EGU2020: Sharing Geoscience Online

Session GI2.5: Data fusion, integration, correlation and advances of non-destructive testing methods and numerical developments for engineering and geosciences applications

Evaluating Resilience of Infrastructures Towards Endogenous Events by Non-Destructive High-Performance Techniques and Machine Learning Regression Algorithms

Nicholas Fiorentini, Pietro Leandri, and Massimo Losa

University of Pisa - Department of Civil and Industrial Engineering,

Largo Lucio Lazzarino, 1, 56122, Pisa, Italy

nicholas.fiorentini@phd.unipi.it

pietro.leandri@ing.unipi.it

losa@ing.unipi.it

AGENDA

A BRIEF SUMMARY

1. KEYWORDS
2. AIM OF THE RESEARCH AND RESEARCH QUESTIONS
3. WORKFLOW
4. MODELING PROCEDURE
5. CASE STUDIES, ROAD SURVEYS, AND DATA COLLECTION

KEYWORDS

DEFINITION OF THE ESSENTIAL WORDS

RESILIENCE

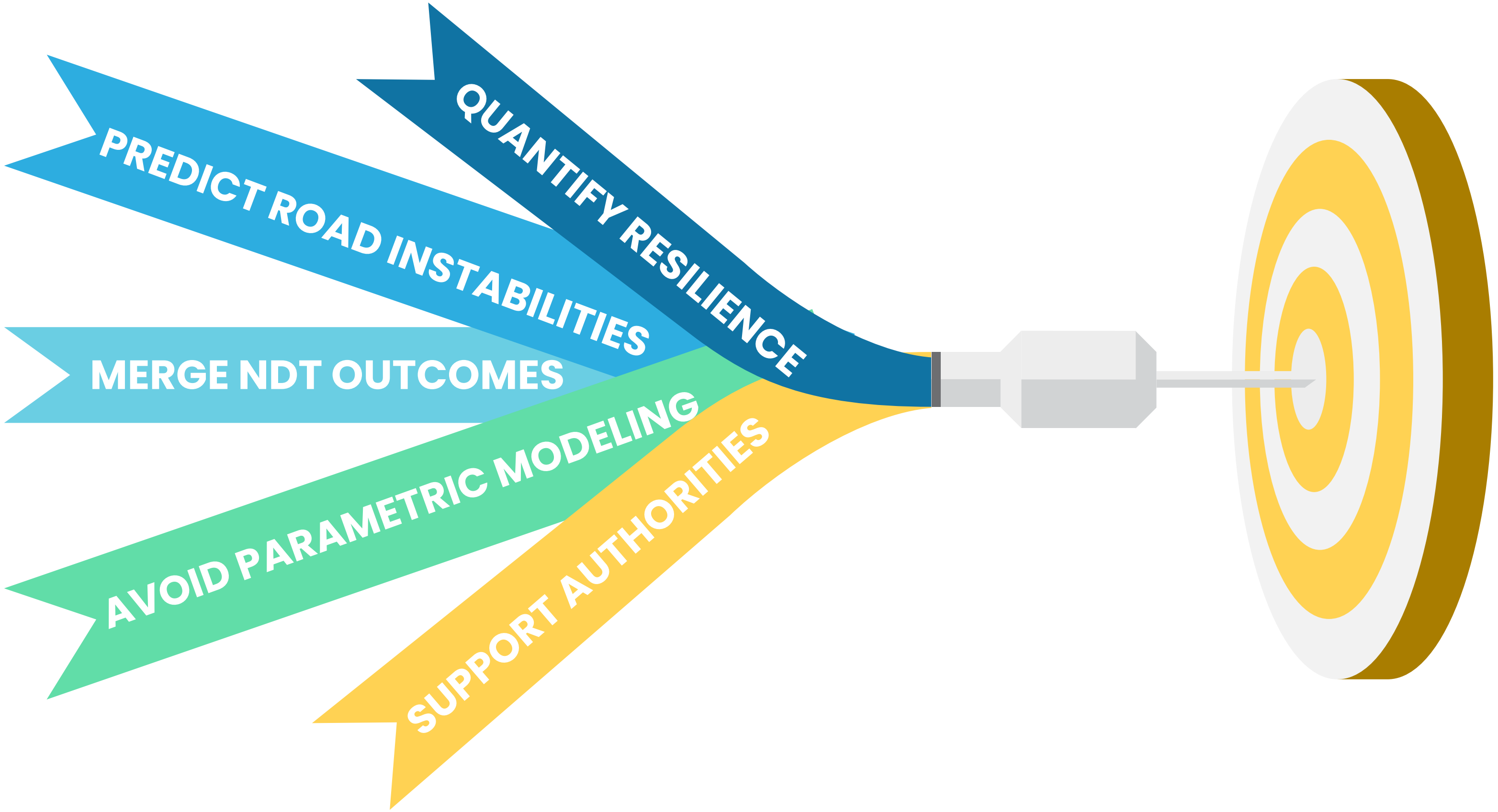
In the field of infrastructures, resilience refers to the intrinsic capacity of the infrastructure to face a significant disruption of its operating status in order to absorb as quickly as possible the change of its internal properties and establish a new condition of equilibrium.

ENDOGENOUS FACTORS

Endogenous factors are events related to the infrastructure itself and can cause decreases in structural performances and in the surface quality. They are design errors, prolonged use, and other human-related events. These events are related to intrinsic features of the infrastructure, such as pavement surface roughness and friction, thickness of the layers, structural features of the pavement, age of the pavement, traffic flow, and so on.

THE AIM

EVALUATION OF HOW MUCH INFRASTRUCTURES ARE RESILIENT TOWARDS **ENDOGENOUS** FACTORS FOR OPTIMIZING MONITORING, MAINTENANCE, AND SAFETY LEVEL.



WHY?

THE MAIN GOAL

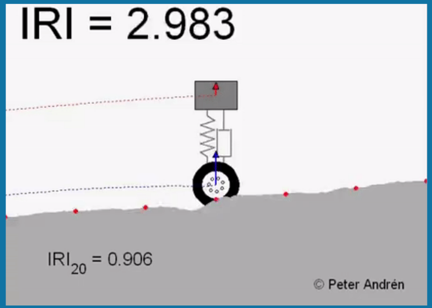
Save Money

Managing and saving as many resources as possible is crucial for the proper functioning of the entire road pavement management cycle, especially in contexts where funds are increasingly limited

Allocate funds properly

Road maintenance planning through reliable predictive tools, developed specifically for improving decision-making processes, leads to an objective allocation of funds over the entire road network. Sites with real or potential criticalities are inspected and maintained with priority.

