

引用格式: LIU Wenbo, SUN Di, CHEN Jiayu, et al. A Systematic Study on Measurement Condition Control and Modeling Methods for Glucose Concentration Prediction Using Near-Infrared Spectroscopy [J]. Acta Photonica Sinica, 2026, 55 (7): 0717001

刘文博,孙迪,陈嘉宇,等. 近红外光谱法糖浓度预测中测量条件控制与建模方法的系统研究[J]. 光子学报, 2026, 55(7): 0717001

近红外光谱法糖浓度预测中测量条件控制与建模方法的系统研究

刘文博¹, 孙迪², 陈嘉宇¹, 姚明飞², 刘雪颖¹, 刘瑾¹, 韩同帅¹, 徐可欣^{1,2}

(1 天津大学精密测试技术及仪器全国重点实验室, 天津大学, 天津, 300072)

(2 先阳科技(天津)有限公司, 天津, 300192)

摘 要:在基于近红外光谱法的糖浓度预测中,由于糖分子吸收与散射信号微弱,预测精度易受测量条件波动影响。本文结合糖水溶液模拟光谱体系与人体口服葡萄糖耐量试验,系统分析线性与非线性测量条件对光谱特征及模型性能的影响,并比较 PLSR 与 MLP 两种建模方法的抗干扰能力,提出“允许变化”与“不允许变化”的测量条件分类策略。仿真结果表明,温度与白噪声是“不允许变化”的,它们呈显著非线性特征,是影响测量精度的关键因素,需将它们分别控制在 $\leq 1^{\circ}\text{C}$ 和 $\leq 0.002\text{ a.u.}$ 范围内,预测误差方可稳定在约 0.2 mmol/L ;而入射角度、光源色温及光源漂移等因素影响近似线性,在较大变化范围内仍可保持较低误差,是“允许变化”的。MLP 在处理单一非线性温度干扰时略优于 PLSR,但在温度波动较大时,两者误差均明显增大。人体口服葡萄糖耐量试验进一步验证了温控条件下光谱与血糖相关性显著增强,留一受试者法交叉验证预测误差为 1.33 mmol/L ,优于未温控组(1.63 mmol/L)。可见,合理的温度控制可有效提升预测精度。本研究为近红外光谱法无创血糖测量中的测量条件控制提供了参考。

关键词:近红外光谱法;无创血糖测量;多变量分析方法;口服糖耐量实验;测量条件

中图分类号:O433.4

文献标识码:A

doi:10.3788/gzxb20265507.0717001

0 引言

近红外光谱法是快速检测物质成分浓度的常用技术,在人体血糖浓度(Blood Glucose Concentration, BGC)测量领域具有广泛的应用潜力^[1-2]。然而,由于糖分子引起的吸收与散射信号较为微弱,为实现 BGC 的有效分辨与准确测量,光谱分析方法需具备较高的灵敏度和抗干扰能力^[3]。在实际采集的光谱数据中,成分信息往往与多种测量条件的干扰共存。当干扰程度较大时,可能淹没或掩盖弱小的糖信号,显著降低测量的准确性。当前近红外光谱分析主要依赖基于化学计量学的多变量分析方法,同时机器学习和人工智能技术也在不断应用于提升光谱预测性能^[4-6]。然而,任何分析方法在抑制测量条件干扰方面均存在能力上限。

在血糖检测领域,测量条件涵盖外部和内部两个方面。外部测量条件包括测量环境(温度、湿度)、测量压力、人机接口状态等。在人体活体测量场景中,通过合理设计,可一定程度上稳定这些条件。例如,在无创血糖检测中,Uwadaira Y^[7]等人通过在探头后侧集成加热装置以稳定皮肤温度;Tronstad C^[8]等人利用绑

基金项目:国家自然科学基金(81971657)

第一作者:刘文博(1999—),男,博士研究生,主要研究方向为无创血糖测量。Email: wenbo_liu@tju.edu.cn

通讯作者:刘雪颖(2001—),女,硕士研究生,主要研究方向为近红外光谱检测。Email: liuxveying@tju.edu.cn;刘瑾(1979—),女,副教授,博士,主要研究方向为光谱检测与光谱分析。Email: liu_jin@tju.edu.cn;韩同帅(1991—),男,助理研究员,博士,主要研究方向为无创血糖测量。Email: hts2014@tju.edu.cn

收稿日期:2026-03-07; **录用日期:**2026-04-19

<http://www.photon.ac.cn>

带固定探头并监测皮肤表面温度; Aloraynan A^[9]等人使用流动氮气控制皮肤微环境湿度; Lundsgaard-Nielsen S M^[10]等人采用氟化镁玻璃作为界面材料以减少界面拉曼信号; Feng R^[11]等人开发的多传感融合血糖仪能监测温湿度对光谱的影响; Caduff A^[12]等人则在探头接口加入玻璃涂层以固定寄生电容,减少状态变化带来的干扰。除了外部条件,皮肤内部还存在不可忽视的生理变化,本文统称为内部测量条件,包括血流灌注、皮肤层厚度、含水量、温度场分布等^[13-16]。这些因素会改变光在皮肤中的光程,进而影响测量信号。此外,内部变化还与外部条件相互耦合,例如:皮肤血流与界面温度密切相关;皮肤层厚度与含水量受接触压力影响显著;皮肤与传感器接触状态影响含水量与温度场;汗液分布则与环境湿度、传感器温度均相关。

上述方法在一定程度上提升了测量条件的稳定性,但从实际应用的便利性出发,仍需进一步实现轻量化改进。本课题组围绕温度控制、传感器重量、测量界面材料及透气性等方面开展了系统研究,研制出一款轻量化可穿戴传感器,其重量仅为6 g^[17]。该传感器集成加热贴并设计有通气孔,在改善皮肤温度控制的同时,有效缓解了出汗问题;同时,其对皮肤施加的压力较小,尽可能减少对皮肤自身状态的影响。在皮肤血流控制方面,课题组的策略是将皮肤加热至接近人体血液温度(约36℃),并利用1050 nm波长的光信号对血流状态进行实时监测,当信号趋于稳定后再开展测量,实验结果表明该方法能够取得较好的效果^[18]。

除上述测量条件需尽可能保持稳定以外,测量过程中的姿势(如手臂弯曲状态)亦需保持一致,以维持传感器与皮肤界面的长期稳定性。要在柔软可变形的活体皮肤上满足这些条件也是具挑战性的。本课题组 Han T^[17]等人研究了双十字的激光定位策略,用于复现统一的手臂姿势,从而保证每次测量时皮肤形态的一致性。在这些测量条件的严格控制条件下,采用该传感器直接获得了与血糖变化呈现高度相关的单波长漫反射光信号。尽管单次口服葡萄糖耐量试验(Oral Glucose Tolerance Test, OGTT)效果良好,但由于测量时人体不能走动,因此难以满足长时间的日常生活中的测量需求。当受测者发生走动、姿势变化等活动时,上述测量条件不可避免地会产生波动。因此,是对测量条件进行严格复现,还是借助多变量分析或机器学习方法对其影响进行校正,成为需要权衡的问题。课题组曾尝试仅依赖算法对长时间、多次OGTT结果进行校正,但效果并不理想。而实施必要的测量条件控制或许仍是难以回避且亟待解决的关键问题。

本文尝试对多种测量条件的变化进行分析,试图找到“允许变化”的测量条件和“不允许变化”的测量条件,以降低测量条件控制的难度。本文分类的依据是:“允许变化”的测量条件是可以借助算法进行修正的,而那些“不允许变化”的测量条件则是难以通过算法有效补偿的关键因素^[19],必需加以控制。在实际应用中,排除那些“不允许变化”的测量条件才能确保测量的可靠性,排除伪相关现象。

本文首先在模拟计算的糖水溶液体系中开展实验,设计了五类外部测量条件变化因素和三类成分变化因素,系统评估了不同条件下光谱分析方法的抗干扰能力。建模方面选取了典型的线性方法和非线性方法作为代表,分别为偏最小二乘回归(Partial Least Squares Regression, PLSR)和多层感知器(Multilayer Perceptron, MLP),用于分析多变量建模方法在不同测量条件范围内的预测性能差异。在此基础上,进一步开展了37例的人体口服葡萄糖耐量试验,其中18例采用局部测量温度控制,另外19例未进行局部温度控制。同时分别构建线性与非线性预测模型,以比较温度控制条件对糖浓度预测精度的影响。本文仅针对人体实验中的温度测量条件进行了研究,尚未涉及其他测量条件的系统分析。然而,本文所提出的研究思路可为后续相关工作的开展提供有益的参考。

1 方法

1.1 朗伯-比尔定律

糖水溶液的光谱测量可采用朗伯-比尔定律进行描述^[20-21],吸光度的表达式如公式(1)。

$$A = \lg\left(\frac{I_0}{I}\right) = \sum_{i=1}^n \frac{\epsilon_i \cdot c_i \cdot L}{\ln(10)} = \sum_{i=1}^n \frac{\mu_{a,i} \cdot L}{\ln(10)} \quad (1)$$

其中, I_0 为入射光强, I 为出射光强;假设溶液体系中有 n 中成分, ϵ_i 为第 i 种成分的摩尔消光系数(单位:1/mg/dL/mm), c_i 为第 i 种成分的浓度(单位:mg/dL), $\mu_{a,i}$ 为第 i 种成分的吸收系数(单位:1/mm); L 为实际行进的光程(单位:mm)。出射光强可以表达为式(2):

$$I = I_0 \cdot 10^{-\sum_{i=1}^n \frac{\epsilon_i \cdot c_i \cdot L}{\ln(10)}} \quad (2)$$

