Precipitation Estimation and Assimilation

G.J. Huffman

NASA/GSFC Laboratory for Atmospheres and
Science Systems and Applications, Inc.

with R.F. Adler, D.T. Bolvin, E.J. Nelkin

Outline

- Introduction
- Microwave-based data sets
- IR-based data sets
- Microwave-IR combinations
- Issues

1. INTRODUCTION – Rain is easy to measure, hard to analyze

The physical process is hard to cope with:

- precipitation is intermittent and not “normal”
- rain is generated on the microscale
- the decorrelation distance/time is short
- point values only represent a small area & snapshots only represent a short time
- a finite number of samples causes problems

Each sensor has characteristic errors

QuickTime™ and a
Animation decompressor
are needed to see this picture.

1. INTRODUCTION – Our Group's Long-term Objective

Find the best global precipitation fields, given several independent estimates of (spatially) varying quality.

Requirements

- "Long-term" "global" precip estimates
- Minimal bias
- Minimal random error

"State-of-the-Art" Combination Approach

- Calibrate microwave data to TRMM data
- Calibrate IR T_b to microwave rainrate
- Merge data after calibration step(s)
- Designed to emphasize strengths and minimize weaknesses of individual data sets

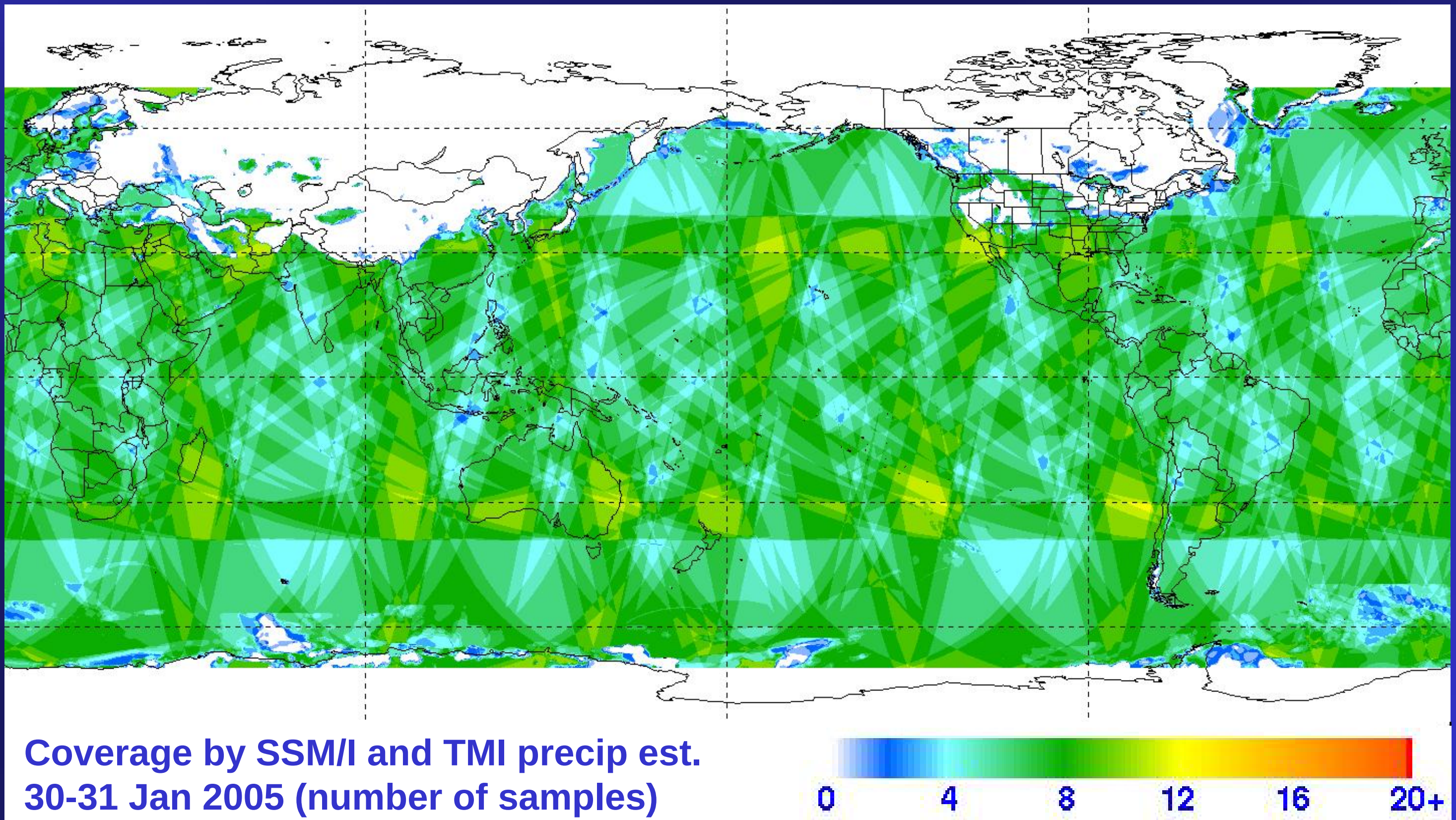
2. MICROWAVE-BASED DATA SETS

Features

- Radiances have the best physical relation to precipitation
- Over-water estimates can use all channels
- Over-land and -coast limited to “scattering” channels – more approximate
- All algorithms have a low-end detectability problem

⇒ All the “classic” channels are knocked out by surface snow/ice/frost

⇒ Only available on low-Earth orbit satellites – sampling issue



Coverage by SSM/I and TMI precip est.
30-31 Jan 2005 (number of samples)

2. MICROWAVE-BASED DATA SETS (cont.)

Specific sensors

- TMI, AMSR-E have the best resolution, most channels
- SSM/I nearly as good – resolution is lower; it has 2 fewer channels
- AMSU-B has lower and variable resolution, fewer channels, larger detectability issue

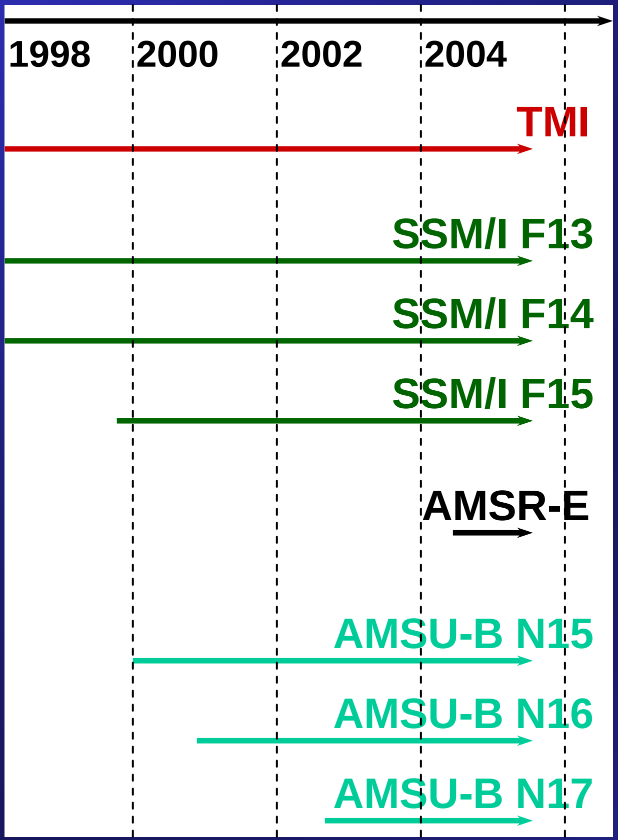
Combined inter-calibrated microwave estimates

- Collecting all the (asynoptic) microwave estimates into 3-hourly maps is the current popular method for achieving quasi-global microwave coverage
- All such collections inter-calibrate the estimates to a standard
  - typically done by histogram matching
  - detectability is accentuated for estimates with low occurrence of precip
  - our group refers to the microwave combination as HQ
- This will continue to be a key concept for GPM

2. MICROWAVE-BASED DATA SETS (cont.)

The satellites record data continuously, and usually not with optimal orbit spacing

A changing complement of satellites modulates the coverage and quality of estimates



QuickTime™ and a
Cinepak decompressor
are needed to see this picture.

3. IR-BASED DATA SETS

Features

- Radiances have the best time/space resolution
- Radiances give cloud-top information, not direct hydrometeor data
- All IR estimates are calibrations – can't do IR rain from first principles
- GPI is a static calibration
- Microwave-calibrated IR
  - histogram matching: VAR, NRL Real-Time
  - spatial structure: CST
  - AI-based (with cloud classification?): PERSIANN

Issues

- Optimum calibration interval?
- A calibrated estimate can't do more than perfectly estimate the thing providing calibration
  - pixel-level IR estimates resulting from 0.25°x0.25° microwave estimates are best thought of as giving 0.25° estimates, even though they're 4 km apart

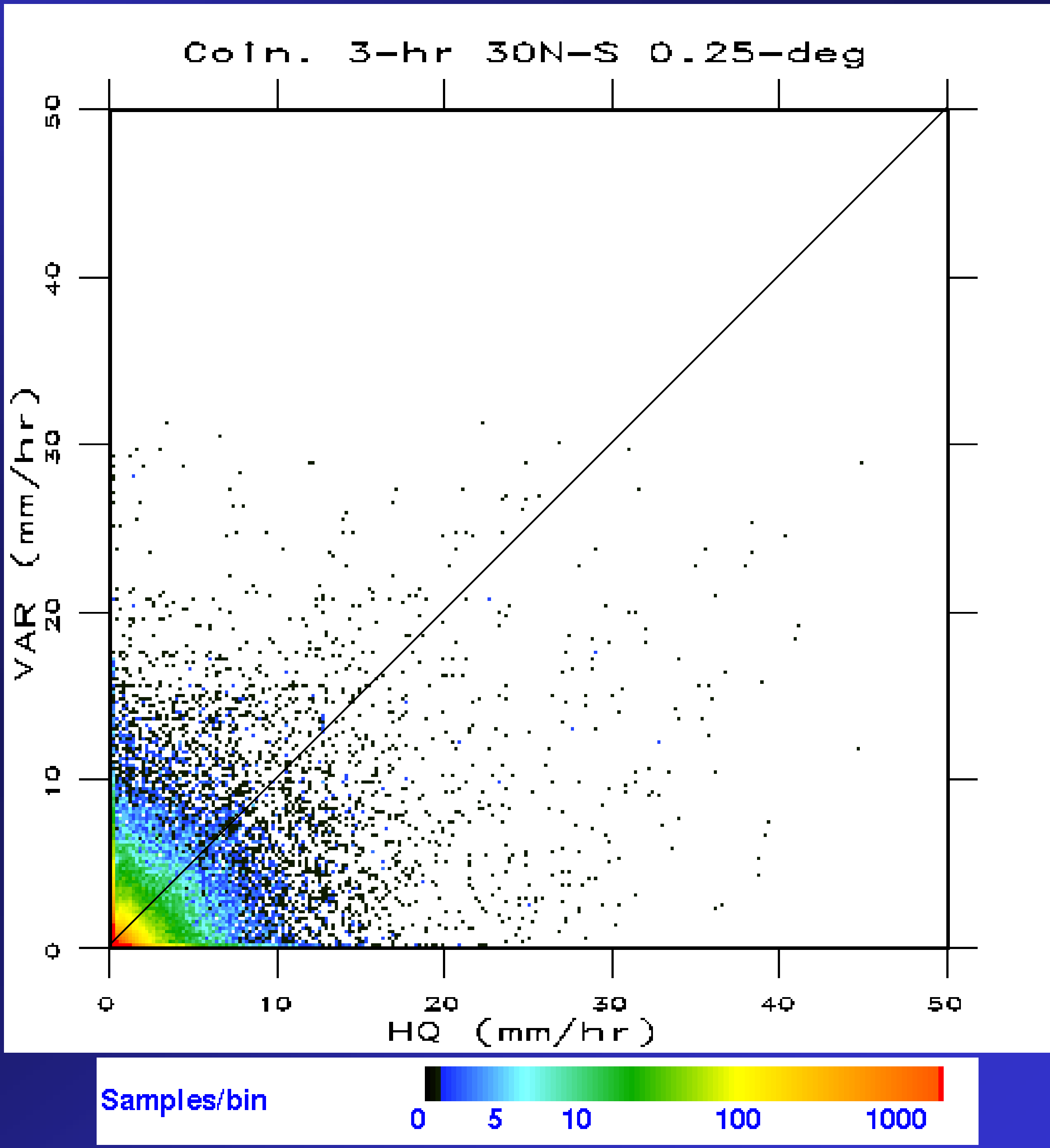
3. IR-BASED DATA SETS – Scatter Between "Coincident" HQ and VAR

