ICX-FIELDHAWK TECHNICAL PAPER
BNC-AN-102 REV A
AUGUST 2026

Wide-area spectrum monitoring with automated classification

Why the capture, not the classifier, sets the ceiling on what a monitoring system can identify.

ABSTRACT

Monitoring systems increasingly hand the identification problem to a trained model: sweep the band, detect every emitter, capture the unknowns as IQ, and let a classifier name them. This paper derives why a swept receiver cannot supply that classifier with honest data, quantifies what gap-free capture guarantees instead, works the storage arithmetic that forces a triggered architecture, and shows through the data processing inequality that no model can recover what the capture destroyed.

WHO THIS PAPER IS FOR

Monitoring and SIGINT system architects designing a survey-to-identification pipeline. **Machine learning engineers** working on RF classification who need to know what their front end is doing to their data. **Regulators and spectrum managers** specifying monitoring equipment.

CONTENTS

1 The problem with sweeping a wide band

Spectrum monitoring used to end at detection. An operator saw an unknown carrier, wrote down a frequency, and went looking for someone who could say what it was. The systems being built now are asked to answer that question themselves, and they answer it with a trained model.

That changes what the receiver is for. It is no longer a display feeding a human; it is a sensor feeding a statistical estimator, and the estimator inherits every property of the data it is given. The question this paper asks is therefore not how fast an analyzer can sweep, but whether what it produces is fit to train on.

The instinct that follows is to buy a faster analyzer, and it is worth killing here rather than in section 2, because everything after this depends on its being dead. The fraction of each pass spent inside any one resolution cell is RBW divided by span, and that fraction does not care how fast the pass is. A receiver that sweeps twice as fast revisits sooner and hears less each time, in exactly the proportion that leaves the product unchanged. **The quantity that decides what a monitoring system sees is not sweep speed. It is how much of the band is being listened to at once**, and no amount of speed moves that number.

Start with the sweep itself. A swept receiver moves a fixed intermediate-frequency filter across the span. That filter is a physical resonator with a settling time of roughly the reciprocal of its bandwidth, and the signal is inside it for only the fraction of the sweep given by the ratio of resolution bandwidth to span. Requiring the filter to settle gives the classical relation:

$$T_{\text{sweep}} \approx k \cdot S \,/\, \text{RBW}^2 \qquad (1)$$

where

S is the span, in hertz

RBW is the resolution bandwidth, in hertz

k is a filter shape constant, roughly 2 to 3 for a near-Gaussian filter

The resolution bandwidth appears squared for two reasons at once, and this is the structural fact worth remembering. Halving it doubles the number of resolution cells to visit, and it also doubles the settling time in each one. The two effects multiply.

Table 1 works that through for a full-span survey.

Table 1. Filter-limited sweep time for a 9 kHz to 40 GHz span. Improving resolution by a factor of ten costs a factor of one hundred in time. Real systems are slower still, because a 1,001-point trace across 40 GHz assigns 40 MHz to each point, so full-resolution coverage at 100 kHz needs hundreds of segmented acquisitions, each with its own retune overhead.

RBW	Resolution cells	Sweep time at k = 2
10 MHz	4,000	0.8 ms
1 MHz	40,000	80 ms
100 kHz	400,000	8 s
10 kHz	4,000,000	13.3 minutes
1 kHz	40,000,000	22.2 hours

2 What a swept receiver actually intercepts

Sweep time is not the interesting number. The interesting number is how long you wait before you see a given emitter at all. For a periodic emitter of pulse duration tau and repetition interval T at a fixed frequency, observed by a receiver that revisits that frequency every T_r and dwells there for t_d, the per-revisit intercept probability is the overlap of the dwell and the pulse:

$$p \approx \min[1, (t_d + \tau) \,/\, T], \quad t_d = T_r \cdot \text{RBW} \,/\, S \qquad (2)$$

