Suno v5 vs v5.5: What Actually Changed in the Audio?

15 metrics · 5 genres · Real spectral data

We measured it so you don't have to guess.

By Petri Korhonen · March 2026 · ~15 min read
masterforge.app/blog/suno-v5-vs-v5-5/

5 Tracks	**Identical Prompts**	**scipy STFT**	**15 Metrics**	**M/S Analysis**
Piano → Metal	v5 & v5.5	nperseg=4096	Per track	Stereo decomp

Introduction

Suno v5.5 arrived with Voices, Custom Models, and My Taste. The community consensus was immediate: "it sounds better." Forums filled with praise. Comparisons were based on vibes, memory, and enthusiasm.

We wanted something more concrete. So we generated the same track on both versions — five times, across five genres — using identical prompts, identical settings, and zero seeds. Then we ran spectral analysis on every pair: 15 metrics per track, scipy STFT, middle 60% of each clip, M/S decomposition.

The result? v5.5 is not simply "cleaner." It's doing something far more interesting. It's adapting its output to match the genre — making intelligent decisions about which frequencies to emphasize and which to discard. But it also introduces new challenges that your mastering chain needs to handle.

This is what 15 metrics and 5 genres actually tell you.

How We Tested

We generated five tracks spanning a wide range of complexity and genre conventions: Glass River (piano ballad, 2 voices, 72 BPM), District (indie rock, 3-4 voices, 128 BPM), Iron Doctrine (thrash metal, 5+ voices, 168 BPM), Reactor Core (hard techno, 3-4 voices, 148 BPM), and Count the Days (dark hip-hop, 2-3 voices, 85 BPM).

For each track, we used identical style prompts, exclude lists, Weirdness, and Style settings on both v5 and v5.5. No seeds — we wanted to test the model's own generation capability. Four generations were made per version, and the closest-matching pair was selected for analysis.
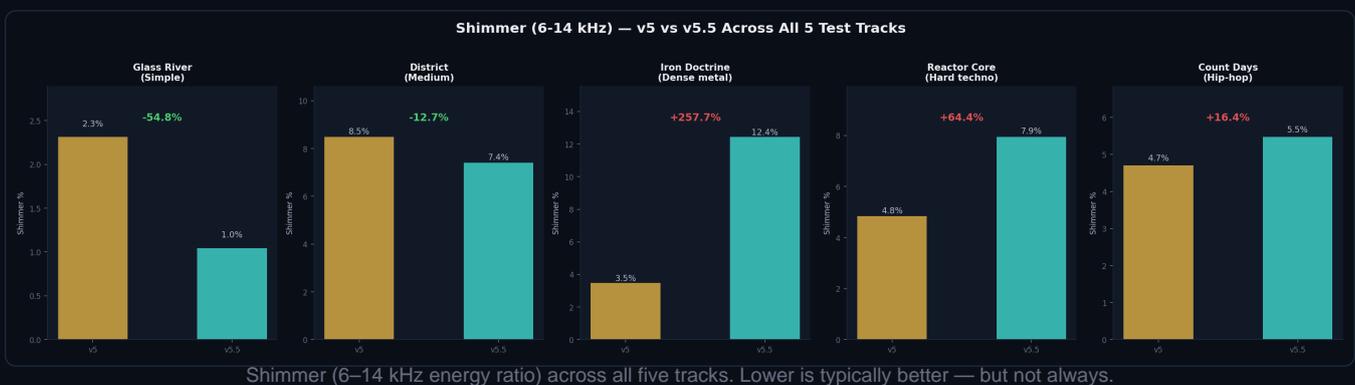
> **IMPORTANT LIMITATION**
>
> Different generations are different performances. These are not controlled A/B codec tests — they're two AI performances of the same brief. Arrangement differences between generations are expected and documented. We compensate by analyzing spectral characteristics (which reflect model behavior) rather than waveform alignment (which reflects arrangement).

