Advanced PV performance analysis on modules and power plants using cloud-based processing

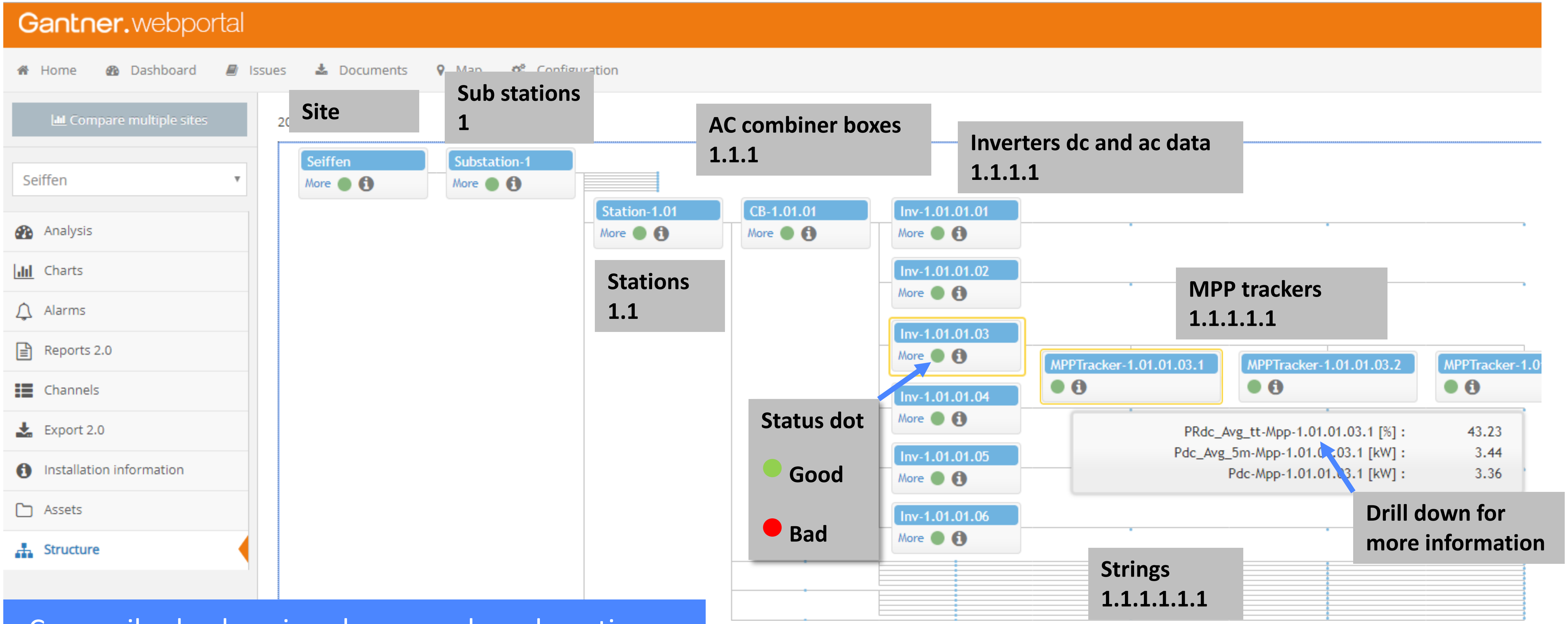
Juergen Sutterlueti[1] and Steve Ransome[2]

[1]Gantner Instruments GmbH, Austria

[2]Steve Ransome Consulting Ltd, U.K.

Do we already utilize available PV power plant data sufficiently?

Checking performance at 7 different levels on a power plant



Can easily check various losses such as downtime, mismatch, shading, tracking per component …

Monitoring a large array – looking at different weather type days

Variable Day	Clear Sky Day	Dull Day

Some effects differ by weather type e.g.

1) clipping only at high irradiance
2) overheating/mis tracking only at high temperatures/irradiance
3) shading at clear sky/low solar elevation
4) …

Compare individual P_{AC} with mean of >150 Inverters (usually look for faulty or worst ones for more analysis)

Gantner instruments

Gantner.webportal

Count	Mean	Std. deviation (k=2)	Rel. deviation [%] min/max	OK range (from mean) [%] [abs]	Median
156/156	6.98	0.35355	-20.53 / 5.31	90.0000	7.08

Component	Pac [kW]	Pac Difference [kW]	PacRel. deviation [%]	Range [%]	PacStatus []
Inv-1.02.19.05	5.549	-1.433	-20.53		low
Inv-1.01.08.04	5.603	-1.379	-19.75		low
Inv-1.02.19.02	5.608	-1.374	-19.68		low
Inv-1.02.19.04	5.673	-1.309	-18.75		low
Inv-1.02.19.03	5.764	-1.218	-17.45		ok
Inv-1.02.19.06	6.031	-0.951	-13.63		ok
Inv-1.02.19.01	6.233	-0.749	-10.73		ok

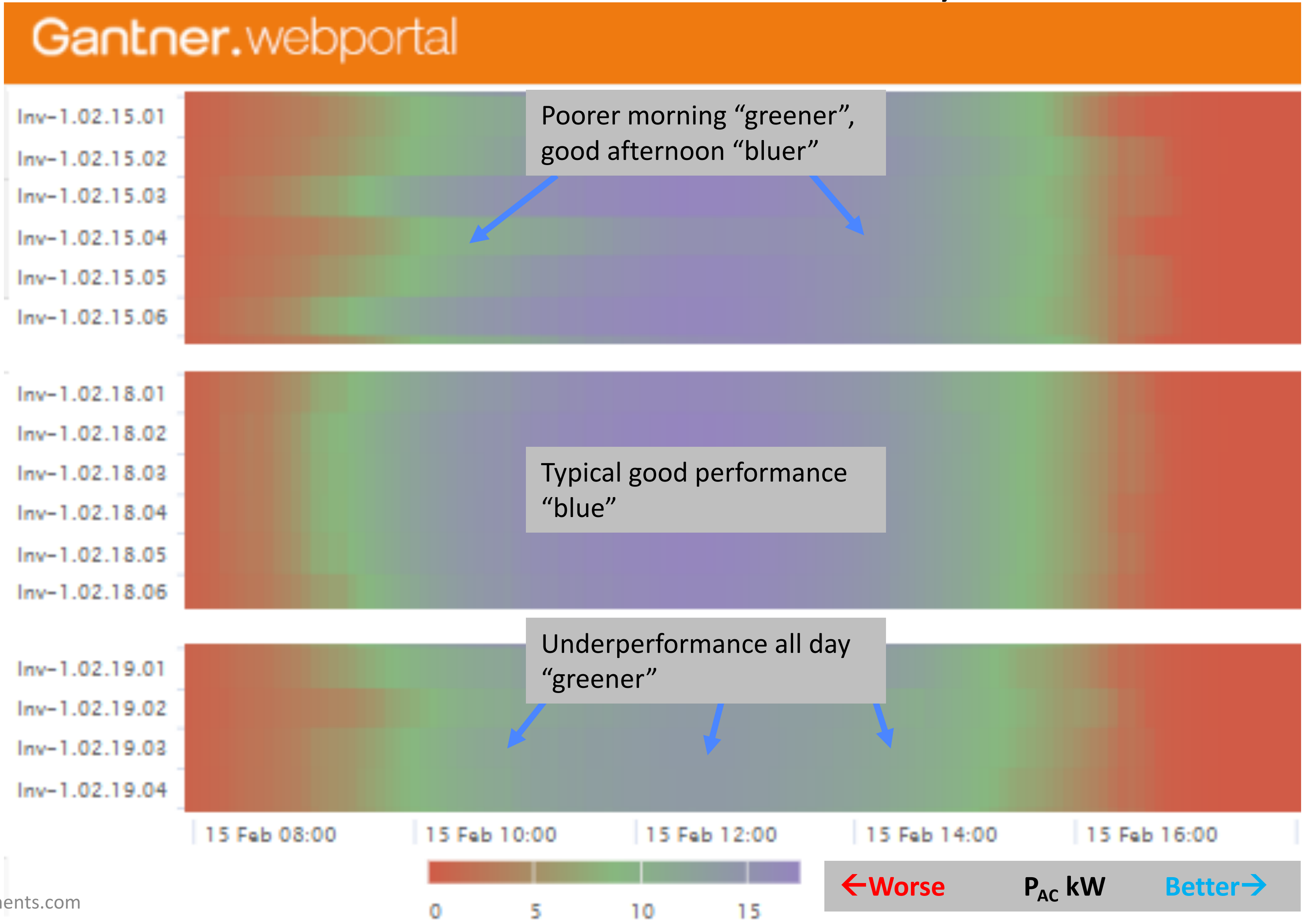
P_{AC} Inverter kW Heat Maps for a clear day (Good performance should be similar for all inverters)

Level 4: Inverter 1.02.15 .. 19

#15 Variable

#18 All good

#19 Poor all day



Poorer morning "greener", good afternoon "bluer"

Typical good performance "blue"

Underperformance all day "greener"

Time →

←Worse P_{AC} kW Better→

P_{AC} Inverters 1.01.12#1..6

Level 4: Inverter 1.1.12.xx

#6 is worse in the morning
than #1..5,
better afternoon

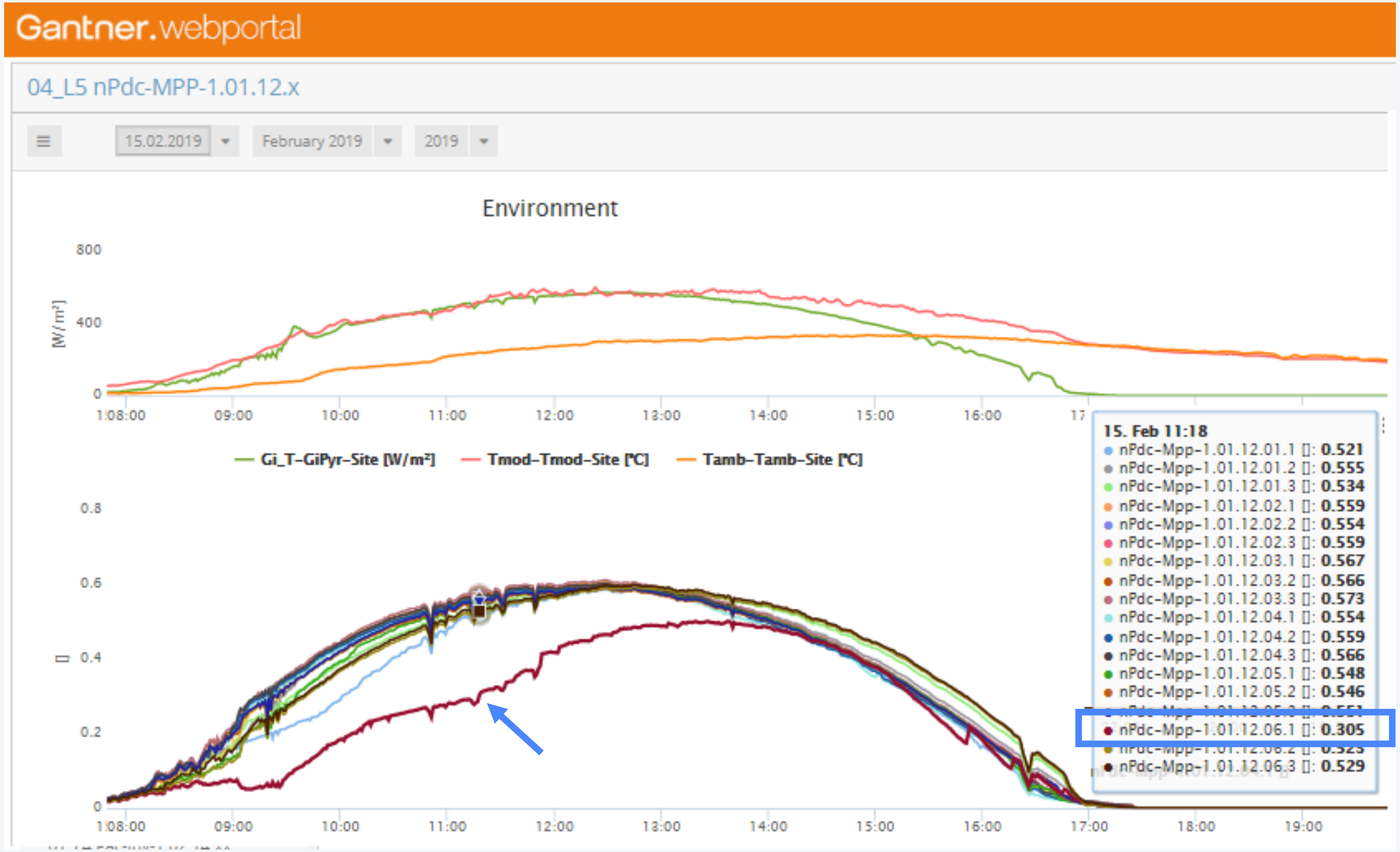
Why is inverter #6 different?

