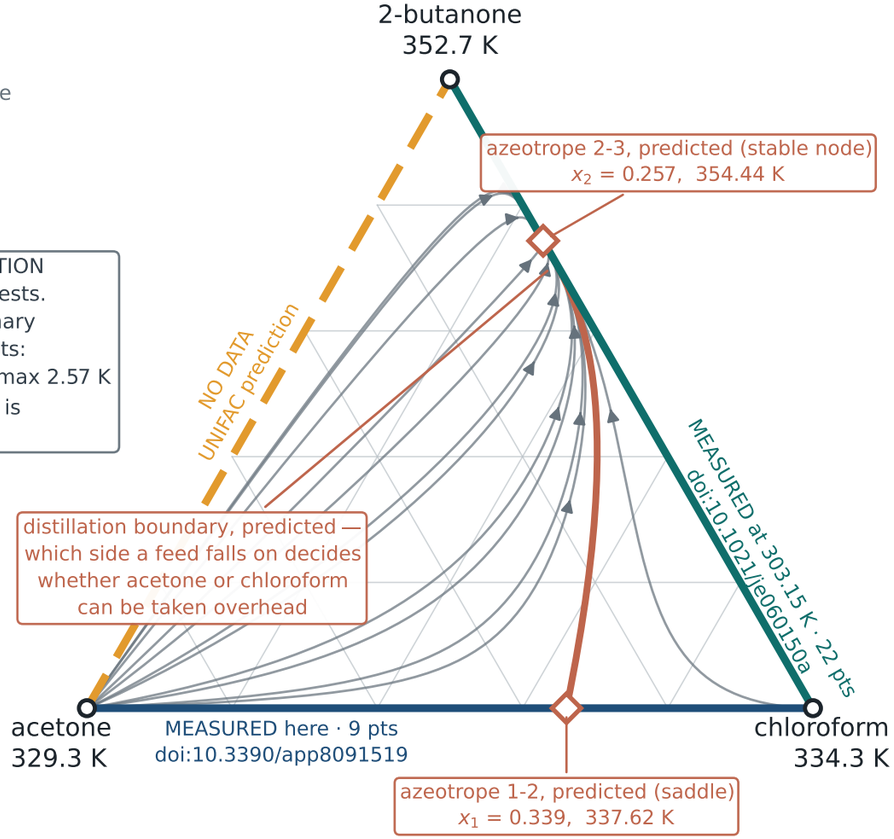
The reason NRTL and UNIQUAC replaced Margules and van Laar: a ternary needs no ternary parameter — but a prediction still needs a measurement

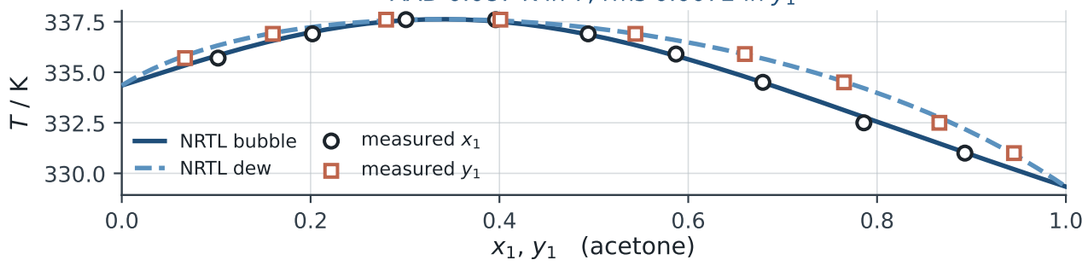
(a) acetone + chloroform + 2-butanone at 101.3 kPa, assembled from three binary parameter sets

residue curves, $dx_i/d\xi = x_i - y_i$ from binary τ_{ij} only — NRTL and UNIQUAC have no ternary parameter

