PREDICTIVE SCHEMES IN GEOHAZARDS MANAGEMENT OF HYDROCLIMATOLOGICAL ORIGIN

Jaime Aristizábal Carlos Motta
Carlos Capachero Nelson Obregón Leonardo Real

cenit
TRANSPORTE Y LOGÍSTICA DE HIDROCARBUROS

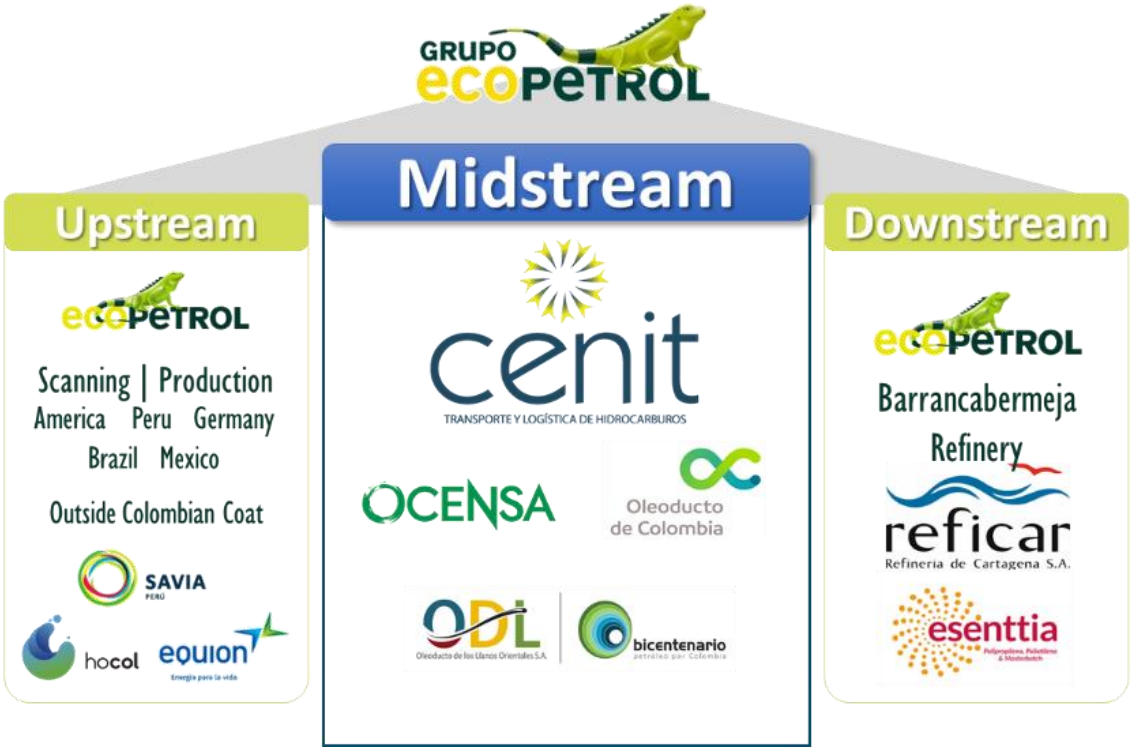
canalclima
entiende tu mundo

"We naturally connect the energy opportunities of Colombia towards the future"

AGENDA

- Introduction
- Cenit's Zoning – an overview
- Methodology
- Conclusion

INTRODUCTION



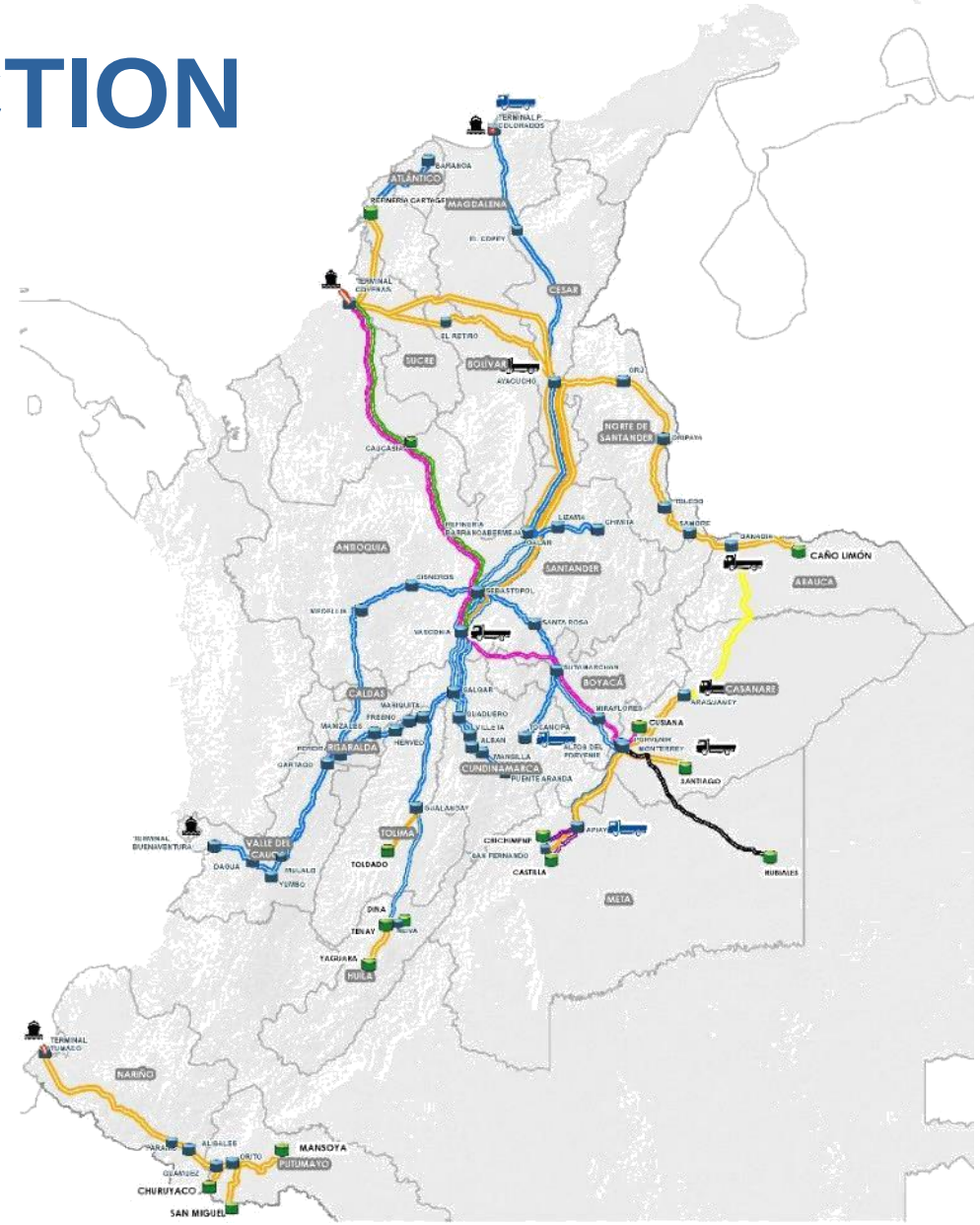
Cenit is part of Ecopetrol Business Group

Guarantee the industry chaining as the key link between production, exportation, refining and commercialization

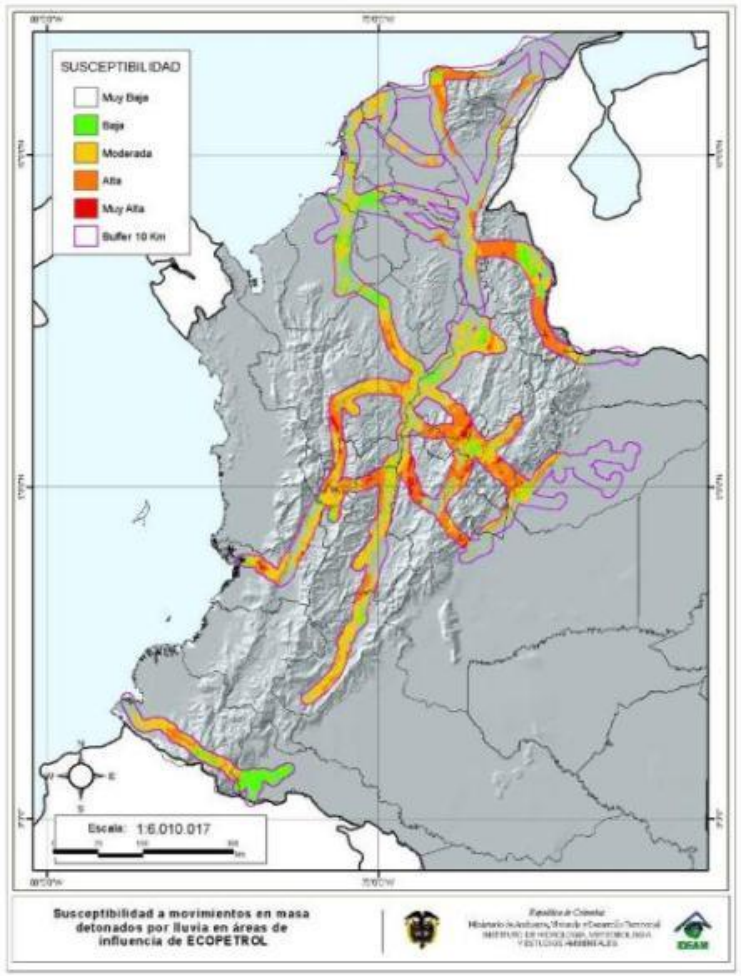
INTRODUCTION

Cenit network composition:

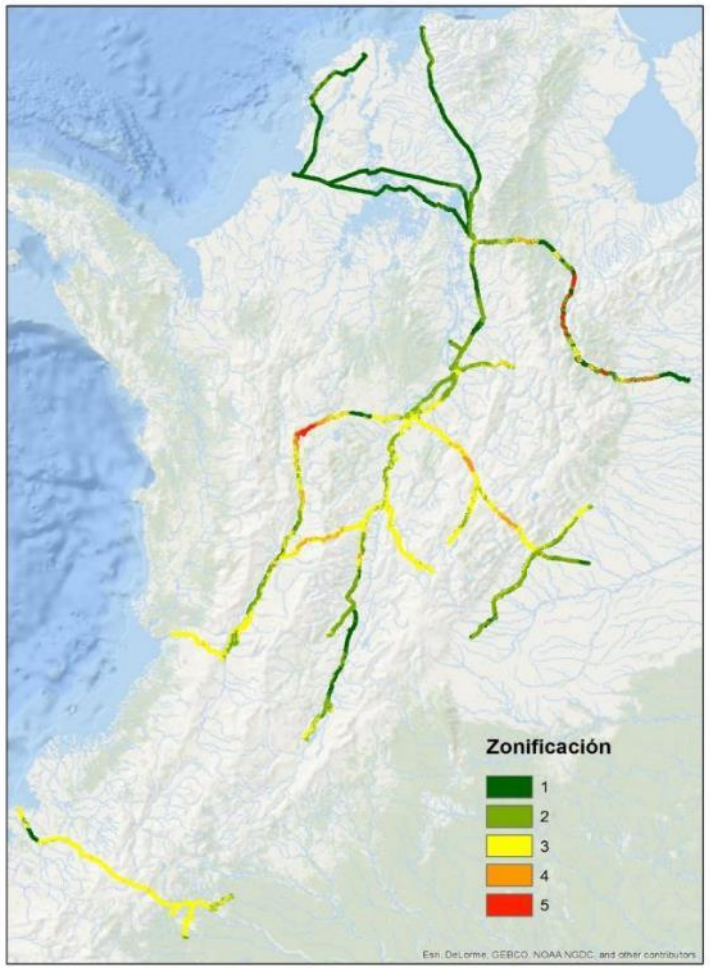
- 8500 km of pipelines
- 54 Pump stations
- 4 Oil Terminals
- 7 Oil Loading and Unloading facilities



Geotechnical Zoning in Cenit



2010 (%)	Susceptibility	2021 (%)
51.2	Very Low or Low	56.2
29.2	Moderated	36.5
18.3	High	5.2
1.3	Very High	2.2



CENIT´S ZONING – AN OVERVIEW

2014
1st Cenit Zoning

Geotechnical Zoning

2016
2d Zoning

Geotechnical + Climate Zoning

2018-2019
3rd Zoning

Implementation of an unsupervised machine learning algorithm

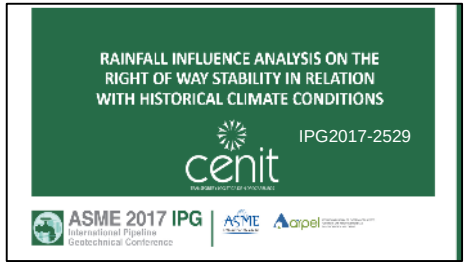
K-means algorithm

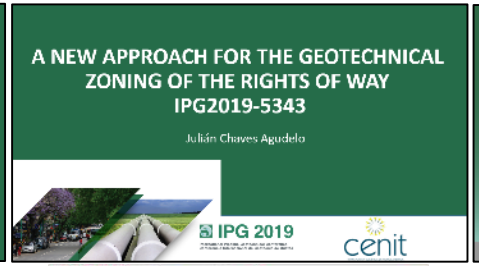
2021
Adaptative Zoning

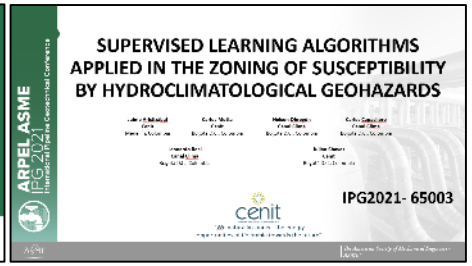
Adaptive geotechnical susceptibility zoning using supervised learning algorithms

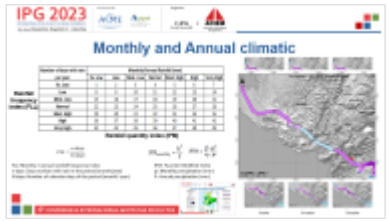
2023
Predictive Schemes

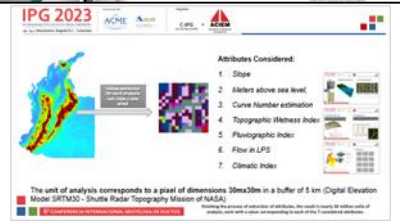
Adaptive geotechnical susceptibility zoning

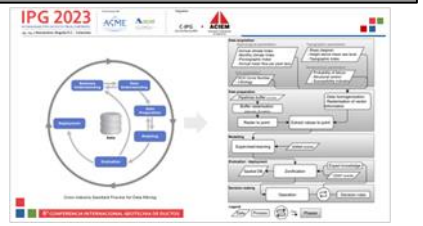












Monthly and Annual climatic

Rainfall frequency index (FLL)

Number of days with rain per year	Monthly/Annual Rainfall (mm)						
	Ex. Low	Low	Mod. Low	Normal	Mod. High	High	Very High
Ex. Low	1	2	3	4	5	6	7
Low	8	9	10	11	12	13	14
Mod. Low	15	16	17	18	19	20	21
Normal	22	23	24	25	26	27	28
Mod. High	29	30	31	32	33	34	35
High	36	37	38	39	40	41	42
Very High	43	44	45	46	47	48	49

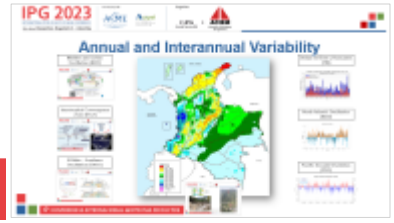
Rainfall quantity index (IFM)

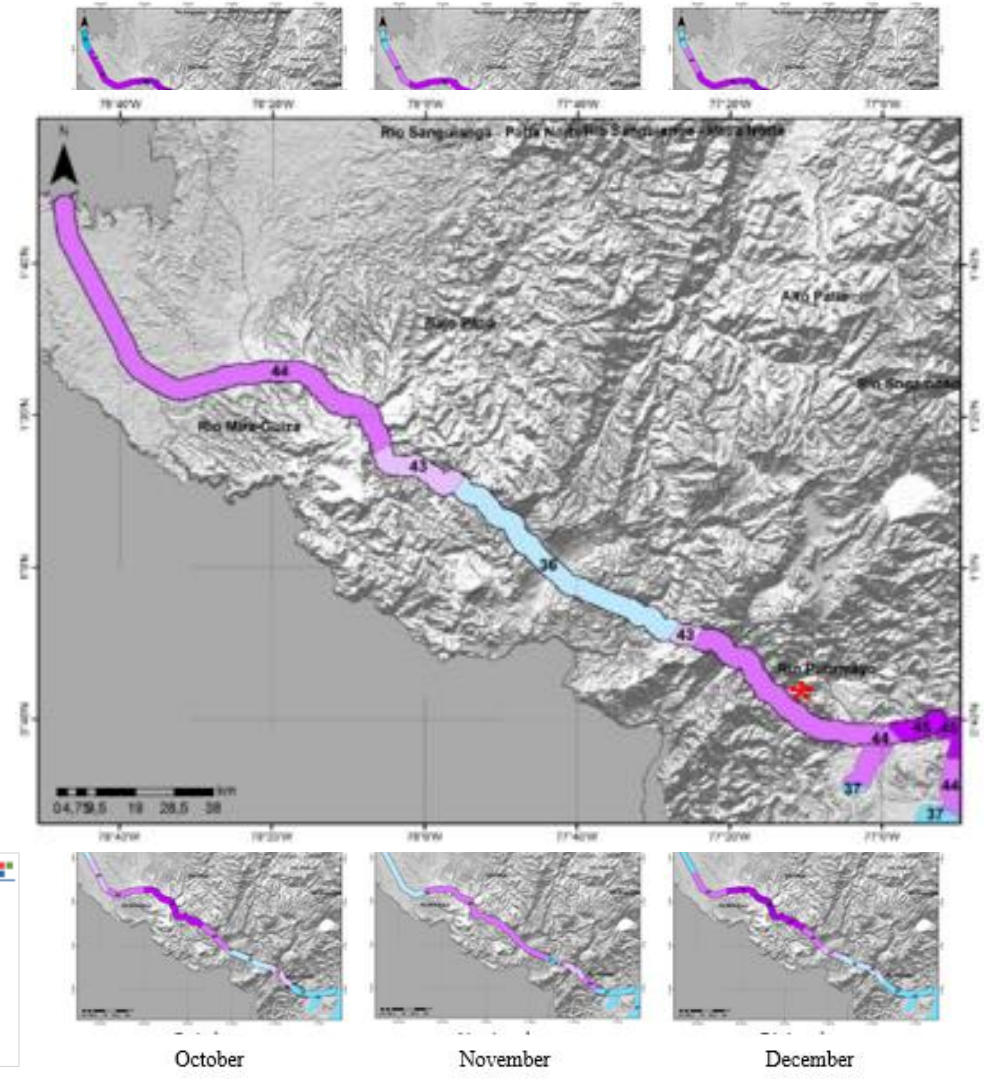
$$FLL = \frac{n\ days}{N\ days}$$

$$IFM_{monthly} = \frac{p_i^{\ 2}}{P} \qquad IFM = \sum_{1}^{12} \frac{p_i^2}{P}$$

FLL: Monthly / annual rainfall frequency index
n days: Days number with rain in the period (month/year)
N days: Number of calendar days of the period (month/ year)

IFM: Fournier Modified Index
pi: Monthly precipitation (mm)
P: Annual precipitation (mm)