Analysis pipeline: Python 3, scipy.signal.stft with nperseg=4096 (~11.7 Hz frequency resolution at 48 kHz), middle 60% of each clip to avoid intro/outro artifacts, and M/S decomposition for stereo analysis. Same methodology as our Guide A artifact analysis.

The Simple Story (And Why It's Wrong)

If you only looked at shimmer — the metallic high-frequency artifact we measured in Guide A — you might conclude that v5.5 is a mixed bag. Two tracks improved, three got worse:

Shimmer (6-14 kHz) — v5 vs v5.5 Across All 5 Test Tracks

| Glass River (Simple) | District (Medium) | Iron Doctrine (Dense metal) | Reactor Core (Hard techno) | Count Days (Hip-hop) |

-54.8% · 2.3% / 1.0% — -12.7% · 8.5% / 7.4% — +257.7% · 3.5% / 12.4% — +64.4% · 4.8% / 7.9% — +16.4% · 4.7% / 5.5%

Shimmer (6–14 kHz energy ratio) across all five tracks. Lower is typically better — but not always.

Glass River's shimmer dropped 54.8% — a massive cleanup. District improved by 12.7%. But Iron Doctrine's shimmer increased 258%, Reactor Core went up 64%, and Count the Days rose 16%.

When Better Numbers Mean Worse Sound

Here is the single most important finding in this entire analysis.

Iron Doctrine is a dense thrash metal track: 5+ simultaneous voices, 168 BPM, dual distorted guitars, double bass drum, screamed vocals. On v5, the shimmer metric read 3.48% — suspiciously low for the densest track in our test. For context, the simpler District (3-4 voices) measured 8.49%.

How can denser metal be "cleaner" than mid-complexity rock?