THE MAIN ASSUMPTION

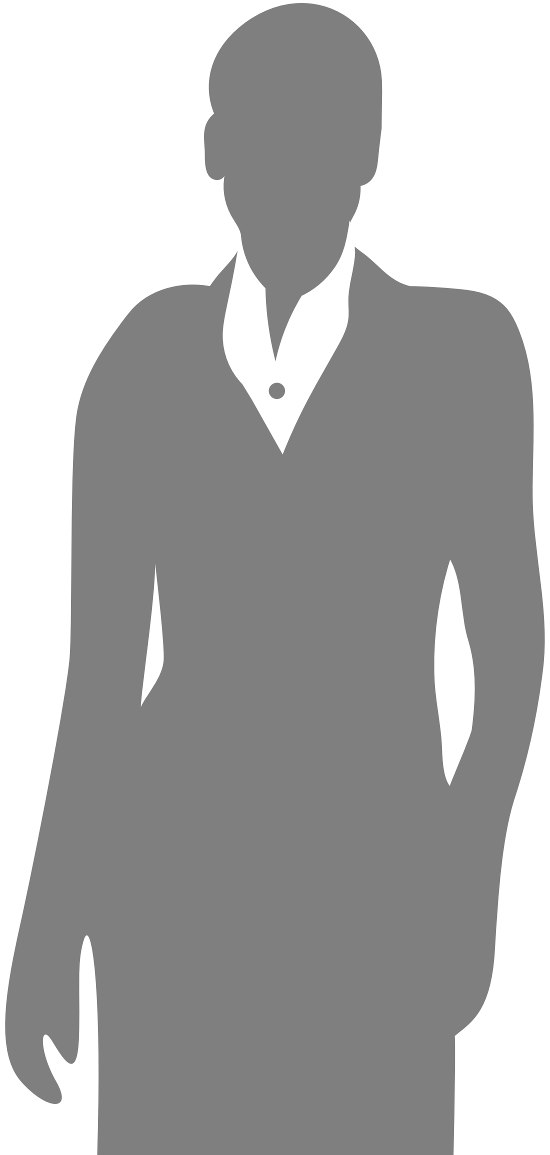


INTERNATIONAL ROUGHNESS INDEX (IRI)

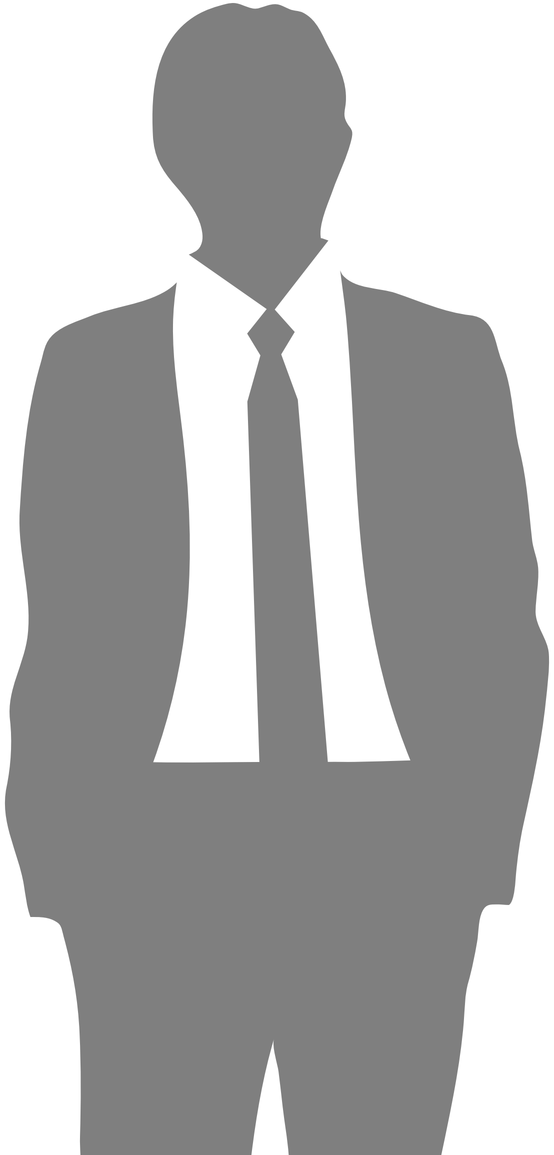
ROAD DAMAGES AND DEFINCIENCIES

WE ASSUMED THE INTERNATIONAL ROUGHNESS INDEX (IRI) **AS PROXY OF ROAD DAMAGES**.
IRI HAS BEEN ASSUMED AS THE PARAMETER FOR REPRESENTING THE CONDITION OF A ROAD SITE. THEREFORE, THE HIGHER IS THE **IRI,** THE HIGHER IS THE PROBABILTY OF IDENTIFYING A DAMAGED ROAD SITE (IN TERMS OF DAMAGE OF THE PAVEMENT STRUCTURE AND THE SUBGRADE).

RESEARCH QUESTIONS



1. Can IRI be predicted correctly merging various NDT-based outcomes?
2. What is the set of endogenous conditioning factors that most affects IRI?
3. Are Machine Learning Algorithms suitable for such purposes?
4. How can Authorities use Machine Learning Algorithms for planning maintenance?