where

t_d is the dwell in one resolution cell per revisit, in seconds

τ is the pulse duration, in seconds

T is the pulse repetition interval, in seconds

T_r is the revisit interval, in seconds

The substitution is worth doing on the page rather than asserting. Expected time to first intercept is the revisit divided by the per-revisit probability, T_r / p. Where the pulse is short against the dwell, $p \to t_d / T$, and substituting $t_d = T_r \cdot \text{RBW} / S$ gives $T_r \cdot T \cdot S / (T_r \cdot \text{RBW})$. **The revisit cancels**, which is the algebra behind the sentence in section 1, and what is left surprises most people the first time they see it:

$$\text{E}[T_{\text{intercept}}] \to T \cdot S \,/\, \text{RBW} \qquad (3)$$

THE ONE RESULT WORTH REMEMBERING

Expected time to first intercept is the pulse repetition interval multiplied by the number of resolution cells. The revisit interval is not in the expression, which is why buying sweep speed buys nothing.

Two worked cases, both for a 40 GHz span. A radar with a 1 microsecond pulse at 1 kHz repetition, observed at 1 MHz resolution with a 100 ms revisit, gives a 2.5 microsecond dwell and a per-revisit probability of 0.0035: expected time to first intercept 28.6 seconds, and a 3.4 percent chance of seeing it in a one-second look. Narrow the resolution to 100 kHz for a better look and the revisit stretches to 8 seconds, so expected time to intercept becomes 381 seconds. **Improving resolution by ten made time to first intercept thirteen times worse.**

Worse still, an intercept is not a measurement. A pulse shorter than the reciprocal of the filter's impulse bandwidth cannot develop full response, and the displayed amplitude is low by approximately:

$$\alpha_{\text{dB}} \approx 20 \log_{10}(\tau \cdot B_i), \quad B_i \approx 1.5 \cdot \text{RBW} \qquad (4)$$

where

B_i is the impulse bandwidth of the resolution filter, in hertz

That 1 microsecond pulse at 100 kHz resolution reads 16.5 dB low. At 100 nanoseconds it reads 36.5 dB low. So with a swept receiver the setting that makes an emitter findable and the setting that makes it measurable are different settings, and no single sweep provides both.

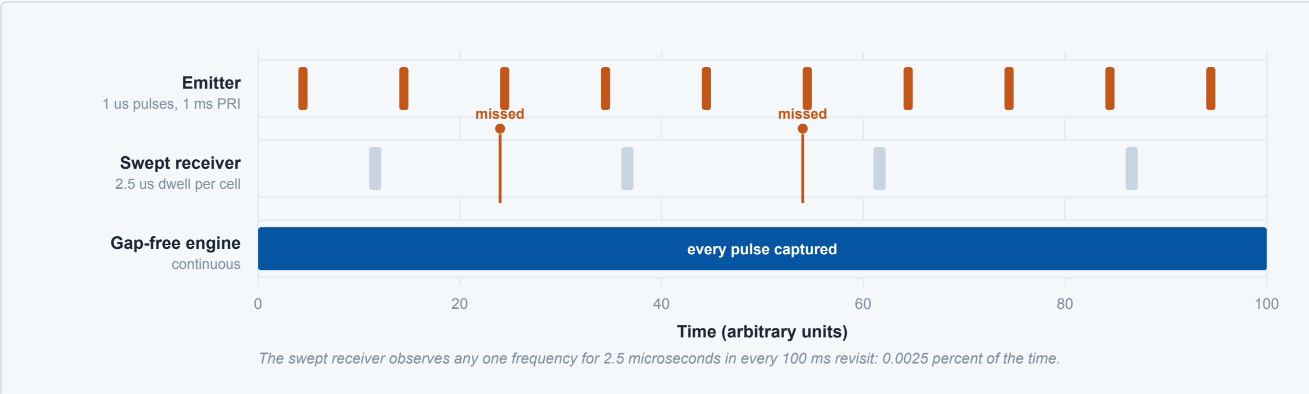

Figure 1. At 1 MHz resolution across 40 GHz a swept receiver is looking at any given frequency for two and a half microseconds in every hundred milliseconds. Counting the pulse width, the window in which a 1 microsecond pulse can land is 3.5 microseconds per revisit, so about one pulse in twenty-eight thousand from a 1 kHz emitter arrives while the receiver is looking at that frequency.

