[L4.5] POC IMPLEMENTATION IN REAL CONDITIONS

Mise en œuvre d'un POC

Main authors: Thibaut Jonville (CEREMA), Leo Maisonobe (STRMTG), Omar Hafa (RATP), Youssoufa Tahirou (VALEO), Christophe Bohn (IRT-SystemX), Emmanuel Arbaretier (APSYS-AIRBUS), Marie-Françoise Guyonnaud (ATC France)

Keywords: Real-world test and validation, Criteria, Methods, Infrastructure, POC, Assessment, Results

Abstract. This deliverable (Task 4.5) is the application of the Task 4.4 of the PRISSMA project. The main objective is to implement on open roads the evaluation methodology defined in previous tasks using existing prototypes, and to conclude on the methodology pertinence. The methodology defined in Task 4.2 & Task 4.3 aims to assess the behaviors of AI-based autonomous/automated vehicles and AI-based supervision systems in real conditions.

In order to ensure the methodology adequacy with real world use cases, a common urban setting is chosen, the itinerary is offered by the Ville de Paris municipality and equipped with the Paris2Connect infrastructure. This setting is used to run two prototype vehicles: [redacted], and the second is RATP Supervision system Evaluation.

The deliverable has been structured in three main parts: first we summarise the base elements required for the tests implementation, in this case AI-based systems, plus a representative road setting and infrastructure, scenarios and evaluation criteria, as described in related deliverables L4.2, L4.3. Second, describe the testing equipment and protocols necessary. Lastly reports the data obtained and post-testing analysis conducted on the data.

Résumé. Ce livrable (Tâche 4.5) est l'application de la Tâche 4.4 du projet PRISSMA. Le principal objectif est de mettre en pratique sur route ouverte la méthodologie définie dans les tâches précédentes à travers des prototypes existants, et de conclure quant à la pertinence de cette méthodologie. La méthodologie définie par les Tâche 4.2 & 4.3 vise à évaluer les

comportements de véhicules automatisés/autonomes intégrant une IA et de systèmes de supervision avec IA en conditions réelles.

Afin d’assurer le déroulement de cette méthodologie sur des cas d’usage concrets, le choix est fait d’un parcours commun proposé par la Ville de Paris et équipé de l’infrastructure Paris2Connect. Deux véhicules prototypes sont ciblés pour valider la méthodologie : le premier est le prototype Drive4U L4 Urban Pilot de Valeo, le second l’évaluateur du système de Supervision RATP.

Le livrables est structuré en trois parties. La première partie nous rappelle les prérequis pour le déroulement de la méthodologie : avec d’une part un véhicule, un parcours, une infrastructure et d’autre part, des scénarios et des KPI. La seconde partie nous montre la mise en place des tests autour du parcours Paris2Connect. Et la troisième partie est la restitution de l’expérimentation et l’aboutissement de la méthodologie.

Table des matières

Definitions and Acronyms

Definitions:

Dynamic Driving Task (DDT)	All operational and tactical functions performed in real time required for vehicle movement, including: 1) control of the lateral and longitudinal movement of the vehicle, 2) monitoring of the road environment 3) reaction to events in the roadway environment, 4) preparation and reporting of manoeuvres, 5) activation of visibility functions. NOTE: Excluded are strategic functions such as trip scheduling, definition of times and positions of departure and arrival points.
Operational Domain (OD)	Real-world conditions that an ADS may experience [1] Set of operating conditions, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics [2]
Operational Design Domain (ODD)	Operating conditions under which an ADS is designed to operate safely [2] Specific conditions under which a given driving automation system is designed to function [3]
Object and Event Detection and Response (OEDR)	DDT subtasks that include monitoring the driving environment (detecting, recognizing and classifying objects and events), as well as executing an appropriate response to such objects and events.
Planned event	A situation that is known in advance, for example, at the time of activation, such as a crossing point (e.g., highway exit, etc.) and requires a transition request.
Unexpected event	A situation that is not known in advance, but which is assumed to be highly likely to occur, e.g., road work, inclement weather, approach of an emergency vehicle, lack of lane markings, falling truck load (collision), and which requires a transition request.
EGO	Name used for automated/autonomous vehicle. In the analysis of a driving situation involving multiple vehicles, the EGO vehicle is the subject of the study, the one whose behaviour we seek to understand or control in interactions with the other vehicles involved in this situation

Acronyms

ADS: Automated Driving System
AI: Artificial Intelligence
ALKS: Automated Lane Keeping System
ARTS: Automated Road Transport System
CAM: Cooperative Awareness Message
CPM: Collective Perception Message
DDT: Dynamic Driving Task
DENM: Decentralized Environmental Notification Message
GRVA: Working Party on Automated/Autonomous and Connected Vehicles (at UNECE)
IVIM: Infrastructure to Vehicle Information Messages
LOM: Loi d'Orientation des Mobilités (French Law: Mobility Orientation Law)
KPI: Key Performances Indicators
MAPEM: Map-data Messages
MRC: Minimal Risk Condition
MRM: Minimal Risk Maneuver
NHTSA: US National Highway Traffic Safety Administration
POC: Proof of Concept
ODD: Operational Design Domain
OEDR: Object and Event Detection and Response
POI: Point of interest
RSU: Road Side Unit
SPATEM: Signal Phase and Timing Messages
UNECE: United Nations Economic Commission for Europe
VRU: Vulnerable Road User

1. Introduction

Purpose of the L4.5 is to apply, develop and experiment the L4.4 procedure deployed with testing and driving methodology in real conditions. Objective is to evaluate behaviours of AI-powered systems through two POCs: one POC for mobility (with vehicle) & one POC for supervision.
The procedure has been applied on existing resources available thanks to partners of the consortium with: infrastructure, pathway and vehicle.
The experimentation was the opportunity to collect a maximum of data during the planning dedicated for the project.
The experimentation has used the scenarios and KPI defined in previous deliverables.
Here are the deployment of the procedure and results associated.

2. Context

The overarching goal of the PRISSMA project is to contribute to a methodology for evaluating the behavior of automated and autonomous vehicles within a territory, and to assess the supervision capability of infrastructure with event alerts from the environment (e.g., crowds that may disrupt traffic flow). The questions posed by the coexistence of these increasingly automated and connected vehicles with conventional vehicles and vulnerable road users such as motorized two-wheelers, bicycles, and pedestrians are fundamental for public policymakers, manufacturers, infrastructure specialists, and road safety experts.

The experimentation of the prototype L4 Valeo robot taxi in Paris was built upon the existing inter-station experimentation between Gare de Lyon, Gare de Bercy, and Gare d'Austerlitz. It constitutes a loop within the Paris metropolitan area and amid traffic. This route follows the experimentation of a regular automated shuttle line (Milla, RATP), with stops and frequencies, connected to the existing public transport network. It notably complements the service for "occasional" trips, operating during weekdays.

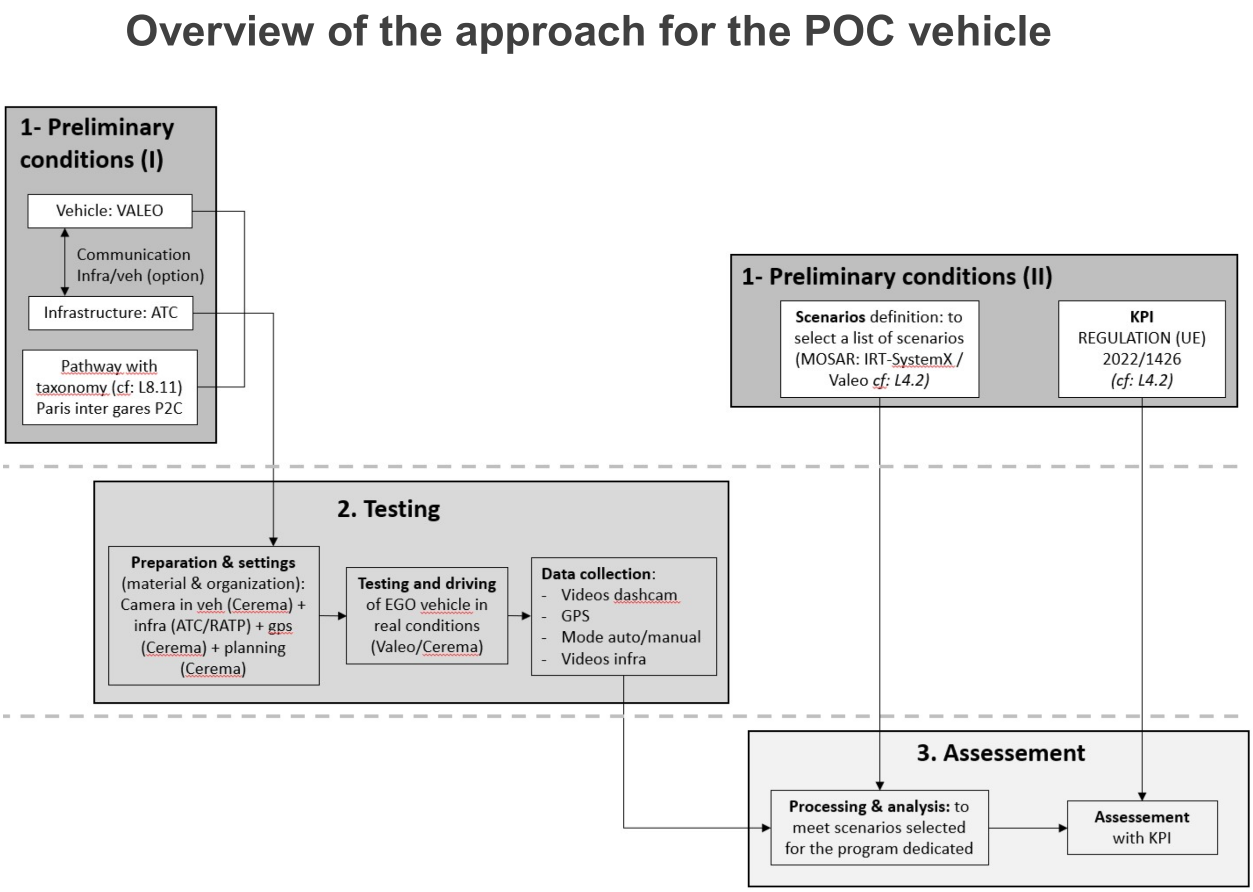
This route is equipped with infrastructure collecting data via smart poles. This infrastructure can identify traffic, crowd movements, and raise alerts to a control center. The utilization of these environmental events would eventually allow communication of alerts to the EGO.

In the implementation of PRISSMA POCs, with the maturity of current equipment, the EGO operated independently of infrastructure and supervision, unfortunately not benefiting from external infrastructure information regarding its augmented environment. The EGO thus functioned with its perception of the nearby environment based solely on the EGO's sensor vision.

However, infrastructure videos were able to be leveraged in scenario analysis as the EGO passed through, providing different perspectives and complementing the data collected by the vehicle.

3. Methodology

Before delving into the details of the methodology, here is a diagram of the overall approach summarizing the process: from prerequisites to evaluation.



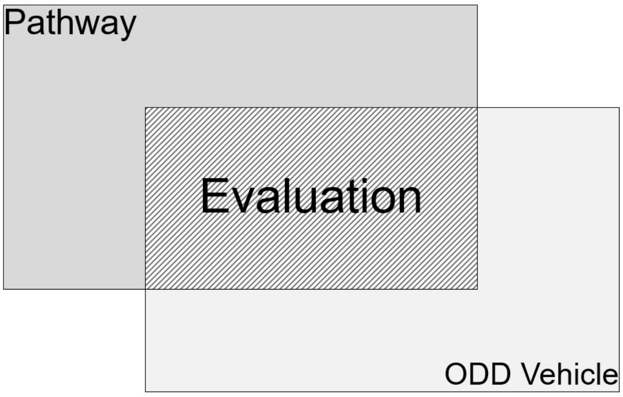
3.1. *Preliminary conditions*

As mentioned in L4.4, safety criteria must be the top priority in real-world tests, which is why safety is the primary prerequisite throughout WP4.

The evaluation focuses on the vehicle's ability to operate within the defined route and environment. In the PRISSMA experimentation, taking a bottom-up approach, the various elements of the experimentation—vehicle, route, and infrastructure—naturally emerged due to their maturity and the state-of-the-art knowledge at this stage. In this context, we know that the vehicle's Operational Design Domain (ODD) is partially capable of accommodating the proposed route.

It is anticipated that a portion of the route will lie outside the vehicle's ODD, and conversely, that some aspects of the vehicle's ODD may extend beyond the route domain, meaning its capabilities may exceed or be partially adapted to the defined route. As we will see later in the document, within the POC framework, for experimentation safety, a safety driver was present on board to take control in the event of the vehicle's ODD being exceeded.

Within the scope of PRISSMA, this initial experimentation focuses on a limited evaluation with 15 scenarios and a route of approximately 2km, yet allowing for the identification of the evaluation boundaries with a vehicle capable of autonomously and automatically navigating the prescribed route and a broad range of situations. We have succeeded in establishing an evaluation framework within the PRISSMA experimentation, albeit restricted in scope.

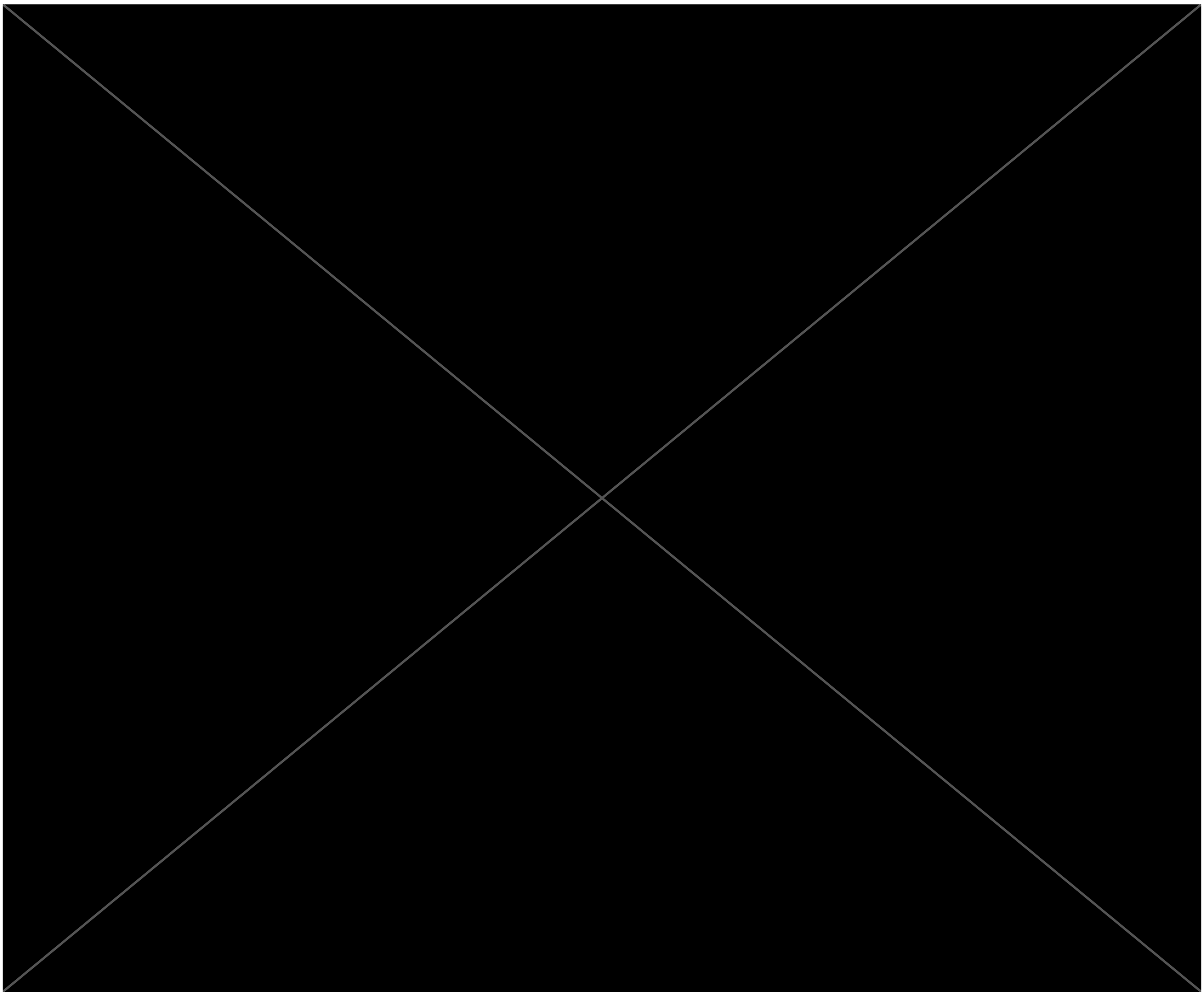


The prerequisites necessary for the methodology implementation (Test setup, data collection, analysis, etc.) are listed as questions classified in the table below. These prerequisites are identified by importance to facilitate the methodology setup: "Mandatory," "Recommended," or "Not important."

The list of questions is not exhaustive and should be completed and "calibrated" according to the intended purposes in the case of an evaluation for homologation.

Questionnary preliminary conditions	Mandatory	Recommended	Unimportante
Vehicle			
Is there a safety-driver? (cf. LOM regulation)	X		
Is it a level 4 designed vehicle?			
Is the timestamp (gps) calibrated with the infrastructure and other devices?		X	
Is veh. ODD adapted with pathway?	X		
Is AI used in the system? What softwares are used in the system?	X		
What are data sources? Videos, CAM, GPS...			
- Videos	X		
- CAM		X	
- GPS: csv			X
Infrastructure			
Is the timestamp (gps) calibrated with the vehicle and other devices?			
What softwares are used?			
What are data sources? Videos, Spatem/Mapem...			
- Videos		X	
- Spatem/Mapem		X	

Pathway			
Is the pathway detailed through a taxonomy?	X		
Does the pathway cover the entire expected assessment?	X		
Scenarios			
Are scenarios are described in a support?	X		
How many scenarios have been defined?	X		
Does the list of scenarios cover the entire expected assessment?	X		
KPI & metrics			
Are metrics from EU regulation 2022/1426?	X		
Does metrics cover the entire expected assessment?	X		
Other			
What are weather conditions?	X		



[redacted]

[redacted]

3.1.2. Infrastructure

For the POC vehicle, the infrastructure used is that of ATC described in deliverable L4.3 (§4; p.9-23) and operated by RATP.

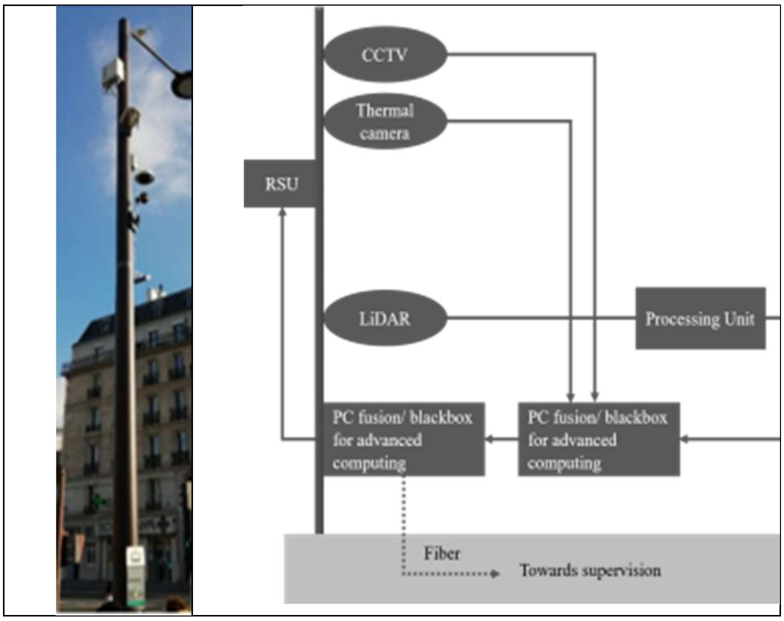


Figure 3: Type and layout of connected equipment in smart pole no.6 at "1 Quai de la Gare" (December 2020)

3.1.3. Pathway

For the POC vehicle, the pathway taken is the Paris2Connect inter-station route described in deliverable L4.3 (§4; p.8-9).

Paris2Connect (P2C) is a fully operational urban infrastructure on a 3.5-kilometre route in the 12th and 13th arrondissements of Paris, connecting the Austerlitz, Lyon and Bercy stations.

Pathway is characterized through the taxonomy defined in the deliverable 8.11.

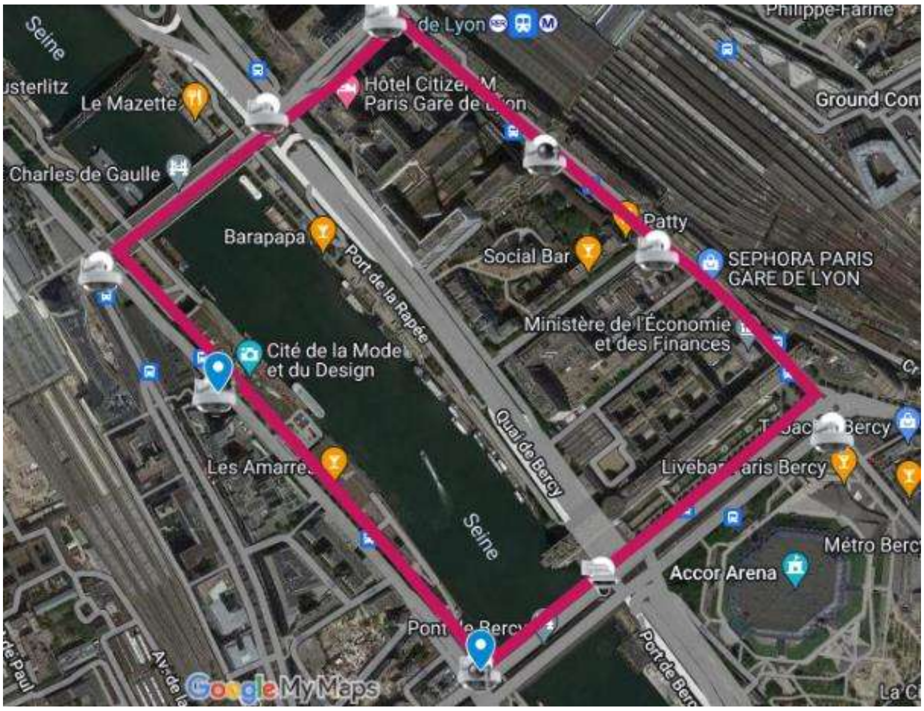


Figure 2: Paris2Connect pathway (other view in appendix n°3)

3.1.4. Scenarios

A- POC Vehicle (Valeo)

The list of scenarios is as described in deliverable L4.3 (§5.2; p.25-47). The selected scenarios have been pre-mapped to certain sections of the P2C pathway.

