

3. Methodology

3.6 Simulation Procedure

The simulation procedure was designed to systematically evaluate leak detection capabilities under controlled conditions while mimicking real-world pipeline operations. The process followed a structured approach from system initialization through data collection, ensuring reproducibility and consistency across all test scenarios.

Step 1: System Setup and Configuration

The initial phase involved configuring the simulation environment to represent a typical natural gas transmission pipeline. A 20-mile pipeline segment was selected as it represents a common distance between compression stations in natural gas networks. The pipe diameter of 12 inches was chosen based on typical medium-capacity transmission lines operating in Nigeria's gas distribution infrastructure.

Pipeline Physical Properties:

- **Length:** 20 miles (32.19 km) - representing a typical inter-station segment
- **Internal Diameter:** 12 inches (304.8 mm) - standard NPS 12 pipeline
- **Material:** Carbon steel (API 5L Grade B)
- **Absolute Roughness:** 0.0006 m (0.6 mm) - representing moderate pipeline aging with light corrosion and scale deposits

The roughness value was selected based on typical operational pipelines rather than new construction, as most leak detection challenges occur in aging infrastructure. This value sits between new pipe roughness (0.045 mm) and heavily corroded pipes (1.5 mm), representing realistic field conditions.

Operating Conditions: The pressure boundary conditions were set to simulate normal operational parameters:

- **Inlet Pressure (P):** 100 psi (689.5 kPa) - typical gathering system pressure
- **Outlet Pressure (P):** 98 psi (675.6 kPa) - representing a 2 psi drop across 20 miles
- **Operating Temperature:** 60°F (15.6°C / 288.7 K) - standard reference temperature
- **Ambient Pressure:** 14.7 psi (101.3 kPa) - sea level atmospheric pressure

These conditions were deliberately kept within moderate ranges to ensure the simulation remained applicable to both onshore and offshore Nigerian gas operations, where pressures typically range from 50-150 psi in distribution networks.

Fluid Properties: Methane (CH_4) was selected as the working fluid to represent natural gas in its simplest form. While actual natural gas contains varying compositions of ethane, propane, and heavier hydrocarbons, methane constitutes approximately 85-95% of Nigerian natural gas composition. Using pure methane simplified thermodynamic calculations while maintaining physical realism.

The CoolProp thermodynamic library was integrated to provide accurate real-gas properties:

- **Density (ρ):** Calculated dynamically based on pressure and temperature using the Peng-Robinson equation of state
- **Dynamic Viscosity (μ):** Temperature and pressure-dependent, obtained from CoolProp correlations
- **Compressibility Factor (Z):** Real gas deviation from ideal behavior, critical for accurate flow calculations
- **Molecular Weight:** 16.04 g/mol for methane

The advantage of using CoolProp over simplified ideal gas assumptions was particularly evident at higher pressures where gas compressibility becomes significant. For our operating conditions, the compressibility factor Z ranged from 0.92-0.95, representing approximately 5-8% deviation from ideal gas behavior. This level of accuracy is essential for realistic flow predictions.

Step 2: Baseline Run (No-Leak Reference Condition)

Before introducing any leaks, a baseline simulation was conducted to establish reference operational parameters. This step served multiple critical purposes: validating the computational model against expected flow equations, providing a comparison benchmark for leak scenarios, and ensuring the system reached steady-state equilibrium.

Baseline Simulation Process:

The system was initialized with the configured pipeline geometry and fluid properties. The flow solver automatically selected the Modified Panhandle A equation based on the pipeline length (>20 miles) and diameter (12 inches). This selection followed the criteria outlined in Section 3.4 of the methodology, ensuring industry-standard calculations appropriate for long-distance gas transmission.

The steady-state flow rate was calculated using:

$$Q_{baseline} = 435.87 \cdot \left(\frac{T_{sc}}{P_{sc}} \right)^{1.0788} \cdot \left(\frac{P_1^2 - P_2^2}{SG^{0.8539} \cdot L \cdot T \cdot Z} \right)^{0.5394} \cdot D^{2.6182} \cdot E$$

With pipeline efficiency $E = 0.95$ (accounting for minor fittings, bends, and surface irregularities), the baseline flow rate stabilized at approximately 12.8 million standard cubic feet per day (MMscfd) or 8.5 ft³/s at operating conditions.

