SPRS
EVALUATION CRITERIA
(Revised JUL 2019)

1. Delivery Performance:

A contractor's delivery performance will be based by Product Service Code (PSC) or North American Industry Classification System (NAICS). The contractor's delivery performance will be based on a weighted combination of the percentage of contract line items with on-time deliveries and the average days late for all deliveries. The formula and weights for delivery performance are:

Delivery Score = (On-time Weight X On-time Score) + (Average Days Late Weight X Average Days Late Score)

- ✷ On-time Weight = 0.6
- ✷ On-time Score = 100 X (age-weighted number of lines shipped on-time during rating period / age-weighted number of lines shipped during rating period)
- ✷ Average Days Late Weight = 0.4
- ✷ Average Days Late Score = maximum of ((100-(total age-weighted days late during rating period / age-weighted number of lines shipped during rating period)) or 0, whichever is higher)

Age Weighting: Delivery and quality records are weighted based on the number of calendar days since the appropriate date for the record type. Age weighting is applied to the data being scored (numerator) and the total lines shipped (denominator) during calculation.

- ✷ Age-Weight Multiplier = $(1095^\dagger - $ age of record in days$) / 1095^\dagger$

  † 1,096 days, if it is a leap year

As a record gets older, its contribution to the delivery score or quality classification diminishes, vanishing at three years

Days Late Exceptions: Some cancellations are counted as delinquencies. These involve records that have been terminated due to the fault of the contractor. Contracts Cancelled due to Contractor Caused Deficiency (Not T4D) (records with Termination Indicator Code K) will be assessed at 180 days late and Terminations for Default (Termination Indicator Code D) will be assessed at 360 days late. As contractor responsible terminations record types these are the only delinquencies counted in the formula. Those terminated at the convenience of the Government are not included.

> **Definition**: A Delivery Record details when a shipment is due based on the contract.

Classification: The DoD System will evaluate CLOSED (shipped, received, or terminated), and OPEN (not yet received) deliveries. No duplicates will be allowed. If there is no delivery information provided by the source databases, the field will be blank (i.e. listed as 'No Classification Records').

2. Quality Performance

Contractor's quality performance will be based by PSC. The formula for quality is:

(Positive weighted data minus Negative weighted data) / Delivery Line Item Total by PSC

If there is no delivery data available, a value of one (1) would be used for the bottom quotient. Each PSC for which there is quality performance data will be assigned a color. Color is based on the high five percent in the commodity (Dark Blue), next 10 percent (Purple), next 70 percent (Green), next 10 percent (Yellow), and last five percent (Red). In this calculation, the companies are classified based on quality performance comparisons for all competitors within an PSC.

**NOTE**: If there is only one percentage group for an entire PSC, the group will be classified as Green.

**NOTE**: If a contractor has delivery data but no quality data for a given PSC, that contractor will be given an automatic Green rating (Delivery Green).

The following are the Quality Performance Records to be used and the weight factors for each:

Record	Service	Positive Weight	Negative Weight
Bulletins	Navy	N/A	-1.0 (critical) - 0.7 (major)
GIDEP Alerts	All	N/A	-1.0 (critical) -0.7 (major) -0.2 (minor)
Material Inspection Records (MIRs)	Navy	+ 1	-1.0 (critical) -0.7 (major) -0.2 (min)
PQDRs - Category 1 or 2	ALL	N/A	-1.0 (Cat 1) -0.7 (Cat 2)
Surveys (excluding Pre-Award Surveys)	DCMA and Navy	+0.7	-0.7 (others)
Test Reports (1st Article, Production, etc.)	Navy	+0.5	-0.5
Supply Discrepancy Reports (SDR)	ALL	N/A	-0.5

Table 1: Quality Performance Records and Weight Factors

3. Calculation of the Average Price, Expected Range, and Average Price Confidence Score

Three data repositories are required to calculate the Average Price and Expected Range for an item.

- A **History Table** is a data table built from contract award and delivery data in the Department of the Navy's Product Data Reporting and Evaluation Program (PDREP). This table holds contract information and unit prices for items with national stock numbers (NSN) such that:
  - The contract award date is 01 January 2010 or later,
    And
  - The combination of 13-digit contractor number, NSN and unit price is unique (counted once in the Average Price calculation)

- A **Price Table** retains the Average Price, the upper and lower control limits of the Expected Range and the Average Price Confidence score for each NSN found in the History Table.

