capstract:2611.00055v1 [eess.SY] 20 Nov 2026

Regulation of Remote Meeting Temperature Using an Unoccupied Conference Room

Mara Voss
Department of Systems Engineering
Westmere Institute of Technology
Renswick
United Kingdom

Leonie Serrat
Department of Systems Engineering
Westmere Institute of Technology
Renswick
United Kingdom

***Abstract*—We study a building controller that regulates the temperature of a fully remote meeting by cooling an assigned, unoccupied conference room. The controller combines a vocal arousal estimate from the video call with the room air temperature and lowers the room setpoint when this composite meeting-temperature signal is high. A two-state model shows that the remote discussion is uncontrollable from the room HVAC input, although the composite signal responds to cooling. In 1,000 paired simulated meetings, increasing the controller gain from zero to 2 °C per arousal unit reduces the fraction of composite samples above a fixed threshold from 60.69% to 23.63%. The corresponding vocal-channel fraction remains 60.69%, and the latent arousal trajectory is unchanged. Normalized cooling input rises from 4.500 to 5.161 °C, while the mean minimum room temperature falls from 22.000 to 21.410 °C. Removing the room-temperature term from the composite signal eliminates the apparent improvement. An independent vocal channel detects the discrepancy but does not make the remote meeting controllable. The study uses synthetic meetings and an illustrative thermal model.**

***Index Terms*—building control, networked control, controllability, vocal arousal, measurement feedback**

I. Introduction

A building controller can respond to occupancy forecasts, temperature measurements and individual comfort requests [1], [2]. Meetings introduce another variable of interest: the intensity of the discussion. Continuous vocal estimates of arousal have been studied in affective computing [3], and the development of emotion during conflict has been examined at the individual and group levels [4]. These measurements are normally kept outside the building automation system.

We consider a controller that uses them to regulate meeting temperature. The meeting is held entirely by video call. All six participants join from separate locations, and none is in the conference room assigned to the meeting. The building nevertheless keeps the room available so that its cooling system can respond to the discussion. When the meeting becomes heated, the controller lowers the room temperature.

The controlled signal combines an estimated vocal arousal score with the temperature of the assigned room. This allows a conventional cooling loop to act on the signal even when the participants are remote. It also raises a question about what the loop has regulated. If the room is empty, lowering its temperature cannot change the participants' physical environment. Yet the composite meeting-temperature signal can fall because it includes the temperature of the room itself.

We model this arrangement as a two-state system and establish a controllability result for the remote case. We then compare the composite signal with an independent vocal channel in paired simulations. No meeting recordings, occupants or building measurements were used; the meeting and estimator parameters are illustrative. The aim is to identify what can be inferred from the closed-loop signals, rather than to validate an emotion recognizer.

II. System Model

A. Remote discussion and assigned room

Let x_k be a dimensionless deviation of meeting arousal from its neutral level at five-minute sample k. It is a scalar aggregate for the six participants, not a temperature in °C. We use the linear model

$$x_{k+1} = a x_k + d_k + w_k, \tag{1}$$

where $a = 0.90$, d_k represents an agenda-related impulse and w_k is an innovation. The vocal channel is $s_k = x_k + \epsilon_k$. It provides a noisy measurement of the simulated state; it is not a fitted speech model. In particular, arousal is not the same as disagreement, and a high value need not mean that a meeting has gone badly.

The assigned room has air temperature T_k in °C. With outdoor temperature $T^{\text{out}} = 28$ °C, an illustrative single-state thermal model is

$$T_{k+1} = T_k + h\left[\frac{T^{\text{out}} - T_k}{\tau} + u_k\right], \tag{2}$$

where $h = 1/12$ h, $\tau = 4$ h and u_k is signed heating input in °C/h. The nominal physical setpoint is $T_0 = 22$ °C. We define the room's cooling displacement as $z_k = T_0 - T_k$.

The participants remain at their remote locations for the entire meeting. Thus u_k affects T_k but does not enter Eq. (1). A connection between conference-room air and the voice channel could exist if a participant were in the room, or if the physical temperature were itself read over the call; neither is assumed here. We return to these restrictions in Section VI.

