ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

High performance computing in geostatistics

Big Data & Environment Workshop

Buenos Aires, November 13th 2015

Fabián Soto F.
fsoto@alges.cl

Ingeniería de Minas
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS UNIVERSIDAD DE CHILE

amtc
ADVANCED MINING TECHNOLOGY CENTER

Álvaro Egaña

Xavier

Fabián

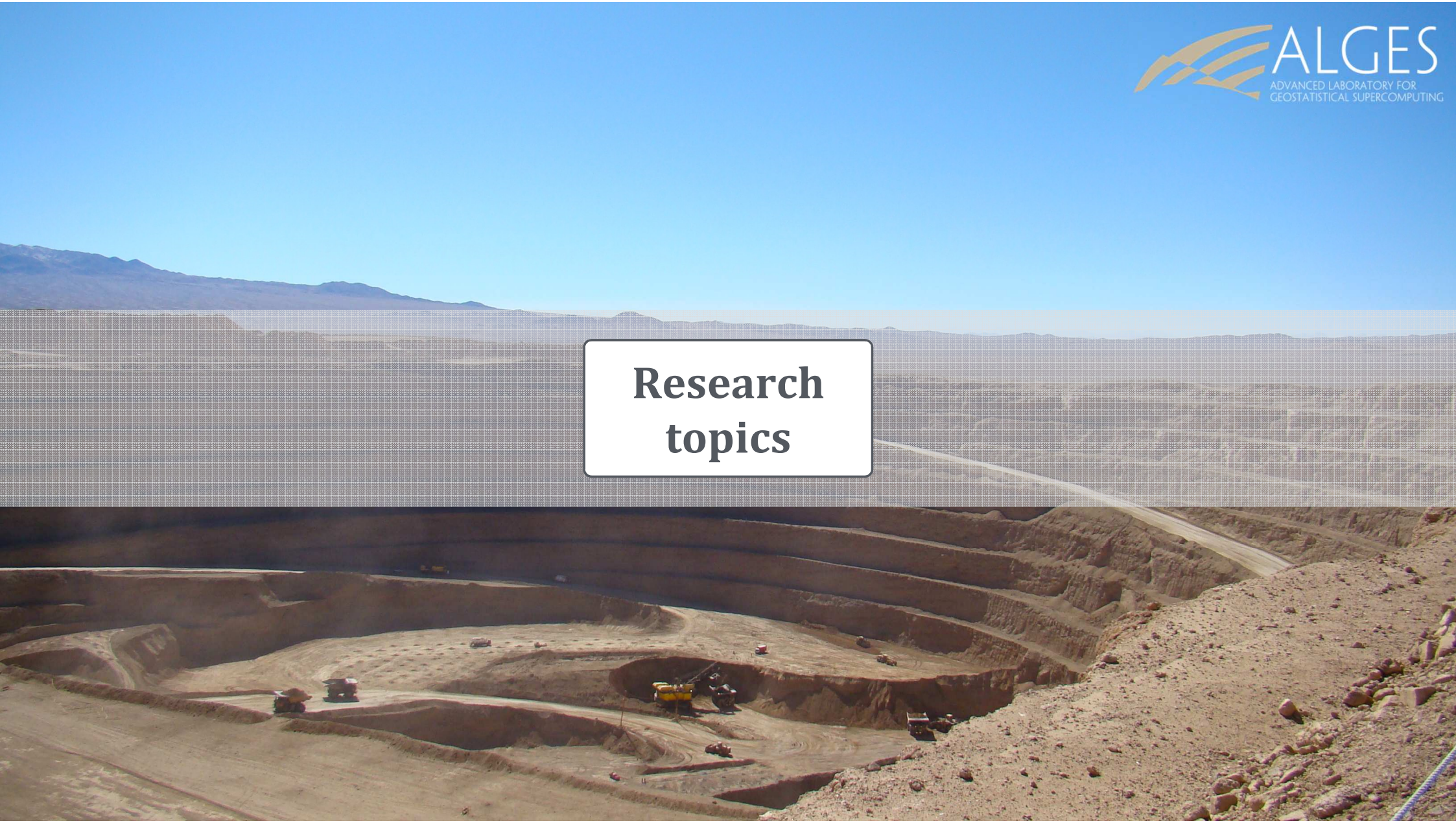
ALGES

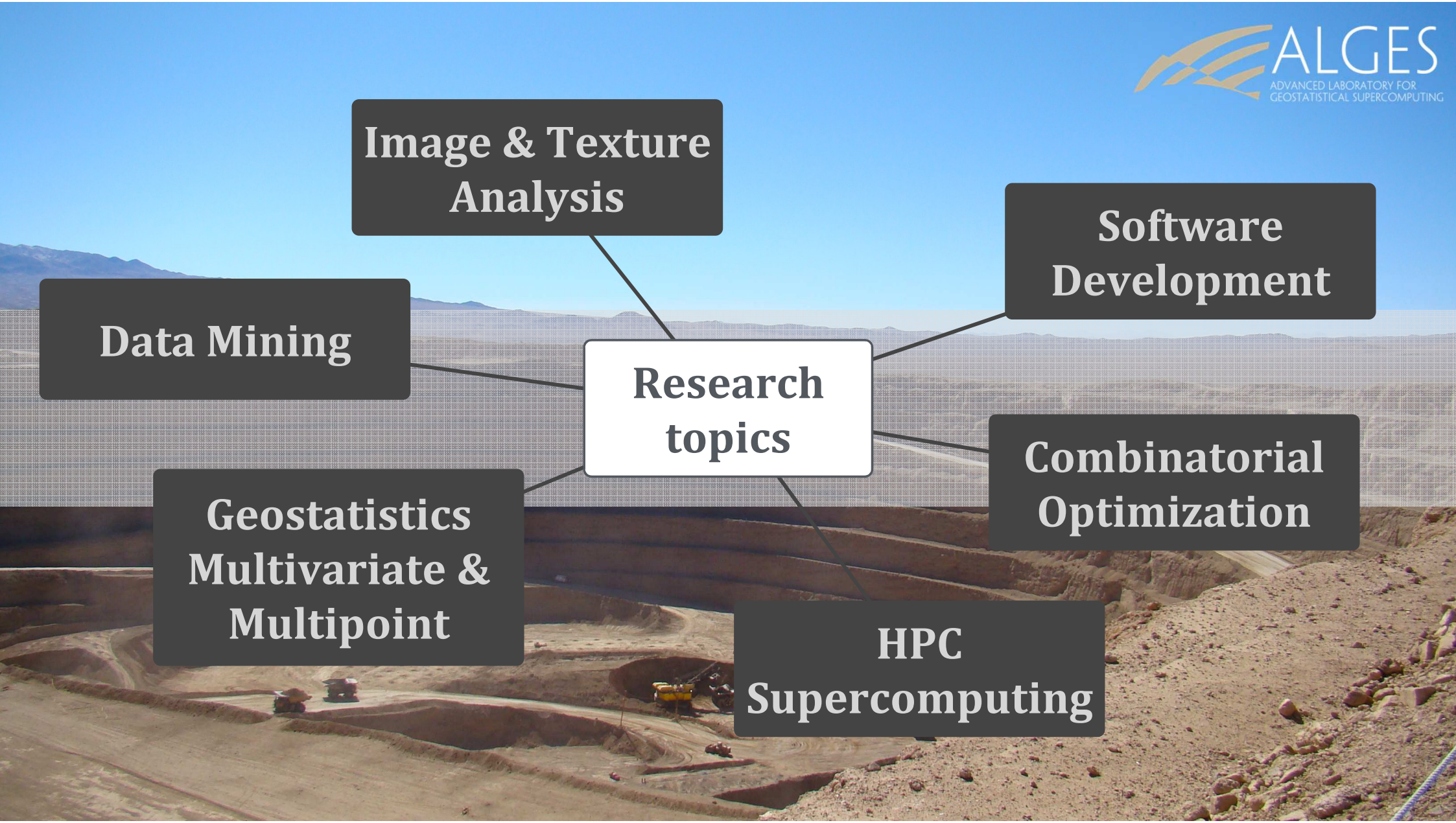
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