Parameters Recorded During Baseline:

A comprehensive dataset was captured to establish the reference state:

Flow Measurements:

- **Inlet Volumetric Flow Rate:** 8.5 ft³/s (0.241 m³/s)
- **Outlet Volumetric Flow Rate:** 8.5 ft³/s (identical to inlet, confirming no leaks)
- **Mass Flow Rate:** 12.3 lb/s (5.58 kg/s) - calculated as $\rho \times Q$
- **Reynolds Number:** 4.2×10^6 - confirming fully turbulent flow regime

Pressure Distribution:

- **Inlet Pressure:** 100.0 psi (as set)
- **Outlet Pressure:** 98.0 psi (as set)
- **Pressure Drop (ΔP):** 2.0 psi across 20 miles
- **Pressure Gradient:** 0.1 psi/mile - typical for moderate flow rates

Temperature Measurements:

- **Inlet Temperature:** 60°F (15.6°C)
- **Outlet Temperature:** 59.8°F (15.4°C) - slight cooling due to Joule-Thomson effect
- **Average Gas Temperature:** 59.9°F for property calculations

Thermodynamic Properties:

- **Gas Density at Inlet:** 0.0734 lb/ft³ (1.176 kg/m³)
- **Gas Density at Outlet:** 0.0720 lb/ft³ (1.153 kg/m³) - slight expansion due to pressure drop
- **Dynamic Viscosity:** 1.09×10^{-4} Pa · s
- **Compressibility Factor (Z):** 0.937 - indicating real gas behavior

Validation Against Analytical Models:

The baseline results were cross-checked against the Weymouth equation and Darcy-Weisbach formulation to ensure computational accuracy. The Modified Panhandle A result matched Weymouth predictions within 3%, which is acceptable given differences in efficiency factors and friction models. The Darcy-Weisbach equation, calculated with a friction factor of 0.0089 (from Colebrook-White), yielded similar pressure drops, confirming model consistency.

SCADA Interface During Baseline:

The visual SCADA interface displayed the baseline condition with:

- All pressure gauges showing steady values (no fluctuations)
- Flow meters at inlet and outlet displaying identical readings
- Leak rate meter reading zero (confirming perfect flow balance)
- Pipeline visualization showing uniform green coloring (normal pressure throughout)
- No leak indicators or alarm conditions present

This baseline state became the reference against which all subsequent leak scenarios were compared. The flow balance equation was verified:

$$Q_{in} = Q_{out} + Q_{leak,total}$$

With $Q_{leak,total} = 0$, the equation held perfectly: $8.5 = 8.5 + 0 \text{ ft}^3/\text{s}$.

Step 3: Leak Simulation (Systematic Failure Scenarios)

The core experimental phase involved introducing controlled leaks of varying sizes and locations to systematically evaluate detection capabilities. This approach allowed us to map the relationship between leak characteristics (size, position) and detection metrics (flow imbalance, pressure drop, response time).

Leak Size Selection:

Four leak diameters were chosen to represent the spectrum of possible pipeline failures:

1. **3 mm diameter (0.118 inches) - Small leak**
 - Equivalent hole area: 7.07 mm^2 (0.011 in^2)
 - Representative of: Pinhole corrosion, small weld defects, minor valve seat leakage
 - Expected detection challenge: High (requires sensitive flow measurement)
 - Typical field scenario: Slow corrosion in aging pipelines
2. **5 mm diameter (0.197 inches) - Moderate leak**
 - Equivalent hole area: 19.6 mm^2 (0.031 in^2)
 - Representative of: Stress corrosion cracking, gasket failures, threaded connection leaks
 - Expected detection challenge: Moderate (noticeable flow imbalance)
 - Typical field scenario: Mechanical joint failures, vibration-induced cracks
3. **10 mm diameter (0.394 inches) - Large leak**
 - Equivalent hole area: 78.5 mm^2 (0.122 in^2)
 - Representative of: Corrosion pit through-wall, impact damage, failed repair welds
 - Expected detection challenge: Low (significant flow loss)
 - Typical field scenario: Third-party damage, severe corrosion, equipment failure
4. **15 mm diameter (0.591 inches) - Critical leak**
 - Equivalent hole area: 176.7 mm^2 (0.274 in^2)
 - Representative of: Rupture initiation, catastrophic fitting failure, major impact damage
 - Expected detection challenge: Very low (immediately obvious)
 - Typical field scenario: Excavation accidents, seismic activity, major corrosion failure

These sizes were selected based on API 1160 pipeline integrity management guidelines, which classify leaks by consequence severity. The range from 3-15 mm covers approximately 95% of reported leak incidents in natural gas distribution systems.

