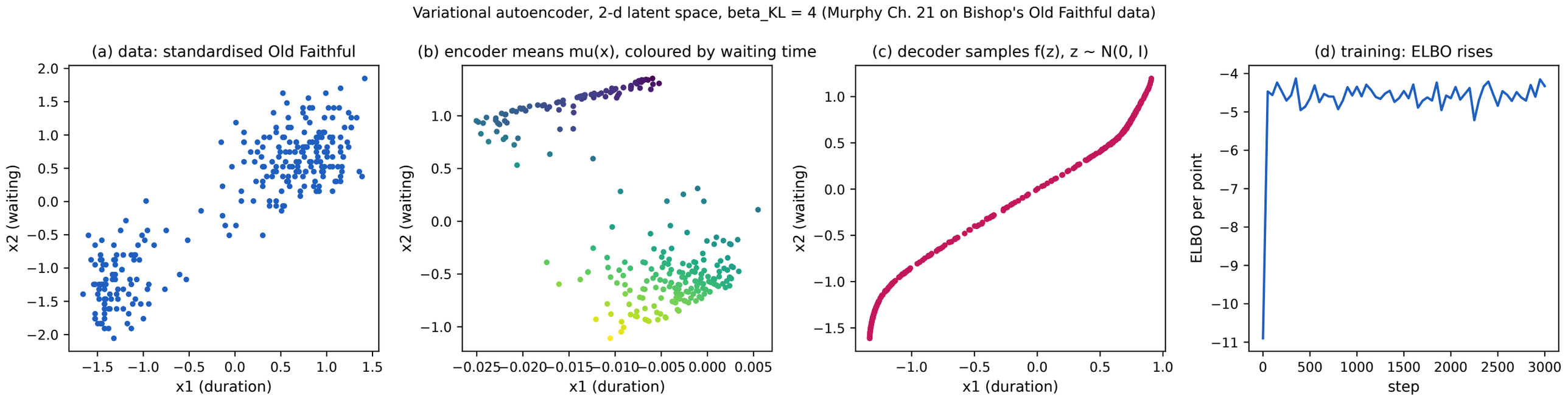

Variational autoencoder, 2-d latent space, beta_KL = 4 (Murphy Ch. 21 on Bishop's Old Faithful data)
(a) data: standardised Old Faithful
x1 (duration)
x2 (waiting)
(b) encoder means mu(x), coloured by waiting time
x1 (duration)
x2 (waiting)
(c) decoder samples f(z), z ~ N(0, I)
x1 (duration)
x2 (waiting)
(d) training: ELBO rises
step
ELBO per point