litho- and organo-facies), as the first order in the model. Once the adsorption capacity is surpassed (e.g., at the 20% expulsion factor), any expulsion factor of >20% will no longer significantly affect the calculated amount of retained HC portion. Thus, the simulation proceeds with only a 50% expulsion factor applied to the p90/50/10 model.

4.4.3.2. Results

Table 4.4 summarizes the results of calculation 3. The predicted overall EE is 16/33/64 %, while the estimated remaining recoverable HC from the Wolfcamp Fm is 57/87/107 BBOE. It can be observed that the higher the initial HI-TOC, the higher the predicted EE. This is because a smaller fraction of the total generated HC is required to surpass the critical saturation threshold for expulsion (Sandvik et al., 1992).

It can be observed that the calculated remaining recoverable HC from the p50 model is comparable to USGS-estimated remaining recoverable HC. On the other hand, the calculation results from p90 and p10 models are significantly different. Compared to USGS's estimated remaining recoverable HC, the p10 models significantly overestimate the EE and, thus, underestimate the calculated remaining recoverable HC by 23 %.

Table 4.4. Hydrocarbon mass balance calculation based on 3D hydrocarbon expulsion simulation prediction of remaining recoverable HC and expulsion efficiency.

HC Volumes are in BBOE (Billion Barrels of Oil Equivalent)	**Generated HC** from the Wolfcamp SR	**Expelled HC** out from the Wolfcamp Fm	**Retained HC** in the Wolfcamp Formation		**Expulsion Efficiency (%)** (derived from PetroMod calculation)	**Produced HC** from the Wolfcamp Fm	**Remaining Recoverable HC** in the Wolfcamp Fm	**USGS Estimates -** Wolfcamp Fm (2018)
				Recoverable HC				
p90	705.28	112.83	592.45	59.25	**16%**	2.11	**57.14**	34.74
p50	1,332.59	442.03	890.56	89.06	**33%**	2.11	**86.95**	77.66
p10	3,016.52	1,926.47	1,090.05	109.01	**64%**	2.11	**106.90**	139.91

4.5. Discussion

4.5.1. Petroleum System Analysis – Wolfcamp as Unconventional Play and Primary Source Rock for the Permian Strata

4.5.1.1. Insight on Modeling Source Rock and Unconventional Play of Interbedded Model: Moderate Expulsion Efficiency is Observed

Based on the mass balance calculation, Wolfcamp unconventional play could contribute to significant remaining recoverable HC potential in the Delaware Basin. It indicates production is still within the early stage as the amount of produced HC is only a tiny portion compared to the recoverable HC. The mass balance calculation also implies that Wolfcamp unconventional play is self-sourced from Wolfcamp SR intervals.

Based on calculations 1 and 2, the mean overall EE is within 40 to 53%. The overall EE is considered moderate for the unconventional play model of “interbedded” source-reservoir like the Wolfcamp Play. In reference to this calculated EE, EE in other unconventional play models, including “massive” source (e.g., Marcellus Fm in Appalachian Basin) and “sandwiched” source-reservoir-source (e.g., Bakken Fm in Williston Basin), which typically considered lower than “interbedded” model (Katz et al., 2017), may even be lower than 40%, and thus leading to limited conventional accumulation outside the source rock unit.

The mass balance calculation can be a comparison or reference in modeling unconventional play. As illustrated in calculation 3, the current 3D model geological setting and more reliable models at various scales must be accurately adjusted to predict retained HC portion and expulsion efficiency to accommodate moderate overall expulsion efficiency and significant remaining HC resources. However, these efforts require more geological interpretation, high-resolution input and modeling, and computational resources.