- A set of **Escalation Tables** translates an item's federal stock class (PSC) into an element of the U.S. Bureau of Labor Statistics' Producer Price Index (PPI) for the purposes of "escalating" (i.e., adjusting for inflation) historical unit prices to the current month.

A computer algorithm called the SPRS **Price Engine** populates the Price Table.

- For a particular item, the Price Engine starts by determining how many times the item has been purchased (i.e., how many transaction records exist in the History Table for the NSN). Call this number *N*.
- If any of *N* prices is stored in fractions of cents, the Price Engine retains the greatest number of significant digits found in any transaction. Price data will be displayed to this number of digits, indicating to the user that fractions of cents have been used in past purchases of the item.
- The Price Engine accounts for the effects of inflation by escalating all *N* prices according to the PPI:
  - The PPI escalation factor (Esc_{PPI}) is based on the month in which the risk is calculated (current month) and the month of the contract award (base month).

  - Using the item's PSC, the Price Engine uses the Escalation Tables to determine the PPI values for the current and base month. Esc_{PPI} is the quotient of these two values.

$$Esc_{PPI} = \frac{\text{PPI value for current month}}{\text{PPI value for base month}}$$

Equation 1: PPI escalation factor calculation

o The price paid is multiplied by the escalation factor to derive the adjusted price, P.

$$P = \text{price} \times \text{Esc}_{PPI}$$

Equation 2: Adjusted Price calculation

o If N ≤ 2, the Price Engine includes all adjusted prices regardless of magnitude.

o If N > 2, the Price Engine employs a tiered, two-pass procedure to remove outliers from the population which would tend to skew the Average Price calculation. The process is as follows:

→ **Pass 1** – performed on all N adjusted prices

⇒ Find the median (M_1) of the population of adjusted prices. ___Note___: when the population is sorted in order of magnitude, the median is the middle value (if N is odd) or the average of the two middle values (if N is even).

⇒ Exclude all prices which are greater than four times the median ($4M_1$).

o Let n be the number of remaining prices. To determine if Pass 2 is required, calculate the mean (μ), standard deviation (σ) and coefficient of variation (c_v) for the population of remaining prices. ___See equations 3, 4 and 5, respectively.___

$$\mu = \frac{1}{n} \sum\nolimits_{i-1}^{n} P_i$$

Equation 3: Population mean calculation

$$\sigma = \sqrt{\frac{\Sigma(P - \mu)^2}{n}}$$

Equation 4: Population standard deviation calculation

$$c_v = \frac{\sigma}{\mu}$$

Equation 5: Coefficient of variation calculation

→ **Pass 2** – performed on the remaining prices only if $c_v > 0.2$.

⟹ Find the median (M_2) of the remaining prices.

⟹ Given the values of M_2 and c_v determine how many remaining prices lie outside the exclusions limits from the median, shown below.

Pass 2 Exclusion Limits		
If the Median is	***And* Coefficient of Variation is**	**the exclusion limits are**
$M_2 \le \$2.00$	$c_v > 0.4$	$M_2 \pm 0.8\,M_2$
	$c_v > 0.2$	$M_2 \pm 0.6\,M_2$
$\$2.00 < M_2 \le \$7.00$	$c_v > 0.4$	$M_2 \pm 0.7\,M_2$
	$c_v > 0.2$	$M_2 \pm 0.5\,M_2$
$\$7.00 < M_2 \le \$100.00$	$c_v > 0.4$	$M_2 \pm 0.6\,M_2$
	$c_v > 0.2$	$M_2 \pm 0.4\,M_2$
$\$100.00 < M_2 \le \$1000.00$	$c_v > 0.4$	$M_2 \pm 0.5\,M_2$
	$c_v > 0.2$	$M_2 \pm 0.3\,M_2$
$M_2 > \$1000.00$	$c_v > 0.2$	$M_2 \pm 0.4\,M_2$

Table 2: Pass 2 Exclusion Limits

o If the number of remaining prices outside the exclusion limits is less than $0.32n$, exclude those prices. Otherwise, the prices are not considered to be outliers and will not be excluded. See the following examples.