因此,出射光强度受入射光强度、摩尔消光系数、被测物浓度、实际行进的光程四种因素共同决定。糖水溶液体系模型如图1。当考虑到光纤相对于比色皿存在一定的角度偏差 α_0 ,通过几何关系可以得出在糖水溶液中实际行进的光程表达式为式(3)。

$$L = L_0 / \cos(\alpha_2) = L_0 / \cos \{ \arcsin [\sin(\alpha_0) / n_0] \} \quad (3)$$

在糖水溶液的光谱测量中,光在溶液中的实际光程受入射角与溶液折射率两项参数共同影响。为了考虑参考测量条件,可将同一样品分别置于厚度分别为 L_1 与 L_2 的两只样品池中,可分别获得出射光强 I_1 与 I_2 ,当光源存在随时间变化的漂移 $f(t)$ 时,两者之间的差分吸光度可表示为:

$$A_D = A_2 - A_1 = \lg \left(\frac{I_0 \cdot f(t)}{I_2} \right) - \lg \left(\frac{I_0 \cdot f(t)}{I_1} \right) = \lg \left(\frac{I_1}{I_2} \right) \quad (4)$$

经过差分计算后,可以有效地消除光源漂移对出射光强的影响。同时,在漫反射测量中差分吸光度表示为两光源-探测器距离(Source-Detector Separation, SDS)下吸光度之差^[17]。

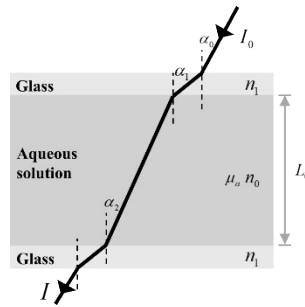


图1 糖水溶液体系模型

Fig.1 Model of aqueous glucose solution system

另外,本文引入“单位差分吸光度”的概念。其定义为在某一测量条件扰动量 Δp 下,差分吸光度变化量与扰动幅度之比,即:

$$S_p(\lambda) = \frac{\Delta A_D(\lambda)}{\Delta p} \quad (5)$$

其中扰动量 Δp 表示对应测量条件(如温度、入射角等)的变化量。

1.2 偏最小二乘回归(PLSR)

PLSR方法^[22]首先对浓度矩阵 Y 及样品光谱矩阵 X 分解,效果为:

$$Y = UQ^T + E_Y \quad (6)$$

$$X = TP^T + E_X \quad (7)$$

其中 U 与 T 分别为 Y 与 X 的得分矩阵; Q 与 P 分别为 Y 与 X 的载荷矩阵; E_Y 与 E_X 为 Y 与 X 的残差矩阵。进一步对得分矩阵作线性回归,即:

$$U = TB \quad (8)$$

$$B = (T^T T)^{-1} T^T Y \quad (9)$$

根据(6)-(9)式,先求解出未知样品的得分矩阵,再推倒出未知样品浓度的预测值,即: $Y_{\text{未知}} = T_{\text{未知}} BQ$ 。为了评价模型的预测能力,并且观察预测数据的异常程度,提供两种评价参数,即:校正集均方根误差(Root Mean Square Error of Calibration, RMSEC)、预测集均方根误差(Root Mean Square Error of Prediction, RMSEP)。一般地,均方根误差(Root Mean Square Error, RMSE)的表达式如下:

$$RMSE = \sqrt{\frac{\sum_{i=1}^m (y_{i, \text{actual}} - y_{i, \text{predicted}})^2}{m - 1}} \quad (10)$$

其中 $y_{i, \text{actual}}$ 为第 i 个样品的参考值, $y_{i, \text{predicted}}$ 为模型对第 i 个样品的预测值, m 为校正集或验证集的样品数^[23]。并且使用相关系数 R 这一参量来评价预测值与测定值的相关程度。相关系数的表达式如下:

$$R = \frac{Cov(y_{i, \text{actual}}, y_{i, \text{predicted}})}{\sqrt{Var[y_{i, \text{actual}}] \cdot Var[y_{i, \text{predicted}}]}} \quad (11)$$

1.3 多层感知器 (MLP)

多层感知器 (MLP) 是一种由多层节点构成的非线性映射模型, 通常包括输入层、隐含层和输出层^[24]。各节点之间通过权重相连, 当输入信号经加权求和后达到传递函数的激活阈值, 信息才会被传递至下一层。通过不断调整节点间的权重参数, 网络能够迭代逼近真实的输入-输出关系。本文采用的 MLP 模型架构及参数空间如图 2 及表 1 所示。输入层使用 1000–1700 nm 范围内的光强值或差分吸光度数值, 隐含层设置为 N 层, 每层包含 M 个神经元; 输出层为目标糖浓度值。激活函数使用 leakyRelu 函数。

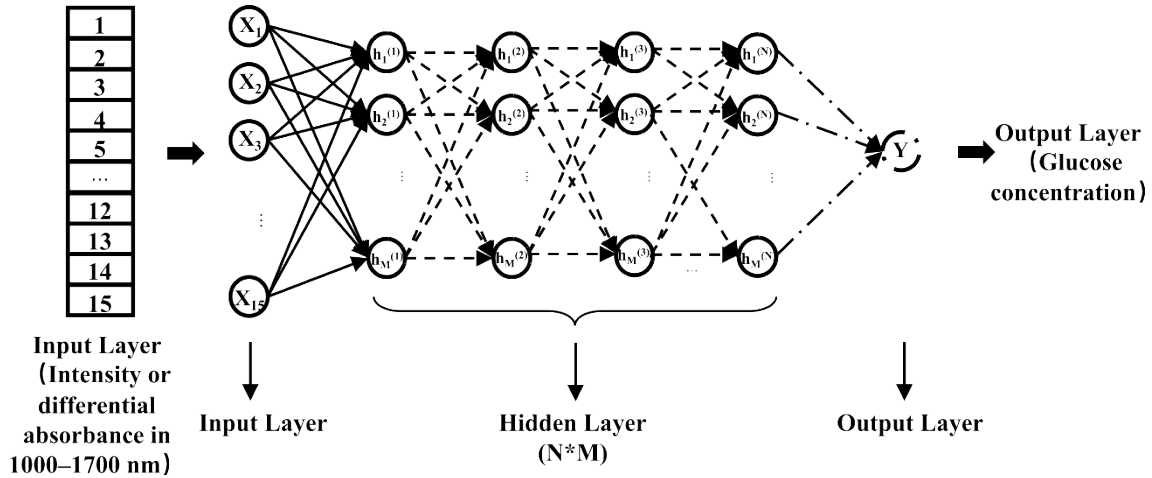


图 2 MLP 模型架构

Fig.2 MLP model architecture

表 1 MLP 模型参数搜索空间

Table 1 MLP model parameter search space

Parameter	Search Range	Sampling Method
Number of layers (N)	[2, 3]	Uniform
Number of neurons per layer (M)	[8, 256]	Log-uniform
Learning rate	[0.0001, 0.002]	Log-uniform
L2 regularization	[0.00001, 0.01]	Log-uniform

1.4 光谱仿真的参数设置

模拟卤素灯光源经准直后入射样品, 使用的光源范围为 1000–1700 nm, 波长间隔为 50 nm, 共选取 15 个波长。其中, 短波段 (1000–1300 nm) 主要反映血红蛋白的吸收特征, 而 1400–1600 nm 波段则对应水、葡萄糖的吸收特征, 其中 1408 nm 和 1536 nm 为葡萄糖的特征吸收峰^[25]。因此, 所选波段能够同时兼顾葡萄糖、水及其他组织成分的吸收特性。为贴近人体实验受试者 (健康人群) BGC 的生理变化范围, 水溶液中葡萄糖浓度设定为 4–12 mmol/L。该范围覆盖了正常空腹 BGC 水平 (约 4.1–5.9 mmol/L) 以及 OGTT 过程中 BGC 升高的典型区间 (可达 7.8–11.1 mmol/L 以上)。体系包含四种成分: 水 (Comp. I)、白蛋白 (Comp. II)、血红蛋白 (Comp. III) 和葡萄糖 (Comp. IV), 其光学参数如图 3 所示^[26–28]。其中白蛋白与血红蛋白的变化仅考虑其吸收系数的影响。

在测量体系中加入五类可能的外部测量条件干扰: 溶液温度 (Cond. I)、光纤相对于比色皿的角度 (Cond. II)、光源色温 (Cond. III)、光源漂移 (Cond. IV) 以及光电检测器白噪声 (Cond. V)。其中, 溶液温度变化 (Cond. I) 主要影响吸收系数与折射率^[29]; 光纤入射角变化 (Cond. II) 主要影响光在溶液中的实际传播光程; 光源色温 (Cond. III) 与光源漂移 (Cond. IV) 主要改变入射光强; 光电检测器白噪声 (Cond. V) 则直接影响出射光强。

本文设置了两种测量模型: 严格控制测量条件的模型 A, 以及未控制测量条件的模型 B。五类外部测

量条件与三类内部成分变化的具体范围见表 2。两种模型分别对应严格控制测量环境与日常非严格控制测量环境两类实际场景。样本量设定为 1000 个,标准样品池厚度为 1 mm,参考样品池厚度为 2 mm。建模方式包括单光程建模与参考测量建模,即基于单样品池(1 mm 光程)的出射光强建模,及基于双样品池(1 mm 与 2 mm 光程)的差分吸光度建模。校正集与验证集中各成分及干扰参数均在表 2 所给范围内随机生成,并保证变化分布一致。校正集中考虑葡萄糖的测量误差为 $\pm 10\%$ 。RMSEP 的计算取五次验证集预测结果的平均值作为最终输出。