B. Meeting-temperature signal

The building system forms a composite signal

$$y_k = s_k + \gamma(T_k - T_0) = x_k + \epsilon_k - \gamma z_k, \tag{3}$$

with $\gamma = 0.5$ arousal units per °C. The first term comes from the call and the second from the assigned room. The coefficient converts the two scales; it is a design parameter, not an empirical relationship between air temperature and emotion. For the neutral room, $y_k = s_k$. For a room cooled below 22 °C, the composite signal is lower than the vocal signal even if the call is unchanged.

The cooling displacement requested by the meeting controller is

$$q_k = \mathrm{clip}(K \max(y_k, 0), 0, 1.5)\,, \tag{4}$$

where K is in °C per arousal unit and q_k is in °C. The physical setpoint is $T_0 - q_k$. A feedforward proportional controller follows it:

$$u_k = \frac{T_k - T^{\mathrm{out}}}{\tau} + 1.8(T_0 - q_k - T_k). \tag{5}$$

Substitution into Eq. (2) yields $z_{k+1} = 0.85 z_k + 0.15 q_k$. The input is sampled once per five minutes; there is no occupancy override. The clipping in Eq. (4) limits the requested cooling displacement to 1.5 °C. The composite signal can be negative, in which case the meeting controller requests no additional cooling.

III. Control Properties

A. Controllability of the remote meeting

With the state ordered as (x, z), and taking q as the control input, the disturbance-free open-loop system has

$$A = \begin{pmatrix} a & 0 \\ 0 & 0.85 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0.15 \end{pmatrix}. \tag{6}$$

The controllability matrix $[B\ AB]$ has rank one. Its first row is zero. No sequence of room-temperature commands can move x from the trajectory fixed by its initial condition and disturbances. This conclusion does not depend on the particular feedback gain or on saturation: by induction, Eq. (1) contains no q or T. The result applies to the specified remote configuration; it is not a statement about meetings whose participants occupy the controlled space.

B. Apparent regulation

Suppose for the moment that measurement noise is zero, $y > 0$ and the actuator is unsaturated. The closed-loop matrix is

$$A_{\mathrm{cl}} = \begin{pmatrix} a & 0 \\ 0.15K & 0.85 - 0.15\gamma K \end{pmatrix}. \tag{7}$$

Its eigenvalues are a and $0.85 - 0.15\gamma K$. The first is unaffected by K. With $\gamma = 0.5$, the linearized system is stable for $0 \le K < 24.67$, assuming $|a| < 1$. This is a local, unsaturated result; the simulations below use the clipped law in Eq. (4).

For a constant positive arousal state x with an unsaturated equilibrium, $z = Ky$ and therefore

$$y = \frac{x}{1 + \gamma K}, \qquad z = \frac{Kx}{1 + \gamma K}. \tag{8}$$

Increasing gain reduces the composite score while leaving x fixed. The equilibrium also shows the tradeoff: as the score approaches zero, the required room displacement approaches x/γ. This calculation assumes $Ky < 1.5$; it does not describe saturated operation. Since the room temperature is measured, an independent observer can recover the vocal channel as $s_k = y_k + \gamma z_k$. The discrepancy is therefore detectable without observing the latent x_k.

IV. Simulation Design

We simulate 1,000 independent three-hour meetings at five-minute intervals, using seeds 27000 through 27999. All participants remain remote. Initial arousal is normal with mean 0.08 and standard deviation 0.03. In Eq. (1), independent innovations have standard deviation 0.025. At samples 6, 18 and 26, the agenda adds an independently drawn uniform impulse between 0.38 and 0.62. Other d_k are zero. The independent vocal measurement noise has standard deviation 0.04. All Gaussian terms have mean zero except the initial condition.

Each meeting begins with $T = T_0$. We compare gains $K = 0$, 0.5, 1, 2 and 4 with $\gamma = 0.5$. The $K = 0$ case is the baseline: the room is held at 22 °C and the composite signal equals the vocal signal. An additional ablation uses $K = 2$ with $\gamma = 0$, so the same controller reacts to the vocal channel without mixing in room temperature. All policies use the same arousal trajectory and vocal noise for each seed.

Our primary score is the fraction of samples with $y_k > 0.3$. The threshold is an arbitrary design choice and is applied to the vocal and latent channels as well. We report mean room minimum temperature per meeting and an integrated cooling-input proxy,

$$C = h \sum_k \max(-u_k, 0), \tag{9}$$

in °C. This is normalized thermal input for the model, not electricity or energy in kWh; thermal capacitance and equipment efficiency are unspecified. With $K = 0$, the model requires $C = 4.5$ °C over three hours to hold the room at 22 °C against the outdoor load. For each metric, 95% Monte Carlo intervals are calculated from the 1,000 independent meeting-level values as $\bar{v} \pm 1.96 s/\sqrt{1000}$. Differences between policies use paired seeds.

