군집 별 표준곡선 매개변수를 이용한 치밀오일 생산성 예측 순환신경망 모델

한동권 · 김민수 · 권순일*

동아대학교

Recurrent Neural Network Model for Predicting Tight Oil Productivity Using Type Curve Parameters for Each Cluster

Dong-kwon Han · Min-soo Kim · Sun-il Kwon*

Dong-A University

E-mail : dkhan@donga.ac.kr / minsoo0821@naver.com / sikwon@dau.ac.kr

요 약

치밀오일 미래 생산성 예측은 잔류오일 회수량 및 저류층 거동 분석을 위해 중요한 작업이다. 일반적으로 석유공학적 관점에서 감퇴곡선법을 이용하여 생산성 예측이 이루어지는데, 최근에는 데이터기반의 머신러닝 기법을 이용한 연구도 수행되고 있다. 본 연구에서는 딥러닝 기반 순환신경망과 LSTM, GRU 알고리즘을 이용하여 미래 생산량 예측을 위한 효과적인 모델을 제안하고자 한다. 입력변수로는 치밀오일 생산 시 산출되는 오일, 가스, 물과 이와 더불어 다양한 군집분석을 통해 산출된 표준곡선이 주요 매개변수이고, 출력변수는 월별 오일 생산량이다. 기존의 경험적 모델인 감퇴곡선법과 순환신경망 모델들을 비교하였으며, 모델의 예측성능을 향상시키기 위해 하이퍼파라미터 튜닝을 통해 최적 모델을 도출하였다.

ABSTRACT

Predicting future productivity of tight oil is an important task for analyzing residual oil recovery and reservoir behavior. In general, productivity prediction is made using the decline curve analysis(DCA). In this study, we intend to propose an effective model for predicting future production using deep learning-based recurrent neural networks(RNN), LSTM, and GRU algorithms. As input variables, the main parameters are oil, gas, water, which are calculated during the production of tight oil, and the type curve calculated through various cluster analyzes. the output variable is the monthly oil production. Existing empirical models, the DCA and RNN models, were compared, and an optimal model was derived through hyperparameter tuning to improve the predictive performance of the model.

키워드

Tight oil, LSTM, GRU, Clustering analysis, Type curve

Ⅰ. 서 론

유전의 생산성 예측은 의사결정 및 개발방법, 투자규모에 상당한 영향을 주는 유전 탐사 및 개발계획에 있어 필수적이고 중요한 영역이다. 석유공학에서 미래 생산예측은 시뮬레이션, 감퇴곡선법, 물질평형법, 부피법 등 다양한 방법이 존재하는데, 이중 감퇴곡선법은 시간에 따른 생산량의 관계를 외삽법으로 방정식을 통해 간단히 예측이 가능하한데 전통적인 저류층과 비전통 저류층에서 생산되는 탄화수소 생산 메커니즘이 상이하여 적용되는 감퇴곡선법 종류가 각각 다르다[1]. 최근 머신러닝을 이용하여 치밀오일 저류층의 생산거동을 예측하는 연구들이 많이 수행되고 있다[2]. 이 기술은 비전통 저류층의 미래개발 계획 수립 시 다양한 목적에 부합하는 프록시 모델(proxy model)로 입력 및 출력의 선정에 따라 변경이 용이하여 공

* corresponding author

학적 기반의 모델링을 대체 및 보완할 수 있는 방법으로 부각되고 있다. 미래 생산거동 예측 문제에 시계열 자료 딥러닝 알고리즘인 LSTM(Long Short-Term Memory) 과 GRU(Gated Recurrent Unit)를 이용하여 오일 및 가스 생산량 예측 연구들이 수행되었다[3][4]. 본 연구에서는 수평시추가 적용된 치밀오일 저류층에서 미래 생산거동을 예측하기 위해 순환신경망인 LSTM과 GRU를 이용한 딥러닝 프록시 모델을 제안하였으며, 모델의 정확도를 향상시키기 위해 입력자료로 비지도학습 중 하나인 군집분석을 통해 군집 별 표준곡선(Type curve)을 입력 변수로 추가하여 모델의 정확도와 변수의 적용성을 검증하였다.

