支持向量机（Support Vector Machine，SVM）最先由Cortes和Vapnik提出，它是一种有监督的模式识别方法。它的主要思想是建立一个分类决策面。SVM利用核函数将数据映射到高维空间，使其尽可能的线性可分。常用的核函数包括线性核函数、多项式核、径向基核（RBF）、傅里叶核、样条核和Sigmoid核函数等。通过比较这些核函数适用的数据特点，无论样本数据特点是高维还是低维，数据量大还是小，RBF核函数展现了很好的分类性能。因此，选择RBF作为SVM的分类核函数。

SVM数据处理过程如下：

设特征数据为*N*维，共*L*组数据，即。

决策面可表示为

 （1）

式中—决策面的权重系数

—非线性映射函数

*b*—阈值

为了最小化结构风险，最优分类超平面应满足以下条件

 （2）

引入非负松弛变量 ，这样分类误差就在一个规定的范围内。因此，优化问题就被转变为

 （3）

式中*c* —惩罚因子，控制模型的复杂程度和泛化能力

引入拉格朗日算法，优化问题被转换为对偶形式

 （4）

其中

 （5）

本文，引入RBF核函数

 （6）

式中g —核函数参数，控制输入空间的范围

上述优化问题转变为

 （7）

可以看出，优化问题取决于两个重要参数c和g，这两个参数会影响SVM的预测性能。SVM预测问题取决于两个重要参数c和g，这两个参数会影响SVM的预测性能。为了提高模型的预测性能，引入网格式搜索法（GS）优化模型建立过程中的两个重要参数。同时避免模型过学习和欠学习的现象发生，采用5倍交叉验证法以训练集最小均方根误差为适应度函数来进行参数寻优。当达到最小均方根误差时，所得到的c和g为最佳参数。GS中，以0.5为间隔进行全局搜索，c和g的范围均是（2-10, 210）

总之SVM预测过程为：

（1）输入数据，规定训练输入、训练输出、预测输入和预测输出

（2）为加快网络收敛速度，进行数据归一化处理

（3）参数寻优，网格数搜索开始

（4）得到最优参数建立预测模型，避免模型过学习和欠学习的现象发生，采用5倍交叉验证法以训练集最小均方根误差为适应度函数来进行参数寻优。当达到最小均方根误差时，所得到的c和g为最佳参数。

（5）预测数据输入

（6）得出预测结果

利用均方根误差（RMSE）评价预测效果好坏，RMSE越小越好。在显示面板结果直接显示了。

以12℃解释为例，当训练集在五倍交叉验证下获得最小均方误差为0.041678时，获得最佳参数c为0.43528，最佳参数g为6.6944。测试集的预测值和真实值之间均方根误差为14.8600。

