A New Generation of Aviation HF Receivers

Direct-to-Digital Technology Eliminates IF Filters, Local Oscillators and Minimizes RF Filtering

Bob Lombardi, Rockwell Collins, Inc.

Radio communications in the high frequency (HF) 2-30 MHz range date back to the earliest days of radio. Originally given to the amateur operators as useless spectrum [1], amateurs soon discovered the most useful characteristic of HF: this is the only place in the electromagnetic spectrum where worldwide communication is routinely possible without the need for man-made infrastructure. While spectrum usage is changing with time, as all spectrum usage has, HF continues to be used worldwide by civil and military aircraft when beyond the range provided by the short range VHF network.

For most of the history of HF communications, dating back to the work of Edwin Armstrong in 1918, HF receivers have been multiple conversion superheterodynes, with multiple local oscillators (LOs), mixers, RF filters, and narrow-band IF filters to set the channel bandwidth. Recent generations of HF radios have generally included digital demodulation and processing of voice audio, data, aircraft selective calling (SELCAL) and other traditionally analog functions (see Fig. 1).

Today, however, an interesting alternative is coming into existence. By optimizing the RF portion of the design and taking advantage of advances in analog to digital converter (ADC) technology, with digital signal processing (DSP), it is possible to create a receiver that reaches extremely high levels of performance with no LO or IF, no IF channel filtering and only minimal RF filtering.

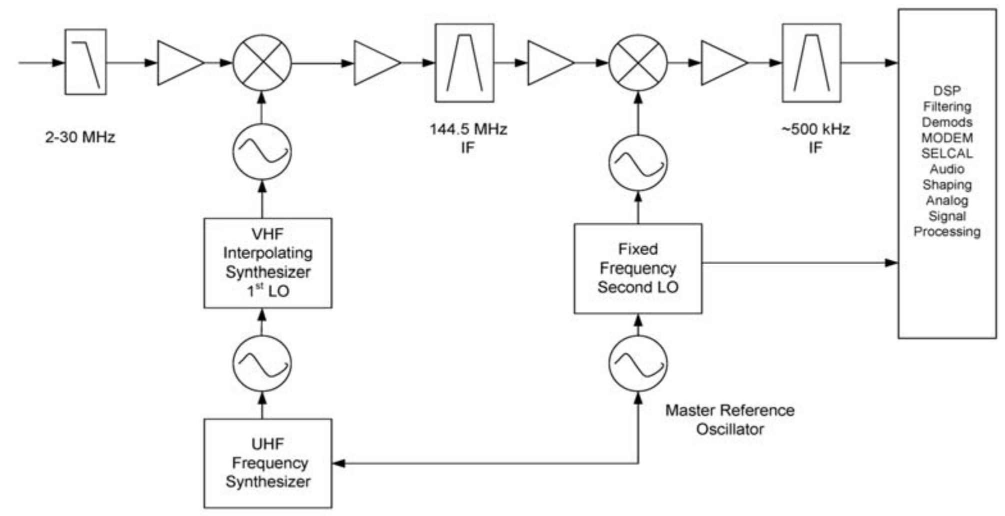

FIGURE 1

A recent (~15-year-old) HF receiver block diagram which uses DSP demodulation and signal processing instead of discrete analog components.

Why would a designer want to risk such a design? How can any design eliminate the filters that we have relied on for more than 70 years?

How Do We Change?

The "why" question is easier, and many designers already know the answer, so let's start there for review and introduction. Modern superheterodynes are plagued with several common problems that arise from nonlinearities in the signal chain. Mixers are the only intentionally non-linear parts in a receiver and the source of many of these problems. The most commonly thought of problem is spurious signal responses. For any non-linear process with two input tones, LO and RF, the output of the circuit has responses at frequencies described by the Diophantine equation M*RF + N*LO where M and N are any integers. For example the first IF is often defined by 1*RF + 1*LO, common high side injection with the IF above the received band (seen in Fig. 1), but there are responses at any set of integers M and N, and the third-order products are notorious causes of trouble in receivers because they are found at small frequency offsets from the IF output. (The order denotes the sum of the two integers; third order is, for example, 2*RF–1*LO or 2*LO–1*RF). For practical analog mixers, these responses go down in severity as the order increases, typically ceasing to be a concern as the order goes over eight or nine (i.e., fourth and fifth harmonics of LO and RF). All experienced receiver designers, however, have horror stories of very high order spurious responses that unexpectedly bit them.

