**Zandvoort Model Technical Documentation**

**Parameter Sources and Calculations**

**Fixed Parameters (Regulatory/Factual)**

* **Fuel consumption rate**: 110kg total fuel ÷ 72 laps = 1.528 kg/lap (F1 regulation)
* **Points system**: Standard F1 points allocation (25-18-15-12-10-8-6-4-2-1)
* **Base pace**: 71.0s (Zandvoort representative lap time)
* **Race distance**: 72 laps (Dutch GP regulation distance)

**Zandvoort-Specific Parameters (source in parenthesis)**

* **Rain probability**: 30% (via Open Weather API, to be updated closer to race day)
* **SC probability**: 67% (Formula 1 official estimate)
* **VSC probability**: 67% (Formula 1 official estimate)
* **Pit time loss**: 16.5s (Down from 21.5 in 2024 due to change in pitlane speed limit)
* **Fuel effect**: 0.035s/kg (standard F1 estimate)

**Extracted Parameters (Data-Driven via F1\_Parameter\_Extractor.py)**

* **Position penalties**: Calculated by comparing qualifying pace vs race pace (laps 8-30), accounting for compound differences
* **Tire performance**: Linear regression on stint lap vs lap time, using laps 2-25 of each stint
* **Driver error rates**: Residual analysis after fitting trend lines to stint data, defining errors as laps >2σ slower than expected
* **DRS effectiveness**: Attempts to measure lap time improvements when cars run close together, falls back to 0.25s theoretical estimate

**Practice-Based Tire Models (Primary Source)**

* **FP1+FP2 tire modeling**: Bayesian MCMC models fitted to combined practice session data
* **Compound degradation**: Posterior distributions from actual track data using NUTS sampling
* **Track-specific parameters**: Base pace and degradation rates extracted from Zandvoort practice sessions
* **Data sources**: Both FP1 (morning) and FP2 (afternoon) sessions for comprehensive track condition coverage
* **Quality assessment**: Model uncertainty quantification and confidence scoring per compound
* **Fallback hierarchy**: Practice models → Extracted parameters → Default parameters

**Fallback Parameters (Used when extraction unavailable)**

* **Position penalties**: Linear approximation (0.12 × (position-1) seconds)
* **Tire offsets**: SOFT baseline: 0.0s, MEDIUM: +0.35s, HARD: +0.65s
* **Degradation rates**: SOFT: 0.18s/lap, MEDIUM: 0.09s/lap, HARD: 0.05s/lap
* **Error rates**: 0.012 dry, 0.030 wet (base probability per lap)

**Parameter Extraction Technical Details**

**F1ParameterExtractor.extract\_position\_penalties()**

* Loads qualifying and race sessions via FastF1
* Extracts Q3/Q2/Q1 times for grid positions
* Gets race laps 8-30 (excludes start effects and tire degradation)
* Groups by driver and compound to normalize strategy differences
* Calculates expected race pace as quali time + 0.5s
* Measures actual penalty as (actual race pace - expected race pace)
* Groups by grid position and takes median penalty

**F1ParameterExtractor.extract\_tire\_performance\_soft\_baseline()**

* Filters race laps for green flag, non-pit conditions
* Creates stint lap counter (cumulative within each stint)
* Uses laps 2-25 of each stint (excludes out lap and extreme degradation)
* Fits LinearRegression(stint\_lap, lap\_time) for each compound
* Removes outliers beyond 2.5 standard deviations
* Calculates offset relative to SOFT compound base time

**F1ParameterExtractor.extract\_driver\_error\_rates()**

* Groups laps by driver and stint
* Fits linear trend line to remove tire degradation effects
* Calculates residuals (actual - predicted lap time)
* Defines errors as laps where residual > 2σ from stint trend
* Counts error rate as (error laps / total laps) per stint
* Separates dry compounds (S/M/H) from wet compounds (I/W)

**F1ParameterExtractor.extract\_drs\_effectiveness()**

* Attempts to find lap time improvements for cars running close together
* Compares current lap vs median of previous 5 laps for same compound
* Identifies improvements between 0.1-1.0s as potential DRS usage
* Removes outliers using IQR method
* Falls back to 0.25s theoretical estimate if insufficient data

