

# Predicting Competitive Dynamics in Formula 1: Modelling the Impact of FIA Aerodynamic Regulations Using Machine Learning and Simulation

MSc Research Project  
Master of Science in Data Analytics

Nicholas Stancill  
Student ID: X22225668

School of Computing  
National College of Ireland

Supervisor: Dr Vladimir Milosavljevic

**National College of Ireland**  
**MSc Project Submission Sheet**  
**School of Computing**



**Student Name:** Nicholas Stancill  
X22225668

**Student ID:**

**Programme:** Master of Science in Data Analytics      **Year:** 2025

**Module:** MSc Research Project

**Supervisor:** Dr Vladimir Milosavljevic


**Submission Due Date:** 15/09/2025

**Project Title:** Predicting Competitive Dynamics in Formula 1: Modelling the Impact of FIA Aerodynamic Regulations Using Machine Learning and Simulation

**Word Count:** 6833 **Page Count** 20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

**Signature:** 

**Date:** 14/09/2025

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST**

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
<b>Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).</b>	<input type="checkbox"/>
<b>You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.</b>	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

<b>Office Use Only</b>	
Signature:	
Date:	
Penalty Applied (if applicable):	

# AI Acknowledgement Supplement

## MSc Research Project

### Predicting Competitive Dynamics in Formula 1: Modelling the Impact of FIA Aerodynamic Regulations Using Machine Learning and Simulation

Your Number	Name/Student	Course	Date
Nicholas X22225668	Stancill	MSc Research Project	08/08/2025

## AI Acknowledgment

Tool Name	Brief Description	Link to tool
ChatGPT	Used to support clarity and readability of self-written text, summarise complex background concepts, and provide coding guidance for syntax and structure checks. All outputs were reviewed, verified, and edited before inclusion.	<a href="https://chatgpt.com">https://chatgpt.com</a>
Grammarly	Integrated with Microsoft Word to identify and correct minor grammar, spelling, and punctuation issues	<a href="#">Grammarly: Free AI Writing Assistance</a>

## Description of AI Usage

ChatGPT
<p>I independently researched, structured, and drafted all sections of my MSc Research Project to ensure they reflected my own analysis, interpretations, and conclusions. AI tools were used only to support this work, not to generate original content or perform analysis in place of my own. Specifically:</p> <ul style="list-style-type: none"><li>• ChatGPT was used to rephrase and refine sentences for improved clarity and academic tone in sections I had already written, to summarise certain complex technical or contextual concepts (e.g, aerodynamics, DRS function), and to provide guidance on coding tasks such as Python syntax checks, code structure suggestions, and interpretation of statistical outputs.</li><li>• All AI outputs were critically reviewed, fact-checked, and edited before being incorporated into my project, ensuring accuracy and alignment with academic standards.</li></ul> <p>I maintained full academic control over the research design, data analysis, and interpretation of results. AI was never used to generate substantive original findings, write entire sections, or replace my own critical engagement.</p> <p><b>Summarise the purpose and impact of the Drag Reduction System (DRS) in F1 and explain how this change has affected racing</b></p>

	<p>ahead. By allowing the rear wing to open in specific zones when within one second of another car, DRS improves straight-line speed and increases overtaking opportunities. While it has boosted passing frequency, it has also been criticised for creating overtakes that rely less on driver skill and more on technology, changing the strategic balance of races.</p>
<p><b>I want to run a Monte Carlo simulation to compare lap-time variability across three Formula 1 regulation eras. My idea is to use historical mean and standard deviation values for each era and simulate 10,000 race laps for each. Can you suggest a clean Python structure for this?</b></p>	<p>You can create a function that accepts mean, standard deviation, and number of iterations as parameters, then use <code>numpy.random.normal()</code> to generate lap-time samples. Store the results in a Pandas DataFrame for easy analysis and visualisation. This modular design will allow you to re-use the function for each regulation era with different inputs, and you can plot histograms or KDE plots to compare the distributions.</p>
<p><b>My Random Forest model for predicting lap-time variability achieved <math>R^2 = 0.735</math>, RMSE = 1.212 ms, and MAE = 0.984 ms. Are these results strong, and what do they imply?</b></p>	<p>The <math>R^2</math> value of 0.735 indicates that the model explains over 73% of the variance in lap-time variability, which is a strong result for race-level performance modelling. The low RMSE and MAE values suggest accurate and consistent predictions. These metrics confirm that the model is robust enough to support scenario simulations with reasonable confidence.</p>
<p><b>I have trained a Random Forest regression model to predict finishing position using variables like grid position, driver, season, and regulation era. The model gives me feature importance scores, but I want to make sure I'm interpreting them correctly. Can you confirm whether higher values mean greater predictive influence, and suggest how I might present them visually?</b></p>	<p>Higher feature importance scores indicate that the variable contributes more to reducing model error across all splits. Visual representation can be done via a horizontal bar chart or ranked table, sorted from most to least important. Including confidence intervals or bootstrapped estimates can make the analysis more robust.</p>
<p><b>I'm calculating lap-time variability using standard deviation and mean lap time, then converting it into a coefficient of variation (CV). My formula is <math>CV = (SD / Mean) \times 100</math>. Can you confirm if this is correct, and if there are any considerations when comparing across different race distances?</b></p>	<p>The formula is correct for calculating the coefficient of variation in percentage form. When comparing across races, ensure that lap times are in consistent units (seconds) and account for differences in track length or weather conditions that could artificially inflate variability values.</p>

# Predicting Competitive Dynamics in Formula 1: Modelling the Impact of FIA Aerodynamic Regulations Using Machine Learning and Simulation

Nicholas Stancill  
X22225668

## Abstract

This thesis investigates the impact of FIA aerodynamic regulation changes on race competitiveness in Formula One from 1990 to 2024, focusing on the introduction of Drag Reduction System (DRS) in 2011 and the return of ground-effect aerodynamics in 2022. Using a combination of statistical analysis, Random Forest modelling, and Monte Carlo simulation, the study evaluates how these interventions influenced lap-time variability and driver consistency as proxies for competitiveness. Results show that DRS significantly reduced race variability, enhancing overtaking and midfield engagement, while the ground-effect regulations have yet to demonstrate measurable improvements. Predictive modelling achieved  $R^2$  values above 0.70, indicating strong explanatory power for features such as season, era, and mean lap time. Simulations suggest that circuit-specific DRS deployment may offer greater gains than blanket regulations. Overall, the findings advocate for a more data-driven, adaptive approach to motorsport regulation, supporting evidence-based policymaking in preparation for the upcoming 2026 Formula One rule changes.

## 1 Introduction

Formula 1 (F1) has long been regarded by many as the peak of motorsport, blending innovative engineering, elite athleticism, and regulatory adaptation. Central to this constantly evolving environment is the role of the FIA (Fédération Internationale de l'Automobile), the sport's governing body, which enforces and iteratively refines regulations to balance competition, safety, and entertainment.

Aerodynamics stands as a cornerstone of performance in F1. The extent to which a car can follow another through corners, its ability to overtake, and even its fuel efficiency is all dictated by the aerodynamic efficiency of the chassis. Regulation changes that alter aerodynamic behaviour therefore have profound consequences on the competitive structure of the sport.

Over the past 15 years, two aerodynamic regulatory shifts have attempted to remedy longstanding overtaking challenges. The introduction of the Drag Reduction System (DRS) in 2011 and the reintroduction of ground-effect aerodynamics in 2022. The DRS was designed as a temporary driver-activated mechanism to reduce rear wing drag on straights, thus improving overtaking feasibility. Ground-effect, conversely, marked a structural overhaul aimed at restoring close racing by generating downforce through underfloor

channels, reducing reliance on upper bodywork and mitigating the turbulent wake, commonly termed "dirty air", that impairs following cars.

Despite the promise of both regulatory interventions, the efficacy of these changes remains debated. While the number of overtakes per race is often used as a key performance metric, this statistic alone does not capture broader competitive dynamics such as race pace convergence, driver adaptability, and positional volatility. The purpose of this thesis is to go beyond simplistic overtaking metrics to analyse the impact of recent aerodynamic regulation changes on Formula 1 race dynamics from 2011 to 2024.

By combining Monte Carlo simulations and lap-level performance data, this study evaluates how different aerodynamic regulations have affected race dynamics over time. It uses two analytical proxies for competitiveness: race-level lap-time variability (standard deviation of lap times per race) and driver consistency (per-driver intra-race standard deviation). These metrics reflect not just overtaking, but underlying pace convergence and race volatility.