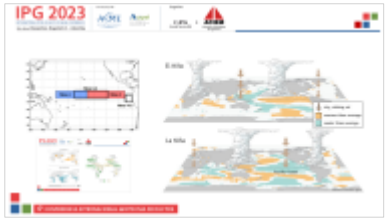


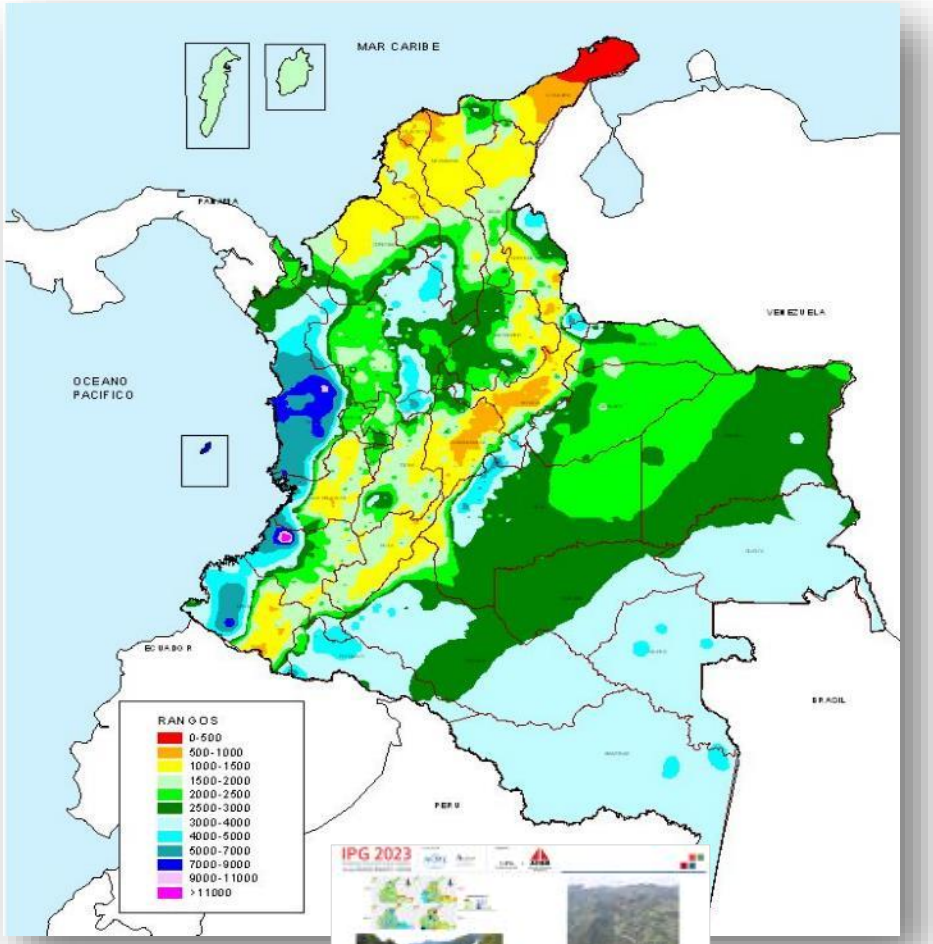
Annual and Interannual Variability

Madden and Julian Oscillation (MJO)

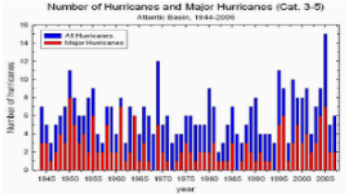
Intertropical Convergence Zone (ITCZ)

El Niño – Southern Oscillation (ENSO)

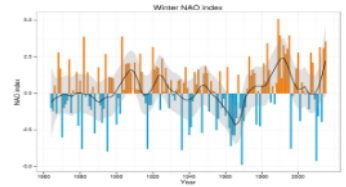




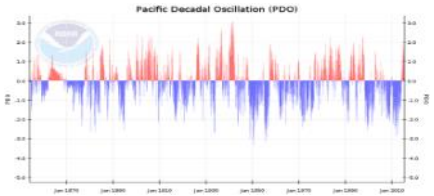
Ondas del Este y Huracanes (TW)

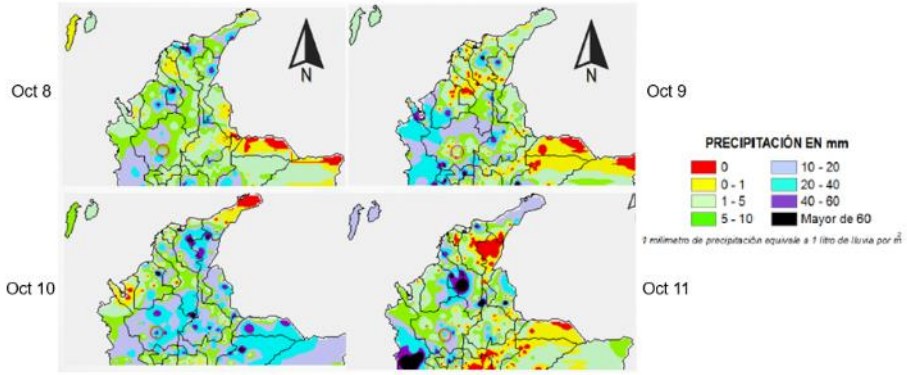


North Atlantic Oscillation (NAO)



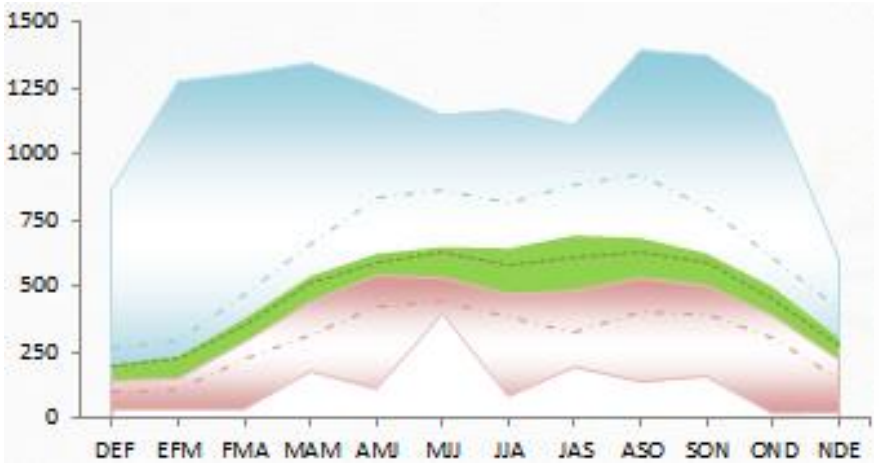
Pacific Decadal Oscilation (PDO)

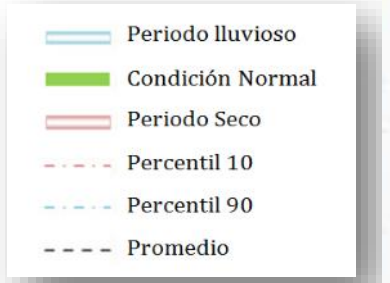


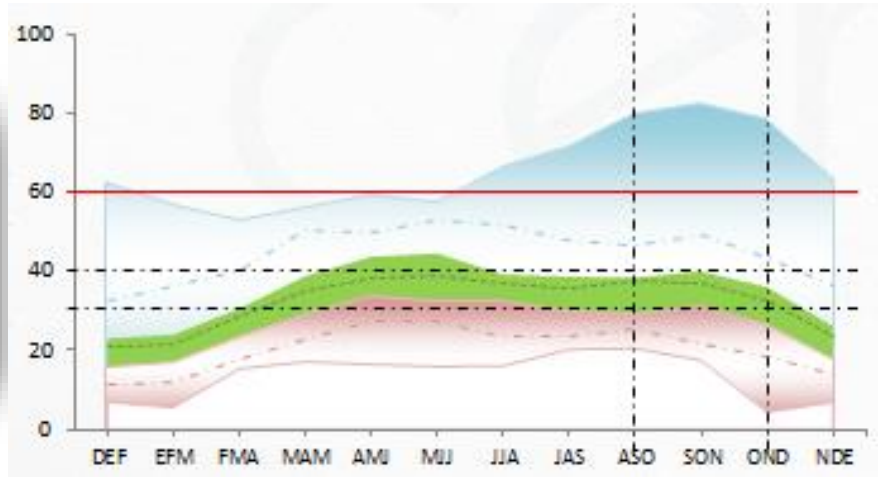


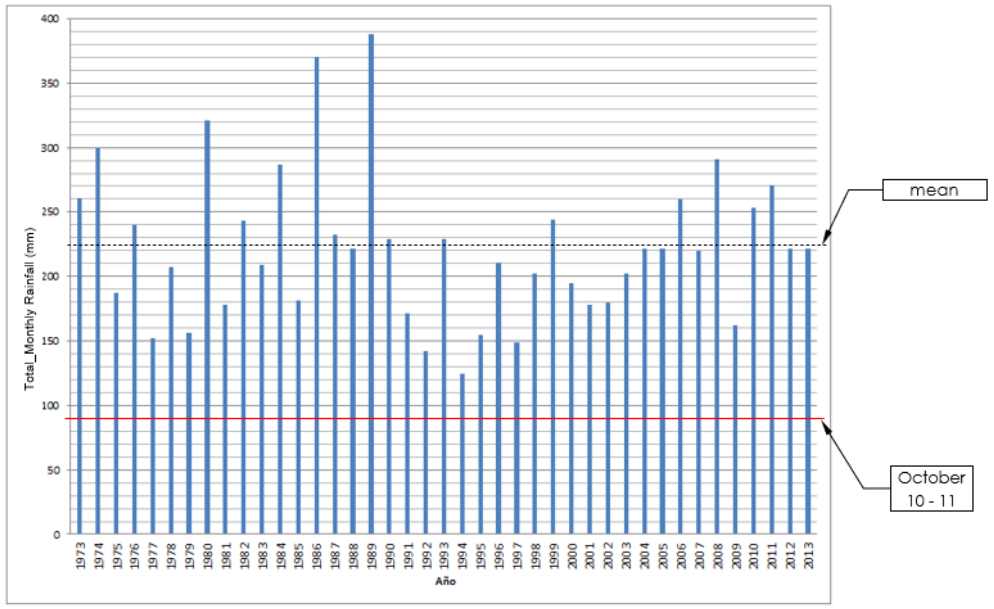
Avalanche´s deposit details area and its effect in pipeline integrity

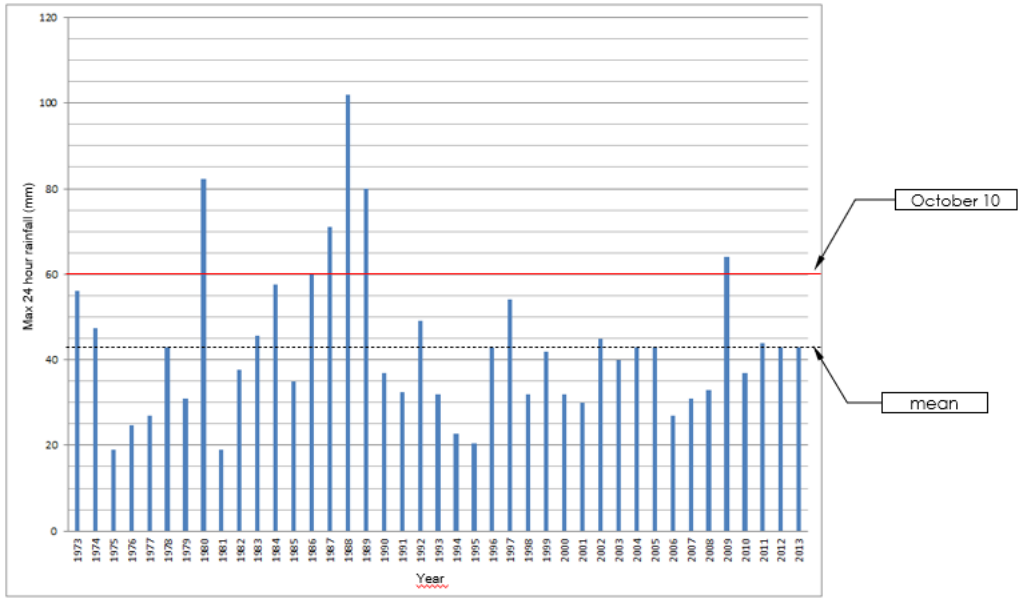
Avalanche crossing ROW Note the changes in land use


1500
1250
1000
750
500
250
0
DEF EFM FMA MAM AMJ MJJ JJA JAS ASO SON OND NDE


Periodo lluvioso
Condición Normal
Periodo Seco
Percentil 10
Percentil 90
Promedio


100
80
60
40
20
0
DEF EFM FMA MAM AMJ MJJ JJA JAS ASO SON OND NDE


Total_Monthly Rainfall (mm)
Año
mean
October
10 - 11


Max 24 hour rainfall (mm)
Year
October 10
mean

CENIT´S ZONING – AN OVERVIEW

2014
1st Cenit Zoning

Geotechnical Zoning

2016
2d Zoning

Geotechnical + Climate Zoning

2018-2019
3rd Zoning

Implementation of an unsupervised machine learning algorithm

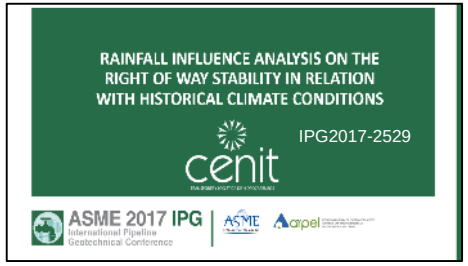
K-means algorithm

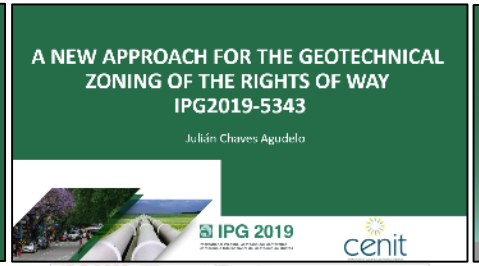
2021
Adaptative Zoning

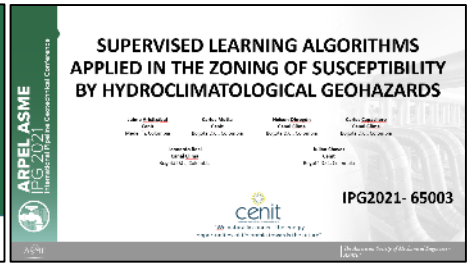
Adaptive geotechnical susceptibility zoning using supervised learning algorithms

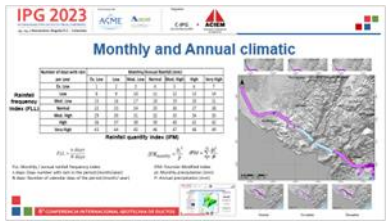
2023
Predictive Schemes

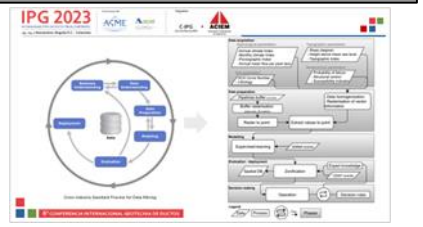
Adaptive geotechnical susceptibility zoning using supervised learning algorithms

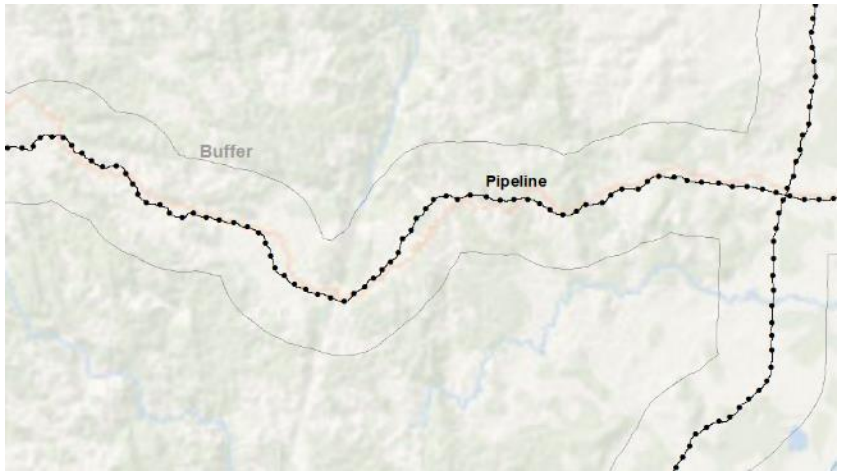


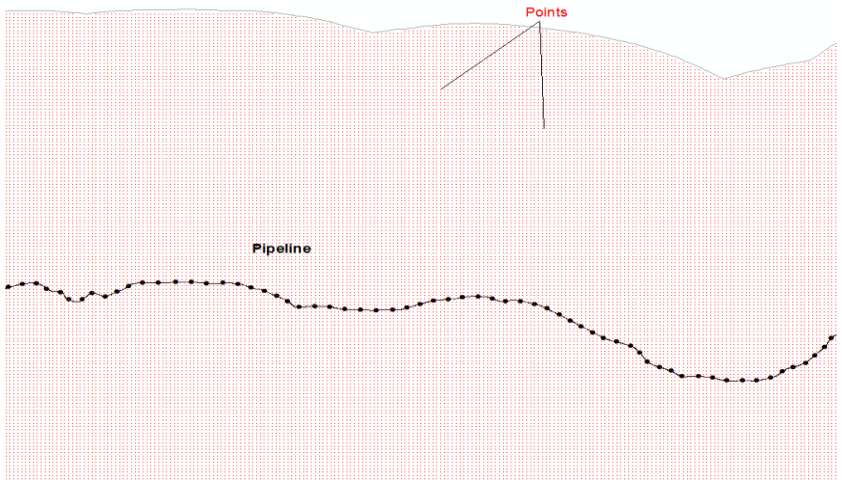






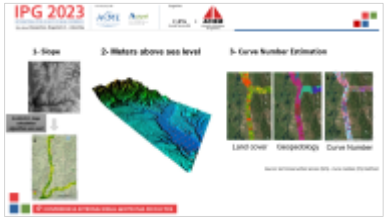


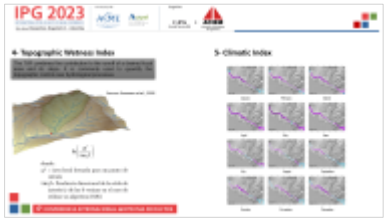


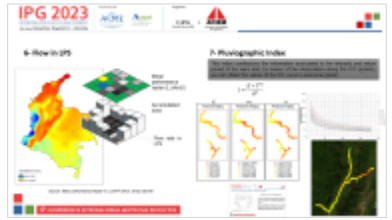


Attributes Considered:

1. *Slope*
2. *Meters above sea level,*
3. *Curve Number estimation*
4. *Topographic Wetness Index*
5. *Climatic Index*
6. *Flow in LPS*
7. *Pluviographic Index*



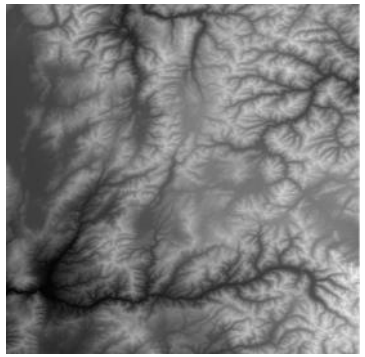




The **unit of analysis corresponds to a pixel of dimensions 30mx30m** in a buffer of 5 km (Digital Elevation Model SRTM30 - Shuttle Radar Topography Mission of NASA)

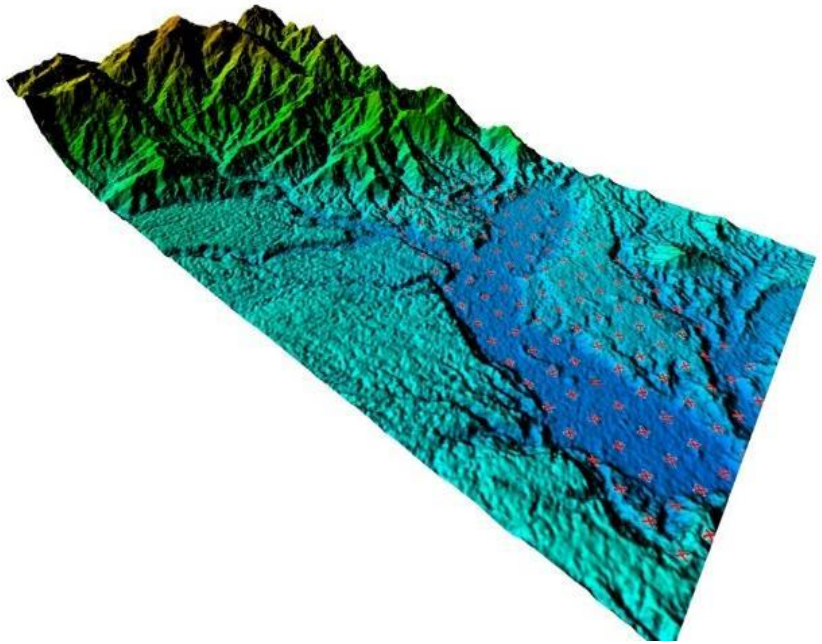
The result is nearly 30 million units of analysis, with values corresponding to each of the 7 considered attributes.