Example 1: For a given population of 24 prices, $n = 24$, $M_2 = \$1.39$ and $c_v = 0.592$. The exclusion limits would be $\$1.39 \pm (0.8 \times \$1.39)$. Therefore, prices greater than $\$2.50$ and less than $\$0.28$ are outside the exclusion limits. If 7 prices are outside the exclusion limits, these prices would be excluded in Pass 2 (since $7 < .32 \times 24$)

Example 2: For a given population of 79 prices, $n = 11$, $M_2 = \$6.42$ and $c_v = 0.312$. The exclusion limits would be $\$6.42 \pm (0.5 \times \$6.42)$. Therefore, prices greater than $\$9.63$ and less than $\$3.21$ are outside the exclusion limits. If 28 prices are outside the exclusion limits, these prices would not be excluded in Pass 2 (since $28 > .32 \times 11$)

Example 3: For a given population of 14 prices, $n = 14$, $M_2 = \$859.81$ and $c_v = 0.193$. Since $c_v \le 0.2$, Pass 2 is not performed.

⟹ After all exclusions are performed, the Price Engine calculates μ, σ and c_v for the final remaining prices.

μ = the Average Price for the item

The Expected Range for the item is $\mu \pm 2\sigma$. Bids in this interval are rated WITHIN RANGE.

Upper Control Limit: UCL = $\mu + 2\sigma$. Bids higher than this value are rated HIGH.

Lower Control Limit: LCL = $\mu - 2\sigma$. Bids lower than these values are rated LOW.

o The Average Price Confidence score is based on the number and statistical distribution of final remaining prices.

$\Rightarrow$ If there are more than 72 final remaining prices, the confidence score is HIGH. The Average Price is highlighted in green.

$\Rightarrow$ If there are one or two final remaining prices, the confidence score is LOW. The Average Price is highlighted in red.

$\Rightarrow$ If there are between three and 72 final remaining prices and $c_v < 0.15$ (i.e. most prices are closer to the Average Price), the confidence score is HIGH. The Average Price is highlighted in green.

$\Rightarrow$ If there are between three and 72 final remaining prices and $c_v \geq 0.15$ (i.e. most prices are farther from the Average Price), the confidence score is MEDIUM. The Average Price is highlighted in yellow.

4. Calculation Supplier Risk Score (numerical and color scores)

A computer algorithm called the SPRS **Supplier Engine** uses 10 factors of past performance information (PPI) to calculate Supplier Risk Scores. These factors are weighted (based on age and relative importance) and summed to produce a numerical and color score for every company which has either a contract or factor data within the **last three years**. Older data is excluded by the Supplier Engine. (Data discussed below is assumed to be no older than three years, unless specified otherwise.)

The Supplier Risk Score is an overall assessment contractor performance, regardless of federal stock class or business sector. Vendors' scores are stored and retrieved according to commercial and government entity (CAGE) code; PPI is gathered by CAGE and by data universal numbering system (DUNS) numbers, as appropriate.

SPRS receives contract award and delivery data from the Electronic Data Access (EDA), Wide-Area Workflow (WAWF) and the Mechanization of Contract Admin Services (MOCAS) systems, and other sources through the Department of the Navy's Product Data Reporting and Evaluation (PDREP) system. Program performance data is received from the Defense Contract Management Agency (DCMA), Government-Industry Data Exchange Program Alerts (GIDEP) and the Contractor Performance Assessment Reporting System (CPARS), (formerly PPIRS-RC) systems.

The Supplier Risk Score described herein is the score reported to awardee/contractor and government users through the standalone Supplier Risk Report.

Vendors which have scored factor data receive both a numerical and color score.

Vendors which have no scored factor data but have at least one contract reported in PDREP do not receive a numerical score but have a GREEN color score. The system will display an asterisk (*) in place of a numerical score. This is a neutral rating, similar to how SPRS displays quality classifications for companies with no quality records but some delivery records.

Vendors which have no scored factor data and have no contracts reported to PDREP receive no numerical or color score. The system will report this as NO SCORE.