4.5.1.2. Revisiting Petroleum System of Permian Strata: Insight on Accumulation Efficiency and Contribution of Multiple Source Rocks toward Significant Remaining Recoverable HC in Leonardian and Guadalupian Fms

Inferred AE from calculation 1 may be considered too high as most well-studied basins typically have AE of less than 15% (e.g., Baur et al., 2011; Duran et al., 2013; Lewan et al., 2002; Magoon and Valin, 1994). The inferred high AE from calculation 1 may indicate that other source rocks besides Wolfcamp, including the Bone Spring, Avalon and probably some of the Guadalupian source rocks, were most likely responsible for charging the Leonardian and Guadalupian Fms. This possibility is also indicated in calculation 2 with an assumed AE of 10%. Wolfcamp expelled-accumulated HC balance was insufficient to accommodate the USGS estimates of Leonardian-Guadalupian remaining recoverable HC (Gaswirth et al., 2018; Schenk et al., 2008), even when assuming a high EE of 80%. Based on our 3D BPSM, the source rocks of the Guadalupian Fm are predominantly within the early to peak oil generation window, while the source rocks of the Leonardian Fm are within the peak to late oil generation window (**Figure 4.10**), indicative of charge potential from these source intervals.

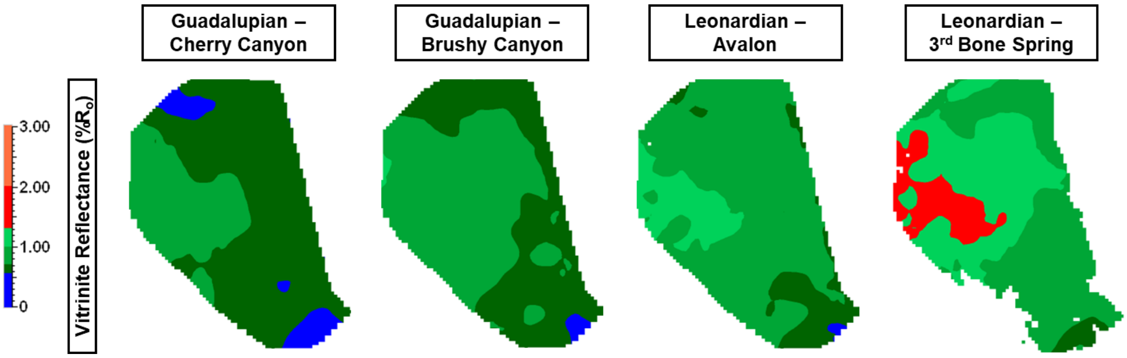


Figure 4.10. Calculated thermal maturity (easy %R_o) for source rock intervals in Leonardian and Guadalupian Formations.

There is a limited number of oil-source correlation studies between Leonardian/ Guadalupian Fms and Wolfcamp SR or in the Leonardian-Guadalupian Fms within the Delaware Basin. A few examples of oil-source correlation studies of Permian strata in the Delaware Basin are Curtis and Zumberge (2017, 2018) and Echegu et al. (2021). However, their analyses are limited to the Wolfcamp-Lower Bone Spring (marine and mixed marine-terrigenous shale source rock) petroleum system within the Delaware Basin and Wolfcamp marine carbonate source rock, which is primarily responsible for HC accumulations outside DB. Oil-source correlation studies for Leonardian and Guadalupian petroleum systems are performed for the Northwest shelf's marine carbonate oil accumulations outside the DB. The oil accumulations in Leonardian and Guadalupian Fms within the Delaware Basin are not yet well studied for oil-source correlation. A recent study by Echegu et al. (2021) also suggests that Wolfcamp could have charged younger Permian strata, assuming vertical migration without a strong top seal was available until the Ochoan Evaporite. Thus, performing migration simulation is necessary to assess the expelled-accumulated HC balance and AE from multiple source rocks, which is beyond the scope of this study.

4.5.2. Highlights of the Mass Balance Calculation Workflow: Limitations and Advantages

4.5.2.1. Limitations of the Workflow

This mass balance calculation is highly dependent on the recovery factor to calculate recoverable HC; however, uncertainty in this parameter is considerable. This limitation becomes prominent in this workflow calculation due to the assessment scale. The whole basin is assumed as a unified fetch area as HC (secondary) migration simulation is not currently considered and is not the focus of this study. However, even calculation at a field scale will show significant uncertainty relative to basin-scale assessment. This is due to the complex lateral heterogeneity of the Fm of interest. For example, EE calculated for one field may not apply to another; likewise, one lithology pattern may differ from another. Thus, the mass balance calculation will also be different.

