Introduction to Machine Learning:

Kernel Extensions

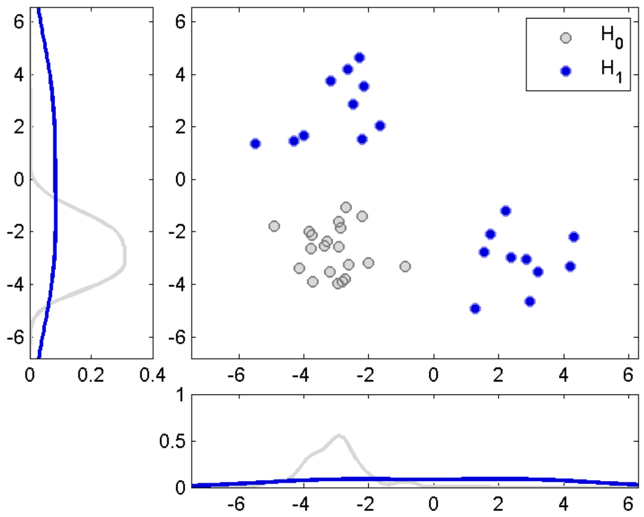
ECE 580
Spring 2022

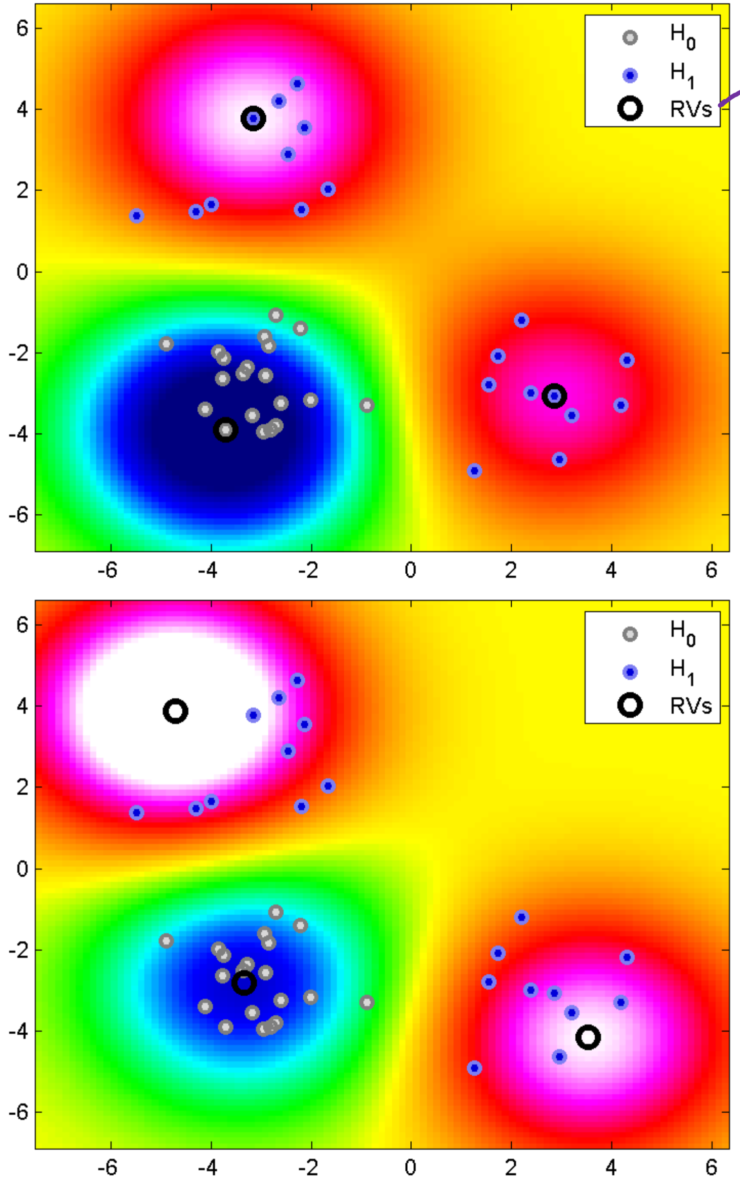
Stacy Tantum, Ph.D.

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Alternate Kernels

Kernels → basis functions





Kernels don't necessarily have to be parameterized by the training data

RBFs don't have to be centered @ training points, they can be located anywhere!

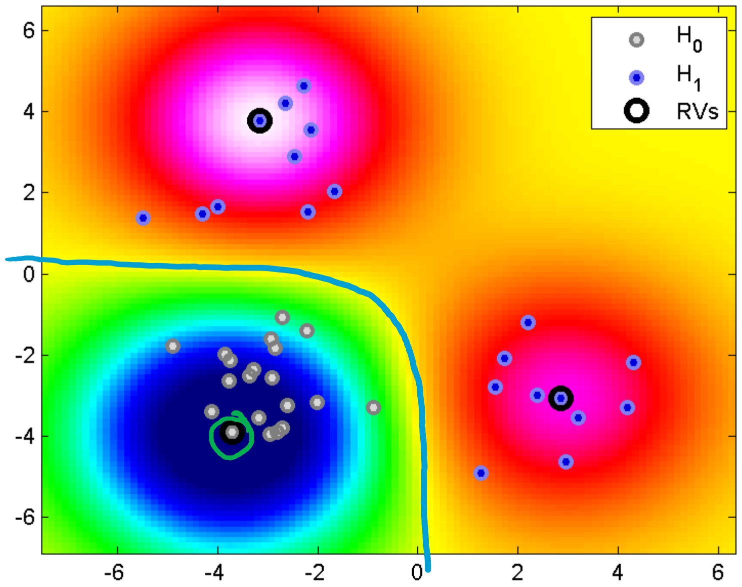
Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Choosing Kernels

