**HSPiP Command Line Interface (CLI)**

The following commands are available. Each column item for a row is a parameter in the command line. Although quotation marks are not required in general, they *are* required for any item containing a space which, of course, is the command line separator. Commands are not case sensitive.

All input files are to be placed in the same folder as the .exe file and the license file. It is the user’s task to fetch input files from other locations. It is probably good practice for the input file to be deleted once the request has been processed. By avoiding access by HSPiP to any folder on the user’s system many potential problems are averted.

Output is always in Out.dat in the same folder as the .exe. There is no benefit to providing an option for a user-named output file for the calculation. The user’s system will, of course, want to take Out.dat and place it in a specific location with a user-defined name.

Because HSPiP is written under DotNet, the DotNet Framework 4.x must be available to the .exe when running on the server system.

Where relevant, the output files start with a Header Row with a label for each column

|  |  |  |  |
| --- | --- | --- | --- |
| Command | Input | Extra | Comment |
| Y-MBS or Y-MBSX | SMILES |  | A **S**ingle Y-MB calculation (with e**X**tra data) |
| Y-MBF or Y-MBFX | Input File1 |  | Y-MB on a **F**ile of SMILES (with e**X**tra data) |
| Sphere | Input File2 | [1-3GA] | Do a Sphere fit to a HSD set [with Options] |
| SO2 or SO3 | Input File3 | dD dP dH4 | Find the best 2 or 3 solvents |
| SOPW or SOTW or SOQW§ | Input File3 | dD dP dH4 | Get a PW/ TW/QW list of solvents |
| SOBD | Input File3 | dD dP dH4D Step6 |  |
| SOE | Input File3\* | dD dP dH4h T V s A %7 | Get a 30-step evaporation curve |
| Distance-S | Input File1 | Input File1 | Calculate the Distance & RED of a list from target(s) in SMILES format (Radius=8) |
| Distance-H | Input File5 | Input File5\* | Calculate the Distance & RED of a list from target(s) HSP + Radius |

Notes

eXtra data means a full output of Y-MB predictions in addition to the HSP and MVol values

Sphere Options are: Nothing is the standard fit with 1=Good. If 1, 2 or 3 appears in the Options then Good is set to that value. If GA is included then the Classic GA fit is done using 1=Good if no number appears. So 2 would mean standard fit with Good=2. 1GA (or just GA) would be GA with Good=1

1 A simple text file with one SMILES per row or one Name+Tab+SMILES. There should be no Header line as there is no simple way to distinguish a Header entry from a bad SMILES. Blank lines are ignored.

2 This must be in standard .hsdx file format. It is a good idea to use the .hsdx file extension so the file can be loaded directly within HSPiP if required, though there is no requirement by the CLI for the extension

3 This must be in standard .sofx file format. It is a good idea to use the .sofx file extension so the file can be loaded directly within HSPiP if required, though there is no requirement by the CLI for the extension. \* File must include Antoine A, B, C in mm Hg units (as per HSPiP)

4 i.e. three extra command line parameters representing the Target dD, dP and dH respectively

5 Whereas Distance-S uses two lists of Names and SMILES (either or both lists could contain just one chemical), Distance-H uses two lists of Name and tab-separated dD, dP and dH. In both cases Out.dat contains a list of the Distances with the Rows being Input1 and the Columns being Input2. If Input2 contains a 4th value after dH this is considered to be the Radius, otherwise Radius=8. The output file for 3 Input1 and 2 Input2 looks like:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Target1:Distance | Target1:RED | Target2:Distance | Target2:RED |
| Solvent1 | 4.3 | 0.51 | 8.2 | 1.22 |
| Solvent2 | 5.2 | 0.62 | 7.8 | 0.91 |
| Solvent3 | 1.3 | 0.23 | 6.8 | 0.73 |

6 Distance below which you want to see all binary combinations at Step X%. Clearly 1% steps will give you lots of data, 10% will be less data but lower accuracy.