3-hr, 0.25°x0.25°

00, 03, …, 21Z 15 Feb 2002

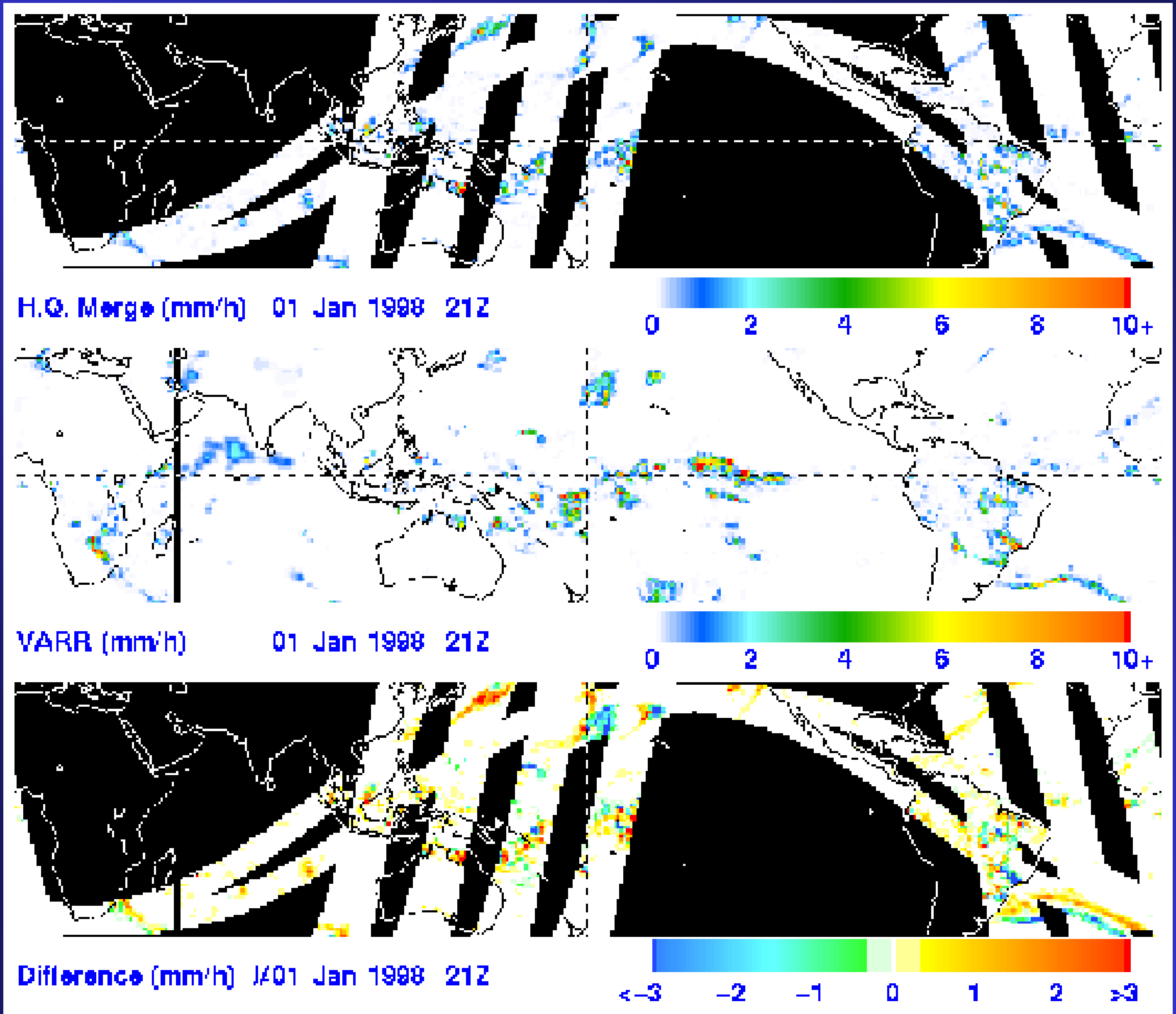
Latitude band 30°N-S

The match-up is bad for 2 reasons



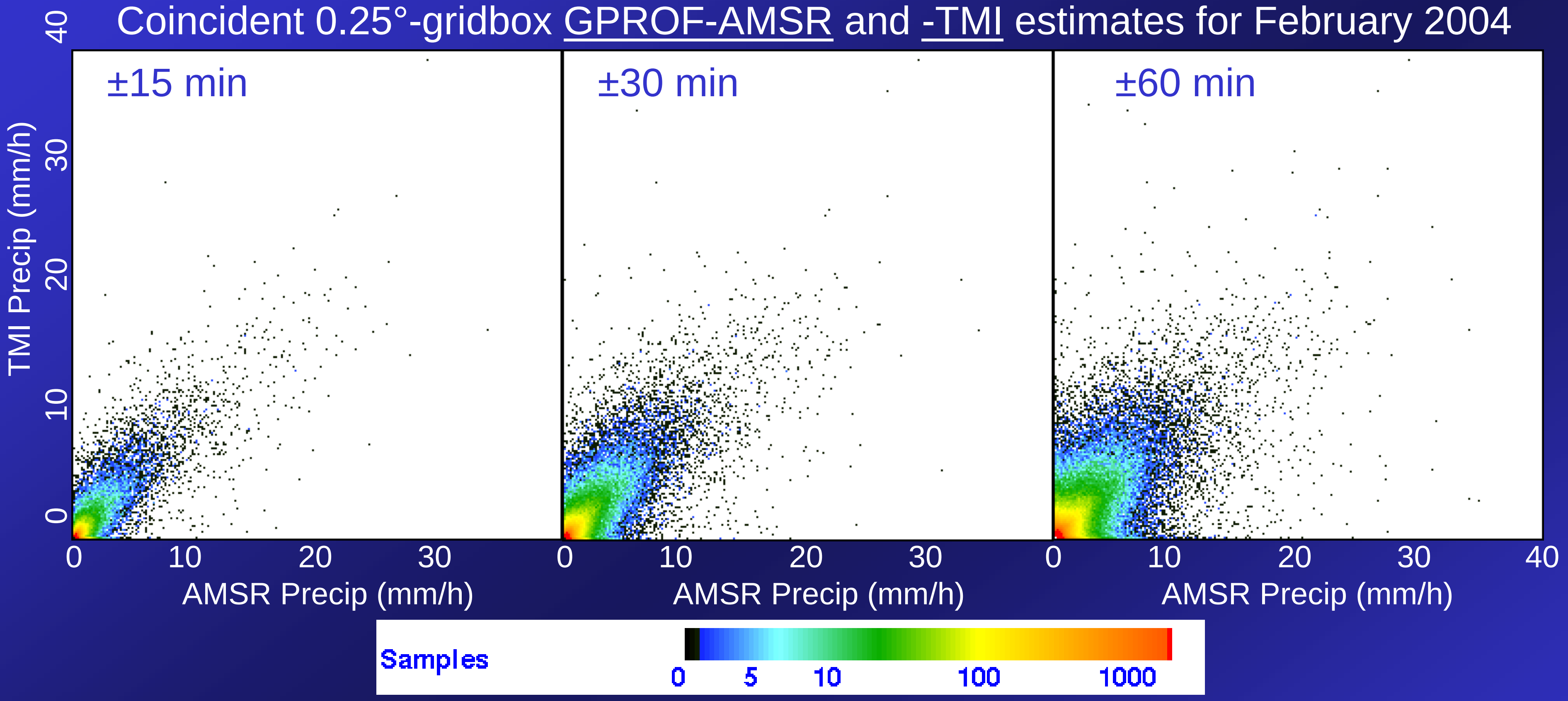
3. IR-BASED DATA SETS – Scatter Between "Coincident" HQ and VAR (cont.)

There are systematic <u>offsets</u> between what the microwave and IR sensors "see"



3. IR-BASED DATA SETS – Scatter Between Coincident HQ and VAR (cont.)

Coincident 0.25°-gridbox GPROF-AMSR and -TMI estimates for February 2004

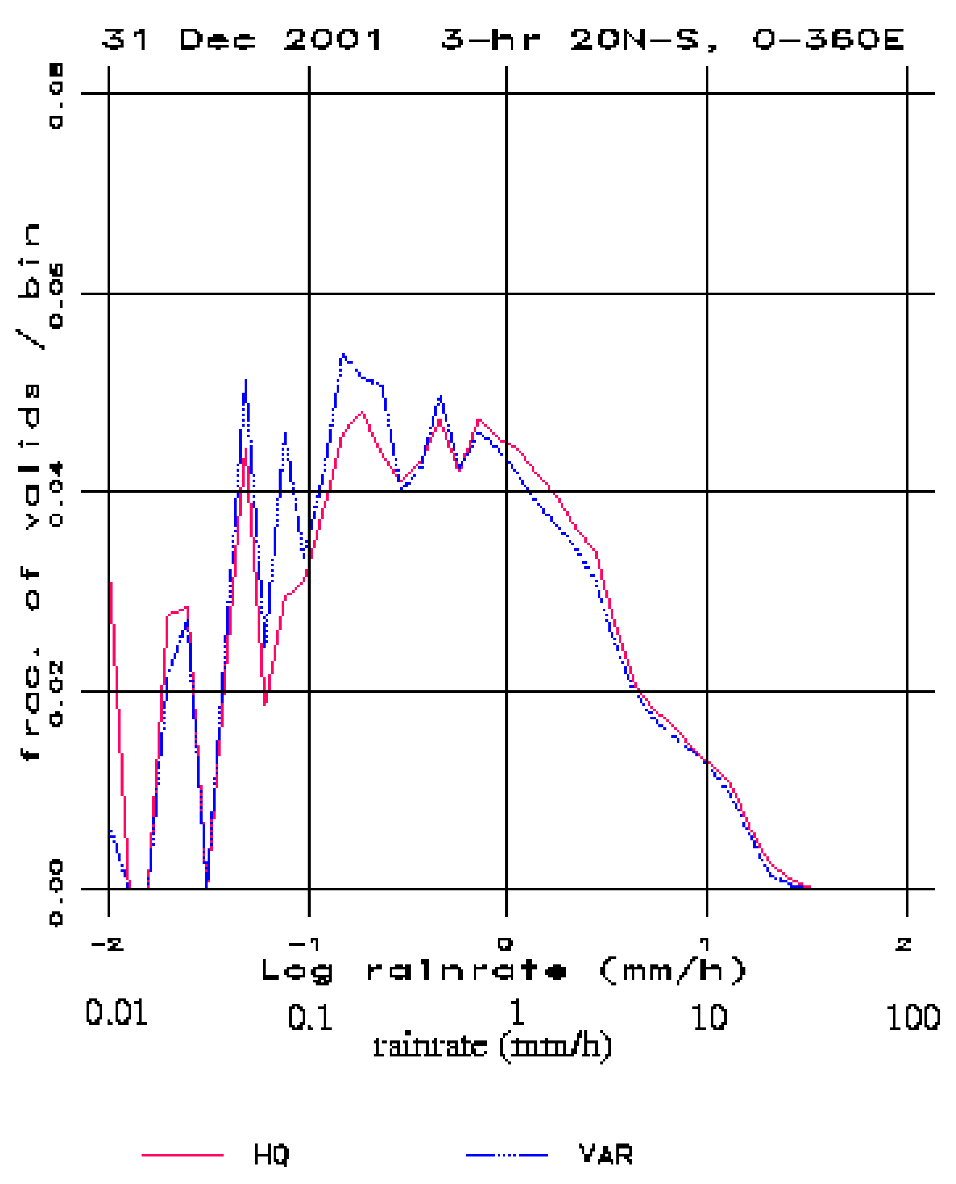


The "standard" 3-hr time window for microwave data introduces error

- same grid box for spatial coincidence
- ±15-, ±30-, ±60-minute windows of time coincidence
- points near axes at ±60 result from advection into/out of box, and/or growth/decay
- limiting the window decreases the microwave sampling in each period

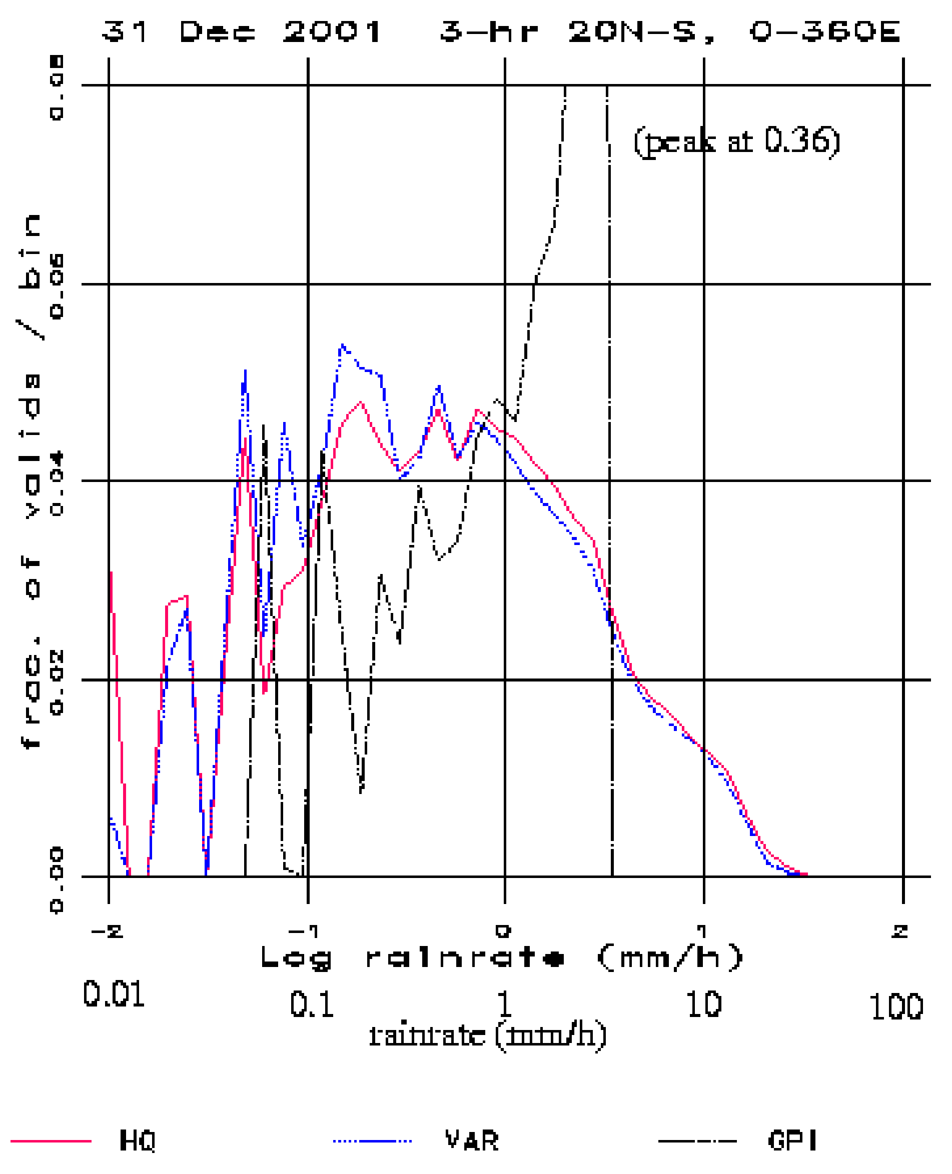
3. IR-BASED DATA SETS –Scatter Between Coincident HQ and VAR (cont.)

