A Process Chain for End-to-End Sensing in Disruptive Environments

Jan Beutel, ETH Zurich

The Issue

- WSNs as a new tool for distributed sensing
- E.g. in environmental science
  - High spatial coverage through many sensors
  - Good temporal resolution/coverage (high rates, long-term observations)
  - Autonomous (disconnected) operation

- Initially it was thought that imperfections in the data are eliminated by heavy oversampling and use of aggregates [c.f. Smart Dust, Pister et al., 1999 and others]
- However this is only theory (so far)

PermaSense

- Consortium of several projects, start in 2006
- Multiple disciplines (geo-science, engineering)
- Fundamental as well as applied research
- More than 20 people, 9 PhD students

http://www.permasense.ch

Competence in Outdoor Sensing

- Wireless systems, low-latency data transmission
- Customized sensors
- Ruggedized equipment
- Data management
- Planning, installing, operating (years) large deployments

Understanding Root Causes of Catastrophes

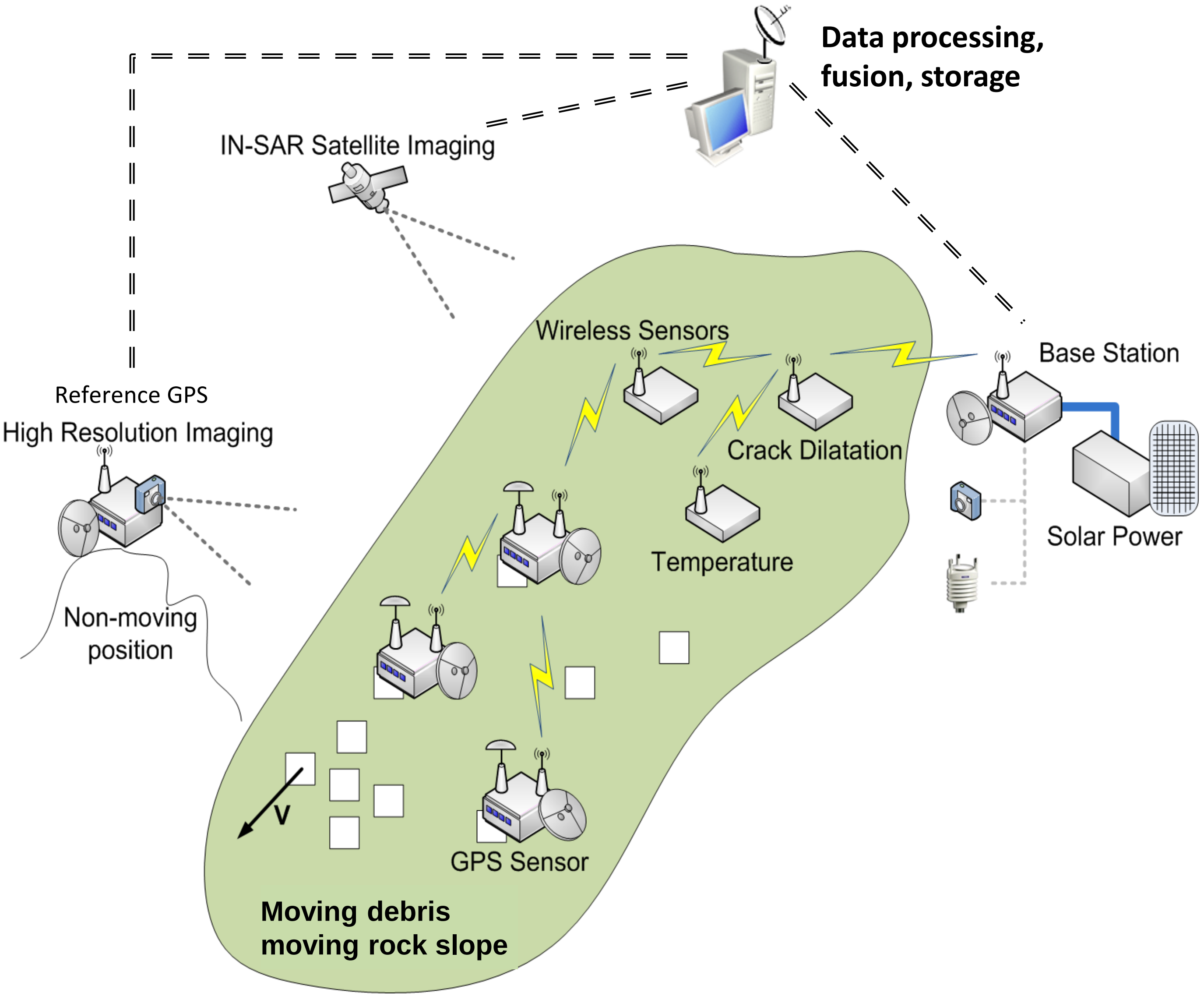
Eiger east-face rockfall, July 2006, images courtesy of Arte Television

What is the Role of Ice Filled Clefts?


Our patient does not fit
into a laboratory


So the laboratory has to go on the mountain

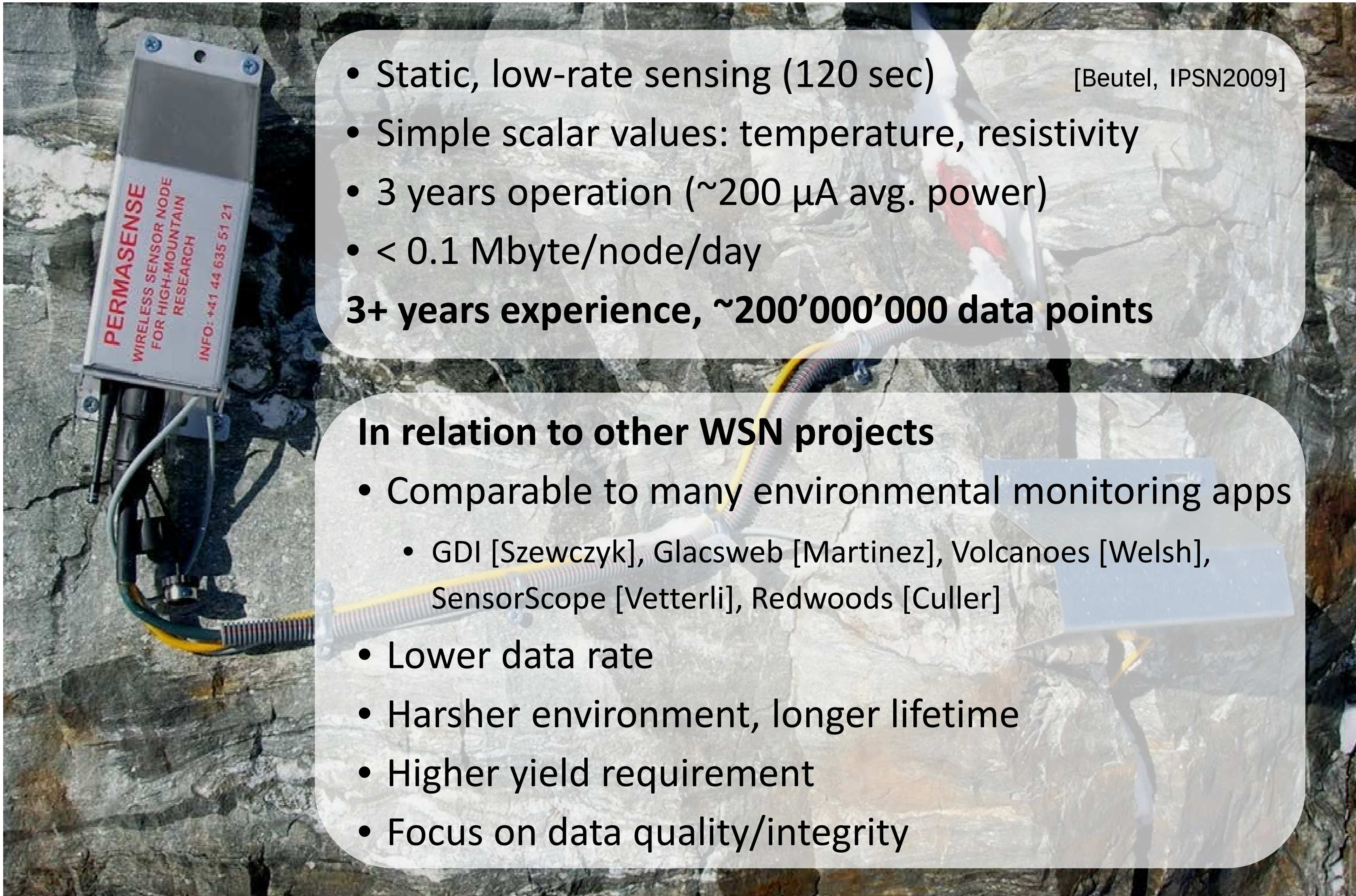
Example: The X-Sense Platform



Current Practice

- A single sensing point is still expensive despite high integration and high-volume, lower-cost hardware
  - Customization/heterogeneity
  - Low volume (of customized units)
  - Infrastructure requirements
  - Considerable end-to-end system complexity
  - Adequate protection (enclosures, connectors)
  - Installation/maintenance effort
- Substantial contribution of installation/maintenance effort to the TCO of WSNs [c.f. Stankovic, Vetterli, Welsh, Culler]
  - Installation = 1 man-day/sensor
  - In most cases much more

Simple Low-Power Wireless Sensors



- Static, low-rate sensing (120 sec) [Beutel, IPSN2009]
- Simple scalar values: temperature, resistivity
- 3 years operation (~200 µA avg. power)
- < 0.1 Mbyte/node/day

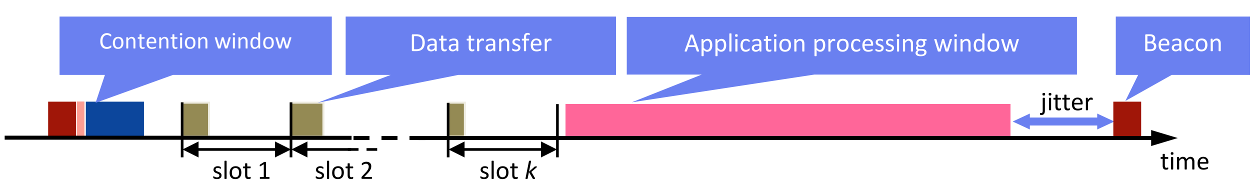
3+ years experience, ~200'000'000 data points

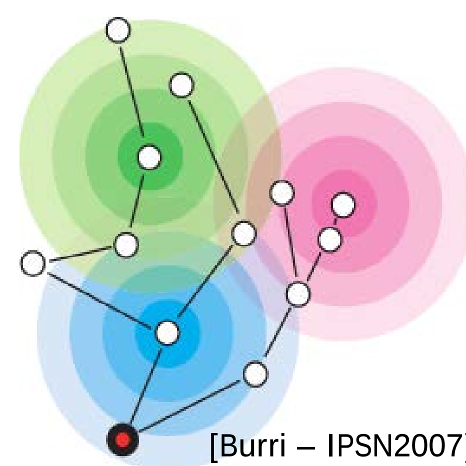
In relation to other WSN projects

- Comparable to many environmental monitoring apps
  - GDI [Szewczyk], Glacsweb [Martinez], Volcanoes [Welsh], SensorScope [Vetterli], Redwoods [Culler]
- Lower data rate
- Harsher environment, longer lifetime
- Higher yield requirement
- Focus on data quality/integrity

Low-power WSN Technology

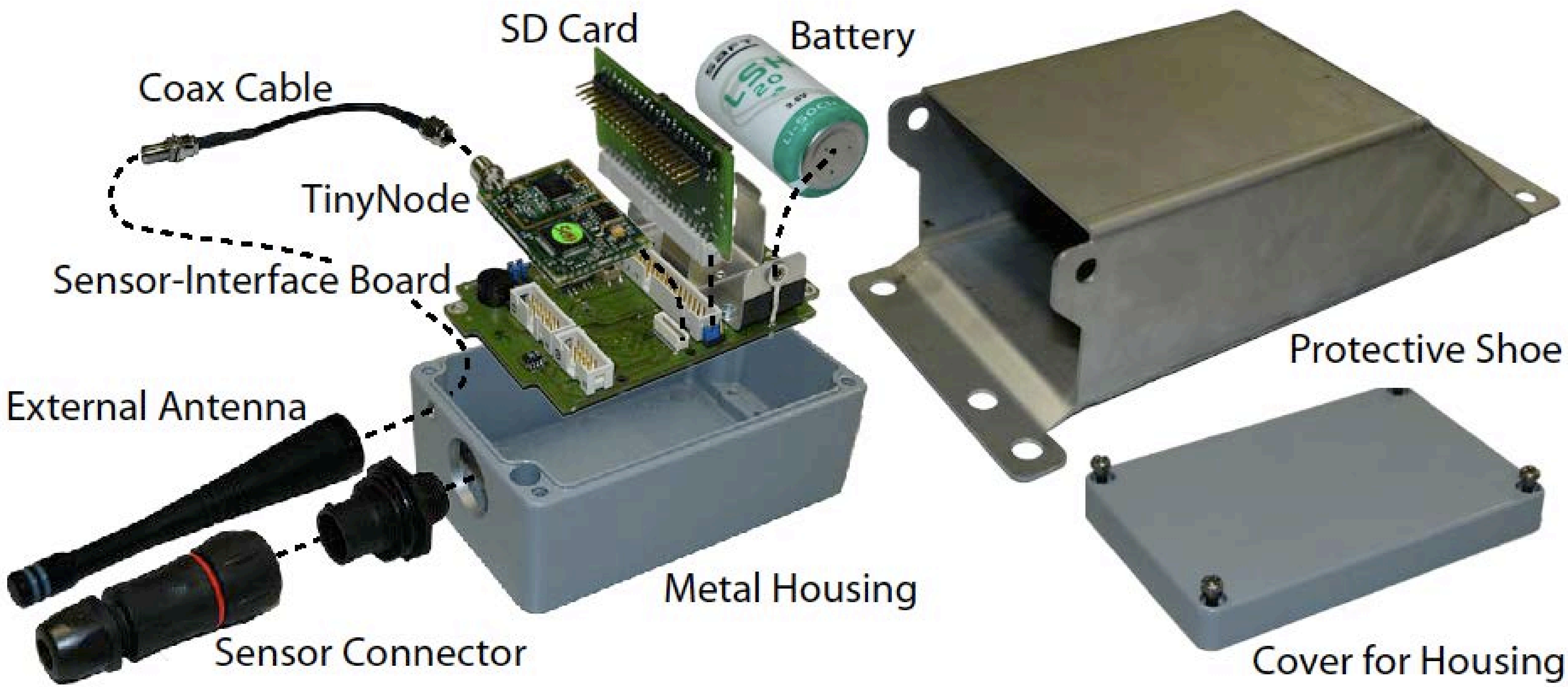
- Shockfish TinyNode184
  - MSP430, 16-bit, 8MHz, 8k SRAM, 92k Flash
  - LP Radio: SX1211 @ 868 MHz
- Sensor interface board
  - 1 GB storage
- 3-year life-time
- Dozer - ultra low-power data gathering system
  - Multi-hop, beacon based, 1-hop synchronized TDMA
  - Optimized for ultra-low duty cycles
  - **0.167%** duty-cycle, **0.032mA**





