arXiv:2302.04063v1 [stat.AP] 8 Feb 2023

行为系统理论与传统线性模型在长期原位测量中预测建筑区域温度的比较

Manuel Koch[a, b] 和 Colin N. Jones[a]
[a] *EPFL, 1015 Lausanne, Switzerland*
[b] *FHNW, 4132 Muttenz, Switzerland*
manuelpascal.koch@epfl.ch/manuel.koch@fhnw.ch, colin.jones@epfl.ch

摘要

模型预测控制在建筑物中的潜力已多次得到证明，并已成功用于实现各种目标，例如降低能源消耗或提高热舒适度。然而，迄今为止，模型预测控制尚未得到大规模部署，部分原因是获取并维护足够准确模型的工程成本高昂。采用自适应数据驱动方法可以解决这一问题。近年来，在学术界，利用行为系统理论实现这一目标的理念逐渐受到关注。本研究比较了该方法的不同变体，包括使用不同数量的数据、采用不同的正则化权重以及使用不同的数据选择方法。我们采用带外生输入的自回归模型（ARX）作为成熟的基准方法。我们对来自真实有人居住建筑的两个长期数据集进行迭代系统辨识，评估所有方法；这两个数据集均未包含为系统辨识而施加的人工激励。研究发现：（1）所有方法均达到了足够的预测准确度。（2）ARX模型的表现略好，同时还具有调参数量更少、计算速度更快的优势。（3）自适应方案与非自适应方案的表现相近。（4）行为系统理论方法的正则化权重呈现出预期的权衡特性，存在一个最佳中间值。（5）使用最近的数据比选择天气状况与待预测日期相似的数据效果更好。（6）更多数据能够提升模型性能。

关键词：建筑控制、数据驱动、模型预测控制、系统辨识

术语表

α 遗忘因子λ 正则化权重$\mathcal{B}$ 系统$\mathcal{H}$ 汉克尔/轨迹矩阵I_{sol} 太阳辐射k 离散时间步长n_{sys} 系统阶数P_{cool} 制冷功率P_{heat} 制热功率T 样本数量t 时间T_{amb} 环境温度T_f 预测步数T_{ini} 初始化步数T_z 区域温度u 控制变量w 扰动W_H 轨迹矩阵宽度W_I 初始化加权矩阵y 观测变量z 输入输出对自适应 自适应 ARX 带外生变量的自回归模型

输入

BST 行为系统理论 DeePC 数据赋能预测控制 生活热水 生活热水 热泵 热泵

HRV 热回收 LTI 线性时不变
MPC 模型预测控制 PE 持续激励
PV 光伏 RMSE 均方根误差
STD 标准差

1. 引言

建筑物占全球能源总消耗的四分之一，在发达国家则高达40%；其中大部分用于供暖（包括热水）、制冷和通风。采用热泵的电力供暖正逐渐取代石油和天然气 [1, 2]。与此同时，制冷的总能源消耗预计将大幅增长，尤其是在发展中国家 [3, 4]。

模型预测控制（MPC）已多次成功应用于建筑物的仿真和实验中 [5, 6, 7, 8]。这些研究最常见的目标是降低能源消耗或提升热舒适度，而目前人们也越来越关注利用建筑物实现电力需求响应。一方面，供暖电气化仍在持续推进，制冷需求也在增长；另一方面，风能和太阳能等不可调度电力来源所占比例不断提高，因此这一趋势预计将持续下去。与前两个目标不同，后两个目标可通过预测或非预测方法实现，而需求响应通常需要较为准确地预测未来数小时的电力需求 [9]。

然而，模型预测控制尚未在建筑领域得到广泛的商业应用。常见的原因是，识别足够准确的建筑动态特性模型需要较高的前期工程成本 [5, 10, 11, 12]。

建筑具有若干特性，会影响合适模型结构的选择：与大多数工业产品不同，建筑通常各不相同，这意味着必须为每栋建筑建立不同的模型。此外，建筑的规模和复杂程度差异很大，从单个热区到数百个热区不等。我们通常测量并控制有人区域的空气温度，其时间常数以分钟计；但大部分热能储存在墙体、地板和家具中，其时间常数可能以小时计。此外，建筑呈现缓慢的时变行为，包括季节之间的周期性变化，以及因隔热层性能下降或室内陈设改变而导致的长期变化。然而，只要精度足够，建筑的瞬时动态就可以用线性时不变系统建模 [5, 8]。系统辨识实验的实施

在有人时段对供暖和制冷系统施加高强度激励，受到维持居住者舒适度以及尽量减轻暖通空调部件负担的要求限制。尽管存在这些挑战，我们仍希望在 $24h$ 预测时域内获得足够准确的预测，以便参与需求响应。我们计划最终将这项工作推广至学术界以外的实际场景，这也带来了更多挑战：建筑中的传感器往往不准确且存在偏差，而且可能无法更换或校准传感器。此外，还应避免过高的计算需求，以尽可能降低系统硬件成本。

为解决这些问题，我们提出采用自适应线性黑箱模型 [13, 14]。具体而言，我们研究一种近期广受关注、基于行为系统理论的方法的性能，详情见第 2.1 节和第 3.1 节。这类方法完全省去了模型辨识步骤。不过，它们在理论上等价于线性时不变模型。因此，我们选择外生输入自回归模型（ARX）作为成熟的基准方法。此外，我们还评估了这两种方法的非自适应变体，以考察自适应性的优势。

现有文献中还采用了其他方法对建筑物建模，但这些方法不在本文的讨论范围内。神经网络和高斯过程模型能够处理非线性，但它们需要大量训练数据，而且还需进行计算成本高昂的非线性优化。这与我们的目标相悖：以有限的计算开销辨识一个基本呈线性的系统。我们曾考虑将状态空间模型与用于状态估计的卡尔曼滤波器结合，作为 ARX 的替代方案，但由于其复杂度更高，且在此情境下没有明显优势，最终未予采用。

本研究评估所选建模方法能否根据第 4 节详述的两栋实际有人居住的住宅建筑的长期测量数据，以足够的精度预测区域温度。这些数据来自常规运行，也就是说，并未为系统辨识进行主动激励。此外，测量数据存在缺口，且过程噪声和测量噪声的水平未知。在这些具有挑战性的条件下进行测试，对于得出未来大规模部署于实际建筑的结论至关重要。

2. 理论背景

2.1. 数据赋能预测控制

数据赋能预测控制（DeePC）是一种基于行为系统理论的模型预测控制（MPC）表述形式。该方法由 Coulson 于 2019年首次发表并命名[15]。与经典模型预测控制不同，它不需要对被控对象进行建模。预测控制问题的解直接由被控对象的测量轨迹线性组合生成。本节将简要介绍该方法：

假设一个阶数为 n_{sys} 的线性时不变系统 $\mathcal{B}$，其离散时间输入输出样本为 $z_k = \begin{bmatrix} u_k & y_k \end{bmatrix}^\top$，并且测量值为 T 个样本 $z_{[1:T]}$。

Willems 基本引理 [16] 表明，系统 $\mathcal{B}$ 的任何长度为 T_f 的可能轨迹，都可以构造为深度为 $L = n_s\ _s + T$ 的汉克尔矩阵各列的线性组合：

$$\mathcal{H}_L(z_{[1:T]}) = \begin{bmatrix} z_1 & \dots & z_{T-L+1} \\ \vdots & \ddots & \vdots \\ z_L & \dots & z_T \end{bmatrix} \begin{matrix} y \\ \\ f \end{matrix} \text{ if } \mathcal{H}_L(u_{[1:T]}) \text{ has full row rank.}$$

现在考虑汉克尔矩阵 $\mathcal{H}_L(u_{[1:T]}) = \begin{matrix} \mathcal{H}^u_{T_{ini}} \\ \mathcal{H}^u_{T_f} \end{matrix}$ 和 $\mathcal{H}_L(y_{[1:T]}) = \begin{bmatrix} \mathcal{H}^y_{T_{ini}} \\ \mathcal{H}^y_{T_f} \end{bmatrix}$，其中每个矩阵都被分为用于初始化和预测的两个组成部分，并合并为 $\mathcal{H}_L(z_{[1:T]}) = \begin{bmatrix} \mathcal{H}^u_{T_{ini}} \\ \mathcal{H}^y_{T_{ini}} \\ \mathcal{H}^u_{T_f} \\ \mathcal{H}^y_{T_f} \end{bmatrix}$。

以及右侧向量 $v = \begin{pmatrix} u_{ini} \\ y_{ini} \\ u \\ y \end{pmatrix}$，它由用于初始化的 T_{ini} 最近样本 u_{ini} 和 y_{ini}，以及优化变量 u 和 y 组成。