Around the tropics VAR's rainrate histogram provides a good match to the calibrating microwave ("HQ")



3. IR-BASED DATA SETS –Scatter Between Coincident HQ and VAR (cont.)

Around the tropics VAR's rainrate histogram provides a good match to the calibrating microwave ("HQ"), <u>better than GPI</u>



4. MICROWAVE-IR COMBINATIONS

Features of the Multi-satellite Precipitation Analysis (MPA)

- Use both microwave and IR estimates
- At present, the “combination” scheme is a selection – HQ where available and VAR elsewhere
- Provides best local instantaneous estimate

Issues

- Statistics are inhomogeneous
  - features can “jump” between data sets due to changes in data source
  - statistical summaries can vary depending on the relative contributions of each data source

Products

- The Real Time MPA (MPA-RT) is produced 7 hours (and soon also 4) after nominal observation time
  - microwave is calibrated to GPROF-TMI
  - HQ, VAR, and HQ/VAR are provided as 3B40RT, 3B41RT, 3B42RT
- A research product is produced a few days after the end of the month
  - microwave is calibrated to the TRMM TMI/PR combined product
  - the month of HQ/VAR is summed and combined with gauge to produce the TRMM 3B43 satellite-gauge monthly product
  - the individual HQ/VARs are scaled to sum to 3B43 and provided as the TRMM 3B42 multi-satellite 3-hourly product

4. MICROWAVE-IR COMBINATIONS (cont.)

Features of the CPC Morphing algorithm (CMORPH)

- Compute advection velocities from pairs of half-hourly IR images
- Use the IR-based velocities to move microwave estimates back and forth in time
- Do a linear "fade" between successive microwave images of "the same" system following the system

Issues

- The IR velocities are typically too fast
- The linear fade typically produces rain fields that are too wide-spread and too weak (but the integrated rain volume might be close)
- CMORPH starts losing to microwave-calibrated IR when the time-interpolated values are more than 2 hours away from the nearest microwave overpass

Products

- CMORPH is produced some 18 hours after nominal observation time
  - need to wait for following microwave overpasses
  - working on improved advection velocity, IR fill-in for missing microwave

5. ISSUES

Now and into GPM, microwave satellites are widely and irregularly spaced

- But what IR approaches yield useful data for assimilation systems?

Light and snowy precip detection is problematic

- For microwave as noted above
  - greatly improved with sounding channels (150, 183 GHz)
- Therefore for calibrated IR
- For TRMM PR
  - greatly improved for GPM Dual-frequency Precip Radar (DPR)
- TOVS and AIRS retrievals show promise
  - how much model support for retrievals will assimilation scientists tolerate?

Random error estimates

- The successful monthly methodology can't be used
- Faisal Hossain and Manos Anagnostou (U. Conn.) SREM2D shows early promise
  - incorporates time/space scale dependence
  - sensitivity to climate regime unknown
- How much detail do assimilation systems need in the error estimate?

5. ISSUES (cont.)

Bias error estimates

- No operational global estimator exists for bias
- Tom Smith (NCDC) is making a reasonable attempt
- Wes Berg (CSU) is working toward regime-dependent bias in specific areas for specific instruments

Is there a place for model estimates?

- There is some evidence that models out-perform current observation-based estimates in cold-season mid-latitude cyclonic systems

Data "archaeology"

- Essentially the entire global archive of geo-satellite data was rescued in the late 1990's, except the Japanese data before 1987 were discarded; also the ISCCP B1 archive covers the full period with a 3-hrly 10-km subsample
- NCDC also holds a more or less complete set from the NOAA leo orbiters
- Step 1: NCDC has beta tests of the necessary navigation/format conversion software to process the B1 for IR data
- These data can stretch the detailed merged global IR record from 2000 back to 1983!
- Other channels could be similarly processed, as available
- As above, we need creative algorithms for IR (or other old geo channels)

5. ISSUES – Condensed List of Discussion Points

Importance of the light precip?
Importance of winter land estimates?
Required/desired time/space interval?
Importance of long-term retrospective analyses?
Format/content of error estimates at the fine scale?
Tolerance for heterogeneous statistics?
Tolerance for model assistance in retrievals?
Place of model estimates?

QuickTime™ and a
Animation decompressor
are needed to see this picture.

NASA/GSFC Code 613.1 Precipitation Web Site: http://precip.gsfc.nasa.gov
Contact: huffman@agnes.gsfc.nasa.gov

1. INTRODUCTION – Rain is easy to measure, hard to analyze (cont.)

Instruments have characteristic errors:

- raingauge
  wind losses
  splashing
  evaporation
  side-wetting
- radar
  raindrop population changes
  anomalous propagation
  beam blockage by surface features
  sidelobes
- satellite
  physical retrieval errors
  beam-filling errors
  time-sampling
- numerical prediction models
  computational approximations
  initialization errors
  errors in other parts of the computation

Tropical Rainfall Measuring Mission (TRMM)

TRMM Sensors

Precipitation radar (PR)

13.8 GHz

4.3 km footprint

0.25 km vertical res.

215 km swath

Microwave radiometer (TMI)

10.7, 19.3, 21.3, 37.0

85.5 GHz (dual polarized

except for 21.3 V-only)

10x7 km FOV at 37 GHz

760 km swath

Visible/infrared radiometer (VIRS)

0.63, 1.61, 3.75, 10.8, and 12 m

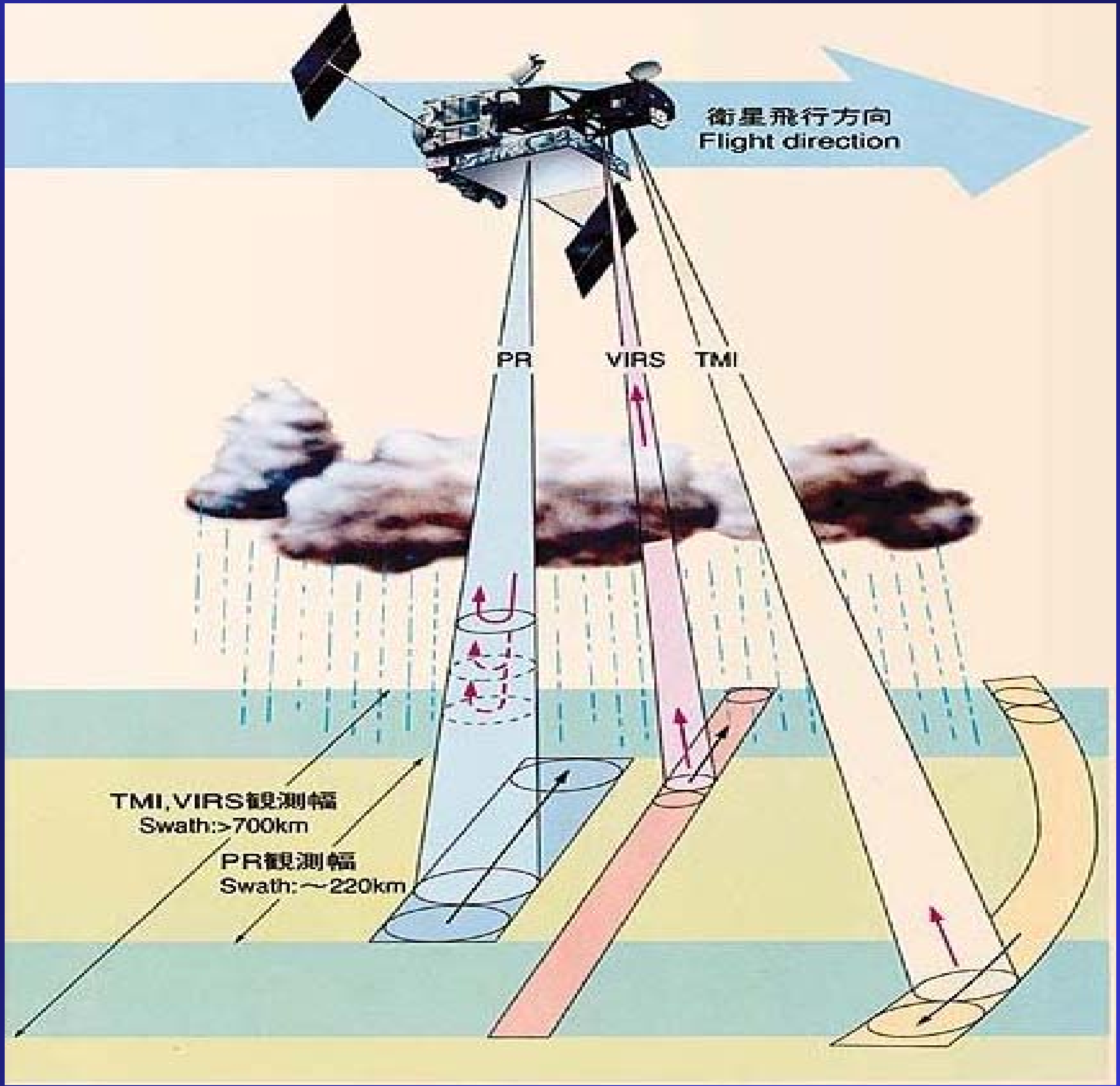
at 2.2 km resolution

Lightning Imaging Sensor (LIS)

Cloud & Earth Radiant

Energy System (CERES)

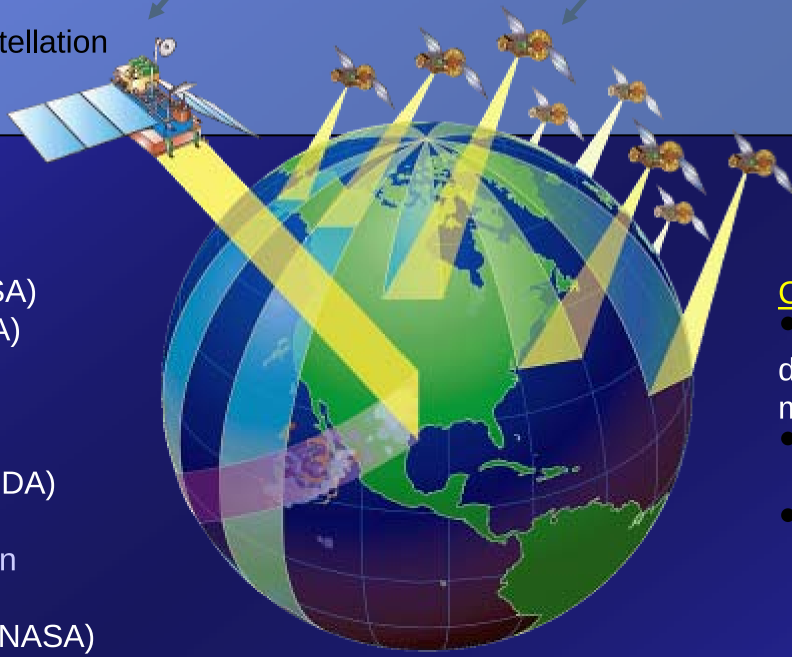
Nov. 1997 launch, 35° inclination; 402 km



GPM Concept

<u>OBJECTIVES</u>
- Understand horizontal & vertical structure of rainfall, its macro- & micro-physical nature, & its associated latent heating
- Transfer Standard for constellation radiometers

<u>OBJECTIVES</u>
- Provide sufficient global sampling to significantly reduce uncertainties in short-term rainfall accumulations
- Extend scientific and societal applications



<u>Core Satellite</u>
- TRMM-like spacecraft (NASA)
- H2-A rocket launch (NASDA)
- Non-sun-synchronous orbit
  ~ 65° inclination
  ~400 km altitude
- Dual frequency radar (NASDA)
  Ku-Ka Bands (14-35 GHz)
  ~ 5 km horizontal resolution
  ~250 m vertical resolution
- Multifrequency radiometer (NASA)
  10, 19, 23, 36, 89, (150/183 ?) GHz V&H

<u>Constellation Satellites</u>
- Pre-existing operational-experimental & dedicated satellites with microwave / millimeter-wave radiometers
- Revisit time
  3-hour goal at ~90% of time
- Sun-synch & non-sun- synch orbits
  600-900 km altitudes