[Burri – IPSN2007]

Ruggedized for Alpine Extremes



Field Site Support

- Base station
  - On-site data aggregation
  - Embedded Linux
  - Solar power system
  - Redundant connectivity
  - Local data buffer
  - Database synchronization
- Cameras
  - PTZ webcam
  - High resolution imaging (D-SLR)
- Weather station
- Remote monitoring and control

Installation/Maintenance Effort

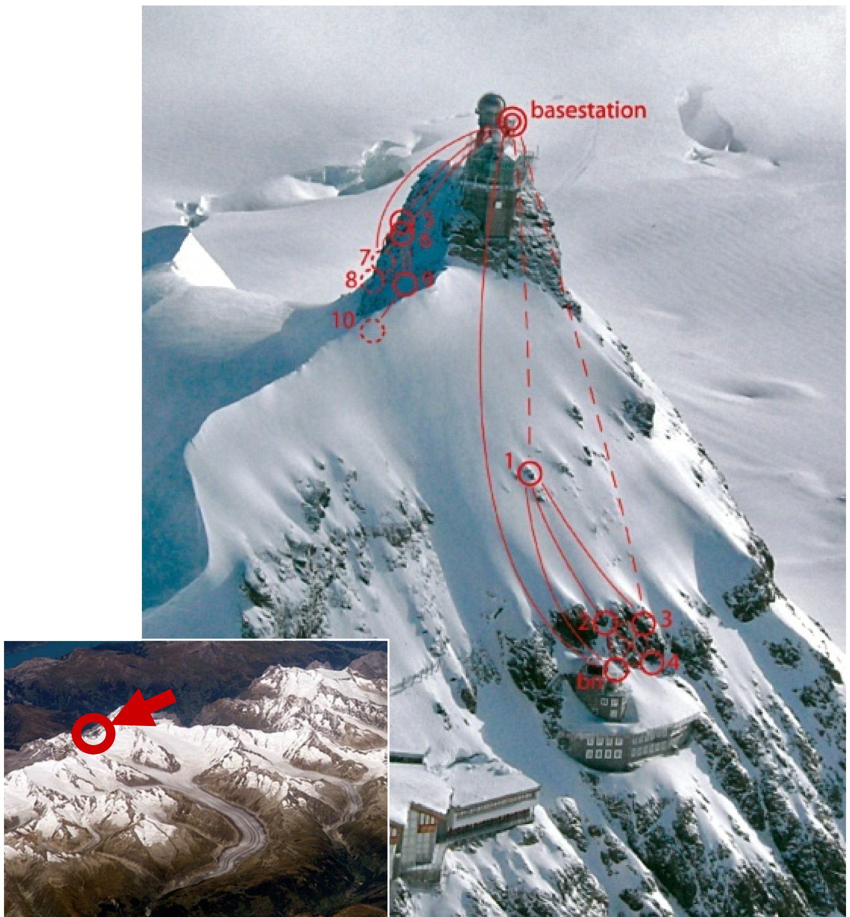
Towards Higher Reliability

- Many applications require “accuracy”
  - Accuracy at the sample level (calibration, repeatability)
  - Accuracy at the ensemble level (deterministic behavior)
  - Specific knowledge of the sensing “location”

- Users require homogeneous data quality, e.g. uniform rate primary data without holes
  - It’s a long time from theory to practice for ideas like stochastic sampling to be accepted by domain users
  - Accurate timing is a must have
  - It is next to impossible to quantify performance & maintain quality operation if failures are acceptable behavior

Deployment Sites 3500 m a.s.l.

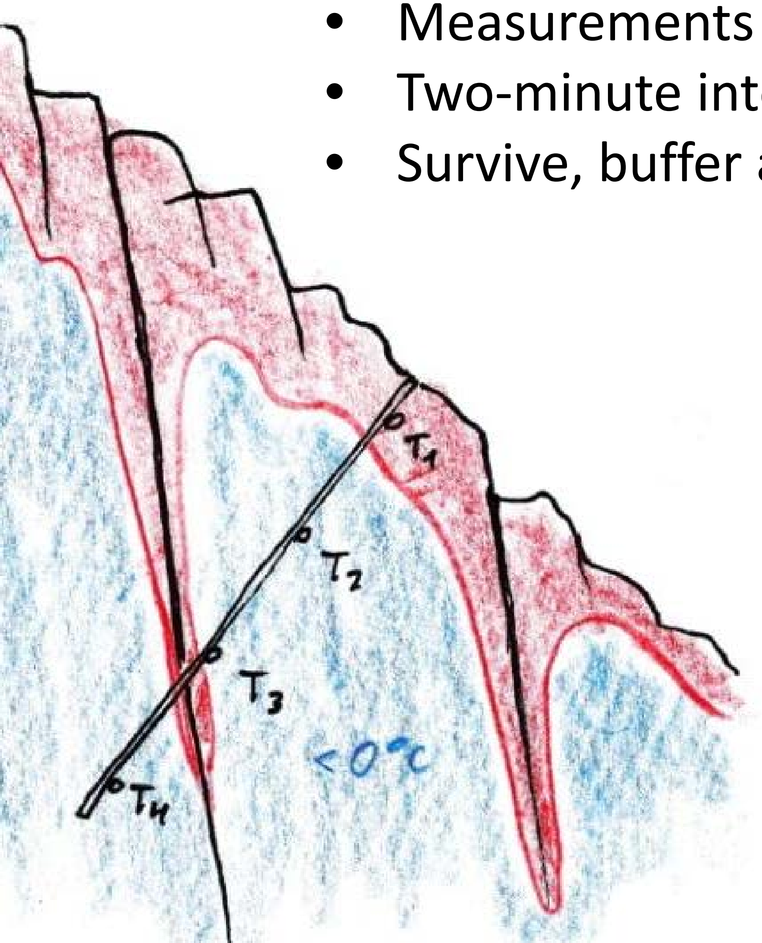
A **scientific instrument** for **precision sensing** and data recovery in environmental extremes

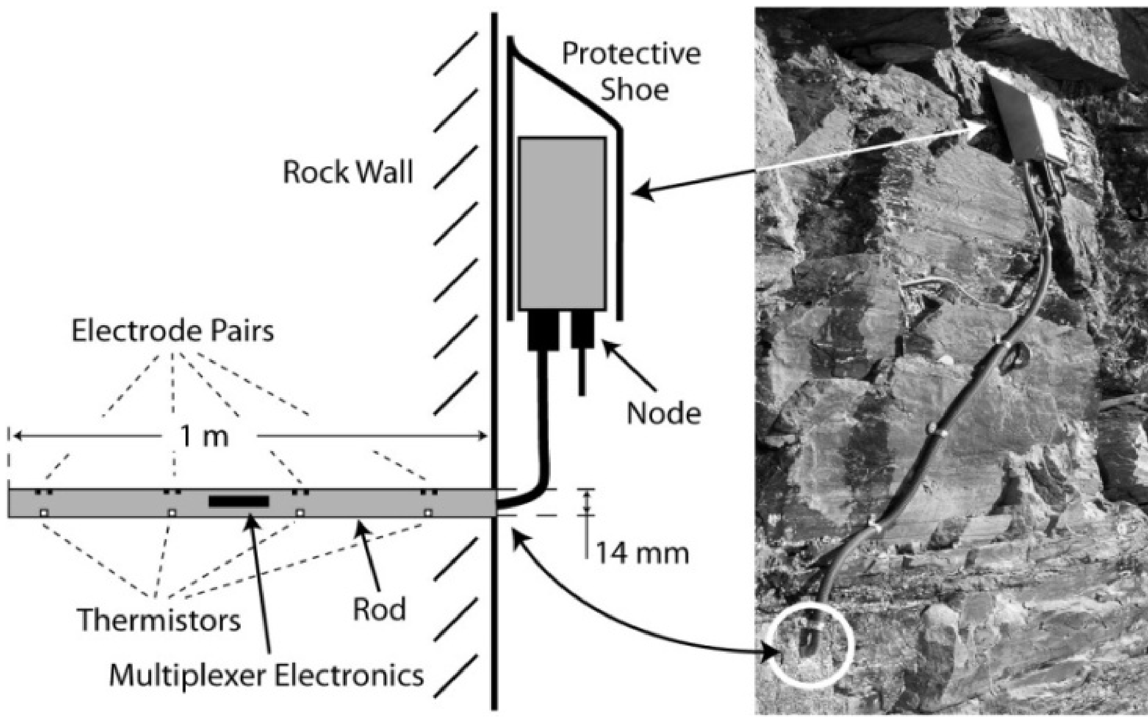


Established: Rock/ice Temperature

Aim: Understand temperatures in heterogeneous rock and ice

- Measurements at several depths
- Two-minute interval, autonomous for several years
- Survive, buffer and flush periods without connectivity





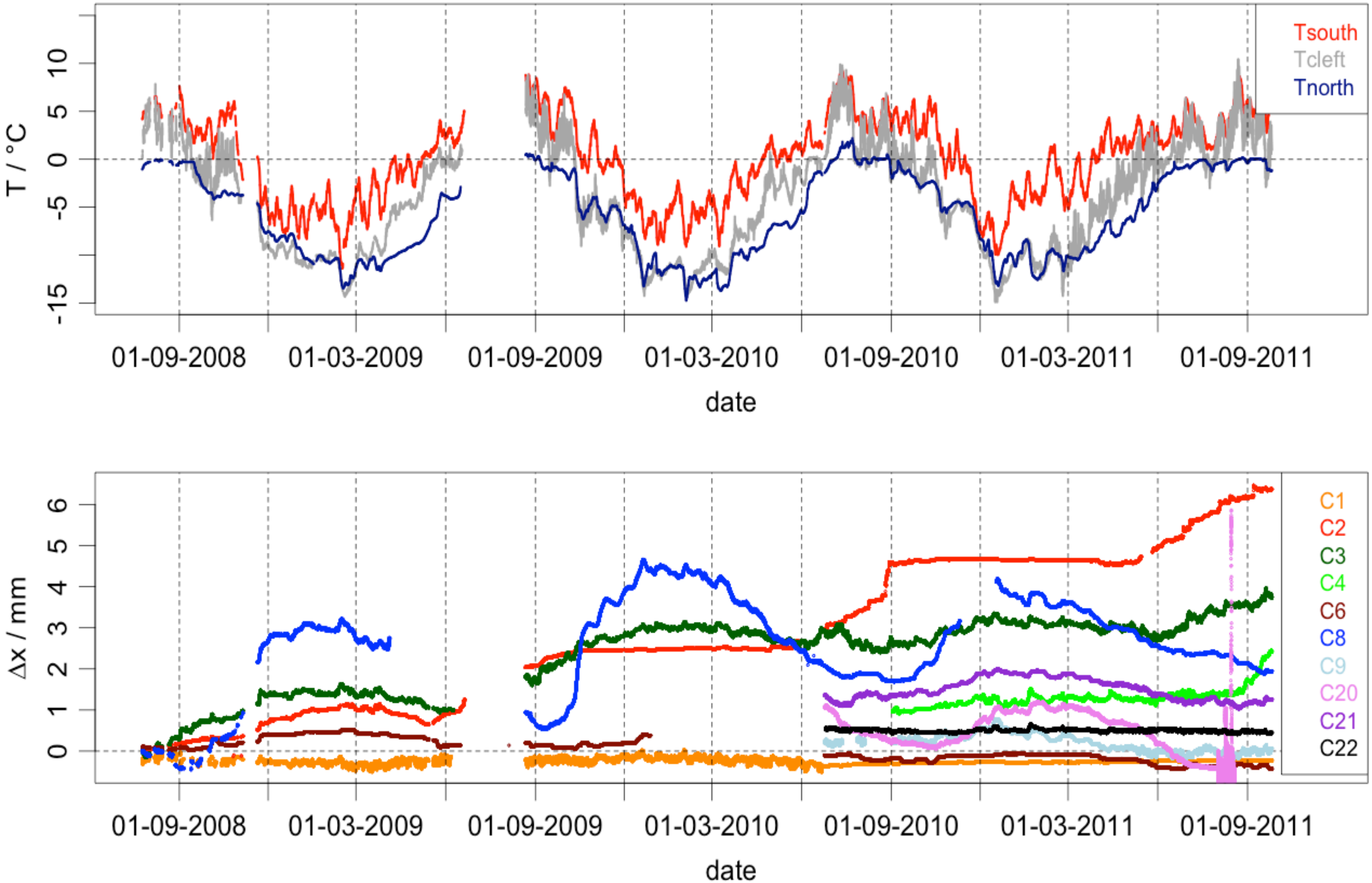
[Hasler 2011]