1- Slope

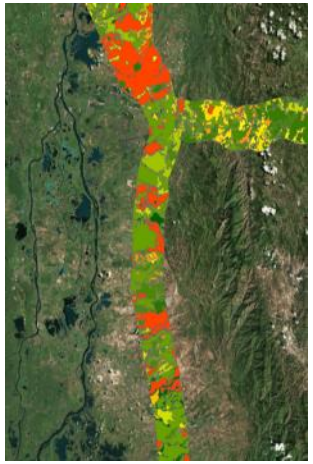


ArcGis10.1 slope calculation algorithm was used

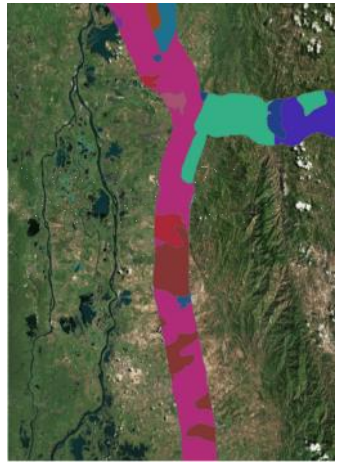
2- Meters above sea level



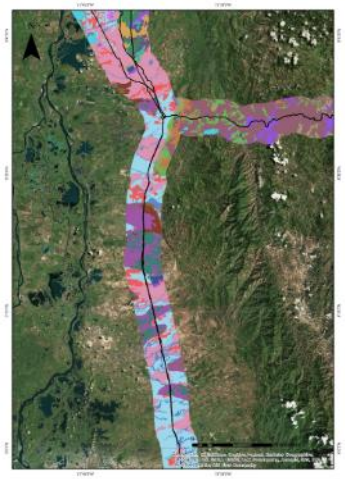
3- Curve Number Estimation



Land cover



Geopedology

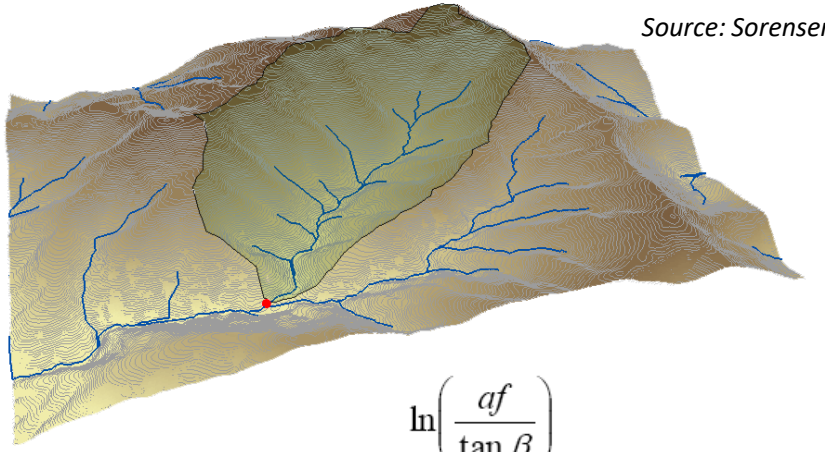


Curve Number

Source: Soil Conservation Service (SCS) - Curve number (CN) method.

4- Topographic Wetness Index

The TWI combines the contribution to the runoff of a drained local area and its slope. It is commonly used to quantify the topographic control over hydrological processes



Source: Sorensen et al., 2005.

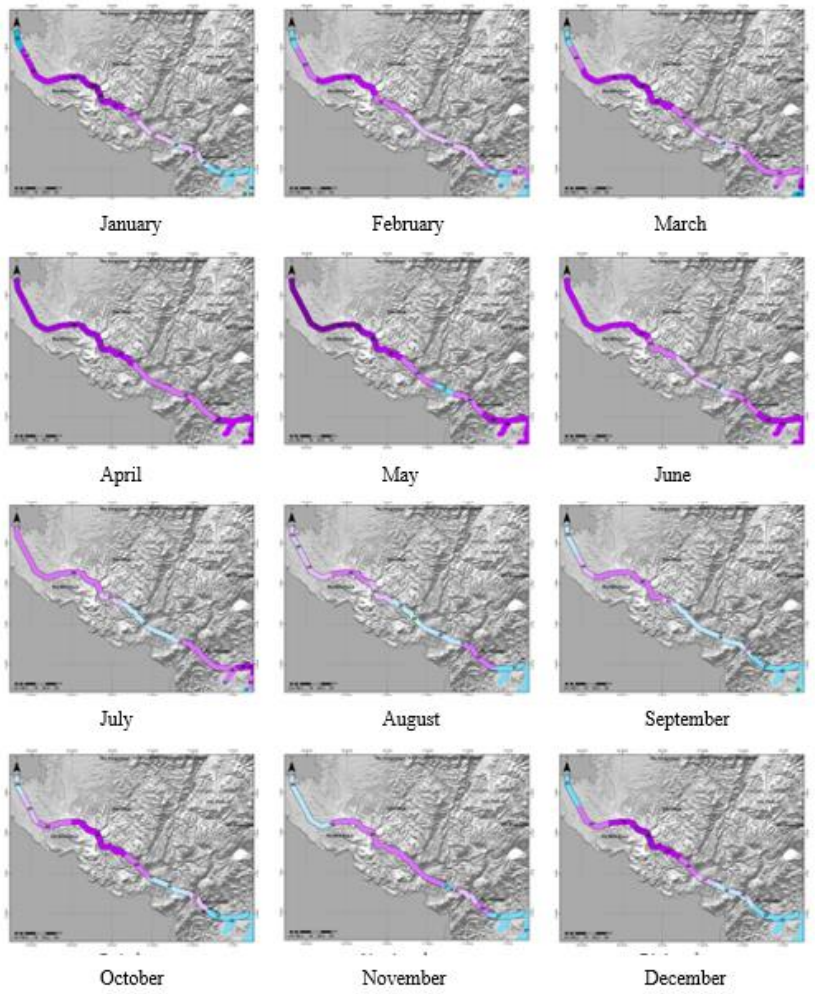
$$\ln\left(\frac{af}{\tan\beta}\right)$$

donde:

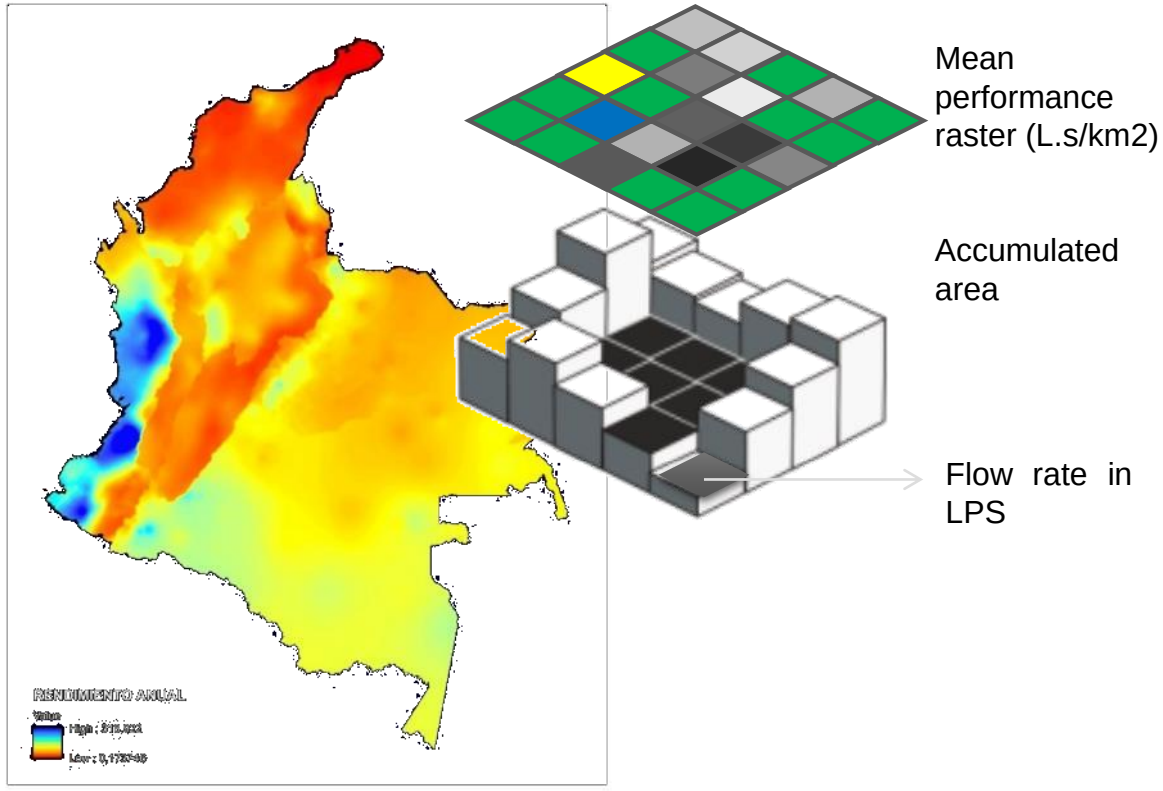
af = área local drenada para un punto de cálculo

$\tan\beta$= Pendiente direccional de la celda de interés (y de las 8 vecinas en el caso de utilizar un algoritmo D(8))

5- Climatic Index



6- Flow in LPS



Mean performance raster (L.s/km2)

Accumulated area

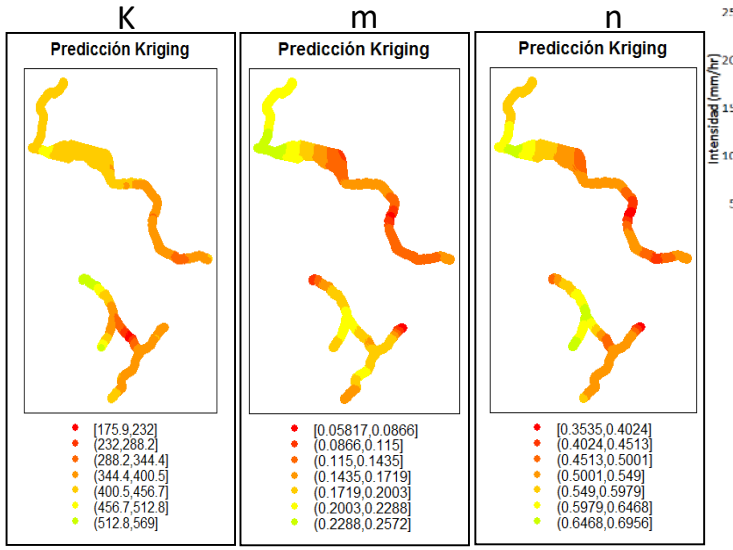
Flow rate in LPS

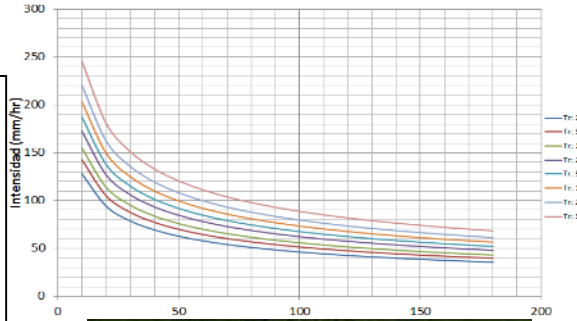
Source: Mean performance Raster in L.s/Km² (ENA, 2010) IDEAM

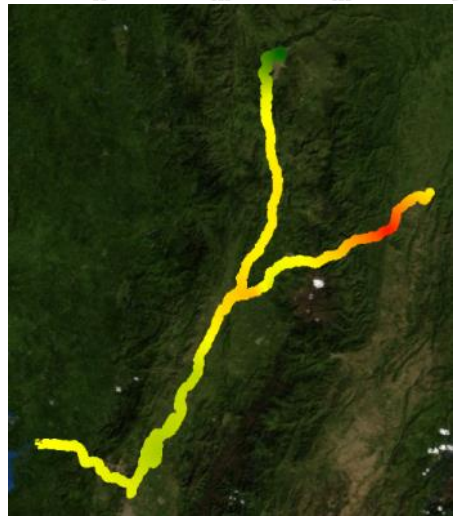
7- Pluviographic Index

This value synthesizes the information associated to the intensity and return period of the rains and, by means of the interpolation along the VIT systems, we can obtain the values of the IDF curve in everyone (pixel).

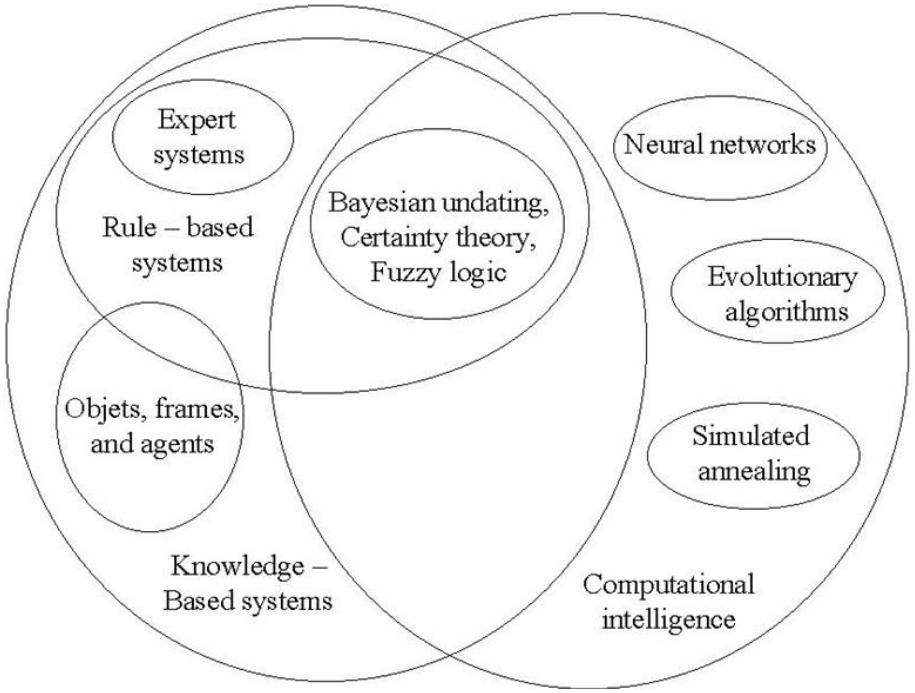
$$I = \frac{K * T^m}{d^n}$$







APPLIED ALGORITHM OF ZONING – Multivariate Analysis



Categories of analysis methodologies
Source: Adrian Hopgood. Intelligent systems for scientist and engineers

Application of the K-means Algorithm

This is an unsupervised machine learning algorithm, by which the units of analysis are grouped in clusters according to the similarity between their attributes.

Once the algorithm is applied, the list of Cenit's geotechnical findings is compared to find the most representative combination of attributes that best represents the existence of geotechnical events.

- 72.5% of the events reported in the influence area are in the categories or cluster 1 and 7.
- When filtering the information sources by type of event, landslides were presented in 71% of the clusters 1 and 7.
- The events reported in clusters 1 and 7 are characterized by their presence in sectors with high curve numbers, moderate slopes 11-30 ° and climatic index between 16-48.

CENIT´S ZONING – AN OVERVIEW

2014
1st Cenit Zoning

Geotechnical Zoning

2016
2d Zoning

Geotechnical + Climate Zoning

2018-2019
3rd Zoning

Implementation of an unsupervised machine learning algorithm

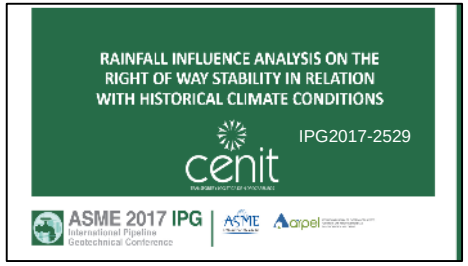
K-means algorithm

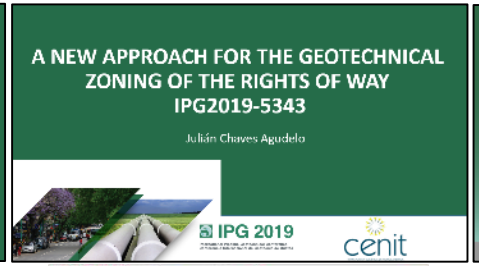
2021
Adaptative Zoning

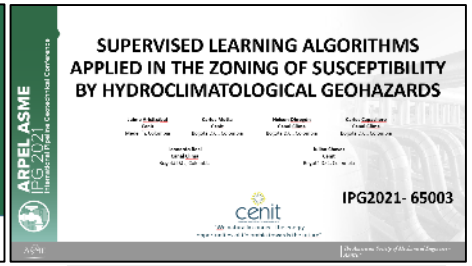
Adaptive geotechnical susceptibility zoning using supervised learning algorithms

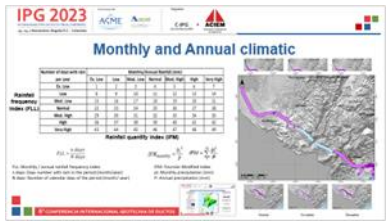
2023
Predictive Schemes

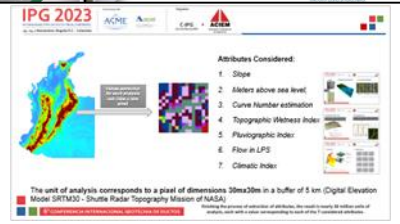
Adaptive geotechnical susceptibility zoning using supervised learning algorithms