7 h=thickness of formula in μm, T=temperature in °C, V=air velocity in m/s, s=seconds for simulation, A=True or False for Activity Coefficient correction, %=%RH used for water-based formulations (even if you are not using a water-based formulation, please provide a value such as 50)

§ This can create an enormous number of combinations, so use with caution.

**HSDX and SOFX File Formats**

The original CLI used the old .hsd and .sof formats which were convenient to generate as simple tab separated files, but were becoming increasingly unmanageable. The .hsdx and .sofx formats are XML. If you open any of the standard files from HSPiP into an XML editor you will see how to generate your own files if they haven’t already been created in the HSPiP GUI.

**Output formats**

Because each command elicits very different types of response, the data in Out.dat has to be parsed differently for each command. See below for a detailed description.

**YMBS**

A header column: SMILES dD dP dH dHD/A MVol

The values or else the dD column contains SMILES Error and the remaining columns are unpopulated

**YMBSX**

A header column: HCode Name dD dP dH dHD/A - - MVol CAS SMILES Formula MWt Density MPt BP RI VP@25°C RER ExpA ExpB LogOHR MIR LogKow LogS LogKsoil AA AB AC Tc Pc Vc Visc@25°C FGList InChIKey FlashPt MCI

The values or else SMILES SMILES Error

**YMBF**

A header column: Name/No SMILES dD dP dH dHD/A MVol

For each Chemical in the original list a line with those data or else the dD column contains SMILES Error and the remaining columns are unpopulated

**YMBFX**

By analogy to YMBSX

**Sphere**

A multi-line output of variable content because of possible extra information such as a bad δD fit. For historical reasons the output is not in a consistent spacing format so parsing is less than ideal. A typical example is:

Sphere Fit data in standard HSPiP format

Possible δD bad fit

In= 20 Out= 7 Total= 27

D= 19.2 P=10.1 H=6.2

Tot = 22.6

R= 11.1

Fit= 1.000

Core= ±[0.20, 0.35, 0.85]

Wrong In= 0

Wrong Out= 0

A parser with a relaxed attitude to spaces and = signs will readily extract the key values

**SO2 and SO3**

Tab separated data giving the following results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Solvent | Vol% | dD | dP | dH |
| 1st solvent | 71.0 | 18.3 | 5.6 | 5.5 |
| 2nd solvent | 29.0 | 14 | 0 | 0 |
| 3rd solvent if SO3 | x | x | x | x |
| Blend | 100 | 17.1 | 4.0 | 3.9 |

**SOPW**

Tab separated data giving the following results, a header column followed by n.(n-1)/2 lines of data for a solvent list with n solvents. The final column is the HSP Distance between the two solvents

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Solvent 1 | Solvent 2 | Vol% 1 | Vol% 2 | Distance | dDmix | dPmix | dHmix | Miscible? | Solve-Distance |
| Aliphatic hydrocarbons | Butyl Benzoate | 29.0 | 71.0 | 0.144 | 17.1 | 4.0 | 3.9 | False | 11.6 |
| Methyl Ethyl Ketone (MEK) | Xylene | 38.0 | 62.0 | 0.146 | 17.0 | 4.0 | 3.9 | False | 8.9 |
| Etc.  |  |  |  |  |  |  |  |  |  |

**SOTW**

This can give order n³ lines so should be used with caution. There are 3 extra columns (not shown) showing the HSP distances between Solv1-Solv2, Solv2-Solv3 and Solv1-Solv3