> **THE ANSWER**
>
> v5 didn't clean up the high frequencies. It killed them. The v5 spectrum drops off sharply above 2 kHz, losing 10-15 dB compared to v5.5 across the entire 2-14 kHz range. v5 "solved" shimmer by making the metal sound muffled — removing artifacts AND legitimate high-frequency content together. Pick attack, cymbal presence, vocal edge — all gone.
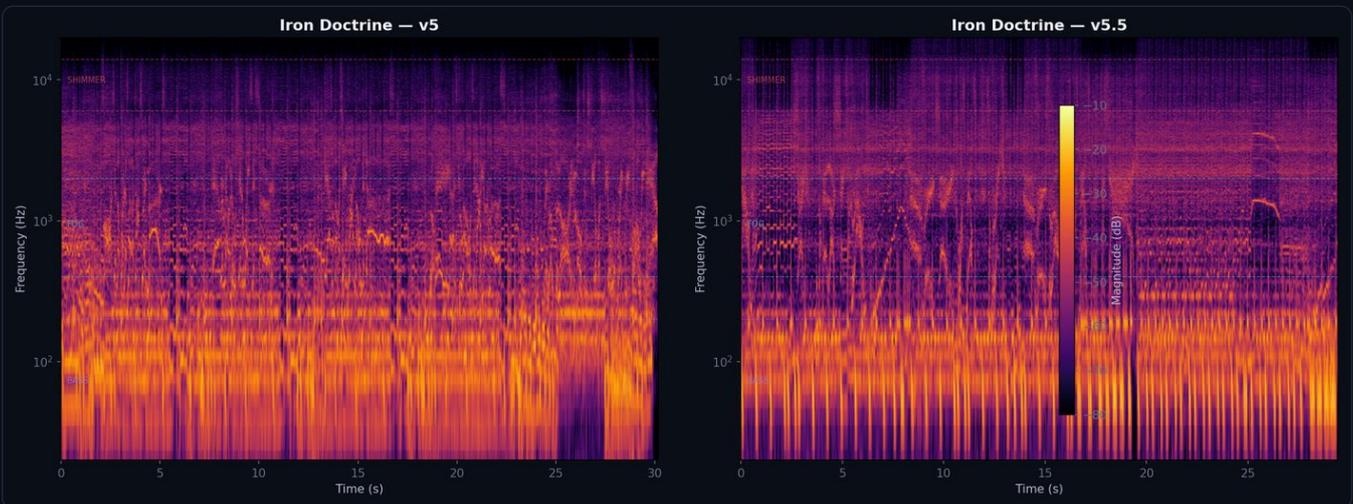
Iron Doctrine spectrograms. Left: v5 (dark above 2 kHz — suppressed HF). Right: v5.5 (visible energy across full range — restored HF).
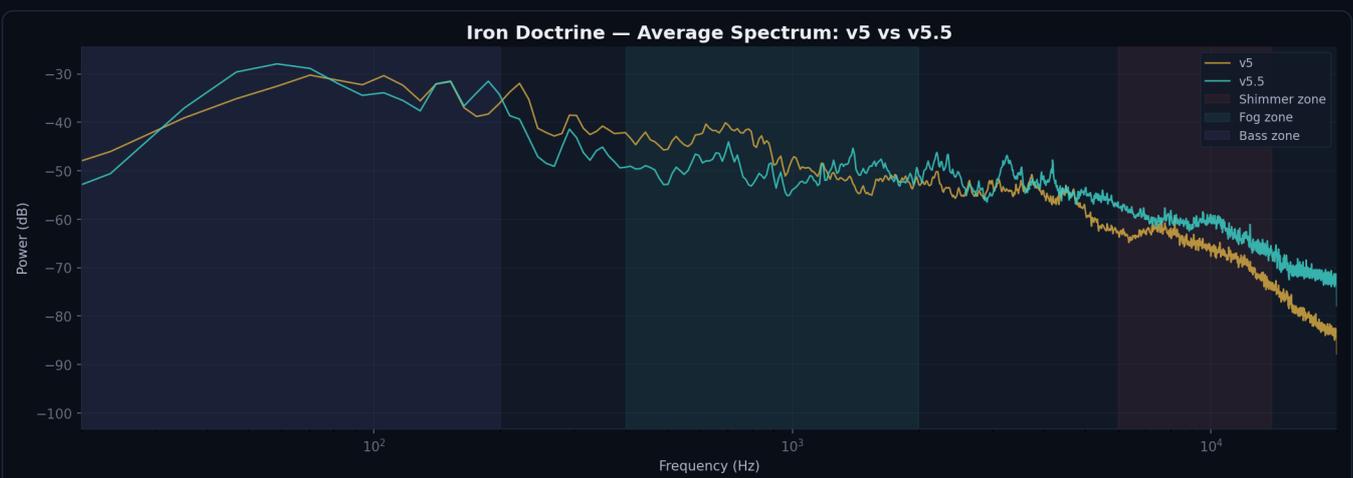
Iron Doctrine average frequency spectrum. The two curves cross at ~2 kHz. Below: v5 has more energy (mid-range mud). Above: v5.5 restores the full high-frequency range.

v5.5's shimmer reads 12.45% — nearly four times higher. But when we put on headphones, the v5.5 version was dramatically better. More detail, better bass definition, more aggressive — actually sounding like metal instead of metal heard through a pillow.

This is the blog's core argument: a metric alone doesn't tell you quality. The shimmer number measures energy in 6-14 kHz, but it cannot distinguish between codec artifact noise and legitimate musical content. In genres with naturally high HF energy — metal, electronic, bright pop — a higher shimmer reading may actually mean the model is doing its job better.

This is why "measurement + listening" beats both "just vibes" and "just numbers."

The Real Story: v5.5 Is Genre-Adaptive

Once we looked past shimmer alone, a clear pattern emerged across all five tracks. v5.5 doesn't apply a single processing change to everything. It adapts its spectral profile to match genre conventions.