Scenarios

Functional		Logical		Sections
1	Crossing Intersection	1	Crossing while no other road users	18
1	Crossing Intersection	2	Crossing intersection while 2-wheelers does not yield	18
1	Crossing Intersection	3	Cyclist crosses EGO path while right turn	25-28
1	Crossing Intersection	4	Pedestrian crossing while green light for EGO	17 / 26 / 42 / 49
2	Cut-in	5	Bus Cut-in in front of ego	22-23-24
2	Cut-in	6	Bicycle cut-in in front of the ego	29-36
2	Cut-in	7	Motorcycle cut-in from the right	9-14
3	Drive-off	8	Drive-off at intersection when surrounded by VRU	17 / 26 / 42 / 49
4	Ego lane change	9	Ego lane change	9
5	Lead vehicle braking	10	Front vehicle harsh braking (due to an occluded bicycle crossing the road)	22-23-24
6	Occluded actors	11	Occluded pedestrian crossing out of crosswalk	1-17
6	Occluded actors	12	Partially occluded pedestrians	TbD
6	Occluded actors	13	Occluded cyclist violating the right of way while right turn	25-28
7	Pedestrians expected	14	No occluded pedestrians crossing in front of the ego	1-17
7	Pedestrians expected	15	Group of pedestrians close to the lane edge (on the right)	41 - 42

Table 10: List of scenarios selected

B- POC Supervision (RATP)

The list of scenarios is as described in deliverable L4.3 (§5.4; p.49-53).

Scenarios

Functional		Logical	
1	Gathering of people “Bercy esplanade”	1	Gathering of people
1	Gathering of people “Bercy esplanade”	2	No gathering of people
2	Congestion at “Austerlitz”	1	Congestion
2	Congestion at “Austerlitz”	2	No congestion
3	Bike at "CdG / Van Gogh" bridge	1	Crossing

Table 10: List of scenarios selected

3.1.5. KPI & metrics

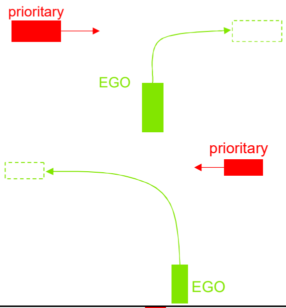
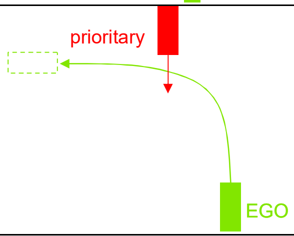
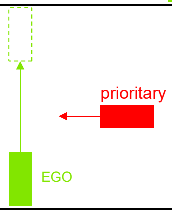
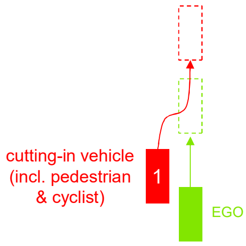
The KPIs and metrics are derived from REGULATION (EU) 2022/1426. Therefore, we consider both quantitative and qualitative metrics for evaluation.

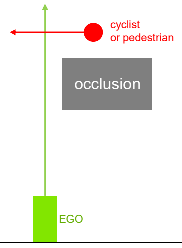
These metrics are selected as part of the experimentation process to build the methodology.

A- POC Véhicule (Valeo)

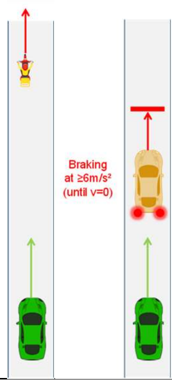
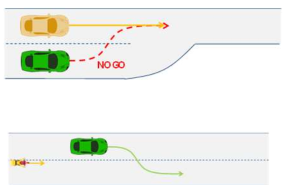
The list of metrics is as described in the deliverable L4.2 (§2.8; p.22-26) and as below:

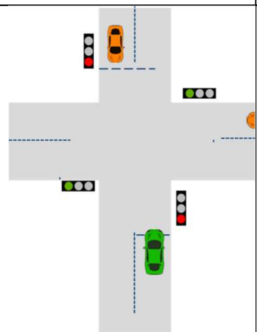
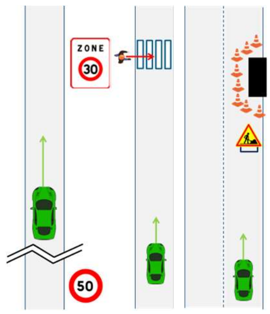
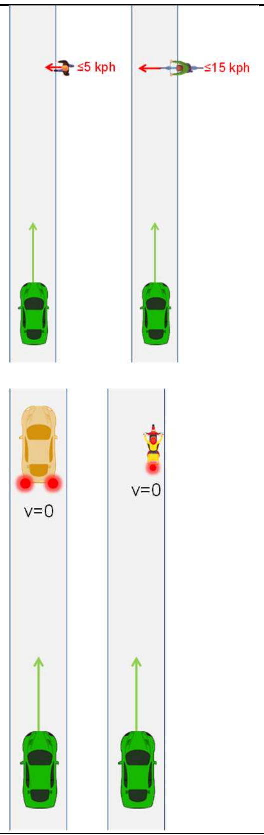
- **Quantitative metrics**

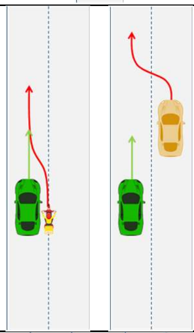
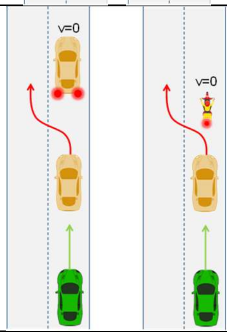
ID	Type of interaction/maneuver		Illustration	Pass criteria
QN_KPI_1	**Turning & Crossing at intersections**	Merging with privileged traffic during **turning (right or left)**		$$TTC \geq TTC_{dyn} = \frac{(v_e + v_a)}{2 \cdot \beta} + \rho$$ With: v_e : speed of the fully automated vehicle v_a : speed of the privileged approaching traffic β : maximum admissible deceleration for the privileged approaching traffic (3 m/s^2) ρ : reaction time of the privileged approaching (1.5s)
QN_KPI_2		Turning maneuver crossing the opposite traffic direction		$$TTC \geq TTC_{int} = \frac{v_c}{2 \cdot \beta} + \rho$$ With: v_c : speed of the privileged conflicting traffic β : maximum admissible deceleration for the privileged crossing traffic (3 m/s^2) ρ : reaction time of the privileged crossing traffic (1.5s)
QN_KPI_3		Crossing with privileged traffic		$$TTC \geq TTC_{int} = \frac{v_c}{2 \cdot \beta} + \rho$$ With: v_c : speed of the privileged conflicting traffic β : maximum admissible deceleration for the privileged crossing traffic (3 m/s^2) ρ : reaction time of the privileged crossing traffic (1.5s)
QN_KPI_4		Cutting in vehicles, pedestrians and cyclists travelling in the same direction		AV avoids collision with cutting-in road user at least when below conditions are fulfilled : $$TTC_{cut} \geq \frac{v_{rel}}{2 \cdot \beta} + \rho + \frac{1}{2}\tau$$ With: TTC_{cut-in} : time to-collision at the moment of the cut-in of the vehicle or cyclist by more than 30 cm in the lane of the fully automated vehicle v_{rel} : relative speed [m/s] between the fully automated vehicle and the cutting-in vehicle (positive if the ADS is faster than the cutting-in vehicle). β : maximum deceleration of the fully automated vehicle (assumed to be equal to 2.4 m/s^2 for fully automated vehicles transporting standing or not fastened vehicle occupants; 6 m/s^2 for other fully automated vehicles) ρ : time required by the fully automated vehicle to initiate an emergency braking and assumed to be equal to 0.1 s τ : time to reach the maximum deceleration β (assumed to be equal to 0.12 s for fully automated vehicles transporting standing or not fastened vehicle occupants; 0.3 s for other fully automated vehicles)

				The compliance with this equation is required only for road users cutting in, and only if the inserting road users were visible at least 0,72 seconds before cut-in
QN_KPI_5	Collision mitigation	Obstructed pedestrian/cyclist crossing in front of EGO vehicle		At impact, speed (v) reduced such as : $\lvert\Delta v \geq 20km/h\rvert$

- ## Qualitative metrics

ID	Type of interaction/maneuver			Illustration	Pass criteria
QL_KPI_1	Lane keeping	a) with a passenger car target as well as a power-two-wheeler (PTW) target as the other vehicle b) with a lead vehicle swerving in the lane c) with another vehicle driving close beside in the adjacent lane			AV does not leave its lane and maintains a stable motion inside its lane (across the speed range and different curvatures within its system boundaries)
QL_KPI_2	Follow a lead vehicle	a) across the entire speed range b) using a passenger car target a PTW target as well as a bicycle target as lead vehicle, provided standardized PTW targets suitable to safely perform the test are available c) for constant and varying lead vehicle velocities (realistic speed profile); d) for straight and curved sections of road; e) for different lateral positions of lead vehicle in the lane; f) with a deceleration of the lead vehicle of at least 6 m/s² mean fully developed deceleration until standstill			- AV maintains and restores a stable motion and a safety distance to a vehicle in front - AV avoids a collision with a lead vehicle which decelerates up to its maximum deceleration.
QL_KPI_3	Lane change	a) with the AV performing lane change to the adjacent (target) lane b) merging at lane end c) merging into an occupied lane	a) with different vehicles, including a power two-wheeler (PTW) approaching from the rear b) in a scenario where it is possible to execute a lane changing maneuver in regular operation c) in a scenario where a lane changing maneuver in regular operation is not possible due to a vehicle approaching from the rear d) with an equally fast vehicle following behind in the adjacent lane, preventing a lane change e) with a vehicle driving beside in the adjacent lane preventing a lane change f) in a scenario where a lane change maneuver during a minimal risk maneuver is possible and executed g) in a scenario where the AV reacts to another vehicle that		- AV does not cause an unreasonable risk to safety of the vehicle occupants and other road users during a lane change procedure - AV is able to assess the criticality of the situation before starting the lane change manoeuvre throughout the entire operational speed range

			starts changing into the same space within the target lane, to avoid a potential risk of collision		
QL_KPI_4	Intersections (T,X, multi-ways or roundabout)	a) with and without traffic lights b) with different rights of way	a) without a lead vehicle b) with a passenger car target as well as a PTW target as the lead vehicle/other vehicle c) with and without approaching or passing vehicles		AV detects and adapts to a variation of different road geometries which can occur within the intended ODD across its whole speed range
QL_KPI_5	National traffic rules & changes in road infrastructure	a) different speed limit signs, so that the AV has to change its speed according to the indicated values b) signal lights and/or stop instructed by a road safety officer/enforcement agents with situations of going straight, turning left and right c) pedestrian and cyclist crossings with and without pedestrians/cyclist approaching/on the road d) temporary modifications: e.g., road maintenance operations indicated by traffic signs, cones and other signalization, access restrictions	a) without a lead vehicle b) with a passenger car target as well as a PTW target as the lead vehicle/other vehicle		- AV complies with national traffic rules - AV adapts to a various permanent and temporary changes of the road infrastructure (e.g. road construction sites) in the entire speed range
QL_KPI_6	Collision avoidance	a) with a stationary passenger car target b) with a stationary PTW target c) with a stationary pedestrian target d) with a pedestrian target crossing the lane with a speed of 5 km/h, also in the presence of other objects relevant in the ODD (e.g. a ball, a shopping bag, etc.) e) with a pedestrian target moving a speed of up to 5 km/h within and partially occupying the lane of the AV and following the same or the opposite direction of the fully automated vehicle f) with a pedestrian target swerving in the same lane of the fully automated vehicle g) with a cyclist target crossing the lane with a speed of 15 km/h h) with a cyclist target which is travelling in the same direction with a speed of 15 km/h i) with the fully automated vehicle turning right and crossing the path of the cyclist travelling in the same direction with a speed of 15 km/h j) with a target representing a blocked lane k) with a target partially within the lane l) with one or more different types of unpassable objects relevant in the ODD (e.g., a dustbin, a fallen bicycle or scooter, a fallen traffic sign, a stationary or moving ball, etc.) m) with multiple consecutive obstacles blocking the lane relevant in the ODD (e.g., in the following order: ego- vehicle / motorcycle / car) n) on a curved section of road			AV avoids a collision with a stationary vehicle, road user or fully or partially blocked lane up to the maximum specified speed of the AV

QL_KPI_7	Response to Passable objects	"Passable object" in AV trajectory (passable object = such an object, that may be rolled over without causing an unreasonable risk to the vehicle occupants or other road users, e.g., a manhole lid or a small branch)	a) without a lead vehicle b) with a passenger car target as well as a PTW target as the lead vehicle/other vehicle		AV is not initiating an Emergency Braking with a deceleration demand greater than 5 m/s² due to a passable object in the lane relevant (for the ODD up to the maximum specified speed of the AV)
QL_KPI_8	Response to cut-in	a) with different TTC, distance and relative velocity values of the cut-in maneuver, covering types of cut-in scenarios in which a collision can be avoided and those in which a collision cannot be avoided b) with cutting-in vehicles travelling at constant longitudinal speed, accelerating and decelerating c) with different lateral velocities, lateral accelerations of the cut-in vehicle d) with a passenger car, PTW as well as bicycle targets as the cutting-in vehicle, provided standardized PTW targets suitable to safely perform the test are available			AV is capable of avoiding a collision with a vehicle or other road user cutting into its lane up to a certain criticality of the cut-in maneuver (criticality threshold defined in ***QN_KPI_4***)
QL_KPI_9	Response to cut-out	a) with a stationary passenger car target centered in lane b) with a PTW target centered in lane c) with a stationary pedestrian target centered in lane d) with a target representing a blocked lane centered in lane e) with multiple consecutive obstacles blocking the lane (e.g. in the following order: ego-vehicle – lane change vehicle – motorcycle – car).			AV is capable of avoiding a collision with a stationary vehicle, road user or blocked lane that becomes visible after a preceding vehicle avoided a collision by an evasive maneuver

In a further step, new metrics/KPI that are more specific to the AI performance assessment may be added (provided by Confiance.AI project, for example).

3.2. Test implementation

3.2.1. Preparatory phase & test plan

3.2.1.1. Onboard equipment (original equipment)

A- POC Véhicule (Valeo)

In our experimentation, we conducted tests using a 'black box' approach to vehicle composition. This means that the vehicle operates under the responsibility of the operator, and the physical and software architecture is neither known nor communicated. Indeed, within the framework of the PRISSMA project, the objective is not to evaluate the autonomous/automated vehicle of the participating partner in the project, but rather to construct the methodology.

With this in mind, it was not deemed necessary to have complete details of the onboard equipment. However, the necessary onboard data was collected and communicated for analysis (autonomous or manual mode of the vehicle with timestamps).

3.2.1.2. Additional equipment (dashcam, gps…)

A- POC Véhicule (Valeo)

As part of the PRISSMA experimentation, Cerema requested and implemented the installation of the Valeo POC in the test vehicle, comprising:

- Front and rear cameras:
  - Blackvue DR750S-2CH

- GPS
  - Mouchard GPS 770

3.2.2. Real conditions driving phase

A- POC Véhicule (Valeo)

The VA operated on the existing Paris2Connect route described in deliverable L4.3 (§4; p.8-9) and reiterated in this deliverable (see: annex n°3). The experimentation thus leveraged existing elements (vehicle and route) to drive in contexts closely resembling what would be necessary for a comprehensive evaluation. This allowed for driving in real-world conditions with a vehicle ODD adequately and sufficiently adapted to the proposed route, within an urban environment as demanding, complex, and dense as that of Paris.

In real conditions, the primary criterion was to ensure the safety of the external environment surrounding the vehicle. Additionally, a safety driver could take over if needed.

3.2.3. Data: compilation and exploitation

A- POC Véhicule (Valeo)

The combination of vehicle and infrastructure equipment allowed for sufficient data collection in constructing the methodology for evaluating the vehicle's behavior in its environment.

This quantity of data formed a database for compiling, processing, and analyzing the subsequent results by Cerema.

It's important to note that compiling data requires different data collections between vehicle videos, infrastructure data, and vehicle driving mode data (automatic/manual). Therefore, the final analysis is only possible if and only if the data were collected simultaneously to be able to perform necessary processing correlations. Thus, there were instances of collecting videos without being able to utilize them because we lacked the vehicle mode. The table below summarizes the available data from the experimentation.

Data availability matrix

	10 days testing									
	October		November							December
Data	16	17	7	8	9	13	15	16	23	18
Dashcam	X	X	X	X	X	X	X	X	X	X
Infrastructure	X	X			X	X	X	X	X	
Vehicle mode		X	X	X	X	X	X	X	X	
Operational days for analysis		X			X	X	X	X	X	

6 days data usable

Table 10: Identification of data available for analysis by source

The data collection allowed for the establishment of a selection of events described in the form of sheets (see: Annex n°5). From Cerema's perspective, these events were sorted and cross-referenced to identify situations corresponding to scenarios defined in PRISSMA (detailed in the Mosar tool and deliverable L4.2). Based on these situations and scenarios, we can proceed, in the remainder of the document, with a detailed evaluation of the events recorded during the test drives.

Below is a graph showing the quantity of sheets created from the volume of collected data.

Table 11: Volume of data available for analysis by source

These 35 sheets are the selected ones for the relevance of their event in meeting the sought-after scenarios for evaluation.

Other event sheets could have been selected, but the constraint remains definitively the analysis time.

3.3. Assessment implementation

3.3.1. Vocabulary

For the remainder of the document, it is important to define the vocabulary used.

- **Stop Time:** The duration during which the vehicle is stationary (speed equal to 0 km/h) and no interaction occurs in autonomous mode. For example, during stops at bus stations, traffic lights, pedestrian crossings, and yielding to the right. This time is excluded from the temporal analysis of driving times.
- **Driving Time:** The duration of automated vehicle circulation in the observed area excluding stop times.
- **Incident Concepts**

  There are 2 types of incidents:
  - **"Stop":** The automated vehicle stops in the middle of the road (excluding stop times mentioned above);
  - **"Slowdown":** The automated vehicle significantly slows down in the middle of the road (speed less than or equal to 5 km/h in urban areas), with its cruising speed averaging 30 km/h.

  These incidents may have several possible causes, which are analyzed later in this document:
  - Vulnerable User Cause: All vulnerable road users, such as pedestrians, cyclists, electric scooter riders, etc.;
  - Motorized User Cause: All motorized road users, such as motorcycles, passenger cars, trucks, etc.;
  - User Cause: A collective term comprising both vulnerable and motorized road users;
  - Without Apparent Reason: No evidence around the vehicle identifies the cause;
  - Other: Causes such as a bird, a branch on the road, an unintentional modification of the infrastructure, etc….

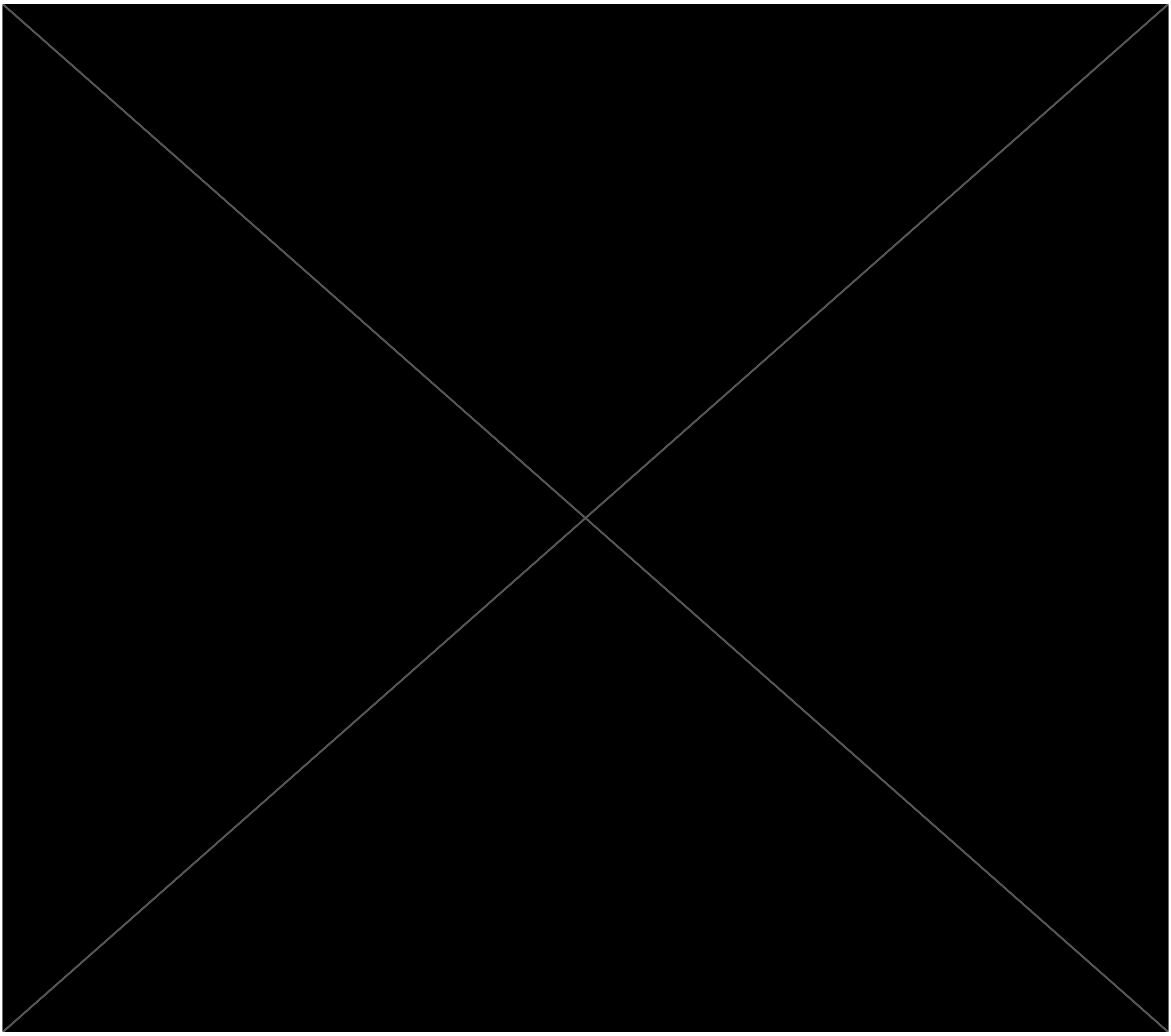
3.3.2. Review of video sequences from dashcams

The analysis was conducted by Cerema.

