

Способ темпоральной интерполяции толщины подвергающейся коррозии стенки газопровода согласованной с физической моделью

Р.Р. Габбасов

Самарский национальный исследовательский университет
им. академика С.П. Королева
Самара, Россия
rklug@mail.ru

Р.А. Парингер

Самарский национальный исследовательский университет
им. академика С.П. Королева
ИСОИ РАН
Самара, Россия
rusparinger@gmail.com

Аннотация—Анализ развивающихся во времени процессов играет с развитием вычислительных мощностей всё более важную роль в современном мире. В данной работе рассматривается процесс коррозионного износа стенки газопровода, а именно, задача регрессии значения толщины стенки трубы. Предлагается новый способ интерполяции во времени значений толщины стенки, производимой в согласовании с физическими показателями транспортируемого газового конденсата. Проводятся эксперименты по машинному обучению регрессионных моделей с использованием алгоритма RANSAC, вводятся определения двух мер соответствия обученных моделей физической реальности. Результаты экспериментов показали, что использование предлагаемого способа интерполяции вместо интерполяции сплайнами позволяет добиться увеличения значения первой меры в среднем в 2 раза, а значения второй меры – в 3 раза.

Ключевые слова— **ВРЕМЕННЫЕ РЯДЫ, КОРРОЗИОННЫЙ ИЗНОС, ЗАДАЧА РЕГРЕССИИ, RANSAC**

1. ВВЕДЕНИЕ

Актуальной во многих областях деятельности человека, в частности, в контексте анализа временных рядов [1, 2, 3] является задача предсказания (регрессии) значений. В данной работе речь идет о регрессии уменьшающейся вследствие коррозии толщины труб, привязанных к двум скважинам одного газового месторождения. Обвязки скважин, рассматриваемых в нашей работе, оборудованы датчиками, снимающими показания физических характеристик проходящего конденсата. Данные с этих датчиков используются в качестве признаков для регрессионных моделей. Замеры толщины производятся в разных местах в разное время на обвязке скважины с помощью ультразвуковой диагностики. Мы предлагаем способ интерполяции толщин во времени, учитывающий физические характеристики потока наряду с данными о промежуточных значениях толщин. Для оценки эффективности применения предлагаемого способа наряду с ним мы используем интерполяцию сплайнами.

2. ОПИСАНИЕ И ПОДГОТОВКА НАБОРА ДАННЫХ

В данной работе использовались данные с двух скважин (с названиями «2-2» и «3-1») одного месторождения. Для каждой из них имелись значения 17-ти изменяющихся со временем параметров, снятые с шагом в 1 час с датчиков, расположенных на трубной обвязке скважины: значения давлений и температур в разных местах, содержания в конденсате CO₂ и его pH. Данные параметры выступали в качестве входных признаков в процессе обучения. Процесс сглаживания

выбросов и заполнения пропусков в этих данных был осуществлён с использованием скользящего окна [4]. Также для каждой скважины имелись значения целевого параметра – сильно разнесенные по времени результаты замеров толщины стенки компонентов обвязки в разных местах обвязки в разное время. В данной работе было рассмотрено два способа интерполяции этого признака (с целью соответствия частоте дискретизации признаков): с использованием квадратичных сплайнов [5] и **предлагаемый нами способ интерполяции**, производимой в согласовании с **физической моделью**, т.е. с **физическими показателями** транспортируемого газового конденсата. Для обоих способов интерполяции соответственно были подготовлены два набора данных для дальнейших экспериментов.

А. Интерполяция квадратичными сплайнами

Характеристика интенсивности коррозионного процесса связана со **скоростью изменения толщины стенки**, поэтому модели обучались регрессировать значения этой скорости. Так как скорость изменения параметра является производной функции параметра, использование линейной интерполяции привело бы к вырождению скорости изменения толщины между двумя известными исходными временными метками в константу, поэтому используется квадратичная интерполяция.

Б. Интерполяция, согласованная с физическими показателями

Данный способ интерполяции, который мы предлагаем, основан на использовании расчетов по стандарту NORSOK M-506 [6]. Данный стандарт позволяет рассчитать теоретическую скорость коррозии трубопровода (в мм/год) в фиксированный момент времени на основе значений давления, температуры и pH потока, содержания в нем CO₂, а также значений диаметра, шероховатости и напряжения сдвига стенки трубы. Имея эти данные, мы по данной методике получили значения теоретической скорости утонения стенки, а затем использовали их в качестве весов интенсивности истончения между двумя исходными замерами.

3. УСЛОВИЯ ЭКСПЕРИМЕНТОВ

Так как рассматриваемые данные являют собой временной ряд, то для каждого целевого значения рассматривалось временное окно признаков размера 3600 (3600 часов = 150 суток = 5 месяцев). В качестве регрессионной модели в данной работе используется модель линейной регрессии с дополнительным использованием алгоритма RANSAC [7] со следующими гиперпараметрами: минимальная доля выбираемых случайных элементов – 0,1, функция ошибок –

квадратичная. Также рассматривались модели регрессора с функцией потерь Хьюбера [8] и регрессора Тейла-Сена [9], однако было установлено, что их обучение занимает слишком много времени относительно времени обучения регрессора с использованием RANSAC при сравнимых с точки зрения точности результатах. Для каждой точки на трубе была обучена собственная регрессионная модель.

4. РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

В данном разделе представлено сравнение результатов обучения регрессионных моделей для двух рассматриваемых в работе скважин («2-2» и «3-1») в случае использования двух представленных способов интерполяции значения толщины стенки.