将 $\mathcal{H}_L(z_{[1:T]})$ 简化为 $\mathcal{H}$，并引入优化变量 g（表示 $\mathcal{H}$ 各列的线性组合）后，预测控制器的公式如下：

$$\begin{aligned} &\min_{g,u,y} J(u,y) \\ &\text{subject to} \\ &\mathcal{H}g = v \\ &u \in \mathcal{U}, y \in \mathcal{Y} \end{aligned}$$

自该方法首次提出以来，已有许多采用这一方法的研究发表。[17]给出了最近的一篇文献综述。

2.2. 不同轨迹矩阵的持续激励

持续激励（PE）是使底层系统可辨识的测量数据集所需满足的条件。Willems 基本引理针对汉克尔矩阵形式的数据阐述了这一条件。后续文献将其扩展至镶嵌汉克尔矩阵和 Page矩阵，如表 1 所示。镶嵌汉克尔矩阵是多个汉克尔矩阵的水平拼接。Page矩阵与汉克尔矩阵类似，但不含重复条目。它包含一段连续的样本序列，类似页面上的文字，因此得名。我们注意到，Page矩阵的条件非常严格，虽是充分条件，却不是必要条件。事实上，尽管严重违反了该论文提出的持续激励条件，仍成功使用 Page矩阵控制了

矩阵	持续激励条件	最小 T
汉克尔 [18]	$\mathcal{H}_L(u_{[1,T]})$ 满行秩	$T \geq L(m+1) - 1$ 的阶数为 L
镶嵌汉克尔 [19]	$\begin{bmatrix}\mathcal{H}_k(u^1_{[0,T_1-1]}) & \mathcal{H}_k(u^2_{[0,T_2-1]}) & \dots & \mathcal{H}_k(u^q_{[0,T_q-1]})\end{bmatrix}$ 具有满行秩	$\sum_{i=1}^{q} T_i \geq k(m+q) - q$ 阶数为 k $len(T_i) \geq k$ 必要条件
第 [18] 页	$\begin{pmatrix}\mathcal{P}_L(u_{[1,T-(M-1)L]}) \\ \mathcal{P}_L(u_{[L+1,T-(M-2)L]}) \\ \vdots \\ \mathcal{P}_L(u_{[L(M-1)+1,T]})\end{pmatrix}$ 具有满行秩	$T \geq L((mL+1)M-1)$ 阶数为 k 充分条件

表1：不同轨迹矩阵结构的持续激励条件及最小采样点数

在 [18], 中的一架四旋翼无人机。这一例子说明，实验研究对于补充该领域近期发表的理论研究十分重要。

在我们的研究中，我们使用由单独、互不相连但彼此重叠的轨迹组成的类Hankel矩阵。据我们所知，尚未针对这种非结构化矩阵制定持续激励条件。鉴于上述研究结果，在没有可用的适配持续激励条件的情况下，我们对真实数据开展了研究。

2.3. 双重控制中的持续激励

在采用自适应方案的情况下，汉克尔矩阵（或任何其他数据结构）会持续更新，因此持续确保持续激励是一项挑战。已发表的提高自适应数据驱动控制器激励水平的方法可分为三类：第一类是从较长的数据日志中选择合适的数据；第二类是在求解得到的预测控制轨迹中加入扰动；第三类是在代价函数中纳入激励水平 [20]。然而，建筑物会受到天气的显著扰动，而天气无法控制，因此第一种方法更具优势。此外，人为提高控制信号的激励水平可能会增加执行器的磨损，也可能导致约束违反。对于建筑物而言，由此产生的温度波动还可能降低居住者舒适度。

2.4. 建筑控制中的 *DeePC*

DeePC 是针对确定性线性时不变系统推导出来的。由于建筑物并不能完全满足这些条件，因此需要通过实验和高保真仿真来评估实际可达到的性能水平。

现有文献较少。Schwarz [21], O'Dwyer [22] 和 Chinde[23] 开展了关于 DeePC 在建筑物中应用的仿真研究。然而，这些研究并未采用自适应方法。Kerkhof [24] 开展了一项关于温室中自适应 DeePC 的仿真研究，但采用的预测时域很短，只有一小时。此外，温室的动态特性可能与传统建筑大不相同。Lian [25] 在一座教育建筑中开展了关于自适应鲁棒双层 DeePC 的实验研究。Kerkhof 和 Lian 均在控制信号中加入人工噪声，以确保持续激励。

3. 方法

由于 DeePC 指的是一种控制方法，而我们关注的是系统辨识，因此下文将基于行为系统理论的方法称为"BST"，而非"DeePC"。对于每栋建筑，共评估 69 种建模方法：一种非自适应 ARX、三种采用不同遗忘因子的自适应 ARX、五种采用不同正则化权重的非自适应 BST，以及 60 种采用不同轨迹矩阵宽度、正则化权重和数据选择方法的自适应 BST。非自适应变体在评估阶段之前进行辨识。自适应变体则在每一步更新，并在评估阶段之前经过 30 天的初始化阶段。对于所有方法，初始化使用 12 个时间步，预测使用 96 个时间步，分别对应 $3h$ 和 $24h$。

预测所用的天气预报是理想的。然而，自适应 BST 变体的数据选择采用了经过操控的预报，详见第 3.1.2 节。

与建筑物长期现场测量中的典型情况一样，数据中存在许多缺口，原因是传感器未返回测量值或返回了物理上不可能的值。这些缺口会在预处理中被识别。长度为 $T_{ini} + T_f$ 或更长的不间断轨迹会被标记为可用于系统辨识和验证。不符合条件的片段将被跳过。对于自适应 ARX，遇到缺口时会暂停模型更新。对于非自适应 ARX，在存在缺口的情况下会采用多批次辨识方法。

*3.1. BST*变体详解

每个时间步都会根据当前状态和天气预报预测区域温度。控制变量为供暖和制冷功率 $u = \begin{bmatrix} P_{heat} & P_{cool} \end{bmatrix}$（巴塞尔建筑仅供暖）。扰动变量为环境温度和太阳辐射 $w = \begin{bmatrix} T_{amb} & I_{sol} \end{bmatrix}$。输出变量为区域温度 $y = T_z$。加入扰动变量后，第 2.1 节中的 DeePC 公式变为：

$$\mathcal{H}_L(z_{[1:T]}) = \begin{bmatrix} \mathcal{H}^u_{T_{ini}} \\ \mathcal{H}^w_{T_{ini}} \\ \mathcal{H}^y_{T_{ini}} \\ \mathcal{H}^u_{T_f} \\ \mathcal{H}^w_{T_f} \\ \mathcal{H}^y_{T_f} \end{bmatrix} \text{ and } v = \begin{pmatrix} u_{ini} \\ w_{ini} \\ y_{ini} \\ u \\ w_{forecast} \\ y \end{pmatrix}$$

我们进一步定义 $\hat{\mathcal{H}} = \begin{bmatrix} \mathcal{H}^u_{T_{ini}} \\ \mathcal{H}^w_{T_{ini}} \\ \mathcal{H}^y_{T_{ini}} \\ \mathcal{H}^u_{T_f} \\ \mathcal{H}^w_{T_f} \end{bmatrix}$ 为 $\mathcal{H}_L(z_{[1:T]})$ 的前五个元素，并定义 $\hat{v} = \begin{pmatrix} u_{ini} \\ w_{ini} \\ y_{ini} \\ u \\ w_{forecast} \end{pmatrix}$ 为 v 的前五个元素。

最小化由最小二乘拟合和对 g 的正则化组成的以下代价函数：

$$J = (\hat{\mathcal{H}}g - \hat{v})^\top W_I (\hat{\mathcal{H}}g - \hat{v}) + \lambda g^\top g \tag{1}$$

其解析解为：

$$g^* = (\hat{\mathcal{H}}^\top W_I \hat{\mathcal{H}} + \lambda I)^{-1} \hat{\mathcal{H}}^\top W_I \hat{v} \tag{2}$$

$$y^* = \mathcal{H}^y_{T_f} g^* \tag{3}$$

W_I 是一个加权矩阵，其赋予初始化步数的权重是预测步数的 100 倍。这样做是为了提高较早预测步的准确性，这些预测步对于预测控制更为关键，因为通常只应用优化轨迹的第一个元素，随后便会重新进行整个优化。

自适应变体的轨迹矩阵宽度记为 W_H，分别为 181、373 和 661。在适当的汉克尔矩阵中，这对应于 3、5 和 8 天的数据（$W_H = T - T_{ini} - T_f + 1$）。对于非自适应 BST，则使用辨识阶段的全部数据来构建一个宽达数千列的矩阵。