The video sequences from the dashcams were initially sorted by the vehicle's driving periods and then manually analyzed by an operator. On this occasion, the operator completed two tables. The first one concerns the situations encountered in the vicinity of the AV as well as the general context of the observations. The second one details the observed AV incidents and their respective causes.

The next phase involves checking the coherence of the encountered situations with the 15 selected scenarios.

Finally, with the scenarios identified in the situations, it is possible to proceed to the evaluation phase according to the criteria identified in REGULATION (EU) 2022/1426.



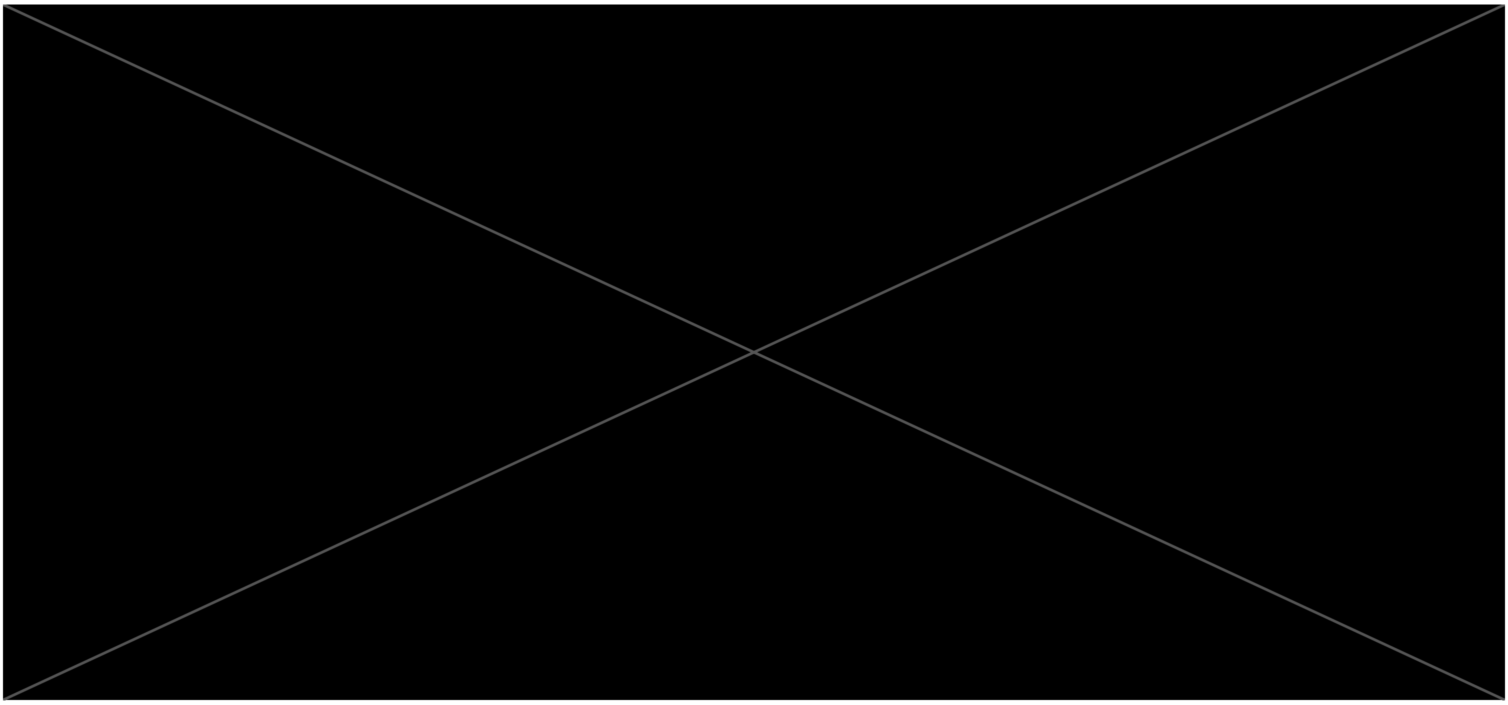
- **<u>First sorting: selection of driving phases within the domain</u>**

The initial sorting aims to reduce the amount of data to be analyzed and focus on specific data that may reveal situations to be evaluated.

- **<u>Second sorting: recording and searching for events during driving</u>**

A sorting table is completed regarding the events.
The following are recorded in the sorting table:

- Stops, driving modes, and events around the AV.
- Contextual information: Weather (sunny, cloudy, or rainy), maximum AV speed on the route, and passage times.

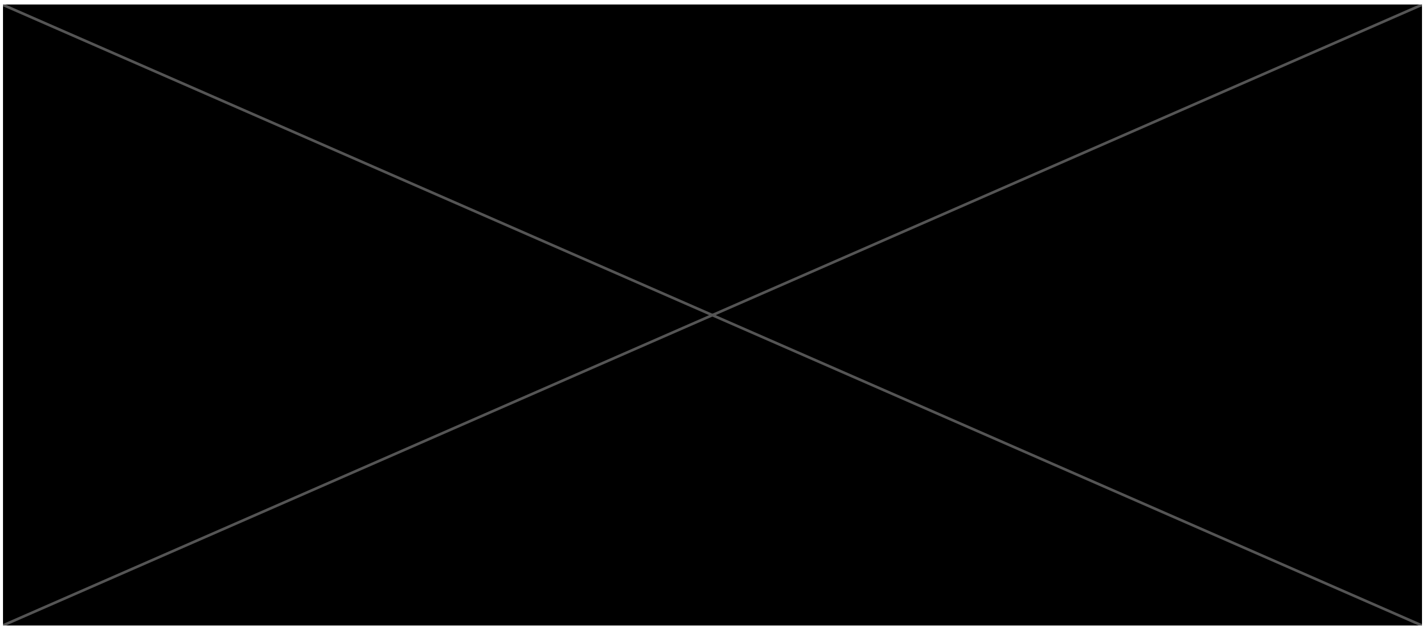


➢ **Compilation of observed scenarios and evaluation**

Finally, the next phase involves compiling the observed scenarios to proceed with the evaluation of the vehicle's behavior in the environment.

3.3.3. Review of infrastructure video sequences

In the context of the PRISSMA experimentation, the video sequences are obtained from infrastructure shots by ATC and processed by RATP.



B- POC Supervision (RATP)

The aim of the supervision POC is to anticipate events on the route of the automated/autonomous vehicle and to enhance the quality of work for supervisors, aiming to reduce the workload of continuous video monitoring. In this regard, the objective is to have relevant algorithms, meaning those with a high detection rate and a low percentage of false alarms.

For the supervision POC, analyses are focused on events. The first step is therefore to extract video sequences corresponding to events within an interval [T0-40s, T0+20s].

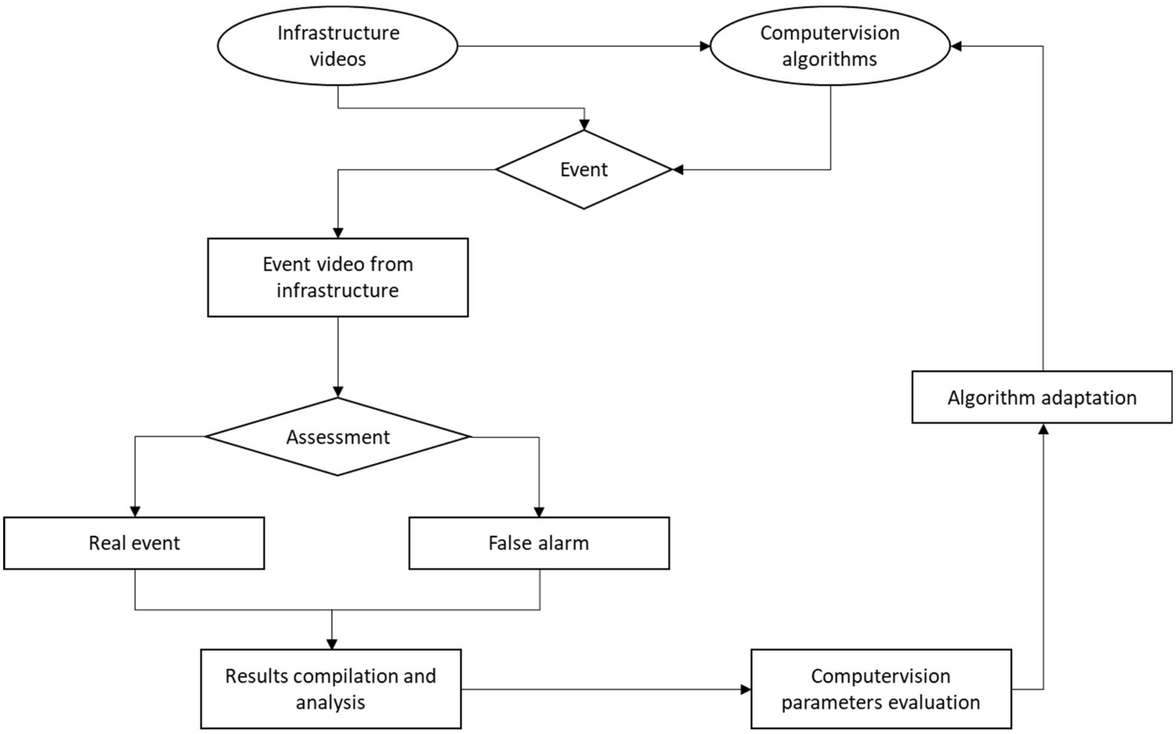


Figure 1: Videos analysis approach & assessment (POC supervision RATP)

In order to simplify the analysis of video sequences, a dedicated interface has been provided to the agents, enabling them to:

- View the video
- Access metadata (location, time, alarm type, confidence level, etc.)
- Determine whether it is a true or false alarm

The evaluations are then compiled into indicators (by date, alarm type, and connected pole) of the percentage of accurate alarms.

This indicator helps to target improvement actions and assess their relevance (the evolution of the KPI following an adaptation).

Consequently, our study has focused on ensuring that a detected event is indeed confirmed, either because a video or photo capture was available, or because a video verification was possible. In cases where it was not possible to determine the confirmed nature of a detection, either because video evidence was not available or when a known detection error occurred, we did not make a judgment on the nature of the event. Here is an example for the latter point:

- The obstacle detection by Computer Vision, in the method we used, was faulty in rainy weather when the wet ground reflected light, such as a lamppost. For pole 1, at the corner of Quai d'Austerlitz and Pont Charles de Gaulle, systematic obstacle detection occurred at the same location on rainy days.
- The detection of a special vehicle was consistently inaccurate for a specific vehicle; for example, a construction vehicle parked consistently within the camera view of smartpole 11.

Furthermore, when a camera made a false detection, such as misidentifying a special vehicle, we did not count successive false detections of the same vehicle if it remained stationary or within the camera's view. However, if it returned minutes or hours later, we counted the new false detection. Similarly, when a correct detection was repeated for hours (e.g., stationary vehicle), we also did not count the repetitions as they inflated the numbers and skewed the percentages.

3.3.4. Video data analysis

Before selecting events for analysis, we first ensured that the vehicle was in autonomous/automated mode.

To do this, we used the vehicle's onboard data. We transcribed the Auto/Manual modes based on the date and time.

Here is an excerpt from the file that allowed the identification of the driving mode.

MODE :		17/10/23	17/10/23		07/11/23		08/11/23	08/11/23	08/11/23
		10:29:36	*10:51:00*		*11:10:39*		10:46:39	10:57:05	12:00:38
White = OFF		*10:29:37*	*10:51:01*		*11:10:40*		10:46:40	10:57:06	12:00:39
		10:29:38	*10:51:02*		*11:10:41*		10:46:41	10:57:07	12:00:40
Grey = ON		*10:29:39*	*10:51:03*		*11:10:42*		10:46:42	10:57:08	12:00:41
		10:29:40	*10:51:04*		*11:10:43*		10:46:43	10:57:09	12:00:42
		10:29:41	*10:51:05*		*11:10:44*		10:46:44	10:57:10	12:00:43
		10:29:42	*10:51:06*		*11:10:45*		10:46:45	10:57:11	12:00:44
		10:29:43	*10:51:07*		*11:10:46*		10:46:46	10:57:12	12:00:45
		10:29:44	*10:51:08*		*11:10:47*		10:46:47	10:57:13	12:00:46
		10:29:45	*10:51:09*		*11:10:48*		10:46:48	10:57:14	12:00:47
		10:29:46	*10:51:10*		*11:10:49*		10:46:49	10:57:15	12:00:48
		10:29:47	*10:51:11*		*11:10:50*		10:46:50	10:57:16	12:00:49
		10:29:48	*10:51:12*		*11:10:51*		10:46:51	10:57:17	12:00:50
		10:29:49	*10:51:13*		*11:10:52*		10:46:52	10:57:18	12:00:51
		10:29:50	*10:51:14*		*11:10:53*		10:46:53	10:57:19	12:00:52
		10:29:51	*10:51:15*		*11:10:54*		10:46:54	10:57:20	12:00:53
		10:29:52	*10:51:16*		*11:10:55*		10:46:55	10:57:21	12:00:54
		10:29:53	*10:51:17*		*11:10:56*		10:46:56	10:57:22	12:00:55
		10:29:54	*10:51:18*		*11:10:57*		10:46:57	10:57:23	12:00:56
		10:29:55	*10:51:19*		*11:10:58*		10:46:58	10:57:24	12:00:57

Figure 2 – Extract from the driving log table following the auto mode (shaded background) / manual mode (white background)

We were able to deduce the very satisfactory quantity of periods in automatic/automated mode, accounting for 79% of the driving time, which allowed us to select scenarios.

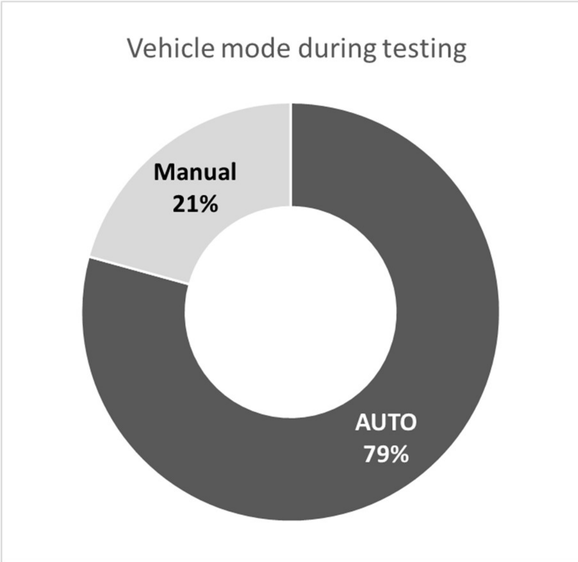


Figure 3 – Chart showing distribution by driving mode Auto/Manual

		Total
Vehicle testing	100%	**08:47:56**
Mode Auto	79%	**06:58:31**
Mode Manual	21%	**01:49:25**

Figure 4 – Table of distribution by driving mode Auto/Manual (in hours and %)

From the periods of driving mode, we were able to identify events in automated mode only.

Then, we assigned these 'validated' events to one of the 15 logical scenarios identified, when these events could correspond to them.

Thus, we established a matrix containing events that met the identified scenarios. We obtained a distribution by scenario, and thus, we were able to evaluate them.

Finally, for each event, we re-examined them step by step to evaluate them KPI by KPI with the metrics.

In the end, we obtained the compilation matrix below.

Compilation matrix of evaluations

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

B- POC Supervision (RATP)

Results for Computer Vision (CV) Figures, period 14/11/2023 to 4/12/2023:

- Special vehicles detection CV => 96% accurate detections before reaching 100% (between 96% and 99%, values between 14/11 and 4/12/2023)
  [Total 25, OK 24, KO 1]

- Crowd detection CV => 96.55% accurate detections
  [Total 29, OK 28, KO 1]

- Traffic jam detection CV => 98% accurate detections
  [Total 151, OK 148, KO 3]

- Obstacle detection CV => 65.6% accurate detections
  [Total 29, OK 19, KO 10]

Conclusion

In this deliverable L4.5, we applied the methodology described in L4.4 under real conditions. Thanks to the existence and provision of a route, infrastructure, and autonomous/automated vehicle by partners, the methodology could be fully implemented and followed through all planned and scheduled phases.

The limitations of the experimentation are the dedicated time and resources. In the case of PRISSMA, these time and resources already provide perspectives. In terms of time, we managed to deploy, plan, and implement available resources within a fairly constrained timeframe.

This resulted in conducting an experiment with a "step-by-step" approach to follow a coherent thread enabling result extraction.

[redacted]

[redacted]

Regarding the supervision POC, two levels of analysis overlap: the relevance of the sought function and the relevance of the algorithm to perform this function. The evaluation focused solely on the latter. Indeed, the algorithm itself is only one link in a complex chain composed of several subsystems, and being able to "ensure" that an algorithm performs a function reduces uncertainty and enables distinguishing a functional chain less dependent on the technique. Evaluating the algorithm is a tedious manual task, only feasible in situations of low occasional activity.

One step of the experimentation (in the ambition of PRISSMA) which is the communication of alerts (detected by supervision) to the AV, unfortunately could not be realized.

Regarding analysis, do the results suffice to define the behaviour of an AV?

What minimum quantity of data would be necessary?

At what level of success can the acceptability of an AV be defined? Who determines this level?

What probability of "good conduct" and what probability of a "black swan" event, even if statistically rare?

Should another methodology be deployed?

Recommendations

The approach pursued within PRISSMA has enabled the establishment of directions for evaluating AI-based autonomous mobility.

However, this approach provides the foundation for evaluation. The density and quantity of data required for a large-scale evaluation are much greater than the data collected. Therefore, following the construction of the various deliverables of WP4, we propose recommendations for a deeper evaluation.

We have observed that it is essential to define the necessary instrumentation for data collection as early as possible, as evaluation depends on it. The architecture of the entire functional evaluation must be developed very early in the process, alongside devising the solution, to add dedicated sensors for measurement if necessary, in addition to the sensors required for the system's proper functioning. For example, depending on the quantitative or qualitative evaluation criteria, it may be necessary to use more specific instrumentation to obtain a finer ground truth (especially for quantifying environmental objects: speed, direction).

Additionally, in constructing the scenarios to be implemented, it would be necessary to establish a hierarchy of scenarios, for example: by family, by complexity, or otherwise.

In the context of safety demonstration for the commissioning of an Automated Driving System (ADS), the real-world testing methodology proposed in this deliverable will need to be adapted, particularly to address the system-of-systems aspect as well as operational and maintenance constraints.

Another way in the evaluation would be the assessment of the AV by the human factor in parallel with video analysis. The evaluation would be purely qualitative and subjective, validated by human knowledge and experience. Thus, the experimentation could involve an evaluator passenger (for example, a driving license inspector).