Established: Crack Dilatation

Aim: To understand temperature/ice-conditioned rock kinematics

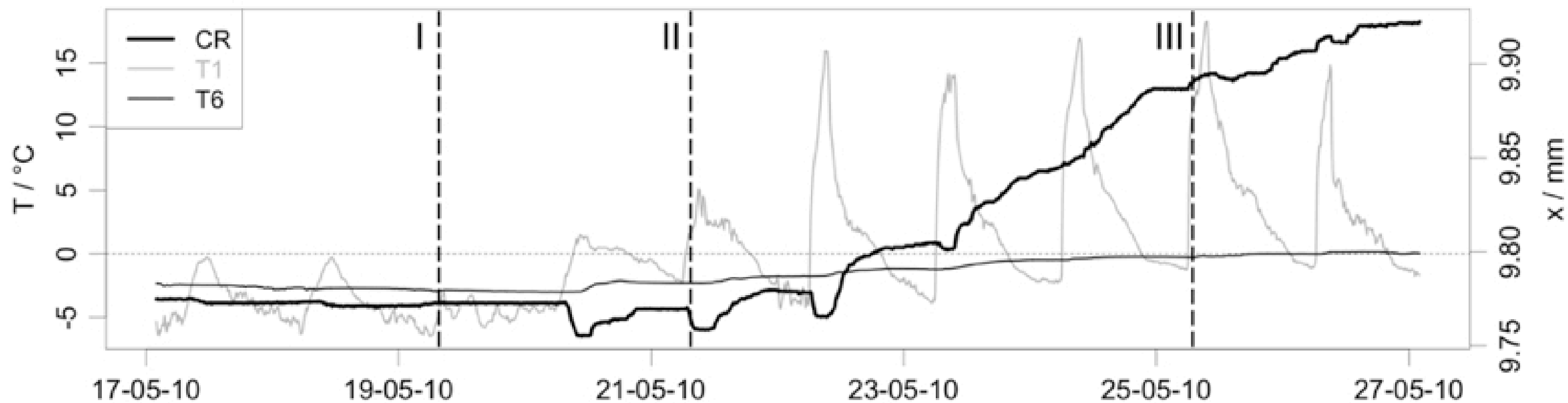
- Temperature-compensated, commercial instrument
- Auxiliary measurements (temperature, additional axes,…)
- Two-minute interval, autonomous for several years
- Protection against snow-load and rock fall

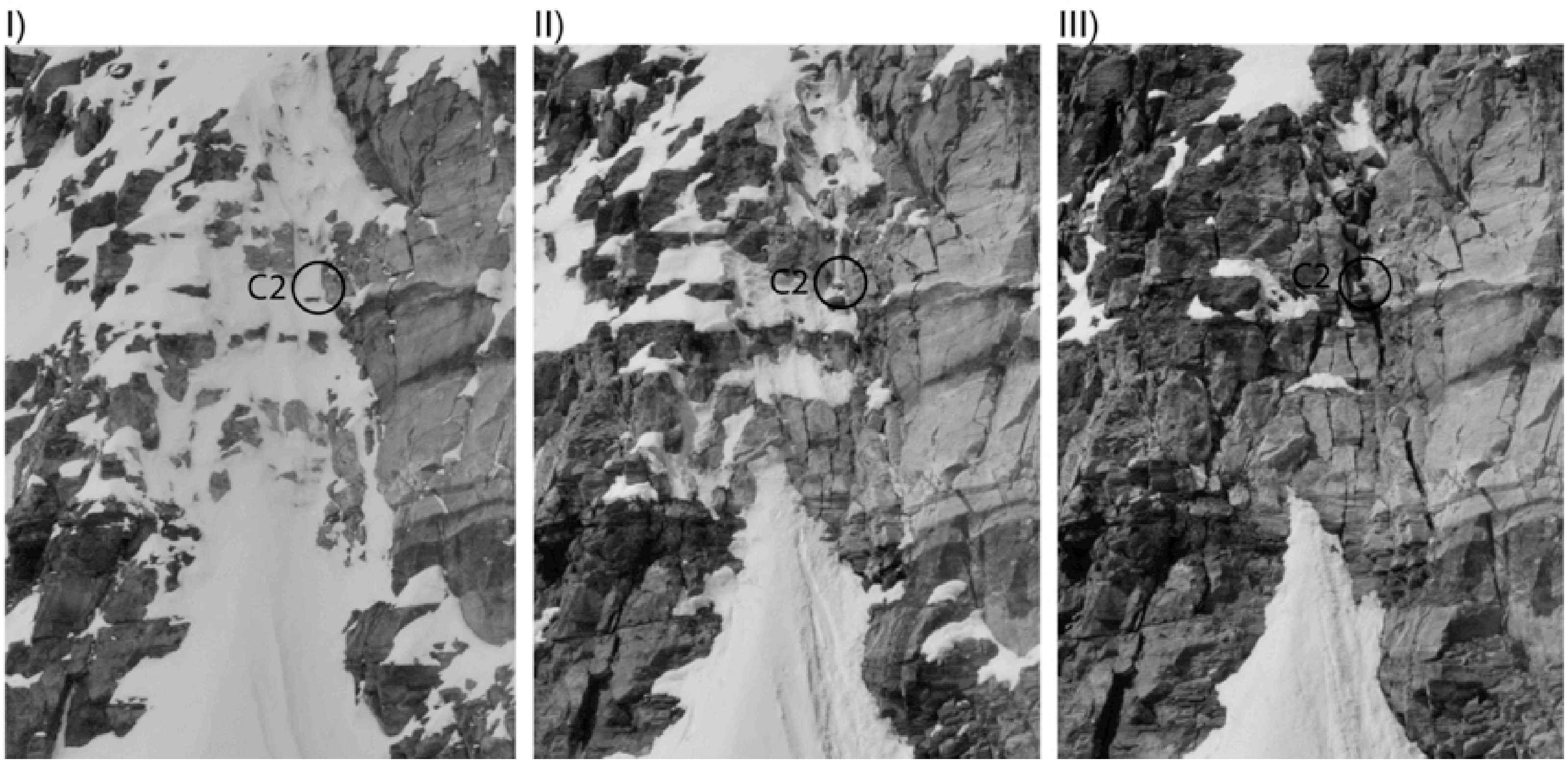
Results: Rock Kinematics



[**Hasler, A., Gruber, S. & Beutel, J.** Kinematics of steep bedrock permafrost, *Journal of Geophysical Research*]

Observation: Acceleration Behavior



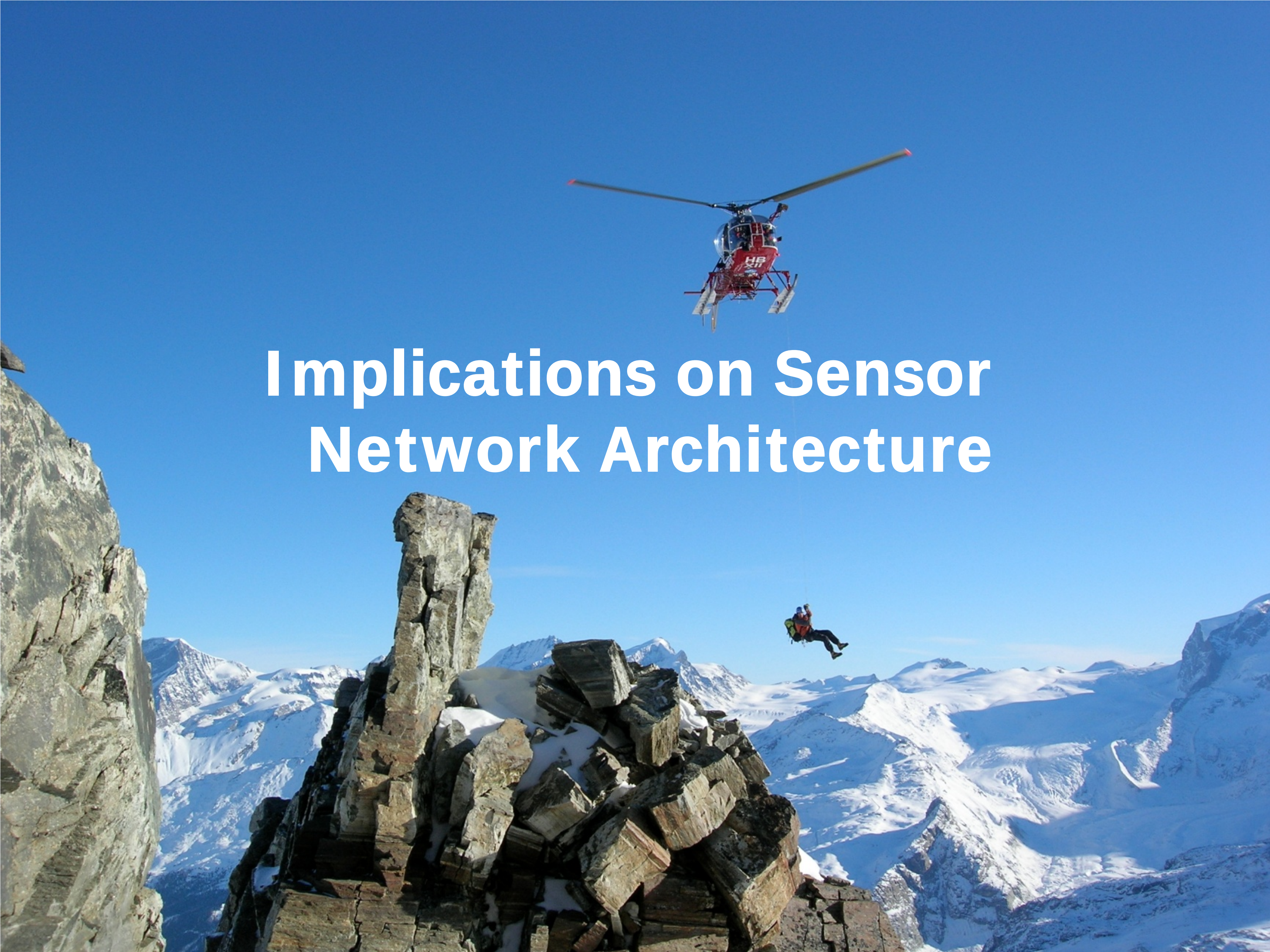


Assumptions/Hypothesis

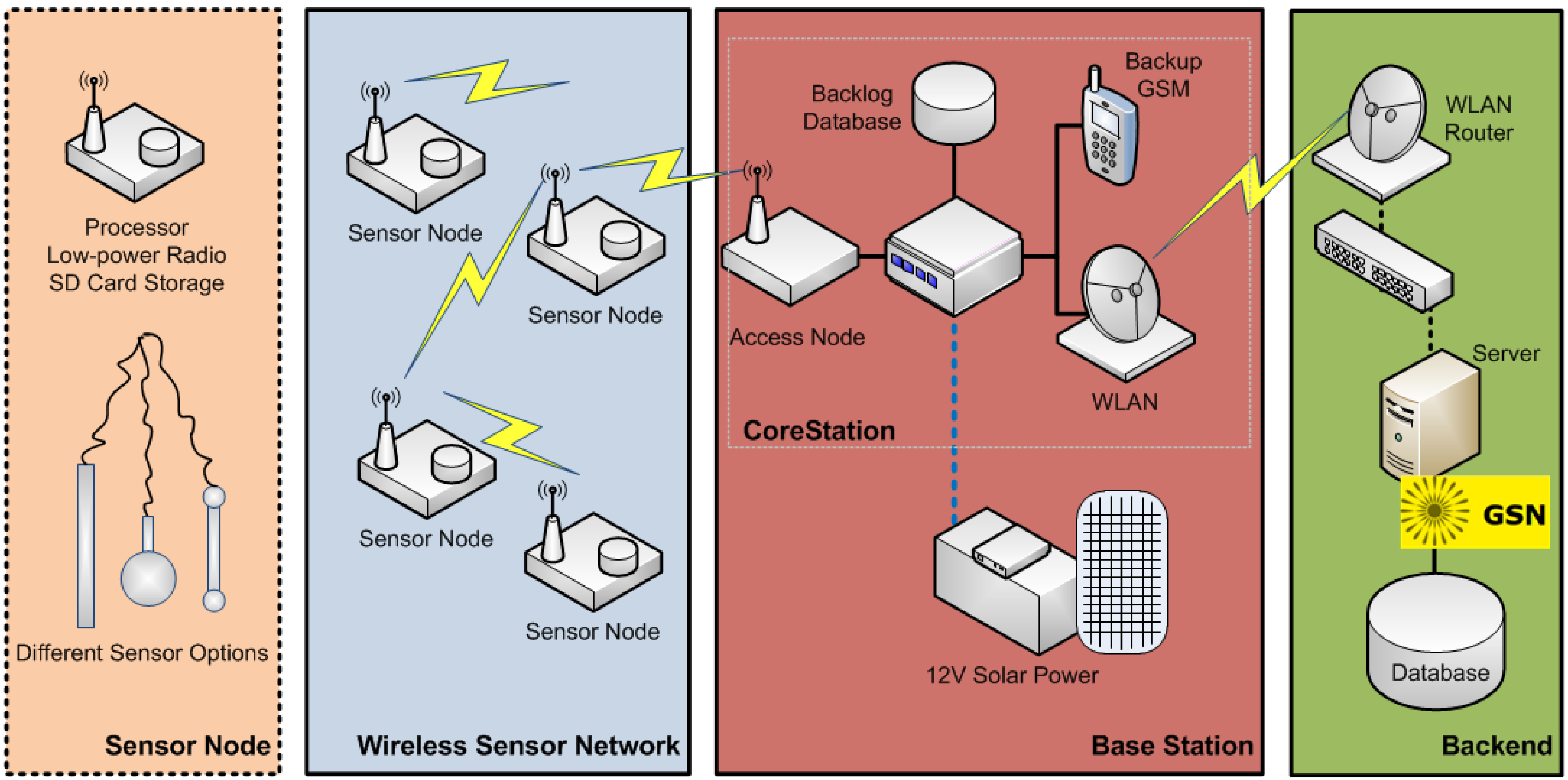
- High up-front investments call for **reliable interaction** of **all system components** at **all layers**
  - Local buffer storage
  - Data synchronization, acknowledgements
  - No single points of failure, redundancy (also in access networks and servers)
  - Timing integrity
  - Data validation

- Knowledge about the “origin” nature of all **primary data along the whole processing chain** is key
  - Traceability, quality metrics, data integrity
  - Accounting for human-in-the-loop