正则化权重 λ 为 $10^0, 10^1, 10^2, 10^3$ 和 10^4。

3.1.1. 数据选择方法详解

我们研究了构造轨迹矩阵的不同方式所产生的影响。一种方法尽可能接近规范的汉克尔矩阵，另外还有三种替代方法。下文将介绍这四种不同的方法。图 1 给出了简化示意图。

最近：选取最近的 W_H 条可行轨迹。如果数据中没有缺口，则得到一个规范的汉克尔矩阵。如果存在缺口，则得到一个马赛克汉克尔矩阵。

相关性最高：选择天气与预报相关性最高的 W_H 可行轨迹。此方法旨在研究，与前述方法相比，选择形状类似于目标轨迹的轨迹是否更有优势。

最小均方根误差：选择天气相对于预报的均方根误差最小的 W_H 可行轨迹。此方法与相关性方法类似，但对缩放和偏移不具不变性。它类似于参数变体模型，其中天气是所关注的参数。由于 BST方法具有多步特性，可能存在的参数方差会被隐式捕捉，而非显式建模。由于环境温度和太阳辐射通常呈现日变化规律，该方法还包含时变建模的要素。

最接近均值：选取平均天气与预报对应值最接近的 W_H 可行轨迹。该方法旨在兼顾选择最近的数据和选择相似的数据。由于平均环境条件通常以周和月为尺度变化，因此预计这种方法与选择最近的数据相似。

所有方法的数据选择范围均限于过去 365 天。由于计算以大约 30 天的数据初始化，可用数据量会随时间增加，直至填满一年的时间窗口。虽然在整个计算期间保持数据量恒定更为理想，但受限于数据集，这无法实现。

天气数据由归一化后的环境温度和太阳辐射组成。计算时使用两个相关系数、均方根误差和差值之和。环境温度归一化至 $\hat{}T_{amb} = \frac{T_{amb}-10[K]}{20[K]}$ 。这会将范围 $[-10°C, 30°C]$ 压缩至 $[-1, 1]$。对于太阳辐射，归一化方式根据第 4 节详述的具体传感器情况进行调整。苏黎世为 $\hat{} \ I_{sol} = I_{sol}/(500[W/m^2])$，巴塞尔为 $\hat{} \ P_{PV} = P_{PV}/(3[kW])$。选择这些值是为了将数值重新缩放至约 $[0, 1]$。不同的归一化范围可抵消这样一个事实：在 $24h$期间，环境温度通常只处于归一化范围内较窄的区间。另一方面，在一个夏日内，太阳辐射通常会跨越其归一化范围的大部分。

尽管这些方法并未产生实际的汉克尔矩阵，但为简便起见，我们仍将它们统称为符号 $\mathcal{H}$ 。

3.1.2. 用于数据选择的操纵后的天气预报

基于选择相似天气的变体与其他变体之间存在一个关键差异：它们会两次受到天气预报不准确性的影响。一次是在选择步骤中，另一次是在预测步骤中。虽然我们在预测中对所有变体都使用理想预报，以实现公平比较，但我们会操纵数据选择步骤中的预报，以模拟这种双重依赖关系的影响。

据我们所知，目前没有现成的工具能够基于测量数据生成人工预报。因此，我们定义了自己的方法。虽然真实的（或至少逼真的）预报更为理想，但我们认为，就本研究的范围而言，该方法能够合理地模拟真实预报的不准确性，从而达到目的。出于同样的原因，

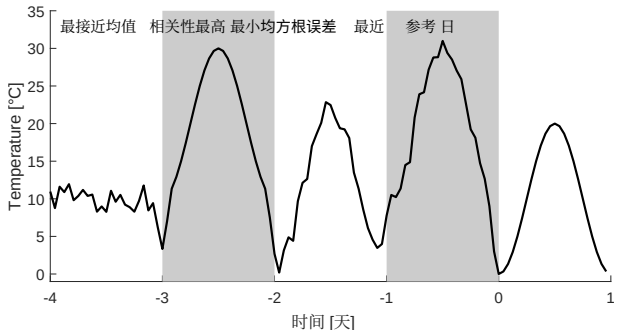

图1：数据选择变体的简单示意图。将过去四天与一个预测参考日进行比较，并将其归类为最接近均值、相关性最高、最小均方根误差和最近。

[26] 也依赖一种自行 разработ的方法来扭曲天气数据。

生成一个随机正弦函数，使其在 $24h$的预测时域内线性增长。对于环境温度，该函数以加法方式使用，从 $0K$ 增长至最大值 $\pm 4K$。对于太阳辐射，该函数以乘法方式使用，从1增长至最大值 $\pm 15\%$。图2展示了乘法情形下的一些示例轨迹。

$$
\begin{aligned}
&f_{scale} = 1 + 0.15\frac{t}{24}sin(at+b) \\
&\text{with} \\
&a = \mathcal{U}(0.5\frac{\pi}{12}, 1.5\frac{\pi}{12})(\text{Uniform distribution}) \\
&b = \mathcal{U}(-\pi, \pi) \\
&t = (0, 24)(\text{Hour of day})
\end{aligned}
$$

*3.2. ARX*变体详解

ARX模型具有以下结构 [27]:

$$y(k) = a_1 y(k-1) + ... + a_{n_a} y(k-n_a) = b_1 u(k-n_k) + ... + b_{n_b} u(k-n_b-n_k+1) \quad (4)$$

其中， n_a 为极点数量， n_b 为零点数量， n_k 为输入影响输出之前出现的输入样本数量。参数通过最小二乘拟合进行估计。整个预测时域内的预测通过将单步预测模型迭代串联 T_f 次来完成。

对于自适应 ARX，采用遗忘因子法 [28]:

$$\hat{\Theta}(k) = \hat{\Theta}(k-1) + K(k)(y(k) - \hat{y}(k)) \quad (5)$$

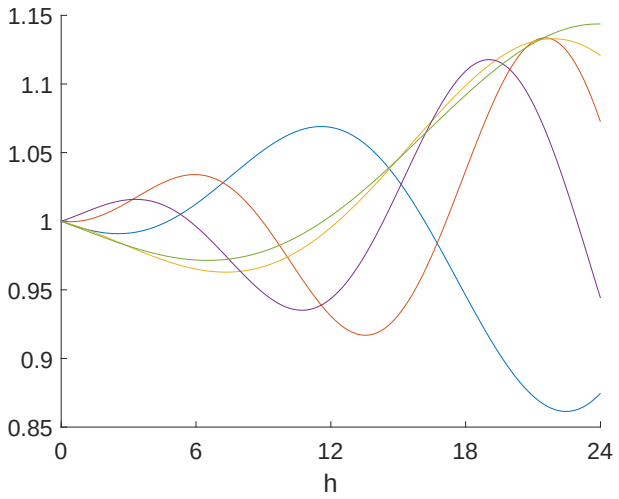

图 2：天气预报偏差的五个示例轨迹

其中，$\hat{\Theta}(k)$ 是时间 k 时的参数估计值，$y(k)$ 是测量值，$\hat{y}(k)$ 是对该测量值的预测值。增益形式如下：

$$K(k) = Q(k)\Psi(k) \tag{6}$$

其中，$\Psi(k)$ 是预测输出 $\hat{y}(k)$ 关于 Θ 的梯度。我们进一步定义：

$$Q(k) = \frac{P(k-1)}{\alpha + \Psi^{\top}(k)P(k-1)\Psi(k)} \tag{7}$$

and

$$P(k) = \frac{1}{\alpha}\left(P(k-1) - \frac{P(k-1)\Psi(k)\Psi(k)^{\top}P(k-1)}{\alpha + \Psi(k)^{\top}P(k-1)\Psi(k)}\right) \tag{8}$$

$$P(0) = 10000 \tag{9}$$

遗忘因子为 α。所使用的 α 具体取值根据关系式 $\alpha = 1 - 1/T$ 选取，使时间常数分别为 3、5 和 8 天。$P(0)$ 是所用 MATLAB 函数的默认值。

3.3. 分析

预测质量通过预测值与实际值之间的均方根误差（RMSE）进行量化。对于误差轨迹，采用每个预测步的绝对误差的均值和标准差（STD）。

4. 数据集

使用了两个数据集。第一个数据集是位于苏黎世附近的一套公寓，于 2018 [29]竣工。它位于一个小型垂直堆叠式研究社区的二楼。其建造注重使用木材和再生材料，以及从真菌中提取的实验性建筑材料。公寓由一间客厅、两间卧室、两间浴室和一处小型入口空间组成。本研究仅考虑客厅和卧室，这些房间的总面积为 $93.8m^2$。测量并取这些房间的温度均值。供暖和制冷由辐射板提供，冷热水由中央系统供应。在公寓的连接点测量热功率。在当地测量环境温度和太阳辐射。用于测量太阳辐射的日射强度计与窗户的朝向一致。公寓有一整面朝东北的大窗户，夹在两块向外挑出的混凝土板之间。因此，太阳得热很大程度上取决于太阳的位置，与全球水平辐射可能存在显著差异。