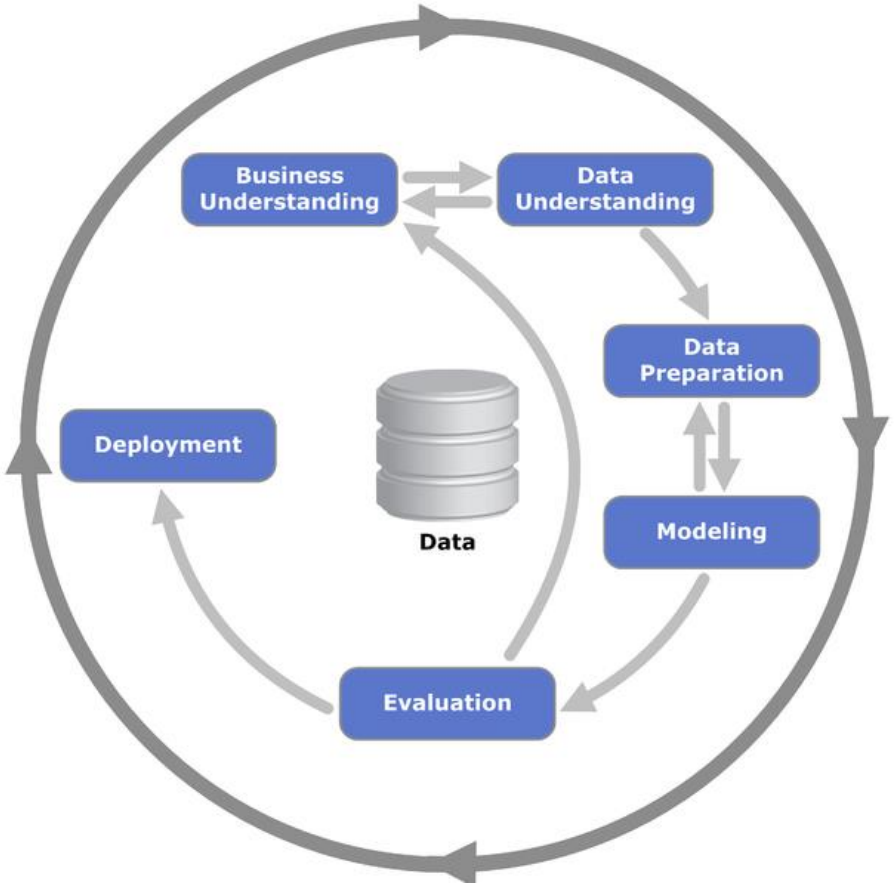




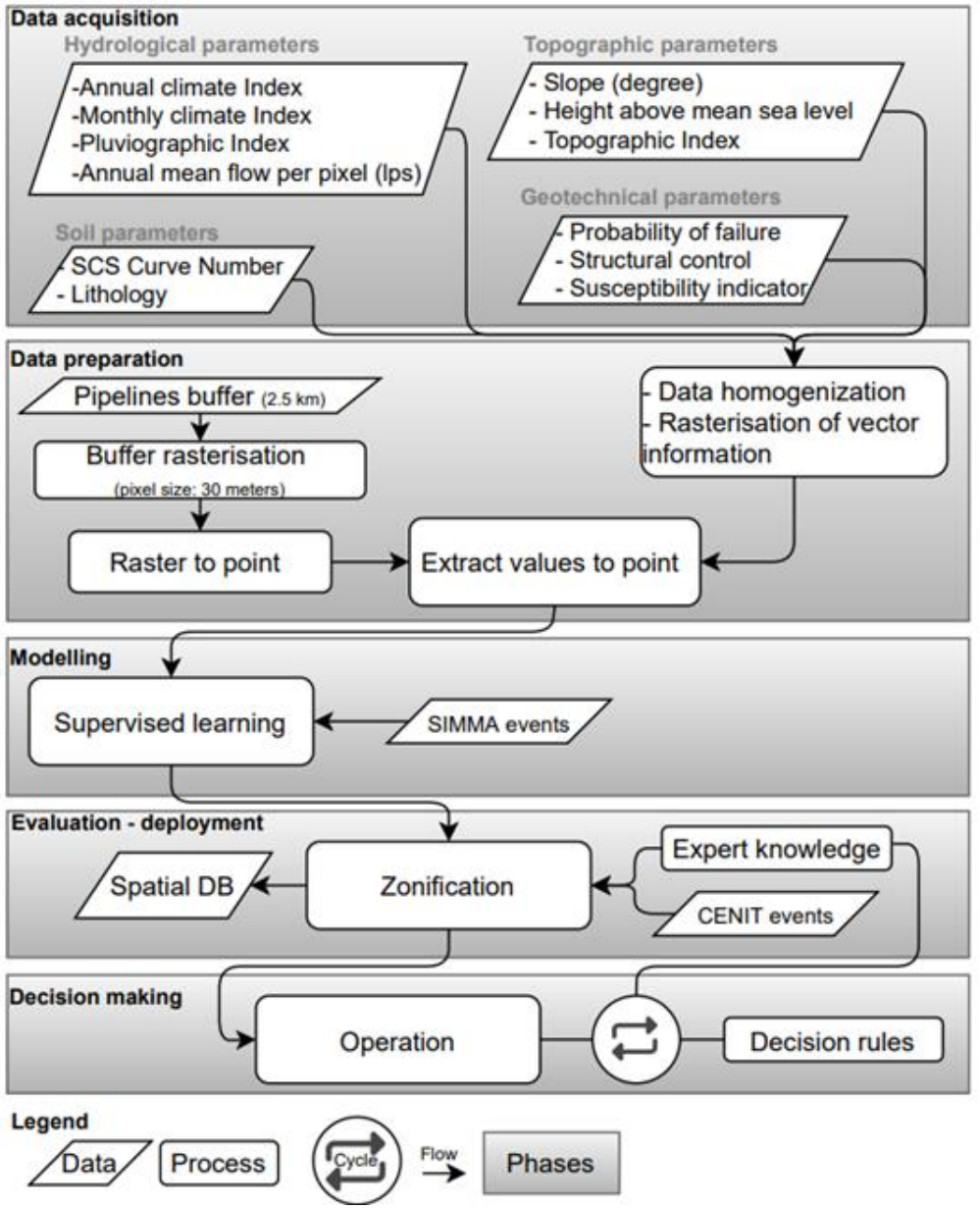


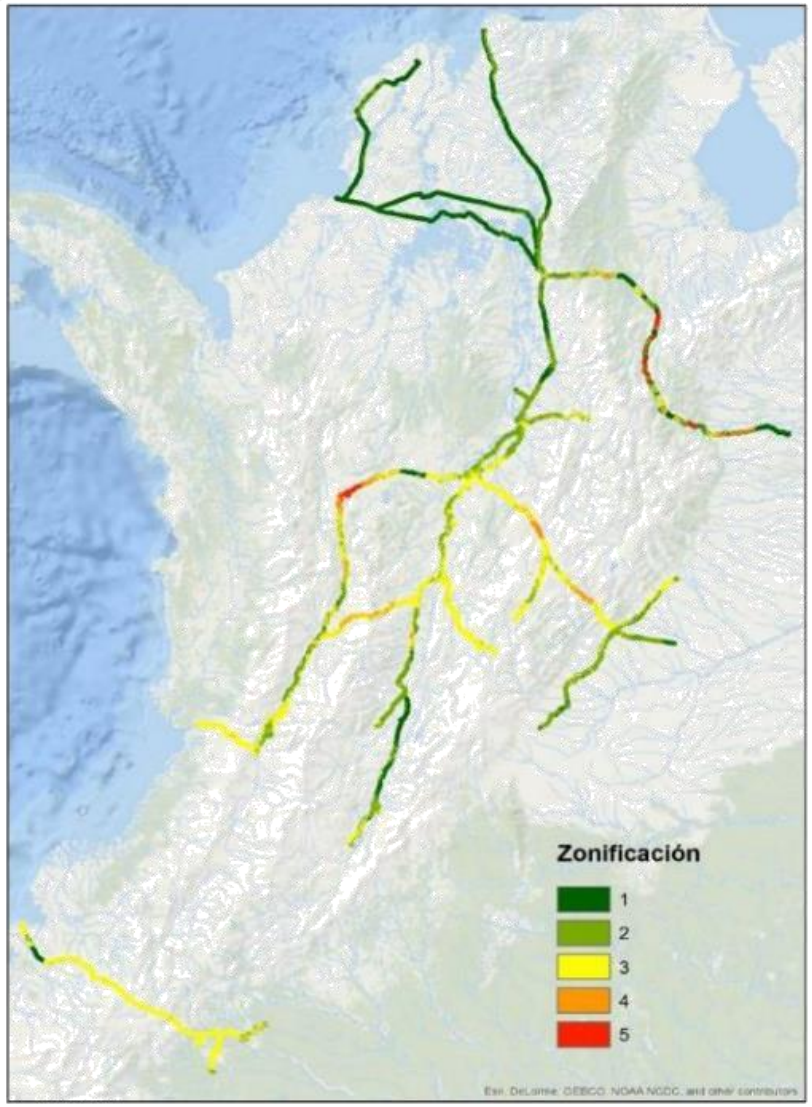






Cross Industry Standard Process for Data Mining



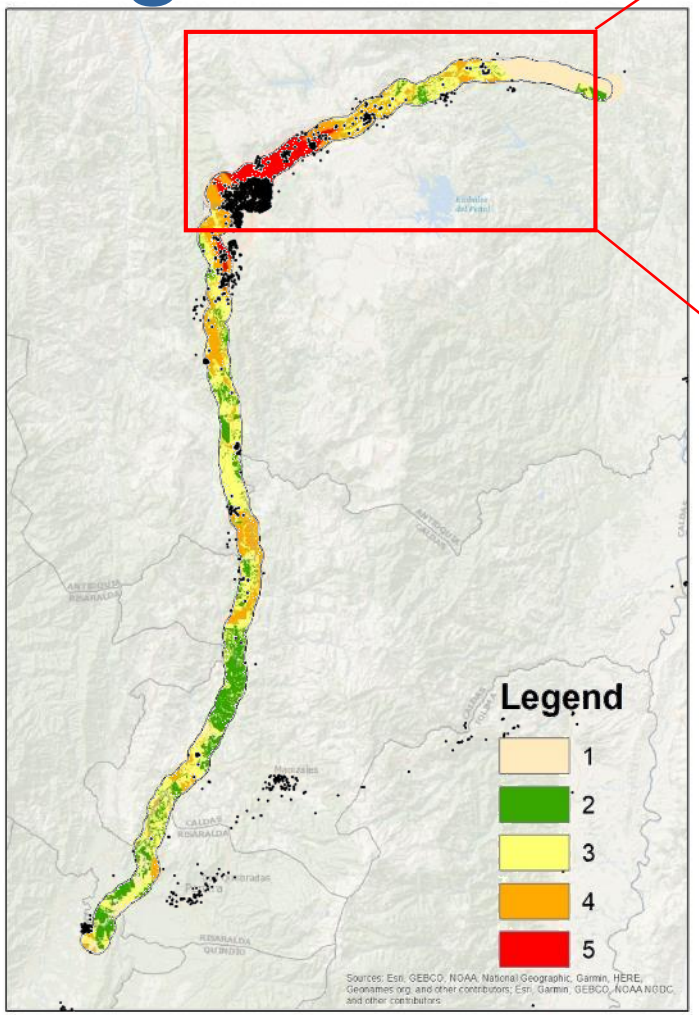


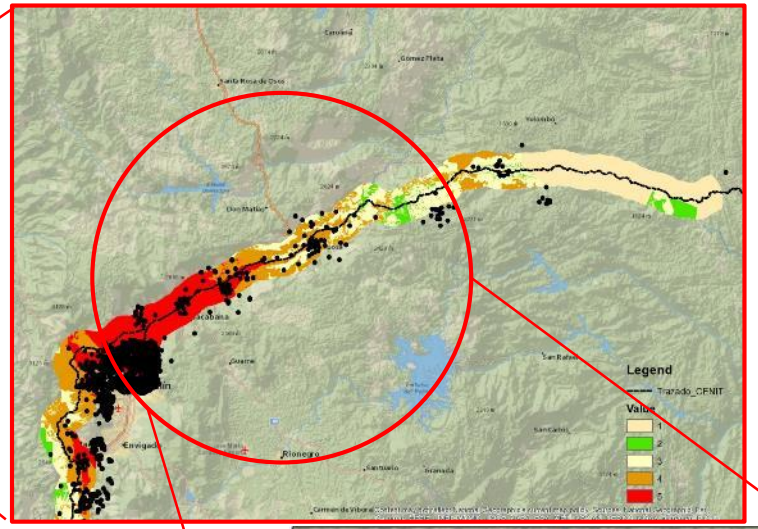
CLUSTER	DESCRIPTION
1	Lamellar erosive processes
2	Severe erosive processes, gullies, furrows
3	High geotechnical susceptibility. some landslides, falls and overturns occur.
4	Erosion in areas with high and very high geotechnical susceptibility, with landslides and falls.
5	Severe processes, occurs in these areas more than 40% of the landslides and flows of the entire ROW.

The methodology implemented for this zoning via fuzzy logic and supervised learning represents a valid and pertinent approach to optimize infrastructure intelligence management processes in pipelines.

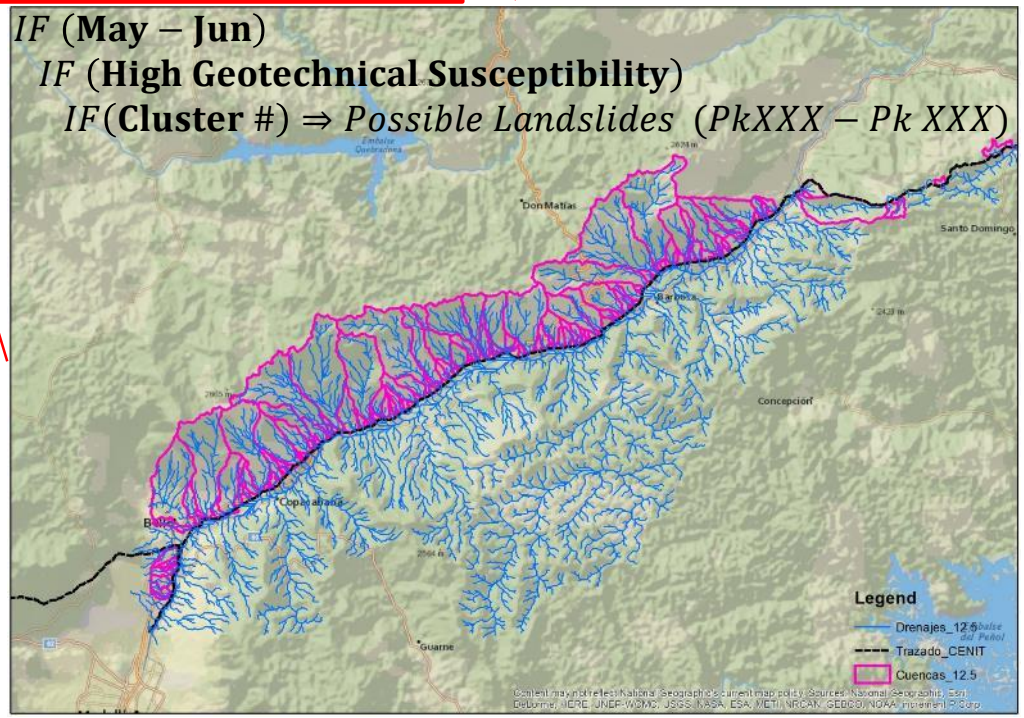
At the same time, it allows to highlight the importance of data-driven-modeling and the design and management of robust information systems that involve descriptors and chronology of potentially damaging events for the analyzed infrastructure.

Rules generation

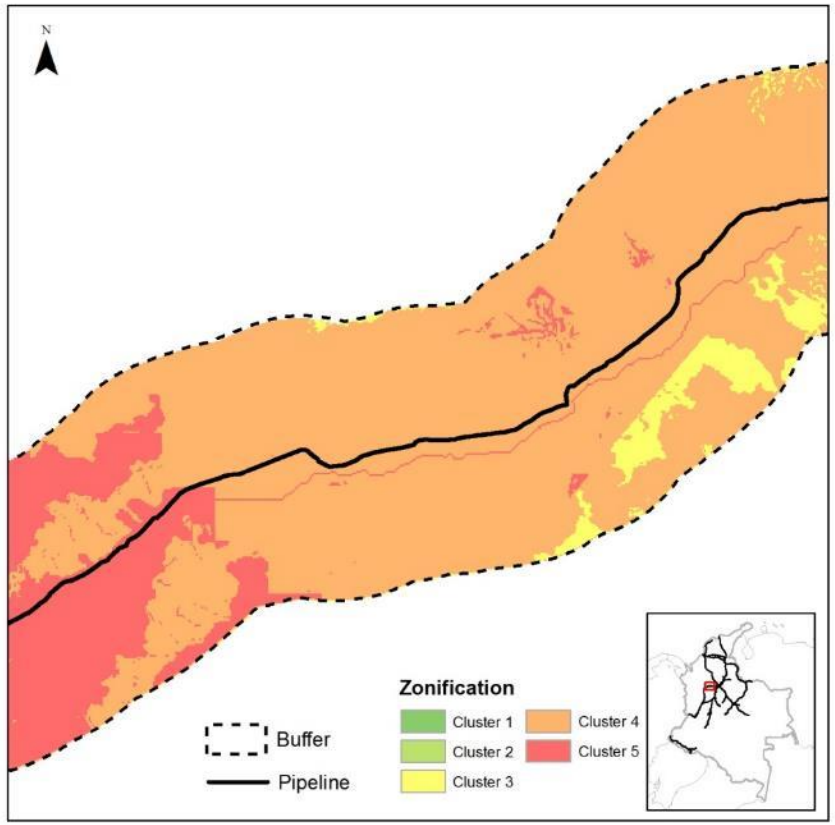


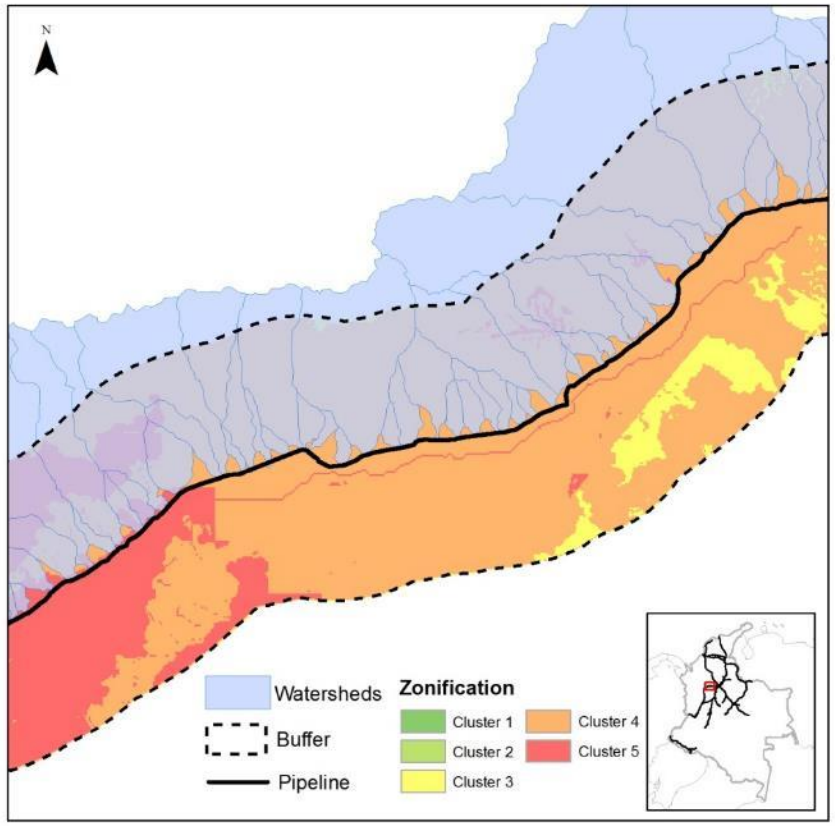


- Cluster Number
- High Geotechnical Susceptibility
- Geotechnical events
- Classification of Findings
- Relevant dates
  - Municipalities
  - Pipelines
- Prioritized areas

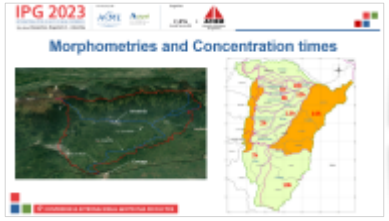


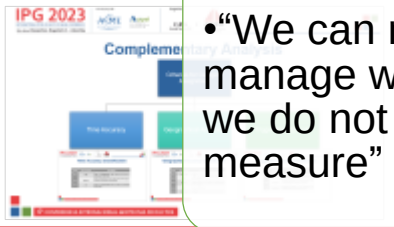
Incorporation of basins in zoning





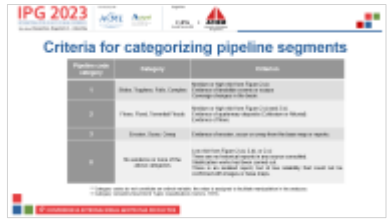
METHODOLOGY

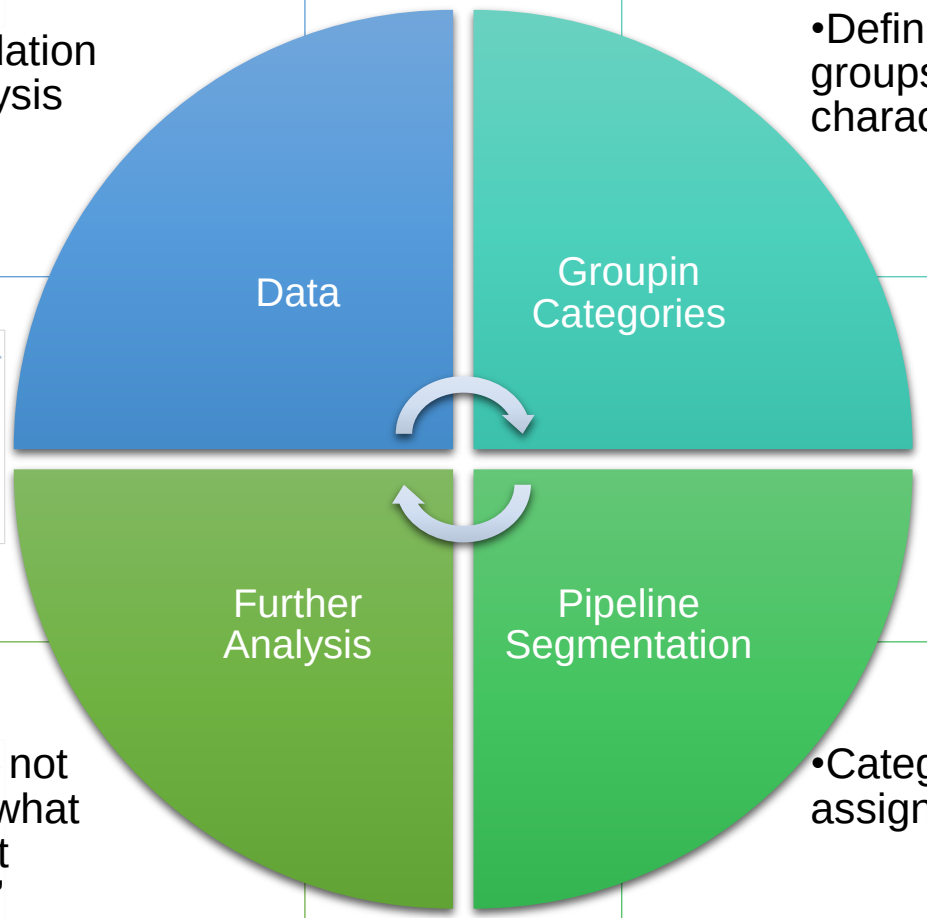




•Consolidation and Analysis

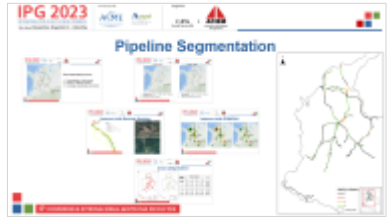
•Definition of groups and characteristics

