# $CO_2$ Injection Rate Predictive Analysis

## Resources

- Validation and Submission Notebook: GitHub Notebook
- SPE GCS Event Details: Event Page
- CO2ML Challenge Repository: GitHub Repo

## 1 Problem Description

This challenge aims to use time series injection information and monitoring data on a carbon capture well to predict carbon capture well injection rate deltas. Correlating the change in injection rate to the behavior of other parameters in the well can be used to provide a checkpoint against carbon migration from the well or other losses during the process. The code developed to predict injection rate deltas based on monitoring well data can be used to validate carbon containment throughout the injection of the well.

## 2 Feature Selection

The following approach will help identify features directly related to predicting `inj_diff`:

### 2.1 Correlation Analysis

Perform **Pearson correlation** analysis to identify features strongly correlated with `inj_diff`:

$$r_{X,Y} = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2 \, \sum_i (Y_i - \bar{Y})^2}}.$$

Features like `Avg_PLT_CO2VentRate_TPH`, `Avg_CCS1_WHCO2InjPs_psi`, and `Avg_CCS1_WHCO2InjTp_F` are likely important as they directly describe $CO_2$ injection rates and pressure/temperature dynamics.

### 2.2 Domain-Driven Importance

Features likely related to well dynamics and injection deltas include:

#### Pressure-Related Parameters

- `Avg_CCS1_WHCO2InjPs_psi` (wellhead pressure)
- `Avg_VW1_ANPs_psi` (annulus pressure)
- Deep well pressures such as `Avg_CCS1_DH6325Ps_psi` or related subzones (Z11D4917 to Z01D7061)

#### Temperature-Related Parameters

- `Avg_CCS1_WHCO2InjTp_F` (wellhead temperature)

- Subsurface zone temperatures: `Avg_VW1_*`$_Tp\_F$

**Vent Rates**

- `Avg_PLT_CO2VentRate_TPH` as a proxy for $CO_2$ migration or losses

**Additional Steps**

- Drop low-variance features (near-zero variance over time)
- Test feature combinations such as the ratio of pressure to temperature for better physical insights

# 3 Rolling Aggregates Transformations

For time-series data, rolling features capture temporal patterns:

## 3.1 Recommended Aggregates

- Mean and Standard Deviation to capture trends and volatility
- Max/Min for extremes related to sudden injection changes
- Lag Features: Include lags of 1, 3, 6, and 12 hours for past behavior influences

## 3.2 Window Size

If data is recorded hourly, use 6-hour, 12-hour, and 24-hour rolling windows to capture short-term and daily patterns.

## 3.3 Standardization

Perform feature standardization (z-score normalization) on rolling aggregates to normalize scales.

## 3.4 Example Transformed Features

For `Avg_CCS1_WHCO2InjPs_psi`:

- `rolling_mean_6h(Avg_CCS1_WHCO2InjPs_psi)`
- `rolling_std_6h(Avg_CCS1_WHCO2InjPs_psi)`
- `lag_1h(Avg_CCS1_WHCO2InjPs_psi)`, `lag_3h(...)`

# 4 Kernel Variables with Physics Basis

## 4.1 Compressibility Effects

$$\Delta P_{\text{comp}} = P_{\text{wellhead}} - P_{\text{subsurface}}, \tag{1}$$
$$\Delta T_{\text{comp}} = T_{\text{wellhead}} - T_{\text{subsurface}}. \tag{2}$$

## 4.2 Flow Dynamics

Mass flow proxy:

$$\text{mass\_flow} = \frac{\text{vent\_rate} \times \text{pressure}}{\text{temperature}}.$$

Injection efficiency:

$$\text{efficiency} = 1 - \frac{\text{vent\_rate}}{\text{vent\_rate} + \text{injection\_rate}}.$$

### 4.3 Thermal Gradients

Zone-specific gradient:

$$\Delta T_{\text{zone}} = T_{\text{wellhead}} - T_{\text{zone}}.$$

# 5 Recommended Next Steps

1. **Feature Engineering:** Generate rolling, lag, and physics-based features; drop less-correlated variables.
2. **Modeling Strategy:** Train tree-based models (XGBoost, LightGBM) with cross-validation.
3. **Distribution Assessment:** If `inj_diff` is skewed, apply log or Box–Cox transforms.
4. **Dimensionality Reduction:** Consider PCA when multicollinearity is high.
5. **Evaluation Metrics:** Use MAE and RMSE; inspect residuals for heteroscedasticity.