**Practice-Based Tire Modeling Technical Details**

**build\_compound\_models\_from\_practice()**

* **Data sources**: Loads FP1 and FP2 sessions via FastF1 API
* **Data filtering**: Green flag laps only, excludes pit in/out laps, removes outliers (>3σ)
* **Session combination**: Merges FP1 and FP2 data while tracking session source
* **Stint analysis**: Calculates stint lap position within each practice stint
* **Sample requirements**: Minimum 10 laps per compound for reliable modeling

**Bayesian MCMC Tire Model**

* **Framework**: Uses numpyro with NUTS sampler for MCMC inference
* **Prior distributions**:
  + alpha ~ Normal(base\_pace, 2.0) for intercept (track-specific base pace)
  + beta ~ Normal(expected\_degradation, compound\_variance) for degradation slope
  + sigma ~ HalfNormal(1.0) for observation noise
* **Compound-specific priors**:
  + SOFT: beta ~ Normal(0.15, 0.05) - faster degradation expected
  + MEDIUM: beta ~ Normal(0.08, 0.03) - moderate degradation
  + HARD: beta ~ Normal(0.04, 0.02) - slower degradation
* **Sampling**: 800 warmup samples + 1200 posterior samples
* **Model**: LapTime ~ Normal(alpha + beta × StintLap, sigma)

**get\_tire\_performance\_from\_practice\_models()**

* **Primary method**: Uses posterior median values (alpha, beta) from practice models
* **Weather handling**: Simplified wet model for INTERMEDIATE compound
* **Fallback cascade**: Practice model → Extracted parameters → Hardcoded defaults
* **Non-linear effects**: Additional degradation penalty for stints >25 laps
* **Track evolution**: Incorporates track surface improvement over session

**Model Quality Assessment**

* **Parameter uncertainty**: Standard deviation of posterior samples
* **Quality tiers**:
  + High: alpha\_std < 0.5s, beta\_std < 0.01s/lap
  + Good: alpha\_std < 1.0s, beta\_std < 0.02s/lap
  + Moderate: Higher uncertainty or limited data
* **Session coverage bonus**: Higher confidence when both FP1 and FP2 available
* **Sample size weighting**: Quality score increases with practice lap count

**Monte Carlo Simulation Technical Details**

**simulate\_race() Function**

* **Weather generation**: Uses numpy.random with specified rain probability, creates lap-by-lap weather array
* **SC/VSC generation**: Random selection based on probability thresholds, determines start lap and duration
* **Lap time calculation**: **Now uses practice-based models as primary source**
* **Fuel effects**: Decreases fuel load by consumption rate per lap, applies fuel\_effect\_per\_kg
* **Position penalties**: Applies traffic factor that decreases over race distance
* **Driver errors**: Random probability check per lap, increases with stint length
* **Pit stop execution**: Normal distribution around pit time loss, reduced during SC/VSC periods

**Enhanced Tire Performance Model**

* **Primary calculation**: get\_tire\_performance\_from\_practice\_models(compound, lap\_in\_stint, compound\_models, base\_pace, weather, track\_evolution, extracted\_params)
* **Model hierarchy**:
  1. Practice-based MCMC models (if available)
  2. Extracted historical parameters (if parameter extraction successful)
  3. Hardcoded fallback values based on historical data
* **Dry conditions**: Uses practice model posterior medians for base time and degradation
* **Wet conditions**: Simplified model with INTERMEDIATE compound (+9s penalty)
* **Non-linear degradation**: Additional penalty for laps >25 in stint: 0.04 × (lap-25)^1.3
* **Temperature effects**: Compound-specific penalties for SOFT overheat, HARD cold tire phases
* **Error handling**: Fallback when practice models fail

**Position Change Mechanics**

* **Undercut probability**: 25% first stop, 12% subsequent stops
* **Position changes**: Discrete random selection from [-2, -1, 0, +1, +2] with defined probabilities
* **Strategic aggressiveness**: Bonus position changes for strategies using 2+ aggressive compounds
* **Bounds checking**: Positions clamped between 1 and 20

**Validation Model Technical Details**

**generate\_simulation\_results()**

* **Integration**: Imports run\_monte\_carlo\_with\_grid from main simulation with practice models
* **Simulation count**: 1000 simulations per grid position using practice-based tire models
* **Output format**: Converts to validation structure (final\_positions, points, times arrays)
* **Fallback behavior**: Graceful degradation to approximation method if practice models unavailable