Spectral centroid tells the story

The spectral centroid — the "center of gravity" of a track's frequency content — moved in opposite directions depending on genre:

TRACK	GENRE	CENTROID V5	CENTROID V5.5	DIRECTION
Glass River	Piano ballad	2,414 Hz	1,809 Hz	-25% darker
District	Indie rock	4,381 Hz	3,781 Hz	-14% warmer
Iron Doctrine	Thrash metal	3,578 Hz	5,157 Hz	+44% brighter
Reactor Core	Hard techno	3,909 Hz	4,499 Hz	+15% brighter
Count the Days	Dark hip-hop	3,264 Hz	3,817 Hz	+17% brighter

Piano ballad gets warmer. Rock gets warmer. Metal gets brighter and more aggressive. These are exactly the directions a human mix engineer would take each genre. v5.5 isn't just applying a blanket EQ — it's making genre-informed spectral decisions.

The Memphis proof

The strongest single piece of evidence came from Count the Days, our dark hip-hop track. Given the same prompt, v5 generated a modern dark trap production. v5.5 generated a Memphis / Three 6 Mafia-style interpretation — a fundamentally different sound world from the same words.

Count the Days — Average Spectrum: v5 vs v5.5
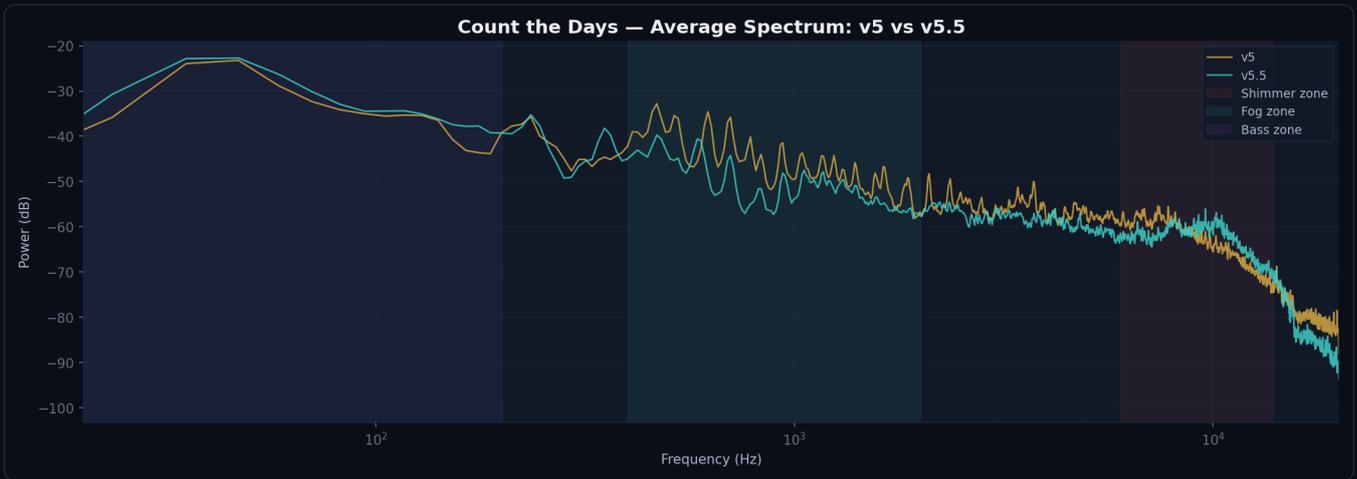
Count the Days spectrum overlay. v5.5's Memphis interpretation has dramatically more sub-bass energy and a different spectral profile than v5's modern trap version.

The numbers confirm the listening experience: sub-bass increased 56.5%, presence dropped 53.9%, stereo width collapsed to near-mono (-85.9%), and dynamic range exploded by 75.5% (15.1 -> 26.5 dB). Every one of these changes is genre-authentic for Memphis hip-hop: heavy 808s, dark frequency balance, mono imaging, empty space between hits.

What Actually Improved Across the Board

Dynamic range: 4 out of 5 tracks

The most consistent genuine improvement was in dynamic range. v5.5 generates music with more breathing room — less aggressive compression, more natural transients.