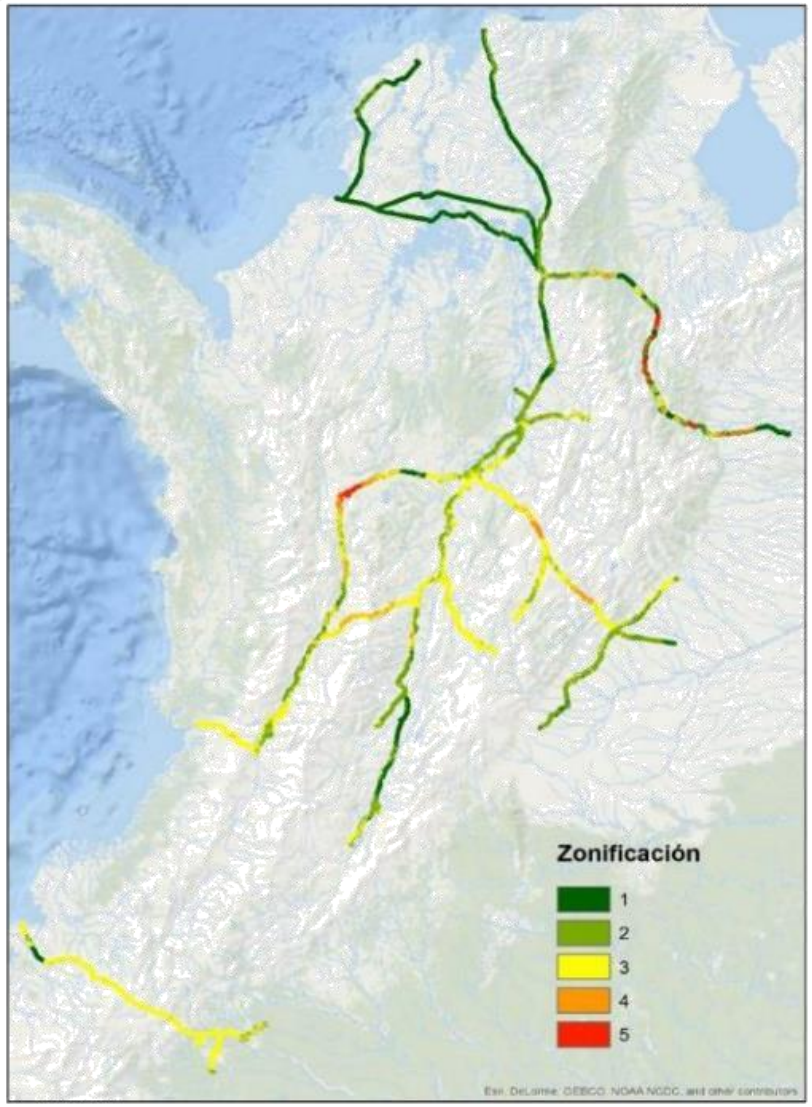




•"We can not manage what we do not measure"

•Category assignment



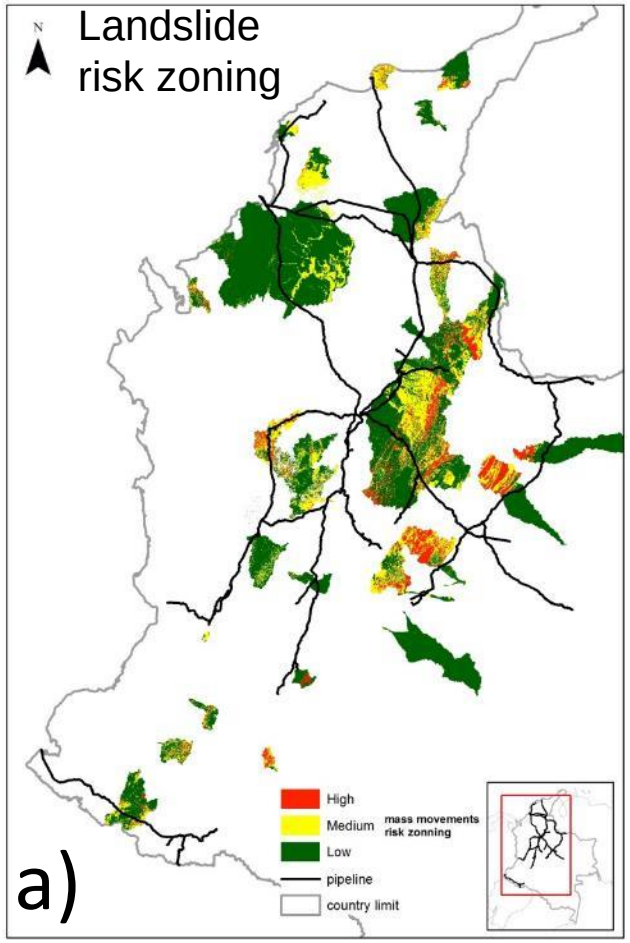


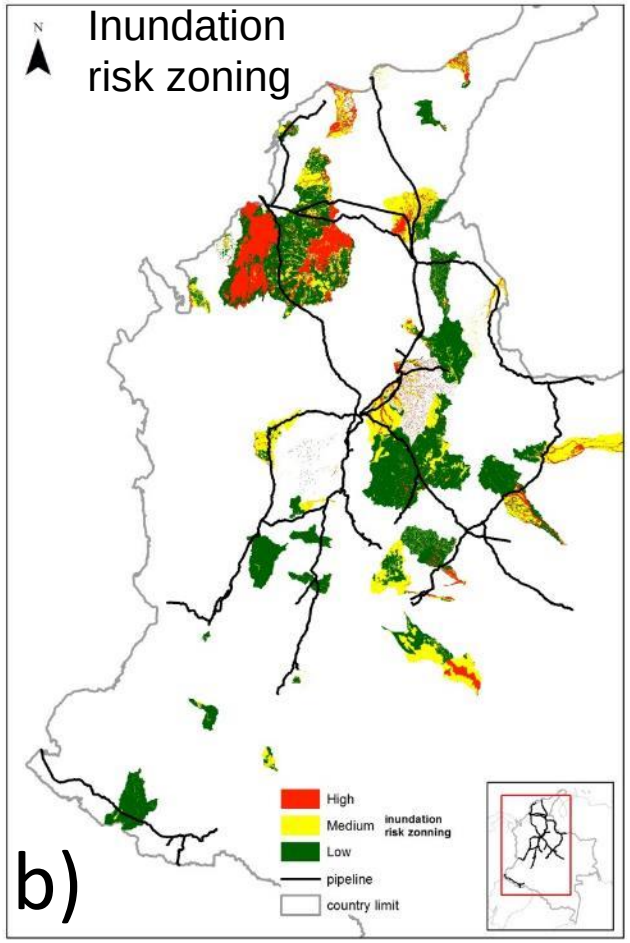
CLUSTER	DESCRIPTION
1	Lamellar erosive processes
2	Severe erosive processes, gullies, furrows
3	High geotechnical susceptibility. some landslides, falls and overturns occur.
4	Erosion in areas with high and very high geotechnical susceptibility, with landslides and falls.
5	Severe processes, occurs in these areas more than 40% of the landslides and flows of the entire ROW.

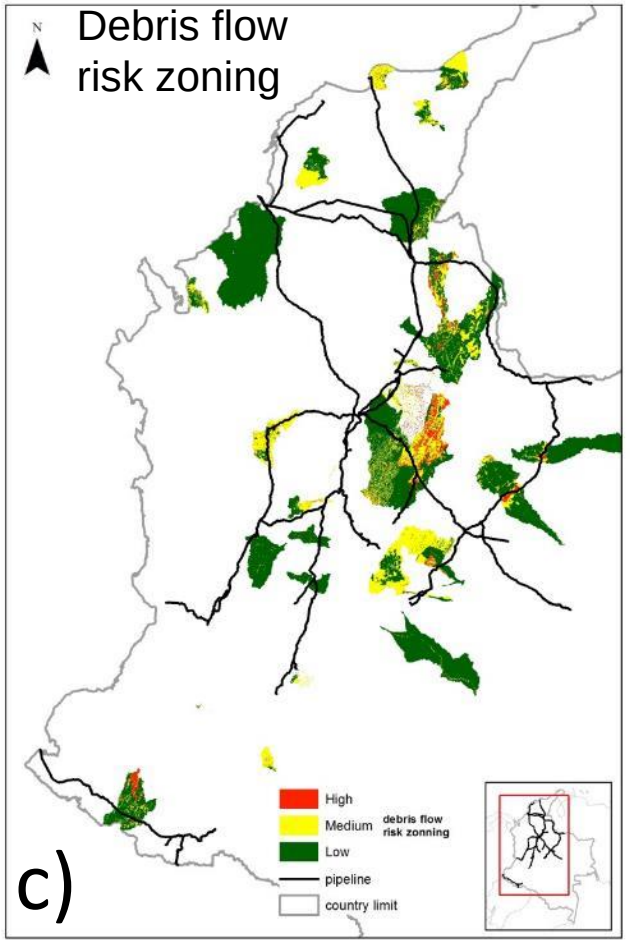
The methodology implemented for this zoning via fuzzy logic and supervised learning represents a valid and pertinent approach to optimize infrastructure intelligence management processes in pipelines.

At the same time, it allows to highlight the importance of data-driven-modeling and the design and management of robust information systems that involve descriptors and chronology of potentially damaging events for the analyzed infrastructure.

Plans for the Ordering and Management of Hydrographic Basins - POMCAS

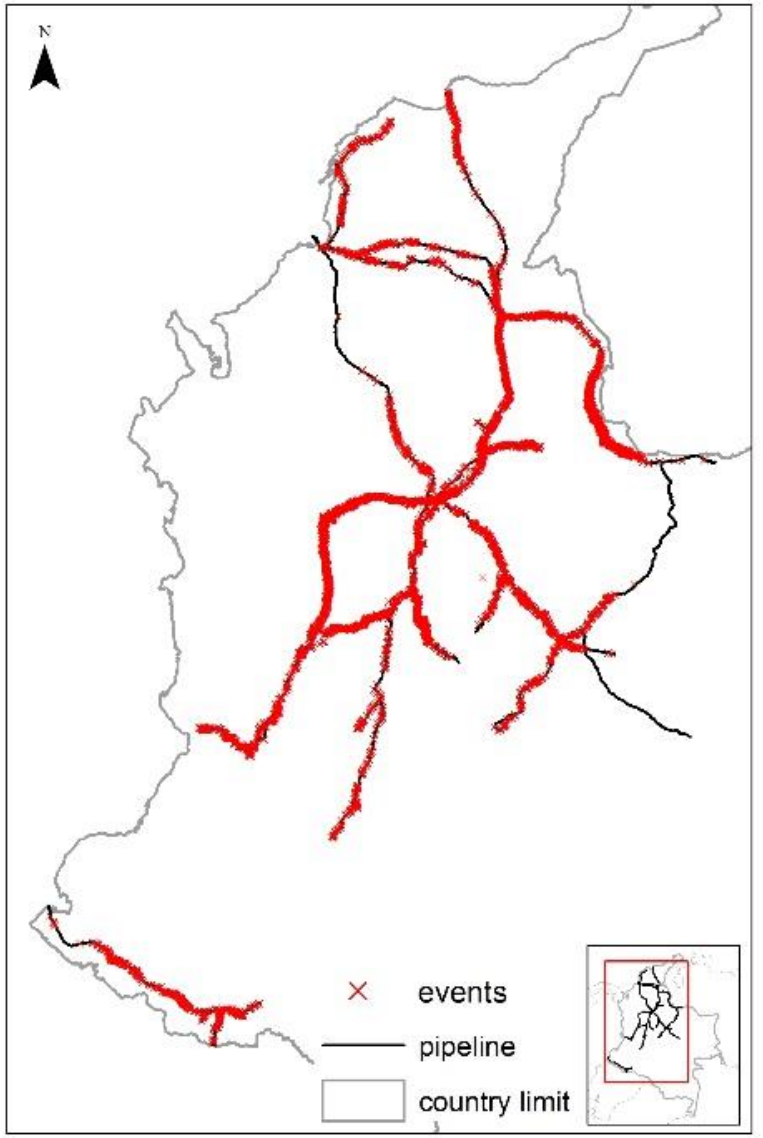




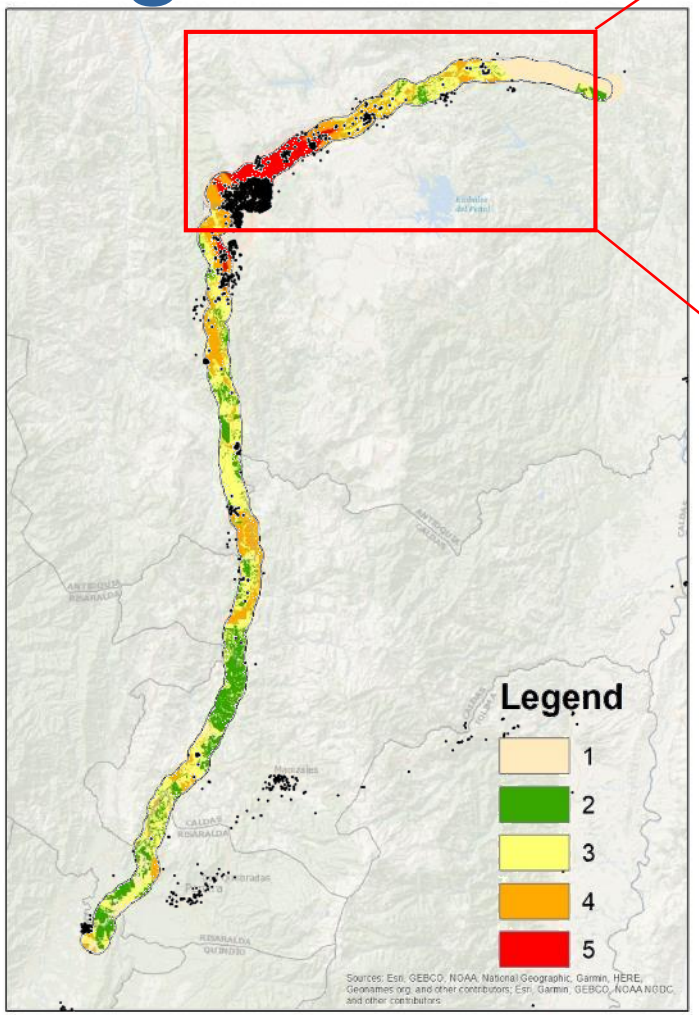


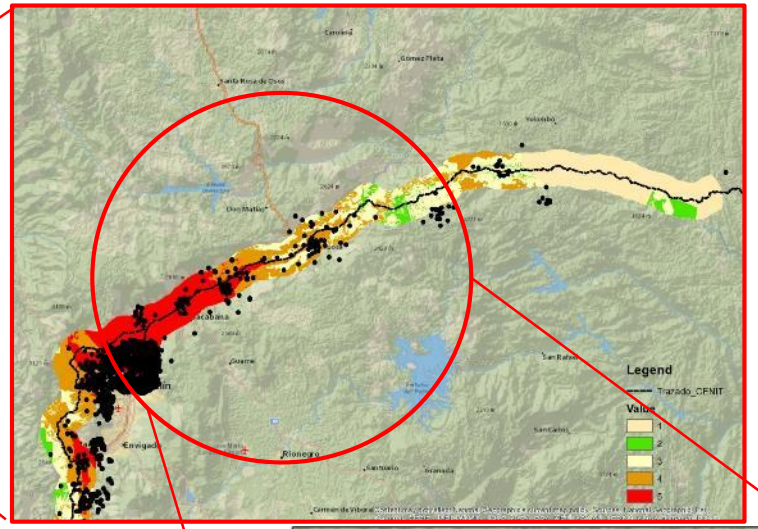
Historic events

Source – Data Partitioning	Records	Dates	Objetive
ECOPETROL	2375	NA	Training
SIMMA SGC	3164	1916 - 2022	Training
Baseline findings CENIT (Dec 2020)	2741	2020 - 2021	Validation
Baseline findings CENIT (Jul 2022)	2362	2021 - 2022	Validation
Baseline findings CENIT (May 2023)	2616	2022 – 2023	Validation