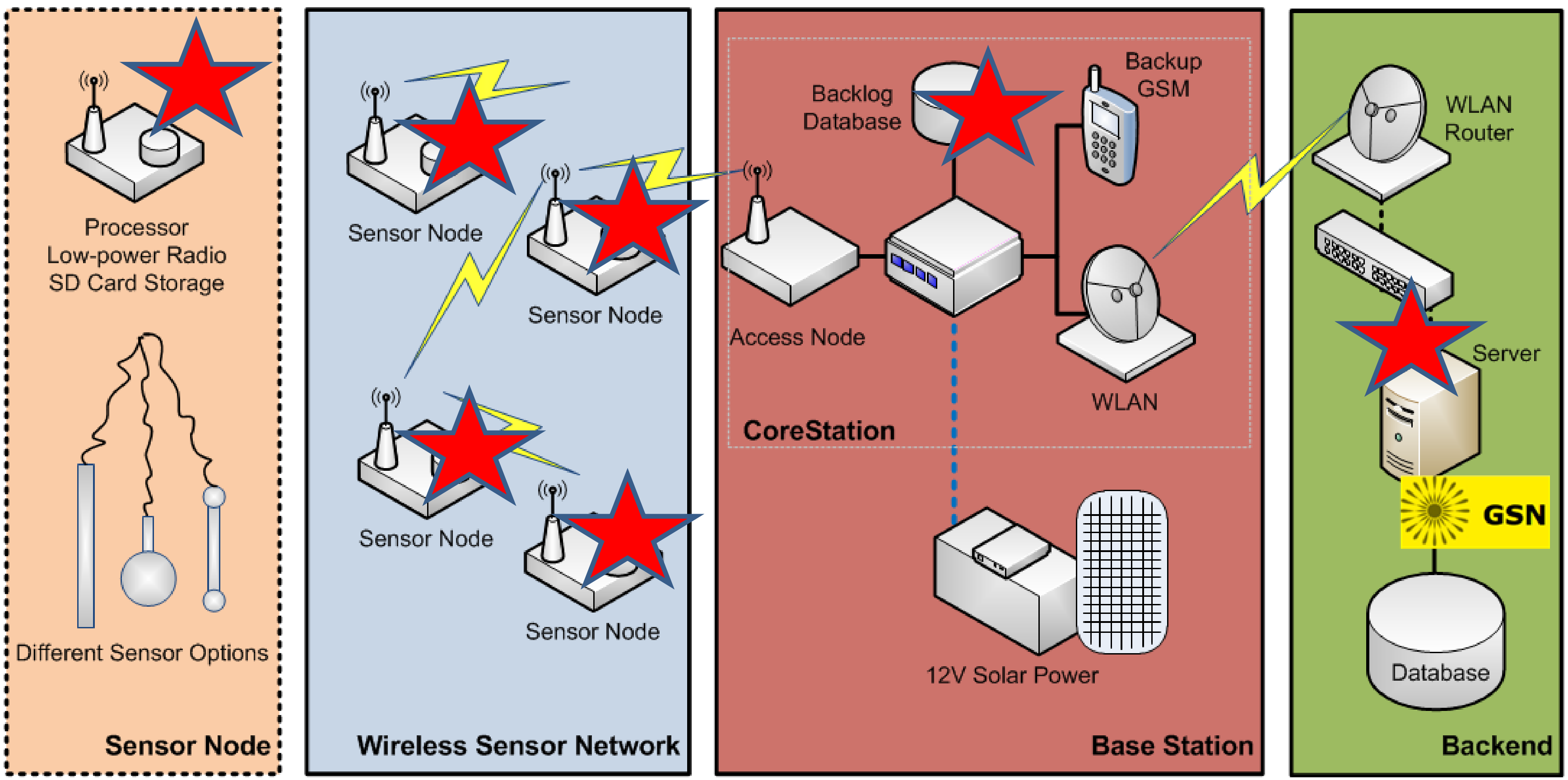
Implications on Sensor Network Architecture



PermaSense System Architecture



Local Data Storage on Every Layer



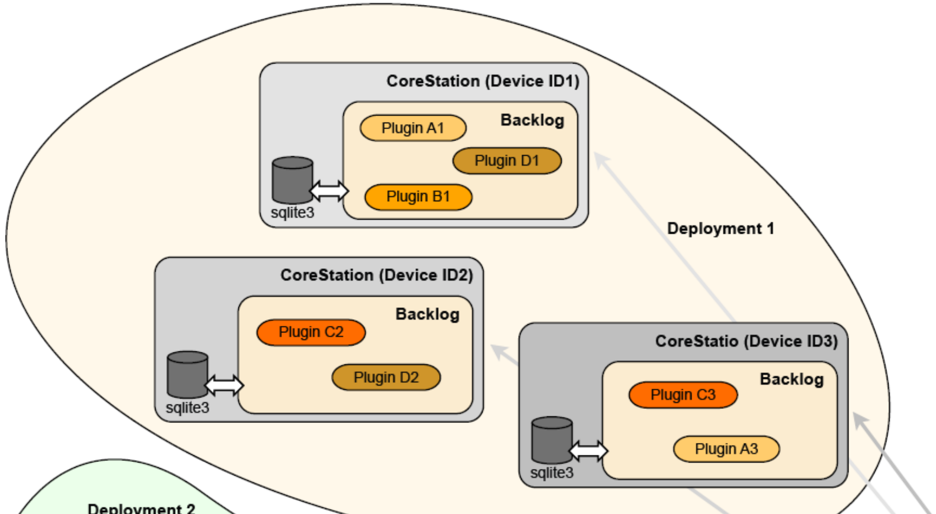
WSN On-Node Storage Layer

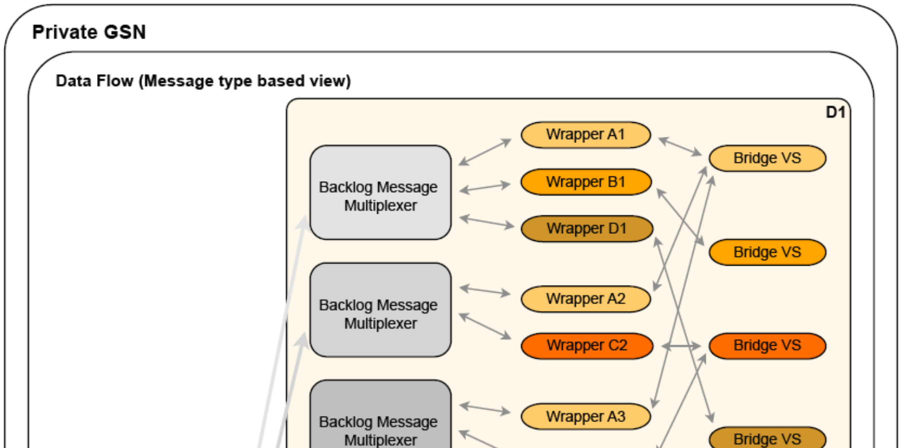
- On-node flash based storage (SD-Card)
  - Integrated with Dozer queuing mechanism (beacon traces & per-link ack’s with backpressure)
  - All generated packets are stored on local flash memory
  - Packets not yet sent are flagged for sending later
  - Bulk access optimized for flash memory (no single packet transfers)
- Enables both delayed sending (disruptions) and post-deployment validation

Mitigating Post WSN Data Loss

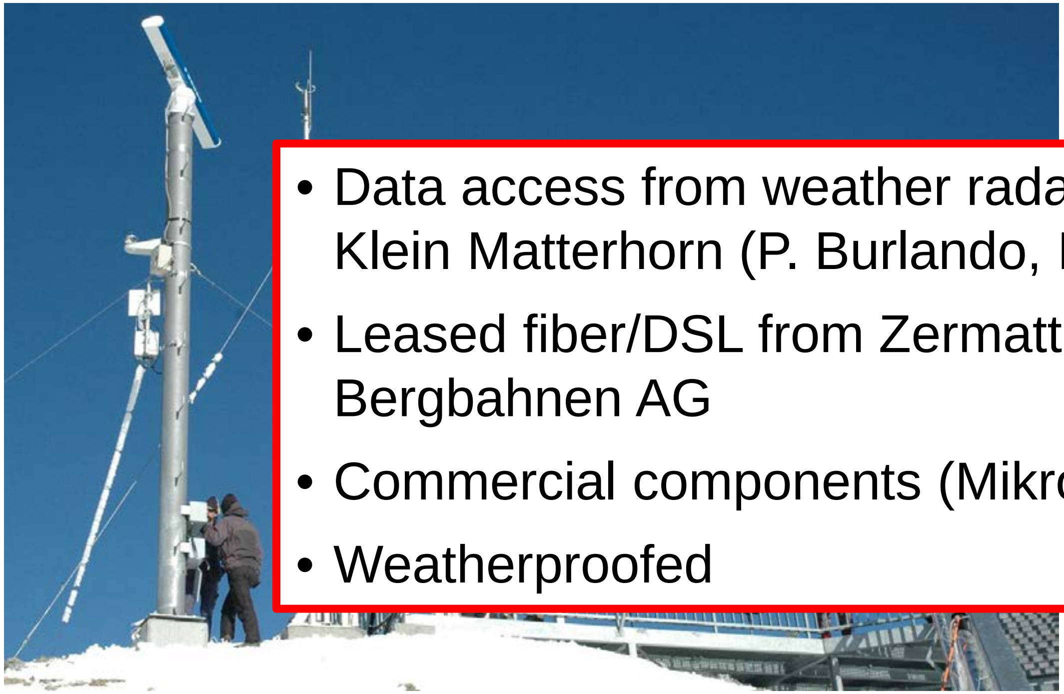
- BackLog = Auxiliary data aggregation layer at device level
  - Remote storage and synchronization layer for Linux systems
  - Python based, designed for PermaSense CoreStation
  - Plugin architecture for extension to custom data sources
  - Data multiplex from plugin to GSN wrapper over one socket

- Reliable (flow controlled) synchronization
- Schedulable plugin/script execution, remote controlled

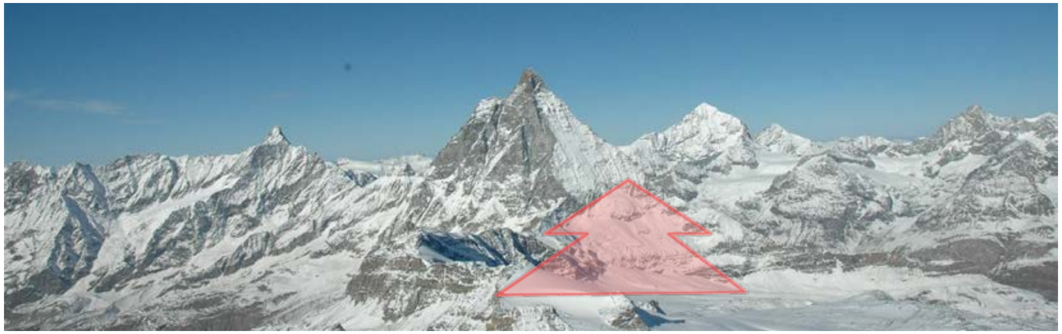




WLAN Long-haul Communication

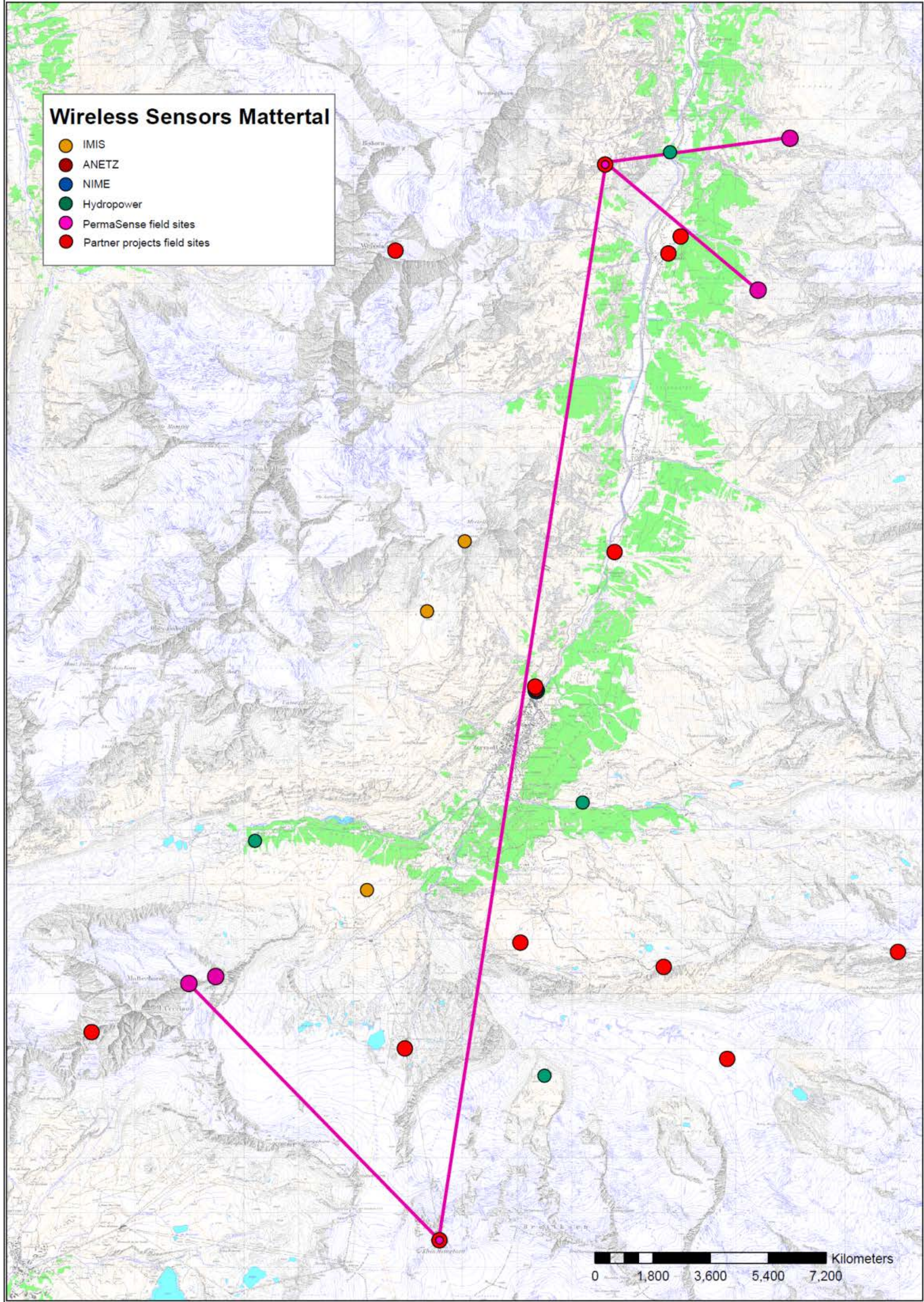


- Data access from weather radar on Klein Matterhorn (P. Burlando, ETHZ)
- Leased fiber/DSL from Zermatt Bergbahnen AG
- Commercial components (Mikrotik)
- Weatherproofed