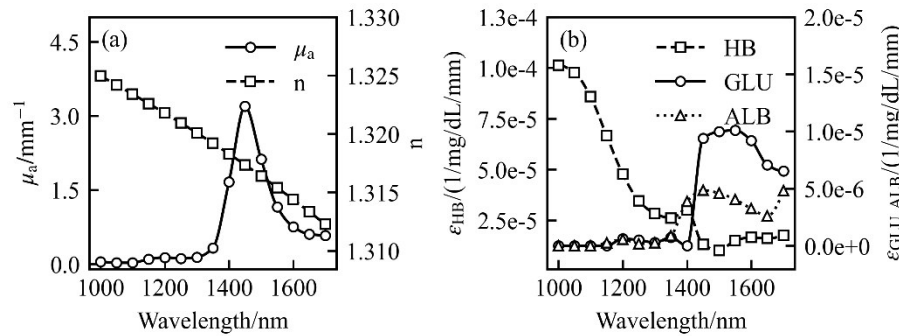


图3 四成分光学参数。(a) 纯水的吸收系数及折射率;(b) 1 mm 光程 1 mg/dL 浓度的白蛋白(Albumin, ALB)、血红蛋白(Hemoglobin, HB)及葡萄糖(Glucose, GLU)引起的摩尔消光系数变化

Fig. 3 Optical parameters of the four-component system. (a) Absorption coefficient and refractive index of pure water; (b) Variations in molar extinction coefficients induced by albumin (ALB), hemoglobin (HB), and glucose (GLU) at a concentration of 1 mg/dL over a 1 mm optical path.

表 2 外部测量条件与内部成分变量的设定范围
Table 2 Defined ranges of external measurement conditions and internal component variables

Category		Baseline Value	Variation under Strictly Controlled Conditions (Model A)	Variation under Uncontrolled Conditions (Model B)
External Measurement Conditions	Liquid temperature (Cond. I)	30℃	0~0.1℃	0~5℃
	Incident angle (Cond. II)	0°	0~0.5°	0~5°
	Light source color temperature (Cond. III)	3000 K	0~10 K	0~100 K
	Light source drift (Cond. IV)	0	0~1%	0~3%
	Detector white noise (Cond. V)	0	± 0.001	± 0.001
Internal Components	Water (Comp. I)	55.6 mol/L	0	0
	Albumin (Comp. II)	35 g/L	0~3.5 g/L	0~3.5 g/L
	Hemoglobin (Comp. III)	120 g/L	0~12 g/L	0~12 g/L
	Glucose (Comp. IV)	4 mmol/L	0~8 mmol/L	0~8 mmol/L

1.5 人体实验设置

本研究共纳入 37 例受试者(健康,无糖尿病病史)开展 OGTT 实验。受试者基本特征如表 3 所示。所有受试者均签署了知情同意书,研究方案已通过天津大学伦理委员会审批。受试者在每次测试前均保持空腹状态,实验环境温度控制在 25~26℃。其中 18 例在温控条件下进行测试,另 19 例在非温控条件下进行测试。试验过程中同步采集受试者前臂皮肤光谱数据。此外,在每名受试者 OGTT 过程中采集 3~4 次指尖血样作为参考 BGC,用于后续对比分析与模型验证。实验中,温控组采用 PID 控制模式,将局部温度稳定在 36℃。该温度接近人体血液温度,经一段时间加热后,可使皮肤血流达到相对稳定状态(通过 1050 nm 波段光信号进行观察),随后开展 OGTT 实验。该策略有助于降低血流生理波动对测量结果的干扰^[18]。

表3 受试者的人口统计学和基本特征
Table 3 Demographic and baseline characteristics of the subjects

Group	Sample Size	Age	Sex (Male/Female)	BMI (kg/m ²)	Fasting Blood Glucose (mmol/L)	Fitzpatrick Skin Type
Temperature—Controlled Group	18	32.6±7.4	10/8	23.4±2.6	5.1±0.4	III
Non—Temperature—Controlled Group	19	31.8±6.9	11/8	22.9±2.8	5.0±0.5	III

实验系统如图4所示。光谱测量系统主要包括:多环铟镓砷检测器,6个超发光二极管SLD1-SLD6(中心波长分别为1050 nm、1219 nm、1314 nm、1409 nm、1550 nm及1609 nm,含3dB带宽),光开关、数据采集处理单元。其中多环检测器内部包含5个分立的环形检测器(环宽为0.2 mm),与入射光纤同心排布,SDS分别为:1.7、2、2.3、2.6、2.9 mm。本文使用2.6 mm和2.9 mm的差分吸光度的数值作为光谱数据。实验中,传感器通过亲肤型无纺布双面胶带与皮肤接触,并采用双十字激光定位,以保证长时间测量的再现性。在人体皮肤测量条件下,采用1050 nm和1550 nm波长进行1 h连续评估,结果表明仪器噪声水平约0.0005–0.0007 a.u.的差分吸光度变化。Han T等人采用该传感器进行的大量人体实验表明,1 mmol/L的血糖变化可引起约0.001 a.u.的差分吸光度变化^[17],因此,该测量系统至少已具备分辨1 mmol/L血糖变化所对应光谱信号的能力。

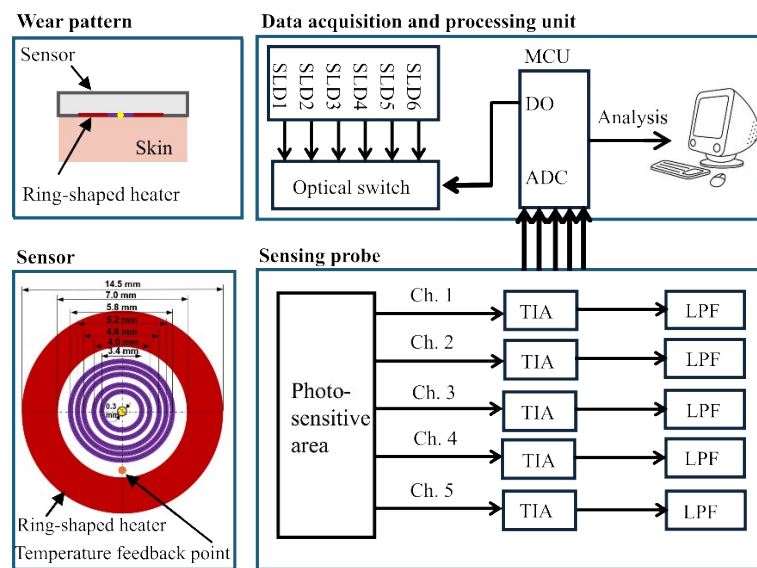


图4 实验系统设置
Fig. 4 Experimental system setup

2 结果

2.1 模拟样本的原始光谱与干扰光谱

模拟样本构造的原始光谱以及不同测量条件变化所引起的单位干扰光谱如图5所示。不同色温变化、光源漂移程度、ALB水平及HB水平所对应的干扰光谱整体表现出较好的一致性。对1350 nm单波长处的单位差分吸光度变化量进行分析(图6)可见:随着液体温度升高及白噪声水平增加,单位差分吸光度呈现出明显的非线性分布特征;而其余影响因素则主要表现为线性分布特征。

2.2 模拟样本的预测结果

表4展示了PLSR与MLP的建模及预测结果。可以看到,在PLSR模型中,单光程测量方式的预测性能较差:预测相关系数绝对值均低于0.2, RMSEP约为糖浓度变化范围的0.21倍。相比之下,参考测量方法显著改善了模型表现,其预测相关系数均超过0.85, RMSEP约为糖浓度变化范围的0.02–0.09倍,说明