V. Results

A. Composite and independent measurements

Table I summarizes the gains. With no meeting control, the composite and vocal channels both exceed 0.3 in 60.69% of samples. At $K = 2$ the composite exceeds the threshold in 23.63% of samples, a paired reduction of 37.05 percentage points (95% Monte Carlo half-width 0.43 points). The vocal

channel remains at 60.69%, and the latent state remains above threshold for 60.94% of samples. The two invariant quantities are identical across gains by construction, not merely close within the simulation error.

TABLE I

MEAN RESULTS OVER 1,000 PAIRED SIMULATED MEETINGS. UNCERTAINTIES ARE 95% MONTE CARLO HALF-WIDTHS; THRESHOLD FRACTIONS ARE IN PERCENT. COOLING INPUT IS A NORMALIZED THERMAL PROXY, NOT ELECTRICAL ENERGY.

Policy	Composite above 0.3 (%)	Vocal above 0.3 (%)	Latent above 0.3 (%)	Cooling input (°C)	Minimum room temp. (°C)
$K = 0$	60.69 ± 0.51	60.69 ± 0.51	60.94 ± 0.50	4.500	22.000
$K = 0.5$	47.09 ± 0.52	60.69 ± 0.51	60.94 ± 0.50	4.767 ± 0.002	21.784 ± 0.002
$K = 1$	36.19 ± 0.43	60.69 ± 0.51	60.94 ± 0.50	4.945 ± 0.003	21.627 ± 0.003
$K = 2$	23.63 ± 0.28	60.69 ± 0.51	60.94 ± 0.50	5.161 ± 0.005	21.410 ± 0.004
$K = 4$	16.32 ± 0.21	60.69 ± 0.51	60.94 ± 0.50	5.339 ± 0.007	21.199 ± 0.005
$K = 2, \gamma = 0$	60.69 ± 0.51	60.69 ± 0.51	60.94 ± 0.50	5.802 ± 0.009	20.984 ± 0.007

For $K = 2$, mean cooling input is 0.661 ± 0.005 °C above the baseline and the average minimum temperature is 21.410 °C. The extra input is expended in the assigned room, which has no attendees. Fig. 1 shows one simulated meeting. After each agenda impulse the vocal signal rises; the composite signal subsequently falls below it as the empty room cools. The two channels do not measure the same output once the controller is active.

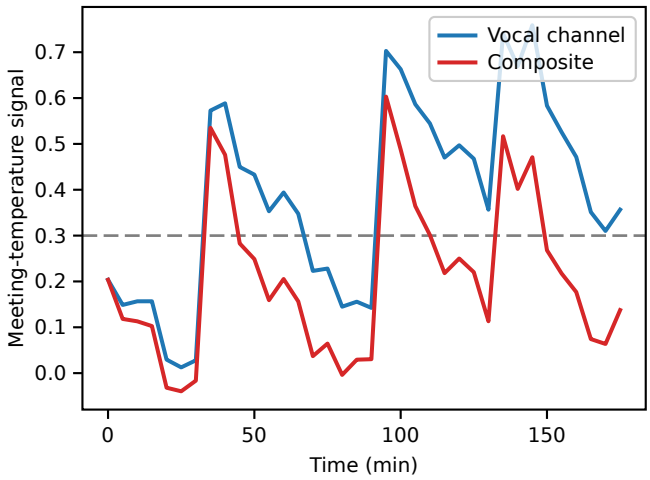


Fig. 1. Vocal and composite meeting-temperature signals in the first simulated meeting with gain K = 2. The dashed horizontal line marks the 0.3 reporting threshold.

B. Gain and signal mixing

Fig. 2 plots the threshold fractions against gain. The composite rate decreases with gain in this experiment, while the independent vocal rate is unchanged. At $K = 4$ the composite rate is 16.32%. The linear equilibrium in Eq. (8) predicts the direction of this change, although the numerical results include transient dynamics, noise and saturation.