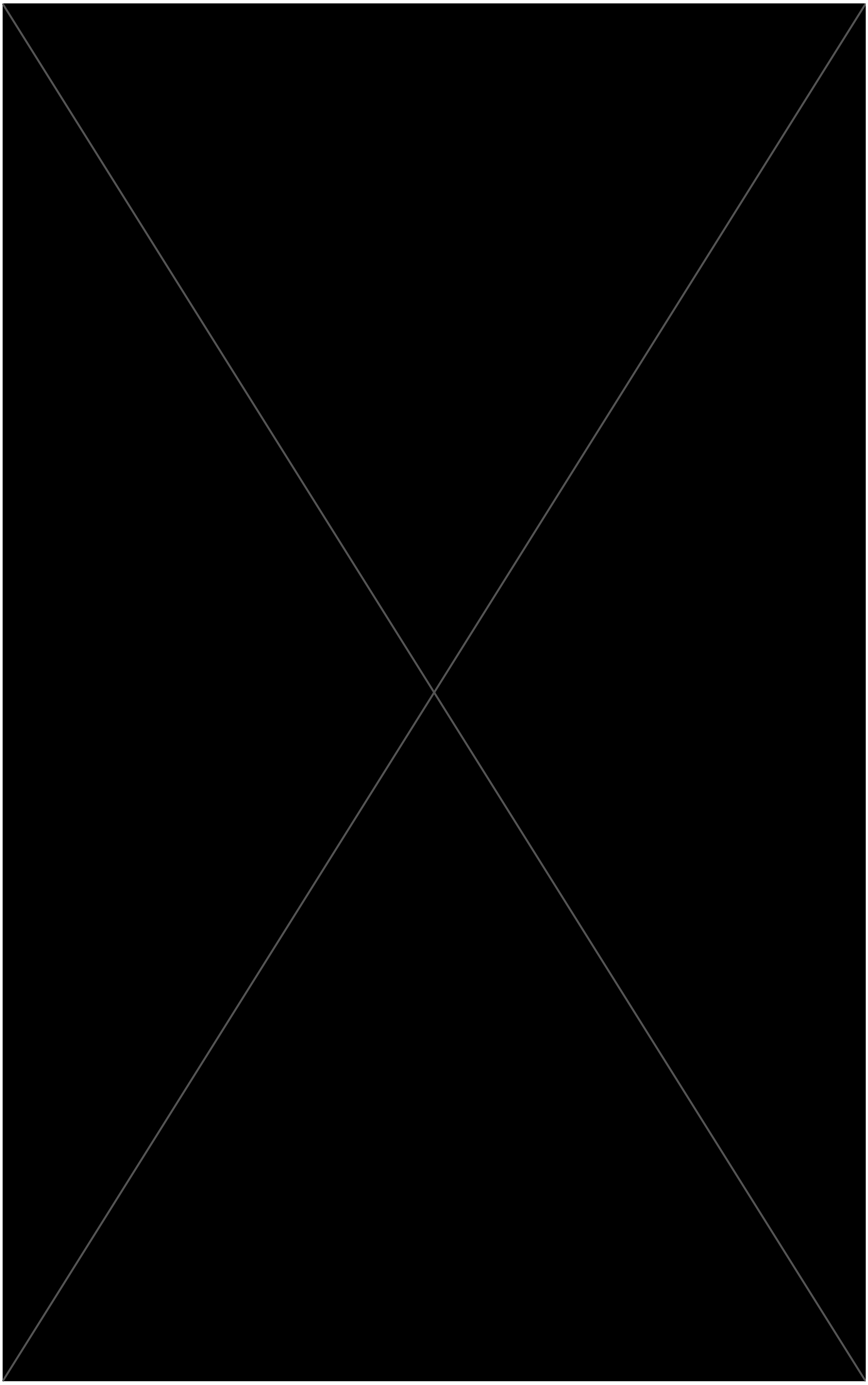
Furthermore, the experimentation has shown that evaluating a VA's overall system proves to be particularly complex. Therefore, the first step would be to assess the subsystems in order to separately evaluate the vehicle functions (braking, ADAS subsystems, ABS, AFU, ESP, LCA, ISA, DAW, DTS, NVS, ACC...). In this case, we would start with a functional evaluation without being able to assume the software architecture. But in this case, would it be possible to evaluate the “fine” and accurate perception of an autonomous/automated (AI) vehicle in a global approach?

Given this level of complexity and the amount of data required, automated or semi-automated means (development of learning blocks) must be found for a comprehensive application of this methodology to achieve qualification or homologation.

REFERENCES

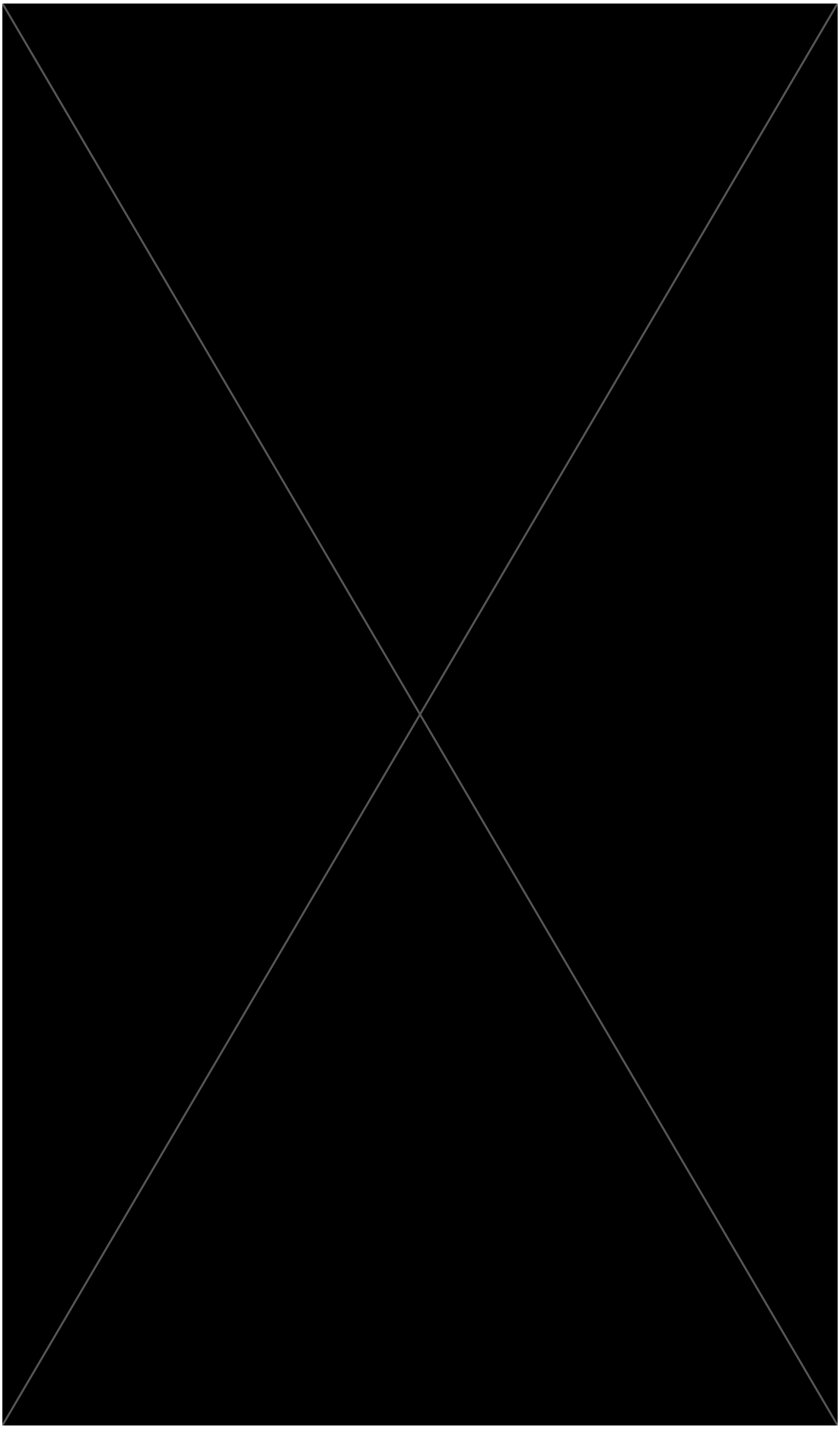
[1] Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles - Official Journal of the European Union - 26.8.2022

APPENDIX 1 – WP4 OVERVIEW

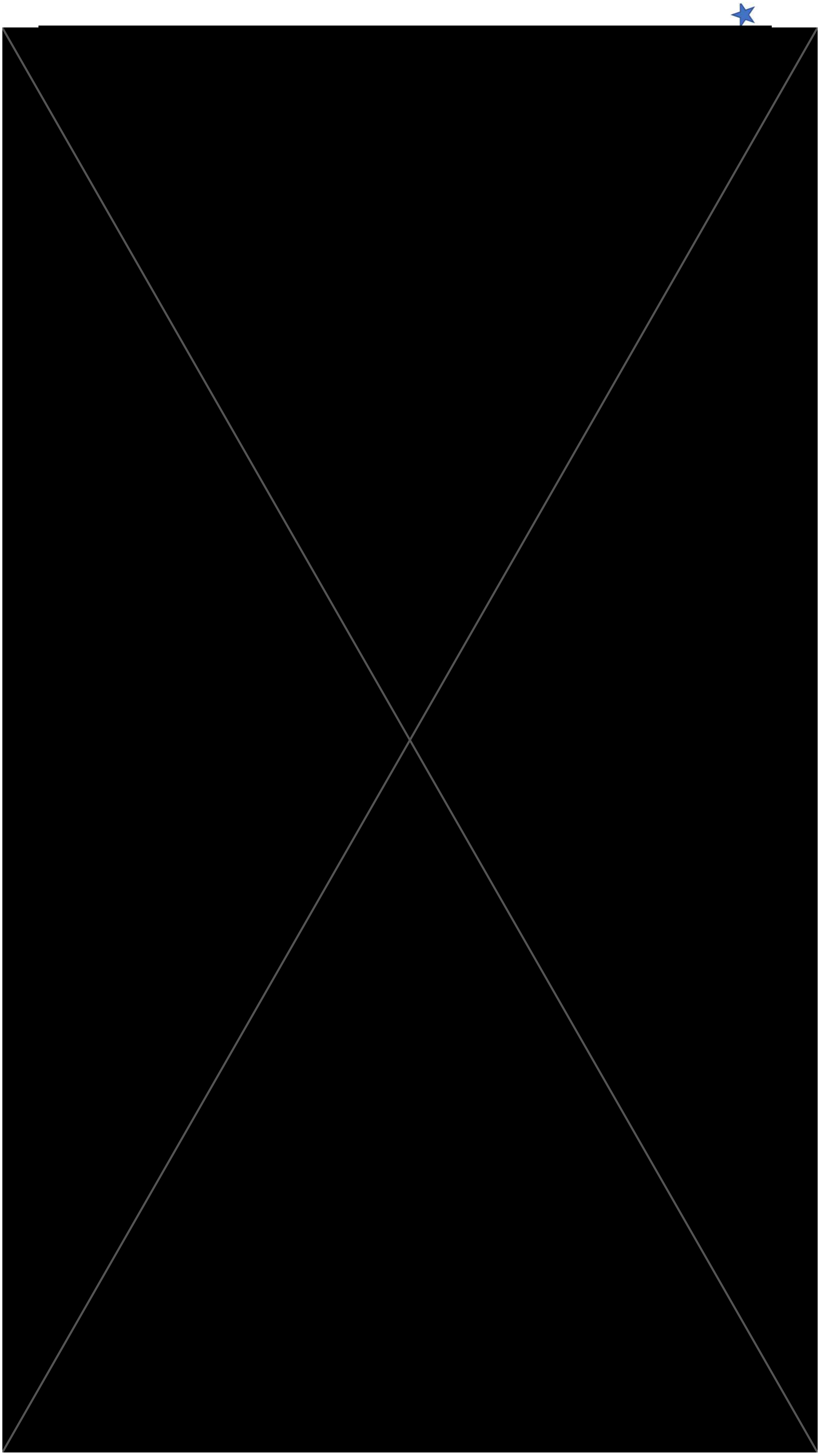


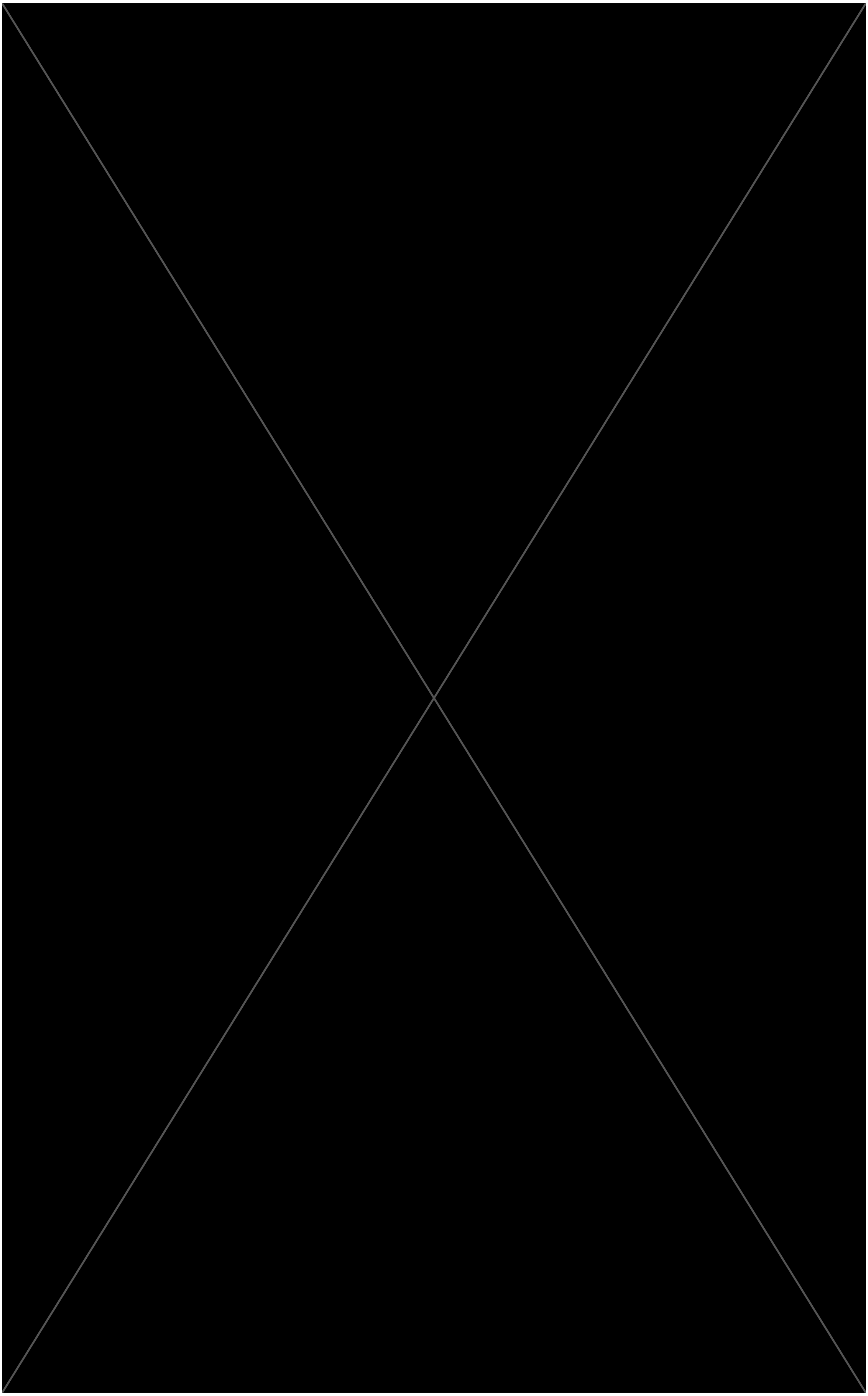
APPENDIX 2 – PLANNING OF EXPERIMENTATION

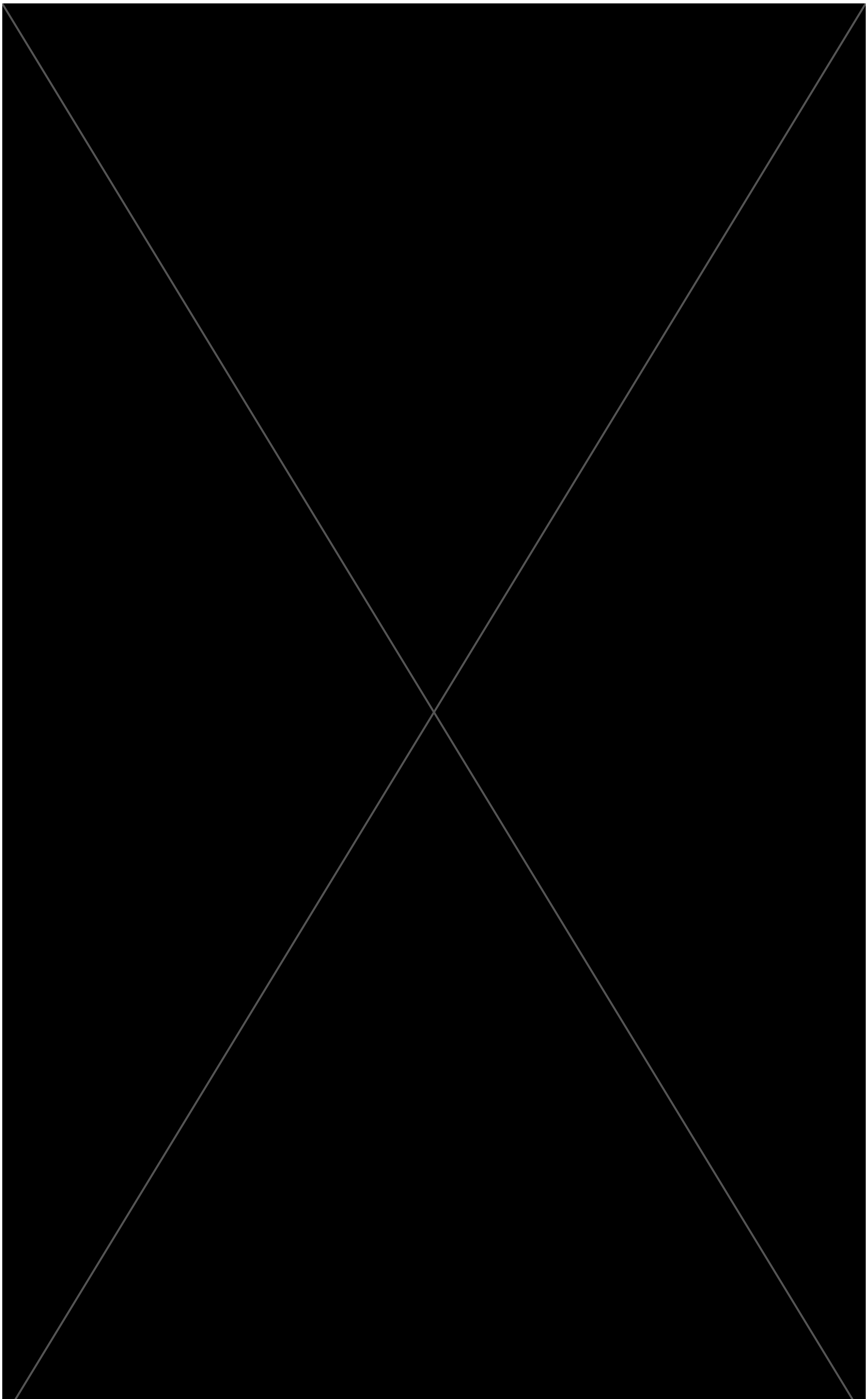
Planning - Part 1/2 (2023)

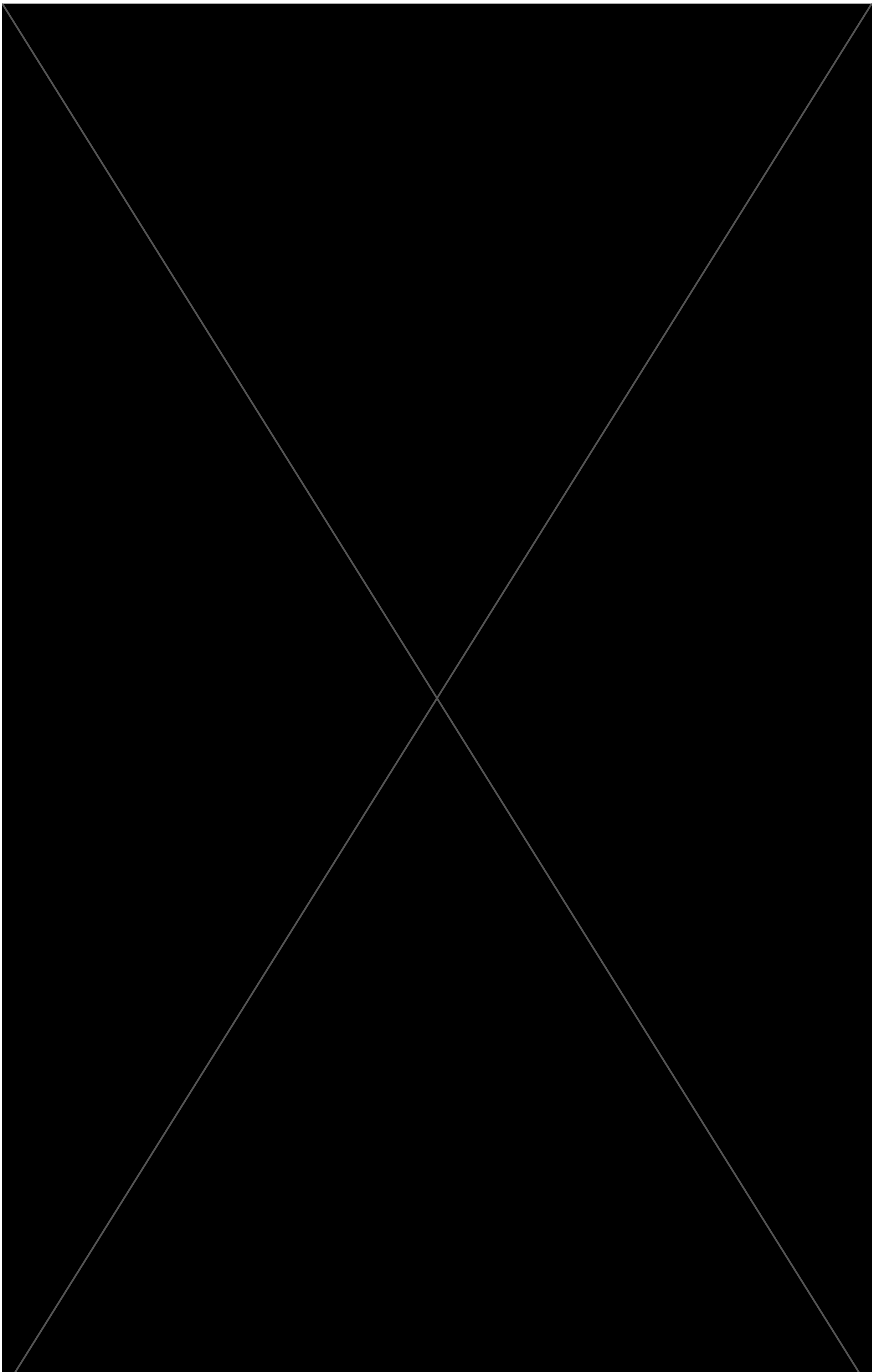


Planning - Part 2/2 (2024)