第二处地点是位于巴塞尔的一栋联排公寓楼，于2015 [30]竣工。该楼共有七套公寓，户型为2.5至3.5个房间。这是一栋低能耗建筑，测得的供暖热能需求为 $17kWh/（m^2a）$。建筑采用辐射板供暖，由地源热泵供能。由于仅测量热泵的总电力消耗，因此使用热回收装置前集中通风排风中的测量值作为整栋建筑的集中室内温度，并在图6中标记为 T_z 。环境温度在当地测量。由于没有当地太阳辐射测量值，因此使用屋顶光伏装置的功率作为替代，该装置朝向西北/东南。考虑到其规模、分向布局以及大部分无遮挡的位置，假定它能很好地近似全球水平辐射。窗户尺寸适中且遮阳良好，以尽量减少夏季过热。

两个数据集均涵盖一年，并以 15 分钟为时间步进行采样。原始数据以更高频率测量后再进行重采样。这可以在预处理中消除传感器噪声。该方法也可用于实时场景。

评估选取了大约 365 天的数据；此外，对于自适应变体，还需要大约 30 天的初始化数据；对于非自适应变体的系统辨识，则分别需要 30 天的供暖（和制冷）运行数据。根据供暖和制冷的情况，并为弥补数据中的缺口，这些数值会进行小幅调整，具体如表 2 所示。缺失数据指的是评估周期内的数据。需要注意的是，根据初始化时长和预测时域的长度，一个缺失数据点会导致多个时间步无法进行评估。

	苏黎世¨建筑	巴塞尔建筑
供暖识别	2020年2月17日—2020年3月17日	2016年10月2日—2016年11月22日
制冷识别	2020年8月21日—2020年9月19日	无制冷
初始化	2020年8月21日—2020年9月20日	2016年10月2日—2016年11月22日
评估	2020年9月21日—2021年9月21日	2016年11月23日—2017年11月23日
缺失数据	$\sim 10\%$	$\sim 2.5\%$

表2：数据选择和缺口



图3：苏黎世建筑

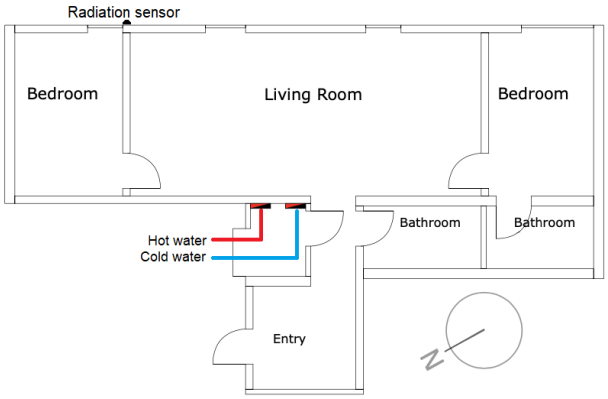

图4：苏黎世建筑平面图



图 5：巴塞尔建筑

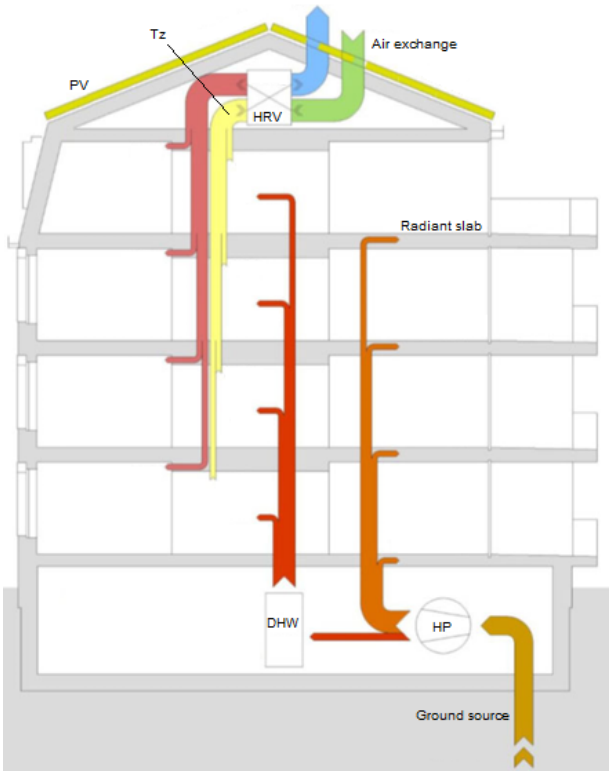

图 6：巴塞尔建筑平面图

5. 结果

图 7 和图 8 展示了 Z¨urich 和巴塞尔建筑的部分特定区域温度预测结果。需要注意的是，后者的温度分辨率较低。以下各节将对所有数据进行多项统计分析。由于基线不同，这两栋建筑将分别进行处理。