4.5.2.2. Advantages of the Workflow

The current mass balance demonstration has already considered the uncertainty in the generated HC, demonstrated through p90/50/10 values derived from the extensive studies of source rock restoration (Baskoro et al., 2021a) and lithofacies (Baskoro et al., 2021b) to characterize the uncertainty in the starting point of the mass balance calculation.

This workflow has some advantages, especially for quantifying remaining recoverable HC or deriving expulsion efficiency as part of a petroleum system analysis study. Firstly, the proposed workflow is an efficient way of taking a quick look at the petroleum system to assess whether the quantified generation balance is sufficient for charging specific reservoirs (both unconventional and conventional) by applying assumed possible expulsion-accumulation efficiency and estimated remaining recoverable HC. Secondly, the workflow can be applied on a smaller scale in a mature production stage, and the more reliable remaining recoverable HC can be estimated. Thirdly, EE of the Fm of interest can be derived by incorporating production data and accurately estimating the remaining recoverable HC. Accurate EE can be derived more confidently in a highly mature production area. Finally, this calculation provides a reference model for performing basin and petroleum system modeling for source rock and/ or unconventional play.

4.6. Conclusion

Using the restored source rock initial TOC, HI, and facies, the total p90/50/10 generated HC from the Wolfcamp SR is 705/1,332/3,016 BBOE. The inversion calculation of expulsion efficiency indicates that the p90/50/10 of 48/40/53 % overall expulsion efficiency is required for the Wolfcamp Formation to achieve the USGS estimated remaining recoverable HC. Compared to the forward and inverse calculations, calculation with expulsion simulation illustrates the current 3D model geological setting and expulsion parameters must be accurately adjusted to predict unconventional HC accumulations better.

The mass balance calculations demonstrate that only a tiny portion of the oil and gas volume has been produced from Wolfcamp reservoirs in the Delaware Basin. Based on the calculated mean expulsion efficiency of 50%, forward calculation shows a large volume of oil and gas of 33/65/149 BBOE remains to be produced from the Wolfcamp Play.

Finally, the mass balance calculation also reveals that multiple other source rocks from the Leonardian and Guadalupian Formations were most likely responsible for significant remaining recoverable HC in the Leonardian and Guadalupian Formations beside the Wolfcamp SR.

4.7. References

Adams, J. E., 1965, Stratigraphic-tectonic development of Delaware Basin: AAPG Bulletin, v. 49, p. 2140-2148.

Amer, A., R. di Primio, and R. Ondrak, 2015, 4-D petroleum systems modelling of the Haynesville shale play—understanding gas in place: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 935-949.

Arango, I., and B. J. Katz, 2019, Understanding expulsion capacity and organic porosity in unconventional petroleum systems: AAPG Hedberg Conference, The Evolution of Petroleum Systems Analysis.

Bai, H., X. Pang, L. Kuang, H. Pang, X. Wang, X. Jia, L. Zhou, and T. Hu, 2017, Hydrocarbon expulsion potential of source rocks and its influence on the distribution of lacustrine tight oil reservoir, Middle Permian Lucaogou Formation, Jimsar Sag, Junggar Basin, Northwest China: Journal of Petroleum Science and Engineering, v. 149, p. 740-755.

Baskoro, A. S., F. Baur, A. Z. Yu, and E. L. Grossman, 2021a, Source rock restoration using a multi-well inversion approach for the Wolfcamp Play in the Permian Delaware Basin of West Texas and New Mexico: AAPG/SEG International Meeting for Applied Geoscience & Energy.

Baskoro, A. S., E. Ganguly, A. Z. Yu, and S. Misra, 2021b, End-to-end machine learning workflow for lithofacies classification and source rock organo-facies identification in the Permian Delaware Basin, southwest USA: AAPG/SEG International Meeting for Applied Geoscience & Energy.