Leak Location Selection:

Three positions along the pipeline were tested to evaluate location-dependent effects:

1. **25% Position (5 miles from inlet) - Upstream location**
 - Distance from inlet: 5 miles (8.05 km)
 - Local pressure: $\sim 99.5 \text{ psi}$ (estimated by linear interpolation)
 - Rationale: Tests detection near high-pressure inlet where leak rates are maximized
 - Expected behavior: Highest leak rate due to maximum pressure differential
2. **50% Position (10 miles from inlet) - Midpoint location**
 - Distance from inlet: 10 miles (16.09 km)
 - Local pressure: $\sim 99.0 \text{ psi}$ (mid-point pressure)
 - Rationale: Represents typical leak location, equidistant from monitoring points
 - Expected behavior: Moderate leak rate, balanced impact on inlet/outlet measurements

3. 75% Position (15 miles from inlet) - *Downstream location*

- Distance from inlet: 15 miles (24.14 km)
- Local pressure: ~98.5 psi (approaching outlet pressure)
- Rationale: Tests detection near low-pressure outlet where leak rates are minimized
- Expected behavior: Lowest leak rate due to reduced pressure differential

The location parameter was expressed as a fraction (0.25, 0.50, 0.75) of total pipeline length, consistent with the simulator's input format. This provided a systematic way to evaluate whether leak position affects detection sensitivity - a critical consideration for sensor placement in real pipelines.

Leak Implementation Process:

For each combination of leak size and location (4 sizes \times 3 positions = 12 scenarios), the following procedure was executed:

Configuration Phase:

1. Load baseline pipeline configuration
2. Navigate to leak management interface in SCADA control panel
3. Select target pipe segment (single 20-mile pipe in this case)
4. Add new leak with parameters:
 - Location: 0.25, 0.50, or 0.75 (fractional distance)
 - Diameter: 3, 5, 10, or 15 mm
 - Discharge coefficient (Cd): 0.60 (sharp-edged orifice assumption)
 - Active status: True

Leak Rate Calculation:

The simulator computed leak rate using the orifice flow equation:

$$Q_{leak} = C_d \cdot A_{leak} \cdot \sqrt{\frac{2(P_{leak} - P_{ambient})}{\rho}}$$

Where:

- $A_{leak} = \frac{\pi D_{leak}^2}{4}$ (circular orifice area)
- P_{leak} = Local pressure at leak position, interpolated from inlet/outlet pressures
- $P_{ambient}$ = 14.7 psi (atmospheric pressure outside pipe)
- ρ = Gas density at leak conditions (from CoolProp)

Example Calculation (10 mm leak at 50% position):

Local pressure: $P_{leak} = 100 - (100 - 98) \times 0.5 = 99$ psi = 682.6 kPa

Pressure differential: $\Delta P = 682.6 - 101.3 = 581.3$ kPa

Leak area: $A_{leak} = \frac{\pi(0.01)^2}{4} = 7.85 \times 10^{-5}$ m²

Gas density at 99 psi: $\rho = 1.165$ kg/m³ (from CoolProp)

Leak rate:

$$Q_{leak} = 0.60 \times 7.85 \times 10^{-5} \times \sqrt{\frac{2 \times 581300}{1.165}}$$
$$Q_{leak} = 4.71 \times 10^{-5} \times 1001.4 = 0.0472 \text{ m}^3/\text{s} = 1.67 \text{ ft}^3/\text{s}$$

This represents approximately 19.6% of the baseline flow rate - a highly detectable leak.