THE METHODOLOGY

THE MAIN STEPS FOLLOWED

NDT-BASED SURVEY

DATA INTEGRATION

MODELING

MAKE PREDICTION AND RANK ROAD INSPECTIONS

STEP 1

Carry out the NDT-based surveys and collect all the required data for a comprehensive representation of the infrastructures' condition

STEP 2

Integration of all different types of data into only one suitable platform (e.g., the use of spreadsheets)

STEP 3

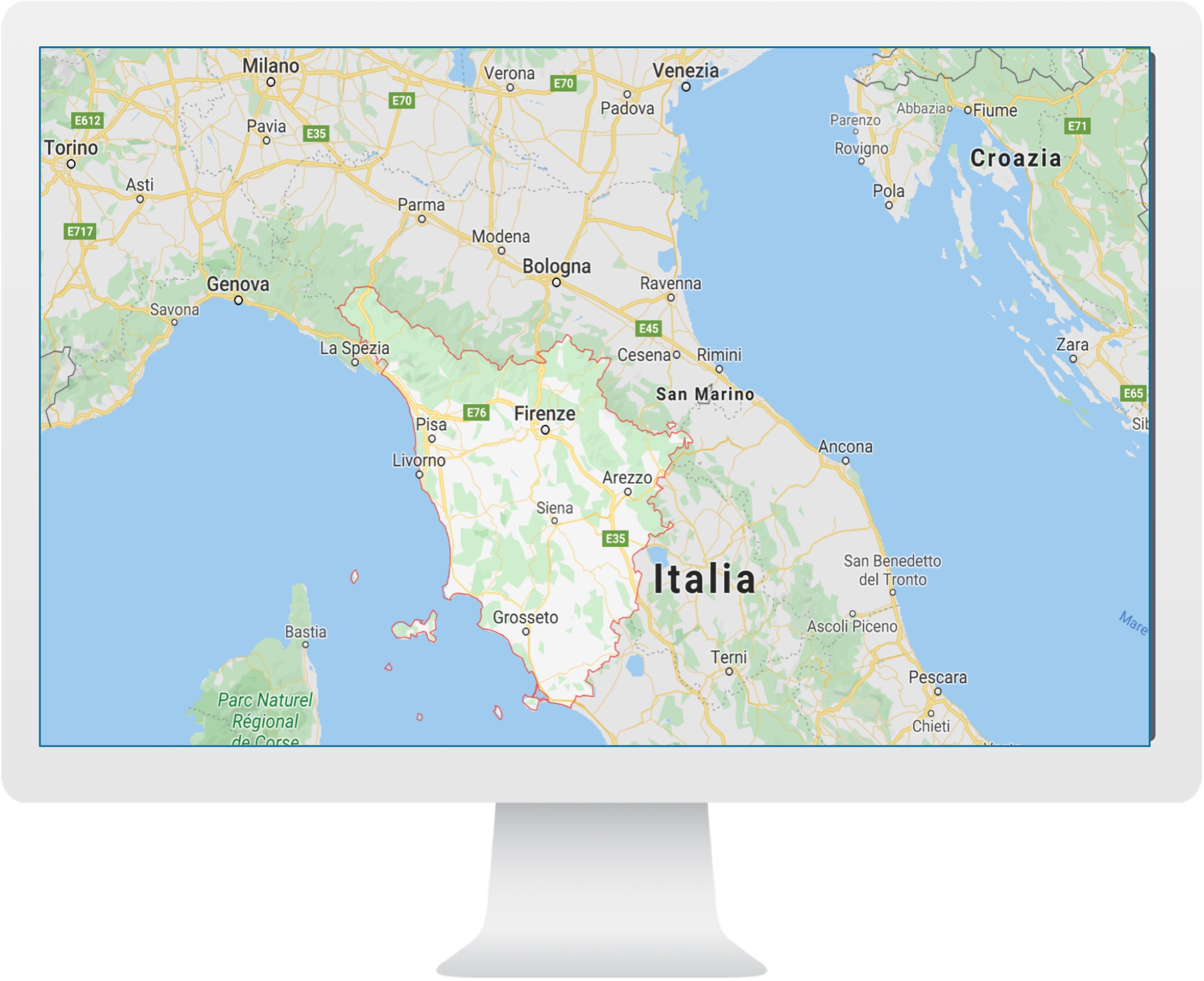
Calibration of the machine learning algorithms: training, validation, evaluation, and comparison of the algorithms for ensuring the prediction reliability

STEP 4

Use the algorithms for supporting decision-making processes: make predictions where data are not available and ranking all road sites from most criticals to those safer.

THE STUDY AREA

THE TUSCANY REGION, CENTRAL ITALY



1 ROAD NETWORK

The road network analyzed is the one managed by the Tuscany Region Road Authority. It is composed mainly by two-lane rural roads. These roads usually cross both urban and rural areas.

2 ROAD DATABASE

The Tuscany Region Road Authority provided a massive amount of information (geometrical and environmental data) related to this network.

3 TEST SITES

We selected four two-lane rural road stretches for carrying out the NDT-based surveys. The stretches extend for more than 10 km. Two sites are located in rural areas, while the other ones are located in urban areas.

TEST SITES

FOUR TWO-LANE RURAL ROAD STRETCHES

Test Site 1
Castiglion Fiorentino
(province of Arezzo)
Length: 5930 m
Urban Area

Test Site 2
Massarosa
(province of Lucca)
Length: 2610 m
Urban Area

Test Site 3
Empoli
(province of Florence)
Length: 1820 m
Rural Area

Test Site 4
Volterra
(province of Pisa)
Length: 1000 m
Rural Area

DATA COLLECTION

NDT-BASED OUTCOMES

The NDT techniques employed are Falling Weight Deflectometer, Ground Penetrating Radar, Skiddometer BV11, and Laser Profiler