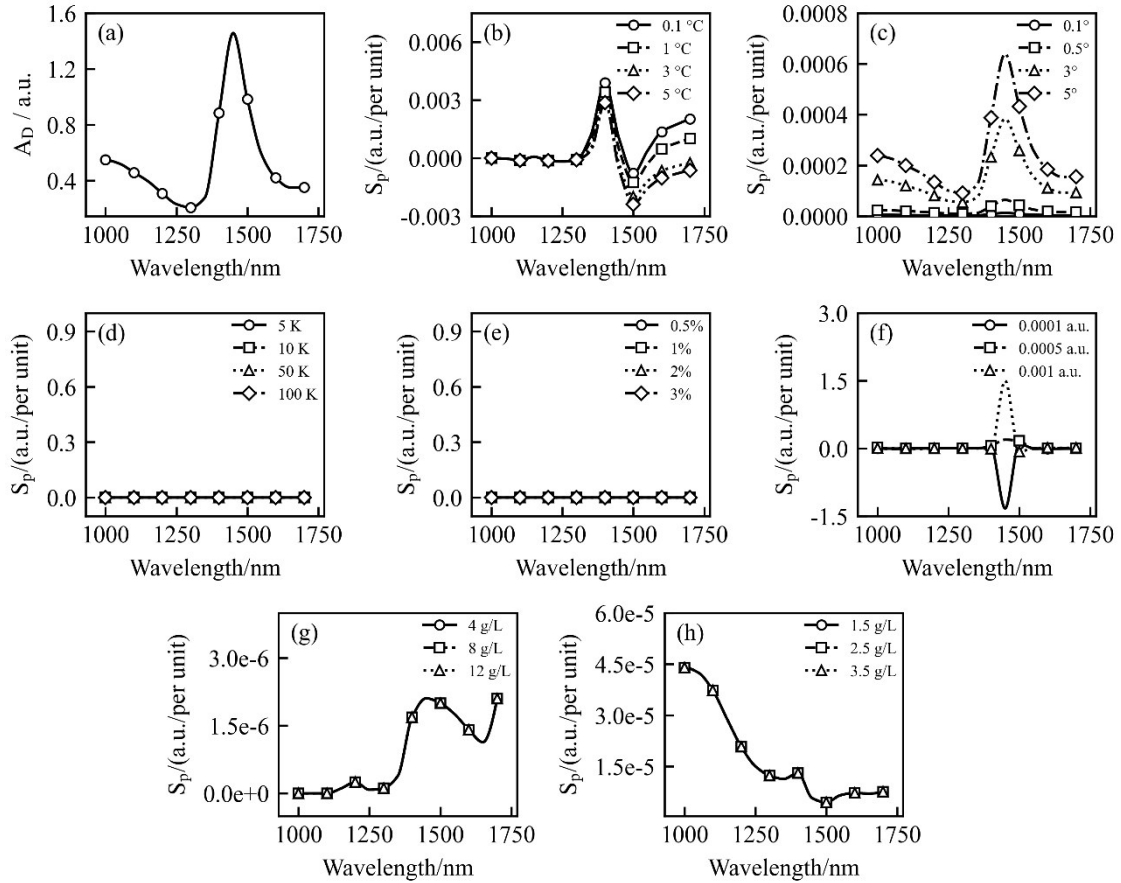


图5 原始光谱与不同测量条件变化时引起的单位干扰光谱。(a) 原始无干扰时的光谱;(b) 不同液体温度变化范围;(c) 不同入射角度变化范围;(d) 不同色温变化范围;(e) 不同光源漂移变化范围;(f) 不同白噪声变化范围;(g) 不同ALB浓度变化范围;(h) 不同HB浓度变化范围

Fig. 5 Original spectra and unit interference spectra induced by variations under different measurement conditions. (a) Spectra without interference (original spectra); (b) Different ranges of liquid temperature change; (c) Different ranges of incident angle change; (d) Different ranges of color temperature change; (e) Different ranges of light source drift change; (f) Different ranges of white noise levels; (g) Different ranges of ALB concentration change; (h) Different ranges of HB concentration change.

参考测量方式明显优于单光程测量方式。进一步比较参考测量下的模型 A 与模型 B 可以发现,模型 A(严格控制测量条件)在 R 和 RMSEP 两项指标上均优于模型 B(未控制测量条件)。其中,模型 A 的 RMSEP 约为糖浓度变化范围的 0.03 倍,其预测精度约为模型 B 的 3.4 倍,表明严格控制测量条件能够显著提升预测相关性和预测精度。

在 MLP 模型中虽然也呈现出与 PLSR 类似的趋势,但其预测误差整体偏高于 PLSR。但在实际测量中,对所有测量条件逐一严格控制往往意味着较高的实施成本。因此,本文进一步系统分析了多项测量条件在不同变化范围下对糖浓度预测精度的影响,以明确哪些条件必须严格控制,哪些条件的变化在可接受范围内不会显著影响模型性能,从而为实际应用提供更具可行性的测量策略。

不同测量条件变化范围对应的 PLSR 预测效果如图 7 所示。在改变任一测量条件的情况下,假设检测器白噪声为固有的 0.001 a.u. 水平,其余测量条件的变化量均设为 0。从结果可以看出,液体温度变化范围与检测器白噪声变化范围越大,模型预测性能越差。当水溶液体系中存在 10 °C 的温度变化或 0.01 a.u. 的白噪声干扰时,糖浓度的预测均方根误差分别达到 0.72 mmol/L 与 0.74 mmol/L,约占糖浓度变化范围的 0.09 倍。本文将此类必须严格限制的条件定义为“不允许变化的测量条件”。相比之下,其他测量条件即使变化较大,其糖浓度 RMSEP 均未超过 0.2 mmol/L(约占变化范围的 0.025 倍)。液体温度及白噪声水平变化时,单位差分吸光度呈现明显的非线性分布特征,削弱 PLSR 的建模能力。而入射角度、色温、光源漂移等“允许变化的测量条件”,呈现更清晰的线性分布,PLSR 能较好地在该类干扰下解析糖信号。

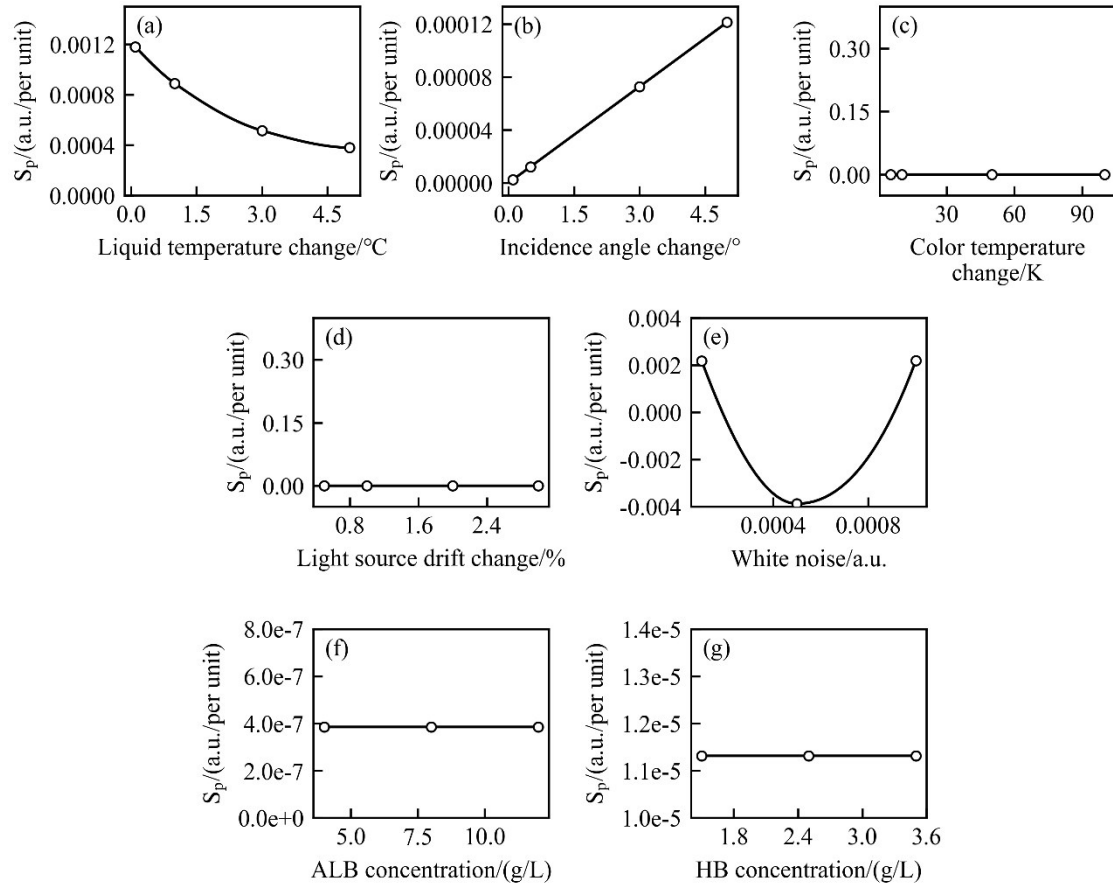


图6 不同测量条件变化范围时1350 nm波长下引起的单位差分吸光度。(a)不同液体温度变化范围;(b)不同入射角度变化范围;(c)不同色温变化范围;(d)不同光源漂移变化范围;(e)不同白噪声变化范围;(f)不同ALB浓度变化范围;(g)不同HB浓度变化范围

Fig. 6 Unit differential absorbance at 1350 nm under different ranges of measurement conditions. (a) Different ranges of liquid temperature change; (b) Different ranges of incident angle change; (c) Different ranges of color temperature change; (d) Different ranges of light source drift change; (e) Different ranges of white noise levels; (f) Different ranges of ALB concentration change; (g) Different ranges of HB concentration change.