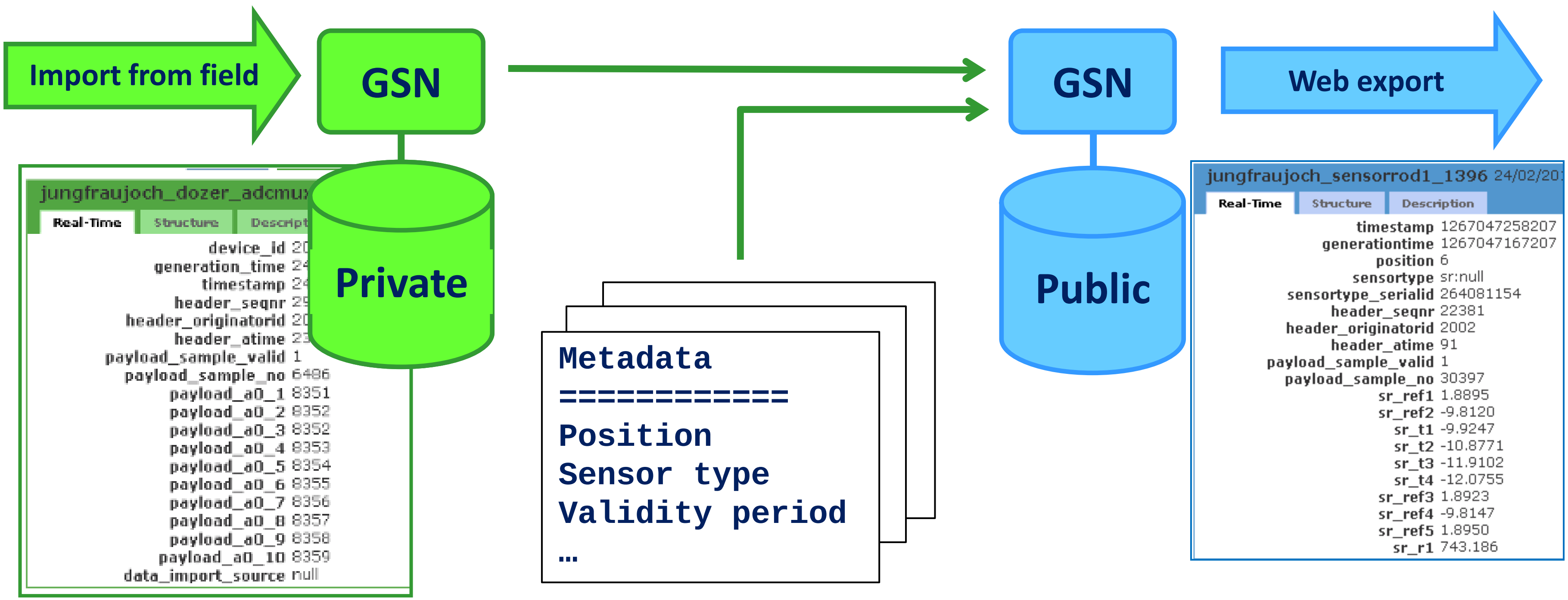
Redundant Access & Monitoring

- Dual WLAN & 3G access network
- Redundant base stations (DH/GG/RD)
- Distributed monitoring infrastructure

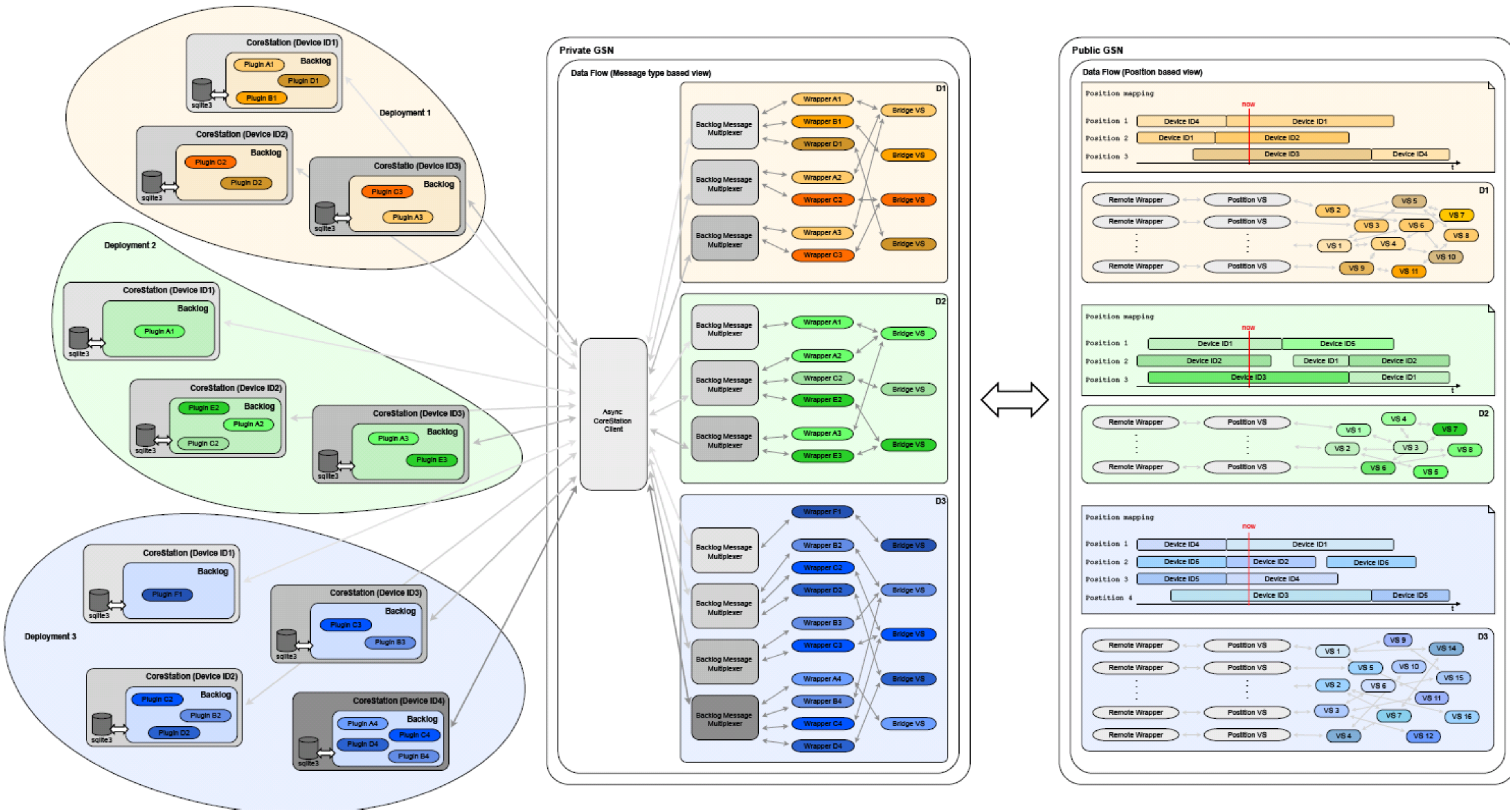


Hierarchical Online Data Processing

- Global Sensor Network (GSN)
  - Data streaming framework from EPFL (K. Aberer)
  - Organized in “virtual sensors”, i.e. data types/semantics
  - Hierarchies and concatenation of virtual sensors enable on-line processing
  - Dual architecture translates data from machine representation to SI values, adds metadata



Multi-Site, Multi-Station, Multi-Revision Data…

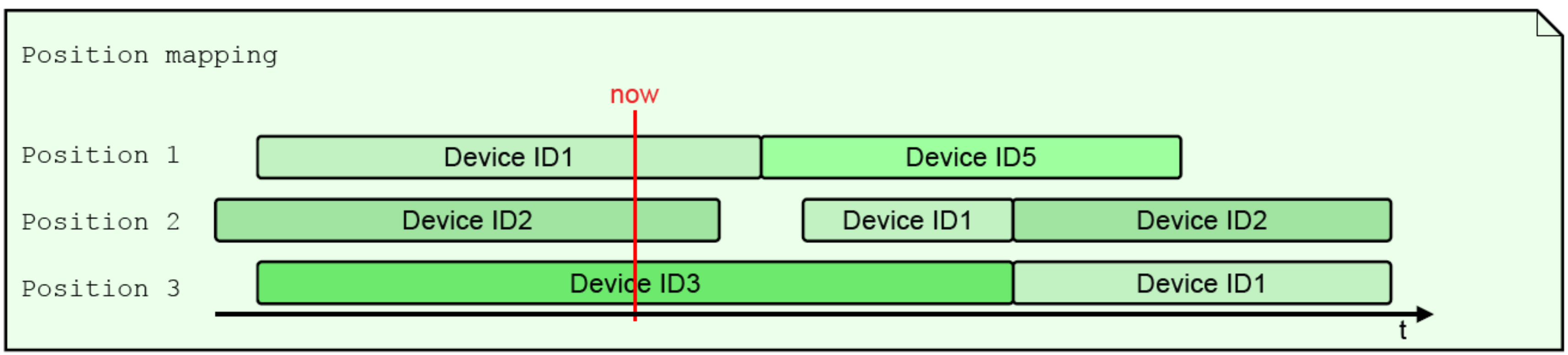


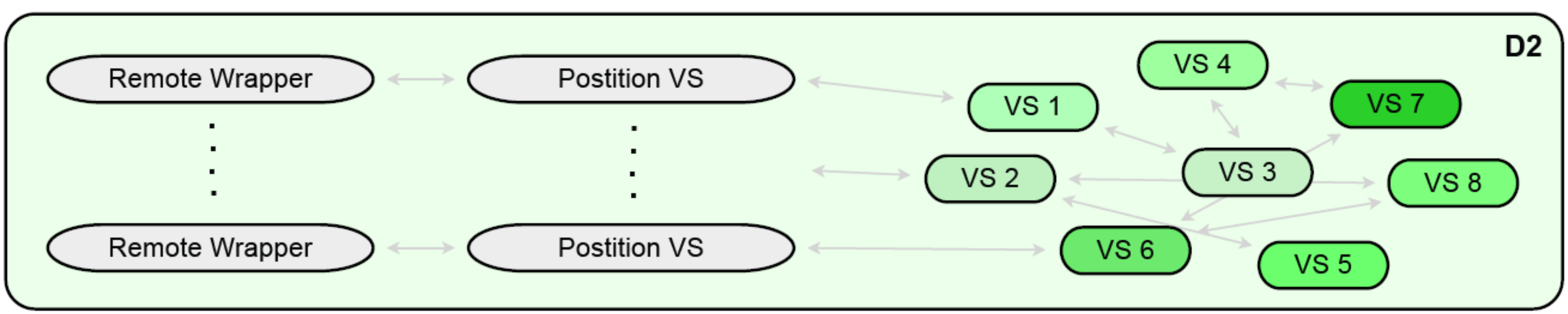
Metadata Mapping Architecture

- Based on 2 GSN instances
  - Separation of load/concern across two machines
  - “Private” GSN instance, raw data, protected, high availability
  - “Public” GSN instance, mapped and converted data, open, non-critical
- Metadata stored in version control system (CSV, SVN)
- Mapping of
  - Positions, coordinates, sensor types, conversion functions, sensor calibration…
- Conversion of
  - Time formats, raw to SI values…
- Replay of metadata/mapping possible, e.g. on errors
- Change management

Metadata Change Management

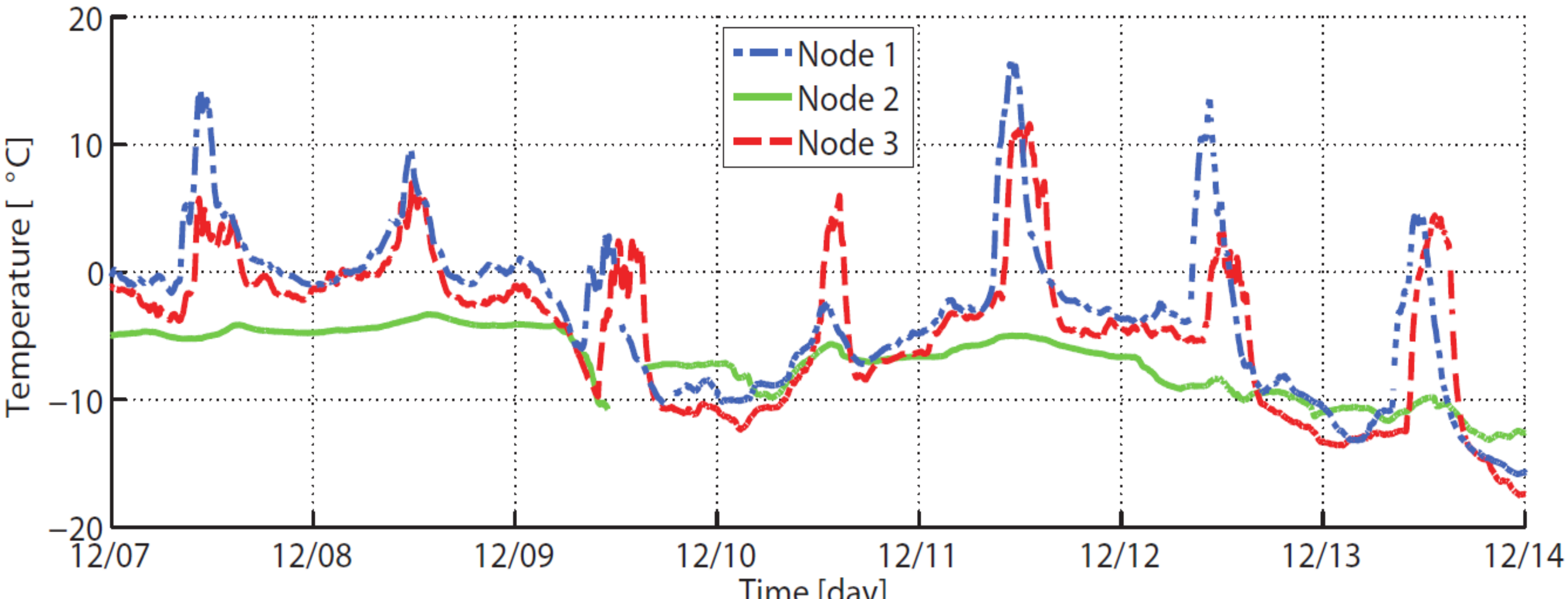
- Allows simple exchange of sensor hard-/software at runtime
- Post-deployment annotation
  - Stop GSN– deployment change – annotate metadata – restart GSN
- Automatic synchronization with 1 day change boundaries