nP_{DC} MPP Trackers 1.01.12 #01.1-06.3

Gantner instruments

Level 5: MPP Tracker
1.1.12.6.xx

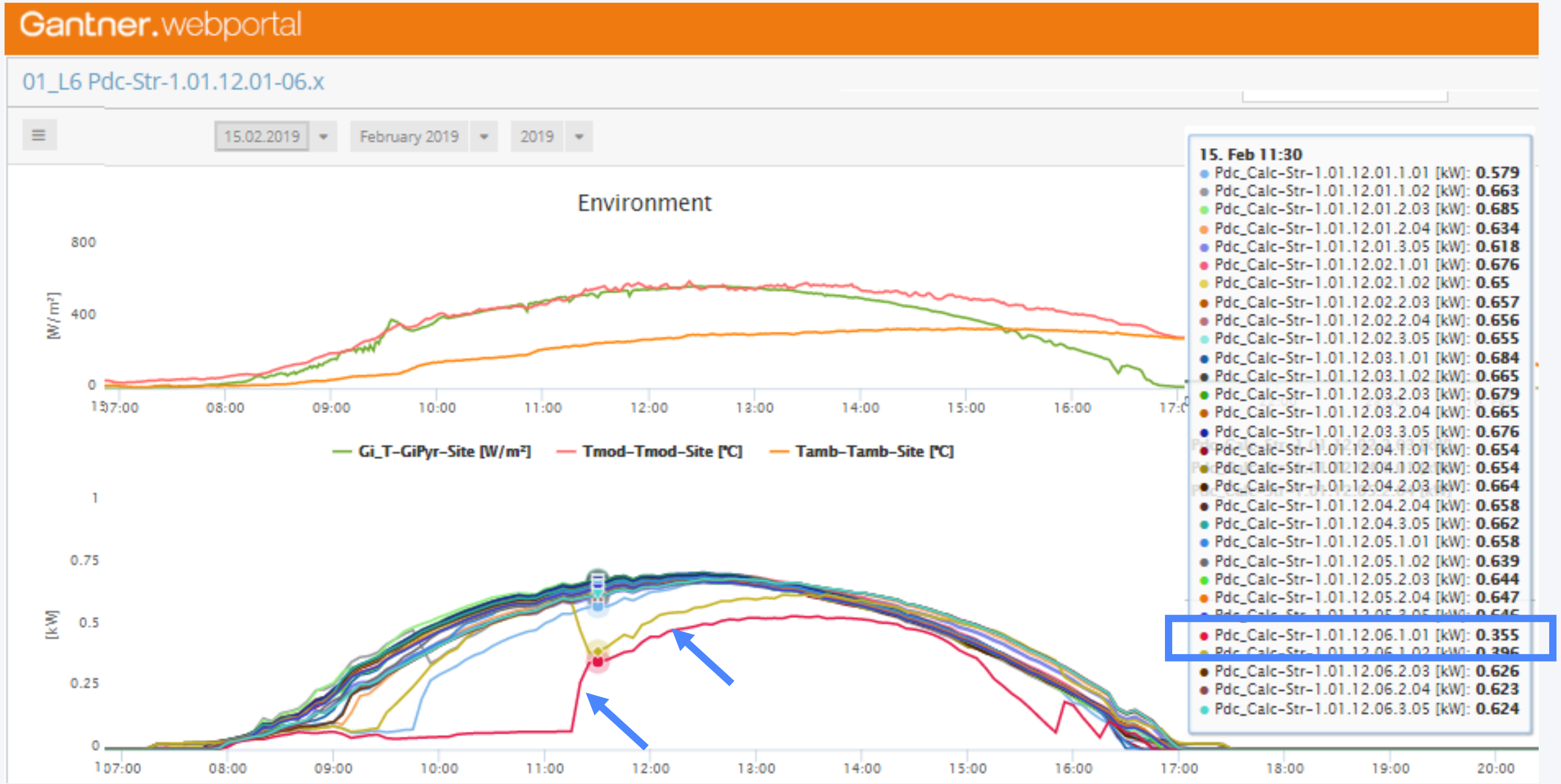
MPP #06.1 poor in the morning



P_{DC} Strings 1.01.12 #01.1.01-06.3.05

Gantner instruments

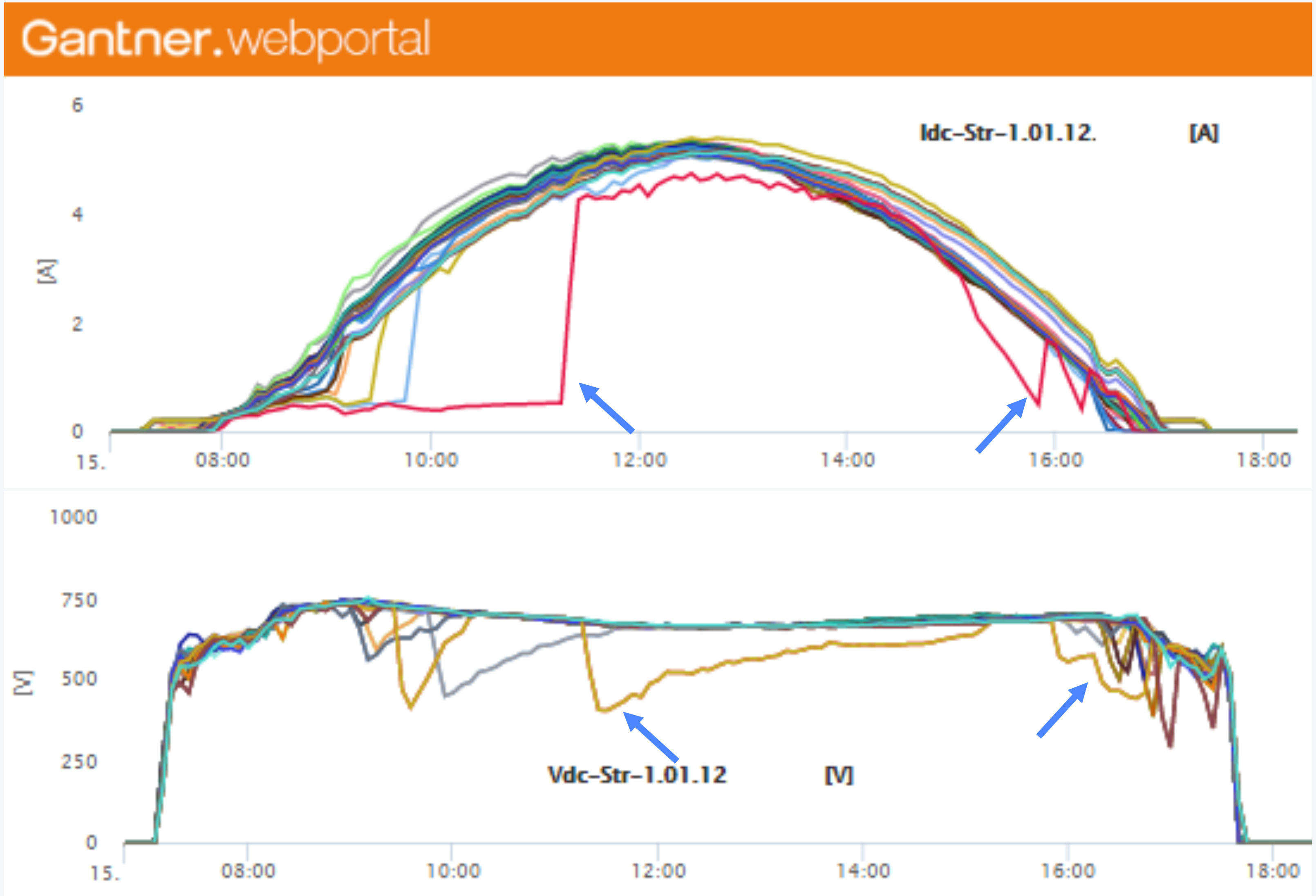
Level 6: String
1.1.12.6.01.xx



Showing two misbehaving

Idc(top) and Vdc(bott) for Strings in Inverter 1.01.12

Level 6: String
1.1.12.6.01.xx



Faults get more apparent as we zoom in closer with fewer modules

Can now investigate at string level reason for unusual behaviour

How to find faults automatically

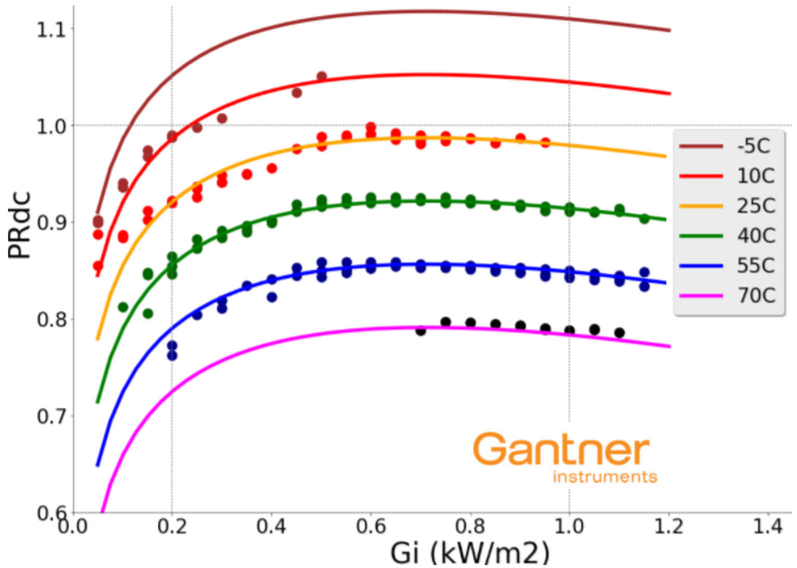
... use the Mechanistic Performance Model (MPM)

- ... to fit measured PR vs. Irradiance and Tmodule
- look for discrepancies or poor fit coefficients

$$\mathbf{PR_{DC} = C_1 + C_2 \times dT_{MOD} + C_3 \times Log_{10}(G_I) + C_4 \times G_I + C_5 \times WS}$$

MPM Coefficients are meaningful, orthogonal, robust and normalised

	Usual approx. Range	Coefficient Dependency	Comment	Unit
C_1	80% to 100%	Performance Tolerance	Actual/Nominal	%
C_2	-0.2% to -0.5%/K	(Tmod-25)C	Temperature Coefficient	%/K
C_3	0 to 20%	$log_{10}(G_I)$	Low light fall (~Voc, Rshunt)	%
C_4	-20% to 0%	G_I	High light fall (~Rseries)	%
C_5	-2% to 2%	Windspeed	Windspeed correction	$\%/(ms^{-1})$



Fitted lines to measured PR_{DC} for c-Si module at GI OTF in AZ with good agreement

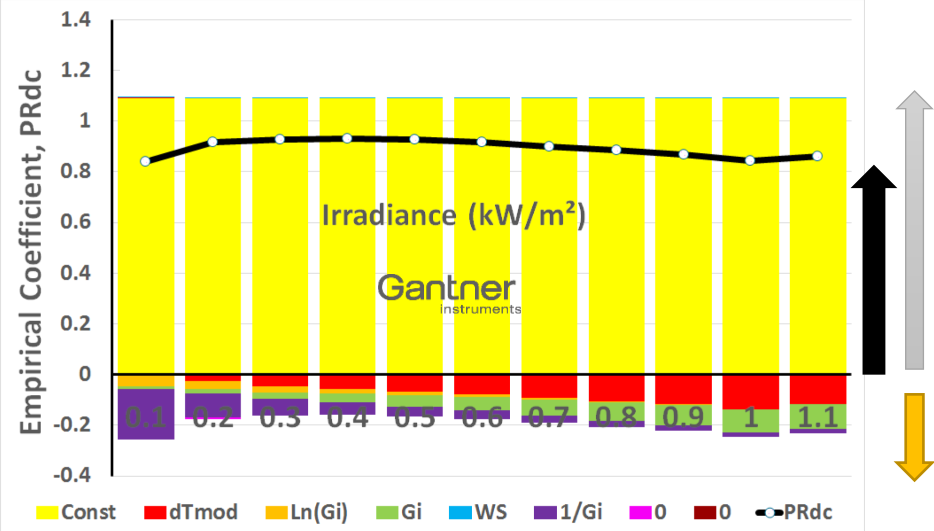
... use the Mechanistic Performance Model (MPM)

- ... to fit measured PR vs. Irradiance and Tmodule
- look for discrepancies or poor fit coefficients

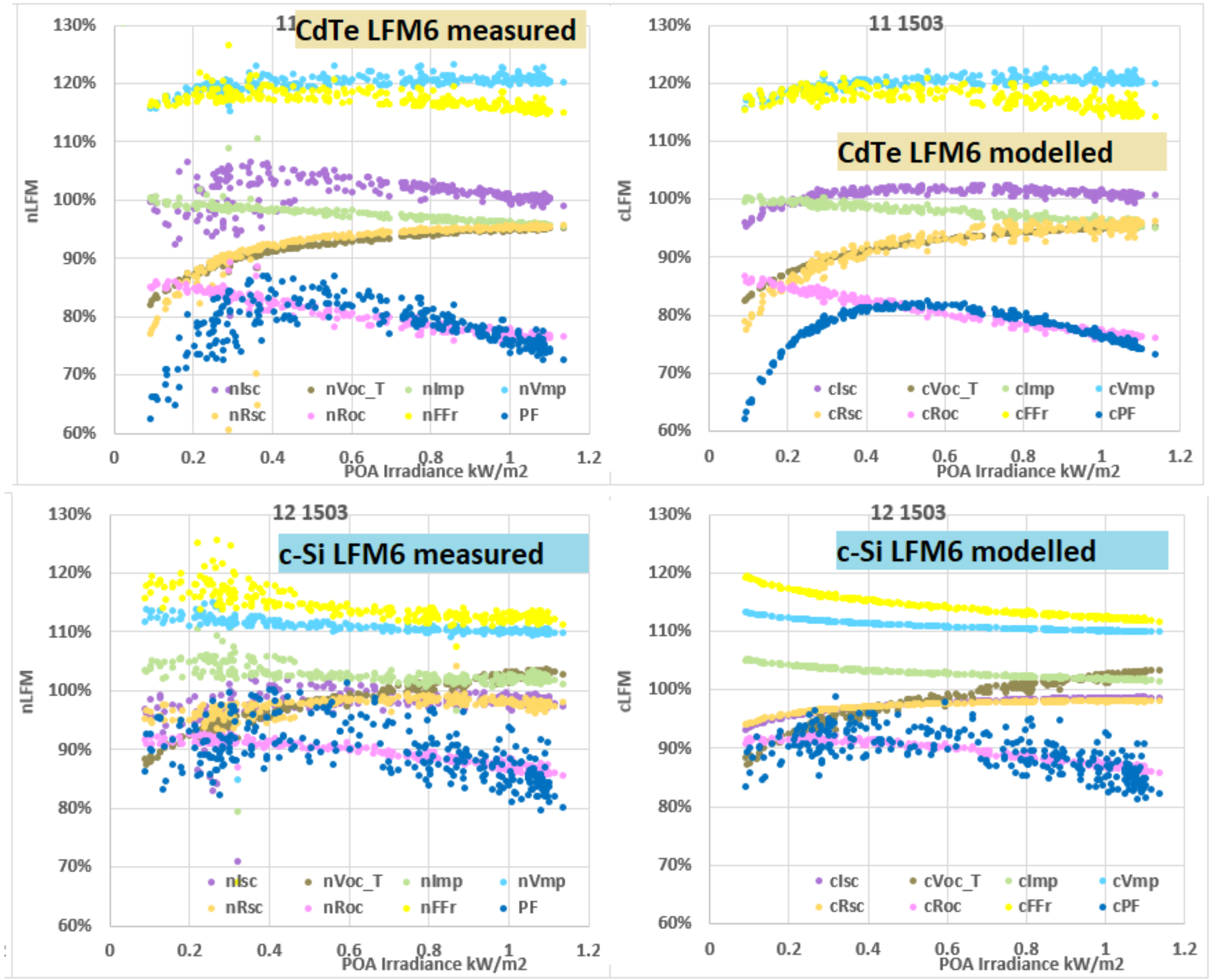
$$\mathbf{PR_{DC} = C_1 + C_2 \times dT_{MOD} + C_3 \times Log_{10}(G_I) + C_4 \times G_I + C_5 \times WS}$$

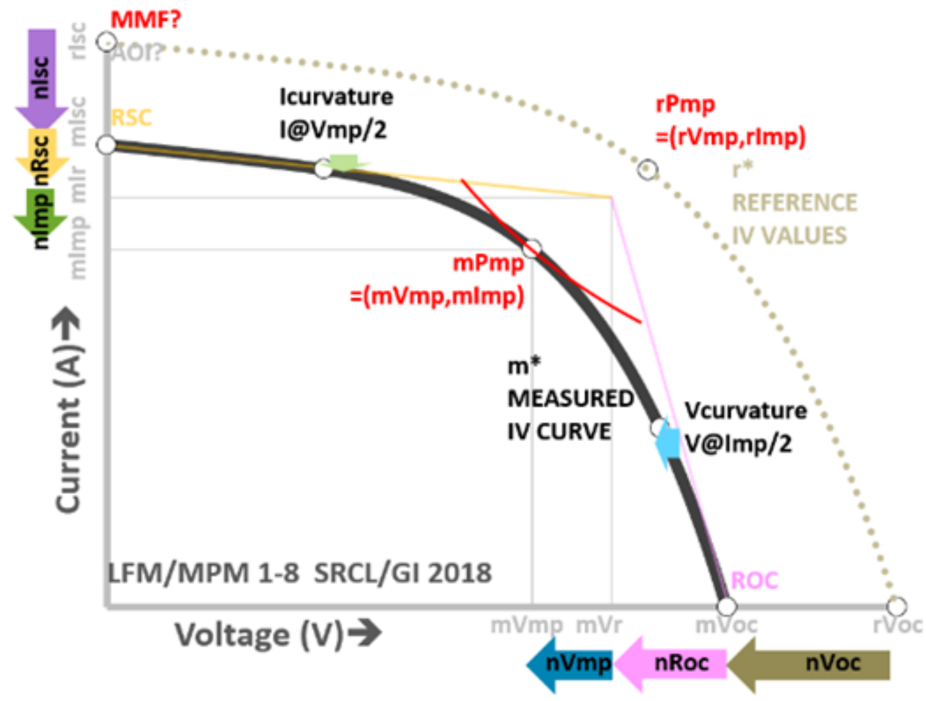
MPM Coefficients are meaningful, orthogonal, robust and normalised

	Usual approx. Range	Coefficient Dependency	Comment	Unit
C_1	80% to 100%	Performance Tolerance	Actual/Nominal	%
C_2	-0.2% to -0.5%/K	(Tmod-25)C	Temperature Coefficient	%/K
C_3	0 to 20%	$\log_{10}(G_I)$	Low light fall (~Voc, Rshunt)	%
C_4	-20% to 0%	G_I	High light fall (~Rseries)	%
C_5	-2% to 2%	Windspeed	Windspeed correction	$\%/(ms^{-1})$



Fitting all 6 LFM parameters (nIsc, nRsc, nImp, nVmp, nRoc & nVoc) with the MPM $nLFM = C_1 + C_2 \times dT_{MOD} + C_3 \times Log_{10}(G_I) + C_4 \times G_I$





MPM fits all LFM coefficients for both Thin Film (top) and c-Si (Bottom) but with values which differentiate their performance at low (Rshunt) and high (Rseries) light levels

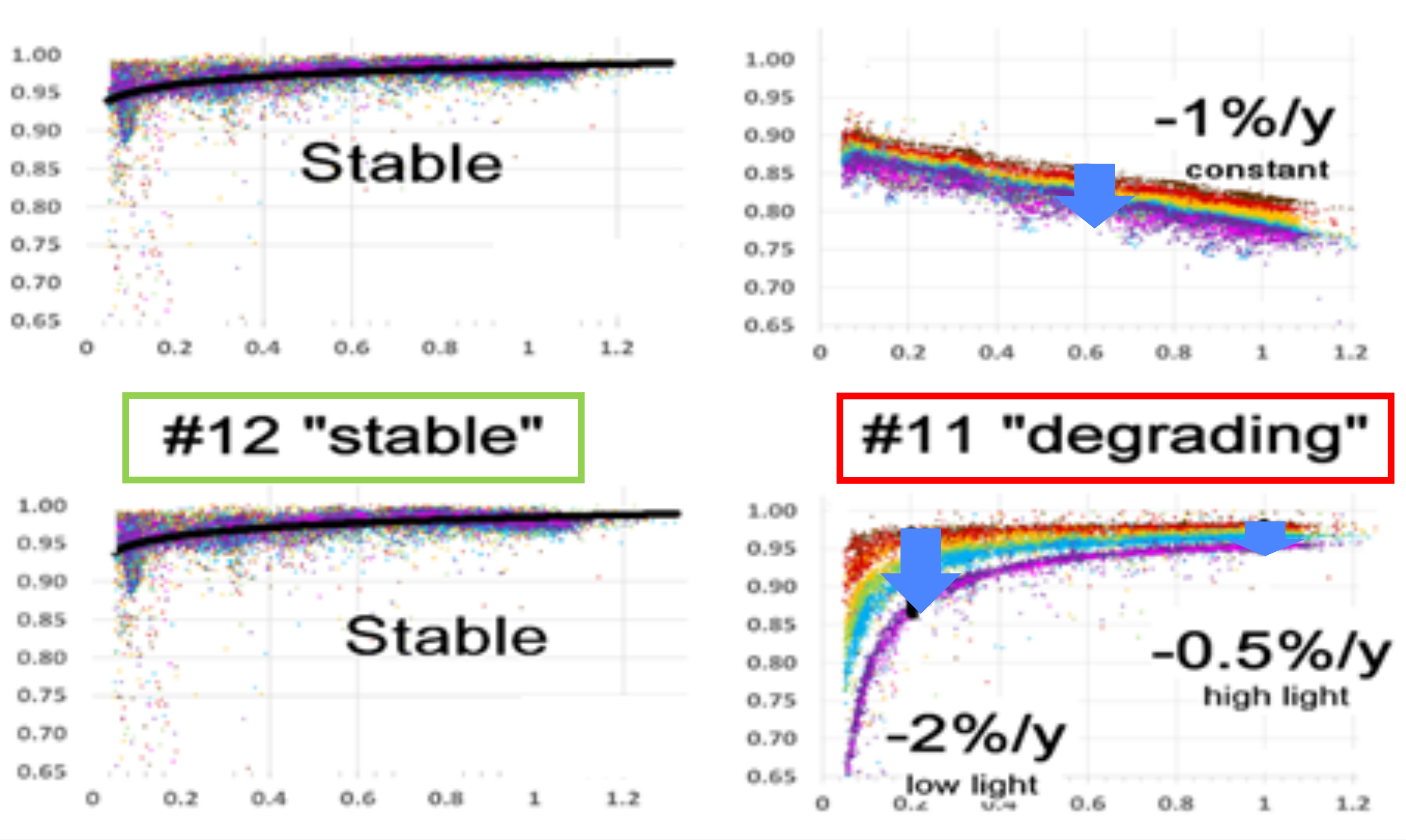
Determining degradation causes and rates from MPM fits to LFM parameters over 7 years

Gantner instruments

nRoc
~Rseries

2010
2011
2012
2013
2014
2015
2016

nRsc
~Rshunt

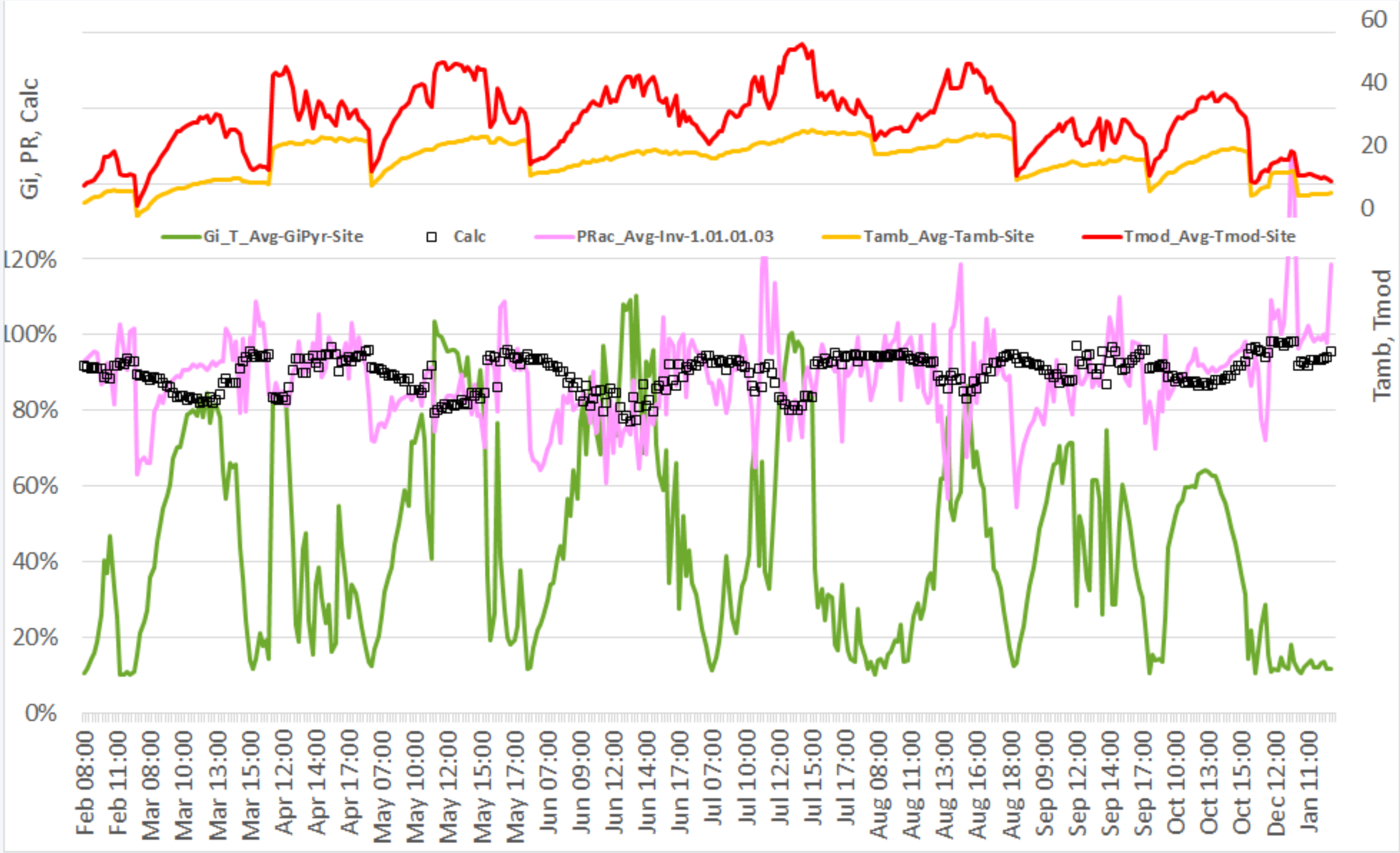


Fitting models to measured data vs. time (1year, UK)