SCHEMATIC

Figure 1 draws the two timelines against each other. The strips are drawn wide enough to see: at true scale a dwell would be 25 micrometers on a timeline one meter long.

Read it for the order of the two rates rather than their ratio. Pulses arrive a thousand times a second and the receiver returns to that frequency ten times a second.

3 What gap-free capture guarantees instead

A real-time engine transforms a continuous stream rather than tuning across it. Where a vendor publishes the governing relation rather than a headline figure, the guarantee can be computed rather than trusted:

$$\mathrm{POI} = 2 \times N \times D \times 8\ \mathrm{ns} \tag{5}$$

where

N is the transform size, in points

D is the decimation factor

POI is the duration for which a signal must be present for its amplitude to be measured correctly

The factor of two is the arrival-phase penalty. A burst landing anywhere inside a single transform window is detected, but its amplitude is wrong, because only part of the window held energy. Spanning two consecutive frames guarantees at least one saw it whole. The published operating points are 0.512 microseconds at a 32-point transform and 32.768 microseconds at 2048 points, at frame rates of 3,906,250 and 61,035 per second respectively.

The density capture later in this section was taken at neither of them. Its display reports a 65.54 microsecond probability of intercept, and 2 × 4096 × 8 ns is 65.536, which is 65.54 to the four digits the display carries. Its 60.3 kHz resolution is 1.98 bins of the 30.518 kHz spacing a 4096-point transform gives at 125 MSPS. Five readings off one screen agree with each other and with the relation above, at a transform size this paper never quotes, which is what a published relation buys over a published figure: the reader can price any size, not only the two on the datasheet.

Note what kind of claim this is. It is not a probability that improves with observation time. It is a boundary: below it, nothing is guaranteed; above it, everything is measured, whatever the arrival time. That difference in kind is what makes a specification meaningful, and it is what to hold on to when the next figure draws both on one pair of axes: the step is a bound, the curve is an average, and only one of the two can be written into a requirement.

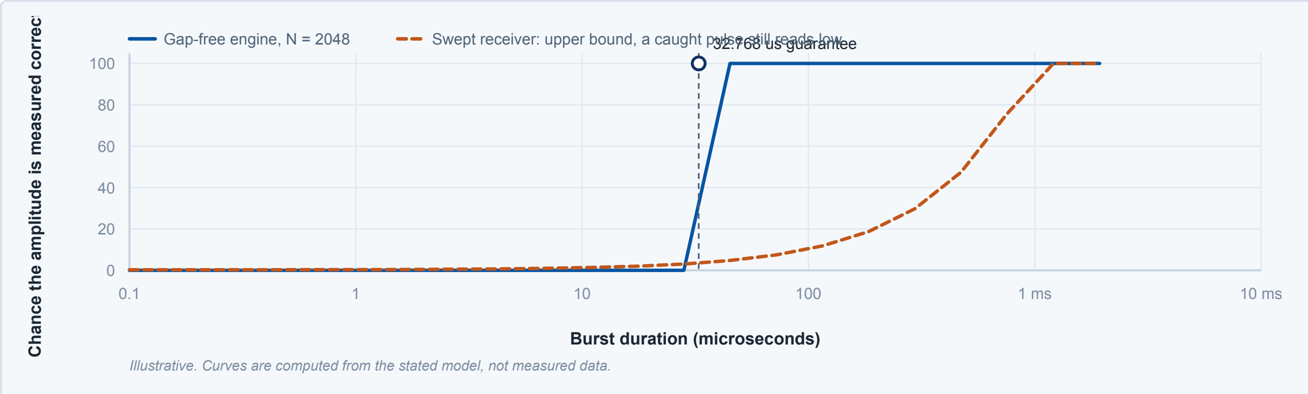

Figure 2. The gap-free engine's intercept figure is a guarantee boundary, not a probability. Every burst longer than two transform frames is measured at its true amplitude regardless of when it arrives. The swept curve is drawn for the section 2 case, a 100 ms revisit at 1 MHz resolution across 40 GHz against a 1 ms repetition interval, and even where it catches the burst the amplitude is still low by equation 4 until the burst outlasts the filter. DERIVED

Figure 2 plots both on the same axes. The shapes matter more than the values: a step a specification can be written against, and a curve whose height depends on three settings the emitter knows nothing about.

