Refiners that rely only on FCC feed hydrotreating or mild hydrocracking1,2 (i.e., use a cat feed hydrotreater [CFHT]) to meet 30-ppmw gasoline pool sulfur specifications must minimize CFHT feed contaminants and ensure that their FCC units have the flexibility to undercut gasoline to meet the sulfur limit.

Refiners that use a CFHT to produce low-sulfur gasoline must balance the degree of desulfurization and run length and closely monitor gasoline end point. Unplanned CFHT shutdowns, high CFHT product sulfur, and the inability to undercut the high-sulfur fraction of FCC gasoline will limit these refiners’ ability to blend low-sulfur gasoline consistently.

Every 100-ppm increase in CFHT product sulfur raises FCC gasoline sulfur by about 10 ppm if a refiner is producing full boiling range (FBR) 80-430°F gasoline. Once the gasoline pool reaches its sulfur specification limit, the refiner must reduce FCC gasoline yield or shut down the CFHT to replace deactivated catalyst.

Upstream distillation units feeding the CFHT must minimize microcarbon residue (MCR), nickel, vanadium, and asphaltenes to optimize CFHT run-length and desulfurization. FCC main fractionator and associated equipment should be flexible enough to maximize fractionation and reduce FCC gasoline yield when the CFHT produces a high-sulfur product.

It is more cost effective for refiners to correct inadequate distillation unit performance upstream of the CFHT and have the flexibility to undercut gasoline compared to a CFHT design based on high-contaminant feed resulting from poor distillation unit performance.

**Refiners must optimize FCC feed hydrotreating when producing low-sulfur gasoline**

**CFHT feed quality**

The CFHT must operate reliably for the targeted run-length and must maintain the FCC unit feed sulfur levels to meet the pool limits. CFHT run-length depends on many variables, including feed contaminants.

The upstream distillation unit process design, equipment design and operation, and crude oil blend properties determine contaminant levels. Poor upstream equipment design, even with a CFHT MCR of 0.5 wt % and total metals less than 1 ppmw, can easily increase CFHT feed contaminants by 50-100% compared to the design value.

Cost-effective upstream process and equipment design improvements can
minimize CFHT feed contaminant levels. For example, vacuum-unit revamps have reduced heavy vacuum gas oil (HVGO) product vanadium more than 30% given the same crude charge rate, crude blend, and HVGO product yield.

Coking and fouling are common in CFHT distillation units; operators must eliminate coking and fouling to reliably control CFHT feed contaminants during a typical 4-5 year crude unit run. Once the CFHT approaches end-of-run, FCC feed and gasoline product sulfur will increase rapidly. FCC main fractionator improvements that lower the gasoline distillation “tail” will reduce the FBR gasoline sulfur content.

**FCC product sulfur**

Fig. 1 shows the sulfur distribution in a FBR gasoline sample produced from a severely hydrotreated feed. Sulfur distribution varies with hydrocarbon true boiling point (TBP) boiling temperatures. Whereas FCC gasoline product sulfur varies with feed sulfur content, the cumulative total sulfur is a function of hydrocarbon boiling temperature.

Fig. 2 shows that total sulfur increases linearly in the 225-390°F TBP boiling range. Sulfur distribution is nearly constant throughout the same boiling range.

At greater than about 390°F TBP, sulfur begins to increase rapidly due to the presence of benzothiophenes. Fractionation improvements that reduce the amount of high boiling-point hydrocarbons can therefore materially lower FCC gasoline sulfur content at the same product yield.

CFHT product sulfur must be low enough so that a refiner can blend the resultant FCC gasoline and still meet the 30-ppmw maximum sulfur specification. For example, a decrease in CFHT desulfurization from 98% to 97% can increase FCC feed sulfur from 550 to 780 ppmw. This would raise the FCC gasoline sulfur from 45 to 75 ppm.