TRAFFIC FLOW DATA

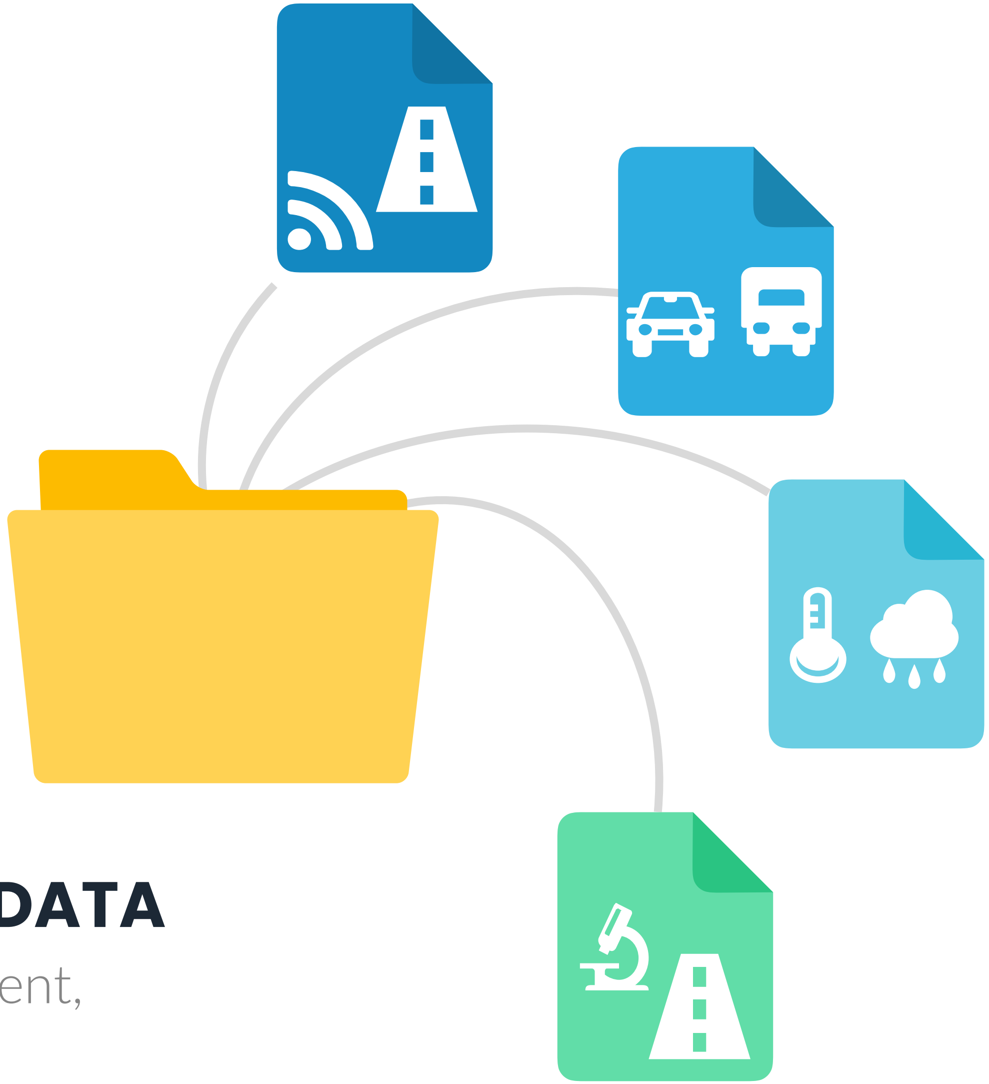
The information on traffic flows have been collected in terms of heavy and light traffic

ENVIRONMENTAL INFORMATION

With “Environmental” we refer to the information available on temperatures and rainfalls

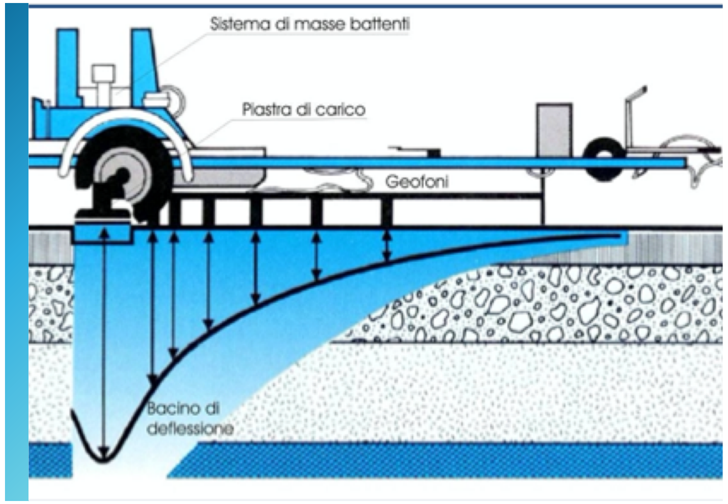
ADDITIONAL PAVEMENT STRUCTURES DATA

Any other relevant information related to the pavement, collected by laboratory test, has been considered



NDT-BASED ROAD SURVEYS

FALLING WEIGHT DEFLECTOMETER

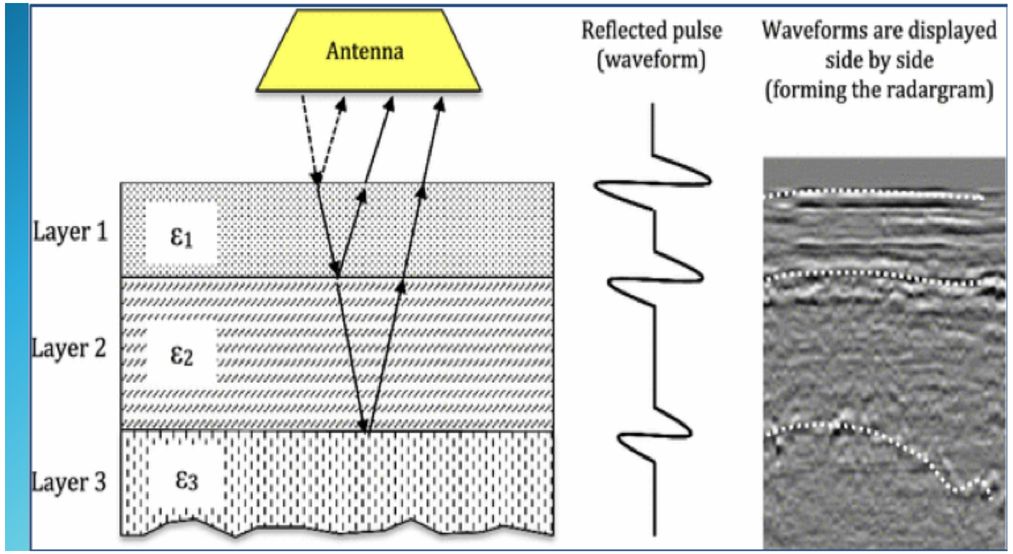


FWD is a testing device used to evaluate the physical properties (Elastic Modulus of the pavement layers and Modulus of the Subgrade reaction) of a pavement structure. The FWD is designed to impart a load pulse to the pavement surface which simulates the load produced by a rolling vehicle wheel. The load is produced by dropping a large weight, and transmitted to the pavement through a circular load plate.

Deflection sensors (geophones), mounted radially from the center of the load plate, measure the deformation basin of the pavement in response to the load.