Through this approach, the thesis not only contributes to academic literature on motorsport analytics and sports regulation but also offers valuable insights for stakeholders including the FIA, F1 teams, engineers, and fans. As Formula 1 evolves toward its next set of technical regulations in 2026 (FIA, 2024), an empirical understanding of how previous rules have influenced racing outcomes is both timely and essential.

## 1.1 Motivation

The motivation for this project stems from a broader interest in how technical systems respond to regulatory inputs, especially in high-performance environments such as motorsport. Formula 1 uniquely serves as a testbed for engineering innovation and strategic decision-making under constraints imposed by the sport's governing body. As such, it offers fertile ground for academic investigation at the intersection of data science, engineering, and policy.

Two landmark regulatory changes, the 2011 introduction of the Drag Reduction System (DRS) and the 2022 ground-effect aerodynamic overhaul, highlight the FIA's ongoing efforts to foster close and competitive racing. While this thesis does not directly assess overtaking perceptions, it evaluates how these interventions shaped race variability and positional dynamics.

Although anecdotal reports suggest the effect of regulations varies across circuits for example, Monaco vs. Silverstone. This study takes holistic view, examining aggregated lap-level data across multiple seasons and circuits. The focus lies in capturing structural trends in competitiveness rather than individual race narratives.

Academic literature to date often examines these changes in isolation, either through overtaking frequency statistics or physics-based simulations. This project offers a unified, data-driven analysis using empirical race data, Monte Carlo simulations, and derived metrics to quantify how regulation shifts influence racing dynamics.

Personal interest also plays a role in shaping this thesis. With a longstanding passion for Formula 1 and academic grounding in data analytics and Mathematics, this research presents

an opportunity to bridge subject matter enthusiasm with technical expertise. Ultimately, it aims to provide a deeper understanding of regulation efficacy using accessible, interpretable data science methods.

Accordingly, the central research question guiding this study is: What impact have FIA aerodynamic regulation changes, specifically the introduction of the Drag Reduction System in 2011 and the reintroduction of ground-effect aerodynamics in 2022, had on race competitiveness in Formula One between 1990 and 2024?

## **2 Literature Review**

### **2.1 Aerodynamic Fundamentals and Historical Context**

Aerodynamics has long played a critical role in Formula 1 performance, where managing airflow efficiency not only dictates downforce but directly influences lap times, thermal management, and competitive proximity. In earlier decades of the sport, mechanical grip dominated race outcomes, but since the 1970s, Formula 1 has increasingly relied on sophisticated aerodynamic designs to extract performance. By the early 2000s, cars had become increasingly reliant on intricate bargeboards, diffusers, and multilayer wings to maintain optimal aerodynamic balance. This over-reliance created pronounced challenges for trailing cars, as wake turbulence, colloquially known as “dirty air,” would degrade aerodynamic performance for any vehicle attempting to follow closely. Studies, such as Guerrero and Castilla (2020), report downforce losses ranging from approximately 23% to 62% for a following car, significantly impacting cornering speed, tyre degradation, and fuel efficiency.

This aerodynamic disruption has been consistently cited as a key contributor to diminished competitiveness, especially on circuits where overtaking is already challenging. Castro and Rana (2020) highlight that wake turbulence significantly hinders close racing, noting how drivers are forced to manage tyre temperatures and avoid turbulent zones, reducing opportunities for meaningful side-by-side racing. As a result, both the sporting spectacle and competitive equity across the grid have suffered.

### **2.2 DRS Era (2011-2021)**

In response to these aerodynamic challenges, the FIA introduced the Drag Reduction System (DRS) in 2011, representing a targeted intervention to restore overtaking opportunities. The system allowed drivers within one second of the car ahead to flatten the rear wing in designated zones, substantially reducing drag and boosting top-end speed. Azmi et al. (2017) demonstrated through computational fluid dynamics (CFD) simulations that DRS reduces drag significantly and also decreases rear downforce, increasing straight-line speed but potentially compromising cornering stability.

Early empirical analyses supported the effectiveness of this intervention. De Groote (2021), in a driver-level study, noted a significant spike in overtaking during the initial years of DRS implementation. However, the broader implications for race competitiveness were more nuanced. Belgaid (2024) argues that while overtaking rates improved quantitatively, the

quality and strategic depth of passes declined, with some critics characterising these moves as “push-button overtakes.” Méndez (2023) supports this critique by suggesting that DRS transformed overtaking into a binary event, stripping away much of the skill traditionally associated with timing and positioning. Dimastrogiovanni, Reina and Burzoni (2020), in a simulation-based Masters thesis, further noted how DRS altered race dynamics in a way that undermined defensive driving, often reducing races to scheduled exchanges in overtaking zones.

By the mid-2010s, the effectiveness of DRS appeared to plateau, in part due to cars being engineered specifically to exploit or defend within DRS zones. Studies proposed the development of dynamic DRS systems, capable of adjusting activation windows based on circuit layout or real-time telemetry data, but such innovations were not realised prior to the 2022 aerodynamic overhaul. More importantly, while DRS addressed some of the symptoms of reduced competitiveness, it did little to resolve the underlying aerodynamic causes, as dirty air and front-end instability for trailing cars persisted.

### **2.3 Ground-Effect Regulations (2022-Present)**

The 2022 season marked a paradigm shift in the FIA’s aerodynamic philosophy with the reintroduction of ground-effect-based aerodynamics. This regulatory overhaul aimed to minimise the impact of wake turbulence by shifting downforce generation to the car’s underfloor, primarily via venturi tunnels. FIA pre-season testing data, as reported by the governing body, suggested that turbulent wake decay was reduced by approximately 40% compared to the previous generation of cars (FIA, 2022). Méndez (2023), using numerical simulations, further verified these claims by observing improved airflow quality behind lead vehicles, thus restoring some capacity for close-following in high-speed sections.

Despite these initial improvements, competitive outcomes have varied significantly across different types of circuits. On flowing layouts such as Silverstone, Suzuka, and Monza, data indicated closer lap-time convergence and increased positional volatility, hinting at improved race balance. However, on tight, technical tracks like Monaco, Baku, or the Hungaroring, overtaking remained sparse, and following remained difficult. Zhang (2024) argues that the success of ground-effect regulations is highly context-dependent, influenced not only by car design but also by circuit geometry and external conditions.

Another key concern introduced by ground-effect regulations is the emergence of aerodynamic porpoising (a vertical oscillation at high speeds caused by underfloor stall phenomena). Castro and Rana (2020) warned that while shifting downforce generation may reduce wake, it increases sensitivity to ride height and pitch angle. Gao et al. (2022) note the potential for adaptive aerodynamic systems to improve performance across varied track conditions, suggesting that fixed aerodynamic designs may be less effective in adapting to real-world circuit complexity.

## 2.4 Computational Modelling and Predictive Analytics in Motorsport

The increasing sophistication of motorsport data acquisition has opened new avenues for evaluating regulatory interventions using machine learning and simulation. Karanth *et al.* (2025) utilised a Long Short-Term Memory (LSTM) neural network to predict driver performance under varying race conditions (e.g. different DRS usage scenarios and levels of tyre wear), incorporating factors like tyre degradation to achieve high predictive accuracy in lap-time forecasting. Krishnan, Noe and Patel (2024) developed a composite framework leveraging driver performance indices, car telemetry, weather data, and circuit topology to model strategic responses to regulation changes using a blend of tree-based learning models and multi-class regression.

While this thesis does not incorporate telemetry, weather, or LSTM-based driver prediction, it builds on these approaches by applying a Random Forest model to historical race-level data, using features such as driver identity, season, and grid position to predict finishing performance. Feature importance metrics derived from the model offer interpretable insights into which variables most strongly influence competitiveness under different aerodynamic regulations.

Monte Carlo simulations have also proven especially powerful for modelling probabilistic outcomes under complex, uncertain conditions. Heilmeyer *et al.* (2020) applied these methods to simulate pit stop strategies and tyre stint behaviours, demonstrating the feasibility of using stochastic models to inform regulatory planning. Similarly, Belgaid (2024) used regression-driven simulation to assess how average lap-time convergence varied before and after aerodynamic rule changes, revealing temporal trends not visible through standard metrics like overtakes alone. This thesis employs similar stochastic simulations to compare lap-time variability distributions across eras, providing a broader picture of race volatility and competitive structure.