# 6 Definition and Importance of `inj_diff`

## 6.1 What is `inj_diff`?

- `inj_diff` represents **changes in the $CO_2$ injection rate** over time.
- It tells us if the amount of $CO_2$ being injected into the well is increasing, decreasing, or remaining stable.

## 6.2 Why Predict `inj_diff`?

Predicting `inj_diff` is important for **real-time monitoring** and **problem prevention**. Here's why:

### 6.2.1 Ensuring Carbon Containment

- When $CO_2$ is injected into a well, we need to ensure it stays underground and doesn't migrate or escape to the surface.
- Sudden, unexpected changes in `inj_diff` could indicate problems like:
  1. **Leakage:** $CO_2$ might escape from the well into nearby rock formations or the atmosphere.
  2. **Pressure Imbalances:** The well might be experiencing issues that could damage its structure or cause leaks.

### 6.2.2 Process Optimization

- If `inj_diff` changes too much, it could mean the injection process isn't efficient, leading to:
  1. **Wasted Energy:** Pumping $CO_2$ at inconsistent rates wastes energy and resources.
  2. **Suboptimal Storage:** The well might not be storing as much $CO_2$ as it should.

### 6.2.3 Early Warning for Failures

- Predicting changes in injection rates helps detect potential failures or risks **before they occur**.
  - For example, a drop in `inj_diff` might indicate clogging in the injection system.
  - A sudden spike might suggest cracks or leaks in the well casing.

# 7 How Do We Benefit?

- **Environmental Protection:** CCS operations are designed to reduce $CO_2$emissions and fight climate change. By monitoring inj_diff, we can ensure $CO_2$remains safely stored underground.
- **Operational Safety:** Avoiding leaks and pressure imbalances reduces the risk of accidents and protects nearby ecosystems.
- **Cost Efficiency:** Predicting and addressing problems early saves money by preventing expensive repairs or system failures.
- **Regulatory Compliance:** Accurate monitoring ensures the operation meets legal and environmental standards.

# 8 How Does This Relate to Your Work?

The purpose of the analysis is to:

- Use historical injection data to predict future changes in injection rates (inj_diff).
- Correlate these predictions with pressure, temperature, and vent rate data to identify **why injection rates are changing.**
- Provide real-time feedback to operators to adjust the process and avoid issues like $CO_2$leakage or system inefficiency.

## 8.1 Wellhead Parameters:

- **Avg_CCS1_WHCO2InjPs_psi** (Wellhead Injection Pressure)
- **Avg_CCS1_WHCO2InjTp_F** (Wellhead Injection Temperature)

These are **wellhead** measurements, recorded at the surface of the well where $CO_2$injection begins. They provide insights into the pressure and temperature at the point where $CO_2$enters the well system.

## 8.2 Downhole (Completion System) Parameters:

- **Avg_CCS1_ANPs_psi** (Annulus Pressure)
- **Avg_VW1_ANPs_psi** (Annulus Pressure for another well or zone)

These parameters are measured in the **annulus**, the space between the well casing and the tubing (or between different casings). Changes in annulus pressure can indicate:

- Well integrity issues (e.g., leaks).
- Reservoir behavior if communication between zones exists.
- **Avg_CCS1_DH6325Ps_psi**, **Avg_CCS1_DH6325Tp_F**
  These are likely **downhole (DH)** measurements in the completion system at a specific depth (e.g., 6325 feet). These parameters monitor wellbore pressure and temperature as $CO_2$moves deeper into the well.