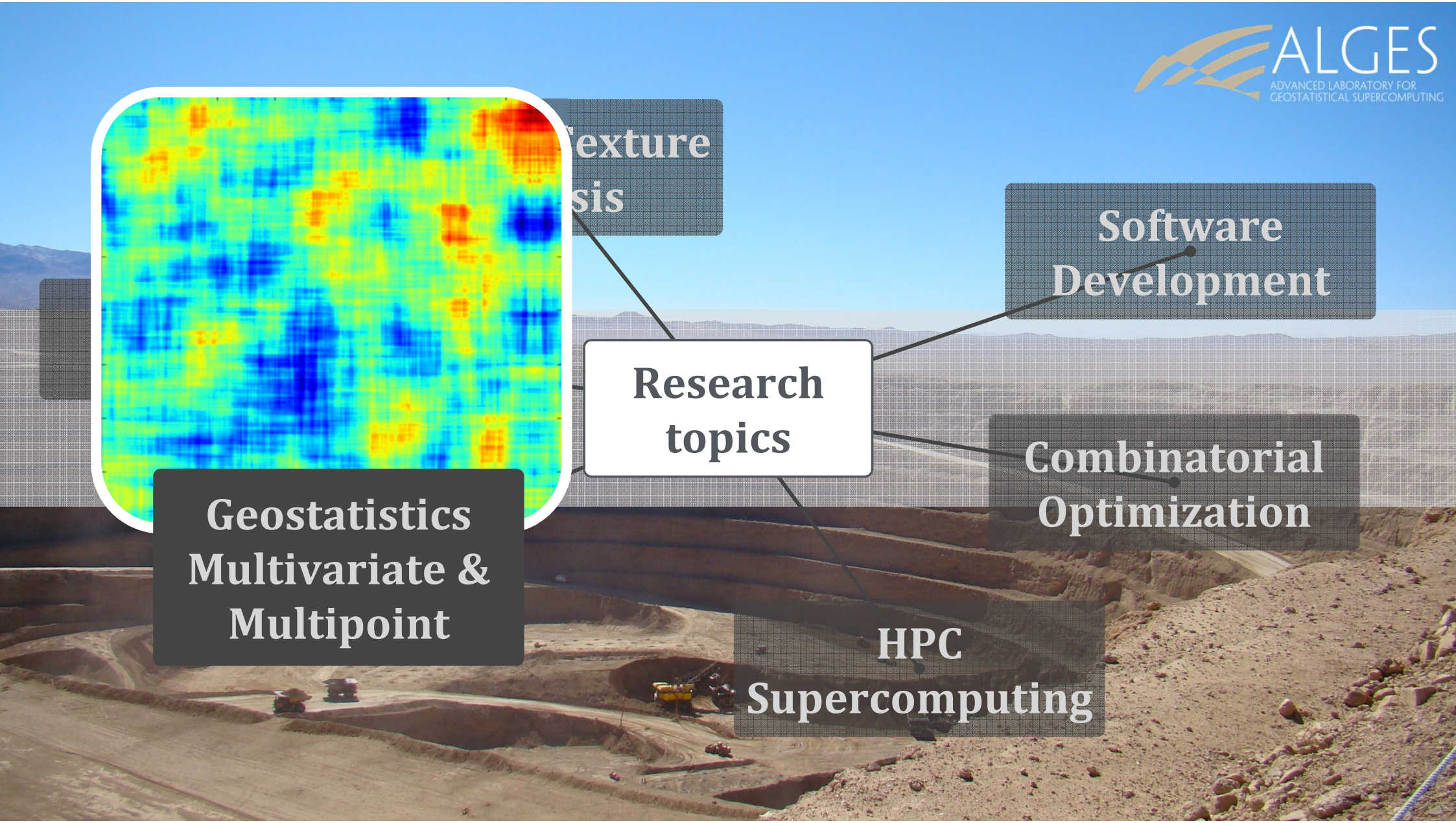
Felipe

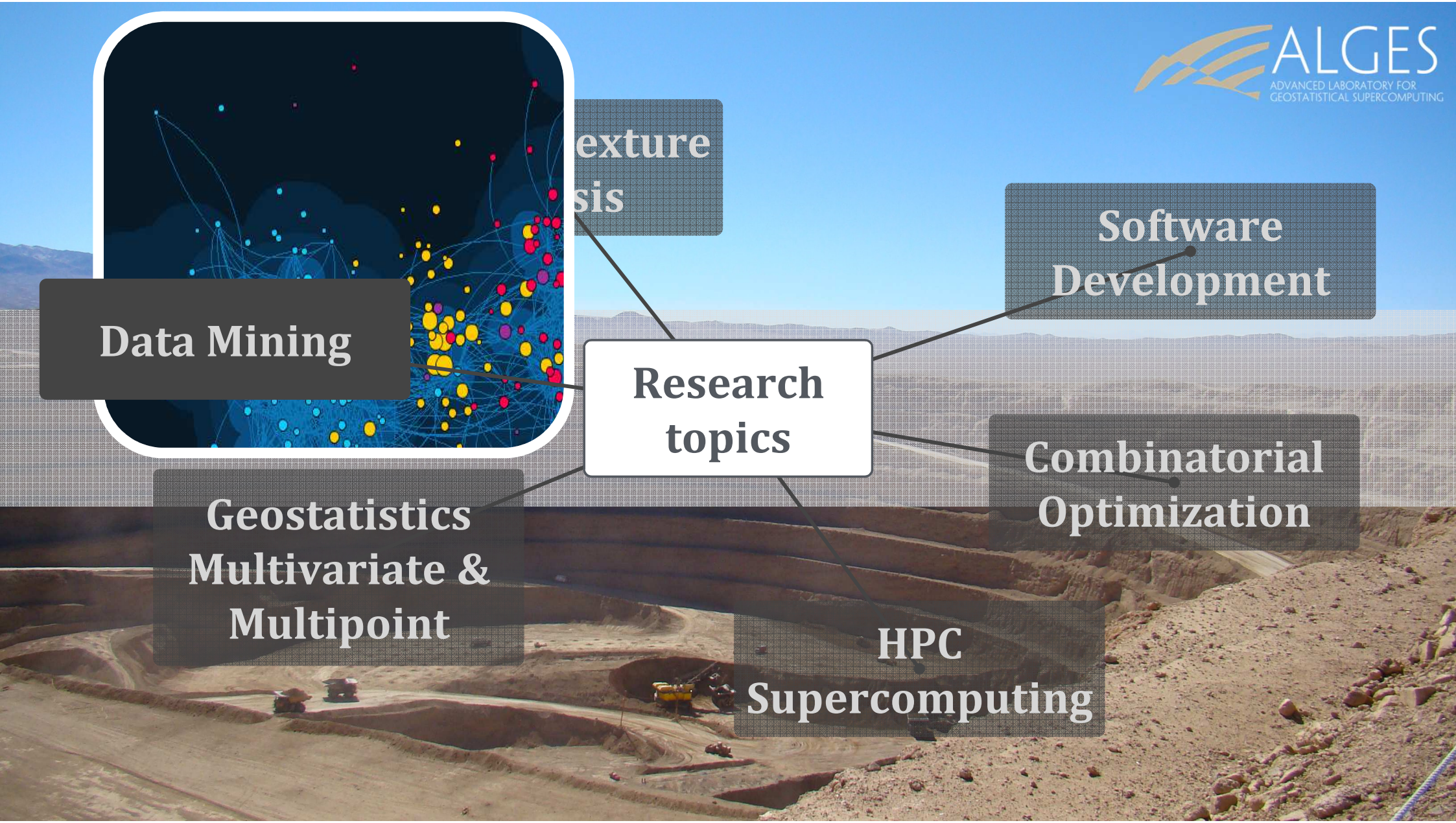
Daniel

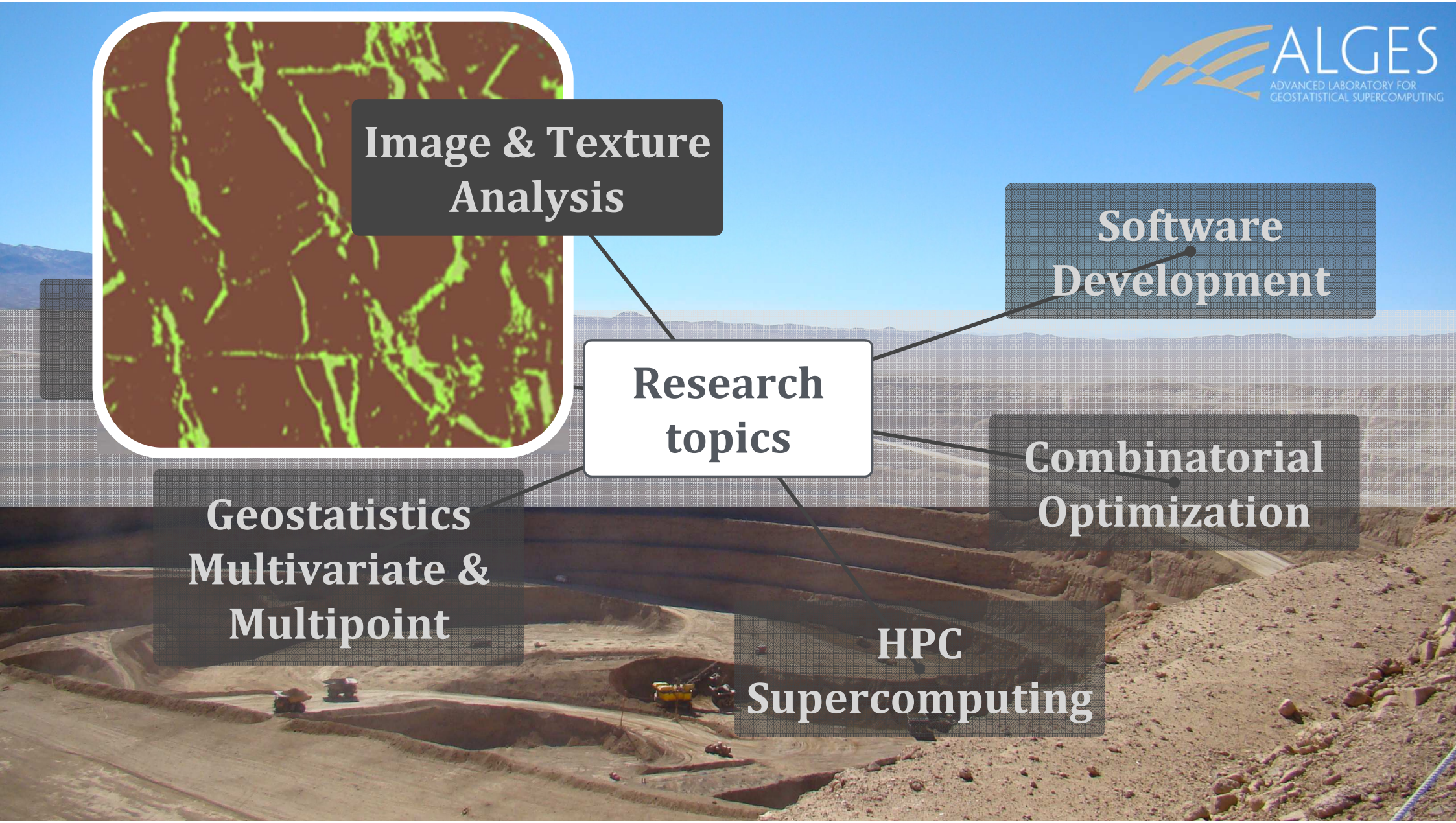
- Geology: Discovery
- Mining: Explotation
- Metallurgy: Processing


ALGES
ADVANCED LABORATORY FOR
GEOSTATISTICAL SUPERCOMPUTING
Research
topics


ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING
Image & Texture Analysis
Software Development
Data Mining
Research topics
Combinatorial Optimization
Geostatistics Multivariate & Multipoint
HPC Supercomputing


ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING
exture
sis
Software Development
Research topics
Combinatorial Optimization
Geostatistics Multivariate & Multipoint
HPC Supercomputing


ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING
exture
sis
Software Development
Data Mining
Research topics
Combinatorial Optimization
Geostatistics Multivariate & Multipoint
HPC Supercomputing


ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING
Image & Texture Analysis
Software Development
Research topics
Combinatorial Optimization
Geostatistics Multivariate & Multipoint
HPC Supercomputing

ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

Research topics

- Image & Texture Analysis
- Data Mining
- Geostatistics Multivariate & Multipoint
- HPC Supercomp
- Software Development

