Bayesian quantification of thermodynamic uncertainties in dense gas flows

Xavier MERLE
Paola CINNELLA

DynFluid Laboratory - Arts et Metiers ParisTech Paris
151, Bvd. de l’Hôpital - 75013 Paris, France

7th SAMO International Conference
01 - 04 July 2013, Nice, France
Flows of complex organic fluids close to saturation conditions are encountered in many engineering applications:
- high-Reynolds wind tunnels to chemical transport,
- refrigeration,
- energy conversion cycles,
- ...

For compressible single-phase flows occurring in thermodynamic conditions close to the liquid-vapor coexistence curve, the fluid thermodynamic behaviour differs significantly from that of a perfect gas and can no longer be represented by the polytropic perfect gas law.

Consequences

We have to find other Equations of State (EOS) to represent this particular thermodynamic behaviour!
Main Issues

- Countless EOS have been proposed in the literature, diversified according to the substance to be modelled.
- EOS based on theoretical and analytical criteria,
  - van der Waals,
  - Redlich-Kwong,
  - Peng-Robinson,
  - Martin-Hou,
  - ...

  provided that some thermodynamic inputs are available for the substance of interest.
- However, such data are typically affected by more or less significant experimental errors.

Consequences

Several uncertainties related to the use of such complex EOS coming from:
- the values taken by the substance-specific coefficients,
- the functional form of the model.
Main Issues

In these conditions, an interesting work to do is:

- to evaluate the response of the system to the uncertainties in the input parameters (Cinnella et al. [1]),
- to calibrate the input parameters, thanks to some experimental data, to improve the way an EOS represents the thermodynamic behaviour (this work).

Problem: no experimental data available up to now!

Consequences

Reference data generated by using a more complex EOS ⇒ investigation of the feasibility of the calibration procedure.

Side effect: calibration of simple, cheap EOS on a more complex, accurate and expensive one.

For this study, we use:

- Span-Wagner EOS as the reference,
- Redlich-Kwong (RKS), Peng-Robinson (PRSV) and Martin-Hou EOS to be calibrated.

The Bayesian inference framework is employed to perform the calibration. In our case, the Bayes rule takes the form:

\[
p(\theta | d, y) \propto p(d | y, \theta) p(\theta)
\]

where

- \(\theta\) is a random vector of parameters,
- \(y\) is the dense gas solver output quantity of interest (model),
- \(d\) is the experimental (numerical) data.

The quantity of interest is the criterion used for the calibration and is therefore chosen according to the experimental data.

- The calibration is performed from simulations of transonic dense gas flows (D5, siloxane) around a NACA0012 airfoil,
- quantity of interest: the pressure coefficient \(C_p = \frac{p - p_\infty}{\frac{1}{2}\rho_\infty U_\infty^2}\), at 17 locations along the airfoil wall (simulated pressure taps).
Bayesian Inference - simulations

Figure: Mesh around a NACA0012 profile.

Figure: Iso-contours of pressure coefficient. ⚫: numerical pressure sensors.
Bayesian Inference - prior and likelihood

A common practice, when no particular informations are available about the parameters, is to impose uniform distributions for prior:

\[ p(\theta) \sim U([a, b]) \]

For the likelihood, we modelled the data \( d \) as:

\[ d(x_i) = \hat{d}(x_i) + e_i \]

- \( e_i \), the experimental noise; \( e_i \) is assumed to be independant and normally distributed with mean zero and a standard deviation of 10% of the observed value.

- \( \hat{d}(x_i) \), the true pressure coefficient value at \( x_i \). \( \hat{d}(x_i) \) is assumed to be equal to the modelled value \( y(x_i) \), multiplied by an error coefficient \( \eta_i \):

\[ \hat{d}(x_i) = \eta_i y(x_i, \theta) \]

which takes into account the discrepancy between the simulation and the actual system.

\( \eta \) is assumed to be well represented by a correlated Gaussian model of the form: \( \eta \sim N(1, K_M) \)
Finally, the likelihood can be written under the form:

\[
p(d|y, \theta) = \frac{1}{\sqrt{(2\pi)^n|K|}} \exp \left[ -\frac{1}{2} (d - y)^T K^{-1} (d - y) \right]
\]

\[K = K_e + K_M,\] with \(K_e\) a diagonal matrix with corresponding variance and \(K_M = \sigma^2 \exp \left[ -\left( \frac{x - x'}{10\alpha X} \right)^2 \right]\)

where \(x\) and \(x'\) are two subsequent observations abscissa separated by the length scale \(10^\alpha X\) (\(X \approx 1\)).