In addition to the spurious frequency responses, nonlinearity in the mixer causes gain compression and other distortions such as intermodulation and cross-modulation. Here, the amplifiers in the receiver chain can get into the distortion act along with the mixer. Even passive components in filters or antennas can cause passive intermodulation distortion (PIM) if the signal levels are high enough. This will make the third order products higher in amplitude, and more troublesome. In my experience, I find this issue overstated, and you must reach impractically high powers for filter components to cause PIM.

The local oscillator itself is not usually considered when looking at distortion issues, but a couple of aspects of the LO spectrum can cause system problems. First, the noise spectrum of the LO will mix with strong adjacent channel signals and produce additional noise in the channel, a phenomenon called reciprocal mixing. Second, spurious signals on the LO can also mix with off-frequency carriers back into the desired channel. In addition to problems we generate, transmitters sharing the HF spectrum will have a broadband noise floor that can be troublesome, either by producing noise on the desired channel (which we obviously can't control) or by combining with—and raising—our LO's noise floor (which we can control). These problems go away when our LO goes away.

To summarize the answer to our first question, then, you would want to eliminate mixers, local oscillators and intermediate frequencies because this set of circuits is the root of many common problems in superheterodyne designs. I can imagine some readers saying, "that's not new; receivers with no LO, mixer or IF have been done before." This, of course, is the tuned RF (TRF) receiver that requires tunable filters and has all of its gain over a relatively small RF bandwidth: this was tried early in receiver history and largely discarded. Today its only use in aviation is for the 75 MHz Marker Beacon receivers; TRFs are also found in some remote car door locks and other remote control equipment. The TRF approach has selectivity that varies depending on the tuned frequency, because of the variation of filter Q with frequency, and it's prone to oscillate because large amounts of gain are needed to produce usable audio output (for a human) for the microvolts of RF input.

The conventional approach to the reduction of receiver non-linearities is to (1) plan the frequency mixing scheme to minimize in-band signals, (2) use narrow filters as early as possible in the signal processing to reduce the number of signals that can cause intermodulation distortion, and (3) run high current in the signal amplifiers and large signals in the LO so that signals encountered while operating are small compared to the stage conditions.

We can now begin to answer the second question: how can we eliminate filters? To answer that, we first need to know about the nature of the RF environment, how large the signals are to which we will be exposed, and how to handle them. In which cases will they help, and in which cases are filters simply security blankets that make us feel better? We need to establish just how good we need to be because we can always envision an interference situation that nothing could survive.

One of the most obvious characteristics of the HF spectrum is its noisiness. The HF spectrum (1.8-30 MHz) is over 4 octaves wide, and the noise floor varies from perhaps 90 dB higher than thermal (kTB) noise at 1.5 MHz, down to 40 dB above ther-

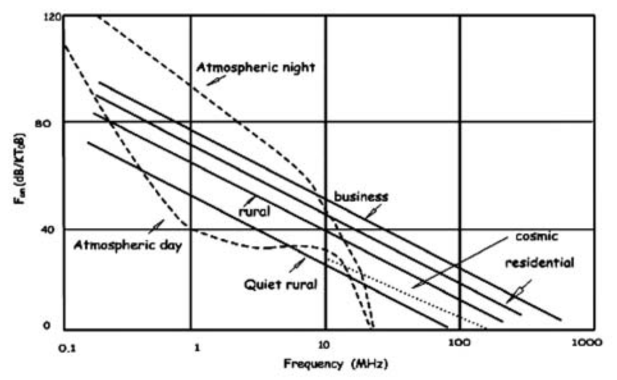

FIGURE 2

Noise levels at HF in dB above kTB—manmade noise classed by population density along with various natural sources. [6]

