OPSD PowerDesk: Day-Ahead Forecasting, Anomaly Detection, Live Monitoring & Dashboarding

Objective:

Using Open Power System Data (OPSD) hourly Time Series data, you will select three European countries and build a day-ahead (24-step) electric load forecasting system. You will analyze trend and seasonality via STL, use ACF/PACF and AIC/BIC to justify ARIMA/SARIMA orders, evaluate with MASE, sMAPE, MSE, RMSE, and MAPE (plus 80% PI coverage when available), detect anomalies using (i) residual z-scores and optional CUSUM, and (ii) a small machine-learning classifier based on labeled examples. You will simulate a live feed with ONE online adaptation strategy (choose Rolling SARIMA refit or Tiny neural fine-tune), and present results on a compact dashboard.

1. **Data (use OPSD; pick three countries)**
   - 1.1. Source: OPSD Time Series (hourly CSV).
     - Pick three countries (e.g., DE, FR, ES; or IT, GB, …).
     - For each country, select:
       - utc_timestamp → rename to timestamp (keep UTC).
       - <CC>_load_actual_* → rename to load (e.g., DE_load_actual_entsoe_transparency). Here **"CC" refers to Country code**
       - Optional: <CC>_wind_generation_actual_* and <CC>_solar_generation_actual_* as exogenous variables
       - Drop rows with missing load; sort by timestamp.
   - 1.2. One tidy DataFrame per country
     - Structure per country: timestamp, load[, wind][, solar].
   - 1.3. Basic sanity plot (per country)
     - Plot the last 14 days to confirm hourly cadence and realistic magnitudes.
   - 1.4. Decomposition & seasonality/trend finding (per country)
     - i. STL (Seasonal-Trend decomposition): daily seasonality period = 24. Save a figure with Trend, Seasonal, and Remainder (Observed optional).
     - ii. Stationarity and differencing: if trend is present, try d=1; if strong daily seasonality, try seasonal differencing D=1 with s=24.
     - iii. ACF/PACF on appropriately differenced series up to lag 48 to guide (p,q) and (P,Q).
     - iv. Information criteria (AIC/BIC): search a small SARIMA grid and select the lowest BIC (break ties with AIC).
     - v. Document your chosen order briefly for each country.
2. **Forecasting (day-ahead, 24 steps)**

2.1. Splits and backtest

- For each country: Train = first 80%, Dev (val) = next 10%, Test = final 10% (chronological).
- Backtest: expanding origin with stride = 24h and horizon = 24h; warm-up ≥ 60 days of history.

2.2. Models (per country)

i. Required classical: SARIMA/SARIMAX using chosen orders from section 1.4.
ii. Optional exogenous: hour-of-day and day-of-week one-hots; optional wind/solar series.
iii. Optional neural (for comparison/bonus): GRU/LSTM direct multi-horizon (last 168h → next 24 steps).

2.3. Save forecasts

- Create per-country CSVs: outputs/<CC>_forecasts_dev.csv and outputs/<CC>_forecasts_test.csv with columns:
- timestamp, y_true, yhat, lo, hi, horizon, train_end (80% PI if available for lo/hi).

2.4. Metrics you MUST report (per country, Dev (val) & Test)

- MASE (seasonality = 24) — PRIMARY
- sMAPE
- MSE
- RMSE
- MAPE
- 80% PI coverage (if intervals available)

Also provide a Test comparison table across the three countries.

3. Anomaly detection (two parts)

3.1. Residual z-score + optional CUSUM (unsupervised; z-score required)

i. Compute 1-step-ahead residuals on Test: $e_t = y_t - \hat{y}_t$
ii. Rolling z-score with window = 336h (14d), min_periods = 168: $z_t = (e_t - \mu_{roll}) / \sigma_{roll}$.
iii. Flag anomaly if $|z_t| \geq 3.0$ → flag_z ∈ {0,1}.
iv. Optional CUSUM on z_t: k = 0.5, h = 5.0; alarm when $S^+ > h$ or $S^- > h$ → flag_cusum.

Save outputs/<CC>_anomalies.csv: timestamp, y_true, yhat, z_resid, flag_z, [flag_cusum].

3.2. ML-based anomaly classifier (with labeling) — REQUIRED

i. Create "silver labels": positive if ($|z_t| \geq 3.5$) OR (y_true outside [lo,hi] AND $|z_t| \geq 2.5$); negative if $|z_t| < 1.0$ AND y_true inside [lo,hi].
ii. Human verification: per country, randomly sample ≈100 timestamps (≈50 positives, 50 negatives) and confirm labels by visual check (±24h).

iii. Train a simple classifier (Logistic/LightGBM) on features from the last 24–48h (lags/rollups, calendar, forecast context).

Report PR-AUC and F1 at fixed precision (e.g., P=0.80). Save anomaly_labels_verified.csv and anomaly_ml_eval.json.