\(\sigma\) and \(\alpha\) become new parameters (known as hyper-parameters)

\[\theta = (\theta_p, \theta_h).\]

The inference is done using sampling techniques of the prior and the likelihood:

- \(pymc\) python library, based on a Markov-Chain Monte-Carlo sampler and the Metropolis-Hastings algorithm,
- samples of 200 000 draws are used, the first 50 000 of which are rejected.
Dense gas solver

Dense gas effects essentially influence the inviscid flow behaviour:
- analysis of single-phase compressible inviscid flows, governed by the Euler equations and completed by a real-gas thermodynamic model.

More specifically:
- 2-D flows ($100 \times 32$ grid),
- cell-centred finite volume scheme for structured multi-block meshes of third-order accuracy,
- scalar dissipation (to reduce the computational costs),
- local time stepping, implicit residual smoothing and multigrid are used to efficiently drive the solution to the steady state.

The computation is about 10 minutes long on a classical personal computer:
- need for a surrogate model,
- in this work: piecewise multidimensional Lagrange interpolations.
Equations of state

- thermal equation: functional relation between thermodynamic variables

\[ p = \frac{RT}{\nu - b} - \frac{a}{T^{0.5}} \frac{1}{\nu(\nu + b)} \]

- caloric equation: temperature dependence of internal energy or heat capacity

\[ c_{\nu,\infty}(T) = c_{\nu,\infty}(T_c) \left( \frac{T}{T_c} \right)^n \]

- these relations involve a lot of parameters and variables: pressure, temperature, specific volume...

- for the EOS we are interested in:

  Redlich-Kwong and Peng-Robinson: \{ \begin{align*}
  & \text{the exponent } n \\
  & \text{the reduced ideal-gas isocoric heat } c_{\nu,\infty} \\
  & \text{the acentric factor } \omega
  \end{align*} \}

  Martin-Hou: \{ \begin{align*}
  & \text{the critical temperature } T_c \\
  & \text{the reduced ideal-gas isocoric heat } c_{\nu,\infty}(T_c) \\
  & \text{the critical pressure } p_c
  \end{align*} \}
Preliminary analyses

- Sobol analysis (Monte-Carlo):

<table>
<thead>
<tr>
<th></th>
<th>Strong influence on $C_p$</th>
<th>Weak influence on $C_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKS</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>PRSV</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>MAH</td>
<td>$T_c, p_c, c_{\nu,\infty}(T_c)$</td>
<td></td>
</tr>
</tbody>
</table>

- Correlations:

Redlich-Kwong

Peng-Robinson
Preliminary analyses

- **Sobol analysis (Monte-Carlo):**

<table>
<thead>
<tr>
<th></th>
<th>Strong influence on $C_p$</th>
<th>Weak influence on $C_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKS</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>PRSV</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>MAH</td>
<td>$T_c, p_c, c_{\nu,\infty}(T_c)$</td>
<td></td>
</tr>
</tbody>
</table>

- **Correlations:**

  ![Redlich-Kwong graph]

  ![Martin-Hou graph]
The previous observations allow to reduce the number of parameters:

- **Redlich-Kwong**: the calibration is performed with only one parameter ($\omega$) whereas $n$ and $c_{\nu,\infty}(T_c)$ are taken at fixed values
  $$\begin{cases} 
  n = 0.5 \\
  c_{\nu,\infty}(T_c) = 180.0
  \end{cases}$$

- **Peng-Robinson**: the calibration is performed with only one parameter ($\omega$) whereas $n$ and $c_{\nu,\infty}(T_c)$ are taken at fixed values
  $$\begin{cases} 
  n = 0.5 \\
  c_{\nu,\infty}(T_c) = 150.0
  \end{cases}$$

- **Martin-Hou**: the calibration is performed with the two parameters $T_c$, $p_c$ whereas $c_{\nu,\infty}(T_c) = 78.0$
Calibration - Redlich-Kwong