Scattered PR_{AC} due to transient weather but good average fit

PR_{AC}.fit =
MPM function (Irradiance, Tmodule, Windspeed …)

PRac_Avg-Inv-1.01.01.03

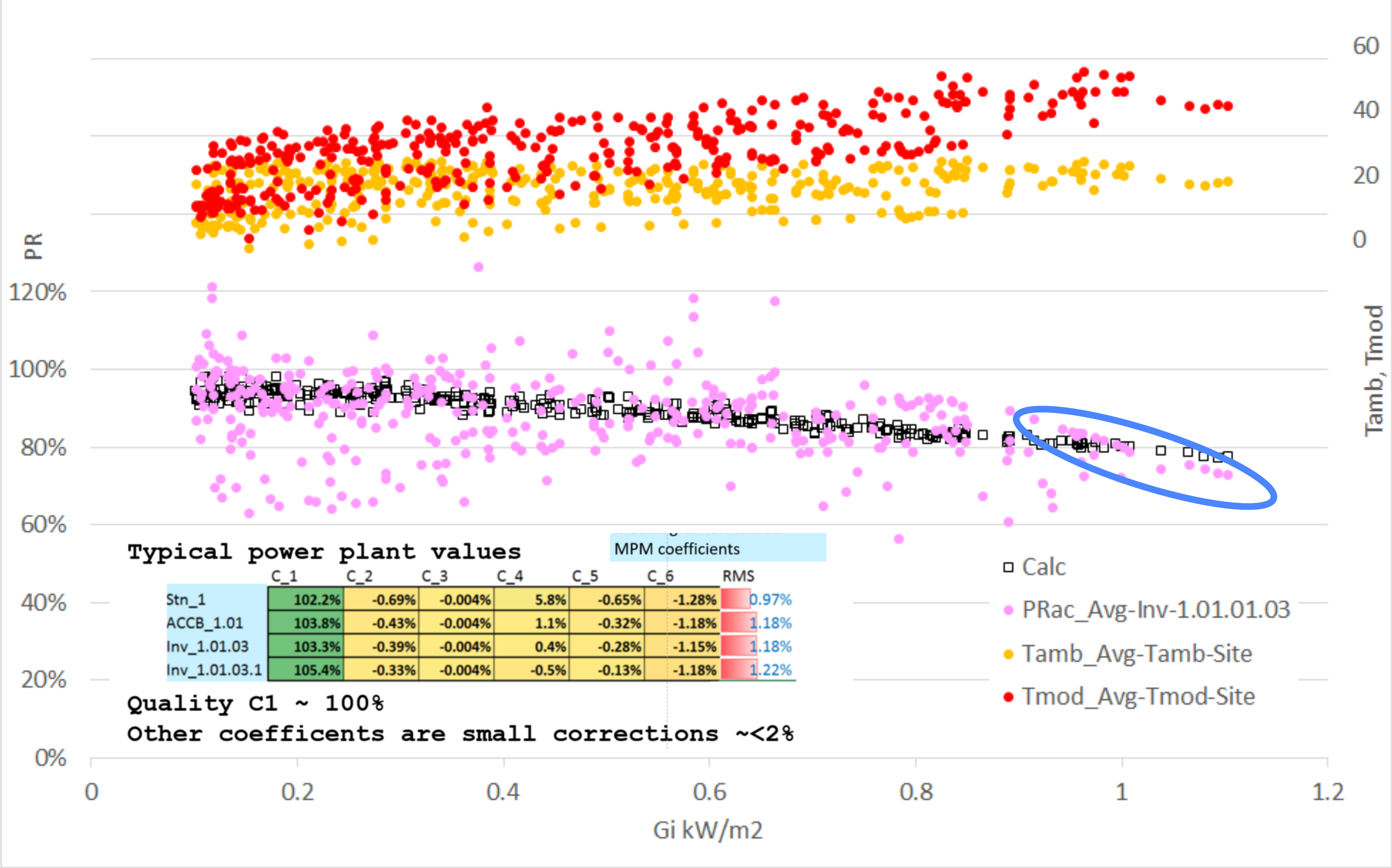


Fitting models to measured data vs. irradiance (1year, UK)

Scattered PR_{AC}, but note clipping at high irradiance > 900W/m^2 needs to be modelled

PR_{AC}.fit = MPM function (Irradiance, Tmodule, Windspeed …)

PRac_Avg-Inv-1.01.01.03



Typical power plant values

MPM coefficients

	C_1	C_2	C_3	C_4	C_5	C_6	RMS
Stn_1	102.2%	-0.69%	-0.004%	5.8%	-0.65%	-1.28%	0.97%
ACCB_1.01	103.8%	-0.43%	-0.004%	1.1%	-0.32%	-1.18%	1.18%
Inv_1.01.03	103.3%	-0.39%	-0.004%	0.4%	-0.28%	-1.15%	1.18%
Inv_1.01.03.1	105.4%	-0.33%	-0.004%	-0.5%	-0.13%	-1.18%	1.22%

Quality C1 ~ 100%

Other coefficents are small corrections ~<2%

Conclusions on fitting models to measured data

POWER PLANT LEVELS

0) Site

1) - Station

2) -- SubStation

3) --- AC Combiner box

4) ---- Inverter

5) ----- MPPT

6) ------ String

7) ------- Module

LARGE POWER PLANTS – MULTIPLE MODULES

- **AC SIDE:** summing power from multiple inverters means “a few” faults at inverter level just reduce output slightly
- **INVERTER:** when working will they have similar performance, so search for outliers with poor performance
- **DC SIDE:** study MPPT and String data on “poorer inverters” to identify problems (and fix them)
- Need to check for shading, clipping, component downtime etc. when fitting power plants
- Check distance from irradiance sensor to array for cloud front delays in variable weather

SINGLE MODULE IV DATA DC

- Can understand and identify any poor behaviour and changes but must allow for variability and wiring losses

Utility Scale Example

Gantner instruments

- Ben Ban, Egypt; Monitoring and grid control
- Project size: 230MW, 1-axis tracking
  - 1400 Combiner boxes, 60 x 2.5MW Inverters, 24 Irradiance sensors/weather stations
  - DC side: 30,000 String currents, 1,400 voltages
- Parameters:
  - 70,000 measured
  - 140,000 Normalized and calculated parameters
    - E.g. PR per component, Aggregation of I, V for each level, limit checks, warning for degradation, temperature limits, …
    - Each parameter has different aggregation methods per time interval (Average, Min, Max, Standard deviation) for direct Loss Factor Model use
- Raw data volume:
  - 400 GB per year
  - 8TB for 20years

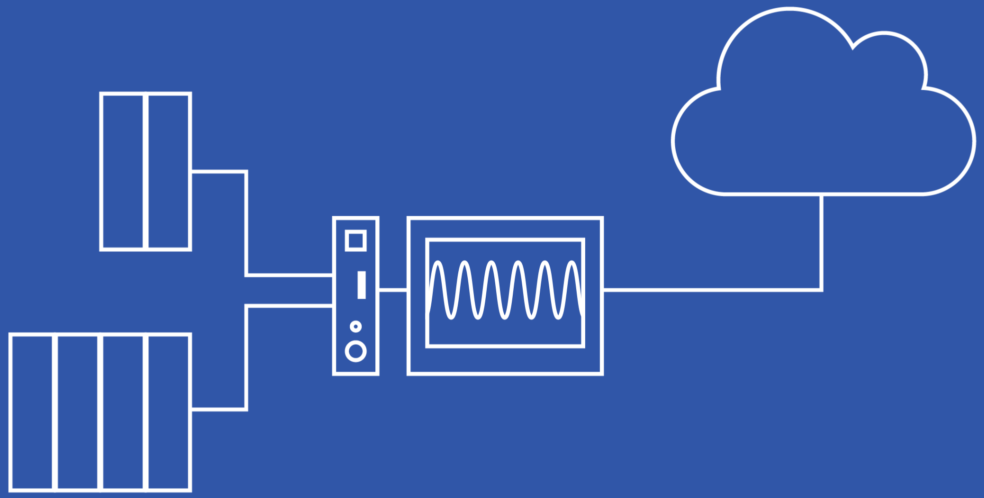
How to store and process very large datasets cost efficiently?

Substantial demand for high performance edge computing in a growing market

Drivers:

- Industry 4.0, big data, AI, ML all need high quality data sources
- More distributed and adaptive monitoring and control applications
- Requires better and faster utilization of data streams
- Flexible data architecture to meet customer needs

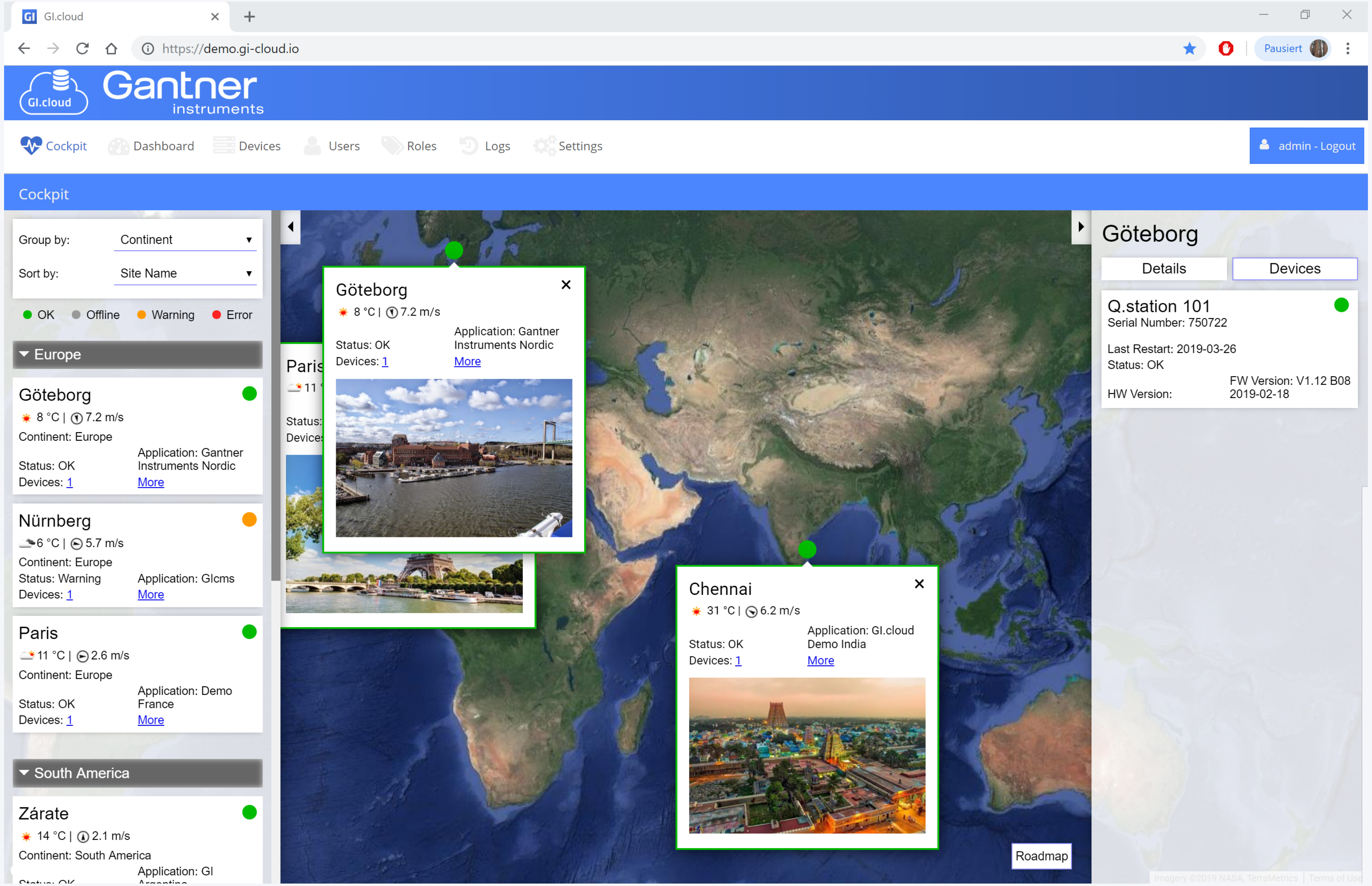
Gantner instruments



Technology and market designers understand the need for powerful cloud and edge computing in combination with adaptable high resolution measurement down to micro-seconds (µs).

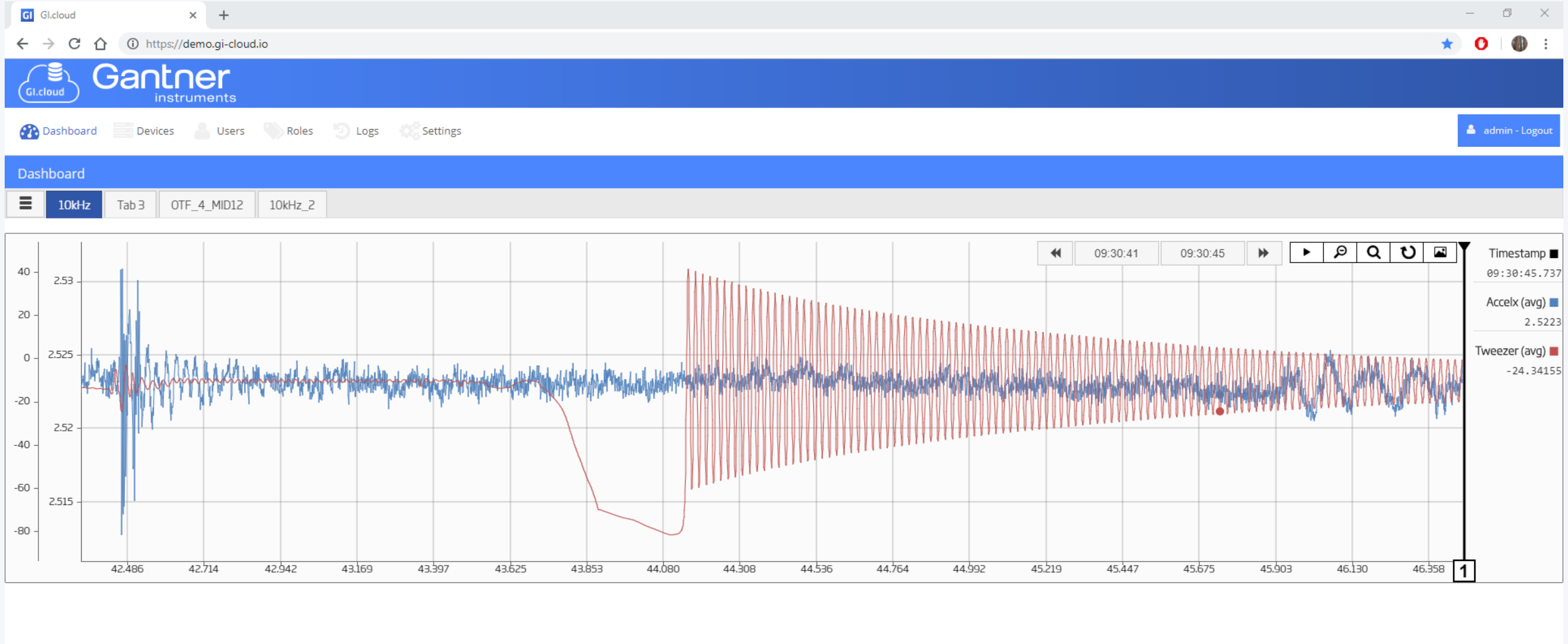
Real time data visualization

- From controller to cloud platform
- Create our own dashboards (read only, read/write)
- Display Real time charts (different controllers, merged data streams)
- Zoom into the single event/Raw data
- Cockpit: see device location, status, warnings, meta data