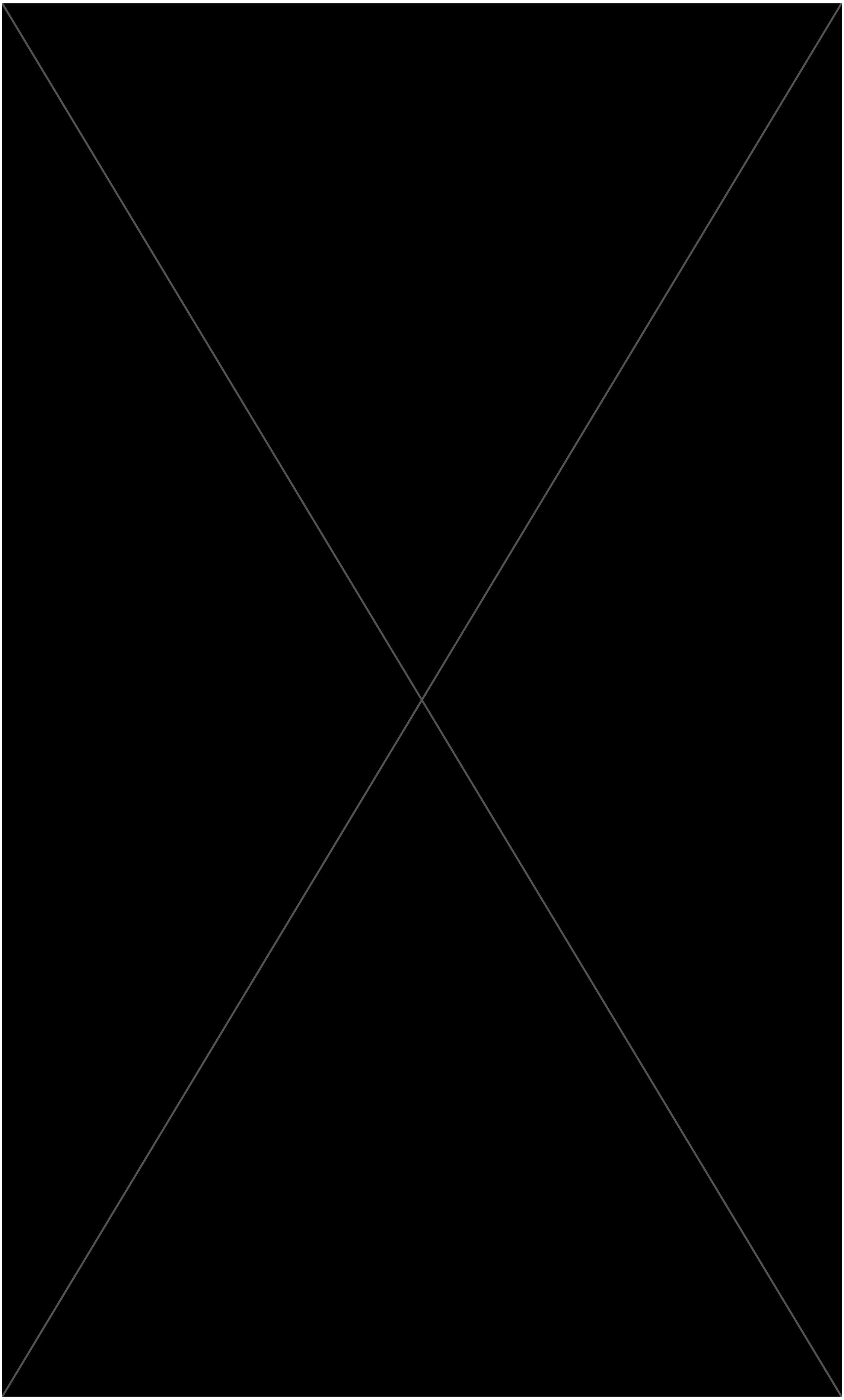


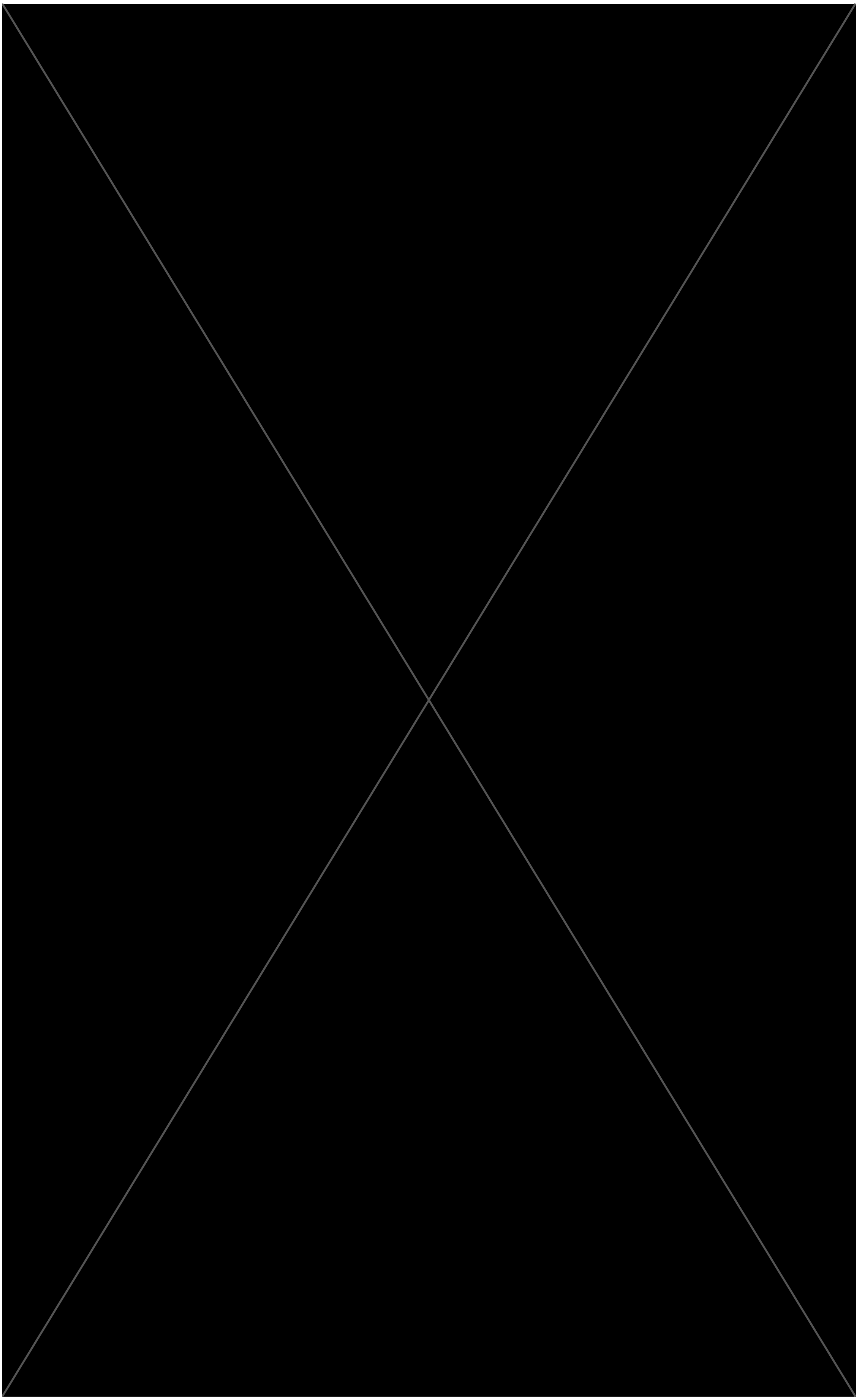




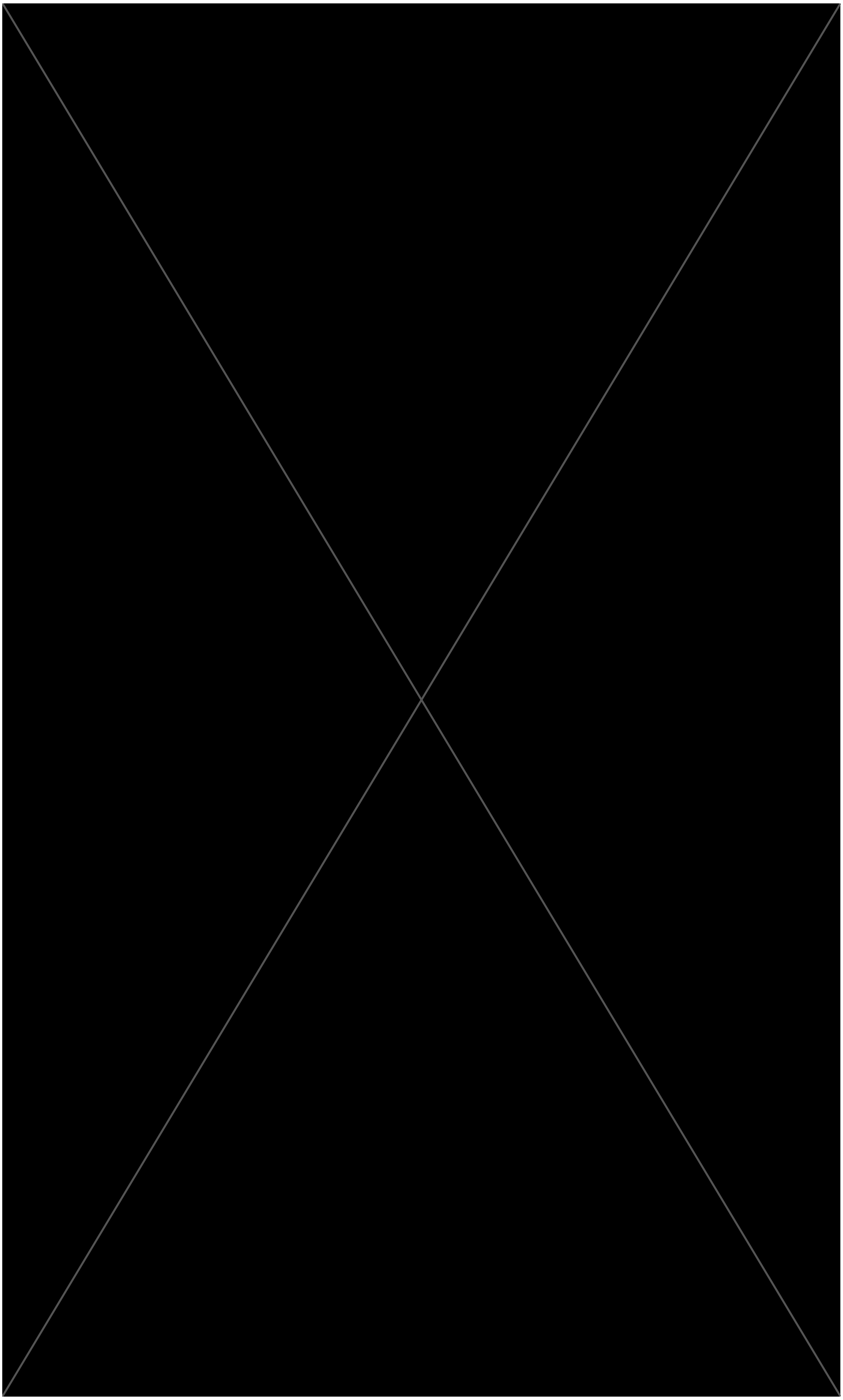


[redacted]

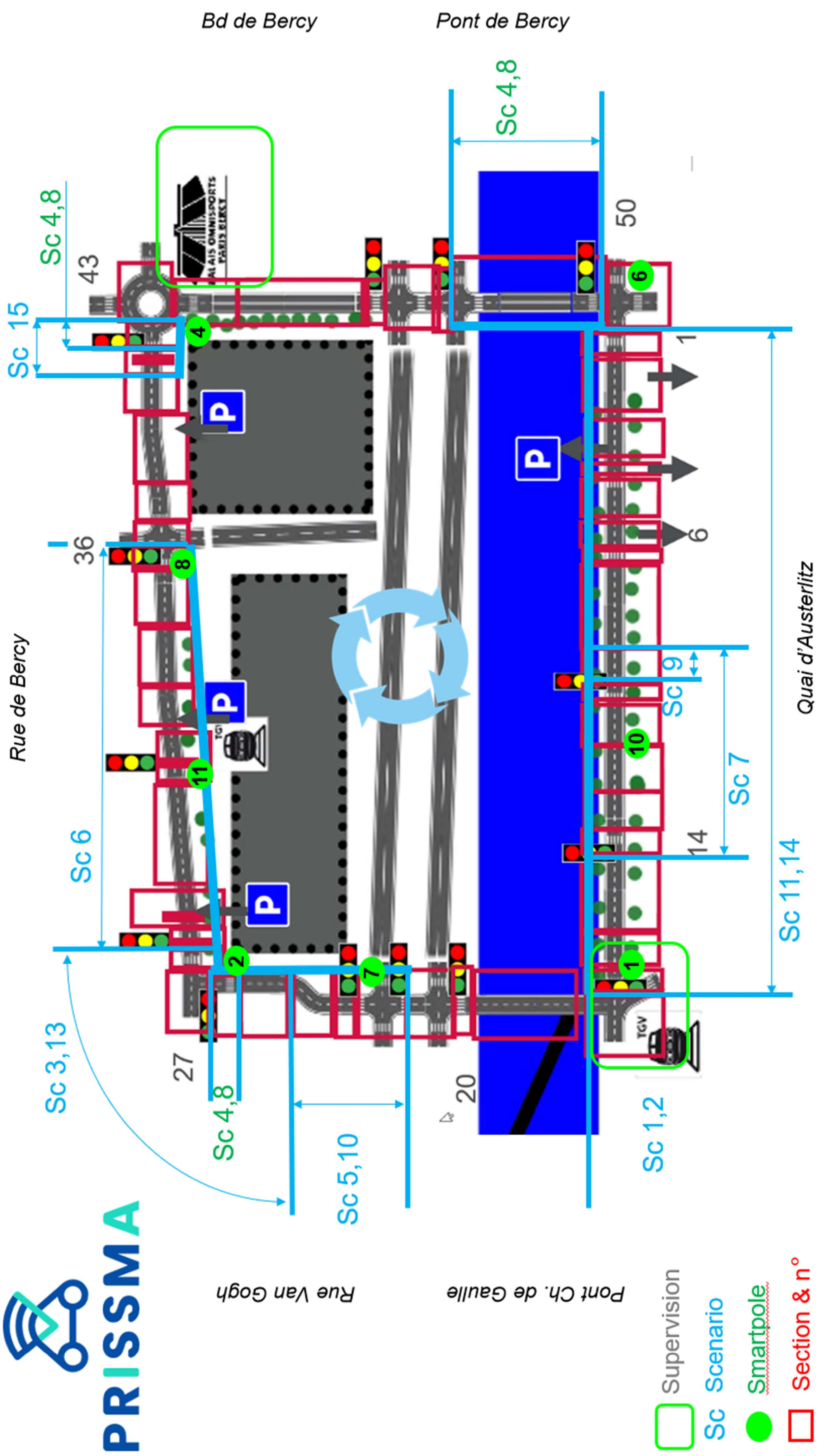


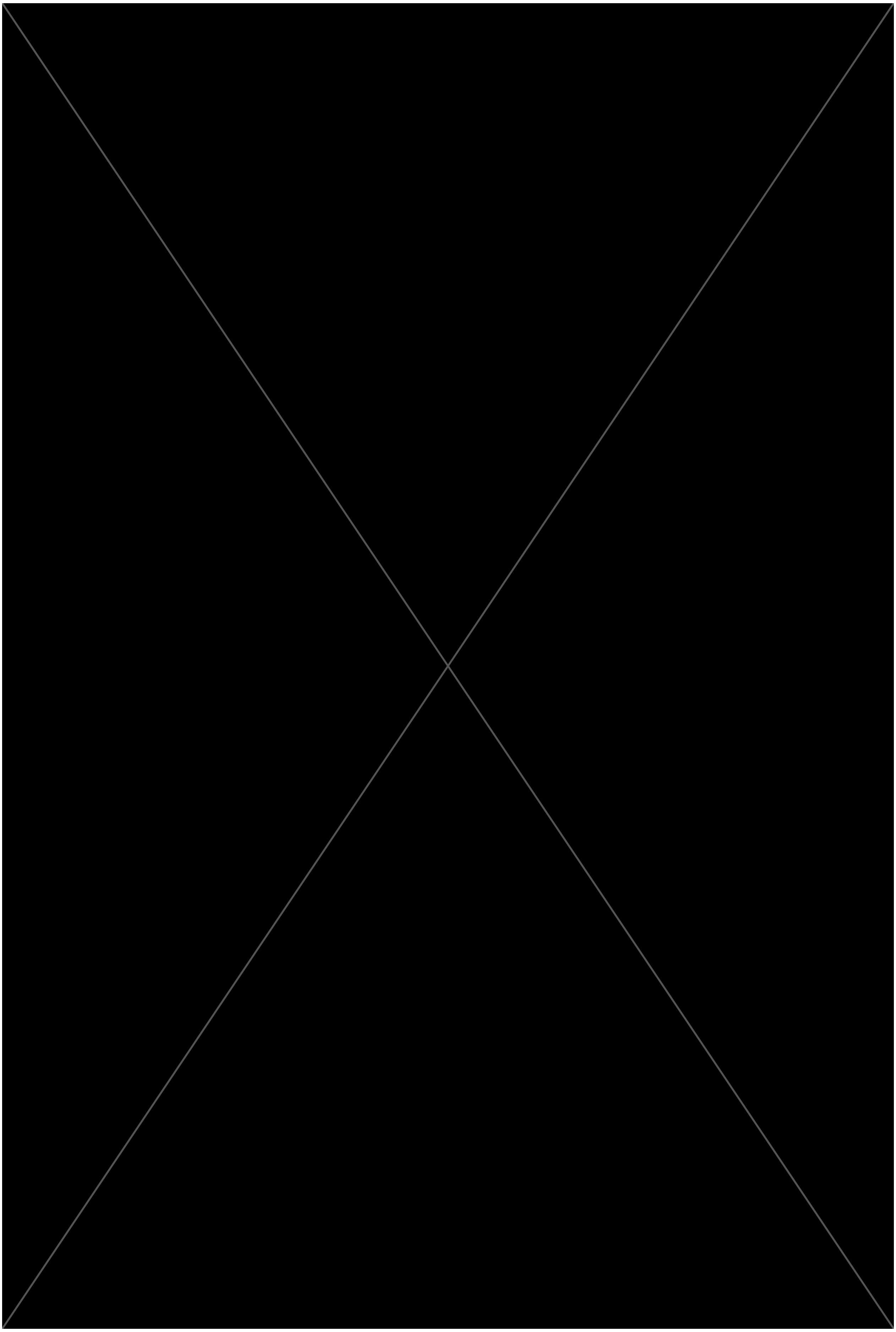


Page 7/7



APPENDIX 4 – PATHWAY





APPENDIX 6 – CASE OF "DRY RUN"

PRISSMA/CEREMA */DterIdF/DM/ITS*

Parcours sans intervention manuelle

EGO réalise un tour de parcours en mode ON
Départ & arrivée: section 27 (au feu à l'angle de la rue Van Gogh et de la rue de Bercy)

Références vidéo:
20231108_112729
20231108_113029
20231108_113329
20231108_113629

Horodatage:
Le 08/11/2023 de 11h30:04 à 11h38:18 (heure Dashcam)
Mode: On-Autonome le 08/11/2023 de 11:29:55 à 11:38:38 (données informatiques)

Météo:
Le 08 novembre 2023 à Paris, la t° maximale était de 13°C et la t° minimale de 8°C. Il y a eu un peu de précipitations, avec en tout 3mm sur la journée. Ce jour-là a été assez nuageux dans son ensemble. L'humidité dans l'air était très importante ce jour-là (en moyenne à 85%).