<u>Precipitation Processing System</u>
- Produces global precipitation products
- Products defined by GPM partners

<u>Precipitation Validation Sites for Error Characterization</u>
- Select/globally distributed ground validation “Supersites” (research quality radar, up looking radiometer-radar-profiler system, raingage-disdrometer network, & T-q soundings)
- Dense & frequently reporting regional raingage networks

MPA Algorithm

"Best" in HQ will be TRMM PR (currently TMI)

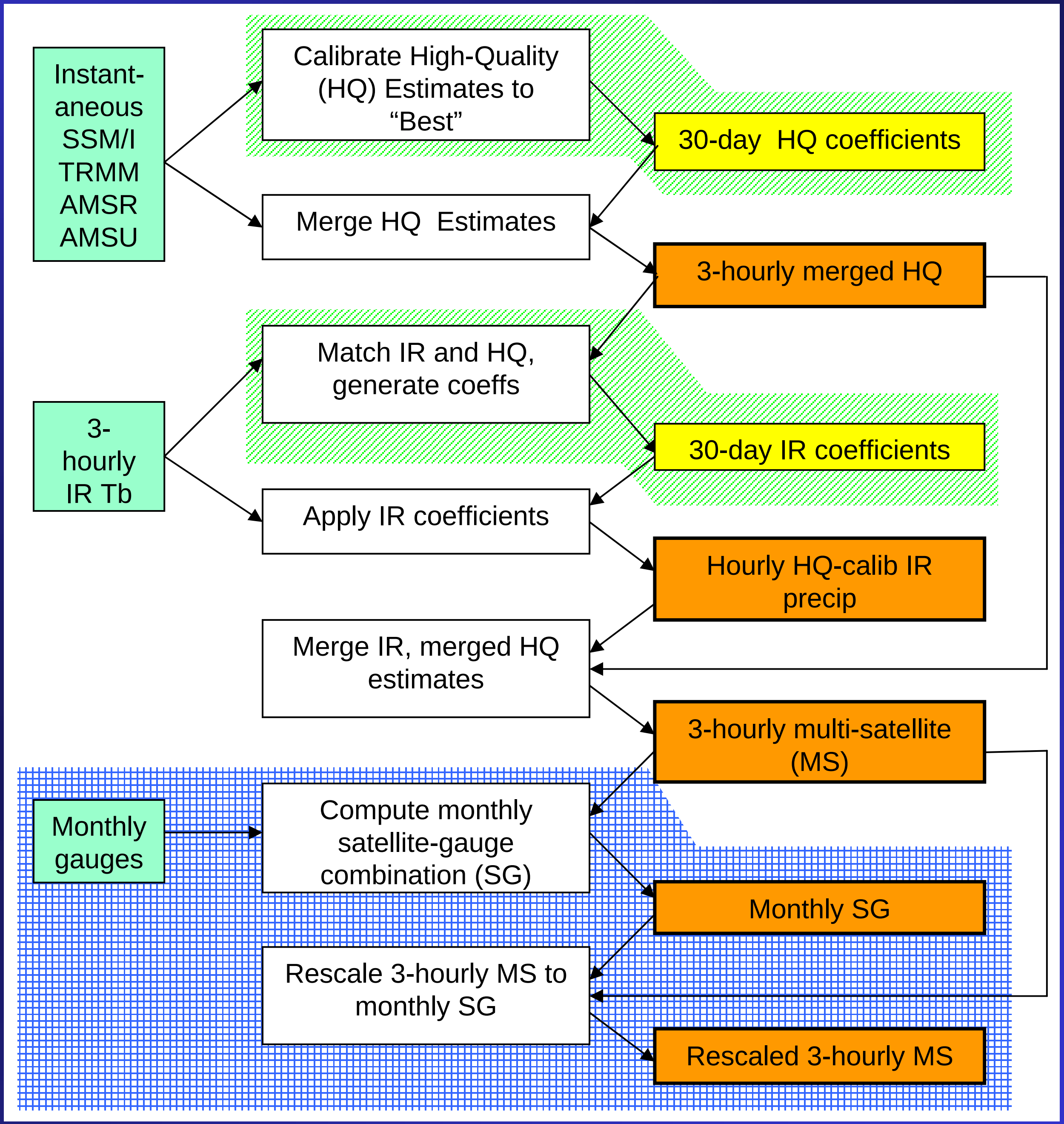
Green shading done async in real time, trailing avg.

Blue shading only done non-real time, adds value

Cyan boxes are inputs

Yellow boxes are calibration coefficients

Orange boxes are products



4. DETAILS – TRMM-BASED MPA APPROACH

Variable Rain Rate (VAR)

- IR product with spatially varying T_b threshold and variable rain rate determined locally for each month
- Rain rate depends on T_b departure below threshold
- Calibrate SSM/I by TMI and merge to form High-Quality (HQ)
- Use HQ to calibrate VAR
- T_b – HQ calibration done on gridbox-average values, not pixels

Combined 3-Hourly Estimate (VAR+HQ)

- Fill holes in HQ with VAR

Issues

- Simple replacement is not the optimal combination scheme
- Differences in datasets introduce unphysical boundaries
- Biases in HQ and VAR are not necessarily consistent

At full resolution the correlation of rain between VAR and HQ is poor; averaging quickly improves the picture

Notice: cover eyes of young children and impressionable students before showing next slide