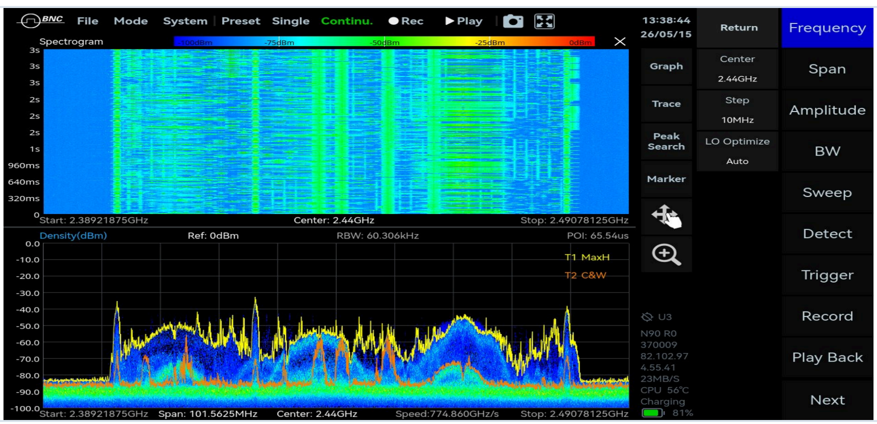

Figure 3. Real-time density with its spectrogram, on an ICX-FieldHawk: 2.389 to 2.491 GHz in a 101.56 MHz span at 60.3 kHz resolution, with the display reporting a 65.54 microsecond probability of intercept, which is 2 × 4096 × 8 ns. The density surface, which the real-time engine computes from every transform rather than from the peak of a sweep, means an intermittent emitter appears as a persistent statistical trace rather than as an occasional spike that may or may not be there on the next look. MEASURED

Figure 3 is the density surface with persistence set to accumulate rather than decay. Read the color as occupancy rather than as amplitude: a faint continuous band is an emitter present at low duty cycle, and a bright narrow line is one that is always there. That distinction is not available from a max-hold trace, which reports both as a line.

4 The pipeline, and where the data goes

A monitoring system is a sequence of volume reductions. Naming them in order, with their data rates, is the fastest way to see why the architecture is shaped as it is. The first two stages ship in SpectraCore; everything past the parameter record is the integrator's.

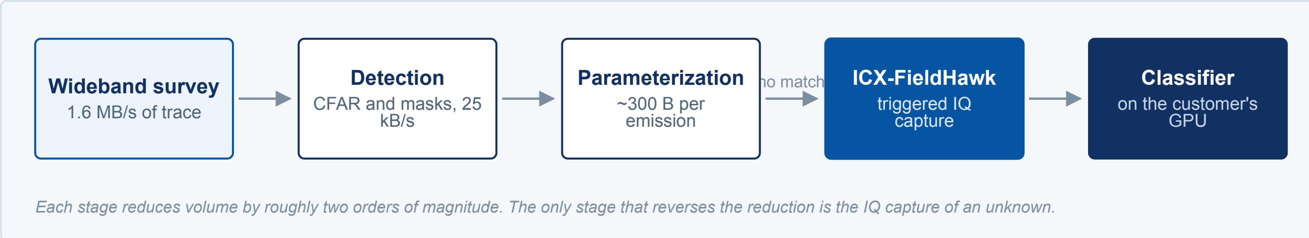

Figure 4. The classifier's real job is to keep the IQ capture path narrow. A 300-byte parameter record is twenty thousand times cheaper than the 6 Mbyte capture that replaces it, which is forty-eight milliseconds of IQ at 25 MHz, so every emitter the system can name from parameters alone is an emitter it does not have to record. SCHEMATIC

Figure 4 sets out the stages. The arithmetic behind them is unforgiving:

$$r = f_s \times 2 \times b/8 \tag{6}$$

where

f_s is the complex sample rate, in samples per second

b is the bits per component, 16 in a conventional format

Table 2 turns that rate into storage.