mal at 30 MHz (see Fig. 2). The atmospheric noise varies from day to night and with the season. The solar noise itself is very variable with season as well as with the solar cycle. There is wide variation due to man-made devices; in an urban environment, the noise may be 40 to 50 dB above the noise in a quiet, rural setting. To generalize from these numbers, noise figure is not generally a difficult design constraint at HF, and a 15 dB NF is quite adequate. Lower NFs can even be troublesome if other parameters are sacrificed for NF. Even on a remote island, it's difficult to think that 15 dB NF is not adequate at 30 MHz. In aviation, the minimum sensitivity requirements allow for a 26 dB NF (1 microvolt in 50 ohms, SSB, 6 dB SNR). It has been decades since sensitivity (or NF) has been an important criterion in choosing an HF radio.

Another obvious characteristic of the HF spectrum is that signals vary over wide amplitude ranges. Different user systems typically specify minimum signals that are met with these sorts of NF values, and maximum strength signals that may vary over a large range as well. We have measured the strength of received signals on typical, full-sized antennas (resonant or nearly-resonant half wave dipole) in typical installations on a number of occasions. We find that signals on the order of –10 dBm are not unusual in a few places across the HF spectrum, usually the lower frequency shortwave broadcasters (the 6 MHz shortwave broadcast band was a notorious example). We find signals at the feed of typical aircraft shunt-fed antennas to be lower than this.

Digital Techniques Meet RF

Much has been written about digital techniques, but to the HF receiver designer, the analog to digital converter is perhaps best thought of as a detector—although one with remarkable properties. The input to the converter can be digitized properly and produce usable audio at input levels far below levels convenient with analog detectors. How far below? Usable output can be obtained without fully turning on one count in the ADC, around 1.0 nanoWatt (–90 dBm) with a 16 bit converter. Experience has shown us that while we can produce a functional receiver with an input of –90 dBm, we achieve better performance with –78 dBm; two full bits above the bottom of the converter's range. Because of the ability to use such low input levels and produce usable output, low enough gain can be used that the receiver is not prone to oscillate, overcoming that objection to the receiver with all of its gain at RF. Once digitized, the inputs are digital signals that can be numerically downconverted to baseband and passed through mathematically ideal filters, overcoming the problems with implementing tunable RF filters or wide range PLL synthesizers to cover the entire HF spectrum with small tuning steps.

Digital signals can be processed in ways we cannot easily implement in the analog domain. One of the strongest advantages of digital filtering is that additional selectivity is almost "free." Adding more poles and zeroes to a digital filter can be as easy as fetching a different equation we already derived, and adding user-selectable filters can be as easy as adding a software procedure. Contrast that to an analog filter where more poles and zeroes mean more components to inventory, more board area and more steps of alignment. In either the analog or DSP domain, more filter complexity will add some time delay, but that's not usually a concern at HF because the propagation delay is so variable. As a result, we can produce a receiver based on a limited amount of RF processing, the A/D and digital processing.

There is much about the digital processing that I will skip over here. Signal decimation, filtering and demodulation are major subjects of their own, and space is limited. I'll just say there is much more beyond the output of the A/D converter that follows the RF circuitry in the block diagram.

Do we need to filter out those large, –10 dBm, signals in the HF spectrum? In general, no. Because these signals don't exceed full-scale at the ADC's input, if they don't cause other distortions, they don't matter. They are filtered out in the DSP. If the signal is a sub-harmonic of our desired frequency, it raises an interesting problem. Our receiver's RF circuitry will generate a harmonic of the input. In the real

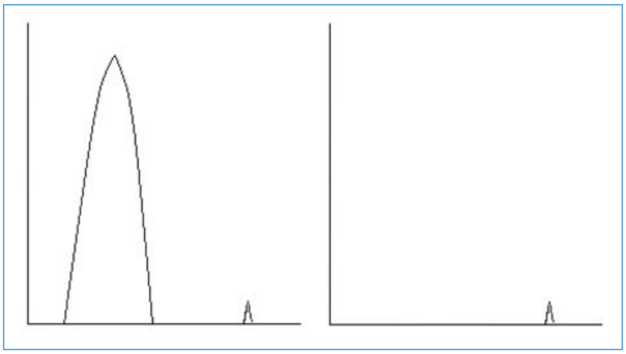

FIGURE 3
How the presence of a transmitter's harmonics sets our design limits.

world, though, the transmitter we're concerned with has its own harmonic. As long as any harmonic we generate is considerably lower than that transmitter's own harmonic, we will not degrade the reception. Figure 3 summarizes the situation graphically.