www.historique-meteo.net/france/ile-de-france/paris/2023/11/

08/11/2023 Températures : **8°C/13°C**
Précipitations : **2.5mm**

APPENDIX 7 – EXAMPLES OF SHEETS OF VIDEO ANALYSIS

PRISSMA /CEREMA /DterIdF/DM/ITS/ILM

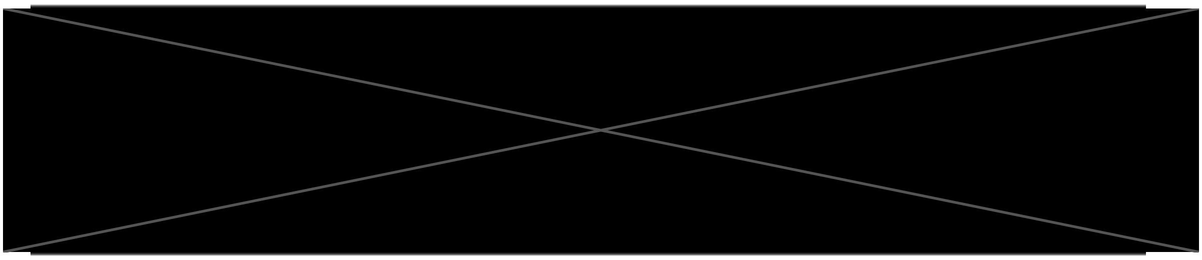
SITUATIONS & SCENARIO

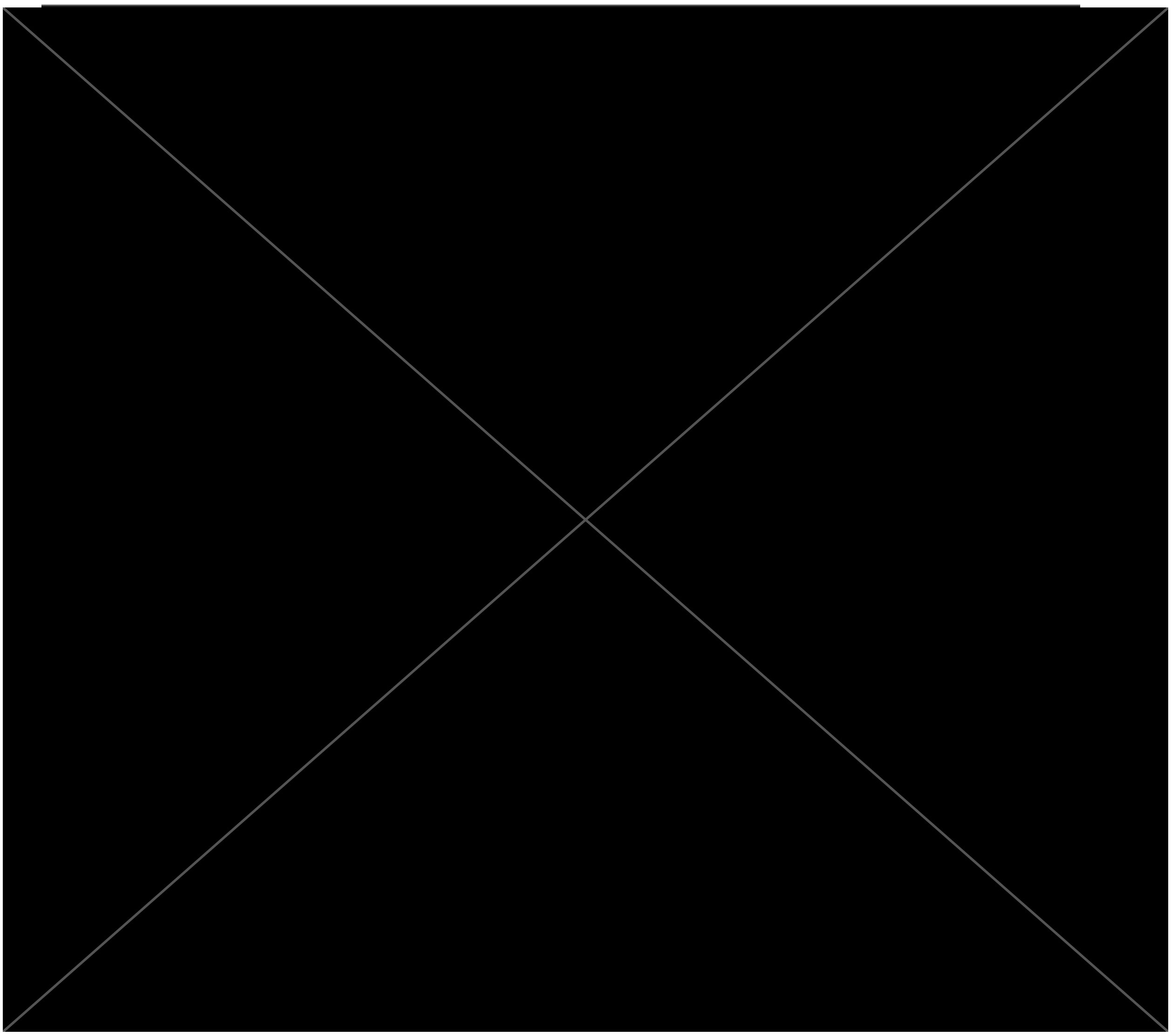
[redacted]

[redacted]

PRISSMA/CEREMA /DteridF/DM/ITS

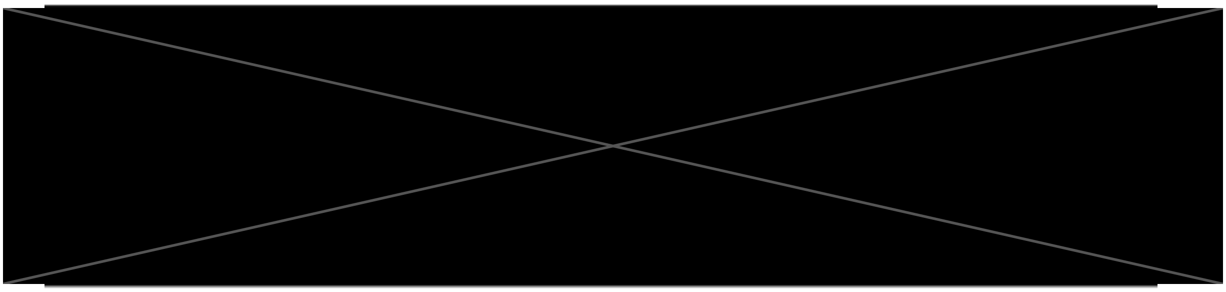
SITUATION & SCENARIO

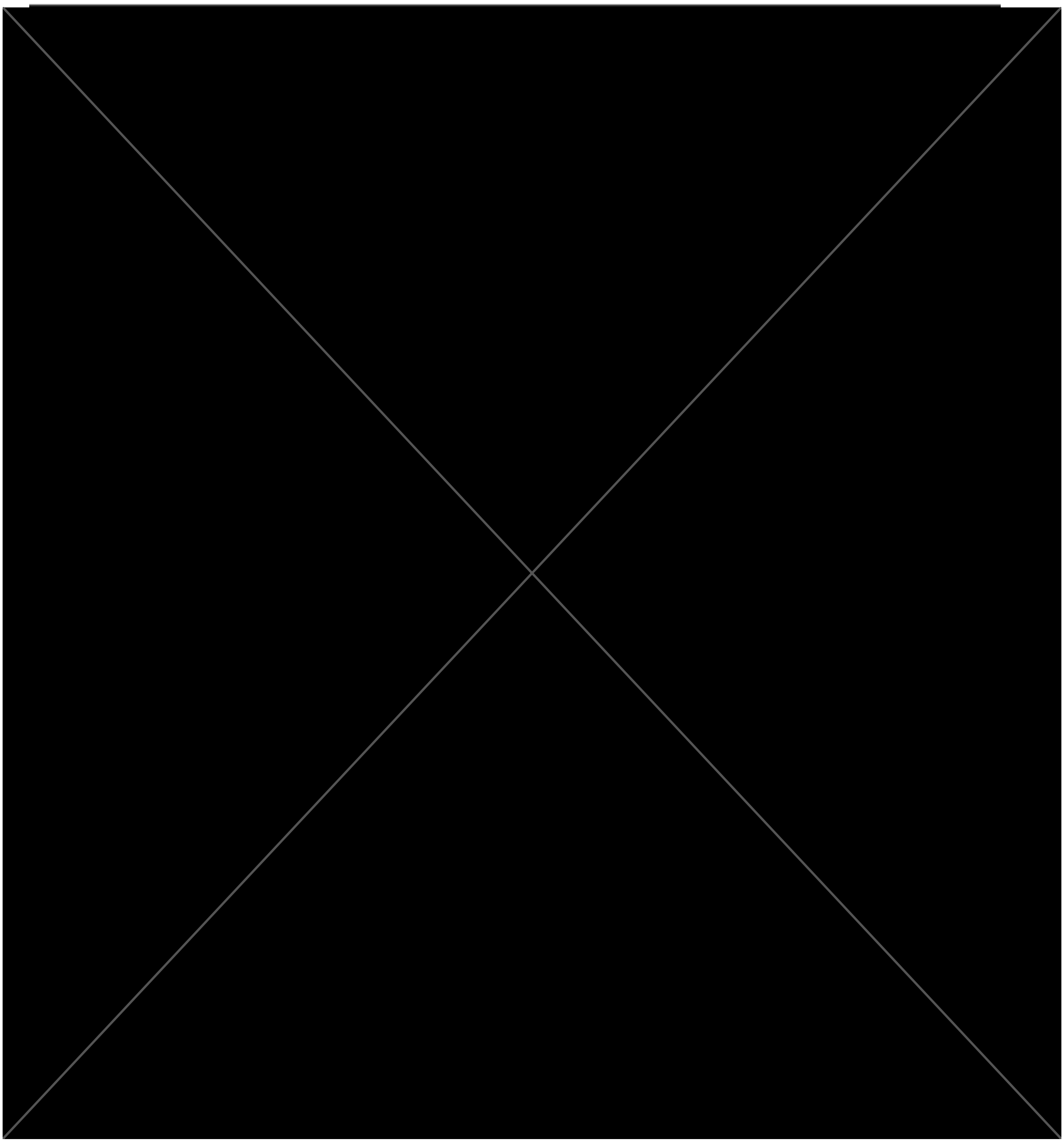




PRISSMA/CEREMA /DteridF/DM/ITS

SITUATION & SCENARIO





PRISSMA/CEREMA /DterIdF/DM/ITS

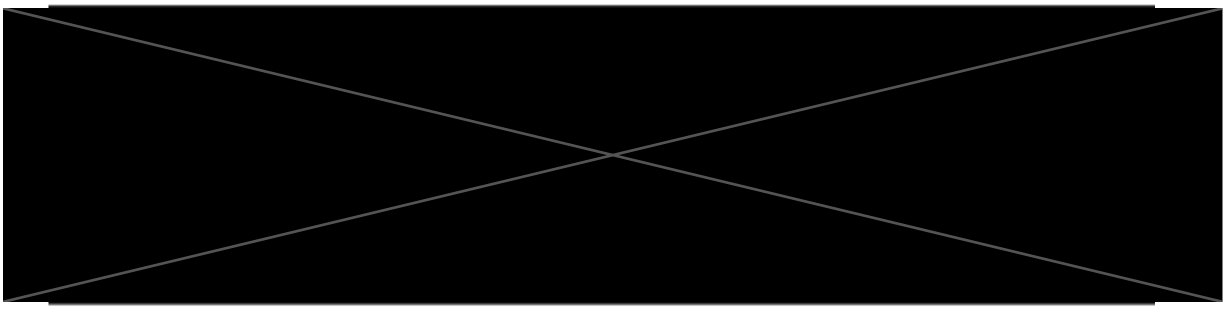
SITUATION & SCENARIO

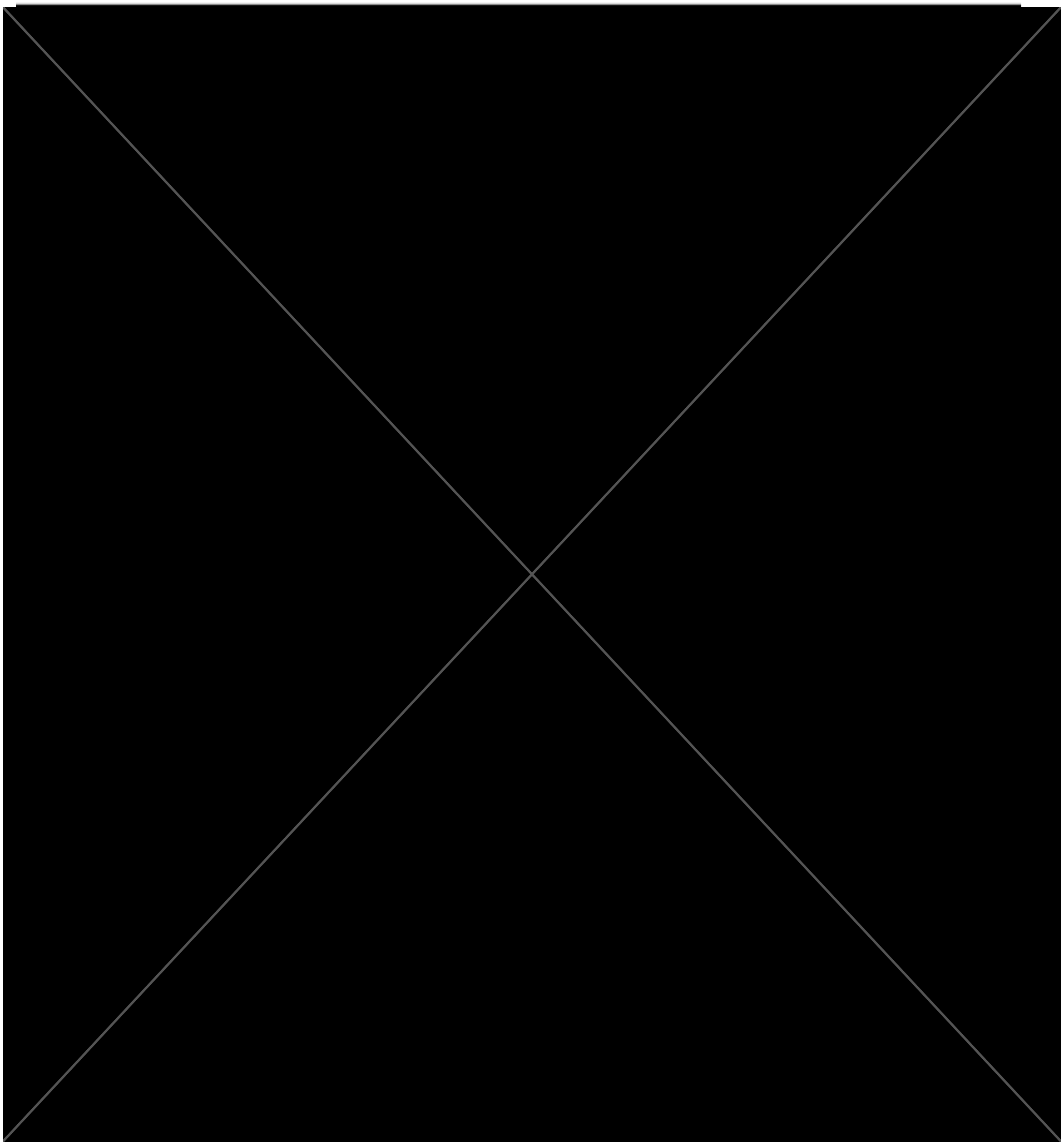
[redacted]

[redacted]

PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO





PRISSMA/CEREMA /DteridF/DM/ITS

SITUATION & SCENARIO

[redacted]

[redacted]

PRISSMA/CEREMA /DteridF/DM/ITS

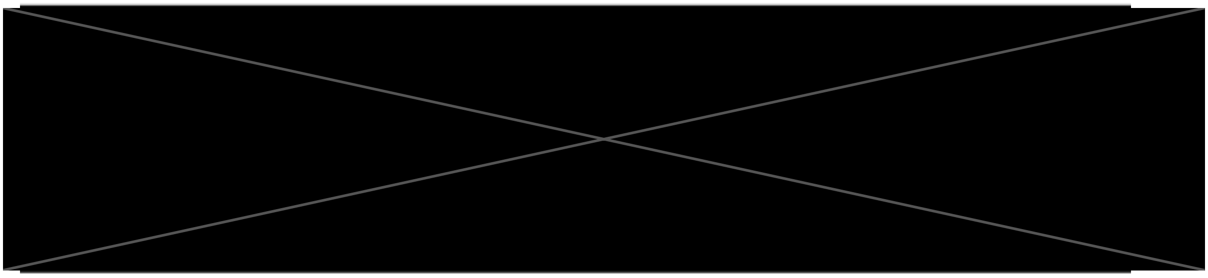
SITUATION & SCENARIO

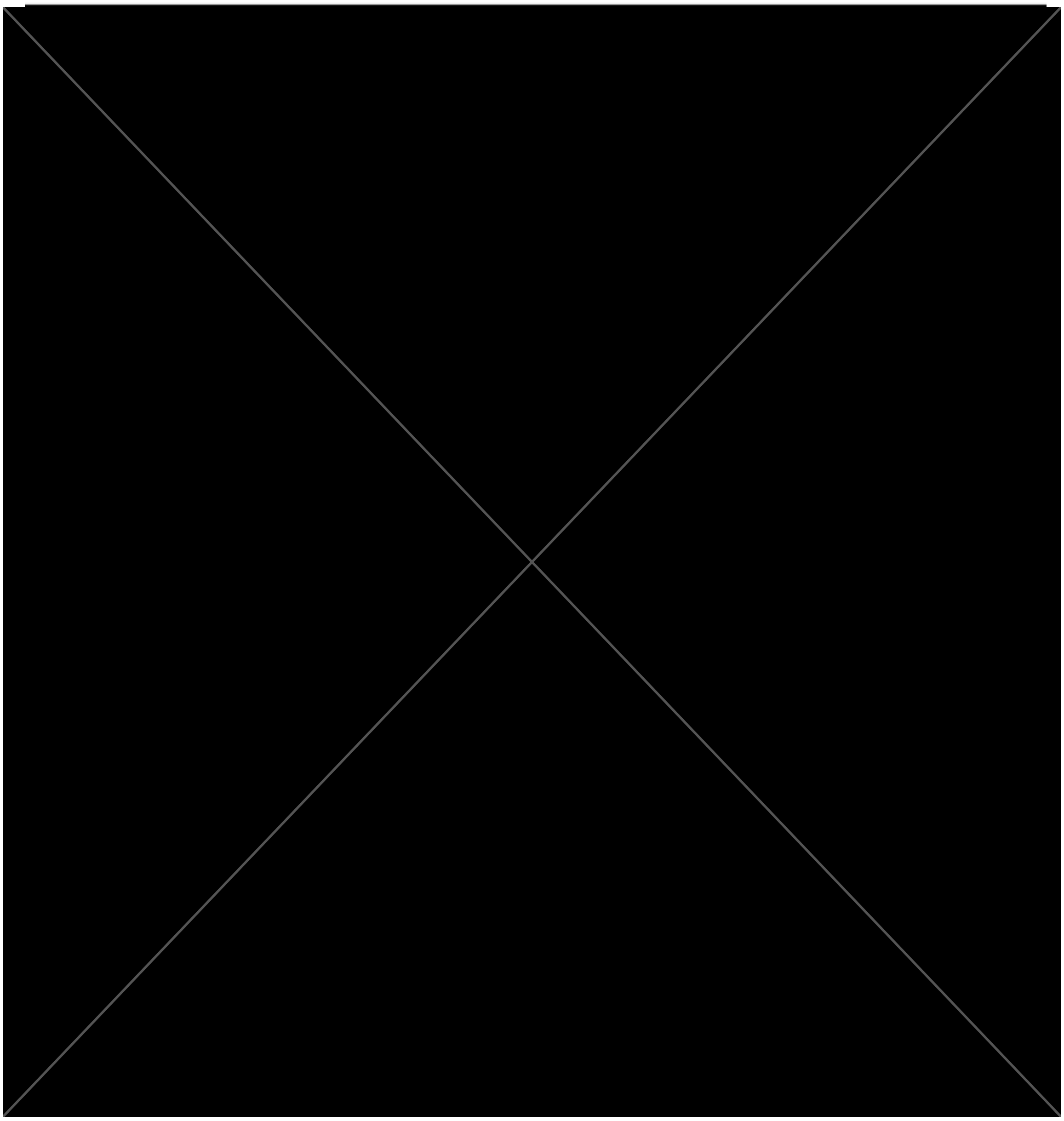
[redacted]

[redacted]