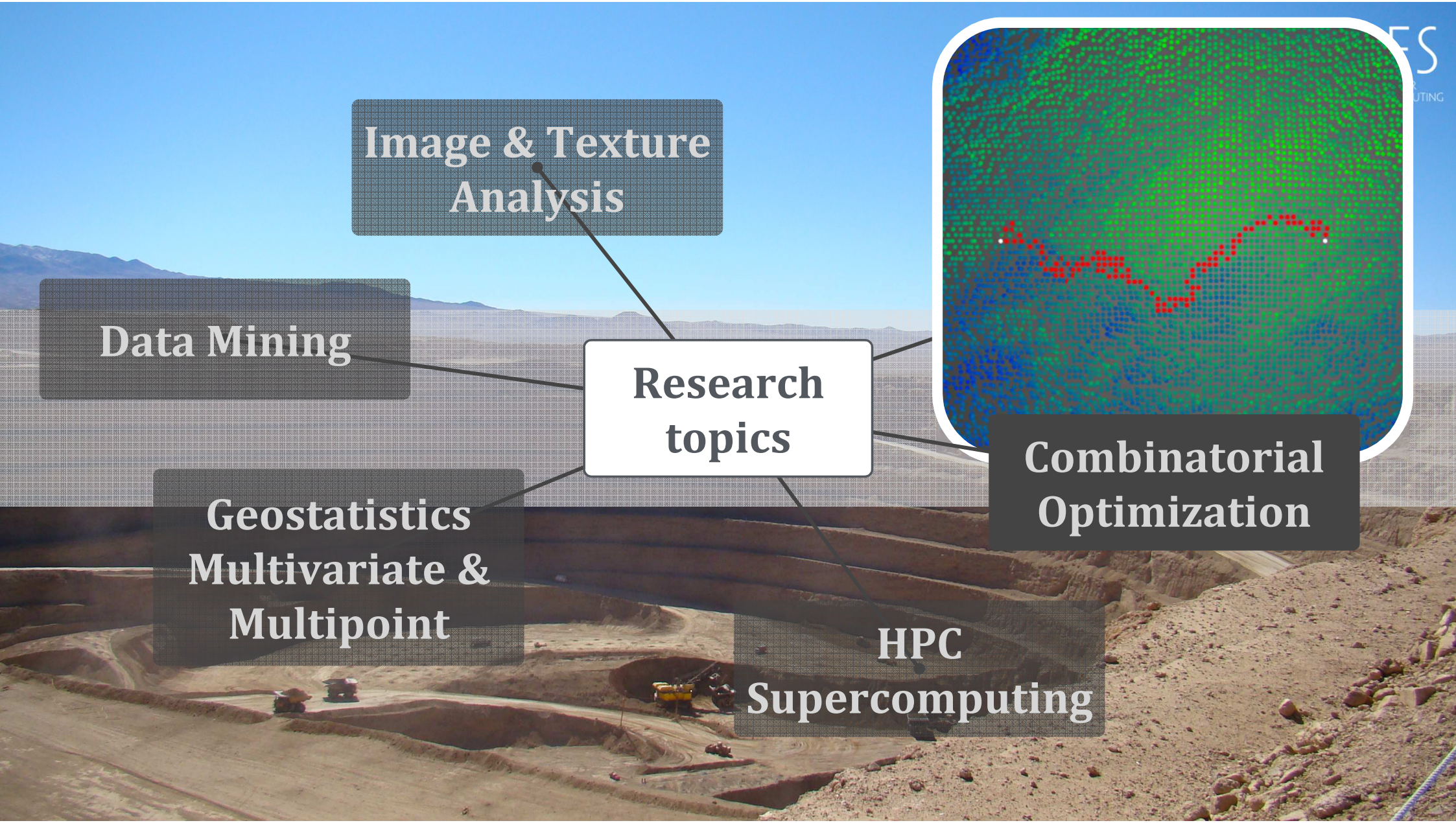


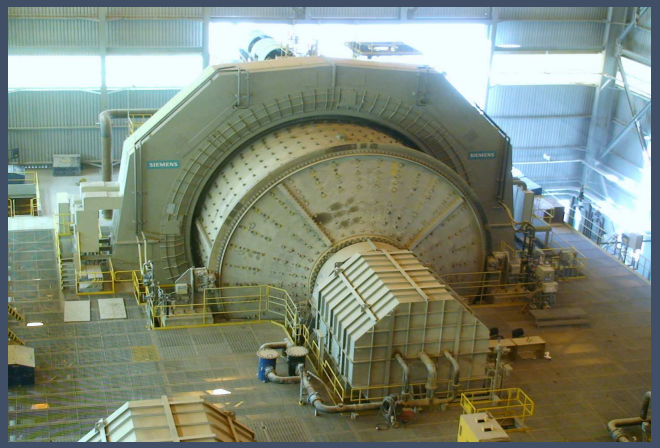
Image & Texture Analysis
Data Mining
Research topics
Combinatorial Optimization
Geostatistics Multivariate & Multipoint
HPC Supercomputing

ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

Research topics

- Image & Texture Analysis
- Software Development
- Data Mining
- Geostatistics Multivariate & Multipoint
- HPC Supercomputing

Environmental challenges

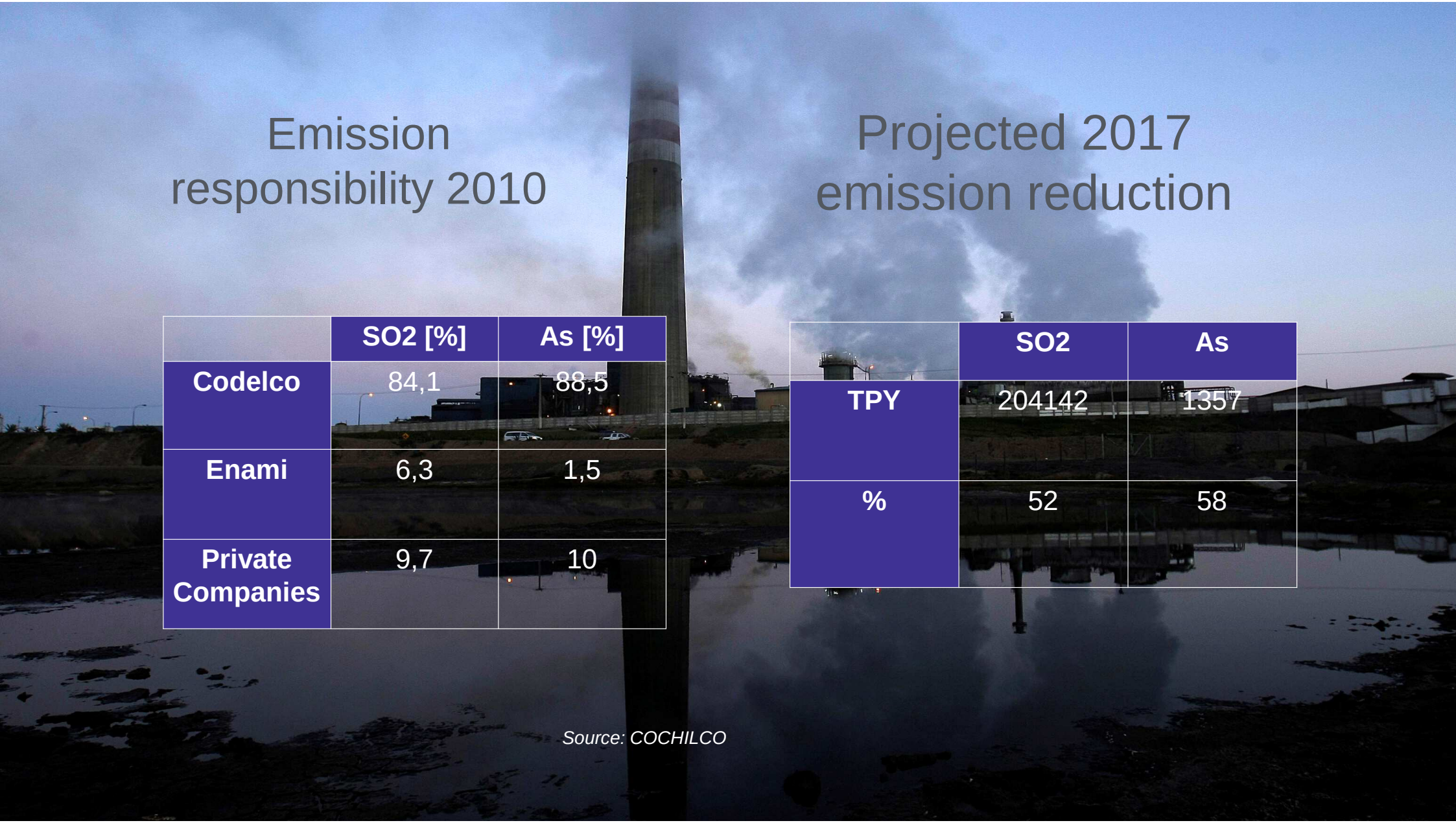


Environmental considerations copper pyrometallurgy

New standart: capture of SO_2 and As > 95% by weight for existing resources

	Chuquicamata	Caletones	Potrerillos	Paipotes	Ventanas	Altonorte	Chagres
Concentrate [kTPY]	1650	1600	680	357	450	1160	610
SO_2 Emissions [kTPY]	96500	124500	89500	24500	24500	24000	13950
As Emissions [kTPY]	812	365	790	35	35	128	107
Capture of SO_2 (%)	91	88	83.5	89,4	89,4	**93,7**	**95,7**