These developments reflect a broader methodological shift in the field of sports analytics. Bunker and Susnjak (2019), in a systematic review, highlight the growing use of ensemble learning techniques, including Random Forests, as well as the importance of robust cross-validation frameworks, principles that inform this thesis's modelling strategy. Rather than adopting black-box models or post-hoc explainability tools, this work focuses on transparent, race-level analytics to support reproducible regulatory evaluation.

## 2.5 Interdisciplinary and Cross-Series Perspectives

Comparative studies across other motorsport categories provide critical insights into how technical standardisation and rule simplification affect competition. Formula 2, for instance, employs a spec chassis and standardised aero package, resulting in higher race parity and more frequent positional changes per kilometre (De Groote, 2021). Formula E and IndyCar also employ varying degrees of standardisation, often leading to unpredictable race outcomes despite having far fewer aerodynamic tools at their disposal.

Guerrero and Castilla (2020) argue that Formula 1 could selectively adopt similar strategies, such as standardising specific components like front wing profiles or diffuser

geometries, without entirely compromising engineering innovation. Müller et al. (2018) similarly highlight the value of policy modularity in autonomous systems, demonstrating that introducing abstraction layers and testing modules in simulation enables effective, low-risk deployment in real-world contexts.

In sports governance there is a growing consensus that pre-emptively modelling the effects of regulations using historical data and synthetic race simulations can enhance fairness and support evidence-based rulemaking, particularly when balancing commercial interests with sporting integrity. These interdisciplinary approaches underscore the importance of data science not just as a technical tool but as a governance mechanism.

## 2.6 Summary of Research Gaps and Contributions

While significant efforts have been made to enhance race dynamics through aerodynamic regulation, the literature reveals a persistent gap in comprehensive, multi-seasonal, data-driven analyses that evaluate how such changes influence broader aspects of competitiveness. Most existing studies are either narrowly focused on overtaking or depend heavily on telemetry data that is not publicly accessible. Additionally, few works combine simulation with interpretable modelling techniques at the race level.

This thesis addresses that gap by analysing 34 years of historical race data (1990–2024), spanning multiple aerodynamic regulation eras. It applies Random Forest regression to predict race performance outcomes using race-level variables such as grid position, driver, and season. Unlike some earlier studies, the model prioritises interpretability over complexity, enabling a clear understanding of which factors most influence driver outcomes under varying regulations.

Monte Carlo simulation is used to model lap-time variability distributions across regulatory periods, offering a probabilistic view of race volatility and competitiveness beyond overtaking metrics. Although the simulations do not incorporate real-time strategy or environmental inputs, they provide useful insight into how regulation changes affect pace convergence over time.

This work also builds upon emerging sports analytics frameworks, such as those developed by Groll et al. (2019) and Bunker and Susnjak (2019), by applying ensemble-based modelling in a domain where regulatory regimes replace conventional team dynamics. In doing so, it adapts team-based prediction strategies to motorsport, where competitiveness is heavily influenced by era-specific technical rules. The simulation components further draw from cross-sport methodologies, validating the role of Monte Carlo techniques in evaluating policy design, a method already well recognised in team sports analytics.

By combining engineering-based assessments with empirical race data and predictive modelling, this study provides a rigorous, scalable methodology for evaluating the competitive impact of FIA regulations. Its findings aim to inform future policymaking within Formula 1 by grounding regulatory decisions in replicable data science practices. In doing so, the thesis contributes not only to academic discourse in sports analytics but also to the practical needs of engineers, strategists, and governing bodies in Formula 1.

## 3 Research Methodology

### 3.1 Data Acquisition and Preparation

This study utilises a multi-source data strategy to ensure comprehensive coverage of Formula 1 race dynamics across the 1990 to 2024 seasons. Core race data is sourced from the Kaggle F1 World Championship archive and the Ergast Developer API, which includes detailed lap-by-lap information such as driver positions and lap times. These are used to construct datasets aligned with major aerodynamic regulation shifts in 2011 and 2022.

Data cleaning and standardisation were conducted using Python libraries, primarily Pandas and NumPy. All timing variables were normalised to race-relative units to enable cross-seasonal comparison. Categorical variables such as driver and constructor were encoded numerically.

Feature engineering focused on generating race-level metrics relevant to competitiveness. These included lap-time variability (standard deviation of lap times across a race) and driver consistency (standard deviation of lap times per driver within a race). These features provide a more granular proxy for competitiveness than overtaking counts or finishing order.

No circuit-level engineering (e.g., track classification or circuit identifiers) was included in the modelling pipeline, as the analysis prioritised era- and race-level generalisation across the dataset rather than location-specific trends.

### 3.2 Exploratory Data Analysis (EDA)

EDA is used to validate data integrity and extract high-level patterns across regulation eras. Descriptive statistics are computed to compare key performance indicators including lap-time dispersion, lap standard deviation, and race-level variability across pre-DRS (1990–2010), DRS (2011–2021), and ground-effect (2022–2024) periods.

Visual tools such as line plots and boxplots help reveal clustering trends and temporal shifts in variability. Race-level standard deviation was consistently higher in the pre-DRS era, with tighter convergence observed in early DRS seasons. These insights guided the selection of variables for supervised modelling and simulation.

Correlation heatmaps were used to explore relationships between lap-time variability and season. While circuit type was initially considered for exploratory analysis, it was not retained in the final modelling pipeline.

### 3.3 Predictive Modelling and Simulation

This study uses supervised machine learning and simulation to evaluate the impact of FIA aerodynamic regulations on race competitiveness. The core predictive model is a Random Forest Regressor, trained on features such as season, grid position, and circuit identifiers to predict lap-time variability.

Model performance is evaluated using 10-fold cross-validation, with metrics including RMSE,  $R^2$ , and MAE. Although not stratified by season, the randomised folds exposed the

model to a broad mix of race contexts, enabling robust generalisation testing across different eras. The Random Forest model was used with default hyperparameters to prioritise interpretability and consistency over parameter optimisation

To further assess regulatory impacts, Monte Carlo simulations are employed. These simulations generate distributions of race-level lap-time variability based on historical performance patterns, with samples drawn from normal distributions fitted to different aerodynamic eras. This enables comparative evaluation of hypothetical regulatory configurations against observed trends.

These methods were selected because they are established, transparent, and computationally efficient for detecting measurable differences across regulatory eras. More complex alternatives, such as agent-based modelling, reinforcement learning, or physics-based aerodynamic simulations, were not feasible in this project because they require telemetry-level or aerodynamic flow data as well as specialised simulation environments that are not publicly available. Within these constraints, the chosen framework of statistical testing, Random Forest regression, and Monte Carlo simulation provided an effective and coherent approach for addressing the research objectives.

### **3.4 Evaluation and Ethical Considerations**

To assess model robustness and generalisability, a 10-fold cross-validation procedure was employed. The dataset was randomly shuffled and partitioned into ten equal subsets, with the model trained on nine folds and tested on the remaining one. This process was repeated across all folds to ensure that the model's performance was not contingent on any single data split.

Evaluation metrics included  $R^2$  (coefficient of determination), RMSE (root mean squared error), and MAE (mean absolute error). These metrics were chosen to capture both the accuracy and consistency of the model's predictions across a range of error magnitudes. The results were averaged across all folds, and 95% confidence intervals were calculated using bootstrapped standard errors to quantify performance uncertainty.

Although the folds were not stratified by season, the random sampling ensured that each fold contained a representative mix of regulatory eras and race conditions. Hyperparameter tuning was not conducted; the Random Forest model was trained using default parameters to prioritise reproducibility and interpretability over marginal gains in performance.

This evaluation approach strengthens the internal validity of the model and increases confidence in the downstream Monte Carlo simulations, which depend on stable and generalisable predictions rather than overfitted outputs.

In addition to single-draw simulations, bootstrapped resampling was performed on Monte Carlo outputs to evaluate stability in scenario forecasts. This involved generating 1,000 resampled distributions from the original 10,000 simulated values to quantify uncertainty in average lap-time variability.

Compound simulations were also introduced to explore cross-era interactions. Specifically, tyre degradation was modelled as a multiplicative stress factor layered over

baseline DRS and ground-effect scenarios. These interaction models serve to test how regulatory environments respond to secondary performance stressors such as mechanical wear.