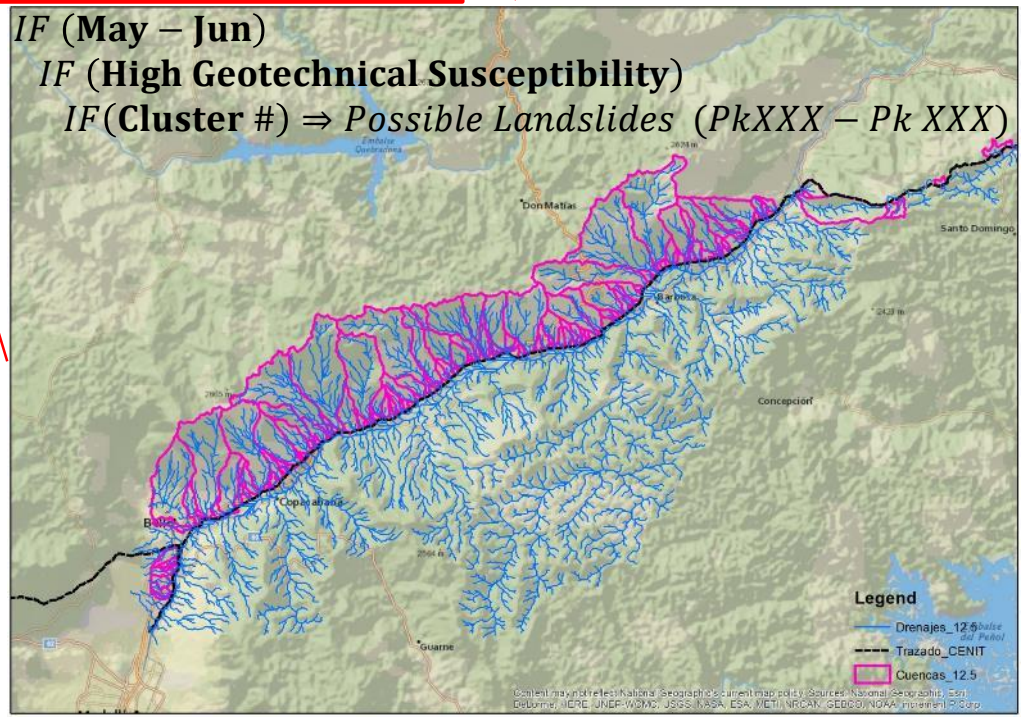


Rules generation

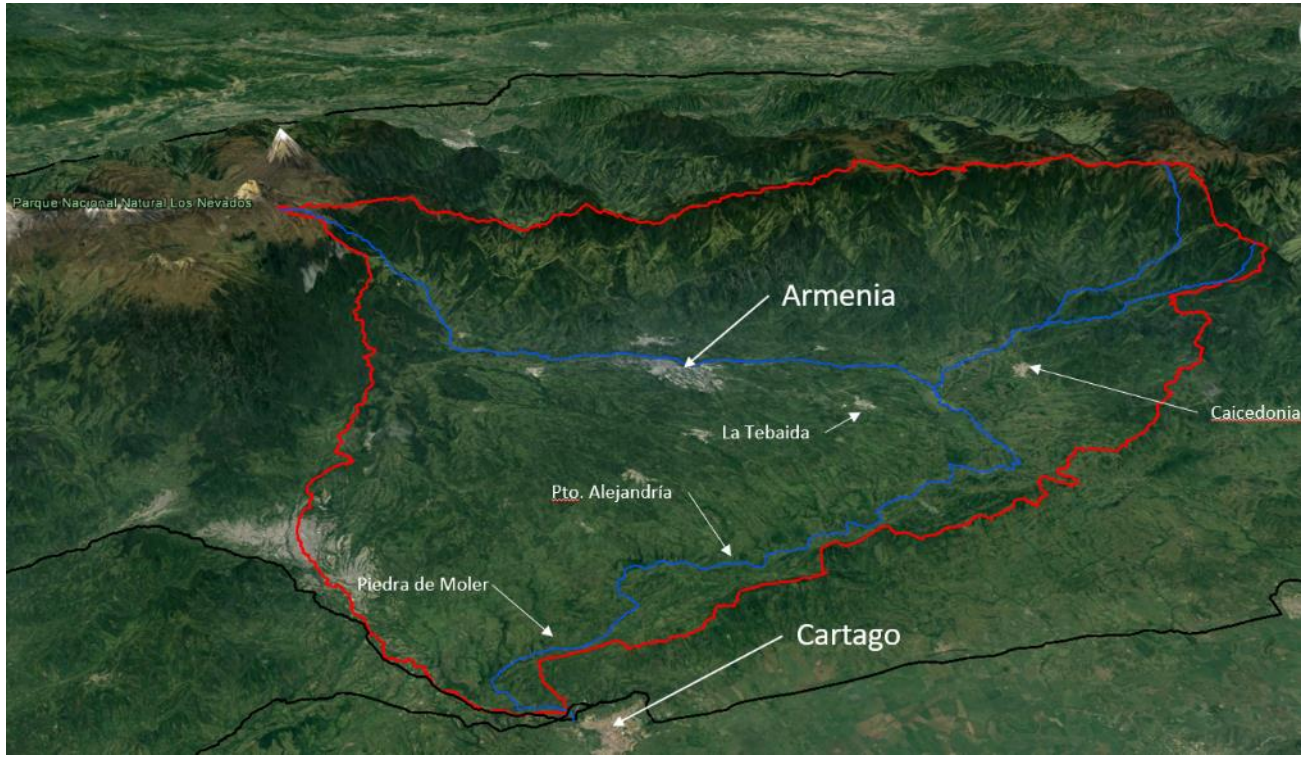


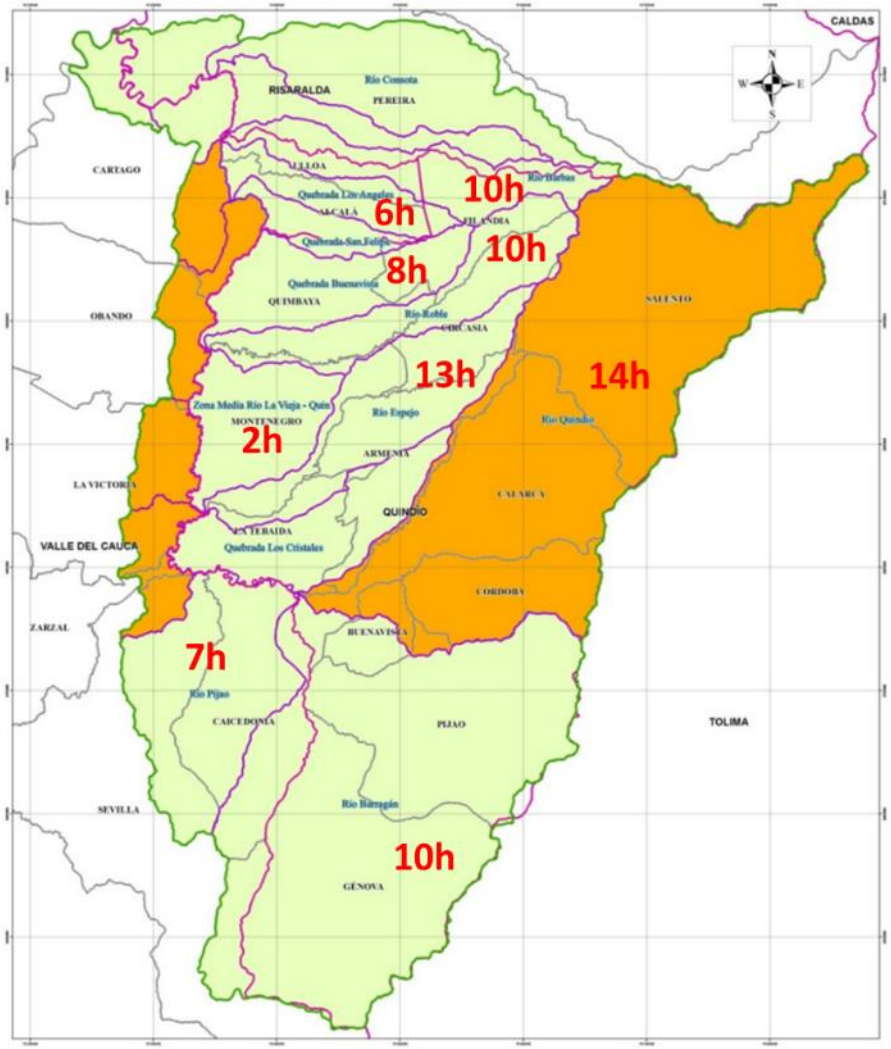


- Cluster Number
- High Geotechnical Susceptibility
- Geotechnical events
- Classification of Findings
- Relevant dates
  - Municipalities
  - Pipelines
- Prioritized areas



Morphometries and Concentration times



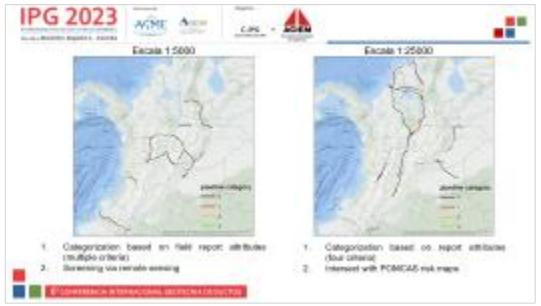


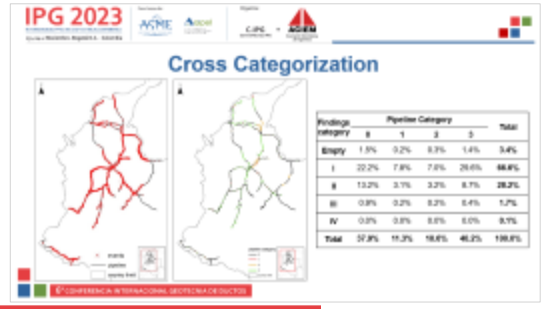
Criteria for categorizing pipeline segments

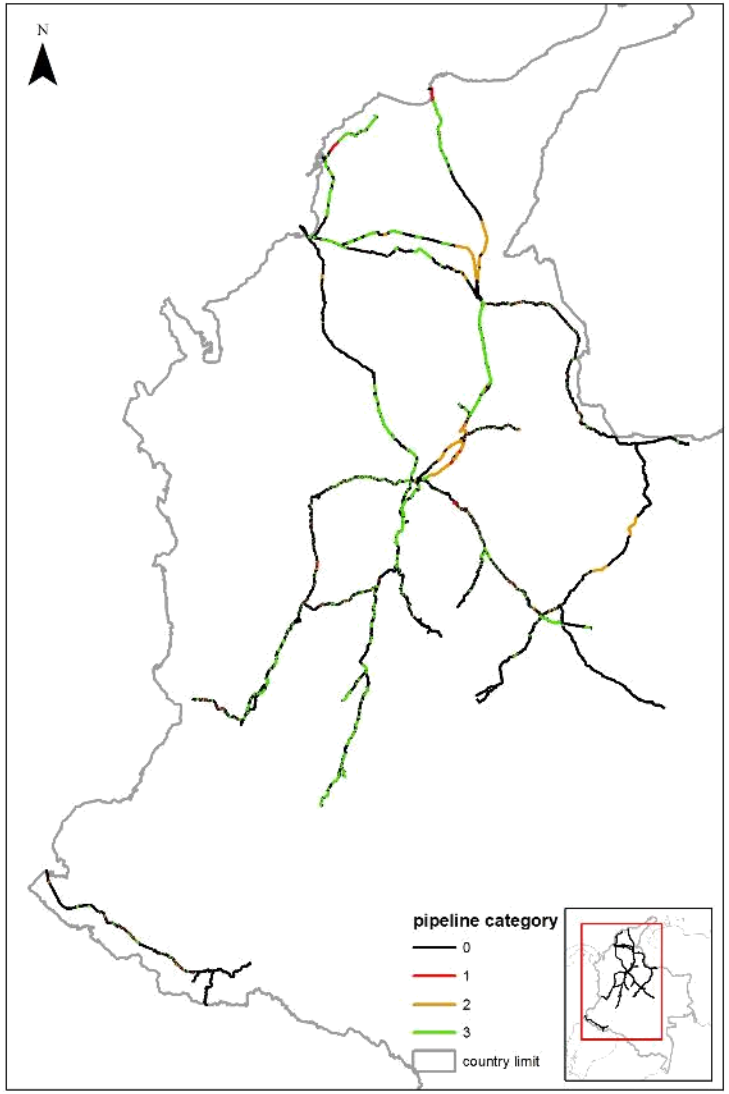
Pipeline code category	Category	Criterion
1	Slides, Toppless, Falls, Complex	Medium or high risk from Figure 2-a) Evidence of landslide crowns or scarps Coverage changes in the basin.
2	Flows, Flood, Torrential Floods	Medium or high risk from Figure 2-a) and 2-b). Evidence of quaternary deposits (Colluvium or Alluvial). Evidence of flows.
3	Erosion, Scour, Creep	Evidence of erosion, scour or creep from the base map or reports.
0	No evidence or none of the above categories	Low risk from Figure 2-a), 2-b), or 2-c) There are no historical reports in any source consulted. Stabilization works had been carried out. There is an isolated report, but of low reliability that could not be confirmed with images or base maps.

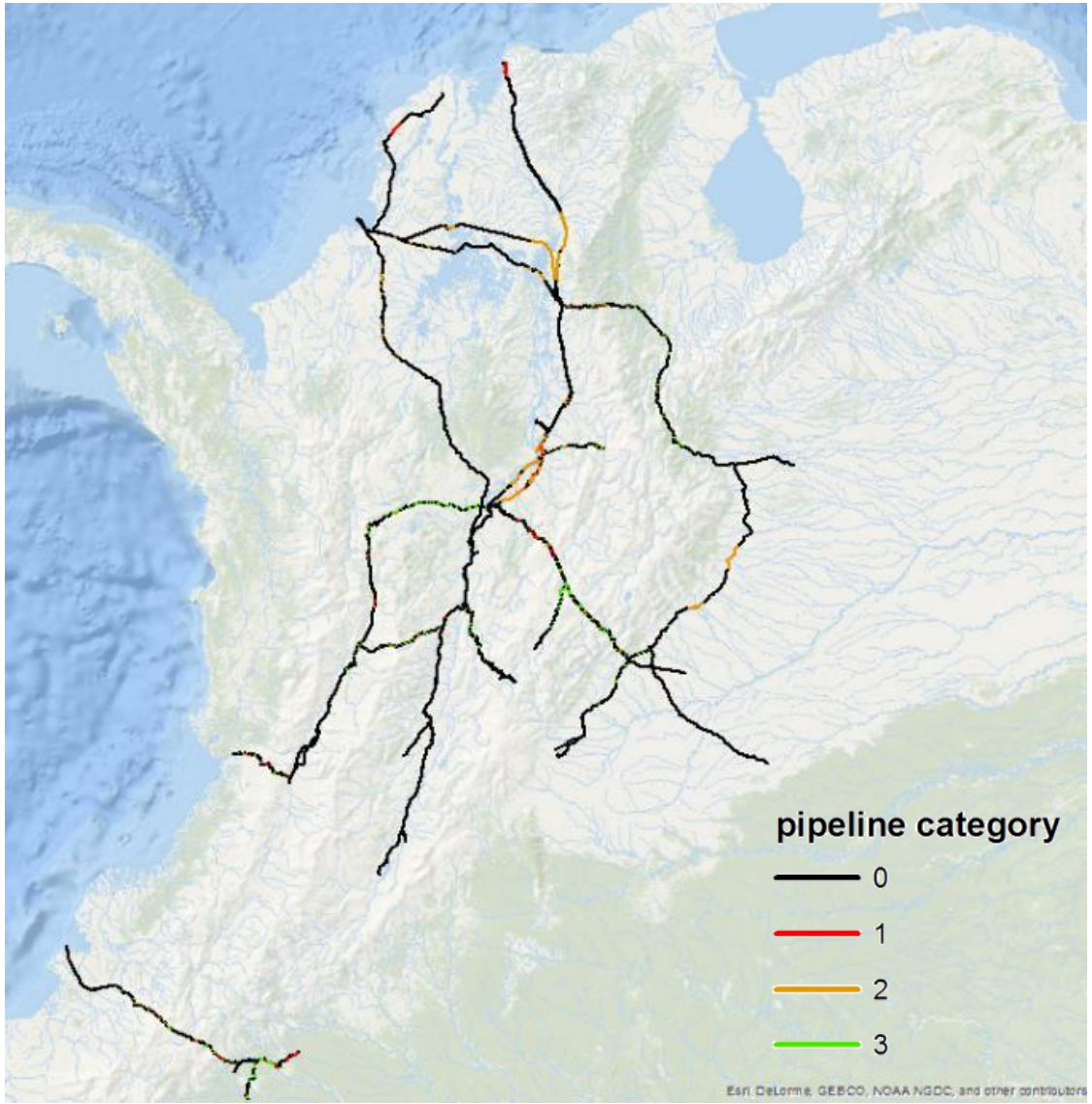
[1] Category codes do not constitute an ordinal variable; the value is assigned to facilitate manipulation in the analyzes.
[2] Category considers Movement Types Classification (Varnes, 1978).

Pipeline Segmentation





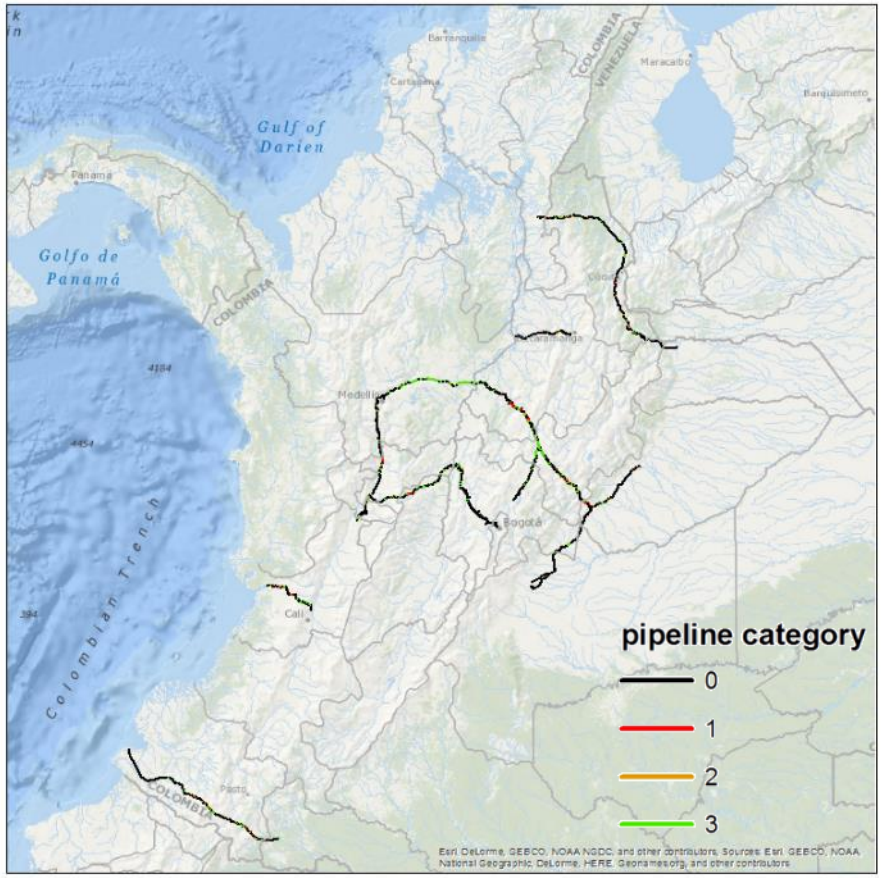




Semi-automated process

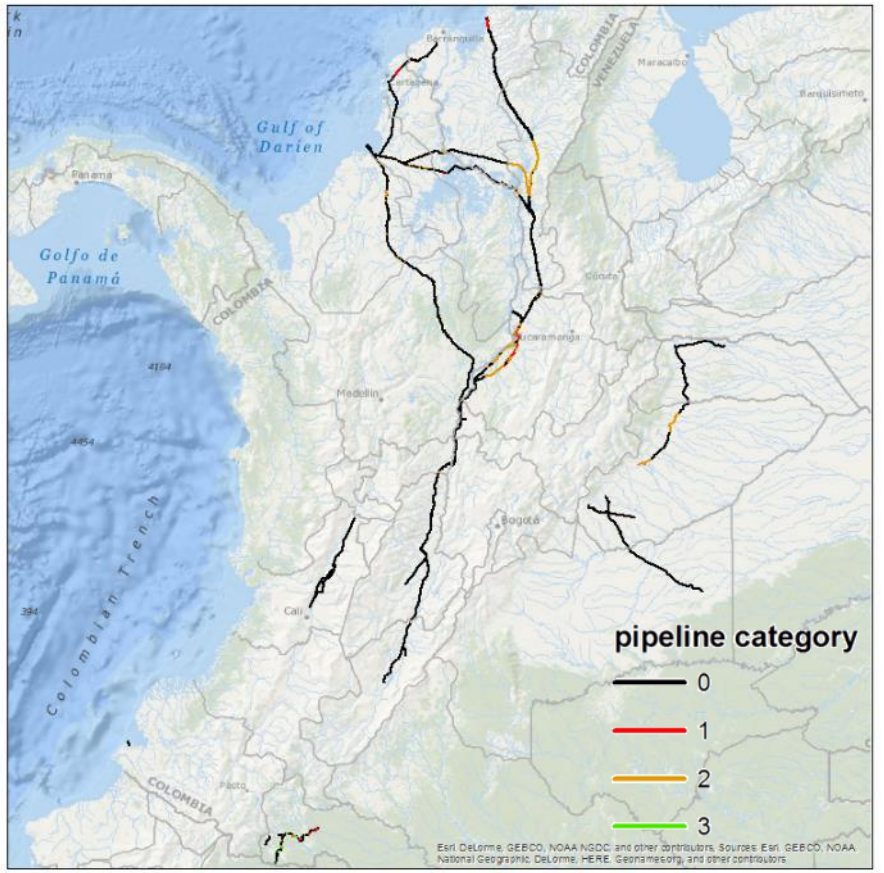
1. Consolidation of all events
2. Review of afferent basins
3. Crossings (subfluvial and aerial)

Escala 1:5000



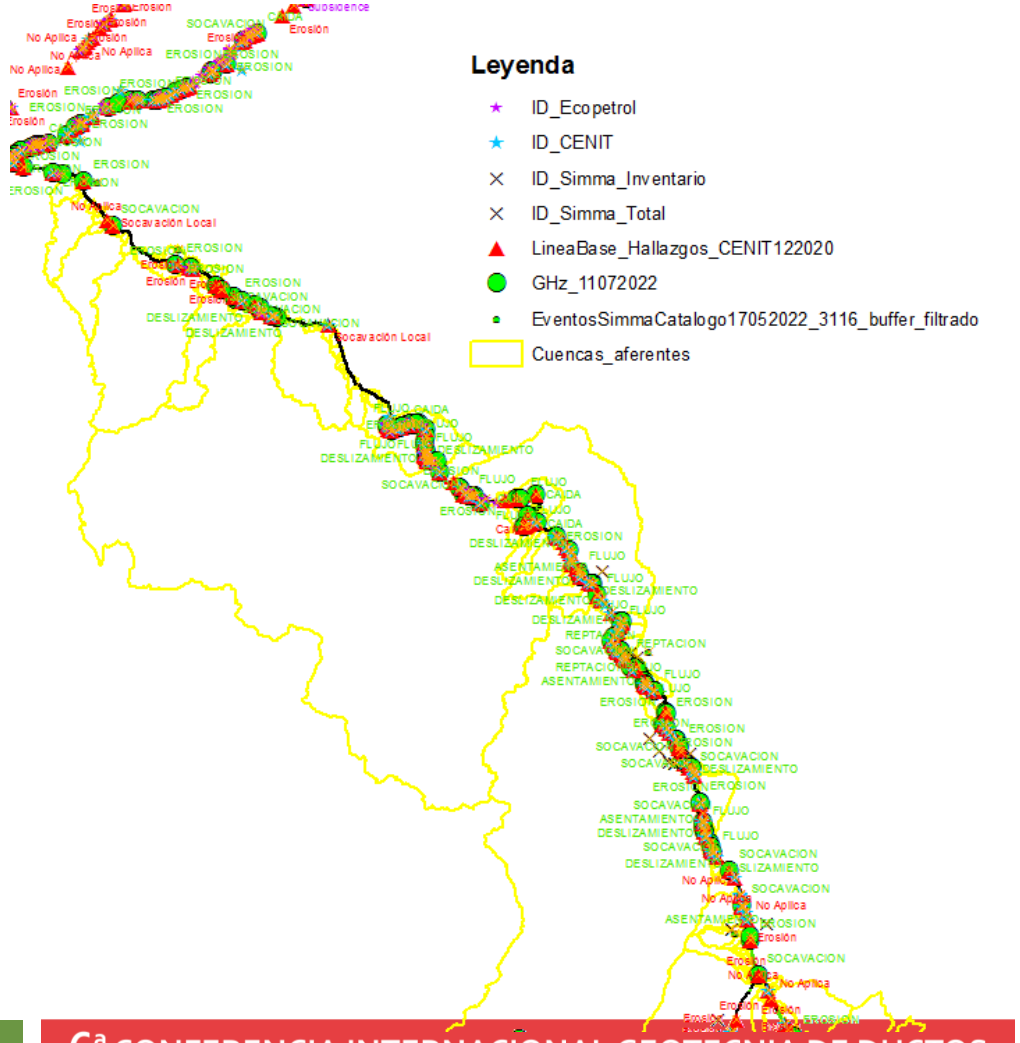
1. Categorization based on field report attributes (multiple criteria)
2. Screening via remote sensing