Source: COCHILCO



Emission responsibility 2010

	SO2 [%]	As [%]
Codelco	84,1	88,5
Enami	6,3	1,5
Private Companies	9,7	10

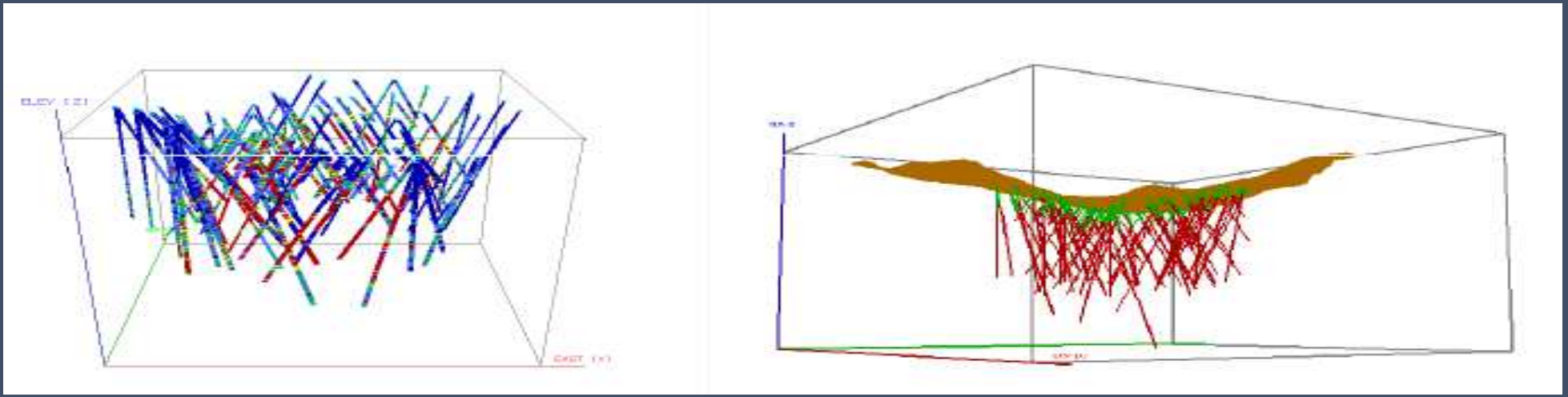
Projected 2017 emission reduction

	SO2	As
TPY	204142	1357
%	52	58

Source: COCHILCO

The Mining Process

Geological resource assesment



How can we know the amount of resources in a deposit?

Standart procedure.

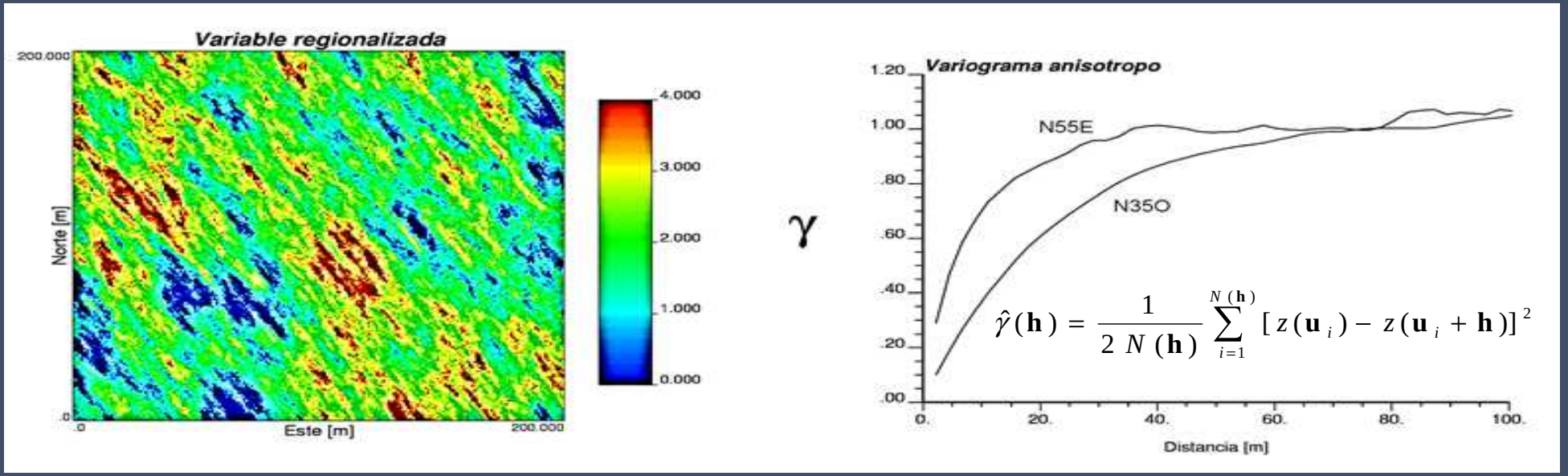
1. Interpretation of Deposit
2. Exploratory data analysis,
3. **Spatial continuity analysis**
4. Estimation
5. Estimation errors
6. Models validation

Data continuity

- Geostatistical estimation and simulation is based in spatial data continuity.
- The variogram is a tool to measure spatial distribution.
- The basic idea is to find the relationship between pairs of data at distance h

Variogram

- We can find **anisotropy**



Variogram and Estimation

- To estimate a value at a position **u** we use a linear combination of known values:

$$Z^{*}(\mathbf{u}) = a + \sum_{i=1}^{n} \lambda_i \cdot Z(\mathbf{u}_i)$$

- **Kriging** method incorporate spatial continuity and anisotropy criteria using the variogram:

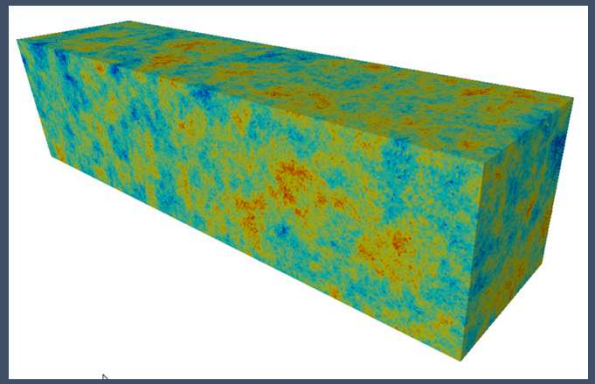
$$\begin{pmatrix} \gamma(\mathbf{u}_1 - \mathbf{u}_1) & \cdots & \gamma(\mathbf{u}_1 - \mathbf{u}_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(\mathbf{u}_n - \mathbf{u}_1) & \cdots & \gamma(\mathbf{u}_n - \mathbf{u}_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ -\mu \end{pmatrix} = \begin{pmatrix} \gamma(\mathbf{u}_1 - \mathbf{u}) \\ \vdots \\ \gamma(\mathbf{u}_n - \mathbf{u}) \\ 1 \end{pmatrix}$$