Effect of Averaging on Scatter Between Coincident HQ and VAR

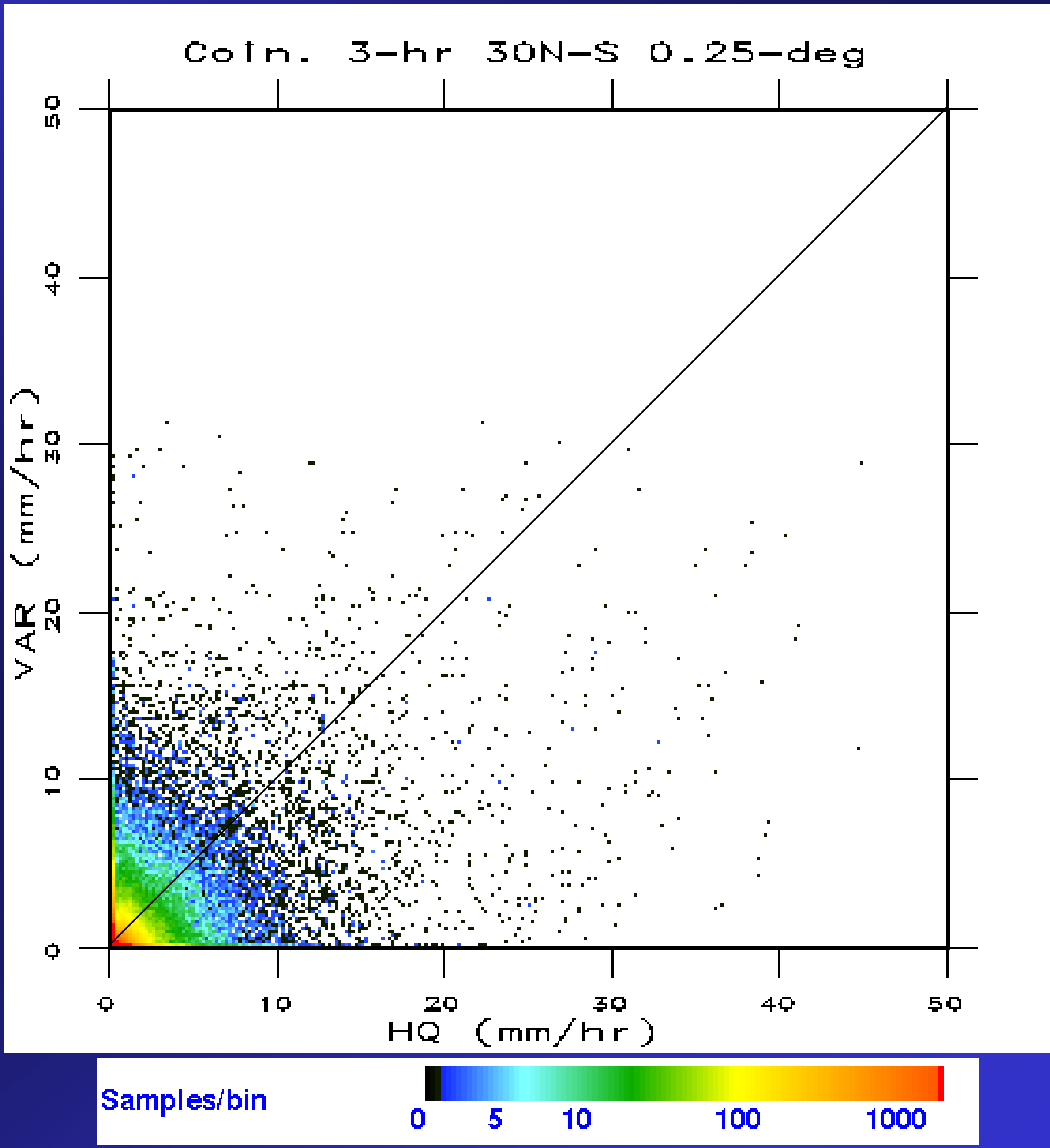
3-hr

0.25°x0.25°

00, 03, …, 21Z 15 Feb 2002

Latitude band 30°N-S

Original data; the next 3 slides show progressively longer time averages



Effect of Averaging on Scatter Between Coincident HQ and VAR

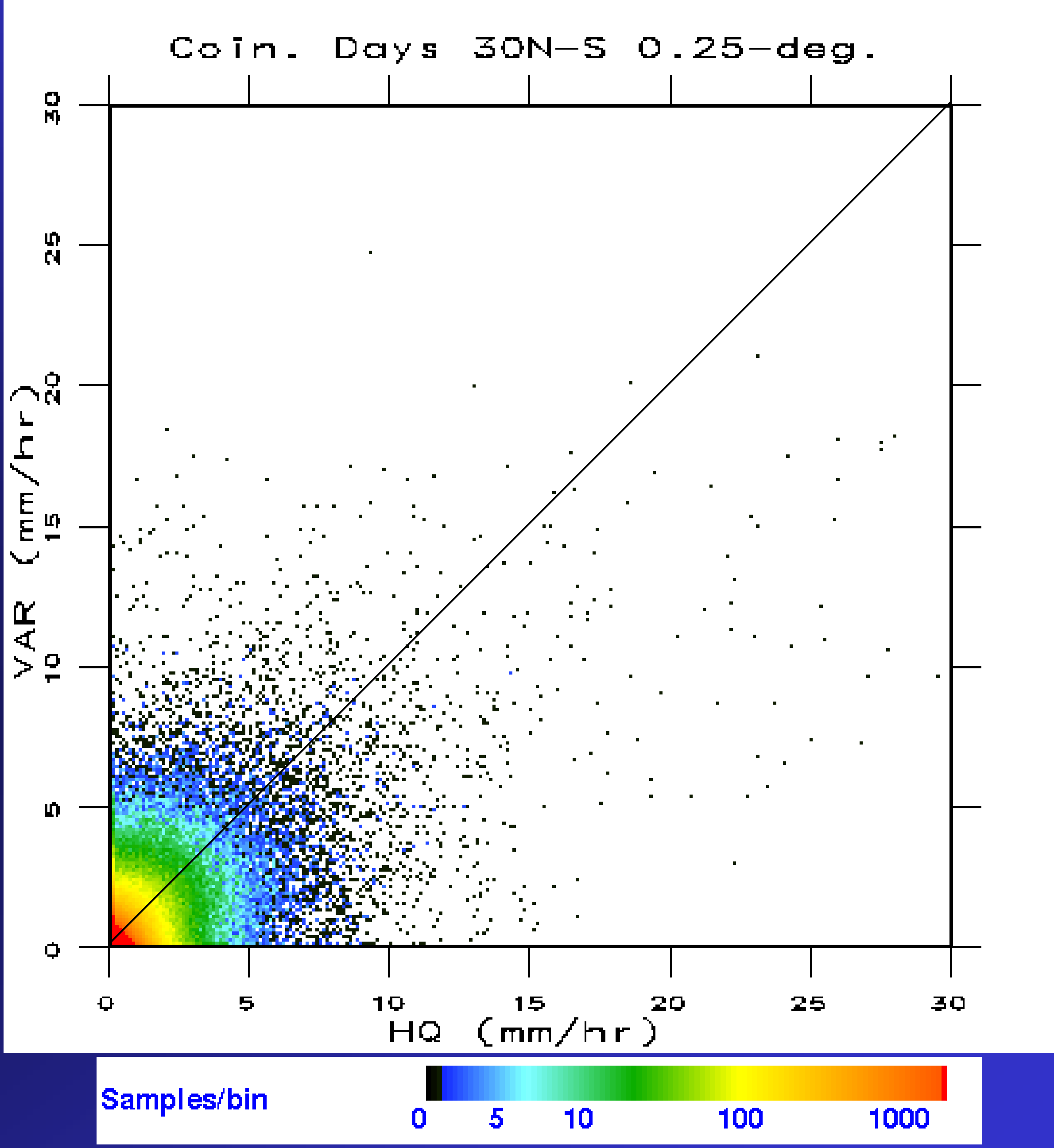
Days

0.25°x0.25°

11, 12, 13, 14, 15 Feb 2002

Latitude band 30°N-S

The “day” has 3 samples for many points



Effect of Averaging on Scatter Between Coincident HQ and VAR

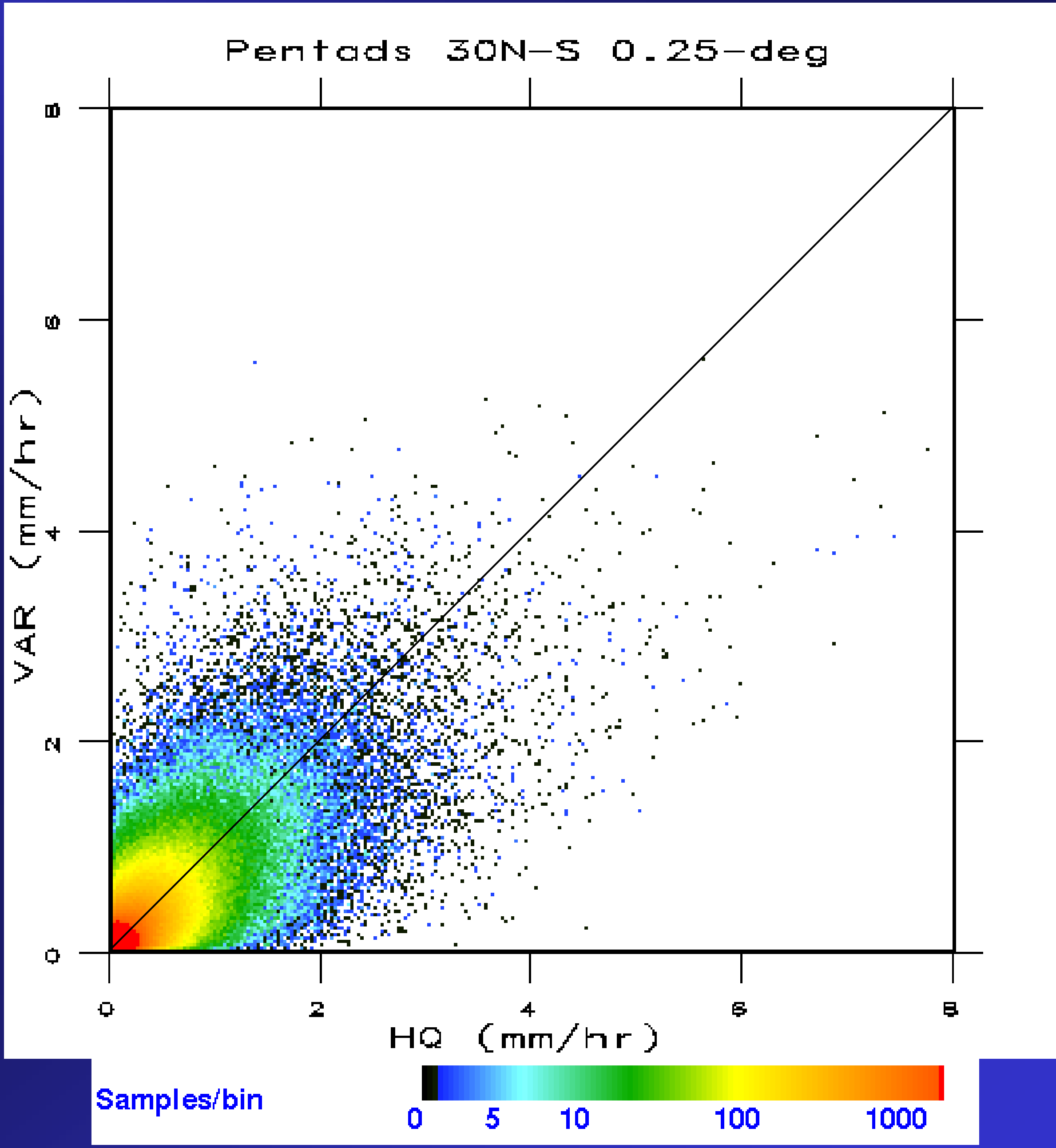
Pentads

0.25°x0.25°

1-5, 6-10, 11-15, 16-20, 21-25 Feb 2002

Latitude band 30°N-S

Note improved behavior along axes



Effect of Averaging on Scatter Between Coincident HQ and VAR

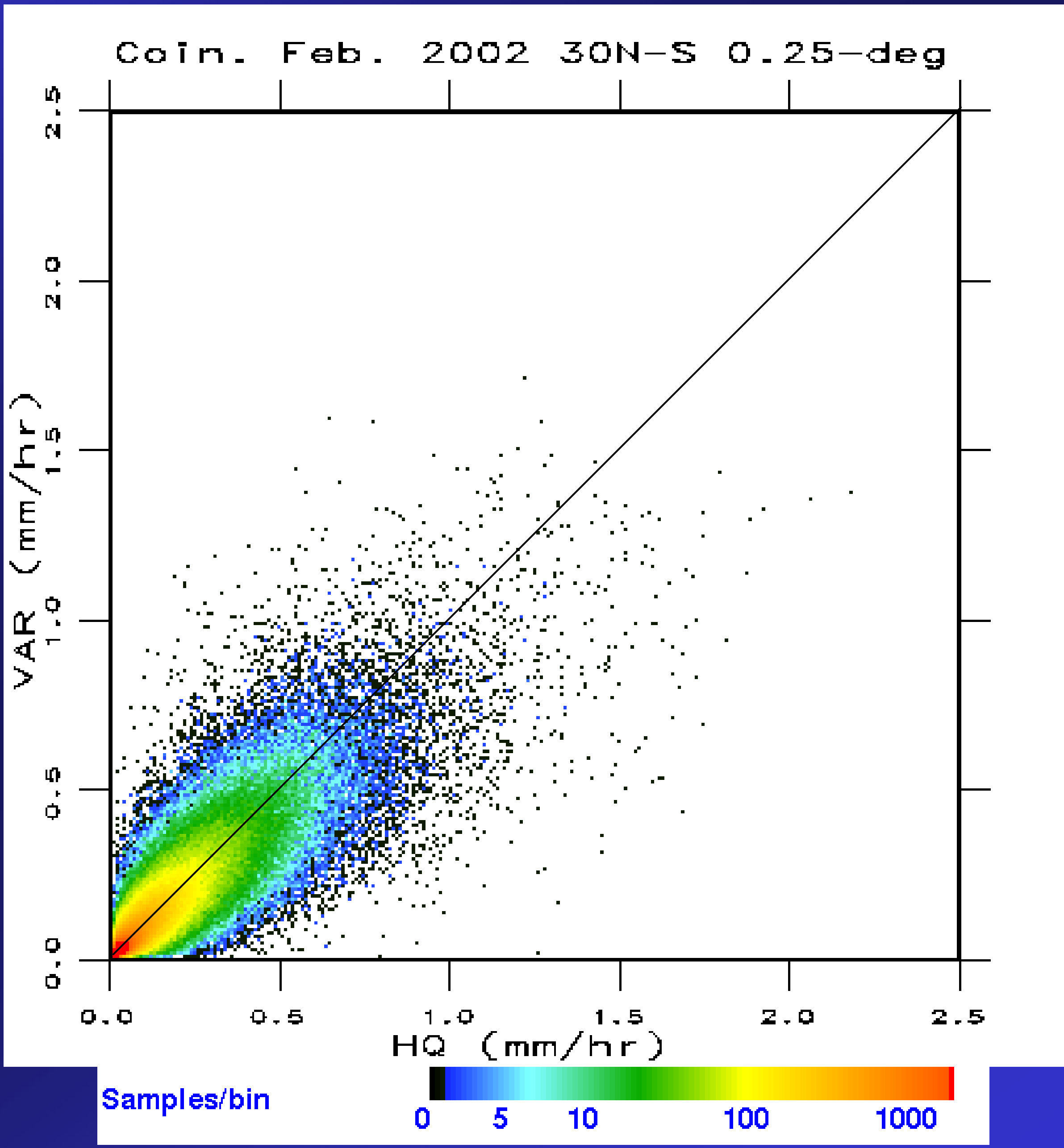
Month

0.25°x0.25°

Feb 2002

Latitude band 30°N-S

Further time-averaging continues to improve the relationship