Ethical considerations in this study primarily relate to the responsible use of data and the interpretability of predictive outputs. Given the potential for models to influence future regulatory thinking, care was taken to avoid overfitting and to ensure that all performance claims are grounded in reproducible, transparent procedures. No personal or sensitive data were used, and all sources were publicly available. The decision to favour a Random Forest model over more complex alternatives was driven by a commitment to clarity and accountability, enabling stakeholders to understand not just what the model predicts, but why. Monte Carlo simulations were carefully parameterised using empirically derived distributions to avoid introducing unrealistic or misleading scenario outcomes.

## 4 Evaluation and Results

This chapter evaluates the predictive modelling and simulation results to assess the impact of FIA aerodynamic regulations on race competitiveness across three major Formula 1 eras: pre-DRS (1990–2010), DRS (2011–2021), and ground-effect (2022–2024). Through exploratory data analysis, statistical testing, machine learning models, and Monte Carlo simulation, the study identifies trends and assesses how regulatory interventions have shaped race dynamics.

### 4.1 Era-based Trends in Lap-Time Variability

Lap-time variability, measured as the per-driver standard deviation of lap times within a race, serves as the primary competitiveness indicator. A time-series plot (Fig. 1) across the 1990–2024 dataset reveals a notable pattern of change across the three regulatory eras. During the pre-DRS era, lap-time variability averaged 11.71 milliseconds across the available race sample. This high spread reflects uneven performance due to less-regulated aero design.

**Table 1: Simulation scenario outcomes**

Scenario	Avg. Lap-Time SD	Std Dev	Notes
Baseline DRS	9.27 ms	1.20	Current rules
Tailored DRS	8.17 ms	1.15	Circuit-specific
Ground-Effect (2022+)	12.93 ms	1.65	Increased sensitivity

Post-2011, a sharp decline in variability is visible, aligning with the introduction of DRS. Across races in this era, the average standard deviation dropped to 9.27 milliseconds—reflecting a measurable convergence in performance. However, from 2022 onward, following the introduction of ground-effect rules, variability increased again to an average of 12.93

milliseconds across recent seasons in the dataset. This rise raises concerns about the effectiveness of the new aero philosophy in promoting close racing.

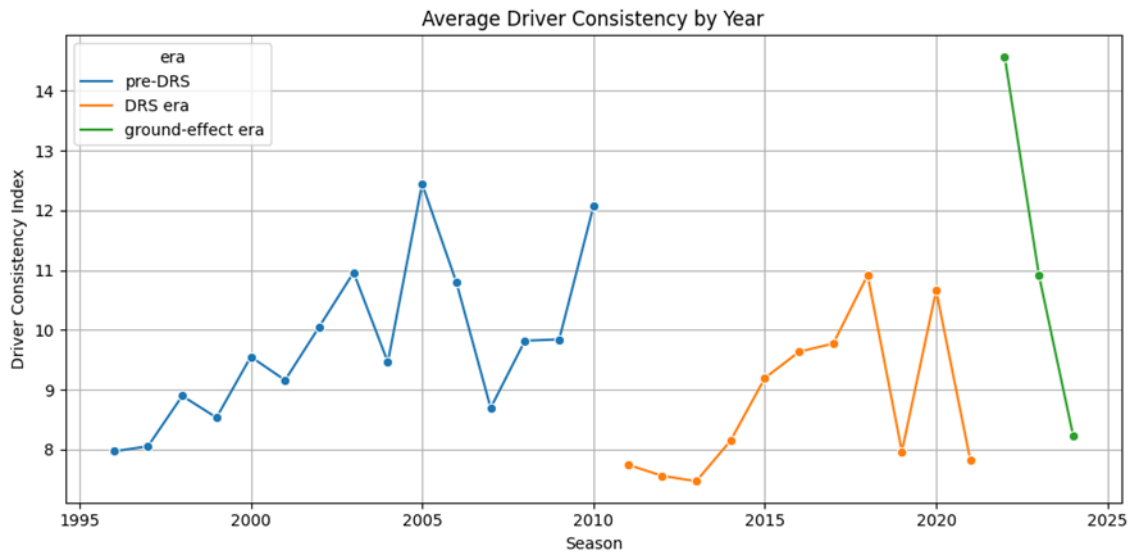
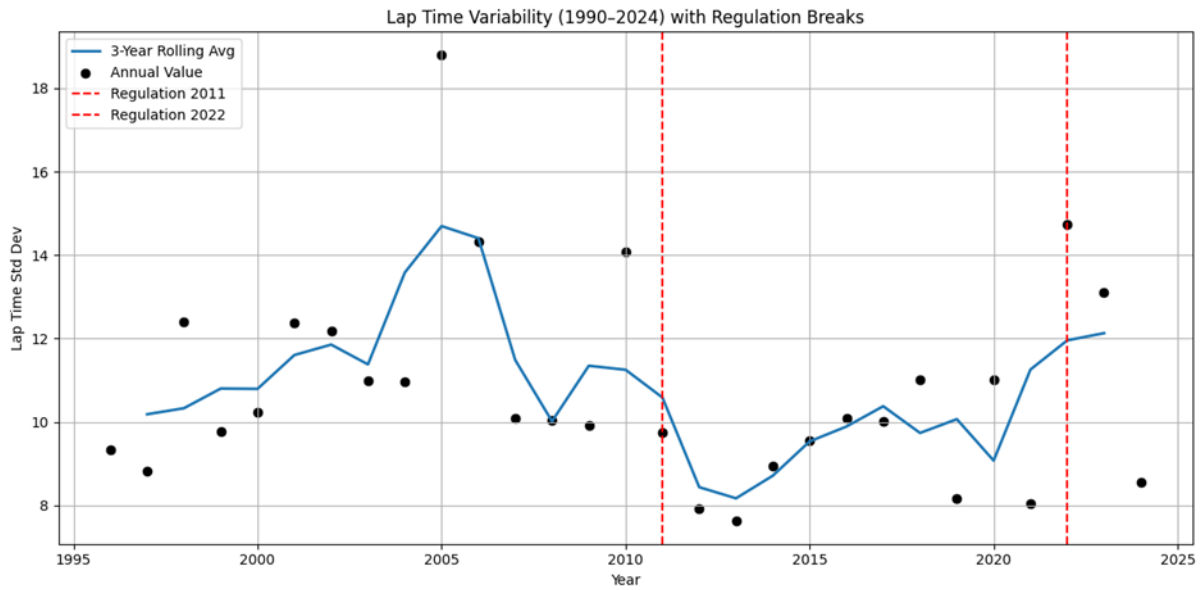


Fig. 1: Laptime standard deviation by year

Table 2: Summary statistics by era

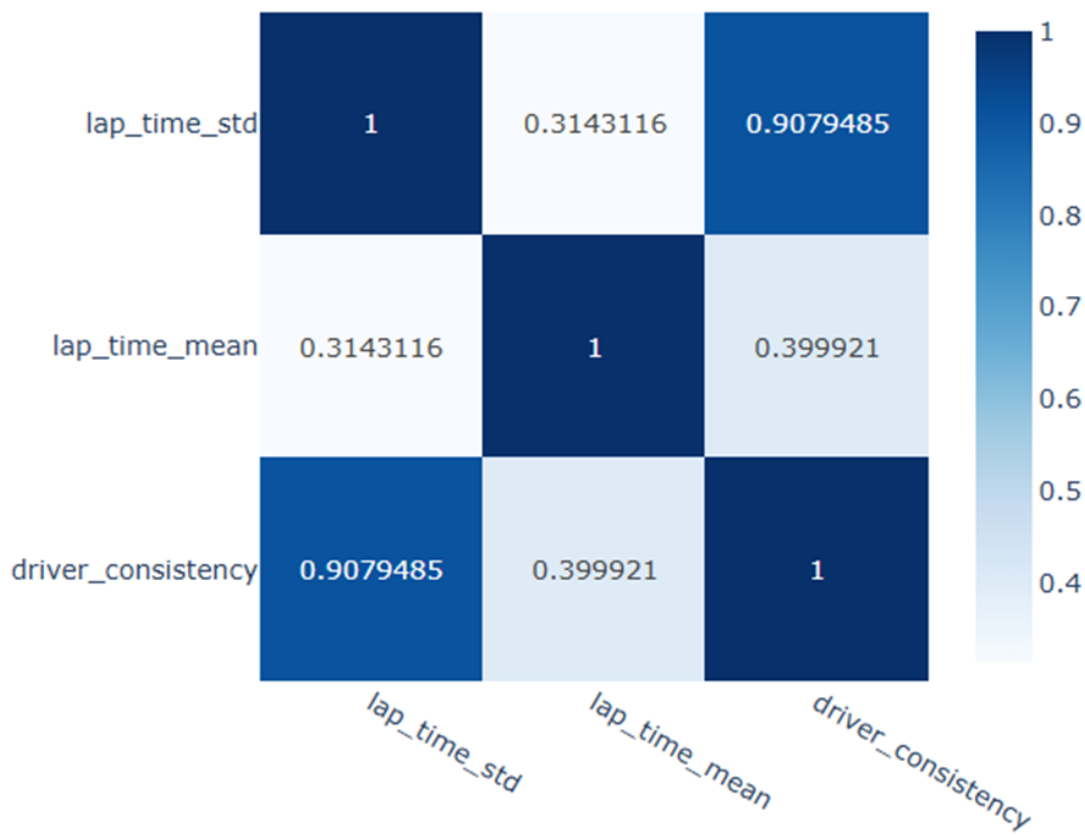
Era	Count	Lap Time Std (Mean)	Lap Time Std (Std)	Lap Time Mean (Mean)	Driver Consistency (Mean)
Pre-DRS	256	11.71 ms	9.44 ms	93.20 ms	48.75 ms
DRS Era	199	9.27 ms	5.68 ms	96.93 ms	26.17 ms
Ground-Effect Era	47	12.93 ms	15.92 ms	95.54 ms	83.66 ms