Baur, F., 2019, Predicting petroleum gravity with basin modeling: New kinetic models: AAPG Bulletin, v. 103, p. 1811-1837.

Baur, F., R. di Primio, C. Lampe, and R. Littke, 2011, Mass balance calculations for different models of hydrocarbon migration in the Jeanne d'Arc basin, offshore Newfoundland: Journal of Petroleum Geology, v. 34, p. 181-198.

Becker, M., A. M. Romero, and A. Yu, 2019, 3-D basin modeling of the Permian Delaware Basin: tectonic evolution assessment to improve definition of boundary conditions: AAPG Annual Convention and Exhibition.

Bievenour, A., and S. Sonnenberg, 2019, Reservoir characterization of the Bone Spring and Wolfcamp Formations, Delaware Basin, Ward County, West Texas: SPE/AAPG/SEG Unconventional Resources Technology Conference.

Blakey, R., 2019, Greater Permian Basin Series, Colorado Plateau Geosystems. Inc. https://deeptimemaps.com/greater-permian-basin/ *(accessed on May 4th 2022)*

Bryant, I. D., C. B. Stabell, and M. Neumaier, 2013, Evaluation of unconventional resources using a petroleum system modeling approach: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 1394-1401.

Cander, H., 2012, Sweet spots in shale gas and liquids plays: prediction of fluid composition and reservoir pressure: AAPG Annual Convention and Exhibition, p. 12-15.

Coutinho, L. F. C., 2008, Analysis of the petroleum mass balance in a mature zone for exploration-Reconcavo Basin, Brazil, Université Pierre et Marie Curie-Paris VI.

Curtis, J. B., and J. E. Zumberge, 2017, An applied geochemical look at the Delaware Basin petroleum systems: AAPG Division of Professional Affairs, Delaware Basin Playmaker's Forum.

Curtis, J. B., and J. E. Zumberge, 2018, Permian Basin petroleum systems–geochemical insight into hydrocarbon generation, migration and well performance: SPE/AAPG/SEG Unconventional Resources Technology Conference, p. 3980-3985.

Duran, E. R., R. di Primio, Z. Anka, D. Stoddart, and B. Horsfield, 2013, 3D-basin modelling of the Hammerfest Basin (southwestern Barents Sea): A quantitative assessment of petroleum generation, migration and leakage: Marine and Petroleum Geology, v. 45, p. 281-303.

Dutton, S. P., E. M. Kim, R. F. Broadhead, W. D. Raatz, C. L. Breton, S. C. Ruppel, and C. Kerans, 2005, Play analysis and leading-edge oil-reservoir development methods in the Permian basin: Increased recovery through advanced technologies: AAPG Bulletin, v. 89, p. 553-576.

Echegu, S., A. K. Bissada, and L. Elrod, 2021, Geochemical characterization and classification of crude oils of the Permian Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 105, p. 223-246.

Enverus, 2019, DrillingInfo Database. Drillinginfo and RigData | Enverus *(accessed on August 2019)*

Ertas, D., S. R. Kelemen, and T. C. Halsey, 2006, Petroleum expulsion part 1. Theory of kerogen swelling in multicomponent solvents: Energy & Fuels, v. 20, p. 295-300.

Fairhurst, B., T. Ewing, and B. Lindsay, 2021, West Texas (Permian) Super Basin, United States: Tectonics, structural development, sedimentation, petroleum systems, and hydrocarbon reserves: AAPG Bulletin, v. 105, p. 1099-1147.

Fryklund, B., and P. Stark, 2020, Super basins—New paradigm for oil and gas supply: AAPG Bulletin, v. 104, p. 2507-2519.

Gaswirth, S. B., K. L. French, J. K. Pitman, K. R. Marra, T. J. Mercier, H. M. Leathers-Miller, C. J. Schenk, M. E. Tennyson, C. A. Woodall, and M. E. Brownfield, 2018, Assessment of undiscovered continuous oil and gas resources in the Wolfcamp Shale and Bone Spring Formation of the Delaware Basin, Permian Basin Province, New Mexico and Texas, 2018, U.S. Geological Survey.