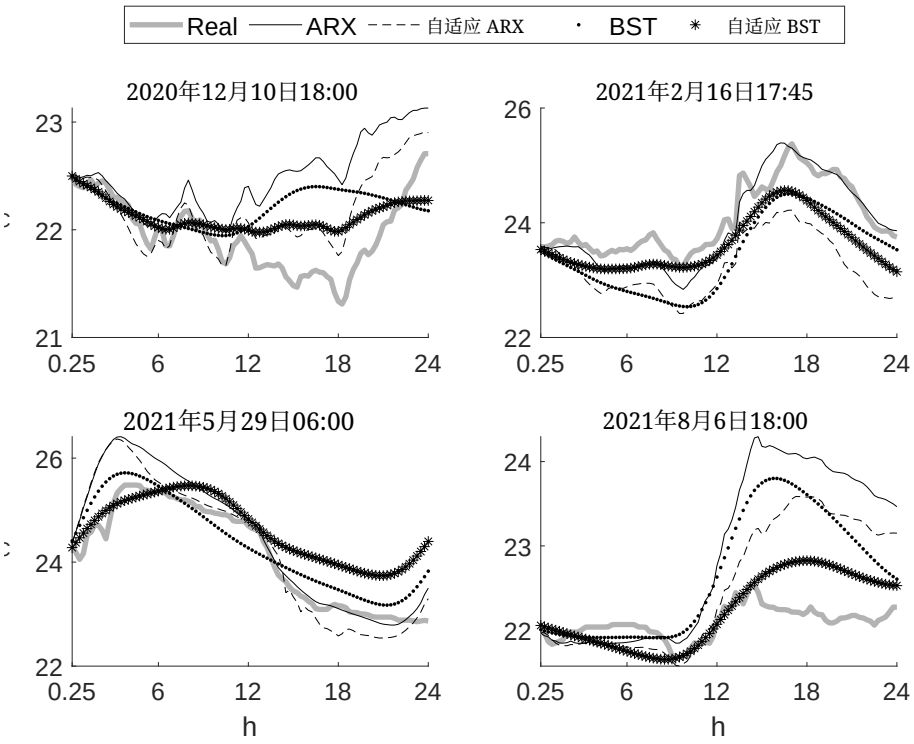

图 7：Z¨urich 建筑在预测时域内指定时间点的选定区域温度预测

5.1. Z¨urich 建筑

图 9 比较了自适应 BST 所有变体在评估阶段的平均预测误差。矩阵越大，性能始终越好。最近 和最接近均值 变体的表现优于相关性最高 和最小均方根误差 变体。

图 10 展示了所有参考方法。具有更高遗忘因子（即更长迹）的自适应 ARX 表现最佳。根据正则化权重的不同，BST变体的表现与 ARX变体相近。

为进一步比较，从四个类别（ARX 与 BST；自适应与非自适应）中分别选出表现最佳的方案。对于自适应 BST，表现最佳的

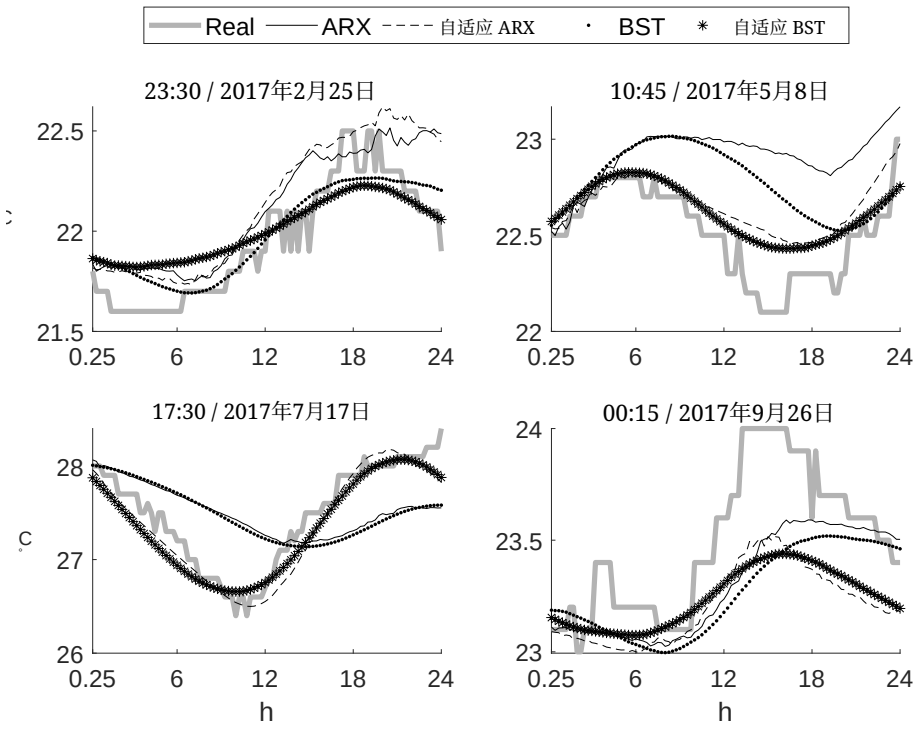

图8：巴塞尔建筑在预测时域内指定时间的部分区域温度预测

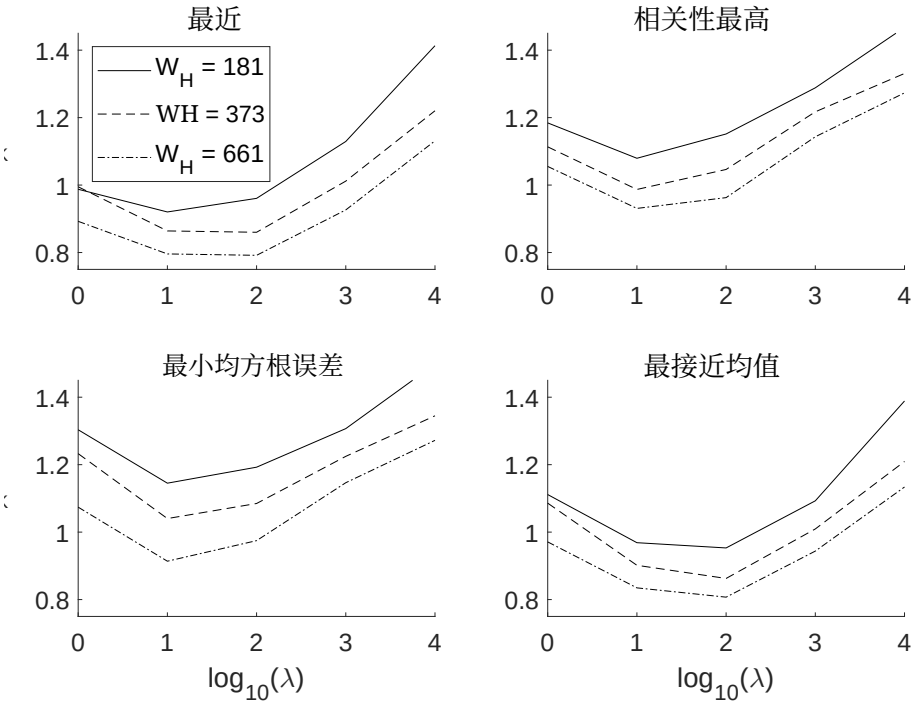

图9：Z¨urich 自适应 BST 平均温度预测误差的比较

子变体是最近的，具有最宽的轨迹矩阵和中等正则化权重。

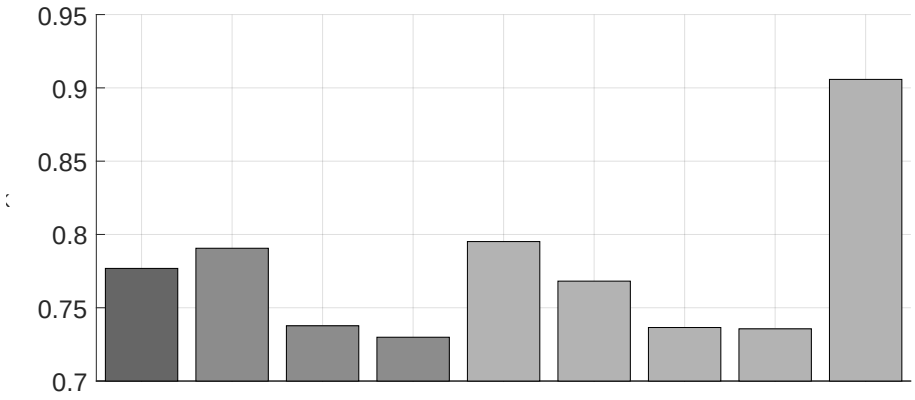

图10：苏黎世参考控制器平均温度预测误差的比较

图11和图12显示了预测时域内的平均预测误差及相应的标准差。所有变体的表现相似。

图13显示了全年预测误差及三次拟合。图14将所有拟合与环境温度进行了比较。未发现明显规律。

5.2. 巴塞尔建筑

图15比较了自适应BST所有变体在评估阶段的平均预测误差。与之前一样，轨迹矩阵越宽，性能越好。中等正则化权重的最近变体表现最佳。其他所有变体的表现都很接近。

图16展示了所有参考方法。两种自适应 ARX 变体的表现显著优于非自适应变体。BST变体的表现与 ARX 相近，具体取决于正则化权重。

与之前一样，从四个主要类别中各选出表现最佳的变体进行比较。

图17和图18展示了不同预测时域下的平均预测误差及相应的标准差。总体而言，在这种情况下，自适应变体优于非自适应变体，其中 ARX方法略占优势。

图19展示了全年预测误差及三次拟合。图20将所有拟合结果与环境温度进行了比较。在仿真时段中段附近，自适应和非自适应变体的误差均出现明显尖峰。ARX与BST之间的相似性表明