Ongoing Research Projects

HPC & Supercomputing

GPGPU Computing

- Granted as the first Chilean CUDA Research Center since 2012
- *Stochastic simulations:*
  - Turning Bands : **54x (88x dual GPU)**
  - MPS : **30x**
- *Variogram calculation:*
  - GAMV : **48x**

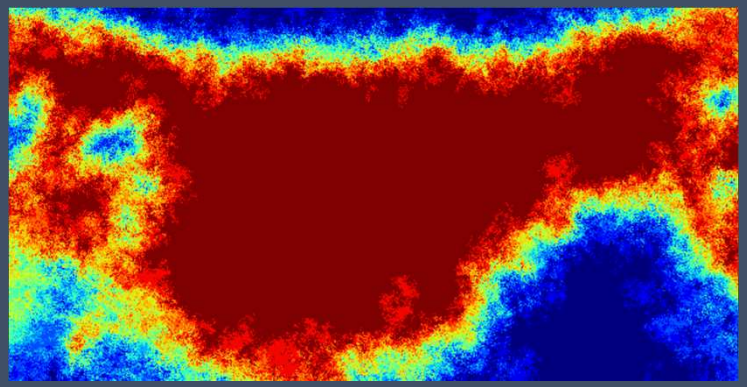


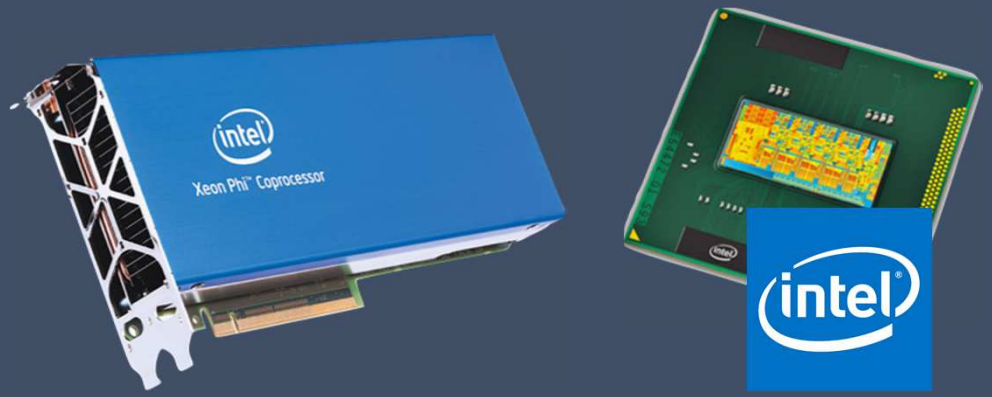
NVIDIA
CUDA™
RESEARCH
CENTER

(*) Timings using a NVIDIA Tesla C20 series GPUs

Multicore Computing

- GSLIB Resurrected: Legacy geostatistical code adapted for the multicore era
- *Stochastic simulations:*
  - SISIM : **15.5x**
  - SGSIM : **18x**
  - MPS : **26x**
- *Variogram calculation:*
  - GAMV : **19.3x**