Hantschel, T., and A. I. Kauerauf, 2009a, Migration and accumulation, Fundamentals of Basin and Petroleum Systems Modeling, Springer, p. 247-340.

Hantschel, T., and A. I. Kauerauf, 2009b, Petroleum generation, Fundamentals of Basin and Petroleum Systems Modeling, Springer, p. 151-198.

Hills, J. M., 1984, Sedimentation, tectonism, and hydrocarbon generation in Delaware Basin, west Texas and southeastern New Mexico: AAPG Bulletin, v. 68, p. 250-267.

Jarvie, D. M., 2017, Geochemical assessment and characterization of petroleum source rocks and oils, and petroleum systems, Permian Basin, US: The Houston Geological Society Bulletin, v. 60.

Jarvie, D. M., R. J. Hill, T. E. Ruble, and R. M. Pollastro, 2007, Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment: AAPG Bulletin, v. 91, p. 475-499.

Justman, H., and R. Broadhead, 2010a, Petroleum source rock data for the Bone Spring Formation, Delaware Basin, southeastern New Mexico, Open File Report 529, New Mexico Bureau of Geology and Mineral Resources A Division of New Mexico Institute of Mining and Technology.

Justman, H., and R. Broadhead, 2010b, Petroleum source rock data for the Brushy Canyon Formation, Delaware Basin, southeastern New Mexico, Open File Report 528, New Mexico Bureau of Geology and Mineral Resources A Division of New Mexico Institute of Mining and Technology.

Katz, B., I. Arango, and F. Frasse, 2017, Expulsion and migration associated with unconventional petroleum systems: The Houston Geological Society Bulletin.

Kuchinskiy, V., K. Gentry, and R. Hill, 2012, Source rock evaluation technique: a probabilistic approach for determining hydrocarbon generation potential and in-place volume for shale plays: AAPG Annual Convention and Exhibition.

Lewan, M., M. Henry, D. Higley, and J. K. Pitman, 2002, Material-balance assessment of the New Albany-Chesterian petroleum system of the Illinois basin: AAPG Bulletin, v. 86, p. 745-777.

Leythaeuser, D., M. Radke, and R. Schaefer, 1984, Efficiency of petroleum expulsion from shale source rocks: Nature, v. 311, p. 745-748.

Liu, S., V. Sahni, J. Tan, D. Beckett, and T. Vo, 2018, Laboratory investigation of EOR techniques for organic rich shales in the Permian Basin: SPE/AAPG/SEG Unconventional Resources Technology Conference.

Magoon, L. B., and Z. C. Valin, 1994, Overview of Petroleum System Case Studies: Chapter 20: Part V. Case Studies--Western Hemisphere, The Petroleum Systems--From Source to Trap, v. Memoir 60.

Muscio, G. P. A., M. A. Everett, L. P. Marpaung, K. T. McDannell, A. A. Miceli Romero, and I. Yurchenko, 2016, Insights into migration efficiency based on stochastic mass balance calculations and full-physics forward basin modeling, AAPG Hedberg Conference, The Future of Basin and Petroleum System Modelling, Santa Barbara, California, AAPG Datapages/ Search and Discovery Article.

Pepper, A., A. Doebbert, J.-M. Laigle, and L. Laigle, 2020, Greater Permian Basin petroleum systems–what are we learning in the transition from conventional to unconventional?: The Houston Geological Society Bulletin.

Pepper, A. S., 1991, Estimating the petroleum expulsion behaviour of source rocks: A novel quantitative approach: Geological Society, London, Special Publications, v. 59, p. 9-31.

Pepper, A. S., and P. J. Corvi, 1995, Simple kinetic models of petroleum formation. Part I: oil and gas generation from kerogen: Marine and Petroleum Geology, v. 12, p. 291-319.