Table 2. Storage for continuous 16-bit complex capture. One hour of gap-free 100 MHz is roughly 1.8 terabytes, and capturing the whole 40 GHz span this way would be seventeen petabytes a day. That number is the reductio that forces triggered, decimated capture.

Analysis bandwidth	Sample rate	Rate	Per hour	Per day
25 MHz	31.25 MSPS	125 MB/s	450 GB	10.8 TB
100 MHz	125 MSPS	500 MB/s	1.8 TB	43.2 TB
Full 40 GHz span	n/a	n/a	impossible	17 PB

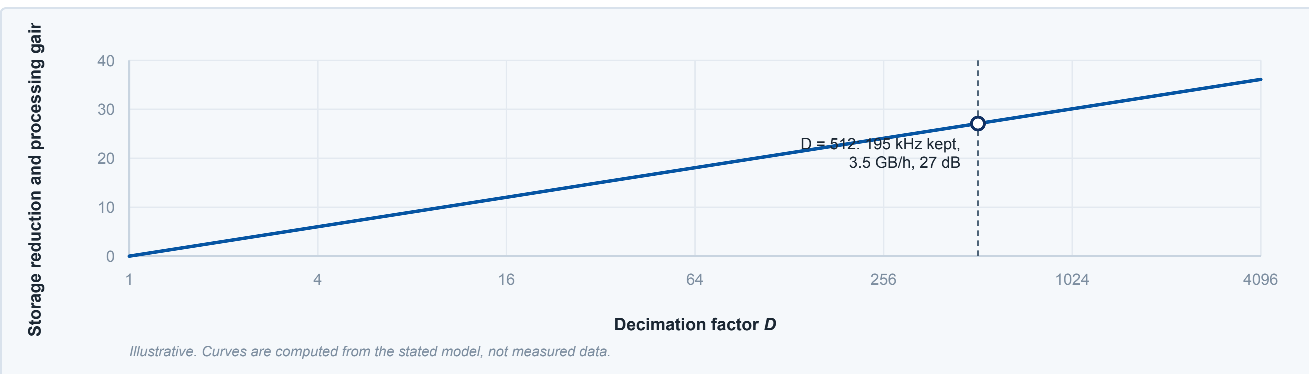

Figure 5. Decimation is the only lever that moves storage by orders of magnitude without discarding a signal you meant to keep. Decimating by 512 keeps 195 kHz of the survey's 100 MHz, takes an hour of one channel from 1.8 terabytes to 3.5 gigabytes, and returns 27 dB of processing gain. DERIVED

Figure 5 plots it. The difference between recording what was found and recording where it was found is the difference between a rack of disks and a shelf of them.

This is why the ICX-FieldHawk datasheet separates its two recording figures instead of quoting the flattering one twice: burst recording covers the full 100 MHz into a 128 Mbyte buffer, which is about 0.26 seconds, while continuous recording to a host is specified at 25 MHz. Filling a buffer and sustaining a transfer are different constraints, and an honest datasheet says so.