The following data and sources are used for the **Supplier Risk Score**:

Supplier Risk Score		
Factor	**Source**	**Category**
Contractor Performance Assessment Reports (CPARS)	CPARS (formerly PPIRS-RC)	Positive/Negative
Corrective Action Plans (CAP)	DCMA	Positive/Negative
Corrective Action Requests (CAR)	DCMA, PDREP	Always Negative
Federal Awardee Performance and Integrity System (FAPIIS)	CPARS (formerly PPIRS-RC)	Always Negative
Government-Industry Data Exchange Program Alerts (GIDEP)	GIDEP	Always Negative
Performance Assessment Reports (PAR)	DCMA	Positive/Negative
SPRS Delivery	SPRS	Always Positive
SPRS Quality	SPRS	Always Positive
Pre-/Post Award Surveys	DCMA, PDREP	Positive/Negative
Suspected Counterfeit	GIDEP	Always Negative

Table 3: Supplier Risk Score

Factor data is divided into three categories: "Always Negative," "Always Positive" and "Positive/Negative."

Always Negative data are factors which always lower the resulting score. Each factor is given a starting score; additional data lowers that score for that factor. The lowest possible score for each Always Negative factor is zero, so once a factor score is reduced to zero, additional data has no effect on the factor or the resulting score.

Example: A vendor has 13 GIDEP alerts for failure experience (non-counterfeit). The starting score for GIDEP alerts is 10, and each such GIDEP alert is worth -1 point. This vendor's GIDEP factor score is zero, because the first 10 alerts reduce the score to zero, and the additional three alerts are ignored. *Note*: as older alerts are excluded due to age; more recent alerts will continue to affect the factor score.

If a vendor has no Always Negative data, it receives the highest score possible for these factors.

Always Positive data are factors which always increase the resulting score.

If a vendor has no Always Positive data, the resulting score is not affected.

Positive/Negative data are factors which can increase or decrease the resulting score based on the nature of the factor. For example, laudatory CPARS report cards would likely result in a positive score for this factor and thus increase the resulting score. Critical CPARS would likely lower the resulting score.

If a vendor has no Positive/Negative data, the resulting score is not affected.

The **Supplier Engine** calculates the score for each factor based on the age of the data, the number of contracts for the vendor in PDREP and the assigned weight of each factor's contribution to the Supplier Score.

Always Negative and Positive/Negative data are scaled based on the number of contracts in PDREP. This tends to normalize (but not equalize) Supplier Scores based on the government's experience with vendors. (The underlying algorithms for Always Positive data already take this into account, so this scaling is not applied.)

Always Negative and Positive/Negative data are age-weighted on a daily basis. As a result, each newer (i.e., more recent and relevant) data point contributes more to the Supplier Score than older data. (The underlying algorithms for Always Positive data already take this into account, so this age-weighting is not applied.)

A description of each factor's calculation and weight of importance is listed below.