A typical FCC gasoline will have about 50% of the sulfur in the heaviest 15 vol % of its FBR. Reducing sulfur content from 75 to 50 ppmw requires undercutting about 8-10% of the FBR gasoline. Many refiners cannot undercut 10% of the gasoline and maintain feed rate because the main fractionator and associated equipment are not designed with this flexibility.

FCC gasoline and light cycle oil (LCO) product fractionation influences sulfur content because the gasoline 95-end point (EP) tail is a function of the main fractionator heat balance and column internals design. Some units produce gasoline with a 25°F tail while others have a 65°F tail. Improved fractionation reduces this tail, which lowers the amount of sulfur.
Poor fractionation also shifts light gasoline into the LCO product; some LCO product samples can contain 5 vol % of <400° F. material. Fig. 1 shows that these hydrocarbons have dramatically lower sulfur than the 400°+ F. portion. Refiners that rely solely on CFHT to meet sulfur specifications must consider improved FCC main fractionator performance to help manage CFHT run-lengths.

**CFHT product, FCC gasoline sulfur**

FCC gasoline product sulfur must be 50 ppmw or less to meet low-sulfur gasoline specifications. CFHT product sulfur must be 550 ppmw or less to make 50-ppmw sulfur FCC gasoline.

Start-of-run (SOR) and end-of-run (EOR) product sulfur content will inevitably vary with operating conditions, feed conditions, etc. Large variations cannot be tolerated after 30-ppmw pool specifications take effect. Yet, SOR to EOR increases of 100% are common.

In one refinery, the FCC gasoline product sulfur ranged from 150 to 350 ppmw for SOR and EOR CFHT conditions, respectively. The corresponding CFHT SOR and EOR product sulfur ranged from 1,800 to 4,000 ppmw, respectively.

Increases or intermittent excursions in feed contaminants reduce CFHT run-length or increase the CFHT capital investment needed to treat the contaminants.

Table 1 shows predicted and actual CFHT feedstock vanadium levels associated with high-metals crudes. In this case, vanadium in the CFHT feed was more than three times the design level, resulting in rapid catalyst poisoning and deactivation, which reduced the run-length to less than half of what was originally anticipated.

Shutting down the CFHT is a major
logistical problem when low-sulfur gasoline is required. The value of high-sulfur gasoline depends on other refiners’ excess capacity in their gasoline sulfur-removal units.

Ultimately, upstream process and equipment performance determine CFHT feed contaminants levels (Fig. 3). Using crude assay data alone to predict CFHT feed properties does not account for actual distillation unit performance.

As refiners process heavier crude oils containing highly volatile vanadium and nickel, and high asphaltene and MCR residues, distillation unit performance will play an increasingly important role in meeting CFHT run-length.

Distillation performance

Atmospheric and vacuum columns and the delayed coker’s main fractionator produce CFHT feed streams in a typical high-conversion US refinery (Fig. 3).

Product contaminants are either physically entrained or are volatile based on boiling range. Better fractionation can remove up to 40% of the volatile contaminants.

For example, a vacuum unit processing Merey 16° API gravity crude oil will produce 30-40% less vanadium in the HVGO product with a properly designed damp vacuum unit vs. a typical dry-unit design. An optimal distillation unit design can completely eliminate entrained contaminants, including MCR, metals, and asphaltenes.

Contaminants in residue streams will poison the high-activity desulfurization catalysts needed for severe hydrotreating. In some instances, refiners have installed guard-bed reactors to remove these unexpected contaminants upstream of the hydrotreating catalyst beds to attain design run-lengths.

Cases in which high entrainment and poor fractionation increase CFHT feed contaminants, correcting the distillation unit’s design and operation is more cost-effective than treating the symptoms in the CFHT unit.

Process design, flash zone, wash section, and stripping section performance largely determine product quality for a given crude oil. These design considerations have historically received little focus in controlling contaminants.