Peters, K. E., D. J. Curry, and M. Kacewicz, 2012, An overview of basin and petroleum system modeling: Definitions and concepts: AAPG Hedberg Series: Basin Modeling New Horizons in Research and Applications, v. 4, p. 1-16.

Popova, O., 2019, Permian Basin Wolfcamp and Bone Spring shale plays Geology review, U.S. Energy Information Administration.

Qian, Y., F. Male, S. A. Ikonnikova, K. Smye, G. McDaid, and E. Goodman, 2020, Permian Delaware basin Wolfcamp a formation productivity analysis and technically recoverable resource assessment: SPE/AAPG/SEG Unconventional Resources Technology Conference.

Romero-Sarmiento, M.-F., M. Ducros, B. Carpentier, F. Lorant, M.-C. Cacas, S. Pegaz-Fiornet, S. Wolf, S. Rohais, and I. Moretti, 2013, Quantitative evaluation of TOC, organic porosity and gas retention distribution in a gas shale play using petroleum system modeling: Application to the Mississippian Barnett Shale: Marine and Petroleum Geology, v. 45, p. 315-330.

Rushing, J., A. Chaouche, and K. Newsham, 2004, A mass balance approach for assessing basin-centred gas prospects: integrating reservoir engineering, geochemistry and petrophysics: Geological Society, London, Special Publications, v. 237, p. 373-390.

Sandvik, E., W. Young, and D. Curry, 1992, Expulsion from hydrocarbon sources: The role of organic absorption: Organic Geochemistry, v. 19, p. 77-87.

Schenk, C. J., R. M. Pollastro, T. A. Cook, M. J. Pawlewicz, T. R. Klett, R. R. Charpentier, and H. E. Cook, 2008, Assessment of undiscovered oil and gas resources of the Permian Basin Province of west Texas and southeast New Mexico, 2007, U.S. Geological Survey.

Schmoker, J. W., 1994, Volumetric calculation of hydrocarbons generated, The Petroleum System--From Source to Trap, AAPG Special Volumes, p. 323-326.

Silver, B. A., and R. G. Todd, 1969, Permian cyclic strata, northern Midland and Delaware basins, west Texas and southeastern New Mexico: AAPG Bulletin, v. 53, p. 2223-2251.

Yu, A. Z., M. Becker, A. S. Baskoro, and M. Bhatia, 2020, Fluid property variations and its relationship with the geo-history of the Permian Delaware Basin: AAPG Annual Conference and Exhibition.

5. SUMMARY AND CONCLUSIONS

5.1. Summary

5.1.1. Chapter 2 - Restoring Source Rock Initial Quality and Quantity with Kinetic-Based Inversion - Applied to the Wolfcamp Play in the Permian Delaware Basin

Based on geochemical interpretation, the organofacies within the Wolfcamp interval vary from predominantly marine clastic (OF “B”) to mixed marine-terrigenous (OF “D/E”) facies. The kinetic-based inversion results show a considerable range of restored HI_i and TOC_i, values indicating source rock heterogeneity of the Wolfcamp Play. Generally, Wolfcamp A and B intervals are excellent source rocks with P50 HI_i of ca. 530 and 429 mgHC/gTOC and P50 TOC_i of 2.64 and 2.27 wt%, respectively. Wolfcamp C and D intervals are also good-very good source rocks with relatively lower potential than Wolfcamp A and B.

Despite the limitation and uncertainty as a consequence of using TR derived from the basin model, the kinetic-based inversion procedure has the advantages of requiring no HI_i assumption and restoring the source rock according to thermal maturity, source rock organofacies, kinetic model, and measured data, among many others. Preliminary calculation of the restored HI_i and TOC_i and the calculated volumes of ultimate expellable hydrocarbon potential support the massive hydrocarbon volume predicted and produced for the Wolfcamp Play in the Delaware Basin.