The net score for each factor is summed to create the numerical score.
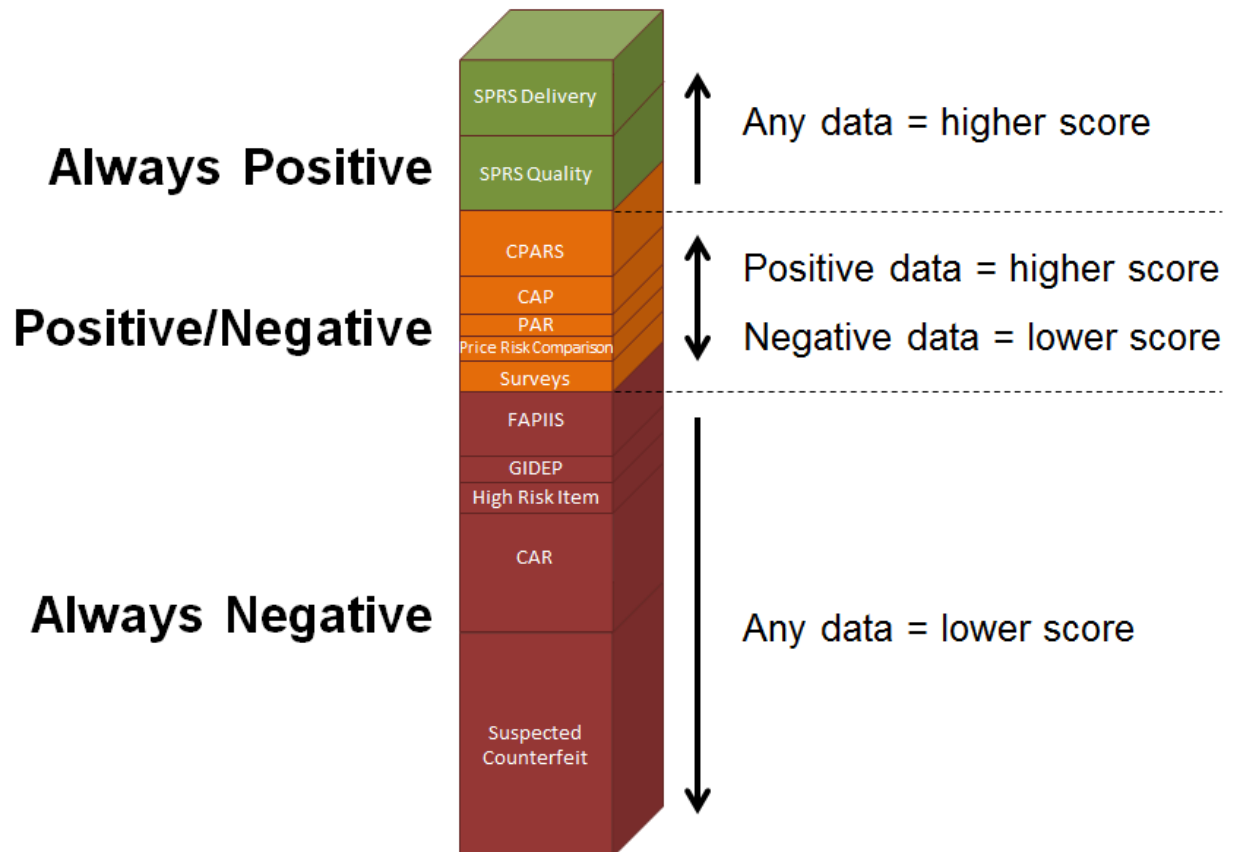
The Supplier Engine ranks the numerical scores by magnitude then assigns a color score to each vendor according to the SPRS 5-color scoring standard.

QUALITY COLOR LEGEND	
Blue:	Top 5%
Purple:	Next 10%
Green:	Next 70%
Yellow:	Next 10%
Red:	Lowest 5%

As the color score is a percentile ranking, the threshold values between colors can change each time the Supplier Engine is run. As a result, vendors with numerical scores near the threshold values may have a change in color score without a corresponding change in numerical score.

SPRS stores the numerical and color scores, contract counts and factor scores in a **Supplier Table** for quick retrieval by the system.

SPRS displays the numerical and color scores as the Supplier Risk Score. Vendors with higher Supplier Risk Scores are considered to have a lower risk of poor contract performance.

SPRS Quality factor:

If vendor has *any* Quality records in any PSC
Score = Average of rankings in each PSC
Top third = +100 pts
Middle third = +50 pts
Bottom third = 0 pts

If vendor has *no* Quality records
Score = 0 pts

Example
Quality rankings for Contractor X

PSC	Ranking	Value
1450	Top	100
1910	Middle	50
5335	Top	100
5910	Bottom	0
6820	Top	100

Average of five rankings = +70.0

Score = +70 pts

SPRS
Quality
(0-100)

SPRS Delivery factor:

If vendor has *any* Delivery records in any PSC
Score = overall Score for all PSCs

If vendor has *no* Delivery records
Score = 0 pts

SPRS
Delivery
(0-100)

SPRS
Quality
(0-100)

Pre-/Post-Award Survey factor:

- +1 point per evaluation factor
    - "Acceptable"
    - "Awarded"
    - "Complete-" and "Partial Award"
    - "Satisfactory"
- -1 point per evaluation factor
    - "Unacceptable"
    - "Unsatisfactory"
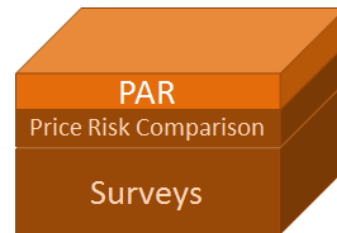   - "No Award"