表4 PLSR和MLP的建模及预测效果

Table 4 Modeling and prediction performance of PLSR and MLP

Model	Metric	Model A — Single-Path Measurement	Model A — Reference Measurement	Model B — Single-Path Measurement	Model B — Reference Measurement
PLSR	RMSEC (mmol/L)	1.67	0.21	1.69	0.77
	R (Calibration Set)	0.18	0.99	0.07	0.89
	RMSEP (mmol/L)	1.67	0.22	1.73	0.75
	R (Validation Set)	0.16	0.99	-0.03	0.90
MLP	RMSEC (mmol/L)	1.65	0.26	1.71	0.9
	R (Calibration Set)	0.24	0.99	0.05	0.85
	RMSEP (mmol/L)	1.67	0.27	1.75	1.22
	R (Validation Set)	0.17	0.99	-0.05	0.70

尽管线性干扰一般影响较小,但在不同组分的变化范围内仍可产生预测差异。例如,对于白蛋白浓度与血红蛋白浓度均在3 g/L的变化范围内,前者导致的预测误差约为0.20 mmol/L,后者约为0.09 mmol/L,差异约为2倍。这是由于二者与葡萄糖的差分吸光度光谱的向量夹角不同:白蛋白与葡萄糖的夹角为 0.43° ,血红蛋白为 1.37° 。白蛋白光谱形状更接近葡萄糖,因此更容易造成模型误判。但总体来看,预测性

能受限的主要原因仍是温度与白噪声带来的强非线性干扰。

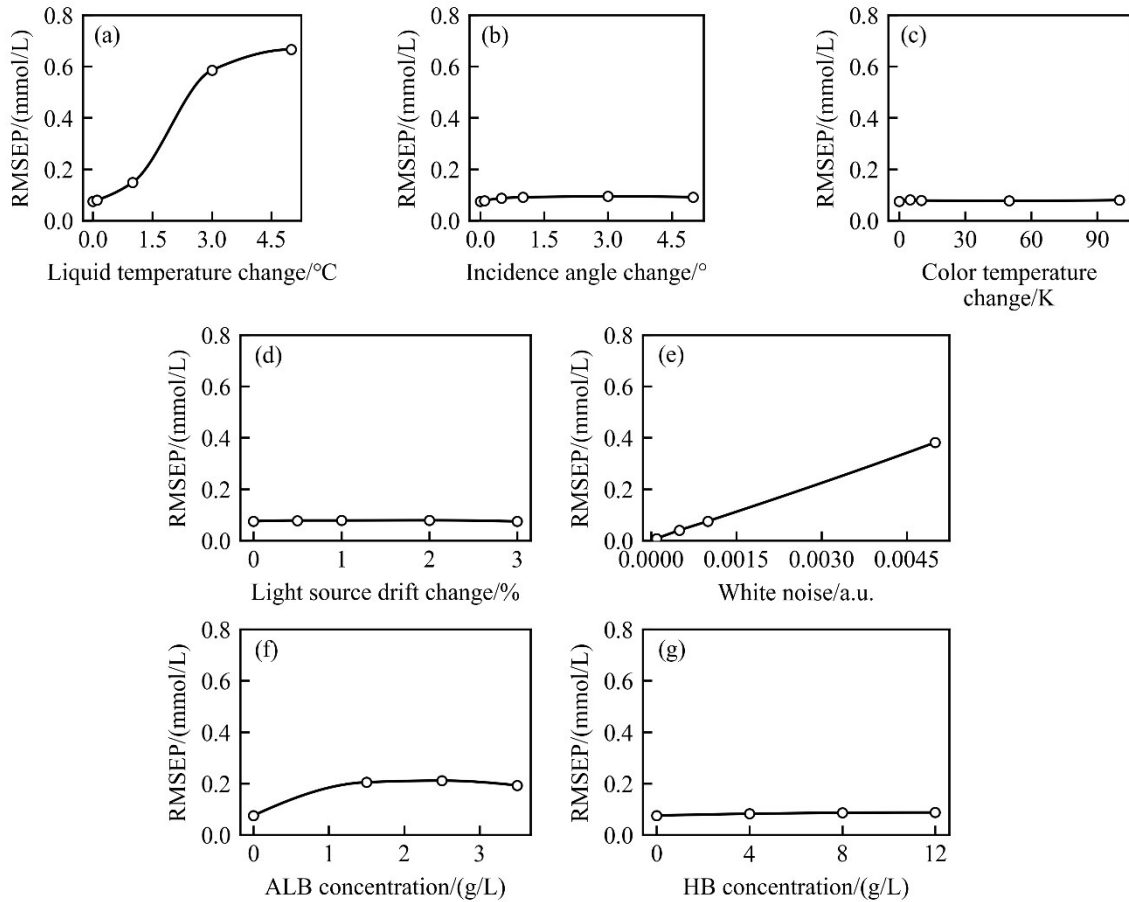


图7 不同测量条件变化范围下的PLSR预测效果。(a)不同液体温度变化范围;(b)不同入射角度变化范围;(c)不同色温变化范围;(d)不同光源漂移变化范围;(e)不同白噪声变化范围;(f)不同ALB浓度变化范围;(g)不同HB浓度变化范围
Fig. 7 PLSR prediction performance under different ranges of measurement conditions change. (a) Different ranges of liquid temperature change; (b) Different ranges of incident angle change; (c) Different ranges of color temperature change; (d) Different ranges of light source drift change; (e) Different ranges of white noise levels; (f) Different ranges of ALB concentration change; (g) Different ranges of HB concentration change.

针对温度变化与白噪声这两类非线性干扰,本文进一步采用MLP模型评估其对预测精度的改善效果,结果如图8所示。结果表明,在不同白噪声变化范围内,两种模型的预测性能基本相当,说明白噪声需优先加以抑制。相比之下,在温度干扰条件下,PLSR和MLP模型的预测误差均显著增大,且随温度变化幅度的增加而逐渐上升。其中,MLP模型的预测误差略低于PLSR,这与其在处理非线性关系方面具有一定优势有关,但仍无法完全消除温度干扰的影响。具体而言,当温度变化范围达到5℃时,MLP模型的预测误差增至0.54 mmol/L,而PLSR模型增至0.67 mmol/L。由此可见,相较于其他近似线性变化的因素,温度是影响测量精度的关键因素之一。

将各类线性与非线性干扰因素同时纳入模型进行比较,结果如表5所示。其中,温度变化范围分别设定为较小范围(0-0.1℃)和较大范围(0-5℃)。结果表明,当温度控制在较小范围内时,模型预测精度显著提升;而当温度在较大范围内波动时,两种模型的预测精度均明显下降。在该条件下,MLP模型的预测性能略低于PLSR,这可能与干扰以线性因素为主有关。

2.3 人体实验的测量结果

图9展示了温控组与未温控组在人体OGTT实验中的典型光谱测量结果。以OGTT开始时刻的光谱作为基准光谱进行分析。结果表明,温控组单波长信号与血糖变化之间的相关性及其灵敏度均显著高于未温控组。在温控条件下,可在一定程度上维持皮肤一定深度范围内温度及生理状态的相对稳定,从而减少外

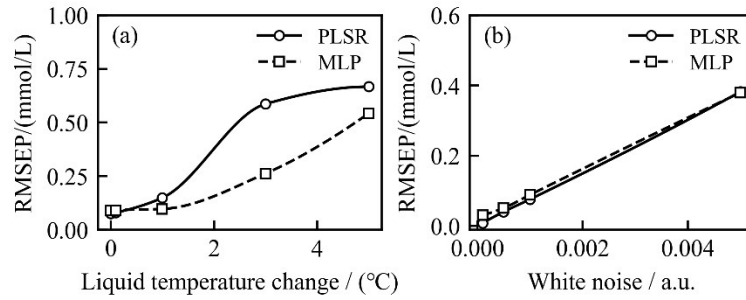


图8 两种建模方法(PLSR及MLP)的预测效果。(a) 不同液体温度变化范围;(b) 不同白噪声变化范围
Fig. 8 Prediction performance of the two modeling methods (PLSR and MLP). (a) Different ranges of liquid temperature change; (b) Different ranges of white noise levels

表5 未严格控制测量条件(模型B)改变温度控制范围的预测结果

Model	Metric	Model B — Reference Measurement (Temperature Variation: 0 – 5 °C)	Model B — Reference Measurement (Temperature Variation: 0 – 0.1 °C)
PLSR	RMSEP (mmol/L)	0.75	0.12
	R	0.90	0.99
MLP	RMSEP (mmol/L)	1.22	0.28
	R	0.70	0.99

界及生理波动对光谱信号的干扰。而在未温控条件下,受试者摄入葡萄糖后,局部温度变化可能引起血管反应、血流灌注和汗腺活动,皮肤生理状态会随血糖升高产生一系列复杂的生理响应。对于光谱信号而言,这些响应通常表现为非线性特征,增加了信号解析的难度^[30-31]。其中最为显著的是代谢水平增强所带来的代谢热增加,这不仅会打破局部组织的热平衡状态,还可能改变组织的光学特性,进而导致光谱信号产生持续性漂移,削弱其与血糖变化之间的对应关系。此外,温控组在 1550 nm 波长附近表现出对血糖变化较高的灵敏度,该结果与 Han T^[17]等人报道的血糖灵敏度水平相当。

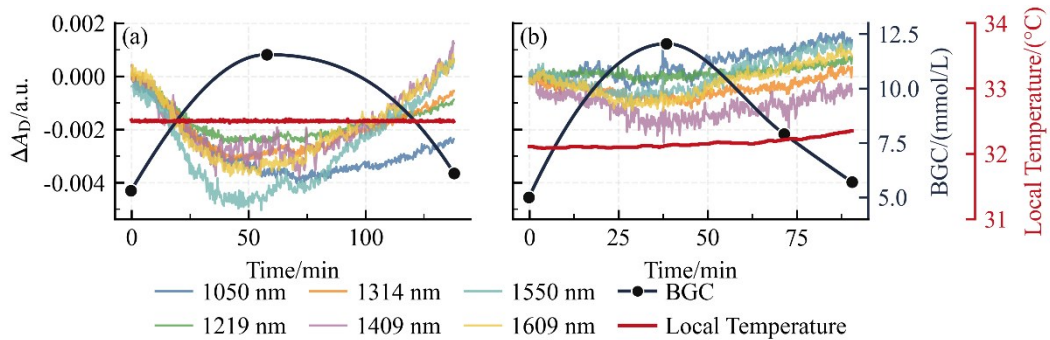


图9 人体 OGTT 实验中的典型光谱测量。(a) 温控组;(b) 未温控组
Fig. 9 Typical spectral measurements in the human OGTT. (a) Temperature-controlled; (b) Non-temperature-controlled.