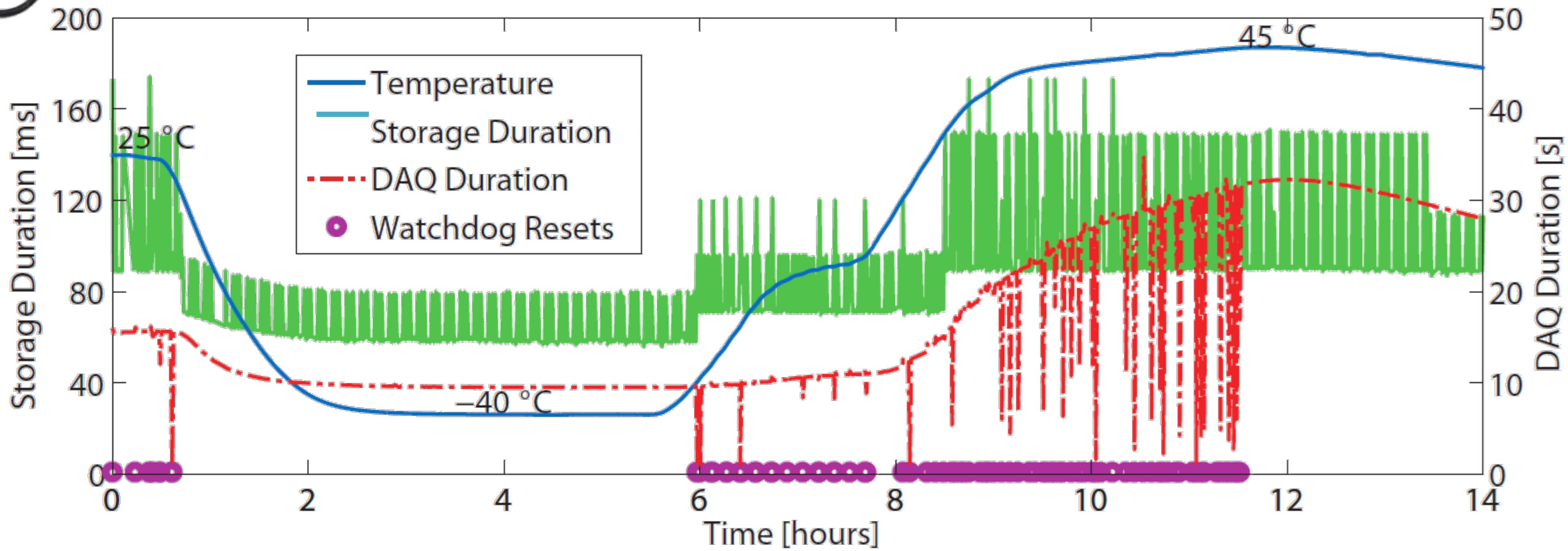
Challenge: The Physical Environment

- Lightning, avalanches, rime, prolonged snow/ice cover, rockfall
- Strong daily variation of temperature
  - −30 to +40°C
  - ΔT ≦ 20°C/hour



Impact of Environmental Extremes

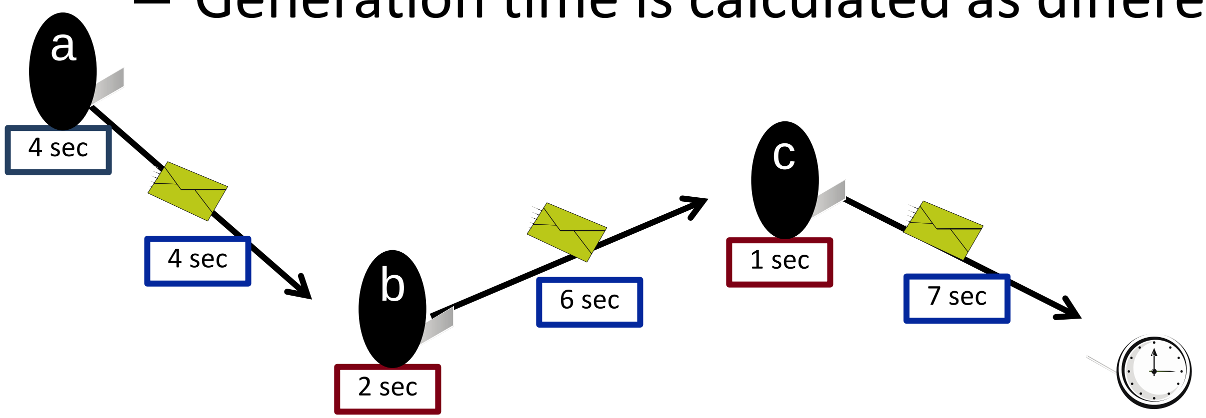
- Tighter guard times increase energy efficiency
- Software testing in a climate chamber
  - Clock drift compensation yields ± 5ppm
- Validation of correct function



[Beutel, DATE 2011]

Reconstructing of Global Time Stamps

- WSN do not have network-wide time synchronization
  - Implications on data usage
- Elapsed time on arrival
  - Sensor nodes measure/accumulate packet sojourn time
  - Base station annotates packets with UTC timestamps
  - Generation time is calculated as difference $\tilde{t}_g = t_b - \tilde{t}_s$

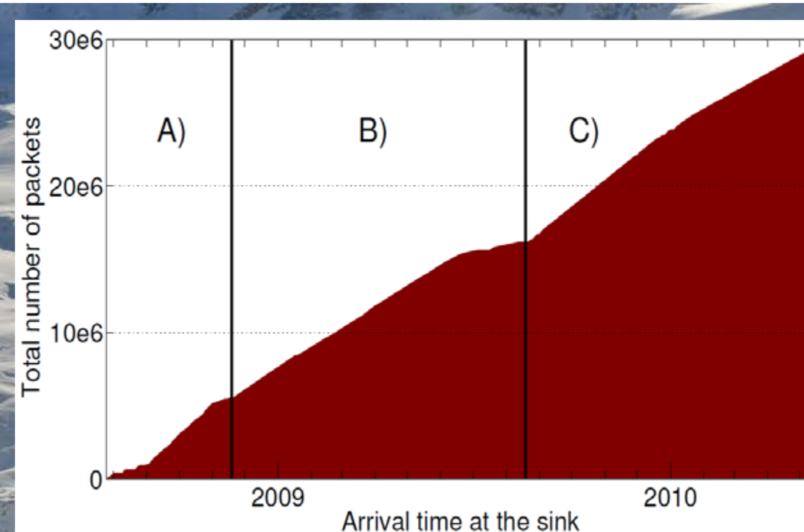


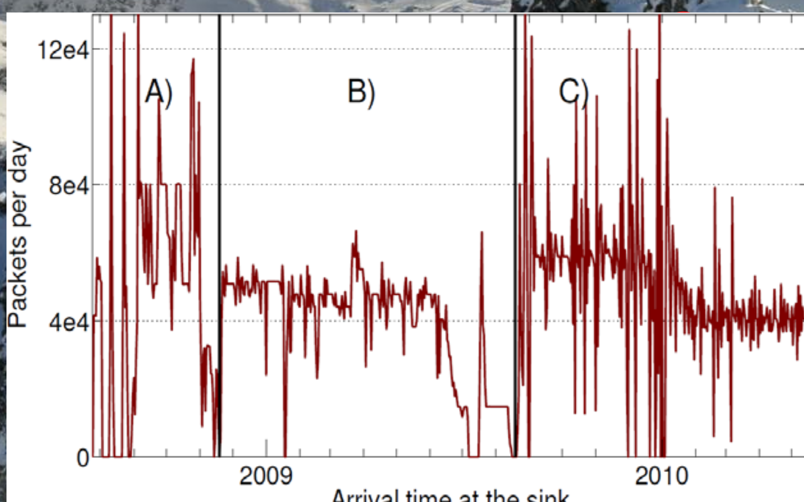
2011/04/14 10:03:31 – 7 sec
= 2011/04/14 10:03:24

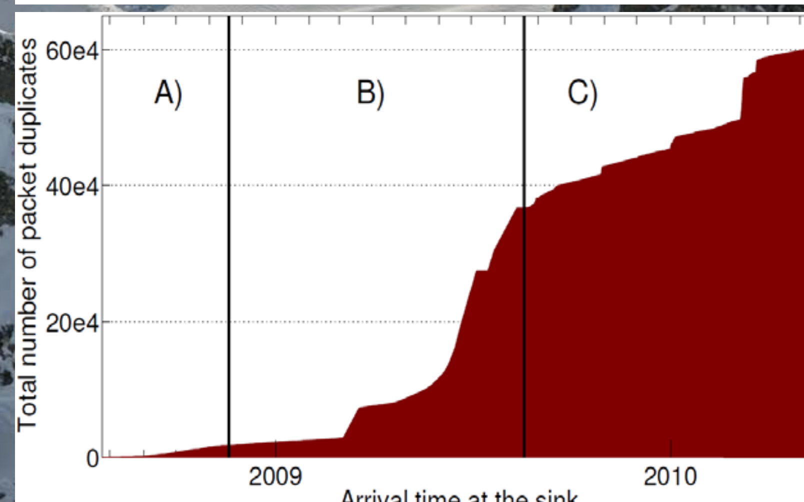
[Keller, IPSN 2011]

Resulting Challenge: Data Integrity

- Long term deployment
- Up to 19 sensor nodes
- TinyOS/Dozer [Burri, IPSN2007]
- Constant rate sampling
- < 0.1 MByte/node/day

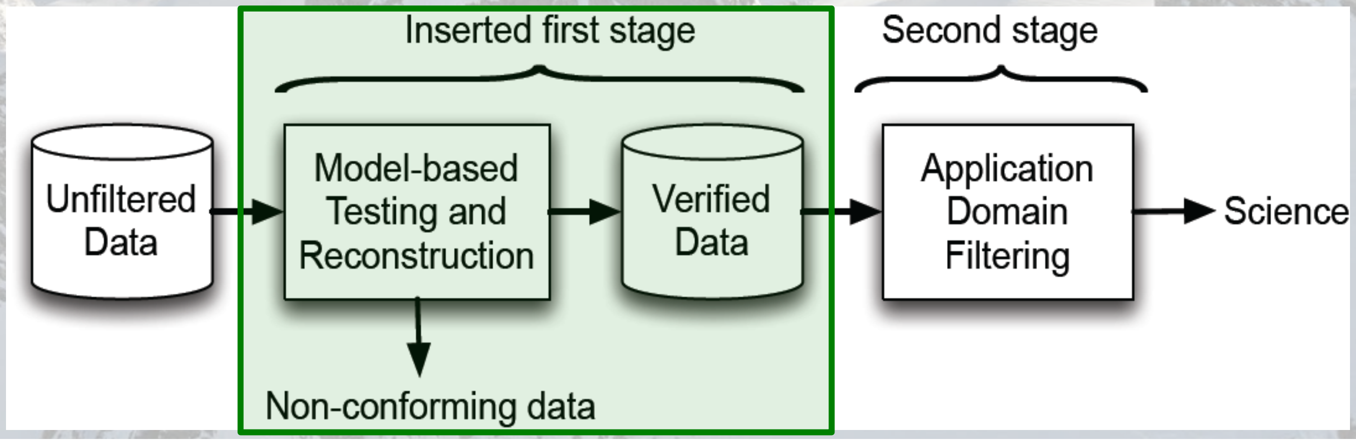






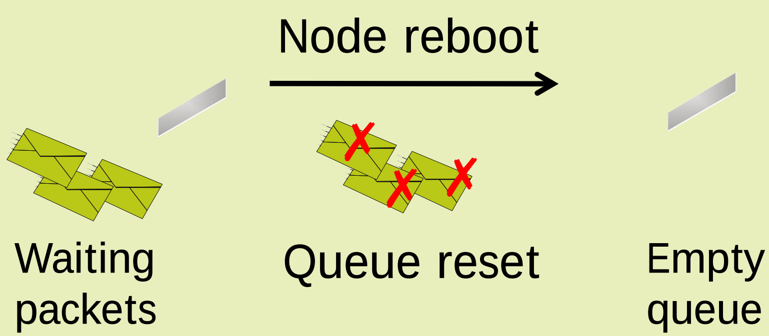
Data is Not Correct-by-Design

- Artifacts observed
  - Packet duplicates
  - Packet loss
  - Wrong ordering
  - Variations in received vs. expected packet rates
- Necessitates further data cleaning/validation



Sources of Errors Included in Model

Data Loss



Clock Drift $\rho \in [-\hat{\rho}; +\hat{\rho}]$

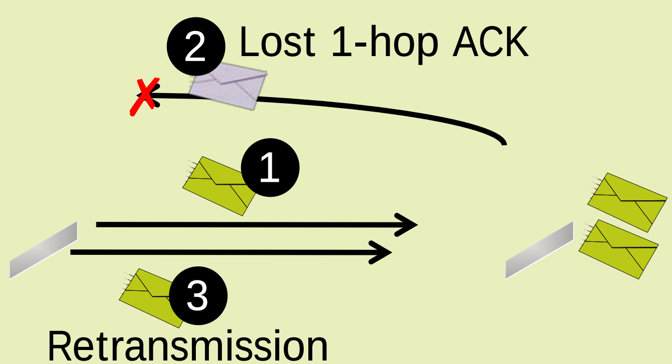
Directly affects measurement of

- Sampling period T
- Contribution to elapsed time t_e

Indirectly leading to inconsistencies

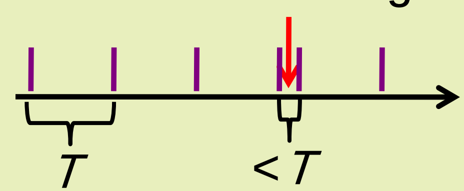
- Time stamp order t_p vs. order of packet generation s