NDT-BASED ROAD SURVEYS

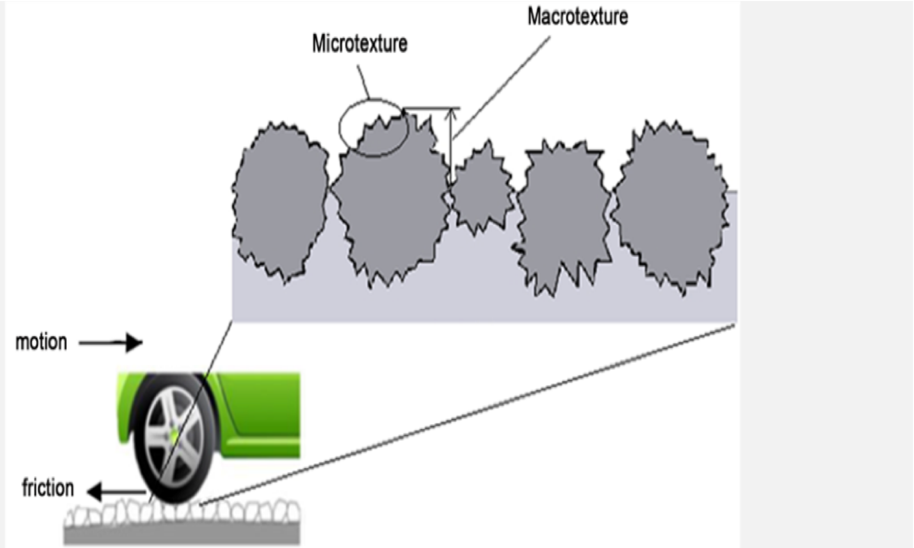
GROUND PENETRATING RADAR



GPR is a geophysical device that uses radar pulses to map the subsurface.
GPR uses electromagnetic radiation in the microwave band, and detects the reflected signals from subsurface structures.
Through the GPR, producing the radargram, it is possible to collect information on each layer of the pavement and the subgrade; for this application GPR is used to detect number, type, and thickness of the layers of the pavement structure.

NDT-BASED ROAD SURVEYS

SKIDDOMETER BV11



SKIDDOMETER BV11 is a NDT road survey able to detect the friction of the surface of the pavement structure.

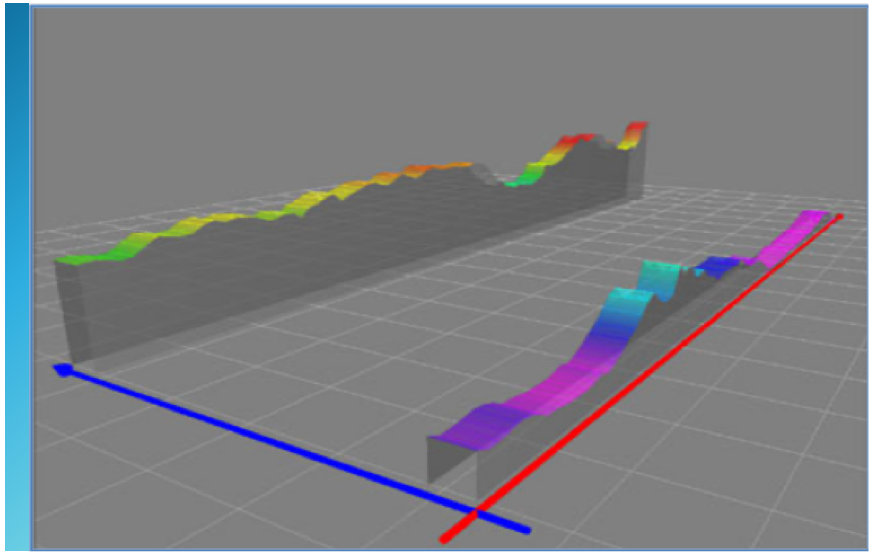
The survey is suitable for checking the safety conditions of particular road sections in which the friction (curves and junction) is particularly mobilized. Indeed, it is known that a lack of friction can be a primary reason for the generation of accidents.

The BV11 provides the use of standardized tires and the use of an integrated system to achieve wet road conditions with different water film thicknesses.

Moreover, the equipment has been integrated with a device that allows measuring the macrotexture of the surface.

NDT-BASED ROAD SURVEYS

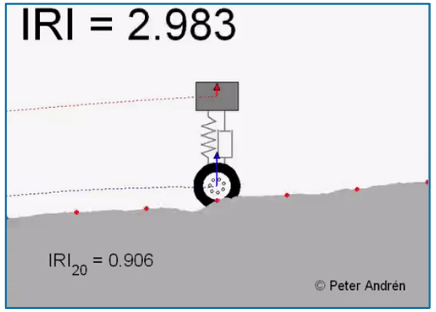
LASER PROFILER VAN



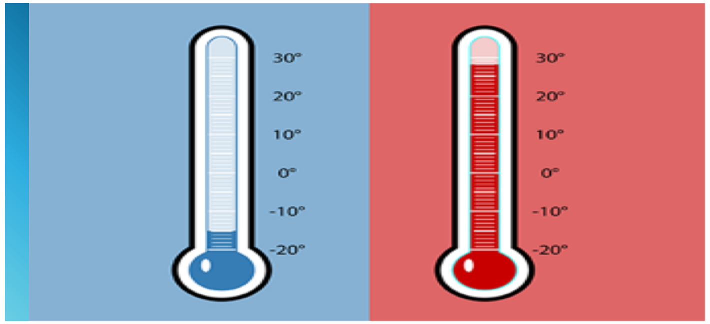
The laser profiler van is able at detecting the longitudinal profile of a road stretch. It can compute the International Roughness Index (IRI).

IRI is used to define a characteristic of the longitudinal profile of a roadway and constitutes a standardized roughness measurement. The commonly units are millimeters per meter (mm/m) or meters per kilometer (m/km).

The index measures pavement roughness in terms of the number of millimeter per meter that a laser, mounted on a standard quarter car, should jump as the van is driven along the roadway.