А. Схожесть моделей

Мера схожести между двумя моделями высчитывалась следующим образом. С помощью обеих моделей на временном участке между двумя соседними исходными моментами замера толщины (входящих в тренировочную выборку) по 17-ти признакам были регрессированы валидационные значения скорости изменения толщины стенки. Мера схожести моделей определялась как максимум между 0 и значением метрики R^2 между двумя получившимися векторами значений. Далее модели со схожестью более 0,9 между собой объединялись в отдельные группы. Первая мера соответствия обученных моделей физической реальности (M_1) определялась как средний размер таких групп моделей (так как физическая связность обвязки ведет к схожести в поведении моделей для разных точек на ней).

Таблица I. РЕЗУЛЬТИРУЮЩИЕ ЗНАЧЕНИЯ МЕРЫ M_1

Скважины	2-2	3-1
Интерполяция квадратичными сплайнами	6,87	4,32
Интерполяция, согласованная с физическими параметрами	10,23	10,57
Соотношение	1,49	2,45

Б. Относительная важность признаков

Была оценена относительная значимость признаков для регрессии утонения стенок в точках на различных участках обвязки: около устья, до штуцера, после штуцера. Исходные признаки были сформированы в шесть наборов в зависимости от местоположения соответствующих датчиков на обвязке скважины. Затем для каждой точки замера на обвязке был обучен набор из шести регрессионных моделей, обученных на соответствующих наборах признаков. Далее модели были поделены на три класса: около устья, до штуцера, после штуцера – в зависимости от местоположения соответствующей точки замера на обвязке. Затем для каждого класса для каждого набора признаков было подсчитано среднее по соответствующим моделям значение метрики L_1 между регрессированными моделями и исходными значениями целевого параметра на тестовой выборке. Полученные значения нормализовывались и инвертировались таким образом, чтобы наименьшее значение L_1 соответствовало 1, а большие значения L_1 – числам меньше 1. Полученные числа и являются значимостью признаков. Вторая мера

соответствия обученных моделей физической реальности (M_2) определялась на основе полученных значений значимости признаков следующим образом:

$$M_2 = (L_1(\mathbf{i}, \mathbf{i}'))^{-1},$$

где $L_1(*,*)$ – метрика L_1 , $\mathbf{i} \in \mathbb{R}^{18}$ – вектор значимостей признаков для различных классов (6 наборов признаков \times 3 класса моделей), $\mathbf{i}' \in \mathbb{R}^{18}$ – вектор, состоящий из одного и того же значения, являющегося медианой вектора \mathbf{i} .

Таблица II. РЕЗУЛЬТИРУЮЩИЕ ЗНАЧЕНИЯ МЕРЫ M_2

Скважины	2-2	3-1
Интерполяция квадратичными сплайнами	0,31	0,24
Интерполяция, согласованная с физическими параметрами	0,77	0,83
Соотношение	2,48	3,46

5. ЗАКЛЮЧЕНИЕ

В данной статье был предложен способ интерполяции во времени значений толщины стенки компонентов обвязки газовых скважин, производимой в согласовании с физическими параметрами проходящего конденсата. Были проведены эксперименты по обучению моделей линейной регрессии с использованием алгоритма RANSAC. Были введены определения двух мер соответствия моделей физической реальности. Результаты показывают, что использование предлагаемого способа интерполяции вместо интерполяции квадратичными сплайнами приводит к повышению меры соответствия моделей физической реальности: значение меры M_1 повысилось в 1,49 и 2,45 раза для скважин 2-2 и 3-1 соответственно, а для M_2 соответствующие повышения составили 2,48 и 3,46 раз. Стоит отметить, что согласованность результатов для обеих предложенных мер говорит о высоком уровне их достоверности.

ЛИТЕРАТУРА

- [1] Boori, M.S. Crop growth monitoring through Sentinel and Landsat data based NDVI time-series / M. S. Boori, K. Choudhary, A. V. Kupriyanov // Computer Optics. – 2020. – Vol. 44(3). – P. 409-419.
- [2] Терехин Э. А. Индикация многолетних изменений в растительном покрове залежных земель лесостепи на основе рядов вегетационного индекса NDVI / Э. А. Терехин // Компьютерная оптика. – 2021. – Т. 45, №. 2. – С. 245-252.
- [3] Plotnikov, D. Daily surface reflectance reconstruction using LOWESS on the example of various satellite systems / D. Plotnikov [et al.] // 2022 VIII International Conference on Information Technology and Nanotechnology (ITNT). – IEEE. – 2022. – P. 1-5.
- [4] Yu, Y. Time series outlier detection based on sliding window prediction / Y. Yu [et al.] // Mathematical problems in Engineering. – 2014. – Vol. 2014.
- [5] Sharma, A. Quadratic splines / A. Sharma, J. Tzimbalaro // Journal of Approximation Theory. – 1977. – Vol. 19(2). – P. 186-193.
- [6] NORSOK M-506. CO2 corrosion rate calculation model.
- [7] Derpanis, K.G. Overview of the RANSAC Algorithm / K.G. Derpanis // Image Rochester NY. – 2010. – Vol. 4(1). – P. 2-3.
- [8] Huber, P.J. Robust regression: asymptotics, conjectures and Monte Carlo / P.J. Huber // The annals of statistics. – 1973. – P. 799-821.
- [9] Wilcox, R. A note on the Theil-Sen regression estimator when the regressor is random and the error term is heteroscedastic / R. Wilcox // Biometrical Journal: Journal of Mathematical Methods in Biosciences. – 1998. – Vol. 40(3). – P. 261-268.