System Response Monitoring:

Once the leak was activated, the simulation automatically recalculated flow distribution:

Flow Balance:

- Inlet flow remained at baseline: $Q_{in} = 8.5$ ft³/s (upstream boundary condition)
- Leak flow extracted from pipe: $Q_{leak} = 1.67$ ft³/s
- Outlet flow reduced: $Q_{out} = 8.5 - 1.67 = 6.83$ ft³/s

Pressure Distribution: The pressure profile along the pipeline was affected by the leak:

- Upstream pressure (inlet to leak): Slightly higher gradient due to higher flow
- Downstream pressure (leak to outlet): Reduced flow means lower pressure drop
- Net effect: Pressure at leak point decreased, affecting leak rate iteratively

SCADA Display Updates:

The visual interface immediately reflected the leak condition:

Quantitative Indicators:

- **Inlet Flow Meter:** 8.5 ft³/s (unchanged - as expected)
- **Outlet Flow Meter:** 6.83 ft³/s (reduced - clear indication of leak)
- **Expected Flow Meter:** 8.5 ft³/s (theoretical no-leak value for comparison)
- **Leak Rate Meter:** 1.67 ft³/s (**LEAK DETECTED**)
- **Flow Imbalance:** $\Delta Q = 8.5 - 6.83 = 1.67$ ft³/s (19.6% loss)

Visual Indicators:

- **Leak Marker:** Orange circle appeared at 50% position on pipeline diagram
- **Leak Spray Animation:** Particle effects showed gas escaping from pipe
- **Severity Classification:** “MODERATE” label based on 10mm diameter
- **Pressure Color Gradient:** Pipeline color shifted from green (normal) to yellow-orange near leak
- **Leak Summary Panel:** Expandable panel showed leak details:

ACTIVE LEAKS: 1

Pipe 1 | Location: 50.0% | Ø 10mm | Moderate | Rate: 1.67 ft³/s

[Click to deactivate]

Data Collection for Each Scenario:

A comprehensive dataset was recorded for each of the 12 leak scenarios:

Primary Flow Metrics:

Parameter	Symbol	Units	Measurement Method
Inlet flow rate	Q_{in}	ft ³ /s	SCADA inlet flow meter
Outlet flow rate	Q_{out}	ft ³ /s	SCADA outlet flow meter
Leak flow rate	Q_{leak}	ft ³ /s	Calculated from orifice equation
Flow imbalance	ΔQ	ft ³ /s	$Q_{in} - Q_{out}$
Percent flow loss	%	%	$(Q_{leak}/Q_{in}) \times 100$

Pressure Measurements:

Parameter	Symbol	Units	Location
Inlet pressure	P_1	psi	Pipeline entrance
Outlet pressure	P_2	psi	Pipeline exit
Leak local pressure	P_{leak}	psi	Interpolated at leak position
Pressure drop (total)	ΔP_{total}	psi	$P_1 - P_2$

Leak Characteristics:

- Leak diameter (mm)
- Leak location (fraction of pipeline length)
- Discharge coefficient (dimensionless)
- Equivalent hole area (mm²)
- Severity classification (pinhole/small/moderate/large/critical)

Detection Metrics (SCADA):

- **Detection Time:** Effectively instantaneous (<0.5 seconds) - the simulator updates in real-time
- **Detection Threshold:** Minimum detectable leak calculated as 1% of baseline flow (0.085 ft³/s)
- **Measurement Accuracy:** Assumed $\pm 0.5\%$ for flow meters (typical SCADA-grade instruments)
- **False Positive Rate:** Zero (simulation has no measurement noise)

Thermodynamic State:

- Gas density at leak point (kg/m³)
- Gas temperature (°F)
- Compressibility factor Z
- Reynolds number (to confirm turbulent flow)

Reproducibility and Sensitivity Analysis:

Each scenario was run multiple times with slight parameter variations to assess sensitivity:

- **Roughness variation:** $\pm 20\%$ (0.00048 - 0.00072 m) $\rightarrow <2\%$ impact on leak rate
- **Temperature variation:** $\pm 10^\circ\text{F}$ (50-70°F) $\rightarrow \sim 3\%$ density change, minor leak rate impact
- **Pressure variation:** ± 5 psi inlet \rightarrow Proportional leak rate change ($\sqrt{\Delta P}$ relationship)
- **Discharge coefficient:** 0.55-0.65 \rightarrow Linear scaling of leak rate

The reproducibility confirmed that leak rate calculations were stable and consistent with theoretical predictions.