TRAFFIC AND ENVIRONMENT



Information on traffic flows have been collected in terms of:

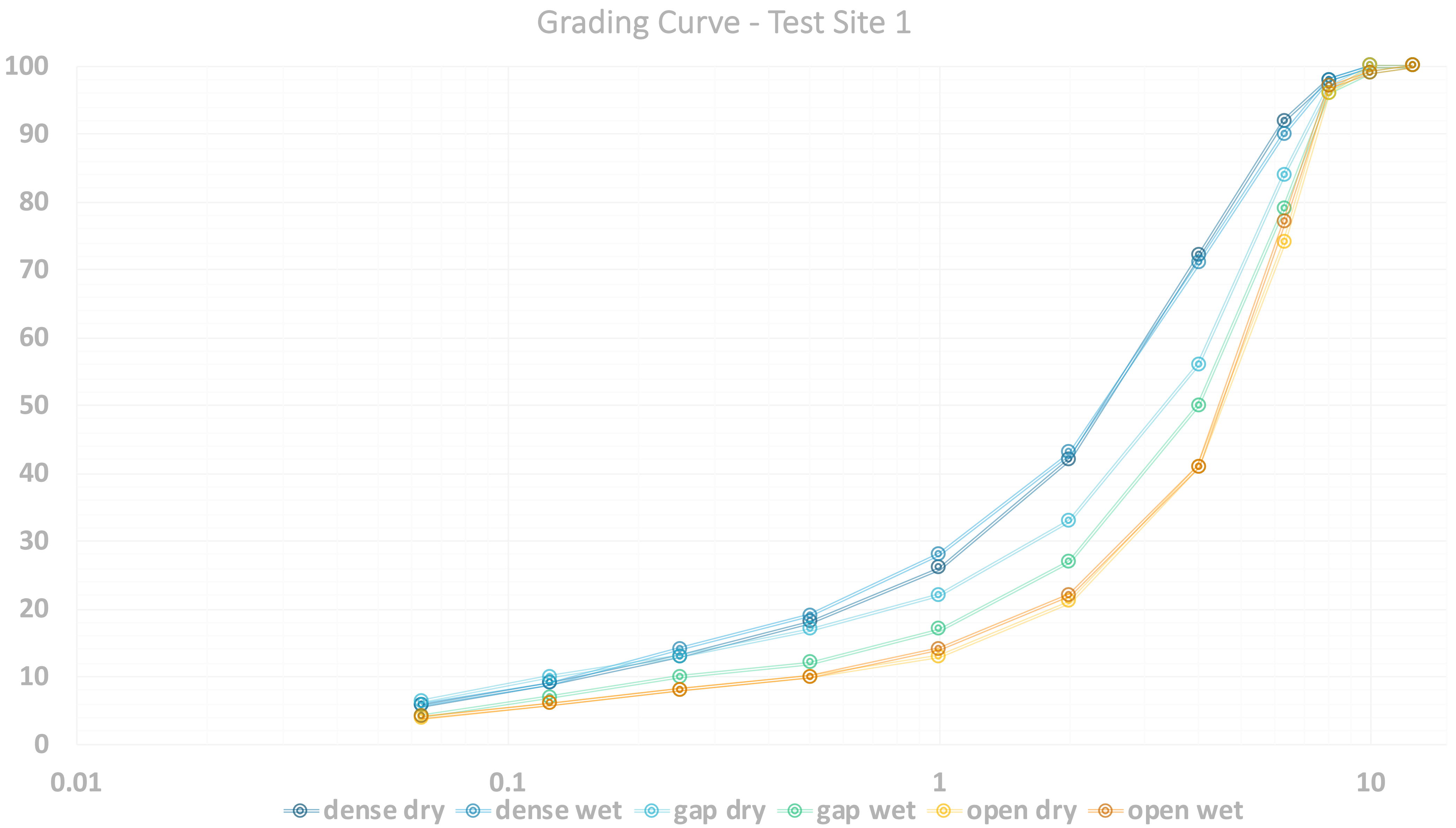
- Average Annual Daily Light Traffic
- Average Annual Daily Heavy Traffic

The following environmental parameters, related to the external temperature and rainfalls, have been collected:

- Average Yearly Maximum Temperature
- Average Yearly Minimum Temperature
- Average Yearly Temperature
- Extreme Maximum Temperature
- Extreme Minimum Temperature
- Average Yearly Freezing Days

- Average Yearly Rainy Days
- Average Cumulative Yearly Rainfall
- Average Yearly Continuous Rainy Days

GRADING CURVES



WHY MACHINE LEARNING?

MODELING THROUGH MACHINE LEARNING ALGORITHMS HAVE SIGNIFICANT STRENGTHS:

01
To account for high non-linear relations between conditioning factors and response

02
To avoid parametric-based models. We have not to assume specific distributions for the variables

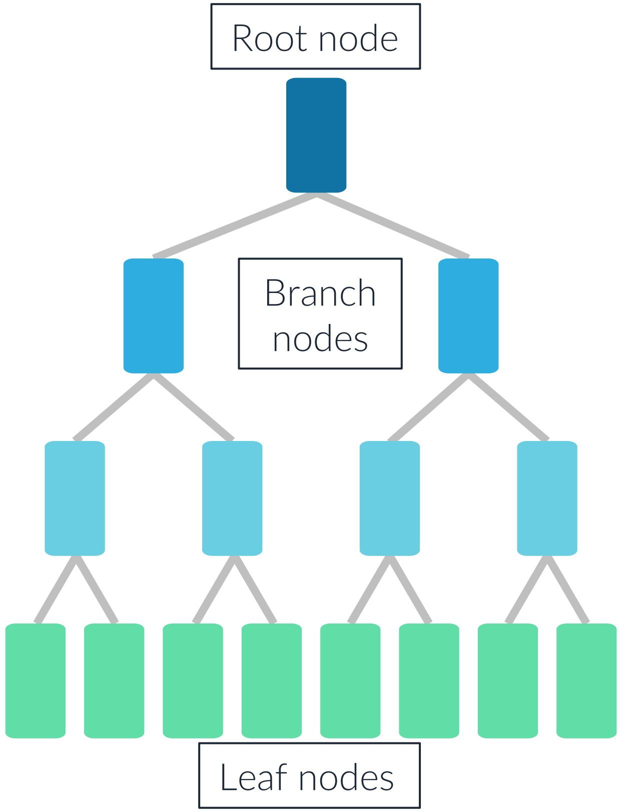
03
To merge a large number of variables (or factors), of different types, that come from different sources

04
To compute the importance of each input factor that affects the target response

05
To exploit Cross-Validation and the test set for evaluating how much a model can generalize on unknown new data

MACHINE LEARING ALGORITHMS (1/2)