Real time data visualization

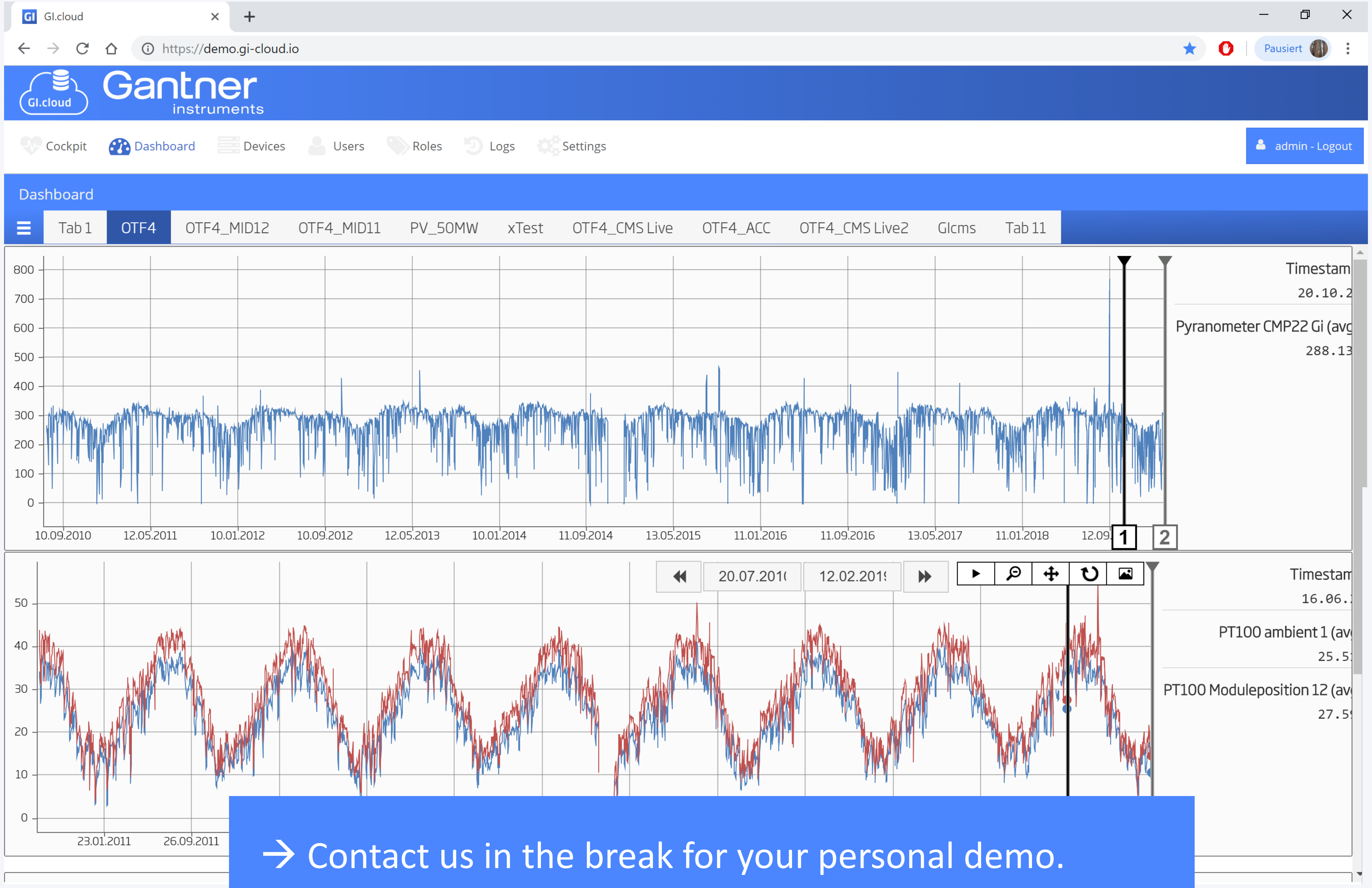
- From controller to cloud platform
- Create our own dashboards (read only, read/write)
- Display Real time charts (different controllers, merged data streams)
- Zoom into the single event/Raw data
- Cockpit: see device location, status, warnings, meta data



Lifetime data

Gantner instruments

- Efficient storage of full life cycle data streams, e.g. 30 years of device data
- Easy data access thanks to hot and cold data storage from 30 years to microseconds
- Store and visualize triggered data and different sampling intervals



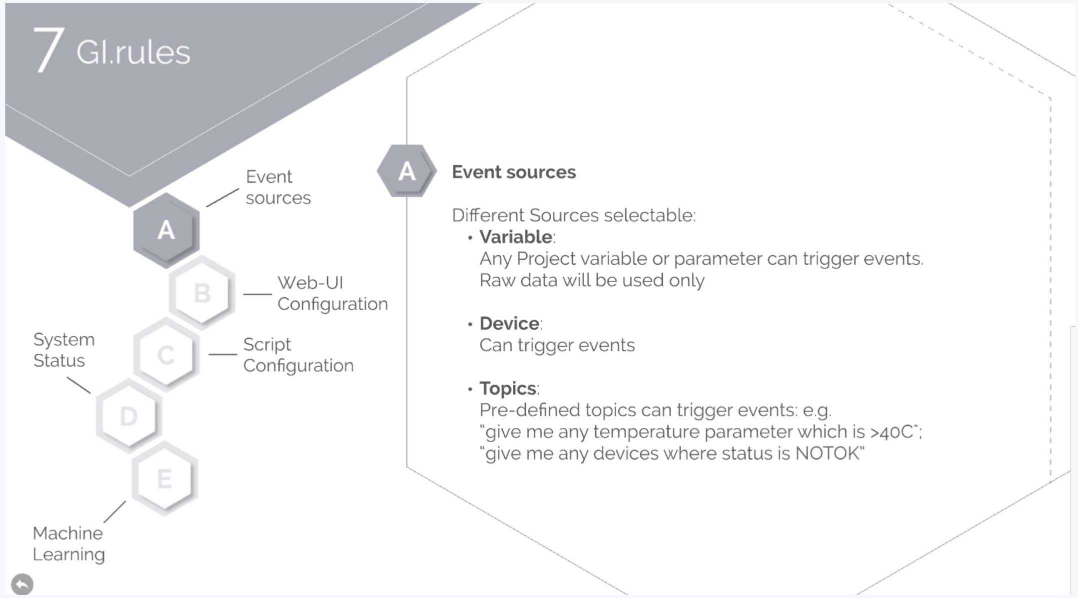
→ Contact us in the break for your personal demo.

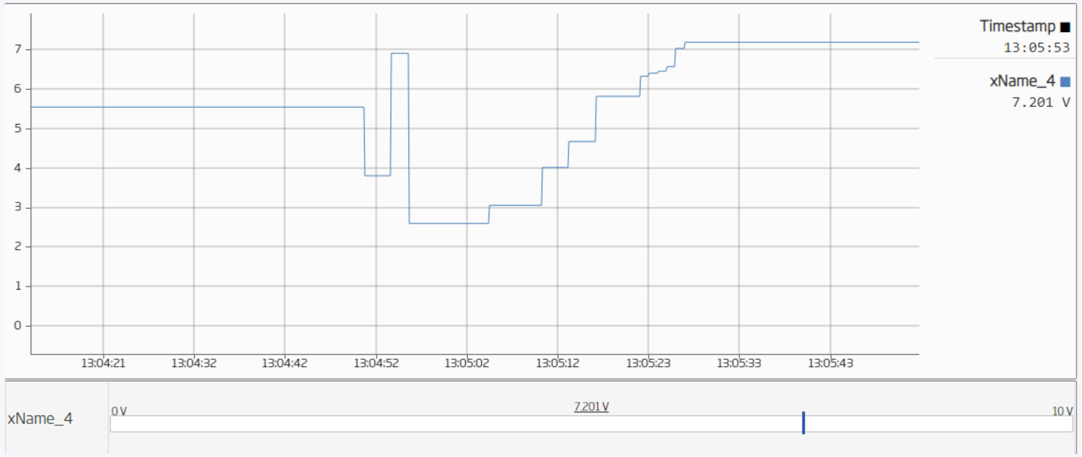
Instant reaction to data

Rules

- Configure predefined rules that react to full data stream (also MHz and non time synchronized)
- Different event sources:
  - Variable
    Each Raw data time stamp will be checked
  - Device
    Can trigger events
  - Topics
    - “pre defined topics can trigger event. E.g. “give me any temperature parameter which is >40C
    - Give my any device where status is “NOT OK”
- Configure and view notifications
- Reporting: triggered report creation with real time data

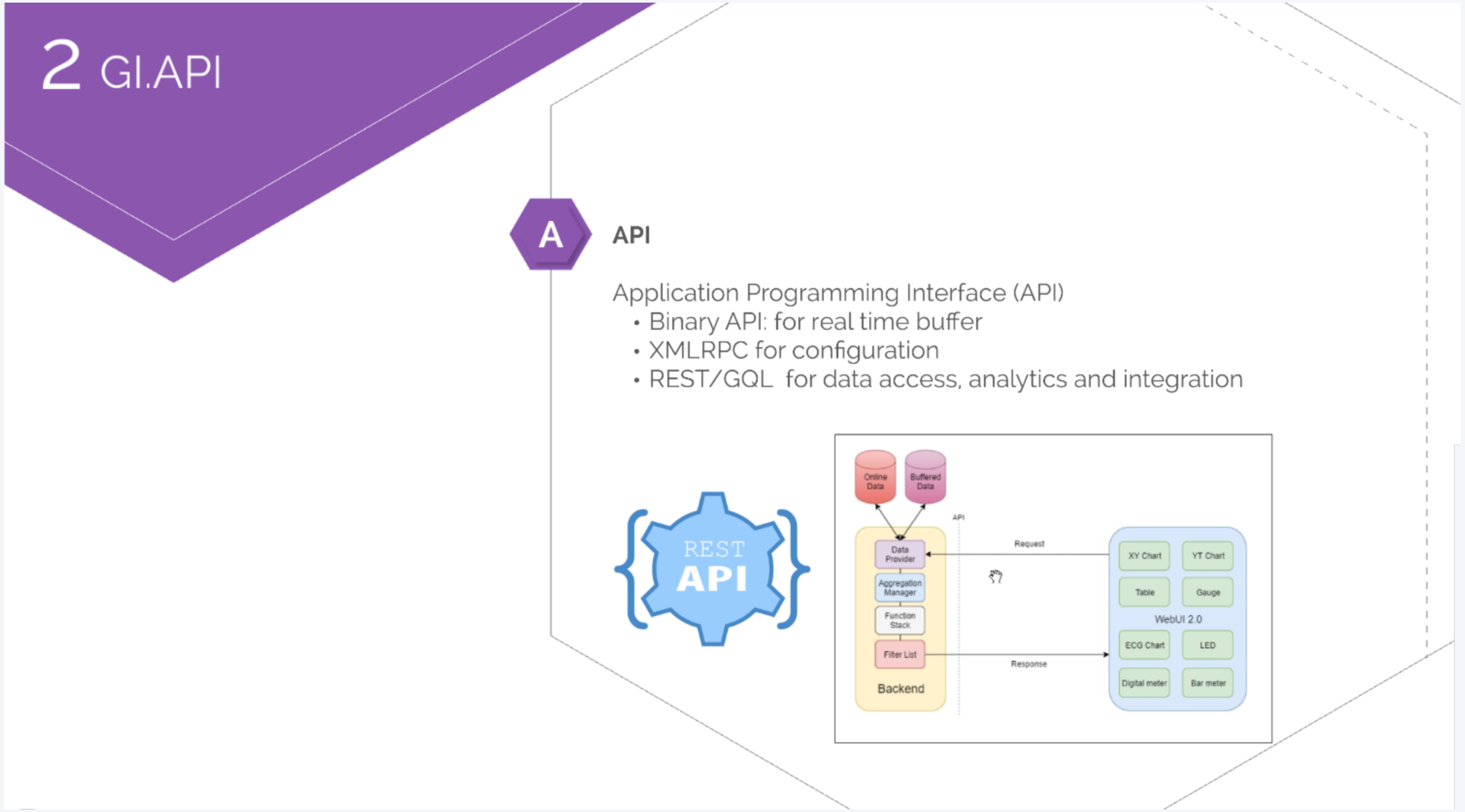
- Bidirectional communication with devices





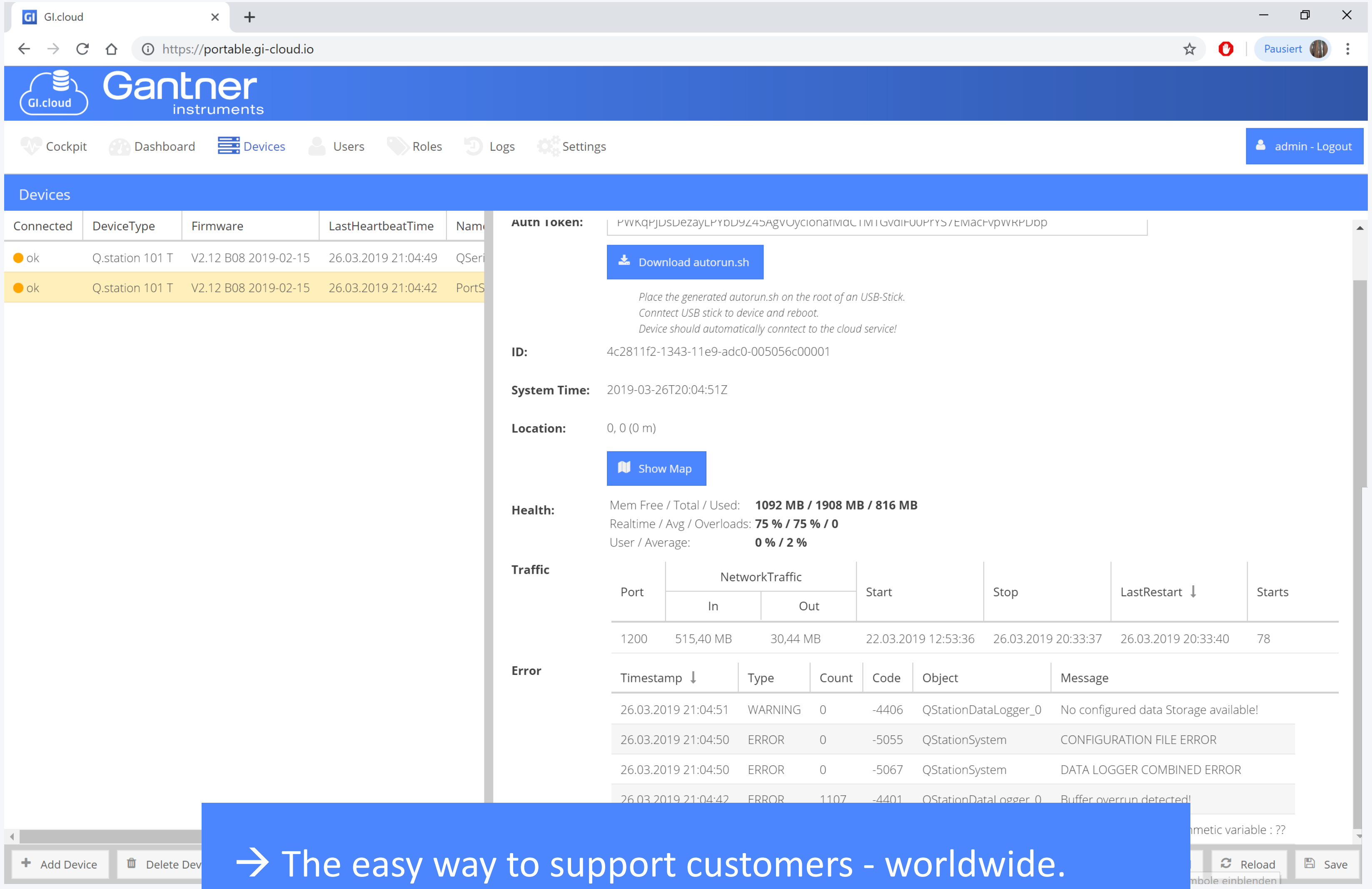
Integration

- APIs: customer & partners can integrate GI.cloud backend into their IT stack
- Build customized Front end applications (java script)
- 3rd party data integration API; visible on the fly on frontend [e.g. 50MW PV data]
- Cloud API Examples:
  - Machine Learning performance prediction (Cyprus, Project IPERMON)
  - VKW (Austria)



Device Management

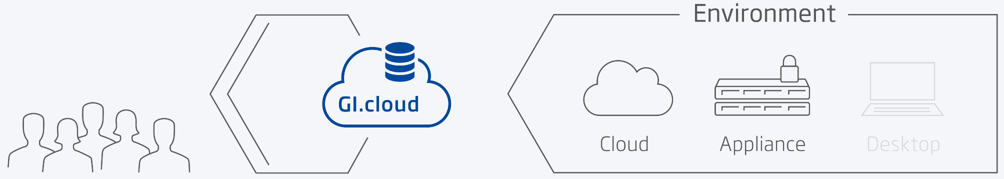
- Access your controller from GI.bench anywhere, anytime with GI.cloud remote service
- Manage all your devices from one cloud instance
- Show status, data traffic, alerts, logs – from all remote devices
- Only internet port needed (https, port: 443, encrypted)
- Forget about VPN, additional fees, port discussion with customers IT, ..



→ The easy way to support customers - worldwide.

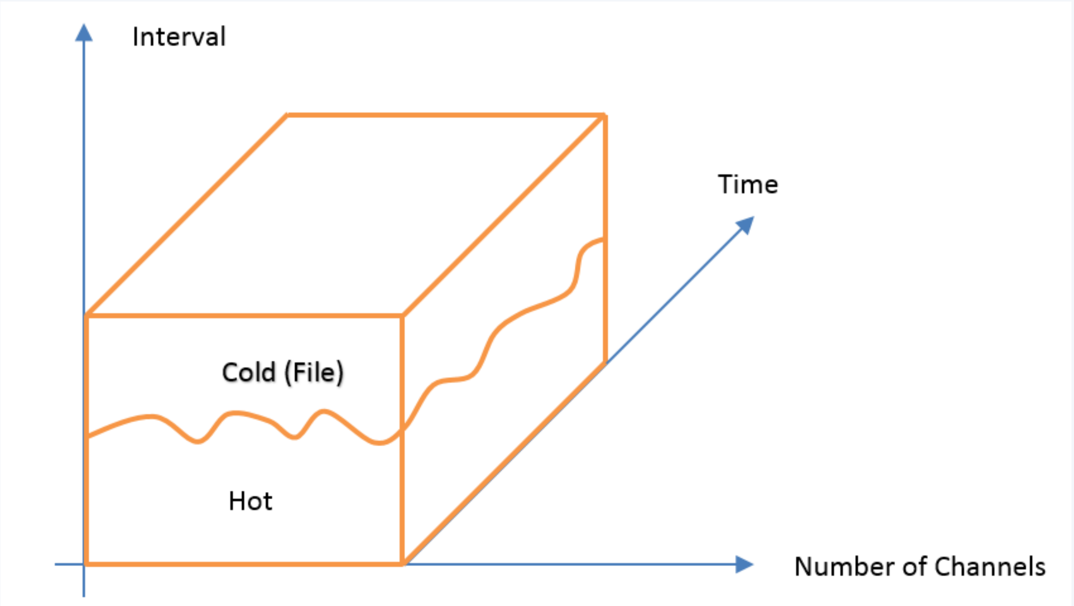
Backend Architecture

- Requirements:
  - Build a universal efficient data backend for MHz and asset life time data
  - Browser based only
  - Single / Multi-Tenant
  - No time resolution limitation (micro-seconds to minutes)
  - (Re-) configurations of Hardware visible on the fly
- Generic, normalized Backend scalable from MHz sampling speed to long term monitoring over decades.
- Efficient data storage and data access for optimize running cost
  - Raw data ("cold data") in distributed streaming platform ("Ringbuffer") by Apache Kafka
  - "hot data" in fastest time series database by Crate.io
  - Project specific aggregation (mean, maximum, minimum, standard deviation), data replay
  - Data backup included
- ONE API for data access
- Enrichment for KPIs or customer specific
- Provides scalability, redundancy and resiliency