5.1.2. Chapter 3 - Basin-Wide Lithofacies Identification Using Unsupervised Machine Learning: Applied to the Unconventional Wolfcamp Play, Permian Delaware Basin

Seven basin-wide lithofacies are identified within the Wolfcamp Play by applying two levels of unsupervised machine learning workflows using a determined subset of wells with four types of well-log measurements (PEF, GR, DT, and DEN). The gross depositional environment interpretation based on identified lithofacies concludes that there are three predominant facies of Wolfcamp intervals: the shallow marine, shallow-to-deeper marine transition, and deep marine. The shallow marine facies is dominated by limestone lithofacies with marine carbonate/ marl organofacies (OF “A”). The shallow-to-deeper marine facies is generally dominated by mixed argillaceous mudstone and other lithofacies with mixed marine-terrigenous organofacies (OF “D/E-B”). The deep marine facies is dominated by siliceous mudstone and other lithofacies with marine clastic organofacies (OF “B”). The vertical and lateral variability of lithofacies and depositional environments is recognized throughout Wolfcamp intervals. Other minor lithofacies are also identified, including calcareous-argillaceous mudstone, siltstone, mixed carbonate-siliciclastic, and heavy mineral (probably ankerite) bearing intervals.

Despite the limitation and challenges, a consequence of a data-driven method, the procedure has the advantages of being a straightforward workflow and using only four types of well-log measurements. The interpretation of lithofacies and depositional environments for Wolfcamp intervals from this workflow is comparable with the interpretation from other studies based on core data.

5.1.3. Chapter 4 - Mass Balance Calculation for the Wolfcamp-Sourced Hydrocarbon in Permian Delaware Basin: Insight on Remaining Recoverable Resource and Expulsion-Migration Efficiency

Calculated generated HC volumes (p90/50/10) from the Wolfcamp whole source rock (SR) intervals based on the assumed 3D geological model and source rock properties are 705/1,332/3,016 Billion Barrels of Oil Equivalent (BBOE). The mass balance is performed with three calculation scenarios. Firstly, mass balance with inverse calculation scenario indicates the back-calculated overall expulsion efficiency for p90/50/10 of Wolfcamp Formation is 48/40/53 %. Secondly, mass balance calculation with a wide range of assumed expulsion efficiencies with a mean expulsion efficiency of 50% determines the mean p90/50/10 remaining recoverable HC in the Wolfcamp Formation is 33/65/149 BBOE. Finally, calculation with expulsion simulation illustrates the current 3D model geological setting, and expulsion parameters must be accurately adjusted to match the prediction of unconventional HC accumulations with the first two calculations as the reference.

Despite the limitation of this workflow due to its dependency on recovery factor and availability of production data, this workflow is an efficient tool for taking a quick look at the petroleum system, especially related to possibilities of generated HC distribution in the system. The calculated mass balance can be a comparison or reference in modeling unconventional systems.

5.2. Conclusion

A new alternative workflow of source rock reconstruction has been proposed. The proposed workflow includes kinetic-based inversion to restore initial quality and quantity and unsupervised machine learning to identify basin-wide litho-organo-facies and gross depositional environments. The proposed workflow reveals that Wolfcamp SR initial quality and quantity are predominantly good to excellent, with initial hydrogen index of 416 ± 198 mgHC/gTOC and initial TOC of 2.45 ± 0.85 wt%. The identified basin-wide gross depositional environment varies from shallow marine in the basin margin to shallow-to-deeper and deep marine in the basin depocenter. Higher quality and quantity of source rock units are found in deep marine facies associated with predominantly siliceous mudstone with marine clastic-typed source rock.

The reconstructed Wolfcamp source rock is integrated into the 3D Delaware basin model and the calculated p90/50/10 generated hydrocarbon volumes is 705/1,333/3,016 BBOE. Based on mass balance calculations with mean expulsion efficiency of 50%, significant mean p90/50/10 recoverable HC of 33/65/149 BBOE remains in the Wolfcamp Formation. The proposed source rock reconstruction and hydrocarbon mass balance workflow can be applied in any source rock or unconventional play to understand source rock distribution and estimate remaining recoverable resources and expulsion efficiency to reduce the hydrocarbon exploration risk.