2.4 人体实验的预测结果

在实验光谱的血糖预测分析中,分别统计了组内建模预测结果(图10和图11)、数据集混合建模预测结果(图12)以及留一受试者法交叉验证(Leave-One-Subject-Out Cross-Validation, LOSOCV)预测结果(图13)。组内建模采用单次 OGTT 实验中获取的光谱数据进行模型训练与预测;数据集混合建模则将多次 OGTT 实验的数据合并后进行统一建模,以评估模型的跨实验泛化能力。建模过程中统一采用 90% 的光谱数据进行模型训练,剩余 10% 的数据用于预测与验证,以保证结果的可靠性与可比性。在组内建模预测结果中,温控组整体表现优于未温控组,且 MLP 模型的预测精度整体高于 PLSR 模型。具体表现为:在相关

系数R分布区间0.95–1.0内,PLSR模型在温控组中有13/18达到该区间,而未温控组仅为5/19;MLP模型在温控组中达到17/18,未温控组为11/19。从整体预测精度来看,在温控条件下,MLP模型的RMSEP(0.20 mmol/L)明显低于PLSR模型(0.37 mmol/L),其预测精度约为后者的1.9倍;在未温控条件下,MLP模型(0.28 mmol/L)同样优于PLSR模型(0.44 mmol/L),预测精度提升至约1.6倍。这主要是因为人体生理因素引起的光谱响应通常具有非线性特征,而MLP在处理非线性方面具有一定优势,因此能够更有效地刻画光谱信号与血糖变化之间的复杂映射关系。

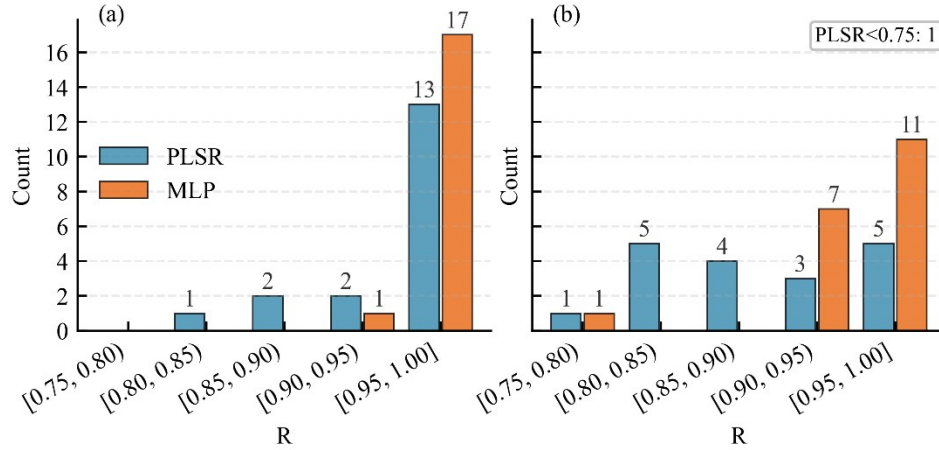


图10 组内建模的不同温控条件下PLSR与MLP模型R值分布。(a) 温控组;(b) 未温控组

Fig.10 Distribution of R values for PLSR and MLP models under different temperature conditions based on within-group modeling. (a) Temperature-controlled; (b) Non-temperature-controlled.

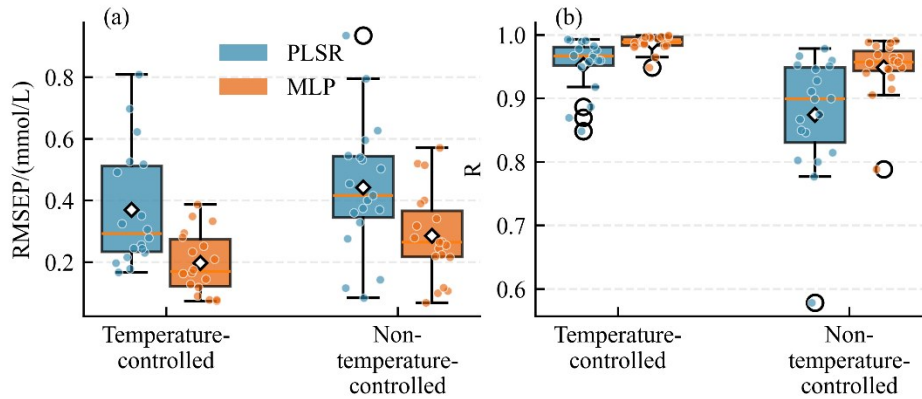


图11 组内建模的不同温控条件下PLSR与MLP模型预测性能比较。(a)为RMSEP的分布结果;(b)为R的分布结果

Fig. 11 Comparison of prediction performance between PLSR and MLP models under different temperature conditions based on within-group modeling. (a) Distribution of RMSEP; (b) Distribution of R.

此外,为评估模型的泛化能力,进一步开展了数据集混合建模预测分析与LOSOCV(图12、图13)。上述两种建模方式均基于参数搜索得到的最优MLP结构与超参数(见表6)。结果表明,在混合建模条件下,PLSR模型在温控组和未温控组中的相关性均较低,R值均低于0.65,整体预测性能明显下降;而MLP模型仍表现出较好的稳定性和较强的泛化能力。具体而言,在MLP模型中,温控组的RMSEP为0.21 mmol/L,R为0.99;未温控组的RMSEP为0.32 mmol/L,R为0.98。两组之间RMSEP相差约1.5倍,表明温度控制对预测误差具有显著影响。在LOSOCV条件下,模型整体预测性能有所下降,但MLP模型仍表现出较好的稳定性。具体而言,在MLP模型中,温控组的RMSEP为1.33 mmol/L,R为0.66;未温控组的RMSEP为1.63 mmol/L,R为0.31,两组之间RMSEP相差约1.2倍。因此,相较于线性模型PLSR,MLP模型在跨数据集混合建模条件下具有更强的鲁棒性和泛化能力;同时,良好的温控条件有助于进一步降低预测误差及提升模型性能。

表 6 MLP 模型最优结构与超参数设置
Table 6 Optimal architecture and hyperparameters of the MLP model

Modeling Method	Group	Number of Layers	Learning Rate	L2 Regularization	Neurons in Hidden Layer 1	Neurons in Hidden Layer 2	Neurons in Hidden Layer 3
Combined Dataset Modeling	Temperature—controlled group	3	0.001652	0.000589	78	255	59
	Non—temperature—controlled group	3	0.001844	0.000354	91	248	86
LOSOVCV	Temperature—controlled group	2	0.000288	0.000058	72	65	—
	Non—temperature—controlled group	3	0.00114	0.001138	10	34	20

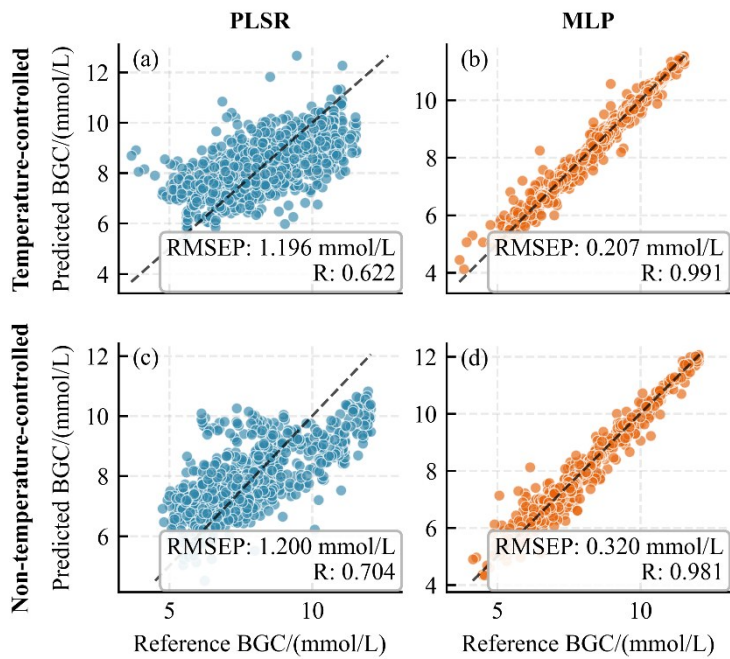


图 12 数据集混合建模下不同温控条件 PLSR 与 MLP 模型预测结果比较。(a)和(b)为温控条件下的 PLSR 模型与 MLP 模型；(c)和(d)为未温控条件下的 PLSR 模型与 MLP 模型
Fig. 12 Comparison of prediction results between PLSR and MLP under different temperature conditions based on combined dataset modeling. (a) and (b) show the PLSR and MLP model under temperature-controlled conditions; (c) and (d) show the PLSR and MLP model under non-temperature-controlled conditions.