Distributed Streaming Platform and Database

- Scalable
- Clustered



APACHE kafka

CrateDB

Analytic Snapshot based on GI.cloud data backend

How to compare PV Performance for many different components over time

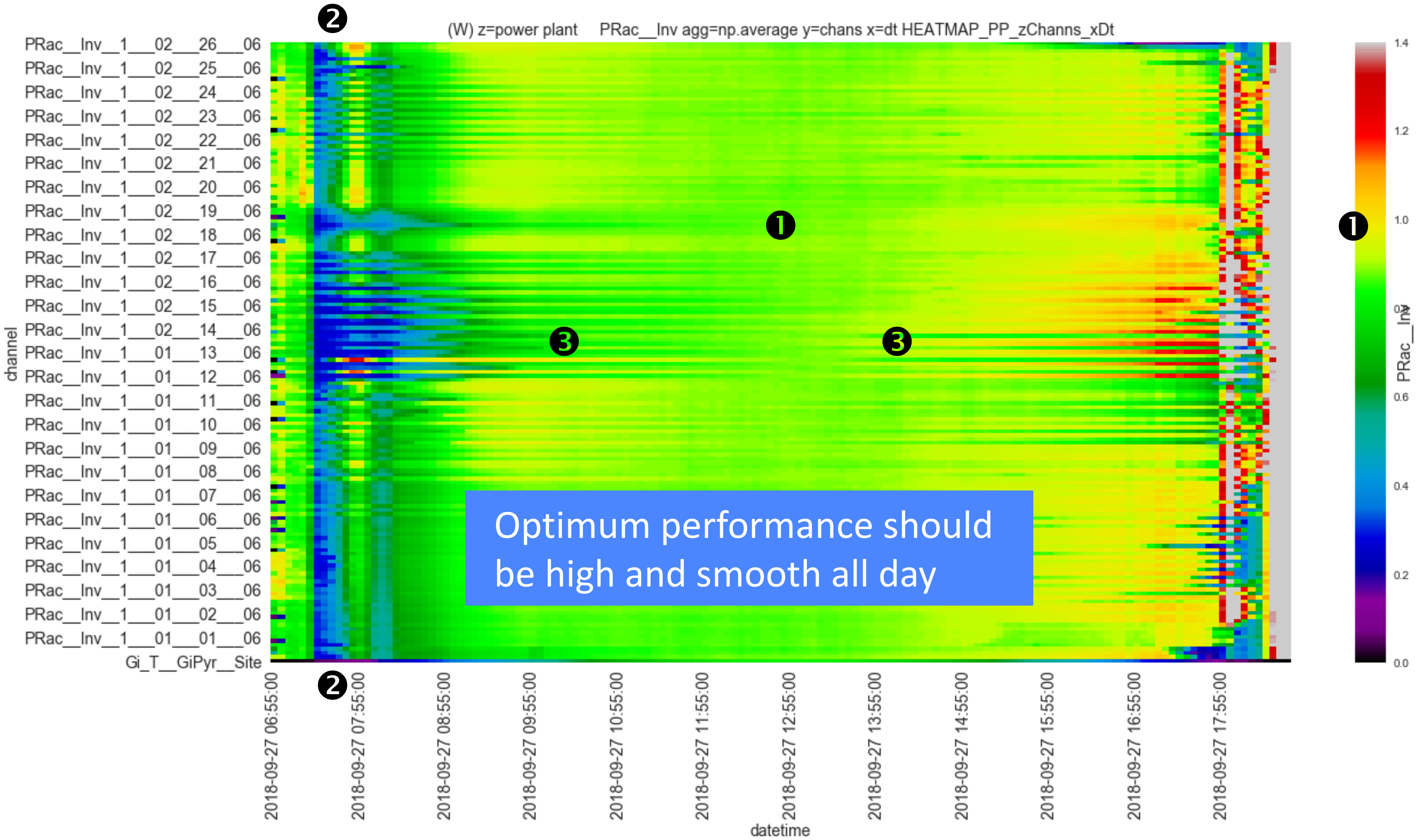
Performance ratio (colours red=best blue=worst) for 156 inverters↑ and time→

PR = Pmeas/Pnom/G

❶High performance ratio (near 100%) is light green to yellow

❷Early morning <08:00 there may be some problems of shading or turn on (blue)

❸ Some inverters that are worse in the morning are better in the afternoon>15:00 – it's likely that these arrays are facing westwards



Comparing standard deviation of similar components – use to find faults

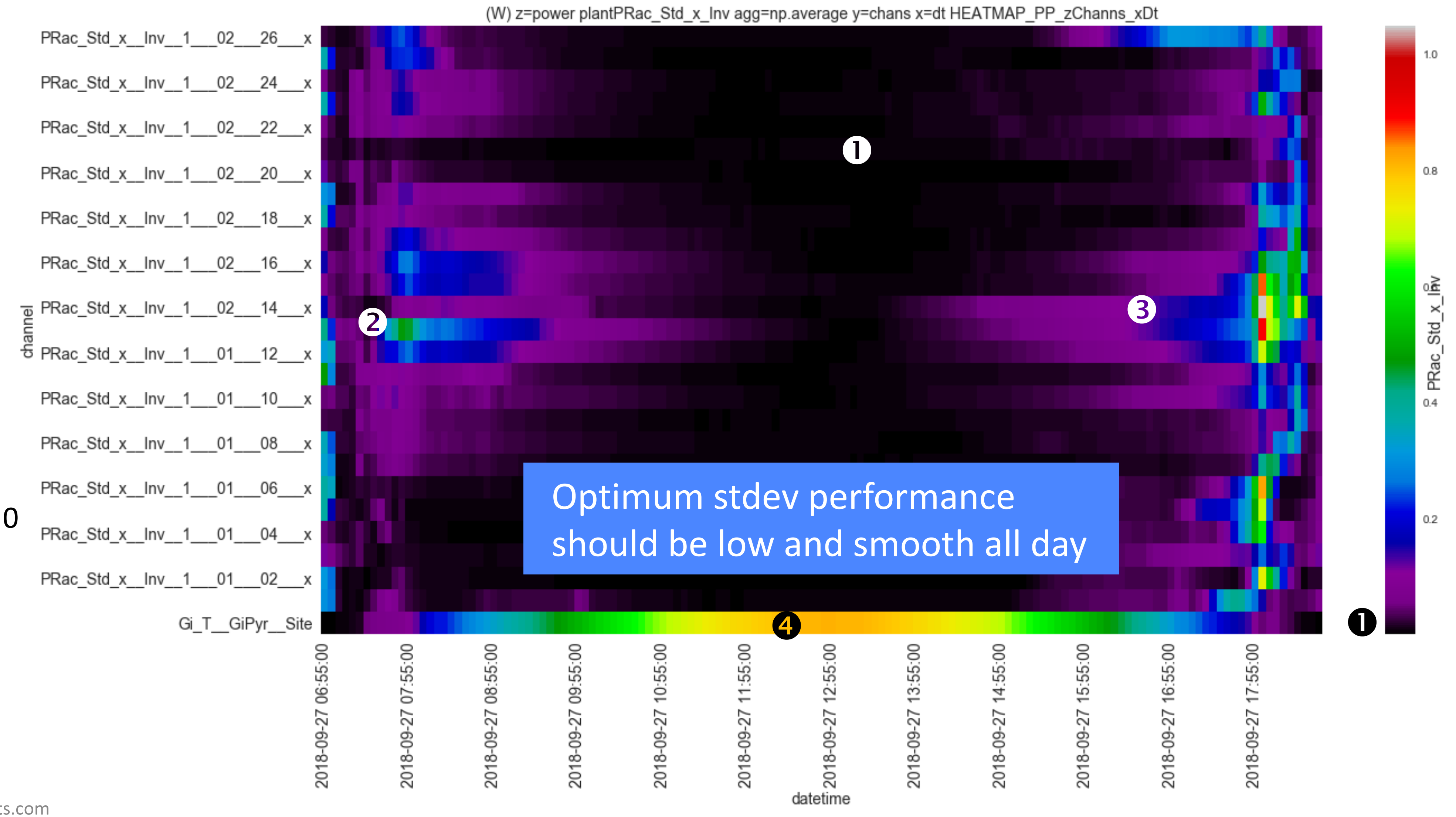
StdDev of PR (colours blue=worst black=best) for 6x26 inverters ↑ and time →

❶ Uniform performance has low stdev (black)

❷ Higher stdev morning from the inverters including some west facing

❸ Higher stdev from same inverters in the afternoon

❹ Lowest trace shows normalised irradiance 0 to 0.8kW/m2



How to compare Inverter Efficiency for many different components over time

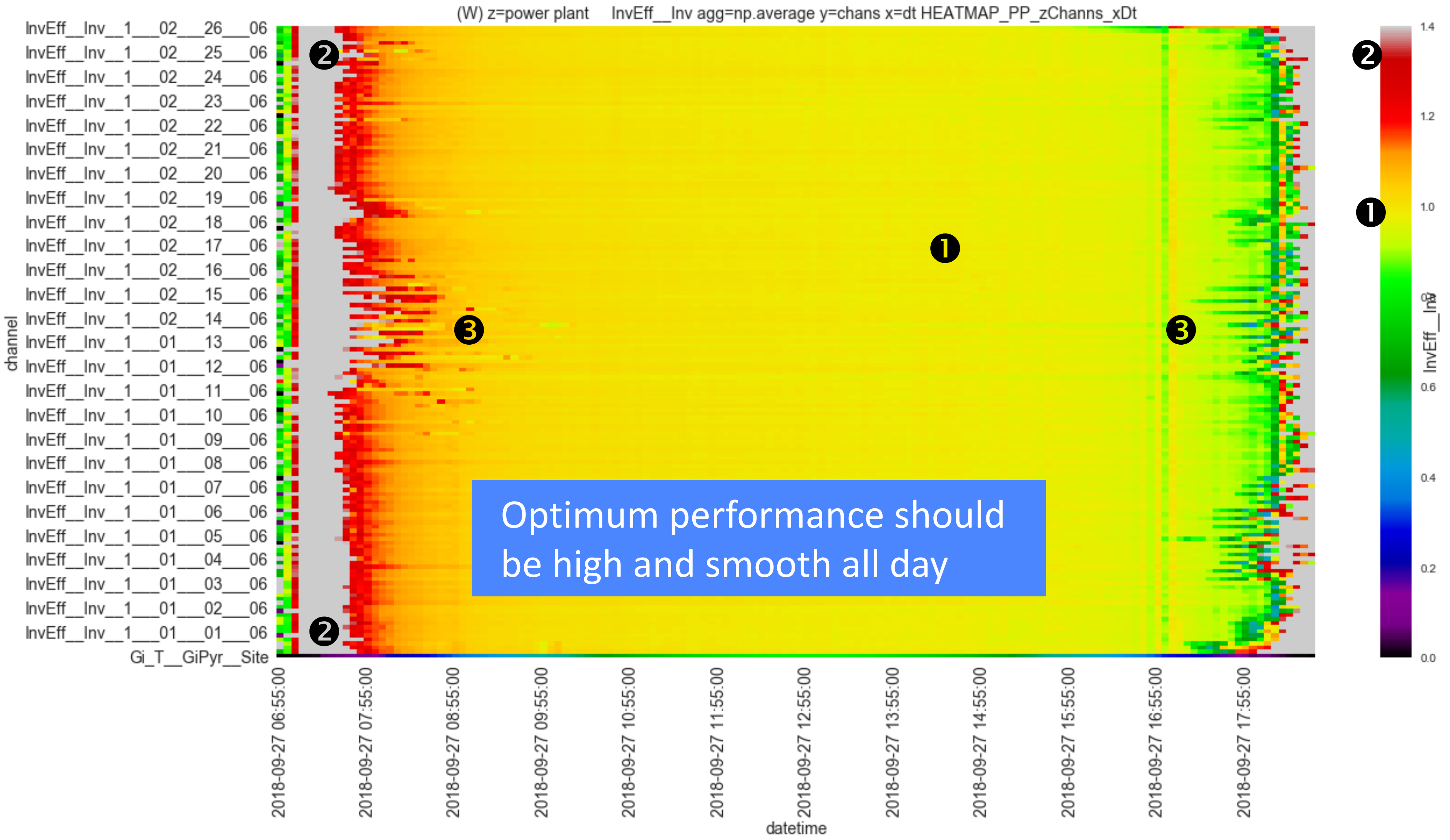
Inverter Efficiency (colours red=best blue=worst) for 156 inverters🡅 and time🡆

❶High Inverter efficiency (near 100%) is yellow

PR = Pmeas/Pnom/G

❷Very early morning there appear to be reports of <100%

❸ Some inverters that are better in the morning are worse in the afternoon



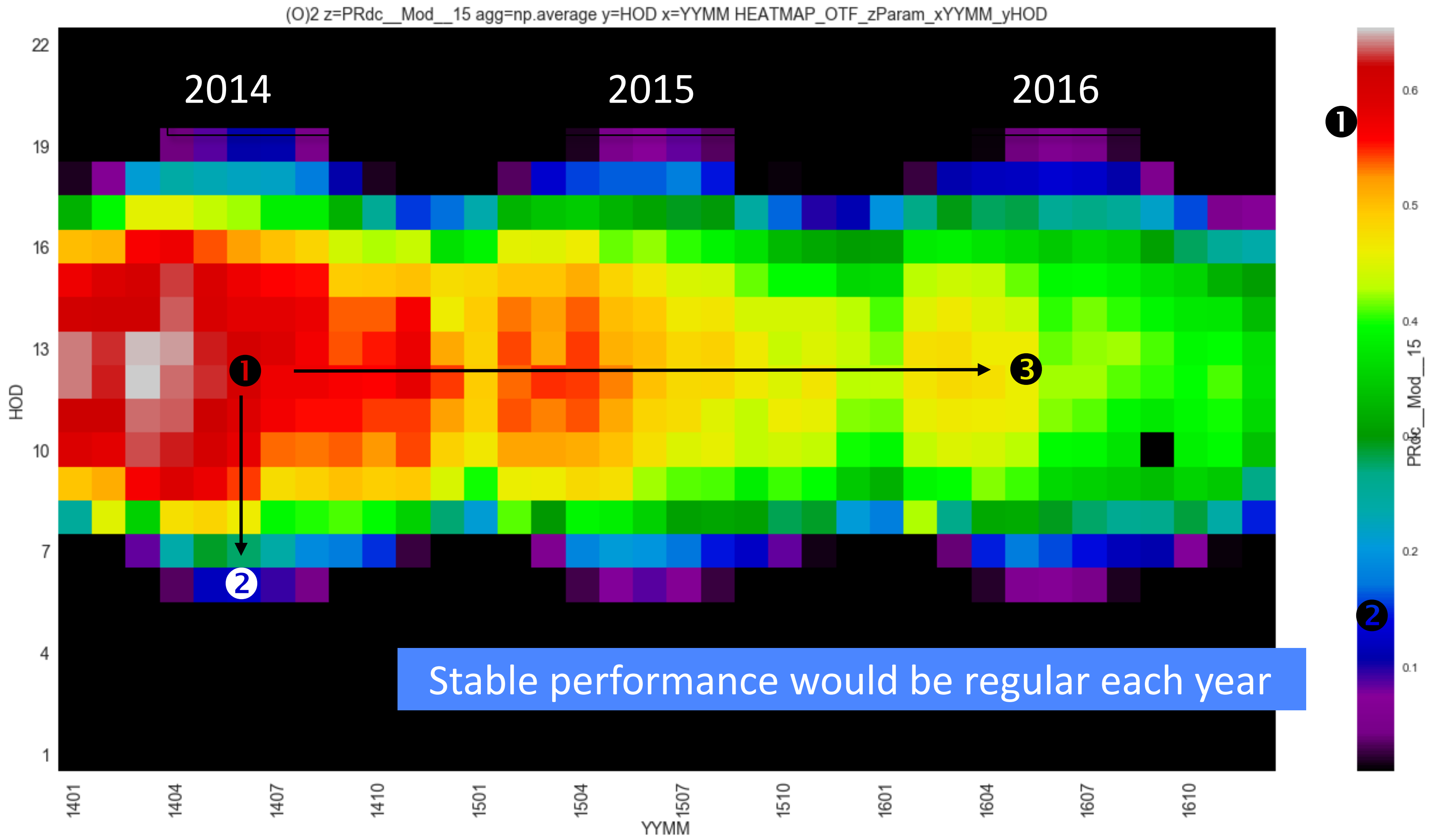
Determining performance stability PR_{DC} by time of day and month

Average PRdc by hour of day 1…24 HOD🡅 and YearMonth 1401…1612 MOY🡆

- ❶1st Summer 1406 Module performance although poor was highest during the day ~0.6
- ❷ it was worse at lower irradiance ~0.2
- ❸>2 years later 1606 this module has degraded badly and is below 0.45

Degradation rates can be obtained by the fall per year from ❶ to ❸ e.g. 0.6 to 0.45

Note longer summer days give "taller" datasets 06:00 to 19:00



Determining performance stability PR_{DC} by irradiance and month

Average PR_{DC} by Irradiance 0 … 1200W/m² G_i🡹 and YearMonth 1401…1612 MOY🡺

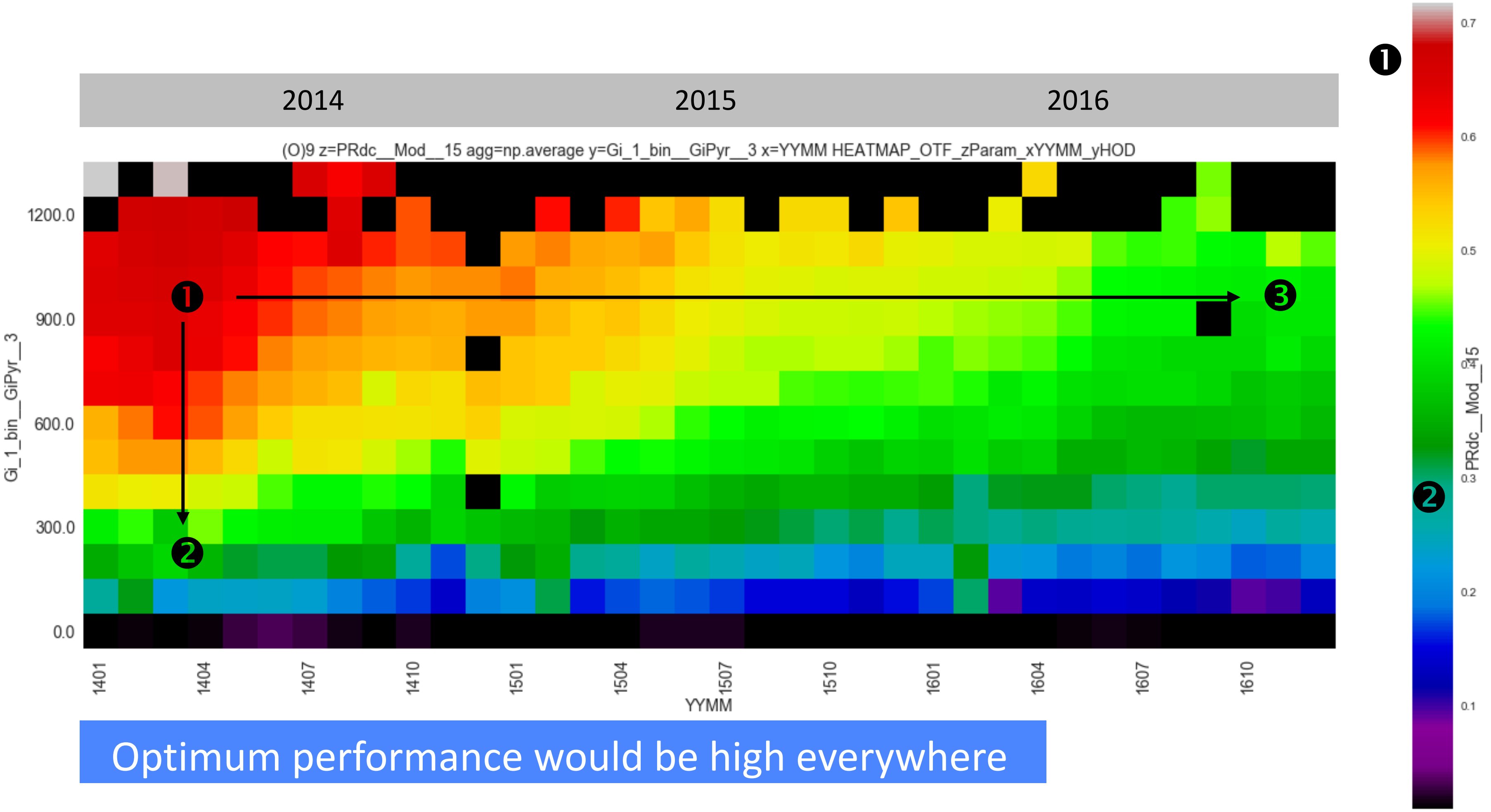
➊Initially Module performance was quite good PR~0.7 at high irradiance 1000W/m2