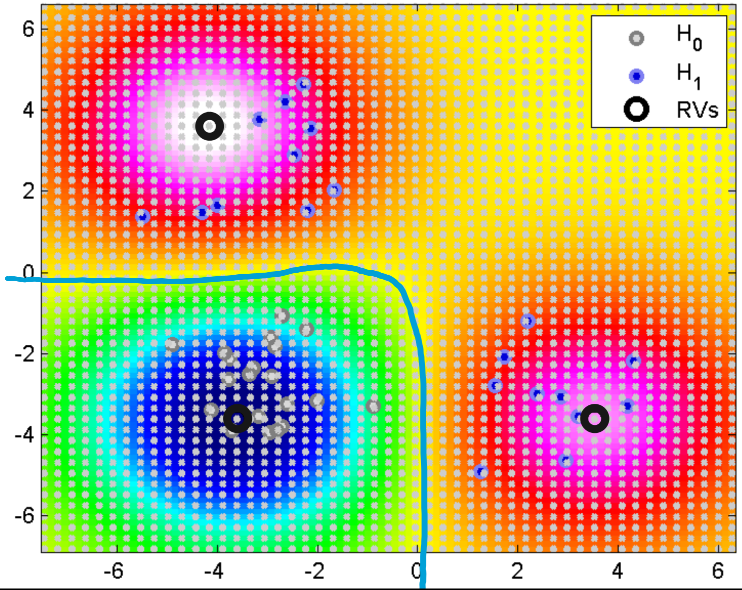
If training data doesn't define kernels, what does?

We can define parameters of candidate basis functions (rather than accepting parameters specified by training data)

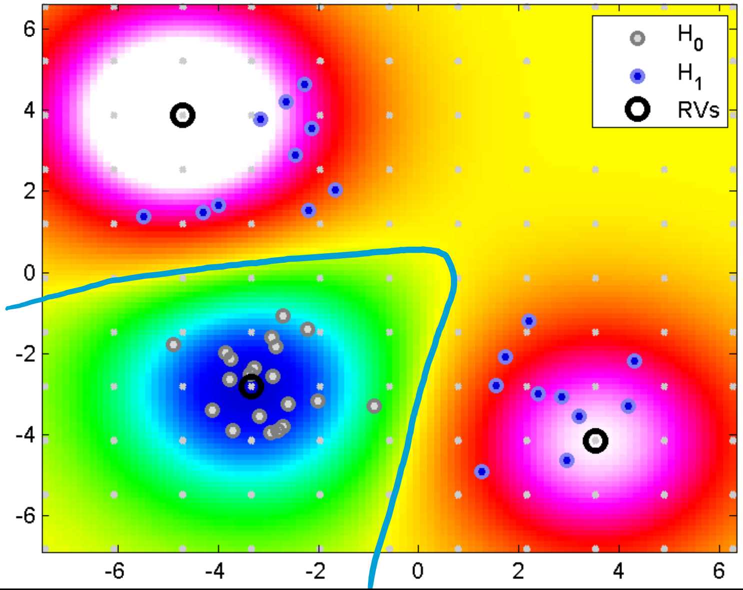
training data



fine grid



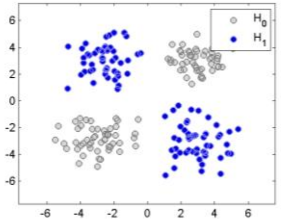
coarse grid



Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Choosing Kernel Parameter(s)

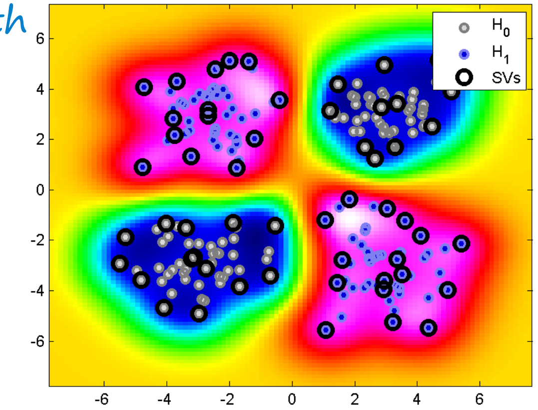
$$\begin{bmatrix} \underbrace{b_{x1}\; b_{x2}\; b_{x3}\; \cdots\; b_{xN}}_{N \text{ linear}} & \underbrace{b_{1,\gamma=1}\; \cdots\; b_{N,\gamma=1}}_{N \text{ RBF } (\gamma=1)} & \underbrace{b_{1,\gamma=2}\; \cdots\; b_{N,\gamma=2}}_{N \text{ RBF } (\gamma=2)} \end{bmatrix}$$