Effect of Averaging on Scatter Between Coincident HQ and VAR

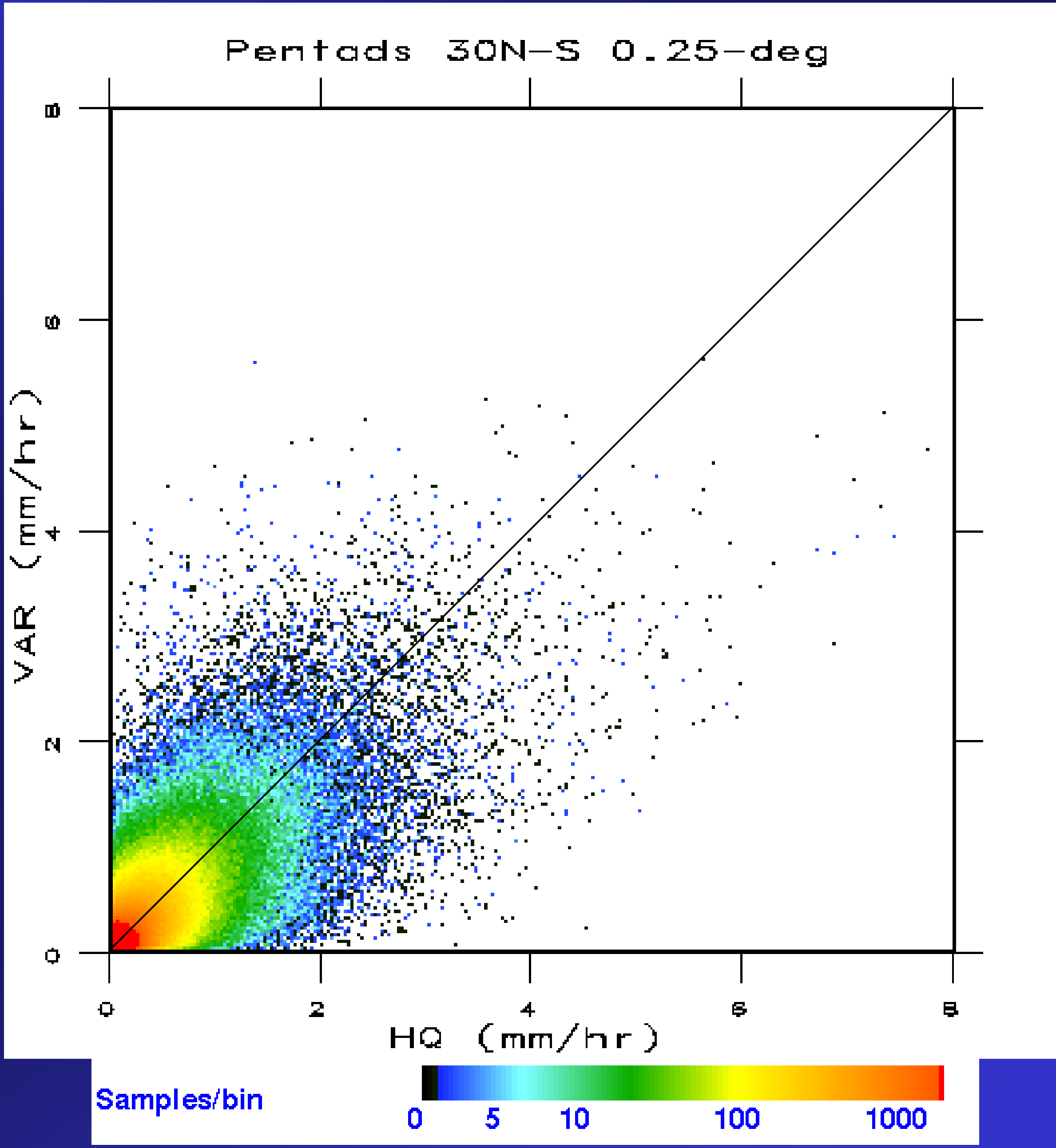
Pentads

0.25°x0.25°

1-5, 6-10, 11-15, 16-20, 21-25 Feb 2002

Latitude band 30°N-S

Now we show the effect of spatial averaging on the pentad time-averages



Effect of Averaging on Scatter Between Coincident HQ and VAR

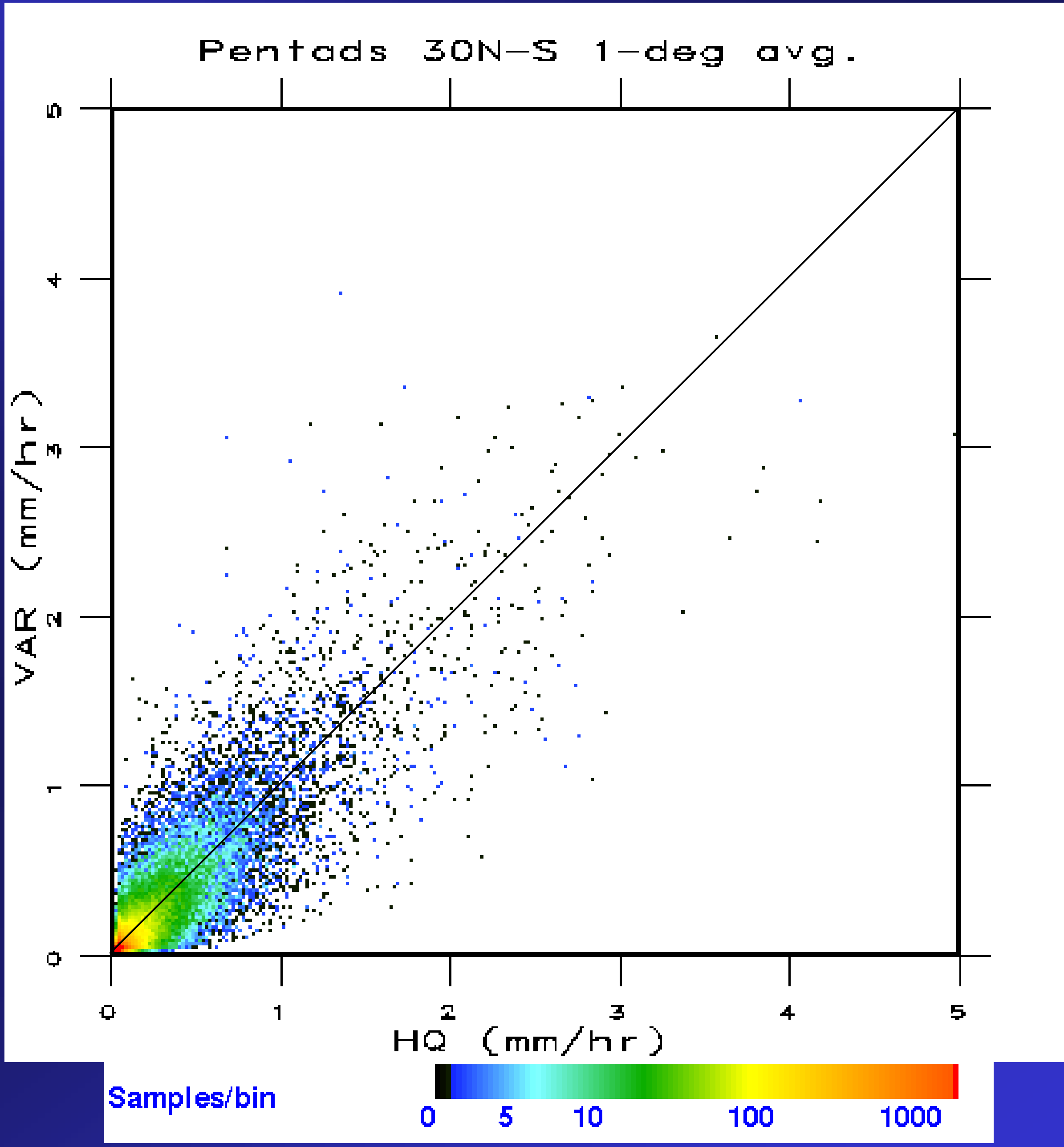
Pentads

1°x1°

1-5, 6-10, 11-15, 16-20, 21-25 Feb 2002

Latitude band 30°N-S

The move to 1°x1° reduces scatter, as well as reducing the number of grid boxes by a factor of 16



Effect of Averaging on Scatter Between Coincident HQ and VAR

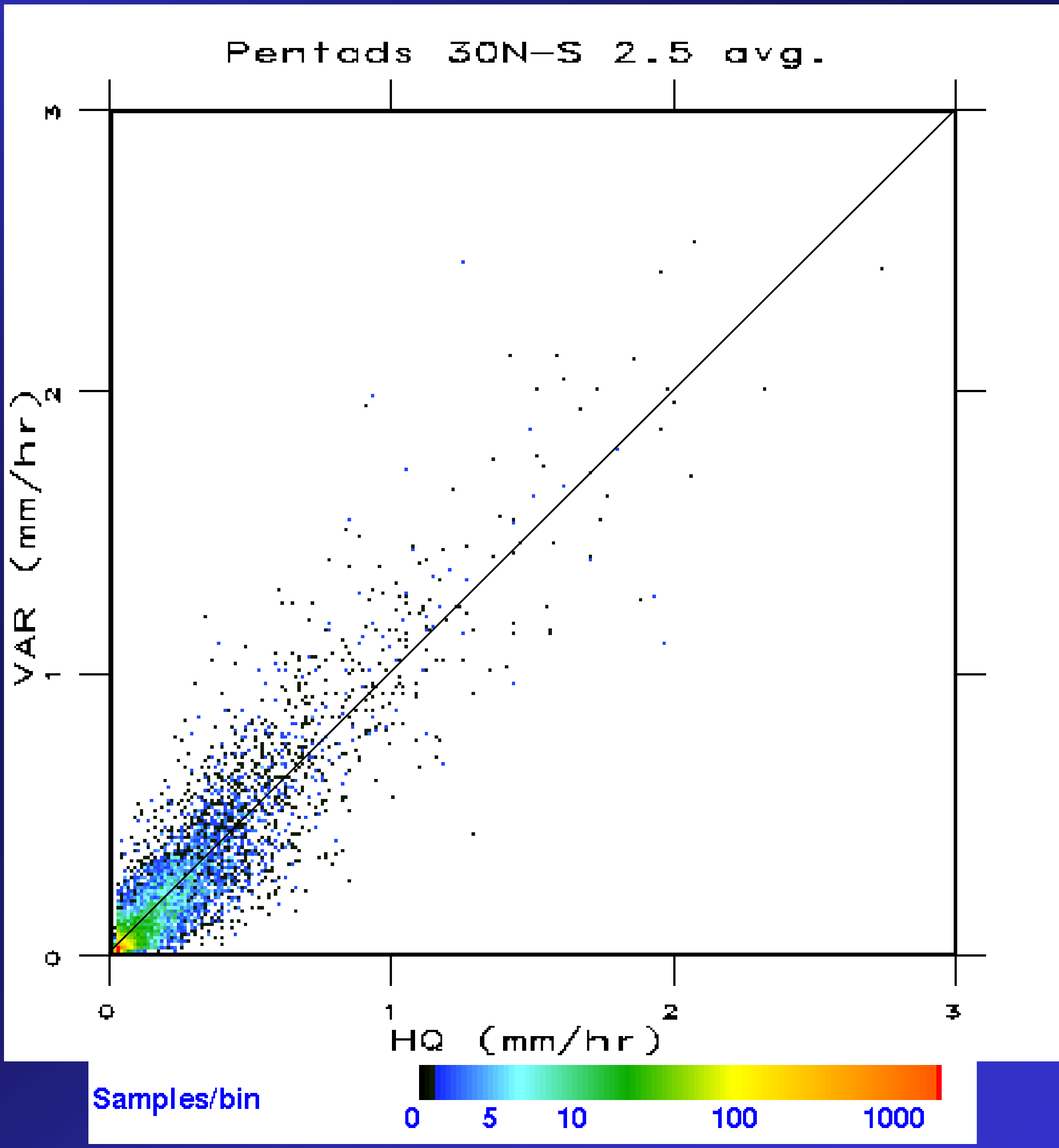
Pentads

2.5°x2.5°

1-5, 6-10, 11-15, 16-20, 21-25 Feb 2002

Latitude band 30°N-S

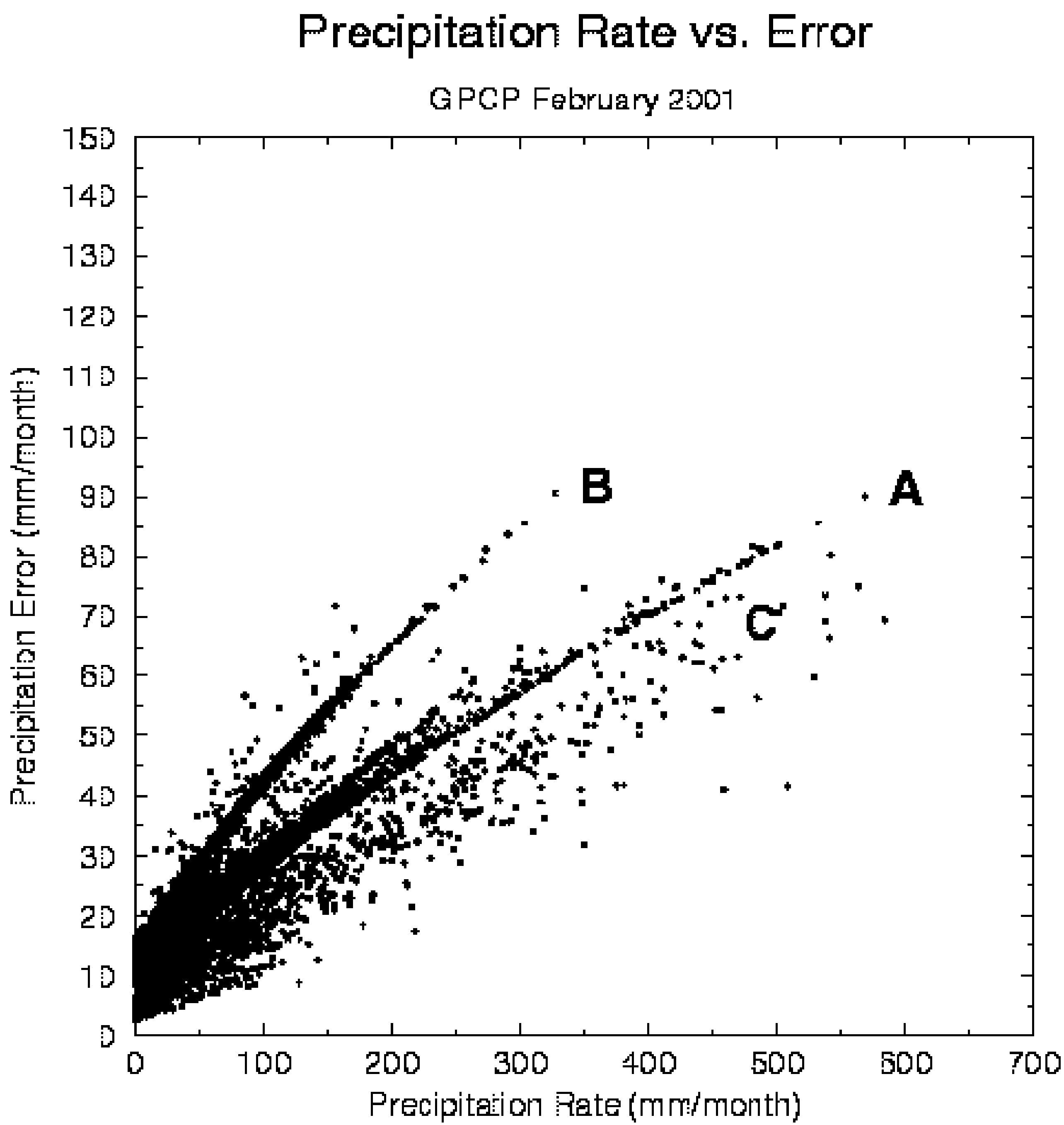
At 2.5°x2.5° there are 100 times fewer grid boxes and considerably less scatter

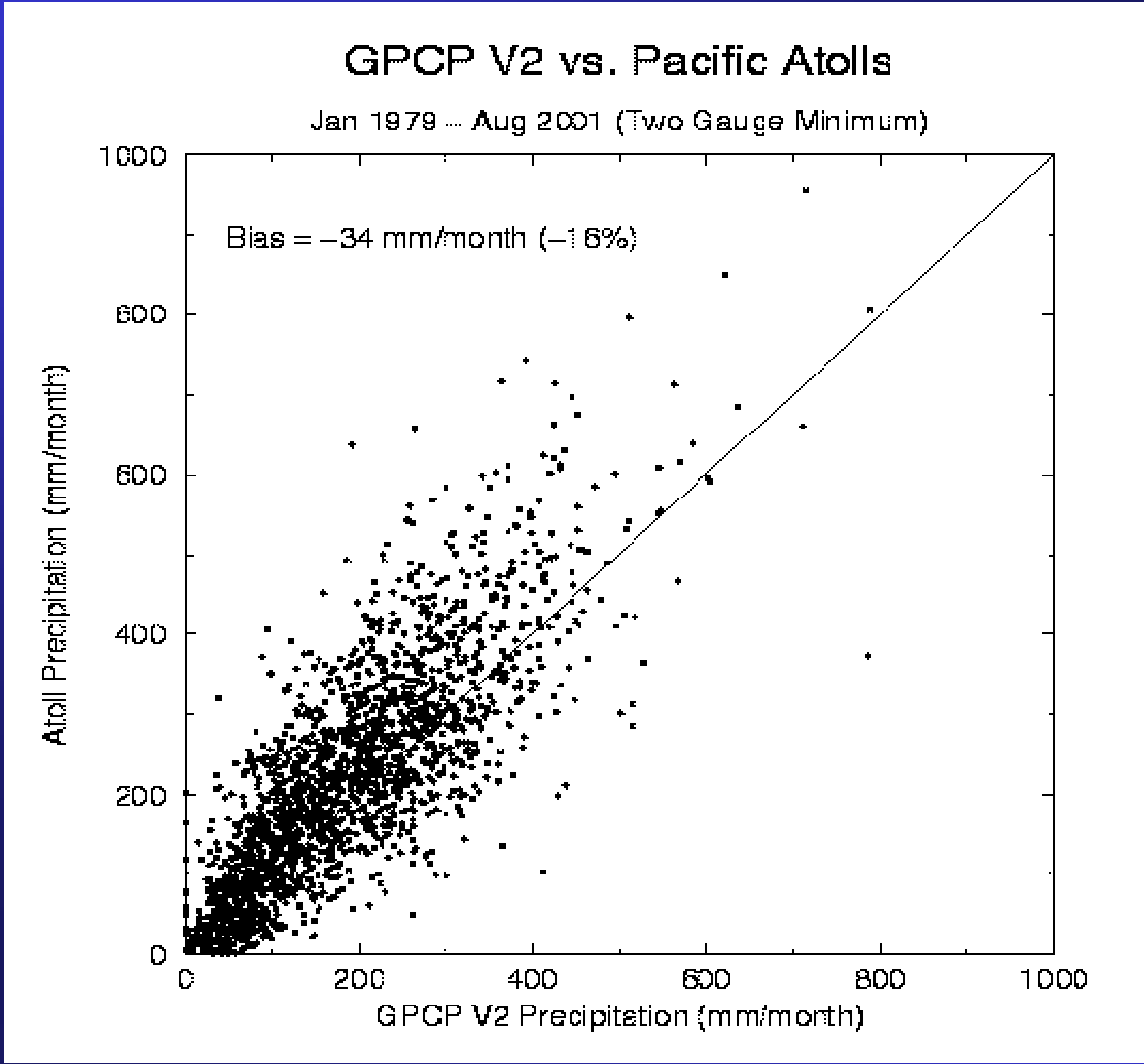


Version 2 Satellite-Gauge analysis and random error field (Feb. 2001)