DECISION TREE



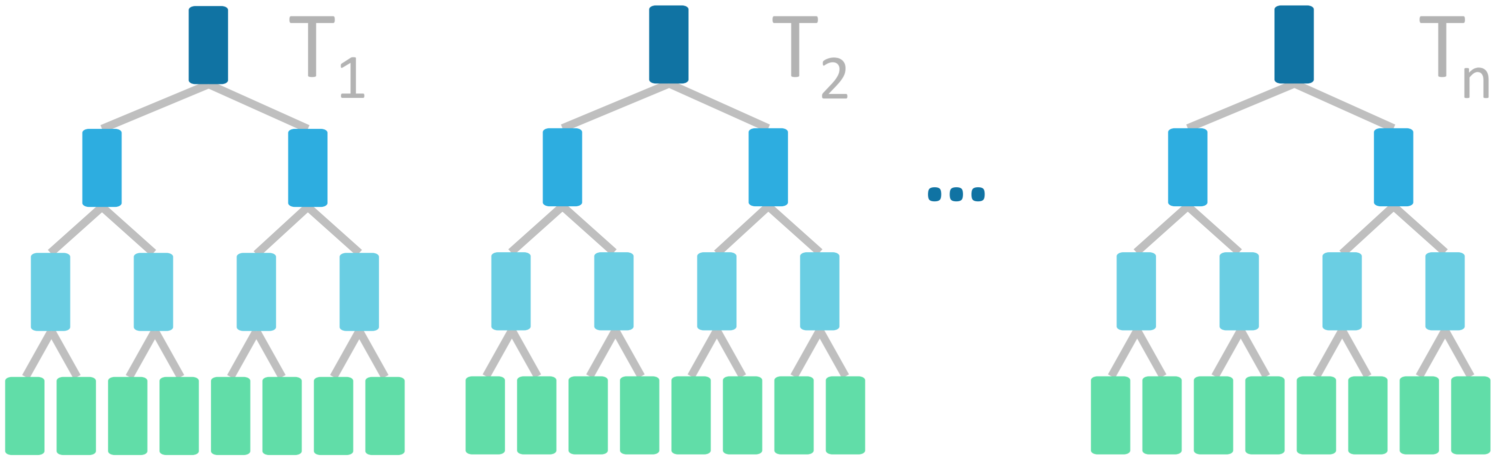
n leaf nodes = n possible predictable values of the target response

Decision Tree is a hierarchical, non-parametric, supervised approach that grows a tree-based model by repeatedly splitting the dataset into homogeneous zones. The decision rules for splitting each node (from the root node to the leaf nodes) are learned by inferring directly from the available data. The Recursive Partitioning algorithm is used for defining the decision rules. Once the Decision Tree is trained, the decision rules can be used for predicting the class of new unknown observations.

- They provide an interpretable solution of the predictions by a tree-graph visualization;
- They provide an automatic variable selection making them insensitive to irrelevant variables, outliers, and the scales of predictors;
- They are computationally efficient even in large problems (generally, low time for training is required)
- They allow handling missing values of the input factors, and both numerical and categorical predictors;
- Their outcomes are unaffected by monotone transformations of the input factors.

MACHINE LEARING ALGORITHMS (2/2)

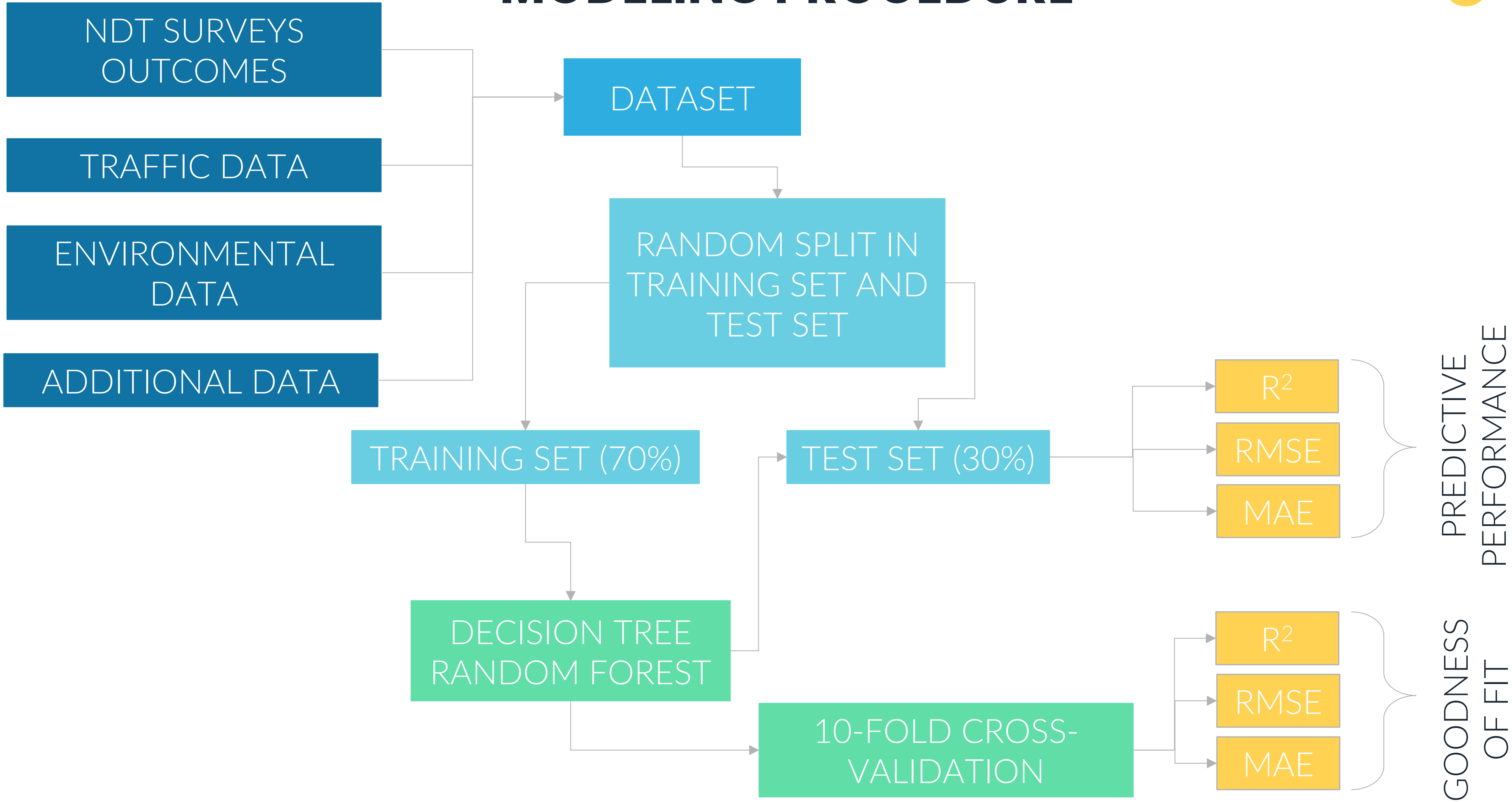
RANDOM FOREST



Random Forest is an Ensemble of Learners. It can be used both for regression and classification task, depending on the target response. The base learner is a Decision Tree for regression. Exploiting the Bootstrap Aggregation (Bagging) and the Feature Randomness, the Random Forest grows a huge sample of n uncorrelated Trees. When a new unknown observation passes through the Random forest, each of the n Trees works alone, thus providing an amount of n different predictions. The arithmetic mean of all the predictions gives the final prediction of the Random Forest.