Step 4: Manual Detection Scenario (Time-Delay Modeling)

While the SCADA-based system provides instantaneous leak detection through continuous flow monitoring, real-world manual detection methods introduce significant time delays. To fairly compare the two approaches, we modeled manual detection scenarios based on industry data and field practices common in Nigerian pipeline operations.

Manual Detection Methods in Practice:

Manual leak detection in natural gas pipelines typically relies on three primary mechanisms:

1. **Physical Inspection (Pipeline Patrol):**
 - **Method:** Personnel walk or drive along pipeline right-of-way, visually inspecting for signs of leaks
 - **Detection indicators:** Dead vegetation, soil discoloration, gas odor, hissing sounds, visible gas plumes
 - **Patrol frequency:** Varies by regulation and operator practices
 - High-consequence areas (HCAs): Weekly to monthly patrols
 - Standard areas: Monthly to quarterly patrols
 - Remote areas: Quarterly to semi-annual patrols
 - **Effectiveness factors:** Weather conditions, terrain accessibility, time of day, inspector experience
2. **Pressure Gauge Monitoring:**
 - **Method:** Operators manually read pressure gauges at stations/valve sites
 - **Detection indicators:** Unexpected pressure drop between measurement points
 - **Reading frequency:** Daily to weekly, depending on location importance
 - **Effectiveness factors:** Gauge accuracy (typically ± 1 -2 psi), baseline pressure variations due to demand changes
 - **Limitation:** Small leaks may not produce detectable pressure changes, especially in high-flow systems
3. **Odor Reporting (Public Detection):**
 - **Method:** Members of public smell mercaptan odorant added to natural gas
 - **Detection indicators:** Rotten egg smell near pipeline corridor
 - **Reporting method:** Emergency hotline calls to operator
 - **Effectiveness factors:** Wind direction, population density, odorant concentration, public awareness

- **Response time:** Varies from immediate (urban) to delayed (remote areas)

Time-Delay Modeling Framework:

For this research, manual detection time was modeled as a probabilistic function dependent on leak size and location. The model incorporated three components:

$$T_{detection} = T_{patrol} + T_{recognition} + T_{reporting}$$

Where:

- T_{patrol} = Time until next scheduled patrol reaches leak location
- $T_{recognition}$ = Time for inspector to identify and confirm leak
- $T_{reporting}$ = Time to report and initiate response

Leak Size-Dependent Detection Times:

Based on data from PHMSA (Pipeline and Hazardous Materials Safety Administration) incident reports and Nigerian gas distribution operator surveys, we established the following detection time ranges:

Leak Category	Diameter	Typical Detection Time	Detection Method	Assumptions
Small	3-5 mm	2-4 hours (average: 3 hours)	Physical patrol, pressure monitoring	Requires close inspection; minimal pressure impact; may not be immediately visible
Moderate	5-10 mm	1-2 hours (average: 1.5 hours)	Physical patrol, odor detection	More visible/audible; noticeable pressure drop; stronger odor
Large	10-15 mm	30-60 minutes (average: 45 minutes)	Odor detection, public reports, pressure alarms	Loud hissing; strong odor; significant pressure drop; visible vegetation damage
Critical	>15 mm	15-30 minutes (average: 22 minutes)	Public reports, audible from distance	Extremely loud; wide-area odor; immediate pressure alarms

Location-Dependent Factors:

Detection time was further adjusted based on leak location:

Urban/Populated Areas (e.g., 25% position near inlet station):

- **Patrol frequency:** Weekly (7 days between patrols)
- **Public proximity:** High (many potential observers)
- **Odor detection probability:** High (mercaptan effective)
- **Detection time multiplier:** 0.6× (faster detection)

Suburban/Moderate Access (e.g., 50% position in mixed-use area):

- **Patrol frequency:** Bi-weekly (14 days between patrols)
- **Public proximity:** Moderate (occasional observers)
- **Odor detection probability:** Moderate (depends on wind)
- **Detection time multiplier:** 1.0× (baseline)

Rural/Remote Areas (e.g., 75% position in undeveloped terrain):

- **Patrol frequency:** Monthly (30 days between patrols)
- **Public proximity:** Low (few observers)
- **Odor detection probability:** Low (sparse population)
- **Detection time multiplier:** 1.5× (slower detection)