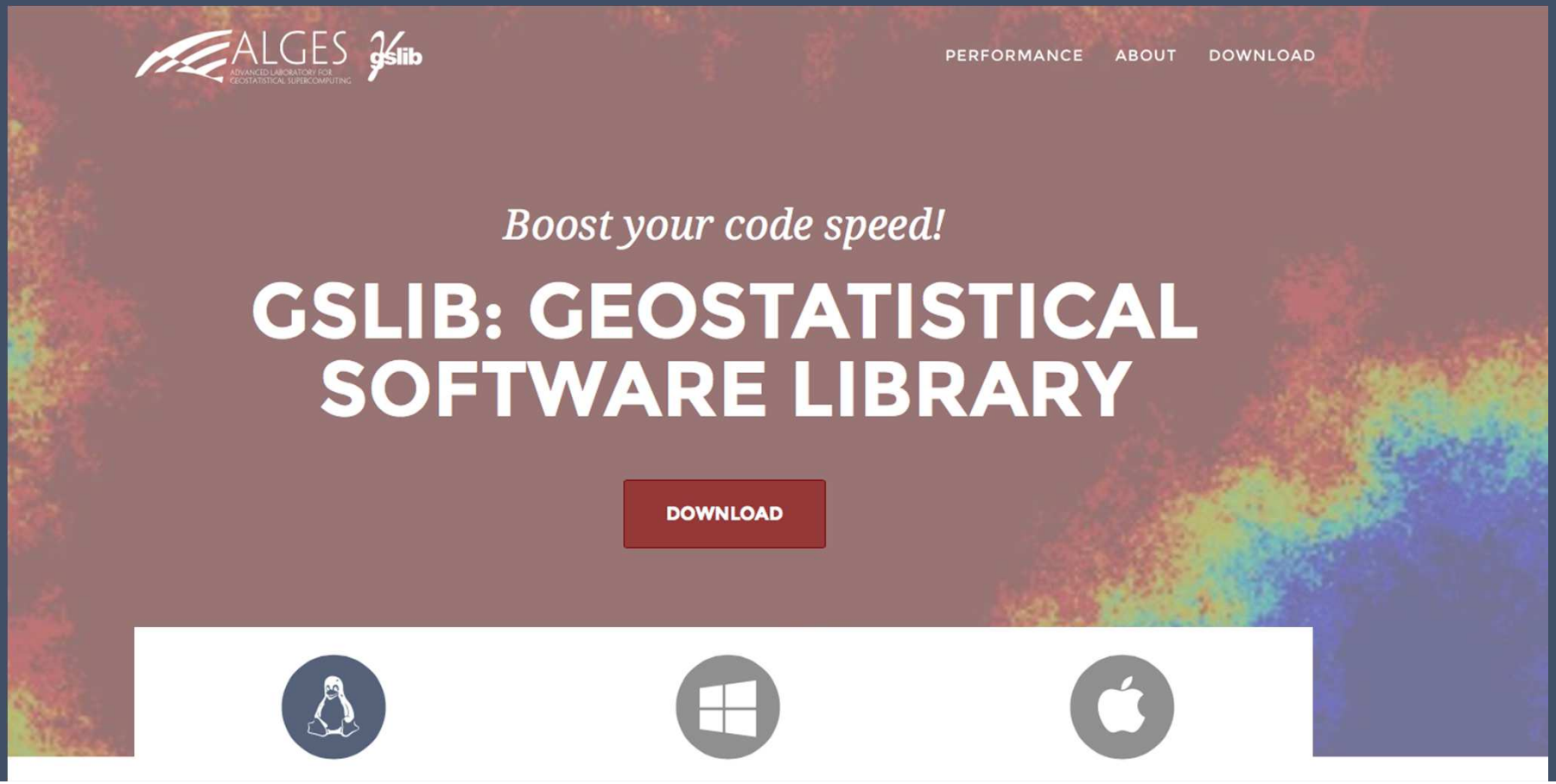




(*) Timings using a 2x8-core Intel Sandy Bridge-EP E5-2670

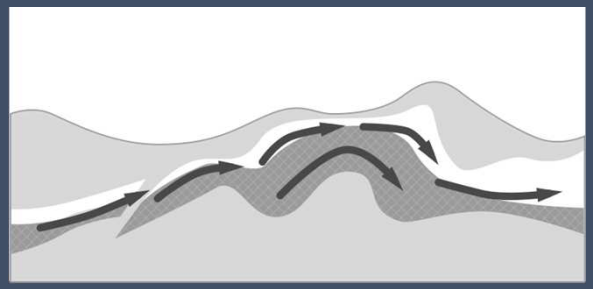
Multicore Computing

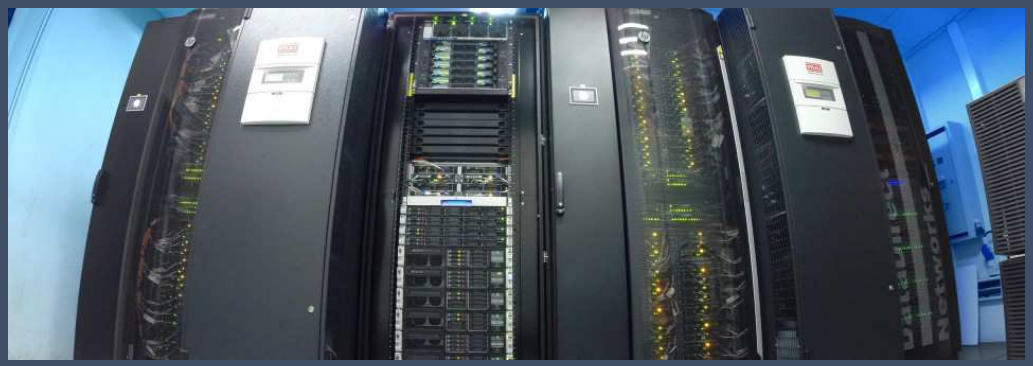
- Multicore GSLIB http://gslib.alges.cl



Distributed Computing

- Collaboration with BSC-Repsol
- Granted one of the first users of *Leftraru* (2nd fastest Latam supercomputer), 2640 cores
- *Stochastic simulations:*
  - SISIM : **32x**
  - SISIM-GFS : **80x**
- *Variogram calculation:*
  - GAMV-LVA: **2.2x (+150Gb distributed RAM)**

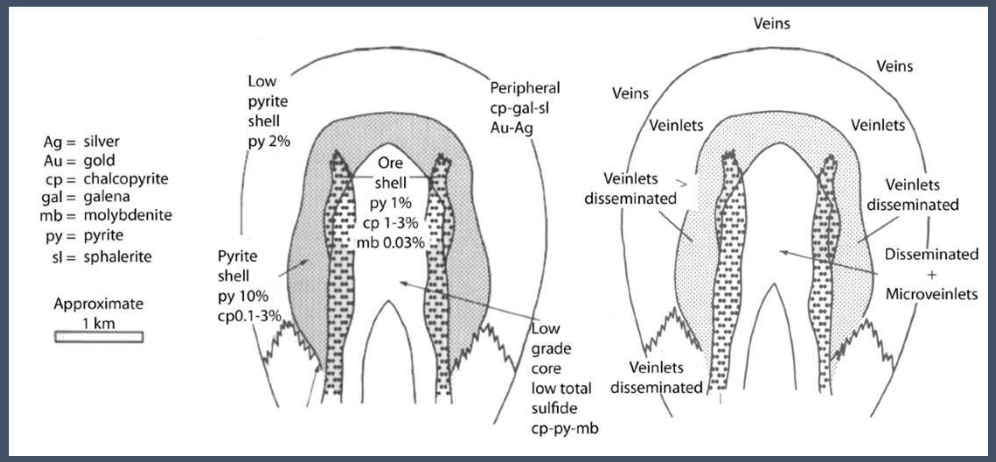


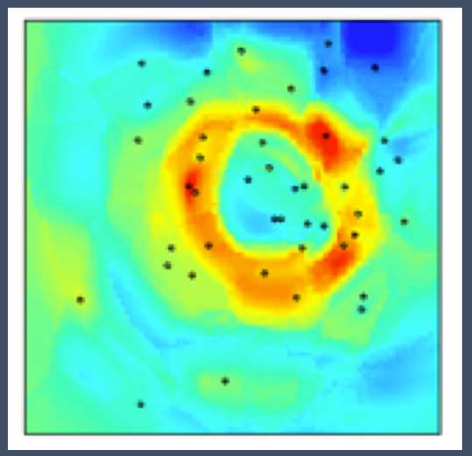


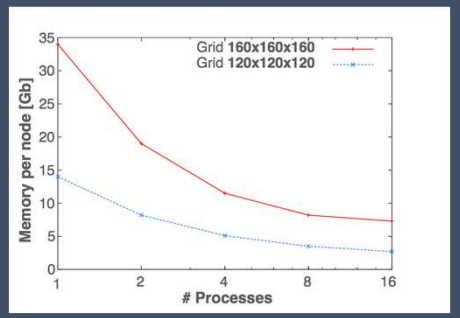
(*) Timings using up to 10 nodes with 2x10-cores Intel Ivy Bridge E5-2660V2