- **Legend:**
  - prior
  - posterior

- **Statistics:**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td>0.68609</td>
<td>0.07180</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.15079</td>
<td>0.07028</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-0.97095</td>
<td>0.19811</td>
</tr>
</tbody>
</table>
Calibration - Peng-Robinson

- legend:
  - prior
  - posterior

- statistics:

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega$</td>
<td>0.82925</td>
<td>0.06900</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.03050</td>
<td>0.02872</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-1.58707</td>
<td>0.56775</td>
</tr>
</tbody>
</table>

$p_{\text{posterior}}(\omega)$

$p_{\text{posterior}}(\sigma)$

$p_{\text{posterior}}(\alpha)$
Calibration - Martin-Hou

- Legend:
  - prior
  - posterior

- Statistics:

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_c$</td>
<td>555.20456</td>
<td>1.37210</td>
</tr>
<tr>
<td>$p_c$</td>
<td>14.33245</td>
<td>0.64496</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.07523</td>
<td>0.02481</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-1.60219</td>
<td>0.33949</td>
</tr>
</tbody>
</table>
Calibration Method

-/Legend:
  - Blue: prior
  - Red: posterior

Statistics:

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_c$</td>
<td>555.20456</td>
<td>1.37210</td>
</tr>
<tr>
<td>$p_c$</td>
<td>14.33245</td>
<td>0.64496</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.07523</td>
<td>0.02481</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-1.60219</td>
<td>0.33949</td>
</tr>
</tbody>
</table>
Calibration

Redlich-Kwong

Peng-Robinson

Martin-Hou

- legend:
  - red: reference
  - green: classical values
  - blue: calibrated solution
  - black: prior solution
Good agreements between calibrated and reference data,

The study seems to show that there is not only one parameter set which best fit the experimental data since some parameters are found to be correlated to each other,

⇒ to be confirmed,

Future work: perform Bayesian model averaging to take into account several model forms and scenarios.
Bayesian inference

e_i is assumed to be independant and normally distributed with mean zero and a standard deviation of 10% of the observed value. Then we can write that:

$$p(d|\hat{d}) = \frac{1}{\sqrt{2\pi}^n |K_e|} \exp \left[ -\frac{1}{2} (d - \hat{d})^T K_e^{-1} (d - \hat{d}) \right]$$

$\eta$ is assumed to be well represented by a correlated Gaussian model of the form:

$$\eta \sim \mathcal{N}(1, K_M)$$

A common used choice for $K_M$ is:

$$K_M = \sigma^2 \exp \left[ - \left( \frac{x - x'}{10^\alpha X} \right)^2 \right]$$

where $x$ and $x'$ are two subsequent observations abscissa separated by the length scale $10^\alpha X$ ($X \approx 1$).

$\sigma$ and $\alpha$ become new parameters (known as hyper-parameters) to be calibrated and representing the magnitude of the error variance and the magnitude of the correlation length.
multi-dimensional lagrange interpolations

\[ \prod f(v_1, ..., v_m) = \sum_{i_1=0}^{n_1} \cdots \sum_{i_m=0}^{n_m} \alpha_{i_1 \ldots i_m} l_{i_1}(v_1) \cdots l_{i_m}(v_m), \]

where the \( l_i \)'s are the 1-D characteristic lagrange polynomials,

Gauss-Lobatto grids to minimize the interpolation error,

in this study \( m \leq 3 \).
Preliminary analyses

- Interpolation: 10 random points inside the mesh for each EOS (total: 170 points)

<table>
<thead>
<tr>
<th>EOS</th>
<th>Parameter range</th>
<th>&lt; 1%</th>
<th>&gt; 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKS</td>
<td>$\omega \in [0.5, 1.1]$, $c_{\nu,\infty}(T_c) \in [30.0, 200.0]$, $n \in [0.0, 1.0]$</td>
<td>163</td>
<td>1</td>
</tr>
<tr>
<td>PRSV</td>
<td>$\omega \in [0.5, 1.1]$ (9), $c_{\nu,\infty}(T_c) \in [100.0, 200.0]$ (9), $n \in [0.0, 1.0]$ (9)</td>
<td>166</td>
<td>1</td>
</tr>
<tr>
<td>MAH</td>
<td>$\omega \in [a, b]$, $T_c \in [a, b]$, $p_c \in [a, b]$</td>
<td>X</td>
<td>Y</td>
</tr>
</tbody>
</table>

Main discrepancies are in the vicinity of the shock.