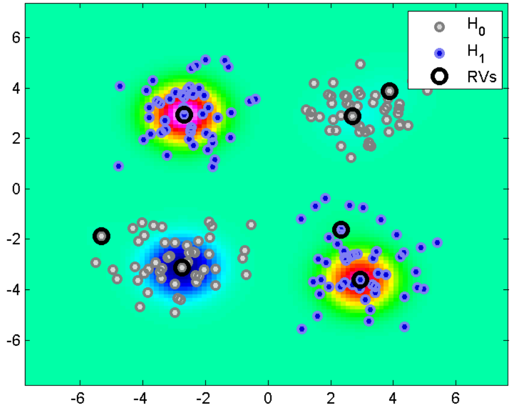


SVM

RVM

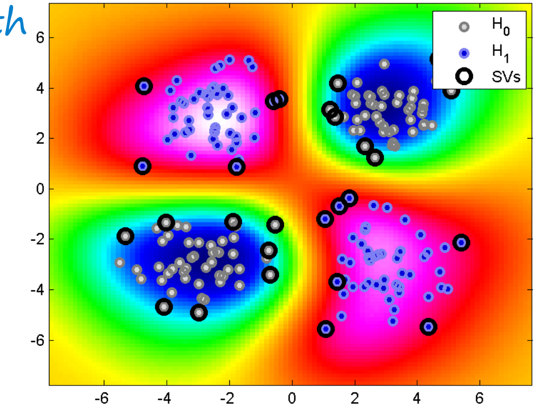
RBF width = 1

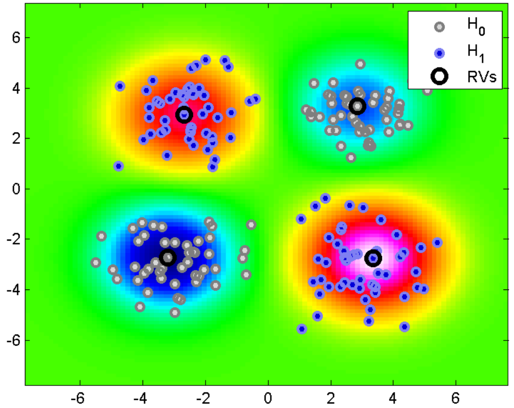




Kernel parameters don't necessarily have to be chosen by a human (a priori, or through cross-validation)

RBF width = 2





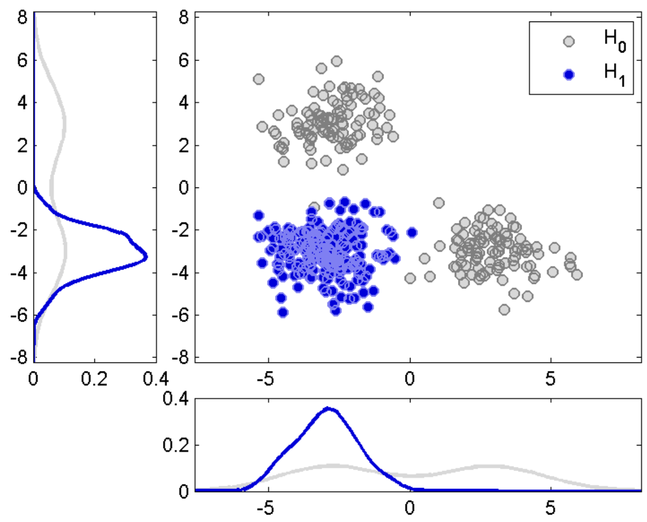
We can use all possible kernels (i.e., define more basis functions) and let the SVM/RVM decide which are SVs/RVs

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

RVM Feature Selection

Use linear (or direct) kernel

- Each kernel (basis function) corresponds to a feature
- Relevant kernels are the selected features (and they come with associated weights!)



μ_{0a} = [-1 1 0 0 … 0]

μ_{0b} = [1 -1 0 0 … 0]

μ_1 = [-1 -1 0 0 … 0]

RVM Relevant Kernels:
1 (w = -0.55)
2 (w = -0.53)

μ_{0a} = [0 0 … 0 -1 1]

μ_{0b} = [0 0 … 0 1 -1]