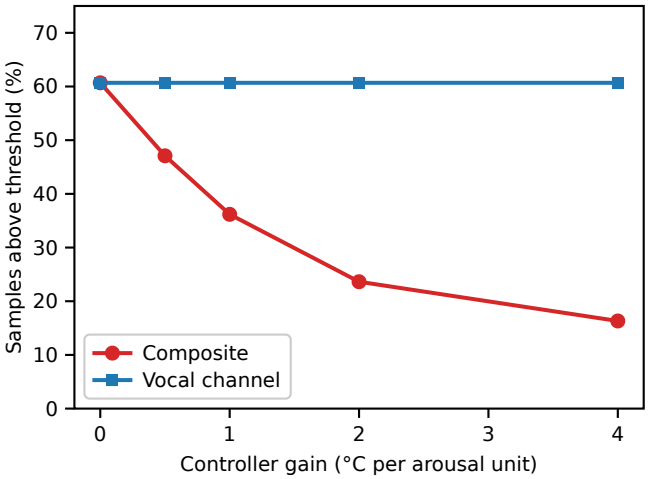


Fig. 2. Fraction of samples above the threshold against controller gain. Points are means across 1,000 paired meetings; the vocal curve is common to all gains.

The $\gamma = 0$ ablation makes the distinction visible. With $K = 2$, the controller still cools the room in response to the vocal channel, and its mean cooling input is 5.802 °C. The composite and vocal channels then agree exactly and both exceed the threshold 60.69% of the time. Removing the temperature term removes the apparent improvement, although it does not stop the controller from acting.

C. A measurement correction

For the mixed controller, $y_k + \gamma z_k$ equals the vocal measurement at every sample, including when the actuator saturates. Computing the score from this corrected channel recovers the 60.69% threshold rate at every gain. It does not recover the latent state exactly because vocal measurement noise remains. This distinction matters: a second meter in the room would confirm the temperature, but it would not provide another measurement of the participants.

VI. DISCUSSION

The meeting controller has an actuator and a sensor, and its composite signal responds to its command. The affected component of that signal is the temperature of the assigned room. The model gives the actuator no path to the remote partici-

pants. Consequently a stable reduction in the controller’s error need not imply regulation of the discussion. In the absence of the independent vocal channel, the building dashboard would report fewer high-temperature meeting samples as the room is cooled more aggressively.

The rank calculation depends on the physical separation specified in the experiment. An attendee joining from the assigned room, an open audio channel carrying the fan noise, or a notification announcing the setpoint could introduce a path from u to the meeting. The sign and size of such a path would need to be measured. The scalar arousal process, linear room, fixed outdoor temperature and Gaussian vocal noise are also deliberately limited. No psychological effect of cold air, and no performance of a real speech model, is estimated here.

The gain comparison isolates the measurement effect because all policies see the same simulated call. It does not predict how often real meetings exceed an arousal threshold. Arousal can accompany engagement or agreement as well as conflict, and a threshold on a synthetic scalar is not an assessment of meeting quality. The mixing coefficient is assumed; a room thermometer cannot calibrate the vocal component.

A separate vocal measurement is sufficient to identify the discrepancy in this model, but it cannot make the empty room an effective actuator. For a fully remote meeting the room can be made colder, the meeting-temperature display can be brought closer to its target, and the participants can continue exactly as before.

Data and Code Availability

All data are synthetic. Model equations, distributions, policy parameters and random seeds are specified in Section II and Section IV.

References

[1] J. R. Dobbs and B. M. Hencey, “Model Predictive HVAC Control with Online Occupancy Model,” *Energy and Buildings*, vol. 82, pp. 675–684, 2014, doi: 10.1016/j.enbuild.2014.07.051.

[2] S. Purdon, B. Kusy, R. Jurdak, and G. Challen, “Model-Free HVAC Control Using Occupant Feedback,” in *Proc. 2013 IEEE 38th Conference on Local Computer Networks Workshops*, 2013, pp. 84–92.

[3] D. Bone, C.-C. Lee, and S. Narayanan, “Robust Unsupervised Arousal Rating: A Rule-Based Framework with Knowledge-Inspired Vocal Features,” *IEEE Transactions on Affective Computing*, vol. 5, no. 2, pp. 201–213, 2014, doi: 10.1109/TAFFC.2014.2326393.

[4] G. A. van Kleef and S. Côte, “Emotional Dynamics in Conflict and Negotiation: Individual, Dyadic, and Group Processes,” *Annual Review of Organizational Psychology and Organizational Behavior*, vol. 5, pp. 437–464, 2018, doi: 10.1146/annurev-orgpsych-032117-104714.