- **A** (Ocean – SSM/I, IR)
- **B** (Ocean – SSM/I)
- **C** (Land – gauges, satellite)

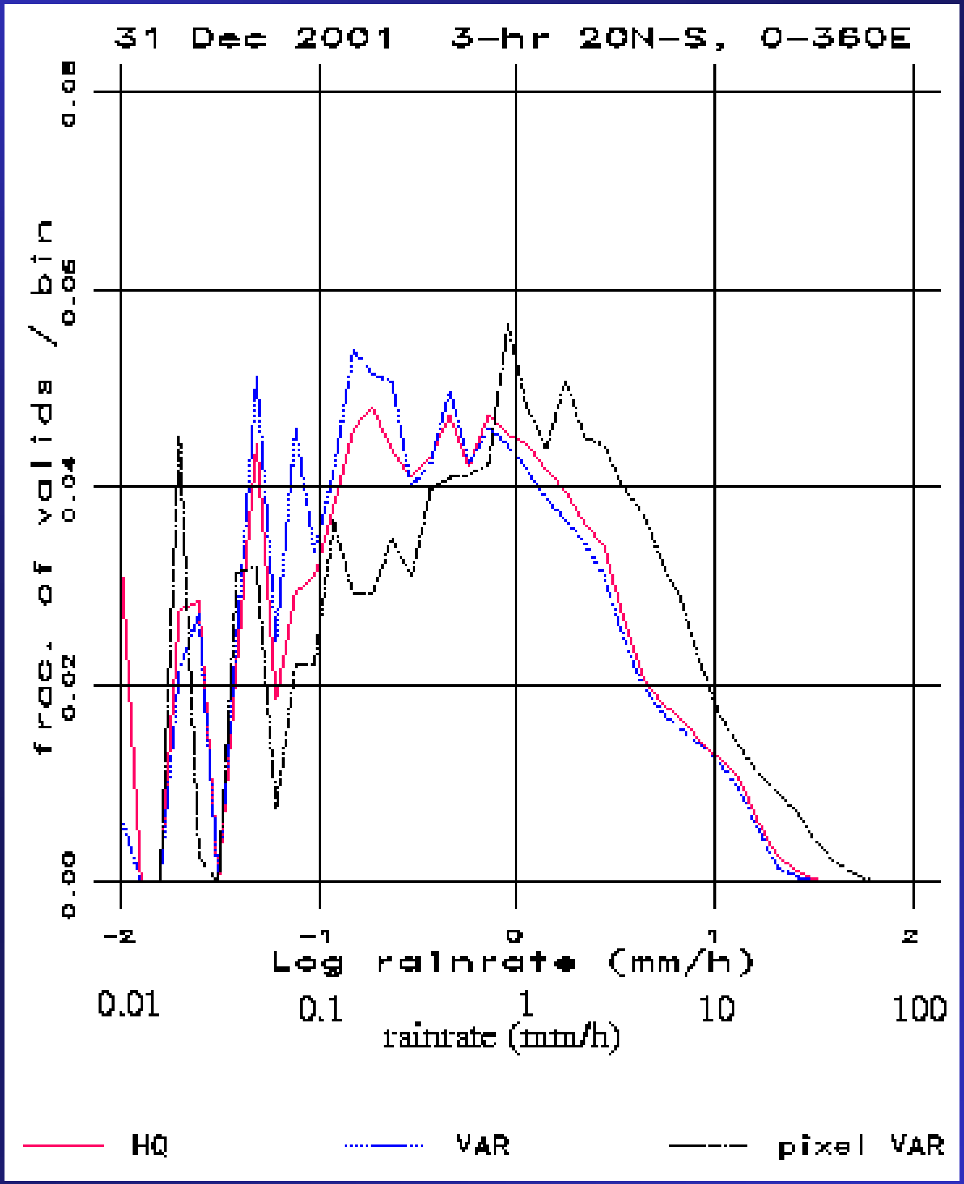
Huffman (1999)



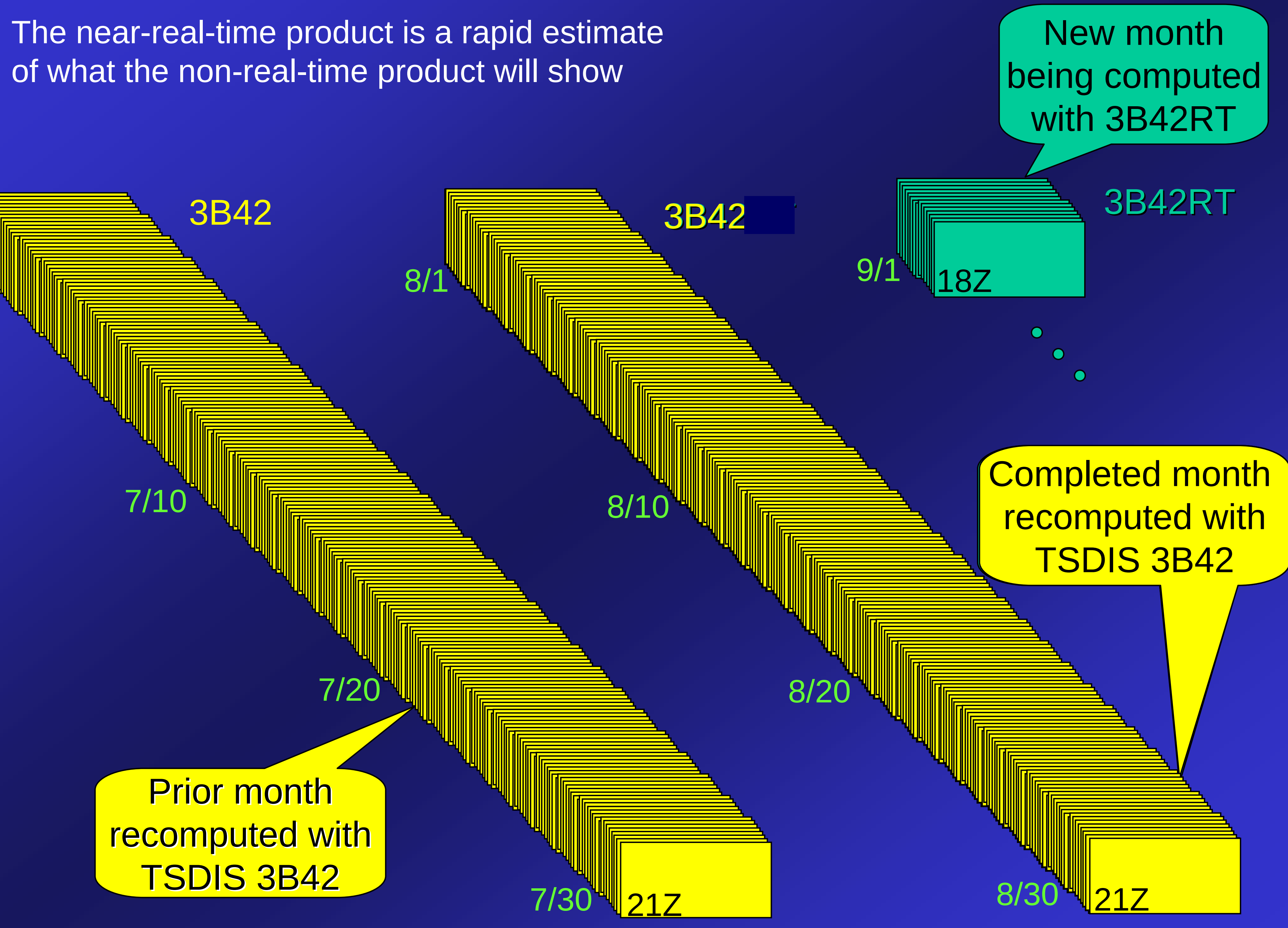

GPCP V2 vs. Pacific Atolls
Jan 1979 – Aug 2001 (Two Gauge Minimum)
Bias = –34 mm/month (–16%)
Atoll Precipitation (mm/month)
GPCP V2 Precipitation (mm/month)
0
200
400
600
800
1000

3. IR-BASED DATA SETS –Scatter Between Coincident HQ and VAR (cont.)

Around the tropics VAR's rainrate histogram provides a good match to the calibrating microwave ("HQ"), better than pixel-based VAR

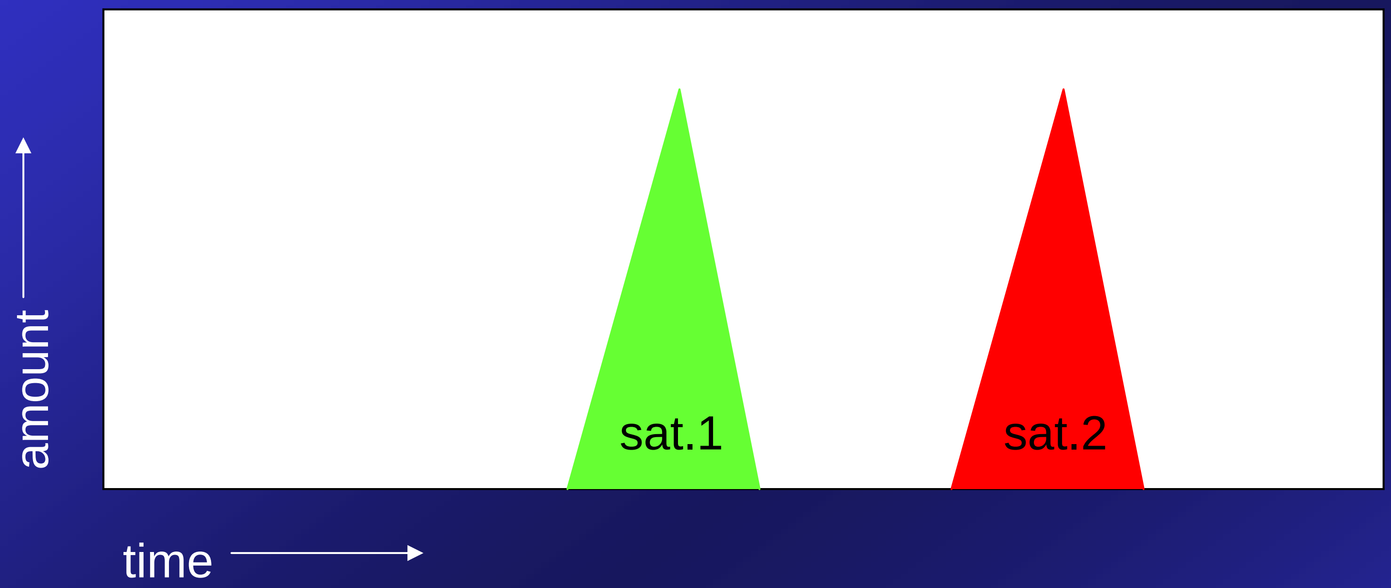


The near-real-time product is a rapid estimate of what the non-real-time product will show



5. CURRENT ISSUES, FUTURE PLANS (cont.)

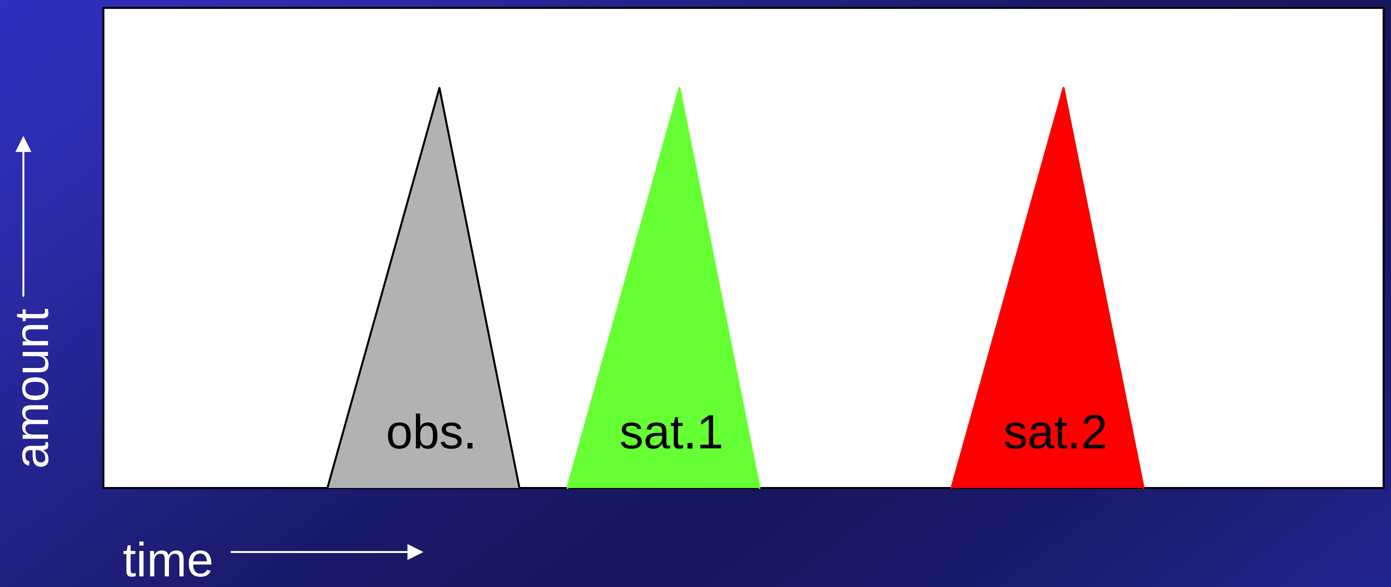
How are these two "satellite" estimates best merged?