Dynamic Range (dB) — v5 vs v5.5 Across All 5 Test Tracks
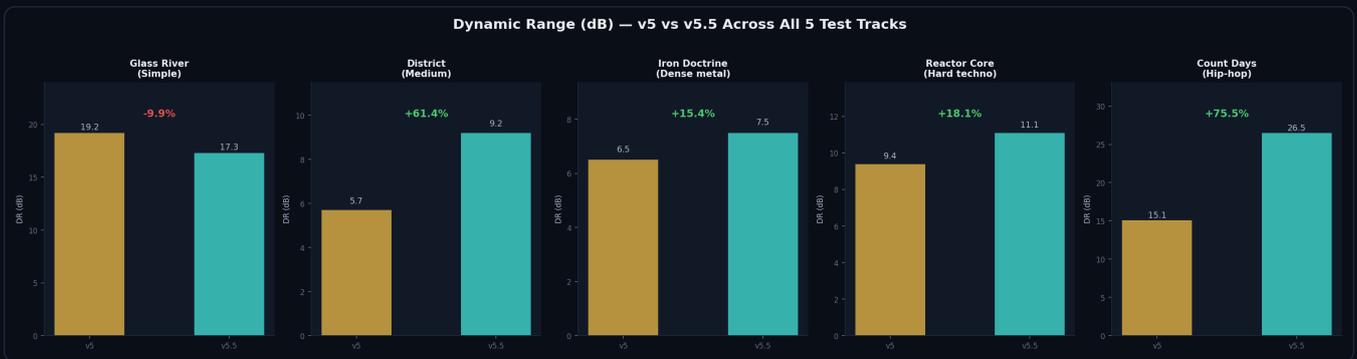
Dynamic range (P95–P5 of RMS frames) across all five tracks. v5.5 improved dynamics in four out of five cases.

TRACK	DR V5	DR V5.5	CHANGE
Glass River	19.2 dB	17.3 dB	-10.1%
District	5.7 dB	9.2 dB	+61.0%
Iron Doctrine	6.5 dB	7.5 dB	+15.4%
Reactor Core	9.4 dB	11.1 dB	+18.1%
Count the Days	15.1 dB	26.5 dB	+75.5%

For mastering, this means less work fighting over-compressed source material. v5.5 tracks have more natural dynamics that respond better to limiting and loudness optimization.

Bass leak improved in electronic and metal genres

TRACK	GENRE	BASS LEAK V5	BASS LEAK V5.5	CHANGE
Iron Doctrine	Metal	0.260	0.114	-56.1%
Reactor Core	Techno	0.003	0.002	-37.8%
Count the Days	Hip-hop	0.004	0.003	-21.7%
District	Rock	0.043	0.046	+6.4%
Glass River	Piano	0.199	0.390	+96.0%

v5.5 keeps bass centered for genres where mono compatibility matters (club, PA systems, phone speakers). The Glass River exception may reflect more realistic piano stereo imaging in the low register.
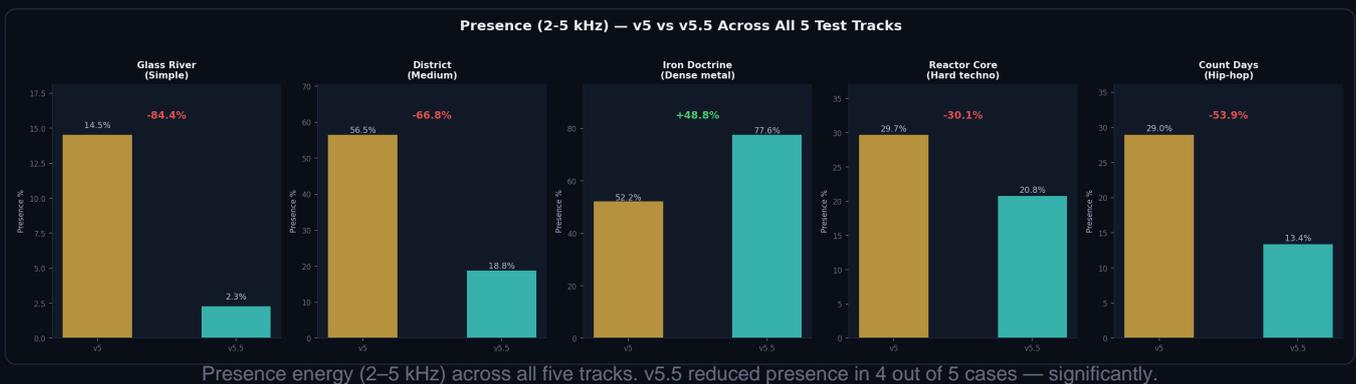
What Didn't Change (And What Got Worse)