Escala 1:25000

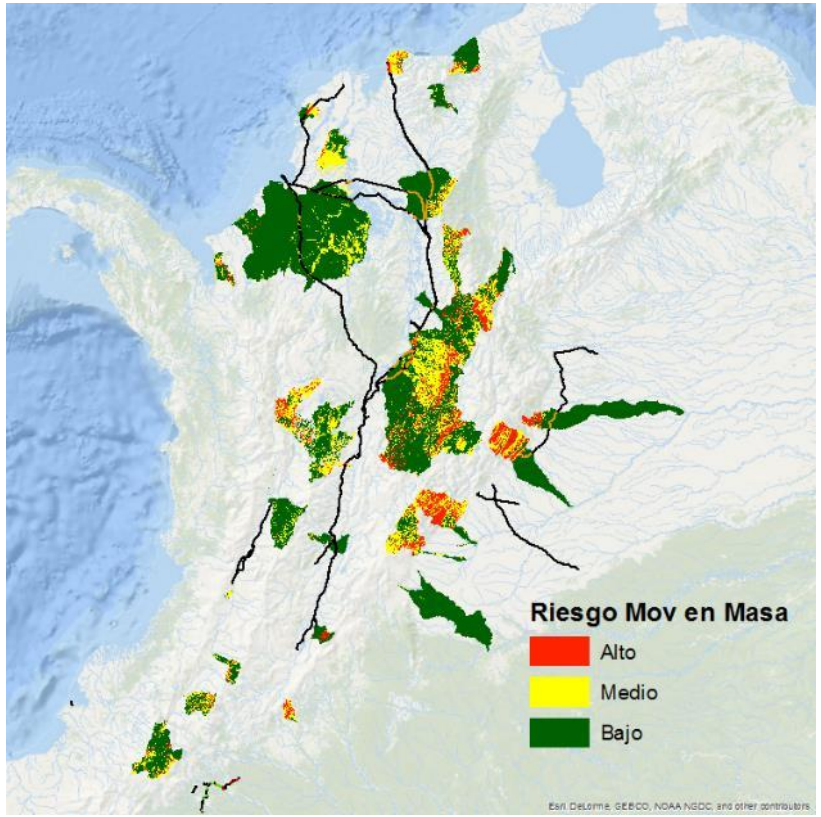


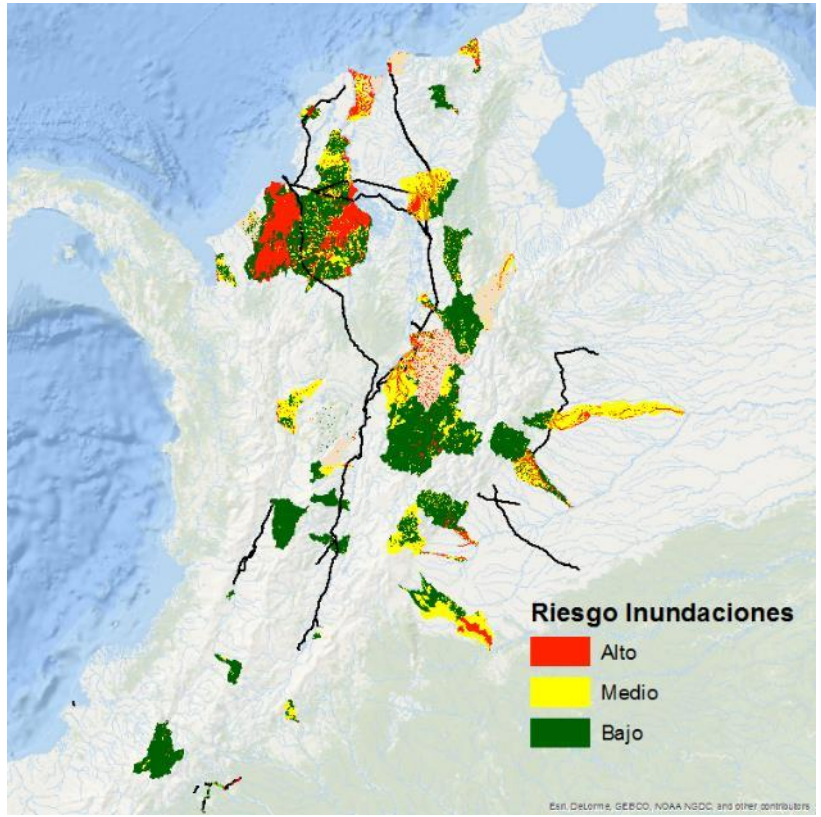
1. Categorization based on report attributes (four criteria)
2. Intersect with POMCAS risk maps

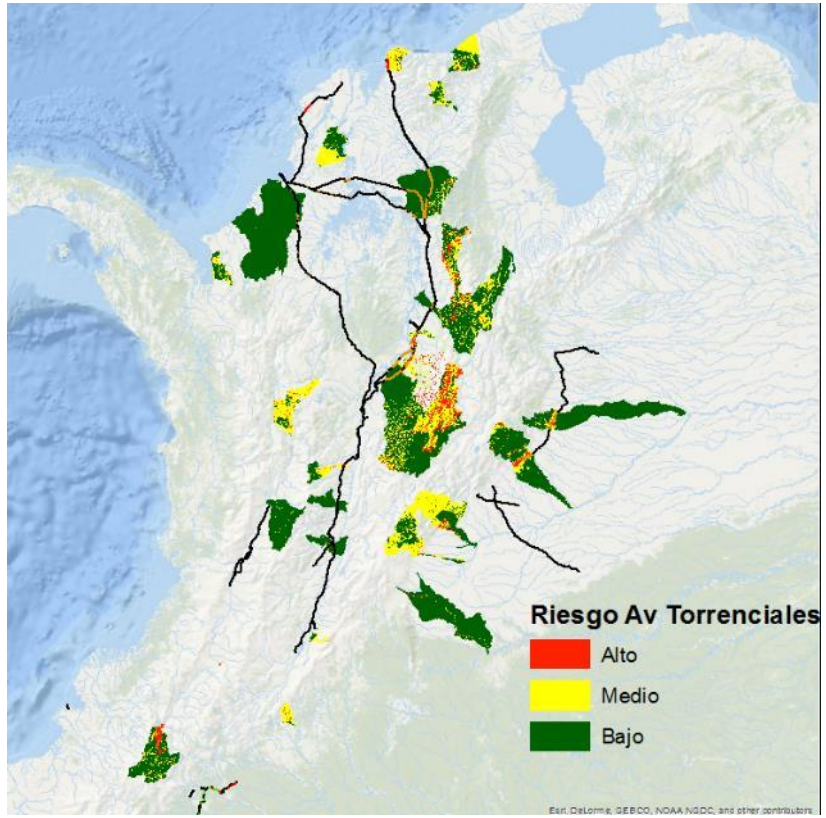
Intersect with Remote Sensing



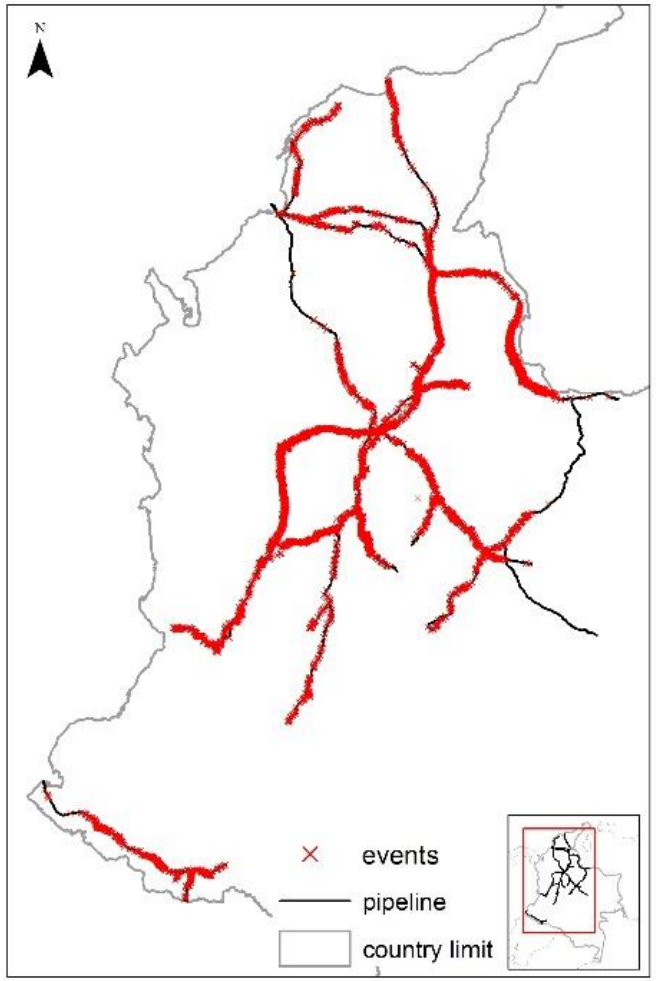
Intersect with POMCAS

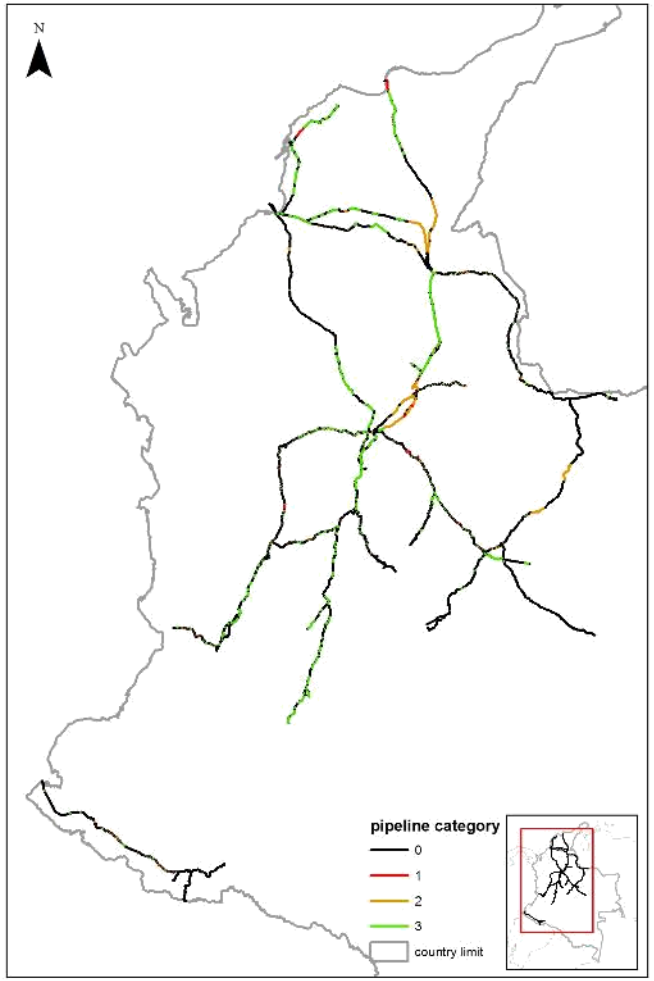






Cross Categorization



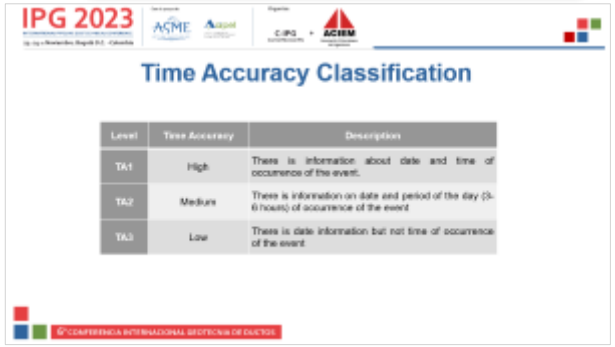


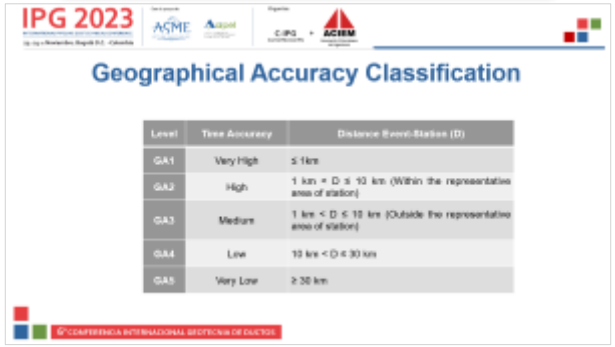
Findings category	Pipeline Category				Total
	0	1	2	3	
Empty	1.5%	0.2%	0.3%	1.4%	**3.4%**
I	22.2%	7.8%	7.0%	29.6%	**66.6%**
II	13.2%	3.1%	3.2%	8.7%	**28.2%**
III	0.9%	0.2%	0.2%	0.4%	**1.7%**
IV	0.0%	0.0%	0.0%	0.0%	**0.1%**
Total	**37.9%**	**11.3%**	**10.6%**	**40.2%**	**100.0%**

Complementary Analysis

Criterias for Category Assignment

- Time Accuracy
- Geographical Accuracy
- Trigger Rain Episode





Time Accuracy Classification

Level	Time Accuracy	Description
TA1	High	There is information about date and time of occurrence of the event.
TA2	Medium	There is information on date and period of the day (3-6 hours) of occurrence of the event
TA3	Low	There is date information but not time of occurrence of the event

Geographical Accuracy Classification

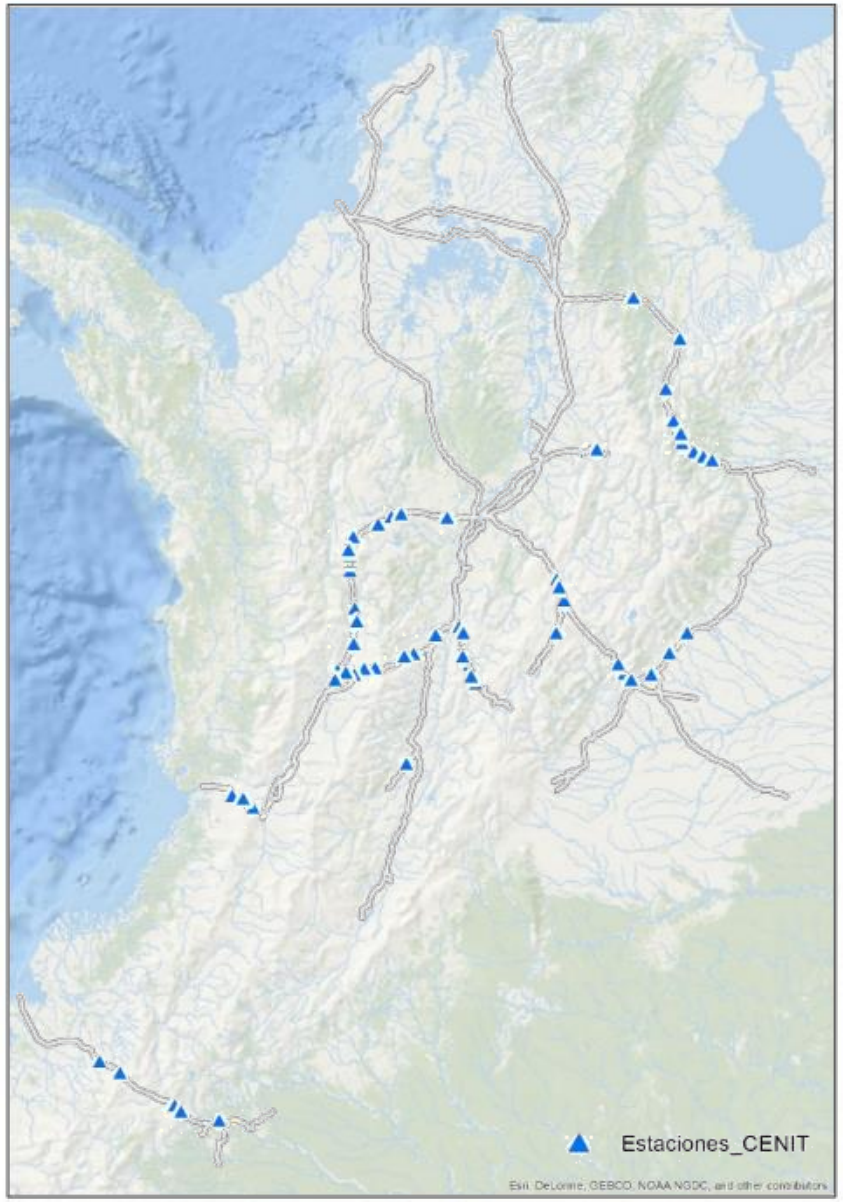
Level	Time Accuracy	Distance Event-Station (D)
GA1	Very High	≤ 1km
GA2	High	1 km < D ≤ 10 km (Within the representative area of station)
GA3	Medium	1 km < D ≤ 10 km (Outside the representative area of station)
GA4	Low	10 km < D ≤ 30 km
GA5	Very Low	≥ 30 km

Trigger Rain Episode Classification

Level	Degree of Definition	Characteristics of the Episodes
TR1	Defined	Well-defined episodes or multi-episodes of rain occurring within the previous 96 hours
TR2	Uncertain	Uncertain episodes or multi-episodes of rain occurring within the previous 96 hours
TR3	Undefined	Undefined episode or no significant rain episode within the previous 96 hours

Monitoring Network

90 measurement points of hydrometeorological variables in real-time that allows to know, quantify, and qualify the extreme precipitation events and its physical effects in the areas of interest.



Example

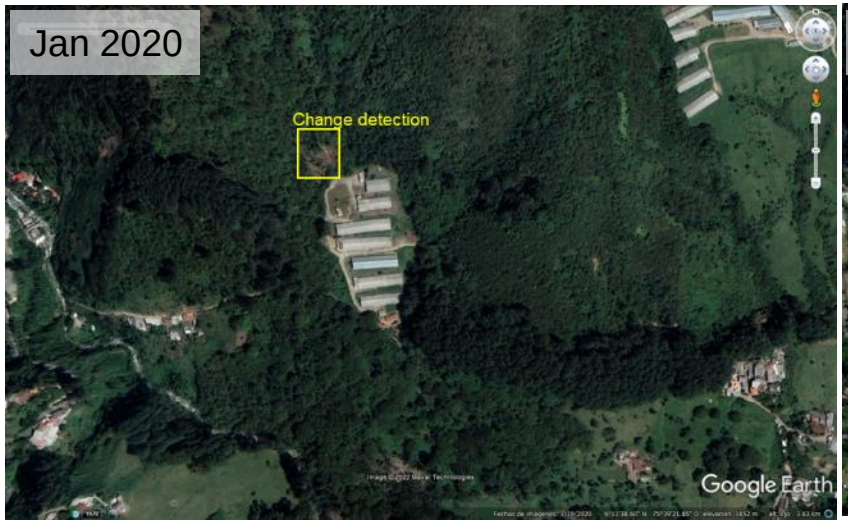

Latitud: 6.188465
Longitud: -75.65733
Elevación: 1902.45±10 m
Precisión: 5.0 m
Tiempo: 13-07-2022 16:38
Powered by NoteCam