4. "Live" ingestion + online adaptation (simulate stream)
    - Pick one of your three countries for live simulation (others remain offline). Run ≥ 2,000 hours of simulated updates.
    - Loop each hour: append next row; at 00:00 UTC forecast next 24h; update z-score and optional CUSUM; check drift; if triggered, adapt; log update.

Log file: outputs/<CC>_online_updates.csv with columns: timestamp, strategy, reason (initial/scheduled/drift), duration_s.

"Live ingestion + online adaptation"?

- **Live ingestion** = pretending your data arrives **hour by hour** (like a real grid feed), not all at once. You "reveal" one new row at a time from the historical OPSD CSV and immediately update your forecasts/alerts.
- **Online adaptation** = after you ingest new data, you **optionally update your model** so it stays calibrated when patterns drift (season changes, demand surges, holidays, etc.). You pick **one** adaptation strategy and implement only that.

Think of it as: **ingest → forecast → detect anomalies → (maybe) adapt → log → repeat** every simulated hour.

- Choose ONE online adaptation strategy (exactly one)

    i. Rolling SARIMA refit (simple & robust): daily at 00:00 refit on last 90 days; also refit on drift trigger.
    ii. Tiny neural fine-tune (only if GRU/LSTM used): every 6h, 1 epoch on last 14 days, LR=1e-4, update output layer only; also on drift trigger.

- Drift trigger & after-update snapshot
- Drift trigger: EWMA (|z|; α=0.1) > 95th percentile of |z| over last 30 days → trigger adaptation.
- After each update: record rolling-7d MASE and rolling-7d 80% PI coverage before vs after (small table in report).

5. Dashboard (Streamlit or equivalent)

    Required elements for the live country:

i. Country selector (preselect live country).
ii. Live series: last 7–14 days of y_true & yhat (line chart).
iii. Forecast cone: next 24h mean with 80% PI (shaded).
iv. Anomaly tape: highlight hours with flag_z=1 (and flag_cusum=1 if present).
v. KPI tiles: rolling-7d MASE, 80% PI coverage (7d), # anomaly hours today, last update time.
vi. Update status: last online update timestamp + reason.

6. What you submit
   6.1. Repository layout (preferred)

Path / File	Purpose
README.md	How to run; countries; environment.
requirements.txt	Python dependencies.
config.yaml	Countries, column names, thresholds, horizons.
data/	Local path to OPSD CSV (do not commit large data).
src/load_opsd.py	Read CSV; build tidy per-country frames.
src/decompose_acf_pacf.py	STL plots; ACF/PACF; AIC/BIC grid & summary.
src/forecast.py	Expanding-origin backtest; save <CC>_forecasts_dev/test.csv.
src/anomaly.py	Z-score (+CUSUM); save <CC>_anomalies.csv.
src/anomaly_ml.py	Silver labels → sample → train; save labels + PR-AUC/F1 metrics.
src/live_loop.py	Simulated stream + chosen adaptation; save <CC>_online_updates.csv.
src/dashboard_app.py	Streamlit dashboard (or in-notebook equivalent).
src/metrics.py	MASE, sMAPE, MSE, RMSE, MAPE, coverage helpers.
outputs/	All generated CSV/JSON artifacts.

notebooks/	Optional exploratory notebooks.

6.2. **Single Colab notebook (acceptable alternative)**
If submitting one Colab only, it must include the same steps and export the same artifacts into /content/outputs/. Required sections: Config; Data Ingest; STL + ACF/PACF + AIC/BIC; Backtesting; Anomaly (rules + ML); Live + Adaptation; Dashboard (inline plots); Pack outputs.

6.3. **Report (≤ 7 pages**)

- Data & STL: list three countries; include one STL figure; 1-line takeaway per country.
- Order selection: ACF/PACF figure + top-5 AIC/BIC table (one country); list final orders for all three.
- Forecast results: Dev/Test metrics (MASE, sMAPE, MSE, RMSE, MAPE, coverage).
- Anomalies: top-10 z-score hours; 1–2 example plots with notes.
- ML anomaly: PR-AUC and F1@P=0.80; brief feature importance commentary.
- Live + adaptation: chosen strategy; one before/after mini-table (rolling-7d MASE, coverage).
- Limitations: 3–5 bullets.

7. Defaults & knobs (use unless strong reasons to change)
   - Seasonality s=24; differencing d∈{0,1}, D∈{0,1}.
   - SARIMA grid (BIC): (p,q)∈{0,1,2}, d∈{0,1}; (P,Q)∈{0,1}, D∈{0,1}, s=24.
   - Backtest: warm-up 60d, stride 24h, horizon 24h.
   - Metrics: MASE, sMAPE, MSE, RMSE, MAPE, Coverage(80%).
   - Anomaly: z-score window 336h; |z|≥3.0; CUSUM k=0.5, h=5.0 (optional).
   - Live: start history 120d; simulate ≥2,000 hours.
   - Adaptation: A) Rolling SARIMA 90d daily (also on drift) OR C) Tiny GRU/LSTM 1-epoch/6h on last 14d (also on drift).
   - Drift trigger: EWMA(|z|; α=0.1) > 95th percentile of |z| over last 30d.