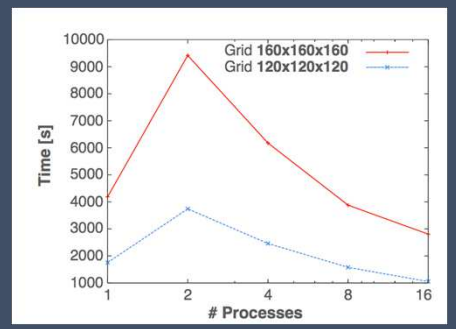
Distributed Computing

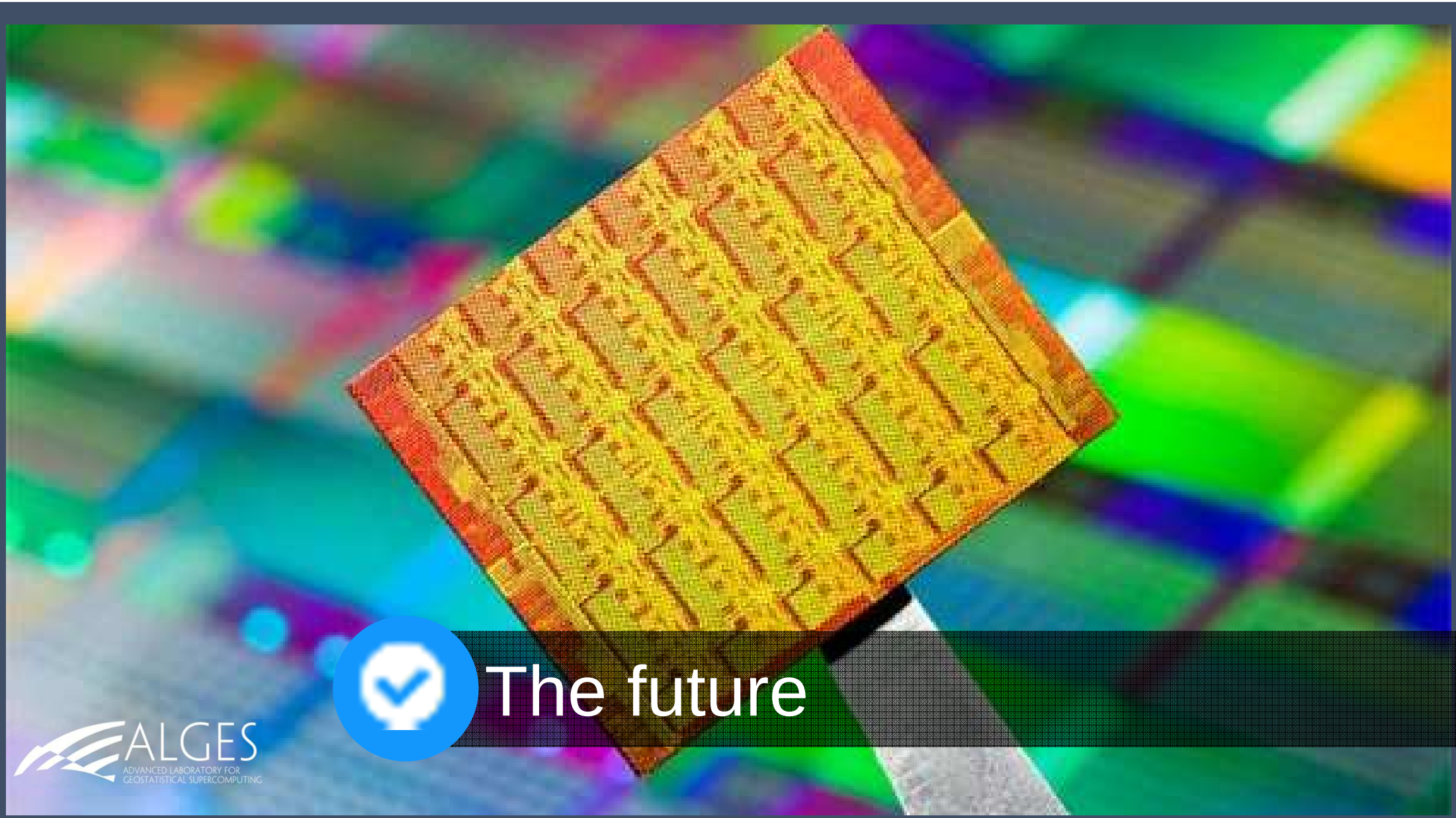
Locally Varying Anisotropy:
an opportunity for HPC in resource evaluation
Porphiry copper case









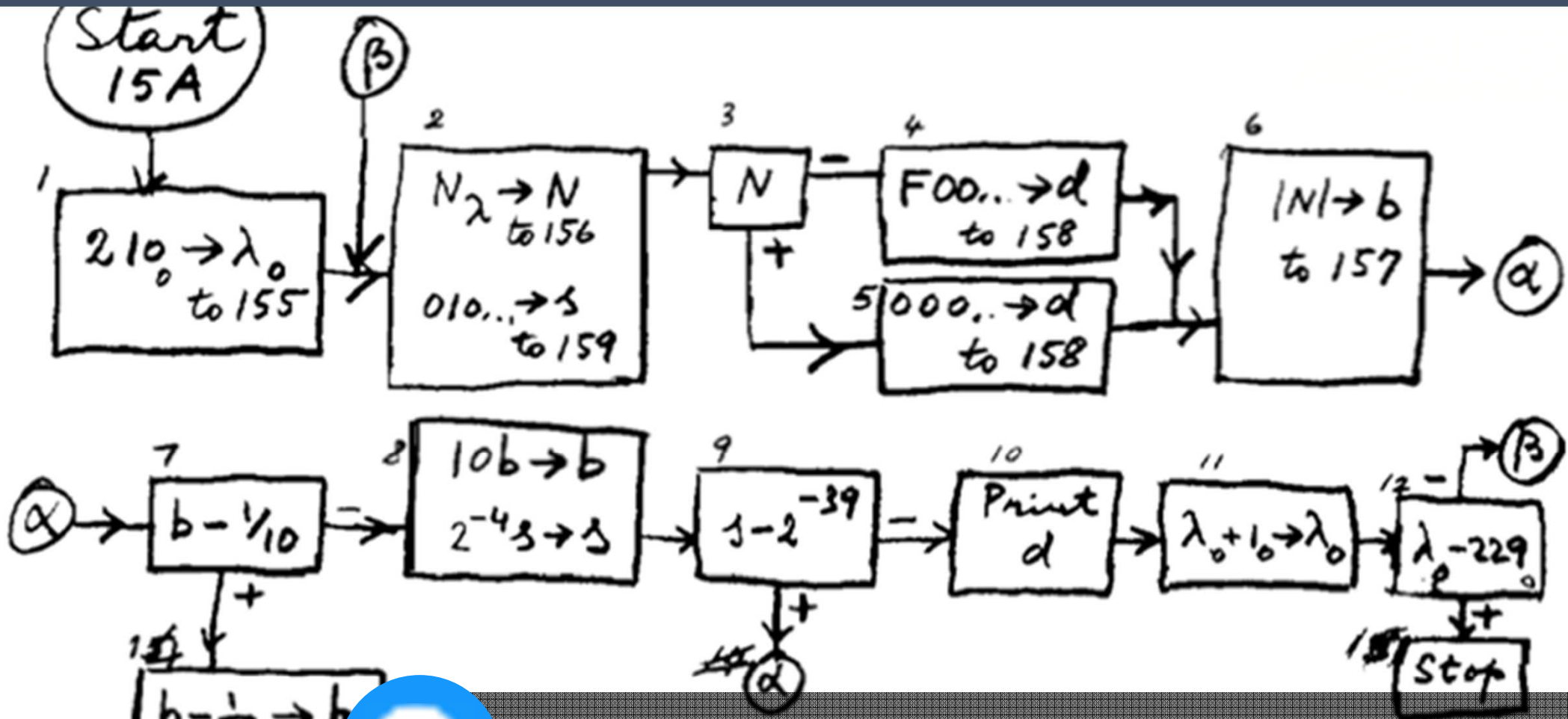


The future

ALGES
ADVANCED LABORATORY FOR GEOSTATISTICAL SUPERCOMPUTING

The Future

- Explore larger architectures with OpenMP, MPI, CUDA and Giulietta framewfork (accelerators, clusters, cloud, supercomputers…)
- To be a reference in the mining landscape enlightening about the “multicore era” and HPC/Supercomputing new technologies
- Collaborate with other mining groups to accelerate new applications (geological modelling, mine planning, exploration and others)
- Explore geostatistics application beyond mining (agriculture, geographic data, environment and others)


Start
15A
β
1
$210_0 \rightarrow \lambda_0$
to 155
2
$N_\lambda \rightarrow N$
to 156
$010.. \rightarrow s$
to 159
3
N
4
$F00.. \rightarrow d$
to 158
5
$000.. \rightarrow d$
to 158
6
$|N| \rightarrow b$
to 157
α
7
$b - 1/10$
8
$10b \rightarrow b$
$2^{-4}s \rightarrow s$
9
$s - 2^{-39}$
10
Print
d
11
$\lambda_0 + 1_0 \rightarrow \lambda_0$
12
$\lambda_0 - 229_0$
β
Stop
$b - \frac{1}{10} \rightarrow b$
$d + s \rightarrow d$

Questions?