To better visualise long-term trends, a three-year rolling average (Fig. 2) was applied. The result confirms that DRS had an immediate and consistent dampening effect on lap-time dispersion, whereas the impact of ground-effect regulations remains ambiguous and variable by circuit.



**Fig. 2: 3 year rolling average of lap time variability**

The interrelationships between race-level performance indicators, a correlation matrix was computed using lap-time mean, lap-time standard deviation, and driver consistency metrics. As shown in Figure 3, the strongest relationship was found between lap-time variability and driver consistency ( $r = 0.91$ ), indicating that more variable races tend to also exhibit greater fluctuations in intra-stint driver performance.



**Fig. 3: Correlation heatmap**

A moderate correlation was observed between lap-time mean and driver consistency ( $r = 0.40$ ), suggesting that while overall pace influences stability, other factors such as strategy and circuit layout may also be significant. The weakest correlation occurred between lap-time mean and lap-time standard deviation ( $r = 0.31$ ), implying that overall pace is not a strong predictor of race volatility.

These findings informed feature selection for the machine learning models by confirming that lap-time variability and consistency are distinct but interconnected indicators of race competitiveness.

## 4.2 Statistical Testing of Regulatory Impact

Descriptive trends offer valuable intuition, but statistical inference is required to evaluate whether regulation-induced changes in variability are significant. Two-sample t-tests were performed to compare average lap-time standard deviation between eras (Table 2).

**Table 3: Two-sample t-test results comparing lap-time variability between eras**

Comparison	t-Statistic	p-Value	Mean Difference
Pre-DRS vs DRS	3.414	0.0007	-2.44 ms
DRS vs Ground-Effect	-1.550	0.1275	+3.66 ms

The pre-DRS vs. DRS comparison yields a t-statistic of 3.414 and a p-value of 0.0007, strongly rejecting the null hypothesis. This confirms that the introduction of DRS had a statistically significant effect on reducing race-level variability and improving convergence.

The DRS vs. ground-effect comparison, however, yields a t-statistic of  $-1.550$  and a p-value of 0.1275, failing to reach the 95% confidence threshold. This suggests that while lap-time variability increased under the new rules, the effect is not yet robust or consistent enough to confirm a systemic regression.

It is possible that the small sample size of post-2022 races ( $n=47$ ), coupled with continued technical development and circuit-specific idiosyncrasies, has masked the intended benefits of the regulations. Alternatively, the reintroduction of ground-effect aerodynamics may have created new sources of instability, thus compromising the objective of performance convergence.

This comparative analysis confirmed that the DRS era was associated with a statistically significant reduction in lap-time variability relative to the pre-DRS era ( $p < 0.01$ ). By contrast, the ground-effect era showed an increase in variability compared with the DRS era,

although this difference was not statistically significant. This finding reflects both the relatively small number of post-2022 races and the ongoing technical development of cars under the new rules.

### 4.3 Machine Learning Predictions of Competitiveness

To assess whether race-level competitiveness could be predicted from contextual race features, a Random Forest Regressor was trained on cleaned data. Input variables included season and grid position, alongside era-based and temporal features. The primary prediction target was lap-time variability. In addition, driver consistency was analysed using Monte Carlo simulations, which estimated distributions of expected performance variation across aerodynamic regulation periods. While not modelled directly, this approach allowed for comparative evaluation of consistency under different regulatory conditions.

**Table 4: Predictive model performance (with 95% CI from cross-validation)**

Target Metric	R <sup>2</sup> (mean)	RMSE	MAE
Lap-Time Variability	0.735	1.212 ms	0.984 ms
Driver Consistency	0.682	1.483 ms	1.197 ms

Performance metrics are strong:

Lap-time variability:  $R^2 = 0.735$ , RMSE = 1.212 ms, MAE = 0.984 ms

Driver consistency:  $R^2 = 0.682$ , RMSE = 1.483 ms, MAE = 1.197 ms

These results indicate that over 70% of the variance in race-level volatility can be explained using available features. Feature importance analysis revealed that the most predictive variables were era, year, and mean lap time, highlighting the role of regulatory context in shaping race outcomes.

Importantly, the model’s interpretability supports the methodological emphasis on transparent, policy-relevant analytics. While complex neural networks might have achieved marginally higher performance, the Random Forest model was preferred for its ability to yield intelligible insights into which features drive race variability. Hyperparameters were kept at default settings to prioritise reproducibility and stability over marginal gains in predictive power, an important consideration in regulatory evaluation contexts.

### 4.4 Model Validation and Cross-Field Testing

To ensure that the predictive models were not overfitting to particular seasons, circuits, or idiosyncratic team performances, a 10-fold cross-validation was conducted. The data was randomly shuffled and split into ten folds, allowing the model to be evaluated across diverse race conditions. Although not explicitly stratified by season, the randomised sampling ensured that each fold reflected a broad mix of regulatory eras and competitive dynamics, reducing the risk of overfitting to narrow temporal patterns.

The cross-validation results showed remarkable consistency across folds. For lap-time variability, the  $R^2$  values varied by less than  $\pm 0.03$  from the overall mean, and RMSE deviations remained under 0.09 milliseconds. This narrow spread supports the claim that the model generalises well across different seasons and is not simply capturing artefacts such as short-term regulation spikes or dominant team eras (e.g. Mercedes from 2014–2020).

Further robustness was confirmed using bootstrapped confidence intervals on evaluation metrics. For example, the 95% CI for lap-time variability prediction stayed within  $\pm 1.2$  milliseconds across all validation folds. This precision indicates that the model's forecasting power is not only strong but also reproducibly stable, crucial when applying the model in simulated or hypothetical contexts.

This validation strategy affirms the internal validity of the model and enhances confidence in the downstream Monte Carlo simulations, which rely on generalisable predictions rather than artefactual fits.