By taking advantage of some relationships between harmonics and intercept points we can design an amplifier for a high output intercept point, OIP3, and keep its output harmonics below the transmitter harmonics that are also received.

A Proposed Architecture

Before we do that, we need to know more about our receiver. It's time to propose a straw man receiver architecture and determine the largest signals likely to be at its output. Figure 4 is the RF processing portion of our receiver. Compare this to the receiver shown earlier; each of the synthesizers in Figure 1 is roughly equal in component count, power consumption and complexity to this entire RF chain.

The heart of the receiver is the A/D converter. Today, typically it is a 16-bit converter sampled at approximately 100 MHz, which provides 105 dB spurious free dynamic range. This converter has a full scale input equal to a 50 ohm input of +12 dBm. We must decide what level signal we will limit at the input of the ADC, and choose 0 dBm, or –12 dBFS. The reason for this choice is that signals in the HF spectrum can be AM or SSB-AM, and two such signals at –12 dBFS (RMS) can produce peaks at full scale, due to the high peak to average ratio of speech.

We add digital AGC attenuators, to be driven by the DSP portion of our receiver, to keep large signals from exceeding this number. Note that signals at the antenna of –40 dBm will reach 0 dBm at the ADC input due to the cascade gain of the analog chain. We have mentioned that –10 dBm signals are encountered in the HF spectrum. This architecture will handle multiple high level signals to +5 dBm, and individual signals until the ADC fills up.

Unlike a conventional analog receiver, this one needs to distinguish between on-channel and off-channel signals in how it implements its AGC function. Off-channel signals should not engage the front end AGC because that degrades the NF and on-channel signals; therefore, for off-channel signals, the second stage AGC is adjusted first. For on-channel signals, we adjust the first attenuator because if they are that strong, it doesn't matter if we degrade the NF, and adjusting the first attenuator can keep the amplifiers from distorting. It ends up being trivially easy to discriminate on- from off-channel signals in the digital domain.

Remember that this chain is processing the entire 2-30 MHz band. If we are listening to a weak station, then strong off-channel signals at the input of the differential amplifier are not limited, and we can suffer gain compression and other distortions. Gain compression can reduce the amplitude of a threshold signal, making it inaudible in the worst case, so we need an RF chain with a high 1 dB compression point. The decision to engage the off-channel attenuator is made if the highest order bits from the ADC go over a pre-determined threshold: the converter's output is a sample of the instantaneous energy in the entire band. The on-channel signals have been processed through decimation, filtering and demodulation, as in conventional radio architectures. This combination gives the ability to handle the extremely large range of signals typically found at HF.

With these concepts in place, we are now ready to decide how good its OIP3 needs to be in order to not produce more harmonic than the transmitter itself. To do that, we need to determine the suppression of

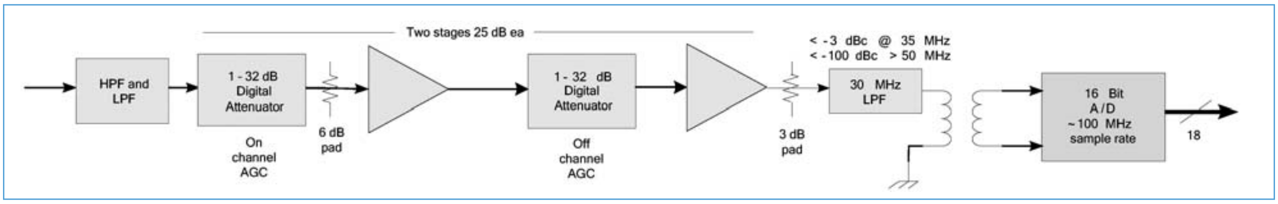

FIGURE 4
Receiver block diagram (RF processing portion).