Fog remains: 3 worse, 1 unchanged, 1 improved

Fog — spectral flatness in the 400-2 kHz range, the "muddy blanket" over the mid-range — was not addressed by v5.5. In fact, most tracks got slightly foggier:

TRACK	FOG V5	FOG V5.5	CHANGE
Glass River	0.037	0.038	+1.9%
District	0.107	0.126	+17.6%
Iron Doctrine	0.211	0.302	+43.0%
Count the Days	0.088	0.108	+23.9%
Reactor Core	0.339	0.301	-11.1%

If your track sounds muddy in the mid-range, v5.5 won't fix it. Arrangement is still the primary defense against fog.

Presence consistently reduced: 4 out of 5 tracks

Presence energy (2–5 kHz) across all five tracks. v5.5 reduced presence in 4 out of 5 cases — significantly.

v5.5 systematically pulls back 2-5 kHz energy — the vocal presence and clarity range. Glass River: -84.4%. District: -66.8%. Count the Days: -53.9%. Reactor Core: -30.1%. Only Iron Doctrine got a presence boost (+48.8%) — the genre that needs aggression.

This is the single most consistent spectral change across all five tracks. Your mastering chain likely needs a presence boost around 2-5 kHz when working with v5.5 material.

Sub-bass increased in 4 out of 5 tracks

v5.5 produces significantly more low-end weight in most genres: Glass River +192%, Iron Doctrine +123%, Count the Days +57%, District +33%. Only Reactor Core saw a decrease (-27%).

More sub-bass means more energy fighting for headroom during mastering. High-pass filtering and careful low-end management become more important with v5.5 sources.

Stereo: narrower by default, but prompt-responsive

Four out of five tracks came out narrower on v5.5 — some dramatically so (Count the Days: -85.9%, Glass River: -55.4%, Iron Doctrine: -50.9%). The single exception was District, whose prompt explicitly requested "wide stereo image" — and v5.5 delivered (+41.9%).

This suggests v5.5 defaults to tighter stereo imaging but responds better to stereo instructions in the style prompt. If you want width, ask for it explicitly.