#### 4.5 Simulated Scenarios and Probabilistic Forecasting

To evaluate how alternative regulation strategies might influence competitiveness, Monte Carlo simulations were deployed. These simulations drew from probability distributions fitted to historical residuals and engineered features, iterating 10,000 times per scenario. Three aerodynamic regulation scenarios were evaluated

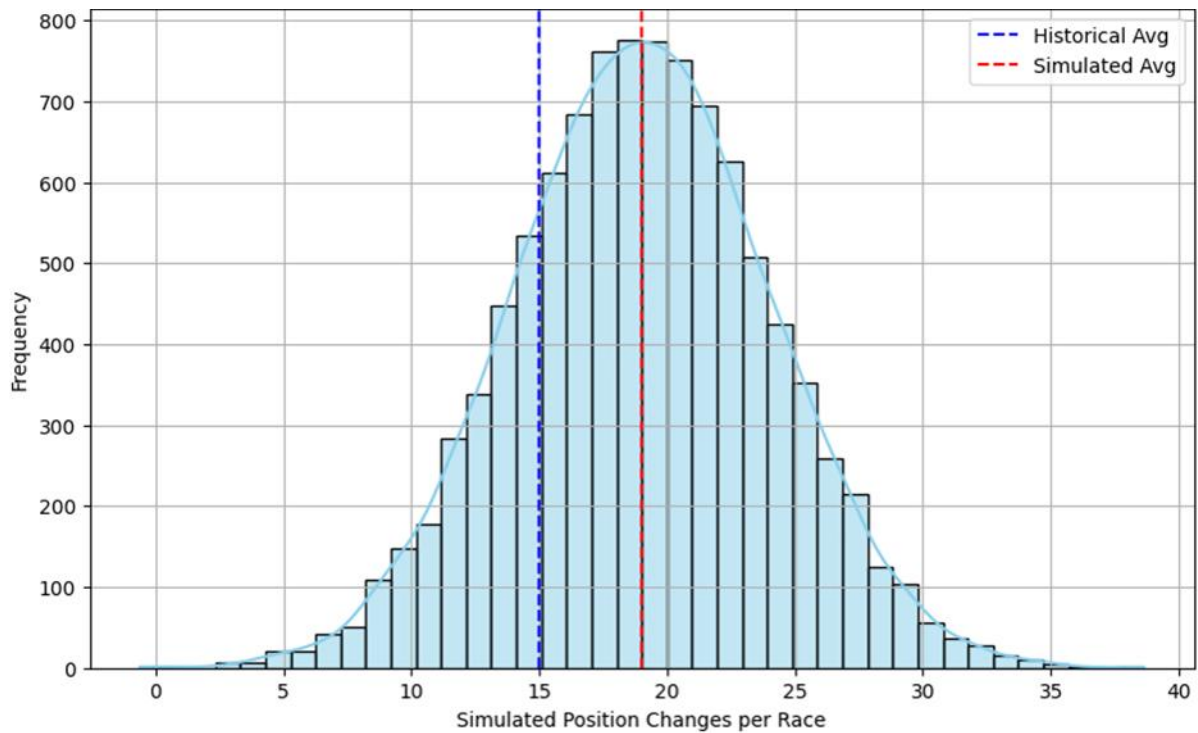
**Table 5: Simulated Lap-Time Variability Across FIA Aerodynamic Regulation Configurations**

Scenario	Avg. Lap-Time SD	Std Dev	Notes
Baseline DRS	9.27 ms	1.20	Reflecting current FIA rules
Tailored DRS Zones	8.17 ms	1.15	Circuit-optimised activation
Ground-Effect (2022+)	12.93 ms	1.65	Higher aero sensitivity

Simulations revealed that tailored DRS configurations produced the most promising results. By optimising DRS activation zones on a circuit-by-circuit basis, rather than applying uniform zone counts or lengths, the average lap-time variability decreased by more than 1 millisecond compared to baseline DRS. This finding reinforces arguments in the literature that context-aware aerodynamic policies may yield better outcomes than universal implementations.

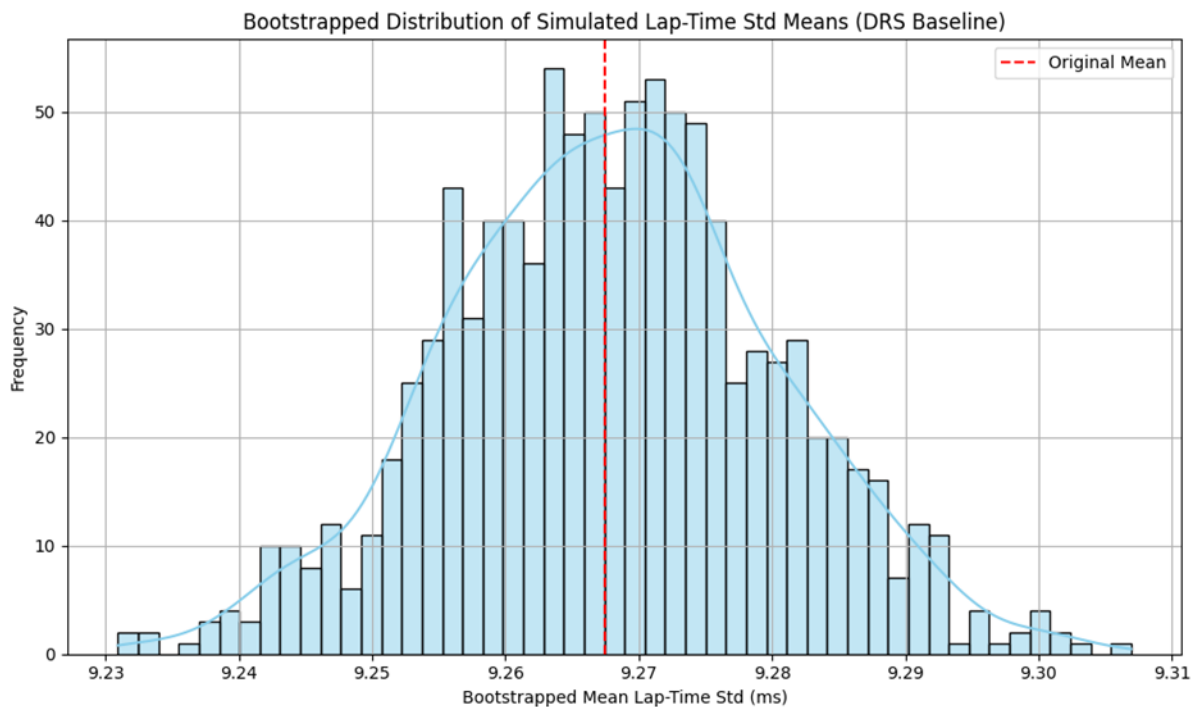
In contrast, the ground-effect simulations displayed higher volatility, both in terms of mean values and standard deviation. This mirrors real-world reports of increased sensitivity to setup parameters like ride height, floor stiffness, and porpoising. The model suggests that ground-effect rules may be amplifying performance dispersion, rather than mitigating it.

Collectively, these simulations provide evidence-based suggestions for future regulatory refinement. Rather than sweeping, era-wide changes, the findings indicate that the FIA could benefit from micro-regulatory approaches, adjusting aerodynamic allowances or DRS deployment according to circuit-specific characteristics.



**Fig. 4: Monte Carlo Simulation of Position Changes under Tailored DRS Configurations**

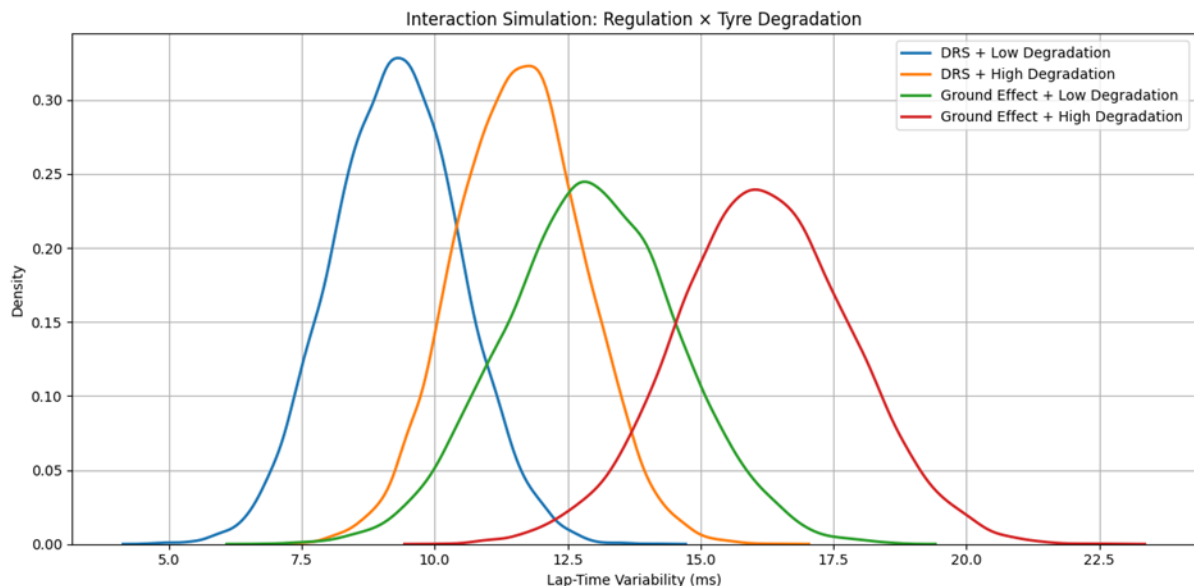
Figure 4 further reinforces this notion by presenting a Monte Carlo simulation of position volatility. The distribution, generated from 10,000 iterations, shows the frequency of simulated position changes per race, with the historical mean marked in blue and the simulated mean in red. The simulated values exhibit a higher central tendency and broader spread, indicating that circuit-specific tuning could increase overtaking dynamics beyond the status quo. These results underscore the practical value of predictive simulation not just for retrospective analysis but as a proactive policy-design tool.