➋ and worse PR~0.3 at lower irradiance 200W/m2

➌>2 years later this module has degraded badly

Degradation rates can be obtained by the fall per year from ➊ to ➌

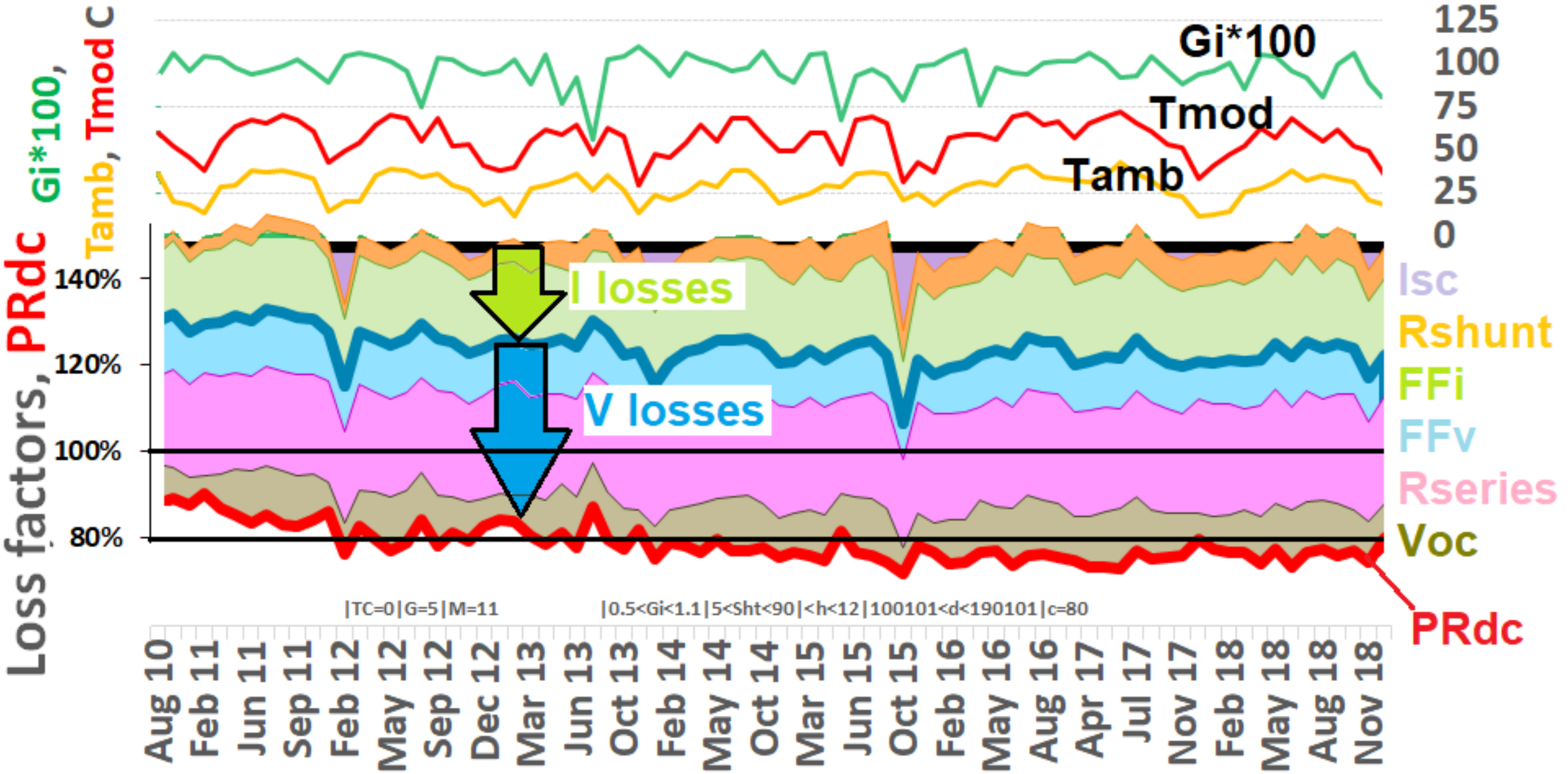
Low light performance LLEC can be got from ➋ / ➊



Optimum performance would be high everywhere

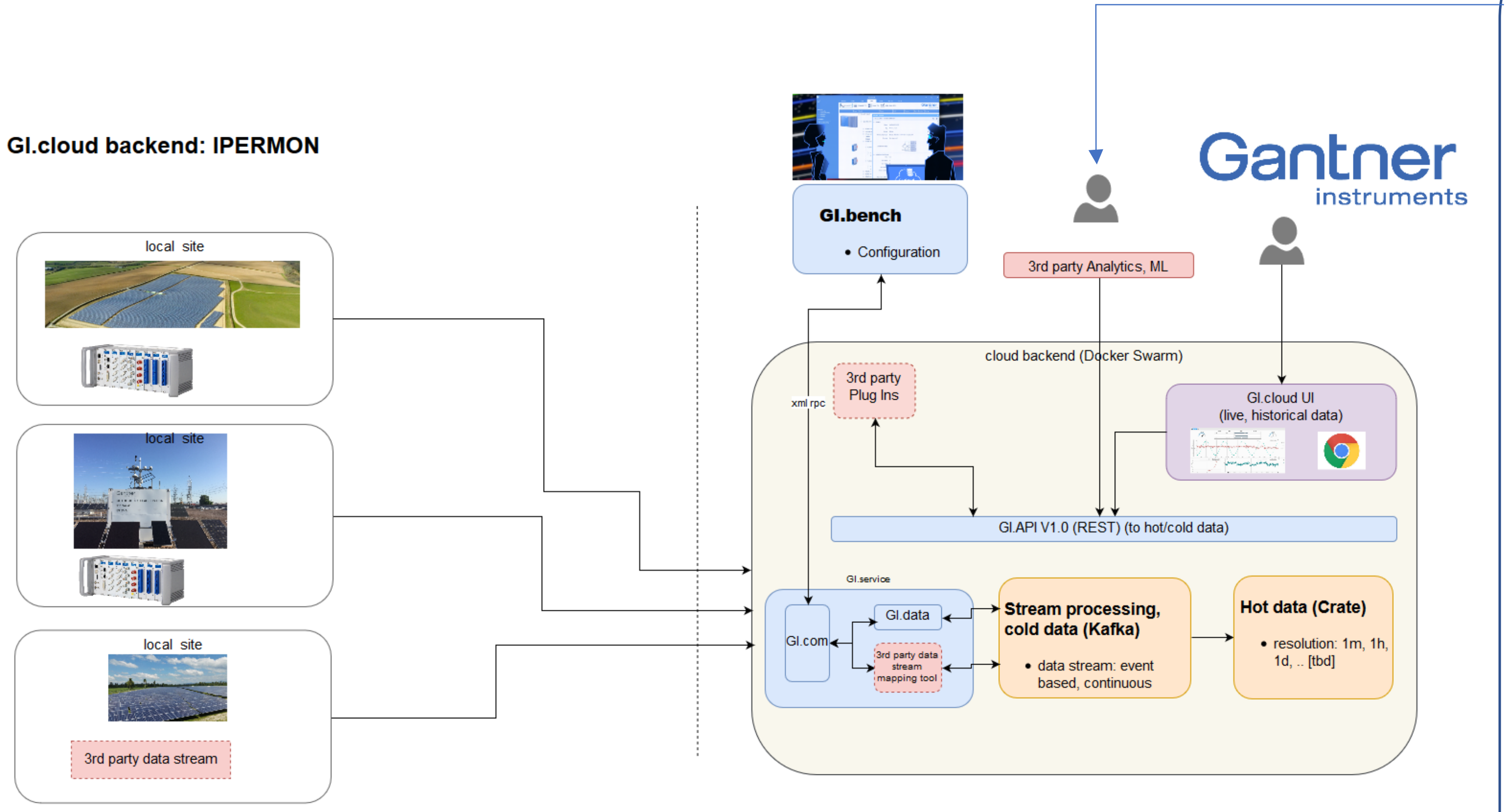
Degradation analysis

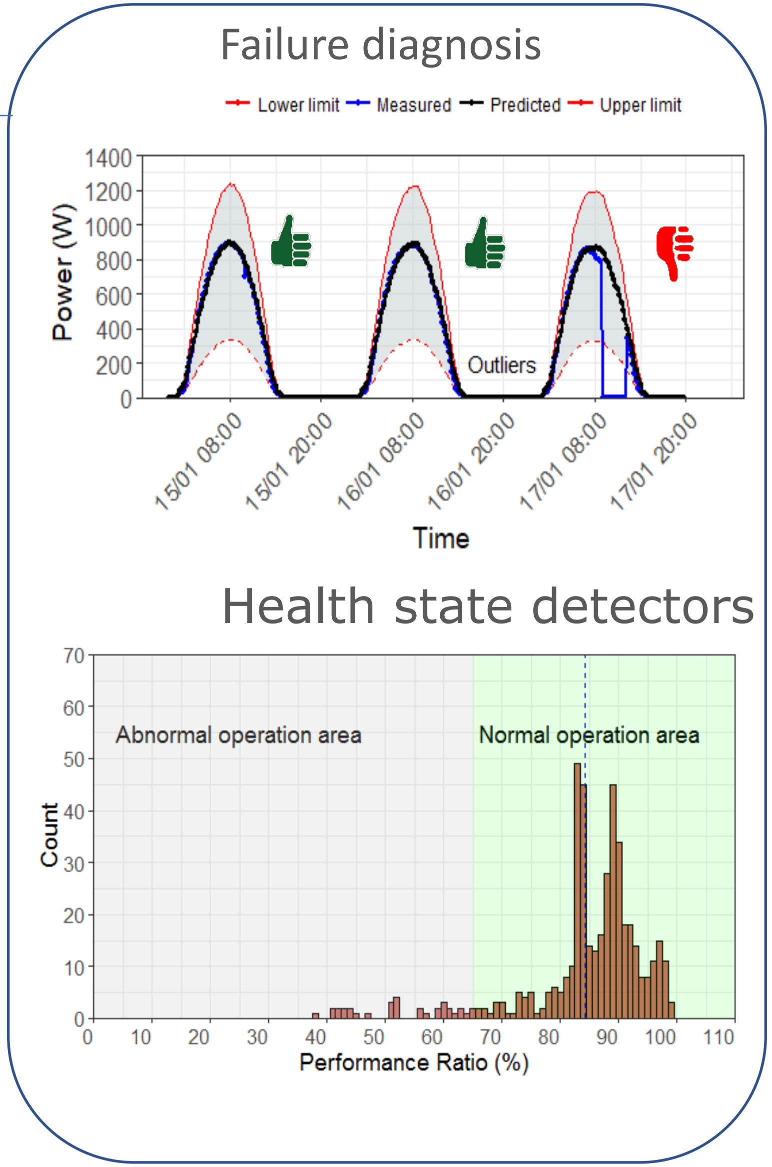
Drill down to responsible parameter to get meaningful conclusions



Integration of diagnostic plus predictive analytics

Machine learning with Project IPERMON





See next talk from Marios Theristis (University of Cyprus):
"Performance Evaluation of PV Power Predictive Models for Realtime Monitoring"

For more information on using MPM/LFM with the GI OTF data checking IEC61853 1-4 :

Gantner instruments

46th IEEE (PVSC) Thursday, June 20th
ORALS - AREA 8: SYSTEM PERFORMANCE AND MODELS @09:45
650 "**Checking the new IEC 61853.1-4 with high quality 3rd party data to benchmark its practical relevance in energy yield prediction**"
Steve Ransome, Juergen Sutterlueti.

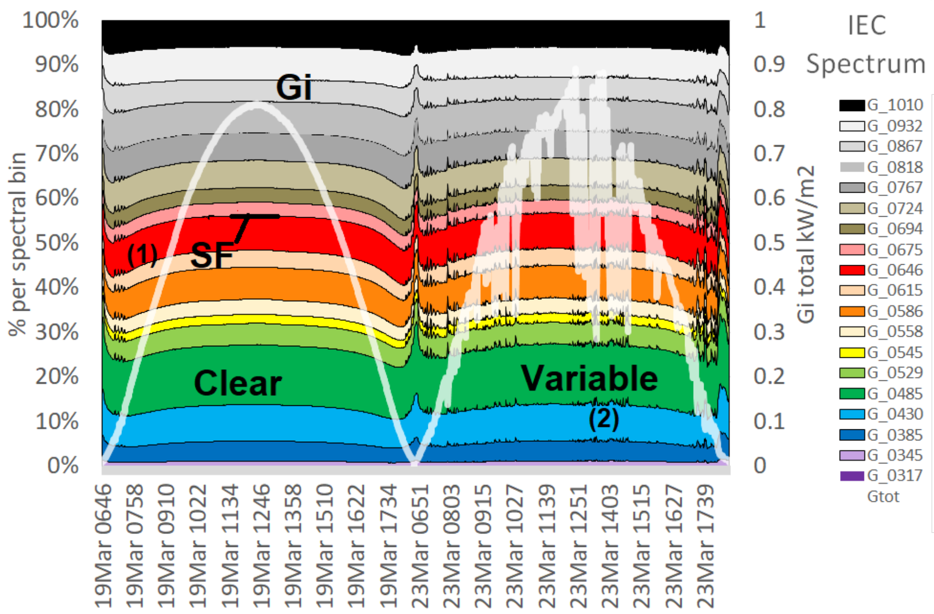
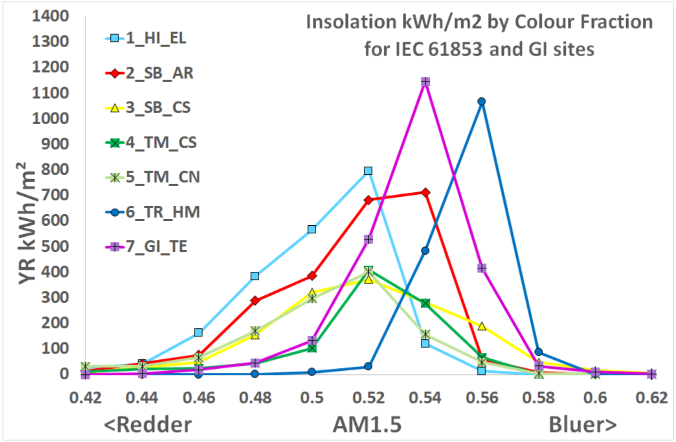
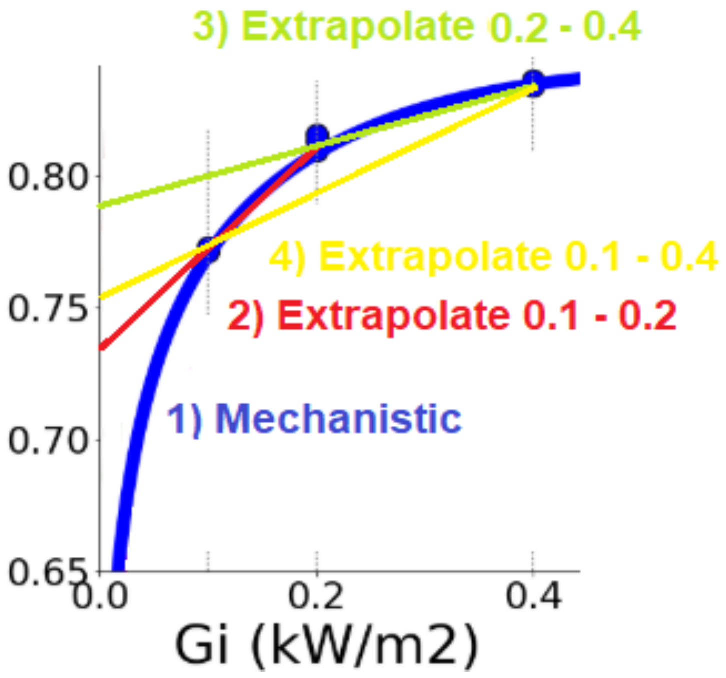


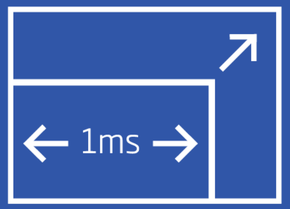
Fig. 8. Gantner Instruments' measurements of spectrum bands as defined in IEC 61853-3 for clear and variable days in Tempe, AZ.