The Cleanest Improvement: Glass River

For a straightforward "did it get better" story, Glass River is the clearest example. A simple piano ballad with two voices — the kind of track where codec artifacts are most audible.
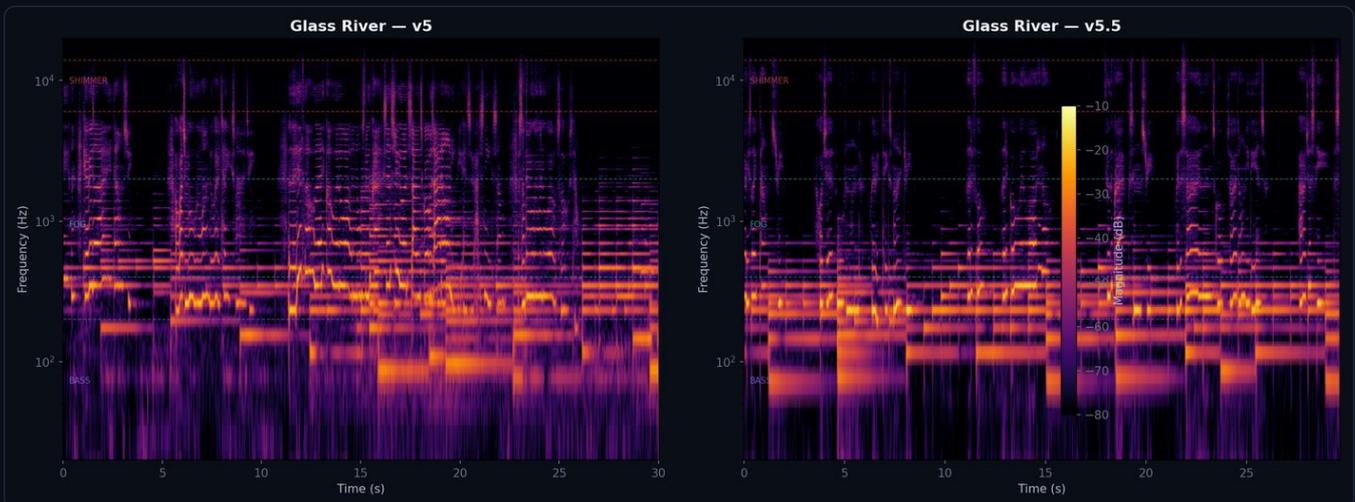
Glass River spectrograms. v5 (left) shows diffuse energy above 6 kHz — shimmer haze. v5.5 (right) has cleaner harmonic separation.
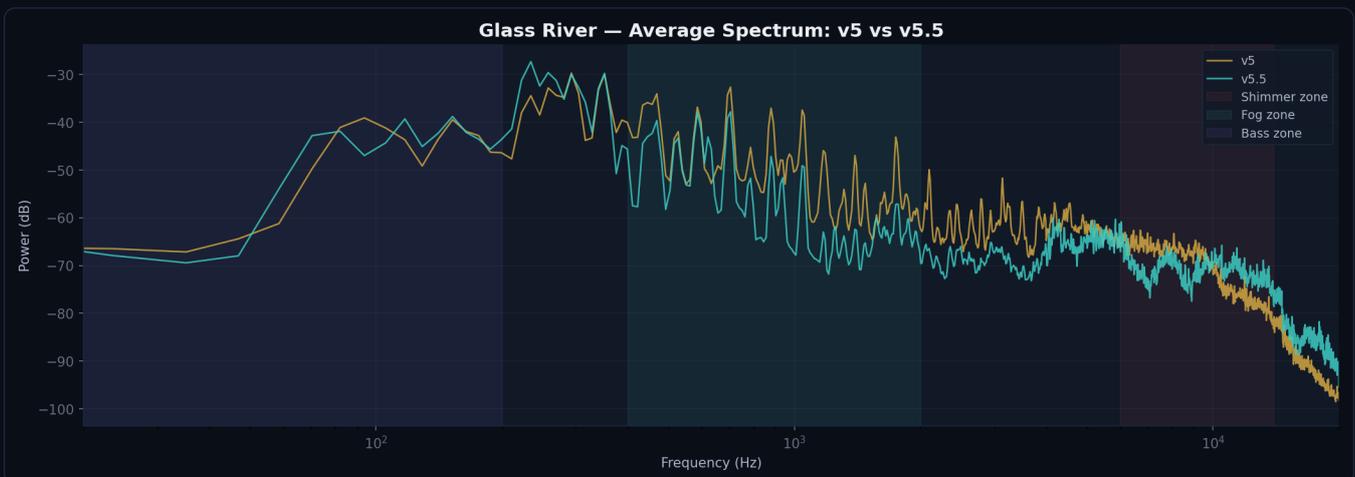
Glass River average spectrum. v5 is 10+ dB louder in the 2–5 kHz presence range. v5.5 is warmer, with more sub-bass and less brightness.

The data confirms the ears: shimmer dropped 54.8% (2.31% to 1.05%), which is actually below the 1.5% we measured on our Guide A reference track Resonance. v5.5 piano ballads are approaching "mastered" levels of high-frequency cleanliness straight out of the generator.