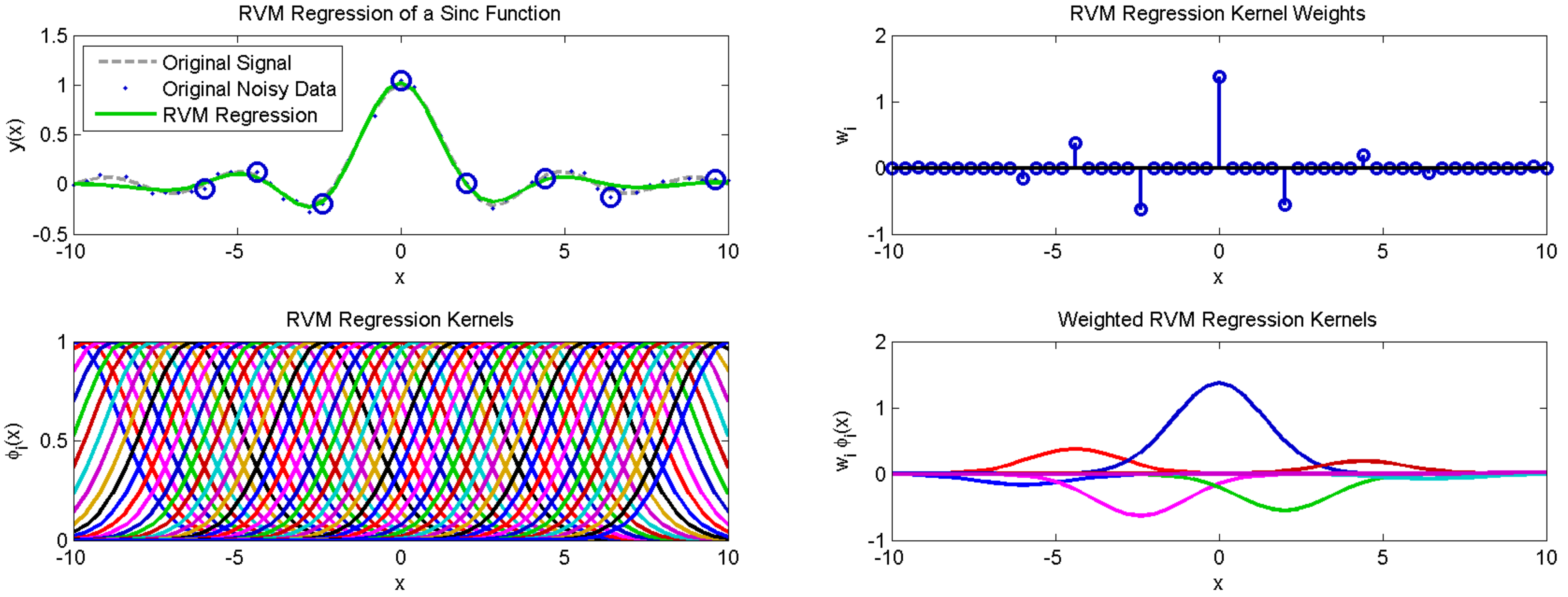
μ_1 = [0 0 … 0 -1 -1]

RVM Relevant Kernels:
D-1 (w = -0.52)
D (w = -0.49)

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

RVM Regression

Model the signal as a weighted sum of kernel functions, with the priors on the weights designed so that the majority of the weights go to zero

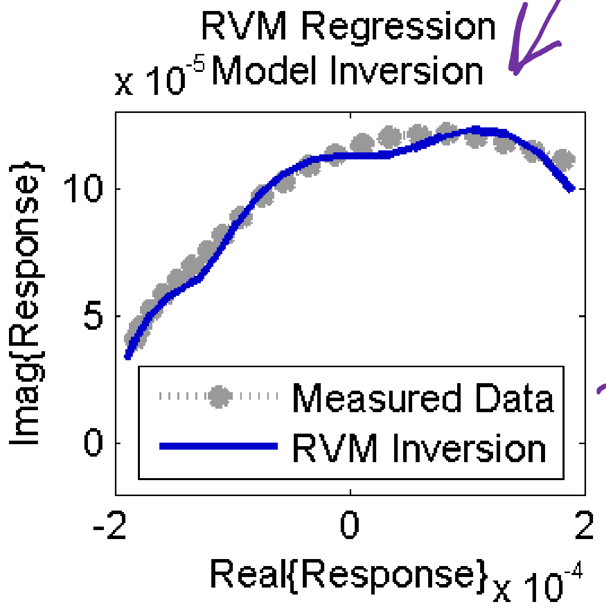


Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Problem-Specific Kernel Functions

Goal: Find a sparse representation for a wideband Electromagnetic Induction (EMI) sensor response to a target



Discrete Spectrum of Relaxation Frequencies*(DSRF)

$$H(\omega) = c_0 + \sum_{k=1}^{K} \frac{c_k}{1 + j\omega/\zeta_k}$$

Determine the K parameter sets (c_k, ζ_k) such that the signal is well-represented with minimal K

* Mu-Hsin Wei, Waymond R. Scott, Jr., and James H. McClellan, "Robust estimation of the discrete spectrum of relaxations for electromagnetic induction responses," *IEEE Transactions on Geoscience and Remote Sensing* **48**(3):1169-1179 (March 2010)