Packet Duplicates



Node Restarts

- Cold restart: Power cycle
- Warm restart: Watchdog reset

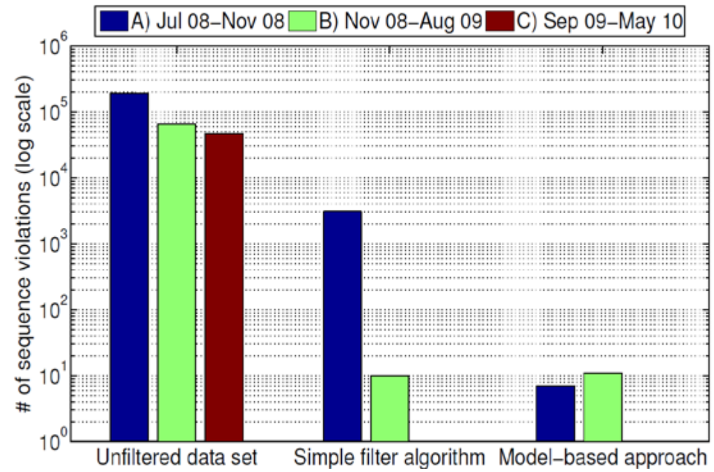


- Shortens packet period
- Resets/rolls over certain counters

Model-based Data Validation Case Studies

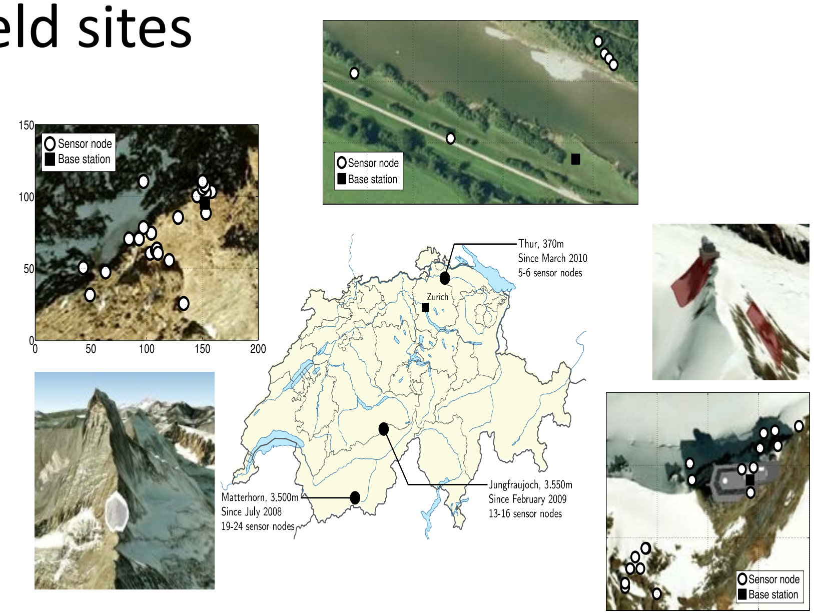
- Validation of correct system function

Counter	A) Jul 08-Nov 08	B) Nov 08-Aug 09	C) Sep 09-May 10
Accepted packets	632,058 (59.4%)	2,110,855 (96.8%)	2,579,444 (95.4%)
Discarded packets	432,826 (40.6%)	69,829 (3.2%)	124,554 (4.6%)
Packet duplicates	4,020 (0.4%)	69,422 (3.2%)	44,601 (1.7%)
$t_s(i) > t_s^{max}$	0 (0.0%)	0 (0.0%)	0 (0.0%)
Failed epoch assignment	235,927 (22.2%)	277 (0.0%)	2,466 (0.1%)
Invalid interval $t_g^{u,l}(i)$	192,879 (18.1%)	130 (0.0%)	77,487 (2.9%)
Total packets	1,064,884 (100.0%)	2,180,684 (100.0%)	2,703,998 (100%)



- Long-term comparison of three field sites

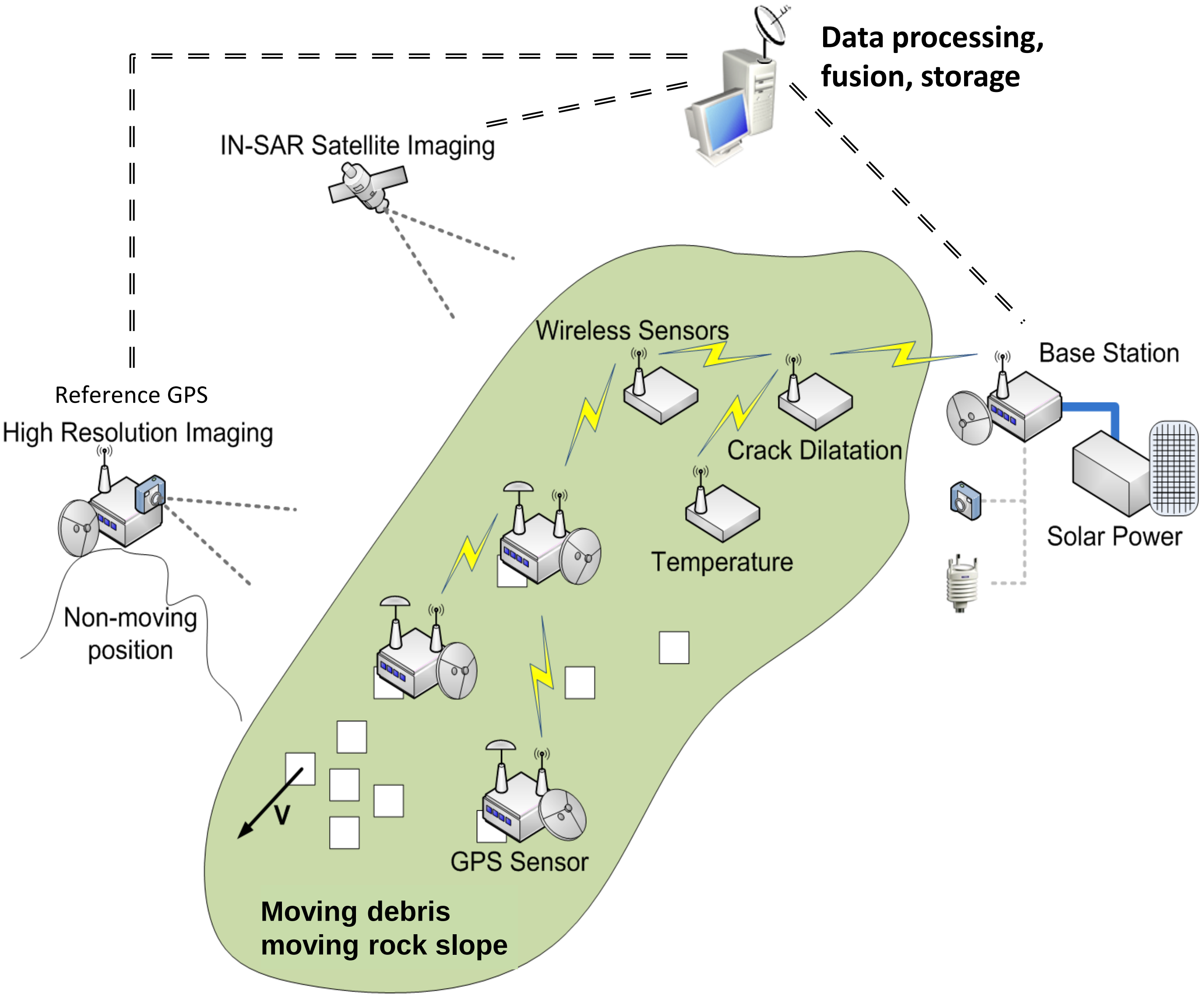
Deployment	Matterhorn	Jungfraujoch	Thur
Packet counters, mean per node			
Unique	1,117,659	1,117,338	915,903
Missing	667	5,368	29
Duplicates	59,333	85,901	11,293
Data yield			
∅	99.94%	99.53%	100.00%
min	99.88%	96.28%	99.99%
max	100.00%	100.00%	100.00%
Radio duty cycle			
max	0.26%	1.31%	0.12%
Median	0.14%	0.34%	0.07%



[Keller, SenseApp 2011, IPSN 2011]

An Example of Fusing Sensor Data

Example: The X-Sense Platform

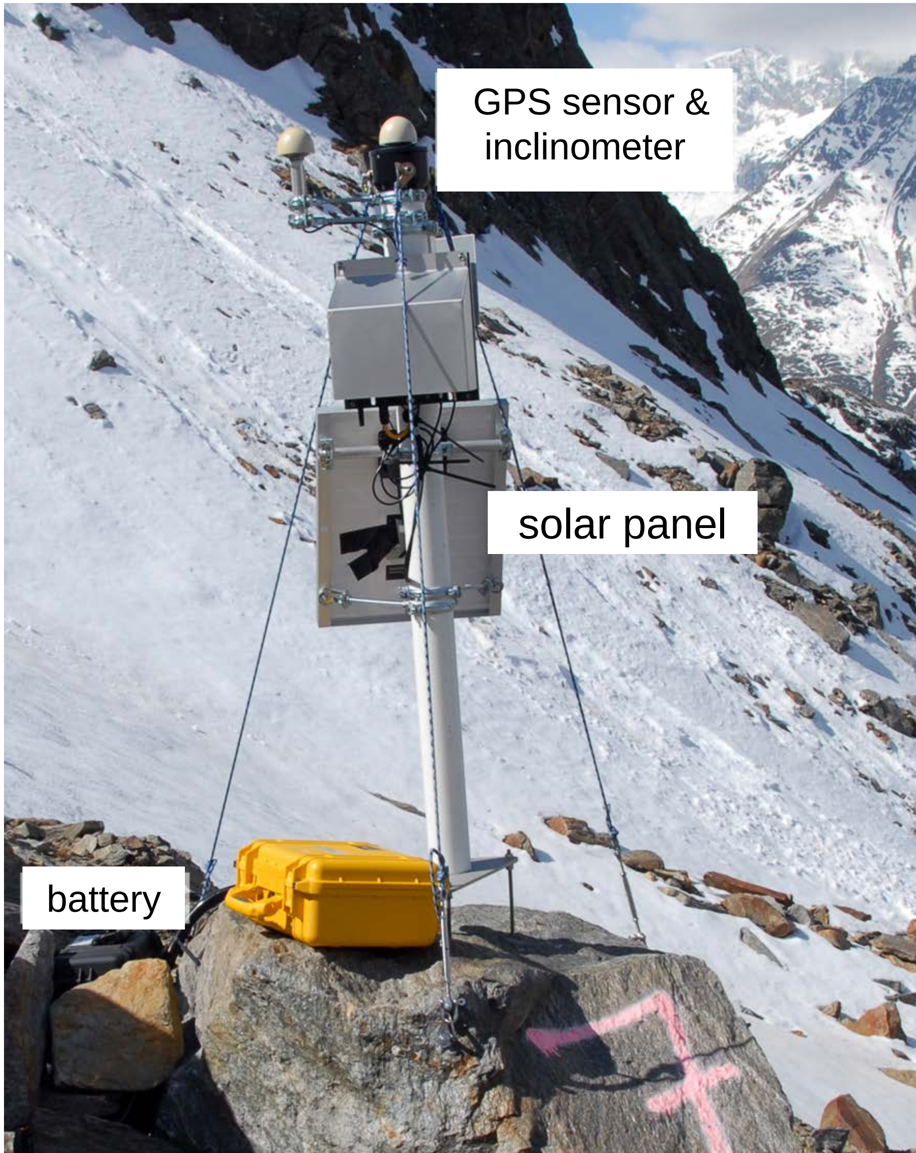


GPS Measurement Devices

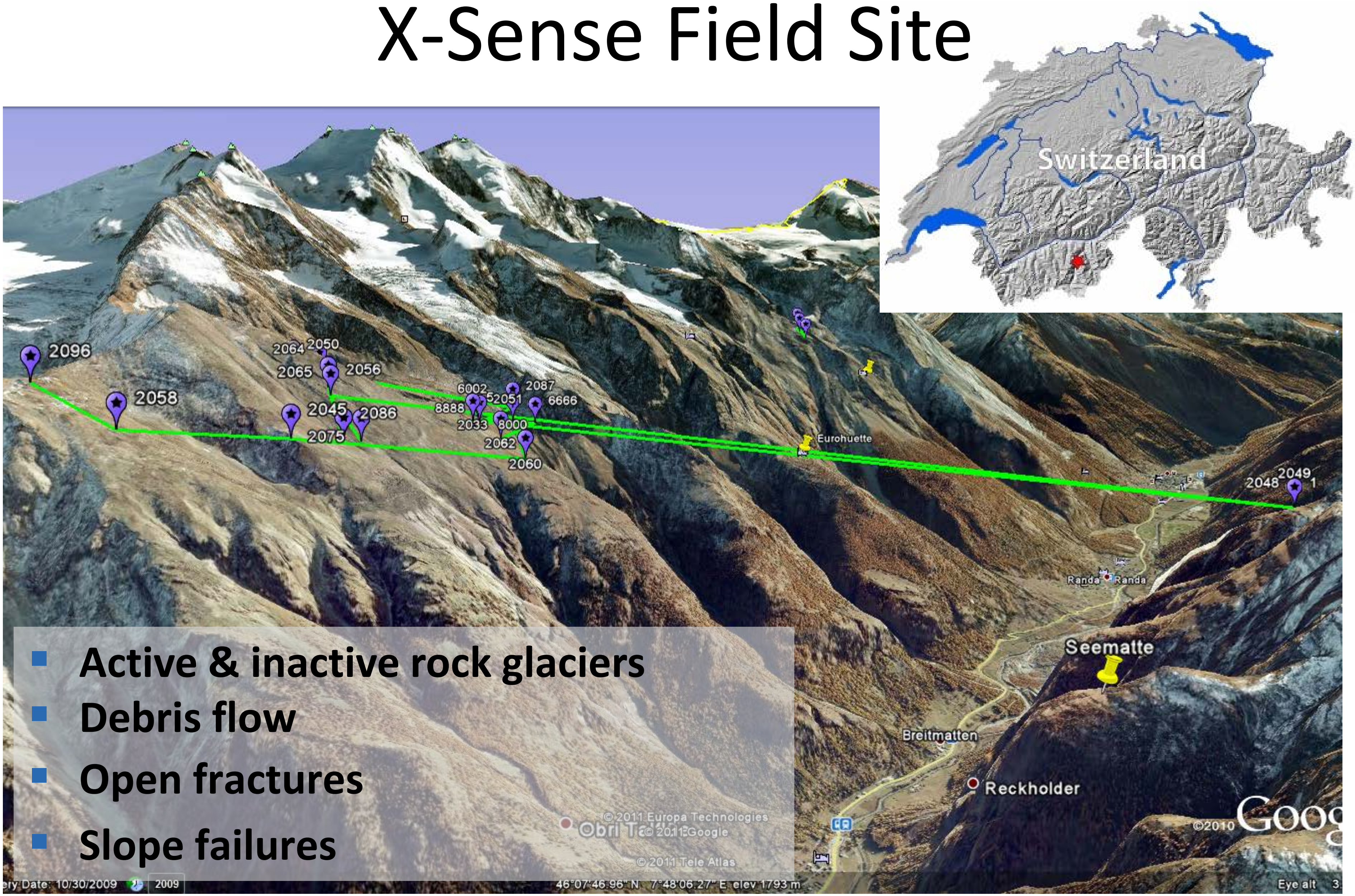
Low-cost L1 GPS Devices

- Dual strategy: Logging units & wireless sensors
- High temporal resolution
- Accurate displacement-rate of a boulder (mm-cm accuracy for daily position)