Open Questions

District's quality degradation: The v5.5 version of District (indie rock) sounded noticeably worse toward the end of the track — more artifacts, more hiss. Our middle-60% analysis didn't capture this. Does v5.5's codec "budget" run out in long, dense arrangements? This needs further investigation.

Reactor Core's minimal improvement: Hard techno showed the smallest subjective difference between v5 and v5.5. Techno may be the genre where v5 to v5.5 matters least — or synthetic signals were already well-served by v5's codec.

What This Means for Your Mastering Chain

If you're mastering v5.5 material, your processing needs to adapt. Based on our five-track analysis:

1 **Boost presence (2-5 kHz)**

v5.5 pulls back the clarity range in 4 out of 5 genres. A gentle shelf or bell boost around 3 kHz will restore vocal presence and instrument definition.

2 **Manage sub-bass**

More low-end energy means more headroom competition. High-pass filtering below 30-40 Hz and careful sub-bass compression will keep the bottom end tight without sacrificing weight.

3 **Check stereo width**

v5.5 defaults to narrower stereo imaging. If your track feels too centered, a mid-side widener on the high frequencies can open it back up. But check mono compatibility first.

4 **Don't trust shimmer numbers blindly**

In bright genres (metal, electronic), higher shimmer may mean the model preserved legitimate high-frequency content that v5 was suppressing. Listen before you de-ess.

5 **Leverage the better dynamics**

v5.5's improved dynamic range means less need for dynamic expansion. Your limiter can work more naturally with source material that already breathes.

Honest Conclusion

v5.5 is a genuine improvement over v5 — but not in the way most people think. It's not just "cleaner audio." It's a smarter model that makes genre-informed decisions about spectral balance, dynamics, and stereo imaging.

For simple, acoustic arrangements (ballads, singer-songwriter, minimal productions), v5.5 delivers a clear and measurable quality improvement. Less shimmer, cleaner backgrounds, fuller sound.

For dense, bright genres (metal, aggressive electronic), v5.5 restores high-frequency content that v5 was suppressing.

What v5.5 doesn't fix: fog in the mid-range. If your track sounds muddy between 400 Hz and 2 kHz, the answer is still arrangement — fewer simultaneous voices, better frequency separation between instruments. The codec's bitrate budget is still finite, and v5.5 spends it more wisely but doesn't expand it.

> **THE BOTTOM LINE**
>
> Upgrade to v5.5 — it's better. But don't expect magic. The rules from our Guide A still apply: arrangement drives quality. v5.5 makes smarter decisions with what you give it, but what you give it still matters most.

Appendix: Complete Metrics

Core metrics across all 5 tracks

TRACK	VOICES	BPM	SHIMMER V5	SHIMMER V5.5	DELTA
Glass River	2	72	2.31%	1.05%	-54.8%
District	3-4	128	8.49%	7.41%	-12.7%
Iron Doctrine	5+	168	3.48%	12.45%	+257.7%
Reactor Core	3-4	148	4.83%	7.95%	+64.4%
Count the Days	2-3	85	4.71%	5.48%	+16.4%

Dynamics and stereo

TRACK	DR V5	DR V5.5	DELTA	WIDTH V5	WIDTH V5.5	DELTA
Glass River	19.2 dB	17.3 dB	-10%	0.251	0.112	-55%
District	5.7 dB	9.2 dB	+61%	0.123	0.175	+42%
Iron Doctrine	6.5 dB	7.5 dB	+15%	0.460	0.226	-51%
Reactor Core	9.4 dB	11.1 dB	+18%	0.021	0.014	-31%
Count the Days	15.1 dB	26.5 dB	+76%	0.057	0.008	-86%

Ready to Master Your AI Tracks?

MasterForge was built for AI-generated music — with tools designed specifically for the artifacts, dynamics, and spectral characteristics of neural codec audio.

masterforge.app

Before You Master — Guide Series