Conclusion

- Accurate and fast PV measurements using a scalable, distributed cloud backend enable an easy understanding of performance for optimization, loss analysis and fault finding using mechanistic models (LFM and MPM) and ML techniques
- System should be able to run as cloud or local installation and allows fast and flexible data access
- APIs should allow read/write access to the backend and connection of own analysis features (not visible to platform vendor)
- PVPMC libraries are included and can be applied to any data source



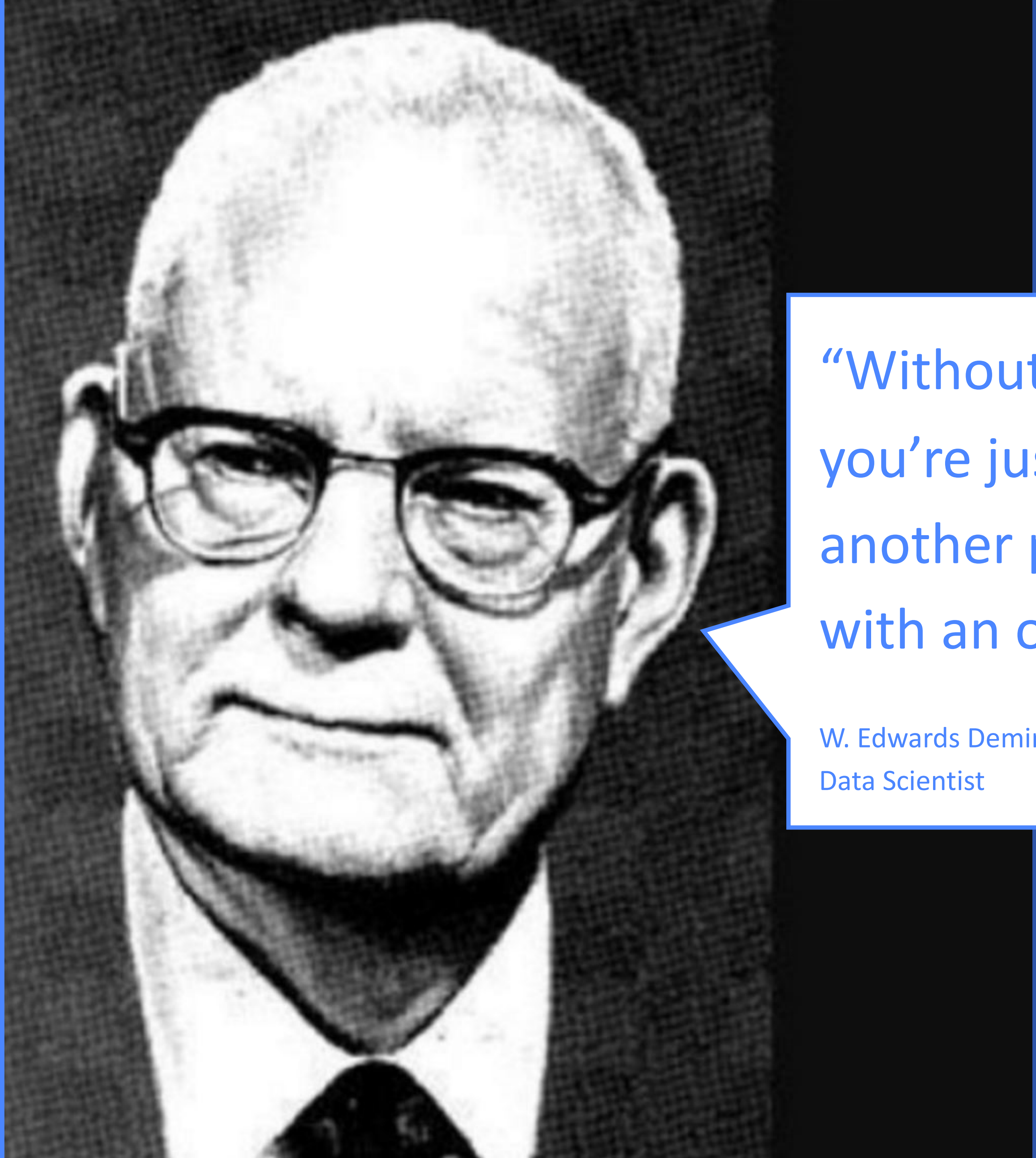
Efficient time series data processing (down to micro-seconds) due to scalable platform

Clustered hot/cold data storage for minimum running cost

APIs designed for fast integration into customers‘ platform and data access

PVlib

PVLib inside



“Without data you’re just another person with an opinion.”

W. Edwards Deming,
Data Scientist

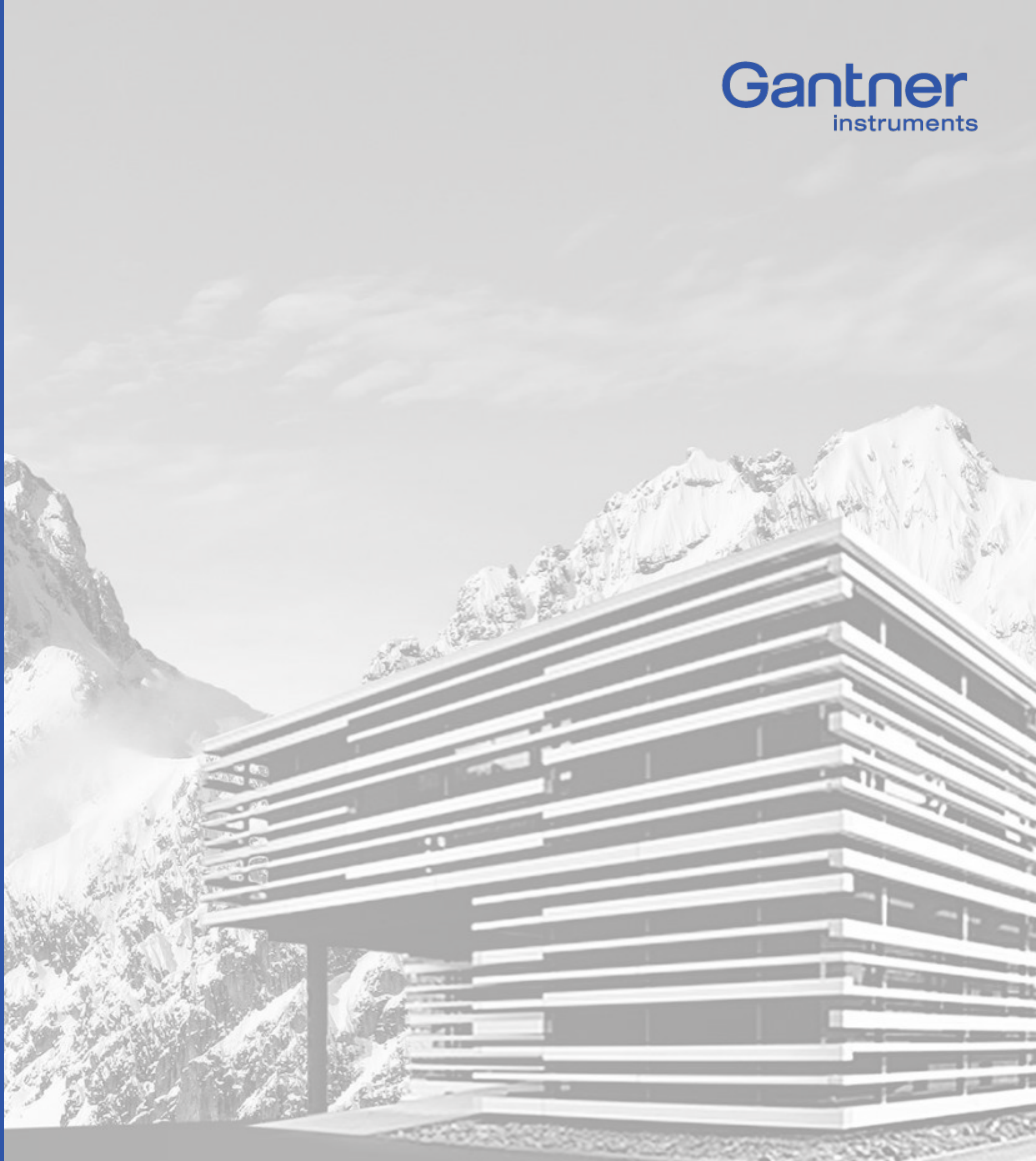
Gantner instruments

Turn Data Into Information.
Turn Information Into Customer Benefits

Thank you very much!

j.sutterlueti@gantner-instruments.com

Appendix



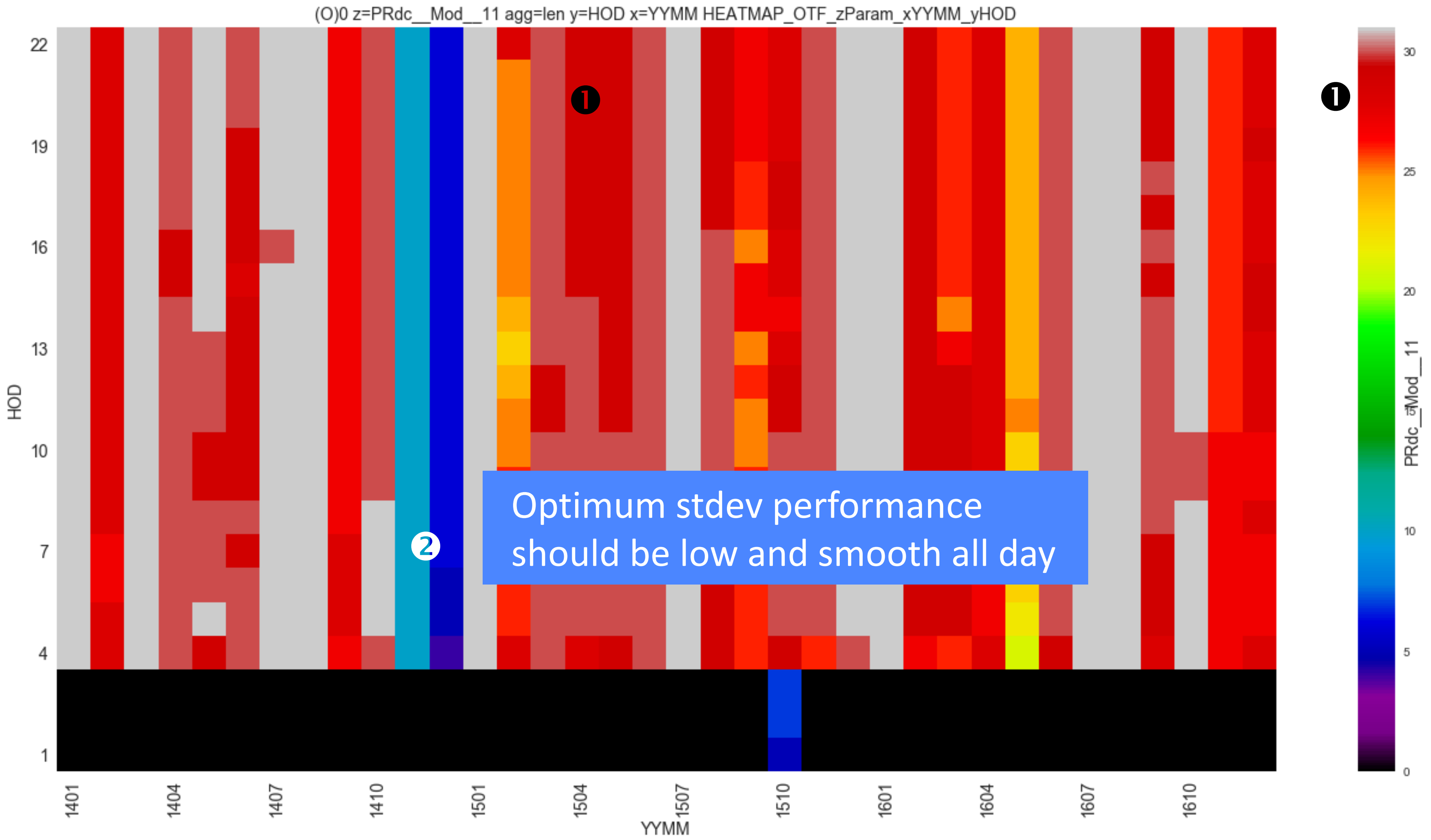
Evaluating uptime by counting measurements each hour and month

Count of measurements (colours) by HOD🡅 and YearMonth YYMMday🡆

🡅HOD 4:00 to 22:00
🡆YYMM 1401-1612

❶Full uptime will have 28-31 measurements per calendar month (red – grey)

❷ Vertical lines of other colours indicate some downtime e.g. blue = only ~10 days measured / month



Determining climate data e.g. average insolation by month and hour

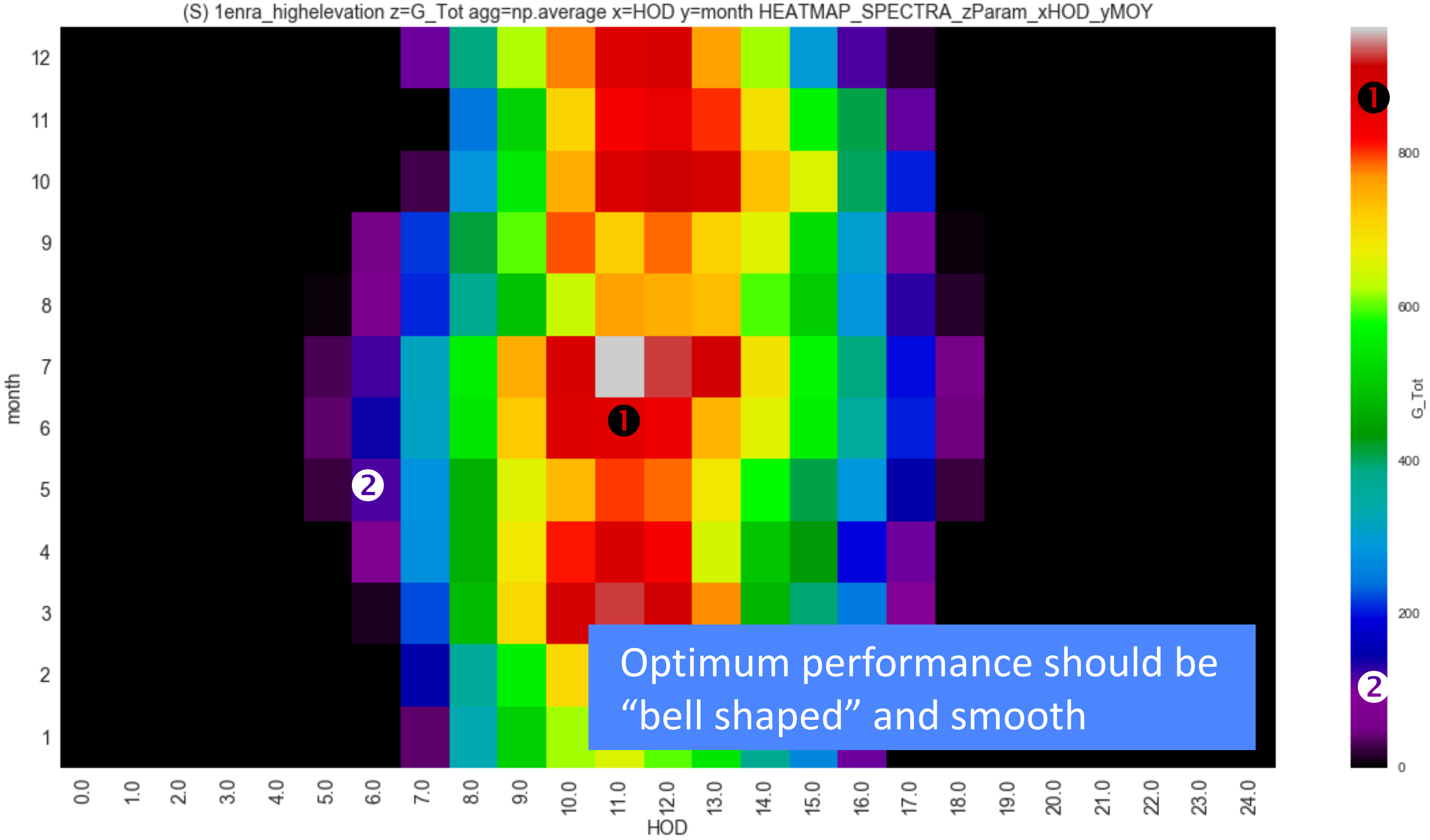
Average Insolation per MonthOfYear 1…12 MOY↑ and HourOfDay 0…24h HOD→

Predicted system performance needs to evaluate climate data e.g. G, Tamb, Tmod, WS, RH …

❶ Maximum average irradiance here is ~900W/m² 12:00 in month 7 July Summer

❷ Average irradiance falls in winter (Months 1 and 12) and also morning and evening (e.g. 6:00 and 18:00)

Can also check irradiance sensors or good vs. bad weather which affects power output.

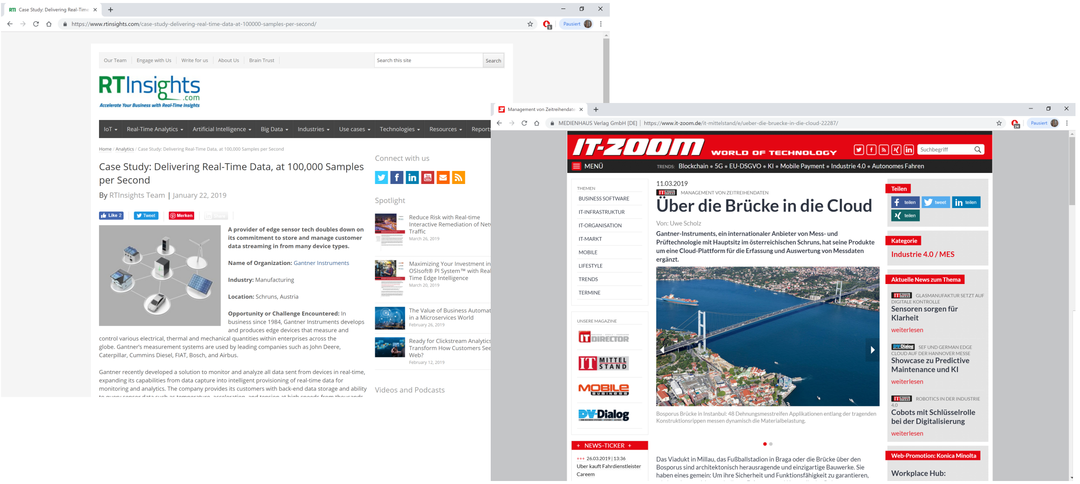


Optimum performance should be "bell shaped" and smooth

Futher Information

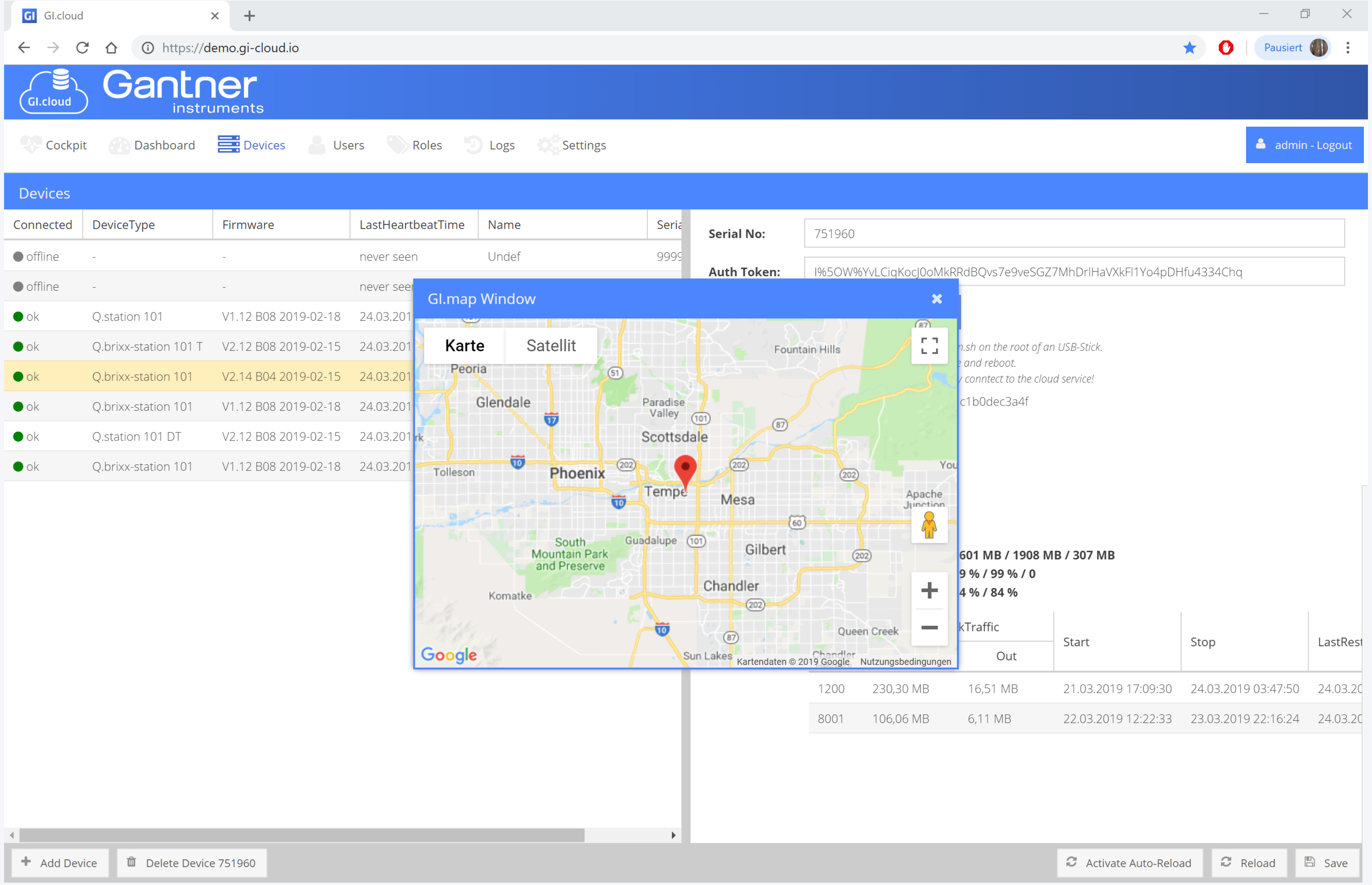
- Ask for demo log in at:
  - gi.cloud@gantner-instruments.com
  - Subject: [Demo Log in GI.cloud]