Calculation Example (10 mm leak at 50% position):

Base detection time for 10 mm leak: 45 minutes (moderate-large category)

Location factor (50% = suburban): 1.0× multiplier

Detection time: $T = 45 \times 1.0 = 45$ minutes

If leak occurred 2 days after last patrol:

- Days until next patrol: 14 - 2 = 12 days (but leak likely detected sooner by odor)
- Expected detection: 45 minutes via public odor report

Pressure Monitoring Detection Model:

For leaks potentially detectable by pressure monitoring, we modeled detection based on gauge reading frequency:

$$\Delta P_{detectable} = \text{Gauge accuracy} + \text{Normal variation} = 1.0 \text{ psi} + 2.0 \text{ psi} = 3.0 \text{ psi}$$

Only leaks causing >3 psi pressure drop would be reliably detected by manual gauge readings. For our 20-mile pipeline with 2 psi baseline drop, this meant:

- **3 mm leak:** ΔP 0.2 psi → Not detectable by pressure monitoring
- **5 mm leak:** ΔP 0.6 psi → Not detectable by pressure monitoring
- **10 mm leak:** ΔP 2.4 psi → Marginally detectable (requires careful comparison)
- **15 mm leak:** ΔP 5.1 psi → Clearly detectable by pressure monitoring

Limitations of Manual Detection Modeling:

Several important caveats apply to our manual detection time estimates:

1. **Idealized Conditions:** Assumes inspectors are properly trained and alert
2. **Weather Independence:** Actual detection may be delayed by poor weather (rain, fog, darkness)
3. **No Seasonal Variation:** Detection times may increase in winter (fewer patrols, less vegetation indicators)
4. **No Human Error:** Assumes inspector recognizes leak when present (not always guaranteed)
5. **Odorant Effectiveness:** Assumes mercaptan odorant added at proper concentration and not degraded
6. **Public Awareness:** Assumes public knows to report gas odors (not always true in rural areas)

Comparison Basis:

The manual detection times established here serve as a realistic baseline for comparing SCADA effectiveness. While SCADA provides sub-second detection, manual methods require minutes to hours, representing a 1000-10,000× difference in response capability. This time difference directly translates to:

- **Greater gas loss volume:** $V_{lost} = Q_{leak} \times T_{detection}$
- **Increased safety risk:** More time for gas accumulation and ignition
- **Higher environmental impact:** Larger methane emissions (greenhouse gas)
- **Greater economic cost:** More product loss and potential damage

The quantified time delays provide concrete data for cost-benefit analysis of SCADA installation versus continued manual monitoring.

Step 5: Data Collection and Storage

A systematic data collection protocol ensured comprehensive documentation of all simulation results for subsequent analysis.

Data Collection Structure:

For each of the 12 leak scenarios (4 sizes \times 3 locations), a structured dataset was compiled:

Simulation Metadata:

Scenario ID: [LEAK-SIZE-LOCATION] (e.g., LEAK-10MM-50PCT)

Timestamp: [Date and time of simulation]

Pipeline Configuration: [Saved JSON reference]

Operator: [Researcher name]

Software Version: NiceGUI 1.4.x, Python 3.11

Input Parameters:

- Pipeline geometry (length, diameter, roughness, elevation)
- Operating conditions (P , P , T, fluid type)
- Leak characteristics (diameter, location, Cd)
- Ambient conditions (P_ambient, temperature)

Computed Outputs - Flow Measurements:

Parameter	SCADA Value	Manual Estimate	Units
Inlet flow rate	[Real-time data]	[Estimated from pressure]	ft ³ /s
Outlet flow rate	[Real-time data]	[Not measured]	ft ³ /s
Leak flow rate	[Calculated]	[Unknown until detection]	ft ³ /s
Flow imbalance	$Q_{in} - Q_{out}$	[Not calculated]	ft ³ /s
Percent loss	$(Q_{leak}/Q_{in}) \times 100$	[Estimated post-detection]	%

Computed Outputs - Pressure Measurements:

- Inlet pressure (psi)
- Outlet pressure (psi)
- Leak local pressure (interpolated, psi)
- Total pressure drop (psi)
- Pressure drop change vs baseline (psi)

Detection Metrics:

- **SCADA Detection Time:** <0.5 seconds (real-time)
- **Manual Detection Time:** [Modeled based on leak size/location]
- **Detection Method (Manual):** Patrol/Odor/Pressure
- **Detection Confidence:** High (SCADA) / Variable (Manual)

Thermodynamic Properties:

- Gas density at inlet, leak, outlet (kg/m³)
- Gas viscosity (Pa \cdot s)
- Compressibility factor Z
- Reynolds number
- Friction factor (Darcy-Weisbach)

Visual Data:

- **SCADA Screenshots:** Captured PNG images of interface showing:
 - Flow meter readings
 - Pressure gauge displays
 - Leak indicator position

- Pipeline color-coded pressure map
- Leak rate meter
- **Pipeline Diagrams:** SVG exports showing leak location and severity
- **Time-series plots:** Flow rate vs time (for scenarios with leak activation during simulation)

Data Storage Format:

All data was stored in structured formats for analysis:

Primary Data File (CSV):

```
Scenario,Leak_Diameter_mm,Leak_Location_pct,Qin_cfs,Qout_cfs,Qleak_cfs,Percent_Loss,Pin_psi,Pout_psi
BASELINE,0,0,8.50,8.50,0.00,0.00,100.0,98.0,99.0,N/A,N/A,N/A
LEAK-3MM-25PCT,3,25,8.50,8.38,0.12,1.41,100.0,98.0,99.5,0.3,180,Patrol
LEAK-3MM-50PCT,3,50,8.50,8.39,0.11,1.29,100.0,98.0,99.0,0.3,180,Patrol
...
```

Configuration Files (JSON): Each scenario's complete configuration was exported:

```
{
  "pipeline": {
    "pipes": [{
      "length_ft": 105600,
      "diameter_in": 12,
      "roughness_m": 0.0006,
      "elevation_change_ft": 0,
      "efficiency": 0.95
    }],
    "fluid": {
      "type": "Methane",
      "inlet_pressure_psi": 100,
      "outlet_pressure_psi": 98,
      "temperature_f": 60
    },
    "leaks": [{
      "pipe_index": 0,
      "location": 0.50,
      "diameter_mm": 10,
      "discharge_coeff": 0.60,
      "active": true
    }]
  },
  "timestamp": "2024-03-15T14:30:00Z",
  "scenario_id": "LEAK-10MM-50PCT"
}
```

Analysis Scripts: Python scripts were developed for automated data processing:

- `calculate_leak_rates.py` - Computes leak rates from orifice equation
- `plot_results.py` - Generates comparison charts
- `detect_time_analysis.py` - Analyzes detection time differences
- `cost_benefit.py` - Calculates economic impact of detection delays

Data Quality Assurance:

Several checks ensured data integrity:

1. Mass Balance Verification:

$$Q_{in} = Q_{out} + Q_{leak,total}$$

Any deviation >0.1% flagged for review

2. Physical Constraints:

- $Q_{leak} \geq 0$ (no negative leaks)
- $P_{leak} > P_{ambient}$ (leak only if pressure differential exists)

- $Q_{out} < Q_{in}$ (outlet always less than or equal to inlet)
3. **Reynolds Number Check:**
- Confirmed $Re > 4000$ (turbulent flow assumption valid)
 - If $Re < 4000$, flagged for laminar flow corrections
4. **Compressibility Factor Range:**
- Verified $0.8 < Z < 1.0$ (reasonable for methane at these conditions)
 - CoolProp calculations cross-checked against AGA-8 charts
5. **Reproducibility:**
- Each scenario run 3 times to confirm consistent results
 - Standard deviation $< 0.5\%$ for all flow measurements

Dataset Summary:

The complete dataset comprised:

- **Baseline runs:** 1 scenario
- **Leak scenarios:** 12 scenarios (4 sizes \times 3 locations)
- **Total simulations:** 13 distinct configurations
- **Repeat runs:** $3\times$ each for reproducibility = 39 total simulation runs
- **Data points per run:** ~ 25 parameters
- **Total data points:** 975 measurements
- **Storage size:** ~ 15 MB (including screenshots and configuration files)
- **Documentation:** Lab notebook entries, screenshot annotations, analysis notes

This comprehensive dataset formed the foundation for comparative analysis between SCADA-based and manual detection methods, presented in the Results section.