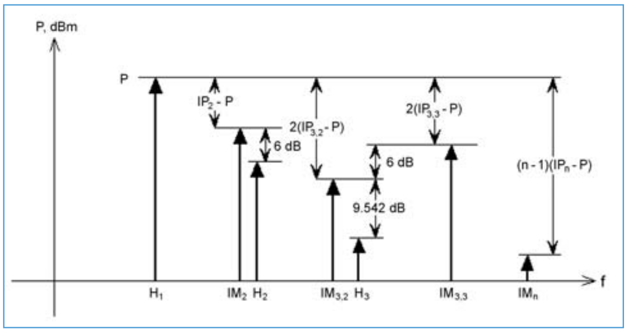

FIGURE 5

Relationship of Harmonics and Intermodulation (IM) Products (5).

the harmonics in the transmitter output: how good are the transmitted signals coming into the receiver? In the US, these requirements are encoded in the FCC regulations and vary somewhat by transmitter type. For most commercial/industrial transmitters, a typical value of attenuation from the fundamental of 43 + 10 log (Py) (where Py is the average power in watts) seems consistently used. This is structured to create a radiated power limit, not an attenuation limit (that is, higher power transmitters attenuate more than lower power transmitters, creating a common power limit at the radiating source). For example, a 1 W transmitter has its harmonics limited to –43 dBc or –13 dBm (30 dBm - 43), and a 1 kW transmitter (30 dB higher power) is required to provide 30 dB more attenuation to limit its harmonics to –13 dBm. Without knowing a distance to transmitter and the frequency, to calculate path loss, an absolute power level is not very helpful to us.

For simplicity and to put a boundary on the problem, let's assume harmonics are limited to –70 dBc. Consider an example in which we're looking at a single strong signal in the RF chain, amplified such that at the input to the converter, the amplitude is 0 dBm (the level above which we'll start reducing gain). In other words, the input signal is amplified and at the input of the ADC, the desired signal is +0 dBm with its third harmonic at –70 dBm. If we want the 3rd harmonic we generate of this input to be 10 dB below the transmitter's contribution, or –80 dBm, what should our OIP3 be (see Fig. 5)?

In general, an amplifier's output harmonics and intermodulation distortion products have established relationships, although phase shifts (FM or PM) can cause these amplitudes to appear differently on a spectrum analyzer. Figure 5 [5] shows that the IM3 product will be 9.54 dB higher—call it 10 dB higher—than the 3rd harmonic we produce. If the transmitter's harmonic is –70 dBm, we want our third harmonic to be at –80 dBm, that 10 dB delta says our IM3 (not our third harmonic!) is –70 dBm, or –70 dBc to our amplifier's output. For our IM3 to be –70 dBc to a +0 output, the OIP3 needs to be 70/2 or 35 dB greater: +35 dBm. Experienced receiver designers will know that an OIP3 of +35 dBm is an aggressive number in a conventional receiver but certainly not impossible. Not having a mixer makes our life easier though, and an amplifier chain with an OIP3 of +35 dBm is not difficult to obtain with modern parts.

The 2nd order intercept is generally quite a bit higher than the 3rd, and isn't usually a concern in systems designed for the IP3 levels we are working with. Fifth and higher order IMD products are weaker and drop off faster as the input decreases, so they are not usually designed for.

There is a special case of intermodulation distortion that receiver designers work with, known as cross modulation. The name comes from the appearance of the modulation of an undesired signal being imposed on the desired signal; in a bench test for example, an adjacent channel might be modulated with a 400 Hz test tone, while the desired channel is modulated with 1000 Hz and the effect is to produce a 400 Hz tone on the desired signal. Numerically, the levels of the cross modulation products are the same as the third order products and cross modulation can be designed for by determining how far below the desired signal the cross modulation may be and calculating the IP3 required. In aviation HF, cross modulation is tested by applying a much stronger signal in a nearby channel (third adjacent), so the analysis performed is for a third order test with unequal tones.