Operators use delayed-coker spray chambers to decrease recycle and improve liquid volume yields; however, the resultant high-MCR and high-metals heavy coker gas oil (HCGO) product has received little attention.

In several instances, the refiner removed the vacuum column vapor horn, collector trays below the wash bed, or wash-section packing to eliminate coking. These changes eliminate coking; but the HVGO produced has much higher contaminants over the whole run-length than in a properly designed column.

In most cases, poor process and equipment design causes coking, which can be eliminated.

Many US refiners process heavy Canadian, Mexican, and Venezuelan crude oils that contain high levels of MCR, vanadium, and asphaltenes. The atmospheric crude tower, vacuum crude tower, and coker main fractionator all contribute contaminants to the CFHT feed.

The crude processed has a major influence on product quality, but the distillation equipment and process design most often dictate product contaminants.

Atmospheric crude unit

Many crude units produce a black atmospheric gas oil (AGO) product. Process and equipment designs are largely responsible for entrainment, which causes the black color.

Most atmospheric towers have indirect overflash control, use trays in the wash section, and have no flash-zone
vapor horn. When operators adjust the AGO product yield, the amount of overflash varies as a result.7

Typically, operators maximize the AGO product draw to minimize load on the vacuum unit. Because reflux is low, the trays do not operate properly due to insufficient liquid.

Furthermore, the two-phase feed velocity is usually 150-300 fps and the feed energy does not disperse due to the absence of flash-zone internals. Atmospheric residue is therefore carried up the column with the rising vapor causing a black AGO product.

A well-designed atmospheric crude column produces a yellow-orange AGO product with no entrained residue and few volatile contaminants. A well-designed column has direct wash oil flow control, a vapor horn, reduced-diameter wash section, and structured packing.

Structured packing has good de-entrainment capabilities even at reduced wash oil flow rates. A smaller wash section cross-sectional area maximizes the liquid rate in gpm/sq ft while minimizing overflash.

Table 2 shows the AGO product contaminants produced from a poorly performing wash section vs. an optimized design (Fig. 4).

Vacuum unit

Vacuum unit process and equipment design influence HVGO product contaminants. The four different types of vacuum units are dry, damp, damp with stripping, and wet. The damp design with stripping uses coil steam and stripping steam, and produces the lowest contaminant HVGO product for a given TBP cutpoint.

Dry units produce the most contaminants. Vacuum column feed enters tangentially or radially; larger-capacity units have multiple feed entries. Column-internals’ designs range from fundamentally sound to poor. Bad designs result from a failure to appreciate the importance of column-internals’ design on product quality and the typical practice of buying the cheapest solution.

The vacuum column flash-zone and wash-section designs are important for operators attempting to minimize HVGO product contaminants while operating at high capacity factors, Cr.8

Two-phase feed enters at 300-400 fps. Critical mass velocity often occurs somewhere in the transfer line, which creates finely atomized residue droplets that are dispersed in the rising vapor. The flash zone must efficiently separate this high velocity two-phase mixture, reduce the vapor velocity, and efficiently distribute the vapor into the wash section.9

The wash section consists of a collector tray and vapor distributor, packed bed, and a liquid distributor. The wash section must remove essentially all the entrained residue and fractionate any volatile metals; otherwise, CFHT feed contaminants will be higher than necessary.

Vapor horns used in most vacuum columns provide little liquid-vapor separation and essentially no vapor distribution. With tangential feed entries, therefore, the flash zone vapor tends to swirl with the vapor concentrated at the vessel wall.

Radial entries have high velocity rates opposite the feed nozzle. Common radial-type vapor horns force the two-phase mixture left and right of the radial nozzle centerline, which causes high-velocity rates. High-velocity areas prevent the wash-section packing from removing entrained residue or fractionating properly.

Typical designs of the collector tray or vapor distributors below the wash bed include a few large-diameter risers that do not distribute vapor. Liquid distributor designs are usually spray heads, although some refiners successfully use gravity distributors.