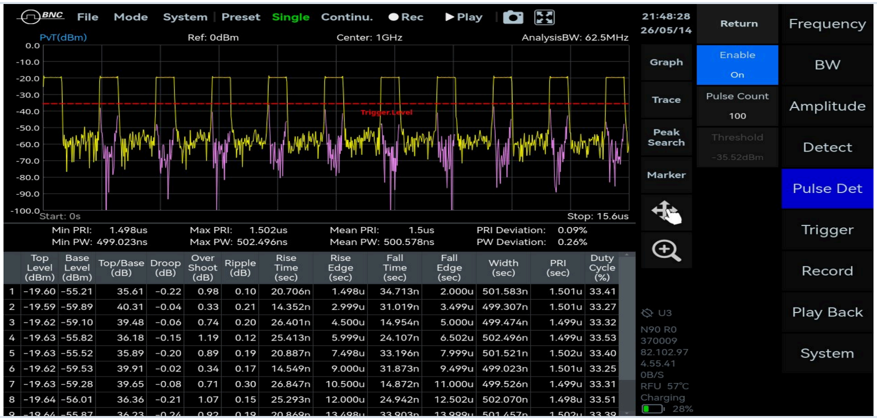

Figure 6. The parameterization stage doing its job: 500 ns pulses at a 1.5 microsecond interval, centered at 1 GHz in a 62.5 MHz analysis bandwidth, with a per-pulse table of level, droop, overshoot, ripple, rise and fall times, width, interval and duty cycle. The displayed 15.6 microsecond record holds about ten pulses and the analyzer accumulates the requested hundred into the table, where the measured interval deviation of 0.09 percent is the kind of feature a classifier keys on. It exists because pulse detection measures in the time domain with a dedicated detector rather than through the transform path, so the intercept boundary of section 3 does not apply to it; what the gap-free front end guarantees is that no pulse in the record is missing. **MEASURED**

Figure 6 is the middle of that pipeline made concrete: each row is a few hundred bytes standing in for a burst that cost megabytes to record. One question remains about the last box in Figure 4, which is how the customer's classifier receives the samples. The ICX-FieldHawk presents itself through SoapySDR, the vendor-neutral hardware abstraction layer, so it appears to GNU Radio and any other SoapySDR application as an ordinary software-defined radio device, and the instrument calibration files install with the driver. The samples reaching the classifier are corrected before the first processing block sees them. A system built this way inherits an existing ecosystem of detectors and demodulators instead of commissioning them, with the amplitude under every detection still traceable. Demonstrated rather than specified; see the verification note.

5 Derived requirements

Sections 2 to 4 were an argument about architecture. Collapsed into requirements they are six things a reader can score any vendor against, this one included. It is a scoring sheet assembled from the argument rather than a hinge that precedes it, which is the honest description: the reasoning is in sections 2 to 4 and Table 3 is its index.

Table 3. Six requirements, derived rather than asserted. R1 and R6 eliminate most candidates before any classifier is discussed, which is the reverse of the order these systems are usually specified in.

#	Requirement	Why not something looser	From
R1	Probability of intercept published as a guarantee, with its transform size	An average intercept figure cannot be written into a specification, because it does not say what was missed.	§3
R2	Gap-free capture across the recorded segment, with no stitching	A seam manufactures wideband energy that a feature extractor cannot distinguish from an emitter.	§3
R3	Bounded amplitude accuracy under stated conditions	Features are compared across sites and across years; an unbounded amplitude makes two detections incomparable.	§6
R4	Raw IQ with published sample rate, depth and decimation range	The classifier consumes samples. A vendor trace is a conclusion somebody else already drew.	§4
R5	One documented interface across every form factor in the fleet	A monitoring program outlives the model it started on, and revalidating per model is the cost that ends programs.	§4
R6	Sustained throughput matched to the decimated bandwidth, not the converter	Burst depth and sustained rate are different specifications, and only one of them survives a long record.	§4

6 Why the capture sets the ceiling

The classifier never sees the emitter. It sees it after the channel and after the instrument. Class, signal and observation form a Markov chain, so the mutual information between the class and what the model sees can never exceed that between the class and the signal. The data processing inequality turns a marketing sentiment into a theorem: **no architecture, no quantity of training data and no amount of compute recovers information the capture destroyed.**

> **THE FAILURE THAT INFLATES YOUR ACCURACY**
>
> There is a worse case than degradation. If an instrument artefact varies with class, because classes were collected on different days, gain states or bands, the model can read the instrument instead of the signal and report excellent accuracy that is entirely non-transferable. Artefacts do not merely weaken a classifier. They flatter it while destroying its field performance.

Four capture properties feed directly into that risk.