- Any linear weighting scheme will damage the statistics:
  - fractional coverage will be too high
  - maximum and conditional rainrates will be too low

5. CURRENT ISSUES, FUTURE PLANS (cont.)

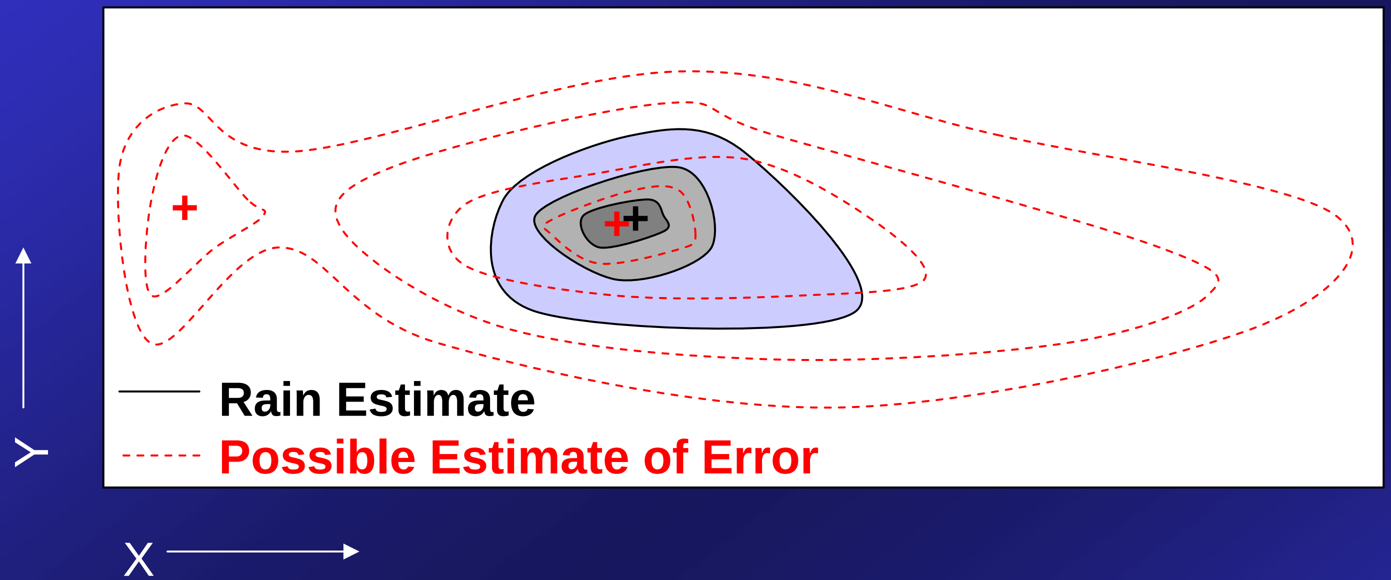
Which “satellite” estimate matches the “observations” better?



- Sat.1 is better than sat.2.
- The usual $\sigma^2 = (\text{sat} - \text{obs})^2$ yields the same bad score for both.
- The improvement can be revealed with “some” averaging, but how much? The answer depends on the averaging.
- What does the user want to know?

5. CURRENT ISSUES, FUTURE PLANS (cont.)

Are there other parameters that can be used to estimate error in zero-rain areas?



- Error is certainly not zero everyplace that zero rain is estimated
- Some locations are very certain not to contain rain, while the no-rain analysis is much less certain in others
- Error estimates in zero rain areas might be helpful in merging different rain estimates
- What does the user want to know?

REFERENCES

Literature:

Adler, R.F., G.J. Huffman, D.T. Bolvin, S. Curtis, E.J. Nelkin, 2000: Tropical Rainfall Distributions Determined Using TRMM Combined with other Satellite and Raingauge Information. *J. Appl. Meteor.*, **39**(12), 2007-2023.

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hdrometeor.*, **4**(6), 1147-1167

Huffman, G.J., R.F. Adler, M. Morrissey, D.T. Bolvin, S. Curtis, R. Joyce, B McGavock, J. Susskind, 2001: Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations. *J. Hydrometeor.,* **2**(1), 36-50.

Huffman, G.J., R.F. Adler, E.F. Stocker, D.T. Bolvin, E.J. Nelkin, 2003: Analysis of TRMM 3-Hourly Multi-Satellite Precipitation Estimates Computed in Both Real and Post-Real Time. *Combined Preprints CD-ROM, 83rd AMS Annual Meeting*, Poster P4.11 in: 12th Conf. on Sat. Meteor. and Oceanog., 9-13 Feb. 2003, Long Beach, CA, 6 pp.

Xie, P., J.E. Janowiak, P.A. Arkin, R.F. Adler, A. Gruber, R. Ferraro, G.J. Huffman, and S. Curtis, 2003: GPCP Pentad precipitation Analyses: An Experimental Data Set Based On Gauge Observations And Satellite Estimates. *J. Climate*, **16**, 2197-2214.

Web sites:

RT: ftp://aeolus.nascom.nasa.gov/pub/merged or http://precip.gsfc.nasa.gov

RT imagery: http://trmm.gsfc.nasa.gov

3B42RT subsetting: http://lake.nascom.nasa.gov/tovas/

Version 6 (as it becomes available):
http://lake.nascom.nasa.gov/data/dataset/TRMM/index.html

Table 1. Summary of publicly available, quasi-operational, quasi-global precipitation estimates from a single sensor. Where appropriate, the algorithms applied to the individual input data sets are mentioned. The TMI GPROF and PR are also available as separate products from the GDAAC. ÒLatencyÓ gives the typical interval between the end of the observational period and release of the product. *Encycl. of Hydro. Sci.*

Algorithm	**Input data**	**Space/time scales**	**Areal coverage/start date**	**Update frequency**	**Latency**	**Archive location**
GPROF	SSM/I	0.5ū/orbit segments	Global Š 70□N-S/July 1987	Monthly	1 pentad	GDAAC (1)
	SSM/I	0.25ū/orbit segments	Global Š 70□N-S/September 2002	3 hours	3 hours	GDAAC (1)
GPROF (3G68)	TMI	0.5ū/hourly	Global Š 37□N-S/December 1997	Daily	4 days	NASA/GSFC TSDIS (1)
GPI	GEO-IR, LEO-IR	2.5□/pentad	Global Š 40ūN-S/ January 1986 Š March 1997	N/A	N/A	NOAA CPC (1)
	GEO-IR, LEO-IR	1□/3-hourly	Global Š 40ūN-S/October 1996	Monthly	1 week	NOAA CPC (1)
NESDIS High Frequency	AMSU	0.25ū/daily 1.0ū/pentad, monthly 2.5ū/pentad, monthly	Global/January 1999	Daily	4 hours	NESDIS ORA
NESDIS/FNMOC SI	SSM/I	0.25ū/daily 1.0ū/pentad, monthly 2.5ū/pentad, monthly	Global/July 1987	Daily	6 hours	NESDIS ORA
OPI	LEO-IR	2.5ū/daily	Global/January 1979	Daily	1 day	NOAA CPC (2)
PR Precip (3G68)	PR	0.5ū/hourly	Global Š 37□N-S/December 1997	Daily	4 days	GDAAC (2)
Wilheit-Chang Statistical	SSM/I	2.5□/monthly	Global ocean Š 60□N-S/July 1987	Monthly	1 month	NASA/GSFC Code 614.3
Wilheit-Chang Statistical (3A11)	TMI	5□/monthly	Global ocean Š 40□N-S/ January 1998	Monthly	1 week	GDAAC (2)

Table 2. Summary of publicly available, quasi-operational, quasi-global precipitation estimates that are produced by combining input data from several satellite sensors. Many of the input data sets are pre-processed into precipitation estimates. The TCI is also available as a separate Level 2 (satellite swath coordinates) product from the GDAAC. ŅLatencyÓgives the typical interval between the end of the observational period and release of the product. *Encycl. of Hydro. Sci.*

Algorithm	**Input data**	**Space/time scales**	**Areal coverage/ start date**	**Update frequency**	**Latency**	**Archive location**
EURAINSAT/A	SSM/I, GEO-IR	4-km/30-minute	Global Š 60N-S/ September 2002	Daily	3 days	Univ. of Birmingham
NOAA CPCP CMORPH	TMI, SSM/I, AMSU, GEO-IR	8-km/hourly	Global Š 60N-S/ December 2002	1 hour	3 hours	NOAA CPC (3)
NRL Real Time	SSM/I, GEO-IR	0.25ū/hourly	Global Š 40ūN-S/July 2000	1 hour	3 hours	NRL Monterey
PERSIANN	TMI, SSM/I, GEO-IR	0.25/6-hourly	Global Š 50N-S/March 2000	6 hours	2 days	Univ. of California, Irvine
TCI (3G68)	PR, TMI	0.5ū/hourly	Global Š 35N-S/ December 1997	Daily	4 days	NASA/GSFC TSDIS (1)
TOVS	HIRS, MSU	1/daily	Global/1979	Daily	1 month	NASA/GSFC Code 613
TRMM Real-Time HQ (3B40RT)	TMI, GPROF-SSM/I	0.25ū/3-hourly	Global Š 70ūN-S/29 January 2002	3 hours	6 hours	NASA/GSFC TSDIS (2)
TRMM Real-Time VAR (3B41RT)	HQ, GEO-IR	0.25ū/hourly	Global Š 50ūN-S/29 January 2002	1 hour	6 hours	NASA/GSFC TSDIS (2)
TRMM Real-Time MPA (3B42RT)	HQ, VAR	0.25ū/3-hourly	Global Š 50ūN-S/29 January 2002	3 hours	6 hours	NASA/GSFC TSDIS (2)

Table 3. Summary of publicly available, quasi-operational, quasi-global precipitation estimates that are produced by combining input data from several sensors, including rain gauges. Many of the input data sets are pre-processed into precipitation estimates. ŅLatencyÓ gives the typical interval between the end of the observational period and release of the product. *Encycl. of Hydro. Sci.*

Algorithm	**Input data**	**Space/time scales**	**Areal coverage/ start date**	**Update frequency**	**Latency**	**Archive location**
CAMS/OPI	CMAP-OPI, gauge	2.5ű/daily	Global/January 1979	Monthly	6 hours	NOAA CPC (2)
CMAP	OPI, SSM/I, GPI, MSU, gauge, model	2.5ű/monthly	Global/January 1979	Seasonal	3 months	NOAA CPC (2)
FEWS Daily Combination	GPI, FNMOC/NESDIS SSM/I, gauge	10 km/daily	Africa/April 2000 South Asia/April 2001	Daily	6 hours	NOAA CPC (2)
GPCP Version 2 SG	*1/79-6/87, 12/87*: GPCP-OPI, gauge *7/87-present except 12/87*: SSM/I, GEO-, LEO-IR, gauge, TOVS	2.5ű/monthly	Global/January 1979	Monthly	3 months	WDC-A
GPCP pentad	OPI, SSM/I, GPI, MSU (*1/79-12/94*), gauge, GPCP SG	2.5ű/5-day	Global/January 1979	Seasonal	3 months	WDC-A
GPCP 1DD	SSM/I, GEO-, LEO-IR, TOVS, GPCP SG	1ű/daily	Global/October 1996	Monthly	3 months	WDC-A
PREC	CMAP-OPI (1979-1998 for development of oceanic EOFs), gauge	2.5ű/monthly (1□,0.5□land)	75□N-60□S/ January 1948	Monthly	10 days	NOAA CPC (4)
TRMM Plus Other Satellites (3B42 V.6)	TCI, TMI, SSM/I, GEO-IR, TRMM 3B43	0.25ű/3-hourly	Global Š 50□N-S/ January 1998	Monthly	1 week	GDAAC (2)
TRMM Plus Other Data (3B43 V.6)	TCI, TMI, SSM/I, GEO-IR, gauge	0.25ű/monthly	Global Š 50□N-S/ January 1998	Monthly	1 week	GDAAC (2)