3 讨论

本研究结合模拟计算与人体口服葡萄糖耐量实验,系统分析了不同测量条件对近红外光谱法糖浓度预测精度的影响,并对测量条件进行了“允许变化”与“不允许变化”的分类。结果表明,非线性干扰因素是限制预测性能提升的关键来源,而单纯依赖多变量建模方法难以完全弥补测量条件波动所带来的误差。

在模拟光谱体系中,液体温度变化与白噪声扰动在单波长下呈现显著非线性特征,即单位差分吸光度与干扰幅度之间不满足线性关系。这类非线性干扰会改变光谱基线形态及吸收峰强度分布,使模型难以准确分离糖浓度信息。虽然 MLP 相较于 PLSR 在处理非线性关系方面具有一定优势,但当干扰幅度超出一定范围时,预测误差仍然显著增加。这说明,算法优化虽可提升模型对复杂关系的拟合能力,但无法替代对关键测量条件的物理控制。从误差控制结果来看,将温度波动与噪声水平等非线性因素限制在较小范围内,是实现高精度预测的必要前提。相比之下,入射角度、光源色温及光源漂移等因素在单波长下呈线性分布,光谱变化规律较为稳定,模型能够通过多变量分析有效提取糖浓度信息。这表明,光谱分析方法对于线性

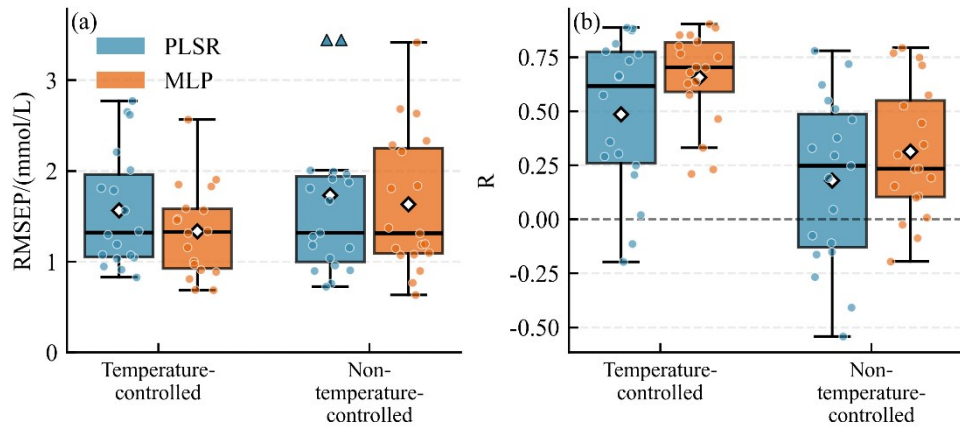


图 13 留一受试者法交叉验证(LOSOVCV)下不同温控条件的PLSR与MLP模型预测结果比较。(a)为RMSEP的分布结果；(b)为R的分布结果

Fig. 13 Comparison of prediction results of PLSR and MLP models under different temperature conditions based on leave-one-subject-out cross-validation (LOSOVCV). (a) Distribution of RMSEP; (b) Distribution of R.

扰动具有较强的鲁棒性,因此在实际系统设计中,可优先关注非线性干扰源的控制。

本研究的人体实验结果进一步印证了前期模拟分析所得出的结论。与未温控条件相比,在温控条件下,光谱信号与血糖变化之间的相关性显著增强,且在 1550 nm 附近表现出较高的血糖敏感特征,该结果与文献^[17]报道的敏感波段基本一致。这说明,适当的温度控制不仅有助于提升信号质量,还能够增强特征波段对血糖变化的响应能力。从建模结果来看,无论是组内建模还是数据集混合建模,温控条件下模型的预测精度与泛化能力均明显优于未温控条件。尤其是在跨数据集混合建模和留一受试者法交叉验证中,温控组仍能保持较高的相关性和较低的预测误差,进一步表明温度控制有助于提升模型的稳定性与鲁棒性。在未温控情况下,受试者摄入葡萄糖后,可能诱发机体代谢水平升高所引起的代谢热等一系列生理变化,会打破局部组织的平衡状态,并可能改变组织的光学特性(如吸收与散射系数)。相关研究表明,摄入葡萄糖后机体代谢水平升高,可能引起局部组织温度变化、血流灌注增加^[31-32]。由此产生的光谱信号漂移及非线性生理响应,使光谱与血糖之间的对应关系变得更加复杂,从而显著增加预测误差。对于温控组,通过将皮肤加热至约 36 °C 并持续维持该温度,可在一定程度上稳定皮肤血流状态,从而减弱由此引发的非线性生理响应^[18]。而在未温控组中,由于同时存在温度波动与血流波动等干扰因素,可能共同导致模型性能下降。本文未同步测量血流灌注等生理参数,相关机制仍需在未來研究中通过多模态测量手段进一步验证。

需要指出的是,本研究人体实验样本量较小,尚不足以全面反映个体差异对模型泛化性能的影响。同时,为简化实验流程,本研究在每次 OGTT 过程中仅采集 3-4 次指尖血作为参考血糖值。由于参考数据点数量有限、时间分辨率较低,可能在一定程度上影响光谱信号与血糖动态变化之间相关性分析的准确性。未来研究应扩大受试者数量,并提高参考血糖采样频率,例如采用连续葡萄糖监测系统(CGM)获取高时间分辨率的血糖数据,以增强模型训练与验证的可靠性。此外,还可结合更多生理参数监测手段,进一步验证测量条件控制策略的普适性,并探索更先进的非线性建模方法与自适应校正策略,在保证关键物理条件稳定的基础上持续提升模型的鲁棒性与泛化能力。

4 结论

本文围绕近红外光谱法糖浓度预测中测量条件对模型精度与稳定性的影响,结合模拟计算与人体口服葡萄糖耐量实验开展研究。结果表明,测量条件稳定性是实现高精度预测的关键,尤其是非线性干扰对模型性能具有决定性作用。在模拟体系中,液体温度变化与白噪声扰动在单波长下呈显著非线性特征,属于“不允许变化”的测量条件。仅在温度 $\leq 1\text{ }^{\circ}\text{C}$ 、白噪声 $\leq 0.002\text{ a.u.}$ 条件下, RMSEP 才能稳定在 0.2 mmol/L 以内。相比之下,入射角度、光源色温及光源漂移等因素影响近似线性,在较大范围变化下仍可将误差控制在 0.2 mmol/L 以下,因此它们是“允许变化”的测量条件。MLP 在处理仅温度干扰方面相较 PLSR 在一定范围内具有优势,但在强扰动条件下仍存在较大误差。进一步的人体实验结果表明,温控条件下的测量性

能略优于未温控条件:在温控组中,采用MLP模型的LOSOCV预测时, RMSEP为1.33 mmol/L;而在未温控组中,MLP模型的LOSOCV预测RMSEP为1.63 mmol/L。

综上,本研究表明,模型优化与测量条件控制在提升预测精度方面均具有重要作用。其中,对于某些“不允许变化”的关键测量条件,对其进行有效控制可能是实现高精度、可靠测量的重要前提。本文的研究结果可为无创血糖测量中测量条件的控制策略提供参考。

参考文献

- [1] Li T, Wang Q, An Y, et al. Infrared absorption spectroscopy-based non-invasive blood glucose monitoring technology: A comprehensive review[J]. *Biomedical Signal Processing and Control*, 2025, 106: 107750.
- [2] Davison N B, Gaffney C J, Kerns J G, et al. Recent progress and perspectives on non-invasive glucose sensors[J]. *Diabetology*, 2022, 3(1): 56-71.
- [3] Liu W, Han T, Chen W, et al. Design key points of high-performance diffuse reflectance optical sensors for non-invasive blood glucose measurement[J]. *Sensors*, 2025, 25(4): 998.
- [4] Wang Z, Zhou L, Liu T, et al. Development of non-invasive blood glucose regression based on near-infrared spectroscopy combined with a deep-learning method[J]. *Journal of Physics D: Applied Physics*, 2022, 55(21): 215401.
- [5] Yu Y, Huang J-P, Zhu J, et al. An accurate noninvasive blood glucose measurement system using portable near-infrared spectrometer and transfer learning framework[J]. *IEEE Sensors Journal*, 2020, 21(3): 3506-3519.
- [6] Fu Y, Huang M, Chen X. Fingertip capillary dynamic near infrared spectrum (DNIRS) measurement combined with multi-variate linear modification algorithm for noninvasive blood glucose monitoring[J]. *Vibrational Spectroscopy*, 2021, 113: 103223.
- [7] Uwadaira Y, Ikehata A, Momose A, et al. Identification of informative bands in the short-wavelength NIR region for non-invasive blood glucose measurement[J]. *Biomedical Optics Express*, 2016, 7(7): 2729.
- [8] Tronstad C, Elvebakk O, Staal O M, et al. Non-invasive prediction of blood glucose trends during hypoglycemia[J]. *Analytica Chimica Acta*, 2019, 1052: 37-48.
- [9] Aloraynan A, Rassel S, Xu C, et al. A single wavelength mid-infrared photoacoustic spectroscopy for noninvasive glucose detection using machine learning[J]. *Biosensors*, 2022, 12(3): 166.
- [10] Lundsgaard-Nielsen S M, Pors A, Banke S O, et al. Critical-depth raman spectroscopy enables home-use non-invasive glucose monitoring[J]. *PLOS ONE*, 2018, 13(5): e0197134.
- [11] Feng R, Geng Z, Li A, et al. Noninvasive blood glucose monitor via multi-sensor fusion and its clinical evaluation[J]. *Sensors and Actuators B: Chemical*, 2021, 332: 129445.
- [12] Caduff A, Dewarrat F, Talary M, et al. Non-invasive glucose monitoring in patients with diabetes: A novel system based on impedance spectroscopy[J]. *Biosensors and Bioelectronics*, 2006, 22(5): 598-604.
- [13] Caduff A, Talary M S, Zakharov P. Cutaneous blood perfusion as a perturbing factor for noninvasive glucose monitoring[J]. *Diabetes Technology & Therapeutics*, 2010, 12(1): 1-9.
- [14] LIU Wenbo, LIU Jin, HAN Tongshuai, et al. Simulation of the effect of dermal thickness on non-invasive blood glucose measurement by near-infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2023, 43(9): 2699-2704.
刘文博, 刘瑾, 韩同帅, 等. 皮肤真皮层厚度对近红外光谱法无创血糖测量影响的仿真[J]. *光谱学与光谱分析*, 2023, 43(9): 2699-2704.
- [15] Mayrovitz H N. Skin tissue dielectric constant: Time of day and skin depth dependence[J]. *Clinical Physiology and Functional Imaging*, 2025, 45(1): e12916.
- [16] Okabe T, Fujimura T, Okajima J, et al. Non-invasive measurement of effective thermal conductivity of human skin with a guard-heated thermistor probe[J]. *International Journal of Heat and Mass Transfer*, 2018, 126: 625-635.
- [17] Han T, Jin Liu, Rong Liu, et al. In vivo near-infrared noninvasive glucose measurement and detection in humans[J]. *Applied Spectroscopy*, 2022, 76(9): 1100-1111.
- [18] Ge Q, Han T, Liu X, et al. Effects of skin blood flow fluctuations on non-invasive glucose measurement and a feasible blood flow control method[J]. *Sensors*, 2025, 25(4): 1162.
- [19] Small G W. Chemometrics and near-infrared spectroscopy: Avoiding the pitfalls[J]. *TrAC Trends in Analytical Chemistry*, 2006, 25(11): 1057-1066.
- [20] Oshina I, Spigulis J. Beer - lambert law for optical tissue diagnostics: Current state of the art and the main limitations[J]. *Journal of Biomedical Optics*, 2021, 26(10): 100901.
- [21] 徐可欣, 高峰, 赵会娟. 生物医学光子学[M]. 北京: 科学出版社, 2011.
- [22] Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics[J]. *Chemometrics and intelligent laboratory systems*, 2001, 58(2): 109-130.
- [23] Faber N M, Song X-H, Hopke P K. Sample-specific standard error of prediction for partial least squares regression[J]. *TrAC Trends in Analytical Chemistry*, 2003, 22(5): 330-334.