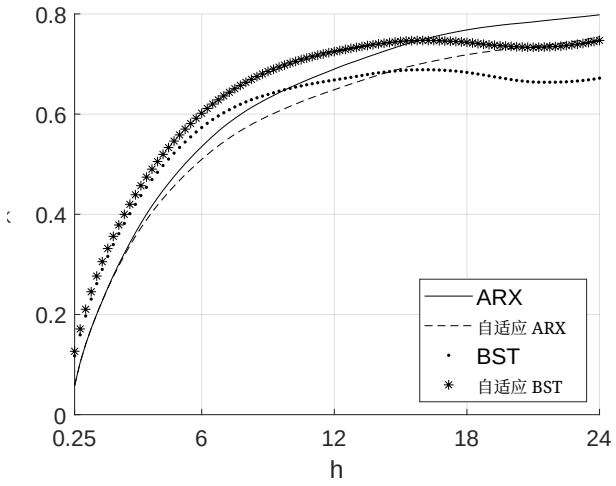

图11：Z¨urich 随时间变化的平均预测误差

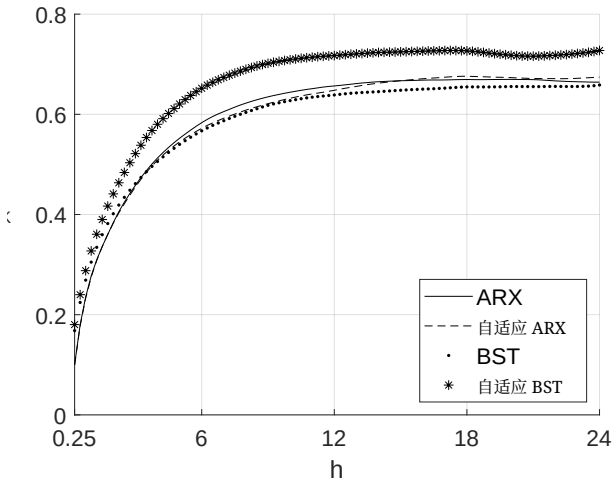

图12：Z¨urich 随时间变化的预测误差标准差

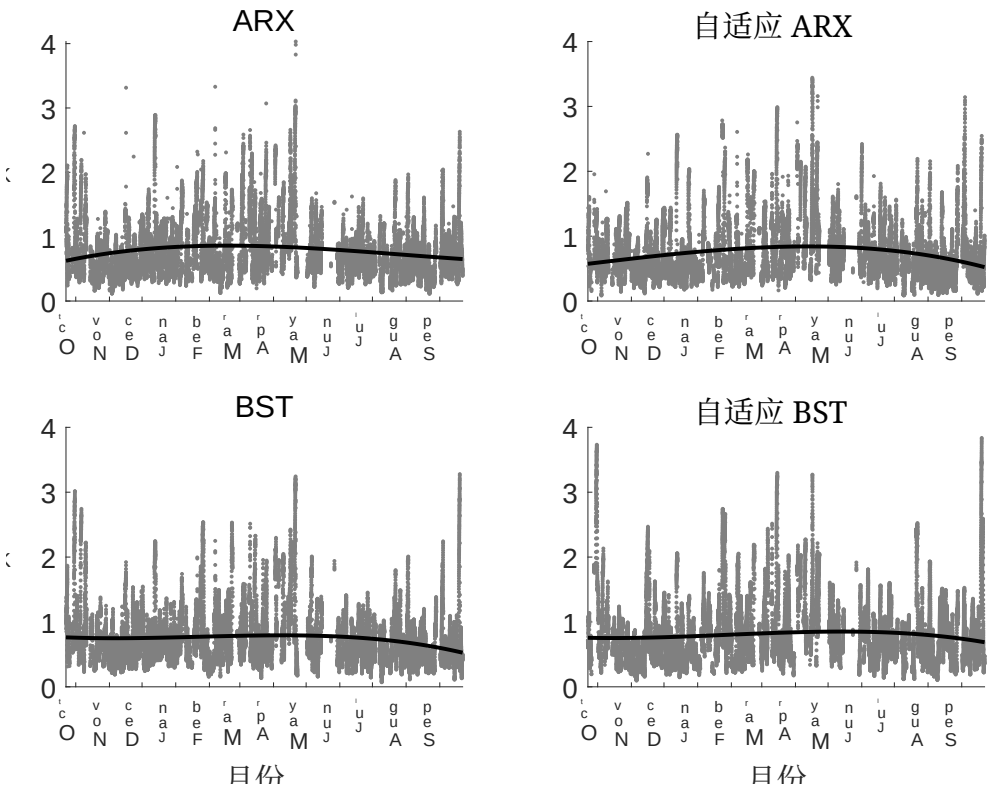

图13：Z¨urich 的季节性预测误差

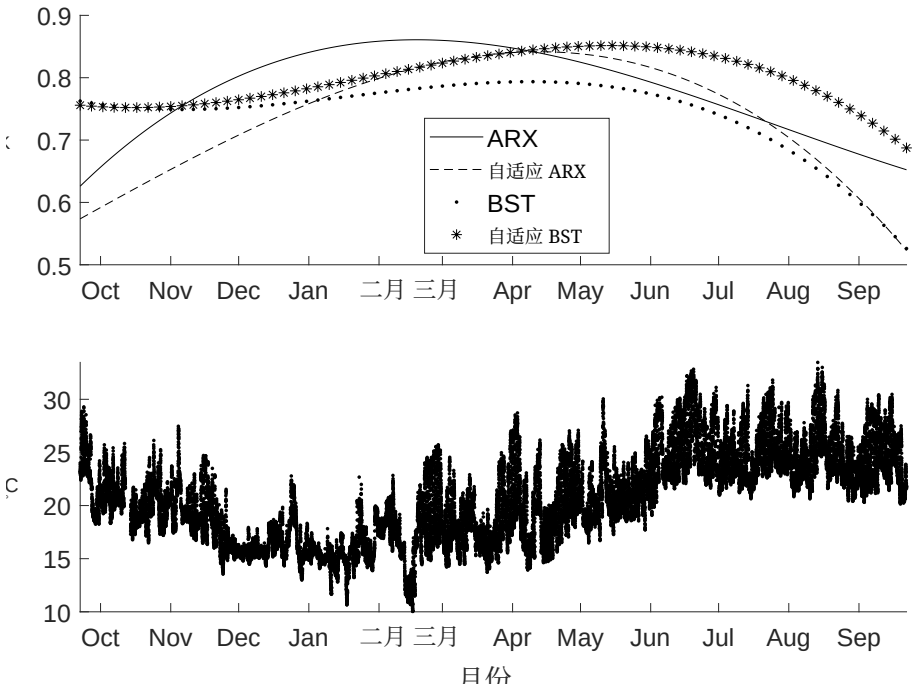

图14：苏黎世的季节性预测误差拟合比较和环境温度

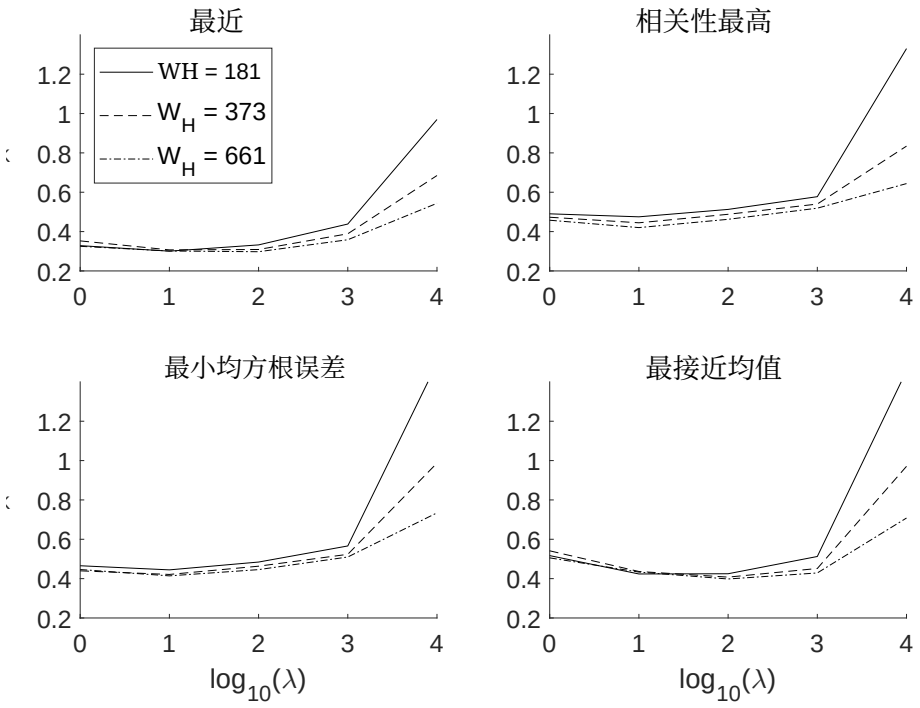

图15：巴塞尔自适应 BST 平均温度预测误差的比较

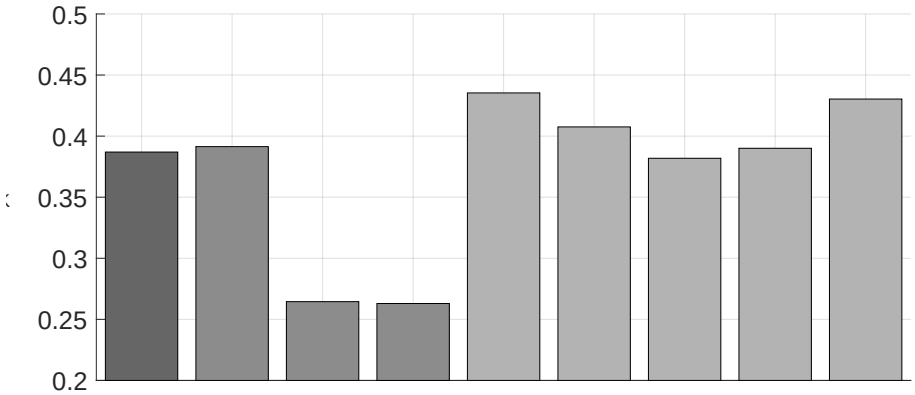

图16：巴塞尔参考控制器的平均温度预测误差比较

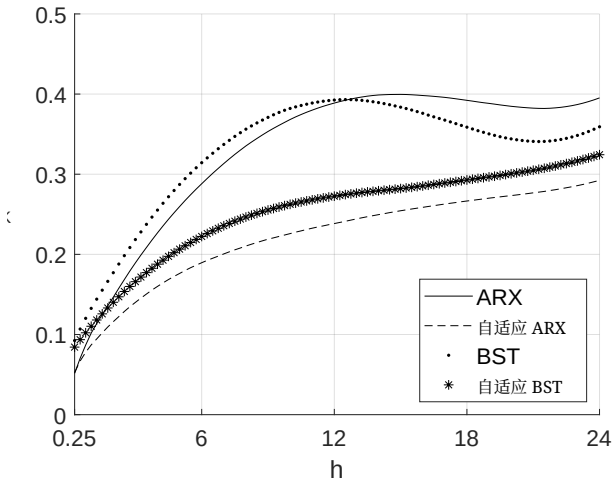

图17：巴塞尔随时间变化的平均预测误差

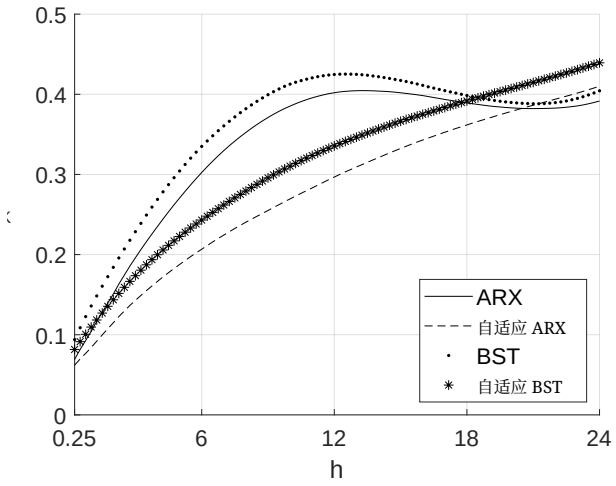

图18：巴塞尔随时间变化的预测误差标准差

（未知）物理原因，而非方法本身的问题。图20显示，自适应变体在夏季月份的表现显著更好。

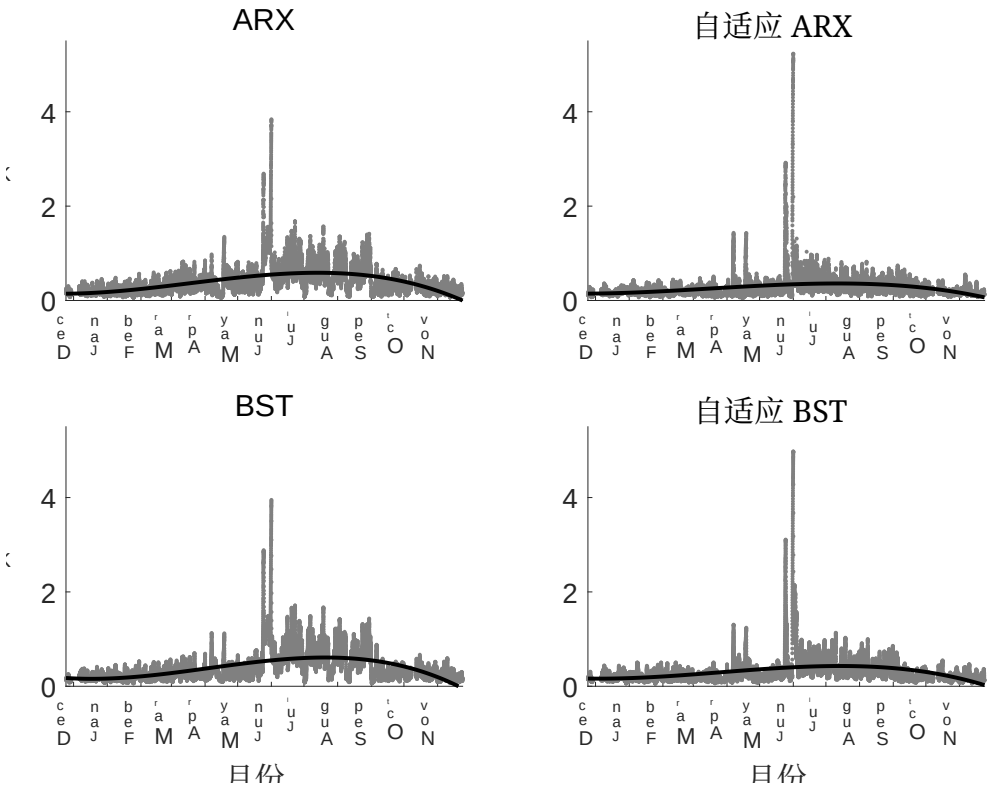

图19：巴塞尔的季节性预测误差

5.3. 关于轨迹矩阵奇异值的说明

在双重控制中，代价函数通常包含一项，用于最大化所得轨迹 [31, 32, 33, 34, 35, 36] 中的激励水平。根据 Willems 引理，失去满行秩意味着汉克尔矩阵失去持续激励。矩阵的最小奇异值可以视为衡量矩阵距离这种状态有多近的指标。因此，我们假设轨迹矩阵的最小奇异值与所得预测准确度之间存在相关性。然而，我们并未发现这种相关性。

6. 讨论与结论

本研究的一个局限是数据集较小，仅包含两栋气候相似的建筑物。若能涵盖更多样的建筑类型和气候条件，将会更理想，但当时没有可用的相应数据集。

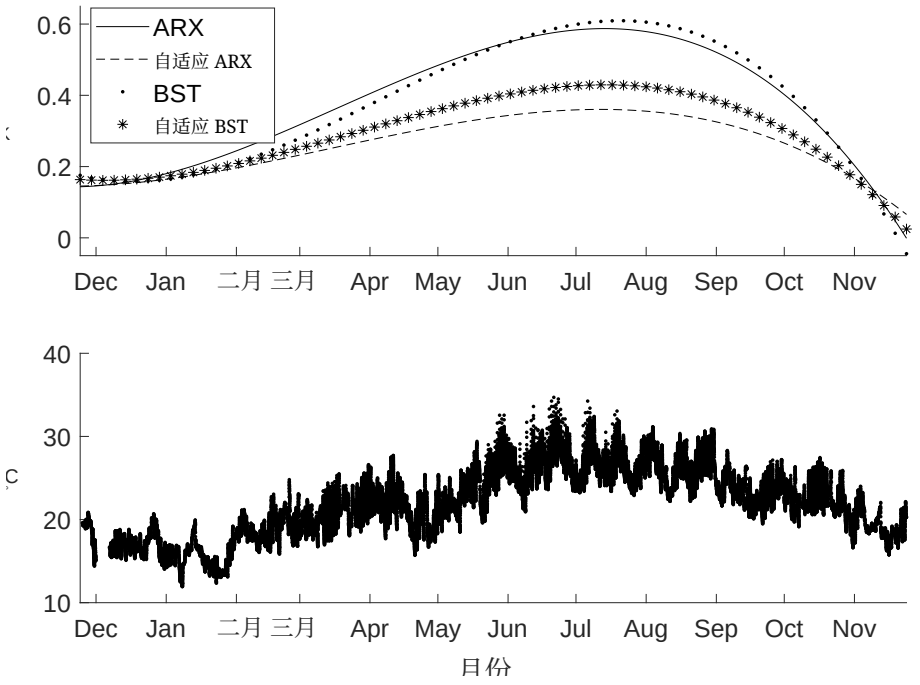