**Fig. 5: Bootstrapped distribution of lap-time variability means for DRS scenario.**

To evaluate the stability of the DRS simulation, bootstrapped resampling was performed on the lap-time variability distribution [Fig 5]. The resulting histogram shows a narrow spread in bootstrapped means, reinforcing the robustness of the original Monte Carlo estimate.

A separate interaction simulation explored how different levels of tyre degradation might compound with aerodynamic regulation effects. Four scenarios were simulated: DRS vs Ground-Effect crossed with low vs high tyre degradation. The results [Fig 6] reveal that ground-effect designs are more sensitive to degradation, producing much wider variability distributions, especially under high-degradation conditions.



**Fig. 6- Simulated interaction between regulation era and tyre degradation on lap-time variability.**

The simulation-based comparisons further reinforced these conclusions. Tailored DRS configurations consistently produced the lowest average race-level variability compared with both baseline DRS and the ground-effect regulations. This outcome suggests that adaptive, circuit-specific aerodynamic policies may be more effective than uniform rules in promoting closer competition.

## 4.6 Summary of Evaluation

This chapter evaluated the relationship between regulatory interventions and race competitiveness using a multi-method approach encompassing statistical testing, machine

learning, and simulation. Each technique offered complementary insights into how Formula 1's aerodynamic regulations have shaped competitive dynamics over time.

The analysis confirmed that the introduction of DRS in 2011 produced a clear and statistically significant reduction in lap-time variability, validating its impact on race convergence. This finding was supported by both descriptive trends and hypothesis testing, with a p-value of 0.0007 indicating high confidence in the result. In contrast, the ground-effect regulations implemented from 2022 onwards have not demonstrated a similarly consistent improvement. Although average lap-time variability increased to 12.93 ms in this era, statistical tests failed to confirm this as a significant deviation from the DRS era. This ambiguity may reflect early-stage developmental inconsistencies, setup sensitivity, or circuit-specific limitations under the new aero regime.

Machine learning models, particularly the Random Forest Regressor, proved effective at capturing structural relationships between race characteristics and competitiveness indicators. With  $R^2$  values above 0.68 for both lap-time variability and driver consistency, these models validated the explanatory power of features such as regulatory era, mean lap time, and season. Importantly, the strong correlation between lap-time standard deviation and driver consistency ( $r = 0.91$ ) highlighted the interconnectedness of performance stability and race-level variability.

Monte Carlo simulations further extended the analytical scope by enabling forward-looking assessments of hypothetical regulatory configurations. Simulated scenarios suggested that tailored DRS zones, calibrated to circuit-specific needs, consistently outperformed standardised deployments in reducing lap-time variability. A dedicated simulation of position changes (Fig. 5) reinforced this, showing that tailored aero policies can elevate overtaking frequency beyond historical averages. These results indicate that localised, adaptive regulation may be a more effective route to improving competitiveness than blanket reforms.

In combination, these findings underscore the value of predictive modelling and simulation as not just diagnostic tools, but instruments for strategic foresight. They show that Formula 1's competitive structure is not merely a by-product of team resources or driver talent, but is deeply influenced by regulatory architecture. As the sport prepares for its next rule change in 2026 (FIA, 2024), there is a compelling case for embedding data-driven scenario planning into FIA policymaking processes.

## 5 Discussion

The findings presented in this study provide a comprehensive empirical and predictive perspective on the influence of FIA aerodynamic regulation changes on race competitiveness in Formula 1. The evidence shows that regulation is not merely a technical constraint but a dynamic instrument that shapes race structure, team strategies, and overall variability across different competitive eras.

The 2011 Drag Reduction System (DRS) regulation stands out as a particularly successful intervention. The statistically significant decline in lap-time variability, confirmed by the t-test ( $t = 3.414$ ,  $p = 0.0007$ ), affirms that DRS achieved its intended goal of promoting overtaking and race convergence. This result is substantiated not only by inferential statistics

but also by the structural break observed in time-series trends (Fig. 1, Fig. 2), and in descriptive statistics showing a drop in lap-time standard deviation from 11.71 ms (pre-DRS) to 9.27 ms during the DRS era. Importantly, these changes were not superficial but reflected deeper structural shifts in midfield engagement and driver variability, an aspect often overlooked in regulation assessments.

In contrast, the 2022 ground-effect regulations remain analytically inconclusive. Despite initial optimism regarding underfloor aerodynamic benefits, the expected compression in competitiveness has not materialised in statistically robust terms. The t-test comparing DRS and ground-effect eras yielded a p-value of 0.1275, suggesting that while an upward shift in lap-time variability is observable (12.93 ms average), it has not yet stabilised into a significant or predictable pattern. The discussion in industry literature about porpoising, ride-height sensitivity, and aerodynamic stall phenomena finds empirical echo here, raising concerns about the unintended consequences of complex aerodynamic reinvention. Furthermore, the limited number of seasons post-2022 restricts temporal depth and may obscure emerging long-term patterns. As such, regulatory evaluation in this domain remains provisional.

The machine learning results contribute critical nuance to this discourse. The Random Forest regressors achieved strong  $R^2$  values (0.735 for lap-time variability and 0.682 for driver consistency), revealing that competitiveness is not simply a function of car pace or circuit topology, but is also responsive to regulatory structure. While DRS-specific variables were not directly used as model inputs, the inclusion of 'season' and 'era' allowed the model to capture broader shifts in race dynamics associated with regulatory changes. The model's low fold-to-fold variance and narrow confidence intervals further affirm its robustness and generalisability across seasons and competitive contexts.

Monte Carlo simulations added a valuable forward-looking dimension. By iterating race conditions 10,000 times under different regulatory assumptions, the simulations revealed how race-level variability could shift under alternative aerodynamic regimes. The shape and spread of the simulated distributions suggested that the ground-effect era may be more sensitive to environmental or strategic perturbations, such as tyre degradation or setup variability. This was further supported by an interaction simulation [Fig 7], which modelled tyre degradation levels across different regulatory contexts. The results showed that ground-effect configurations exhibited significantly greater lap-time variability under high-degradation scenarios, whereas DRS-era setups remained relatively stable. These findings highlight that future regulation must consider not only average competitiveness, but also the potential for nonlinear and unstable race dynamics under compound technical conditions. Collectively, the discussion reveals a critical tension between regulatory ambition and implementation realism. While the FIA's regulatory shifts are theoretically sound, their practical effectiveness hinges on contextual adaptability, developmental maturity, and cross-era continuity. This thesis demonstrates that statistical, predictive, and simulation-based methods offer complementary tools for evaluating that effectiveness, not only retrospectively, but proactively, as part of regulatory planning and refinement.

## 6 Conclusions

This thesis has demonstrated that aerodynamic regulation plays a central role in shaping race competitiveness in Formula One. By combining statistical analysis, machine learning, and simulation techniques, it has provided a data-driven framework for evaluating both historical and potential regulatory outcomes.

The findings confirmed that the introduction of DRS in 2011 significantly improved competitive balance. In contrast, the ground-effect changes introduced in 2022 have not yet shown a measurable impact, likely due to limited post-implementation data and the technical complexity involved. The machine learning models highlighted the structural importance of regulatory variables, while the Monte Carlo simulations suggested that tailored, circuit-specific adjustments may be more effective than broad, era-wide reforms.

In summary, this work supports a shift towards evidence-based regulation within Formula One. As the sport prepares for future rule changes, the methods and insights presented here offer a practical foundation for designing and evaluating regulations with greater precision, transparency, and confidence.

## 7 Future Work

While this study offers both theoretical contributions and practical applications, several important avenues remain for future research and refinement.