- **Gaps.** A capture with gaps is not a shorter capture, it is a different signal. A burst straddling a gap has the wrong duration, and periodic gating convolves the spectrum with a comb set by the instrument, which a cyclostationary feature extractor cannot tell from an emitter's own framing.
- **Amplitude flatness.** If the front end is 1.5 dB hotter at 2.4 GHz than at 5.8 GHz, and one class lives in each band, the model can score well by reading the front-end response. Of the ways a real dataset acquires a shortcut, this is the one that survives review.
- **Phase noise.** Integrating a flat −100 dBc/Hz from 1 kHz to 1 MHz gives 0.81 degrees of residual phase error, or 1.41 percent of error vector magnitude before the signal arrives, against the 2.5 percent a 256QAM signal has to stay inside. The local oscillator has spent most of the budget, and classifiers confuse themselves precisely between adjacent QAM orders. **The receiver caps modulation-order accuracy before the model is trained.**
- **Quadrature imbalance.** One tenth of a decibel of gain error and one degree of phase error give about 40 dB of image rejection. The resulting mirror is spectral asymmetry manufactured by the receiver, and spectral asymmetry is a genuine modulation feature. The artefact is not merely noise on the feature; it is indistinguishable in kind from the feature.

> **DESIGN RULE**
>
> Every instrument parameter that is not the label must vary independently of the label across the training set, and must vary less than the smallest inter-class difference you intend to resolve. Evaluate with a leave-one-collection-condition-out split, never a random split. A random split cannot detect the failure that decides field performance.

7 Measurement practice

None of this depends on which analyzer you buy, and all of it determines whether the resulting dataset is worth training on.

1. **Set the false-alarm budget first.** Forty gigahertz at 100 kHz resolution is 400,000 cells. At one revisit per second, holding false alarms to one per hour needs a per-cell threshold about six standard deviations above the noise. Requiring three detections in five looks reaches the same rate at roughly three standard deviations, buying back most of the sensitivity at the cost of five times the latency and every emitter that appears fewer than three times.
2. **Randomize everything that is not the label.** Gain state, band, antenna, cable, site, day and temperature. Where a variable cannot be randomized, record it so its contribution can be tested afterward.
3. **Record the SpectraCore state with every capture.** Center frequency, sample rate, reference level, attenuation, decimation, transform size, temperature and a disciplined timestamp. An IQ file without a sidecar is a pile of integers.
4. **Decimate at the point of capture.** Record the emission's bandwidth, not the survey's. Keeping 200 kHz of an emission found in a 100 MHz window is a 500 to 1 reduction before any other measure is applied, and the decimation filter returns the matching processing gain.
5. **Never transform across a recorded gap.** Mark overflows in the record and skip those frames rather than letting the seam manufacture wideband energy.
6. **Calibrate the classifier's confidence.** Modern networks are systematically overconfident, and in a monitoring system the confidence number drives operator workload, so an uncalibrated score is an operational defect rather than a statistical footnote.

Field evidence

The pipeline in this paper is not a proposal. Systems integrators serving defense and national security customers have built exactly this architecture, buying the capture guarantee and the documented interface and writing every layer above the IQ line themselves, which is the arrangement Figure 4 draws. What comes back from those programs is which decision turned out to be load-bearing, and it is not the classifier. A system that changes its detector can reprocess its archive and arrive at a comparable answer; a system that changes its receiver cannot compare across the change at all, so the capture is the only part of the pipeline that has to be right the first time.

8 Honest limits

- **Sweeping is still the right tool for some jobs.** Against a stationary carrier a swept receiver at a narrow resolution bandwidth reaches a lower noise floor and resolves closer-spaced tones than a real-time engine of the same cost, because it spends its entire dwell in one cell instead of dividing it across a band. If every emitter of interest is always on, the intercept arithmetic in section 2 never bites and the extra hardware buys nothing.
- **A classifier cannot name what it was never trained on.** A closed-set model assigns every input to a known class, confidently. Open-set methods reduce this; they do not remove it.
- **Reported accuracy is conditional.** Figures around 95 percent are achieved on synthetic data with a known class set and a fixed record length; the same model over the air loses a substantial part of it. An accuracy quoted without record length, class set and collection conditions is not a specification.
- **The operator stays in the loop.** The output that matters most operationally is the admission that the system has never seen something before, and that admission is by construction one the system cannot resolve. The metric that matters is not accuracy but operator-hours per confirmed identification.
- **Externals and content are different activities, legally as well as technically.** Frequency, bandwidth, timing, modulation and bearing are monitoring. Recovering content is not, and the architecture should reflect that: classify externals by default and gate demodulation behind an explicit, logged, authorized action. That happens to be the same architecture the storage arithmetic already forced.

