

Grid-Forecast Feature Set

This document outlines the recommended features for building the Delhi Transco energy demand forecasting model using the Holidays, Weather, and Electricity datasets.

1 Holidays Dataset

Columns: ['day', 'holiday', 'holiday_type', 'month']

Useful Features: - `holiday`: Boolean, 1 if holiday, 0 otherwise - `holiday_type`: Categorical (festival, public holiday, city event) – encode as one-hot or label encode - `day`: Can derive `is_weekend` - `month`: Capture seasonal effects

Derived Features: - `days_to_next_holiday`: Optional, number of days until next holiday - `holiday_flag`: 1 if major event or festival (scale if needed)

2 Weather Dataset

Columns: ['datetime_utc', '_conds', '_dewptm', '_fog', '_hail', '_heatindexm', '_hum', '_precipm', '_pressurem', '_rain', '_snow', '_tempm', '_thunder', '_tornado', '_vism', '_wdird', '_wdire', '_wgustm', '_windchillm', '_wspdm', 'year', 'month', 'hour', 'dayofweek']

Useful Features: - `_tempm`: Temperature - `_hum`: Relative humidity - `_precipm`: Rain / precipitation - `_fog`, `_hail`, `_thunder`, `_tornado`: Extreme weather flags (boolean) - `_windchillm`, `_heatindexm`: Comfort indices - `_wspdm`, `_wgustm`: Wind speed and gusts

Derived Features: - `is_extreme_weather`: 1 if conditions like hail, thunder, tornado present - Lagged features: `_tempm_lag1`, `_hum_lag1`, etc. - Rolling averages: `_tempm_rolling3h`

3 Electricity Dataset

Columns: ['timestamp', 'demand_mw', 'temp_c', 'dewpoint_c', 'rel_humidity_pct', 'wind_dir_deg', 'wind_speed_ms', 'pressure_hpa', 'year', 'month', 'day', 'hour', 'minute', 'demand_ma3']

Target Variable: - `demand_mw`

Useful Features: - `temp_c`, `dewpoint_c`, `rel_humidity_pct`: Weather proxies - `wind_speed_ms`, `wind_dir_deg`, `pressure_hpa`: Weather effects - `hour`, `day`, `month`, `year`: Time-based features

Derived Features: - `lag_demand_1h`, `lag_demand_24h`: Previous demand trends - `demand_ma3`: 3-hour moving average - `is_weekend`: From day - `season`: Derived from month

4 Combined Feature Set

Feature	Type	Notes
timestamp	datetime	Merge key
hour	int	From timestamp
day	int	From timestamp
dayofweek	int	0-6
month	int	1-12
is_weekend	bool	Derived from dayofweek
holiday_flag	bool	From holidays dataset
holiday_type	categorical	One-hot encode
temp_c / _tempm	float	Temperature
rel_humidity_pct / _hum	float	Humidity
dewpoint_c / _dewptm	float	Dew point
wind_speed_ms / _wspdm	float	Wind speed
wind_gust	float	Optional from _wgustm
precipitation	float	_precipm or rain
extreme_weather_flag	bool	If hail / tornado / thunder present
lag_demand_1h / 24h	float	Past demand
rolling_demand_ma3	float	Already available

Target: - `demand_mw`