PRISSMA/CEREMA /DterIdF/DM/ITS

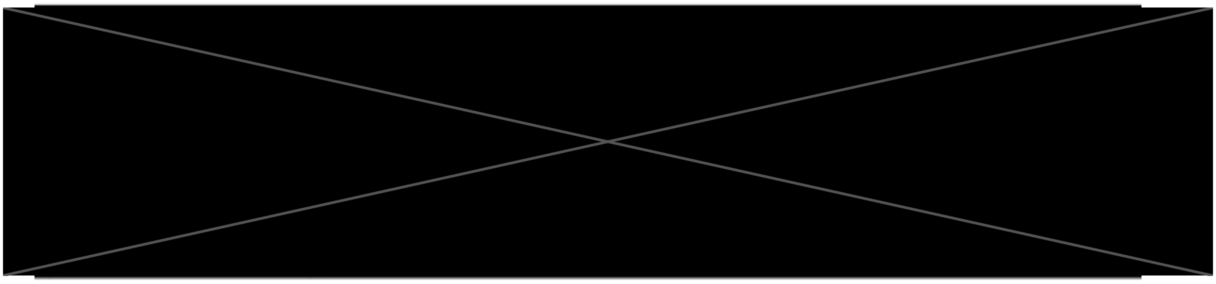
SITUATION & SCENARIO

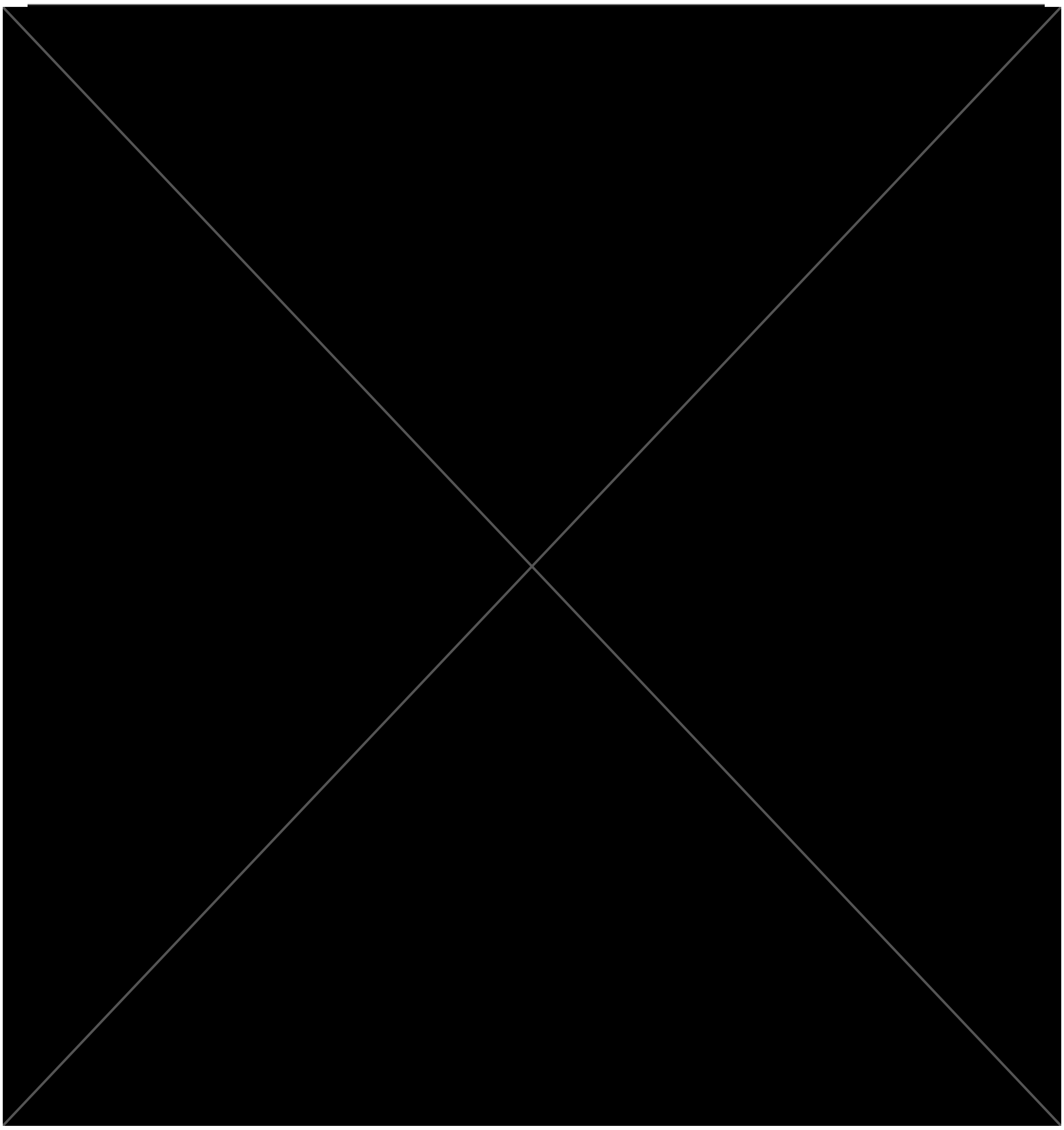




PRISSMA/CEREMA /DteridF/DM/ITS

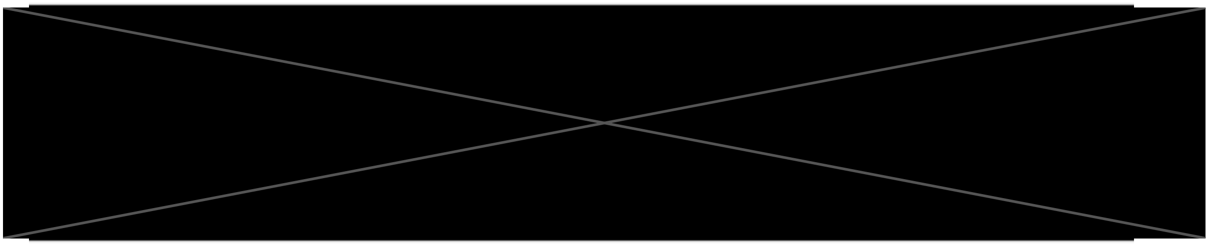
SITUATION & SCENARIO

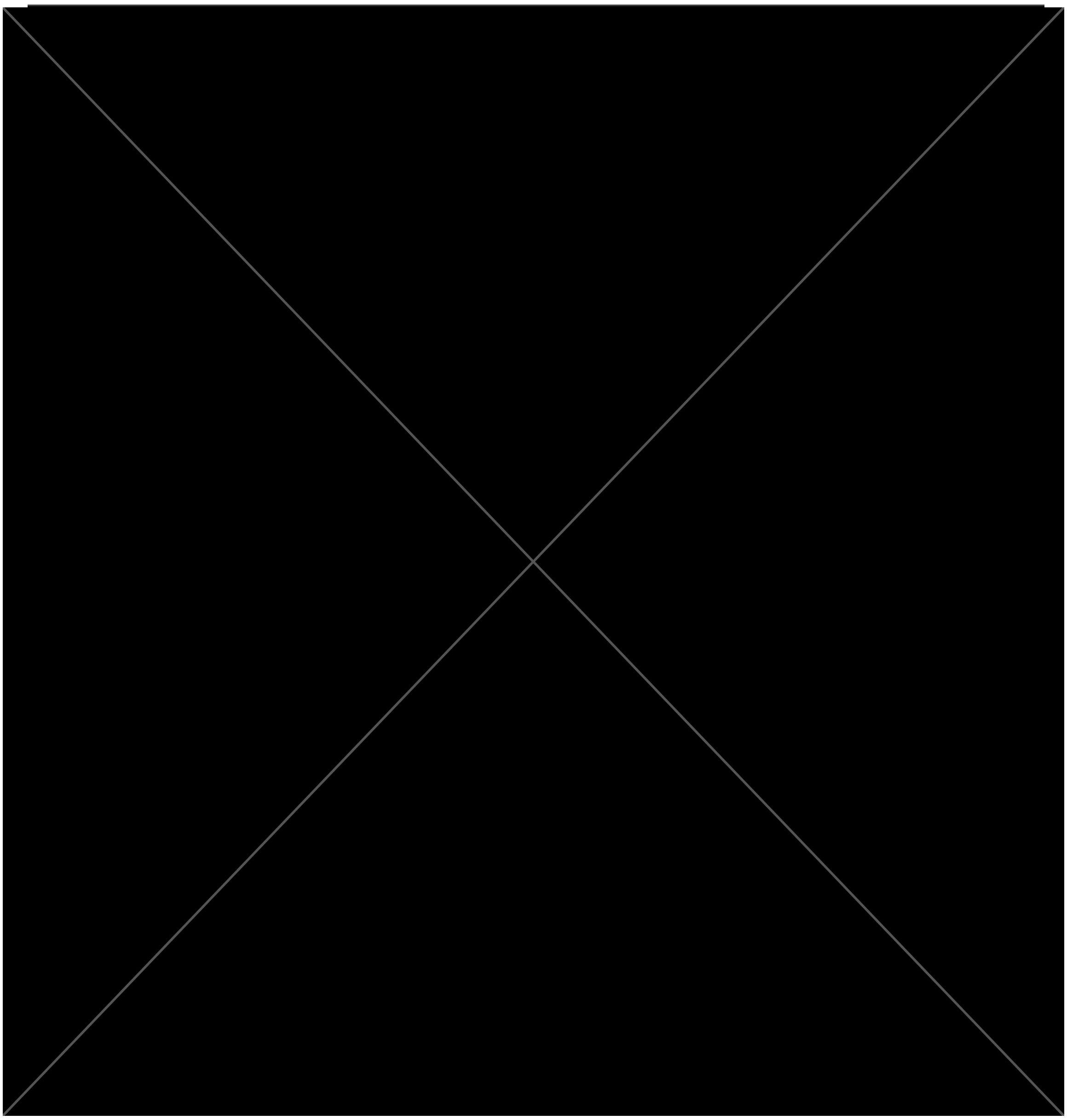




PRISSMA/CEREMA /DterIdF/DM/ITS

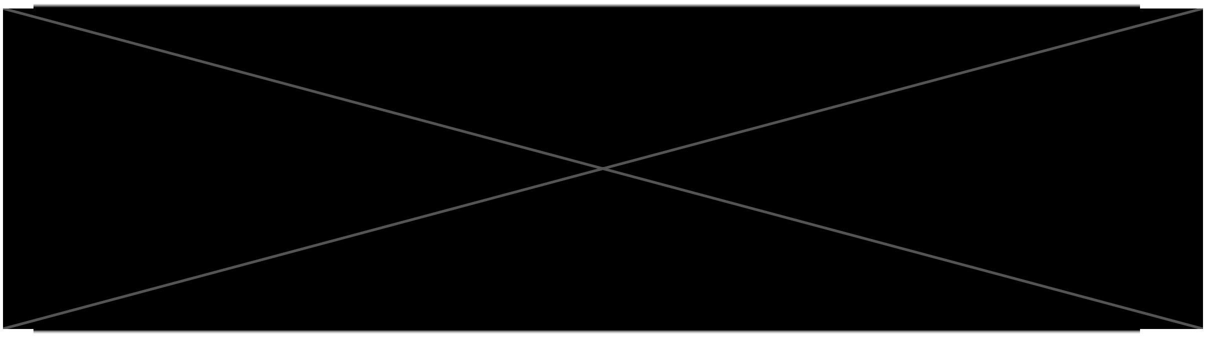
SITUATION & SCENARIO

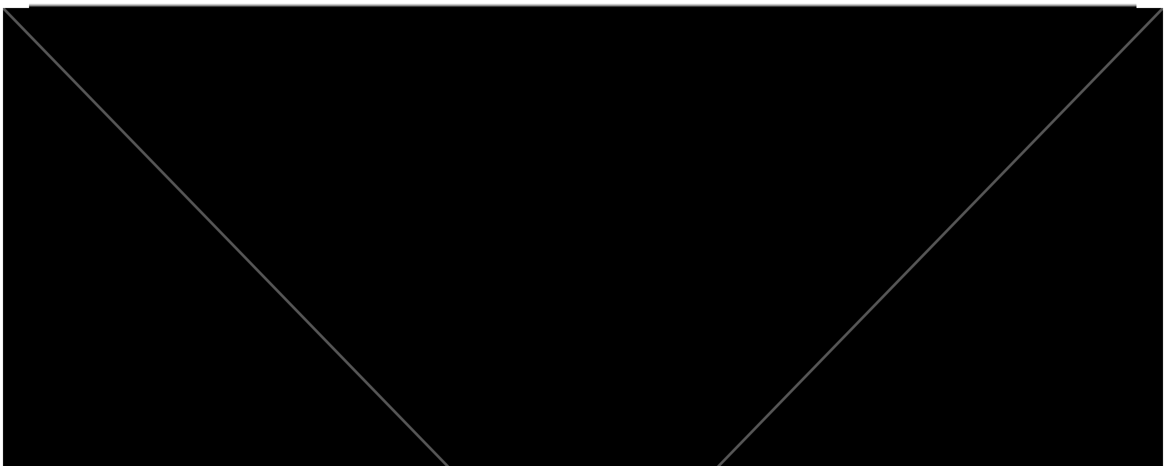




PRISSMA/CEREMA /DteridF/DM/ITS

SITUATION & SCENARIO





***PRISSMA/CEREMA** /DterIdF/DM/ITS*

SITUATION & SCENARIO

SITUATION 16 *Équivalent Scénario:* *Smartpole: 7* *Lieu:* ***Rue Van Gogh***

H début d'action : 11:04:20 *Section: 22* *Références vidéo: 20231017-110229* *Timing vidéo: 11h04:15 <> 11h04:30*

Rabattement d'un véhicule devant EGO

1/ 11h04:19 EGO roule à bonne allure rue Van Gogh

2/11h04:20 Devant, un véhicule change de voie et tente de se rabattre «prudemment» devant EGO

3/ EGO ralentit. Le véhicule se rabat donc entièrement devant EGO

4/ Derrière il n'y a pas de véhicules tout proche. Le trafic est concentré plus loin

5/ 10Le véhicule est passé devant EGO sans heurts et profite du feu vert pour accélérer

6/ EGO passe au feu vert à son tour sans heurts.

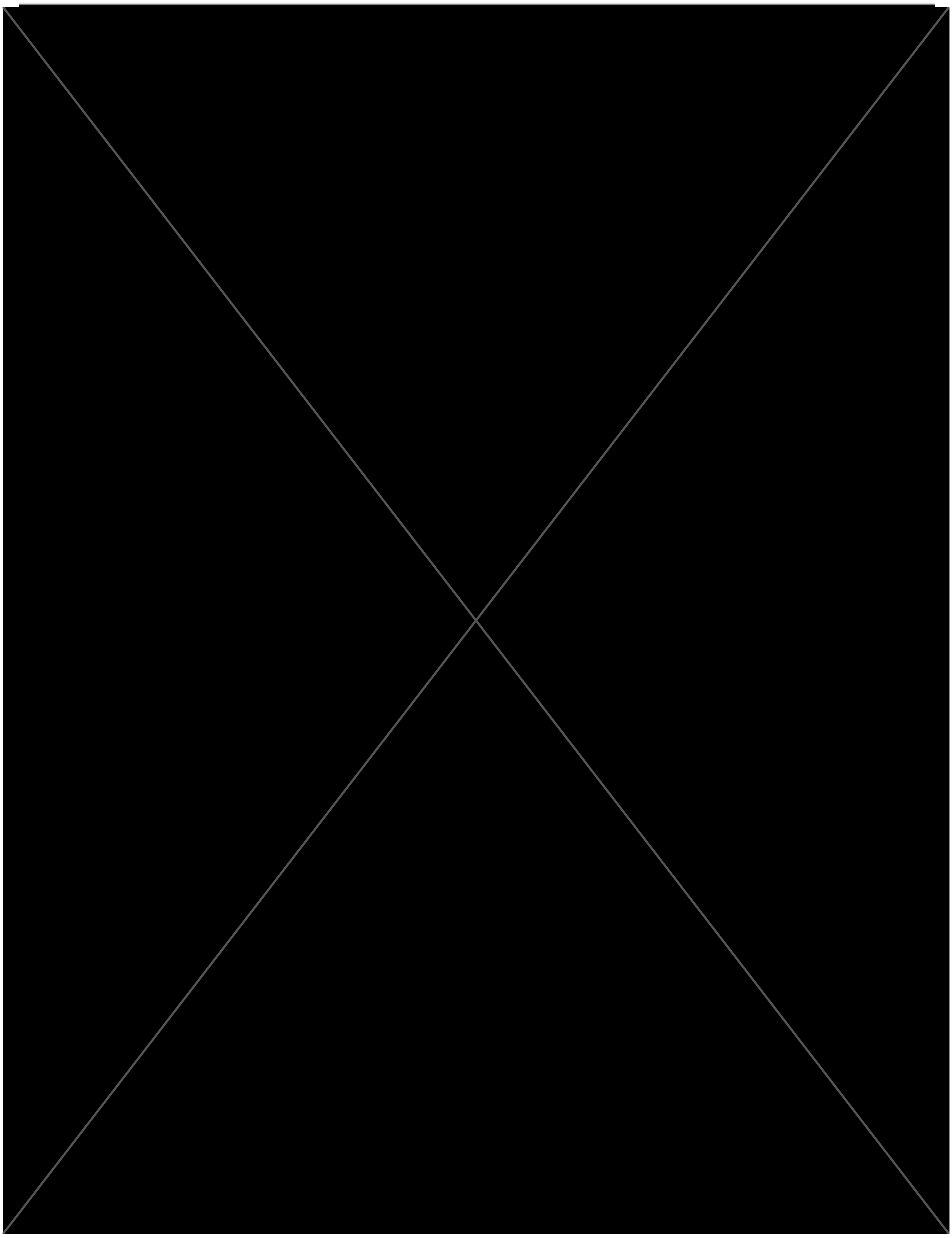
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO

[redacted]

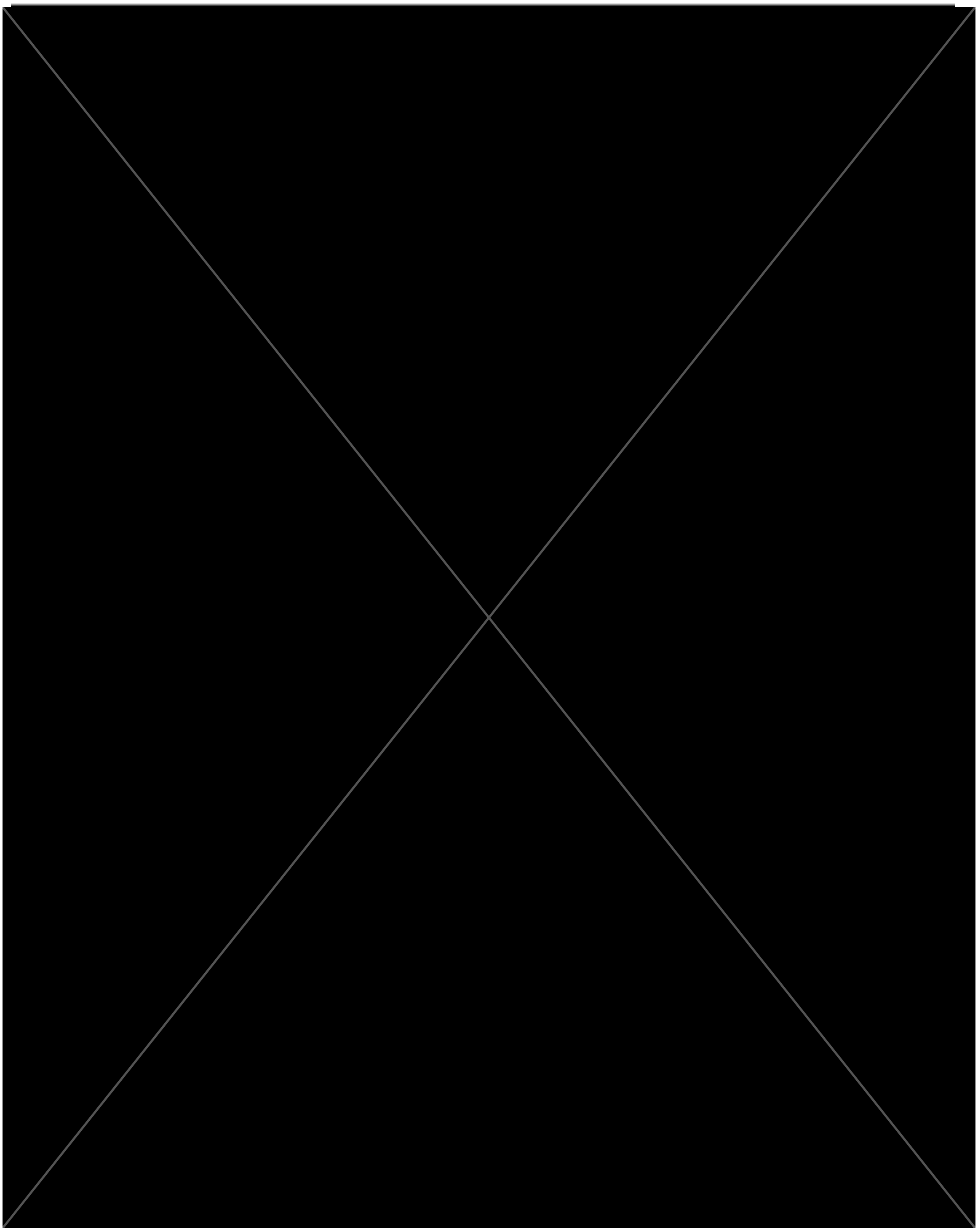
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



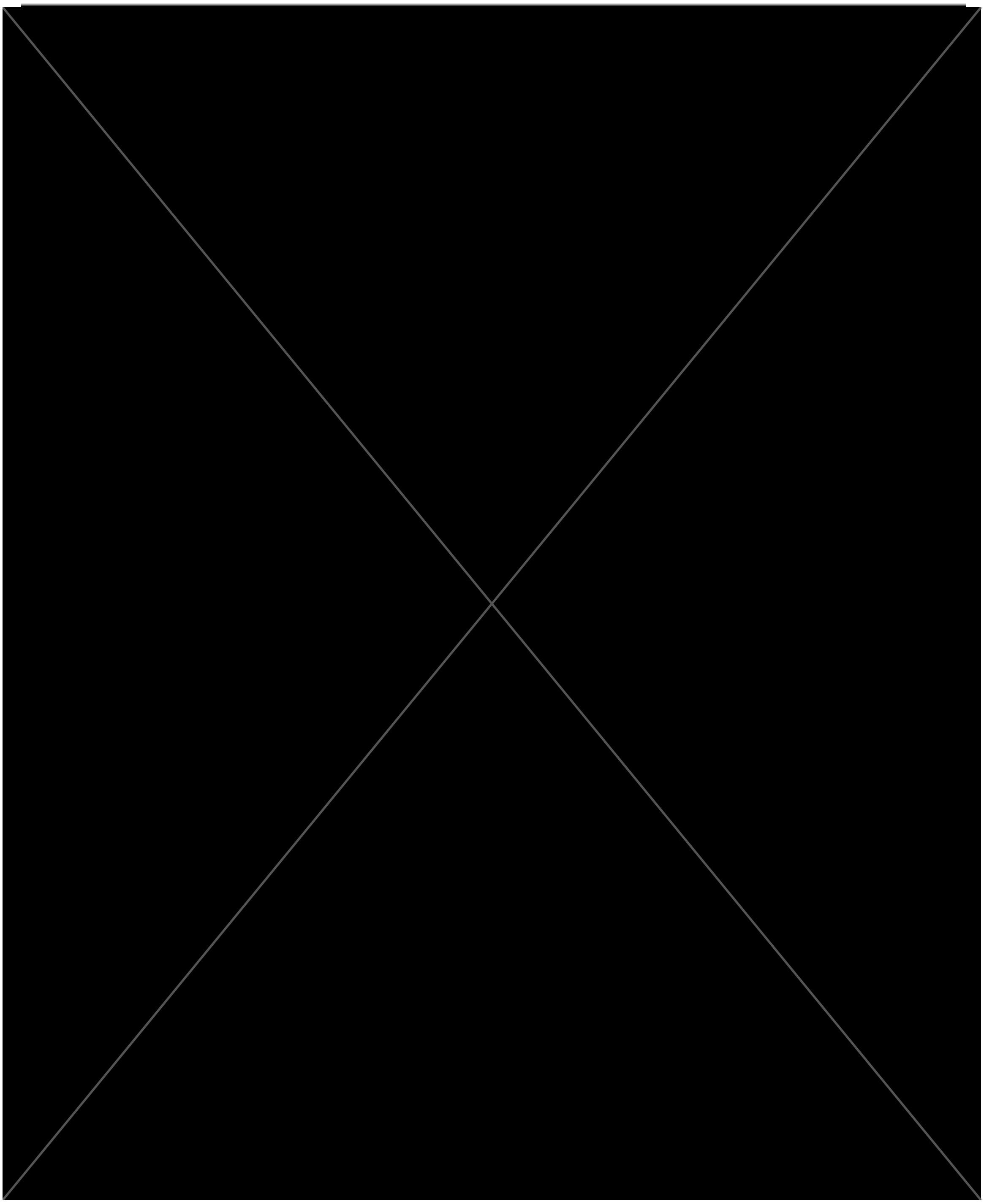
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