At higher feed rates and lower operating pressures, vapor velocities begin to reach the fundamental limit where de-entrainment of small droplets cannot occur.

Vacuum columns operating at capacity factors greater than 0.42-0.45 fps are near the limits of grid or structured packing beds.10 11 In these cases, the operator must closely control column flash-zone pressure to avoid residue entrainment caused by low-pressure operation.12 Without good pressure control, high entrainment will occur despite well-designed flash zone and wash section internals (Fig. 5).

Delayed-coker main fractionator

Delayed cokers reject most of the MCR to coke and maximize liquid production. When operators reduce coker-drum operating pressures and increase liquid-volume yield, however, the quantity of MCR and metals in the HCGO product increases. The main fractionator’s wash section should be designed to minimize HCGO product metals and MCR without reducing run length.

In recent years, delayed coker licensors have designed spray chambers to minimize recycle and maximize liquid volume yield. These spray chambers are

This tower packing is completely plugged with coke (Fig. 7).
designed with 1-3 levels of spray headers.

Although spray chambers minimize recycle, they often do not remove the tail on coke drum vapor that contains high MCR and metals. The HCGO product contaminants levels, therefore, have increased.

At low column-capacity factors ($C_f = 0.22-0.25$ fps), spray chambers work reasonably well. Performance, however, degrades as operators reduce coke-drum operating pressures to increase liquid-volume yield and increase charge rates. This is especially true when refiners try to minimize HCGO product MCR and vanadium while processing heavy Canadian, Venezuelan, and Mexican feedstocks.

A packed section can replace spray chambers, which, if designed properly, will operate trouble-free throughout the run (Fig. 6). If designed incorrectly, the packing will rapidly coke up and cause high HCGO product contaminants.

**CFHT contaminants**

An increase in product contaminants and a decrease in CFHT run-length result from damaged wash-section internals or the formation of coke. The internals design should meet the mechanical requirements of the service and minimize or eliminate coking.

Wash sections are subject to severe uplift forces from rapid water vaporization or high liquid levels; these forces necessitate a robust mechanical design. In most cases, an internals design for severe service, start-up procedures for properly drying out stripping steam, and the avoidance high liquid levels can prevent damage. Wash-section internals will not effectively remove contaminants once they are damaged.13

Many refiners have coked their vacuum column and delayed-coker main fractionator wash sections; a few have badly fouled the atmospheric column’s wash section. Coking can occur in the packing (Fig. 7) or on the collector trays below the packing or spray chamber.

Poor equipment design or operations cause coking in wash-section internals. Many refiners decrease the wash oil’s flow rate to increase yield, which results in a low liquid flow rate through the middle of the packing.

Coke begins to form where the oil residence time is high; once coke is formed it tends to catalyze the formation of more coke. Eventually product qualities deteriorate so much that the unit must be shut down.14 Fundamentally sound equipment design and a clear understanding of the consequences of incorrectly operating the unit can prevent coking.

**Cat naphtha, LCO fractionation**

The heaviest portion of the FBR gasoline contains 50% of the sulfur. Gasoline...
sulfur increases as the gasoline tail increases and as more low-boiling-range hydrocarbons are included in the LCO product.

Improved fractionation will recover more of the low boiling point, low-sulfur material from the LCO product and reduce the gasoline tail, which lowers the sulfur in FCC.

There are approximately 175 FCC units in the US. Many were built years ago based on minimum capital investment; therefore, they poorly fractionate the gasoline and LCO products.

A few FCC main fractionators have 15 trays between the gasoline and LCO product draws, 10-12 trays are more typical, and many units have only 6-8 trays. In the last 10 years, several refiners have replaced trays with structured packing. Some of these have been successes, others have required changes to correct poor liquid and vapor distribution.