## 8.3 Zone-Specific Parameters:

The **VW1_Z...** parameters seem to represent measurements at specific zones (or depths) within the well or reservoir:

- **Avg_VW1_Z11D4917Ps_psi**, **Avg_VW1_Z11D4917Tp_F**, etc.
  - These are zone-specific measurements (e.g., Z11, Z10, Z09, etc.), capturing the **pressure (Ps)** and **temperature (Tp)** at various depths.
  - These depths could correspond to perforated intervals or distinct reservoir zones.
- These measurements are critical to understanding:
  - How pressure and temperature vary between zones.
  - Whether CO2injection is affecting multiple zones differently.

## 8.4 Near-Wellbore Reservoir Interaction:

- **Avg_VW1_WBTbgPs_psi**, **Avg_VW1_WBTbgTp_F**
  These seem to be **near-wellbore reservoir conditions** (e.g., at the wellbore-tubing interface). They measure:
  - **Pressure:** Indicates how CO2injection impacts the reservoir near the wellbore.
  - **Temperature:** Provides information on thermal effects due to CO2injection.

## 8.5 How This Data is Useful:

1. **Surface Monitoring:**
   Wellhead parameters (WHCO2InjPs, WHCO2InjTp) are useful for operational monitoring and ensuring the injection rate is consistent.
2. **Well Integrity:**
   Annulus pressure (ANPs) measurements help detect leaks or abnormal pressure buildup that could indicate well integrity issues.
3. **Reservoir Behavior:**
   Zone-specific data (Z11D4917Ps, Z10D5001Ps, etc.) help monitor the injection's effect at different depths or reservoir layers. For example:
   - If pressure rises too quickly in a specific zone, it could signal **fracturing** or **poor permeability**.
   - A sudden drop in pressure in one zone might indicate **leakage** into an unintended zone.
4. **Thermal Effects:**
   Temperature measurements (Tp) are critical because injecting CO2 (often at lower temperatures) can create significant thermal gradients, affecting rock mechanics and CO2containment.

## 8.6 Conclusion:

These variables are indeed measured within the well system and reservoir, offering a detailed view of how CO2injection impacts the **wellbore**, **completion system**, and **reservoir zones**. By analyzing these measurements, you can:

- Predict changes in injection rates (inj_diff).
- Correlate these changes with well conditions to detect potential problems like **leakage**, **blockages**, or **inefficiencies**.
- Gain insights into reservoir performance and CO2containment efficiency.

# 9 Pressure Difference Kernels

## 9.1 (a) Injection Pressure Drop in the Completion System

$$\Delta P_{\text{completion}} = \text{Avg\_CCS1\_WHCO2InjPs\_psi} - \text{Avg\_CCS1\_DH6325Ps\_psi}. \tag{3}$$

This measures the pressure drop between the wellhead and the downhole sensor at 6325ft. A large drop may indicate flow resistance, blockage, or leaks.

## 9.2 (b) Zone-Specific Pressure Differences

$$\Delta P_{\text{zone}} = \text{Avg\_VW1\_WBTbgPs\_psi} - \text{Avg\_VW1\_Z11D4917Ps\_psi}. \tag{4}$$

This measures the pressure drop from the wellbore to specific zones (e.g., Zone 11). Substitute other zone pressure features to compute for each zone:

# 10 Thermal Gradient Kernels

## 10.1 (a) Temperature Gradient: Wellhead to Completion

$$\Delta T_{\text{completion}} = \text{Avg\_CCS1\_WHCO2InjTp\_F} - \text{Avg\_CCS1\_DH6325Tp\_F}. \tag{5}$$

A sharp drop suggests rapid heat exchange or cooling during injection.