**validate\_position\_predictions()**

* Filters strategies based on actual weather conditions (wet vs dry)
* Calculates Mean Absolute Error between predicted and actual final positions
* Computes accuracy within ±2 positions as percentage
* Selects best strategy based on lowest MAE
* Records position error as (predicted - actual) for bias analysis
* **Enhanced reporting**: Shows tire model source (practice vs extracted vs fallback)

**validate\_race\_conditions()**

* Compares simulation rain probability (30%) against actual weather occurrence
* Detects safety cars by identifying laps >25% slower than rolling median baseline
* Compares simulation SC probability (67%) against detected incidents
* Reports parameter quality based on extraction sample sizes
* **Practice model validation**: Assesses quality of FP1+FP2 tire models vs actual degradation

**Weather/SC Detection Methods**

* **Rain detection**: Counts INTERMEDIATE/WET tire usage vs dry compound usage
* **SC detection**: Rolling 5-lap median baseline, identifies anomalous slow laps
* **Strategy classification**: Counts pit stops and identifies compound sequences

**Model Architecture Flow**

**1. Parameter Extraction Phase**

Historical Data (FastF1) → Data Cleaning → Statistical Analysis → Parameter File

**2. Practice Model Building Phase (New)**

FP1+FP2 Sessions (FastF1) → Data Filtering → Bayesian MCMC → Tire Models → Quality Assessment

**3. Simulation Phase**

**Practice Models** + Parameters + Strategy → Monte Carlo Loop → Weather/SC Generation → Lap Simulation → Position Tracking → Results

**4. Validation Phase**

Actual Race Data → Simulation Results → Comparison Analysis → Accuracy Metrics → Validation Report

**Key Model Limitations**

**Parameter Extraction**

* Limited to data available in FastF1 (relatively low resolution)
* Position penalty calculation assumes quali-race delta is purely traffic-related
* DRS effectiveness measurement is proximity-based, not direct telemetry
* Small sample sizes for some parameters reduce reliability

**Practice-Based Tire Modeling**

* **Practice vs race conditions**: FP1/FP2 may not represent full race degradation
* **Limited stint lengths**: Practice stints typically shorter than race stints
* **Weather dependency**: Models trained on practice conditions may not apply to wet races
* **Track evolution**: Practice track evolution may differ from race day progression
* **Compound availability**: Limited to compounds used in practice sessions
* **Car setup differences**: Practice setups may not match race configurations

**Simulation Model**

* Binary weather states (dry/wet) vs gradual transitions
* Simplified position change mechanics vs complex race dynamics
* Fixed car performance gaps vs variable competitive windows
* No driver-specific performance modeling
* **Practice model dependency**: Simulation accuracy limited by practice session data quality

**Validation Model**

* Only validates against single race outcome vs statistical expectations
* Cannot validate weather/SC probabilities from single event
* Strategy filtering assumes perfect hindsight about race conditions
* Position prediction accuracy depends heavily on parameter quality
* **Practice model validation**: Difficult to assess practice model accuracy until race completion

**Data Quality Indicators**

**Practice Model Quality (New)**

**Excellent (R² > 0.8, >50 practice laps)**

* High confidence degradation curves
* Low posterior uncertainty
* Both FP1 and FP2 data available

**Good (R² > 0.6, 20-50 practice laps)**

* Reliable degradation trends
* Moderate posterior uncertainty
* Single session data or limited multi-session coverage

**Moderate (R² < 0.6, <20 practice laps)**

* High uncertainty in degradation rates
* Limited practice data
* Fallback to extracted parameters recommended

**Traditional Parameter Quality**

**High Quality (>100 samples)**

* Position penalties for grid positions 1-10
* Tire performance for major compounds (S/M/H)
* Driver error rates in dry conditions

**Moderate Quality (20-100 samples)**

* Position penalties for grid positions 11-20
* Driver error rates in wet conditions
* DRS effectiveness measurements

**Low Quality (<20 samples)**

* Track-specific parameters for rarely visited circuits
* Weather-specific degradation rates
* Qualifying vs race pace correlations in mixed conditions

**Tire Strategies**

Based on historical averages from 2021-2024 Dutch Grands Prix (via FastF1 API), enhanced with practice session insights for optimal stint length determination.