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Solvent 1 | Solvent 2 | Solvent 3 | Vol% 1 | Vol% 2 | Vol% 3 | Distance | dDmix | dPmix | dHmix |
| Acetonitrile | Aliphatic hydrocarbons | n-Amyl Acetate | 5.0 | 7.0 | 88.0 | 3.184 | 15.6 | 3.8 | 5.7 |
| Acetonitrile | n-Amyl Acetate | n-Amyl Alcohol | 4.0 | 96.0 | 0.0 | 3.221 | 15.8 | 3.9 | 6.1 |
| Acetone | Acetonitrile | n-Amyl Acetate | 0.0 | 4.0 | 96.0 | 3.221 | 15.8 | 3.9 | 6.1 |
| Acetone | Aliphatic hydrocarbons | n-Amyl Acetate | 6.0 | 6.0 | 88.0 | 3.233 | 15.7 | 3.5 | 5.8 |
| Etc. |  |  |  |  |  |  |  |  |  |

**SOQW**

The output can be imagined by extrapolation from SOPW and SOTW. It was thought unnecessary to add the 6 extra Solv-Solv distance columns.

**SOBD**

The aim is to find all solvent pairs with a Distance less than or equal to the specified value, at the given %X step values. So if you like a pair with a distance 4 (which is not the optimal pair) and want to know all other pairwise options at steps of 5% with a distance less than 4 then you would have 4 and 5 as the D and Step parameters. One CLI user especially wanted this option in order to explore extra options where different solvent ratios might be desired for other reasons. D is constrained to a maximum of 8 (the search makes little sense with higher values) and a minimum of 0.5, Step is best entered as 1, 2, 5, 10 or (max) 20; any other steps will not go right to 100%. Be careful with D and Step if you have a lot of solvents as it is easy to have 100’s of 1000s of results. The output is similar to SOPW:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Solvent 1 | Solvent 2 | Vol% 1 | Vol% 2 | Distance | dDmix | dPmix | dHmix | Miscible? | Solve-Distance |
| Aliphatic hydrocarbons | Butyl Benzoate | 10 | 90 | 0.144 | 17.1 | 4.0 | 3.9 | False | 3.4 |
| Aliphatic hydrocarbons | Butyl Benzoate | 15 | 85 | 0.146 | 17.1 | 4.0 | 3.9 | False | 3.6 |
| Etc.  |  |  |  |  |  |  |  |  |  |

**SOE**

The output contains a number of setting values that can be ignored. The data start on line 11 (this may change) with a header row (this will always start with Times s) then there are a maximum of 30 time steps – which may be fewer if the evaporation is complete before the given time. The first 6 columns are always there, followed by however many solvents are being evaporated – always a minimum of 2.

In the following output there were 3 solvents and the total evaporation time was 40s, with 6.2% of the original solvent mix remaining at the end of that time.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Time s | Total | dD | dP | dH | Ra | S1 | S2 | S3 |
| 0 | 100 | 15.5 | 12.0 | 6.6 | 5.2 | 60.0 | 30.0 | 10.0 |
| 1.3 | 94.1 | 15.5 | 12.0 | 6.6 | 5.2 | 58.3 | 31.1 | 10.6 |
| … |  |  |  |  |  |  |  |  |
| 40.0 | 6.2 | 15.8 | 4.1 | 6.1 | 5.0 | 0.0 | 5.4 | 94.6 |

**Distance-S and Distance-H**

In the example below there were 6 un-named chemicals (so given numbers) and their Distances and REDs from the 2 named chemicals are given. So in general for N chemicals in the first file and M chemicals in the second there is a table of N+1 rows and 2M+1 columns

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Ethanol:Ra | Ethanol:Red | Acetone:Ra | Acetone:Red |
| 1 | 10.8 | 1.34 | 8.7 | 1.08 |
| 2 | 10.4 | 1.30 | 9.3 | 1.16 |
| 3 | 14.3 | 1.79 | 8.0 | 1.00 |
| 4 | 14.3 | 1.79 | 8.0 | 1.00 |
| 5 | 14.1 | 1.76 | 9.9 | 1.24 |
| 6 | 11.1 | 1.39 | 6.4 | 0.80 |

**Failure**

If there is an error in the Command line then Out.dat is returned as a blank file.