- Sobol analysis:

<table>
<thead>
<tr>
<th>EOS</th>
<th>Strong influence on $C_p$</th>
<th>Weak influence on $C_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RKS</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>PRSV</td>
<td>$\omega, c_{\nu,\infty}(T_c)$</td>
<td>$n$</td>
</tr>
<tr>
<td>MAH</td>
<td>$T_c, p_c, c_{\nu,\infty}(T_c)$</td>
<td></td>
</tr>
</tbody>
</table>
Redlich-Kwong

Thermal equation:

\[ p = \frac{RT}{\nu - b} - \frac{a/\nu^{0.5}}{\nu(\nu + b)} \]

- \( p \) pressure,
- \( T \) absolute temperature,
- \( \nu \) the specific volume,
- \( R \) the gas constant,
- \( a \) and \( b \) two material dependant parameters.

The Soave modification[1]:

- \( a(T) = a_c\alpha(T); a_c = 0.42747R^2\frac{T_c^2}{p_c}, \alpha(T) = [1 + m(1 - T_r^{0.5})]^2, T_r = \frac{T}{T_c} \)
  
  and \( m = 0.480 + 1.57\omega - 0.176\omega^2 \)

- \( b = 0.08664R\frac{T_c}{p_c} \)

where the subscript \( c \) denotes critical-point values and \( \omega \) is the substance acentric factor.

Caloric equation:

\[ c_{\nu,\infty}(T) = c_{\nu,\infty}(T_c) \left( \frac{T}{T_c} \right)^n \]

where \( c_{\nu,\infty}(T) \) is the ideal-gas-limit isocoric specific heat.

Those equations can be written in a normalized form:

\[
\begin{align*}
\frac{p_r}{\rho_r} &= \frac{T_r/Z_c}{\nu_r - b_r} - \frac{a_r/T_r^{0.5}}{\nu_r(\nu_r + b_r)} \\
\frac{c_{\nu,\infty}(T)}{R} &= \frac{c_{\nu,\infty}(T_c)}{R} (T_c)^n
\end{align*}
\]

in such a way that the RKS model only depends on the following three factors:
- the acentric factor \( \omega \),
- the exponent \( n \),
- the reduced ideal-gas constant-volume specific heat at the critical temperature \( c_{\nu,\infty}(T_c) \).
Peng-Robinson

Thermal equation:

\[ p = \frac{RT}{\nu - b} - \frac{a}{\nu^2 + 2b\nu - b^2} \]

with

- \( a = 0.457235R^2\frac{T_c^2}{p_c}\alpha(T) \),
- \( b = 0.077796RT_c\frac{T_c}{p_c} \).

To improve the results, the recorrelated \( m \) as a function of \( \omega \) of Stryjek and Vera can be used:

\[ m = 0.378893 + 1.4897153\omega - 0.17131848\omega^2 + 0.0196554\omega^3 \]

Thanks to same normalization technique as in the RKS case, the PRSV model depends only on three parameters:

- the acentric factor \( \omega \),
- the exponent \( n \),
- the reduced ideal-gas constant-volume specific heat at the critical temperature \( c_{\nu,\infty}(T_c) \).
Martin-Hou

Thermal equation:

\[ p = \frac{RT}{\nu - b} + \sum_{i=2}^{5} \frac{f_i(T)}{(\nu - b)^i} \]

with

- \( b = RT_c \frac{1 - \beta/15}{pc} \); \( \beta = 20.533 - 31.883Z_c \), \( Z_c = \frac{pc\nu_c}{RT_c} \) the critical compressibility factor,
- \( f_i(T) = A_i + B_i T + C_i \exp \left( -\frac{kT}{T_c} \right) \); \( k = 5.475 \).

\( A_i, B_i \) and \( C_i \) can be expressed in terms of \( T_c, pc, Z_c \), the Boyle temperature (function of \( T_c \)) and one point on the vapour pressure curve.

This equation can be reduced in such a way that the MAH model depends only on the following parameters: \( pc, T_c, Z_c \), the normal boiling temperature \( T_b \), \( n \) and \( c_{\nu,\infty}(T_c)/R \).

The work of Cinnella et al. shows that only \( T_c, pc \) and \( c_{\nu,\infty}(T_c)/R \) have a great influence on the outcome.