## 10.2 (b) Zone-Specific Temperature Gradients

$$\Delta T_{\text{zone}} = \text{Avg\_VW1\_WBTbgTp\_F} - \text{Avg\_VW1\_Z11D4917Tp\_F}. \tag{6}$$

Similar to pressure differences, these gradients can indicate uneven thermal behavior across zones.

# 11 $CO_2$ Flow Efficiency Kernels

## 11.1 (a) Injection Efficiency

$$\text{Efficiency}_{\text{CO2}} = 1 - \frac{\text{Avg\_PLT\_CO2VentRate\_TPH}}{\text{Avg\_PLT\_CO2VentRate\_TPH} + \text{Avg\_CCS1\_WHCO2InjPs\_psi}}. \tag{7}$$

This compares the venting $CO_2$ rate to the injection pressure, indicating how efficiently $CO_2$ is being injected into the reservoir.

## 11.2 (b) Flow Resistance Proxy

$$R_{\text{flow}} = \frac{\Delta P_{\text{completion}}}{\text{Avg\_CCS1\_WHCO2InjPs\_psi}}. \tag{8}$$

Higher values can indicate blockages or suboptimal flow dynamics.

# 12 Annulus Integrity Kernels

## 12.1 (a) Pressure Imbalance in Annulus

$$\Delta P_{\text{annulus}} = \text{Avg\_CCS1\_ANPs\_psi} - \text{Avg\_VW1\_ANPs\_psi}. \tag{9}$$

A large imbalance might suggest a leak or communication between the tubing and annular spaces.

## 12.2 (b) Annulus Thermal Gradient

$$\Delta T_{\text{annulus}} = \text{Avg\_CCS1\_WHCO2InjTp\_F} - \text{Avg\_VW1\_WBTbgTp\_F}. \tag{10}$$

Changes in annulus temperature can indicate fluid movement or thermal anomalies.

## 13 Reservoir Communication Kernels

### 13.1 (a) Pressure Differences Between Zones

$$\Delta P_{\text{zones}} = \text{Avg\_VW1\_Z11D4917Ps\_psi} - \text{Avg\_VW1\_Z10D5001Ps\_psi}. \tag{11}$$

This measures the communication (or lack thereof) between adjacent zones. Extend this to other pairs of zones.

### 13.2 (b) Temperature Differences Between Zones

$$\Delta T_{\text{zones}} = \text{Avg\_VW1\_Z11D4917Tp\_F} - \text{Avg\_VW1\_Z10D5001Tp\_F}. \tag{12}$$

Similar to pressure differences, this can highlight thermal anomalies between zones.

## 14 Combined Thermal–Pressure Kernels

### 14.1 (a) Wellhead Pressure–Temperature Ratio

$$\frac{P}{T}_{\text{wellhead}} = \frac{\text{Avg\_CCS1\_WHCO2InjPs\_psi}}{\text{Avg\_CCS1\_WHCO2InjTp\_F}}. \tag{13}$$

### 14.2 (b) Zone-Specific P/T Ratio

$$\frac{P}{T}_{\text{zone}} = \frac{\text{Avg\_VW1\_Z11D4917Ps\_psi}}{\text{Avg\_VW1\_Z11D4917Tp\_F}}. \tag{14}$$

## 15 Leak Detection Kernels

### 15.1 (a) Mass Flow Proxy

$$\text{Mass Flow} = \frac{\text{Avg\_PLT\_CO2VentRate\_TPH} \cdot \text{Avg\_CCS1\_WHCO2InjPs\_psi}}{\text{Avg\_CCS1\_WHCO2InjTp\_F}}. \tag{15}$$

A sharp drop in this proxy might indicate leakage or inefficiency in the system.

### 15.2 (b) Zone-Specific Leak Detection

$$\Delta \text{Leak}_{\text{zone}} = \text{Avg\_VW1\_Z11D4917Ps\_psi} - \text{Avg\_VW1\_Z10D5001Ps\_psi}. \tag{16}$$

Large pressure discrepancies between zones might indicate unintended fluid migration.