DSRF Signal Model for Regression

DSRF signal model is a linear combination of nonlinear functions

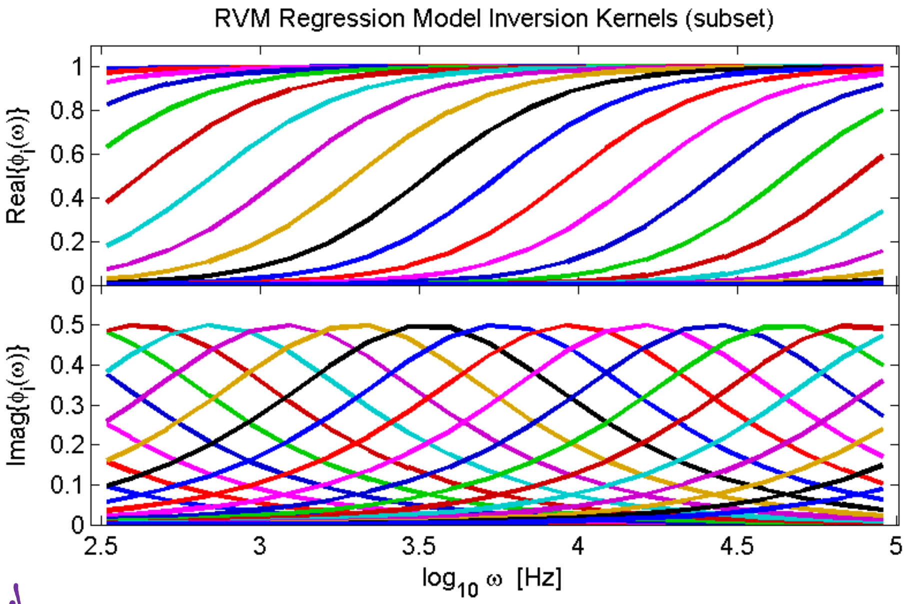
$$H(\omega) = c_0 + \sum_{k=1}^{K} c_k d(\omega, \zeta_k)$$

with

$$\boxed{d(\omega, \varsigma_k) = \frac{1}{1 + j\omega/\zeta_k}}$$

?

1st principles E/M modeling

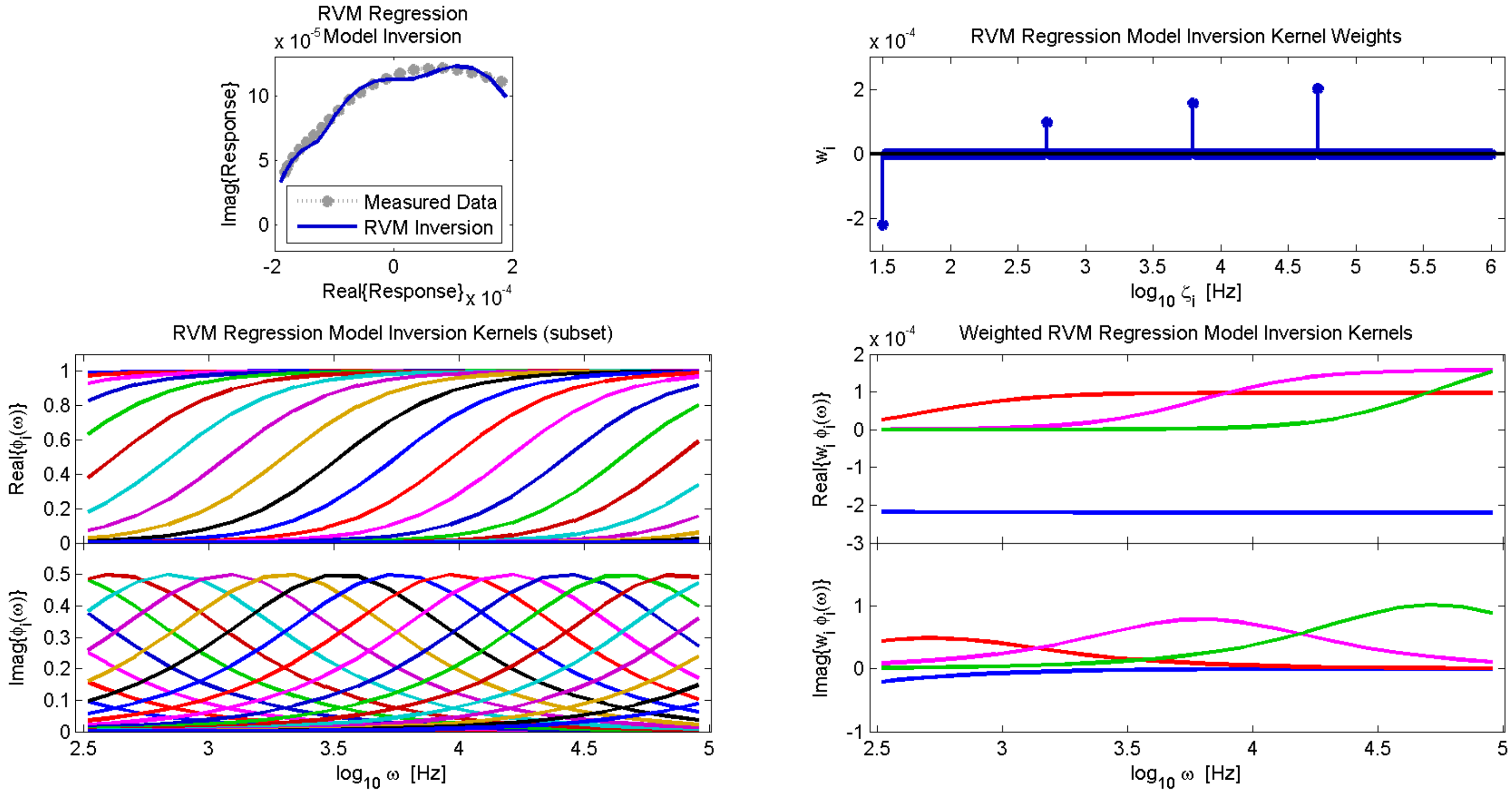


Regress out the coefficients c_k for the nonlinear DSRF basis functions

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

RVM Regression for Model Inversion

Model the signal as a weighted sum of kernel functions, and use RVM regression to find a sparse solution*