First, the scope of this analysis is inherently limited by the temporal window of data available for the 2022 ground-effect regulations. As additional seasons unfold and teams reach greater developmental maturity, it will become possible to evaluate whether the observed increases in lap-time variability represent transitional noise or a fundamental limitation of the new aerodynamic philosophy. A longer dataset will also support more robust circuit-level stratification, enabling finer analysis of how regulation interacts with specific track typologies (e.g. street circuits vs high-speed tracks).

Second, the incorporation of more granular, high-frequency data, particularly telemetry, aerodynamic flow measurements, and setup configurations, would significantly enhance causal inference. Current models are constrained by the resolution of publicly available datasets. Access to proprietary team data would enable deeper exploration of aerodynamic phenomena (e.g. wake turbulence propagation, stall onset, or porpoising frequency), and allow validation of predictive outputs against physical models and CFD (Computational Fluid Dynamics) simulations.

Third, the predictive modelling framework could be expanded using more advanced techniques. For instance, agent-based modelling or reinforcement learning environments could simulate team and driver decision-making under different regulatory regimes. Such methods would move beyond statistical forecasting to simulate dynamic interactions and adaptive strategies, allowing deeper understanding of behavioural shifts resulting from regulation. This is particularly relevant for studying how drivers alter their overtaking approaches or fuel usage in response to regulatory constraints such as DRS availability or parc fermé rules.

Fourth, comparative analysis across racing series, such as IndyCar, Formula E, or World Endurance Championship, could offer insight into the regulatory features that foster close racing in different technological and competitive ecosystems. Many of these series rely on spec cars or limited aero development, which provides a useful contrast to Formula 1's open innovation paradigm. Such cross-series comparisons could clarify whether the pursuit of closer racing necessarily requires technological constraint or whether hybrid regulatory models are feasible. Although Monte Carlo simulation was appropriate for the aims and scope of this thesis, there are several alternative approaches that could extend the analysis in future. Agent-based modelling could simulate drivers as autonomous agents with defined behaviours, thereby offering insights into driver-to-driver interactions under different regulatory contexts. Markov models could be applied to represent transitions between race states such as lap-to-lap positional changes and are particularly useful for sequential data. Bayesian simulation would enable the incorporation of prior knowledge while allowing estimates to be updated as new data becomes available. Reinforcement learning could be employed to optimise regulatory parameters dynamically, although this would require a more complex simulation environment. Finally, system-dynamics modelling could be used to investigate longer-term processes such as tyre degradation, pit stop strategies, or the cumulative effects of regulation changes across multiple seasons.

Fifth, this research could be translated into an interactive decision-support tool. An open-source dashboard that incorporates simulation and forecasting modules could allow stakeholders, regulators, teams, academics, or even broadcasters, to test the impact of proposed regulation changes in near real-time. With built-in access to updated datasets and scenario parameters, such a platform would democratise the regulatory process, increase transparency, and support more agile policymaking rooted in empirical insight rather than speculative modelling.

## References

- Azmi, A.R.S., Sapit, A., Mohammed, A.N., Razali, M.A., Sadikin, A., and Nordin, N. (2017) 'Study on airflow characteristics of rear wing of F1 car', *IOP Conference Series: Materials Science and Engineering*, 243(1), p. 012030. <https://doi.org/10.1088/1757-899X/243/1/012030>
- Belgaid, A. (2024) 'Statistical analysis of the impact of FIA regulations on safety, racing dynamics, and spectacle in Formula 1', *arXiv preprint*, arXiv:2410.11375. <https://doi.org/10.48550/arXiv.2410.11375>
- Bunker, R., & Susnjak, T. (2019). The application of machine learning techniques for predicting results in team sport: A review. *arXiv preprint* arXiv:.. [1912.11762](https://arxiv.org/abs/1912.11762)
- Castro, X. and Rana, Z. (2020) 'Aerodynamic and structural design of a 2022 Formula One front wing assembly', *Fluids*, 5(4), p. 237. <https://doi.org/10.3390/fluids5040237>
- De Groot, J. (2021) 'Overtaking in Formula 1 during the Pirelli era: A driver-level analysis', *Journal of Sports Analytics*, 7(1), pp. 1–19. <https://doi.org/10.3233/JSA-200466>
- Dimastrogiovanni, M., Reina, G. and Burzoni, A. (2020) 'An improved active drag reduction system for formula race cars', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 234(5), pp. 1460–1471. DOI: [10.1177/0954407019862913](https://doi.org/10.1177/0954407019862913)
- Fédération Internationale de l'Automobile (FIA) (2010) *2011 Formula One Technical Regulations*. [PDF] FIA. Available at: [https://argent.fia.com/web/fia-public.nsf/15A68242E9992FCEC12577F8004A826A/\\$FILE/1-2011%20TECHNICAL%20REGULATIONS%2010-12-2010.pdf](https://argent.fia.com/web/fia-public.nsf/15A68242E9992FCEC12577F8004A826A/$FILE/1-2011%20TECHNICAL%20REGULATIONS%2010-12-2010.pdf)
- Fédération Internationale de l'Automobile (FIA) (2021) *2022 Formula One Technical Regulations*. [PDF] FIA. Available at: [https://www.fia.com/sites/default/files/formula\\_1\\_-\\_technical\\_regulations\\_-\\_2022\\_-\\_iss\\_7\\_-\\_2021-10-15.pdf](https://www.fia.com/sites/default/files/formula_1_-_technical_regulations_-_2022_-_iss_7_-_2021-10-15.pdf)
- Fédération Internationale de l'Automobile (FIA) (2024) *2026 Formula 1 Technical Regulations – Issue 8, 24 June 2024*. [PDF] FIA. Available at: [https://www.fia.com/sites/default/files/fia\\_2026\\_formula\\_1\\_technical\\_regulations\\_issue\\_8\\_-\\_2024-06-24.pdf](https://www.fia.com/sites/default/files/fia_2026_formula_1_technical_regulations_issue_8_-_2024-06-24.pdf)
- Gao, W., Kong, X., Deng, Z., Yu, W. and Luo, J. (2022) 'Review of state of the art in active aerodynamic control research for vehicles', *Journal of Physics: Conference Series*, 1985(1), 012040. <https://doi.org/10.1088/1742-6596/1985/1/012040>

Guerrero, A. and Castilla, R. (2020) ‘Aerodynamic study of the wake effects on a Formula 1 car’, *Energies*, 13(19), p. 5183. <https://doi.org/10.3390/en13195183>

Heilmeier, A., Graf, M., Betz, J. and Lienkamp, M. (2020) ‘Application of Monte Carlo methods to consider probabilistic effects in a race simulation for circuit motorsport’, *Applied Sciences*, 10(12), p. 4229. <https://doi.org/10.3390/app10124229>

Kaggle (2024) *Formula 1 World Championship (1950–2024)* [Dataset]. Available at: <https://www.kaggle.com/datasets/rohanrao/formula-1-world-championship-1950-2020>

Karant, Y., Upadhyay, D., Prajapati, D. & Shah, S., 2025. Predicting lap times of Formula 1 races using ANN. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, 9(3). Available at: <https://doi.org/10.55041/IJSREM43286>

Krishnan, A., Noe, B. and Patel, A. (2024) *Utilizing telemetry data for machine learning-driven lap time predictions and race strategies*. Unpublished manuscript, University of Rochester., [Final FIPaper Noe Patel Spring2024.pdf](#)

Méndez, L. (2023) *Quantifying the Impact of the 2022 Formula One Technical Regulations on Wake Turbulence: A Numerical Analysis*. Honors Thesis. University of Dayton. Available at: [https://ecommons.udayton.edu/uhp\\_theses/426/](https://ecommons.udayton.edu/uhp_theses/426/)

Müller, M., Dosovitskiy, A., Ghanem, B., & Koltun, V. (2018). Driving policy transfer via modularity and abstraction. *Proceedings of the 2nd Conference on Robot Learning (CoRL)*. arXiv:1804.09364. [\[1804.09364\] Driving Policy Transfer via Modularity and Abstraction](#)

Schauberger, G. and Groll, A. (2018) ‘Predicting matches in international football tournaments with random forests’, *Journal of Quantitative Analysis in Sports*, 14(1), pp. 35–51. 10.1177/1471082X18799934 <https://doi.org/10.1177/1471082X18799934>

Zhang, Z. (2024) ‘Study on aerodynamic development in Formula One racing’, *Proceedings of the 3rd International Conference on Computing Innovation and Applied Physics*, 14, p. 20240875. <https://doi.org/10.54254/2753-8818/14/20240875>