II. 생산성 예측 기술

1945년 Arps가 제안한 감퇴곡선법은 생산감퇴 경향을 모사한 경험식이다. 전통적인 감퇴곡선법은 운영조건이나 생산제약에 대한 변수를 반영하지 못하거나 숙련된 엔지니어의 경험에 의해 분석을 수행해야 하는 단점이 있다. 시뮬레이션 모델은 저류층의 정적, 동적 자료를 활용하여 지질 모델을 구현하고 다양한 조건에 따라 생산이 변화하는 양상을 모사할 수 있어 활용성이 높으나 경제적인 문제와 모델 입력자료의 획득 문제로 모든 저류층에 활용하기 힘든 제약이 존재한다. 최근 4차산업혁명 기술의 발달과 데이터의 폭증으로 데이터에 기반한 생산성 예측 모델 연구가 활발히 진행되고 있다. 이 중 일부는 디지털오일필드라는 개념으로 여러 가지 석유공학 연구에 머신러닝과 인공지능 기법이 도입되고 있다. 초기 머신러닝 지도학습 기반의 랜덤포레스트, 서포트벡터머신, 인공신경망 모델을 다양한 석유공학에서 구하고자 하는 목적함수를 출력자료로 하여 프록시 모델에 활용하였다. 딥러닝 알고리즘에 제안된 이후 다층 신경망(MLP), 순환신경망(RNN), 컨볼루션신경망(CNN) 등이 연구에 적용되었다. 생산성 예측 및 물리검층 자료의 경우 시간이나 심도에 따라 매개변수들이 취득되는데 순환신경망의 개념이 적용가능하며 본 연구에 예측기법으로 그림 1과 같은 LSTM, GRU 딥러닝 알고리즘을 활용하였다.

III. 학습 및 예측 결과

본 연구에서 사용한 데이터는 미국 노스다코타주 Bakken 치밀 오일 생산 저류층의 2011년부터 생산중인 9년 이상의 399개의 생산정 자료를 학습자료로 활용하였다. 입력자료 이상치 분석이나 결측치 대체 및 data smoothing, 분석 활용 가능여부를 확인결과 388개 생산정이 최종 선정되었다. 딥러닝 모델의 입력변수는 유정의 오일, 가스, 물, 그리고 생산거동 예측성능을 더욱 향상시키기 위해 군집분석을 통한 전체 학습자료의 군집 별 표준곡선을 입력자료로 활용하였다. 학습자료의 조건을 동일하게 하기 위해 90개월의 자료를 분석자료로 하고 이 중 63개월을 학습자료로 27개월을 테스트자료로 하여 실제 생산량과 비교하였다. 제안한 LSTM/GRU 기반 모델의 예측성능을 비교하기 위해 감퇴곡선법 및 LSTM, GRU 모델을 비교분석하였다. 테스트자료의 실제 생산량과 LSTM, GRU 알고리즘의 예측값을 비교하기 위해 회귀모델 평가지표를 사용하였다.

군집분석은 K-mean, K-medoids(PAM), 밀도기반 군집법(Density based spatial clustering: DBSCAN)을 이용하였으며 최적 군집개수를 산출하기 위해 silhouette 기법을 적용하였다. K-mean 군집의 경우 4개, PAM은 2개, DBSCAN 군집은 개수를 자동적으로 선정해주므로 요구되지 않으며, 분석결과 군집의 개수가 4개로 산출되었다. 하지만 DBSCAN은 silhouette 계수가 현저히 낮아 본 학습자료의 군집분석으로 용이하지 않은 것으로 판단되었다. 본 연구에서는 PAM 기법을 이용하여 군집을 수행하였으며 군집의 개수는 2개로 설정하였다.

표 1은 입력변수의 다양한 조합에 따른 5개의 케이스 모델이며, 각각의 생산 예측 모델의 평가지표 비교를 통해 최적의 입력변수 조합 케이스를 찾는 분석을 수행하였다.

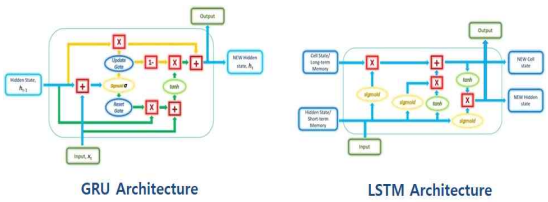

그림 1. 딥러닝 기반 순환신경망

표 1. 입력변수 종류에 따른 학습 모델 케이스

Type	Type of input variable
Case 1	Oil
Case 2	Oil+Gas
Case 3	Oil+Water
Case 4	Oil+Water+Gas
Case 5	Oil+Water+Gas+ Type curve(clustering)

그림 2는 Test 1 well의 생산거동 예측 결과 비교 결과이다. 군집 1의 표준곡선을 입력변수로 추가한 Case 5번의 경우 예측성능이 가장 우수하였으며, LSTM, GRU 알고리즘 비교결과 MAE가 각

각 32.60, 17.40 bbl이며, MAPE는 6.33%, 3.36%이다. 그림 2는 Test 2 well 비교 결과이고 군집 2의 표준곡선을 입력변수로 추가하였을 때 예측성능이 가장 우수한 것으로 확인되었다. LSTM, GRU 알고리즘 비교결과 MAE가 각각 67.22, 26.31 bbl이며, MAPE는 10.05%, 3.95%으로 두 테스트 자료 비교결과 군집분석을 통해 구현한 표준곡선 매개변수를 추가하고 GRU 알고리즘을 사용한 경우 모델의 예측성능이 MAE의 경우 최소 15 bbl, 최대 41 bbl, MAPE의 경우 최소 3%, 최대 7%이상 향상되는 것으로 나타났다.