The receiver RF system needs to reject signals in the A/D converter aliasing bands because digital processing is not capable of determining intended from interfering signals that appear to be on the same frequency once they are converted [2, 3, 4]. These bands occur spaced by harmonics of the clock and can provide full strength signals from the system if not filtered properly. The majority of this anti-aliasing filtering is provided by the 30 MHz LPF at the right of the receiver chain. The first low pass filter, top left in the diagram, assists with attaining over 120 dB rejection of alias frequencies.

To a receiver designer, another way to look at A/D aliases is as M*LO ± 1*RF spurious, where M is a harmonic of the LO, the sample clock. For example, for a 100 MHz clock, a valid input for HF receivers is the first Nyquist band, where 5 MHz, appears as 5

MHz. But 105 MHz appears as 5 MHz after sampling as well. So do 205, 305, 405, and so on, or M*LO + 5 MHz. The A/D alias 5 MHz below the sample clock, 95 MHz, appears as 5 MHz, as do 195, 295, 395, etc., or M*LO – 5MHz.

I have tested A/D converters for the strength of this response, and it degrades very slowly, much more slowly than for a conventional analog double-balanced mixer. Extremely good filters are needed for alias suppression, with good shielding on the PWB. In any realistic HF installation, the antenna and coupler (antenna tuning network, if used) can offer additional rejection of possible alias frequencies and help keep them out of the receiver. On the A/D converter data sheets you'll find simple driver circuits, where there is typically only a simple RC or one pole LC filter. These are usable in only the most casual receivers, and you won't be using a converter like these for casual receivers. These filters are useful for limiting broadband noise, but will not be adequate for alias rejection.

This architecture brings immediate benefit to aeronautical HF communications systems. They can receive multiple channels simultaneously, the number limited by the processor throughput in the system. Civil Aeronautical HF communications have both voice and data networks operating on different frequencies throughout the HF band. Rather than needing two dedicated HF transceivers, one for voice and one for data, one direct sampled transceiver can simultaneously receive on widely different voice and data frequencies with no loss of performance. How much of an improvement is available? A transceiver built using this approach, also using modern, 6th generation LDMOS FETs for the power amplifier, is roughly 2/3 the volume, 1/2 the weight and 1/3 the power of previous generation transceivers.

Lessons Learned

The development of this receiver proceeded as analyzed, with one exception. The A/D converter generates harmonics of its input, and its own harmonics are substantially worse than the ones provided by the RF chain. They don't behave the same way as harmonics in an analog system, which increase with increasing fundamental signal level; rather, they behave non-monotonically as the input increases and can actually be observed going up and down with a lab spectrum analyzer connected to the A/D's input. With careful attention to levels and control of the signals in the digital domains, we can manage those harmonics. In some extreme cases, a bank of semi-octave high pass filters that reduce the fundamental by at least 40 dB would remove all traces of the harmonics, at the expense of increasing size, weight, power and cost.

I made reference earlier to situations where filters are security blankets. The discussions here have shown a way to think about the problem, which is that as long as we don't generate products that are worse than those the environment itself brings, we get no benefit from the filters. Part of that is managing the level of undesired signals with AGC, but the essential requirement is a very linear RF chain. As children, we gave up our security blankets when we realized there really were no monsters under the bed. As engineers we can do no better than give up our security blanket filters as we realize they offer no benefit a well-designed system doesn't offer. This is the dawn of a new age, where we can replace large, expensive, tricky-to-align, analog/RF filters with RF amplifier linearity.

References

1. Desoto, Clinton, *200 Meters and Down—the Story of Amateur Radio*, American Radio Relay League, 1936, 1981, 2001.

2. Brannon, Brad, "Basics of Designing a Digital Radio Receiver (Radio 101)," Analog Devices, Inc. App Note, http://www.analog.com/static/imported-files/tech_articles/480501640radio101.pdf

3. Richard Groshong and Stephen Ruscak, "Undersampling Techniques Simplify Digital Radio," *Electronic Design,* May 23, 1991, pg. 67-78.