9 Summary, definitions and further reading

Table 4 collects the governing relations and published values with their conditions.

Table 4. Governing relations and published values with their conditions.

Quantity	Relation or value	Condition
Swept sweep time	$k \cdot S / RBW^2$	k about 2 to 3
Expected time to intercept	$T \cdot S / RBW$	dwell longer than the pulse
Pulse desensitization	$20 \log_{10}(\tau \cdot 1.5\ RBW)$	$\tau \cdot B_i$ well below 1
100 percent POI	2 × N × D × 8 ns	full amplitude accuracy
Published POI points	0.512 us at N = 32; 32.768 us at N = 2048	D = 1
Gap-free bandwidth	100 MHz where fitted	the 4.5 to 9 GHz handheld models ship with 50 MHz as standard, 100 MHz optional
Burst capture	128 Mbyte, about 0.26 s at 100 MHz	16-bit components
Continuous capture	25 MHz	sustained to host

Source: the ICX-FieldHawk handheld, rugged and USB datasheets. Values on the preliminary portfolio and overview pages should be verified prior to application deployment.

CFAR
Constant false alarm rate. A detection threshold estimated from surrounding cells so that the false alarm rate stays fixed as the noise floor varies.

Cyclostationarity
Periodicity in a signal's second-order statistics, tied to symbol rate, carrier or frame structure. Stationary noise has none, which is why cyclostationary detection works below the energy detection floor.

Probability of intercept
The minimum duration a signal must be present for the analyzer to measure its amplitude correctly. A stronger claim than the duration required merely to detect it.

Revisit interval

The time between successive observations of the same frequency by a swept receiver. The quantity that governs what it misses.

Symbols used in this paper.

Symbol	Meaning	Units
S	Span	Hz
RBW	Resolution bandwidth	Hz
T_r	Revisit interval	s
t_d	Dwell per resolution cell	s
τ	Pulse or burst duration	s
T	Pulse repetition interval	s
N, D	Transform size, decimation factor	points, dimensionless

Further reading

- Recommendation ITU-R SM.1600, *Technical identification of digital signals*. The anchor standard for identifying digital emissions from signal externals and IQ recordings.
- Recommendation ITU-R SM.1880, *Spectrum occupancy measurement and evaluation*.
- ITU Radiocommunication Bureau, *Handbook: Spectrum Monitoring*.
- T. J. O'Shea, T. Roy and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, 2018. The source of the synthetic-versus-over-the-air accuracy gap.
- T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley, 2006, section 2.8, for the data processing inequality.

If you are specifying a monitoring sensor and want to work through the intercept and storage arithmetic against your own emitters of interest, our application engineers would be glad to do that with you.

VERIFICATION NOTE

The following values in this paper are not yet confirmed against a published Berkeley Nucleonics datasheet and are marked verify in the text. They must be confirmed before this paper is released.

- ITU-R revision suffixes should be confirmed at itu.int before publication. SM.1794 on wideband monitoring was suppressed in 2018 and must not be cited as current guidance.
- Onboard AI acceleration appears only on preliminary pages, on no datasheet. This paper deliberately places classification on the customer's own compute.
- The SoapySDR and GNU Radio path in section 4 is documented in a platform integration guide and demonstrated with captured results, but does not yet appear on a datasheet. The distribution route is unsettled, so no package name, repository or driver string is printed here.
- The deployment described in the field evidence section is reported from account history rather than a published source, and no end user, vendor or country is named.

Berkeley Nucleonics Corp.

Test, Measurement and Nuclear Instrumentation since 1963

For details: info@berkeleynucleonics.com or 800-234-7858

BNC-AN-102 Rev A

August 2026