图20：巴塞尔的季节性预测误差拟合比较和环境温度

巴塞尔建筑的预测误差明显小于苏黎世建筑。这可能是因为前者规模更大，窗墙面积比更小，因此巴塞尔建筑的惯性与扰动之比也高得多。

BST和ARX方法的平均预测准确度相近，ARX略占优势。不过，ARX的计算速度快了几个数量级。 [21] 和 [24] 报告的DeePC与传统MPC计算时间之比也相似。此外，ARX需要调整的参数更少。除去BST中正则化权重 λ 的最差取值，从应用角度来看，所有方法的准确度都足够高，在 $24h$ 的预测时域内，预测误差远低于 $1K$ 。更普遍地说，线性时不变方法足以预测所研究建筑物的区域温度。值得注意的是，在未为系统辨识而有意施加激励的情况下，仍能达到这一准确度。

ARX在预测的前几个预测步中具有明显优势，但这一优势在后续预测步中逐渐减弱。对于大多数预测控制应用而言，前几个预测步的准确度更为重要。不过，对于某些特定应用而言，长期预测时域内保持更稳定的准确度可能更为理想。例如，某些需求响应形式所需的电力消耗日前预测 [9]。

比较选择轨迹矩阵数据的不同方法，使用

最近的数据与根据相似天气选择数据相比，准确度相当或更高。这也是最容易实现且计算速度最快的方法。

尽管数据集包含季节性过渡期，自适应和非自适应方法的表现仍相近。然而，相对于建筑材料发生显著劣化所需的时间尺度而言，一年仍然很短[37, 38]。此外，数据集不包含建筑物的建筑占用情况、用途或周边环境的任何重大变化。考虑此类长期变化并不需要采用本研究使用的连续模型更新方法。采用 [11] 中所用的定期更新应当足够。

对于数据中的缺口，研究发现，BST 情况下直接忽略缺口、ARX 情况下跳过缺口，是最简单可行且效果良好的方法。

7. 利益冲突声明

作者声明，据其所知，他们不存在可能影响本文所报告研究工作的财务利益冲突或个人关系。

8. 数据可用性

本研究所用数据属于机密信息。

9. 致谢

我们谨向 Interreg 计划致谢，感谢其提供支持。此外，我们还要感谢 EMPA 的 Loris Di Natale 和 FHNW 的 Gregor Steinke 提供本研究所用的数据集，并感谢 FHNW 的 Ralph Eismann 提供咨询建议。

