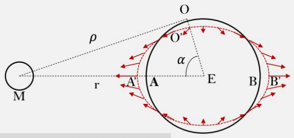
Extraction and Analysis of Earth Tide Signals

Xuhua Gao, Department of Energy Resources Engineering, xuhuag@stanford.edu

Introduction

- The gravitational forces between the Earth and the Sun and between the Earth and the Moon causes tidal deformation of the earth body.
- Both Earth tide and ocean tide are measurable, and this study focuses on Earth tide, i.e. tidal deformation of the solid Earth
- If the downhole pressure is measured in a closed well, then the periodical earth tide signals could be extracted from the pressure measurement.



We are interested in extracting earth tidal signals because it provides valuable information about the properties of subsurface formation.

$$p(t) = p_d(t) + \Delta p(t) + e(t)$$

- $p(t)$ is the measured downhole pressure
- $p_d(t)$ is the long-term trend (non-tidal part)
- $\Delta p(t)$ is the periodic pressure change that is caused by the earth tide, i.e. the tidal part
- $e(t)$ is the error part

- The left equation is a simplified equation that illustrates the relationship between the measured data and the tidal signal
- All the four variables are function of time and are treated as time-series data
- The objective is to separate the tidal part from the non-tidal part, or extract the tidal part from the raw measurements

Baseline

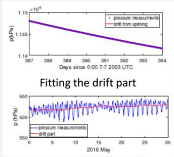
- The baseline result could be provided by traditional data filtering methods including:
  - Data cubic spline interpolation

$$p_o = a p_{ra}(t - t_{sh}) \qquad p_d(t) = \sum_{k=0}^{n} \theta_k(t) y_k$$

  - Low-pass filtering (LCF)
  - Savitzky–Golay filter (suitable for well test data)

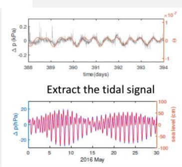
$$p_d(t) = \sum_{t'=t-M}^{t'=t+M} h(t-t')p(t')$$

Length of approximation interval: 2M+1



- Fitting the drift part: the appropriate long-term trend determined by traditional filtering methods for location 1 and location 2

- The residual tidal part is compared with theoretical tides.
- Baseline result performance:
  - Location 1 RMS error: 5.57
  - Location 2 RMS error: 1.42



Methods

- Different RNN models and architecture are tested for the purpose of extracting tidal signal & separating tidal part from the non-tidal parts
  - RNN with LSTM + dropout regularization
  - Nonlinear autoregressive neural network with external input (NARX)

$$y(t) = \alpha_0 + \alpha_1 y(t-1) + \lambda \tanh(\gamma(y(t-1) - c)) + \varepsilon_t$$

$$y(t) = f(y(t-1), \ldots, y(t-d))$$ Without external features (NAR)

$$y(t) = f(x(t-1), \ldots, x(t-d), y(t-1), \ldots, y(t-d))$$ w. external input (X(t))

- The training and hyperparameter search process follows the steps:

pre-train	error analysis	fine tune	error analysis
Train the model to learn theoretical tidal variation pattern and predict theoretical tide	Compare training error, validation error and test error	Input real data and fine tune the model learn to separate tidal part from non-tidal part	Compare training error, validation error and test error

first cycle → second cycle

References

[1] Burbey, T.J. (2010) Fracture characterization using Earth tide analysis *Journal of Hydrology* 380(3-4):237-246.

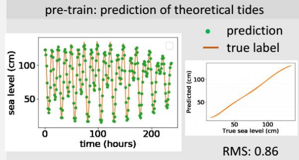
[2] Mentes, G. (2015). Artificial neural network model as a potential alternative for barometric correction of extensometric data. *MARÉES TERRESTRES: BULLETIN D INFORMATIONS*, 149, 12001-12012.

[3] Tian, C., & Horne, R. N. (2017). Recurrent neural networks for permanent downhole gauge data analysis. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

[4] Wada, Y., Pecharanin, N., Taguchi, A., Iijima, N., Akima, Y. and Sone, M. (1995). Application of recurrent neural network for active filter. In *Neural Networks Proceedings*., IEEE International Conference(Vol. 1, pp. 488-491).

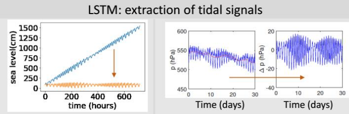
Results

pre-train: prediction of theoretical tides



RMS: 0.86

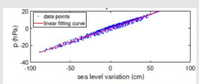
LSTM: extraction of tidal signals



input / output — Simplest case: tide + linear trend

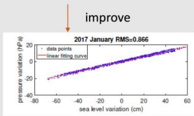
Input → output — Real data case: complex trend

Error analysis: LSTM



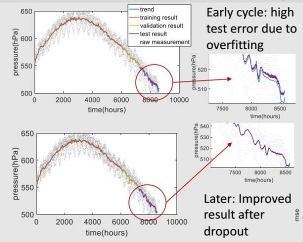
Base line RMS: 1.42

improve



LSTM RMS: 0.88

NARX: extraction of tidal signals



Early cycle: high test error due to overfitting

Later: Improved result after dropout

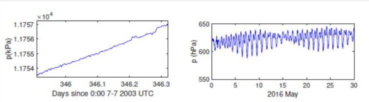
Comparison between NARX and LSTM:

	LSTM	NARX
Pre-train	good	good
Simple case	good	good
Complex case	Mode-rate	good

NARX is generally better, but need to avoid overfitting

Dataset and Preprocessing

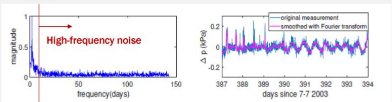
- Data come from two locations, tidal variation at location 1 has a very small amplitude and larger change in long-term trend while tidal variation at location 2 are large and the long-term trend is smoother



(a) sample data from location 1 (b) sample data from location 2

- The difference in amplitude is due to that location 1 is on-shore reservoir while location 2 is off-shore reservoir. The effect of ocean tide is more significant for location 2.

- The data from location 1 is relatively more noisy compared with data from location 2, and the high-frequency noise could be removed by transforming the data into frequency domain through Fourier transform.



(a) frequency domain (b) smoothed data with Fourier transform

- The smoothed data (ingestible format) demonstrate a clearer cyclic tidal variation pattern, making it more convenient to compare with theoretical tides.