Surveys
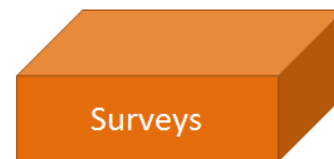
Program Assessment Report (PAR) factor:

- Score = sum of Category ratings (not Overall)
    - Cost , Schedule, Performance
- Green = 1 pt
- Yellow = 0 pts
- Red = -1 pt
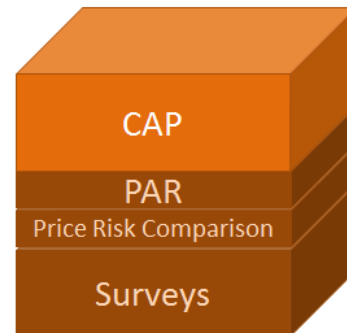

Example
PARs for Contractor X

Evaluation Area	Green	Yellow	Red
Cost	3	5	1
Schedule	1	5	1
Performance	2	1	0
Total	6	11	2
Value	**6**	**0**	**-2**

Score = 6 + 0 − 2
= 4 pts

Corrective Action Plan (CAP) factor:

- +1 point per record
    - Accepted
    - On time
- -1 point per record
    - Rejected
   - Late

CPARS factor:

- Score = Average of Rated Areas per record
    - Exceptional = +10 pts
    - Very Good = +5 pts
    - Satisfactory = +1 pt
    - Marginal = -5 pts
    - Unsatisfactory = -10 pts

Example
CPAR record for Contractor X

Evaluation Area	Rating	Value
Quality	VERY GOOD	5
Schedule	VERY GOOD	5
Cost Control	SATISFACTORY	1
Management	MARGINAL	-5
Utilization of Small Business	N/A	
Regulatory Compliance	EXCEPTIONAL	10

Average of five rated areas = +3.2

Score = +3.2 pts

CPARS
CAP
PAR
Price Risk Comparison
Surveys
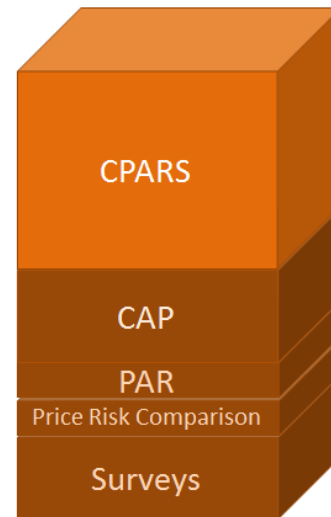
Suspected Counterfeit factor:

- Starting Score = 100 pts
- -25 pts
    - First incident
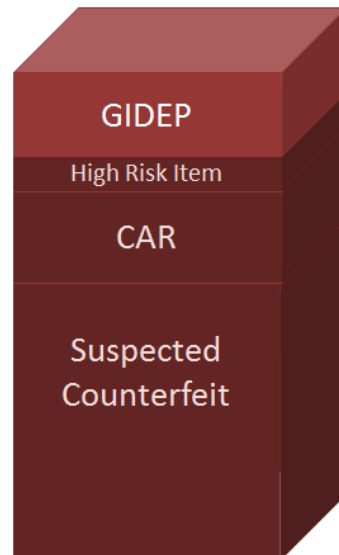- -50 pts
    - Each subsequent incident

Suspected
Counterfeit

Corrective Action Plan (CAR) factor:

- Starting Score = 60 pts
- -1 pt ea
  - Level 1
- -10 pt ea
  - Level 2
- -50 pts ea
  - Level 4

Level 3 CARs	
Occurrence	Percent Deducted
First	30%
Second	65%
Third	100%

CAR

Suspected Counterfeit

GIDEP alert (non-counterfeit) factor:

- Starting Score = 10 pts
- -1 pt
  - each record

GIDEP

High Risk Item

CAR

Suspected Counterfeit

FAPIIS factor:

- Starting Score = 50 pts
- -10 pts per record
  - DoD Determination of Contractor Fault
  - Defective Pricing
- -25 pts
  - Termination for Cause
  - Termination for Default

FAPIIS

GIDEP

High Risk Item

CAR

Suspected Counterfeit