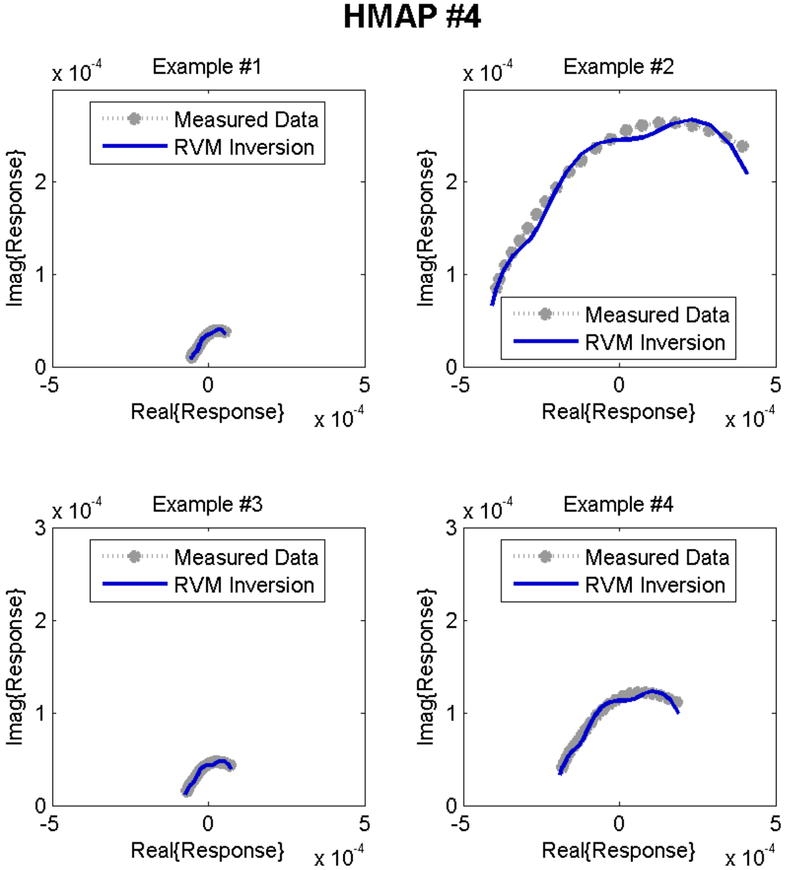


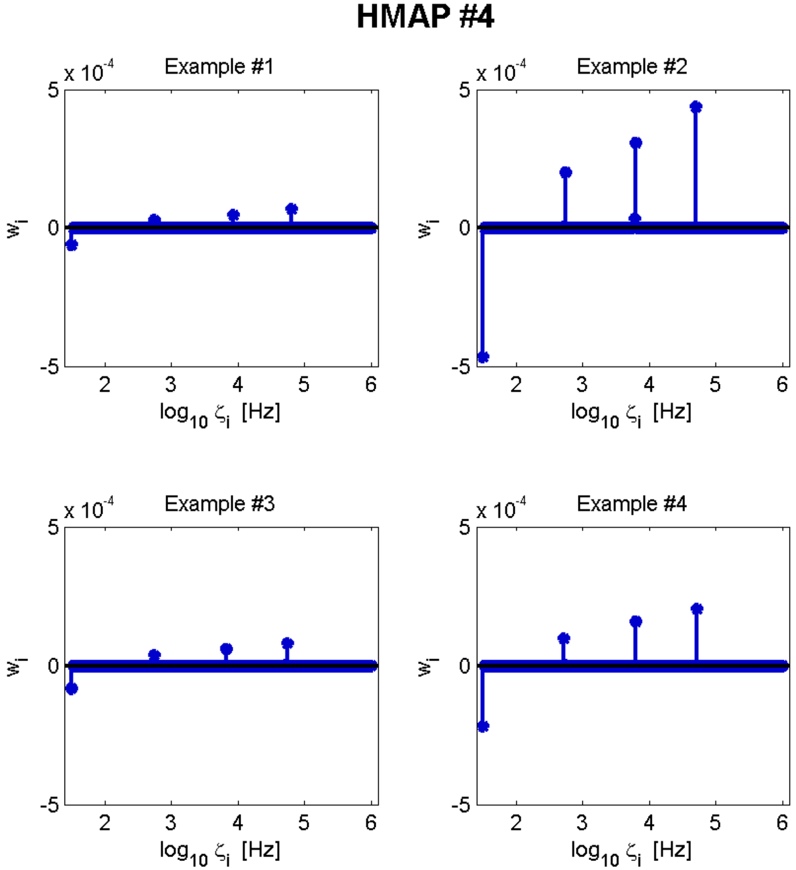
* Stacy L. Tantum, Waymond R. Scott, Jr., Kenneth D. Morton, Jr., Leslie M. Collins, and Peter A. Torrione, "Target classification and identification using sparse model representations of frequency-domain electromagnetic induction sensor data," IEEE Transactions on Geoscience and Remote Sensing, **51**(5): 2689--2706 (May 2013)

Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Example RVM Regression Inversions

Relaxation frequencies (ζ_k) are quite consistent despite considerable variations in the measured data

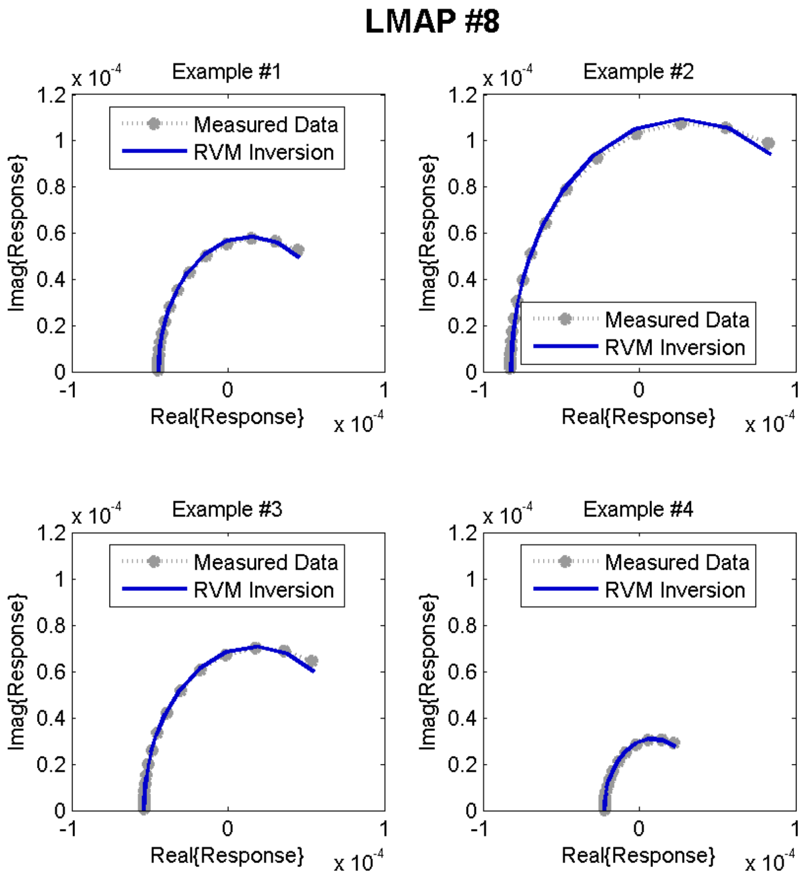


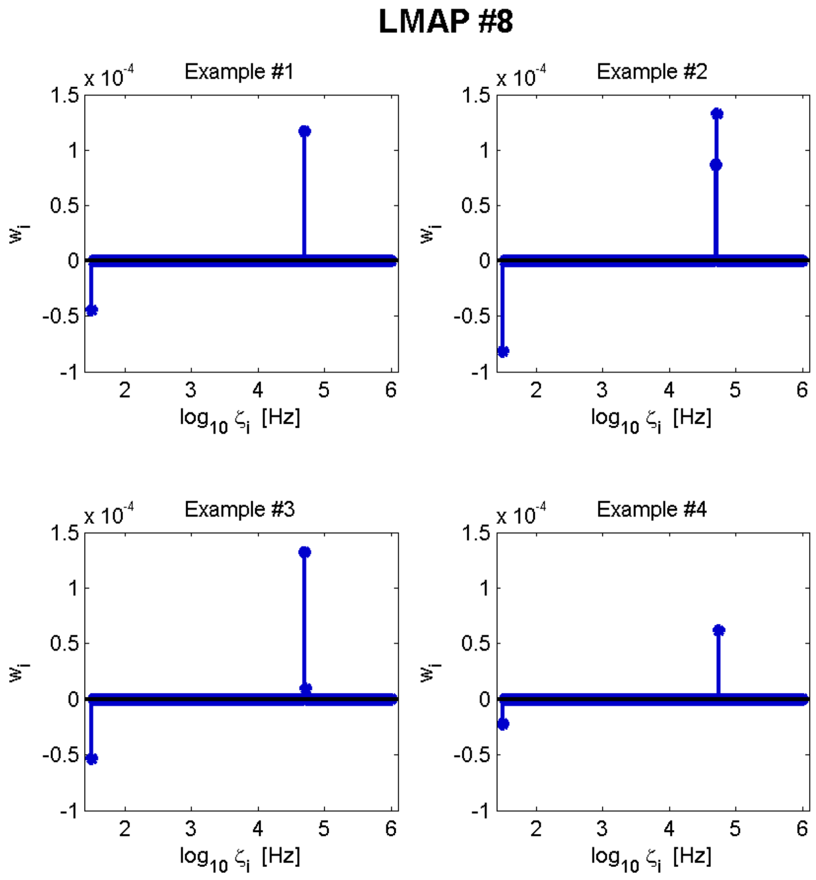


Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Example RVM Regression Inversions