Jan 2001
Google Earth


Jun 2006
New Roads
Google Earth

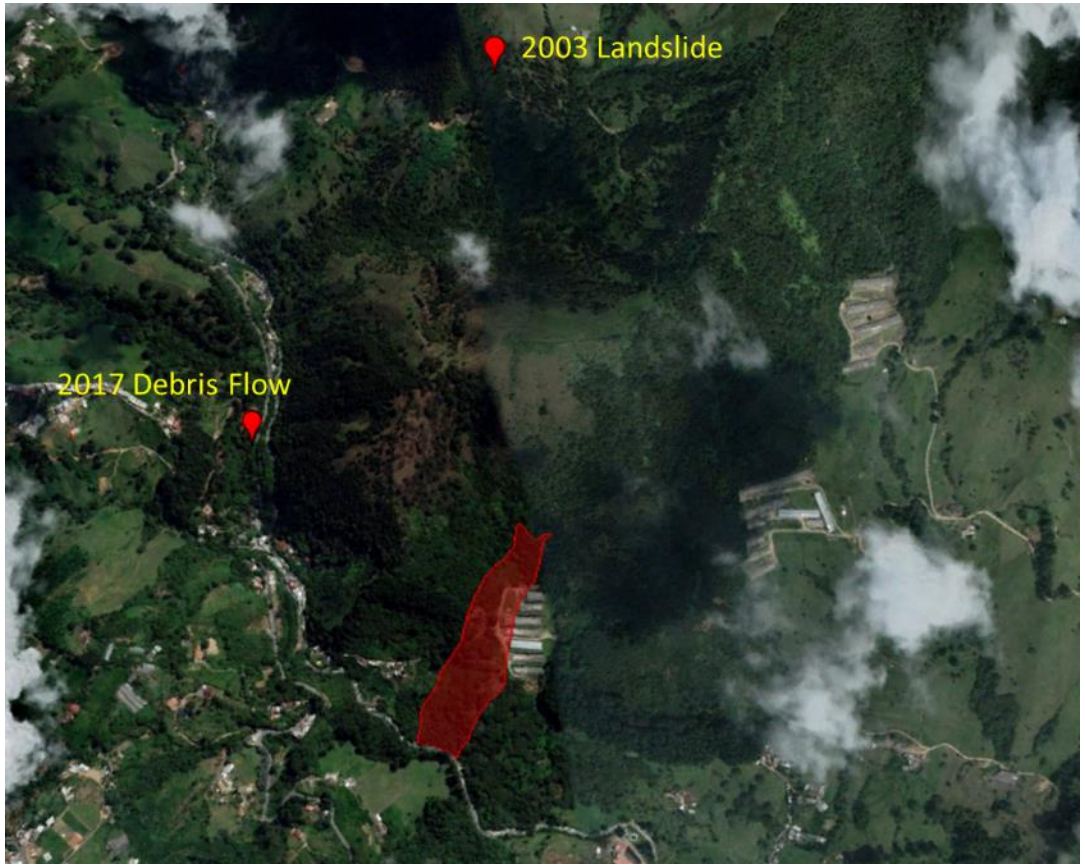

Feb 2012
Landslide Crown
Google Earth

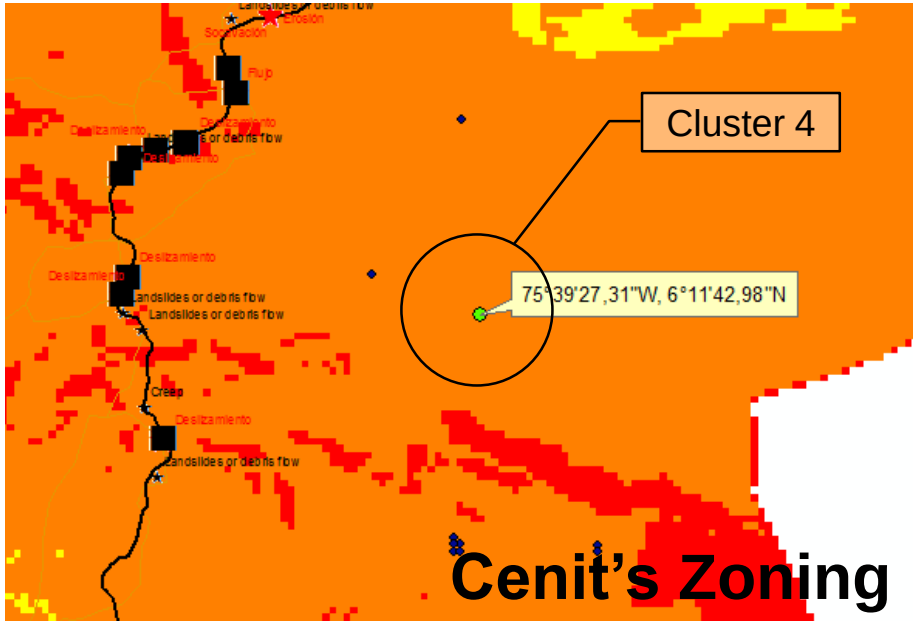

Jun 2017
Change detection
Google Earth

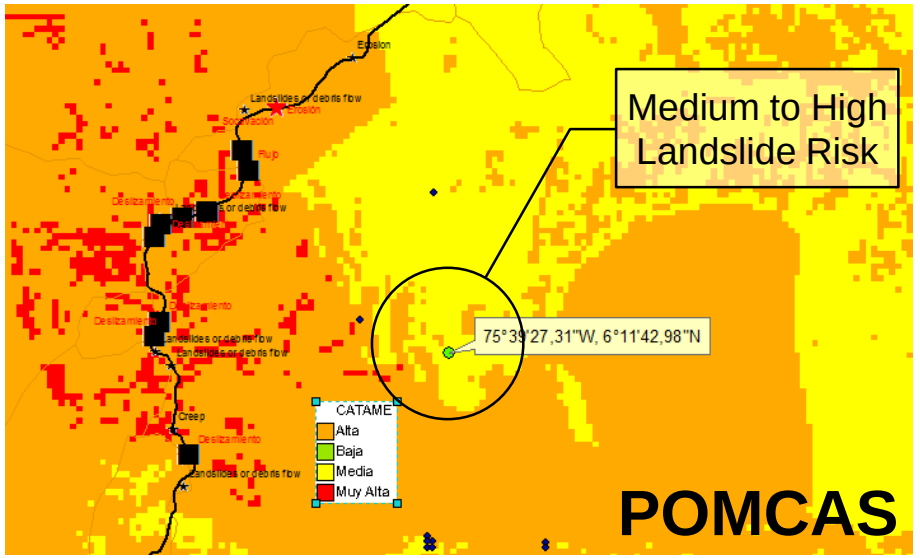

Jan 2020
Change detection
Google Earth

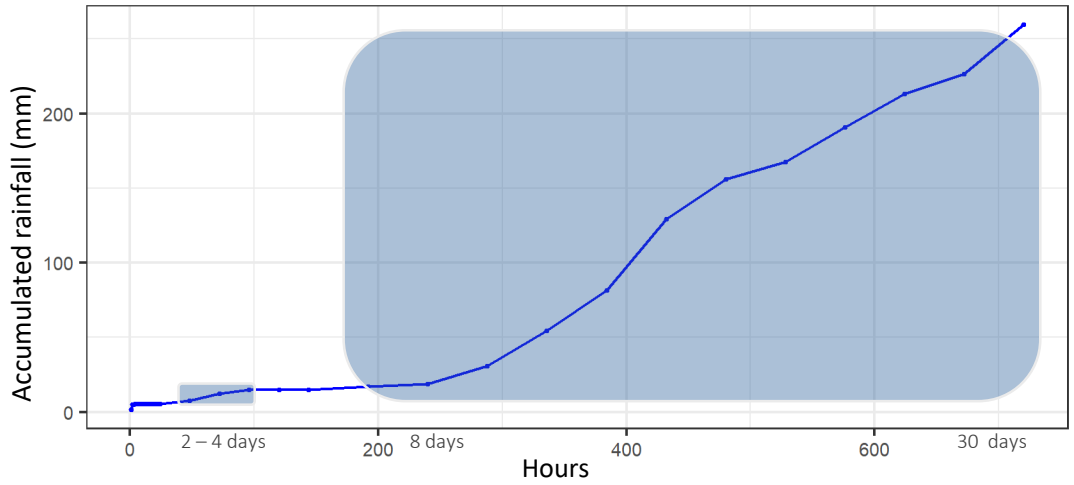

Nov 2021
Change detection
Google Earth

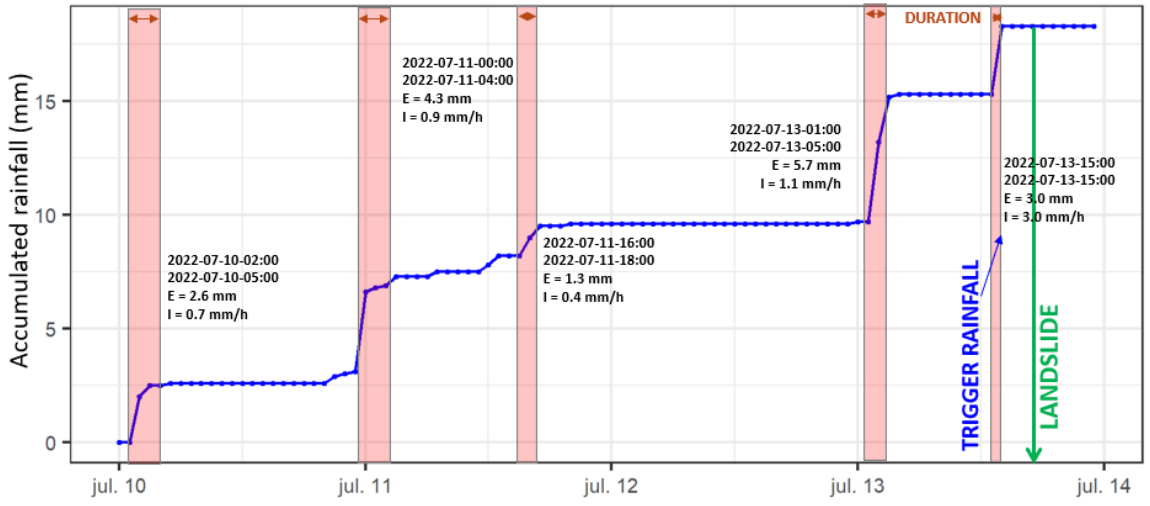
SIMMA Reports

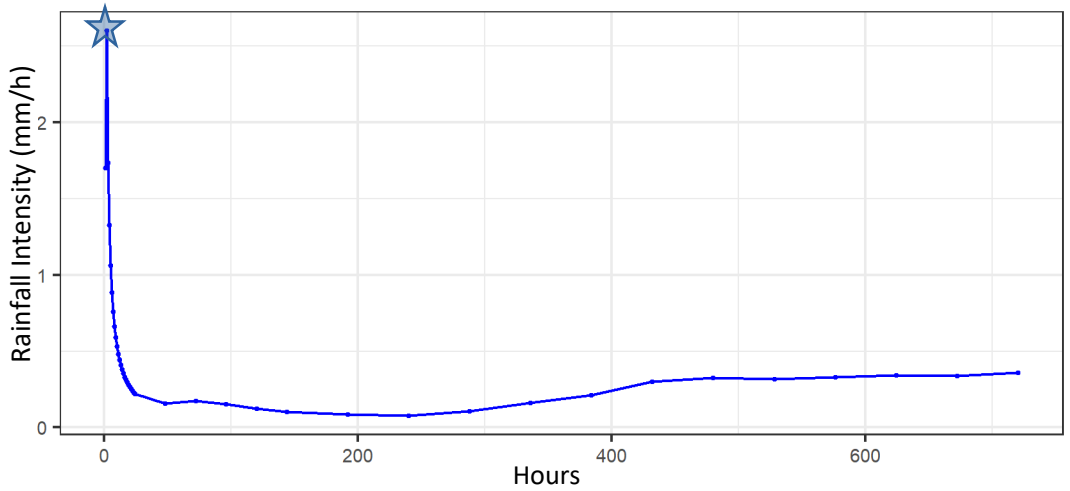




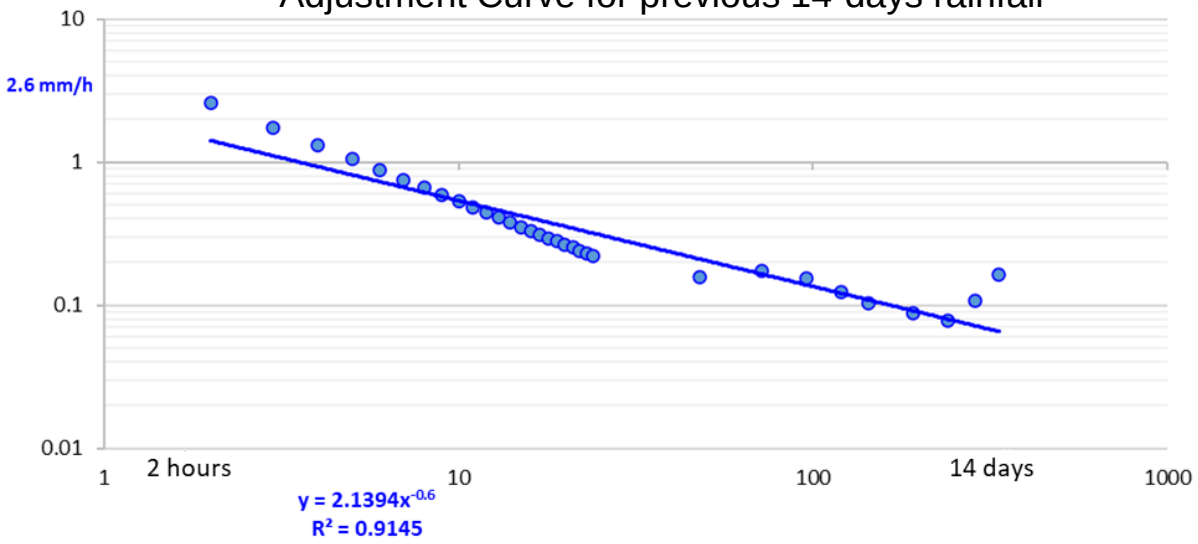


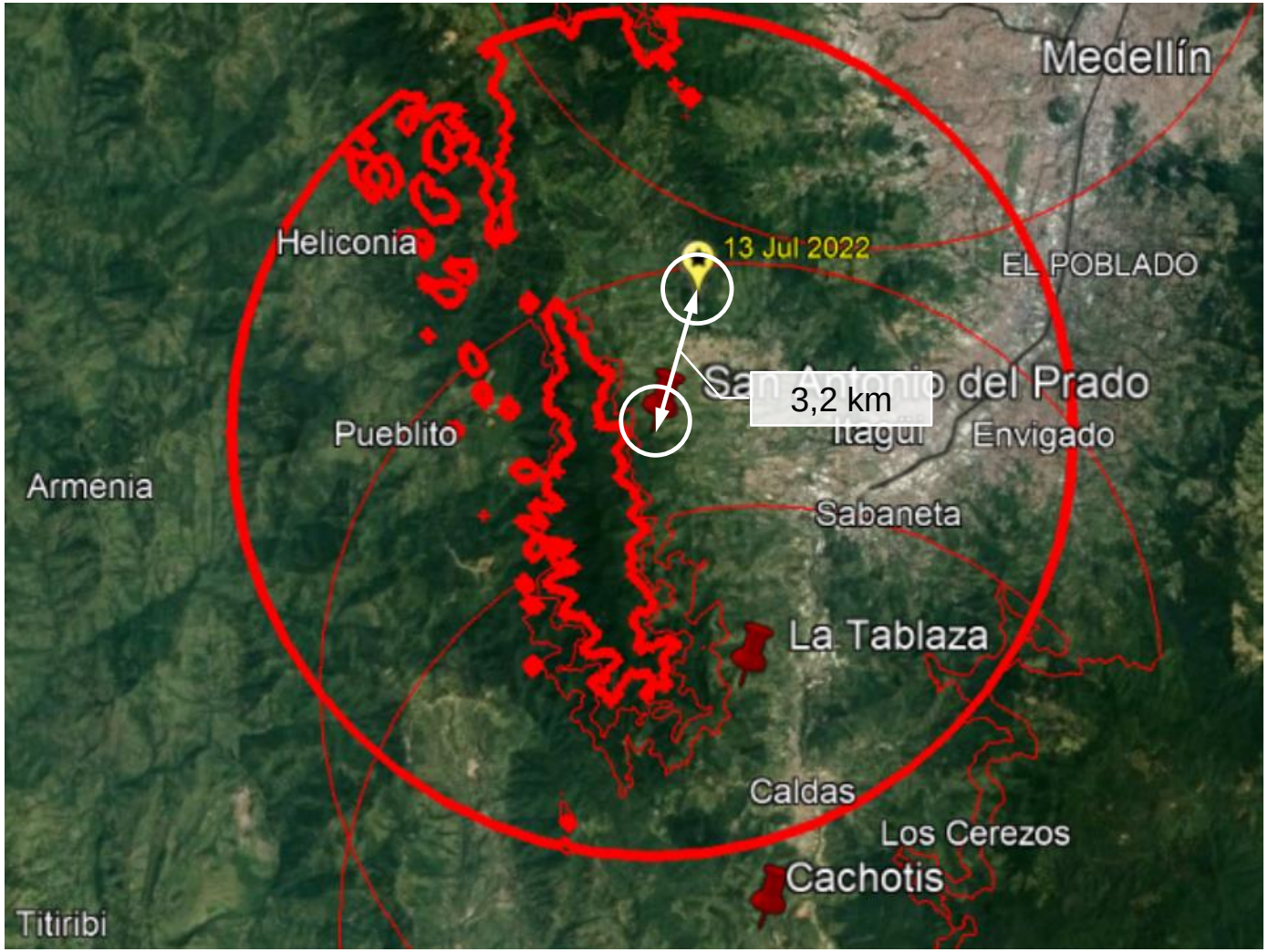






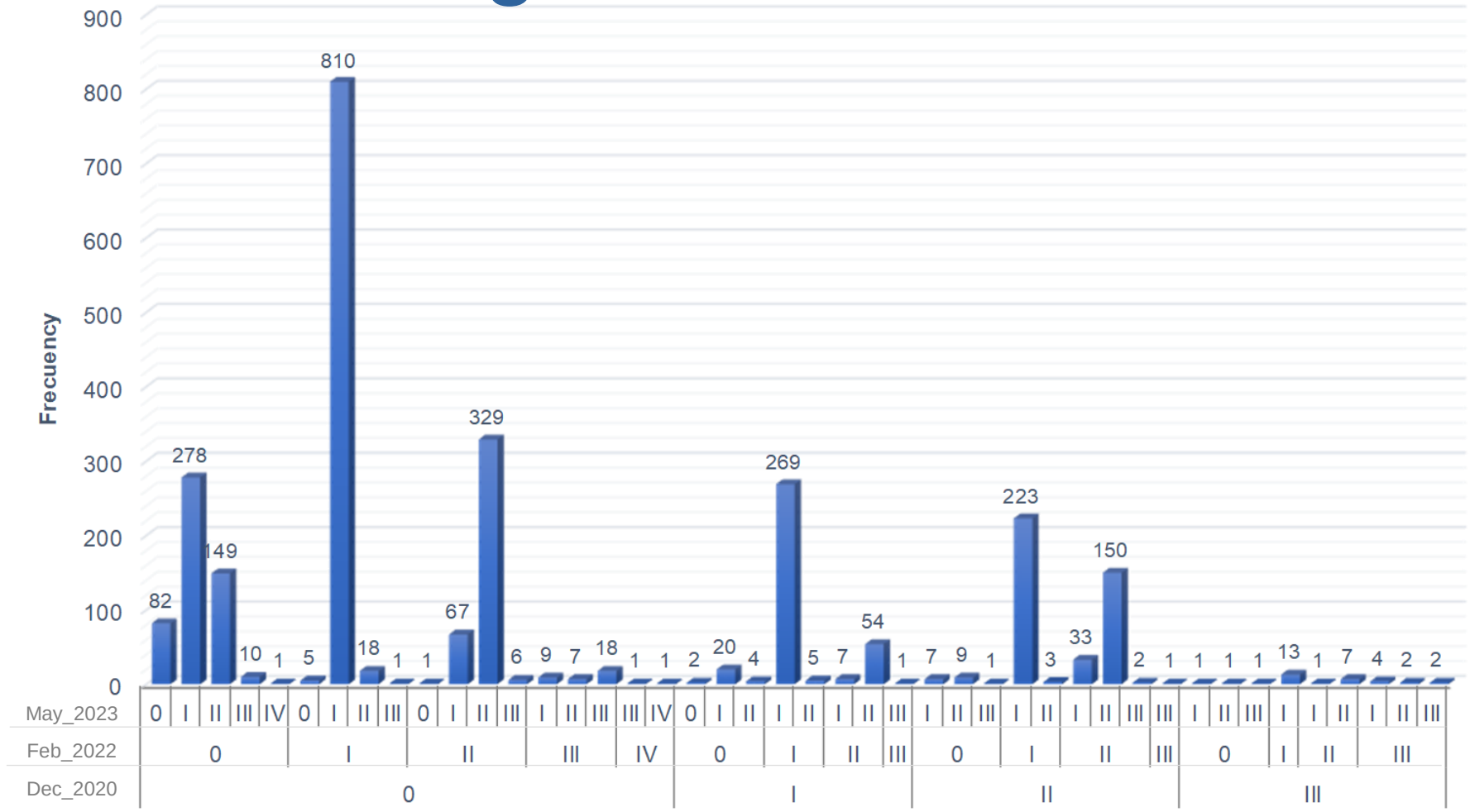
Adjustment Curve for previous 14-days rainfall





Time Accuracy	Geographical Accuracy	Trigger Rain Episode Classification
TA2	GA2	TR1

Findings evolution in time



CONCLUSION

For this analysis process to contribute adequately to decision-making, the importance of having sources of information (hydrometeorological and ROW patrol) in real time has been identified in which the quality and continuity of the measurements is guaranteed during sufficiently representative periods to be able to understand the dynamics of the instability processes and their climatic triggers.

This enables the transition from a real-time information to a right-time information.