[Wirz, WLF 2011, Buchli SGM 2011]



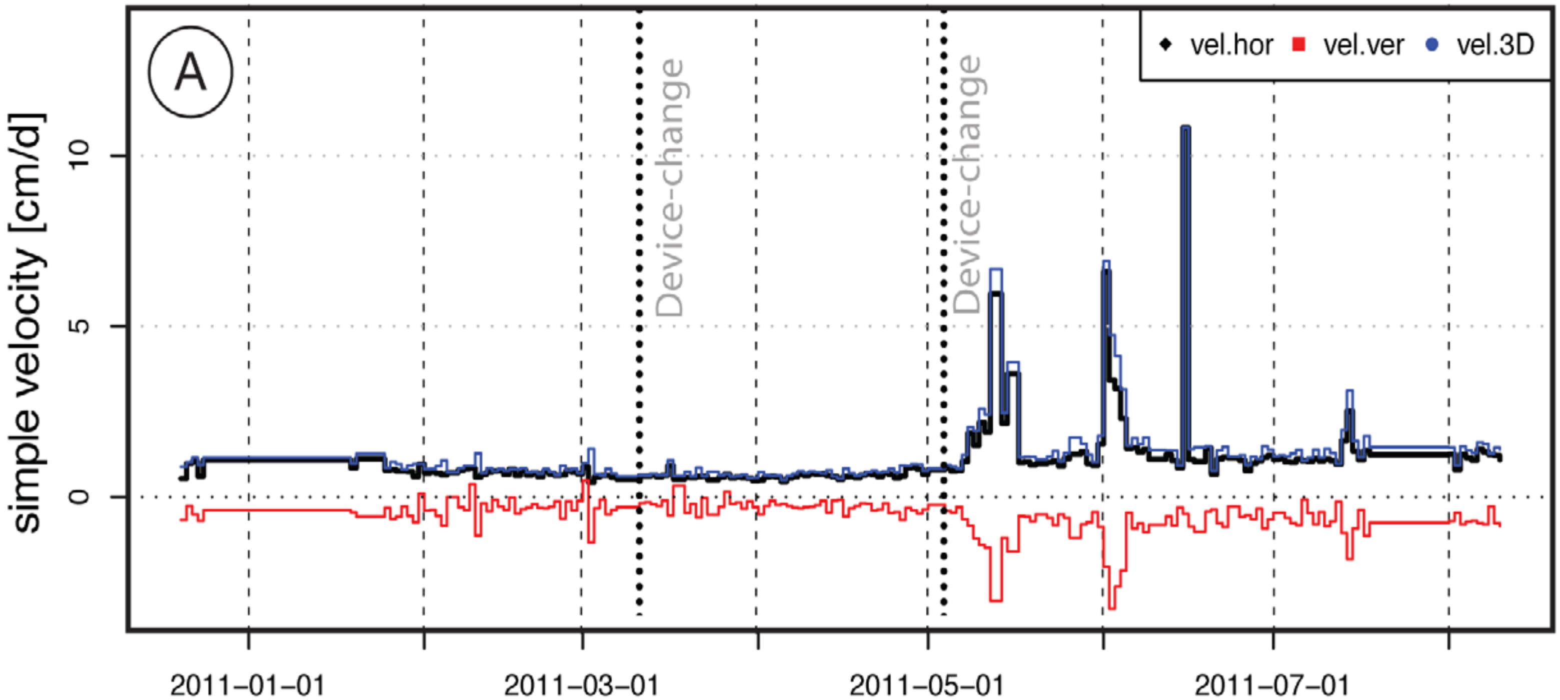
X-Sense Field Site



- Active & inactive rock glaciers
- Debris flow
- Open fractures
- Slope failures

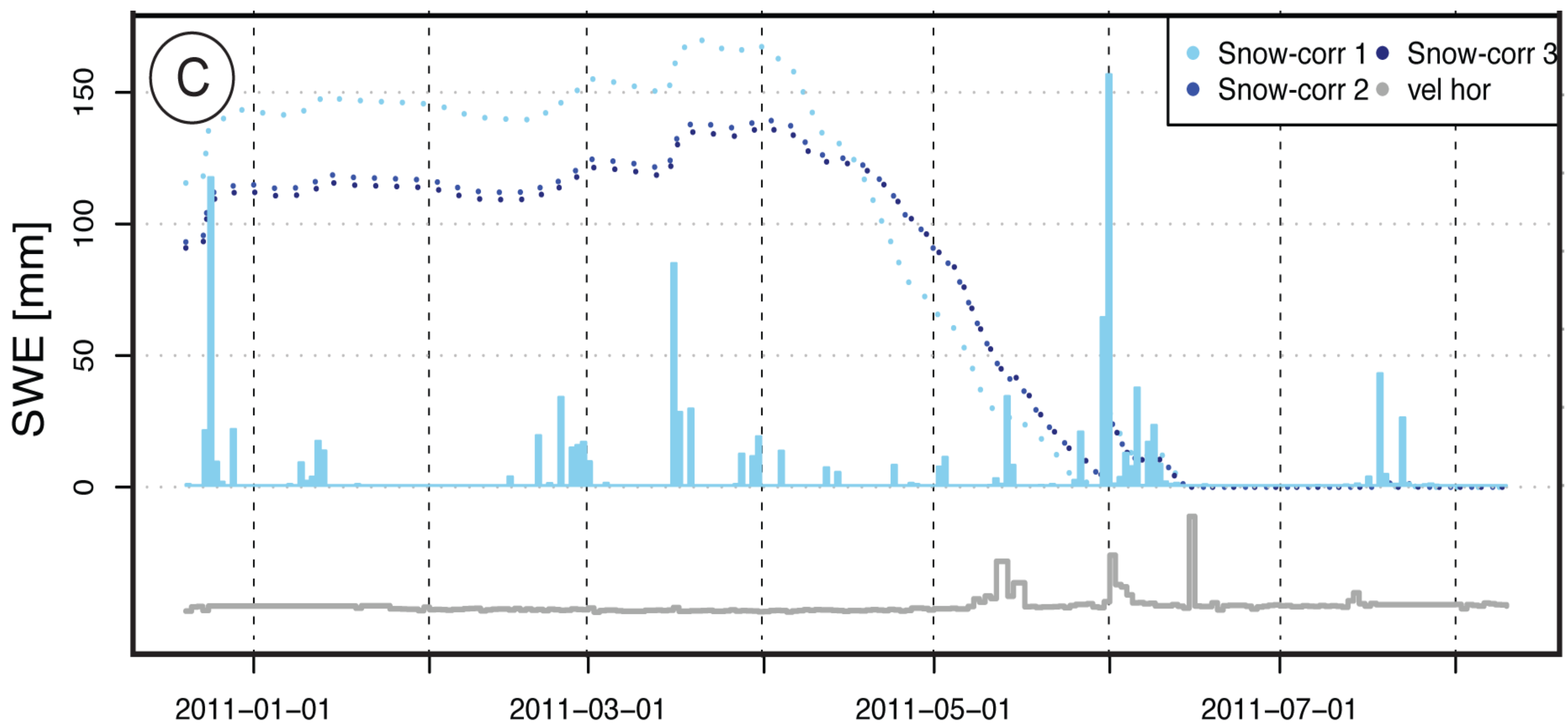
GPS Data Analysis

- Post-processing of GPS time series
  - Correction to coordinates at ground level
  - Derivation of differing measures of velocity

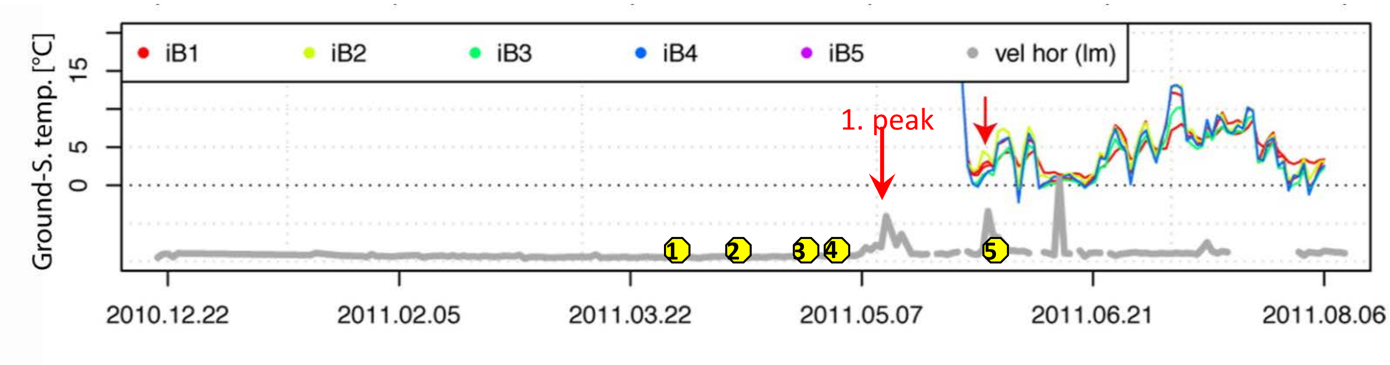


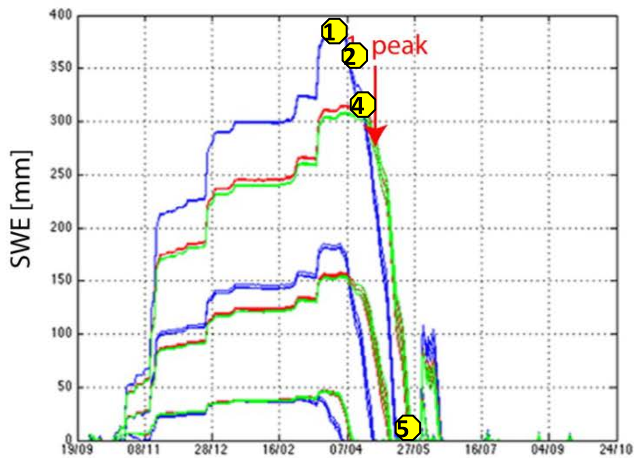
GPS Data and Simulation Combined

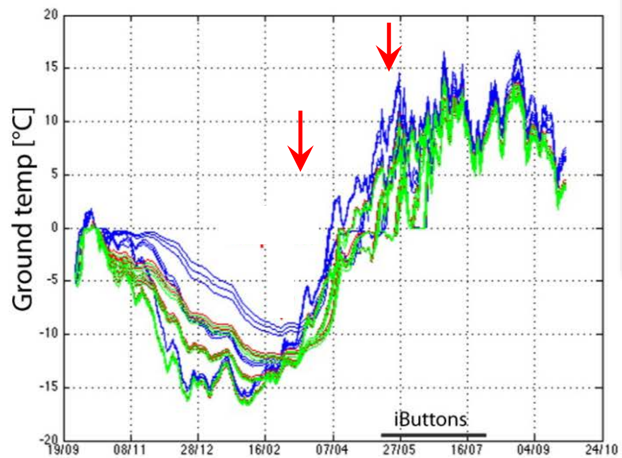
- Comparison with environmental data
  - First peak during snow melt, second during heavy precipitation
  - Third peak has no apparent correlate



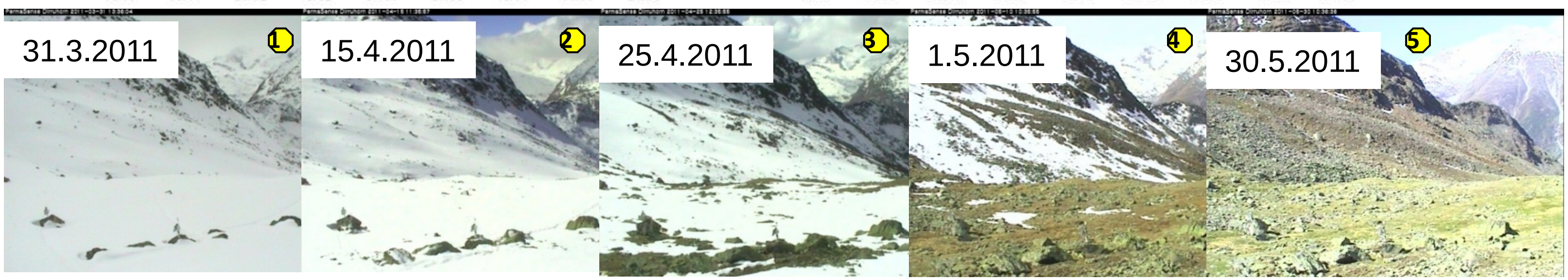
Data Fusion and Interpretation







V. Wirz, P. Limpach, J. Beutel, B. Buchli and S. Gruber: Temporal characteristics of different cryosphere-related slope movements in high mountains. *Proc. 2nd World Landslide Forum*, Springer, Berlin, October, 2011.



- ETH Zurich
  - Computer Engineering and Networks Lab
  - Geodesy and Geodynamics Lab
- University of Zurich
  - Department of Geography
- EPFL
  - Distributed Information Systems Laboratory
- University of Basel
  - Department Computer Science

nano-tera.ch

NCCR MICS
National Competence Center in Research
Mobile Information and Communication Systems

cces

FNS NF
Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

Bundesamt für Umwelt BAFU
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

HILTI Art of Technology evoNET Internet & Computer Solutions ARC'TERYX

Interested in more?

http://www.permasense.ch