参考文献

[1] L. P´erez-Lombard、J. Ortiz、C. Pout，建筑能耗信息综述，Energy and Buildings 40 (3) (2008) 394–398。doi: 10.1016/j.enbuild.2007.03.007。.[2] D. Urge-Vorsatz、L. F. Cabeza、S. Serrano、C. Barreneche、K. Petrichenko，建筑物供暖与制冷能耗趋势及其驱动因素，《可再生与可持续能源评论》41 (2015) 85–98。doi:10.1016/j.rser.2014. 08.039。[3] M. Isaac、D. P. van Vuuren，气候变化背景下全球住宅部门供暖与空调能源需求建模，Energy Policy 37 (2) (2009) 507–521。doi:10.1016/j.enpol.2008.09.051。

[4] 制冷的未来——分析，https://www.iea.org/reports/the-future-of-cooling。[5] G. Serale、M. Fiorentini、A. Capozzoli、D. Bernardini、A. Bemporad，模型预测控制（MPC）提升建筑与暖通空调系统能效：问题表述、应用与机遇，Energies 11 (3) (2018) 631。doi:10.3390/en11030631。[6] A. Afram、F. Janabi-Sharifi，暖通空调控制系统理论与应用——模型预测控制（MPC）综述，建筑与环境72 (2014) 343–355。doi:10.1016/j.buildenv.2013.11.016。[7] A. Mirakhorli、B. Dong，基于占用行为的建筑室内气候模型预测控制——批判性综述，Energy and Buildings 129 (2016) 499–513。doi:10.1016/j.enbuild.2016.07.036。[8] J. Drgoˇna、J. Arroyo、I. Cupeiro Figueroa、D. Blum、K. Arendt、D. Kim、E. P. Oll´e、J. Oravec、M. Wetter、D. L. Vrabie、L. Helsen，关于建筑模型预测控制，你需要了解的一切，Annual Reviews in Controldoi:10.1016/j.arcontrol.2020.09.001。[9] F. A. Qureshi、T. T. Gorecki、C. N. Jones，用于参与基于市场的需求响应的模型预测控制，IFAC会议论文集47 (3) (2014) 11153–11158。doi:10.3182/20140824-6-ZA-1003.02395。[10] M. Killian、M. Kozek，关于节能建筑模型预测控制的十个问题，建筑与环境105 (2016) 403–412。doi:10.1016/j.buildenv.2016.05.034。[11] H. Wolisz、T. M. Kull、D. M¨uller、J. Kurnitski，用于动态激活住宅建筑结构热质量的自学习模型预测控制，Energy and Buildings 207 (2020) 109542。doi:10.1016/j.enbuild.2019.109542。[12] D. Sturzenegger、D. Gyalistras、M. Morari、R. S. Smith，瑞士办公楼模型预测气候控制：实施、结果与成本效益分析，IEEE控制系统技术汇刊24 (1) (2016) 1–12。doi:10.1109/TCST.2015.2415411。[13] E. T. Maddalena、Y. Lian、C. N. Jones，建筑控制的数据驱动方法——综述与未来发展方向，Control Engineering Practice 95 (2020) 104211。doi:10.1016/j.conengprac.2019.104211。[14] T. Hilliard、M. Kavgic、L. Swan，商业建筑的模型预测控制：趋势与机遇，Advances in Building Energy Research 10 (2) (2016) 172–190。doi:10.1080/17512549.2015.1079240。[15] J. Coulson、J. Lygeros、F. D¨orfler，数据赋能预测控制：探索 DeePC 的浅水区，载于：2019年第18届欧洲控制会议（ECC），2019年，第307–312页。doi:10.23919/ECC.2019.8795639。

[16] J. C. Willems, P. Rapisarda, I. Markovsky, B. L. M. De Moor，关于持续激励的说明，Systems & Control Letters 54 (4) (2005) 325–329. doi:10.1016/j.sysconle.2004.09.003.[17] I. Markovsky, F. D¨orfler，数据驱动分析、信号处理与控制中的行为系统理论，Annual Reviews in Control 52 (2021) 42–64. doi:10.1016/j.arcontrol.2021.09.005.[18] J. Coulson, J. Lygeros, F. Dorfler，分布鲁棒机会约束数据赋能预测控制，IEEE Transactions on Auto- matic Control (2021) 1–1doi:10.1109/TAC.2021.3097706.[19] H. J. van Waarde, C. De Persis, M. K. Camlibel, P. Tesi，状态空间系统的 Willems 基本引理及其对多个数据集的扩展，IEEE Control Systems Letters 4 (3) (2020) 602–607. doi:10.1109/LCSYS.2020.2986991.[20] J. Berberich, J. K¨ohler, M. A. M¨uller, F. Allg¨ower，非线性系统的线性跟踪模型预测控制 第二部分：数据驱动方法, arXiv:2105.08567 [cs, eess, math]arXiv:2105.08567.[21] J. Schwarz，建筑与能源枢纽的数据驱动控制doi:10. 3929/ethz-b-000493171.[22]E. O'Dwyer, E. C. Kerrigan, P. Falugi, M. A. Zagorowska, N. Shah，利用分段轨迹降低计算量并提升性能的数据驱动预测控制，arXiv:2108.10753 [cs, eess] arXiv:2108.10753.[23] V. Chinde, Y. Lin, M. J. Ellis，建筑暖通空调系统的数据赋能预测控制，Journal of Dynamic Systems, Measurement, and Con-troldoi:10.1115/1.4054314.[24] L. Kerkhof，自主温室的最优控制：一种数据驱动方法，博士论文，代尔夫特理工大学 (2020).[25] Y. Lian, J. Shi, M. P. Koch, C. N. Jones，通过双层重构实现自适应鲁棒数据驱动建筑控制：实验结果，arXiv:2106.05740 [cs, eess]arXiv:2106.05740.[26] P. Scharnhorst, B. Schubnel, C. Fern´andez Bandera, J. Salom, P. Taddeo, M. Boegli, T. Gorecki, Y. Stauffer, A. Peppas, C. Politi，Energym：用于控制器基准测试的建筑模型库，《应用科学》11 (8) (2021) 3518. doi:10.3390/app11083518.[27]估计 ARX、ARIX、AR 或 ARI 模型的参数 - MATLAB arx - MathWorks Switzerland, https://ch.mathworks.com/help/ident/ref/arx.html。

[28]创建用于 ARX 模型在线参数估计的系统对象 - MATLAB recursiveARX - MathWorks Switzerland, https://ch.mathworks.com/help/ident/ref/recursivearx.html?s tid=doc ¯ ta.[29] L. Di Natale, B. Svetozarevic, P. Heer, C. N. Jones，用于建筑热建模的物理一致神经网络：理论与分析，arXiv:2112.03212 [cs, eess] arXiv:2112.03212.[30]Minergie-A-Eco MFH Aescherstrasse 12 监测, https://www.fhnw.ch/de/die-fhnw/hochschulen/architektur-bau-geomatik/institute/ineb/ineb-forschung/nachhaltiges-bauen-und-betreiben/aescherstrasse.[31] E. D. Klenske, P. Hennig, B. Scholkopf, M. N. Zeilinger，保持信息价值的近似双重控制及其在建筑控制中的应用，载于：2016 年欧洲控制会议（ECC），2016，第 800–806 页。doi:10.1109/ECC.2016.7810387.[32] A. Ebadat, D. Varagnolo, G. Bottegal, B. Wahlberg, K. H. Johansson，面向应用的房间占用率估计算法输入设计，载于：2017 年 IEEE 第 56 届决策与控制年会（CDC），2017，第 3417–3424 页。doi:10.1109/CDC.2017.8264159.[33] B. A. Hernandez Vicente, P. A. Trodden，针对受约束线性系统、具有持续激励的稳定预测控制，Systems & Control Letters 126 (2019) 58–66. doi:10.1016/j.sysconle.2019.03.003.[34] X. Lu, M. Cannon, D. Koksal-Rivet，鲁棒自适应模型预测控制：性能与参数估计，International Journal of Robust and Nonlinear Control 31. doi:10.1002/rnc.5175.[35] E. ˇ Z´aˇcekov´a, M. Pˇcolka, M. ˇSebek，在可预测扰动存在时具有保证可辨识性的区域模型预测控制，Journal of the Franklin Institute 357 (2) (2020) 978–1001. doi:10.1016/j.jfranklin.2019.10.016.[36] S. Br¨uggemann, R. R. Bitmead，模型预测控制中的前瞻性持续激励，Automatica 136 (2022) 110033. doi:10.1016/j. automatica.2021.110033.[37] 双层密封中空玻璃窗板的加速老化与耐久性及其对建筑供热需求的影响，Energy and Build- ings 116 (2016) 395–402. doi:10.1016/j.enbuild.2016.01.015.[38] 建筑围护结构和机械部件退化对整栋建筑性能的影响：综述，Energy Procedia 132 (2017) 321–326. doi:10.1016/j.egypro.2017.09.739。