Refiners can lower FCC gasoline sulfur using fractionation because the heavy end of the gasoline contains a large portion of the sulfur. Fractionation depends on the number of trays, amount of packing, reflux rate, and efficiency of the mass transfer device. All these variables influence the gasoline tail.

As fractionation efficiency decreases, gasoline sulfur increases. For those refiners using CFHT to meet the 30-ppmw sulfur limit, good fractionation is essential.

The main fractionator heat balance affects the gasoline-LCO reflux rate. At higher FCC unit feed rates and reactor steam rates, the fractionator vapor-liquid rates exceed the column gasoline-LCO section tray capacity. Refiners, therefore, adjust heat balances by increasing the slurry and HCO pumparound to keep the upper column section from flooding.

This allows higher feed rates but reduces reflux rates; thus, many refiners do not have good separation between the gasoline and LCO streams.

Fractionation tray or packing efficiency also impacts gasoline and LCO product fractionation. A well-designed tray has an efficiency of about 65% in the gasoline-LCO fractionation section; a 0.5-in. crimp structured packing has a 24-in. height equivalent to a theoretical plate (HETP) when used with a high-quality liquid distribution system.

Tray efficiency drops dramatically when the trays flood. For example, 10 flooded trays may generate only 1-2 theoretical fractionation stages. Poor liquid and vapor distribution reduced the efficiency of a 120-in. deep bed of structured packing to almost zero in another refinery.

Table 3 shows the improvement in distillation tail when packing has high efficiency vs. a lower efficiency that results from poor liquid distribution.

Sulfur distribution is highly nonlinear. The 390° F.-EP hydrocarbons contain 50% of the sulfur in the form of benzothiophenes.

Table 3 shows an FCC gasoline D86 95%-EP tail resulting from varying gasoline-LCO fractionation performance. In this case, the refiner lowered FCC gasoline sulfur content by more than 10% simply by improving fractionation.

This represents about 5-ppmw sulfur in a 50-ppmw sulfur FBR gasoline. Fractionation between FCC gasoline and LCO will become increasingly important for controlling product sulfur.

Separation is a function of reflux, number of trays, amount of packing, and the fractionation device efficiency. Refiners must address all these factors to improve gasoline and LCO product fractionation.

**Undercutting FCC gasoline**

The refiner must reduce gasoline yield if the FCC unit’s feed sulfur exceeds the level required to produce a blend-grade FBR gasoline. Refiners can undercut the heaviest portion of the gasoline to LCO product or draw a heavy naphtha stream from the main fractionator to reduce gasoline sulfur.

The unit, however, must have the flexibility to handle the consequences of undercutting. If not, a lower FCC-unit feed rate is required, which will further reduce the amount of FCC gasoline for blending.

Some potential consequences of undercutting are: increased wet-gas production, reduced main-fractionator overhead temperature, decreased propylene recovery, less LCO pumparound heat removal, and higher LCO product rundown cooling.

One downstream effect is that refiners will have to process a higher naphtha boiling-range material through the diesel hydrotreater and subsequently fractionate it in the back end of this unit.

Another method that reduces FCC gasoline sulfur is to produce heavy naphtha from the main fractionator (Fig. 9). The CFHT EOR can yield a heavy naphtha product. Operators can blend heavy naphtha into the gasoline pool during CFHT SOR when sulfur may not be limiting, or simply sell it.

**References**


**The authors**

Tony Barletta is a chemical engineer for Process Consulting Services Inc., Houston. His primary responsibilities are conceptual process design and process design packages for large capital revamps. Barletta previously worked as a production planner and process specialist for BP Oil’s Alliance refinery and a process engineer for Hess Oil Virgin Islands Corp. He holds a BS in chemical engineering from Lehigh University.

Scott W. Golden is a chemical engineer with Process Consulting Services Inc. His previous experience includes refinery process engineering, and distillation troubleshooting and design. He holds a BS in chemical engineering from the University of Maine and is a registered professional engineer in Texas.