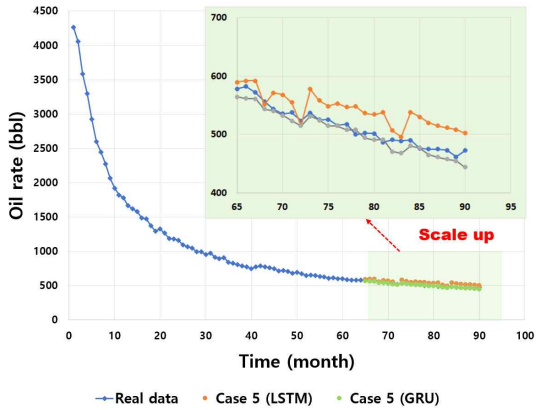

그림 2. Test 1 well의 생산거동 예측 결과 비교

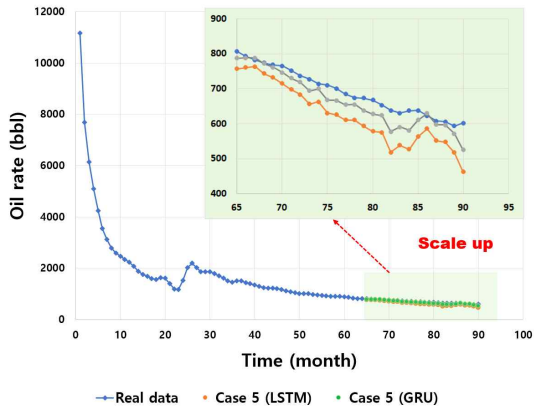

그림 3. Test 2 well의 생산거동 예측 결과 비교

Ⅳ. 결 론

이 연구에서는 군집분석을 통해 도출된 표준곡선을 입력변수로 추가하여 순환신경망인 딥러닝 기반 LSTM, GRU 알고리즘을 이용하여 미래생산거동을 예측할 수 있는 지도학습 모델을 제안하였다. 미국 Bakken 치밀오일의 현장자료를 데이터 전처리 기법을 통해 최종 388개의 유정을 학습자료로 선정하였다. PAM 군집분석을 통해 2개의 군집이 최적인 것으로 확인되었고, test data는 군집 별 상이한 표준곡선 입력변수를 적용하였다. 다양한 케이스의 입력변수 및 LSTM, GRU 알고리즘 민감도 분석을 통해 군집 별 표준곡선을 적용한 Case 5일 경우, GRU 알고리즘을 사용할 경우 모델의 예측성능이 가장 우수한 것으로 나타났다.

본 연구에서 개발한 모델을 더욱 발전시키기 위해 딥러닝 모델의 Regularization이나, CNN 알고리즘과의 결합모델, 주성분 분석과 같은 다른 비지도 학습 적용에 따라 을 통한 데이터 전처리를 적용하여 예측성능을 더욱 높이는 연구를 수행할 예정이다.

Acknowledgement

이 논문은 2021년도 정부(산업통상자원부)의 재원으로 해외자원개발협회의 지원(데이터사이언스 기반 석유·가스 탐사 컨소시엄) 2021년도 교육부와 한국연구재단의 지원(No. 2020R1I1A1A01060571)을 받아 수행한 연구 과제입니다

References

[1] D. Han, S. Kwon, H. Son, and J. Lee, “Production forecasting for shale gas well in transient flow using mahcine learning and decline curve analysis,” in *Proceeding of the SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference,* Brisbane, Australia, 18-19, Nov. 2019.

[2] D. Han, and S. Kwon, “Application of Machine Learning Method of Data-Driven Deep Learning Model to Predict Well Production Rate in the Shale Gas Reservoirs,” *Energies,* Vol. 14(12), 3629, 2021.

[3] K. Lee, J. Lim, D. Yoon, and H, Jung, “Prediction of Shale-Gas Production at Duvernay Formation Using Deep-Learning Algorithm,” *SPE Journal,* Vol. 24(06), pp. 2423-2437, 2019.

[4] X. Li, X. Ma, F. Xiao, S. Wang, S, and S. Zhang, “Application of Gated Recurrent Unit (GRU) Neural Network for Smart Batch Production Prediction,” *Energies,* Vol. 13, 6121, 2020.