- gi-cloud.io
- Media Articles
  - Case Study: Delivering Real-Time Data, at 100,000 Samples per Second, in: Real time insight, Big data news platform in Silicon valley
  - „Über die Brücke in die Cloud“, Deutscher Mittelstand IT
  - Going open-source to handle IIoT real-time data <u>Link</u>



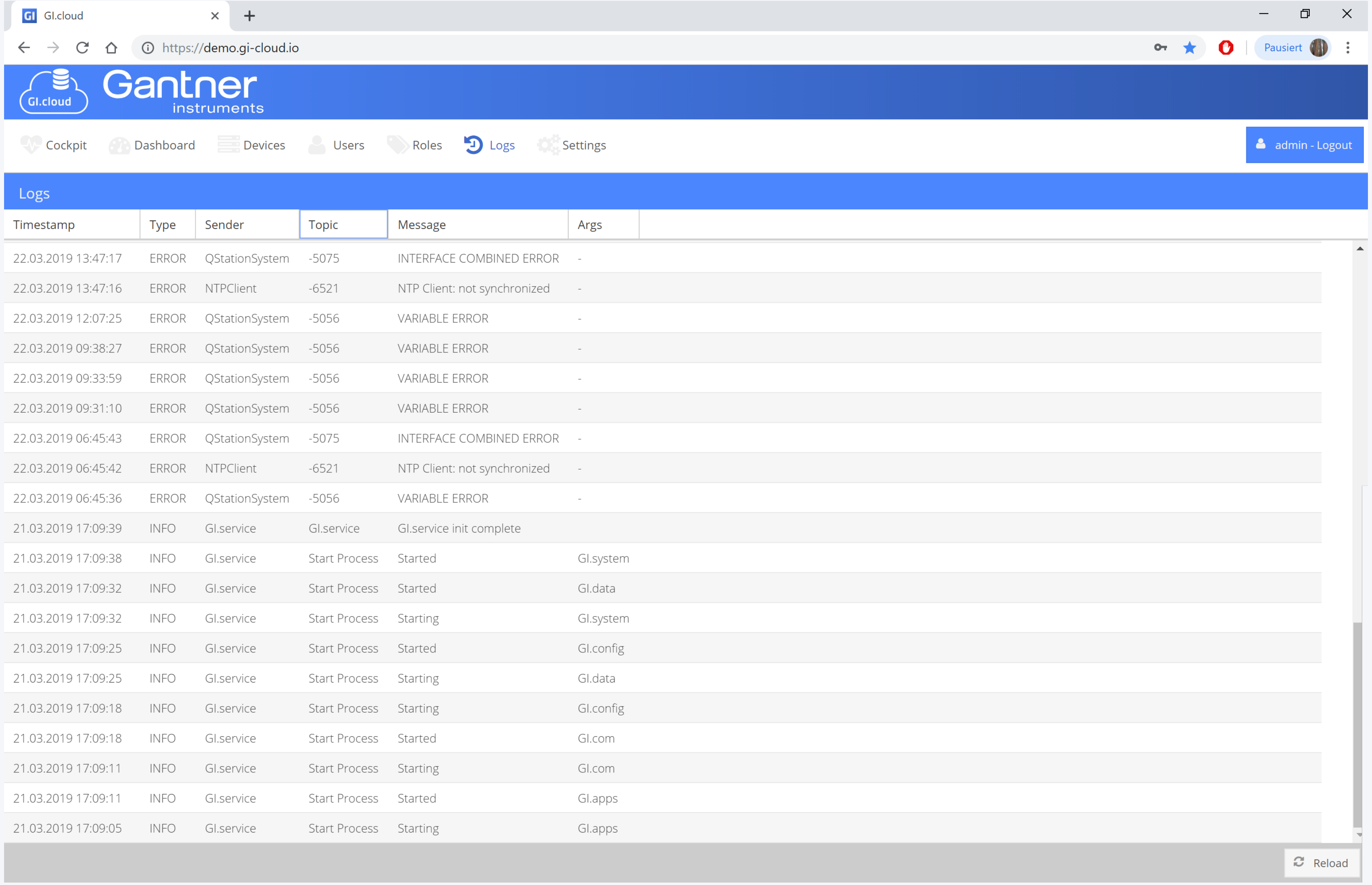
Device Management

- Devices & Meta data



Device Management

- Logs



Timestamp	Type	Sender	Topic	Message	Args
22.03.2019 13:47:17	ERROR	QStationSystem	-5075	INTERFACE COMBINED ERROR	-
22.03.2019 13:47:16	ERROR	NTPClient	-6521	NTP Client: not synchronized	-
22.03.2019 12:07:25	ERROR	QStationSystem	-5056	VARIABLE ERROR	-
22.03.2019 09:38:27	ERROR	QStationSystem	-5056	VARIABLE ERROR	-
22.03.2019 09:33:59	ERROR	QStationSystem	-5056	VARIABLE ERROR	-
22.03.2019 09:31:10	ERROR	QStationSystem	-5056	VARIABLE ERROR	-
22.03.2019 06:45:43	ERROR	QStationSystem	-5075	INTERFACE COMBINED ERROR	-
22.03.2019 06:45:42	ERROR	NTPClient	-6521	NTP Client: not synchronized	-
22.03.2019 06:45:36	ERROR	QStationSystem	-5056	VARIABLE ERROR	-
21.03.2019 17:09:39	INFO	GI.service	GI.service	GI.service init complete	
21.03.2019 17:09:38	INFO	GI.service	Start Process	Started	GI.system
21.03.2019 17:09:32	INFO	GI.service	Start Process	Started	GI.data
21.03.2019 17:09:32	INFO	GI.service	Start Process	Starting	GI.system
21.03.2019 17:09:25	INFO	GI.service	Start Process	Started	GI.config
21.03.2019 17:09:25	INFO	GI.service	Start Process	Starting	GI.data
21.03.2019 17:09:18	INFO	GI.service	Start Process	Starting	GI.config
21.03.2019 17:09:18	INFO	GI.service	Start Process	Started	GI.com
21.03.2019 17:09:11	INFO	GI.service	Start Process	Starting	GI.com
21.03.2019 17:09:11	INFO	GI.service	Start Process	Started	GI.apps
21.03.2019 17:09:05	INFO	GI.service	Start Process	Starting	GI.apps

Connectivity

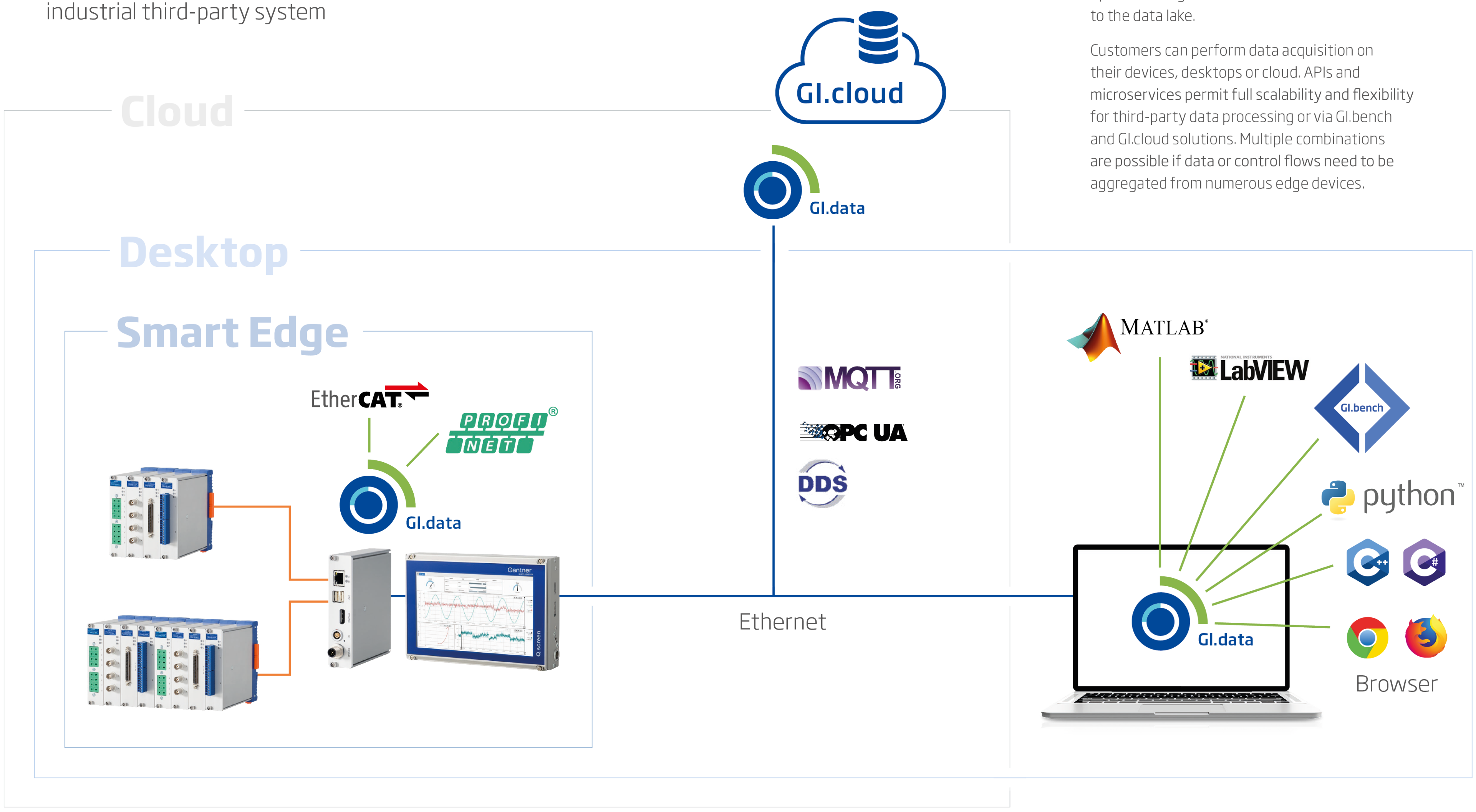
GI.connectivity

Access, store and handle your edge device data with high performance interfaces on GI proprietary platforms or any industrial third-party system

Run Your Data Acquisition on Your Devices, Desktop or Cloud

GI.connectivity fully integrates data storage, security, configuration, authentication and update management from the sensor interface to the data lake.

Customers can perform data acquisition on their devices, desktops or cloud. APIs and microservices permit full scalability and flexibility for third-party data processing or via GI.bench and GI.cloud solutions. Multiple combinations are possible if data or control flows need to be aggregated from numerous edge devices.

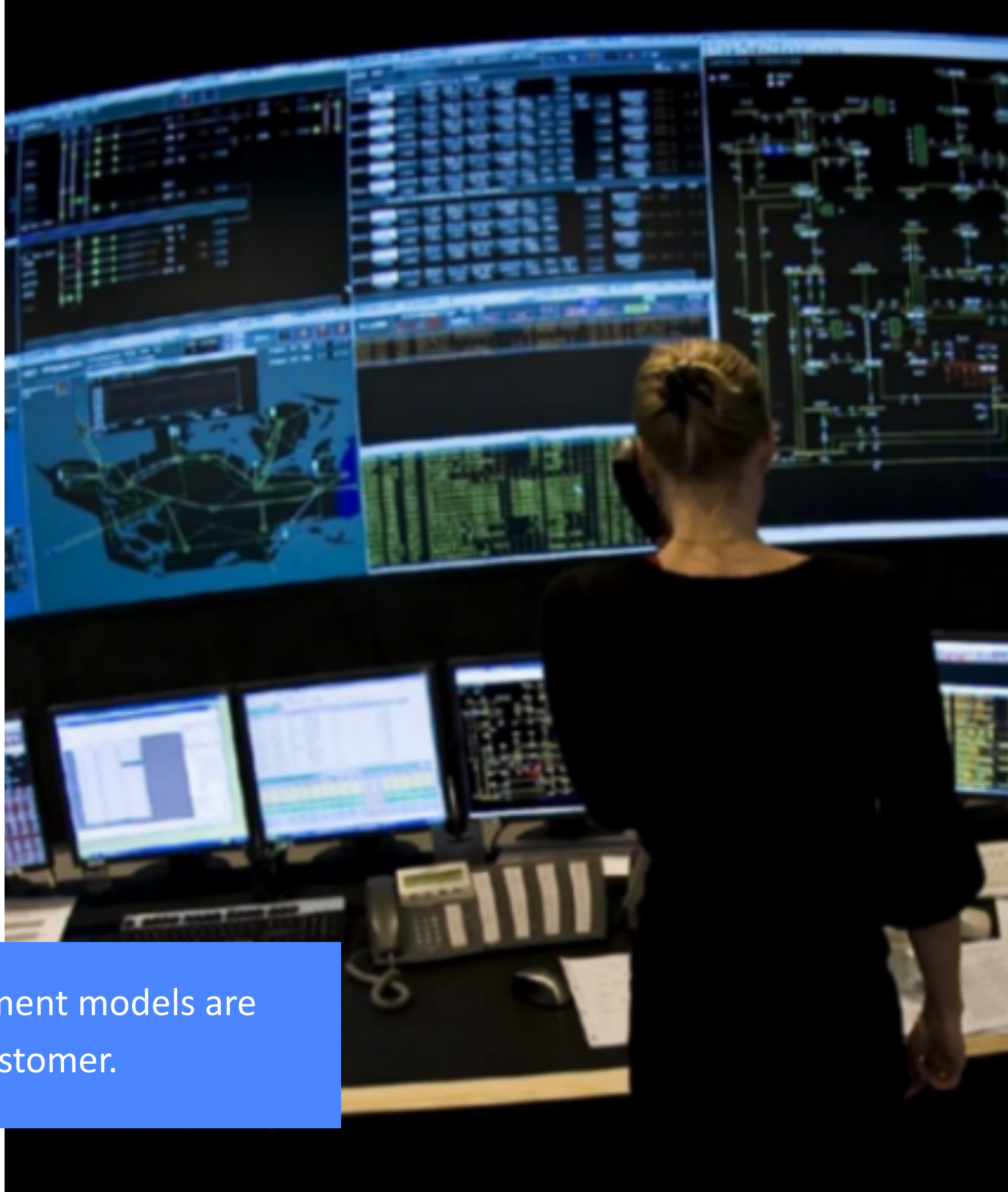


Edge Computing
Enablers & requirements

- Information and communications technology (ICT)
- Automation technology
- Signal processing
- Control systems
- Data processing & analysis

Requirements:

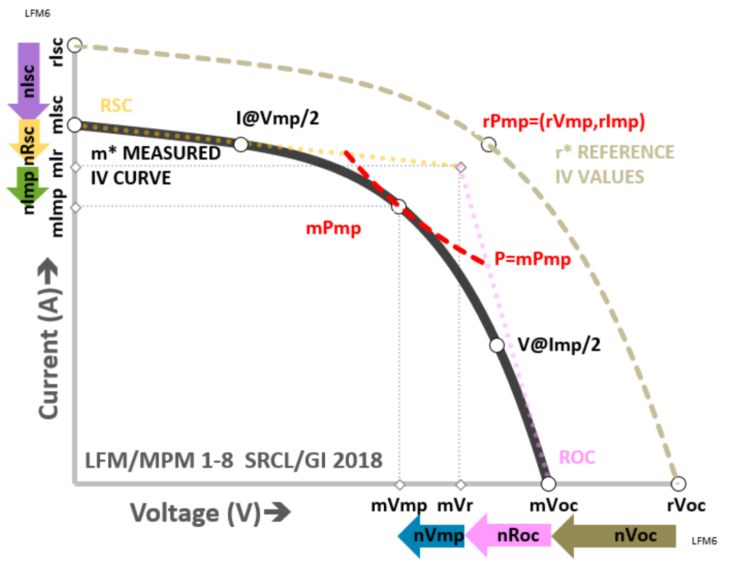
- Platforms: Distributed, Scalable, resilience, …
- Handle fast data, big data
- Open interfaces; Simple & Flexible



New Business and investment models are emerging closer to the customer.

The LFM : Defining loss parameters 1,2,4 (and 6 for full IV curve)

Gantner instruments



The LFM PR_{DC} definitions for 2, 4 and 6 parameters respectively are given in equations below (where $PR_{DC}x$ means the equation is for x parameters).

$$PR_{DC2} = [\quad nI_{DC} \quad] \times [\quad nV_{DC} \quad] \text{❷}$$
$$PR_{DC4} = [nI_{SC} \times \quad nFF_I \quad] \times [\quad nFF_V \quad \times nV_{OC}] \text{❹}$$
$$PR_{DC6} = [nI_{SC} \times nR_{SC} \times nI_{MP}] \times [nV_{MP} \times nR_{OC} \times nV_{OC}] \text{❻}$$

Figure 3 illustrates visually how these values all multiply to give the usual PR_{DC}.

(Note the 6+2 version has two extra parameters to quantify the curvature from cell mismatch or rollover.)

Optional corrections : Thermal, Soiling, AOI/Beam Fraction, Spectral

❶	P_{MP} or PR_{DC} or Eff						**Array**
❷	nI_{DC}			nV_{DC}			No IV
❹	nI_{SC}	nFF_I		nFF_V		nV_{OC}	
❻	nI_{SC}	nR_{SC}	nI_{MP}	nV_{MP}	nR_{OC}	nV_{OC}	IV

Module

❻-❷	nIc = Mismatch	nVc = Rollover