- The Random Forest algorithm has all the advantages of the Decision Tree, but it allows avoiding also overfitting issues.

MODELING PROCEDURE



RANGE OF VARIATION (1/4)

FWD AND GPR

FEATURE	MINIMUM	MAXIMUM	MEAN	ST.DEV.
Bituminous layers thickness [mm]	132	316	188.42	25.57
Sub-base layer thickness [mm]	188	653	382.53	148.46
Bituminous layers Modulus [MPa]	1061	13050	5746.80	2779.24
Sub-base layer Modulus [Mpa]	85	1405	464.39	215.57
Subgrade Modulus [Mpa]	34	943	313.66	220.12
Stress (FWD)	1189	1786	1400.80	242.28
Load (FWD)	84.05	126.21	99.02	17.12
D1 (FWD geophone 1)	161	1512.40	517.77	272.11
…	…	…	…	…
D9 (FWD geophone 9)	2	172.20	58.97	42.24

RANGE OF VARIATION (2/4)

ENVIRONMENTAL INFORMATION

FEATURE	MINIMUM	MAXIMUM	MEAN	ST.DEV.
Average Yearly Maximum Temperature [°C]	19.33	20.81	19.89	0.61
Average Yearly Minimum Temperature [°C]	9.73	13.01	11.30	1.30
Average Yearly Temperature [°C]	14.84	16.25	15.76	0.55
Extreme Maximum Temperature [°C]	35.90	41	38.77	2.11
Extreme Minimum Temperature [°C]	-9.4	-3.4	-6.8	2.5
Average Yearly Freezing Days [days]	1.25	18.82	11.22	7.29
Average Yearly Rainy Days [days]	104	115.80	122.50	7.48
Average Cumulative Yearly Rainfall [mm]	792	1444	1033	297
Average Yearly Continuous Rainy Days [days]	25.6	28.2	26.2	0.9

RANGE OF VARIATION (3/4)

TRAFFIC FLOW, SKIDDOMETER BV11, AND LASER PROFILER

FEATURE	MINIMUM	MAXIMUM	MEAN	ST.DEV.
Average Annual Daily Light Traffic [veic/day]	4488	13632	10695	3531
Average Annual Daily Heavy Traffic [veic/day]	464	930	732	156
Estimated Texture Depth [mm]	0.38	1.49	0.69	0.12
Braking Force Coefficient	0.53	0.92	0.72	0.05
IRI [mm/km]	1.5	5.7	2.6	0.69

RANGE OF VARIATION (4/4)

CATEGORICAL VARIABLES

FEATURE	CLASSES
Type of pavement	Flexible, semirigid
Type of asphalt mixture	Open graded, dense graded, gap graded
Use of Reclaimed Asphalt Pavement (RAP)	Yes, no
Use of Crumb Rubber (CR)	Yes, no
Type of CR process	Dry, wet

ADDITIONAL DATA

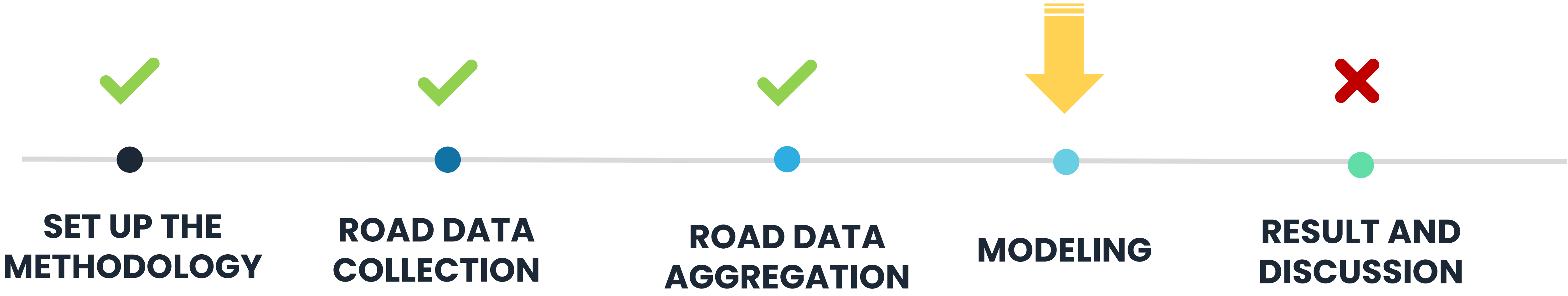
The Grading curves of the asphalt mixtures, from UNI Sieve 0.063 mm to UNI Sieve 12.5 mm are available (11 additional independent variables);

The volumetric mix-design of the asphalt mixtures are available (10 additional independent variables);

Moreover, Indirect Tensile Strain (ITS) test outcome, in dry and wet condition, is available (2 additional independent variables);

Therefore, in order to predict IRI (target response), the algorithms should consider 55 numerical factors and 5 categorical factors.

MILESTONES



Due to the COVID-19, we have not been able to complete the calibration and validation of the models for presenting the findings at the EGU2020.

We apologize to the readers, and we hope to show you all the findings of this research somewhere, as soon as possible.

Evaluating Resilience of Infrastructures Towards Endogenous Events by Non-Destructive High-Performance Techniques and Machine Learning Regression Algorithms

Thank You!

Nicholas Fiorentini, Pietro Leandri, and Massimo Losa

Department of Civil and Industrial Engineering, Engineering School of the University of Pisa,

Largo Lucio Lazzarino, 1, 56122, Pisa, Italy

nicholas.fiorentini@phd.unipi.it

pietro.leandri@ing.unipi.it

losa@ing.unipi.it