Independent Statistics and Analysis
U.S. Energy Information Administration

April 2026

AEO2026 Fact Sheet: EIA Improves Hydrocarbon Production Data Transparency and Precision

The Hydrocarbon Supply Module (HSM), which we first deployed only one year ago in AEO2025, underwent a major modernization of its Lower 48 onshore data preprocessing workflow during model development for the *Annual Energy Outlook 2026* (AEO2026). This update was a critical effort: Lower 48 production accounts for over 80% of total U.S. crude oil output and significantly influences hydrocarbon projections.

Background and previous methodology

The AEO2025 version of HSM relied on two legacy SAS-based preprocessors (leftover from the legacy Oil and Gas Supply Module) that generated the core assumptions for currently producing wells and future wells in known continuous formations.[1] The legacy preprocessors lacked documentation, relied on hard-to-maintain code and static input files, required manual updates, and provided limited ability to validate key inputs—particularly for Lower 48 onshore projects and production on federally owned lands, a growing source of hydrocarbon production.

New methodology

For AEO2026, we rebuilt the Lower 48 onshore preprocessing workflow using a modernized architecture based on a PostgreSQL database and a Python preprocessing module. The system was re-engineered end-to-end, incorporating documentation, a structured update process, and improved validation workflows. We also implemented geospatial mapping using the U.S. Geological Survey Protected Areas Database to estimate production shares on federally owned lands, replacing the previous insufficient identification method.

Impact to projections

The modernized data processing pipeline significantly improves transparency, reproducibility and validation as the new pre-processing tool supports the publication of projections for Lower 48 onshore crude and natural gas production on federal and non-federal lands. Through clearer outputs, structured documentation, and faster runtime, we now have enhanced analytic capabilities. Although granular production data existed previously, the introduction of county-level identifiers provides essential geographic context, enabling a deeper understanding of assumptions and more accurate validation of results. This enhanced geospatial precision

[1] Continuous formations are widespread hydrocarbon accumulations in low-permeability rock (for example, shale and tight sands). Unlike conventional reservoirs, they require advanced techniques like hydraulic fracturing to produce and account for most of the growth in U.S. oil and natural gas production over the past decade.

materially improves our ability to project hydrocarbon production, particularly on federally owned lands, for AEO2026 and future outlooks.

Running the data processing pipeline is the first step in preparing HSM for each AEO. Although previous AEOs used the old methodology for this step, for AEO2026, we only ran the data processing pipeline using the new methodology that provided enhanced geospatial mappings. As a result, we cannot say how the results from the new methodology compare to the old.