- [24] Wu Y, Feng J. Development and application of artificial neural network[J]. *Wireless Personal Communications*, 2018, 102(2): 1645-1656.
- [25] Yadav J, Rani A, Singh V, et al. Prospects and limitations of non-invasive blood glucose monitoring using near-infrared spectroscopy[J]. *Biomedical Signal Processing and Control*, 2015, 18: 214-227.
- [26] Jensen P S, Bak J, Andersson-Engels S. Influence of temperature on water and aqueous glucose absorption spectra in the near- and mid-infrared regions at physiologically relevant temperatures[J]. *Applied Spectroscopy*, 2003, 57(1): 28-36.
- [27] Amerov A K, Chen J, Arnold M A. Molar absorptivities of glucose and other biological molecules in aqueous solutions over the first overtone and combination regions of the near-infrared spectrum[J]. *Applied Spectroscopy*, 2004, 58(10): 1195-1204.
- [28] LIU Jin, CAI Zijin, ZHANG Ziyang, et al. Evaluation on noninvasive blood components measurement based on photoplethysmography[J]. *Optics and Precision Engineering*, 2016, 24(6): 1264-1271.
刘瑾, 蔡子晋, 张紫杨, 等. 光电容积脉搏波无创测量人体血液成分的评估[J]. *光学精密工程*, 2016, 24(6): 1264-1271.
- [29] Ge Q, Liu W, Lu Y, et al. Skin temperature measurement based on diffuse reflection light at scattering variation independent source-detector separation[J]. *Infrared Physics & Technology*, 2024, 139: 105295.
- [30] Liu W, Chen J, Liu X, et al. Analysis on direct and additional temperature spectra of skin: A near-infrared diffuse reflectance spectroscopy study with temperature control[J]. *Infrared Physics & Technology*, 2026, 152:106187.
- [31] Reminy K, Hue O, Antoine-Jonville S. Effect of warm environment on the skin blood flow response to food intake[J]. *International Journal of Hyperthermia*, 2020, 37(1), 836-842.
- [32] Iredahl F, Högstedt A, Henricson J, et al. Skin glucose metabolism and microvascular blood flow during local insulin delivery and after an oral glucose load[J]. *Microcirculation*, 2016, 23(7): 597-605.

A Systematic Study on Measurement Condition Control and Modeling Methods for Glucose Concentration Prediction Using Near-Infrared Spectroscopy

LIU Wenbo¹, SUN Di², CHEN Jiayu¹, YAO Mingfei², LIU Xueying¹, LIU Jin¹, HAN Tongshuai¹, XU Kexin^{1,2}

(1 *State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072, China*)

(2 *Sunrise Technology Co., Ltd., Tianjin 300192, China*)

Abstract: This study aimed to systematically investigate how measurement conditions affect the accuracy and stability of glucose concentration prediction based on near-infrared (NIR) spectroscopy, and to clarify the roles of physical control and modeling methods in improving predictive performance. By combining a simulated glucose-solution spectral system with human oral glucose tolerance test (OGTT) experiments, we examined the effects of linear and nonlinear disturbances on spectral characteristics and model behavior, compared the robustness of partial least squares regression (PLSR) and multilayer perceptron (MLP), and proposed a classification strategy distinguishing measurement conditions as either “allowable variations” or “non-allowable variations.” A combined simulation-and-human-experiment framework was established. In the simulation study, an aqueous glucose solution was used as a simplified optical system to isolate the effects of controlled disturbances on NIR spectral response and prediction accuracy. Representative factors included liquid temperature variation, white noise, incident angle variation, light-source color temperature, and light-source drift. Their influences were analyzed at both the spectral and modeling levels. In particular, the relationship between disturbance magnitude and unit differential absorbance at a single wavelength was examined to determine whether the induced spectral variation followed an approximately linear or nonlinear pattern. Two modeling approaches were compared: PLSR as a representative linear regression method and MLP as a representative nonlinear learning method. Models were trained to predict glucose concentration under different disturbance scenarios, and performance was evaluated mainly by the root mean square error of prediction (RMSEP). Based on spectral response patterns and error trends, measurement conditions were classified into “allowable variations,” referring to

factors that could vary over a broad range while maintaining low prediction error, and “non-allowable variations,” referring to factors requiring strict control. To verify whether the simulation-based conclusions applied to physiological measurements, human OGTT experiments were also performed. NIR spectra were collected under temperature-controlled and non-temperature-controlled conditions. In the controlled condition, the skin was heated to approximately 36 °C and maintained near this level during measurement. Fingertip blood glucose values were collected intermittently as references. The correlation between spectral signals and blood glucose changes was analyzed under both conditions, with particular attention to features near 1550 nm. Model development included within-group modeling, combined dataset modeling, and leave-one-subject-out cross-validation (LOSOCV). The simulation results showed that different disturbances had markedly different effects on spectral prediction. Liquid temperature variation and white noise exhibited clear nonlinear characteristics at the single-wavelength level; the relationship between disturbance amplitude and unit differential absorbance was non-proportional, indicating that these factors changed spectral response in a way that could not be adequately represented by simple linear correction. Such nonlinear disturbances altered both the spectral baseline and absorption peak distribution, making it difficult for the models to separate glucose-concentration information from interference. As the amplitudes of temperature fluctuation and noise increased, prediction errors rose substantially. These two factors were therefore identified as “non-allowable variations.” Only when temperature fluctuation was limited to no more than 1 °C and white noise was restricted to no more than 0.002 a.u. could RMSEP remain stably around 0.2 mmol/L. In contrast, incident angle variation, light-source color temperature, and light-source drift showed approximately linear effects within the investigated ranges. The induced spectral changes followed stable patterns, allowing the models to recover glucose-related information with higher robustness. Even when these factors varied over a larger range, prediction errors could still be maintained below about 0.2 mmol/L. These factors were therefore classified as “allowable variations.” Comparison of the two modeling methods showed that MLP had a slight advantage over PLSR when only a limited nonlinear temperature disturbance was present. However, once temperature fluctuation became larger, the errors of both models increased significantly, indicating that nonlinear modeling cannot replace physical control of critical measurement conditions. The human OGTT experiments further supported the simulation findings. Compared with the non-controlled condition, the temperature-controlled condition produced a stronger correlation between NIR spectra and blood glucose changes, with more prominent sensitivity near 1550 nm. Model performance also improved under this condition. In LOSOCV, the MLP model achieved an RMSEP of 1.33 mmol/L in the temperature-controlled group, whereas the corresponding RMSEP in the non-controlled group was 1.63 mmol/L. Similar trends were observed in within-group and combined dataset modeling. This study demonstrates that measurement-condition stability is a key determinant of accurate glucose concentration prediction using NIR spectroscopy. The combined evidence from simulation and human OGTT experiments indicates that nonlinear disturbances are the main factors limiting predictive performance. Temperature variation and white noise were identified as critical “non-allowable variations” because their effects were strongly nonlinear and could not be sufficiently compensated for by either PLSR or MLP. By contrast, incident angle variation, light-source color temperature, and light-source drift behaved approximately linearly and could therefore be tolerated within a wider range as “allowable variations”. Although MLP showed a limited advantage over PLSR under nonlinear interference, model optimization alone was insufficient when disturbance amplitude became large. Human experiments further confirmed that proper temperature control strengthened the spectral correlation with blood glucose and improved prediction under cross-subject validation. Overall, effective physical control of key nonlinear factors is a prerequisite for high-precision prediction.

Key words: Near-infrared spectroscopy; non-invasive blood glucose measurement; multivariate analysis methods; oral glucose tolerance test; measurement conditions

OCIS Codes: 300.6340; 280.1415; 170.1580; 170.6935

CSTR: 32255.14.gzxb20265507.0717001