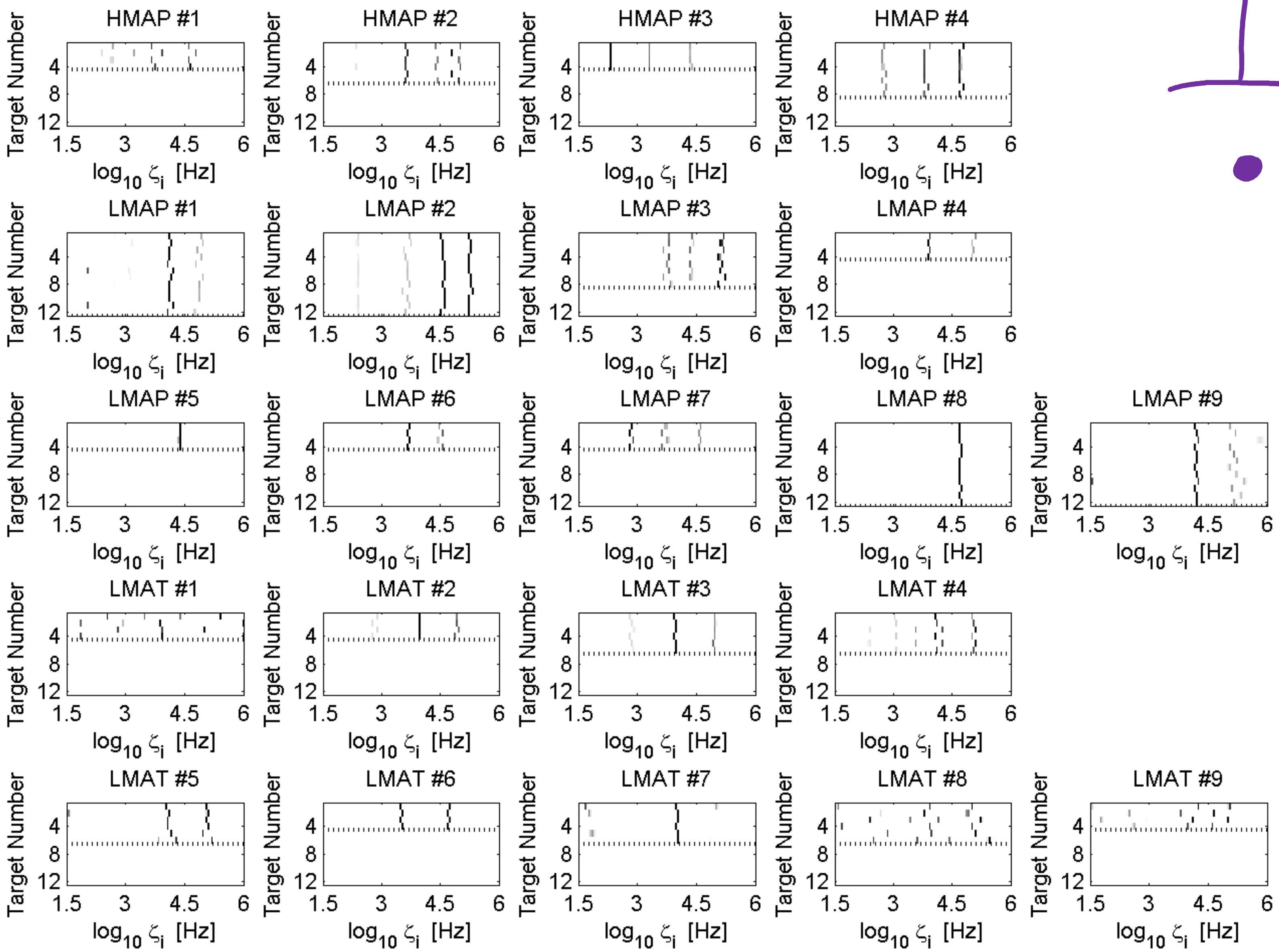
Relaxation frequencies (ζ_k) are quite consistent despite considerable variations in the measured data

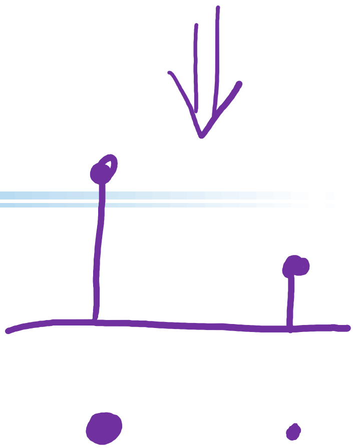




Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.

Estimated DSRFs by Target Type





Copyright © 2014-2022 by Stacy Tantum. All rights reserved.
Not to be distributed without express written consent of the author.