4. Kester, Walt, Editor, *Mixed Signal Design and DSP Techniques*, Analog Devices, 2000; available as text or as part of an online book: http://www.analog.com/library/analogdialogue/archives/39-06/mixed_signal.zip

5. Agilent Advanced Design System, "Intermodulation and Harmonic Basics," online at http://eda-download.soco.agilent.com/docs/genesys2006_10/Sim/SPECTRASYS/Advanced/Intermod_Basics.htm

6. Bianchi, Cesidio and Meloni, Antonio, June 2007, "Natural and Man-Made Terrestrial Electromagnetic Noise: An Outlook," *Annals of Geophysics,* Vol. 50, No. 3.

ABOUT THE AUTHOR

Bob Lombardi is a Principal Engineer in the Commercial Systems division of Rockwell Collins, where he has been a PLL specialist, receiver designer, and RF/analog engineer for over 15 years. He is a 1988 graduate of the Florida Institute of Technology. Bob can be reached at: rlombard@rockwellcollins.com

Some comments by Bob Lombardi W4ATM, the designer: (21.10.2015)

A little more background. When we first started, we were all concerned about that issue that "all the power in the world" (or between our 2 MHz HP and 30 MHz LP cutoffs) was going to hit the ADC, so we needed switchable bandpass preselectors. The first block diagrams had that, as did the first prototypes. Then we did some experiments. We used an
SDR-14 (RFSpace's first high quality band sampler) and put it on a bunch of different antennas, first here in Melbourne and then Cedar Rapids and a couple of places around the country. We found, by and large, that the power in the HF band was fairly low until you hit a few SW BCBs, notably
6 and 7 MHz where you'd find a station, or a few of them at -10 dBm. One of our guys got the enviable assignment to record the antenna input during a test flight of the A380 over continental Europe - and then out to sea. As a bonus, he asked the pilots if they knew of any SW stations they could fly near, and they did. We all said that maybe things could be worse during a sunspot peak, but then we figured it would probably move the problem up in frequency, but the magnitude would stay the same. This got us almost all the way to dropping preselectors, but not quite. I needed to do some more analysis.

The gist of that analysis is in the article, but basically we figured out how bad the signals coming into the aircraft were likely to be; that is, the transmitters' spurious and harmonics, and figured out how good the system had to be in order to not produce anything worse than what was coming from the outside world. Then we made it better than that.
To get certification on the radios, we do a set of industry standard tests for intermod, crossmod, spurious and lots of other things. It was derived that a 3rd order OIP3 of -6 would pass those tests. Our previous generation radios were +11. This radio is +41.

But before I leave the preselector topic, don't forget this is an HF for a civil aircraft. They don't (couldn't) carry very broadband antennas; they carry something that an antenna tuner can make look like an acceptable load, making a high Q system. There's a potent preselector before anything gets to the radio.

You will note that anti-aliasing filter has over 100 dB attenuation once you get to the clock/2 boundary, so things in the 65 MHz band just ain't getting in. Plus, our IP2 is over +100 dBm, so that IM product is going to be pretty small. The entire HF spectrum is open, however, so it is conceivable a plane could be listening on one frequency and be near two absurdly strong transmitters that produce an on-channel intermod.

That AGC trick came about because most of our radios use a delayed AGC, where the back end is attenuated before the front end. A few of us came up with the approach one day while sitting around discussing AGCs. The DSP guy was this really bright new grad, maybe two years out of college. We could propose some idea like that and he'd bang it out in a few minutes. All in all a very good team.

My boss and I always thought that our biggest mistake was not putting a USB interface onto the radio so that we could use them at home. It would have been relatively easy for the DSP guy to code the CI-V command set, and we could use an Icom controller to run the radio! I had wanted to take it home during a DX contest on 40, when it's wall-to-wall big signals in between the Euro broadcasters, but I had no way to run it here.