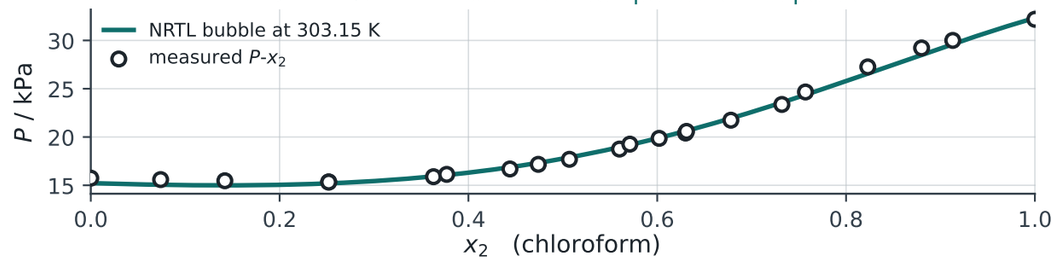


the interior is a PREDICTION
no measurement here tests.
assembled NRTL vs ternary UNIFAC, 55 interior points:
mean $|\Delta T_{bub}|$ = 0.72 K, max 2.57 K
— two models agreeing is not a measurement

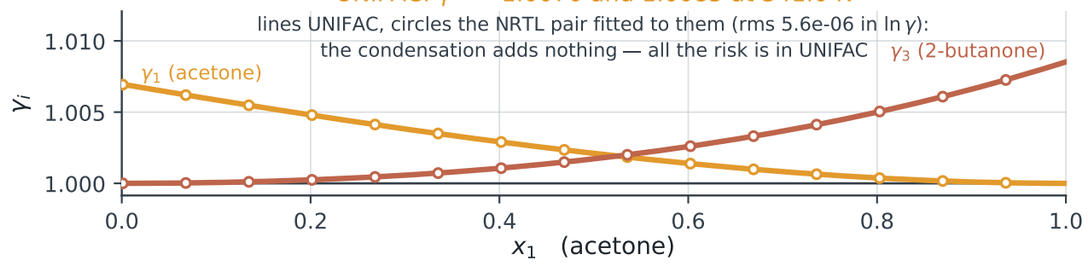
(b) edge 1-2 FITTED — and it reproduces its own data
AAD 0.087 K in T, rms 0.0072 in y_1



(c) edge 2-3 FITTED at 303.15 K, used 50 K higher
AAD 1.27 % in P — but τ held temperature-independent



(d) edge 1-3 PREDICTED — there is no data to fit
UNIFAC: γ^{∞} = 1.0070 and 1.0085 at 341.0 K



acetone (1) + chloroform (2) + 2-butanone (3) at 101.3 kPa. NRTL (α = 0.3) on every pair; the ternary uses the three fitted pairs unchanged and adds NO parameter of its own. Edge 1-2: Barker on T, doi:10.3390/app8091519 (9 points, 101.3 kPa). Edge 2-3: Barker on P, doi:10.1021/je060150a (22 points, 303.15 K). Edge 1-3 has no measurement in this module and is predicted by original VLE UNIFAC (Fredenslund-Jones-Prausnitz form, Hansen et al. VLE table, the subset embedded in vlekit.unifac), then condensed to an NRTL pair at 341.0 K.
ASSUMPTIONS, stated because none of them is tested here: τ is taken temperature-independent, so edge 2-3 is used 50 K above its data; the vapour is ideal and $\Phi_i = 1$; α = 0.3 is assumed, not fitted; and the residue curves, the azeotropes, the boundary and the whole interior are PREDICTIONS. There is no measured ternary datum on this figure.
WHAT WOULD TEST IT: measured ternary P-x-y or T-x-y for this system at 101.3 kPa, a ternary azeotrope search, or a measured residue curve. Absent those, the only handle is a second independent prediction: ternary UNIFAC on the same grid (55 points, every $x_i \geq$ 0.05) sits within 0.72 K on average and 2.57 K at worst of the assembled NRTL — a consistency check between two models, NOT a validation. On edge 1-2, where measurements do exist, that same UNIFAC is off by 0.25 K (F4.20).