Forecasting Optic Fiber Deployment for enhanced Revenue Planning

Author: Archit Handa
Affiliation: BS in Data Science and Applications, IIT Madras

My Profile

Business Problems

- Automate the manual ***Capacity Planning*** process
- Identify potential causes for ***Idle Capacity*** and ***Underperformance (Missed Targets)***
- Formulate strategies to ***improve forecasting*** and ***prevent over-utilisation*** of resources

Acknowledgement :

1. Sterlite Technologies Limited (STL)
2. Mr. Pankaj Singh, PMO Lead at STL
3. Mr. Shailendra Kumar, PMO Analytics Engineer at STL

CODE24 HR SUMMIT BEST PROJECT AWARDS

01 Introduction

Sterlite Technologies Limited (STL), a global data network solutions provider, faces issues in its **Trenching & Ducting (T&D)** process, leading to missed revenue targets, delays in turn around time, and resource wastage. This project aims to resolve this using a ***Machine Learning*** model with ***Regression*** and ***Time Series Analysis*** to ***enhance capacity and revenue planning***.

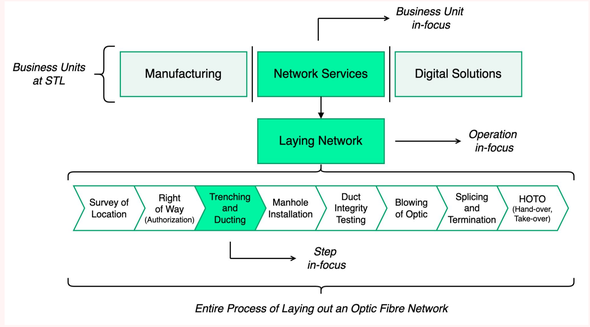

Figure 1: Business Structure at STL

02 Methodology / Approach

1. **Data Collection** using in-house tool **FieldForce** to capture daily statistics for the T&D process
2. **Data Preprocessing** (imputing missing data; aggregating spans for projection; engineering work duration and productivity features)
3. **Data Modelling** and **Model Testing**
4. **Identifying** reason behind the trends exhibited
5. **Devising strategies** to make planning resilient to future shocks and uncertainties

03 Metadata

- Data collected from Oct 2021 to Mar 2024; usable data from Jan 2022 onwards
  - **Training:** Jan 2022 to Jan 2024 (2 years)
  - **Testing:** Feb 2024 to Mar 2024 (2 months)
- Data collected from **390 spans** across **10 states** from ***various soil and weather landscapes***
- Features:
  - **Span Details** (Span, State, and Scope)
  - **Machine Type** (JCB, Poclain, HDD)
  - **Machine Count** for each type
  - **Machine Status** (Active or Idle/Breakdown)
  - **Daily Output** (in meters)
  - **Work Start and End Time**

04 Methods used for Data Analysis

Two main analytical approaches followed:

1. **Regression Analysis:**
   Regularised Multiple Linear Regression (MLR OLS) model with ***α=10*** as a baseline for the monthly output forecasting (Figure 2)
2. **Time Series Analysis:**
   a. **Seasonal Decomposition** using a 30-day moving average to separately understand the effects of ***Seasonality*** and recent ***Trends*** on overall T&D Output
   b. Set model to predict ***T&D output as percentage of overall Scope*** to aid the model to forecast better and eradicate impact of shocks (both upward and downward)
   c. **Auto-Correlation Function (ACF)** and **Partial ACF (PACF)** plots to determine hyperparameter bounds for the model
   d. ***SARIMAX*** model to predict output given the count of machines deployed (Figure 3)

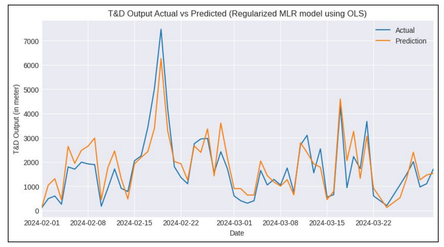

Figure 2: Regularised MLR OLS Model

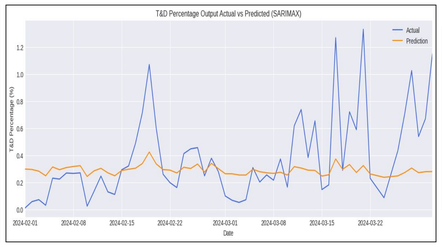

Figure 3: SARIMAX Model

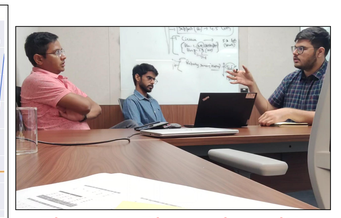

Figure 4: Discussion with People of Contact

05 Results and Findings

Finding 1: Regularised MLR OLS model obtains **better accuracy** (**6.3%** difference from actual value); yet tends to **overestimate**

Finding 2: SARIMAX model captures the **seasonality effect** yet obtains slightly lower accuracy (8.44% difference) and tends to **underestimate**

Finding 3: Ensembling the two models significantly drops the percentage difference to **1.07%**; this model is ***being tested by STL*** for ***final deployment***

Finding 4: Periods of high output immediately followed by low output suggest that current planning at STL overestimates the output, causing **over-utilisation of resources**, **idle capacity** due to machine burnout, and **stressful working environment**

06 Conclusion / Recommendations

Recommendation 1: ML Model achieves ***high accuracy***; ***robust*** to shocks; ***automates*** the capacity planning process; can be tweaked for weekly, monthly, or quarterly predictions; easy to ***incrementally train*** with more data for better results

Recommendation 2: Prevent over-utilisation of resources through ***vocational training*** regarding machine breaks and failure; ***monitor machine usage***; establish a ***fixed work schedule*** with ***maximum operating hours*** and dedicated ***shifts for timely maintenance***

Related Resources:
Access the project reports, project code, summary presentation deck, and recorded discussion by scanning the QR Code