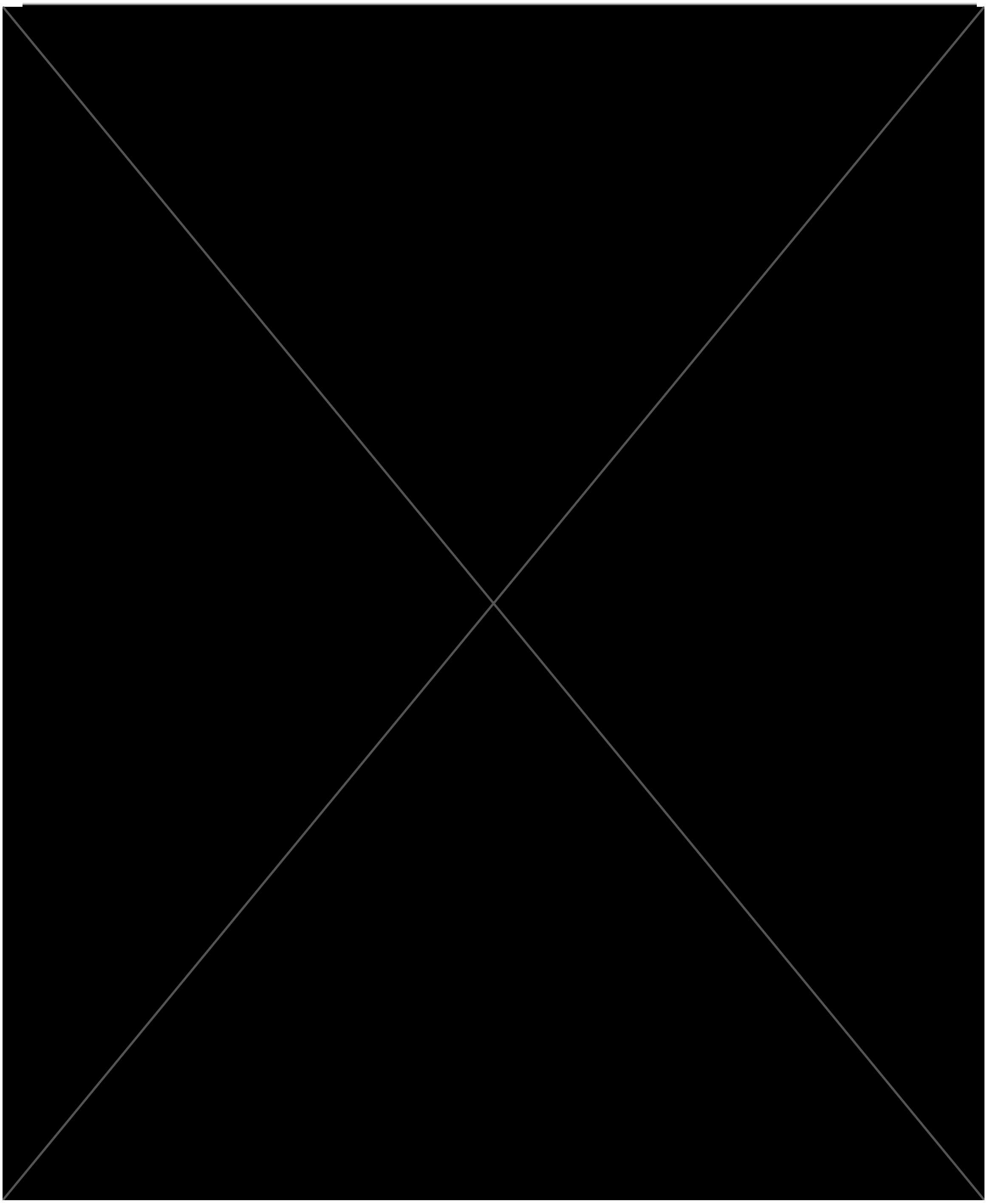
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



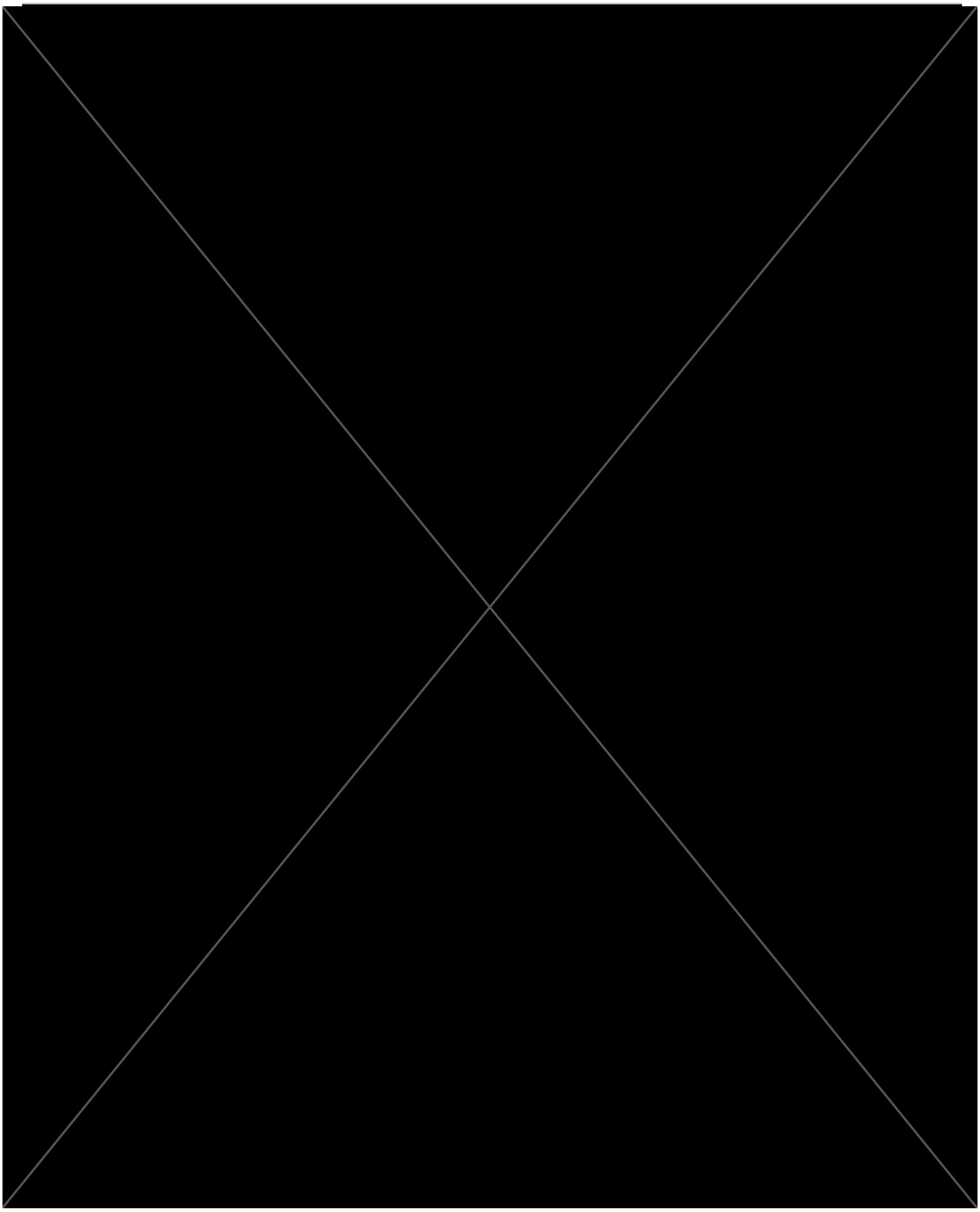
PRISSMA/CEREMA */DterIdF/DM/ITS*

SITUATION & SCENARIO



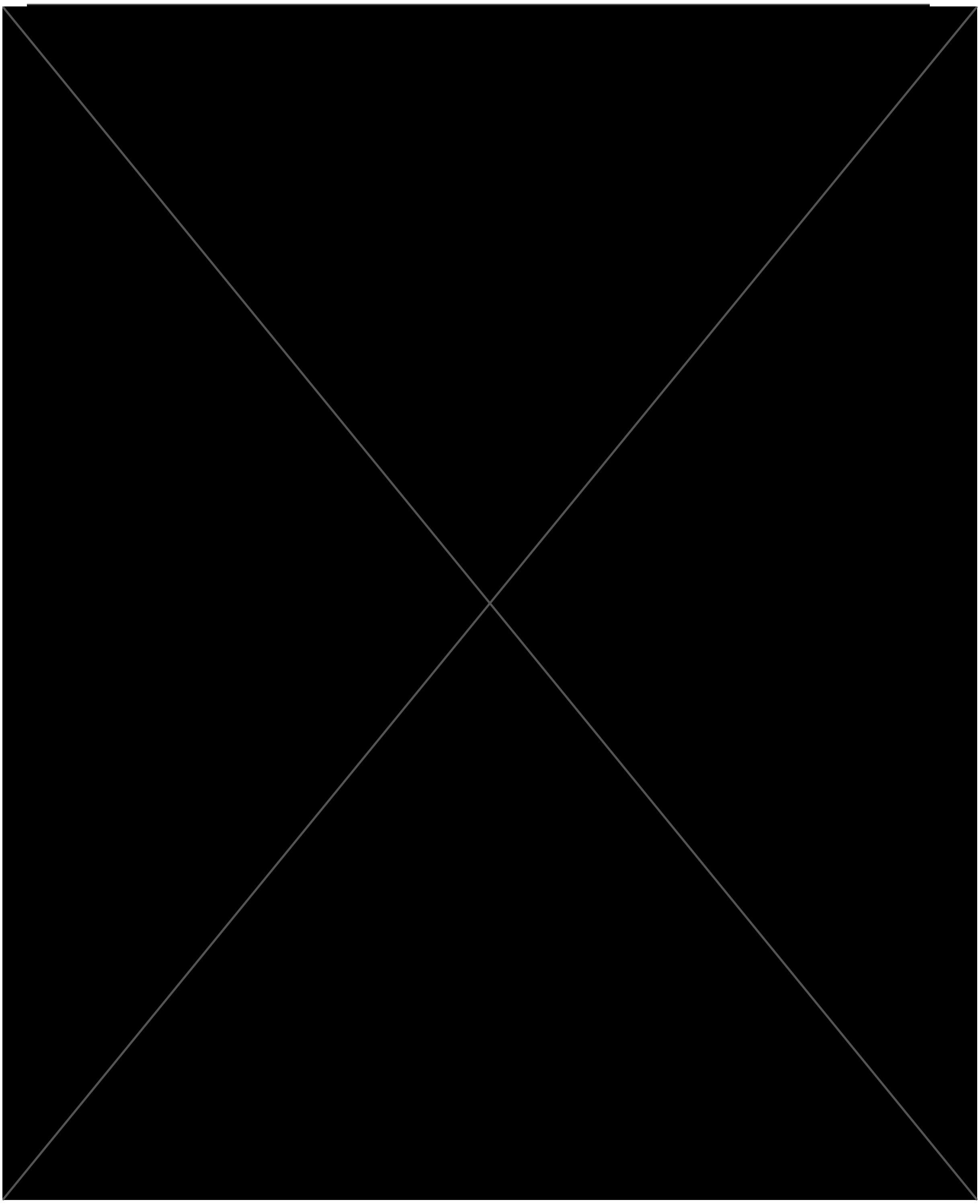
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



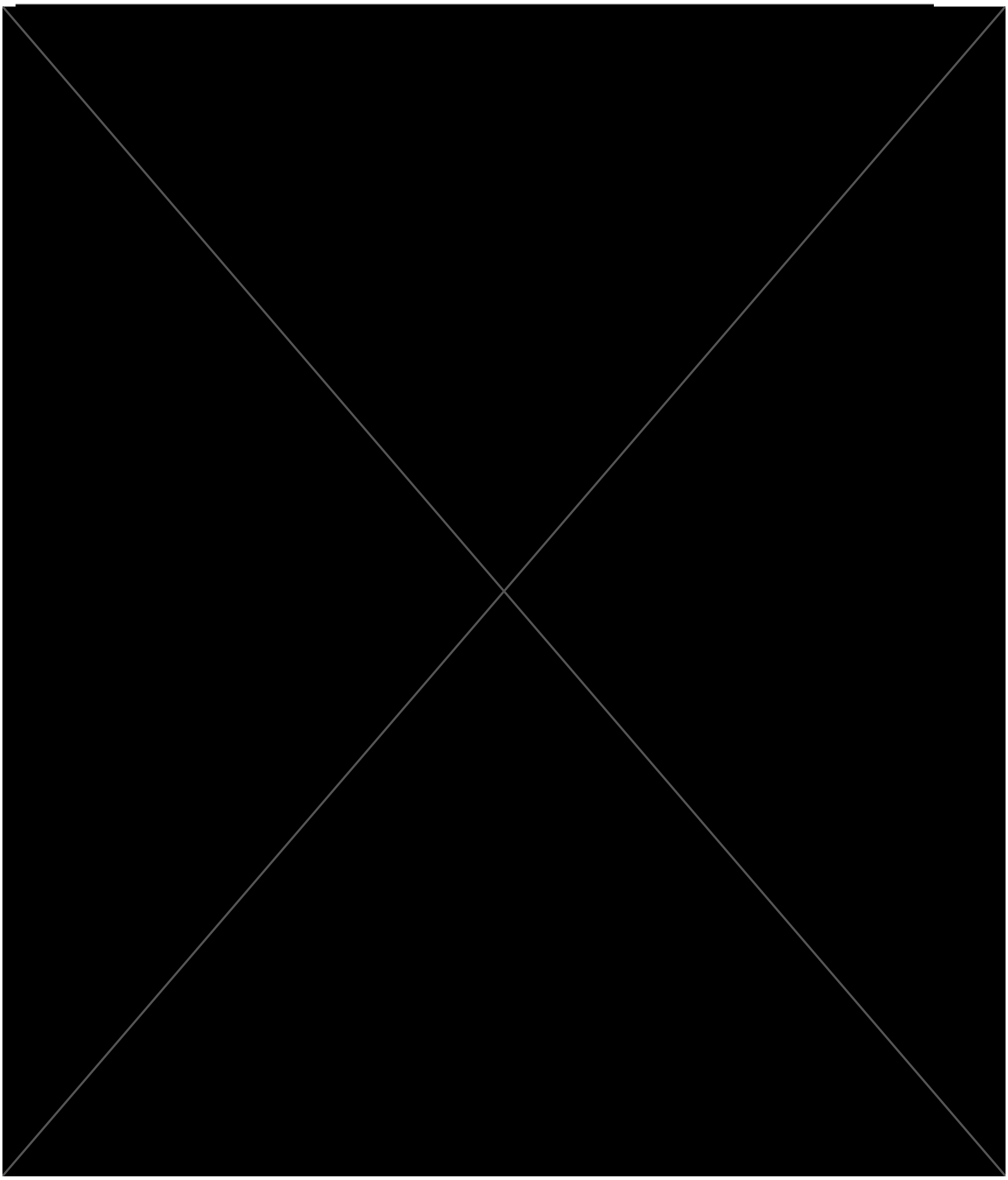
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



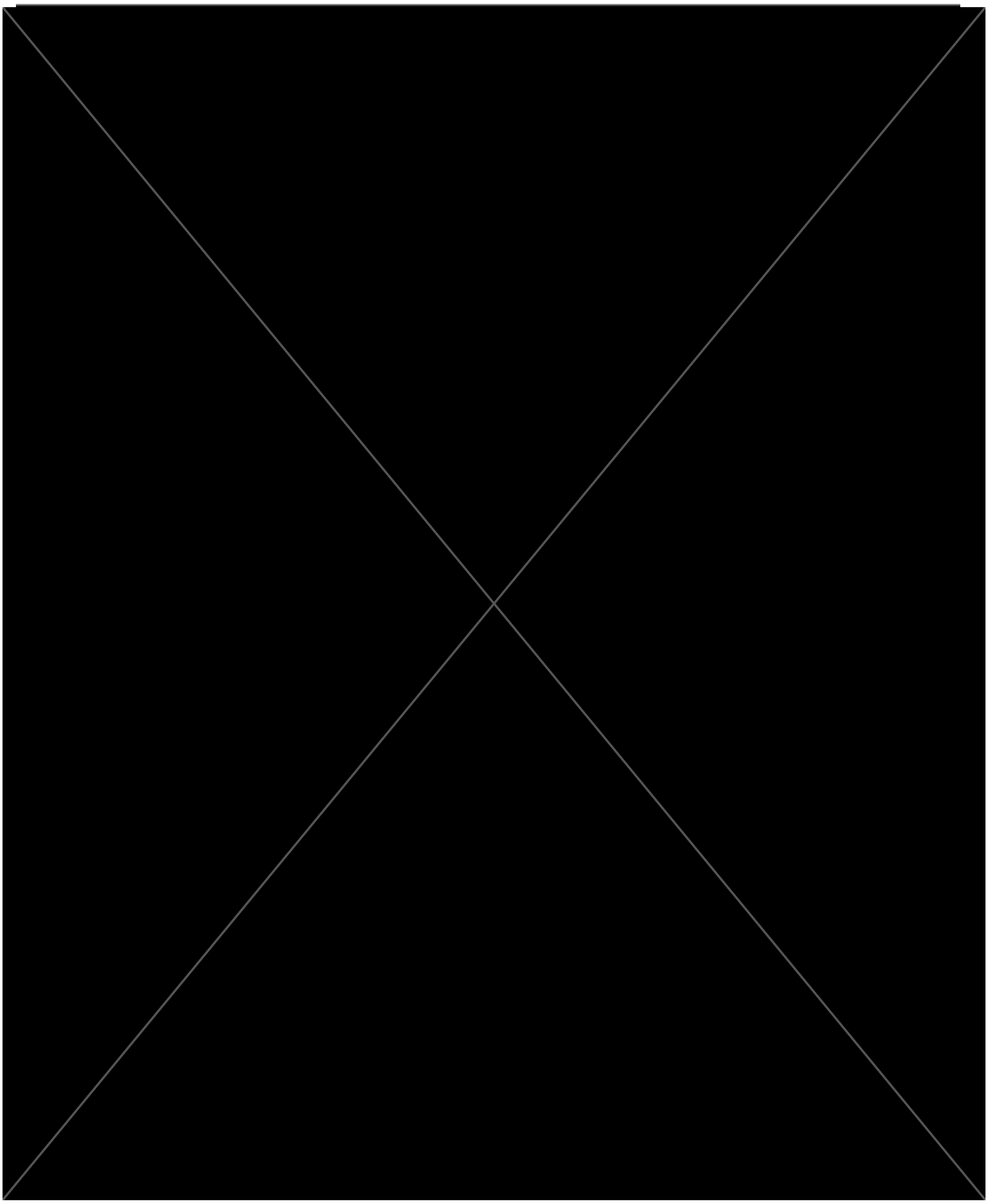
PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



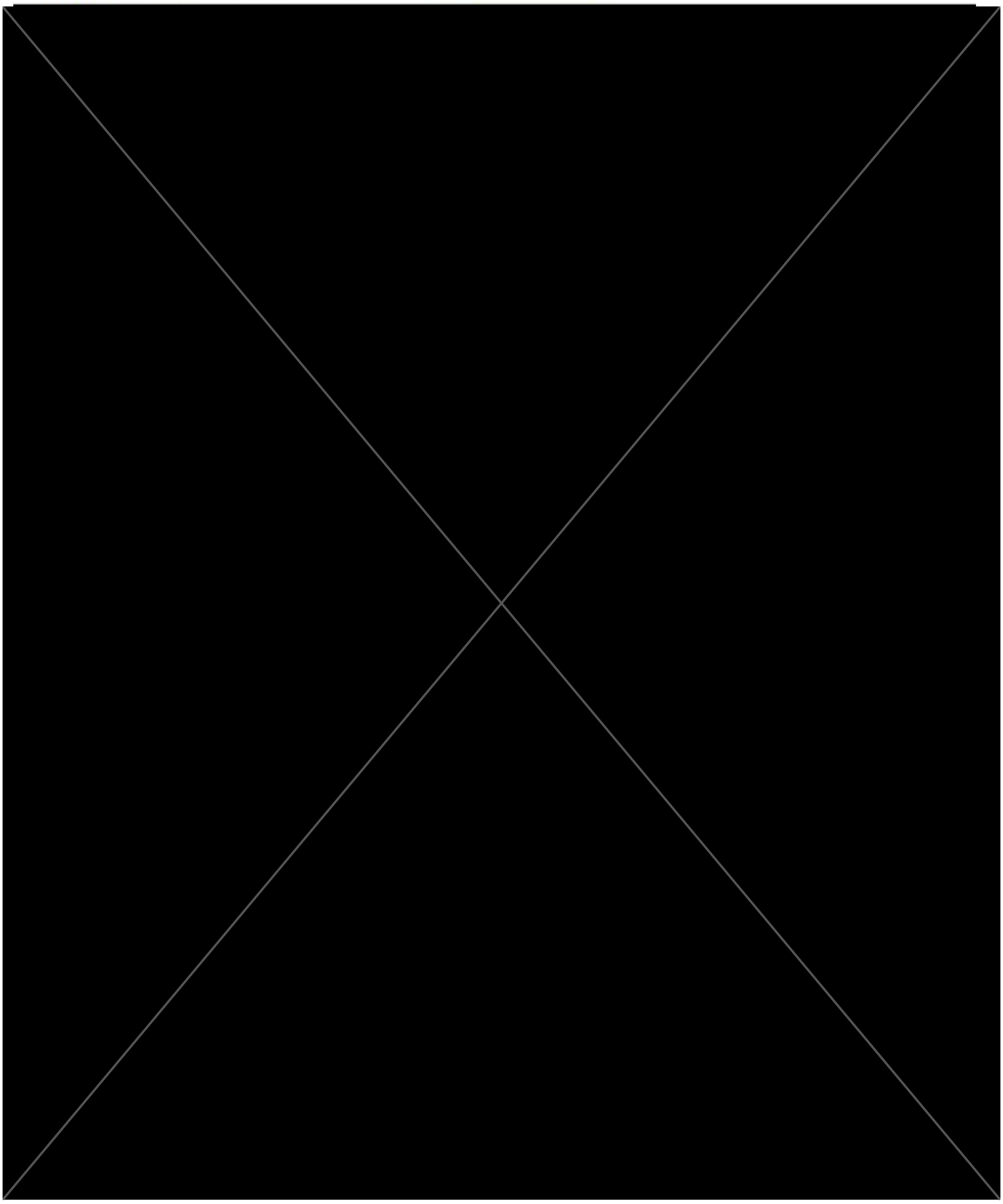
PRISSMA/CEREMA /DteridF/DM/ITS

SITUATION & SCENARIO



PRISSMA/CEREMA /DterIdF/DM/ITS

SITUATION & SCENARIO



PRISSMA/CEREMA /DteridF/DM/ITS

SITUATION & SCENARIO